

Methodology in the Book Publishing Activity Statistics

Madrid, December 2018

Introduction

This statistic is an operation included in the National Statistical Plan. It was first carried out in the INE in 1965 and is now conducted on an annual basis and in accordance with the standards contained in the Revised Recommendation concerning the International Standardization of Statistics on the Production and Distribution of Books, Newspapers and Other Periodicals adopted by the UNESCO General Conference in November 1985.

The aim of the statistics is to provide information on the number and characteristics of titles published in Spain based on the physical titles deposited in the National Library of Spain (NLS).

Among users of the statistics are the INE itself (National Accounts) as well as the Ministry of Culture and Sports (MCD for its acronym in Spanish) and other ministries and national public bodies, Central Statistical Offices of the Autonomous Communities, as well as researchers and universities.

Starting from the 2015 Book Publishing Activity Statistics, the primary data used will be the bibliographic records of the automated catalogue of the National Library of Spain. This replaces the traditional fieldwork and improves the efficiency of its elaboration. The basic information existing in the registers has been completed with the information from the Statistics on Publishing of Books with ISBN.

This change in the process of elaborating the Book Publishing Activity Statistics has been possible thanks to the collaboration with the Ministry of Culture and Sports (through the National Library of Spain, the Sub-Directorate General for the Promotion of Books, Reading and Spanish Letters and the Sub-Directorate General for Statistics and Studies). The feasibility of using administrative registers available to carry out this statistical operation was analysed, respecting the comparability of the time series of the total volume of titles published in comparison with previous years' editions.

The Project of modification in the process of collecting information for the Book Publishing Activity Statistics was presented to the High Council on Statistics in June 2015, and was unanimously approved by the Permanent Commission of that Council at its meeting in September 2015.

The existing legislation on administrative files used by the statistical operation is ***Law 23/2011, of 29 July, on legal deposit.***

This law guarantees the availability of bibliographic records by making reference to the statistics of book production both in its preamble (with the Legal Deposit, instruments are adopted to provide the State with the necessary information to enrich its statistics on book production, which are fed by the materials actually conserved by the National Library of Spain) and in its articles (Article 2. *Objectives of Legal Deposit, 2. Collect the required information to compile official statistics on the bibliographic heritage*).

This law states that the Spanish Bibliography is prepared in the National Library of Spain. This consists of an exhaustive register of publications produced and distributed in Spain, under the legal obligation of publishers to deposit published titles and therefore each title will obtain its corresponding Code of Legal Deposit.

With respect to the other source of the statistics, **the Statistics on Publishing of Books with ISBN**; it is a statistical operation integrated in the National Education Project, prepared by the Sub-Directorate General for the Promotion of Books, Reading and Spanish Letters (MCD).

Its objective is to provide results, relating to books registered with the ISBN and to publishers, which facilitate the analysis of book production and its evolution through the exploitation of the ISBN administrative register. It does not include titles with no commercial purpose and therefore without ISBN.

As pointed out in the report "Panorámica de la edición española" (MCD), this statistic is complemented by the Books Publishing Activity Statistics of the INE. Both operations provide information relating to the publication of titles but referring to two different moments in time and, therefore, their primary data also come from different sources of information: from the ISBN on publishing projects informed by the publisher himself, and from the INE, on library materials already published and deposited in the NLS, so that the information is obtained directly from the physical book itself.

1. Type of operation, analysis unit and scope of research

Type of operation. This is an exhaustive operation that does not generate any burden for the respondent, since the statistics use an administrative record generated from the data provided by the publishers, on a mandatory basis, when depositing the copies, in the Legal Deposit Service of the Department of Acquisitions and Heritage Increase of the National Library of Spain.

The unit of analysis consists of the titles (books and brochures) published nationwide and deposited in the National Library of Spain during the reference year.

The geographical scope includes the whole national territory. All publications published in Spain regardless of where they have been printed.

The reference period is the calendar year since the aim of the statistics is to estimate the book production for a given year (t).

However, in order to retrieve publications not included in previous years that have entered the NLS databases late, the valid reference period is set according to the year and month of cataloguing and year of publication. Therefore, the following publications will be included:

Cataloguing date: from 01/01/year t to 30/04/year t. *Date of edition:* in year t

from 30/04/year t to 31/12/year t

in years t-3 to year t

from 01/01/year t+1 to 15/06/year t+1

in years t-3 to year t

Cataloguing date. Entry date in the NLS database.

Date of publishing. Date on which the publisher issues a publication. They usually coincide. Starting from the statistics referring to 2016, the cataloguing period extends from 30/04 of the reference year to 15/06.

2 Main concepts

TITLE

Title is a term used to designate a publication in which the text is predominant (whether printed or in electronic format) and constitutes a single whole whether it consists of one or more volumes.

Titles or publications for advertising purposes or with an ephemeral nature and those with less than five pages are excluded.

Titles that are the subject of statistical studies can be presented in different formats and media:

Format. This refers to the physical description of the title.

Modern monographs: Titles with only textual content. They can be printed on paper or in electronic format (tangible media).

Electronic resources: Document in electronic format (tangible) in which other formats are included besides text (images, multimedia applications, background music, voice-over, etc.). For example; multimedia courses, power point presentations with music and images / videos. Only those that do not constitute a simple attached material are included, since in that case they would be considered as part of the main title.

Kits: A resource made up of two or more different means or by different forms of the same means, neither of which prevails over the other and intended to be used as a unit. Only kits containing monographs are included.

Atlas. Only cartographic documents corresponding to atlases will be taken into account.

Media (only physical or tangible media).

Only paper format. Only includes monographs.

Paper format along with computer media. Includes Kits.

Only computer media (CD_ROM, DVD, etc.). It includes some monographs and all electronic resources.

All the publications collected in the NLS have a tangible medium. Therefore, **digital books are not included** because they do not have such tangible media as they are electronic books published on the Internet.

ISBN CODE (INTERNATIONAL STANDARD BOOK NUMBER)

This is an international standard book number that identifies the book edition of a specific publisher and a particular edition. It is an indicator of the publishing forecasts.

It should also be pointed out that not all titles must have an ISBN number, as this is a number designed to promote publications trade. If they are not intended for commercial purposes, that is, they are not to be offered for sale, they do not require ISBN.

It is estimated that approximately 20% of published titles do not have ISBN code.

In any case, Royal Decree 2063/2008 of 12 December of the Ministry of Culture by Law 10/2007 of 22 June on Reading, Books and Libraries as regards the ISBN, published in the Official State Bulletin (BOE for its acronym in Spanish) number 10, of 12 January 2009, in its "Single Derogatory Provision a)", repeals the obligation to record the ISBN in publications, established by Decree No. 2984/1972. Therefore, there is no obligation to apply for an ISBN to obtain the Legal Deposit, much less to publish a book.

LD CODE

Legal Deposit Code. It is issued subsequently to the ISBN. It is an indicator of books already published.

It is a mandatory code for all books and is issued by the legal deposit offices.

3. Data collection process

The primary data for the compilation of this statistics are obtained directly from the titles deposited by the publishers, in accordance with their legal obligation, in the Legal Deposit Service of the Department of Acquisitions and Heritage Increase of the National Library of Spain during the calendar year.

The NLS sends to the INE, through the Statistics and Studies Division of the MCD, the cataloguing file from the Integrated Library Management System (SIGB for its acronym in Spanish) in accordance with INE specifications (based on UNESCO recommendations) in terms of information contained and definitions. Likewise, it includes additional information that allows filtering the set of titles catalogued in the NLS according to the INE scope.

The NLS corrects, as far as possible, cataloguing errors by implementing error detection/correction processes prior to data extraction.

The INE, in order to guarantee the quality of the exchange file used for the statistical exploitation, carries out the appropriate coverage controls (analysis of filter fields) and content (invalid values and incorrect relations between different fields) and establishes strategies for cleaning, updating and imputation of invalid data. Likewise, a series of tables have been designed with the object of verifying the quality of the aggregate data and the relationships between the different variables, as well as their comparability with other statistical sources.

Additionally, the Sub-Directorate General for the Promotion of Books, Reading and Spanish Letters sends the INE, through the Statistics and Studies Division of the MCD a file,

generated from the file of Statistics on Publishing of Books with ISBN, which allows to identify the ownership (public or private) of the publishing houses.

4. Variables under study and classification

The variables which are the subject of research in the statistics are the following (all of them except the variable “ownership of the publisher”, may be obtained directly from the bibliographic records file of the NLS):

Province of publishing

This refers to the province where the central headquarters of the publishing house that has published the title is located.

Language (publication and translation)

All publications are classified by publication language, while translations are classified by language of the original text.

As of 2015 disaggregated information is included on the languages of publications issued in two or more languages (national or foreign).

Number of pages

Publications are classified by number of pages in intervals. Based on the number of pages, the INE will codify them in intervals.

Type of publication

According to the number of pages the titles are classified into:

Book. Any non-recurrent publication that consists of at least 49 pages, not counting the cover pages, published in the country and made available to the public.

Brochure. Publication of the same type as the book that consists of 5 to 48 pages.

Edition

Depending on the print run, the titles are classified into:

First edition. First publication of a text by a specific publisher in a specific collection and in a certain format.

Re-edition. Re-editing is understood to be an edition that differs from the previous editions by some modifications made to the content or presentation.

Reprint. A reprint does not contain any content or presentation changes other than typographical corrections made in the text of previous editions. Reprints are not recorded or accounted for in the total count of titles. Since its publication in 2012 and with the entry into force of Law 23/2011, of 29 July, on Legal Deposit the publisher becomes the main depositor (and not the printer as before said law). This implies the loss of information on the number of reprints (copies) published of the same title.

Subject category

Statistical data relating to book production are classified according to the 25 categories of subjects covered by the UNESCO recommendation (the figures in brackets refer, in general terms, to their correspondence with the UDC).

1. Generalities (0)
2. Philosophy, psychology (1)
3. Religion, theology (2)
4. Sociology, statistics (30-31)
5. Political science, economic science (32-33)
6. Law, public administration, social welfare and assistance, insurance (34,351-354,36)
7. Military art and science (355-359)
8. Education, teaching, training, entertainment (37)
9. Trade, communications, transport, tourism (38)
10. Ethnography, cultural anthropology (customs, folklore, habits, tradition) (39)
11. Mathematics (51)
12. Natural sciences (52-59)
13. Medical sciences, public health (61)
14. Engineering, technology, industries, trades and crafts (62, 66-69)
15. Agriculture, forestry, stockbreeding, hunting and fishing (63)
16. Domestic science (64)
17. Management, administration and organization (65)
18. Urban development, town planning, architecture (70-72)
19. Plastic and graphic arts, photography (73-77)
20. Music, performing arts, theatre, film and cinema (78,791-792)
21. Games and sports (793-799)
22. Philology, languages, linguistics (81)
23. History of literature and literary criticism (82)
24. Geography (91)
25. History, biography (92-99)

For certain tables, these categories are grouped into the simplified UNESCO classification:

- Generalities (0)
- Philosophy, psychology (1)
- Religion, theology (2)
- Social sciences (30-39)

Pure Sciences (Mathematics, Natural Sciences) (51-59)
Applied Sciences (Medicine, Engineering, Technology) (61-69)
Arts (70-799)
Philology (81)
History of literature and literary criticism (82)
Geography and History (91-99)

Publication category

Given their particular interest, the following categories of titles are published:3w

Text books. Starting from the 2015 edition, exercise books and compulsory reading books are considered as separate titles from the textbook they accompany.

Children and youth publications. Starting from the 2015 edition, this category replaces the previous "children's books" category.

Titles of public publishing houses.

Year of publication

The aim of the statistics is to estimate the book production for the reference year and, for this reason it covers the titles that are actually deposited in the NLS during the reference year, which mainly coincide with the year of publication t .

However, in order to retrieve the publications that have been deposited with a certain delay in the NLS and that would otherwise be "lost", in addition to the year of publication (t), titles published up to three years earlier (from year t to year $t-3$) will be included. This criterion has been in force every year and therefore the annual series are comparatively valid.

5. Dissemination of the statistics

The publication consists of two sections:

1. Main series since 1993. It covers the evolution of book publishing production classified by subject, edition type, publication class, language of the original text and language of publication, Autonomous Community and Cities where they were published and date of printing.

It also includes the monthly series since 1998, which provides monthly information on the titles published by subject of publication.

2. Detailed results for the year of publication of the statistics. They present information broken down by place of publishing (province and Autonomous Community) and categories of subject and publication.