

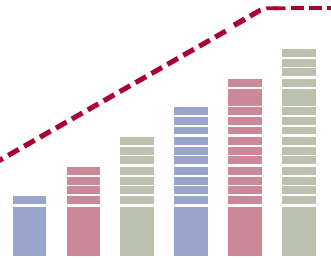


Abstracts



2016

European Conference on
Quality in Official Statistics



Madrid (Spain) 31 May - 3 June



Welcome to Q2016

Dear Colleagues,

The National Statistical Institute of Spain (INE) and Eurostat are pleased to invite you to the European Conference on Quality in Official Statistics (Q2016) which will be held in the “Círculo de Bellas Artes” in Madrid, Spain, on 1-3 June 2016.

In addition, a series of short training courses will take place on the day previous to the start of the Conference (31 May 2016).

Since their creation in 2001, these Conferences have become an excellent framework to present and discuss the progress and development of quality in official statistics, as well as to exchange methods and good practices between experts in different areas (statistical offices, international organizations, researchers and academics).

The Conference aims to cover relevant and innovative topics on quality ranging from the challenges and the new paradigm of quality in an information and knowledge-driven society including big data and multi-source statistics, to governance and management aspects like the ones linked to the ESS Vision 2020 or the lessons learned from 2013-2015 peer reviews in the European Statistical System. Furthermore, as an open forum of debate, it represents an opportunity to introduce innovation in the measurement and management of the quality in statistical domains and in specific statistical products.

The Conference website offers information on the conference venue, accommodation, short courses and leisure-time activities as well as deadlines for submission of papers. Registration is now available.

Looking forward to seeing you in Madrid!

Gregorio Izquierdo
President. INE of Spain

Walter Radermacher
Director General Eurostat. Chief Statistician of the European Union



The National Statistics Institute of Spain (INE), is the central body of official statistics in Spain. INE's mission is to collect, produce and disseminate high quality and relevant official statistical information for society in an effective, independent, professional and complete way, and to make it available to all users. The INE is responsible for the overall coordination of the statistical services of the state, for applying and enforcing rules of statistical confidentiality, and for enhancing the relations with the statistical community worldwide. The activity of INE is governed both by the Spanish Law on the Public Statistical Services and by EU statistical legislation.

eurostat

Eurostat is a Directorate-General of the European Commission. Its mission is to be the leading provider of high quality statistics on Europe.

Together with the Member States, Eurostat implements standards, methods and classifications for the production and dissemination of comparable, reliable and relevant data. Users of European statistics include the Commission and other institutions of the European Union, national governments, international organisations, businesses, universities, media and public at large. Eurostat also supports non-member countries, including the candidate countries, in adapting their statistical systems.

SCIENTIFIC COMMITTEE

- Mr. Walter Radermacher, *Director General, Eurostat, Co-Chair.*
- Mr. Gregorio Izquierdo, *President, National Statistics Institute (INE), Spain, Co-chair.*
- Mr. Eduardo Barredo, *Director, Eurostat.*
- Ms. Alda Carvalho, *Director General, Statistics Portugal.*
- Mr. Miguel Ángel De Castro, *Director, National Statistics Institute (INE), Spain.*
- Mr. Martti Hetemäki, *Chairman, ESGAB.*
- Ms. Mariana Kotzeva, *Deputy Director General, Eurostat.*
- Mr. Vijay N. Nair, *President, International Statistical Institute.*
- Mr. Konrad Pesendorfer, *Director General, Statistics Austria.*
- Mr. David Rios, *High Council on Statistics, Spain.*
- Mr. Aurel Schubert, *Director General, European Central Bank.*
- Ms. Ineke Stoop, *Chairman, ESAC.*

PROGRAMME COMMITTEE

- Mr. Jean Pierre Poncelet, *Eurostat, Co-chair*
- Mr. Agustín Cañada, *National Statistics Institute (INE), Spain, Co-chair*
- Mr. Thomas Burg, *Statistics Austria.*
- Mr. Tasos Christofides, *ESAC.*
- Mr. Kari Djerf, *Statistics Finland.*
- Ms. Maria Dologová, *Statistical Office of Slovak Republic.*
- Ms. Zsuzsanna Kovacs, *Eurostat.*
- Ms. Pilar Martín Guzmán, *International expert.*
- Mr. Miguel Ángel Martínez, *National Statistics Institute (INE), Spain.*
- Mr. Jean-Marc Museux, *Eurostat.*
- Ms. Marina Signore, *Italian National Institute of Statistics.*
- Mr. Peter van de Ven, *Organization for Economic Co-operation and Development.*
- Ms. Maria João Zilháo, *Statistics Portugal.*

09:00-10:30	Opening - Plenary Session 1 Welcome address: Gregorio Izquierdo (INE), Miguel De Castro (INE), Mariana Kotzeva (Eurostat)	
	Producing Good Quality Official Statistics: A Shared Experience Opening Keynote Speech: Wayne Smith (Statistics Canada)	
10:30-11:00	Coffee Break	
	Fernando de Rojas	Columns
11:00-12:30	33 - SPECIAL SESSION: Peer Reviews: A Tool to Enhance Trust in European Statistics?	1 - Enterprise Architecture in Statistical Offices
	The peer review as a main driver for statistics Austria's strategy 2020 - Statistics Austria	Enterprise architecture in the Hungarian Central Statistical Office - Hungarian Central Statistical Office
	Trustworthiness, quality, value: What assurance do independent assessments of code-compliance give? - UK Statistics Authority	Governmental and statistical enterprise architecture frameworks as tools for modernizing the national statistical system - Statistics Finland
	ESS peer reviews: An efficient means to implement the European Statistics Code of Practice - Eurostat, Luxembourg	Re-engineering and re-designing statistical production, modernisation; integrated production methods - Eurostat, Luxembourg
		Statistics Catalonia: Towards total integration of statistics production - Statistical Institute of Catalonia (Idescat)
		Redesign of the statistical information system: Czech experience - Czech Statistical Office
12:30-14:00	Lunch	
13:10-13:55	Speed Talk Session 1	



Opening - Plenary Session 1 Welcome address: Gregorio Izquierdo (INE), Miguel De Castro (INE), Mariana Kotzeva (Eurostat)			
Producing Good Quality Official Statistics: A Shared Experience Opening Keynote Speech: Wayne Smith (Statistics Canada)			
Coffee Break			
Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna	
2 - Quality Improvement Methods in Household Surveys	3 - Quality Challenges in Social Statistics: Preserving Privacy and other Issues	4 - Administrative Data: Cross-Cutting Issues	
Redeveloping the household budget survey - Statistics Norway	Available methods for privacy preserving record linkage on census scale data - City University London (UK), University Duisburg-Essen, (Germany)	The assurance of administrative data - a proportionate approach - UK Statistics Authority	
From output to input/output harmonisation in the EU LFS - Central Statistical Office of Poland	Developing longitudinal statistics on recipients of welfare benefits and their labour market position - Statistics Norway	When is administrative data enough to replace statistical information: A based on census comparison quality indicator - Statistics Portugal	
A standard tool for metadata quality assessment of external sources in the statistical production (QMETATOOL) - National Statistical Institute (INE), Spain	Do respondents answer differently in web survey than in face-to-face interview: Field work experiment from the European Health Interview Survey (EHIS) - National Institute of Public Health, Slovenia	Evaluating the quality of administrative data as input for official statistics - Statistics Denmark	
The use of regression models in labour market flow statistics - Eurostat, Luxembourg	The privacy protecting aspect of indirect questioning designs - Johannes Kepler University JKU, Austria	Mapping potential administrative data for statistics purposes - Statistic Portugal approach within public administration - Statistics Portugal	
Moving towards web data collection in household surveys - Statistics Finland	A latent class model to estimate labour cost from multi-source data - Istat, Italian National Statistical Institute		
Lunch			
Speed Talk Session 2			

	Fernando de Rojas	Columnas	
	5 - Quality Management and Governance	34 - SPECIAL SESSION: Big Data and Official Statistics: Challenges and Opportunities	
	Inter-organisational trust in the context of microdata exchange in the ESS - Eurostat, Luxembourg	Enhancing the foundation of official economic statistics with Big Data - U.S. Census Bureau	
	The commitment on confidence in European statistics as a means for enhancing quality of official statistics and efficiency of the national statistical system: Experience of the Czech Republic - Czech Statistical Office and CERGE-EI	Big Data and the integrated production of official statistics - Statistics Iceland	
14:00-15:30	Reforming information quality assurance at the National Institute of Statistics and Geography (INEGI) in Mexico - INEGI, Mexico	Comparative assessment of three quality frameworks for statistics derived from Big Data: The case of Wikipedia page views and automatic identification systems - Eurostat, Luxembourg	
	Coordinating migration related statistics and analyses at Statistics Norway – bridging users and producers - Statistics Norway	Measuring representativeness of internet data sources through linkage with register data - Poznań University of Economics and Business, Poland	
15:30-16:00	Coffee Break		



	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	6 - Peer Review: Learning from Countries' Experiences	7 - Quality Indicators	8 - Integrated Production and Quality
	The experiences of Turkey as an enlargement country in the second round of peer review - Turkish Statistical Institute	Challenges and discoveries in developing quality indicators for the GSBPM - Turkstat, Statistics Canada, Istat	Harmonisation at the ONS: Comparability of survey and administrative data in conjunction with European influences - Office for National Statistics, UK
	Peer review (2013-2015): Lessons learned, challenges and opportunities to the INE and the Spanish Statistical System - National Statistical Institute (INE), Spain	Quality indicators for the individual level – potential for the assessment of subgroups - Statistics Austria	Facing the challenge to increase quality while working more efficiently using Lean Operational Management and Lean Six Sigma at Statistics Netherlands - Statistics Netherlands
	Perspective of three different countries on peer review on ESCoP: Lessons learned, future challenges and common features on coordination role - Statistics Netherlands, ISTAT, Statistics Austria	Quality assessment of statistics in Eustat - Basque Statistics Office (Eustat), Spain	Validation in the ESS: A member state perspective - Federal Statistical Office (Destatis), Germany
	The 2014/2015 European peer reviews – facilitating for statistical cooperation in the Nordic countries - Statistics Norway, Statistics Finland, Statistics Denmark	Agricultural administrative sources data quality: A proposal for standardized indicators - Istat, Italian National Statistical Institute	Integrated data processing system (EAR) - Hungarian Central Statistical Office
	Second round of peer reviews – experience of (some experts at) Statistics Finland - Statistics Finland	Inclusive growth indicators on regional level - German Institute for Economic Research (DIW)	
Coffee Break			

	Fernando de Rojas	Columnas	
	35 - SPECIAL SESSION: The Statistics Code of Practice for the ENP South Countries	9 - Quality Management Systems 1	
	The development of a cop for the ENP south countries - Eurostat, Luxembourg	Relevance and benefit of commercial quality improvement methodologies at Statistics Canada - Statistics Canada	
	PCBS experience in implementing European Code of Practice (CoP) - Palestinian Central Bureau of Statistics	Quality within ONS – providing a framework for statistical producers and assurance for our users - Office for National Statistics, UK	
16:00-17:30	The implementation of the code of practice: Main challenges: Morocco case - Statistics Directorate, High Commission for Planning, Morocco	How to cope with all those rules - Statistics Netherlands	
	Implementation of Code of Practice for the ENP south countries- example: The access and use of administrative data in Morocco - High Commission for Planning, Morocco	A systematic approach to quality: The development and implementation of a quality management framework in the Central Statistics Office - Central Statistics Office, Ireland	
	The statistics Code of Practice in Jordan: Challenges and future - Department of Statistics Jordan	How to increase quality in the central banks statistical business process? The experience of Banco de Portugal - Banco de Portugal	
	The implementation of the Code of Practice: The Algerian case in the context of using new technologies for the 2018 census - Office for National Statistics, Algeria		
17:35-18:20	Speed Talk Session 3		



Day 1 - Wednesday 1 June Programme

	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	10 - Business Register	11 - Models & Early Estimates	12 - Enhancing Statistical Literacy
	Quality management of cross-border relations in the European system of interoperable statistical business registers - Statistics Austria	Using administrative data and model based estimation for improving Italian agriculture statistics - Istat, Italian National Statistical Institute	Figuring figures: Exploring Europeans' awareness of official economic statistics - University of Oviedo, Applied Economics, Spain
	Quality improvement of the Eurogroups register - Eurostat, Luxembourg	Plausibility assessment of flash estimates of the income distribution - Eurostat, Luxembourg	A knowledge-driven society – challenges for Polish official statistics - Central Statistical Office of Poland
	Applying the generic statistical business process model (GSBPM) to the national business register; the Spanish experience - National Statistical Institute (INE), Spain	Models in official statistics - Statistics Netherlands	Statistical literacy portal as a marketing tool - National Statistical Institute (INE), Spain
	Profiling: A new and better way to apprehend the globalization - Insee, France	GDP flash estimates: Sophistication through simplicity - Statistics Lithuania	European Master in Official Statistics in use: Short-term goals vs. Long-term vision - Eurostat, Luxembourg
	Dated - undated - outdated: The issue of time stamps and reference dates in the statistical business register - Federal Statistical Office (Destatis), Germany	Early estimates of environmental accounts (economy-wide material flow accounts) - Eurostat, Luxembourg	Statistical training as a factor in quality - National Statistical Institute (INE), Spain
	Speed Talk Session 4		

	Fernando de Rojas	Columnas	
	13 - Coordination of Statistical System	36 - SPECIAL SESSION: Competence Management in Statistics	
	After the peer review – steps towards enhancing the co-ordination role of HCSO in the national statistical system of Hungary - Hungarian Central Statistical Office	A competency measurement model - Central Bank of Italy	
	NSS coordination: A tool or a fundamental principle? - National Statistical Institute (INE), Spain	Using a business architecture for identifying the competence needs of a statistical institute - Istat, Italian National Statistical Institute	
09:30-11:00	Tools and methods for national coordination in national statistical system - Statistics Finland	The European statistical system's enterprise architecture reference framework and capability model - Eurostat, Luxembourg	
	Coordination of NSS – requirements and practice - Central Statistical Bureau of Latvia	The competencies for a register based statistical institute - Statistics Sweden	
	Strategic planning in a decentralised statistical system - the German example - Federal Statistical Office (Destatis), Germany		
11:00-11:30	Coffee Break		



Day 2 - Thursday 2 June Programme

	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	14 - Satisfying User's Needs: Communication	15 - Methodology: Linkage and Modelling	16 - Quality Assessment & Audits
	Training for compilers of statistical releases - Statistics Finland	Improved maritime statistics with Big Data - Statistics Sweden	Introducing a framework for process quality in national statistical institutes - Istat, Italian National Statistical Institute
	How to tailor press work if quality standards of official statistics conflict with media interests - Federal Statistical Office (Destatis), Germany	Determining permanent residency status using registers in Estonia - Statistics Estonia	Minding the store: An internal audit program for demographic programs at the U.S. Census Bureau - U.S. Census Bureau
	Communication and quality in official statistics translated into English - Istat, Italian National Statistical Institute	Latent class multiple imputation for multiply observed variables in a combined dataset - Tilburg University	Statistical quality by design: Certification, culture and management - Statistics Netherlands
	Communicating revisions and the timing of revisions to users - Federal Statistical Office (Destatis), Germany	Exploring the statistical matching possibilities for the European quality of life survey - European Foundation for the Improvement of Living and Working Conditions (Eurofound)	Assessing quality control: Evaluating the quality audits - U.S. Census Bureau
			A new process for assessing the quality of an output makes better quality statistics - Office for National Statistics, UK
Coffee Break			

	Fernando de Rojas	Columnas	
	17 - Quality Reporting	18 - Multi-Source Statistics	
	ESS standard quality reporting implementation: The point of view of a national statistical institute - Istat, Italian National Statistical Institute	Quality assessment of multi-source statistical processes - Istat, Italian National Statistical Institute	
	How to fulfill all demands for quality reporting - Statistics Denmark	Spain 2021. Why this census will have more quality than the previous one? - National Statistical Institute (INE), Spain	
11:30-13:00	Quality data sheets - Federal Statistical Office (Destatis), Germany	Designing the integration of register and survey data in earning statistics - Istat, Italian National Statistical Institute	
	ESMS implementation in statistics Estonia - Statistics Estonia	Correction for linkage error in population size estimation - Statistics Netherlands / VU University Amsterdam	
		Improving the statistical process in the hotel occupancy survey - National Statistical Institute (INE), Spain	
13:00-14:30	Lunch		



Day 2 - Thursday 2 June **Programme**

	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	37 - SPECIAL SESSION: Governance and Coordination of the National Statistical Systems in the Enlargement and ENP-East Countries	19 - Human Resources Development: A Quality Culture	20 - Big Data Oriented Systems
	Overview on the set-ups of the institutional framework in the enlargement and ENP east countries - Eurostat, Luxembourg	Development of expertise and multi-skillness at Statistics Finland - Statistics Finland	Assessment of risks in the use of Big Data sources for producing official statistics – results of a stakeholder survey - Eurostat, Luxembourg
	Governance structure of Turkish statistical system within the context of coordination mechanisms - Turkish Statistical Institute	Quality at any cost? – Examining the hard reality of job reductions while promoting the quality agenda. - Office for National Statistics, UK	Using huge amounts of road sensor data for official statistics - Statistics Netherlands
	Implementation of the amendment to regulation 223/2009 and the coordination role of the national statistical institutes - Statistical Office of Montenegro	Enhancing quality and efficiency through participatory processes: Aiming for the ESS Vision 2020 - Statistics Norway	Measuring quality of statistics based on mobile phone data – the case of population, tourism and balance of payments statistics in Belgium - Proximus, Eurostat, Statistics Belgium
	National statistical system of the Republic of Belarus: Coordination issues - National Statistical Committee of The Republic of Belarus	Creating flow and reducing cycle time by reorganizing teams around components in Statistics Estonia - Statistics Estonia	Fuzzy sentiment analysis using Spanish tweets - Instituto de Estudios Fiscales, Spain
		Towards a global education in official statistics - University Complutense, Madrid; National Statistical Institute (INE), Spain	Data ecosystem: A new challenge for official statistics - Istat, (Italian National Statistical Institute), University Tor Vergata, Italy
Lunch			

	Fernando de Rojas	Columnas	
	21 - Administrative Data Systems	38 - SPECIAL SESSION: Synergies for Europe's Research Infrastructures in the Social Sciences and Official Statistics	
	Reviewing the general and centralised system of collecting administrative data and how it responds to the guidelines of the code of practice and the peer review results in 2013-2014 - Statistics Finland	Synergies for Europe's research infrastructures in the social sciences and official statistics - City University London, UK	
	Improving quality in the estimation of "true" economic performance of entrepreneurs by integrating statistical and administrative data: a new method for measuring under-reporting - Istat, Italian National Statistical Institute	Experiment for testing questionnaire translation methods in the European social survey (ESS): ask the same question versus more adaptive approaches - GESIS-Leibniz Institute for the Social Sciences, Germany	
14:30-16:00	All that glitters is not gold! - Statistics Portugal	Measuring occupations: Respondent's self-identification from a large database - University of Amsterdam / AIAS, Netherlands	
	Quality evaluation for a statistical register: The Italian frame SBS - Istat, Italian National Statistical Institute	Administrative data linking in the GGP: Enriching administrative data with surveys - Netherlands Interdisciplinary Demographic Institute	
	Quality assessment and validation of administrative data sources in health statistics - National Institute of Public Health (NIJZ)	New forms of data – legal, ethical and quality issues - NSD (Norway), NIDI (Netherlands) , UKDA (England)	
16:00-16:30	Coffee Break		
16:30-17:30	Plenary Session 2 Challenges of Quality Assurance for Big Data Keynote Speech: Genoveva Ruzic (Statistical Office of the Republic of Slovenia)		



Day 2 - Thursday 2 June Programme

	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	22 - Macroeconomics	23 - Quality Management Systems 2	24 - Metadata Systems
	The asymmetry dilemma with Intrastat – which data is the better one? National experiences out of the ESS VIP “redesign of Intrastat” - Federal Statistical Office (Destatis), Germany	Quality work in statistics – from Q2001 to 2016 - Statistics Norway	Practical experience in the implementation of data and metadata standards in the ESS - Eurostat, Luxembourg
	Automatic balancing of the national accounts - Statistics Sweden	The quality supporting framework of the ESS vision 2020 - Eurostat, Luxembourg	Process metadata development and implementation under the GSBPM v5.0 at Statistics Spain - National Statistical Institute (INE), Spain
	Assessing the quality of national and regional accounts data in the context of ESA 2010 transmission programme - Eurostat, Luxembourg	Enhancing quality practices at the Brazilian Institute of Geography and Statistics - IBGE, Brazil	Metadata management systems, living spirals or flat line? - Statistics Portugal
	Public investment: Recording in EDP statistics & treatment under the SGP - Eurostat, Luxembourg	Quality management of statistical programs at the National Agricultural Statistics Service - U.S. Department of Agriculture, National Agricultural Statistics Service	Integrated metadata system in the Hungarian Central Statistical Office - Hungarian Central Statistical Office
		Assessing and improving quality in official statistics: the case of the French Label Committee - Insee, France	

Coffee Break

Plenary Session 2

Challenges of Quality Assurance for Big Data

Keynote Speech: Genoveva Ruzic (Statistical Office of the Republic of Slovenia)

	Fernando de Rojas	Columnas	
08:30-10:00	39 - SPECIAL SESSION: Quality of International Statistics. The Challenges for International Statistics at Global, National and Local Levels	25 - Methodology: Sampling/Non-Sampling Errors & Calibration	
	The role of international organisations in defining standards that follow the quality requirements and ensure comparability of data - Eurostat, Luxembourg	Uncertainties in the Swedish PPI and SPPI - Statistics Sweden	
	International Civil Aviation Organization (ICAO) quality system for the aviation data collected under the statistics programme. - ICAO, Inter. Civil Aviation Org., Canada	Standard error estimation – how to do it quickly, efficiently and correctly - Statistical Office of The Republic of Slovenia	
	Enhancing statistical capacities of OIC member countries to achieve sustainable development goals: The role of SESRIC - SESRIC, Turkey	Use of calibration in official statistics: Design versus model-based calibration - pros and cons - Statistical Office In Poznań, Poznań University of Economics And Business, Poland	
	Statistics governance and quality assurance - the experience of FAO. - FAO	Inference for statistics based on complete enumerations? - Statistics Austria	
	Quality assurance procedures for estimating purchasing power parities - World Bank, Washington D.C., USA		
10:00-10:15	Coffee Break		



	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	26 - Data Collection & Burden Respondent	27 - Big Data & Web Scraping	28 - Integrated Production & Business Process Model
	Using paradata to assess the quality of an online questionnaire - Statistics Austria	Big data in official statistics: Estimate job vacancies by using web scraping techniques - Federal Statistical Office (Destatis), Germany	Linkage of main components of GSBPM model through integrated statistical information system. - Statistical Office of The Slovak Republic
	Quality in the gathering of online data: Standardising online questionnaires, integration with administrative sources and development of bias control mechanisms (Eustat) - Basque Statistics Office (Eustat), Spain	Quality implications of the use of Big Data in tourism statistics: Three exploratory examples - National Statistical Institute (INE), Spain	Data warehouse for agricultural statistics as a tool for standardization of data processing and improving the quality of EAA calculation - State Statistical Office, the former Yugoslav Republic of Macedonia
	Response burden database and response burden indicators at the statistical office of the Republic of Slovenia - Statistical Office of The Republic of Slovenia	Quality control of web-scraped and transaction data (scanner data) - Statistics Austria	Monitoring quality in the production process model: Experiences from the INE of Spain - National Statistical Institute (INE), Spain
	Managing an integrated respondent communication: Statistics Portugal experience - Statistics Portugal	New data sources to indicate levels of active citizenship - Joint Research Centre (European Commission), Italy	Statistical metadata and the GSBPM - Central Statistics Office, Ireland
	Developments in measuring the burden placed on businesses responding to statistical surveys - Office for National Statistics, UK		Karat: The new integrated data transmission system of the HCSO - Hungarian Central Statistical Office
Coffee Break			

	Fernando de Rojas	Columnas	
10:15-11:30	40 - SPECIAL SESSION: Opportunities and Challenges for Quality of Official Statistics-The Future of the ESS Code of Practice	29 - Data Collection: Cross-Cutting Issues	
	Improving the comparability dimension of European statistics by minimizing unnecessary variation - Eurostat, Luxembourg	Using monetary incentives in face-to-face surveys: Are prepaid incentives more effective than promised incentives? - GESIS- Leibniz Institute for the Social Sciences, Germany	
	Big Data: From noise to evidence - European Statistical Governance Advisory Board, Luxembourg	Quality driven data collection. Towards a system for quality management based on dashboard information - Statistics Norway	
	Indicators. From evidence to decision - European Statistical Governance Advisory Board, Luxembourg	Mixed mode experiments in a panel survey – changes in response behaviour - Statistics Sweden	
		Modernisation of German household surveys: modularisation & mixed- mode – future challenges - Federal Statistical Office (Destatis), Germany	
11:30-12:45	<p align="center">Closing - Plenary Session 3 High Quality Statistics in Changing Landscape - Challenges ahead Chair: Eduardo Barredo (Eurostat) Panelists: Walter Radermacher (Eurostat), Konrad Pesendorfer (Statistics Austria), Alfredo Cristobal (INE), Martti Hetemäki (ESGAB), Juan Peñalosa (Bank of Spain)</p>		
12:45-13:15	Refreshments		



Day 3 - Friday 3 June Programme

	Antonio Palacios	Valle Inclán	Ramón Gómez de la Serna
	30 - Satisfying User's Needs: Dissemination	31 - Administrative Data: Topic Oriented	32 - Quality Management Systems 3
	Changing user needs raise demand for useful indicators - Statistics Sweden	Improving building permit administrative data set for short term analysis: The French case - Ministry of Ecology, France	Quality guidelines - Federal Statistical Office (Destatis), Germany
	Data visualisation - how to ensure understandable and sustainable statistical products - Federal Statistical Office (Destatis), Germany	Evaluating the potential for moving away from a traditional census - Office for National Statistics, UK	Managing survey quality through performance management: Building a performance management capability - U.S. Census Bureau
	Assessing and documenting quality of indicators: Established practices and outstanding issues - Eurostat, Luxembourg	Integrating geo-referenced data from different sources: Livestock surveys and administrative data - Istat, Italian National Statistical Institute	The OECD recommendation on good statistical practice - OECD Statistics Directorate
	Reviewing aspects of user orientated quality reporting - Office for National Statistics, UK	Administrative data and survey data on electricity use in Hungary - Hungarian Energy And Public Utility Regulatory Authority	Data quality assessment methods and tools in SSO - Macedonia - State Statistical Office, the former Yugoslav Republic of Macedonia
Closing - Plenary Session 3 High Quality Statistics in Changing Landscape - Challenges ahead Chair: Eduardo Barredo (Eurostat) Panelists: Walter Radermacher (Eurostat), Konrad Pesendorfer (Statistics Austria), Alfredo Cristobal (INE), Martti Hetemäki (ESGAB), Juan Peñalosa (Bank of Spain)			
Refreshments			

Content

Parallel Sessions

Session 1	Enterprise Architecture in Statistical Offices	24
Session 2	Quality Improvement Methods in Household Surveys	29
Session 3	Quality Challenges in Social Statistics: Preserving Privacy and other Issues	34
Session 4	Administrative Data: Cross-Cutting Issues	39
Session 5	Quality Management and Governance	43
Session 6	Peer Review: Learning from Countries' Experiences	48
Session 7	Quality Indicators	53
Session 8	Integrated Production and Quality	58
Session 9	Quality Management Systems 1	63
Session 10	Business Register	68
Session 11	Models & Early Estimates	73
Session 12	Enhancing Statistical Literacy	78
Session 13	Coordination of Statistical System	83
Session 14	Satisfying User's Needs: Communication	88
Session 15	Methodology: Linkage and Modelling	92
Session 16	Quality Assessment & Audits	97
Session 17	Quality Reporting	102
Session 18	Multi-Source Statistics	106
Session 19	Human Resources Development: A Quality Culture	111
Session 20	Big Data Oriented Systems	116
Session 21	Administrative Data Systems	121
Session 22	Macroeconomics	126
Session 23	Quality Management Systems 2	130
Session 24	Metadata Systems	135
Session 25	Methodology: Sampling/Non-Sampling Errors & Calibration	139
Session 26	Data Collection & Burden Respondent	143
Session 27	Big Data & Web Scraping	148
Session 28	Integrated Production & Business Process Model	152
Session 29	Data Collection: Cross-Cutting Issues	157
Session 30	Satisfying User's Needs: Dissemination	161
Session 31	Administrative Data: Topic Oriented	165
Session 32	Quality Management Systems 3	169

Special Sessions

Special Session 33	Peer Reviews: A Tool to Enhance Trust in European Statistics?	173
Special Session 34	Big Data and Official Statistics: Challenges and Opportunities	176
Special Session 35	The Statistics Code of Practice for the ENP South Countries	180
Special Session 36	Competence Management in Statistics	186
Special Session 37	Governance and Coordination of the National Statistical Systems in the Enlargement and ENP-East Countries	190
Special Session 38	Synergies for Europe's Research Infrastructures in the Social Sciences and Official Statistics	194
Special Session 39	Quality of International Statistics. The Challenges for International Statistics at Global, National and Local Levels	199
Special Session 40	Opportunities and Challenges for Quality of Official Statistics-The Future of the ESS Code of Practice	206

Speed Talk Sessions

Speed Talk Session 1	Dissemination & Addressing Users Needs	209
Speed Talk Session 2	Methodology	215
Speed Talk Session 3	Modernisation & Big Data	221
Speed Talk Session 4	Demographic & Social Statistics	229

Floor Plan Círculo de Bellas Artes	237
---	-----

Chair: Jean-Marc Museux

ENTERPRISE ARCHITECTURE IN THE HUNGARIAN CENTRAL STATISTICAL OFFICE

Csaba Ábry

Hungarian Central Statistical Office, Hungary

In the Hungarian Central Statistical Office (HCSO) the issue of standardisation and integrity of a statistical system has always been taken into consideration at elaborating its statistical system and also when improvement actions took place concerning the activity of the HCSO. As a result the elements of an Enterprise Architecture are elaborated adequately, but there is no uniform structure yet.

Therefore the HCSO decided to elaborate a structure for the Enterprise Architecture based on the international standards in this field. These international standards are as follows: Common Statistical Production Architecture (CSPA), Generic Activity Model for Statistical Organisations (GAMSO) Generic Statistical Business Process Model (GSBPM) elaborated by the UNECE and the European Statistical System Enterprise Architecture Reference Framework (ESS EARF) elaborated by the Task Force on Enterprise Architecture. The HCSO uses these standards in order to ensure comparability and better understanding at international level.

Currently the HCSO is at the conceptual level at elaborating the Enterprise Architecture model. The aim of this action - beside the above mentioned uniform structure- is to give a base for the future work of reviewing the processes and the products of the statistical activity of the HCSO in order to modernise those where it is needed. Quality of statistical production can be assured when the structure of the organisational work and roles are defined.

Keywords: enterprise architecture, standardisation.



GOVERNMENTAL AND STATISTICAL ENTERPRISE ARCHITECTURE FRAMEWORKS AS TOOLS FOR MODERNIZING THE NATIONAL STATISTICAL SYSTEM

Timo Koskimäki, Ville Koskinen
Statistics Finland, Finland

Enterprise Architecture (EA) frameworks such as TOGAF or the Zachmann provide an abstraction of an enterprise by modeling it using architectural perspectives - business, data, applications and technologies – and different levels of detail– abstract, logical, physical. The value of this approach is that it can support a holistic and top down approach to the development of an enterprise.

The statistical community has adopted ideas from Enterprise Architecture frameworks. In some initiatives, such as the Common Statistical Production Architecture (CSPA) and the ESS Enterprise Architecture Framework, there is a strong vision of a common enterprise architecture for official statistics.

On the other hand, national architecture standards governing the architectures of all state agencies are being set up to improve public sector efficiency and service quality. In addition, National Statistical Institutes may have internal needs for standardizing their own activities in areas that are not yet fully covered either by international or national standard frameworks.

Thus, a NSI has to consider three slightly different drivers and frameworks, one stemming from the international statistical community, one stemming from the national governmental initiatives and one stemming from the internal needs of a NSI. There is a need to balance these three drivers by defining the scope of the different architectures.

In the paper we attempt to answer to the question of which areas of a NSI's activities should be addressed by national public sector architectures, which areas should be addressed by the statistical community's general enterprise architecture and the gaps of the two frameworks from a NSI point of view. We also discuss which areas of the international statistical EA frameworks would need more development work to better serve the NSI's and international statistical organizations.

Keywords: Enterprise Architecture, National, International.

RE-ENGINEERING AND RE-DESIGNING STATISTICAL PRODUCTION, MODERNISATION; INTEGRATED PRODUCTION METHODS

Mihaela Vacarasu, Daniel Suranyi
Eurostat, European Commission, Luxembourg

Since several years the official statistical community is working on the modernisation of statistical production. The High Level Group (HLG) is promoting several initiatives and international standards to make statistical production more streamlined and eventually produce higher quality statistics faster and more efficiently.

Starting from domain specific implementations for National Accounts within Eurostat, a new architecture approach for statistical production based on standards and statistical services puts this idea into practice. By implementing standards such as CSPA, GSBPM or SDMX in a complimentary way, the project aims at designing and implementing a National Accounts Production System architecture based on a service oriented approach (NAPS-S).

The paper will show how process and product quality can be improved significantly by using existing standards to design new production architecture for statistics and how this design can be put in practice. The shift from monolithic, domain specific implementations to an open, modular architecture is achieved by small iterations that deliver immediate business value. The business requirements are derived based on a modelled business process. Existing services can be mapped to it and integrated through a process manager. Functionalities not yet covered can be developed into services or kept within existing systems, depending on their business value.

The National Accounts are an early adopter of such architecture; the architecture and its implementation can be further scaled to other statistical production lines.

A modelled statistical production process will inherently lead to more transparency in how statistics are being produced for both the statisticians working within the production process, but also vis-à-vis the statistical community. A transparent process increases the quality of its output through a common understanding of the production steps, a better control and monitoring and an enhanced flexibility for changes and improvements.



STATISTICS CATALONIA: TOWARDS TOTAL INTEGRATION OF STATISTICS PRODUCTION

F. Udina, C. Rovira, J. Sort, M. Perelló

Statistical Institute of Catalonia (Idescat), Barcelona, Spain

The mission of official statistics is to provide good quality statistics about a society to allow for improvement of public policies and the general good as a whole. A modern approach requires the integration of information from many sources, total quality of production processes and use of multiple channels of dissemination to reach a rich variety of users always requiring strict control of data security and confidentiality.

Over the last three years, Idescat has conducted a two track process of innovation to achieve these goals: on the business track, the focus has been the implementation of a production metadata system of documentation based on GSBPM (Qualitas project) that is guiding the modernisation of statistical processes; on the technological track, the focus has been the implementation of a new information management scheme (Cerdà Platform) based on four conceptual stages: Data capture – Integration – Stat product generation – Data access.

The Cerdà Platform is entering the final phases of implementation and is allowing the integration of any incoming data (validated through the first stage) by coding information on people, economic entities and territory using the three core statistical registers. Thus, every piece of integrated information is available to statisticians in all areas of the Institute. Through a process of anonymization the combined information is made available to different levels of users.

The Qualitas project provides methodological guidance, quality control and metadata production to implement the desired statistical products. Starting from some proofs of concept, all current statistical production is being moved to the new model.

Among other things, providing data from several origins (admin data, survey data, big data, etc.) to researchers in the social sciences is one of our main goals. Taking care of data security and confidentiality issues is crucial here, as is developing safe protocols to ensure that good researchers have good quality data to develop safe projects.

Keywords: GSBPM; integrated statistical production processes; modernizing the production and dissemination of official statistics; statistical and admin data for research.



REDESIGN OF THE STATISTICAL INFORMATION SYSTEM: CZECH EXPERIENCE

Marek Rojíček

Czech Statistical Office, Prague, Czech Republic

In the last years, preparation and implementation of the integration of software tools for collection, processing and dissemination of statistical data has been realised. The main tasks of this project were decrease of the administrative burden laid down on respondents together with the increased quality and accessibility of statistical data for every user of the statistical information system (SIS).

The specific targets of this project, realised within the general framework of e-Government implementation and financed mostly by EU funds, were the integration of mathematical-statistical methodologies, procedures, tools and technologies used for evaluation of users' requirements, preparation and processing of statistical tasks, analysing and publication of statistical information. Within this project the development of data warehouse has been completed including the creation of system of secure and protected access to data, as well as the unification of administrative data sources and introduction of mutual electronic contact with respondents.

The main part of the project has been realised in the years 2013 – 2014 and in 2015 the project was transformed into routine statistical production process. After implementation of the new SIS the need for higher cooperation of all statistical as well as IT departments significantly increased. Although most of the roles were anchored into the line organisation structure, it was necessary to keep partly project type of the management of the whole process due to its complexity. We expect that the initial investment in the form of financial resources and human capacities should be returned in the following years by decreased administrative burden of respondents and increased quality and accessibility of statistical indicators for users.

Keywords: metadata, digitalization, integration, data collection.



Chair: Florentina Álvarez

REDEVELOPING THE HOUSEHOLD BUDGET SURVEY

Sverre Amdam, Kristin Egge-Hoveid
Statistics Norway

Statistics Norway is in the process of redeveloping its Household Budget Survey. The paper will present the past, present and future, focusing on quality challenges.

The past design of the survey was based on paper diary reporting, and an initial CAPI interview. Interview costs, manual coding of hand written diaries and enclosed receipts led to high survey cost. Data quality was encumbered by response-rate below 50 percent and under-reporting. This challenge is shared by several statistical offices. Manual editing of data also impacted on data quality and thus the need to modernize the survey.

The present is a digitalized survey design to be implemented in 2017. The design is an initial telephone interview, a web diary with automatic coding into COICOP-categories (Classification of Individual Consumption According to Purpose), and a web questionnaire. Paper receipts will be scanned and automatically coded. The processing of data will be modernized with the use of statistical methods for editing. The development digitalized instruments for data collection and more standardized systems for editing will have an impact on data quality in terms of more effective data collection, more precise coding and better editing. However –more digitalized data collection may lead to lower response rate and more missing or inadequate completion. Thus it places great demands on the technical solutions to be developed. Usability tests and experiments will be conducted in spring 2016. Key results will be presented in the paper.

The future is a design that makes use of alternative sources of consumer data from individuals/households, such as transaction data from banks, retailers and other financial institutions. These sources may be used both in addition to, and partly instead of a survey. The utilization of these data sources can reduce response burden, reduce costs and meet the quality challenge of unit- and item- non-response in a new way. This will set new standards for sampling, and volume of data. However, this opens up new challenges such as privacy concerns, technical issues, and dependency on collaboration with big market actors such as banks and retailers.

Keywords: Household Budget Survey, Web-Survey, Big-Data.

FROM OUTPUT TO INPUT/OUTPUT HARMONISATION IN THE EU LFS

A. Zgierska

Central Statistical Office, Warsaw, Poland

One of the main goals of the regulation of data collection within the European Community is to receive comparable survey results among countries covered and coherent data for aggregates e.g. for the European Union. In the beginning the European Union Labour Force Survey (EU LFS) was concentrated only on output harmonisation of the results elaborated by countries and transmitted to Eurostat according to a common code of data record. Then - taking into account the importance of comparable indicators for monitoring and progress assessment as an effect of the implementation of the Employment Guidelines - a common definition of unemployment in all Member States, combined with a greater harmonisation of LFS questionnaires, was indicated. At present there have been elaborated model questionnaires both for employment/unemployment data collection and hours actually worked as the results of two Eurostat Task Forces. Moreover a model questionnaire has been prepared by the specific Task Force for an ad hoc module (AHM) realised next to the LFS every year. These model questionnaires for LFS AHMs are not obligatory, but they are strongly recommended to be used by Member States.]

After illustrating the evolution of the approach to the EU LFS harmonisation and presentation of the currently discussed elements of the future Integrated European Social Statistics, the paper focuses also on presentation of some advantages and disadvantages of input/output harmonisation.

Keywords: Statistical input harmonisation, labour market statistics, Labour Force Survey.



A STANDARD TOOL FOR METADATA QUALITY ASSESSMENT OF EXTERNAL SOURCES IN THE STATISTICAL PRODUCTION (QMETATOOL)

J. Orche Galindo
National Statistical Institute (INE), Spain

Knowledge and documentation of the source is vital to help us to use external data in the statistical production. i.e.: how the data are collected, why they are gathered, how they are processed, concepts and definitions used, etc...

A tool for a standard quality assessment of these questions gave standard clues to a statistical office about strengths, weaknesses and what could be improved in the use of external sources for statistical purposes.

A practical tool (QmetaTool) has been developed for these purposes and gathers the basic aspects to take into account in any use of external micro-data in the statistical production. The issues are based in the state of clarity in the definition of units, time reference, geographical scope and variables (with their classifications and codifications). Clarity in the communications of changes in these definitions are evaluated too. Furthermore, the metadata assessment is focused on clarity in the description of checks and modifications in the database made by the owner. The last aspect is the state of knowledge in the existence of unique keys or a combination of variables that can be used to identify the populations units.

Finally, the tool has been tested in practice with metadata of administrative registers used in real statistical operations like the Spanish Labour Force Survey. This standard tool offers the possibility to be used in any statistical operation fully or partial based on external sources. The time spent in the implementation of the tool is minimized and there is a graphical output for the overall situation of the metadata in the external data source. Just one assessment is needed by each source (no matter how many uses are made) and it can be implemented from any origin (public or private sector) and even in the future use of big data for statistical purposes.

Keywords: tool, metadata, assessment.

THE USE OF REGRESSION MODELS IN LABOUR MARKET FLOW STATISTICS

H. Kiiver, F. Espelage
Eurostat, Luxembourg

In October 2015, Eurostat published for the first time labour market flow statistics using individual country Labour Force Survey (LFS) data, covering the transitions between the three ILO labour market statuses employment, unemployment and inactivity for the population aged 15-74, broken down by sex. Further work, driven by high demand for additional breakdowns of this data, revealed that even just adding another breakdown by age (15-24, 25-54, 55-74) or education (low, intermediate, high) will result in about a third of data points which cannot be published due to low sample size. Analysts and policy makers are however interested in much more rigorously defined subgroups to understand for example the labour market transitions of long-term unemployed, of individuals close to retirement, or those young individuals neither in employment nor in any education nor training. Given the current sample sizes and rotational patterns, these detailed breakdowns are impossible for many countries.

In these cases, simple regression techniques can be used to extract the desired information from the data, avoiding to a large extent the problems related to small sample sizes.

While the introduction of regressions solves thus an important problem, a whole new set of issues arises. A number of questions has to be discussed before regression results can be published, thereby complementing the already disseminated flow statistics.

This paper discusses these questions in detail, drawing also on existing experience by national statistical institutes, and shows the solutions found for the case of labour market flow statistics. Using the transitions from unemployment to employment as an example, the main results and their implications and interpretations for policy makers are examined; furthermore, the paper proposes a format in which regression based statistics can be presented along the existing statistics without neglecting either the cross-country comparability or the comparability over time.

Keywords: the use of modelling in official statistics.



MOVING TOWARDS WEB DATA COLLECTION IN HOUSEHOLD SURVEYS

P. Taskinen, L. Hulkko
Statistics Finland, Helsinki, Finland

The computer assisted data collection was a step forward years ago. Today, the data collection still relies strongly on the interviewer in the household surveys. The web based digital collection as a self-administered mode is an option for an interviewer administered survey. To be realistic, CATI mode will remain in household surveys but a partial transition to the electronic mode will be welcomed as a part of the digital or electronic development of the whole society. A proper testing is needed before the production environment is ready for the mixed mode collection to meet the standards of the official statistics.

The EU Labour Force Survey (EU-LFS) is one of the most important social and economic statistics measuring the labour market status of the population. The EU-LFS is also a rather complicated survey. E.g. the data production is continuous; the EU-LFS is a panel survey where the same person answers multiple times; the unemployment rate and other vital statistics are published every month, thus the timetable is a tight one.

Statistics Finland intends to introduce a new data collection package concerning the EU-LFS and other household based surveys in a digital environment. In order to ensure the quality of the statistical output, Statistics Finland is carrying out mixed mode data collection pilot projects.

In the presentation, the testing procedure of the EU-LFS mixed mode collection is enlightened. Further, following themes are discussed: does the mixing of modes affects the responding? How to use text messages and emails as communication tools? What is the best way to mix the modes in a complex survey like the EU-LFS?

Keywords: Labour Force Survey, digital environment, mixed mode collection.



Chair: Tasos Christofides

AVAILABLE METHODS FOR PRIVACY PRESERVING RECORD LINKAGE ON CENSUS SCALE DATA

R. Schnell¹, C. Borgs²

1 City University London, United Kingdom

2 University Duisburg-Essen, Germany

The trend towards register based censuses is obvious in Europe. For these kind of censuses, record linkage across administrative data bases is essential. Without unique personal identifiers (PIDs) record linkage is laborious, but routine. If the jurisdiction requires the use of encrypted identifiers for record-linkage, the task is challenging. The set of problems associated with this task has created an academic subfield called Privacy Preserving Record Linkage (PPRL). PPRL requires the use of advanced encryption techniques. Most often, standard encryptions such as MD5 or SHA-1 are applied on phonetic codes of identifiers. In real world settings, such codes miss many true links due to excessive errors in identifiers. Furthermore, missing identifiers usually require the use of hierarchical matching schemes. Therefore, new techniques such as embedding or Bloom Filters have been suggested. However, most encryption techniques can be attacked with cryptographic methods. All applications of PPRL have to demonstrate that their encryptions can resist at least a reasonable amount of cryptographic efforts. Based on recent cryptographic research, recommendations on the practical applications for PPRL will be given. Furthermore, in order to handle census scale datasets, special techniques for finding nearest approximate neighbours of encrypted records (blocking techniques) have to be used. The state of the art on blocking within PPRL will be described briefly.

Keywords: Administrative data, big data, record linkage.



DEVELOPING LONGITUDINAL STATISTICS ON RECIPIENTS OF WELFARE BENEFITS AND THEIR LABOUR MARKET POSITION

Tor Morten Normann, Unni Beate Grebstad
Statistics Norway

A characteristic of the Norwegian welfare state is comprehensive welfare benefits aimed at providing subsistence level, as well as improving labour market participation. This paper will describe a project where Statistics Norway has developed a new and more dynamic statistics on recipients of some central welfare benefits by measuring their position in the labour market and welfare system over a period of time after benefit reception, and also the period before for some benefits. Contrary to ordinary cross sectional statistics, this new statistics has a longitudinal approach, following the recipients over a period of five years. The paper will describe some of the main challenges faced during the process of developing this statistics, focusing on data availability, definition of populations and the construction of indicators. In the Norwegian case, the access to comprehensive register data is an obvious advantage, and the project extracted data from Statistic Norway's events database, containing information on the population's benefit reception and labour market participation as events. The paper describes how this made it possible to measure the duration and sequence of benefits over a period of time, including labour market participation. The definition and construction of the longitudinal statistical populations, unique for each of the welfare benefit included in this new statistics, was a central element in the project, and will be described in the paper. The criteria for inclusion in the population were reception of benefits over a period of time, and eligibility for the labour market in an output period. The benefits in question are work assessment allowance, social assistance, disability pensions and sickness allowance. Indicators must take into consideration the variation over the measurement period. At the same time, when including both labour market position and benefit reception, the challenge of simultaneous events had to be solved. The paper will present the different indicators, now published by Statistics Norway.

Keywords: longitudinal statistical populations, welfare benefits, labour market position.

DO RESPONDENTS ANSWER DIFFERENTLY IN WEB SURVEY THAN IN FACE-TO-FACE INTERVIEW: FIELD WORK EXPERIMENT FROM THE EUROPEAN HEALTH INTERVIEW SURVEY (EHIS)

M. Kralj, M. Zaletel, D. Lavtar, T. Zupanič
National Institute of Public Health, Ljubljana, Slovenia

Expansion of web surveys at the turn of 20th century brought new opportunities and new challenges at the same time. Web surveys enable lower costs and time savings for national data providers, and offer respondents a choice to complete the questionnaire at the time and with pace they wish to. Mixed-mode designs have become a leading way of collecting survey data. Many organisations have experienced that the results from web surveys differ from other modes. There are many reasons for this, maybe the most important is that many respondents provide different answers when completing the survey on their own via web. How different are actually web surveys compared to other modes of data collection, and even more importantly, do they provide the same results? In the paper, reasons for different results between two modes (web survey and face-to-face interview) will be examined, e.g. potential misuse or lack of understanding of the question may lead in respondents giving inaccurate information, while in a face-to-face interview socially desirable answers may be obtained due to the interviewer's presence. The special attention will be given to sensitive data, e.g. health data, social non-desirable behaviours, etc. The case study will be presented: during the field work of European Health Interview Survey (2014) National Institute of Public Health carried out an experiment with re-interviews. The aim was to test whether the respondents answer differently in web survey than they do in a face-to-face interview and to measure the magnitude of differences for different types of variables.

Keywords: mixed-mode design, experiment, re-interview.



THE PRIVACY PROTECTING ASPECT OF INDIRECT QUESTIONING DESIGNS

Andreas Quatember
Johannes Kepler University JKU, Linz, Austria

Talking about indirect questioning designs such as randomized response (RR) techniques, the aspect of privacy protection is the motivation behind. Asking (sensitive) questions indirectly and not directly should have a positive effect on the answering behaviour of the respondents (response/nonresponse, truthful/untruthful) by reducing the individual's fear of an embarrassing "outing". Nevertheless, when discussing, for instance, a certain RR design, this most important aspect is rarely considered.

On the one hand, the decision, which of the possible RR questioning designs shall be used for a certain variable, has to take account of the type of sensitivity of the variable. Is the variable sensitive as a whole or are just certain variable values sensitive? The chosen questioning method has to match with the given variable in this respect. On the other hand, a performance comparison between different RR designs only makes sense when it is done under the same level of the privacy protection that is offered by them. Otherwise, the direct questioning would always be the best strategy. This aspect is not considered sufficiently in the relevant literature on RR techniques.

In the talk, this aspect of indirect questioning designs is presented exemplarily for certain RR techniques, stressing the relationship between data confidentiality and data quality. Furthermore, other aspects of the RR techniques have also to be considered when talking about the privacy protection as perceived by the respondents. Such are the type of the randomization device or the practical realisation of the RR strategy in various stages.

Keywords: Privacy Protection, Questioning Designs, Nonresponse.

A LATENT CLASS MODEL TO ESTIMATE LABOUR COST FROM MULTI-SOURCE DATA

U. Guarnera, S. Pacini, R. Varriale
National Statistical Institute, Italy

In recent years, statistical analysis based on different data-sources has become an active area of research in both theoretical and applied statistics. In particular, due to the increasing availability of administrative data, problems concerning the use of multiple sources for estimation purposes have been receiving an increasing attention in Official Statistics. National Statistical Institutes (NSIs) frequently try to combine data from available sources in order to build “statistical” archives to be used in different phases of the statistical production process.

In this work we describe a procedure to predict the labour cost of enterprises when measurements are available from two different administrative data sources. Contrary to the situations where the information coming from all the sources can be considered a (possibly erroneous) measure of the amount of interest (see, for example, [1], [2], [3], [4]), in this context only one source is considered to be a direct measure of the response variable, while the information coming from the other source is treated as auxiliary information. In particular, we propose to use a latent class model, where latent classes correspond to different error patterns. The proposed model produces individual predictions (estimates) of the response variable, obtained taking expectations of the true data distribution conditional on the observed data. These predictions can be used for different purposes. First, they can be directly used as “smoothed” estimates of the response variable in presence of errors. Alternatively, they can be used for editing activities. This approach can also be used to assess the quality of the data sources in terms of the model parameters.

Keywords: Multi-source statistics, Administrative data, Latent class models.

[1] Bakker, B.F.M., Daas, P.J.H. (2012) Methodological Challenges of Register-Based Research. *Statistica Neerlandica*. 66(1), 8-17

[2] Guarnera U., Varriale R. (in press). Estimation from Contaminated Multi-Source Data Based on Latent Class Models. *Statistical Journal of the IAOS*.

[3] Pavlopoulos D., Vermunt J.K. (2015). Measuring temporary employment. Do survey or register data tell the truth? *Survey Methodology*. Vol. 41, No. 1, 197-214.

[4] Scholtus S., Bakker B.F.M. (2013). Estimating the validity of administrative and survey variables through structural equation modeling. A simulation study on robustness. Discussion paper (201302), Statistics Netherlands, The Hague/Heerlen [Internet].



Chair: Antonio Argüeso

THE ASSURANCE OF ADMINISTRATIVE DATA - A PROPORTIONATE APPROACH

PJ Babb

UK Statistics Authority, London, UK

Administrative data collections provide an attractive means of obtaining generally cheap and comprehensive data for producing official statistics. But they can present substantial challenges, coming from systems with unique definitions and needs. Statistical producers may have little involvement in the data collection.

The UK Statistics Authority has developed a proportionate approach to the quality assurance of administrative data, with three levels of assurance: basic, enhanced and comprehensive. The level selected is dependent on the risk of data quality concerns and degree of public good represented by the statistics.

The approach identifies four practice areas that reflect the need for the assurance of administrative data to extend beyond the checks made by statistical producers on the data they receive. Producers should also have a good understanding of how the data are compiled and why; have established effective communication links with data supply partners; and understand their partners' data quality processes and standards.

The 'QA toolkit' [] helps statistical producers identify appropriate areas of practice to assure the data for their particular statistics. It describes the regulatory standard used in the Assessment of official statistics by the Statistics Authority. This presentation will outline the toolkit and illustrate ways in which statistical producers in the UK have used it to enhance their quality assurance of administrative data and information for users about the associated strengths and limitations.

Keywords: Administrative data, data quality, regulatory standard.

<http://www.statisticsauthority.gov.uk/assessment/monitoring/administrative-data-and-official-statistics/index.html>

WHEN IS ADMINISTRATIVE DATA ENOUGH TO REPLACE STATISTICAL INFORMATION: A BASED ON CENSUS COMPARISON QUALITY INDICATOR

A. Delgado, S. Lagarto, P. Paulino, J. Capelo
Statistics Portugal, Lisbon, Portugal

Statistics Portugal is considering the use of administrative data in 2021 Censuses. To face this challenge, the quality of the available administrative data sets is measured comparing administrative data with Census information. The purpose is to evaluate the risks or impact of replacing part of statistical information with information obtained by administrative sources. Record linkage methods were applied between the 2011 Census results and each administrative dataset. Eighteen variables from seven administrative sources (from Social Security to Education) were selected based on the administrative source contribution to potential replacement of Census collected information. For each matched record pair, contents within variables are compared, producing an equality rate estimate (errors or lack of information are not considered at this time).

The results show very high equality rates between contents from each linked pair of records, to both geographic and demographic variables (municipality of residence, sex, date of birth, legal marital status, country of birth, country of citizenship). When comparing socioeconomic variables, results are less homogeneous: identical contents have less uniform distribution across sources and between Census and administrative sources. However, some data obtained by certain sources, related with labour force characteristics, also got high correspondence rates for compared record pairs.

Additionally, considering that some statistics might be obtain by other sources, some Census microdata (regarding economic and educational characteristics of the population) were compared with data from national Labour Force Survey. These results converge to the general comparison exercise results.

Finally, the results of the Post Enumeration Survey of 2011 Census were used to verify the reliability and consistency of the comparison results.

Keywords: 2021 Portugal Population and housing Census, administrative data, Census microdata, linked data.



EVALUATING THE QUALITY OF ADMINISTRATIVE DATA AS INPUT FOR OFFICIAL STATISTICS

Peter Tibert Stoltze

Head of Division, Statistics Denmark, Copenhagen, Denmark.

Several projects have provided substantial contributions on how to evaluate the quality of input data. However, less effort has been devoted to the question of how to integrate the results of such a quality assessment in a statistical production system.

The use of administrative sources for the production of official statistics varies a lot between statistical domains, e.g. direct tabulation, use in combination with a survey, use for building a statistical register etc. We will present the main challenges in terms of quality measurement for each type of use of the data.

Within this project we have first collected and reviewed existing methods for assessment of input quality. The most promising methods have been tested on actual data in several statistical areas and with different types of administrative sources. Based on the testing we have recommended which approaches are more suitable for the different uses of administrative sources. The results will be made accessible for a wider audience through a commented repository on the CROS portal.

Finally, a consolidated checklist for input quality regarding both dimensions and indicators is constructed. It might be the case that we cannot create a single method fit for all purposes, in which case variations will have to be created for specific purposes. The consolidated checklist will be tested with at least three different administrative data sources.

Note: The proposed talk will present the results from the first work package within the ESSnet Quality of multisource statistics. The ESSnet is organised within ESS.VIP.ADMIN and sees participation from Denmark, Norway, Netherlands, Hungary, Austria, Ireland, Lithuania, and Italy. This first work package will start in January and end in June 2016, just in time for presentation at Q2016. Hence, the content of this abstract will evolve into more specific results in ample time before the conference.

Keywords: Administrative data, multisource statistics, quality.

MAPPING POTENTIAL ADMINISTRATIVE DATA FOR STATISTICS PURPOSES - STATISTICS PORTUGAL APPROACH WITHIN PUBLIC ADMINISTRATION

Magda Ribeiro; Maria João Zilhão; Paulo Saraiva
Statistics Portugal, Lisbon, Portugal

The use of administrative data for statistical purposes has assumed particular importance in the context of the production of official statistics, without prejudice to its quality, as a way to reduce costs and decrease the burden on respondents. Is part of a course which Statistics Portugal (SP) has been following, and for which becomes indispensable the cooperation of the different public entities which owns such data.

This paper will describe the approach taken by SP to: (i) identify some of the constraints associated with obtaining data through administrative sources, even though the legal framework in which SP operates allows it; (ii) summarize the results of a mapping exercise of potential administrative data sources that can be used for statistical purposes, regardless of its usage (eg. appropriation of administrative data for substituting direct data collection; and/or for data validation purposes, already collected by other means); and, (iii) report on the administrative sources that are already being used by INE, focusing more on economic statistics, stressing on the success factors and on the constraints of its usage in relation to the cooperation amongst institutions and the legal frameworks.

A strong focus will be put on the mapping exercise exploring the approach and the means developed for it; the relation with this mapping exercise with all activities at SP production classification activities system; and finally the results of its potential usage classified by the short or long term perspective of its real appropriation, taking into account the diligences that have to be made in order to assure the adequate quality criteria for statistics proposes.

Keywords: Administrative data; Potential administrative Data usage; Quality and administrative data.



Chair: Eduardo Barredo

INTER-ORGANISATIONAL TRUST IN THE CONTEXT OF MICRODATA EXCHANGE IN THE ESS

M. Kotzeva, S. Bielecki

Eurostat, European Commission, Luxembourg

The exchange of identifiable microdata is one of the key elements of the ongoing modernisation of European statistics as set out in the “ESS Vision 2020”. Microdata exchange is seen as an opportunity for the European Statistical System (ESS) to reap the benefits of enhanced quality, increased efficiency and reduced burden. In particular, it should allow properly addressing the measurement of cross-border phenomena, thus increasing the accuracy and relevance of European statistics on globalisation.

Identifiable microdata enjoy strong protection in European and national law. At the same time European statistics law enables National Statistical Institutes to exchange microdata for statistical purposes provided that data confidentiality is preserved. However, trust between different actors of the statistical data life cycle is a key precondition for the exchange of identifiable microdata; at the same time, loss of trust is one of the main risks of such cooperation. While the risks entailed in the exchange of identifiable microdata must not be neglected, they can and should be managed, i.e. analysed and mitigated.

The paper analyses the components, functions and factors of trust within an inter-organisational network and discusses them in the specific context of microdata exchange in the ESS. The authors draw elements from various concepts and models of trust across the spectrum of management research. On this basis, enablers and safeguards of trust within the ESS are considered and recommendations for possible actions and for further research formulated.

Keywords: trust, inter-organisational networks, ESS Vision 2020, microdata exchange.



THE COMMITMENT ON CONFIDENCE IN EUROPEAN STATISTICS AS A MEANS FOR ENHANCING QUALITY OF OFFICIAL STATISTICS AND EFFICIENCY OF THE NATIONAL STATISTICAL SYSTEM: EXPERIENCE OF THE CZECH REPUBLIC

M. Bohatá¹, I. Ritschelová²

1 Czech Statistical Office and CERGE-Economics Institute, Prague

2 Czech Republic Czech Statistical Office, Prague, Czech Republic

The aim of the up-to-date official statistical service - to provide a sophisticated, high quality and efficient information service to a broad variety of users - requires all the assets and resources available in the NSS and not used or underused be deployed. The Commitment on Confidence in European Statistics (CoC) presented in the amended Regulation on European Statistics No 223/2009 can be seen not only as a means fostering the implementation of the European Statistics Code of Practice within the national statistical system but simultaneously also as an instrument paving the way to a whole-off production of national statistics under the fully-fledged coordination and leadership of the NSI.

The paper will analyse various potential benefits of the CoC. Besides the institutional improvements embodied in the broad concept of quality, it will focus mainly on efficiency of the whole national production system. It will also present experience with the development and adoption of the CoC which has turned to be a complicated political process faced with many obstacles. The major one has been linked to the explanation why there is a need for such an instrument if official statistics enjoys a high reputation, and how to avoid a risk of undermining this reputation. Lessons from the Czech experience will be offered in conclusions.

Keywords: quality, efficiency, national statistical system.



GOVERNANCE STRUCTURE OF TURKISH STATISTICAL SYSTEM WITHIN THE CONTEXT OF COORDINATION MECHANISMS

G. Gedik, T. Altun

Turkish Statistical Institute, Ankara, Turkey

Turkish Statistical Institute will present the governance structure of Turkish Statistical System in general and how Turkish Statistical Institute coordinates the System through robust tools and mechanisms. Turkish Statistical System is a multi-partner system with Turkish Statistical Institute as being the main coordinator. In addition to Turkish Statistical Institute, other institutions are intensively involved in the production of official statistics, primarily the Central Bank of the Republic of Turkey, the Ministry of Finance, the Undersecretariat of Treasury, and the Ministry of Food, Agriculture and Livestock.

This paper will also cover Statistics Law of Turkey, adopted in the year 2005, emphasising the sound coordination role of Turkish Statistical Institute in the System by introducing multi-annual Official Statistics Programme and the Statistical Council as well as regulating other coordination tools such as classifications. Additionally, practical tools will be elaborated, as functioning of 50 working groups covering all representatives of Turkish Statistical System, the preparation of a national data release calendar, use of international data transmission system, and protocols and memoranda of understanding signed with the stakeholders and meeting with producers and users of official statistics. Cooperation and communication studies of Turkish Statistical Institute with stakeholders in the context of “public-public” and “public-private” basis will also be explained. The paper will touch upon endeavours of Turkish Statistical Institute regarding the potential use of administrative registers of other national authorities for statistical purposes.

This paper will finally review the main findings of the light peer review and peer review studies carried out for Turkish Statistical System in 2011 and 2015 respectively, taking into consideration the recommendations and improvements actions regarding the coordination structure of Turkish Statistical System and the role of Turkish Statistical Institute within.

Keywords: National Statistical Systems, Coordination, Governance, Turkey.



REFORMING INFORMATION QUALITY ASSURANCE AT THE NATIONAL INSTITUTE OF STATISTICS AND GEOGRAPHY (INEGI) IN MEXICO

G. M. Rubio Soto, N. Torroja
INEGI, Mexico City, Mexico

INEGI has introduced new legal, institutional and technical measures for strengthening information quality assurance over the last year. The approval of the Statistical and Geographical Information Quality Assurance Norm by the Board of Directors sparked a number of actions aimed at institutionalizing quality assurance. First, the Norm entrusts the Quality Assurance Committee with the task of coordinating quality assurance at INEGI. The Committee is integrated by the managing directors and led by INEGI President. Second, the Committee developed a quality assurance framework, including an Institutional Quality Policy. Following international experience, the Policy adopts a set of principles for guiding the statistical and geographical activities at the institutional, process and output level. Each principle is supported by good practices focusing on producing relevant, timely, accessible, and reliable information in a cost-effective way, through sound methodologies, adequate process implementation and an enabling environment fostering professional independence, objectivity and transparency.

Third, a quality assurance self-assessment was piloted to identify areas of improvement. This exercise applied the UN National Quality Assurance Framework (NQAF) at the agency and unit level. At the agency level, managing directors participated in a qualitative assessment. This was complemented with the application of an assessment matrix at the program/project level and two questionnaires for the National Statistical System coordination and information dissemination. Finally, the Committee developed an Annual Quality Assurance Plan based on the self-assessment. The Plan sets out the strategies and the institutional and domain-specific actions for achieving INEGI quality assurance medium-term goals, as well as the challenges ahead.

Keywords: Quality assurance, information quality, institutional reform.



COORDINATING MIGRATION RELATED STATISTICS AND ANALYSES AT STATISTICS NORWAY - BRIDGING USERS AND PRODUCERS

Silje Vatne Pettersen
Statistics Norway, Oslo, Norway

Migration related statistics and analyses are in high demand by government agencies, the media and the public at large. Migration flows, stocks and integration measures such as participation in the labor force and in the educational system are continuously monitored. To meet the demand for relevant, updated and easily accessible information, Statistics Norway has a Coordinator of Migration Related Statistics and Analyses. The position is jointly financed with the Norwegian Ministry of Children, Equality and Social Inclusion through a 5-year Framework Agreement. The Coordinator bridges users and producers of statistics, and serves as a contact point in a field that is otherwise fragmented between different departments at Statistics Norway. Annually, approximately 10 million NOK is funnelled through the Coordinator and earmarked statistics and analyses on migration and integration. Statistics Norway has thus become the largest quantitative research institution on migration issues in Norway.

In this presentation we will describe the Norwegian “model” of organizing the migration field at the national statistical office, and discuss advantages and challenges of having a Framework Agreement with the Ministry. We will briefly present the main points in the Agreement and the role of the Coordinator, and give an outline of production and migration-related topics that are covered annually at Statistics Norway.

Keywords: migration, analyses, coordination, funding.

Chair: Zsuzsanna Kovacs

THE EXPERIENCES OF TURKEY AS AN ENLARGEMENT COUNTRY IN THE SECOND ROUND OF PEER REVIEW

Neslihan ONCUL

Turkish Statistical Institute, Ankara, Turkey

This study aims to relay the experiences of Turkey obtained within a pilot study in an enlargement country right after the Second Round Peer Review carried out in the member countries in 2013-2015 period.

As is known, the process of the first round peer review was not only implemented in ESS member countries, but it was also implemented in the enlargement countries under the definition "light peer review". The second round of peer review, on the other hand, represents a process of great significance in that the enlargement countries should follow the same procedures as member countries in a first-time application. In this sense, the experiences of Turkey will be a lodestar both for Eurostat in respect of future applications and for the prospective six enlargement countries.

In this document the preparations made by Turkish Statistical Institute for the second round of peer review in the light of her experiences she gained in the first round and the related activities carried out in November 2015 as well as the enhancement measures put in place in the aftermath of these activities will be addressed.

In the meantime the significance of CoP and Peer Review activities for the statistical office of an enlargement country and the contributions first and second round peer reviews have made will also be explicated. The study will especially provide special focus on the predictions, and the related plans, on how the second round peer review, which was more comprehensive than the first one, can contribute in respect of inferences as to the future activities in this context.

Keywords: Peer Review, Enlargement Countries, Code of Practice.



PEER REVIEW (2013-2015): LESSONS LEARNED, CHALLENGES AND OPPORTUNITIES TO THE INE AND THE SPANISH STATISTICAL SYSTEM

Agustín Cañada, Luisa Muñoz
National Statistical Institute (INE), Spain

During the period 2013 -2015, a new round of Peer Reviews (PR) was conducted in the European Union and the EFTA countries. As regards the National Institute of Statistics of Spain (INE) the visit of the peer reviewers was held in November 2014; the assessment report was published in February 2015, and in March, a plan for improvement actions suggested by the INE, in response to the report's recommendations, was disseminated.

This paper describes the main features of the process as well as its implications and impacts on the INE's activity: firstly, aspects of the methodology and organization of the process -close to an audit approach- which have posed a major challenge and significant effort for the INE, are discussed; secondly, the conclusions of the evaluation report of the statistical system, including fourteen specific recommendations on various areas of statistical activity (institutional environment, production processes and statistical outputs) are summarized; Finally, we have described the plan of improvement actions proposed by the INE in response to the recommendations of the report.

This plan of improvement actions, which covers the period 2017-2020, is highlighted as an example of PR's relevance and impact on the future of the INE and its statistical system: its set of remarks has provided inputs for the National Statistical Plan 2017-2020, and therefore constitutes one of the guiding elements of the of the INE's activities in the short and medium term.

Among the lessons learned for future rounds of PR, it is worth underlining the need for a shift in approach with regard to the run-up to the review process, infusing greater flexibility and simplicity into questionnaire design and documentation gathering, as well as the possibility of including specific statistical domains within the auditing process.

Keywords: Peer Review, audit-like approach, Statistical Planification.

PERSPECTIVE OF THREE DIFFERENT COUNTRIES ON PEER REVIEW ON EUROPEAN STATISTICS CODE OF PRACTICE: LESSONS LEARNED, FUTURE CHALLENGES AND COMMON FEATURES ON COORDINATION ROLE

C.M. Fransen¹, M. Gandolfo², B. Grandits³

1 Statistics Netherlands, The Hague, the Netherlands

2 ISTAT, Rome, Italy

3 Statistics Austria, Vienna, Austria

In the second round of peer reviews on the implementation of the European Code of Practice (CoP), an audit-like approach was applied covering all contexts of the CoP. In addition special attention was paid to the coordination role of NSIs in their respective national statistical systems.

The coordination role is important for NSIs as their strong position ensures that all other national authorities which contribute to the development, production and dissemination of European statistics comply with the standards of the European Statistical System (ESS) and fulfil the quality requirements for European statistics.

The revisions of Regulation 223/2009 reinforce the coordinating role already attributed to the NSIs in terms of scope, thus achieving more efficient coordination of statistical activities at national level. The NSI is responsible for coordinating statistical programming and reporting, quality monitoring, methodology, data transmission and communication on ESS statistical actions at the national level.

In this joint paper we compare the lessons learned during the process and the outcome of the Peer review, taking into account the different national statistical systems of Austria, Italy and the Netherlands. A comparative analysis of the similarities, differences and best practices allows us to highlight whether they are related to one typical system or have general applicability, and to explain the value of a strong coordination role. One conclusion might be that we should reinforce the coordination role as laid down in the Revised regulation 223/2009 even further and include a new principle on coordination in the European Statistics Code of Practice.

Keywords: peer review, coordination role, regulation 223/2009.



THE 2014/2015 EUROPEAN PEER REVIEWS - FACILITATING FOR STATISTICAL COOPERATION IN THE NORDIC COUNTRIES

Naja Andersen¹, Ari Tyrkkö², Live Margrethe Rognerud³

1 Statistics Denmark - Copenhagen - Denmark

2 Statistics Finland - Helsinki - Finland

3 Statistics Norway - Oslo - Norway

In 2014-2015 Peer Reviews were conducted throughout the European Statistical System (ESS) based on the European Statistics Code of Practice. National Statistical Institutes (NSIs) and National Statistical Systems (NSSs) were evaluated based on self-assessment and visits from peers. The reviews resulted in a list of recommendations and corresponding action plans from all the countries.

For the Nordic countries the result can be looked upon in the light of their similarities in both the development of their societies and their statistical systems. They all have register-based statistical systems, and they have had a formalized Nordic statistical cooperation since 1889. Several of the recommendations from the peer reviews and the planned actions are similar in the Nordic countries which in the authors' opinion open up for a further strengthening of collaboration between the NSIs.

The paper investigates similar recommendations and challenges pointed at in the peer reviews of five countries to see where the countries could benefit from even stronger cooperation. The authors believe there are several unexplored areas for common development projects, and that the Nordic NSIs/NSSs could be even better at sharing experience, tools and methods in statistical productions as well in organisational development.

The exercise emphasizes cooperation between countries to have full benefit of the vast work conducted with the Peer Reviews. It gives examples which hopefully encourage similar efforts between other countries. EU member states and EFTA countries are facing the same increased obligations to deliver results according to the same standards and towards the same ESS Vision 2020. The authors believe a stronger collaboration will both increase quality of the statistics and improve cost-effectiveness.

The article gives

- a basis for initiating cooperation across borders on actions following peer reviews
- inspiration to cross-country projects for development of official statistics
- ideas for collaboration to meet the ESS Vision 2020

Keywords: peer review, cooperation.



THE SECOND ROUND OF PEER REVIEWS – EXPERIENCE OF (SOME EXPERTS AT) STATISTICS FINLAND

Kari Djerf , Sirkku Mertanen
Statistics Finland, Helsinki, Finland

The second peer review of the European Statistics Code of Practise was carried out and the results will be evaluated. We will give some practical experience both from the domestic and the ESS point of view.

The self-assessment is the key element of making the whole process work well as it should lay the basis for evaluation. This time it was much better organized than in the first round, 2006-2007. Still there were parts in the main questionnaire not completely thought and evaluated in advance which caused some confusion. Much of the confusing elements were born from the Quality Assurance Framework. We will comment some of the points. The actual filling process was known to be time consuming. The process will be described and some evaluation of its success and problems will be presented. There will also be an assessment how well the results matched the effort.

Statistics Finland was the first organization to be evaluated in the peer review phase after the pilots and, thus, some practicalities may have not been fully operational at the time. Based on our experience we will comment the role of the consulting organisation, selection of the pool of peers, the actual peer review process and reporting. Similarly the domestic experience will be provided, especially what was learnt, how the improvement actions look like and how the organisation can and will use those in the future work.

Finally, some further ideas will be given on how to make the whole self-assessment and peer review process more consolidated before the next round.

Keywords: Code of Practice, self-assessment, peer review.



Chair: Giorgia Simeoni

CHALLENGES AND DISCOVERIES IN DEVELOPING QUALITY INDICATORS FOR THE GSBPM

Born¹, N. Dorsan², L. Reedman¹, M. Signore³

1 Statistics Canada

2 Turkish Statistical Institute, Ankara, Turkey

3 ISTAT

In 2014, the UNECE Modernization Committee on Standards created a working group to develop quality indicators for all phases of the Generic Statistical Business Process Model (GSBPM). The working group proceeded step by step through the model, drawing inspiration from the ESS Code of Practice, the NQAF, ESMS, ESQRS and work done by NSOs, particularly Statistics Canada's Quality Guidelines. The group mapped quality indicators for surveys to the structure of the GSBPM. The first draft included 168 indicators covering all 43 sub-processes of the model and overarching processes. The international expert community was invited to comment on the utility of the proposed indicators, as well as to propose additional indicators.

The feedback confirmed that there is keen interest in quality indicators. There is demand for indicators to be categorized in terms of their intended audience (internal or external to the NSO), their scope (process versus product) and their level (specific to a particular process or global, pertaining to the entire NSO). There was discussion about the relative merits of quantitative (percentage or count) versus qualitative (unacceptable, acceptable, excellent) indicators. Additionally, there is interest in identifying a sub-set of "key" quality indicators, or aligning the quality indicators with performance measurement indicators. Although the working group started with the survey context, it is clear that there is a need to incorporate indicators pertaining to administrative data. Another challenge is defining indicators that are useful. Usefulness can be characterized in terms of ease of implementation, information value, quality discrimination power, and coverage of the entire statistical process. Usefulness of some indicators can vary from one NSO to another.

This paper describes the challenges encountered in developing generic quality indicators, and highlights areas for further development.

Keywords: GSBPM, quality indicators; performance measurement.



QUALITY INDICATORS FOR THE INDIVIDUAL LEVEL - POTENTIAL FOR THE ASSESSMENT OF SUBGROUPS

E.-M. Asamer, H. Rechta, C. Waldner
Statistics Austria, Vienna, Austria

In the scope of the transformation to a register based census 2011 in Austria, a quality framework for statistical data based on administrative sources was developed. Now, this quality framework is used yearly to evaluate the quality of the register based labour market statistics. These quality indicators offer a wide range of possibilities to analyse the attributes on their own but also in combinations. This paper gives a short overview of the quality framework and its three stages (raw data level, census data base, final data base) as well as the different types of attributes (simple, multiple, derived). In the second part of the paper we present an approach for analysing different subgroups, thus showing the full potential of the quality framework. Crossing specific values of an attribute with other attributes or source registers, offers the possibility to analyse strengths and weaknesses of this register based statistics for these subgroups. Then this approach is applied on real examples from the register based labour market statistics 2013.

Keywords: Register based Statistics, Quality Indicators, Quality of Census.



QUALITY ASSESSMENT OF STATISTICS IN EUSTAT

María Victoria García Olea, Enrique Morán Aláez
Basque Statistics Office (Eustat), Spain

In this paper we will try to show how the quality assessment of statistics is deployed in the Basque Statistics Institute in the framework of quality management. Assessment is considered an important and final step in the statistical production process and we intend to conduct it in a standardised periodic way.

Assessment is seen in Eustat as an internal reflection that the responsible for the statistics makes with his/her team about the quality of them so that some improvements for the new cycle can be introduced.

In this paper we will explain how we have introduced this step of the statistical procedure; first of all the questionnaire for the assessment has been developed as an adaptation of the Eurostat's DESAP; also an assessment protocol and a standard report have been designed to guide the assessment and save the results that will be monitored for the following years. We will show how these tools have been tested three years ago in a subset of ten statistics, from various fields and of different methodologies, and how the preliminary version of the tools has been updated through the opinions and suggestions of the teams in order to comply with the aims and characteristics of our Institute. Then the assessment procedure has been extended to the whole organization.

This paper will detail aspects such as the identification of the statistics teams, the length of the assessment, the types of assessed statistics, the difficulties and positive features found while introducing and adapting the DESAP procedure to Eustat's peculiarities.

Finally we will present our plans for future developments and improvements, such as the migration of the Excel application to a web questionnaire, the introduction of a Basque language version of it and the extension of the assessment to statistics produced by other regional statistical producers.

Keywords: Quality assessment, statistics team, assessment tools, improvement, difficulties.

AGRICULTURAL ADMINISTRATIVE SOURCES DATA QUALITY: A PROPOSAL FOR STANDARDIZED INDICATORS

A. Bruni, S. Cusimano, D. Fusco

Italian National Statistical Institute, Istat

Farm Register is part of harmonized registers of Reg. 2186/93, abolished by Reg. 177/2008; according to the European Community indications, it provides a basic framework widely applicable for many possible uses. For example, it may be put to permit effective sample design with stratification by size, type of activity and location (Commission Decisions of 9 March 2005).

Administrative and statistical sources are used and combined to build up the Italian Farm Register: the main administrative sources handled are Integrated Administration and Control System (ICAS) and System for the identification and registration of Bovine animals. With the extensive use of administrative sources, it is of vital importance to determine the statistical usability of a statistical register on a regular basis. It means to determine the key quality constituents of administrative data in a systematic, objective, and standardized way. So, the aim of the paper is to identify the best fitting quality indicator available for quality assessment of above mentioned administrative sources at microdata level.

The identified set of indicators refers to the processing steps usually needed to pass from an administrative register to a statistical register (i.e. variable and unit harmonization, handling of missing objects and of missing values).

Administrative agricultural sources are affected by the specificities of the sector, characterized by small and very small productive units, strongly aided for the realization of a minimum income. Labour force is mainly based on family labour, often seasonal and part-time. Moreover the sector is strictly integrated with other activities such as transformation, trade, tourism, etc. These items make complex the correct identification of units as well as the estimation of their actual size and their principal activity. Finally there is a lack of a unique benchmark or pivot archive that covers the whole potential universe of agricultural holdings.

Keywords: input quality assessment, quality indicators, farm register, agricultural statistics.



INCLUSIVE GROWTH INDICATORS ON REGIONAL LEVEL

Ulrike Ch. Rockmann

German Institute for Economic Research (DIW), Berlin, Germany

Indicators as a useful reduction of complexity are almost omnipresent. They fulfil the quest for predictability and give the impression that they are a valid fundament for evidence-based policies as well as for policy evaluation. Therefore the need for indicators not only on the global but also on the regional is self-evident.

This paper focuses on the Inclusive Growth dimension Education [1] – especially on indicators that describe inequality with regard to the access to high-quality education. It considers whether the indicators developed on UN/OECD/EU level are recognized in Germany and meet the user needs. Adjustments for national and regional purposes using official statistics data and data from other sources are described. The role and involvement of official statistics in the development processes and the consequences for indicator quality is discussed.

[1] see <http://www.ipc-undp.org/publication/26506>, <http://www.oecd.org/inclusive-growth/All-on-Board-Making-Inclusive-Growth-Happen.pdf>, http://ec.europa.eu/eurostat/statistics-explained/index.php/Europe_2020_headline_indicators [30.10.2015].

Keywords: Inclusive Growth, indicators for policy making, education.



Chair: Jean-Pierre Poncelet

MEASURING THE QUALITY OF THE ITALIAN NATIONAL STATISTICAL SYSTEM: MODEL AND PRACTICE

A. Villa, S. Terracina

Italian National Institute of Statistics, Rome, Italy

The Italian Code of Official Statistics is a best practice across National Statistical Institutes (NSIs) at European level. Istat was indeed the first NSI that adopted a code for the statistics quality at national level in the wake of the European Statistical Code of Practice. In order to monitor compliance to the Code and to disseminate its principles throughout the Italian Statistical System, since 2010, year of approval of the Code, several activities have been planned, designed and tested. All these activities aimed at defining the theoretical monitoring system and identifying the tools to achieve it. This paper presents the activities and results related to the first monitoring cycle of the Italian Code of Official Statistics within the National Statistical System (Sistan), implemented from October 2010 until December 2014. In the following the Code and the model for its monitoring, the testing of its instruments and its outcomes, as well as the main findings of the activities conducted between 2012 and 2015, will be presented.

Keywords: quality, peer review, Italian Code of Official Statistics.



HARMONISATION AT THE ONS: COMPARABILITY OF SURVEY AND ADMINISTRATIVE DATA IN CONJUNCTION WITH EUROPEAN INFLUENCES

C. Wroth-Smith, B. Aquilina
Office for National Statistics (ONS), UK

The ONS Harmonisation Team has been working on harmonising survey questions, concepts and definitions to ensure that the quality of official statistics meets user requirements and supports comparability both within the UK and internationally. The ONS' harmonisation vision is that all inputs, processing and outputs for the Census and surveys and all data from administrative records will be harmonised, so that users can compare data from different sources with confidence and can merge and match data more easily. The Harmonisation work programme has been influenced by the recent ESS Peer Review.

The harmonisation of administrative data is a new and challenging area of work and forms a large part of this harmonisation work programme. The harmonisation team would aim to standardise the process of data matching across the Government Statistical Service (GSS) and produce harmonised inputs, processes and outputs for administrative data.

A further new area of work is the harmonisation of business statistics to comply with EUROSTATs Framework Regulation Integrating Business Statistics (FRIBS). The regulation requires ONS to move to a harmonised set of variables by 2019. Alongside this, and in collaboration, ONS are harmonising business survey questions where possible as Electronic Data Collection is rolled out over the next four years.

The ONS have a long working relationship with the UK Data Service and together plan to improve the Variable and Question Bank to include cognitive information, prominence, citations and version control using detailed metadata about questions, providing the user with more information than previously presented.

The paper sets out what the ONS has achieved to date with harmonisation, including the development of a harmonised question library, harmonised principles for surveys, administrative data and also what remains to be done. It also outlines the benefits of harmonising and details the issues and challenges faced when attempting to harmonise.

Keywords: harmonisation, statistical data quality, administrative data.



FACING THE CHALLENGE TO INCREASE QUALITY WHILE WORKING MORE EFFICIENTLY USING LEAN OPERATIONAL MANAGEMENT AND LEAN SIX SIGMA AT STATISTICS NETHERLANDS

Jeremy Visschers
Statistics Netherlands

Part of the ten point strategic agenda of Statistics Netherlands is to make our processes more effective and efficient and at the same time improve and secure the quality of our processes and output. This seems to be the eternal challenge: to become better with less effort.

In the past, and still on going, Statistics Netherlands has succeeded in meeting this challenge by large methodological and IT-redesigns that allowed breakthroughs in both quality and efficiency. To put more focus on the work processes, instead of on methodology and IT-systems, and to tap into the knowledge and potential of the employees, Statistics Netherlands embraced Lean Six Sigma and Lean Operational Management to realise an increasing part of their goals on efficiency and quality.

This presentation gives an insight in the way Statistics Netherlands introduced and implemented Lean Six Sigma and Lean operational Management, and the challenges and successes we had with it. Also the relation with the aspired ISO9001 certification is addressed.

VALIDATION IN THE ESS: A MEMBER STATE PERSPECTIVE

L. Quensel-von Kalben

Federal Statistical Office of Germany (Destatis), Wiesbaden, Germany

Validation is one key area of the ESS Vision 2020 and still rather neglected. Together with statistical data editing (to which it is strongly connected), its contribution to quality is paramount and its consumption of effort is a major factor in statistical production.

Eurostat launched an ESS.VIP Validation 2013 (continuing an internal project) and focused on three areas in particular, a common methodology, a common language and a common validation infrastructure. The main goal was an improvement of data quality at the interface of the NSIs and Eurostat, i.e. the data transmission process by using IT-services using an international standardized validation language.

To increase Member State involvement an ESSnet “ValiDat Foundation” was launched in December 2014. Main deliverables of the ESSnet were a “Handbook on Validation methodology”, an evaluation of the Validation and Transformation Language (VTL) of the SDMX-community and an analysis of the specific requirements of the Member States concerning a common technical architecture/infrastructure for validation.

The interests of Eurostat and the Member States, but also between the Member States are not always mutual. All agree that a common methodology is in demand and a validation language is of prime importance. Not all agree that VTL is the language to choose. While Eurostat focuses at the transmission level, some Member State head for goals that are more ambitious and look for solutions that can be integrated in their own environments. Other Member States are more eager to use centrally provided services and some would favour solutions in their own premises. The ESS.VIP and the ESSnet propose different scenarios how to deal with these differences.

The presentation will give an overview about the results of the ESSnet, achievements in the ESS in the meantime, and some recommendations for future actions on national and international level.

Keywords: Validation, ESS Vision 2020, ESSnet ValiDat Foundation.

INTEGRATED DATA PROCESSING SYSTEM (EAR)

H Debreceni

Hungarian Central Statistical Office, Budapest, Hungary

The main purpose of the EAR system is to prepare final production outputs based on datasets created in the course of data preparation and data entry processes and to support the data processing activities in general using standard protocols for any statistical domain. The development of the EAR system started in 2008 with an extensive on-demand survey within the Hungarian Central Statistical Office and is now used for several data processing activities in the HCSO. The result of this survey was a collection of methods used for data processing, which was the basis for the development of the EAR system. The main motivation was to increase the efficiency of data processing using standard methods and solutions and to provide common documentation for the data processing phase of the statistical business processes.

Based on our experiences with the EAR system, one of the main outcomes of the EAR is the increase in quality of data processing tasks due to standardization, full integration with the HCSO meta-information system and the automatically generated documentation of data processing activities. The users of the EAR system are statisticians; they design, redesign and manage their data processing activities. There are more than 200 users in the production environment from all subject-matter domains.

The HCSO considers the EAR system as a key importance system in the business processes with a lot of functionality to support statisticians in the design and management of the data processing tasks.

The functionality of the EAR is twofold. It is

- a framework - It supports the design and the management of data processing activities, using standard methods and data process definition structure
- a set of standard elements – the EAR procedures/functions are standalone standards, defined in a common way to support the definition of data processing activities

Keywords: data processing, integrated systems, standardization.



Chair: Thomas Burg

RELEVANCE AND BENEFIT OF COMMERCIAL QUALITY IMPROVEMENT METHODOLOGIES AT STATISTICS CANADA

L. Reedman, A. Ness

Statistics Canada, Ottawa, Ontario, Canada

Developing and maintaining credibility is a particular challenge for National Statistical Organizations. Without the trust of the public and policy makers, statistical products have little value. To achieve this level of trust, not only does quality need to be assured, but also continuously improved. One way to do this is through certification from a recognized source, such as the ISO or Six Sigma. Most commercial quality improvement methodologies were developed for application in manufacturing; however, many of the principles and practices are a good fit with the core business of producing official statistics. Although the stamp of approval is a demonstration that certain clearly defined quality measures are in place, the effort required to achieve and maintain the certification can be significant, and the effectiveness (impact on actual quality) is not guaranteed.

This paper looks at quality improvement methodologies typically used in industry, and compares them to the principles and practices in use at Statistics Canada. The commercial methodologies considered are Total Quality Management, Lean, Six Sigma, the DMAIC improvement model, agile practices and ISO certification. We look at the underlying assumptions, objectives, constraints, application and expected outcomes, as well as challenges and potential disadvantages and benefits.

Keywords: quality, certification, effectiveness.



QUALITY WITHIN ONS – PROVIDING A FRAMEWORK FOR STATISTICAL PRODUCERS AND ASSURANCE FOR OUR USERS

J. Pobjoy

Office for National Statistics, (ONS) Newport, United Kingdom

A requirement of the European Code of Practice is that National Statistics Institutes must define their quality policy and make this available to the public.

To ensure this requirement is met, ONS have a Quality Management Strategy in place which is publicly available to provide assurance to the users of our statistics and includes activities which monitor, improve and report on the quality of statistical products. It also serves as a useful framework for the producers of statistics within ONS. The strategy has recently been updated and relaunched.

The updated strategy reflects the activities that we have in place as an organisation to manage quality and sets out goals for improvement. It is designed to reflect the organisational approach and is such relevant to all areas of the office. The strategy is supported by a statistical quality framework which sets out the quality initiatives in place for quality assurance, quality control, quality improvement and quality reporting.

The quality management strategy also strengthens the governance for quality management within ONS. A network of quality champions provide regular reports to the Quality Centre and any issues relating to quality management are fed up to a senior committee within the organisation twice each year

The updated quality strategy makes a clear connection between day-to-day activities and the organisation's commitment to quality as well as providing a framework for the activities of the Quality Centre. The strategy proposes all areas create a local framework to capture any local activities that are in place in addition to corporate activities.

This paper/presentation will provide more detail on the quality strategy and provide additional information around the quality frameworks. We will report on the how things have progressed since the launch of the strategy and how the quality management culture is being embraced in the ONS.

Keywords: quality management, strategy, framework.



HOW TO COPE WITH ALL THOSE RULES

Douwe Kuurstra, Max Booleman
Statistics Netherlands

International law, international regulations, national law, specific statistical regulations, Code of Practice (CoP), Privacy, ISO 27001 and ISO 9001. Some of the examples of 'rules' we have to work with. How do we manage that, how do we transform rules into practice? For Statistics Netherlands this is a continuing challenge. Until recent we combined all the rules in the Quality guidelines and worked directly with those guidelines, for instance as an audit framework and a self-assessment questionnaires. The peer review was very positive. The Peer Review team concluded that Statistics Netherlands, and by extension the production of European statistics in the Netherlands, is very highly and uniformly compliant with all principles of the CoP. But we are moving on; we want more external proof to show our commitment with quality of process and product, privacy, security (ISO 27001) and continuous improvement.

Employees of Statistics Netherlands are reliable professionals and every employee of our office is responsible for the quality of his own product. But their main task is to produce statistics, not to understand and fore fill all those rules mentioned before. Therefore we, the people from the quality department, have the task to guide and inform them, to translate existing rules; not only from English to Dutch, but also from theory to practise. That's to our opinion the only way to make them involved: to make them real owners of quality. That's our goal for the years to come.

In the proposed paper we explain the system which lead to the positive remark from the peer review. Then we will tell about the next steps: Privacy certification and ISO 9001:2015 certification. The lessons learned from our original approach and the privacy and ISO approach will bring us to the next level: fulfilling rules without knowing them but just by attitude.

Keywords: Code of Practice, ISO Quality Management.



A SYSTEMATIC APPROACH TO QUALITY: THE DEVELOPMENT AND IMPLEMENTATION OF A QUALITY MANAGEMENT FRAMEWORK IN THE CENTRAL STATISTICS OFFICE

S. Portillo, K. Moore

Central Statistics Office, Cork, Ireland

In the context of increasing and more formal scrutiny of official statistics at international and EU level and having regard to our goal of the continuous improvement of our statistical processes and products, the Central Statistics Office launched a Quality Management Framework (QMF) programme of improvement projects in January, 2014.

The development of the QMF is an extensive and long-term programme of activities, which will ensure that statistical production meets the highest standards as regards quality and efficiency. The project focuses on the issues of standardization and storing of documentation in a central repository based on the principle of updating documentation once and reusing it multiple times for different products such as section manuals, quality reports, etc. It includes a system of metrics for data quality and internal processes, metadata standards, strong documentation control and other important internal governance elements. The project extends the use of systematic process improvement programmes to streamline, standardise, document and build robustness into our core processes of data collection, management, compilation, analysis and dissemination.

This paper will tell the story of the development of the QMF from how it was initiated to its current status where it transitions from planning to implementation. It will highlight the challenges faced by the quality team including the importance of identifying and implementing quick wins at an early stage, while keeping in mind critical success factors through the lifespan of the implementation pilot. The planning phase of this project was completed in November, 2015 with the implementation phase starting in January, 2016.

Keywords: Quality Improvement, Documentation, Change Management.



HOW TO INCREASE QUALITY IN THE CENTRAL BANKS STATISTICAL BUSINESS PROCESS? THE EXPERIENCE OF BANCO DE PORTUGAL

António Agostinho, Alexandra Miguel,
Banco de Portugal

Statistical production responsibilities in the field of European Statistics, in particular due to the evolution of European System of Central Banks' statistics requirements, had a considerable increase in the last decade. During this period Banco de Portugal has been developing several initiatives to ensure a statistical production with high quality standards aiming at fully meeting user's needs.

This paper gives an overview of the main initiatives under development in Banco de Portugal' Statistics Department to improve the quality of statistical data results and provide a more efficient data quality management in statistical systems. These new developments comprise: i) reduction of reporting burden for economic operators; ii) implementation of solutions based on elementary data to allow an integrated exploration for different purposes; iii) implementation of systems based on Business Intelligence architecture; and, iv) development of solutions for more efficient statistical communication.

Keywords: data quality management, business intelligence, production integration, statistical communication.

Chair: Steven Vale

QUALITY MANAGEMENT OF CROSS-BORDER RELATIONS IN THE EUROPEAN SYSTEM OF INTEROPERABLE STATISTICAL BUSINESS REGISTERS

N. Rainer¹, H. van der Ven², I. Xirouchakis³

1 Statistics Austria

2 Statistics Netherlands

3 Eurostat, European Commission, Luxembourg

Statistical business registers (SBRs) are the backbone for the production of enterprise statistics. In addition to their main task of providing population frames for enterprise statistics, they serve various other functions, such as linking statistical units to administrative and other basic data sources as well as micro data linking between statistical domains. Furthermore, SBRs are a powerful and detailed data source for the observation of the structure and development of the economy.

However, as SBRs are maintained by the NSIs, the coverage of the register is limited to units resident in their country. Thus, relations between enterprises active in different countries, but belonging to the same enterprise group are not recorded in the national business registers. Therefore, also the unit multinational enterprise group in its totality does not exist in any national SBR.

In order to remedy this draw-back Eurostat developed the EuroGroups Register (EGR) which covers the most important enterprise groups in Europe with all their constituents, irrespective in which country they are resident. The maintaining of the EGR is managed by a division of labour between Eurostat and the NSIs. In order to be able to provide consistent population frames in the European Statistical System (ESS) information between the EGR and the national SBRs needs to be exchanged and the responsibilities of the involved parties have to be fixed.

The current VIP.ESBRs project aims to support the production of consistent and higher quality business statistics through improved interoperability of the SBRs via the EGR. One of the tasks of the ESSnet ESBRs is to model and propose the appropriate Business Architecture that would – when implemented - lead to the project's goals in the coming years successively. The ESBRs will thus contribute to the implementation of the ESS Vision 2020.

The paper will provide an overview of the ESBRs Business Architecture, especially of the processes and information that would need to be maintained and exchanged. As linking of the units in the national SBRs with the proper units in the EGR is of crucial importance for the derivation of consistent frame populations, the paper will focus on the validation and quality management of the cross-border relations.

Keywords: European System of interoperable Statistical Business Registers (ESBRs), Enterprise Architecture, Business Architecture, Quality management, Data validation.

QUALITY IMPROVEMENT OF THE EUROGROUPS REGISTER

A. Liotti, S. Maus, J.-M. Museux, L. Szekely, I. Xirouchakis
Eurostat, European Commission, Luxembourg

The EuroGroups Register (EGR) includes information on multinational enterprise groups and its constituent units when at least one legal unit of a group is located in a European country. The EGR aims to serve as backbone for European business statistics, in particular concerning globalization and cross-border questions.

The paper will discuss quality aspects of the EGR and will consist of three main parts.

After having introduced the relevant questions concerning the quality of the EGR, a selection of key figures to analyse the present (AS-IS) state of the EGR quality are presented in the first part of the paper. These quality indicators will be calculated on the basis of the latest reference year available for the EGR frame and will cover aspects of input, throughput and output quality.

The second part will deal with the target (TO-BE) state of the EGR quality after implementation of a data quality programme (DQP) and will outline the steps needed to achieve the target EGR quality. The DQP, which implementation is envisaged by end of 2016, is based on four main components, following ESS wide standards: Quality reporting, quality standards, quality assessment and quality improvement. The overall objective of the DQP is to achieve a level of quality allowing EGR to fulfil its backbone role.

Using the target business architecture of ESBRs as a basis, the final part of the paper will introduce, position and analyse the impact on quality of the on-going development of EGR 2.0 information system and the profiling methodology for European enterprise groups. Profiling is a specific methodology of collecting high quality information on enterprise group structures that will substantially improve the quality of national statistical business registers and the EGR.

Keywords: Data quality, Enterprise Architecture, EuroGroups Register (EGR), Profiling.



APPLYING THE GENERIC STATISTICAL BUSINESS PROCESS MODEL (GSBPM) TO THE NATIONAL BUSINESS REGISTER; THE SPANISH EXPERIENCE

L.E. Barbado

National Statistical Institute (INE), Spain

Over the last decade, the growing demand for high-quality and comparable business statistics and the need to make production more efficient has put the focus on the management of Business Registers for statistical purposes.

Recent international initiatives aim to reinforce the central role of Business Registers as the backbone for all business statistics. This context involves an increasing range of uses; as the reference for survey frames, a tool for direct statistical production on the business population or the central basis for developing new statistics, through an efficient combination of administrative and statistical data.

The experience in application of the Generic Statistical Business Process Model (GSBPM) to Business Registers is quite new. This innovative action is recognized as a critical step, in order to fulfill the above-mentioned objectives, as well as adopting specific interoperability requirements and implementing coordinated data quality programs. In this scope, the descriptive process also needs to cover the national interactions with the Euro Groups Register (EGR), the Statistical Register of the European Communities on multinational enterprise groups.

The application of the GSBPM to the management of the Spanish Business Register was carried out by the NSI during 2015. This paper provides a first assessment of the work done, focusing on the selected approach for the description of the GSBPM phases and the criteria adopted for a proper assignation of the core parts of our business process. The main restrictions found and the potential value added of this exercise are also pointed out.

Keywords: National Business Register (NBR), Euro Groups Register (EGR), Generic Statistical Business Process Model (GSBPM), management model.



PROFILING: A NEW AND BETTER WAY TO APPREHEND THE GLOBALIZATION

O. Haag

National Institute of Statistics and Economic Studies (Insee), France.

French foreign trade statistics are traditionally calculated by Customs for trade in goods, and by the Banque de France for trade in services. These statistics are prepared on the basis of data on the legal units. INSEE provides the Customs and Banque de France statistical offices with a copy of its statistical business register (SIRUS) to serve as a reference.

The INSEE statistical business register is fed directly by the structural statistics production system (ESANE) with the export turnover figures declared by companies in their profit and loss accounts.

It is therefore possible to calculate export statistics directly on that basis. While these statistics do differ somewhat from the official statistics of the Banque de France and the Customs, mainly due to differences between their sources, they are nonetheless of very high quality. The main contribution of the statistical business register in this respect is that it makes it possible to calculate statistics on the basis of units other than the sole legal units: for profiled enterprises on the one hand, and groups on the other. In these times when globalisation has come to play such a prominent role in the French economy, such figures can prove precious for economic analyses.

This article first describes the French statistical business register system and the way to define the profiled enterprises. Then, it presents foreign trade statistics calculated from the statistical business register at legal unit and enterprise levels showing the impact of moving to the enterprise unit on French trade statistics.

Keywords: business statistics, trade statistics, profiling, business register.



DATED - UNDATED - OUTDATED: THE ISSUE OF TIME STAMPS AND REFERENCE DATES IN THE STATISTICAL BUSINESS REGISTER

R. Sturm

Federal Statistical Office Germany, Wiesbaden, Germany

A crucial feature of the data in statistical business registers is the attribution of time references to the register information. For appropriate use of register data it is crucial to distinguish clearly between the “technical time reference” (the date/period of time an information enters the register or is edited in the register or made available by the register) and the “validity time reference” (the date/period of time in the “real world” for which the information is valid). Without “validity time reference” one may compare apples with oranges. Survey frames and extrapolation based on the register would refer to a mixture of real-world situations.

The problem statement may sound trivial. The presentation will illustrate the relevance and the complexity of the issue using some real-world evidence. An approach to “do things properly” will be presented– which unfortunately leads again to issues of complexity.

Keywords: Business Register, time references, validity.



Chair: Hannah Kiiver

USING ADMINISTRATIVE DATA AND MODEL BASED ESTIMATION FOR IMPROVING ITALIAN AGRICULTURE STATISTICS

Roberto Gismondi, Loredana De Gaetano, Federica Piersimoni
ISTAT (Italian National Statistical Institute), Roma, Italy

In the European Union, on the basis of the Regulation (EC) 543/2009 each Member State must produce estimates on agricultural surfaces according to 127 kinds of cultivations and calculate early estimates concerning the new agricultural year.

Actually, agricultural surfaces are estimated by ISTAT through regional experts evaluations; it is difficult to assess quality and several delays occur. Early estimates are obtained through a sample survey (12,000 agricultural holdings) and design-based Horvitz Thompson weights corrected for non responses. This strategy does not use statistical relationships between crop data concerning consecutive years and differences between forecasted and true surfaces may be large.

As regards agricultural surfaces, IACS data referred to 2014 and 2015 are available. IACS is the Integrated Administration and Control System managing payments to agriculture producers, who fill a yearly administrative dossier providing information on agricultural surfaces by kind of crops. Administrative IACS data have been summarized by regions and Italy as a whole. In both years the average difference between IACS and ISTAT surfaces is 2% only. Moreover, IACS data are available within a shorter time lag if compared with experts estimations, cover a broader set of cultivations and may substitute actual estimates gradually.

As regards early estimates, for the agricultural year 2015/2106 double sampling has been applied: 50% of the sample has been drawn from the subset of early estimates survey's respondents in the previous year (panel units); the other 50% has been selected from the subset of holdings interviewed in the 2013 farm structure survey (FSS, not panel). According to double sampling theory, a model based regression estimator which combines separate crop estimates for panel and not panel units has been introduced, with model weights depending on agricultural surfaces. Results show decrease of estimates model variances and higher degree of coherency between surface use in following years.

Keywords: Administrative data, Agriculture, Crop, Double sampling, IACS, Panel, Surface.

PLAUSIBILITY ASSESSMENT OF FLASH ESTIMATES OF THE INCOME DISTRIBUTION

B. Micu, A. Leulescu, C. Lopez-Vilaplana
Eurostat, European Commission, Luxembourg

EU-SILC indicators on poverty and income inequalities are an important part of the toolkit for the European Semester which is the yearly cycle of economic policy coordination among member states. However, income year N is only available in the autumn N+2 which comes too late for the policy agenda. In order to be able to provide earlier information, one approach pursued in Eurostat is the development of flash estimates of the income distribution.

Several strands are investigated including: re-weighting procedures and modelling techniques in order to account for demographic and labour market changes captured in more timely sources such as LFS, and the use of the microsimulation models that replicate at individual/household level the effects of the different taxation regimes.

The flash estimates are therefore model-based and rely on micro-data. The quality assessment of the flash estimate includes back-testing (measuring the model's ability to forecast the past) supplemented by a plausibility assessment.

Plausibility is NOT the a priori probability; we define a highly plausible estimate as one that fits prior knowledge well: with many different sources of corroboration, without complexity of explanation, and with minimal conjecture. This is a view of plausibility as a subject-related notion, thus our approach is close(r) to the data users' perspective.

As part of the assessment procedure, we look at: (1) continuity of historical trends, (2) coherence with other data sources, and (3) consistency with the evolution of macroeconomic indicators.

The paper describes the approach and some interesting intermediate outcomes.

Keywords: plausibility, income, nowcasting.



MODELS IN OFFICIAL STATISTICS

Kees Zeelenberg, Barteld Braaksma, Bart Buelens, Peter-Paul de Wolf
Statistics Netherlands

In official statistics, models are sometimes used when there enough data for using design-based techniques are available. In those cases, models can provide alternative, in some sense, better estimates. National statistical institutes (NSIs) have always been reluctant to use models, apart from specific cases. Based on the experience at Statistics Netherlands we argue that NSIs should not be afraid to use models. In order to aid statisticians at Statistics Netherlands in the use of models, we have set up a number of guidelines based on the principles of objectivity and reliability of the European Code of Practice, and the idea that the primary purpose of an NSI is to describe, and not to prescribe or judge. In particular we require that the use of models is documented, made transparent to users, and refrains from making forecasts; also the models should rely only on actually observed data and they should be validated extensively.

After a description of these guidelines, we discuss how several examples, such as small-area estimates, seasonal adjustment, capture-recapture for hidden populations, corrections of mode effects with mixed-mode data collection, and corrections for non-response, fit into the guidelines. We also discuss some cases that do not seem to fit the guidelines.

We also look at the relatively new area of big data in official statistics. Here problems of coverage and selectivity may be severe, and we discuss how models may be used to produce more accurate statistical information. We also discuss some examples where modelling of big data was not successful.

Keywords: models, official statistics, big data.

GDP FLASH ESTIMATES: SOPHISTICATION THROUGH SIMPLICITY

Kristina Kiriliauskaitė
Statistic Lithuania, Vilnius, Lithuania

The sustainable development of GDP measurement is very important for the evaluation of the economy. Thus, huge attention is given to producing quarterly GDP estimates, their accuracy, timeliness and accessibility. Qualitative and relevant GDP estimates lead to accurate economic, political and business decision-making. Lithuania is one of the European countries which produce GDP estimates at a very early stage – within 30 days of the end of the corresponding quarter. When calculating actual GDP at such an early stage, a considerable amount of statistical data is still missing. Therefore, mathematical and econometric methods have to be used to derive GDP flash estimates.

A Lithuania's GDP flash estimate based on the production approach is obtained using an indirect method by estimating the value added of each economic activity and intermediate consumption separately. The paper presents Lithuania's experience and good practice of making GDP flash estimates in the context of limited data availability using classical econometric methods, such as linear regression, time series analysis (ARIMA, ARIMAX) and combined nowcasting techniques. It describes the sophisticated modelling procedure for deriving hundreds of models and their incorporation into one GDP estimation system – to produce high quality, timely and accurate estimates through simplicity. The paper explains how the models were built, how the explanatory variables were chosen, and how estimate modelling was performed through the combination of linear regression and time series models. A chain-linking method, used to obtain real GDP volumes, and ways of estimating quality analysis are also covered by this topic. Moreover, the paper discusses estimation system stability and adaptation to the global methodological changes (ESA 2010).

Keywords: GDP flash estimate, regression and time series analysis, estimate quality analysis.

EARLY ESTIMATES OF ENVIRONMENTAL ACCOUNTS (ECONOMY-WIDE MATERIAL FLOW ACCOUNTS)

A. de la Fuente, R. Marra Campanale
Eurostat, Luxembourg

Environmental accounting is a multipurpose conceptual framework that describes the interactions between the economy and the environment. The economy-wide material flow accounts (EW-MFA) is one module of the European environmental accounts. They trace the material extraction from the environment feeding through the economy. The EW-MFA are designed to meet the information needs of institutions, citizens and researchers. Moreover, EW-MFA are used for the lead indicator 'resource productivity' of the Europe 2020 flagship initiative A resource efficient Europe. Eurostat publishes EW-MFA data with a timeliness of T+27 months, i.e., three months after the deadline for country data transmission established in Regulation (EU) No. 691/2011. After key EU policy users called for an improvement of this timeliness for economic analysis, decision-taking and policy-making, Eurostat set up production of early estimates for the EU and Member States at T+9 months and T+18 months. Whereas T+27 estimates are produced by the countries with standard techniques, Eurostat bases the T+9 estimates on regression modelling techniques and predictor variables and the T+18 estimates on more limited national data sources. This approach allows publishing every year in March (Y-3 data), June (Y-2 data) and September (Y-1 data). Users are satisfied with this production cycle and the experience might be extended to early estimates in other environmental accounts. The use of modelling in official statistics raises some questions. Some argue that official statistics should not make use of models. Actually, modelling is everywhere as almost any step of statistical production requires some form of model or assumption. A proper use of modelling can allow official statistics meet the needs of users which otherwise would be unattainable. Requirements are the use of well-established and well-tested models, proper documentation of methods and transparency towards users.

Keywords: environmental accounts, timeliness, data modelling.

Chair: Antonio Salcedo

FIGURING FIGURES: EXPLORING EUROPEANS' AWARENESS OF OFFICIAL ECONOMIC STATISTICS

M. R. Vicente, A.J. López
University of Oviedo, Asturias, Spain

Economic issues have been a major concern for Europeans in the last few years. Unemployment, the economic situation and inflation have been consistently reported as the three most important concerns for European Union's population according to Eurobarometer surveys. Given such worries, it is reasonable to think that people would know about the current economic performance, that they would be aware of the main economic figures. But, do they really know them? Do they know the national rate of unemployment? Or whether prices have increased or decreased compared to last year's?

Research in the United States shows that public awareness of official economic statistics tends to be low and, moreover, that there is variation among socio-economic groups. In particular, income, age, educational attainment and gender have been found to be significantly correlated with people's knowledge of the main economic figures.

Within the European framework, Eurostat considers statistical information to be essential for decisions and evaluation at European level, and also an important objective and a down-to-earth way of measuring how people live. This paper focuses on this second perspective, with the aim of exploring to what extent Europeans are aware of some of the most important economic indicators.

In particular, attention is focused on three key indicators: Gross Domestic Product, unemployment rate and inflation. Using data from a recent Eurobarometer, we are able to explore these issues across the 28 member states of the European Union. Several regression models are specified and estimated in order to identify the relationship between individuals' economic awareness and their socio-economic characteristics. We also explore whether individuals' beliefs about the importance of economic issues influence their awareness of official economic statistics.

Keywords: official statistics, economic indicators, statistical literacy.



A KNOWLEDGE-DRIVEN SOCIETY – CHALLENGES FOR POLISH OFFICIAL STATISTICS

J. Gustyn, A. Mróz

Central Statistical Office, Warsaw, Poland

With rapid progress of societies the awareness regarding the significance of data, information and knowledge has raised in every sphere of life. Information has become the most wanted resource, sometimes taking priority over tangible assets. The development of new information and communication technologies has, however, strengthen not only the value of information, but also its amount. The ease with which vast quantities of information can be gathered and processed nowadays is both a blessing and a curse. The problem we face now regards an overload of data and information, which does not necessarily translate into knowledge.

Accurate and reliable statistics are fundamental to modern knowledge-driven society. Every day, millions of individual and collective decisions are taken on the basis of statistical data. Using objective, reliable and high-quality information in decision-making processes substantially improves their effectiveness. It testifies the unique role of official statistics as a producer, manager and disseminator of huge volumes of data. Managing immense information resources seems to be the challenge for the next decades, also for the official statistics confronting the demand for more, better and faster data.

With the aim to address these challenges, the Central Statistical Office of Poland has taken activities aiming at enhancing not only the production of statistics, but also statistical comprehensibility. Statistical education and communication with users are of particular importance, as they determine perception of information and attitude towards the official statistics. One of such activities, in the face of abundance of strategic documents and information scattering, was gathering all measures significant to the monitoring development policies in one place or providing users with interpretational clues. Such solutions contribute to the fulfilment of a crucial task undertaken by the CSO, being a statistical signpost for the knowledge-driven society.

Keywords: information overload, statistical education, monitoring development.



STATISTICAL LITERACY PORTAL AS A MARKETING TOOL

Adolfo Galvez, Maria Jesus Vinuesa
National Statistical Institute (INE), Spain

The Spanish Statistical Literacy portal Explica from INE takes a wide approach to reach the goal of improving the statistical knowledge in society. In fact, it is not only seen as a learning platform targeted at the academic institutions and actors: teachers and students. It is also seen as a marketing tool in order to spread, in a user friendly way, important statistical concepts and values of official statistics to a wider audience.

In addition to providing resources directly applied in the classroom such as tutorials on basic statistical concepts, how to conduct a survey, how to make data meaningful and so on, Explica presents specific aspects of official statistics, which helps to understand properly these concepts and avoid misinterpretations. Good examples of it are the Statistical Secrecy, Life Expectancy, Threshold Poverty, etc. Through Explica values of official statistics and quality in particular are promoted, but in a less formal way, using visual aids in videos, so it can be easier to grasp.

Moreover, Explica is much more than a website where you can find the contents as they are expanded through social networks such as (Youtube, Twitter...) reaching all kind of public. This project engages with the new ESS Vision 2020 as it provides valuable information about official statistics and their high quality standards and “improves the statistical literacy of European citizens and institutions by guiding them through the deluge of data and information from various origins.”

Keywords: Statistical Literacy portal, marketing tool, high quality standards.



EUROPEAN MASTER IN OFFICIAL STATISTICS IN USE: SHORT-TERM GOALS VS. LONG-TERM VISION

Z. Aleksonyte-Cormier, M. Zwick

Eurostat, European Commission, Luxembourg

The European Master in Official Statistics was launched by Eurostat in 2014. Following the first EMOS call in 2014, the European Statistical System Committee awarded twelve Master programmes with the EMOS label in May 2015. In autumn 2015 the second EMOS call for applications was published and the next round of Master programmes should be awarded the EMOS label in May 2016.

With the running EMOS programmes, a number of activities have been launched or are about to be launched and these should be relatively quick wins. Given that EMOS represents not only a label, but also a growing European network of universities, NSIs, Central Banks and other data producers, networking activities such as workshops, open days, summer schools and thematic webinars are extremely important and other ways of connecting different actors should therefore be further developed in order to facilitate networking, sharing and cross-border exchanges and to enhance the European dimension of EMOS.

Another relevant and perhaps longer-term consideration for EMOS is how to use it for continuous training of staff already working in NSIs. In this respect, and despite the intrinsic differences in the target groups, it will be important to further explore synergies between internal training in statistical offices, the European Statistical Training Programme and EMOS.

With these developments in mind, the first graduates from EMOS Master programmes will be ready for the labour market in spring 2017. Therefore, it is now time for recruiting data producers to start thinking about how to integrate EMOS graduates and bring added-value to EMOS in the long run.

Keywords: Official statistics, statistical education, lifelong learning.



STATISTICAL TRAINING AS A FACTOR IN QUALITY

F. Veira, T. Martínez de Ibarreta
National Statistical Institute (INE), Spain

In Spain, the National Statistics Institute (Spanish initials: INE) has a department called the Public Administration School of Statistics (Spanish initials: EEAP), whose mission is to provide training in statistical methods, techniques and products for civil servants from all public administration bodies, with the aim of being continuously adapted to the new trends in this field.

The intention is to showcase the positive experience of this arrangement, by which the training needs of the National Statistical System, incorporating public administration bodies from national to local level, can be met.

By means of this training unit, the changing needs and approaches required by the strategy behind national and international statistics can be addressed and put into practice.

The EEAP provides both face to face (in various formats) and online courses. This latter approach means more students can be reached.

This initiative provides a window through which products and processes can be correctly explained to users of statistics.

One of the main objectives is to promote the use of classifications and procedures in the production of official statistics.

Keywords: Continuum Learning, Governmental Institutions and Methodology.



Chair: Ana Carmen Saura

AFTER THE PEER REVIEW – STEPS TOWARDS ENHANCING THE CO-ORDINATION ROLE OF HCSO IN THE NATIONAL STATISTICAL SYSTEM OF HUNGARY

J. Kárpáti

Hungarian Central Statistical Office, Budapest, Hungary

The paper/presentation will introduce how the Hungarian Central Statistical Office started developing its organisational structure and the legal framework of official statistics in the country, in order to develop its coordinative function. It describes briefly the related ESS Code of Practice Peer Review recommendations and deals first with the new approach to the organisational position of supporting subject matter (“business”) functions and their alignment to the GSBPM model, and in particular the GAMS0 model. The second part will describe the toolbox of the major new coordination functions and regulatory solutions that got incorporated into a completely new draft of the national statistics act, supporting the strengthening of such activities throughout the system. It will especially deal with tools like the drafting of a national code of practice, the audits of the official statistical system organisations and the establishment of a new board for the members, also with a brief outlook to similar actions in other member states. This part will also highlight some reactions, opinions, proposals received from the member organisations of the official statistical system upon the draft law, and the arguments HCSO gave during the iteration process (the new act will pass the parliamentary debate phase after the submission of this abstract). The third part of the paper will introduce the recent actions that HCSO already implemented and the first experiences in 2016 working by the new operational model.

Keywords: co-ordination, statistics act, organisational development.



NSS COORDINATION: A TOOL OR A FUNDAMENTAL PRINCIPLE?

Antonio Salcedo, Yolanda Gómez
National Statistical Institute (INE), Spain

One of the objectives of the amendment of Regulation 223/2009 has been the reinforcement of the coordination role of the NSIs in the National Statistical Systems. The amendment has fleshed out the coordination role harmonizing some minimum requirements to make this tool effective. Now it's time to implement this amendment and the questions are: Is it enough with a “legal” coordination? What is the impact of an effective national coordination in the ESS?. In this paper we will analyse the scope of the coordination role and its impact at EU level. The need for updating the Code of Practices is a fact and is a consequence of its own nature of self-assessment instrument. The Code is an alive tool that has to be adapted to the reality and the emerging needs. The “Coordination” role has been “peer-reviewed” on a voluntary basis since 2006. At this stage the debate on the importance of coordination is open at EU level. There are some Countries and International organizations that recognize “coordination” as a principle related with the institutional environment. However, in the European Statistical System, there are still some questions to be answered, as, for example, is coordination a fundamental principle under the EU legal system? What would be the consequences of including it as new principle in the CoP? This article will go through these issues focusing in the possible existing options.

Keywords: Coordination, statistical principles, Code of Practices, institutional environment.



TOOLS AND METHODS FOR NATIONAL COORDINATION IN NATIONAL STATISTICAL SYSTEM

Leena Storgårds

National Statistical Institute, Statistics Finland

Updated Regulation 223/2009 strengthens the coordination role of National Statistical Institutes. EU Member States are now implementing the regulation and are in process to consider a) how to define coordination, b) who are the other national statistical authorities (ONAs) and c) what are the tools, methods and ways to do coordination at national level.

To define coordination is not always easy as it has interlinks with cooperation. Coordination is not always cooperation but the good way to coordinate uses always cooperative methods. Aim of coordination should be collaboration and unification of all units in national Statistical System.

To define scope of coordination, one should first have clear understanding who are ONAs in National Statistical System that are producing ESS-statistics. Borderline between ESS-statistics and national statistics will be handled as they have clear connection to clarify the scope of ONAs. Product list and definition of EU-statistics is also important when ESS-statistics' scope is defined.

To understand different ways of doing cooperation, examples are given as Statistics Finland has long experience in coordination and has developed tools and methods to support this task.

Presentation will concentrate on sharing these ideas and experiences with audience.

Keywords: coordination, ONA, list of products.



COORDINATION OF NSS – REQUIREMENTS AND PRACTICE

J.Minkevica

Central Statistical Bureau of Latvia, Riga, Latvia

Member States of European Union now are facing implementation of amended Regulation 223/2009 on European statistics. The amendments concern professional independence, coordination role of National Statistical Institutes (NSIs), Commitments on Confidence (CoC) and access to administrative records. Member States must implement these requirements at the national level in their National Statistical Systems (NSS) in accordance with national characteristics.

The coordination of NSS involves setting up a national framework where main requirements for development, production and dissemination of (official) statistics are defined, implemented and monitored. In fact, the scope of coordination includes all aspects covered by the Regulation 223, and further elaborated in the European Statistics Code of Practice. For statistical authorities of Latvia rules of the game are clear. The new Statistics law guides through the labyrinth of rules to be followed and requirements to be applied. As the law is already adopted, now the challenge is its technical implementation. The presentation will provide an insight into the new law, as well as will reflect next planned steps.

Keywords: coordination, NSS, statistics law.

STRATEGIC PLANNING IN A DECENTRALISED STATISTICAL SYSTEM – THE GERMAN EXAMPLE

Y. Vatter

Destatis, Wiesbaden, Germany

Concerning the statistical processes, the Code of Practice underlines the importance of a reputation for good management and efficiency as they benefit official statistics' credibility. Consequently, strategic planning is of vital importance for assuring and further enhancing the quality of official statistics. It is generally acknowledged that only through systematic planning, i. e. clearly defined objectives, measures and timelines, the envisaged future of being "the most important information provider" can be achieved and secured.

The National Statistical System in Germany is decentralised, following the federal structure of the country. The main producers of official statistics are the Federal Office of Germany and the 14 Statistical Offices of the Länder. This poses a challenge to strategic planning processes, especially with regard to coordination: due to the work-sharing organisation of statistics production and the political requirements at the federal as well as at the Länder level, strategic planning has to account for various demands. The joint strategy has to reconcile and translate these into a shared vision of the German system of statistical offices. Since 2012, decision making processes and strategic planning procedures have been established among the German statistical offices and the Strategic Priority Plan is set up. The commitment to quality plays an important role. Through comprehensive coordination at strategic, management and operative levels, agreements are settled among the different bodies. In the second round of peer reviews in 2014/2015 these achievements have been acknowledged. In order to further strengthen the compliance with the CoP the peer review team recommended that periodic progress should also be published.

This paper presents the established decision making processes and coordination methods in the form of "management by measures" and delegated accountability. It shows its strengths as well as its limits and addresses the upcoming challenges of drafting a consolidated progress report.

Keywords: institutional environment; governance; decision making; coordination role of the NSI; peer review; implementation of the Code of Practice; cooperation; decentralised statistical system; strategic planning; progress report.

Chair: María Jesús Vinuesa

TRAINING FOR COMPILERS OF STATISTICAL RELEASES

K. Huuhko, E. Seppänen

Statistics Finland, Helsinki, Finland

At Statistics Finland, statistical experts produce releases themselves and publish them online. Annually, around 600 releases are published. There are instructions on the structure of the releases but there is considerable variation in the quality of the releases due to the decentralised editing model. During 2014 and 2015, Statistics Finland organised an extensive training round for statistical experts covering all statistics. The authors were invited to the sessions as small groups and examples were used to review the recommendations for compiling statistics.

Many sub-areas (instructions, graphic design, software versions) related to releases have changed in recent years. New instructions and the reasons for changing them were explained in the training. The training also discussed the different release elements (texts, graphs, tables, database tables) and highlighted their specific roles as conveyors of knowledge. It was also emphasised how the use of correct graph types and presentation methods has a significant effect on the understandability and usability of statistical data. The aim was also to inspire experts to consider their role in conveying information: to summarise the most essential information from vast data masses.

Due to decreasing resources, it is important to utilise API solutions and other automatic reprocessing methods in our statistical production. Our customers will also utilise API solutions more often next to the direct use of the PX-Web user interface. Machine processing and interface solutions rely heavily on standardised metadata of structural database tables so the training strongly emphasised their critical importance in this changing data environment.

The training also gave advice on writing release texts. Special attention was paid to the clarity and understandability of the text, as well as language planning. Because new statistical data are published on the web, the writers must consider the requirements of writing for the web.

Keywords: Statistical releases, Presentation of data, Standardised metadata.



HOW TO TAILOR PRESS WORK IF QUALITY STANDARDS OF OFFICIAL STATISTICS CONFLICT WITH MEDIA INTERESTS

S. Sattelberger

Federal Statistical Office (Destatis), Wiesbaden, Germany

Figures produced by official statistics only matter if they arouse the interest of the public. The media play the most important part in distributing statistical information also to further target groups like political and economic leaders. Journalists prefer to make use of official statistics because they regard our data as reliable and objective high-quality sources. The Code of Practice (CoP) helps to maintain this solid reputation. However, its practical implementation in press work is challenging if quality standards do not correspond to the relevant perspective of the media. This paper provides recommendations how to communicate the CoP to journalists and to address obvious misunderstandings in the mutual relationship.

Using recent examples from everyday press work at Destatis, the paper illustrates how CoP principles (e. g. statistical confidentiality, impartiality and clarity) may conflict with media interests (e. g. disclosures, exclusivity and emotionality). Regular press monitoring helps to identify such critical situations:

- The privacy of respondents may be misinterpreted as covering up political or economic scandals.
- Journalists may see violations of users' equality in accessing statistical releases if competing media quote figures not explicitly published in a press release.
- Adding methodological background information to revised figures may not only be regarded as irrelevant, but also as too complicated and boring for the public.

Proposals are given on how to tackle tricky situations in explaining quality standards. In each case, a tailored communication strategy should be applied depending on the seriousness of the misunderstanding. However, there are no ideal solutions and, due to the diversity of statistical activities and press work, lessons can be learned from each case. Therefore, the conference participants are encouraged in the discussion to share their experiences in meeting similar challenges in press work.

Keywords: press and public relations, users' perspective on quality standards, communicating quality standards to users.

COMMUNICATION AND QUALITY IN OFFICIAL STATISTICS TRANSLATED INTO ENGLISH

P. Collesi, F.Ferrante, F.Scambia
Istat- National Statistical Institute, Italy

The language of statistics has been so far very little explored, although it is topical due to the wide use of statistics in many and varied fields. The paper proposes an analysis on the use of English as a Lingua Franca – ELF [Taviano, Stefania (2010). *Translating English as a Lingua Franca*. Firenze: Le Monnier Università.] in statistical communication and with special reference to press releases translated into English. By means of the English translations of national official statistics all the educated European citizens have access to statistics, which are now freely accessible on the web. This fact is leading to new language strategies [UNECE – *Making Data Meaningful* (2006, 2009, 2011)] and to an increasing use of ELF for statistics intercultural communication. Requirements of accessibility and clarity by the Statistics Code of Practice impel statisticians to improve the way of presenting statistics to an international audience as a quality feature. This paper analyzes key issues of the translation process and language effectiveness in presenting Istat integrated press-releases translated into English. The paper is the result of a new experience Istat is carrying on in order to make press releases simultaneously available and readable to national and international public. In a knowledge-driven society official statistics produced by EU member states are crucial in building the general knowledge that should be available to all members of society beyond national borders. The paper deals with a very specific topic, but which is also part of the ESS Vision 2020 when it refers to a “communication strategy that satisfies divergent and ever-changing user needs at both national and European level”.

Keywords: Communication, quality, English translation.



COMMUNICATING REVISIONS AND THE TIMING OF REVISIONS TO USERS

I. Meinke

Federal Statistical Office (Destatis), Germany

Revisions are part of the regular production process of about one quarter of the statistics published by the German Federal Statistical Office. Some statistics-specific revision practices and revision cycles are already well documented and communicated to users. But this information is not easily available for all statistics which undergo revisions, it is scattered on the webpage or in various documents and the terminology used in the context of revisions is not always consistent. Revisions, when not properly explained and communicated to the users, can damage the credibility of official statistics and limit their usability.

Therefore, the German Federal Statistical Office published in accordance with the European Statistics Code of Practice a revision policy, defining what revisions are and explaining the general revision procedures to the users. Subsequently, a revision calendar was published, providing an overview of all statistics which underlie revisions and describing briefly and concisely their revision cycles by using a standardized structure. The revision calendar contains information on the timing of revisions of approximately 85 statistics. As far as known, the comprehensive extent of the revision calendar is unique within the ESS.

The purpose of this paper is threefold: First, it aims to give an overview of the current situation within the ESS concerning key documents on revisions. According to the Peer Review improvement actions published on the Eurostat website, other NSIs besides Germany were asked to elaborate a revision policy and a revision calendar or to clearly distinguish between revisions and errors. Thus, the second part of the paper is focused on drawing a clear distinction between both terms. Finally, following the Q2014-paper on the development of a revision policy, this paper concentrates on the presentation of the recently elaborated revision calendar, its aim, content and standardized structure as well as its interaction with the revision policy.

Keywords: Communicating revisions to users, clear distinction between revisions and errors, revision calendar and its interaction with the revision policy, current situation within the ESS concerning revisions, implementation of the code of practice.

Chair: Jorge Saralegui

IMPROVED MARITIME STATISTICS WITH BIG DATA

A. Ado¹, A. Isaksson², I. Jansson², M. Justesen², J. Moström, F². Söderbaum¹

¹ Transport Analysis, Sweden

² Statistics Sweden, Sweden

In recent years, Big Data as a potential data source for official statistics has been widely discussed and debated. The high volume, high velocity, and wide variety of these data require new tools and methods for capturing, managing, and processing them efficiently.

Big Data sources have already made their way into official statistics production at Statistics Sweden, i.e. scanner data to replace manually collected data for price statistics, and additional sources are currently being investigated. In a joint project during 2015, Statistics Sweden and Transport Analysis (the government agency responsible for official transportation statistics) have explored the potential of AIS data for improving maritime transport statistics. AIS (Automatic Identification System) is a system that enables a vessel to identify and track the movements of other vessels. This is a first attempt to consider automatically retrieved positional data for the purpose of official statistics.

One goal of the project was to develop methods based on AIS data to improve calculation of a distance matrix between ports. AIS data are already used by Eurostat to calculate routes between ports in Europe, but we take a different approach. We will describe the method and important differences compared to the Eurostat method.

We will describe important quality aspects of the AIS data, our experiences of processing large amounts of positional data, and other lessons learned during the project that will be useful for future work also with similar data sources. Since AIS data bear close resemblance to other position data, such as cell phone data and GPS data, the methods that are developed within this project can be used also in other contexts to calculate for instance different types of transport flows and commuting patterns.

Keywords: AIS; positional data, distance matrix.

DETERMINING PERMANENT RESIDENCY STATUS USING REGISTERS IN ESTONIA

K. Lehto¹, E. Maasing¹, E.-M. Tiit^{1,2}

¹ Statistics Estonia, Estonia

² University of Tartu, Estonia

The next population and housing census in Estonia at the end of 2020 is intended to be register-based. At the moment methodologists are analysing different administrative registers to determine the total population of persons, households and dwellings. This presentation will give an overview of the method and the results of determining the total population of persons. This method is also planned to be used in regular population statistics.

All the data about Estonian citizens and foreigners who have registered their usual residence in Estonia or have got an Estonian residence permit are collected in the Population Register (PR). All persons have a unique personal identification code that is also used by other administrative registers in Estonia. By law, everybody is obliged to ensure that their correct usual residential address is entered in the PR. There have been cases where people who have left Estonia do not have their leaving registered in the PR, or people who have returned to Estonia do not provide the PR with this information. So, the Estonian PR is over-covered for the census. It can be assumed that the people who actually live in Estonia are represented in other administrative registers because they are using services and receive payments. Our main idea is to calculate a residency status R for every person j using the formula:

$$R_{(j)}(k+1) = d R_j(k) + g X_j(k),$$

where

k – a year,

R_j – a value of residency status,

X_j – a count of registers where the person was active during the given year (sum of the signs of life),

d – stability parameter, and

g – signs of life parameter.

The higher the value of residency, the greater the probability of being a permanent resident of Estonia.

The presentation will give an overview on the method and results in Estonia for the years 2012–2015.

Keywords: Population, census, registers.



LATENT CLASS MULTIPLE IMPUTATION FOR MULTIPLY OBSERVED VARIABLES IN A COMBINED DATASET

L. Boeschoten^{1,2}, D. Oberski¹, A.G. de Waal^{1,2}

1 Tilburg University, Tilburg, Netherlands

2 Statistics Netherlands, Den Haag, Netherlands

Combining register data and sample surveys has several benefits and is therefore more and more used by National Statistical Institutes. However, both the registers and samples can contain measurement error. Some measurement errors become visible when the logical relations (edit rules) in the data are investigated. Other errors remain invisible. When the same variable is measured in multiple datasets within a combined dataset, we get an indication of errors that we were unable to detect using the edit rules. We propose a new method based on latent class modelling that estimates the number of measurement errors in the multiple sources, and simultaneously takes the edit rules into account. Furthermore, we use the latent class model to multiply impute a new variable. The new variable can be used in several ways. As a better approximation of the true scores, the new variable can be used to enhance the quality of statistics based on the combined data. The new variable can also be used to assess the quality of the underlying data sources or the quality of the combined data. This method is applied to a combined dataset from Statistics Netherlands, and the performance of this method is investigated by a simulation study.

Keywords: Latent class model; multiple imputation; combined dataset.



INTEGRATION OF MULTIPLE LISTS: OVERVIEW AND PROPOSAL FOR LINKAGE ERROR ADJUSTED ESTIMATORS

T. Tuoto¹, L. Di Consiglio²

¹ Italian National Statistical Institute, Rome, Italy

² Eurostat, European Commission, Luxembourg

The increasing amount of data, represented both by the administrative registers and more recently by big data, is an enormous potential for enlarging and detailing the statistical information. The integration and combination of those already existing data with traditional statistical surveys data is a pressing challenge for National Statistical Institutes. Micro-level integration of sources is a standard practice, generally performed by means of record linkage techniques. However, as a matter of fact, the linkage process is not completely error free and statisticians have to take linkage errors into account in the subsequent analyses performed on integrated data [1,2]. Linkage errors appear particularly relevant when the target is to measure the size of a population enumerated in different lists [3]. In this context, a widespread method is based on the capture-recapture model, multi-capture when dealing with more than two lists [4,5,6]. The standard capture-recapture model assumes no linkage errors, and hence their occurrence may strongly affect the bias of the resulting population count estimation. Nevertheless, few contributions [7] have already dealt with this issue, even now that in many countries the integration of several lists is proposed as an alternative to the traditional census. In this work, we extend [3] that provides population size estimation in presence of errors when linking two lists. In the context of multiple recaptures, this study explores adjustments for linkage errors in population size estimators following the approach of [8], analysing their impact on accuracy.

[1] Neter, J., Maynes, E.S. and Ramanathan, R. (1965). The effect of mismatching on the measurement of response error. *Journal of the American Statistical Association*, 60, 1005-1027.

[2] Chambers, R. (2009). *Regression Analysis Of Probability-Linked Data*. Official Statistics Research Series 4.

[3] Di Consiglio L. and Tuoto T. (2015) Coverage Evaluation on Probabilistically Linked Data, *Journal of Official Statistics*, Vol. 31, No. 3, 2015, pp. 415–429

[4] Cormack R. M. (1989) Log-linear models for capture-recapture. *Biometrics*, 45: 395–413.

[5] Fienberg S. E. (1972). The multiple recapture census for closed populations and incomplete 2k contingency tables. *Biometrika*, 59:409-439.

[6] International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology*, 142:1047-1058.

[7] Lee A. J., Seber G. A. F., Holden J. K., and Huakau J. T. (2001) Capture-Recapture, Epidemiology, and List Mismatches: Several Lists, *Biometrics*, 57: 707-713

[8] Ding, Y. and Fienberg, S.E. (1994). Multiple sample estimation of population and census undercount in the presence of matching error. In *Proceedings of 1994 Annual research conference and CASIC technologies Interchange*, Bureau of Census, Virginia, US

Keywords: Capture-recapture, log-linear model, linkage errors.



EXPLORING THE STATISTICAL MATCHING POSSIBILITIES FOR THE EUROPEAN QUALITY OF LIFE SURVEY

I. Rioboo, T. Leonicikas

European Foundation for the Improvement of Living and Working Conditions (Eurofound). Dublin, Ireland

The increasing growth of collected data together with the continuous research interest on the interrelationships of several socio-economic aspects, calls for robust methodologies to combine different sources of information.

One alternative is the statistical matching which aims to link specific variables from independent sample surveys referred to the same target population, using information shared between them as a link.

Despite this is a challenging problem for statisticians, the potential benefits of matching several surveys lie in the possibility to expand the analysis incorporating different angles without increasing response burden or costs.

In this work we explore several matching alternatives that could be used for the European Quality of Life Survey (EQLS) with other related surveys, discussing specific methodological issues related to all the steps of the process.

Keywords: EQLS, statistical matching, combined analysis.



Chair: Maria Dologova

INTRODUCING A FRAMEWORK FOR PROCESS QUALITY IN NATIONAL STATISTICAL INSTITUTES

F. Barbalace, G. Brancato, M. Signore, G. Simeoni
Italian National Statistical Institute (Istat)

One of the key areas of the ESS Vision2020 is the promotion of the efficiency in production processes. Such an objective calls for strengthening and enhancing the quality framework on statistical processes. At the beginning of 2000, the Leadership Expert Group (Leg) on Quality had already acknowledged the importance of quality in statistical processes by defining some characteristics that processes should have. Notably, one area of the ES Code of Practice focuses on statistical processes and introduces the principles they should follow for ensuring the quality of European statistics. However, whereas in the area of statistical output the quality framework is highly structured, this does not appear to be the case for the statistical processes. The aim of this paper is to contribute to the establishment of a framework for process quality in National Statistical Institutes. Processes are first defined and classified according to their different nature, e.g. statistical vs. organizational ones, taking into account the GAMS0 (Generic Activity Model for Statistical Organizations) model recently endorsed by the UNECE High-Level Group for the Modernisation of Statistical Production and Services, including the GSBPM (Generic Statistical Business Process Model) model for statistical processes. Then, the desirable quality dimensions are defined, by analysing those introduced by the Leg on Quality (efficiency, effectiveness, robustness, flexibility, transparency and integration), assessing their suitability to be applied to the different types of processes of a statistical organisation and adding new relevant characteristics, such as “identified”, which means that the process is mapped, the responsibilities are assigned and the procedures are defined, and “controllable”, meaning that it can be governed or ruled so that the errors and their spread can be prevented.

The framework proposed represents a precondition for the definition of quality indicators and assessment methods for process quality at different levels of a statistical organisation.

Keywords: Process quality.



MINDING THE STORE: AN INTERNAL AUDIT PROGRAM FOR DEMOGRAPHIC PROGRAMS AT THE U.S. CENSUS BUREAU

R. Levy, J. Scott

U.S. Census Bureau, Washington, DC, U.S.A.

The U.S. Census Bureau is a leader in collecting data and disseminating information products on people, businesses, and governments. These information products range from micro data sets and table packages to methodological papers and analytical reports.

To ensure the information products are at an acceptable level of quality before dissemination, the Bureau has developed and codified a set of statistical quality standards. These standards govern the processes and procedures for preparing the products. And, to ensure that U.S. Census Bureau programs are following the standards, the agency has developed several internal audit programs. These audit programs periodically review the various programs to gauge compliance with the standards.

The audit program which targets information products pertaining to demographic data was chartered in 2011. As of November 2015, thirty programs have been audited. This paper will review the audit program by patterns of noncompliance with the quality standards and audit burden, both overall and by type of program. The paper will also discuss the implication of these patterns and future directions for the audit program and quality standards.

Initial analysis of audit program data indicates that demographic program areas are most commonly not compliant with the following standards: Completing statistical reviews; conducting policy and sensitivity reviews; developing timetables for reviewing files; and following policies for archiving information products. Also, programs that include data collection tend to have more noncompliant items than other types of programs. Programs average five noncompliant items, out of about two hundred and seventy ratable items.

The importance of this research lies in alerting stakeholders about areas where staff at statistical agencies may need more support and guidance in producing data and information products of high quality.

STATISTICAL QUALITY BY DESIGN: CERTIFICATION, CULTURE AND MANAGEMENT

Kees Zeelenberg
Statistics Netherlands

We argue that to achieve and to communicate quality of official statistics, it is essential that national statistical institutes adopt some system of quality by design, i.e. formal quality certification, e.g. ISO or EFQM, as well as work processes and guidelines based on Total Quality Management (TQM) and plan-do-check-act cycles. We discuss three major elements of statistical quality and show how they are being implemented at Statistics Netherlands.

First, for external users quality is difficult to judge. In essence this means that users will have to trust official statistical information. To maintain this public confidence in official statistics in this age of extreme public openness and public criticism, it is mandatory that this trust is somehow corroborated. We argue that certification of the statistical production processes and outputs by means of an independent standard, such as ISO or EFQM, is necessary for this corroboration.

Secondly, it is not efficient nor would it be acceptable for staff members, if we were to check every action of every statistician in every statistical process. Simply put, we cannot have an auditor standing behind each statistician. So to achieve quality we have to rely to a very large extent on the statistical conscience and statistical expertise of statisticians and managers. We argue that such a quality culture comes not by itself, but needs to be achieved by embodying it in management and work guidelines, based for example on Lean Operational Management (LOM), and needs to be supplemented by additional means such as quality sessions.

Thirdly, the proof of the pudding is in the eating, and so we must ensure that quality is actually embodied in the statistical products that are being disseminated. We argue that for this we need a variety of tools, such as implementation of statistical quality from the start in development of statistical processes, quality guidelines, and quality indicators, supplemented by a plan-do-check-act (PDCA) cycle that is an essential element of various quality systems such as TQM and Lean Six Sigma [1].

Keywords: statistical quality, certification, ISO, EFQM, LOM, TQM, Lean Six Sigma.

[1] Marret Smekens and Kees Zeelenberg, 2015, Lean Six Sigma at Statistics Netherlands, *Statistical Journal of the IAOS* 31(4), 583-586, doi: 10.3233/SJI-150930



ASSESSING QUALITY CONTROL: EVALUATING THE QUALITY AUDITS

Justin D. Nguyen, Carma R. Hogue
U.S Census Bureau, Washington D.C.

Effective quality control systems are the foundation for successful manufacturers and data publishers. For years, statisticians have proposed theories and findings to improve data product quality. The Economic Directorate of the U.S. Census Bureau collects various economic data with a requirement to accurately capture and analyze our data conduct quality audits to ensure programs correctly identified problems to save time and money and ensure quality. Our primary objective is to assess our program areas' compliance with best practices, particular in the area of publications, to ensure that statistically sound practices are used in the collection of data and in the presentation of results to the public. These data have gone through rigorous quality control procedures to assure the highest possible quality and consistency. In this paper, we discuss the first two rounds of quality audit processes that: i) identified inspection goals and inspection plans, ii) deterred and detected data quality, iii) communicated quality expectations and provided recommendations based on the audit results. In the remainder of the paper we discuss measures to correct systems challenges and introduce quality assurance measures through the survey life cycle.

Keywords: U.S Census Bureau, quality control, integrated project models.



A NEW PROCESS FOR ASSESSING THE QUALITY OF AN OUTPUT MAKES BETTER QUALITY STATISTICS

C. Davies

Office for National Statistics, Newport, United Kingdom

There is a requirement under the UK Code of Practice for Official Statistics to ensure official statistics are produced to a level of quality that meets users' needs and to seek to achieve continuous improvement in statistical processes by undertaking regular reviews.

In January 2015 the Quality Centre rolled out a new process for assessing the quality of an output within the Office for National Statistics (ONS). This new process is called a 'Regular Quality Review (RQR)' and replaces the use of a self assessment tool to carry out quality reviews. The decision to develop a new process was made based on feedback from statistical output managers during a user engagement exercise.

The RQR process reduces the burden on the data producer, as instead of completing a large self assessment tool they discuss their output in the context of the five quality dimensions and the Generic Statistical Business Process Model with a methodologist. The process also results in recommendations tailored to improving the quality of the statistical output.

To date 40 RQRs have been successfully completed. Customers feel the new process is proportionate and the recommendations useful. Quality Centre monitors recommendations and ensure they are implemented.

In this paper, we describe the new process, present feedback from business areas, discuss the common themes around recommendations and show how outputs have improved as a result of the RQRs.

Keywords: Improvements, reviews, assessments.

Chair: Agustín Cañada

ESS STANDARD QUALITY REPORTING IMPLEMENTATION: THE POINT OF VIEW OF A NATIONAL STATISTICAL INSTITUTE

G. Simeoni

Istat - Italian National Statistical Institute, Italy

Importance of quality reporting has always been acknowledged in the European Statistical System (ESS). Most European regulations includes requirements on quality information to be provided. However, content and level of detail of required information are extremely different from one regulation to another one.

The ES Code of Practice (CoP) and the Regulation n. 223/2009 fostered harmonisation of quality reports (QRs) in the ESS and promoted their dissemination to users.

Cross-domain harmonisation is desirable for several reasons, e.g. to obtain an acceptable minimum level of documentation for all European Statistics, to improve comparability, to provide users with clear and structured reference metadata.

However, harmonisation of ESS QRs implies several steps to be performed: the definition of conceptual standards and their promotion among domains, the development of technical tools to support implementation, a strong coordination of implementation in statistical domains.

Over the last years great efforts have been spent to improve standard quality reporting at ESS level. Istat has been particularly involved in these initiatives with Eurostat and other NSIs. The paper presents an overview on the activities already performed and on some further planned initiatives.

Efforts at ESS level can not be sufficient. Similar efforts should be spent at national level. Obviously, the actions to be performed can differ from NSI to NSI considering the existing background. The paper focuses on the Istat experience on quality reporting: what already done - mainly software enhancements - and what still to be performed to improve coordination among domains. Finally, a recognition of how QRs are disseminated by NSIs will be reported, as a basis to define how to proceed at Istat. Indeed, national dissemination of ESS QRs is foreseen for 2016 as an improvement action deriving from the recommendation of the 2013-2015 CoP Peer review reports.

Keywords: quality reporting.

HOW TO FULFIL ALL DEMANDS FOR QUALITY REPORTING

Kirsten Wismer and Karin Blix
Statistics Denmark

At Statistics Denmark we have implemented the DDI based software Colectica to organise metadata. The idea was that all metadata should be organised using international standards and metadata should be gathered in one place and be reused. For all published statistics all the fields of SIMS (Single Integrated Metadata Structure) are filled in and stored in Colectica. On our website the user oriented SIMS fields are shown and presented in layers – the overall picture first – and more details for the more dedicated users.

An important feature of our new metadata system is the reuse of data. All ESMS fields can “automatically” be transferred from Colectica to the ESS Metadata Handler for those statistical domains that have chosen this way of quality reporting. For the MIP’s e.g. a generic template for quality reporting has been developed. The MIP’s are also reported by institutions that are not part the ESS. They do not use SIMS, but a compromise was reached so that all the fields in the MIP quality reporting template where mapped to SIMS. Hence for reporting quality on MIP, an extraction from the SIMS fields entered in Colectica can be used.

In their meeting in November 2015 the ESSC endorsed the revised version of SIMS and stated that SIMS will be the reporting standard for quality reporting according to Article 12 of Regulation 223/2009 on European statistics. We foresee that all statistical domains in the ESS in the near future will adapt to using the ESS Metadata handler and that SIMS will be the basis for all quality reporting. In the paper we give examples of the entering of information in the SIMS fields and the different quality reports that can be extracted from the system. Examples of the integration of metadata will also be presented.

Keywords: Quality reports, metadata, SIMS.

QUALITY DATA SHEETS

T. Tümmler

Federal Statistical Office (Destatis), Wiesbaden, Hesse, Germany

Measuring quality is one of the major challenges of quality management in statistics. The European standard structures for quality and metadata reporting increasingly require quantitative quality indicators to underpin qualitative assessment of quality of statistics provided in quality reports.

In a heavily decentralised national statistical system where most statistics are produced in collaboration by 15 independent statistical institutes it is difficult to acquire information from the statistical production process that is essential in order to calculate quality indicators like item response rates, imputation rates and other process indicators. In Germany, the field work is generally carried out by 14 independent state statistical offices that transmit aggregated results on state level to the Federal Statistical Office. The Federal Statistical Office provides methodology, publishes statistical results for Germany as well as relevant metadata like quality reports. So the statistical production process is segmented by two dimensions, geography (each state statistical office is responsible for the field work in its state) and process step (field work by the state statistical offices, methodology, publishing and quality reporting by the Federal Statistical Office).

Attaining the data that is necessary to generate quality indicators and other information on the statistical production process cannot be solved by organisational means alone. In order to keep the workload for producing such data at a minimum level, IT-systems need to fully support compilation, analysis and management of these data.

This paper describes the general idea and the goals of Quality Data Sheets as well as projected costs and benefits regarding the implementation. Finally, the vision for the extension of Quality Data Sheets for purposes of process and resources management will be illustrated.

Keywords: Implementing of quality reporting, quality indicators, decentralised statistical system.



ESMS IMPLEMENTATION IN STATISTICS ESTONIA

Remi Prual

Statistics Estonia, Estonia

Statistics Estonia (SE) decided to implement the Euro SDMX Metadata Structure (ESMS) for better horizontal and vertical integration across processes and organisations. We developed and adopted regular (yearly) cycles that help SE to describe, harmonize and improve reference metadata on statistical activities according to standardised requirements (ESS metadata standards) and make that metadata publicly available on SE's website. Metadata are updated regularly (yearly), in accordance with the compilation of the official statistical programme, which is a list of demographic, social, economic and environmental statistical activities compiled each year for the following five years and approved by the Government of the Republic.

For now we can say that all ESMS metadata have been reviewed and harmonised as much as possible based on the guidelines. Reference metadata are disseminated (automatically and directly from the metadata system) on SE's website with the use of new systems.

The use of the ESMS has raised awareness of statistical production and processes in Statistics Estonia. Survey managers and other employees understand more easily that processes are often very similar or even the same for other statistical activities. The ESMS has also been well recognised as a very useful set of concepts for describing and sharing know-how about statistical activities. It is quite actively used for analysing the efficiency and similarities/differences between statistical activities and for making other analyses about the work programme. ESMS-based metadata are additionally used in metadata and quality reports that are prepared and sent to Eurostat (using Metadata Handler) and other organisations.

Keywords: Standards, quality reporting, metadata standardisation.

Chair: Martina Hahn

QUALITY ASSESSMENT OF MULTI-SOURCE STATISTICAL PROCESSES

G. Brancato, A. Boggia, A. Bruni

Italian National Statistical Institute, Istat

Quality assessment is a key activity in a National Statistical Institute. Istat has a long standing experience on quality assessment, which is one of the pillars of its quality policy. Istat quality assessment strategy is based on auditing and self-assessment procedures, supported by: Quality Guidelines, auditing and self-assessment questionnaires, other documentation useful for the assessment, such as product and process quality reports, and final assessment reports defining improvement actions and identifying best practices. So far, the assessment procedures and the tools have been developed for survey or census based statistics, and partially explored for statistical processes using administrative data. The on-going Istat modernisation process has conveyed to a new paradigm in statistical production, with a prominent role played by the extensive use of administrative sources. The new scenario calls for rethinking the quality model and assessment tools. A tailored quality framework has been developed, harmonised with the international literature, and quality guidelines, stating the principles and methodologies to be followed, have been issued. They are published on Istat website, in Italian. Currently, the suitable assessment instruments in the multi-source context are being studied and developed.

Aim of this paper is to report on the approach, the difficulties encountered and the solutions that are being adopted for the quality assessment of statistical processes based entirely or partially on administrative data, taking into account for the variety of usages of administrative data and the complexity related to different outputs that can be produced, ranging from statistics to be disseminated up to statistical registers. The focus will be on the design of the assessment questionnaires, and how they are developed in order to gain evidence on input, process and output quality as well as on the sources of errors.

Keywords: quality assessment, auditing questionnaire, multi-source statistics.



SPAIN 2021. WHY THIS CENSUS WILL HAVE MORE QUALITY THAN THE PREVIOUS ONE?

J. Vega, A. Argüeso

National Statistical Institute, Spain.

One of the main reasons that is motivating most of developed countries to a change in their census methodology is to carry out a higher quality operation. In order to succeed in this task and taking advantage of the current situation, Spain is considering to increase strongly the usage of administrative registers in 2021 Census project.

This paper will focus on analysing the strengths and weaknesses, from the point of view of quality, of the new strategy as opposed to the previous one.

Apart from that, an analysis from the point of view of each one of the six classical aspects that form the Eurostat quality concept (relevance, accuracy, timeliness and punctuality, accessibility and clarity, comparability and coherence) about how next Census will change from the previous one will also be included.

Keywords: Census, administrative registers, quality.



DESIGNING THE INTEGRATION OF REGISTER AND SURVEY DATA IN EARNING STATISTICS

C. Baldi, M. Ciarallo, S. De Santis, S. Pacini, C. Congia, C. Casciano
National Statistical Institute (Istat), Italy

This paper addresses the redesign of the Structure of Earning Survey 2014 made possible by the availability of the new employer-employee register on wages, hours and labour costs. The redesign has aimed on one side to reduce the burden on the enterprises and improve the quality of the data and on the other side to collect the data not available in the register and to correct the definition of those available in the register. Three main areas will be covered: the survey sampling, the questionnaire redesign and the development of regression models to mass-impute and correct the variables. As for the sampling the availability of the register information has allowed to dramatically reduce the sample size without lowering the target quality measures, while at the same time producing a survey structure more adequate to distributional analysis and micro econometrics. The complexity of the questionnaire has been reduced significantly and its content redefined in a way to be a bridge between the administrative and statistical definitions. Moreover the prefilling of key data on wages and hours should reduce to a large extent the measurement errors. Finally, an in depth analysis of the content of the Social Security data on which the register is based is allowing to specify a new class of regression models to impute the missing data on the register.

Keywords: Response Burden, Register based statistics, integration.

CORRECTION FOR LINKAGE ERROR IN POPULATION SIZE ESTIMATION

B.F.M. Bakker^{1,2}, L. Di Consiglio³, D.J. van der Laan¹, T. Tuoto³, P.-P. de Wolf¹, D. Zult¹

¹ Statistics Netherlands, The Hague, Netherlands

² VU University, Amsterdam, Netherlands

³ ISTAT, Rome, Italy

The availability of data coming from new sources like administrative registers and big data, make it possible to build lists of statistical units (enterprise, resident population,...) as a sampling frame, as a backbone to link all the relevant sources to or for the information itself. Ideally, such lists cover the entire population. To investigate the quality of these lists, the size of the target population has to be measured including the population size that has been missed in all lists. A widespread method is based on the capture-recapture model also called multi-capture when more than two lists are used. They make several assumptions, such as independence of inclusion probabilities, closed population, no erroneous captures and perfect record linkage, which in practice cannot always be met, particularly when data are not collected for statistical purposes.

Recent studies have evaluated the true usual resident population in the Netherlands using multi-capture by combining three administrative registers: the population register, the employment register and the crime suspects registers [1,2]. The sensitivity of the results to violation of the independence, perfect linkage and no erroneous captures assumptions is quite large [3,4]. For the independence assumption and the assumption that there are no erroneous captures they found reasonable solutions, but for the linkage error they did not. Extending their work, we propose to correct for the violation of the assumption of no linkage error in the multi-capture model applied by applying the method described in [5].

References:

[1] Bakker B. F. M., Gerritse., S. C., van der Heijden P. G. M., van der Laan D. J., van der Vliet H. N., and Cruyff M. J. L. F. (2014) Estimation of non-registered usual residents in The Netherlands, Conference of European Statistics Stakeholders, Rome, Italy, 2014.

[2] Gerritse, Susanna C., Bart F. M. Bakker & Peter G. M. van der Heijden, 2015b, Different methods to complete datasets used for capture-recapture estimation, In: Statistical Journal of the IAOS, vol. 31, nr. 3

[3] Gerritse Susanna C., Van der Heijden Peter G.M, and Bakker Bart F.M. (2105a) Sensitivity of Population Size Estimation for Violating Parametric Assumptions in Log-linear Models, Journal of Official Statistics, Vol. 31, No. 3, 2015, pp. 357–379

[4] Gerritse, Susanna C., Bart F. M. Bakker, Daan B. Zult & Peter G. M. van der Heijden, 2016, The effects of linkage errors and erroneous captures on the population size estimation (submitted)

[5] Di Consiglio and Tuoto (2015) Coverage Evaluation on Probabilistically Linked Data, Journal of Official Statistics, Vol. 31, No. 3, 2015, pp. 415–429

Keywords: under coverage, capture-recapture method, linkage error correction.



IMPROVING THE STATISTICAL PROCESS IN THE HOTEL OCCUPANCY SURVEY

Elena Rosa-Perez

National Statistical Institute (INE), Spain

Tourism is one of the most important industries in Spain, fostering the social-economic development of many regions. Therefore, an appropriate tourism statistics system is needed. From the supply side there are several surveys that give a complete picture of the situation. One of them is the Hotel Occupancy Survey that monthly disseminates data about travelers and overnight stays, number of establishments and bed-places and indicators like the Average Daily Rate and the Revenue per Available Room.

This survey benefits from the Tourism Registers to increase the quality of the survey frame. The Registers provide information about new establishments and modifications in the identification variables of the existing hotels. One of these variables is the classification in stars and it is used to stratify the directory and to sample, another variable is the number of bed-places, used as an auxiliary variable to estimate the number of travelers and overnight stays.

Another way to improve the quality of the survey is by using new data collection methods. Like an e-questionnaire that makes possible to include workflows and different types of edits during the collection and a selective editing phase once the questionnaire has been filled in. Another one is the electronic data reporting, in particular, XML files obtained with an application directly from the hotel management system and that contain all the variables requested in the questionnaire. These XML files provide more detailed data than the questionnaire, with a greater level of breakdown, which makes possible the reduction of the bias and more accurate estimations.

Besides, the Hotel Occupancy Survey will face new challenges in the future, like the possible use of Big Data or the standardization of statistical processes that will enhance undoubtedly the quality of the data.

Keywords: tourism, administrative register, multi-mode data collection.



Chair: Eduardo Barredo

DEVELOPMENT OF EXPERTISE AND MULTI-SKILLNESS AT STATISTICS FINLAND

P. Koivuniemi, S. Tolkki
Statistics Finland, Helsinki, Finland

The objective is to describe the professional identity groups of statistical experts, on guidance to multi-skilled small groups, where the basis is the development of expertise in line with data scientist thinking according to the ESS vision. We will also illustrate the steering process of team leaders, where the aim was to support statistical expertise in its various roles.

The Training Programme in Statistical Skills includes one period, which focuses on the professional identity of statistical experts. It is carried out in small groups led by supervisors. The themes of the meetings are: Me and statistical work, Statistical professionalism, Competence and learning, Characteristics of expertise, Organisational culture, and Changing statistical work. The aim of supervision is to reinforce the experiences of meaningfulness of one's work, which ensures motivation and thus guarantees the high quality of work.

In today's statistical work, expertise is needed in a wide variety of areas (cf. data scientist). It is hard to find persons who would have sufficiently extensive competence – that is why we need cooperation and teams. Means of target-oriented supervision can be used to promote cooperation of experts and solve topical challenges. The acute topics are: Promoting the coverage of statistical information in the social media, Utilisation of big data and Production of new services.

One of the roles of a statistician is a team leader, where expert work expands to the management of the whole statistical process and to supervisory tasks. The role has been administered in groups led by a solution-focused supervisor. The themes discussed were the position of teams in the organisation, key tasks and responsibilities of team leaders, cooperation between team leaders and supervisors, functionality of team work, and different kinds of teams and their challenges.

Keywords: Professional identity, target oriented supervision, multi-skillness.



QUALITY AT ANY COST? - EXAMINING THE HARD REALITY OF JOB REDUCTIONS WHILE PROMOTING THE QUALITY AGENDA

B J Best

Office for National Statistics (UK)

In hard times where difficult decisions have to be made, how can we detach personal/ethical considerations from purely evidence based decision making? Numerical data regarding resource requirements clearly identify potential savings and reductions. Can these be translated into hard cash without regard to human feelings and while still encouraging quality data processing and workforce loyalty?

Quality at any cost? - Examining the hard reality of job reductions while promoting the quality agenda. How administration grade staff are protected, exploited and deployed in a competitive, shrinking statistical marketplace.

In an operational business area where the majority of our people are on the basic pay grades, we are required to make efficiency savings while attempting to maintain quality and fulfil our business commitments. Our customers are internal to the department and we have a plethora of management information that suggests that a reduced workforce could still achieve acceptable levels of survey processing. The ethical considerations revolve around the evidence which suggests we could make savings balanced against the moral rights of reducing numbers and potentially having a negative impact on morale, goodwill and the willingness of staff to be innovative and flexible. We rely on the commitment of our staff and this is central to our quality agenda. Without the buy-in of our workforce, new ways of working which enhance quality and promote curiosity would justifiably be perceived as purely cost cutting and would undoubtedly be viewed with suspicion and scepticism.

This paper sets out to put into context, the dilemmas faced by operational managers operating in the real world and juxtaposes quality and cuts.

Keywords: Ethics, Quality, Resourcing.



ENHANCING QUALITY AND EFFICIENCY THROUGH PARTICIPATORY PROCESSES: AIMING FOR THE ESS VISION 2020

D. Roll-Hansen, G. Olsen, I. M. Møller
Statistics Norway, Oslo, Norway

Statistics Norway is working to improve our performance: Streamlining our work with a user perspective in mind, aiming to get more out of limited resources. One of our prime goals is to enhance quality as defined in the European Statistics Code of Practice.

The core of this is to systematically assess and improve work processes and operational control. The overall aim is to strengthen a culture where employees are encouraged to be curious about mistakes and eager to learn from them.

Goals are set by the management and all staff are participating in developing and continuously improving work processes. This is important for the motivation and productivity of staff members, as employees...

- Increase output as a result of more involvement, commitment and affection, resulting from a high degree of autonomy and control.
- Work smarter, they are encouraged to develop their skills and competence in general.
- Work more responsible, when more responsibility is in their hands.

Managers have challenging positions within this paradigm. They need to give staff the autonomy necessary to achieve motivation and high-performance. We will present some recent examples of how we work in Statistics Norway to develop by...

- Systematically going through work processes,
- Facilitating a participatory process,
- Addressing managerial challenges
- Training for lean leaders and
- Keep making improvements.

We hope the article will provide inspiration on...

- How participatory measures to enhance quality and efficiency can be implemented,
- Work to strengthen a culture for continuous improvement,
- Potential gains and
- Pitfalls to avoid.

We are looking forward to discussing the above issues with our European colleagues, learning from their experiences and perspectives.

Keywords: Quality, Participatory Processes, Lean.



CREATING FLOW AND REDUCING CYCLE TIME BY REORGANIZING TEAMS AROUND COMPONENTS IN STATISTICS ESTONIA

Allan Randlepp, Eveli Hübner
Statistics Estonia, Estonia.

There is a constant need for bringing release time closer to the accounting period. To do that, the cycle time of the statistical production process must be reduced. In Statistics Estonia, we have chosen to introduce Lean tools and Agile methods for defining the value stream and letting value flow better, based on customer demands.

Our objectives:

- To reduce work in progress through a systematically arranged management system with defined components (collection instrument, data processing package, cleansed dataset, statistical output etc), component owners.
- To improve work flow in the value stream and information flow between employees by forming teams around the components. Each team has all core competences and tools to produce a certain component and has appointed an Agile coach who helps team to continuously improve.
- To gain better coherence between strategic, tactical and operational planning by entering all plans into JIRA and introducing Kanban boards to visualize the work flow.

Keywords: Statistical production, modernisation, resource management.



TOWARDS A GLOBAL EDUCATION IN OFFICIAL STATISTICS

Javier Montero¹, José M. Robles¹, Francisco Veira²

¹ Complutense University, Madrid, Spain

² National Statistical Institute (INE), Spain

In this paper we want to stress the relevance of designing a global educational strategy towards official statistical, whose quality should be based upon the simultaneous achievement of several objectives. The first one is the support of a network of qualified statisticians, grown from common quality standards as pretended by EUROSTAT [1] with its EMOS (European Masters in Official Statistics) program, and ready to address the new paradigms and opportunities that come along the recent exponential growth of available data. Such EMOS program should be develop by selected Universities within a collaborative network, somehow supervised by National Institutes of Statistics, but also in collaboration with other private and state institutions that produce relevant social and economic data. Moreover, at the same time that human and updated technological resources are properly assured, a specific effort is needed in continuous education, in such a way that professionals in official statistics can periodically update their knowledge, by sharing problems and solutions with similar institutions and other research groups. But technical quality should be accompanied with specific social programs to assure the right influence in Society. This objective implies a strategy with secondary education (particularly to prepare teachers and give them support not only to educate in statistical concepts but to educate students in learning from data, see e.g., the interesting initiative [3]). Finally, a media plan is also needed, to be developed in collaboration with scientific associations, in such a way that official statistics gets closer to Society, both in appreciation (as a reliable direct source of knowledge) and understanding (and avoid the tricky world described in [4]).

Keywords: Statistics in Secondary Education, Undergraduate courses in Science, European Masters in Official Statistics, Continuum Learning, Data Scientist in Society and Governmental Institutions.

References:

[1] <http://ec.europa.eu/eurostat/web/europe-2020-indicators/europe-2020-strategy>

[2] <http://www.cros-portal.eu/content/emos>

[3] http://www.ine.es/explica/olimpiada2015_inicio.htm

[4] J.A. Paulos: Innumeracy, mathematical illiteracy and its consequences (Hill & Wang Pub., 2001)

Chair: David Salgado

ASSESSMENT OF RISKS IN THE USE OF BIG DATA SOURCES FOR PRODUCING OFFICIAL STATISTICS - RESULTS OF A STAKEHOLDER SURVEY

Albrecht Wirthmann, Martin Karlberg, Bogomil Kovachev, Fernando Reis, Loredana di Consiglio Eurostat, European Commission, Luxembourg

An increasing number of statistical offices are exploring the use of Big Data sources for the production of official statistics. For the time being there are only a few examples where these sources have been fully integrated into the actual statistics production. Consequently, the full extent of implications caused by their integration is not yet known.

A first attempt to identify and structure risks related to using Big Data sources in the exploration and production phases of official statistics was made in the paper “Structuring risks and solutions in the use of big data sources for producing official statistics – Analysis based on a risk and quality framework” [1]. The paper follows a systematic approach of defining risks in the context of the suggested quality framework. The main conclusion from the paper is that it is impossible to establish a single likelihood or impact for a given “big data risk” – typically, both measures depend heavily on the utilised Big Data Source as well as on the type of statistical product. In order to gain more insight, a source-specific survey of the identified risks has therefore been conducted among stakeholders. The respondents were asked to quantify likelihood and impact of risks for a Big Data source of their choice (among a set list of eight Big Data sources), to provide a rationale for their assessments, and to suggest measures for prevention and mitigation of the identified risks. In order to be more complete, the respondents were also invited to identify additional risks in the exploration and use of Big Data sources for official statistics.

The paper will analyse and present the results of the stakeholder survey, contrasting the findings to the analysis of [1].

Keywords: Big Data, Risk assessment.

References:

[1] Wirthmann A, Karlberg, M., Kovachev B., Reis F., (2015), “Structuring risks and solutions in the use of big data sources for producing official statistics – Analysis based on a risk and quality framework”, http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2015/mtg1/WP18-Wirthmann_AD.pdf



USING HUGE AMOUNTS OF ROAD SENSOR DATA FOR OFFICIAL STATISTICS

M.J.H. Puts, M. Tennekes, P.J.H. Daas, C. de Blois
Statistics Netherlands, Heerlen, the Netherlands

On the Dutch road network, about 60,000 road sensors are located of which 20,000 sensors are on the Dutch highways. Both vehicle counts and average speed are collected and stored in the National Data Warehouse for Traffic Information. Only vehicle counts were used in this study. To enable the production of official traffic statistics four challenges needed to be solved. The first was processing huge amounts of data. The dataset studied contained all vehicle counts collected during every minute of the day by sensors on the Dutch highways from 2010 until 2014; 80TB of data in total. A highly efficient pre-processing step was implemented that selected the essential records and fields and transformed and stored the data in the most efficient way. The second challenge was checking and improving data quality as quite some sensors lacked data for many minutes during the day. A cleaning and estimation step was developed that enabled a precise and accurate estimate of the number of vehicles actually passing the sensors. To monitor the stream of incoming and outgoing data and control this fully automatic statistical process, quality indicators were defined on the 'raw' and processed sensor data. The next challenge was to determine calibration weights based on the geographic locations of the road sensors on the roads. This was needed because road sensors are not uniformly distribution over the road network. As the number of active sensors fluctuates over time, the weights need to be determined periodically. The last methodological challenge was related to the accurate estimation of the traffic intensity over time. Here, a time series approach was used that coped with fluctuations in the amount of data available. As a result of these steps highly accurate numbers could be produced on the traffic intensity during various periods on regions in the Netherlands.

Keywords: Big Data, methodological challenges, processing data.



MEASURING QUALITY OF STATISTICS BASED ON MOBILE PHONE DATA - THE CASE OF POPULATION, TOURISM AND BALANCE OF PAYMENTS STATISTICS IN BELGIUM

Freddy De Meersman¹, Christophe Demunter², Fernando Reis², Marc Debusschere³

1 Proximus, Brussels, Belgium

2 Eurostat, European Commission, Luxembourg

3 Statistics Belgium, Brussels, Belgium

Mobile phone operators' data has now been used for some years in research in several domains. In more recent years official statistics has also been considering seriously the potential of this big data source. One of the main concerns about the use of data which collection was not specifically designed to produce statistics is their potential bias. One way to assess the accuracy of the statistics produced with a new data source is comparing the results to the ones obtained with the traditional methods and with other alternative sources. In this case we compare statistics on population, tourism (visits) and balance of payments produced with several data sources available at a mobile operator in Belgium with the official population and tourism statistics, with the purpose to assess their accuracy. However, accuracy is only one of the dimensions of quality of statistics. Therefore, we also assess the quality of mobile phone operator data in terms of: relevance, in particular its potential in the production of new statistics not possible with the existing traditional methods; timeliness and punctuality; comparability, in particular between countries; coherence; and accessibility and clarity.

Keywords: Big data, Mobile phone data, Quality.

FUZZY SENTIMENT ANALYSIS USING SPANISH TWEETS

P. Rey del Castillo

Instituto de Estudios Fiscales, Madrid, Spain

Opinion mining and sentiment analysis are two topics with growing interest in artificial intelligence. Last years' research on these areas has evolved in increasing complexity and sophistication with tasks such as opinion retrieval, sentiment extraction, classification and summarization being carried out using text mining and natural language processing techniques [1].

Possible strategies go from concept-based approaches (using web ontologies or semantic networks) and data mining (passing a labeled training corpus of texts to a machine learning algorithm), to the most crude keyword spotting methods [2]. These last approaches use sentiment lexicons defining the prior polarity of particular terms, and aggregate these scores to determine the text's polarity. Recent literature has taken down the text analysis to sentence-level, as different opinions and sentiments towards the same topic can be present across the span of a document [3].

Twitter is a popular social media platform where the number of characters in a message is limited to 140. These messages, called tweets, can convey opinions about a range of topics. On the other hand, fuzzy sets theory is a valuable tool in situations dealing with imprecision and ambiguous or incomplete information. Its use can be of interest to tackle the assignment of polarity to small pieces of text such as the tweets.

This paper describes a sentiment analysis performed to Spanish tweets following a fuzzy sets approach and using an existing Spanish lexicon. A profiling method to evaluate the quality of the resulting indicator is proposed later [4].

Keywords: sentiment analysis, fuzzy sets, tweets.

References:

- [1] Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013), New Avenues in Opinion Mining and Sentiment Analysis, *Intelligent Systems IEEE*, 28 (2), pp. 15-20, DOI: 10.1109/MIS.2013.30.
- [2] Pang, B. and Lee, I. (2008), Opinion Mining and Sentiment Analysis, *Foundations and Trends in Information Retrieval*, 2 (1–2), pp. 1–135.
- [3] Cambria, E. (2013), An introduction to concept-level sentiment analysis. In: Castro, F., Gelbukh, A., González, M. (eds.) *MICAI 2013, Part II. LNCS*, 8266, pp. 478–483, Springer, Heidelberg.
- [4] Buelens, B., Daas, P., Burger, J., Puts, M., and Van den Brakel, J. (2014), Selectivity of Big Data, Discussion paper 201411, Statistics Netherlands.

DATA ECOSYSTEM: A NEW CHALLENGE FOR OFFICIAL STATISTICS

D. Fazio¹, E. Giovannini², M. Signore¹

¹ Italian National Statistical Institute, Rome, Italy

² University Tor Vergata, Rome, Italy

Nowadays, NSIs are called to open the doors to new ways of constructing statistics in an era characterised by an infinity of new data sources available on Internet or coming from satellites, sensors, etc. [1]. The liquid data floating on the web is of different types and supplied by several providers. On one side, the Big and Open data are available as indirectly generated by citizens for many reasons (such as use of Social Media, on-line shopping, mobile phones, etc.) and belong to specific providers (private and public). On the other side, the crowd-sourced data is voluntarily collected by civil society communities through Web2.0 collaborative platforms sharing information for different scopes.

Recently, the use of new data sources is at the forefront of methodological and experimental studies carried out by NSIs, Academia, research institutes and private sector [2]. Most of the experiences developed so far are using the Big and Open data in a way to save time, money and to reduce the response burden, namely as inputs to: i) construct new statistics - as for administrative data; ii) replace existing surveys; iii) produce nowcasts.

The exploitation of the crowd-sourced data has to follow a different direction. This data can have great potential at sub-national level. The grassroots local data generated by communities can provide valuable information for a “closer” picture of the local reality and emerging phenomena, even though its quality cannot simply be assessed following traditional quality frameworks.

This paper explores how non-official local data can be used by official statistics. It argues that they can represent “satellite” information to complement official statistics and stimulate Social Innovation. The construction of Data Ecosystems at local level is a great opportunity to build up a knowledge-driven society where the communities can raise their voices towards a conscious Social Innovation [3].

Keywords: complement official non-official data, web civil society communities, Social Innovation.

References:

[1] Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), chaired by E. Giovannini, A world that counts, November 2014. Available at: www.undatarevolution.org

[2] Eurostat European Commission, NTTS 2015 Reliable Evidence for a Society in Transition, Conference Proceeding, March 2015. Available at: <http://www.cros-portal.eu/sites/default/files/NTTS2015%20proceedings.pdf>

[3] European Commission, FP7 Web-COSI project, Web Communities for Statistic for Social Innovation, coordinator D. Fazio, Italian National Statistical Institute. See www.webcosi.eu



Chair: Sofia Rodrigues

REVIEWING THE GENERAL AND CENTRALISED SYSTEM OF COLLECTING ADMINISTRATIVE DATA AND HOW IT RESPONDS TO THE GUIDELINES OF THE CODE OF PRACTICE AND THE PEER REVIEW RESULTS IN 2013-2014

Sinikka Laurila
Statistics Finland

Background

A new general and centralised system for collecting administrative data was implemented in 2015 at Statistics Finland. It was the outcome of a two year project, where a) a new information system for collecting administrative data was built, and b) the whole process of administrative data collection and its organisation was renewed.

The principles of the general and centralised administrative data collection

The system was built according to the guidelines of the ICT strategy of Statistics Finland. The system utilises existing metadata systems. The process metadata steers the administrative data from outer servers to users at Statistics Finland and the metadata system of data descriptions is used when inspecting the technical and substantial quality of the data. The aim of the system is to be as automatic as possible. The future goal is that all administrative data is collected via the system in 2017.

Code of Practice and the guidelines of the Peer Review in 2013-2014

In the Code of Practice, principle 8 “Appropriate Statistical Procedures” gives guidelines on managing and developing the collection of administrative data. The following goals are currently implemented in collecting administrative data:

The use of administrative data.

- Cooperation with data providers (especially in the public sector) is established and happens regularly.
- Cooperation with data providers is always based on mutual contracts.

The peer review measured both the realisation of the guidelines of CoP at Statistics Finland and gave recommendations for future development in collecting and using administrative data in statistical processes.

IMPROVING QUALITY IN THE ESTIMATION OF “TRUE” ECONOMIC PERFORMANCE OF ENTREPRENEURS BY INTEGRATING STATISTICAL AND ADMINISTRATIVE DATA: A NEW METHOD FOR MEASURING UNDER-REPORTING

R. Filippello, F. Oropallo, A. Puggioni, P. Sassaroli
Istat - National Statistical Institute of Italy

The low response rate in statistical sampling surveys on very small enterprises in Italy represented a big issue considering that they account for a large share in the target population of the statistical business register.

The development of an integrated database with different administrative sources (Frame of Structural Business Statistics) to estimate the economic performance of enterprises, has improved the quality of the SBS and National accounts estimates, eliminating the sampling error and in particular removing bias estimation due to non-response, because concentrated in specific subsets of the universe.

Thanks to the new sources of economic data, in fact, the measurement of the “regular” component of economic performance of entrepreneurs has been improved. Notwithstanding the enhancing of the regular component, the problem of measuring “true” economic performance including non-observed result due to voluntary underreporting by entrepreneurs still remained.

The information available at the micro level allowed, at the recent general revision of national accounts, following the introduction of Regulation ESA 2010, in the year 2014, to define a new method for estimating under-reporting.

In this paper general information about the construction of the new SBS-Frame are given. Moreover the new method of the estimation of under-reporting is presented. The new method has been developed through the following steps: a) analysis and stratification of the universe of enterprises in homogeneous sub-populations in terms of economic and fiscal behaviour; b) selection of units affected and not affected by underreporting for each subpopulation; c) definition of statistical methods to estimate the portion of income not declared in administrative data.

In particular, the paper focuses the under-reporting estimation for marginal and micro entrepreneurs. Cluster analysis and the use of behavioural economic models have been implemented in order to estimate economic results and statistical imputation has permitted to measure the non-observed performance.

Keywords: Administrative data, micro-integration, under-reporting.



ALL THAT GLITTERS IS NOT GOLD!

Pedro Cunha¹, Sónia Quaresma², Jorge Magalhães²

1 Statistics Portugal, Porto, Portugal

2 Statistics Portugal, Lisboa, Portugal

In Portugal since last March using the citizen portal it's possible to change your address not only with public entities like the national registry or finances but also with agencies like the employment agency and even some private companies like the Highways Green Card, whenever you choose to do so. Having this ability is great but it relies on the citizen to provide the correct address thus having a limited impact in the building of a national address database.

In the second quarter of 2015 was inaugurated in France La Base Adresse Nationale which has 25 million addresses already and with the collaboration of the people, public authorities, state authorities, public operators and enterprises, the database provides open data and aims to quickly identify and manage 200-300000 new addresses created each year.

Last April, during the U.S. National Address Database Summit the U.S. recognized the importance of a usable, unified set of addresses as a piece of national infrastructure. But there, as well as in Portugal, a combination of bureaucratic, legal and organizational challenges have stood in the way of fixing this problem. The main concern being the security of personal information.

During the last census operation Portugal also performed a housing census integrating geospatial information which can constitute our initial National Households File but has to be updated to be of use as an instrument for sample design for the surveys of the NSI as well as a base for the next census operation.

We propose to show how we guarantee enhance and measure the quality of the file, through its update and maintenance, using administrative data and surveys.

Furthermore we suggest and discuss models of partnership and collaboration between public and private organizations (ADENE and EDP) to secure and assess the quality management of a National Address Database.

Keywords: National Address Database; Information Management and Maintenance; Partnership between Public and Private Organizations regarding Information collection.

QUALITY EVALUATION FOR A STATISTICAL REGISTER: THE ITALIAN FRAME SBS

O. Luzi, F. Rocci, R. Sanzo, R. Varriale, G. Brancato
National Statistical Institute, Italy

At the Italian National Statistical Institute (Istat), the direct use of administrative data for estimating business statistics has progressively increased, stimulated by the augmented availability and quality of secondary data on both private and public businesses. In this context, in 2013 Istat has implemented a new statistical register (called Frame SBS) for the annual production of economic accounts statistics based on the massive, integrated use of administrative and survey data ([1]).

Given the peculiarities of the target population and the characteristics of the available sources, the development of the system has implied the management of a number of challenging issues, like the harmonization of concepts in the sources (target populations/units, target variables), the evaluation of their quality and usability (coverage, accuracy, etc.), the analysis and treatment of measurement, coverage and response errors.

The transition from the traditional survey-based estimation procedure to the new production strategy has determined a deep change in terms of both the methodological and organizational strategies adopted. Furthermore, the use of integrated administrative data has determined the need of developing new methods and tools for the evaluation of the quality of all the components of the statistical process: input data sources, data processing, and outputs. Starting from the results of the European projects BLUE-ETS and AdminData, and taking into account the theoretical framework proposed by [2], in this document we delineate a first proposal of indicators for measuring and documenting the quality of the Frame SBS. The final goal is to implement a comprehensive control system for the regular monitoring of the quality of the Frame SBS taking into account all the quality dimensions (accuracy, timeliness, coherence, etc.), allowing the identification of the weaknesses of the process, their impact on quality dimensions (accuracy, timeliness, coherence) and supporting the evaluation of quality improvements.

Keywords: Administrative data, Quality.

[1] Luzi, O., Guarnera, U., Righi, P. (2014) The new multiple-source system for Italian Structural Business Statistics based on administrative and survey data. European Conference on Quality in Official Statistics (Q2014). Vienna, 3-5 June.

[2] Zhang, L.-C. (2012) Topics of statistical theory for register-based statistics and data integration. *Statistica Neerlandica*. Vol. 66, 41–63.



QUALITY ASSESSMENT AND VALIDATION OF ADMINISTRATIVE DATA SOURCES IN HEALTH STATISTICS

M. Zaletel, T. Zupanič

National Institute of Public Health, Ljubljana, Slovenia

Sufficient data quality is essential for production of national statistics, both for survey and administrative data, used for statistical purposes. The assessment and validation of later can be achieved by different means; one method is by using external sources, which are also administrative sources with different initial purpose of data collection and are judged to be of better quality. This is often the case for health care statistics, where administrative sources are widely used. The initial administrative database in this case is often composed of information, gathered from different sources, e.g. medical doctors, patients (their educational attainment and information on their health history), coroners at the hospitals, nurses, etc. The starting point in our case study was the hospitalisation database, which is an important data source for estimates on health case utilisation and provides estimates on incidence for some diseases. The analysis and comparison with additional external data sources have been performed. These additional data sources could be collected or compiled for statistical purposes (e.g. population and census data gathered by NSI), external administrative databases (e.g. police transport accidents database) and internal databases (e.g. Causes of Death). Variables of interest are in most cases socio-demographic data (e.g. educational attainment level) which take essential part in majority of analysis, but when of poor quality they can lead to false estimations and consequently to wrong inferences. In the paper, steps necessary to evaluate and compare different data sources are presented, quality analysis of key variables, essential for national and European statistics are shown. Based on detailed quality analysis, external data sources were evaluated and assessed as appropriate or inappropriate data sources.

Keywords: quality assessment, administrative data, data linkage.

Chair: Peter van de Ven

THE ASYMMETRY DILEMMA WITH INTRASTAT - WHICH DATA IS THE BETTER ONE? NATIONAL EXPERIENCES OUT OF THE ESS VIP “REDESIGN OF INTRASTAT”

Ilda Duarte Fernandes
Destatis, Germany

Since the implementation of the Intrastat system in 1993 to display movements of goods between the EU Member States, there was a focus on the administrative burden imposed on the reporting enterprises. By the end of the 1990ies, the European Commission launched the SLIM initiative (Simpler Legislation for the Internal Market) and in 2004 a reduction of the coverage rate for both flows from 99% to 97% was introduced. Two years later the Commission suggested a medium-to-long term implementation of the so-called single flow system, implying the compilation of national imports data from partner countries' export data. The subsequent discussion about the feasibility of this system showed that due to existing quality issues the single flow system is not practicable on a medium-term perspective. Decisive criteria are asymmetries between the EU partner countries' data.

Nevertheless, in November 2011 the European Council decided to envisage further measures on burden reduction, explicitly not excluding the option of a single flow system, while maintaining a “sound level of quality”. Destatis intensively examined the use of mirrored partner data on the quality of intra-Community trade statistics. Special focus was laid on the effects caused by asymmetries: As for 2014 German exports to EU Member States were around 2.5% higher than the respective imports. Broken down by partner country or commodity level those asymmetries might be even higher.

This may be caused by errors in reporting or in the compilation process. But to some extent asymmetries are due to methodology: They often occur in connection with triangular trade or transit businesses. Whenever more than two countries are involved in the transaction, there is a risk for implausible reporting. As these asymmetries partially are inherent in the system and cannot be totally avoided they should not be considered as quality flaws.

Based on these findings it is analysed how far the degree of asymmetries fits as a tool for quality assessment in intra-Community trade statistics with regard to the use of partner data.

Keywords: External trade, Intrastat, asymmetries.



AUTOMATIC BALANCING OF THE NATIONAL ACCOUNTS

M. Fridén, A. Isaksson, D. Lennartsson, A. Lennmalm, M. Saltveit, K. Strandberg, Y. Xie
Statistics Sweden

An important part of the calculations of GDP is the balancing of the supply and use in the economy in the framework of supply and use tables. At Statistics Sweden today, this work is mainly done manually; a quite resource-demanding undertaking. We propose an alternative balancing approach, making use of the existing framework of supply and use tables and utilizing the uncertainties of input data as weights. From the balanced results, the GDP estimates can be induced, along with their uncertainties. The approach, basically a generalized least-square using the uncertainties of input data as weights, origins from Stone et al (1942). It provides an automatic and fairly objective way of conducting the balancing process compared to the manual (and somewhat subjective) methods used today. It also gives valuable insights into the uncertainties of various input data sources. We present empirical results from a recent evaluation of the method in the production environment of Statistics Sweden. The main challenge proved to be the in-house collection of all necessary information on uncertainties.

Keywords: Uncertainty in GDP, Supply and use tables, Sampling and nonsampling errors.



ASSESSING THE QUALITY OF NATIONAL AND REGIONAL ACCOUNTS DATA IN THE CONTEXT OF ESA 2010 TRANSMISSION PROGRAMME

Julio Cabeça, Hans Wouters, Ani Todorova
Eurostat, Luxembourg

The European System of Accounts 2010 (ESA 2010) is a European statistical standard for the compilation of National Accounts data. It has been developed in line with the System of National Accounts 2008 (2008 SNA) and is legally supported by Regulation (EU) No 549/2013 of the European Parliament and of the Council of 21 May 2013 on the European system of national and regional accounts in the European Union.

Article 4(2) states that Member States shall provide the Commission (Eurostat) with a report on the quality of the data to be transmitted; the quality of the data has to be assessed by Eurostat. Article 12 mentions that by 1 July 2018 and every five years thereafter, the Commission shall submit a report on the application of this Regulation to the European Parliament and the Council. This report shall evaluate the quality of data on national and regional accounts.

Eurostat has been working with national statistical offices in working out how quality criteria set out in Article 12 of Regulation (EC) No 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics can apply in the context of ESA 2010. Relevant quality criteria have been identified alongside with appropriate quantitative and qualitative indicators to assess them. This paper takes into account the work being carried out until now and will present the first results of the discussions being held with countries as well as a proposal of quality criteria and indicators applicable to the national accounts framework. Quality assessment and metadata description overlap at many instances. Performing quality assessment and interpreting the results are not trivial tasks and should be done carefully.

Keywords: SNA2008, ESA 2010, Transmission Programme, quality, indicators.

References:

- [1] System of National Accounts, 2008 (SNA2008)
- [2] European System of Accounts, 2010 (ESA2010)

PUBLIC INVESTMENT: RECORDING IN EDP STATISTICS & TREATMENT UNDER THE SGP

L. Prado Ureña
Eurostat, European Commission, Luxembourg

Eurostat is responsible for assessing the quality of the data to be used in the context of the excessive deficit procedure. In recent months, transparent communication to users regarding National Account methodology applied to Public Private Partnerships has become a key issue.

Investment undertaken by government represents an important share of government expenditure and is closely monitored by Eurostat to ensure the correct calculation of government deficits and homogeneous reporting across MS.

Investment can be undertaken directly by government or, via specific contracts involving the private sector known as Public-Private Partnerships (PPPs). By using PPPs, government expenditure can, under specific conditions, be spread over a long period of time.

PPP rules have been introduced in the new ESA2010, which is a regulation to be followed for the compilation of national accounts in EU countries. The correct application of such rules is crucial to obtain a proper measure of government investment and ensure the quality of EDP statistics.

Over the years, governments have used PPP projects for the construction of infrastructure assets needed to render public services. Since the economic and financial crisis, the levels of investment decreased. As an initiative for economic recovery, the EC launched the Investment Plan for Europe. This aims at mobilising 315 bn of investment over three years and at making better use of public money to attract private investors. The cornerstone of this Plan is the European Fund for Strategic Investment (EFSI).

Eurostat has been confronted with questions suggesting whether the treatment of investment and PPP rules in ESA2010 could be in contradiction with the objectives of the Investment Plan for Europe.

In view of better communicating to users, this paper will:

- provide an overview of Eurostat's role in assessing the quality of EDP data;
- provide a simple picture of public investment, PPPs and the Investment Plan for Europe;
- draw the line between statistical recording and treatment under the SGP.

Keywords: Transparent communication to users, EDP quality, Public investment, PPPs, Investment Plan for Europe.



Chair: Kari Djerf

QUALITY WORK IN STATISTICS - FROM Q2001 TO 2016

Hans Viggo Sæbø

Statistics Norway, Oslo, Norway

Quality has been on the agenda in NSIs and Eurostat the last 20 years, following the UN Fundamental Principles on Official Statistics adopted in 1994. The first European conference on quality in statistics, Q2001 in Stockholm, was an important milestone. The basis for this conference was the work and conclusions from the Leadership Expert Group (LEG) on Quality. Later international cooperation initiatives in this area have set the terms for the next Q-conferences, which after 2004 have taken place each second year. Main European developments comprise the Quality LEG Implementation Group, European Statistics Code of Practice, Sponsorship on Quality and Peer Reviews. Similar developments on quality thinking in statistics have taken place outside Europe, directed by the UN or regional cooperation bodies.

Reflecting on these developments is useful for guiding the way forward, on creating and maintaining a culture for continuous improvement in European and national statistics.

The paper considers the main developments in quality work and their impact on statistics. In addition to the establishment of quality frameworks and reviews based on these, the developments comprise a change of focus from product quality to processes and better balancing of quality components and costs. Together with technological improvements and more use of administrative data sources, this has resulted in more statistics produced more efficiently with improved timeliness. New data sources provide new possibilities, but also challenges for quality assurance.

Examples given are believed to be representative for several statistical institutes. Seeing the earlier developments including lessons learnt from the last European Peer Reviews in a longer term perspective may guide the planning of new initiatives, such as further development of the Code of Practice and quality in the Vision 2020. What can we learn that is relevant for the future of quality work in statistics?

Keywords: Quality, quality assurance, continuous improvement.



THE QUALITY SUPPORTING FRAMEWORK OF THE ESS VISION 2020

Zsuzsanna Kovács, Jean-Pierre Poncelet
Eurostat, European Commission, Luxembourg

Members of the European Statistical System (ESS) defined their long term strategy and modernisation programme as the “ESS Vision 2020” and adopted it in May 2014. The area of quality has been identified as one of the strategy’s five key areas – subsequently, it has been identified as one of the “supporting frameworks” of the Portfolio which has been set up to implement the strategy and which is composed of a limited number of projects and supporting frameworks.

The paper describes how the QUAL@ESS Vision 2020 has been constructed in order to reply to and fulfil the ambitious objectives of the Vision’s Quality key area – it then explains how this supporting framework is also a vehicle for taking the quality work of the ESS forward and for modernising the entire ESS common quality framework.

The paper also presents the first results of the implementation of the coordination action of the QUAL@ESS Vision 2020 supporting framework and explains why this coordination role is of high importance with regard to the implementation of the ESS Vision 2020. It then concludes that the first lessons learned not only from the quality coordination of the ESS Vision 2020 Implementation Portfolio, but also from the second round of ESS Peer reviews, might lead to a second, and more in-depth, revision of the European Statistics Code of Practice.

Keywords: Coordination, ESS Vision 2020, Code of Practice.



ENHANCING QUALITY PRACTICES AT THE BRAZILIAN INSTITUTE OF GEOGRAPHY AND STATISTICS

M. L. B. Zacharias

Brazilian Institute of Geography and Statistics, Rio de Janeiro, Brazil.

In the last years, IBGE, the Brazilian Institute of Geography and Statistics, has been focusing on different aspects of the quality of statistical production, while taking into consideration the best practices and the principles of official statistics. Many significant efforts have been implemented in order to improve the accuracy, comparability, accessibility and transparency of the statistical production process: the edition of IBGE Code of Practice, the release of a set of Guidelines and Protocols related to the dissemination of statistics, the increasing adoption of international standards and frameworks, like DDI, SDMX and GSBPM, are some examples of these efforts. In addition, one of the major projects in progress to improve quality and efficiency is the redesign of the corporate statistical metadata system. The intent of the metadata modernization is to establish common semantic structures, robustness of systems and harmonization of the survey questions and of the variable names, concepts and definitions, in a user friendly environment. Furthermore, the metadata structure is being developed in alignment with international standards, in particular, Neuchâtel Model and the Data Documentation Initiative – DDI standard. The aim of this paper is to present the progress and development of some quality management practices implemented recently, as well as the initial steps towards the modernization of the metadata system.

Keywords: Statistical quality, Code of Practice, Metadata, Standardisation.



QUALITY MANAGEMENT OF STATISTICAL PROGRAMS AT THE NATIONAL AGRICULTURAL STATISTICS SERVICE

Pamela D. McGovern, Robyn E. Sirkis
National Agricultural Statistics Service, Washington, DC, USA

The U.S. Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) is committed to providing timely, accurate, and useful statistics in service to U.S. agriculture. NASS conducts hundreds of surveys every year along with the Census of Agriculture every five years, and prepares reports covering virtually every aspect of U.S. agriculture. Production and supplies of food, prices paid and received by farmers, farm labor and wages, chemical use, and changes in the demographics of U.S. producers are only a few examples. An effective quality management program is vital in ensuring the utility, objectivity, and integrity of the statistical information that NASS provides to its customers and stakeholders. Therefore, it is essential to establish and implement strong quality management principles and practices that achieve consistent results and build quality into survey and census processes and products.

To meet its quality objectives, NASS developed a quality management model. Several of the key model components include establishing a quality assurance framework of standards and guidelines, implementing and strengthening quality assurance practices, conducting quality evaluations and assessments, measuring and improving processes through continual improvement, promoting education and outreach, and ensuring a sense of quality commitment.

This paper discusses the quality management model components and provides examples of the work NASS is doing to ensure product and process quality. We will discuss several project initiatives, as well as, the results and challenges encountered. Some examples of these efforts include: establishing a comprehensive set of statistical standards to instill quality in our information products and promote transparency; evaluating Blaise call history files to eliminate inefficiencies and balance respondent burden, cost, and quality; implementing an automated quality control system for telephone data collection; developing an innovative tool to provide data and quality metrics on large establishments; and conducting a data quality evaluation to measure questionnaire performance.

Keywords: quality management model, quality measurement, process and product quality.

ASSESSING AND IMPROVING QUALITY IN OFFICIAL STATISTICS: THE CASE OF THE FRENCH LABEL COMMITTEE.

Marc Christine

National Institute for Statistics and Economic Studies (Insee), France

The French National Statistical Label Committee, which exists since 1994, is competent to examine the projects of collecting social or economic data through statistical surveys, carried out by official statistical authorities. Actually, regarding the law of June 7, 1951, on legal obligation, coordination and confidentiality in the field of statistics, any official statistical survey, after being agreed by the National Council for Statistical Information (Cnis), must obtain the Quality Label delivered by this Committee. The work of the Committee deals with surveys concerning either businesses, households or individuals, agricultural farms or local public establishments. After a favourable opinion on the project, a notice of compliance is given to the survey by the Label Committee.

The examination of surveys lies on an in depth analysis of six criteria, each of them being linked with one or more items of the Code of practice:

- General context: european constraints, legal obligations.
- Statistical methodology, with focuses on two main domains : sampling, sampling frame, data basis, statistical units, first; then : non-response correction, calibration, imputation and other post-data collection methods
- Data collection process, contact with the interviewees, confidentiality protection, security of data transmission.
- Study of the questionnaire, with narrow links with the topics of relevance with respect to the opportunity delivered by the Cnis
- Costs and burden on respondents, length of questionnaires and duration of the survey
- Data dissemination, either aggregated or detailed.

The examination leads to formulate recommendations which are strictly followed to improve the process and make it closer to the best practices.

In its report on the National Statistical French system the Peer Review, in December 2014, has underlined the positive role played by the Label Committee. The paper describes the way it works and how worthy the assessment of surveys is.

Keywords: quality, methodology, official statistics.

Chair: Steven Vale

PRACTICAL EXPERIENCE IN THE IMPLEMENTATION OF DATA AND METADATA STANDARDS IN THE ESS

Bogdan Zdrentu, Hubertus Cloodt, Jan Planovsky
Eurostat

Standardisation of structural, reference and other metadata and making them available to the European Statistical System (ESS) and beyond is becoming more urgent with quickly raising demands for automatized machine to machine data and metadata transmission, exchange, validation and dissemination.

At international level, Eurostat is one of seven sponsoring organisations of Statistical Data and Metadata Exchange (SDMX) initiative to foster standards for the exchange of statistical information.

Within the ESS, reference metadata are presented in files based on a standardised format called Euro SDMX Metadata Structure (ESMS). These metadata represent the instruments used to monitor and implement the quality principles defined by the Code of Practice. Eurostat has also developed a standard for the production and dissemination of quality reports within the ESS Standard Quality Report Structure (ESQRS). While ESMS is used to produce a user-oriented quality report and is focused on output quality, ESQRS is focused to produce a producer-oriented quality report. Finally, "Single Integrated Metadata Structure" (SIMS) represents an inventory for all ESS quality concepts and reunites reference metadata from ESMS and ESQRS.

Practical implementation of standards by using underlying data models (such as Data Structure Definitions – DSD - defined by SDMX Information Model) must be inevitably accompanied by development of supporting IT tools, reference repositories (e.g. codelists), guidelines and training manuals. Two concrete examples are ESS Metadata Handler (ESS-MH) and Euro SDMX Registry. Although these tools are developed to satisfy Eurostat needs, they could be easily adapted and reused by other organisations at ESS, international or national levels.

The objective of this paper is to share our implementation experience at concrete statistical production process in constructing standardised data and metadata structures enabling seamless data collection, validation and exchange using modern IT tools and at the same time providing well-structured metadata of high quality.

PROCESS METADATA DEVELOPMENT AND IMPLEMENTATION UNDER THE GSBPM v5.0 AT STATISTICS SPAIN

D. Salgado, A.I. Sánchez-Luengo
National Statistical Institute (INE), Spain

In 2012 Statistics Spain took the decision to start promoting the use of the GSBPM as the common language to describe the production process of every statistical operation conducted by the office. During 2015, after some preliminary pilot experiences with diverse surveys, a standard for process metadata was developed and approved in a joint internal working group comprising experts with diverse background (methodology, computer science, data collection, subject matters, statistical dissemination, and survey quality). This adaptation of the GSBPM develops a third level of the model accounting for the characteristics of our production though following the principles and guidelines of this international standard. This third-level adaptation shows two main ingredients:

- Description of third-level production tasks.
- Description of business processes using the Business Process Model and Notation (BPMN) with the third-level production tasks as atomic activities.

Following international recommendations on the application of the GSBPM [1] the description of each third-level production task is given by specifying its input, output, throughput, tools, documentation and unit(s) responsible of its execution. The division of each second-level GSBPM subprocess is undertaken following the principles of functional modularity [2].

Regarding process modelling conformance, it is agreed to follow the BPMN descriptive conformance sub-class, possibly complemented with some elements from the analytic sub-class [3].

The adapted standard for the description of process metadata at Statistics Spain can be found on its web [4]. Currently, Statistics Spain is undergoing its implementation for (nearly) all statistical operations under its responsibility.

References:

- Informal Task on Metadata Flows (2013). Metadata Flows in the GSBPM. Work Session on Statistical Metadata, Geneva 6-8 May 2013. Working paper 22.
- J.H. Saltzer and M.F. Kaashoek (2009). Principles of Computer System Design: An Introduction. MIT press.
- OMG (2011). Business Process Model and Notation v2.0. Standard document URL <http://www.omg.org/spec/BPMN/2.0>.
- Instituto Nacional de Estadística (2015). Estándar de documentación de procesos de producción de operaciones estadísticas del INE: los informes estandarizados de los metadatos de proceso. In Spanish.

URL http://www.ine.es/clasifi/estandar_procesos.pdf

Keywords: Process metadata, GSBPM, functional modularity.

METADATA MANAGEMENT SYSTEMS, LIVING SPIRALS OR FLAT LINE?

A. Portugal, S. Nunes, J. Poças
 Statistics Portugal, Lisbon, Portugal

Metadata and quality management systems can support a consistent process through the identification of needs, design and build of statistical operations, proceed to the evaluation phase and restart of a new loop.

At the INE-PT Metadata Management System (SMI) the objects standardization, namely concepts, classifications, variables or collection instruments leads to a more optimized flow to launch a statistical operation in the field and disseminate its results. The multitude of analysis and results that can be produced is stimulus to a more challenging appropriation of data.

One of the challenges of today's information is the ephemeral meaning of data and the need of versioning. The different objects that make part of a metadata system rearranged with different versioning can lead to a multitude of information that multiplies the initial meaning of data. At SMI objects standardization is implemented to achieve a better understanding among all intervenients during the design and building phases. Reuse of objects is crucial to explore the potential of metadata management system. Good search tools minimize the creation of new objects.

Integration with other systems supplies information to the building phase of applications and also reports to statistical operation's analysis and evaluation of processes.

Information in metadata management systems can't be carved in stone. The user friendliness and users recognition as an added value to the process are essential to keep the system alive. Challenges are in place, new standards and procedures are continuously arising demanding an effective answer from the metadata management system.

Keywords: Metadata, Management system, Quality, Standards.

INTEGRATED METADATA SYSTEM IN THE HUNGARIAN CENTRAL STATISTICAL OFFICE

Ágnes Almádi

Hungarian Central Statistical Office, Hungary

In the Hungarian Central Statistical Office (HCSO) the concept of a metadata-driven, integrated statistical system has always been considered as a strategic goal. Therefore the HCSO decided to elaborate a structured and integrated metadata system in the 1970s, and the most recent main improvement actions took place in 2007-2009. Concerning the development of the metainformation system, currently all types of methodological information are provided by seven subsystems. In 2015, the overview of the whole structure is under the loop aiming the optimization of functions required by the new standardization concepts. Quality of metadata can be assured among others when they are up-to-date, relevant, comparable, consistent, available, standardized, complete, and understandable. Ongoing standardization projects (e.g. the Statistical Data and Metadata Exchange (SDMX)) to standardize the structure of data-transfer; the Single Integrated Metadata Structure (SIMS) to describe the metadata of statistical domains; the Data Documentation Initiative (DDI) to describe the metadata of microdata; the Enterprise Architecture confirm the necessity to overview of the current metadata handling methods and metadata storing principles.

The aim of this overview action is to raise the quality of the metadata and the efficiency in the metadata-driven, integrated system of the HCSO. The classification of the statistical domains (the new version is based on SDMX List of statistical domains) is the base of the revision of the organisation of metadata. The purpose is to ensure the controllability of the maintenance through proper coordination, so the main aspects of the overview is the supervision of the technical requirements and the content and the main observation factors are the completeness and the correctness of the metadata.

Keywords: integrated metadata system, standardisation, classification of statistical domains.



Chair: Tasos Christofides

UNCERTAINTIES IN THE SWEDISH PPI AND SPPI

M. Frién, K. Strandberg and A. Norberg
Statistics Sweden, Stockholm, Sweden

In estimating the uncertainties in a sample survey it is easy to concentrate on the sampling error since it often can be quantified numerically. In the Swedish PPI and SPPI there is an established formula for estimating the sampling error. The formula takes into account the multi-stage sampling design as well as the finite population correction. However, a big part of the uncertainties in these surveys are non-sampling errors, such as specification error, measurement error or model error. An effort has been made to estimate the impact of these errors and for each stratum the error contribution from non-sampling sources has been judged to be “Low”, “Medium” or “High”. It is not clear how to combine the sampling error with the non-sampling error (or bias) into an overall measure of mean square error. We propose a method to estimate the MSE and to identify stratum with the biggest total uncertainty.

Keywords: variance estimation, non-sampling error, total survey error.

STANDARD ERROR ESTIMATION - HOW TO DO IT QUICKLY, EFFICIENTLY AND CORRECTLY

R. Seljak, J. Pikelj, P. Blažič

Statistical Office of the Republic of Slovenia, Ljubljana, Slovenia

Standard error estimation is in the case of sample surveys one of the most demanding and challenging tasks during the survey evaluation process. On the other side it is also true that in sample surveys standard error can still be considered as a key quality indicator that predominantly indicates the accuracy of the disseminated results. It is therefore of crucial importance that these indicators are on the disposal for the estimates, they are on disposal quickly and are based on the sound and transparent methodology. It is a special challenge, how to fulfil all these requests in the situation when we are more and more facing demands for more results, shorter timeliness and lower costs.

At the Statistical Office of the Republic of Slovenia we spent last few years to develop generic software for standard error estimation. The developed metadata driven application should enable that the standard errors are on disposal for all the statistical results that are based on the data from random sample and that they are on disposal at the same time as the results themselves. During development of such general tool, which aims to cover a wide range of different input data and different estimators, we faced many methodological and practical challenges. Many times pragmatic solutions to the problems outweighed the more “theoretically clean” ones. In the paper we present the main principles of the general application, describe the main challenges in its development and how we dealt with them.

- L. Lyberg et al. - Survey measurement and Survey Quality, Wiley, 1997
- Sarndal Carl-Erik: Model Assisted Survey Sampling: Springer-Verlag, 1992

Keywords: standard error, estimators, metadata driven application.

USE OF CALIBRATION IN OFFICIAL STATISTICS: DESIGN VERSUS MODEL-BASED CALIBRATION - PROS AND CONS

Marcin Szymkowiak^{1,2}

1 Center for Small Area Estimation, Statistical Office, Poznań, Poland

2 Poznań University of Economics and Business, Poznań, Poland

Calibration is a common method of estimation in survey sampling whereby sampling weights are adjusted to reproduce known population totals [2]. Calibration can be treated as an important methodological instrument, especially in large-scale production of official statistics. Many national statistical institutes all over the world use this technique in different surveys including censuses, sample surveys and surveys based on administrative registers. In the literature calibration is also considered in the context of the modelling approach. A model-based calibration was introduced by Wu and Sitter [3] and its key property is that the design weights are calibrated to match the population total of predictions derived via an assumed model.

Although modelling in general is very popular in statistics, the application of model-based estimation procedures in official statistics is rather limited [1]. National statistical institutes all over the world have been rather skeptical about the use of model-based methods in official statistics because they have relied on censuses, registers and surveys, using mostly design-based and model-assisted techniques. This also applies to model-based calibration, whose validity depends on the correctness of the model.

The main aim of the article is to present the difference between design and model-based calibration and identify areas in official statistics where these techniques can be used. Advantages and drawback of using design and model-based calibration in official statistics will also be discussed. The two kinds of calibration estimators will be compared in a simulation study based on real data.

Keywords: calibration, model-based calibration, modelling in official statistics.

[1] Brakel van den J., Bethlehem J. (2008), Model-Based Estimation for Official Statistics, Statistics Netherlands, Voorburg/Heerlen.

[2] Deville J-C., Särndal C-E. (1992), Calibration Estimators in Survey Sampling, Journal of the American Statistical Association, Vol. 87, 376-382.

[3] Wu C., Sitter R.R. (2001), A model-calibration approach to using complete auxiliary information from survey data, Journal of the American Statistical Association, 96, 185-193.



INFERENCE FOR STATISTICS BASED ON COMPLETE ENUMERATIONS?

J. Klotz

Statistics Austria, Vienna, Austria

It is often assumed in official statistics that statistics based on complete enumerations, such as census records or vital events register data, are 'true values', thus not subject to 'randomness'. Based on that it is concluded that statistical inference procedures such as standard error calculation or hypothesis testing are unnecessary, or even not allowed for such statistics. In my presentation I argue that whether a statistic is random depends not only on the data source, but also on the use to which the statistic is put. For example, if a census result is used to make decisions affecting populations in the future, then it makes sense to interpret the census as a random snapshot of a population over a longer period. I exemplify five general cases of applications in which statistics based on complete enumerations should be interpreted as realizations of random variables: identification of (too) small populations, inference for a superpopulation, correct weighting in ecological regression analysis, error detection, and assessment of uncertainty in projections. Finally, Statistics Austria's method of calculating the standard error of the annual total fertility rate of Political Districts is presented.

Keywords: official statistics, randomness, complete enumerations.



Chair: Pilar Martín-Guzmán

USING PARADATA TO ASSESS THE QUALITY OF AN ONLINE QUESTIONNAIRE

Alexander Kowarik, Marc Plate, Magdalena Six, Matthias Till

Paradata are process metadata data for instance generated during the process of collecting survey data. The conversion to a new, integrated system for survey data collection (EISNeu) in Statistics Austria makes it possible to collect different kinds of paradata for the use of measurement and non-response error evaluation. The Austrian Household Budget Survey (HBS) - providing information about expenditure and living standards of private households - is one of the first surveys at Statistics Austria making use of the new survey software.

The paper will describe indicators for potential measurement errors on the basis of paradata of the Austrian HBS on response times, mouse clicks and check-message appearance. These measurement error-indicators are meant to aid detecting pitfalls in the questionnaire design. Three different types of indicators are considered:

First we take a look at indicators related to the time needed to complete single questions, subparts of the questionnaire and the whole questionnaire. By that it is possible not only to examine if there are questions for which a large share of respondents needs an unusual amount of time for answering but also bivariate relations between time needed and sociodemographic characteristics of the respondent, or the correlation between the time indicators and questionnaire drop-out rates. Secondly indicators about the use of helping texts aiming to detect problems with the design of certain questions are considered. Finally the paper thematises indicators about so called check-messages which appear if the answer to a question contradicts a pre-defined rule such as "A person age 15 is unlikely to have a university's degree as highest education". A high check-message occurrence rate can either reflect a comprehension problem by the respondents or an incomplete view of the survey designer on the reality the concept should measure.

QUALITY IN THE GATHERING OF ONLINE DATA: STANDARDISING ONLINE QUESTIONNAIRES, INTEGRATION WITH ADMINISTRATIVE SOURCES AND DEVELOPMENT OF BIAS CONTROL MECHANISMS (EUSTAT)

Cristina Prado, Patxi Pizarro, Carmen Guinea
Basque Statistics Office (Eustat), Spain

The Basque Statistics Institute has spent years gathering data over the Internet, and has managed to ensure that all surveys, whether economic, demographic or social, now use this collection method, which has led not only to a reduction in costs, but also an improvement in the management of human resources dedicated to these tasks. It also gives all citizens a quick, simple and convenient way of completing surveys.

Furthermore Eustat, aware of the huge importance of data quality in statistical processes, has made a great effort to standardise the design of all web questionnaires, to improve their accessibility, create designs that are more visual, and offer help and controls to avoid the biases that may arise in data from self-completed surveys, which do not have the external help present in other collection methods.

Another course of action undertaken in recent years has been to integrate data from external administrative sources into web questionnaires, to serve a dual function. Firstly, it serves to contrast and verify the data entered in the questionnaires, and secondly, it is used to give respondents suggestions for possible responses. This allows us to improve the quality of the responses and reduce non-responses for certain types of questions.

Finally, we also want to show the Demographic Survey Manager that we have developed a single application that controls and monitors all of the surveys from a single point, and allows us to obtain data on the status of the survey process online, allowing us to establish corrective measures against unexpected situations that affect the quality of the data being collected at that moment.

The aim of the report will be to demonstrate the standardised design of Eustat's web questionnaires from the perspective of data quality, and not only regarding the format of the questions but also regarding the control and error validation system, the help system established to facilitate question comprehension, and the integration of data from administrative sources from external bodies. To do this, we will use various online questionnaires that are currently live, in the areas of economics and demographics, namely "Private Education Expenditure and Financing", "The Information Society for Families" and "Population in Relation to Activity".

Keywords: Integration, Design, Bias.

RESPONSE BURDEN DATABASE AND RESPONSE BURDEN INDICATORS AT THE STATISTICAL OFFICE OF THE REPUBLIC OF SLOVENIA

V. Šegan

Statistical Office of the Republic of Slovenia, Ljubljana, Slovenia

In its continuous efforts to measure and manage the response burden for business surveys, the Statistical Office of the Republic of Slovenia (SURs) plans a series of actions in this area. One of them is the project Measuring Response Burden for Business Surveys (MOPS), which lasts from October 2014 until March 2016.

The purpose of the project was to establish a system for measuring the actual and perceived response burden in a standardised way. While building this system we tried to follow some key principles. Since modes of data collection can be quite different, we had to construct a relatively robust system that can cover all of these differences. The system of measuring response burden should also have a minimal impact on regular statistical processes. The usage of existing methods, programs and databases were given a priority over completely new solutions.

The main product that is the result of the project is a database, which is the basis for calculating different response burden indicators. Response burden indicators will help us build a burden management strategy. Three main types of outputs were considered: the first is a set of predefined indicators of actual response burden calculated at different levels (level of the NSI, level of the individual survey, level of the enterprise). The second is the possibility of using special analysing tools by different users. This would allow ad hoc analysis at different levels and for different purposes. The third is the possibility to prepare the data on the actual burden in the way that could be easily used for coordinated sample design.

The prototype of the response burden database and some examples of the response burden indicators and standardised outputs will be presented.

Keywords: actual response burden, response burden measurement, response burden indicators.



MANAGING AN INTEGRATED RESPONDENT COMMUNICATION: STATISTICS PORTUGAL EXPERIENCE

Paulo Saraiva, Almiro Moreira
Statistics Portugal

Statistics Portugal started ten years ago a process of modernisation of its production system. In this process, focusing firstly on business surveys, three key elements worth to be highlighted:

- The implementation of the Simplified Business Information (an administrative internet based source) that conveys each year almost census accounting data on the corporation sector, covering the information requested by four public institutions;
- The full implementation of an Integrated Survey Management System, which consists on components that support the basic statistical production sub-processes collect, process, analyse and disseminate;
- The promotion of a better relationship between data providers and Statistics Portugal notably by improving the dialogue with the data providers and by designing a service of customised feedback information to them.

This paper provides an overview on the content of these three elements.

Keywords: Official statistics production; process integration; respondent management; data collection; survey management.



DEVELOPMENTS IN MEASURING THE BURDEN PLACED ON BUSINESSES RESPONDING TO STATISTICAL SURVEYS

A.Tucker, D. Williams, M. Pope, R.Sanderson
Office for National Statistics, UK

The UK Code of Practice for Official Statistics (UK Statistics Authority, 2009) specifies the need to report annually on the burden placed on respondents to surveys of businesses and households. Whereas information on the time taken for a household to respond to a questionnaire can be measured at the point of collection, it is more challenging to measure the time and hence the cost to businesses of responding to surveys used to compile official statistics. A traditional approach to such measurement of surveys conducted using paper questionnaires is to send a short review questionnaire to a sub-sample of businesses. This review questionnaire gathers information both on the time taken to respond to the main survey but also who in the business provides this information; this can then be used to estimate the financial costs to the business.

Such reviews ceased at ONS in 2012 and information on respondent burden was collected through a self-assessment tool used by survey managers to assess the quality of statistical outputs. This used information from previous reviews and knowledge of survey changes to estimate respondent burden. However, it proved difficult to collect high quality information on respondent burden without the data from these review surveys. Therefore, motivated by the aim to improve the measurement of the financial costs to businesses of responding to surveys, a shortened review process has been piloted. In an effort to balance the burden placed on respondents by this process and to make the process as efficient as possible, we have tested the use of statistical modelling to estimate respondent burden for surveys with similar characteristics.

In this paper, we report on the pilot exercise carried out, including the methodology, results and conclusions of this work. We also consider the implications for the future measurement of respondent burden placed on businesses.

Keywords: Respondent burden, Code of Practice, future measurement.

Chair: Albrecht Wirthmann

BIG DATA IN OFFICIAL STATISTICS: ESTIMATE JOB VACANCIES BY USING WEBSCRAPING TECHNIQUES

Martina Rengers

Federal Statistical Office, Wiesbaden, Germany

In the last years Big Data has become more and more of high interest. After broad exploratory work organised e.g. by the UNECE [1] the DGINS conference emphasised in the Scheveningen Memorandum of September 2013 [2] the relevance of Big Data for the European Statistical System (ESS) and called for incorporating Big Data in the conceptual design of official statistics.

To this end, an action plan was endorsed in September 2014 [3], leading to an ESSnet action on Big Data. Within the ESSnet BIG DATA at least three pilot studies should explore the potential of selected big data sources for official statistics, with all its attendant challenges. Germany is one of four countries which take part in the pilot study regarding web scraping job portals which is planned to start from 2016. Web scraping is a technique employed to extract large amounts of data from websites. Since many job offers are made available via job portals, the use of these data potentially can be highly interesting to complement job vacancy statistics, as well as labour market statistics in general. The pilot study aims at analysing all relevant aspects to be considered for an assessment of a use in official statistics: Regarding the conceptual aspects, it will have to be analysed which kind of information is available at different job portals and this information fits the existing statistical standards. Regarding the technical implementation a suitable web scraping tool needs to be selected and tested in practice. Additionally, there are a number of legal aspects relating to data access that the project will focus on, e.g. copyright laws and terms of use of the job portals. Only on the basis of these findings, the quality of the data can be assessed and the potential future role of these data can be outlined.

The presentation will detail the approach of the pilot study and present first results.

[1] UNECE 2013: What does “big data” mean for official statistics? Report of the High-Level Group for the Modernisation of Statistical Production and Services (HLG). <http://www1.unece.org/stat/platform/display/hlgbas>

[2] DGINS 2013: Scheveningen Memorandum, Big Data and Official Statistics, <http://ec.europa.eu/eurostat/documents/42577/43315/Scheveningen-memorandum-27-09-13>

[3] ESSC 2014: ESS Big Data Action Plan and Roadmap 1.0, BDAR (ESSC 2014/22/8/EN), http://www.cros-portal.eu/sites/default/files/ESSC%20doc%2022_8_2014_EN_Final%20with%20ESSC%20opinion.pdf

Keywords: big data, webscraping, job vacancy, labour market statistics.

QUALITY IMPLICATIONS OF THE USE OF BIG DATA IN TOURISM STATISTICS: THREE EXPLORATORY EXAMPLES

F. Cortina García, M. Izquierdo Valverde, J. Prado Mascuñano, M. Velasco Gimeno
National Statistical Institute (INE), Spain

Tourism statistics is one of the subject areas which are being considered at present in the ESS as a potential field for the development of big data use in order to improve the relevance, opportunity and punctuality of the products offered under the quality standards of official statistics.

In Spain, data from traffic loops and traffic control cameras are already being used in the estimation of inbound tourists. The paper will also analyze three pilot studies about the use of big data and the integration of multiple sources:

- Credit cards:

A two-year monthly series of average expenditure by credit card holder, country/region of residence and region/NACE of the point of sale terminals (POS) have been provided to INE by one financial operator in Spain. Comparisons with survey data are being conducted.

- Mobile phones:

Flows and average stays per month of both residents and inbound tourists, per NUTS II or country of residence and NUTS II of destination have been produced on demand by a MNO operator.

- Web scraping:

Package tours should be unpacked and treated as the sum of their components in the context of TSA. To carry out this treatment ('net valuation' approach) specific information from the companies that make tour packages is needed, facing the following problems:

- National legislation in most cases do not include the obligation to provide information to non-resident units providing packages to visit our country.

Payments to hotels for concepts such as accommodation, breakfast, Wi-Fi, are often traded globally without the reporting unit available differentiated information.

- The estimation of the percentage that belongs to the home economy and the destination economy.

INE conducts a survey to travel agencies and tour operators which helps to overcome these deficiencies. A pilot on web scraping to collect prices of package tours and of its components if purchased individually has been launched, whose results will be evaluated as a potential source to improve the quality of the procedure of unbundling tourism package expenditure.

Keywords: big data, traffic loops, mobile phones, credit cards, package tours, net valuation, web scraping.



QUALITY CONTROL OF WEB-SCRAPED AND TRANSACTION DATA (SCANNER DATA)

Ingolf Boettcher
Statistics Austria

New data sources such as web-scraped data and business transaction data (e.g. online and scanner data from retailers for price statistics) have the potential to improve official statistics, both in terms of quality (more data) and efficiency (low data collection costs, lower response burden). However, when using new data sources statisticians have to review and eventually replace traditional data quality control procedures to comply with existing quality standards. The challenges to deal with are manifold: How to define and identify outliers in millions of transactions data sets? How to select representative data sets from the internet for official data production? How to validate integrity and completeness of web-scraped data? How to integrate new and diverging data set structures into established statistical production processes?

New kinds of skills (“data science”) are required from statisticians to handle these issues, such as advanced knowledge of data manipulation and programming – but also the right amount of statistical creativity to transform new and ever-changing (big) data sources into high quality official statistics.

Paper and presentation will outline Statistics Austria’s web-scraping and scanner data projects and describe the approaches taken to address the main quality challenges.

Keywords: web-scraping, scanner data, big data quality.



NEW DATA SOURCES TO INDICATE LEVELS OF ACTIVE CITIZENSHIP

C. Rosales Sánchez
Joint Research Centre (European Commission), Italy

Data availability is a general constraint in the generation of indicators for decision-making processes. Web 2.0 technologies offer new potentials of data sources (also considered as big data sources) that needs to be investigated. In this research we focus on indicators of active citizenship as a relevant dimension of social capital that needs to be measured and monitored. Citizens' participation empowers communities, increasing social capital and reducing possibilities of social exclusion, thus beyond-GDP initiatives consider citizen engagement an indicator of societies' quality of life.

Classical measurements of active citizenship are mainly based on surveys, which are costly and time consuming, and on voter turnout, which temporal availability depends on election periodicity.

Considering that the information publicly available by online social networks may offer a sense of how much citizens are engaged in their communities, this research analyses the potential of the microblogging online site Twitter as unofficial data source to measure active citizenship. As a result, this research defines a method to use Twitter data to obtain an indications of active citizenship that complements the results from official data sources offering a more timely, less costly information, and with higher spatial and temporal resolution.

Keywords: Twitter, active citizenship, social media, multi-source statistics.



Chair: Marina Signore

LINKAGE OF MAIN COMPONENTS OF GSBPM MODEL THROUGH INTEGRATED STATISTICAL INFORMATION SYSTEM

Ludmila Ivancikova, Helena Glaser-Opitzova
Statistical Office of the Slovak Republic

Development and implementation of new models, modern procedures and tools led to complex assessment of existing statistical system within production of the Statistical Office of the Slovak Republic. Stimulus was given Vision, existence of models such as GSBPM and need of reaction to requirements of internal and external environment.

In article we are concerned with main components of integrated statistical system, which was developing in the Statistical Office of the Slovak Republic in 2013 and 2014. The system is described not in terms of information technologies, but from the view of its specific using for statistical production and for creation of individual statistics. Article is focused on system architecture in relation to preparation, data collection and data processing and dissemination. The goal of system components is to contribute to reducing burden on respondents, increase of work efficiency and consequently make data of higher quality at input and output level.

Keywords: GSBPM, statistical system, integration.

DATA WAREHOUSE FOR AGRICULTURAL STATISTICS AS A TOOL FOR STANDARDIZATION OF DATA PROCESSING AND IMPROVING THE QUALITY OF EAA CALCULATION

D. Stankov, L. Tasevska
State Statistical Office, Republic of Macedonia

This paper provides information about re-designing of statistical processes in the area of Agricultural statistics, aiming at increased efficiency and quality of the results and outputs produced in the State Statistical Office of Republic of Macedonia (SSORM).

The State Statistical Office of the Republic of Macedonia (SSORM) conducts various agricultural surveys and uses administrative data for statistical purposes. The activities are often decentralized and implemented in various ways, depending on the needs and the dynamic.

In this regard, and having in mind the quality issues, SSORM has decided to establish a Data Warehouse for Agricultural Statistics. The data and metadata definitions are standardized so as to enable integration of historical survey data into a single database. The data model used for the Data Warehouse for Agricultural Statistics provides for the end user a simple and query-centric view of the data. The Data Warehouse enables an automatic calculation of the Economic Accounts for Agriculture data sets using available data from the surveys that are integrated in the Data Warehouse.

This paper discusses important issues related to establishing Data Warehouse for Agricultural Statistics as a tool for standardization of data processing and improvement of quality of data, as well as increasing the efficiency in the working activities as a result of automation of processes related to compilation of data sets on Economic Accounts in Agriculture.

Keywords: Data Warehouse, standardization, metadata, data quality, automation.



MONITORING QUALITY IN THE PRODUCTION PROCESS MODEL: EXPERIENCES FROM THE INE OF SPAIN

Asunción Piñán, Agustín Cañada
National Statistical Institute (INE), Spain

Following the current trend towards standardization promoted by international organisms (UNECE, EUROSTAT), the INE of Spain is currently working in a deeper extension of the GSBPM to its production processes.

The GSBPM recognizes several over-arching processes that apply across statistical business processes. One of these processes is the quality management, which includes quality assessment and control mechanisms. It recognizes the importance of evaluation and feedback throughout the statistical business process.

Together with the documentation of the production processes, the INE of Spain is organizing and enhancing its quality management system. The way forward has been to relate any step of GSBPM to the quality tasks, as they are carried out nowadays in the INE, and the evaluation procedures that should be undertaken. This evaluation procedures stem from the document of UNECE [1] with the proposed quality indicators for the GSBPM, adapted to the INE characteristics.

In a further step, the results of this process will be included in future revisions of the INE quality guidelines.

Keywords: GSBPM, standardization, evaluation procedures, quality management system.

References:

[1] Quality Indicators for the Generic Statistical Business Process Model (GSBPM), 2015, Working paper, United Nations Economic Commission for Europe, Statistical Division



STATISTICAL METADATA AND THE GSBPM

Fiona O'Riordan
Central Statistics Office, Cork, Ireland

Statistical Meta Data in a Statistical Institute is vital to ensure quality at each part of the data life cycle. The definition of meta data can be varied and subjective and may not produce the required result. A strong framework to ensure standardised Meta data is necessary. The GSBPM1 can provide this in conjunction with management of the data life cycle. Available Meta data at key times i.e. process or disseminate can increase the quality of the data and provide a road map for the past and the future of the data lifecycle.

This paper will examine the journey for CSO, examining the maturity level of meta data in the CSO and how to communicate and implement this change Office wide. The issues with the different requirements, methods of storing and accessing and the important of good quality meta data at all stages. It will also review the benefits of the use of the GSBPM

Keywords: GSBPM, Meta Data, Governance.

KARAT: THE NEW INTEGRATED DATA TRANSMISSION SYSTEM OF THE HCSO

Ildikó Györki, Ildikó Szűcs
Hungarian Central Statistical Office, Hungary

The number of administrative data sources (secondary data sources) used for the production of official statistics has increased in the HCSO in the past few years. These data transmissions were managed differently throughout the HCSO with different safeguards in place regarding quality, the accompanying metadata and database-management. In order to address this issue, a new, integrated data transmission system – called KARAT – was developed in 2013-2014 for the transmission of administrative data to the HCSO from the owners of administrative data.

KARAT provides a secure channel for receiving secondary datasets; it applies uniform data transmission procedures and standard documentation. The system also manages not only the transfer of secondary data, but it provides secure channels for the transfer of statistical data sets from the HCSO to users requesting statistical data from the Office.

All data sources (primary and secondary data sources) used at the HCSO are registered and described in the Metainformation System. The new data transmission system is built on metadata, it gives the documentation of secondary sources, and its metadata drive the flow of data transmission and processing of the transferred data sets. The system has a number of automatic, proactive functions to support the task of the data masters and data providers of the secondary sources. The deadlines, the flow and completeness of data transmission are controlled by the system. It ensures automatic functions to load the datasets into database. The receipt of the datasets is also registered in the system. The data providers and the responsible statisticians are noticed about either the reception or any problem occurred. The user requesting statistical data get similar support to download their required data sets.

KARAT also has monitoring functions to manage the whole data transfer procedure. The quality of the data transmission process can be measured by different indicators. The new system contributes to the improvement, standardisation and monitoring of process quality related to the transmission of administrative data and data requests making the KARAT system one of the key systems for the HCSO business processes.

Keywords: integrated metadata system, use of administrative datasets, standardisation.



Chair: Francisco Hernández

USING MONETARY INCENTIVES IN FACE-TO-FACE SURVEYS: ARE PREPAID INCENTIVES MORE EFFECTIVE THAN PROMISED INCENTIVES?

Michael Blohm, Achim Koch

GESIS - Leibniz Institute for the Social Sciences, Mannheim, Germany

Like many other surveys –including voluntary surveys from official statistics –the German General Social Survey (ALLBUS) has been facing an increase in nonresponse in the past years. ALLBUS is a repeated cross-sectional, multi-topic survey with about 3.500 respondents each round. The survey is conducted face-to-face and fielded every other year. Between 1994 and 2010, the response rate has been declining from 54 % to 35%. Against this background, several experiments were mounted in recent ALLBUS surveys to investigate whether or not the provision of respondent incentives might stop this trend. According to the literature, prepaid incentives are more effective than promised incentives in increasing rates [Singer 2013]. This is well documented for mail surveys. For face-to-face surveys, however, only a few experiments are available providing mixed evidence.

In our presentation we report the results of our latest experiment conducted in the ALLBUS 2014. In this experiment a random subsample of target persons was offered a monetary incentive of €10, conditional on the completion of the interview (promised). Another random subsample was sent a €10 note together with the advance letter (prepaid). The remaining part of the sample acted as a control group, receiving no incentive. We look at differential effects of a prepaid/promised incentive on response rates (1), we investigate incentive effects on sample composition/response distributions (2), and response quality (3). Special attention will be paid to the issue of fieldwork efforts and survey costs (4). The money spent to provide incentives can then be weighed against the cost savings resulting from a more cooperative sample. Since prepaid incentives have to be paid to each member of the gross sample of a survey, they usually have to come along with a large reduction in fieldwork effort than promised incentives in order to be cost effective.

Keywords: Face-to-face Survey, Incentives, Data Quality.

QUALITY DRIVEN DATA COLLECTION. TOWARDS A SYSTEM FOR QUALITY MANAGEMENT BASED ON DASHBOARD INFORMATION

N. Bjørnstad, G. Haraldsen, P. Næs and A. Sundvoll
Statistics Norway

An important part of Statistics Norway's current strategy is about streamlining all parts of the statistical production. In order to manage the data collection process in a cost-effective manner, we need to display and analyze coordinated process indicators relevant to data quality during the data collection process. Statistics Norway is presently developing a dashboard system for data collection supervision and management, based on process data and quality indicators.

The data collection management system will yield relevant and complete information about specific surveys, and aggregated information about all ongoing surveys. We want to take necessary actions taken during individual surveys based on responsive management principles. Furthermore a systematic overview over the total, current data collections is necessary in order to perform effective resource allocation. Therefore a dashboard system both needs to give an overview and offer a path towards survey specific details.

The data collection department will use this new dashboard system in day-to-day management and in ongoing quality management work. In addition the data collection dashboard will support the communication between the data collector and sponsors.

The management system currently under construction will be based on data automatically captured from the production systems in combination with paradata and quality assessments collected from the survey instruments. A new version of the questionnaire for the Structural Business Survey (SBS) is used as a pilot for the management system. One of the reasons why this survey was chosen is that it introduces a more proactive approach to error prevention than the traditional error check-and-correct approach traditionally used both in computerized questionnaires.

Our presentation will describe the management system more in detail and invite to a discussion on how different kinds of process and quality information should be presented on a dashboard.

Keywords: Data collection, management system, quality, indicators.



MIXED MODE EXPERIMENTS IN A PANEL SURVEY - CHANGES IN RESPONSE BEHAVIOUR

J Eklund, M Järnbert, T Orusild
Statistics Sweden, Stockholm, Sweden

Over recent decades, response rates in interview surveys have decreased rapidly in Sweden as in many other countries. This has led to higher uncertainty in estimates as well as higher costs of data collection. One way to deal with this problem is to allow different response modes. In this paper we examine what happens with response behavior in a panel survey when web mode is introduced as an additional mode in a traditional single mode telephone interview survey.

Three experiments, with large sample sizes (3,000-12,000), have been carried out in 2014 and 2015 mixing telephone interviews and web mode within a political opinion poll at Statistics Sweden. In all experiments the survey individuals have been randomly assigned into experimental and control groups. The experiments have so far shown positive results on increased response rates and no significant mode effects. Compared to the control group, the response rates are especially higher among earlier non contacts, but there are significant differences also among earlier refusals as well as among respondents from the previous survey round. The effect of offering web is also positive on new panels, where the respondents receive information on login by postal mail instead of e-mail.

Using three survey rounds in the overlap between single mode and mixed mode we seek out to learn more about who answers, when they answer and how they answer. In the evaluation we use data set of the surveys together with register data and para data from the data collection.

Keywords: Response rates, Data quality, mode effects.

MODERNISATION OF GERMAN HOUSEHOLD SURVEYS: MODULARISATION & MIXED-MODE - FUTURE CHALLENGES

K. Blanke/Th. Riede

Federal Statistical Office (Destatis), Wiesbaden, Hesse, Germany

Household surveys in official statistics are faced with a series of social, political and methodological challenges, due to requirements on data quality and data security as well as the availability of appropriate software packages. This also applies in the case of Germany and several other European countries. Considering the methodological perspective survey design is highly discussed by the implementation of mixed-mode data collection and the increasing demand to develop web questionnaires on one hand and the collection of data by the perspective of modules, rather than single surveys (modularisation) on the other hand. Against this background, in 2012 Eurostat introduced a two-year cooperation project on European level. In that ESSnet-Project “Data Collection in Social Surveys Using Multiple Modes” (briefly: ESSnet DCSS) five member states of the European Statistical System (ESS) (Finland, the Netherlands, Great Britain, Norway and Germany) conducted methodological preparatory work regarding the design of surveys using multiple modes. Building on the results of the ESSnet DCSS and further considerations, Destatis is currently preparing a major reform on the German system of household surveys. In this context, the presentation will indicate our understanding of an upcoming modularization in household statistics and how mixed-mode data collection is going to be supported on a methodological level as well as on the perspective of IT.

Keywords: Household surveys, Mixed-Mode, Data Collection.



Chair: Adolfo Gálvez

CHANGING USER NEEDS RAISE DEMAND FOR USEFUL INDICATORS

Monica Nelson Edberg
Statistics Sweden, Stockholm, Sweden

User needs have been a focus for statisticians for many years. Since the financial crisis, a growing complexity and an aim for higher efficiency in all economic processes have evolved. Technical development and a higher capacity of data transformation also offer new possibilities for users of statistics. Therefore, user demand is changing over time – an important fact that has to be taken into account.

The changing pattern differ for different users groups. At the same time as research is increasing in many areas, other areas such as journalism, analytic studies and sectorial analysis have continuous efficiency constraints. Some user categories, such as researchers, demand larger datasets, at a micro data level while others, such as journalists and business analysts demands more timely data, and “storytelling”, as they don’t have the time and resources to do their own analysis.

New users groups have entered the scene as many areas now have a global arena. These groups are politicians, both nationally and locally, and global business leaders. Their needs of statistics are primarily timely indicators, long time series and less data revisions, as decisions have to be based on a solid ground.

The fact that user needs are changing over time will be a challenge for the statistical community. From a user perspective the time-series have always been of the outmost importance, while the estimate is just one observation of this. Methods used should be transparent and easy to understand and new phenomena should be presented as they evolve.

The paper will give examples of how user needs are changing and show some examples of which “useful indicators” there are a need for.

Keywords: User needs, Indicators.



DATA VISUALIZATION - HOW TO ENSURE UNDERSTANDABLE AND SUSTAINABLE STATISTICAL PRODUCTS

S. von Oppeln-Bronikowski, C. Bergmann
Federal Statistical Office (Destatis), Germany

Statistical information is used today by a growing number and a broadening spectrum of users. This places new demands on the presentation of data. On the one hand, the grasp of statistics has to be improved, and on the other, the readers' interest aroused. Figures must be telling to make the abstract concrete. Only then can statistics be grasped by everyone and thus become comprehensible.

Visualisation is an indispensable tool of every statistician. With the right context, even a simple line chart turns an abstract row of figures into a little story revealing the information value of a statistics also to the layman. The visual memory is able to remember information better and for a longer time. Meanwhile, many diverse forms of presentation are used, which is vivid proof that the benefits of graphical representations and their value added are recognized.

As the complexity of the underlying data increases, static forms of presentation are reaching their limits. The huge amounts of data made available by the current information technology can often be visualised only inadequately with traditional means. So far, the potential offered for instance by spatial data has remained largely unused.

The digitalisation of the dissemination channels is opening up new ways of visualisation. Interaction and animation add another dimension to traditional forms of presentation: interactive graphs are 'hands-on' statistics. They do not only make figures tangible but also understandable. The possibility to interact provides for completely new approaches as an instrument of analysis and for understanding correlations, causalities, and references.

Keywords: Statistical Literacy, Accessibility, Code of Practise, Storytelling, Sustainable Products, Visualization, Graphical skills, Resources, Communication, User-friendliness.



ASSESSING AND DOCUMENTING QUALITY OF INDICATORS: ESTABLISHED PRACTICES AND OUTSTANDING ISSUES

Nicola Massarelli, Barbara Kurkowiak, Evangelia Ford-Alexandraki
Eurostat, European Commission, Luxembourg

Statistical indicators often have a close link with policy-making and monitoring and therefore are not just like other statistical products. For this reason, they require a specific metadata different from the ones used for data collections and accounting systems.

In particular, metadata on indicators should include a description of the specific context for which the indicator is used (e.g. monitoring a policy strategy), the exact definition and an explanation of the indicator-specific methodology, information which is normally not included in the metadata of datasets. On the other hand, methodological details on the underlying sources may be less relevant for an indicator-specific documentation.

Metadata for indicators should also provide a concise grading summarising the overall quality of an indicator. Several approaches could be followed for this purpose, which triggers the need to make choices for the quality dimensions to consider, on how to grade their quality and on if and how to summarise the individual grades.

Eurostat has established practices for documenting statistical products in general and for documenting and assessing the quality of indicators. These practices have evolved over time, on the basis of changing needs and with a view to streamline the information for the users. Considerations on possible further improvements are on-going.

The paper addresses the issues above on the basis of Eurostat experience, critically reviewing and discussing the various practices which have been, are or could be followed in the near future.

Keywords: Indicators, quality profile, quality grading.



REVIEWING ASPECTS OF USER ORIENTATED QUALITY REPORTING

S. Tucker

Office for National Statistics, Newport, United Kingdom

National Statistics Institutes are required by the European Code of Practice to provide quality reports to users alongside official statistics.

To meet this requirement, ONS produces a Quality and Methodology Information (QMI) report for each Statistical Bulletin. QMIs report against the European Statistical System (ESS) quality dimensions and other quality characteristics. They also contain information on the strengths and limitations of data which help users decide upon suitable uses for that data.

QMIs have been published for 5 years, so now is a good time to review how quality information is communicated to users. Are there any gaps in what is provided compared to current user needs? What can be done to extend the use of this information?

The primary purpose of quality reporting has been to help users decide upon suitable uses of the data. It was determined that there could be a step before this, namely to reduce the misuse of the data. Research has therefore been undertaken to look at how ONS can first help users understand how to not misuse the data, and then how the data should be used. This paper will discuss some work that has been carried out to reduce the risk of users misusing data. The findings from internal focus groups and meetings will be presented, including further work with statistical producers to create webpages for user testing. The paper will explore the results of user testing and the implications for communicating quality information to users in the future.

Keywords: quality reporting, users, communication, strengths and limitations.



Chair: Sixto Muriel

IMPROVING BUILDING PERMIT ADMINISTRATIVE DATA SET FOR SHORT TERM ANALYSIS: THE FRENCH CASE

François Limousin, Benoît Pentinat
Ministry of Ecology, France

Dealing with administrative data for short term analysis may be very confusing. On the one hand, plenty of information is available but on the other hand, at the same time, there may be severe lacks as the gathering process can be a little bit too long to be completed in e.g. one month. This is the case for building permit indicator in France, where a monthly transmission is defined but local authorities don't always play the game (on purpose or not). Our aim was then to reconcile administrative data and short term analysis.

The old way was to consider the flow of incoming data during the month : if the data collection process is stable, this is very efficient but it's no longer the case today as several reforms radically change the process for granting permits...and so far the statistical scheme.

Several kinds of methods were compared and faced with the set of constraints we have to deal with: producing monthly indicators, at the most detailed geographical level possible, coherent with the micro-data and the method used for annual housing accounts. We finally reached a position, compatible with the principles of the European Statistics Code of Practice, especially timeliness and reliability.

Our method uses time-series modelisation, in order to extract the short-term component and minimize the influence of old information. We challenged our method to assess maximum quality, and explore what will be the main sources of revision. It led to the choice of a full set of parameters. The work is still in progress and we are very careful about the quality of the models underneath. Revisions are monitored and accepted by the main users.

Keywords: administrative data, building permits, short term analysis.

EVALUATING THE POTENTIAL FOR MOVING AWAY FROM A TRADITIONAL CENSUS

B. Tinsley, L. O'Leary

Office for National Statistics, UK

In 2014, the UK's National Statistician made a recommendation(1) that an online census should be conducted in 2021 in England and Wales(2), and that census and population statistics should be enhanced using administrative data and surveys. This would also offer a springboard to the greater use of administrative data and surveys in the future.

The Government accepted this recommendation and stated their ambition that censuses after 2021 would be conducted using other sources of data, provided that the feasibility of such an approach could be sufficiently validated(3).

The Office for National Statistics (ONS) is conducting research and producing new outputs based on administrative data. Each year, these research outputs will be assessed to show the progress in demonstrating whether the Government's ambition can be realised. This will culminate in a recommendation in 2023 about the future of census and population statistics. A number of things would need to be in place in order to recommend moving away from a traditional census. Some pre-conditions have been developed and underpin the evaluation criteria that will be used to assess a move towards an approach based on administrative data and surveys. The evaluation criteria are aligned with the five ESS quality dimensions. Understanding the quality of administrative data, and how administrative data and surveys can be combined to produce quality statistical outputs is a large part of this assessment. Analysing administrative data at record-level has highlighted inconsistencies that lead to statistical quality issues, and we are beginning to take the first steps in feeding these back to data suppliers with the aim of improving quality over time.

Keywords: Administrative data, statistical data quality, administrative data-based census.

[1] <http://www.ons.gov.uk/ons/about-ons/who-ons-are/programmes-and-projects/beyond-2011/beyond-2011-report-on-autumn-2013-consultation--and-recommendations/national-statisticians-recommendation.pdf>

[2] A similar approach is being taken to the 2021 Censuses and use of administrative data in Scotland and Northern Ireland.

[3] <http://www.publications.parliament.uk/pa/cm201415/cmselect/cmpublic/601/60104.htm>

INTEGRATING GEO-REFERENCED DATA FROM DIFFERENT SOURCES: LIVESTOCK SURVEYS AND ADMINISTRATIVE DATA

C.Sermoneta, D.Fusco, V.Moretti

ISTAT, Italian National Statistical Institute, Rome, Italy

The statistical data collection is based, more frequently, on combined information from different sources, both to increase the knowledge and to ensure a higher quality of statistical data.

In particular the geo-referenced database provides the possibility of their integration through the use of geographic information systems.

In models of small area estimation in which it takes georeferenced data, the estimates are more reliable respect to the use of traditional methods. The creation of a national farm register, to use as a reference list as support to surveys, implies the need to have a unique code to identify the farm. For example, in the estimate of the livestock, the administrative data from registry can be associated with survey data to obtain more detailed information about the kinds of farms involved in the study, about a particular animal disease or about the environment (eg. localization of polluting emissions, information on the use of the territory, the presence of harmful substances related to the disposal of organic waste, et).

The availability of the geographical coordinates of the farms provides an accurate and objective linkage between the database collected by administrative sources. In this way it's possible to go beyond some limits of matching procedures done using the identification number or the VAT number like matching variables. In fact, using multiply sources we could have different definition of the frame, so that a farm can be identified as technical economic unit and as a geographical location, with significance in the difference of the estimates.

This paper presents the results of some analysis on geo-referred area using map charts. They will illustrate how the territorial reference improve the integration of data from different sources and provide auxiliary information and estimates ranges suitable implementation techniques of record linkage, extensively used for the administrative data for statistical purposes.

[1] Report of the Australian Bureau of Statistics on developing a statistical-geospatial framework, (2012) <http://www1.unece.org/stat/platform/download/attachments/81297858/2013-2-ProgReview-E.pdf>

[2] ESS Coordination of Geospatial Information and Statistics, (2012) [http://www.czso.cz/dgins2012/dgins.nsf/i/session_iii_eurostat/\\$File/Contribution_Eurostat%20rev.pdf](http://www.czso.cz/dgins2012/dgins.nsf/i/session_iii_eurostat/$File/Contribution_Eurostat%20rev.pdf)

[3] Linking of geospatial information to statistics and other data, (2013) http://ggim.un.org/docs/meetings/3rd%20UNCE/E-C20-2013-9linking%20GI%20to%20Statistics%20Summary_en.pdf

[4] Kuzma I. "An interactive web mapping application for presenting and dissemination of small area statistics"

http://www.cros-portal.eu/sites/default/files/NTTS2013fullPaper_115.pdf

Keywords: geo-referenced, geospatial, integration, administrative data.

ADMINISTRATIVE DATA AND SURVEY DATA ON ELECTRICITY USE IN HUNGARY

A. Gerencser

Hungarian Energy and Public Utility Regulatory Authority

The official energy statistics is compiled solely by Hungarian Energy and Public Utility Regulatory Authority as Other National Authority (ONA) in Hungary.

The annual energy statistics covers energy use statistics broken down by energy products (including electricity) and activity sector by NACE groups. Data are traditionally compiled based on an end-user sample survey.

On the other hand, electricity suppliers have also recently started providing detailed data on electricity supplied to end-users, which can be used for statistical purposes by matching via the General Business Register.

The paper briefly describes the method of the administrative data collection, the overall quality of data obtained compared to survey data, the potential benefits and limits of using these administrative data for statistical purposes.

Keywords: administrative data, energy use, Other National Authority.



Chair: Maria João Zilhão

QUALITY GUIDELINES

M. Herzner

Federal Statistical Office (Destatis), Wiesbaden, Hesse, Germany

The development of Quality Guidelines is a result of an internal analysis concerning Destatis' fulfilment of the ESS Quality Assurance Framework and was further strengthened by the latest Peer Review. Two improvement actions deal with the topic:

- To develop Quality Guidelines specifying quality principles and objectives for each part of the statistical production process according to the GSBPM.
- To develop a strategy for the monitoring of the application of the Quality Guidelines.

In addition to external users of our statistics, the Quality Guidelines are particularly aimed at the staff of the statistical offices of the Federation and the Länder. They identify quality measures for each sub-process of the GSBPM that are to be considered for all statistics. The core elements of these quality measures are then summarised in a verifiable checklist. Thus, the Quality Guidelines have to be as concrete and action-oriented as possible. They go beyond the description of methods and instruments already being used and aim to set the desired quality standards for the process of statistical production and show how these can be met. The Quality Guidelines will not only serve as a reference, but also as an impetus for process improvements and for promoting quality assurance tools not yet being comprehensively used.

Besides the identification of quality principles, the Quality Guidelines outline the institutional and organisational framework as the basis for data quality management and describe how statistics are being produced in the federal statistical system referring also to the Code of Practice and the Quality Assurance Framework.

Even though Quality Guidelines have been in use in different NSIs for a number of years, as far as we know, these do not directly map methods and instruments to the GSBPM sub-processes. This paper presents the German approach, its benefits, challenges and the lessons learnt.

Keywords: Process quality, GSBPM, quality assurance.

MANAGING SURVEY QUALITY THROUGH PERFORMANCE MANAGEMENT: BUILDING A PERFORMANCE MANAGEMENT CAPABILITY

C. Cordes¹, S. Hostetter², R. Livergood¹, A. Squires¹, J. Treat²

¹ The MITRE Corporation, McLean, Virginia, USA

² U.S. Census Bureau, Washington, DC, USA

This paper describes how the U.S Census Bureau implemented Performance Management for measuring and reporting on survey quality in the American Community Survey (ACS or the Survey). The paper provides the rationale for implementing Performance Management; the challenges along the way; and key lessons learned. The ACS is a nationwide survey that collects and produces information on demographic, social, economic, and housing characteristics about U.S. population every year.

Prior to this effort, ACS did not have a formal and cohesive process for reporting and managing the performance of the Survey. As a U.S. federal statistical program, it collected and reported on a variety of measures, but none of these measures were used for the ongoing management of the Survey. In addition, they provided limited information when balancing competing elements of cost, data product timeliness, and respondent burden against survey quality. Senior management recognized the need to establish a broader and deeper understanding of the Survey's execution and performance from a sampling and non-sampling perspective. Performance management allowed them to manage survey execution in near real-time, to keep sight of important mission factors like public support for and perceived value of the survey, to facilitate a richer discussion with stakeholders, to assess the impact program and environmental changes, and to provide data users with a quality product.

Keywords: survey performance; measurement; governance.



THE OECD RECOMMENDATION ON GOOD STATISTICAL PRACTICE

Julien Dupont
OECD Statistics Directorate

The Organisation for Economic Development and Cooperation (OECD) recently adopted a Recommendation on Good Statistical Practice. The Recommendation is intended to provide a common reference to assess the quality of national statistical system and official statistics, which are fundamental for OECD statistical and analytical works. In addition, the Recommendation will also be the main instrument for the assessment of statistical systems of accession countries.

The Recommendation includes twelve recommendations structured in five main areas: (i) the institutional, legal and resource requirements that enable statistical systems to function; (ii) the methods, quality and processes of statistical production; (iii) dissemination; (iv) co-ordination and co-operation; and finally, (v) statistical innovation. Each of the twelve recommendations is presented with a set of indications of good practices that has been compiled based on the OECD's experience in accession reviews.

The Recommendation aims to complement existing code of practice and international standards currently applied by OECD Member countries, but also to be more specifically relevant for OECD statistical activities, for example in including good practices on the use of new sources of statistical information, or on the coordination of the statistical system.

This paper aims at presenting the Recommendation, and at providing an overview on the importance of quality for an international organisation as the OECD.

Keywords: quality, national statistical system, international organisation.

DATA QUALITY ASSESSMENT METHODS AND TOOLS IN SSO - MACEDONIA

Dejan Peeski

State Statistical Office, Macedonia

Quality is of primary importance in the world of statistics. The State Statistical Office of Macedonia (SSO) is committed towards producing timely and internationally comparable statistical data with high quality.

The main aim of this paper is to share SSO's experience in the implementation of different methods and tools for data quality assessment: quality reports, quality indicators, self-assessment checklists and user satisfaction survey.

The production of quality reports in SSO can be distinguished between user-oriented and producer-oriented quality reports. The user-oriented quality reports are based on the Euro-SDMX Metadata Structure and SSO is publishing the reference metadata on Internet in a standardized manner. The producer-oriented quality reports are based on the ESS Standard Quality Report Structure. The ESS standard quality and performance indicators are regularly calculated through the statistical process and are included in our quality reports. In order to further improve the production processes and statistical products for better compliance with the European quality standards, the SSO introduced the European self-assessment checklist for survey managers - DESAP as a process-oriented way to discuss and improve the quality of both: processes and products. Our office is focusing of increasing the use of administrative and other secondary data sources for the production of statistics. This approach makes our office highly dependent on the quality of those sources and this is the reason for introducing a checklist for the quality evaluation of the administrative data sources. User satisfaction survey is conducted regularly and the purpose of the survey is to obtain the opinion of different users about the products and services of the SSO.

As a conclusion we can declare that SSO has a well-defined systematic approach to quality assessment, and the quality is ensured by using different methods and tools which are harmonized with the ESS methods, tools and methodological documents.

Keywords: quality reports, self-assessment, user satisfaction survey.



Chair: Pieter Everaers

Discussant: Marie Bohatá

THE PEER REVIEW AS A MAIN DRIVER FOR STATISTICS AUSTRIA'S STRATEGY 2020

Thomas Burg

Head of Quality Management and Methods

With regard to the second round of Peer Reviews the exercise for Statistics Austria took place in October 2014. Since the edition of peer review reports containing the recommendations for the Austrian statistical system was close to the start for the development of the new strategic concept of Statistics Austria – the so called Strategy 2020 – it seemed evident that the improvement actions based on the findings of the peers became a most relevant input for the forthcoming strategic concept.

Starting by describing the Peer Review process at Statistics Austria, its main findings and the most relevant improvement actions the paper gives subsequently an insight in the process of the developing the Strategy 2020. It is shown how the recommendations of the peers were translated into important strategic measures aiming to improve the current situation and to enhance the compliance to the European Statistics Code of Practice. A second aspect worth to be considered is that not only the measures but also the structure of the process of developing the strategy did gain from the Peer Review by integrating various organizational layers into the process.

Finally taking into account the measures of the Strategy 2020 not triggered by the Peer Review recommendations the paper proposes possible issues or a revision of the Code of Practice and enhancements for the Quality Assurance Framework.

Keywords: 2nd Round of Peer Reviews, European Statistics Code of Practice, strategic development of an NSI.



TRUSTWORTHINESS, QUALITY, VALUE: WHAT ASSURANCE DO INDEPENDENT ASSESSMENTS OF CODE-COMPLIANCE GIVE?

R. Laux
UK Statistics Authority

This paper explores the role that independent assessments of statistical systems – including the ESS peer reviews - have in enhancing trustworthiness. It argues that such reviews are necessary but not sufficient – other steps are also needed (at the European and national level) to enhance trust; and that while a focus on Trust is important, the ESS should be building on the foundations of trust and quality, and concentrating increasingly on ways to enhance the value of European Statistics.

Keywords: Trust, peer-review, assurance.



ESS PEER REVIEWS: AN EFFICIENT MEANS TO IMPLEMENT THE EUROPEAN STATISTICS CODE OF PRACTICE?

Kirsi Karkkainen, Luis Del Barrio
Eurostat, European Commission, Luxembourg

The European Statistical System (ESS) underwent a second round of peer reviews in 2013 - 2015 to monitor the implementation of the European statistics Code of Practice. The main objectives of the peer reviews were to improve the efficiency and credibility of the ESS, to strengthen its capacity to produce European statistics, to provide reassurance to stakeholders both about the quality of European statistics and the trustworthiness of the ESS, and to give an inside view to the producers of statistics on progress made.

A specific ESS Task Force developed a methodology for the peer reviews, based on an audit-like approach. Organisation was outsourced to the private sector, with the contractor finding and hiring independent peer reviewers and managing the exercise. Eurostat monitored the work of the contractor and ensured quality control, but it was not involved in the work of the peer reviewers.

This paper presents the methodology and how it worked in practice. It explains the reasons for opting for an audit-like methodology, sheds a critical look on what has worked best and what less and draws certain conclusions for the future. It will present a view on whether the methodology contributed not only to the efficient run of the exercise as such but also whether it improved the quality of the Peer reviews as a tool for enhancing the quality of the ESS.

Keywords: European Statistics Code of Practice, peer review, audit-like methodology.



Chair: Miguel Ángel Martínez

ENHANCING THE FOUNDATION OF OFFICIAL ECONOMIC STATISTICS WITH BIG DATA

B. Dumbacher, R. Hutchinson
U.S. Census Bureau, United States

Official economic statistics produced by the United States Census Bureau have long served as a high-quality benchmark for data users. To maintain this quality and enhance the foundation of its economic programs, the Census Bureau has begun exploring the potential of Big Data sources such as credit card transaction data, point-of-sale data, and publicly available building permit data. While this type of data may allow the Census Bureau to improve the timeliness, geographic detail, and product-line coverage of its economic data products, they do introduce concerns of transparency and consistency. This paper covers the Big Data findings that the Economic Directorate of the Census Bureau has discovered thus far as well as the Directorate's Big Data vision for the future.

Keywords: Big Data, Official statistics, Economic statistics.



BIG DATA AND THE INTEGRATED PRODUCTION OF OFFICIAL STATISTICS

A.Ö.Karlsson

Statistics Iceland, Reykjavík Iceland

A difficult issue for many NSI's is that their production processes are still heavily influenced by the stovepipe approach to statistical production where each statistic is produced by using its own production process. The result of this is a highly inefficient and costly production process without any form of standardization.

One possible solution for this problem can be sought within the framework of big data. Big data is one of the issues dealt with in the ESS vision for the year 2020 where it is categorized under the heading of new data sources. However the framework of big data is more extensive than simply a new source of data.

One of the keys of the big data framework is its scalability and flexibility, e.g. no assumptions are required for data and theories are irrelevant. This element of the big data movement might be very beneficial for NSI's especially because of the increased focus on processes which are used for different types of statistical products, moving away from the stovepipe approach for the development and production of official statistics.

This presentation will explore to what extent different statistical processes can stand to benefit from the big data framework with the aim of higher level of harmonization of methods across production processes and, therefore, relying less on the stovepipe approach. Examples will demonstrate how this has been done within Statistics Iceland, including (but not limited to) the use of statistical learning techniques, the use of big data sets for different statistics and big data collection methods applied to small data. The focus is on how this can be successfully implemented within the framework of the European statistics code of practice and the current legal framework.

Keywords: Big data, statistical production, efficiency.



COMPARATIVE ASSESSMENT OF THREE QUALITY FRAMEWORKS FOR STATISTICS DERIVED FROM BIG DATA: THE CASE OF WIKIPEDIA PAGE VIEWS AND AUTOMATIC IDENTIFICATION SYSTEMS

Albrecht Wirthmann, Fernando Reis, Loredana di Consiglio, Bogomil Kovachev, Michail Skaliotis
Eurostat, European Commission, Luxembourg

National and international statistical agencies are currently experimenting with the production of statistics derived partly or entirely from big data sources. At the same time there have been initiatives in the official statistics community and elsewhere to extend existing quality frameworks to statistics which production involve the use of this type of data sources. UNECE's suggested framework for the quality of big data and Eurostat's accreditation of big data sources as input data for official statistics are two examples in this regard. The framework proposed in the report on big data of AAPOR (American Association for Public Opinion Research) is an example coming from outside official statistics. These frameworks have been developed based mostly on theoretical considerations, even if early experiments have provided some input. In this paper, we propose to enrich the experience in the application of these frameworks to particular use cases of statistical products based on big data sources in order to assess their suitability, feasibility and completeness. We apply these three quality frameworks in the context of "experimental" cultural statistics based on Wikipedia page views and to data from Automatic Identification Systems (AIS) used for the production of transport statistics.

Keywords: Big Data, Quality framework.

MEASURING REPRESENTATIVENESS OF INTERNET DATA SOURCES THROUGH LINKAGE WITH REGISTER DATA

M. Beręsewicz^{1,2}

1 Poznań University of Economics and Business, Poznań, Poland

2 Statistical Office, Poznań, Poland

New data sources, in particular big data and the Internet have become an important issue in Official Statistics [1]. The non-statistical character of these data sources requires that they should be assessed before they can be incorporated in the statistical system. Moreover, despite their size, new data sources are non-probability samples that often suffer from the self-selection error [2]. Nonetheless, the crucial question from the viewpoint of survey methodology is the concern about the representativeness of these data sources. The answer is often challenging and its measurement is not widely discussed in the literature. This paper aims to bridge this gap by proposing an approach to measure the representativeness of new data sources.

The main goal of the paper is to assess the representativeness of Internet data sources through probability linkage with register data. To achieve this aim selected websites publishing ads for residential real estates available in Poland will be used and linked with the Register of Prices and Values of the Real Estate Market. The main underlying assumption is that an independent source that fully covers a similar target population is available. In addition, the adoption of R-indicators will be discussed [3]. A detailed description and results of the study will be presented using data for the city of Poznań, Poland.

The paper and the research has been financed by the National Science Centre, Poland, Preludium 7 grant no. 2014/13/N/HS4/02999.

Keywords: Internet data sources, representativeness, probabilistic record linkage.

1. Daas, P. J., Puts, M. J., Buelens, B., & van den Hurk, P. A. (2015). Big Data as a source for official statistics. *Journal of Official Statistics*, 31(2), 249-262.
2. Bethlehem, J. (2010). Selection bias in web surveys. *International Statistical Review*, 78(2), 161-188.
3. Schouten, B., Cobben, F., & Bethlehem, J. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, 35(1), 101-113.

Chair: Pieter Everaers

THE DEVELOPMENT OF A COP FOR THE ENP SOUTH COUNTRIES

Claudia Junker

Eurostat, European Commission, Luxembourg

Eurostat will present a description of overall institutional framework and the process of developing a Code of Practice in a different regional context and in a challenging institutional framework as it is present in the ENP South countries. This different institutional framework and context are characterised by a different understanding of the professional independence of an NSI, of the content of official statistics, what it should serve and how it should be compiled. The concept of professional independence is not yet rooted in the society of the countries and statisticians have to face a completely different understanding of the concepts outlined in a Code of Practice from the side of policy makers and the society at large.

The process of developing a Code of Practice in such a different and very challenging institutional context was characterised by an intensive and long-lasting discussion and consultation process, with countries stating that they already implement the Code of Practice without understanding the meaning of each and every principle and indicator of the Code of Practice. The process therefore, started with a very detailed explanation of the overall institutional framework of the ESCoP, of the meaning of each principle and indicator, of the challenges to apply them in every-day work and a thorough check on how far the framework, the principles and indicators were indeed already applied and implemented in the countries. It ended with a detailed discussion on each principle and indicator and how realistic it was to implement them in the ENP South countries, if not in the short run but in the medium and long term. The challenge in this discussion was to find the best balance between being realistic and sufficiently ambitious. This process was finalised with a thorough discussion and review of the adapted text of the principles and indicators at the high level Forum meeting of the EU-Mediterranean statisticians and its final adoption by the same Forum in May 2015.

The paper will finally draw a few lessons learnt from the process of developing such a Code of Practice for a different context and institutional environment in order to provide some ideas and proposals on how such a process could be launched also in other regions.

Keywords: institutional framework, Code of practice, adaptation to a different context, ENP South countries.



PCBS EXPERIENCE IN IMPLEMENTING EUROPEAN CODE OF PRACTICE (COP)

Ali Hussein

Palestinian Central Bureau of Statistics (PCBS)

This paper discusses three main sections about the implementation of the European Code of Practice in PCBS.

First, the historical overview for the implantation of the COP in PCBS, through a light peer review mission, its report and recommendations. The mission's report concluded that PCBS is on the right track, and suggested a wide range of proposals for further development. Afterwards, PCBS has transferred a list of improvement measures to specific mandates, and sets of mechanism for following-ups and implementation during the period 2013-2016. Secondly, the advantages and challenges of code of practice implementation in PCBS such as:

- Increasing the coordination role of the national statistical institutes within the same statistical system.
- Highlighting good practices suitable to foster compliance with the Code of Practice.
- Recommending improvement actions needed in view of fully complying with the Code of Practice.
- Enhancing the quality of the statistical outputs, such as implementing the EFQM.
- Development of a strategy to achieve the long-term financial independence of the PCBS.
- Developing a master plan to increase the contribution of administrative bodies for the production of official statistics of the NSS.
- Developing a master plan to increase the use of data from administrative sources.
- Developing and approving a strategy for improving the statistical methodology used by other producers of official statistics within the NSS.

Finally, this paper will present applying the Code of Practice for the European Neighbourhood of South Countries inside PCBS, and within the National Statistical System, to enhance the user confidence in statistics at the national and international levels.

THE IMPLEMENTATION OF THE CODE OF PRACTICE: MAIN CHALLENGES: MOROCCO CASE

H. Fdhil

Statistics Directorate, High Commission for Planning, Rabat, Morocco

At the end of MEDSTATIII program, Eurostat in a close cooperation with the ENP south countries carried out a reflection on how to build a permanent framework of the UE MED statistical cooperation. Through this exercise, the quality in statistics has emerged as a strong priority. Since then, the process has evolved towards the setting up of the UE MED Quality working group chaired by Eurostat, the adoption of the CoP for the ENP south countries and finally converges on developing approaches to implement it especially, in the challenging institutional contexts. At national levels, a particular dynamic has been settled in this regards. This paper aims to emphasize the ongoing developments in Morocco regarding the review of the statistical law, setting up the mechanisms for a constructive inter-institutional coordination and the development of the NQAF. The legal framework was revised to fit with the challenges that the NSS is facing in terms of multiplication of the statistical sources, strengthening the statistical production in specific fields and establishing the rules for inter-institutional coordination. The addition of the principle 16 in the adopted CoP responds strongly to this national need. The development of the NQAF is being developed using a pilot process aiming the valorization of the practices developed during more than 7th decades of statistical work which meets many CoP principles but also the identification of the remaining gaps. The approach is focusing on an appropriation of the processes, developing the NQAF and planning its implementation in a progressive way. The articulation between the works done or to be developed at regional and national levels is crucial for an optimization of the efforts. The development at regional level of common tools, the strengthening of national expertise, the sharing of UE and MED experiences would help the processes engaged at national levels.

Keywords: Institutional framework, National Quality Assurance Framework (NQAF), Code of Practice, Morocco.



IMPLEMENTATION OF CODE OF PRACTICE FOR THE ENP SOUTH COUNTRIES- EXAMPLE: THE ACCESS AND USE OF ADMINISTRATIVE DATA IN MOROCCO

M. Bouazza Bouchkhar
High Commission for Planning, Morocco

Statistics producers in Morocco are increasingly making use of administrative sources for the production of official statistics. Thus the High Commission for Planning (HCP), since 1975, has created a specific structure for the collection, processing and dissemination of administrative data on different sectors. Because identifying administrative data quality issues is an important aspect of preparing the data for use, HCP has developed metadata on administrative data collected. After its exploitation, several indicators have been developed, which allows having information on different dimensions of the quality.

However, there are some challenges to wider use of administrative data such as: access to certain administrative data, increasing privatization of public functions, growth of private sector data, user interest in new types of data, etc. This accounts for the importance of the review of the regulatory and institutional framework of official statistics in Morocco.

Recently, many quality initiatives were launched. At national level, there are the project of implementation of National Quality Assurance Framework by the HCP and the new law on statistics. At regional level, there is the implementation of the Code of Practice for ENP South countries. These quality frameworks and actions will improve the institutional environment and allow to systemize existing methods and to implement a new ones, such those relating to the use of administrative data.

The paper will present an overview of the institutional, organizational and technical frameworks and the process of using administrative data. Also, the paper will focus on some practices applied and will be applied in Morocco, especially those concerning the use of administrative data, and necessary for the implementation of the newly adopted CoP.

Keywords: Institutional framework, administrative data, Code of Practice.



THE STATISTICS CODE OF PRACTICE IN JORDAN: CHALLENGES AND FUTURE

Mohammad Khalaf
Department of Statistics, Amman, Jordan

Statistical code of practice is considered important tool gives deep evaluation for the status of official statistics. In Jordan, evaluation of CoP application revealed that more than 90% of the principles and indicators are practiced. Despite this fact, the application faces many challenges. The lack of involvement of employees in the process of evaluation and improvement is the first challenge. Even though, low awareness of the need of CoP in statistical work is another serious challenge. Moreover, 'who-what-when' is another challenge faces the application of CoP and its follow up to improve its application in the future. Top and middle management should be aware of CoP as a doctrine of the statistical office. This maximizes the role and the burden of top and middle management. The lack of practical tools to measure CoP extent application is another challenge. Moreover, CoP requires building strategy that explains the methods of application and measurements as well as the methods of improvement per time. The study recommended international cooperation to find out general tools for measurement, application and improvement to help different countries to have first step on the track.

Keywords: Code of Practice, evaluation, challenges.



THE IMPLEMENTATION OF THE CODE OF PRACTICE: THE ALGERIAN CASE IN THE CONTEXT OF USING NEW TECHNOLOGIES FOR THE 2018 CENSUS

Tarik Bourezgue
Office for National Statistics, Algiers, Algeria

Nowadays, to achieve progress in this era of data revolution, it is necessary to make use of the rapid emergence of new technologies. This is, once more, a challenge for official statistics. Algeria is planning to carry out its sixth General Census of Population and Housing (RGPH) in 2018. Based on the context described above, the development of our quality approach relies essentially on the capitalization of cooperative work with Eurostat on the principles of the European Quality Assurance Framework for official statistics and the implementation of the Code of Practice. As ONS intends to use, for the first time, tablets (equipped with mobile chips) as a collection tool instead of the paper questionnaire, we need a system that segregates data as and when data is entered. Another challenge in this method is the accuracy of data entered: the mobile applications environment has a multitude of particularities. This requires an adaptable approach for performance analysis in terms of quality management. In driving the process step by step, to optimize its management, our paper will address the issues related to the implementation of the Code of Practice, and its adaptation to our national context, so as to insure a better quality for the census process in identifying the relevant features (e.g. interaction time, volume of provided information, error management, data security, transaction security, ...) and the properties that should be highlighted for ensuring the implementation of the Quality Assurance Framework in a practical environment.

Keywords: Census Mobile Application, Quality Assurance Framework, Code of Practice, Algeria.



Chair: Roberta Pace

A COMPETENCY MEASUREMENT MODEL

Dario Russo
Central Bank of Italy

A model for Competency Mapping and Measurement allows information to be obtained about the skill appropriateness associated with different processes; such data makes it possible to assess staff, to better use resources, and to optimise training activities and HR management.

In the recent organisational research literature, Competency Mapping is attracting the attention of researchers. Competency mapping is increasingly used in the field of HR development and empowerment: due to the intensive use of technology, knowledge is considered as a key to the achievement of competitive gains, not only in the provision of services, but also in the more traditional sectors of production of goods and industrial products.

In this paper, a model for Competency Mapping and Measurement is presented. This approach allows information to be obtained about the level of appropriateness of the skills associated with different processes. These data make it possible, inter alia, to analyse the morphology of the business processes, to assess the staff, to better use and distribute the resources over the processes, to promote the mobility of employees across the firm and to optimise training activities and HR management (resource-based view of the firm).

Keywords: Business Process Modelling, Competency Mapping, HR Development.



USING A BUSINESS ARCHITECTURE FOR IDENTIFYING THE COMPETENCE NEEDS OF A STATISTICAL INSTITUTE

Piero Falorsi, Nadia Mignolli, Giorgina Brown
Italian National Institute of Statistics (Istat)

This paper shows how a Business Architecture Activity Model is called to play a central role in the modernisation program of statistical information production, and how it can help define the human resource competencies that underpin an organisation's ability to conduct its business.

In the last decade, official statistics has undergone a dramatic shift both in the production model and in its output. The traditional chain, based on the vertical integration of different survey-specific tasks carried out to collect, process, analyse and disseminate statistical data, has become outdated. Together with many other statistical institutes in the world, the Italian National Institute of Statistics – Istat – is moving from a stovepipe model to a new structure consistent with the industrialisation and standardisation of production processes. The new business environment enhances productivity, improves efficiency and effectiveness, supports better quality control, and allows for quicker innovations.

Within this context, Istat adopted a generic Business Architecture (BA) Activity Model which is called to play a central role in the modernisation program of statistical information production by defining in detail the activities required to pass from the as is to the to be situation and to produce statistical outputs. This reference model is being promoted at the international level to become a standard within Eurostat and UNECE.

Human resource competencies are among the most important capabilities that underpin an organisation's ability to conduct its business. This paper shows how such competencies should be defined on the basis of the organisation's business activity model: associating competencies to each BA activity or process can help identify competency gaps in an organisation. Future challenges include finding and developing new skills relating to increasingly complex knowledge demands and to new data sources such as big data, as well as adapting them to the growing use and integration of administrative data in statistical registers. Also essential is ensuring the Institute's employees are well trained to meet the need for high quality, timely, and cost-effective data and to face the new challenges ensuing from modernisation and organisational change, through both managerial and specialised skills development.

Keywords: Business Architecture Activity Model, Capabilities, Competencies.



THE EUROPEAN STATISTICAL SYSTEM'S ENTERPRISE ARCHITECTURE REFERENCE FRAMEWORK AND CAPABILITY MODEL

Emanuele Baldacci
Eurostat, European Commission, Luxembourg

The paper shows how an Enterprise Architecture framework is critical in the process of change management of statistical production that aims at harnessing new data sources to enhance the quality of data-based services provided to users; it applies the framework to the European Statistical System.

Official statistics is undergoing a major change in recent years triggered by the potential benefits stemming from technological innovation and increasingly more limited budget sources. Reaping the benefits of new data sources generated by humans and machines to streamline data production and enhance the quality and quantity of statistical services provided to users requires a radical rethinking of the value chain and the production process in statistical organisations. This paper shows how an Enterprise Architecture framework can be a guiding principle in redesigning the statistical production chain, harnessing the potential of cheaper data sources such as big data to provide more tailored statistical services to meet growing needs of different users. The key enabling factors of these transformation process are information standards, service-oriented production architectures based on interoperability and reuse and integrated metadata systems. The organisations' human capital and culture would need to adapt to the emerging environment to ensure that benefits from the ongoing transformation are fully captured and new capabilities are created to deliver more timely data-rich services with higher value for money.

Keywords: Change management, Enterprise Architecture, Big Data.



THE COMPETENCIES FOR A REGISTER BASED STATISTICAL INSTITUTE

Claus-Göran Hjelm¹, Thomas Laitila^{1,2}

1 Statistics Sweden

2 Örebro University

Discussion on requirements on human resources at NSIs when transforming statistics production towards a statistical survey system utilizing administrative data.

Since the seminal work by Jerzy Neyman in the 1930s, National Statistical Institutes have grown into professional organisation building competencies in producing statistics using sample surveys. This means the organisations are specialised in sample survey design and estimation, collection and processing of data retrieved from contacts with units in the population, and communicating results from sample surveys. This competence profile also affects formulation of research problems and survey designs. Administrative data and building of statistical register systems is by statistical institutes perceived as an option in preparing for future challenges and demands. However, the problem complex in producing statistics based on administrative data is completely different from the one faced with in sample surveys (Wallgren and Wallgren, 2014). Data is differently retrieved and processed, and estimates are derived and communicated in a different way. Present competence profiles at NSIs may therefore not be adequate for a transformation from a sample survey based to a register based statistical institute. This paper contains an analysis and discussion of future requirements on the competence profiles of NSIs when changing the organisation towards a statistical survey system built around a statistical register system. Administrative data held by public authorities are assumed a main source for registers. The analysis and discussion takes recent work on quality aspects of register statistics as a starting point. Work presented within the MIAD project (Di Consiglio and Falorsi, 2014) presents requirements and necessary resources for successful use of administrative data. Adding the frameworks on quality assessment suggested in Daas et al. (2010) and Laitila et al. (2011) gives a structure indicating requirements on competencies in the stages of retrieving data, processing data, analyzing data, estimation and communication.

References:

- Daas, P.J.H., Ossen, S.J.L., Tennekes, M. (2010). Determination of Administrative Data Quality: Recent results and new developments. Proceedings of Q2010 European Conference on Quality in Official Statistics, Statistics Finland and Eurostat, Helsinki, Finland.
- Di Consiglio, L. and P.D. Falorsi (2014). Different contexts for the statistical use of administrative data. Proceedings of Statistics Canada Symposium 2014: Beyond traditional survey taking: adapting to a changing world.
- Laitila, T, Wallgren A. and B. Wallgren (2011). Quality Assessment of Administrative Data, Research and Development – Methodology reports from Statistics Sweden, 2011:2.
- Wallgren and Wallgren (2014). Register-based Statistics – Statistical Methods for Administrative Data, 2nd Ed, Wiley, New York.
- Hjelm, Claus-Göran et al. (2007). Register-based statistics in the Nordic countries - Review of best practices with focus on population and social statistics. UNECE.

Chair: Pieter Everaers

OVERVIEW ON THE SET-UPS OF THE INSTITUTIONAL FRAMEWORK IN THE ENLARGEMENT AND ENP EAST COUNTRIES

Claudia Junker

Eurostat, European Commission, Luxembourg

Eurostat will present an overview on relevant set-ups of the institutional quality framework in the enlargement and ENP East countries, including the application of the principle of professional independence, the development of quality frameworks and coordination tools and mechanisms in the countries and relevant links to the definition of official statistics and the NSS in these countries as well as the challenges of implementing the ESCoP and coordination and how to possibly address them.

Information on the set-up of institutional quality frameworks in the countries as well as on the application of the principle of professional independence and the coordination practice of the national statistical systems has been collected during the light peer reviews and adapted global assessments implemented in the enlargement and ENP East countries during 2009-2013 through the regular monitoring of improvement actions. They have been discussed at length during various high level seminars, workshops and policy meetings.

The paper will thus describe the various institutional quality frameworks highlighting both good practice as well as failures in the application and different coordination tools, structures and mechanisms in the enlargement and ENP East countries, as identified in cooperation with the countries. The description of practices to develop institutional quality frameworks will focus on the statistical laws, their ongoing or recent revisions, measures to increase the professional independence by reviewing the subordination of the NSIs, activities to increase the user focus as well as "standard" coordination practices such as the statistical work programmes, statistical release calendars, websites, Memoranda of Understanding with other producers of official statistics and others. Some innovative practice such as the set-up of specific committees with other producers, specific working groups in the process of accession negotiations, the provision of validation tools to other producers, the use of the statistical society and others will also be described.

The paper will finally also review specific challenges in developing the institutional quality framework, increasing the professional independence of the NSI and in coordinating a statistical system in the context when the term "official statistics" is not yet well defined (or even still seen as "state statistics") and when there is certain misperception of the coverage of official statistics as well as of the professional independence in the countries. These challenges will be linked to some ideas on how to address them in the given context and what could be adequate tools and mechanisms to improve the overall institutional framework of the statistical systems.

Keywords: institutional quality framework, ESCoP, coordination of the national statistical systems, good practice, challenges, enlargement and ENP East countries.



GOVERNANCE STRUCTURE OF TURKISH STATISTICAL SYSTEM WITHIN THE CONTEXT OF COORDINATION MECHANISMS

G. Gedik, T. Altun

Turkish Statistical Institute, Ankara, Turkey

Turkish Statistical Institute will present the governance structure of Turkish Statistical System in general and how Turkish Statistical Institute coordinates the System through robust tools and mechanisms. Turkish Statistical System is a multi-partner system with Turkish Statistical Institute as being the main coordinator. In addition to Turkish Statistical Institute, other institutions are intensively involved in the production of official statistics, primarily the Central Bank of the Republic of Turkey, the Ministry of Finance, the Undersecretariat of Treasury, and the Ministry of Food, Agriculture and Livestock.

This paper will also cover Statistics Law of Turkey, adopted in the year 2005, emphasising the sound coordination role of Turkish Statistical Institute in the System by introducing multi-annual Official Statistics Programme and the Statistical Council as well as regulating other coordination tools such as classifications. Additionally, practical tools will be elaborated, as functioning of 50 working groups covering all representatives of Turkish Statistical System, the preparation of a national data release calendar, use of international data transmission system, and protocols and memoranda of understanding signed with the stakeholders and meeting with producers and users of official statistics. Cooperation and communication studies of Turkish Statistical Institute with stakeholders in the context of “public-public” and “public-private” basis will also be explained. The paper will touch upon endeavours of Turkish Statistical Institute regarding the potential use of administrative registers of other national authorities for statistical purposes.

This paper will finally review the main findings of the light peer review and peer review studies carried out for Turkish Statistical System in 2011 and 2015 respectively, taking into consideration the recommendations and improvements actions regarding the coordination structure of Turkish Statistical System and the role of Turkish Statistical Institute within.

Keywords: National Statistical Systems, Coordination, Governance, Turkey.

IMPLEMENTATION OF THE AMENDMENT TO REGULATION 223/2009 AND THE COORDINATION ROLE OF THE NATIONAL STATISTICAL INSTITUTES

Bojana Radojevic

Statistical office of Montenegro, Podgorica, Montenegro

Law on official statistics and official statistical system of Montenegro (Statistical Law) gives to Statistical office of Montenegro (MONSTAT) explicit responsibility for coordinating the National Statistical System. Montenegro has a comprehensive suite of instruments to operate a National Statistics System (NSS) in an effective manner. This suite includes the Statistical Law, the five-year Development Strategy of official statistics, the five-year Programme of official statistics, the Memorandum of understanding between MONSTAT and other official producers and the MONSTAT Annual Plan of official statistics.

Furthermore, Article 8 of the new Law gives MONSTAT a leading role in the coordination of the NSS, explicitly referring to coordination of official statistics as one of MONSTAT's activities.

During 2015, coordination role is empowered on effective way through follow instruments: Been formed Sector for coordination of statistical data sources, which coordinates with the official statistical system at the national level. Activities in the sector relating to:

- Coordination and cooperation with producers of official statistics;
- Coordination and cooperation with scientific research institutions, universities...;
- Coordination the preparation of legislation, secondary legislation of the statistical system, (the Law on Statistics, Development Strategy, a five-year Program of official statistics, the Memorandum of understanding, the Annual Plan, the Report on the realization of the Annual Plan, Calendar of publishing statistical data) and memorandums;
- Establishment of cooperation with administrative data sources and other relevant institutions in charge of keeping administrative records;
- Coordination of Council of Statistical System;
- Consultation with users in order to expand the user community and
- Other activities related to coordination of statistical system.

Adopted two procedures which regulate coordination of statistical system:

- Procedure of accession of statistics in the system of official statistics of Montenegro;
- Procedure on the use of administrative data sources for the purposes of official statistics.

Keywords: coordination of the NSS, cooperation, produces, quality.



NATIONAL STATISTICAL SYSTEM OF THE REPUBLIC OF BELARUS: COORDINATION ISSUES

Irina Mazaiskaya

National Statistical Committee of the Republic of Belarus

The institutional framework of the national statistical system of the Republic of Belarus comprises:

- the system of state statistics bodies (National Statistical Committee and regional statistical offices);
- other producers of official statistics (i.e. government bodies/ organisations authorised for maintaining state statistics in respect to the organisations subordinated to them or under their jurisdiction as well as on topics that fall under their competence).

The legal framework of the national statistical system functioning is the State Statistics Law of the Republic of Belarus which determines the procedure of organisation and maintenance of state statistics in the country and regulates the relations connected with the implementation of state statistical activities. In 2015 a new version of the Statistics Law was prepared. The draft Law strengthens the coordinating role of Belstat and specifies the system of producers of official statistics. It now clearly defines their mandate for primary data collection as well as the use of administrative data for the compilation of official statistical information.

In order to comply with the principle of independence when implementing state statistical activities as well as to reinforce the coordinating role of the state statistics bodies the membership of the Interagency Council on State Statistics and its charter will be approved by Belstat.

The tools ensuring the coordination of activities of other producers are as follows:

- elaboration of the annual statistical work programme which includes all producers of official statistics;
- endorsement of draft laws and regulations involving the issues of state statistics;
- approval of statistical survey tools;
- approval of official statistical methodologies developed by other producers of official statistics;
- concluding data exchange agreements.

Moreover, in order to reach the balance of interests of producers and users of official statistics, the Interagency Council on State Statistics – an advisory body on the issues of state statistics – was set up.

Keywords: institutional framework, tools for coordination of the national statistical systems, Belarus.

Chair: Rory Fitzgerald and Sarah Butt

SYNERGIES FOR EUROPE'S RESEARCH INFRASTRUCTURES IN THE SOCIAL SCIENCES (SERISS)

R. Fitzgerald, S. Butt
City University London, London, UK

Synergies for Europe's Infrastructures in the Social Sciences (SERISS) is an EC funded project to strengthen and harmonise the collection and curation of social science data across Europe. The project brings together leading European Research Infrastructures in the social sciences – the European Social Survey (ESS ERIC), the Survey of Health, Ageing and Retirement in Europe (SHARE ERIC) and Consortium of European Social Science Data Archives (CESSDA AS) – alongside organisations representing the Generations and Gender Programme (GGP), European Values Study (EVS) and the Wage Indicator Survey. Together, these infrastructures are undertaking a programme of work to address key challenges in cross-national data collection (including representativeness and translation), break down barriers between infrastructures through the creation of shared online tools, and equip social research for the future by exploring new forms of data collection.

Many of the issues tackled under SERISS will be highly relevant to national statistical institutes and other producers of official statistics. To maximise benefits from the project and successfully promote the role social statistics can play in informing debate on the challenges facing European society, it is important that the SERISS project forges links not only between the main academically-driven social science infrastructures but also with bodies involved in the production of official statistics across Europe.

This session will include a short introductory presentation giving an overview of the main strands of work under the SERISS project (www.seriss.eu). There will then be four substantive presentations highlighting some of the key areas where outputs from SERISS could feed into the production of high quality official statistics: advances in translation, automated coding of socio-economic variables, linking administrative and survey data, and issues associated with curating new forms of data.

Keywords: Cross-national surveys; data archiving, translation, socio-economic coding.



EXPERIMENT FOR TESTING QUESTIONNAIRE TRANSLATION METHODS IN THE EUROPEAN SOCIAL SURVEY (ESS): 'ASK THE SAME QUESTION' VERSUS MORE ADAPTIVE APPROACHES

B. Dorer

GESIS-Leibniz Institute for the Social Sciences, Mannheim, Germany

The European Social Survey (ESS) is a biennial academic survey fielded in 25+ countries since 2002. A British-English source questionnaire is translated into all participating language versions. The quality of the questionnaire translations is crucial for the comparability of the resulting data; therefore, the ESS has been using different methods to ensure high quality questionnaire translations since its first round: the source questionnaire is designed to minimize translatability and intercultural problems; the final translation process follows the TRAPD (Translation – Review – Adjudication – Pre-testing – Documentation) model, based on a committee or team approach. Assessment includes an external translation 'verification' and formal checks using the Survey Quality Predictor (SQP).

So far, the ESS has followed the 'Ask-the-Same-Question (ASQ)' approach: all translations should be as close as possible to the source text, ideally 'asking the same question'. Under the new European cluster project, Synergies for Europe's Research Infrastructures in the Social Sciences (SERISS), an empirical experiment will be carried out to test this ASQ approach versus another where national teams are given more leeway to adapt rather than translate the source questionnaire into their national contexts. The exact set-up of the experiment will be decided in 2016: 30 questions will be translated independently using both methods into a small number of languages. These translations will then be administered via a webpanel also developed under SERISS. The analysis will focus primarily on equivalence (using statistical techniques such as IRT) whilst also testing the conceptual space, using correspondence analysis, or a related technique.

This will be the first empirical evidence about the ASQ approach – which has been applied by several surveys assuming it allows for more equivalence – but the scientific community is lacking a proof that it does indeed produce better, i.e. more comparable and equivalent translations than by leaving national experts more room for adaptation.

Keywords: European Social Survey, questionnaire translation, empirical experiment.

MEASURING OCCUPATIONS: RESPONDENT'S SELF-IDENTIFICATION FROM A LARGE DATABASE

Kea Tijdens

University of Amsterdam / AIAS, Netherlands

Most surveys use an open-ended question to measure occupation, followed by office coding. This is expensive and time-consuming, and some texts can be coded only highly aggregated or not at all. Alternatively, in web-surveys or during the interview respondents can self-identify their occupation from a large database of coded occupational titles. For the coding quality the size of the database is important, given that a national labour market easily has 10,000's of job titles. The paper details the database.

For many years, the worldwide WageIndicator websites on work and wages apply this self-identification method. In its Salary Check web-visitors can identify their occupation and view the related salaries. In its web survey respondents are asked to self-identify their occupation. Both applications use the same multilingual database of approximately 1,600 occupational titles, all coded ISCO08 at five digits. Users can navigate the database by means of a 3-level search tree or by text string matching. Nine in ten use the latter.

As part of WorkPackage 8 of SERISS, the database is extended to 5,000 occupational titles, coded ISCO08 5-digit for approximately 35 languages, using the coding indexes from National Statistical Offices. These occupations are translated into English and their codes are compared. The subsequent solutions for 'same occupation-different code' problems are detailed.

An API (Application Programming Interface) is designed for the survey holders (free of charge during the SERISS project). For some countries, the database includes a gender filter, showing (fe)male titles to (fe)male respondents. A life demo of the database will be shown.

WP8 facilitates an occupation>>industry prediction algorithm for respondents' easy self-identification of industry, as the majority of occupations are industry-bound. The API shows respondents a list of the five most likely industries, including an option 'other', which then allows respondents to search an industry database.

Keywords: occupations, API database, web surveys.



ADMINISTRATIVE DATA LINKING IN THE GGP: ENRICHING ADMINISTRATIVE DATA WITH SURVEYS

Tom Emery
NIDI

In social survey research there is a great deal of interest in enriching survey data with administrative data sources such as income, employment or even criminal records. However, the added value of such data linking to administrative data sources is rarely considered. Using data from the GGS in Sweden as an example, this paper outlines ways in which administrative data can be enriched by linking to a social survey. First, the survey process provides an opportunity to attain consent from respondents to link data from normally dispersed administrative records (e.g. employment and birth records). In so doing, social surveys provide the key of consent for complex data linking and the subsequent ability to use administrative data to answer pressing social questions. Second, the social surveys collect data themselves which is wholly absent from administrative records but is nonetheless of interest to both administrative data holders and social scientists. For example, the GGS contains data on the distribution of household work which, when taken in conjunction with administrative data provides key insights into gender roles throughout society. Third, social surveys provide an opportunity to validate the data collection processes within administrative data collections. The analysis presented within the paper examines these three advantages and the degree to which they are evident in the case of the Swedish GGS. Given Sweden's strong administrative data tradition, it represents an example of how social survey data can supplement even a highly developed social statistics system.

Keywords: Administrative Data; Data Linkage; Survey Data.

NEW FORMS OF DATA - LEGAL, ETHICAL AND QUALITY ISSUES

V. Kvalheim¹, K. Utaaker¹, T. Emery², H. L'Hours³

¹ NSD, Bergen, Norway

² NIDI, Den Haag, Netherlands

³ UKDA, University of Essex, England

This paper will discuss an issue of direct relevance to European Social Science research, namely that of the major legal and ethical challenges facing cross national research which rely on access to large scale data on an individual level. In SERISS these issues are particular focused on social surveys and the use of new data types in a survey context, including biomarkers, administrative data and social media data.

Data generated by governments and national statistical agencies, the use of internet and web-based services and the emergence of a digital economy, are potentially valuable resources for researchers in the social sciences. However, their use is rather limited compared to their perceived research potential. The fact that these data are originally and primarily produced for purposes other than research creates some major concerns connected to the re-use of such data by scientists. The challenges are multifaceted and the solutions complex. However, legal and ethical issues are among the key issues raised by the growing amounts of quantitative data and new ways of working with data in empirical social science research. The demand for stronger protection and harmonization of rules and practices is high, particularly related to the use of data generated by or in relation to global communication network such as the internet. This has been an important driving force behind the proposed data protection regulation in EU. The question is how the new legislation balances the interest in privacy against the use for historical, statistical and scientific purposes.

In SERISS the impact of the new European General Data Protection Regulation and national data sharing practices, including that of national statistical agencies, will be described and monitored. Strategic policy guidelines and best practices to solve/overcome legal and ethical barriers to high quality, empirical research based high quality individual data, will be developed.

Keywords: Data content, data protection, big data.



Chair: Peter van de Ven

THE ROLE OF INTERNATIONAL ORGANISATIONS IN DEFINING STANDARDS THAT FOLLOW THE QUALITY REQUIREMENTS AND ENSURE COMPARABILITY OF DATA

P. Everaers

Eurostat, European Commission, Luxembourg

The ownership by countries of outcomes for the indicators of the Sustainable Development Goals, as agreed in the UN General Assembly (25-27 September 2015) results document, clearly increases their responsibility to deliver independently from policies. With this decision, the role of the international organisations in the reporting and monitoring has - for many indicators - changed from preparing the indicators to the role of advisor on methodology and production methods. As a consequence, the influence by international organisations on the quality of the indicators will clearly decrease, and responsibility on the reliability is more in the hands of the national statistical organisations.

European statistics follow a Quality Assessment Framework that covers mathematical, methodological, conceptual, institutional and contextual quality. In this context European statistics are different from statistics made in other European Commission Directorates General but also from statistics produced in many other regions of the world where the quality framework is less developed. Especially the issue of the contextual quality and the commitment of the Member States governments to allow the National Statistical Systems to follow the Code of Practice requirements, as this is introduced via the amended Regulation 223/2009 on European Statistics, should guarantee the production and dissemination of statistics which are transparent and independent from governments.

There are several ways of guaranteeing for other regions a contextual quality level comparable to that in the EU. One way is via a further strengthening of the quality frameworks in other regions of the world by consolidation the Fundamental Principles on this issue. Currently such a commitment is only referred to in the preamble of the Fundamental Principles revised in 2013. However, this will be a process that will take several years. Another approach could be by strengthening the role of quality assessor for specialised UN agencies. This could be done by more explicit checking (peer reviews, audits) of the country specific statistics in combination with a further strengthening of the standards and guidelines.

Keywords: Sustainable Development Goals, statistical standards, quality.

INTERNATIONAL CIVIL AVIATION ORGANIZATION (ICAO) QUALITY SYSTEM FOR THE AVIATION DATA COLLECTED UNDER THE STATISTICS PROGRAMME

Sainarayan, Ananthanarayan, Jessica Claudia Díaz de León Gómez
International Civil Aviation Organization (ICAO), Montreal, Quebec, Canada

ICAO utilizes the Integrated Statistics Database (ISDB) to collect, process, store and disseminate the statistics reported from its Member States through the different Air Transport Reporting Forms (ATR). During the design phase of the ISDB it was decided to build into the system various data quality control functions to assist ICAO's staff with this process by automating all the necessary calculations and producing a report for each ATR. These quality control processes were divided into two main activities:

- verification and validation,
- cross-check within the same data set and across other data sets.

The system verifies arithmetic relationships such as between the partial figures and the totals. If an error is found in any of the verification steps, the system does not allow the statistical officer to proceed to the next step (validation) until the errors are rectified.

Validation is a process whereby the system uses the reported data, and where available, information from the same or other data series, to calculate a number of parameters; for example, a percentage change over the previous period, basic parameters which can be derived from the data themselves and comparisons with similar reported in other form(s), to assess the quality of the data submitted.

The newly reported data are compared with submissions of the previous month and/or previous year, if available. For some data, the system verifies the consistency in the relationship between various data elements. Finally, the integrated nature of the system allows to run quality checks between different data sets ensuring homogeneity. Where the system identifies deviations or "errors" in the data submitted, States or other reporting entities may be asked to provide clarification and, if necessary, to submit amended data.

Keywords: Quality control, data production.



ENHANCING STATISTICAL CAPACITIES OF OIC MEMBER COUNTRIES TO ACHIEVE SUSTAINABLE DEVELOPMENT GOALS: THE ROLE OF SESRIC

M. Kulaklıkaya, Z.Z. Selcuk, A. Karaman
SESRIC, Ankara, Turkey

The strengthening of the National Statistical Systems (NSS), specifically of the National Statistical Offices (NSOs), by improving their capacities is required to monitor and report the achievements for the 169 targets under the 17 Sustainable Development Goals (SDGs). In this paper, the authors analyse the performance of the 57 member countries of the Organisation of Islamic Cooperation (OIC) against the Statistical Capacity Indicator (SCI) scores between 2009 and 2015 to exhibit the strengths and underline the challenges faced by them. The authors also highlight the role of the Statistical, Economic and Social Research and Training Centre for Islamic Countries (SESRIC), an OIC subsidiary organ, in contributing to the efforts of its member countries by organizing the annual sessions of the OIC Statistical Commission (OIC-StatCom), conducting capacity building activities under the OIC Statistical Capacity Building (OIC-StatCaB) Programme based on South-South cooperation approach, and carrying out thematic projects in the area of statistics in close collaboration with the regional and international agencies to bring about synergies for overcoming the “data gap” problem and achieving good quality official statistics. The authors finally show that the inputs provided by SESRIC in the new OIC Ten Year Plan of Action (TYPOA) 2025, a high level document prepared in parallel to the Post-2015 Development Agenda, privileges it to play significant roles as a partner and as a member of the Committee for the Coordination of Statistical Activities (CCSA) in the whole OIC region in facilitating the monitoring and reporting mechanisms of SDGs at the global, regional, national, and thematic levels.

Keywords: SDGs, National Statistical Systems, SCI, Statistical Capacity Building, International Organizations.



STATISTICS GOVERNANCE AND QUALITY ASSURANCE - THE EXPERIENCE OF FAO

S. Katz, P. Gennari

Food and Agriculture Organization of the United Nations

The Statistical System of the Food and Agricultural Organization of the United Nations (FAO) is decentralized across the organization: for many years each FAO technical department has carried out its own statistical programme of work and has maintained ownership of their data. This has resulted in limited statistical coordination and no formal quality assurance procedures at corporate level. To overcome this problem, there has been a major effort to strengthen FAO's Statistical Governance system over the past few years, starting in 2012 with the establishment of the role of Chief Statistician and the creation of an Inter-Departmental Working Group on Statistics, followed by the development of a corporate Statistical Quality Assurance Framework. More recently, this work has culminated in a new accountability framework for FAO Statistics under the Chief Statistician and a proposal to establish a country membership-based Global Commission on Statistics.

This paper presents FAO's experience in implementing this strengthened statistical governance system and in particular it describes the various issues and challenges that have emerged during a gradual transition from more informal coordination mechanisms to the more formal ones in place today. The important role of Quality Assurance in this process will also be emphasized, as well as the many benefits the strengthened governance system has brought to the organization, including the coordinated approach adopted by FAO in its submission of indicators for the Sustainable Development Goals agenda.

Keywords: FAO Statistics, Quality Assurance, Statistics Governance.



QUALITY ASSURANCE PROCEDURES FOR ESTIMATING PURCHASING POWER PARITIES

N. Hamadeh, M. Rissanen

The World Bank Group, Washington, DC, USA

The International Comparison Program (ICP) is a worldwide statistical initiative designed to estimate Purchasing Power Parities (PPPs). PPPs are conversion rates that convert to a common currency while equalizing the purchasing power of different currencies, hence allowing for accurate volume measures of economic activity as well as comparisons of standards of living across countries. Among many other uses, the Millennium Development Goals (MDGs) and Sustainable Development Goals (SDGs) rely on PPP-based indicators for their key goals on economic growth, poverty, energy and environment. To estimate PPPs for the world's economies, the ICP conducts surveys to collect price and expenditure data for all goods and services that make up the gross domestic product (GDP). Significant resources are allocated to data collection and validation in order to ensure PPP estimates of the highest possible quality. Validation procedures are carried out first at the national level to ensure intra-country data quality, subsequently at the regional level to ensure inter-country data quality, and finally at the global level to ensure inter-country data quality across the regions. Moreover, efforts are devoted to validating the resulting regional and global PPPs from an economic point of view to ensure their soundness. The quality procedures require close cooperation and collaboration between participating countries, regional coordinating agencies and the global coordinating agency. Frequent meetings, training sessions and validation workshops are organized to ensure consistency of methodology, comparability of prices and soundness of resulting PPPs. The ICP Book [1] and the ICP Operational Guidelines [2] were also developed as reference materials to ensure a common understanding on PPPs, underlying methodology, quality assurance and data validation procedures.

Keywords: International Comparison Program (ICP), Purchasing Power Parities (PPPs), Sustainable Development goals (SDGs).

References:

[1] World Bank. 2013. *Measuring the Real Size of the World Economy: The Framework, Methodology, and Results of the International Comparison Program—ICP*. Washington, DC: World Bank.

[2] World Bank. 2015. *Operational Guidelines and Procedures for Measuring the Real Size of the World Economy: 2011 International Comparison Program*. Washington, DC: World Bank.



EVALUATION OF UNFPA'S CONTRIBUTION TO THE 2010 ROUND OF POPULATION AND HOUSING CENSUSES

Benoit Kalasa, Rachel Snow, Edilberto Loaiza, Sabrina Juran
United Nations Population Fund (UNFPA)

An independent evaluation of the 2010 census round was conducted in 2014-15, examining UNFPA's support to countries. Compared to the 2000 round, more countries conducted a census in the 2010 round, including several high population countries in Africa, and countries lacking a census for more than 40 years. Hence, overall coverage of the global population increased substantially, and the number of countries or territories missing a census fell from 27 to seven.

The evaluation consisted of three levels of analysis: surveys completed by national statistical offices (NSOs) in 132 countries; semi-structured interviews, discussions and review of documents; and 13 country case studies. The evaluation was designed to gauge the extent and value of UNFPA's support to countries, and inform UNFPA's strategy for the 2020 round. UNFPA's support brought high value to countries, and was well aligned with the needs and priorities of partner governments. Support included: 1) advocacy to promote the census and ensure international standards; 2) assistance transferring UN guidelines into concrete planning and operations; 3) direct financial support and procurement; and 4) convening other partners, including South-South and triangular cooperation.

A key finding was that while support and coordination by UNFPA, governments and other stakeholders were effectively mobilized for generation of censuses in most countries, there was too little investment in utilization of these data. Partnerships were highly responsive at national level, but often ad-hoc and demand driven, suggesting a need for stronger forward-looking collaborative planning. There was diminishing technical support towards later stages of the census, in particular census analysis, dissemination and utilization. The lack of analytical capacities within NSOs and other potential users needs greater investment. A full review of the evaluation results will be shared in the presentation.

Keywords: Census, Population, Data for Development.

K. Osaki-Tomita. The 2010 Round of Population and Housing Censuses: a global review. UNSD



WHAT USERS WANT: THE IMPORTANCE METADATA FOR ASSURING THE QUALITY OF SDG STATISTICS

Jacob Assa, Matthias Reister
UNSD

The experience of monitoring progress towards the Millennium Development Goals (MDGs) has shed light on the importance of coordinating national and international statistical processes and outputs. In particular, the issue of gaps and differences between national MDG data and global estimates came to the fore in various contexts, including the Statistical Commission. The 2030 Agenda and its accompanying Sustainable Development Goals – for which indicators are currently being developed – will only increase the demands on national and international statistical systems to produce high quality, disaggregated, timely and consistent statistics. In terms of quality assurance, this paper focuses on the importance of metadata as an anchor for all dimensions of quality recognized by various quality frameworks. The paper argues that in order to overcome the challenges that the SDG monitoring exercise will present to global and national statistical systems, all efforts must centre on assuring the availability of proper metadata at both levels. Only such complete and consistent metadata, in turn, would be able to communicate to users whether the resulting mix of national and global data on SDGs is relevant (i.e. meets the users' current and potential needs), accurate (close to the true values), timely (minimal gap between the phenomenon described and data availability), punctual (minimal lag between the release date and agreed target date), accessible and clear (easy to access statistics, with the appropriate information and assistance, and to make optimum use of them), comparable (differences between statistics attributed to differences between the true values of the statistical characteristics), and – most importantly – coherent (degree to which statistical processes by which data were generated used the same concepts and methods), which will determine whether the statistical outputs can be validly combined and used jointly.

Keywords: quality, indicators, metadata, users.

Chair: Martti Hetemäki

Discussants: Konrad Pesendorfer and Ed Humpherson

IMPROVING THE COMPARABILITY DIMENSION OF EUROPEAN STATISTICS BY MINIMISING UNNECESSARY VARIATION

E. Baldacci¹, L. Javec², I. Stoop³

1 Eurostat, Luxembourg

2 Statistics Sweden

3 The Netherlands Institute for Social Research (SCP)/ESAC

One quality criterion for European statistics is comparability across countries.

Often there are trade-offs between comparability and other design aspects such as other quality dimensions as well as national best practices.

There are several factors that could affect the degree of comparability across countries, e.g.,

- Social and cultural environments
- Data collection modes and mixes of modes
- Questions and questionnaires
- Translation
- Methodological and financial resources

There are two main methods to increase comparability, input and output harmonization.

Unnecessary variation between countries should be minimised. This can be achieved by input harmonization such as ensuring conceptual equivalence, implementing scientific translation procedures, adapting questions, enforcing identical definitions, using probability sampling, and aiming for high response rates and small measurement errors.

Necessary variation will always remain, e.g. in sampling frames, response rates, experience of interviewers, prevalence of internet access, and the different languages in which the survey will be fielded. This requires detailed documentation, including, among other things, easily accessible questionnaire translations and adaptations, complete information on noncoverage, response rates that have been calculated according to standards, and information on survey modes on an individual level. One additional source of variation in European statistics is the increase in the use of non-survey data, such as register data. The availability, accessibility, quality and use of this type of data varies greatly across countries. A lack of comparability seriously diminishes the value and usefulness of European statistics. To strengthen comparability, surveys such as the European Social Survey (ESS), Survey of Health, Ageing and Retirement in Europe (SHARE), and Programme for the International Assessment of Adult Competencies (PIAAC) should be benchmarked. They all use well developed design strategies involving a mixture of standardization and flexibility and a strong central infrastructure.

Keywords: comparability, variation, benchmark.



BIG DATA: FROM NOISE TO EVIDENCE

Patricia O'Hara¹, P. Everaers²

¹ European Statistical Governance Advisory Board, Luxembourg

² Eurostat, Luxembourg

Big Data are becoming prominent on the agendas of statisticians. The availability of a large variety of Big Data information is expected to substantially influence the process of producing statistics and generating statistical information. Research methodology is heavily affected and requires a rethinking of the basic concepts of validity and reliability. The traditional survey based approach with operationalization from theory and hypothesis via theoretical concepts to the measurement and data collection will be replaced by an approach where more importance is given to an inductive fact and relations finding approach. The availability of a manifold of data, like in Big Data, that are not collected based on a conceptual model that has been developed from a theoretical model for a specific research question, but purely as the by-product of a digital development, creates also a different data environment for the concept of validity. However, also governance issues as the dependencies from data providers and statistical decision making process will be affected. All this can lead to a different production and dissemination process of statistics and therefore will affect the quality framework for statistics. Without doubt the ESS Code of Practice will be affected by this change in nature in the process of translating policy and research questions into the production and dissemination of statistics. As there will be additional stakeholders in this Big Data world also the mandate and tasks of ESGAB will be affected.

Keywords: Big Data, quality framework, Code of Practice.

INDICATORS: FROM EVIDENCE TO DECISION

W.J. Radermacher², Pilar Martín-Guzmán¹, E. Giovannini¹

1 European Statistical Governance Advisory Board, Luxembourg

2 Eurostat, Luxembourg

Evidence based decisions are more objective, not just rational in terms of the logic and opinion of a decision maker; they allow others to participate in the preparation and follow-up of the decision. The two-way-dynamic interaction between policy making and measurement has never been more visible as in the recent post-2015 SDG procedure. 'Indicators' represent a category of metrics, which are designed and produced for one specific purpose, at a particular point of time related to specific decision-making processes. In some cases, like for the analysis of sustainability, indicators may also be based on projections of future pathways.

The interplay between this specific purpose and the corresponding indicator or indicator set (e.g. dashboards) is characterised by a bundle of opportunities and risks, which need to be known and taken into account, if better decisions based on trustworthy evidence should be achieved. Statistical governance and quality principles have to reflect the increased importance and responsibility that is attributed to indicators. Guidance is also necessary for statistical institutions, which are under pressure to provide what policy makers expect: early estimates, models and composite indicators. The ESS Code of Practice could be reviewed in this regard.

Keywords: indicators, governance, Code of Practice.



Chair: Celia Santos

RE-ENGINEERING THE DATA COLLECTION AND DISSEMINATION PROCESS: STRENGTHENING OF DATA QUALITY WITH LIMITED RESOURCES: CASE OF MOROCCAN SDB STRENGTHENING

Mohamed Salimi

Statistics Directorate (High commission of Planning), Rabat, Morocco

This paper will present an innovative project in the field of modernization of the process of the statistical production and dissemination of statistical data. It consists on the reverse of the approach used in the Statistics Directorate to collect, treat, compile and disseminate statistical data.

Concretely, instead of the current approach which adopt the following chronological order: data collection, treatment, dissemination and the update of the Statistical Data Base (SDB), the new approach is based on a new logic : Data collection, treatment and validation, update of SDB then automatic generation of dissemination products.

To establish this model plan required a redesigning of our SDB by, among others improvements, building it on indicators-basis instead of series-basis and developing a module for generating automatically the annual statistical yearbook and an SDMX module for exchanging data on development indicators, including MDGs, with the United Nations Statistics Division. In addition, the new database offers the possibility to disseminate data on the intranet and the web for users.

The paper will explain the problems faced earlier, how they was fixed, the benefits in terms of data quality and in rationalizing used resources. The data quality issues faced will be explained; they are related to the treatment and validation processes and to waste time in repeated and manual tasks.

Furthermore, the paper will discuss the status of the project and the remaining steps to finalize. The other important and challenging specific objective to reach is to use SDMX for collecting data from the maximum of line ministries directly to the new SDB, completing the automation of the overall process in a context of reduced staff and resources and where ITS provide revolutionary solutions.

Keywords: Automation, Data Quality, dissemination.

QUALITY REPORTING FOR SEASONAL ADJUSTMENT WITH JDEMETRA+

G. Bruno¹, D. Buono², A. Ciammola³, D. Ladiray⁴

1 National Statistical Institute, Roma, Italy

2 Eurostat, Luxembourg

3 National Statistical Institute, Roma, Italy

4 National Statistical Institute, Paris, France

“Quality report” is an expression widely used in the seasonal adjustment context but they cover it covers various meanings and concepts. In fact, a different report is needed when you produce the seasonally and/or calendar adjusted series, when you define or try to improve the seasonal adjustment process, when you disseminate the data and when you use the data.

JDemetra+, the European seasonal adjustment software, incorporates a plug-in that allows customizing the quality report that meets the various user needs.

The presentation emphasizes several important points:

- the identification of the quality dimensions specific for seasonal adjustment;
- the choice of the relevant statistical indicators to check the quality of seasonally and/or calendar adjusted data;
- the setting of the weight of each indicator and the definition of a global quality indicator;
- the use of the report for selective editing during production time;
- the dissemination format(s) of the report compliant with European standards (ESS Standard for Quality Reports Structure and Euro-SDMX Metadata Structure).

Keywords: Seasonal adjustment, quality report, JDemetra+.



VISUALISATION OF DATA IN EUSTAT: A PARTNERSHIP WORK WITH THE UNIVERSITY OF THE BASQUE COUNTRY

J. Aramendi, E. Goni, I. Iztueta
Basque Statistics Office (Eustat), Spain

Statistical offices are one of the largest producers of statistical information and they face the challenge of disseminating data clearly, providing relevant data to all sort of users. An alternative to the classic data-table are graphics that try to synthesise the released data. So in 2010, Eustat decided to improve the graphics on the Website, our window. Visualisation can serve as a tool to provide better understanding and to boost the use of our official statistics. The project began with a grant focused on the visualisation of statistical data in collaboration with the university in order to analyse Eustat needs and to review European and international statistics organisations practices. In 2012, visualisation was already a statistical operation within the Statistics Law.

The work resulted in the standardisation and updating of the graphics used in press releases. These graphics summarise the majority of Eustat data series and are interactive to select the series, the period and other preferences. These graphics, developed in open code, are easy to update and can be reused as the data and the graphic elements are separated.

We have developed new products as population pyramids, with a selector for years and geographical area; as well as interactive maps for various subjects at local level. The graphs have become a useful tool that needs updating and maintenance. It is important to have the data in a suitable format and available for required disaggregation. It is also necessary to set up common criteria across the organisation.

Currently, and with a view to create a work group on visualisation Eustat-University, a Master thesis on Computer Engineering and Intelligent Systems is being carried out at the University of the Basque Country. This thesis is a resource to boost visualisation, which is a long-term project that requires continuous attention.

Keywords: Visualisation, Data integration process, Interactive content.

NEW WAYS FOR THE DISSEMINATION IN OFFICIAL STATISTICS: ACCESSING DATA THROUGH GIS

Enrique Morán Aláez, Mariano González Izquierdo, Jaime Garrido Vázquez
Basque Statistics Institute - Eustat

We have prepared a new way of accessing statistical data by means of a GIS. We think this new channel will provide users with a powerful tool to create and access to very detailed territorial data, not only for administrative or statistical areas but also to user-defined areas. This new application may also have other important consequences, such as increasing the number of users, promoting the image of our Institute and improving the satisfaction of our main informants by giving them special statistical products.

So we have created a new statistical database which includes data for each building; that is data about persons, dwellings, companies and establishments and their main characteristics (gender, age, facilities, economic activity, number of employees, etc.) tabulated at this level.

This web tool to access the data allows selecting a geographical area through a GIS; the users will be given the choice of selecting an area from an existing or an own layer, by polygons designed freely by themselves, by distances to a point or an axe or others.

Then they will be able to choose a subject, a table, and some or all the cells in the table, between those offered by the database, which are aggregated to that particular geographic area.

An important question is to check for disclosure control. Due to the free selection, the area can be very small or there can include very few units. In both cases there is a minimum that must be attained to be allowed to download the information. Otherwise the user will be informed that no information is given due to statistical disclosure control.

There is also access to the statistical units from economic directories, with special restrictions to comply with the statistical act. Some of these units are also presented as geographical layers.

Keywords: Geographical information, disclosure control, layer, free selection, territorial data.



TRENDS IN QUARTERLY NATIONAL ACCOUNTS ESTIMATES: TOWARDS A CONTINUOUS RE-ESTIMATION SCENARIO

H.Olaeta

Basque Statistics Office (Eustat), Spain

The need of timely, almost real-time, information on macroeconomic performance has grown substantially in the last years. This trend affects directly to the Quarterly National Accounts Systems, as the main short-term tool to identify the economic cycle turning points and to monitor the effects of different macroeconomic policies being applied in our economies.

The first estimation of the GDP is published within a month from the reference quarter in most statistics offices, often with very limited amount of observed information. As short-term information becomes available, different re-estimations of the GDP and its components are performed over consecutive dates. This continuous possible update of estimates can lead to confusion and might compromise in a way the former more static and robust data assimilation by our users.

In this work a comparative study of the dissemination policy of quarterly accounts in different statistical offices is performed and the size of deviations of subsequent estimates quantified and compared. There is a need of updating standard quality indicators to accommodate to this new scenario where re-estimation is not to be confused with the concept of revision, more related to detecting errors. This leads to the need of a revision of current dissemination formats which usually do not mention the quantity and quality of the information set used in each estimation point.

The advantages and risks of the growing presence of provisional official estimates are addressed together with the need of adapting quality assurance tools in short-term statistics that are subject to planned subsequent re-estimations.

Keywords: Quarterly National Accounts, Re-estimations and Revisions, Quality Control.

¿ACTIVE? WEB LISTENING. WHAT OUR USERS ARE SAYING?

José J. Zurikarai, Mikel Bilbao
Eustat

In the world of official statistics we continuously need to know how our users are, what they needs are and whether we fulfill their expectations with our products and services. It's a long time ago that user satisfaction surveys were launched, but we know better than others about their limitations and boundaries. Our data publication web services have big amounts of users or web-traffic with unknown origin. But that anonymous traffic creates itself more information in the web that can be identified. We can identify not only traditional Media companies making news with our data. Thanks to web 2.0 and social there are more and more web resources where our data can be disseminated. So we have all those intermediaries helping us disseminating our data results. But they also are active agents that may give value to our information or even undervalue it. So, in this scene web listening becomes the tool to be used, the tool that makes possible to identify and assess what is being said about our organization, our statistics, or about our product and services.

With this active listening we have the ¿proper? features to build the image of those users that are using and sharing our information with others, so that we can improve our work in the dissemination areas.

We will show the results obtained for EUSTAT during the last months. First, a brief idea of how we should organize the job, the construction of keywords. This is essential regarding official statistics, because it shows the difference that can be found between our technical use of language and how those concepts are used in the internet. Then some comments about things that were unexpected and that we found out thanks to this technique or about the expansive wave effect of social network.

Keywords: Web listening, Users knowledge, Tool.



Chair: Maria Isabel Lázaro

A NEW URBAN/RURAL CLASSIFICATION, USING GEOSPATIAL INFORMATION DATA IN KOSOVO STATISTICS

Idriz Shala, Burim Limolli

Kosovo Agency of Statistics (KAS)

The main objective of this paper is to analyse the distribution of the urban/rural resident population of Kosovo, as obtained from their 2011 population and housing censuses, according to their administrative criteria and on the basis of a new approach for data classification as well, and to compare the differences at national and regional levels. Selected census variables are tabulated by urban/rural modalities using administrative and non-administrative criteria, including the 1 km² grid-based typology recently adopted by the European Union.

The methodology of this study is organised into four subsequent phases: i) analysis and comparison of definitions and classifications of urban and rural populations, ii) definition of the 1km² grid-based urban population in Kosovo, and generation of a grid covering the entire territory of the country; iii) aggregation of the census population into grid cells on the basis of the place of residence of the enumerated population, and implementation of a grid-based dataset of census data organised in a Geographic Information System environment.

Keywords: population and housing census, census data comparability, Geographic Information System, grid data, population distribution, urban/rural classification, urbanisation.



HOW NSIS CAN CONTRIBUTE TO THE THIRD SECTOR'S EFFORT TO PURSUE SOCIAL AIMS

M. G. Calza, A. Carciotto

Istat, National Statistical Institute, Rome, Italy

The current economic crisis and the state budget constraints for welfare have contributed to the third sector's development in filling the gaps left by the shortage of public social aid. In fact, Civil Society Organisations and Social Entrepreneurs play an important role in solving social problems, promoting social cohesion, as recently recognised at a European level. The availability of data, speeded up by the advance of digital technologies, are crucial factors for their activity, often limited exactly by data shortage. Nevertheless, their request of detailed data, the barriers they encounter to access and easily interpret them often induce CSOs and SE to produce data by themselves, generating a supply of non-official statistics, affected by a lack of specialised skills, and thus changing the landscape where official statistics is operating.

This apparently conflicting situation between official and non-official statistics can be the occasion to develop new forms of partnership. In fact, CSOs and SE would profit from NSIs' expertise in producing data complying with quality issues. Similarly, NSIs would be encouraged to investigate other data sources as a complement to official statistics, through a combination of new and traditional data sources, as recommended by Lisbon memorandum. To take up the challenge, NSIs have to share experiences amongst themselves and invest in IT and methodological skills, if they wish to keep their leading role in producing information at the service of society.

Keywords: third sector, non-official statistics, partnership.



TOWARD IMPROVING QUALITY ADJUSTMENT METHODS FOR PRICE STATISTICS: MEASUREMENT OF PRICING PATTERNS AND QUALITY IMPROVEMENT RATIOS OF DURABLE GOODS IN JAPAN

N. Abe, Y. Ito, K. Munakata, S. Oyama, and K. Shinozaki
Bank of Japan, Tokyo, Japan

In order to secure output quality of price statistics, compilers must take serious effort to implement quality adjustment appropriately based on characteristics of each product. In this connection, we focus our attention on the nature that firms often set a price of newly-introduced product somewhat higher than prices of existing products. We believe the differences in prices between new and old products could be composed of both reflections of differences in qualities between products and opportunistic hikes in product prices to regain profitability. Estimating the quantitative impact of above factors on differences in prices is indispensable so as to make an appropriate quality adjustment; however, little is known about the impact empirically.

This paper examines pricing patterns over product life cycle and quality improvement ratios (fractions of a difference in prices between each pair of new and old products attributable to a difference in qualities between the same pair) with regard to major categories of durable goods in Japan. Those categories consist of eight commodities of home electrical appliances and twelve commodities of digital consumer electronics. Using hedonic regressions with a great deal of data obtained from a well-known internet price comparison service in Japan, we find that firms push prices of their products back in some degree to raise profitability at the timing when they launch a new product. A glance at the pricing patterns over product life cycle reveals that the impact of price pushback diminishes over time; however, the pace of falling in prices varies considerably among different commodities. The quality improvement ratio of each commodity shows a unimodal and fat-tailed frequency distribution and the mode of the distribution will be around 0.5. The results bring significant implication for improving quality adjustment methods for price statistics and understanding price-setting strategies adopted by firms deeply.

Keywords: price index, quality adjustment, hedonic approach.



FRAMEWORK REGULATION COSTS: BUSINESSES' PERSPECTIVE

P. Bordelo, P. Braz, S. Rodrigues
Statistics Portugal, Lisbon, Portugal

The growing integration of businesses in the global economy, poses new challenges and the need of constant attention to factors affecting their competitiveness. Among them, the framework regulation costs stand out. Framework regulation costs are negative effects resulting from rules, procedures, actions or omissions, that hinder businesses, and that are not attributable to the investor, its business or organization.

In 2015, Statistics Portugal launched a survey, focused on nine main areas, identified as potential areas of obstacle to businesses' activities: starting a business, licensing, network industries, financing, judicial system, tax system, administrative burden, internationalization and human resources. Companies were inquired on the levels of obstacle they perceived in the multiple areas and on the evolution between 2012 and 2014.

The global framework regulation cost indicator, which aggregates nine cost domains, scored an intermediate value (3.04 on a 1 to 5 scale), having been slightly higher in small and medium companies than in large and micro enterprises. By business sector, companies from accommodation and food services, construction and real estate, manufacturing and agriculture, forestry and fishing, perceived higher-than-average framework regulation costs. Among the 9 areas of costs analyzed, the main constraints to business activity were identified in the judicial system. Licensing and the tax system followed as the most problematic areas, and these three registered the highest values, regardless of the size and business sector of companies.

Keywords: framework regulation costs; global indicator; constraints to business activity.

ADMINISTRATIVE DATA IN SAMPLING AND CALIBRATING THE BASQUE LABOUR FORCE SURVEY

E. Goni, J. Aramendi, I. Iztueta, E. Morán, A. Serrano
Basque Statistics Office (Eustat), Spain

Our labour force survey, PRA, is a continuous survey that has been carried out since 1985 in the Basque Country, and is a panel with a 1/8 quarterly rotation. This survey was designed to obtain quarterly estimates and evolution of the labour status of the population in the Basque Country and its provinces.

In the most recent review of the survey, Eustat has included several methodological improvements which, in addition to keeping the main objectives, also aim to reach other specific objectives, such as: reconciliation with administrative sources, coherence with other statistical sources of the Basque Statistics Institute, a greater geographical disaggregation and estimates of certain subpopulations.

To achieve this, administrative registers from the labour market have been incorporated in the data framework, and the Cube Method (Déville and Tillé, 2004) has been used to balance the probabilistic sample in accordance with variables in our population and housing register: socio-demographic and labour characteristics. Additionally, the number of strata has been increased, and allocation in each stratum has been fixed based on nonresponse in previous surveys.

The information obtained from administrative registers is also taken into account during the weighting and calibration: population structure and employment and unemployment levels in administrative registers. The methodological change for calibrating survey data has involved a change in the tasks, for example working with record linkage.

We are going to explain the improvements applied in sample design as well as the calibration to the administrative sources. Our estimates are more accurate for various collectives, and moreover, several administrative sources and the survey have been reconciled, meaning we can explain the changes in employment flows.

Keywords: Labour force statistics, administrative data calibration, balanced sampling.

SHORT TERM INDICATORS DURING POLITICAL CRISIS: INDEX OF INDUSTRIAL PRODUCTION IN AS A CASE STUDY

E. Mustafa

Palestinian Central Bureau of Statistics, Ramallah, Palestine

Producing official statistics during unstable political and security conditions is a challenging task for many countries. During such critical conditions, the importance of detailed reliable figures increases to describe what is happening on ground. Index of Industrial Production IIP plays an important role in reflecting the economic situation in the country through observing the industrial activity as well as GDP growth.

In the case of IIP the only data source is the industrial establishments which become unreachable during political crisis like wars. To publish reliable figures that serve the current situation becomes a priority taking into account the indicator quality, accuracy as well as timing.

This paper introduces the Palestinian experience in estimating the IIP during 2014 war on Gaza Strip. The methodology that was followed took several directions. Many sources of information were aggregated together to reach an appropriate estimate for the missing data. The weight of a specific activity compared to other activities in the index, general information about the production process in each establishment, exports and imports of the targeted products and inputs, in addition to the nature of the product itself either basic or luxury good, all together were used in the analysis process. Moreover, following news and reports coming from the field of conflict was of great importance in drawing specific conclusions. Testing the differences between the estimated IIP and the revised IIP after raw data was available the results showed high level of accuracy.

The contribution of this papers is twofold. Firstly, it introduces a methodology to estimate the IIP and similar indicators during unstable political conditions. Secondly, the paper focuses on the accuracy of choosing the right methodology to publish statistical figures of high quality when the internationally recommended methodologies become inapplicable in similar situations.

Keywords: Index weights, IIP, basic and luxury goods.



Chair: Celia Santos

USING TABLETS FOR THE 2018 ALGERIAN CENSUS: MOBILE CENSUS APPLICATION QUALITY ASSESSMENT

Tarik Bourezgue

Office for National Statistics, Algiers, Algeria

Nowadays, to achieve progress in this era of data revolution, it is necessary to make use of the rapid emergence of new technologies. This is, once more, a challenge for official statistics. Algeria is planning to carry out its sixth General Census of Population and Housing (RGPH) in 2018. ONS intends to use, for the first time, tablets (equipped with mobile chips) as a collection tool instead of the paper questionnaire. This paper mainly concentrates on ONS' vision of census data collection through the use of mobile technology, namely 3G and 4G LTE technologies. So we need a system that segregates data as and when data is entered. Another challenge in this method is the accuracy of data entered: the mobile applications environment has a multitude of particularities. This requires an adaptable approach for performance analysis in terms of total quality management for software projects. In this context, quality is analyzed from the viewpoints of developers, of users and, respectively, of one who aims to recover an investment. Thus, our paper will address the issues related to the implementation of this mobile solution so as to present the census data meticulously with minimum hardship through a mobile application quality estimation model that requires the identification of features (e.g. interaction time, volume of provided information, error management, data security, transaction security, ...) and which properties should be highlighted for ensuring usability in a practical environment. The paper proposes to look at a set of indicators normalized on the $[0; 1]$ interval as to measure the application quality level.

Keywords: Census Mobile Application, Total Quality Management, Aggregate Indicator.



THE TRANSITION TO THE CAPI SYSTEM AND ITS IMPLICATIONS AND CONSEQUENCES FOR DATA QUALITY ON THE MOROCCAN LABOUR FORCE SURVEY

J. Guennouni, M. Aaibid

High Commission for Planning, Rabat, Morocco

This paper describes the conversion from the paper and pen interviewing (PAPI) to the computer assisted personal interviewing (CAPI) of the Moroccan Labour Force Survey (MLFS), and examines the implications and the consequences of move to the CAPI mode for the data quality. The MLFS is a national continuous survey of some 60,000 households in Morocco carried out by the High Commission for Planning (HCP), and aims to determine the volume and the main demographic, cultural, and socio-professional characteristics of the labour force.

In 2007, the MLFS went in the field using the CAPI mode. The transition from PAPI to CAPI has presented many challenges including, among others, designing a CAPI system compatible with existing fieldwork practice and data processing procedures, providing and improving high quality data for the wider user community, reducing the data cleaning time, improving the dissemination timeliness, optimizing the human and the material resources, and rationalizing the fieldwork.

To ensure the success of the transition to the CAPI system, a process of decentralization of data entry and data cleaning tasks related to MLFS has been, for several years, gradually launched. This process has successfully led to the implementation of this system. After a decade of using CAPI in the MLFS, and according to a wider user community, the data quality level has increasingly improved. The CAPI has offered many advantages such as: avoiding routing errors, offering a computerised method for coding, in the field, especially questions that need classifications, applying immediately the validity and consistency checks, saving cost in printing, in dispatching, and in data entry, and improving the dissemination timeliness. Moreover, the multitudes of events generated and recorded during the interview process are invested to optimize the fieldwork.

Keywords: CAPI, PAPI, MLFS.



INTEGRATED AUDIT ASPECTS IN THE HUNGARIAN GDP COMPILATION

Bálint Murai
Hungarian Central Statistical Office, Hungary

The paper describes the 3-year modernisation project of the National account's compilation system in Hungary. One of the main driving forces behind the re-engineering of the annual statistical process was to fully harmonise it with the EU's audit standard. As a result of the redesigned system data are produced, stored and validated in the framework of the Process Table which is the tool for the direct verification exercise of the GDP/GNI data in the EU. The paper presents the challenges and the implementation of this giant modernisation project, summarises the main achievements and the outline of the new compilation methodology.

Keywords: National accounts, GDP compilation, Process Table, GNI audits.



THE USE OF TECHNOLOGY IN STATISTICAL PRODUCTION

Jeremy Heng, Eugenia Goh, Ang Boon Heng
Ministry of Manpower, Singapore

In an advanced information age, National Statistical Agencies need to continuously review and improve their processes in order to meet the growing demand for statistics in policy-making and research. It is important that the production and dissemination of official statistics are objective, timely and reliable so that the government and community-at-large can make informed decisions. As the National Statistical Agency in charge of official labour statistics, the Manpower Research and Statistics Department (MRSD) of Singapore Ministry of Manpower has explored and harnessed technology to enhance operational efficiency and disseminate quality information. This enables the government to formulate policies and programmes to secure a better future for its people. It also helps job-seekers, employers and employees to make informed decisions on career and human resource matters. This paper examines several initiatives of the past, present and future in which MRSD has made use of technology in our statistical product cycle – data collection, processing, analysis and dissemination.

Keywords: technology, quality, efficiency.

POSDEM: THE SELECTION PROCESS BETWEEN PROBABILITY SAMPLING PLANS USING A SOFTWARE BASED ON EMPIRICAL SUPERPOPULATION APPROACH

G. Sánchez-Crespo Benítez
National Statistical Institute (INE), Spain

Software, POSDEM, permits the evaluation of sampling methods. The acronym could be translated as "Optimising the Selection Process in Probability Sampling". Sampling theory proves, under certain conditions, that if population has a linear trend, Yates method of end corrections presents a least expected error. If we have a linear trend with periodic variation, then modified systematic sampling is better, and in case of a parabolic trend or autocorrelated population, centered method is the best.

Several alternatives of systematic sampling and unequal probabilities sampling are frequently used. However, both may be distorted by the structure of the investigated population. Trends, cycles and inappropriate relationship between variables are present more often than it seems, which could even lead to losses in accuracy. In both cases, it is necessary to study the relationship between the sampling plans and the population frame before applying these methods.

POSDEM allows, for different methods, answering to, among other questions, what is the best plan for a specific sampling frame; which is the sampling size related for a significance level; what is the mean square error expected; what are the units that must be investigated; what happens if ...

An example, (see references) has been prepared in order to show the possibilities of this software. This example examines several sampling methods to know what happens if in the structure of the population unexpected changes occurs. For instance, it may be expected that population presents a linear trend but population actually has a polynomial trend, because there had been changes in the structure of the population since the last available information. When the hypothesis have unexpected changes, is there any sampling method robust enough? With this software, this question can be answer. Another interesting feature is the behaviour of the sampling methods when the degree of randomness changes.

Keywords: Systematic sampling, Variable sampling interval, Probabilities proportional to size, Superpopulation model, Computer application: POSDEM.

References:

- Bellhouse, D.R. & Rao, J.N.K. (1975); Systematic sampling, in the presence of a trend; *Biometrika* 62. P.694-697;
- Biehler, R (1997); Software for learning and doing statistics. *International Statistics Review*
- Cochran, W.G. (1977): «Sampling Techniques», 3rd edition, New York: Wiley. P.205
- Murthy, M.N. (1967): «Sampling theory and Methods» Statistical Publishing Society, Calcutta. P.127
- Krishnaiah, P.R.; Rao, C.R. (1988): «Sampling», *Handbook of Statistics*, North-Holland. P.131
- Sánchez-Crespo, JI. (1997) *Journal Oficial Statistics*, Vol. 13, Nº 4, Pg 327-339



IMPROVING THE PROCESSING OF IRELAND'S BUSINESS EXPENDITURE ON R&D SURVEY USING GSPBM

K Phelan

Central Statistics Office, Cork, Ireland

Over the past number of years, the Central Statistics Office has been heavily engaged in process mapping the various surveys that it conducts as part of efforts to continually improve our processes and by extension the quality of the outputs. The process mapping exercise was conducted as part of systematic approach to improve the standard of our survey documentation. As part of this work programme, the biennial survey on Business Expenditure on Research and Development (BERD) was process mapped in 2013.

As part of ongoing quality improvements the CSO decided in 2015 to run a pilot on reviewing and updating the existing mapping standards using the template of the GSPBM version 5. This pilot focused on the BERD survey and following the success of this pilot, this GSPBM template will be rolled out across the Office.

This paper will look at the challenges faced and the benefits gained from using the GSPBM as a template for our process mapping standards.

Keywords: Improve, GSPBM, process.



WIRED FOR DATA - THE TRANSFORMATION OF DATA COLLECTION IN THE UK

K. Moore

Office for National Statistics, UK

The Office for National Statistics is the largest producer of Statistics in the UK. The Office is transforming the way it collects data; moving from traditional paper questionnaires to increased use of electronic data collection and new sources of data including that from other UK Government Departments and Big Data. In an uncertain financial climate, the use of these data sources provides the potential for efficiencies in statistical output production. This presentation describes the CORA project, the first step along this road and the challenges facing it to maintain the quality of ONS outputs whilst redesigning their production processes in an attempt to realise these potential benefits.

Keywords: Quality assurance, digital transformation, generic processes.



ON THE USE OF STATISTICAL PROCESS CONTROL IN MONITORING MORTALITY. AN APPLICATION TO EUROPEAN COUNTRIES

A. Debón¹, V. Giner-Bosch¹, M. Cabrerizo-Cabanos²

¹ Centre for Quality and Change Management, Universitat Politècnica de València.

² Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universitat Politècnica de València

The evolution of mortality is a key global concern from both an economical and a social point of view. Particularly, being able to detect and predict changes in mortality compared to its expected behavior as accurately as possible is a desirable goal. In this context, the standard mortality ratio (SMR) is commonly used in order to measure the mortality of a country with regards to its neighbouring countries at a given moment in time.

In this work, we address the study of the evolution of the SMR of a country over time, modeled as a time series. The joint use of time series and statistical process control (SPC) techniques to model and monitor the behavior of the SMR is explored. Both approaches are relevant and complementary. On one hand, time series are an appropriate tool to study and characterize the evolution of SMR over time and to forecast on it. On the other hand, SPC allows detecting significant changes in the trend of the variable being monitored. More precisely, we suggest monitoring the residuals of the adjusted time series model using control charts.

We present and discuss the results of applying our proposal to mortality data of European countries in a 20-year period. These results show the relevance of our approach and allow envisaging our next research steps.

Finally, the use of other approaches combining time series and SPC techniques is also outlined.

Keywords: mortality, standard mortality ratio, time series analysis, statistical process control, control charts.



Chair: Maria Isabel Lázaro

MEASURING ALCOHOL CONSUMPTION IN SOCIAL SURVEYS – A COMPARATIVE STUDY

B. O. Lagerstrøm¹, K-A. Lund¹, K. Ringdal²

¹ Statistics Norway, Oslo, Norway

² Norwegian university of science and technology, Trondheim, Norway

Various methodological issues influence the measurement of alcohol consumption in surveys. One factor is the reference period for which questions are asked—that is, whether respondents are asked for an exact recall of their intake during a short, recent period or for a summary of their drinking behavior over a longer period, such as the past year. Longer recall periods provide sufficient time to link consumption data with concurrently collected data on the prevalence of alcohol-related outcomes. Another factor influencing survey results is the approach used to measure alcohol consumption (Dawson 2003). Respondents might be told to report the factual entity of alcohol which they drank in a specified period or to report units of alcohol in accordance to national guidelines. Finally, features of the overall survey design — such as the mode of interview (i.e., in person versus by telephone) or topics of survey, may influence the reliability and validity of the data.

In this paper we investigate effects on estimates of alcohol consumption by different approaches to measurement of alcohol consumption between different surveys and survey modes. Comparisons are done by comparing measurement of alcohol consumption in Norway from the 7th round of the European Social Survey (ESS) (CAPI/factual entity reporting/show card (visual aid)), the European Health Interview Survey (EHIS/verbal aid) (CATI/unit reporting), and a third national survey on alcohol, drugs and tobacco (CATI/unit reporting/visual or verbal aid). The three surveys differ according to three dimensions a) mode of data collection, b) operationalization of the questions to collect data on alcohol consumption (used to calculate grams of alcohol) and c) to which stimulus respondents are exposed, to aid their comprehension of “unit”. These surveys make it possible to study both mode effects and the effects of different measurement instruments on the estimation of alcohol consumption in Norway.

Keywords: Measurement of alcohol, measurement error, survey mode, show cards.



PROXY RESPONSES TO SUBJECTIVE QUESTIONS: THE INFLUENCE ON THE RESULTS OF THE HEALTH EXPECTANCY INDICATOR

D. Lavtar, M. Zaletel, A. Rogelj
National Institute of Public Health, Ljubljana, Slovenia

Healthy life years (HLY) is a health expectancy indicator that is on yearly basis calculated from life tables (mortality component) and prevalence of limitation because of health problems in activities people usually do (disability component). The disability prevalence data are provided by the GALI (Global Activity Limitation Instrument) question from Statistics on Income and Living Conditions (EU-SILC). The HLY indicator is calculated using the Sullivan's method.

Comparability is mainly pre-determined by the comparability of the data sources. For the GALI question the use of proxy respondents should be limited as much as possible. Even though proxy respondents are beneficial in case of sample units that are unable to answer on their own behalf (especially old and disabled interviewees), there is however a question of quality of proxy response. Studies report fairly good agreement between proxy respondent and target respondent in assessments of functioning, physical health, and cognitive status, but this agreement depends on many circumstances, where survey mode and socio-demographic characteristics of respondents should not be neglected. The paper will address the differences of proxy and non-proxy respondents (self-response) and try to identify the characteristics of respondents for whom most proxy responses are given. The case study will be presented: from the Slovenian EU-SILC survey data from 2008 till 2014 GALI question will be analysed in the view of different survey modes used and proxy/self-response prevalences. Estimates of HLY based only on self-reported limitation prevalences do not differ from those taking in account proxy and non-proxy responses. There are about one fifth proxy responses to the questions on health. Women tend to be proxies more often than men. Male proxy partners seem to under-report to higher extent the limitations of their female partners than the opposite. Conclusions will show the robustness of the results to the under-reported prevalences.

Keywords: proxy respondents, subjective questions, health expectancies.



IMPROVEMENTS FOR THE OFFICIAL STATISTICS ON SOCIAL DETERMINANTS OF HEALTH: AN EXPERIENCE FROM THE USERS' PERSPECTIVE

Cabrera-León Andrés^{1,2}, Daponte Antonio^{1,2}, Mateo Inmaculada^{1,2}

1 Andalusian School of Public Health,

2 Biomedical Research Networking Centre on Public Health and Epidemiology (CIBERESP). Granada, Spain

The objective was to identify and select indicators to assess the impact on health of the social context and the latest economic recession in Spain and its regions.

Based on the Spanish conceptual framework of the determinants of social inequalities in health, we identified indicators sequentially from key documents, Web of Science, and organizations with official statistics. The information gathered resulted in a Directory of Indicators that was reviewed by an expert panel. We then selected a set of these indicators according to geographic (availability of data for regions) and temporal (from 2006 to 2012) criteria.

We identified 203 contextual indicators related to social determinants of health and selected 96 (47%) based on the above criteria; 16% and 35% of the identified indicators did not satisfy the geographic or temporal criteria, respectively. At least 80% of the indicators related to dependence and healthcare services were excluded. The final selection of indicators covered all social determinants of health. 62% of these were not available on the Internet. Around 40% of the indicators were extracted from sources related to the Spanish Statistics Institute.

During the identification and selection of indicators and data, we detected some need for improvements: to include data segmentation by region and age group; to expand the time series for indicators on migrant populations; to offer bookmarks for data queries; and to make public the indicators and data already requested by users.

We propose a Directory with wide information of contextual indicators on social determinants of health and a Database to facilitate assessment of the impact of the latest economic recession on health and health inequalities in Spain and its regions.

The proposals for improvements that emerged during this work may be useful to increase the quality of the statistical processes and products from key sources with official statistics.

Keywords: Social determinants of health, contextual indicators, official statistics.



THE ROLE OF STATISTICS IN PRODUCING QUALITY OF LIFE INDICATORS - THE CASE OF HUNGARY

Máté Farkas

Hungarian Central Statistical Office, Budapest, Hungary

In the last decades, especially from the 1970s and onwards an intensive social scientific discourse has emerged about the use of complex statistical indicators measuring multidimensional concepts, e.g. well-being, development, quality of life etc. While after World War II it was widely accepted by mainstream economists and scholars that quality of life is strongly correlated with objective, economic factors, later the attention has turned to subjective and complex approaches. In my point of view official statistics must have the biggest responsibility to produce comprehensive, longitudinal data on quality of life of a given society considering the adequate theoretical background of these phenomena. In my presentation firstly I introduce a short history of different quality of life approaches in a global context, referring to the different synonym concepts as well (well-being, welfare, happiness etc.). In Hungary researches on the quality of life and social well-being have started to grow only a decade ago, and their “culmination” was the Hungarian Well-being Indicator System worked out by the Hungarian Central Statistical Office in 2014 [Kelemen and Kincses 2015*]. The other part of my presentation focuses on the statistical methodology of this system, and reflects the main results of the survey on social well-being in Hungary.

Keywords: quality of life, social well-being, indicator system, Hungary.

* Kelemen R. – Kincses Á. (2015): A jóllét magyarországi indikátorrendszerének elméleti alapjai. (Theoretical background of the Hungarian Well-being Indicator System.) In: *Gazdálkodás*, 2015/3. pp. 220–236.



MONITORING SUSTAINABLE DEVELOPMENT: LESSONS FROM MDG, CHALLENGES FOR SDG

A.J. López-Menéndez, R. Pérez-Suárez
University of Oviedo, Oviedo, Spain

The Millennium Development Goals (MDG) have been conceived fifteen years ago as combination of human needs and basic rights that every individual around the world should be able to enjoy. According to the United Nations Millennium Declaration a list of eight goals has been agreed together with specific targets to be achieved by 2015.

Once this milestone has come, the monitoring of MDG pose a challenge, not only for the less developed countries (where a “data revolution” is needed), but also for developed countries, since some indicators are reported with very long lags, data coverage remains erratic and national statistical systems lack the capacity to generate high-quality data.

Furthermore, world leaders have recently agreed the new Sustainable Development Goals (SDG) connecting the issues of multidimensional poverty, inequality and exclusion, and sustainability. Although the methodology of SDG monitoring is still under development, a recent report by the Sustainable Development Solutions Network (SDSN) proposes 100 indicators and 10 principles for global monitoring indicators, emphasizing the need of simplicity, high frequency and disaggregation.

In this context, our paper provides an overview of the MDG achievements under the Poverty-Growth-Environment approach. A performance index is proposed with the aim of measuring progress towards the MDG goals in different countries and regions, also analysing the existing relationships between goal performance and economic growth.

Keywords: MDG, Poverty, Sustainability.

EVALUATION OF THE CENSUS ERROR IN SPAIN

M.L. Rodero-Cosano, J.A. Salinas-Pérez, M. Carbonero-Ruz
Universidad Loyola Andalucía, Sevilla, Spain

The population and housing census, carried out every 10 years in Spain, provides data on the main demographic and social aspects of the population, as well as their distribution at any given time. The availability of reliable and updated information is crucial to improve the design of public policies.

So far Spain had conducted censuses by using the state-of-the-art methodologies, which includes two main strands. The first one consists on carrying out a general survey in every home, just as the census of 2001 was conducted [1]. The second one has two elements, a short questionnaire administered to the complete population, and a long questionnaire only for a selected sample about socio-economic issues. This method was employed in previous census [2], [3].

The 17th census was carried out in Spain in 2011 [4], within the 2010 World Population and Housing Census Programme [5], based on the European Regulation 763/2008 that contained the main recommendations of this program [6]. Among the wide range of possible options for collecting information, Spain has opted for a model based on complete administrative records and in situ verification of residence and building information. A survey of around 12.3% of the population completed the statistical operation.

Given that, this work aims to assess the statistical error due to the application of the new methodology and to compare with those were in previous censuses. To do this, we compare the survey microdata of the census with other available data of demographic statistical operations such as Continuous Municipal Register Statistics.

Keywords: Census, statistical error, data collection.

[1] National Statistics Institute, 'Population and Housing Census 2001', 2007. [Online].

[2] National Statistics Institute, 'Population and Housing Census 1991', 1994. [Online].

[3] National Statistics Institute, 'Population and Housing Census 1981', 1985. [Online].

[4] National Statistics Institute, 'Population and Housing Census 2011', 2014. [Online].

[5] United Nations Statistics Division, '2010 World Population and Housing Census Programme', 2008. [Online].

[6] EUROPEAN PARLIAMENT AND OF THE COUNCIL, 'REGULATION (EC) No 763/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 9 July 2008 on population and housing censuses', 2008. [Online].

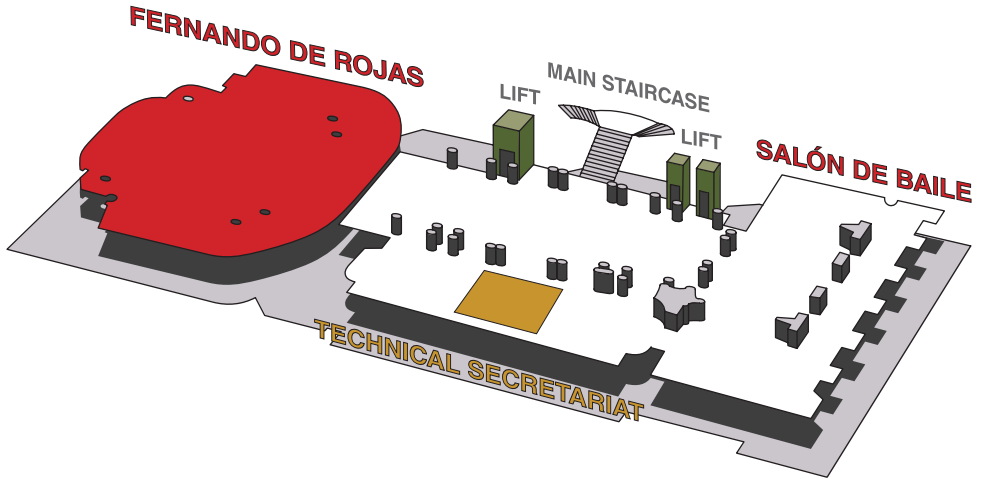


EUROSTAT IMMIGRATION STATISTICS. MISLEADING INFORMATION FOR EVIDENCE-BASED POLICYMAKING

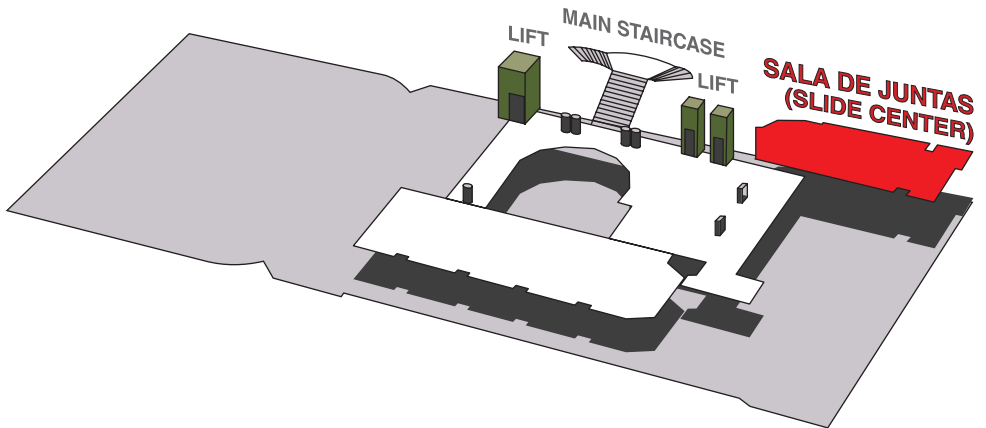
Amparo González-Ferrer
IEGD, CSIC, Madrid, Spain

Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection, and the rest of EU legislation that developed the Regulation since then, have definitely improved the availability of statistical information concerning different aspects of international migration in Member States. However, the collection and production of the requested statistics is done by national authorities, which depart from national-specific immigration concepts and procedures. National immigration figures need to be grouped in different ways to assure that national definitions fit in with the Eurostat statistical categories and definitions. Often, the decisions that need to be made to achieve such adaptation vary across Member States reflecting cross-national variation in immigration laws and policies. The manner in which such a statistical translation has been made remains very poorly documented as metadata in the Eurostat platform. As a result, Eurostat immigration statistics often present as comparable figures that are not because they respond to different concepts in their respective countries, and these concepts were not previously homogenised. The information collected through qualitative interviews with different national experts in charge of the production of the requested tables, as well as our own analyses of some of these figures, will be utilized to illustrate inconsistencies between Eurostat and national figures, as well as cases of misleading cross-national comparability. The focus will be on statistics on residence permits for family reasons, and statistics on residence permits by skill level and duration of the permits. Policy implications of the limitations identified in the analyses will be discussed.

Keywords: immigration statistics, comparability, policy.

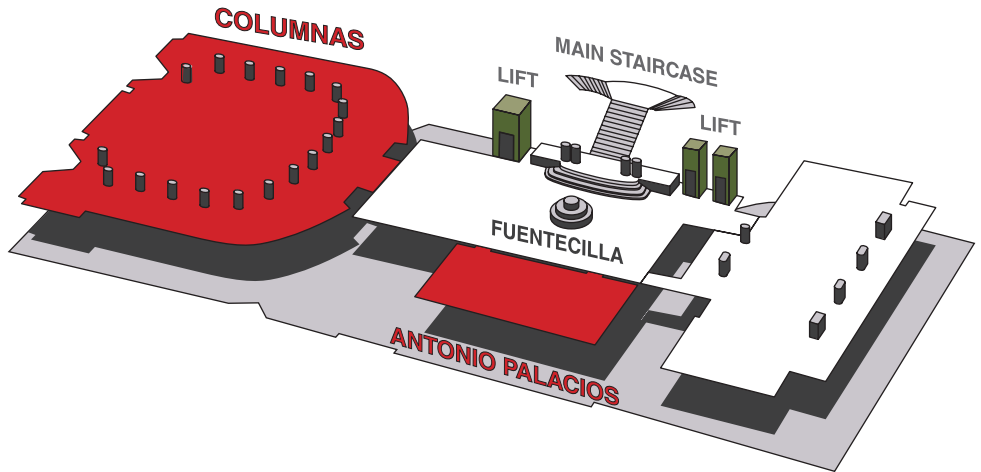


3rd Floor

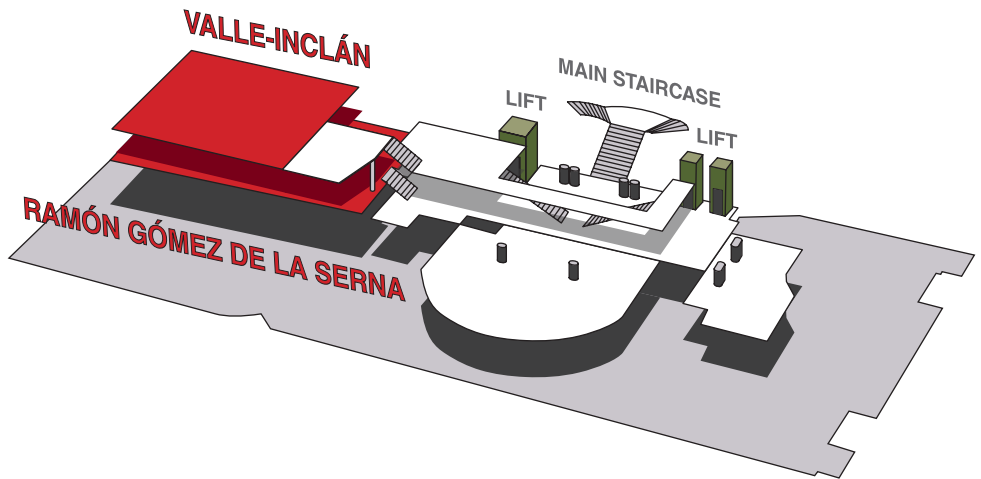




4th Floor



5th Floor





www.Q2016.es