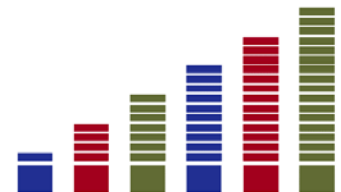# Evaluating the Quality of Input Data for Official Statistics

Session 4: Administrative Data:
Cross-Cutting Issues

June 1, 2016

Regin Reinert

reg@dst.dk

STATISTICS DENMARK

INE Instituto Nacional de Estadística   eurostat
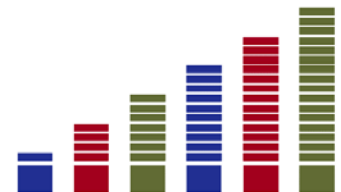
Madrid, May 31 - June 3

ESSnet Project on Quality of Multisource Statistics

- To guarantee the quality of output produced using administrative sources.

- 4 work packages (WP).

- WP1 deals with quality of input data.

- Within WP1 existing methods and current practices.

- Create a checklist using the best from existing methods.

- The task is <u>not</u> to create quality indicators.

Gross List of Methods for Measuring Input Quality of Register Based Statistics

- Denmark initiated a list of methods.
- Other participants of WP1 added methods to the list.
- Methods are identified by origin, web-link, and year.

Gross List of Methods (15)

ESSnet	2013	Use of Administrative and Accounts Data in Business Statistics
		WP2 Usefulness of administrative data for business statistics and initial quality checking
		WP6 Quality Indicators when using Administrative Data in Statistical Outputs
Netherlands	2009	Checklist for the Quality evaluation of Administrative Data Sources
Sweden	2007	Register-based Statistics: Administrative Data for Statistical Purposes
Sweden	2011	Quality Assessment of Administrative Data
Istat	2010	Quality assessment and reporting in a short-term business survey based on administrative data
Istat	2014a	Evaluating administrative data quality as input of the statistical production process
Istat	2014b	Towards a more efficient system of administrative data management and quality evaluation to support statistics production in Istat
UK	2013	Guidelines for Measuring Statistical Output Quality
USA	2013	Data Quality Assessment Tool for Administrative Data
NZ	2011	Evaluation of administrative data sources for subnational population estimates
NZ	2011	Evaluation of alternative data sources for population estimates
Austria	2010	A Quality Framework for Statistics based on Administrative Data Sources using the Example of the Austrian Census 2011
BLUE-ETS	2011	Report on methods preferred for the quality indicators of administrative data sources
Eurostat	2014	Memobust Handbook on Methodology of Modern Business Statistics
Eurostat	2014	ESS Handbook for Quality Reports

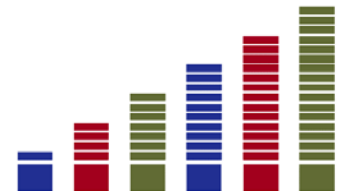# Overlap between Methods

More than 500 indicators

- Some indicators are present in all list, e.g. undercoverage and overcoverage.

However, there are many unique indicators

- It seems like indicators often are constructed to measure certain difficulties that are relevant to particular statistics and not necessarily relevant to other statistics.
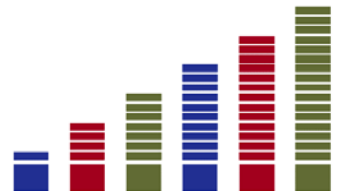
Our task is not to re-invent the wheel

- Our task is to compile the best of **existing** methods into one method.
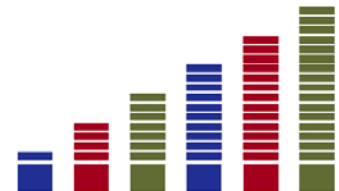
Promising Indicators

- Participants of WP1 were asked to point out indicators well suited.

- Mainly quantitative indicators were selected.

- There was a wish for clear indicators with precise definitions, that could be calculated (or stated) unambiguously.
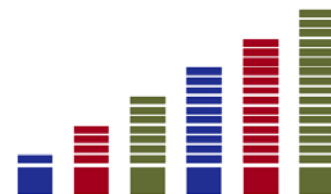
Criteria for Selecting Promising Indicators

- Indicators that were selected by at least three countries were included in the list, 9 indicators in total.

- All 9 indicators having more than two selections belonged to the ESSnet Admin "Accuracy" dimension.

- 6 dimensions are presented in the ESSnet Admin project.

- At least 2 indicators from each dimension.

- 8 additional indicators were selected in order to have all 6 dimensions covered.

Promising Indicators

- Our list relies heavily upon the ESSnet Admin data list of indicators
  - 23 quantitative indicators
  - 46 qualitative indicators
- Our list uses 16 of the quantitative indicators from the ESSnet Admin list.
- Of those not selected we find indicators like
  - %reduction of survey sample size when moving from survey to admin data
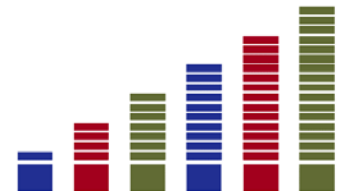 - Output related indicators

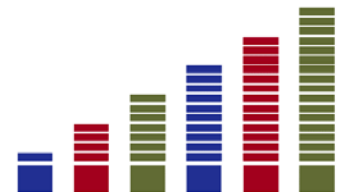| Accuracy | ESSnet 9:
Item non-response
(% of units with missing values for key variables) |
	ESSnet 10: Misclassification rate
	ESSnet 11: Undercoverage
	ESSnet 12: Overcoverage
	ESSnet 14: Size of revisions from the different versions of the admin data RAR – Relative Absolute Revisions
	ESSnet 15: % of units in admin data which fail checks
	ESSnet 16: % of units for which data have been adjusted
	ESSnet 17: % of imputed values (items) in the admin data

List of Promising Indicators - 2

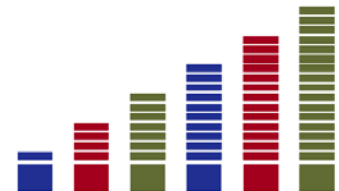| Timeliness and punctuality | ESSnet 4: Periodicity (frequency of arrival of the admin data) |
	ESSnet 18: Delay to accessing / receiving data from Admin Source
Coherence	ESSnet 5: % of common units across two or more admin sources
	ESSnet 21: % of relevant units in admin data which have to be adjusted to create statistical units
Comparability	ESSnet 19: Discontinuity in estimate when moving from a survey-based output to an output involving admin data

List of Promising Indicators - 3

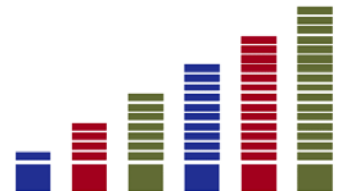|	
Cost and efficiency	ESSnet 7: % of items obtained from admin source and also collected by survey
	CBS 2009, Source 4.1: Cost of using data source
Use of administrative data	ESSnet 2: % of items obtained exclusively from admin data
	ESSnet 3: % of required variables which are derived using admin data as a proxy

Testing Indicators

- Precise definition of indicators is demanded.
- Even with sufficient detailed definitions, calculations of indicators can be challenging.
  - Data entered into registers are often edited as soon as possible in order to have good quality.
  - Input data are only available directly from the source.
- Definition of relevant comparisons.
- Deep knowledge of the structure of the register might be required.
- Calculating quality indicators does not always use many resources, but it might require the right resources.
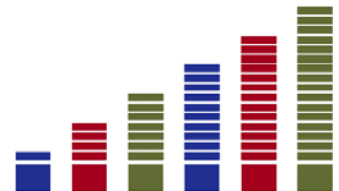
Results

- Test of the checklist has been performed in Austria, Hungary and Denmark.

- The list consists only of 17 indicators.

- Most of the indicators tested are characterized as useful.

- There is agreement between results from the three countries.

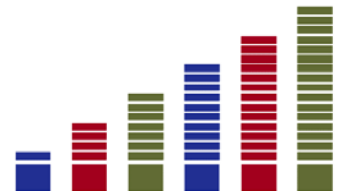- The list will be evaluated and adjusted by other participants of WP1.

Purpose of the workshop:

- to inform about the ESSnet project
- to get a better knowledge of assessment methods used and demanded in different countries.
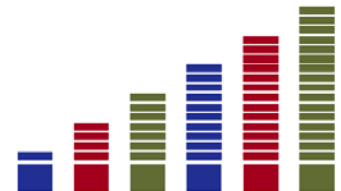
Access to Administrative Data

- In many years focus has been on accessing administrative data for statistical purposes.

- In recent years more and more countries have access to administrative data and by 2015, 1/3 of the UNECE countries base their census at least partially on administrative registers.

- Focus has shifted from accessing administrative data to ensuring the quality of statistics produced using administrative data.

Population Register or Not

- In the workshop in Budapest we experienced a clear distinction between countries with population registers and countries without population register.

- Countries with population registers were eager to improve the quality of their registers and their statistics produced using administrative registers.

- Countries without population register were more concerned about accessing data that could improve their statistics.

Thank you for your attention!

Regin Reinert

reg@dst.dk