BANCO DE **ESPAÑA**
Directorate General Economics, Statistics and Research

04.09.2007

BOLETÍN ESTADÍSTICO TIME-SERIES FILES

User's Manual

Statistics Department

CONTENTS

1 Introduction

The objective of this document is to describe the information contained in the time-series files for the tables in the ***Boletín Estadístico*** and to provide guidance on the possible uses of these files.

For each series, the data covering the whole sample period available and the qualitative information, or metainformation, are provided. Metainformation gives details about the series such as the economic item they represent, the units in which the data are expressed, the sources and the notes.

Although the target audience of this information is as wide-ranging as the network through which it is distributed, two possible user categories can be distinguished:

a Users that download into their computer one or various tables for processing with a spreadsheet.

b Users or institutions that download all or a large part of the information to load it into a database and combine it with other statistical sources or even with the institution's own information.

The information furnished caters for users of both types.

This Manual is organised as follows. Section 2 details the two types of files disseminated with their format and content; Section 3 contains guidance for users that wish to process the information with a spreadsheet; and Section 4 contains guidance for institutions that wish to develop a program for automatically downloading files into a database.

2 Time-series files

The time-series files of the tables in the ***Boletín Estadístico*** have the **csv (comma separated value)** format, in which, as its name indicates, the fields or values of each line or record are separated from each other by commas. The decimal separator is the dot "**.**". Later on two examples of **csv** format files are included.

From the standpoint of content, they can be classified into two types:

a Catalogue file with a list of all the series and the information on the characteristics of each series. There is a single catalogue file for all the ***Boletín Estadístico*** series. Its name is **catalogo_be.csv.**

The catalogue file is updated monthly, coinciding with the release of the print edition.

Example: Records in the file **catalogo_be.csv**

"BE010101",586293,"BE0101","UEM. Consumo privado. Tasa de variacion","INDEFINIDO","Porcentaje",-2,2,"Porcentaje","TRIMESTRAL", ...

"BE010102",586294,"BE0101","UEM. Consumo publico. Tasa de variacion","INDEFINIDO","Porcentaje",-2,2,"Porcentaje","TRIMESTRAL",

"BE010103",586295,"BE0101","UEM. FBCF. Tasa de variacion","INDEFINIDO","Porcentaje",-2,2,"Porcentaje","TRIMESTRAL", ...

"BE010104",586292,"BE0101","UEM. Demanda interior. Tasa de variacion","INDEFINIDO","Porcentaje",-2,2,"Porcentaje","TRIMESTRAL", .

"BE010105",586296,"BE0101","UEM. Exportaciones de bienes y servicios. Tasa de variación" , "INDEFINIDO" , "Porcentaje", ...

b Files with the data series of a ***Boletín Estadístico*** table

These contain the whole sample period of each of the time series in a table. There will be one file for each table, except in double-entry tables, for which there will be one file for each column in the table, when the number of series in the table exceeds 255. These files are updated daily, which means that each csv file will be updated when the data in the related table change.

Example: Records in the file **be0101.csv**

```
"1996 ENE",98.9,99.0,99.2,99.4,99.0,98.9,99.1,98.2,98.7,100.10000,101.10000,100.90000,99.00000
"1996 FEB",99.4,99.4,99.5,99.6,99.5,99.3,99.5,98.6,99.3,100.20000,101.30000,101.00000,99.10000
"1996 MAR",99.7,99.9,100.1,99.8,100.4,99.8,100.0,99.0,99.4,100.20000,101.30000,101.30000,99.20000
"1996 ABR",99.9,100.2,100.4,99.9,101.1,100.0,100.1,100.0,99.4,100.40000,101.60000,101.30000,99.60000
"1996 MAY",100.1,100.3,100.7,99.9,101.9,100.1,100.1,99.9,99.7,100.30000,101.80000,101.30000,99.20000
"1996 JUN",100.1,100.2,100.6,100.0,101.4,99.9,100.1,99.3,100.0,100.00000,101.80000,101.40000,98.40000
"1996 JUL",100.1,99.9,100.4,100.0,101.1,99.6,99.8,99.1,100.6,100.00000,101.90000,101.30000,98.40000
```

The files of time series are named according to the following format: ***ppccaa[oo].*csv**

Where:

pp*:** is the publication code, **be** for all ***Boletín Estadístico tables.

***cc*:** is the chapter number, with two digits. Example: 01 for Chapter 1.

aa: is the table number, with two digits. Example: 01 for Table 1.

oo: is the column number, with two digits; it is used only for double-entry tables. Example: 01 for Column 1.

csv: is the extension.

Examples: the file with the name **be0101.csv** relates to the series of Chapter 1, Table 1, which is a normal table; and the file with the name **be891001.csv** relates to the series of Chapter 89 (8-bis), Table 10, Column 1, which is a double-entry table.

The users that wish to download all the ***Boletín Estadístico*** time-series files are provided with the compressed file **be.zip**[1] containing them. This file is updated monthly together with the catalogue file, coinciding with the release of the print edition.

Also, as regards users interested in downloading all the files of a particular chapter, for each chapter a compressed file has been generated containing all its files. The name of these files has the format **ppcc.zip**.

[1] Files have been compressed with the program **WinZip**.

Where:

pp*:** is the publication code, **be** for all tables of the ***Boletín Estadístico.

***cc*:** is the chapter number, with two digits. Example: 01 for Chapter 1.

zip: is the extension.

Example: the file with name **be01.zip** contains all the time-series files relating to Chapter 1.

These files are generated each time that any csv file of the chapter is updated.

2.1 Content of catalogue file

The catalogue file contains a line or record for each time series of the ***Boletín Estadístico*** tables. When the same series is included in various tables, it will have a line in the catalogue file for each table in which it appears. Each column or field contains one characteristic of the time series. The name and content of each of them are described below:

a Name of series.
The name of the series relating to the ***Boletín Estadístico*** tables has the following format: ***ppccaaoo[ff]***

Where:

pp*:** is the publication code, **be** for all tables of the ***Boletín Estadístico.

***cc*:** is the chapter number, with two digits. Example: 01 for Chapter 1.

aa: is the table number, with two digits. Example: 01 for Table 1.

oo: is the column number in vertical tables and in double-entry tables or the row number in horizontal tables. Example: 01 for Column 1 or for Row 1.

ff: is the row number in double-entry tables. Vertical and horizontal tables will not have these two digits. Example: 01 for Row 1.

Examples: the series name **be060101** relates to the time series of Chapter 6, Table 1 (which is a vertical table), Column 1; and the series name **be89100101** relates to the time series of Chapter 89 (8-bis), Table 10 (which is a double-entry table), Column 1, Row 1.

b Sequential code
The time-series names described in the preceding section are of the topographical type. This means that if the order of a series is changed in a table or if a table changes its number or chapter, then the name of the series will change. In order that a series remains identified after these changes, it is assigned a sequential code that accompanies it unchanged throughout the life of the series.

c Name of file with time-series values

This is the name of the file for the table to which the series belongs and which contains the observations on that series.

d Description

The description of the series consists of a string of characters indicating concisely the economic item represented by that series. It is complemented by the title described later on.

e Type of variable

This indicates whether the economic item contained in the series is a flow, a stock, an average or an annualised rate of change. The type of variable is necessary information for properly converting the series frequency. The following table lists the various types of variable along with an explanation of them.

Type of variable	Explanation
Start	Stock variable observed at the start of the period.
End	Stock variable observed at the end of the period
Average	Average of the variable for the period
Sum	Flow of the variable for the period
Annualised	Annualised variable

f Units code

Code indicating the units in which the series is expressed. The meaning of this code is stated in the units description and exponent field.

g Exponent

This indicates the power of 10 by which the series values would have to be multiplied to obtain the units. The meaning of this value is stated in the units description and exponent field.

h Units description and exponent

This describes or decodes the units and exponent codes.

i Number of decimal places

This indicates the precision with which the series is measured.

j Frequency

The frequency of the series shall be the highest of those with which the series appears in the table. For example, if a series appears in a table with monthly and annual frequencies, the frequency stated in the related file shall be monthly.

k Date of first observation

Date of first observation in the sample period covered by the file with the time-series values. The date of the first observation is expressed in the frequency of the series.

l Date of last observation

Date of the last observation in the sample period covered by the file with the time-series values. The date of the last observation is expressed in the frequency of the series.

m Number of observations

Number of observations that comprise the series sample period and that are contained in the file with the time-series values.

n Title

This describes in detail the economic item measured or represented by the series. It complements the description. It is formed by a set of character strings separated by the character "/".

o Source

This describes the original sources of the series. It is formed by a set of character strings separated by the character "/".

p Notes

This contains general considerations on the preparation of the series or on certain observations in particular. It is formed by a set of character strings separated by the character "/".

2.2 Content of the files with time-series values

The files with the time-series values of a ***Boletín Estadístico*** table contain four header lines identifying the series and a line with values for each date in the sample period.

a Header lines for identifying series

In order to identify the values of each series, four lines with the following main characteristics are included in the file: the first line contains the list of time-series names; the second line contains the sequential codes; the third line contains the description; and the fourth line contains the units. The content of these fields is explained in Section 2.1.

b Value lines

The header lines are followed by value lines. Each line contains the values of one of the dates in the sample period.

The values of each series will have the decimal places specified in the catalogue file, on the line relating to that series. See Section 2.1.

When there are no data for a given series and date, a symbol explaining why no value exists shall be included in its place. The following table lists the symbols used and their meaning.

Symbol	Explanation
–	Inexistence of the phenomenon in question
…	Data not available

3 Downloading csv files into a spreadsheet

This section includes, by way of an example, some indications on how to work with **csv** files in a Windows environment with an Excel spreadsheet and on how the catalogue file can be used to make searches for series.

3.1 Options for downloading files into Excel

In the Windows environment, **csv** files can be opened from the following applications:

a Internet explorer.
b File explorer.
c Excel.

From any of the applications in these three options, Excel will be opened and the file will be downloaded such that each value or field is entered in a cell of the spreadsheet.

For this operation to proceed properly from any of the aforementioned three applications, the **csv** extension has to be associated with the Excel application and the appropriate options have to be specified in the Windows regional settings.

To associate the **csv** extension with the Excel application, the sequence followed is:

a Open My Computer or Windows Explorer and, in the tools menu, click on Folder options.

b Select File Type.

c In the list of registered file types, select the desired action: New, if there is no association for the **csv** extension; Change, if there is one but it is not with the Excel application.

For more details on association, see Windows help.

If the **csv** extension is not associated with the Excel application, to open the file under options a) and b) above, it will be necessary to specify the application with which it is wished to open the file.

For Excel to download properly the files under the three options mentioned above, a further requirement is that the Windows regional settings must have the following specifications:

1 Decimal separator : **.**

2 Thousands separator: (none)

3 List separator: **,**

To set up these options, the steps are as follows:

1 In My Computer, click first on the option Control Panel and then on Regional Settings.

2 Select Number.

3 Change the aforementioned options.

If the aforementioned values are not specified for the decimal separator, thousands separator and list separator, then the behaviour of Excel depends on the version of Windows and of Excel. It is most likely, however, that each line of the **csv** file will download into the first cell of the spreadsheet instead of each value or field downloading into one cell. If this happens, you can try to download the information correctly into Excel[2] as follows:

1 In Excel, select, in the first column, the range of lines with the information of the downloaded **csv** file.

2 In the *Data* menu, click on the option *Text to columns*. The Office Assistant window for converting text to columns will open. Select the option Delimited and indicate that the separator is the dot ".".

However, if the **csv** file contains values with the decimal point, these values will be downloaded incorrectly. In these cases it is recommended to edit the file and replace the separators before opening it directly from Excel or from Internet Explorer integrated with Excel.

3.2 Searches using the catalogue file

The catalogue file can be used to locate time series by making searches in the time-series characteristics. The procedure for these searches is the option *Edit >Search* in Excel. In the search window, indicate:

a The text you wish to search for.

b In the **search** box, specify: **by row**.

c In the box **search in** box, specify: **values**.

[2] For more information on **csv** and Excel formats, see "Microsoft Excel User's Guide 2".

For example, you could locate the series in which “consumer prices” appears literally in the headings column. Once the row relating to a series has been located, the file containing the time-series values is obtained in the Name column.

4 Downloading csv file information into databases

Given the variety of database management systems and of data modelling designs for storing the information provided, it is complicated to indicate in detail how to develop a program to download this information into a database for integration with other statistical or management data of an institution. Therefore, this section gives some ideas on the strategy to be followed in designing a downloading program.

To download all or most of the ***Boletín Estadístico*** series into a database, a file named **be.zip** is provided. As mentioned above, this file contains in compressed form the catalogue file and the files with all the ***Boletín Estadístico*** tables.

Similarly, one of the purposes of the catalogue file is to make it easier to download information into a database. First, the catalogue file contains the characteristics or metainformation of all the series and, second, it contains for each series:

a Name of the data file.

b Date of first and last observation.

c Number of observations.

This enables the program to open the data files automatically and to control the dates and number of observations to be read.

The steps included in the downloading program are as follows:

1 Open catalogue file.

2 Read the lines of the catalogue file with the information relating to the series of a particular table, e.g. those of the table **beccaa**.

3 Open file containing the observations or values of the series of the table **beccaa** that was read in the preceding step.

4 Read the observations of the file opened in the preceding step.

5 Close the file opened in step 3.

6 Update the series of the table **beccaa** in the target database.

7 Repeat steps 2 to 5 until the end of the catalogue file is reached.

8 Close catalogue file.