Enterprises in Europe

Fourth Report

data

Enterprises in Europe

Fourth Report

Small and Medium-sized Enterprises Database

eurostat

Enterprises in Europe

Fourth Report

Small and Medium-sized Enterprises Database

Diskettes, Manual and Methodology

The SME tabular data is structural enterprise data, broken down by employment size classes and economic activities at national and regional levels. Users can find the raw SME data for several reference years at a national level on the diskettes presented here. This enclosed guidebook is divided into two parts. The first part is a technical manual for accessing the SME data. The second part is a methodological supplement. The methodology is indispensable for the right interpretation of the data.

PART 1

CUB.X

for

MS-Windows 3.1

Software

for Manipulating

Arrays of Statistics

Version 1

Mini Reference Guide

CUB.X Presentation

All the data used in the examples are totally fictitious. Some of the screens are designed solely to enhance explanation or document composition and so may be different from reality.

The CUB.X program for Windows is a tool for exploiting statistical multi-dimensional tables.

CUB.X is the property of Eurostat. This software can be used for your needs, free of charge, but cannot be resold. It may not be used for commercial gain or distributed without prior authorisation by Eurostat.

Information in this document and the content of all CUB.X files are subject to change without notice and do not represent a commitment on the part of Eurostat.

Information in this document and the content of all CUB.X files are subject to change without notice and do not represent a commitment on the part of Eurostat. We shall not be liable for editorial omissions made herein ; nor for incidental or consequential damages resulting from the misuse of the CUB.X program.

A full Cub.X manual is available on request.

Conventions

⌐🖰 ➤ means : select using the mouse.

⌨ ➤ means : select using the keys.

🪟 ➤ means : use MS-Windows features

REMEMBER ! Under the MS-Windows environment :

▣ ➤ ⌐🖰 click to maximize a window

▣ ➤ ⌐🖰 click to reduce a window (put the window in background)

▣ ➤ ⌐🖰 click to reduce or restore a window

To exit any program running under MS-Windows :

⌨ ➤ ALT+F4   -or-

⌐🖰 ➤ Double-click the System Menu (▦) in the upper left corner of the active window.

⬆⬇➡⬅ ➤ means : use the arrow keys (to scroll, etc.).

(if you are using the numeric keypad, make sure NUM LOCK is off before using it).

CONTENTS   page

Installation procedure

If MS-Windows is not already running, type **Win** at the DOS prompt.

First, to prevent eventual incidents or errors (like formatting the diskette...), make a copy of the distribution diskette using the MS-DOS Diskcopy command or the Copy Disk option from Windows File Manager.

⊟   Insert the CUB.X for WIN installation diskette in your diskette drive

Setup.exe   This procedure covers the basics for setting up Cub.X
⌁ ➤ From the File menu, choose Run   ⌨ ➤ (ALT, F, R)

Type : **a:\setup.exe**. (or **d:/setup** if you are running setup from the CD-ROM).

When the program asks you for the target directory, click OK to accept **c:\refer.win** directory **by default** (or type any other directory).
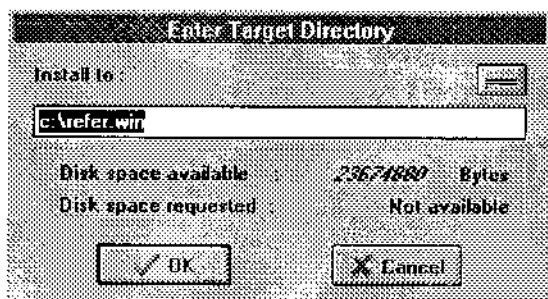
The installation procedure creates the "New Cronos Software" program group.

The icons of the CUB.X Program are the following :

CUB.X for the program
Browser

and   for the Help system.
CUB.X help

Working with Cub.X

If Windows is not already running, at the system prompt, type **win** and then press Enter.

In the Program Manager, double-click the "New Cronos Software" group icon or the group icon that contains Cub.X

New Cronos
Software

Starting CUB.X   To launch Cub.X

➢   Double-click the CUB.X Browse icon

CUB.X
Browser

-or-

➢   Select the CUB.X Browse icon using the arrow keys, then press Enter

---

Exit from CUB.X   To quit Cub.X

➢   From the File menu, choose Exit   ➢   (ALT, F, X)

-or-

➢   Use the "exit" feature of MS-Windows

The data file

Opening a data file	To open an existing data file File extensions : filename.dft and filename.cbx	
🖱 ➤ From the File menu, choose Open	⌨ ➤	(ALT, F, O)

The following dialog box appears :
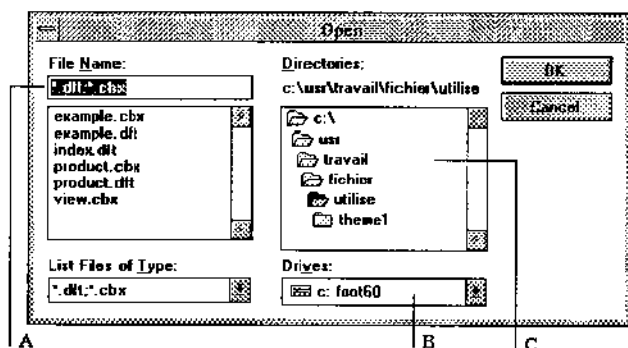
A : type or select the file you want to open.

B : select a new drive, if necessary.

C : select a new directory, if necessary.

🖱 ➤

Choose the OK button.

See also "Many views, one datafile" (Working with views) on page 18-19.

Closing a file	To open another file without quitting Cub.X	
🖱 ➤ From the File menu, choose Close.	⌨ ➤	(ALT, F, C)
Then : open a data file as above.		

Information on the datafile	To open an additional information file File extension : filename.hlp	
🖱 ➤ From the "Information and Help" menu, choose Information on the table	⌨ ➤	(ALT, I, I)
Warning ! Data files may not have this additional information.		

Note : when you open a data file, an "Information and Help" window *may open* automatically, and hide a part of the .DFT window : it gives you *additional information* on the table, the origin of values, dimensions.... *should your table has an additional information file.*

The index file

An index is a file that organizes your data and allows you quicker access to them. If your tables have a lot of data, *it is necessary to use an index* with the table.

If you don't really know how to work with an index,

⇨   **do not** delete the first index you receive, and

⇨   **do not** activate the command confirm index in the Options menu.

CUB.X will automatically read the index when you open the data file.

If you decide not to use the index for a particular reason, you can activate an option that will ask you if you want to use the index, each time you open a table.

Confirming an index	To use an index
	File extension : filename.IDX
🖱 ➤ ***Before opening*** the file *.DFT, click the Options menu and look at Confirm index	

If the command is checked (✔), the *index confirmation* is active. So, *the index is NOT automatically used* : **you will have to confirm its use.** A confirmation message will appear.

If the command is not checked , it means that the index confirmation is not active. So, *the index will be used without notice.*

Consultation of a data file :

screen description

Cub.X is organised like a page of a spreadsheet ; it allows data to be viewed, selected and exported. The window shows only a part of the page that may contain more rows and columns than can be displayed on the screen.
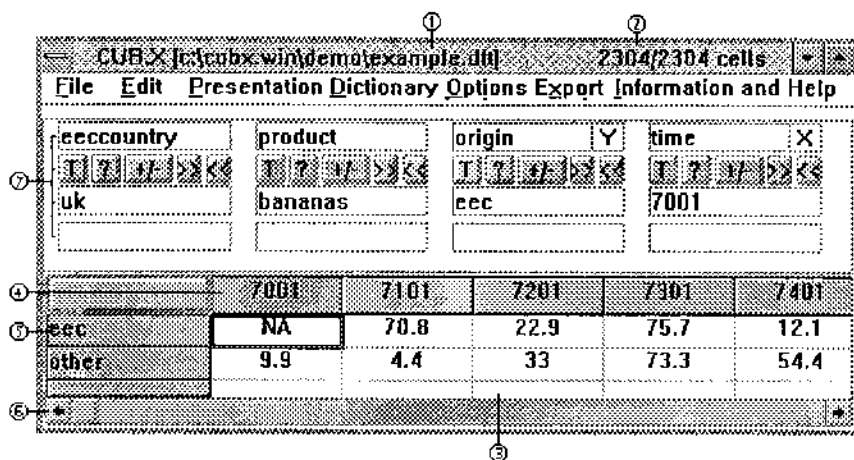
① Complete name of the current file
② Number of cells
③ Data presentation
④ X axis, one of the dimensions
⑤ Y axis, one of the dimensions
⑥ Horizontal scroll bar
⑦ Dimension blocks

① **Current file name**
Drive, directories and file names are between 2 square brackets.

② **Number of cells**
The first number (2304/...) is the number of selected cells with ▦▦ ; the second (.../2304) is the number of all the cells defined in the file.

③ **Data presentation**
The intersection of the X and Y axes, and also the selected positions in each other dimension block (set of positions - see page 9).

④ **X axis**
Horizontal axis. A value of the X axis defines a column.

⑤ **Y axis**
Vertical axis. A value of the Y axis defines a row.

⑥ **Horizontal (vertical) scroll bar**
Appears when the window size doesn't allow the whole table to be displayed.

⑦ **Dimension blocks**

8

```
command line        eeccountry    Y        title of the
for data                                     dimension
presentation        [toolbar icons]
                                             origin of the
label of            UK                       dimension
the origin          United Kingdom
```
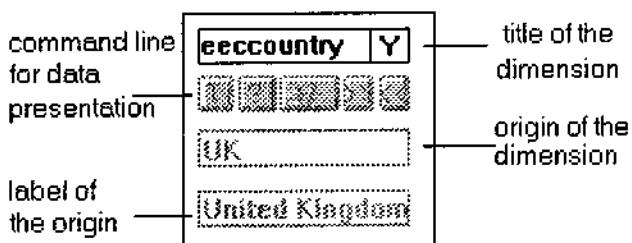
Y	**Axis (X or Y)**	**T**	**Labels**
	Indicates that the dimension is used as X or Y axis.		"Removes" all the dimension labels. To display all the labels, use the command Options/Labels.
[icon]	**Information**	[icon]	**Select / Unselect positions**
	Gives you more information about the dimension and its positions.		Allows you to select or unselect positions of the dimension. Useful for exporting a subset of the data.
[icon]	**Go forward**	[icon]	**Go back**
	Go to the next position(s) of the dimension.		Go to the previous position(s) of the dimension.

eeccountry

Title of the dimension

Name of the dimension. This name corresponds to the associated dictionary (eeccountry.dic, for instance).

UK

Origin of the dimension

This button allows you to select directly any position you want to see. By clicking this button, the list of all the positions appears. Click on one of them to select it.

United Kingdom

Label of the origin
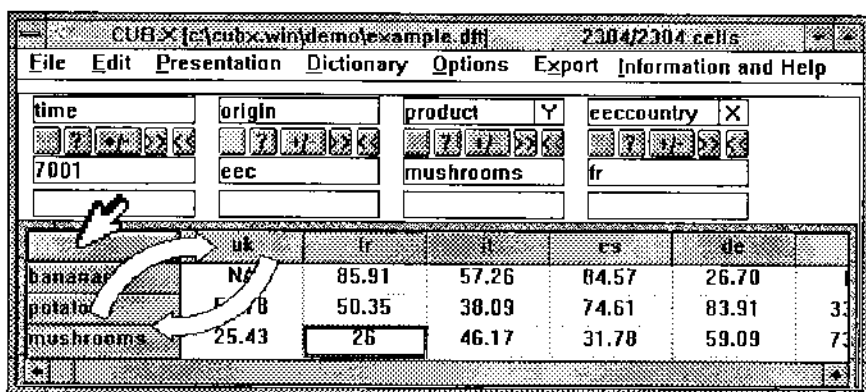
When CUB.X finds the associated dictionary of a data file, it puts the labels in this box. If you don't see the entire label, click on it : the label will appear in a label box.

Displaying data

It is useful to be able to rotate axes to see better some set of values. In the following figure, the arrow shows you where you have to click to swap the two axes.

CUBX [c:\cubx.win\demo\example.dft] 2304/2304 cells				
File Edit Presentation Dictionary Options Export Information and Help				

time	origin	product	Y	eeccountry	X
7001	eec	mushrooms		fr	

	uk	fr	it	es	de	
banana	N/	85.91	57.26	84.57	26.70	
potato	/8	50.35	38.09	74.61	83.91	3:
mushrooms	25.43	26	46.17	31.78	59.09	7:

This command allows you to have only one file with one multi-dimensional table, but many views of the data.

Three ways are provided to do this very important manipulation :

- ☉ drag and drop a dimension title toward an axis.
- ☉ drag and drop a dimension block title towards another one.
- ☉ use keyboard shortkeys to change quickly the place of each dimension.
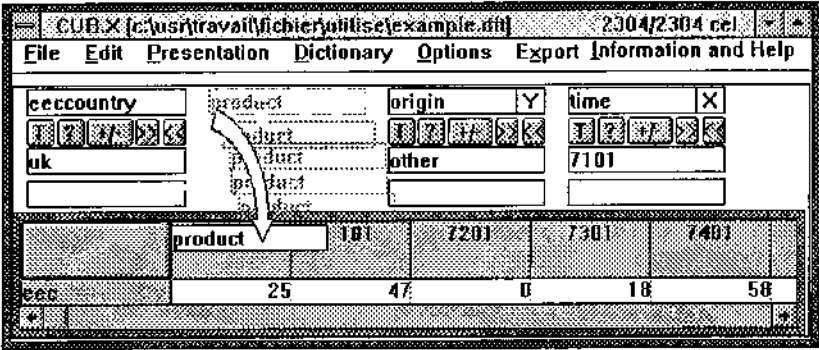
Move cursor to the title of the dimension block. Click and hold down the left button, drag towards an axis, and drop (by releasing the mouse button).

Note : this option is applies only to dimensions other than X and Y.

In this example, product becomes the new X axis, and time will take the previous place of product. The order of the other dimensions is reorganised.

⌨ ➢ **CTRL + R** to shift the dimension blocks to the right

⌨ ➢ **CTRL + L** to shift the dimension blocks to the left.

> Press **CTRL + G**

First, you have to decide what you want to present on the chart.

— The X axis of the table will be the X axis of the chart, in the example below the countries.

— The selected position of the Y dimension will define the Y axis. In the example below, the position mushroom will be the dimension product.

— The current line of the table (the line where the cursor is) will be represented on the chart.
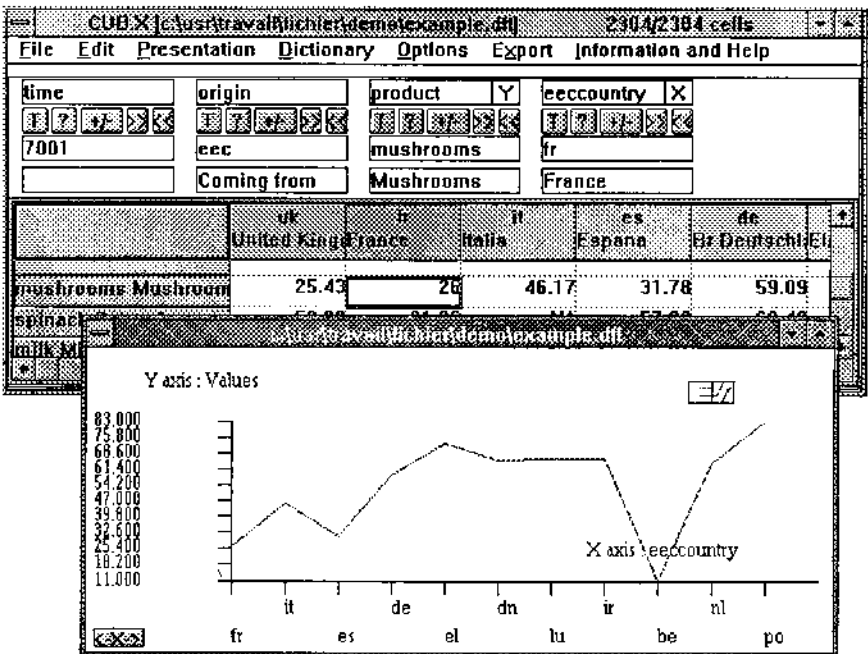
Note : Cub.X is not designed for charting : this is a rough and ready facility for providing a graphical view of the selected data. Making more complex charts, data manipulation and calculation, is left to your preferred software. (See "Export manipulation" page 23).

> Click this button to reverse the axes order on the chart sheet

<img> ➤ From the Options menu, choose Justify. ⌨ ➤ (ALT, O, J)

You can chose between 3 options : Left , Center or Right Justify. The alignment you choose will be kept for the next session of CUB.X whichever file you open.

Displaying a grid
To choose the aspect of the table

<img> ➤ From the Options menu, choose Table. ⌨ ➤ (ALT, O, T)

You can chose between 3 options. The option that you choose will be kept for the next session of CUB.X whichever file you open.

Grid			Panel			None		
	bananas Bananas	potatoes Potatoes		bananas Bananas	potatoes Potatoes		bananas Bananas	potatoes Potatoes
uk United	NA	56.70	uk United	NA	56.70	uk United	NA	56.78
fr France	85.91	50.35	fr France	85.91	50.35	fr France	05.91	50.35
it Italia	57.26	38.09	it Italia	57.26	38.09	it Italia	57.26	38.09
es Espña	84.57	74.61	es Espña	84.57	74.61	es Espña	04.57	74.61

Displaying all the labels
Display the labels associated with each position of all the dimensions

To **activate** all the labels for all the dimension :

<img> ➤ From the Options menu, **check** the ⌨ ➤ (ALT, O, L)
command Labels

Table with labels activated :

	eec Coming from Economic European Communities	other Coming from outside the EEC
uk United Kingdom	NA	9.96
es Espana	84.57	49.13
de Br Deutschland	26.70	26.57
el Elada	NA	44.48
dn Danmark	3.87	81.26
lu Luxembourg	61.04	71.91
ir Ireland	77.17	81.74

Table without labels activated :

	eec	other
uk	NA	9.96
es	84.57	49.13
de	26.70	26.57
el	NA	44.48
dn	3.87	81.26
lu	61.04	71.91
	77.17	81.74

> Click [T] to remove all the dimension labels.
> The [T] becomes [ ].

> Click [ ] to display again all the dimension labels.
> The [ ] becomes [T].

You can selectively activate the labels of only some of the dimensions.

If you want few labels, it is better to remove the check mark of the menu Options/ Labels , whereas if you want a lot of labels, it is better to check this option.

To see the entire label   When labels are too large to be read

> Point to the label on the dimension block, click and hold down the left button of the mouse.

-or-

> Point to the label in the table, click and hold down the left button of the mouse.

> Release the left button and the label disappears.

Example :

Decimals   To choose the number of decimals

> From the Options menu, choose   [⌨] > (ALT, O, D)
> Decimals

Columns/Rows — To size all columns and/or rows

🖱 ➤ From the Options menu, choose ⌨ ➤ (ALT, O, C)
   Columns/Rows

➤ Use the mouse to select a size in one of the dialog boxes

OR ➤ Use the mouse to size **directly one** row and/or **one** column :

	bananas Bananas	potatoes Potatoes	mushroor Mushrooms
uk United Kingdom	NA	56	25
fr France	85	50	26
it Italia	57	38	46

Restore native order — Used to reset the original presentation after making changes in the display of the table

🖱 ➤ From the Presentation menu, choose ⌨ ➤ (ALT, P, N)
   Restore native order

Reset ranges — To cancel everything done with [+/-]

🖱 ➤ From the Presentation menu, choose ⌨ ➤ (ALT, P, E)
   Reset ranges

Refresh screen — To rebuild a disturbed display

🖱 ➤ From the Presentation menu, choose ⌨ ➤ (ALT, P, R)
   Refresh screen

Change of language

🖱 ➤ From the Options menu, choose ⌨ ➤ (ALT, O, L,➔,
   Language ↓ or ↑)

➤ Select one Language Press Enter

Refer also to "Dictionary manipulation".

> Click on the ▓▓▓ button of the corresponding dimension block. This dialog box appears :
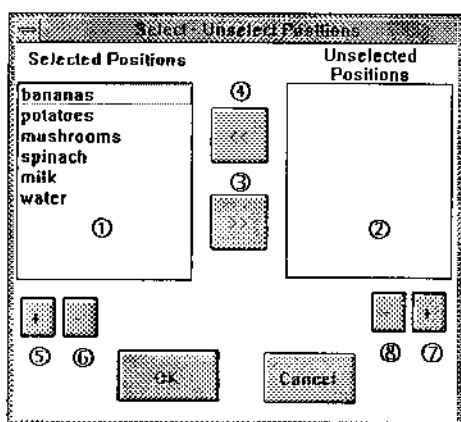
① > This box displays all the selected positions of the dimension (the positions to be shown on the screen).

② > This box displays all the unselected positions of the dimension (the positions not to be seen on the screen).

The button ③ allows you to move the highlighted positions from box ① to box ② : these positions will be **unselected and not shown.**

The button ④ allows you to move the highlighted positions from box ② to box ① : these positions will be (re-)**selected and shown.**

By clicking one of the buttons ⑤ - or ⑦ -, a dialog box appears. In this dialog box, you can enter an expression, called a regular expression, that allows you to activate positions (highlighted positions) to be selected in box ①, - or in box ② respectively -.

By clicking one of the buttons ⑥ - or ⑧ -, a dialog box appears which allows you to deactivate some highlighted positions, from box ①, - or from box ② respectively -. The mechanism of regular expressions is used (cancel some selections made in box ①, - or in box ②-).

This button validates all your changes.

Cancel any previous operations in this dialog box.

Notes :

⇨ These selection(s) do not imply that your values will be deleted from the table.

⇨ You will find a detailed explanation of regular expressions and their use in the Memo on page 28.
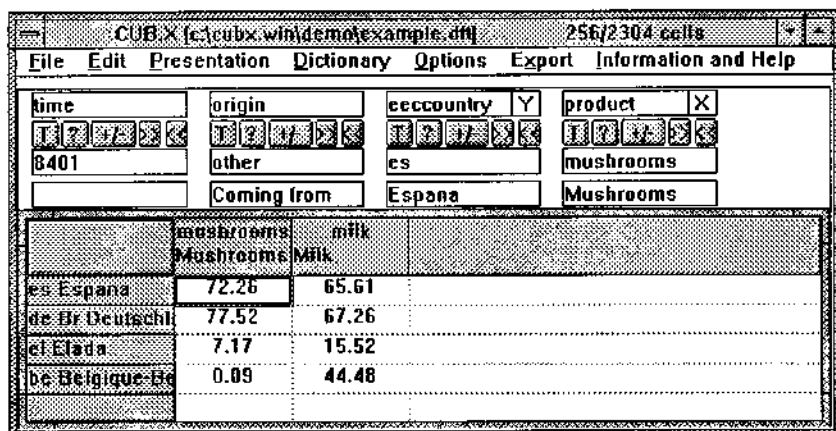
Example

Initially, *all* the dimensions values can be seen. Suppose you may want to see only some of the position.

To **unselect** some positions :

1   Click [+/-] in the dimension block to unselect some positions.

2   In the selected positions box ①, click all the products you don't want to see, to unselect them

3   Click [»] : the highlighted products are transferred automatically to the unselected positions box.

4   Click OK

The result is the same table with only the positions that interest you.

| — | CUB.X [c:\cubx.win\demo\example.dfl] | | 256/2304 cells | ▼ | ▲ |

| File | Edit | Presentation | Dictionary | Options | Export | Information and Help |

time	origin	eeccountry	Y	product	X
8401	other	es		mushrooms	
	Coming from	Espana		Mushrooms	

	mushrooms Mushrooms	milk Milk			
es Espana	72.26	65.61			
de Br Deutschl	77.52	67.26			
el Elada	7.17	15.52			
be Belgique-Be	0.09	44.48			

To re-select those positions, you just have to do the same thing in box ② and click [«]

Many views, one datafile

Introduction

A view allows you to have a certain way of looking at your data, specifying a presentation and selection. It is similar to saving your file with another name ; but saving a view (.CBX) is much more interesting than saving a file (.DFT).

To open a view	File extension : filename .CBX
⬧ ➤ From the File menu, choose Open	⌨ ➤ (ALT, F, O)

Select the view that you want to open (a name followed by the extension .CBX). The view opens the corresponding data file with the data represented as you choose.

To save a view	To save or create a presentation that you would like to work with (to export, to read...) File extension: filename .CBX
To store a view on your hard disk :	
⬧ ➤ From the Presentation menu, choose View. Then choose the option Save a view	⌨ ➤ (ALT, P, V, S)

☝ **Warning !** Make sure that the file keeps the extension .CBX.

Working with views

You have two ways to work with an existing view of a file :

⇨ You open directly the view (.CBX), instead of opening the data file (.DFT).

⇨ You open the data file (.DFT), and then, when you need it, you **load** the attached view.

To load a view	To load an existing view attached to the opened .DFT datafile. File extension : filename .CBX
🖰 ➤ From the Presentation menu, choose View. ➤ Then choose the option Load a view.	⌨ ➤ (ALT, P, V, L)

When you release the mouse button, CUB.X allows you to save current changes in a new view :

Save a view
Do You want to save the current view ?
[Yes] [No]
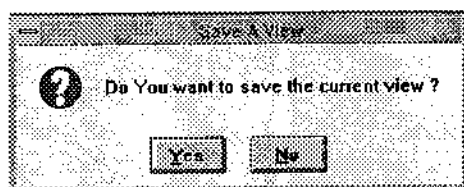
If you want to save it, follow the instructions of the paragraph "To save a view".

A dialog box appears, asking you to choose the view (.CBX) that you want to load in the file list box, for example, PRODUCT.CBX.

You are in the same data file, but with another "View".

You can load many views in a CUB.X session (one at a time).

Note : you can't load a view that is not attached to the open data file because a view is linked to a unique data file. In case of mismatch, you get this message :

Error
Filename c:\cubx.win\demo\example.dft
defined in view file does not match
current filename c:\usr\travail\fichier\demo\example.dft
[OK]
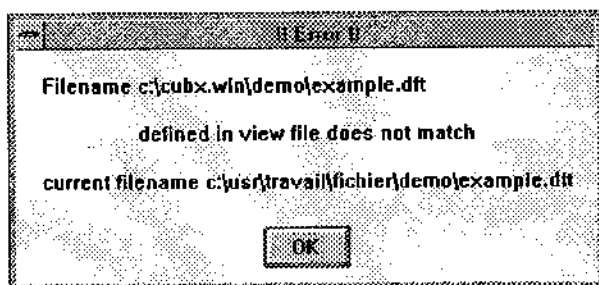
Dictionary manipulation

Introduction

When you see the dimensions with their corresponding labels, it means that CUB.X has found the dictionaries and associated them with each dimension. All the following conditions have to be respected :

⇨ the associated files .DIC exist.

⇨ the dimension and the file .DIC have the same name. The file name is limited to 8 characters. So, for example, for the dimension "eeccountry", the corresponding dictionary will be "eeccount.dic".

⇨ CUB.X finds the dictionaries in the defined directory built automatically from the dictionary root. If CUB.X doesn't find any dictionaries, you have to indicate where they are stored. (See also page 22 : "Change dictionary root").

Checking dictionary

Dictionary and dimension	To verify what each dimension uses as dictionary
⌐ ➤ From the Dictionary menu, choose Dimension	▦ ➤ (ALT, D, I)

This dialog box appears, and you can verify each dimension :

The storing of dictionaries follows the same logical structure as the data files.
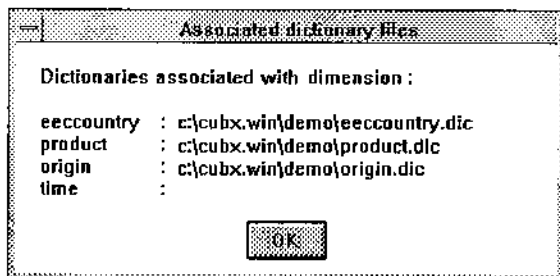
Associated dictionary files

Dictionaries associated with dimension :

eeccountry	: c:\cubx.win\demo\eeccountry.dic
product	: c:\cubx.win\demo\product.dic
origin	: c:\cubx.win\demo\origin.dic
time	:

OK

➤ From the Dictionary menu, choose   ⌨ ➤ (ALT, D, E)
Search path for the labels

There are many paths, because there are a lot of possibilities to store these files. CUB.X builds these paths from the root directory which is the basic directory used to find .DIC file for each dimension ; (to build the root dictionary, see the next page).

Example :

In this figure, you can see all the paths that CUB.X builds and where it searches for the labels (the files .DIC) to associate them with each corresponding dimension.

Search path for labels

Order of searching "labels files" :

root is : c:\cubx.win\demo

0 :
c:\cubx.win\demo\doc\en\base_example\domain_example\collection_cub.x
1 : c:\cubx.win\demo\doc\en\base_example\domain_example
2 : c:\cubx.win\demo\doc\en\base_example
3 : c:\cubx.win\demo\doc\en
4 : c:\cubx.win\demo\base_example\domain_example\collection_cub.x
5 : c:\cubx.win\demo\base_example\domain_example
6 : c:\cubx.win\demo\base_example
7 : c:\cubx.win\demo

[OK]
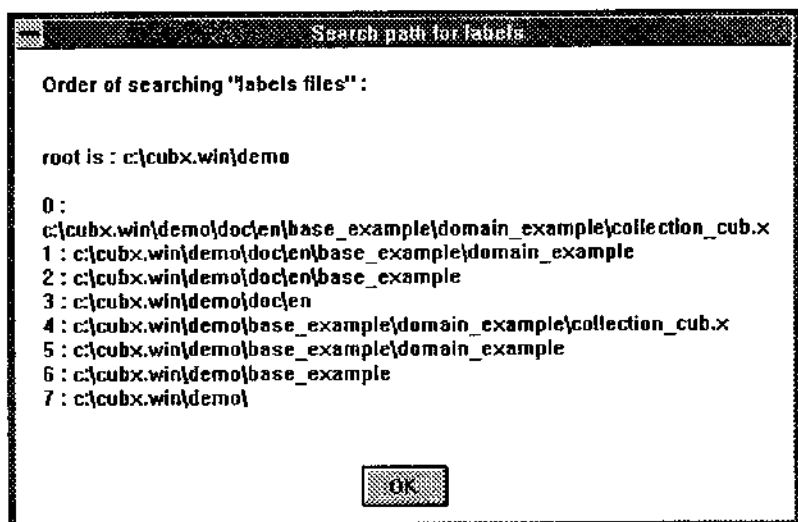
The chosen language is English ; in the path, you can see "en", because the English dictionaries are in the directory "en".

In accordance with the selected language (Options/Language), CUB.X builds the paths and can open the dictionary in the corresponding language.

For example, choose the language French. To check the new path, choose the command *Search path for labels* from the menu Dictionary. The "en" from the previous figure has become "fr".

Change of language

Refer to "Displaying data"

 ➤ From the Dictionary menu, choose
Dictionary root.

 ➤ (ALT, D, D)

➤ Choose the dictionary directory

The dictionary mechanism is activated from the directory root. This top directory is the basic directory which is used to find the .DIC file for each dimension.

Example

① Choose the drive where Cub.X will find the dictionaries. ➔ ② Go on the root of the drive ➔ ③ Choose the "top" directory of the dictionaries
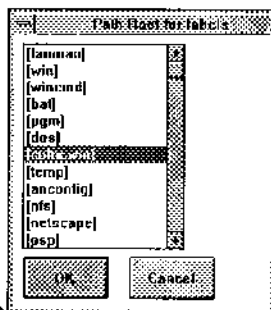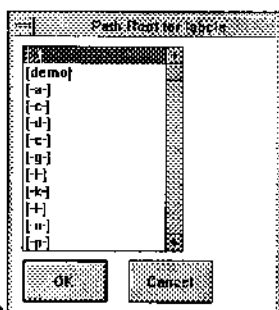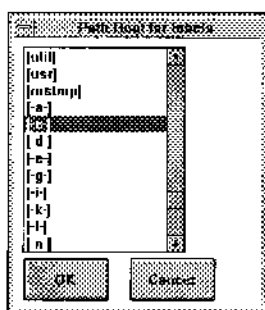
When you are on the root of the drive, and see all the top directories of the drive, select the top directory where the dictionaries are stored, but **don't go further.**.

The screen shows you the directory used for the demo example, which possibly doesn't correspond to the directory of your data files.

1   **Click (only once)** on the top dictionaries' directory, to select it.

2   Click OK

When you select the root directory, you must stay on the main directory **without going further** because Cub.X builds the path automatically. If you should go deeper in the tree structure, Cub.X will build an erroneous path.

22

Export manipulation

☝ *Important before exporting data*

Exporting a lot of data can use up a lot of resources and *risks freezing your PC for a very long time*.

Number of cells to export = number of selected cells

| ═ | CUB.X [c:\cubx.win\demo\example.dfi] | 540000/1200000 | ▼ | ▲ |

It is better and faster to export several two-dimensional tables
(the displayed XY table that interests you),
rather than a large multi-dimensional table.

Introduction

You use the Cub.X command Export to export data or to be able to read data in different formats. Remember that whatever method you choose to export data, only the selected data will be exported.

Two ways to export data to be used in another application are provided :

① Use the Windows' Clipboard
② Use the Export facility

① 🏢 Windows' Clipboard

Copy all the positions	You want to export the displayed table (X and Y) with all its positions
🖱 ➤ From the Edit menu, choose Copy XY all	⌨ ➤ (ALT, E, A)

Copy selected positions	Having unselected some positions in a presentation, you want to export the displayed table (X and Y) with only its selected positions
🖱 ➤ From the Edit menu, choose the command Copy XY selected	⌨ ➤ (ALT, E, S)

② Export facility

Export multi-dimensional tables

To export XY	To export the data displayed in a two-dimensional table

🖱 ➤ From the Options menu, see the command
XY export mode (otherwise All)

⇨ To export XY, the command must be *checked*.

Click the menu Export, and choose the target file format.

Type a new name for the target file.

To export all the possible combinations	To export all the data

🖱 ➤ From the Options menu, see the command
XY export mode (otherwise All)

⇨ To export all the data, the command must be **NOT** checked.

Click the menu Export, and choose the target file format.

Type a new name for the target file.

Note : if labels are not displayed, they will not be exported. Please see "Display data".

☝ Do not use reserved extensions like .XLS, .DOC, etc. It is better to make the distinction between your exportation files and other applications files.

Export formats

Choose the export format you want to use : Table, Stream or Eurostat

➤ select again between the options provided	⌨ ➤ (➔, ↓ or ↑) Press Enter

Table : three different table formats are available.

DIF   Data Interchange Format
Only for XY tables and accepted by most of the well-known spreadsheets on the market.

EXCEL/TAB This option allows tables to be input to the EXCEL system. In fact, this format is a "tab delimited" format, each cell being enclosed by a double quotation mark <"> and separated with a <tab>.

CSV   Comma Separated Values
This format is the same as that of EXCEL except that the delimiter between all fields is the comma <,>.
This kind of file is accepted, for instance, by the IMPROV program.

Stream : three different kinds of file are available.

dBASE   Accepted by dBASE software, this file has X + 1 fields with
X = number of dimensions
1 = value (last field).

Text   This format is a text file with X + 1 columns
X = number of dimensions
1 = value (last column)
The separator between columns is a <space>.

CSV   Same format as Text but using double quotation marks <"> to enclose the contents of each column and a comma to separate the columns.

Eurostat :   the three formats (Fame, Osiris and Gesmes) are internal to Eurostat.

Help

Help Menu

⊞⁄🖰 ➤ From the Information and Help menu, choose Help	⌨ ➤ (ALT, I, H)

Note : you are in a file managed by MS-Windows.

Help Interface
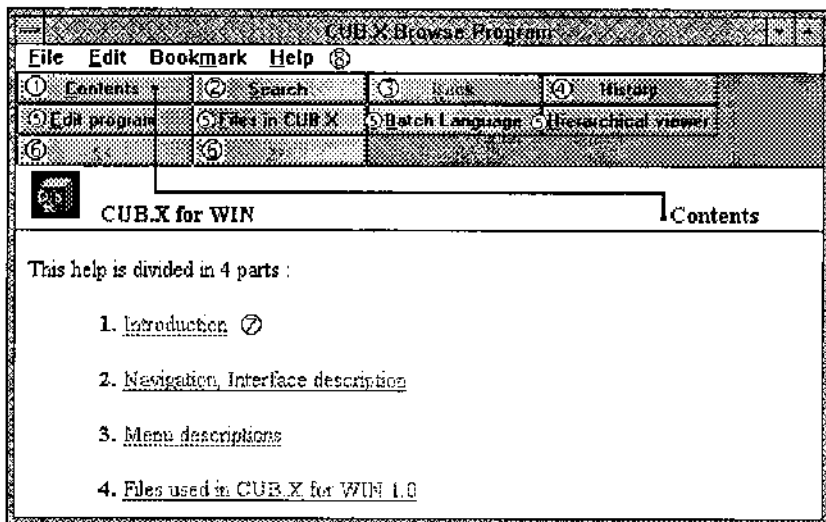
① to ④ : managed by Windows' Help program.

⑤ Opens automatically the specific help files.

⑥ 2 buttons : | < < | to go to the previous page.

   | > > | to go to the next page.

⑦ Jump to a cross-reference : if you click any term or phrase in colour and/or underlined, you go straight to the specific part of the help file.

⑧ Menu Bar : managed by Windows' Help program.

This page is left intentionally blank for your notes...

A memo on

Regular expressions (Unix)

Definition

A regular expression is a form used to specify and match strings. Regular expressions use *metacharacters* (characters with a special meaning) and *nonmetacharacters* (normal characters) as described below.

Introduction

The regular expression allows you to make some selections (to highlight them) or to remove existing selections (to remove the highlighting) in order to Select or Unselect them.

Using regular expressions replaces a series of the actions with the mouse in the dialog box before clicking [   ] or [   ] to effect the transfer between the selection and unselection boxes :

1   Click the select button [   ]

2   In the "select - unselect positions" dialog box that appears, you click [   ] to highlight the item to select or unselect ; and you click [   ] to remove this highlight.

3   The following dialog box appears : you enter a regular expression that will allow you to highlight (or to remove) the existing items.

> **Select Positions**
>
> Enter a pattern (UNIX regular expression)
>
> `N.*`
>
> [OK]   [Cancel]
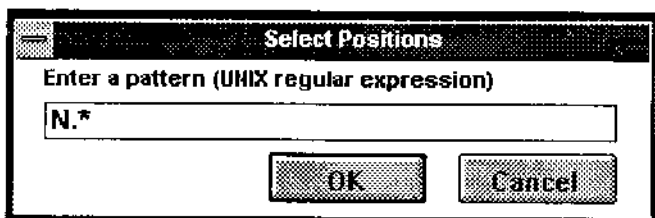
A regular expression is a formalised way of describing a pattern which is then used to search (in this case) the labels with a view to select or unselect automatically.

For example, the regular expressions 'N.*" finds all the labels beginning with a capital N (N followed by zero or more character of any sort - "." meaning 'any single character" and "*" meaning the "." occurring 0, 1, 2... times).

Regular expression Metacharacters

To build a regular expression, you will need some of the following metacharacters. Here are all the metacharacters (MC) supported in CUB.X.

MC	Name	Meaning	MC	Name	Meaning
\	back-slash	protection	(	left parenthesis	begin of expression
^	caret	negation	)	right parenthesis	end of expression
[	left square bracket	begin of class	*	asterisk	repeat 0,1,2,... times the expression to its left
]	right square bracket	end of class	?	question mark	repeat 0 or 1 time the expression to its left
|	pipe	logical OR	+	plus	repeat 1,2,3... times the expression to its left
.	dot	any single character			

Basic expression

A basic *expression* is one the following :

Nonmetacharacter
A   If you write A, you will match the letter A

Quoted metacharacter
*   Matches one of the special metacharacters, in this case the asterisk "*".
With "\", you cancel the effect of the metacharacter.

Character class
[ABC]   Matches any one of the characters A B or C ;

[A-Z0-9]   Matches one letter from A to Z (only in uppercase) or a digit.

Complemented character class
[^0-9]   Matches any character except a digit.
          Note : the ^ (negation) is active throughout the class.

Operators

These operators combine regular expressions into longer ones :

Alternation   A|B   Matches A or B (uppercase)

                (AB) | (CD)   Matches AB or CD, but not A or B or C or D

Concatenation   AB   Matches uppercase A immediately followed by
                                    uppercase B

Closure   A*   Matches zero or more A's : { $\varnothing$; A ; AA ; A...}
                                    (nothing or one or many uppercase A's)

Positive closure   A+   Matches one or more A's : {A ; AA ; A...} (at least
                                    one uppercase A)

Conditional closure   A?   Matches zero or one A's : { $\varnothing$; A} (nothing or
                                    one A)

Parentheses   (r)   Matches the same strings as the expression 'r'
                                    does ; used to group many expressions, and also
                                    to give them priorities.

Use the labels for selection

To make selections with the labels of the dimension	working with the regular expression
↳🖰 ➢ From the Options menu, look at the command +/- **Labels**.	

When you use a regular expression, you can do the selection on the entire label or on the abbreviation ('Belgium" or 'be') : in the selection boxes, you will always have the abbreviations, so you must work with labels from memory, or place the dialog box in a such way that you can see the labels from the displayed table.

> ➢   To scan the entire set of labels with a regular expression, there must be a check mark (✓).

For example : for the dimension 'eeccountry", you can write 'Belg.*" to select Belgium, only if you use the entire label.

No check mark means that you use the abbreviation "b.*"

INDEX

Eurostat contacts

Number of version of CUB.X : ....

Contact :
"BUREAU DE DIFFUSION" of Eurostat ☎ +352 4301 32628
                                    🗐 +352 4301 32594

For a *technical problem* call :

                        Daniel DEVILLE ☎ +352 4301 34940
                                        🗐 +352 4301 32594

Minimum Hardware and System Requirements

- ◆ a 80486 based computer,

- ◆ at least 2 megabytes (MB) of random-access memory (RAM),

- ◆ at least 2 megabytes of hard disk to install CUB.X for WIN,

- ◆ MS-DOS 3.3x and MS-Windows 3.1 (Enhanced mode) or later.

PART II

Methodology

Introduction

The methodology begins with some introductory reflections on concepts and practical difficulties the statistician must face when elaborating statistics on SMEs. The following sections explain the major characteristics of the SME tabular database and the possible country deviations. The reader will find a summary table at the end providing the basic information by country necessary to interpret the results in this publication: main source, analytical unit, size and sectoral coverage, availability of economic variables, exchange rate of the national currency with the ecu.

Detailed information on the country methodologies necessary to analyse and compare the data correctly is provided next. Although the structure of the data is simple, it does not necessarily allow automatic processing, but often requires the consideration of exceptions, due to data availability problems and differing collection systems in the EU Member States.

Main elements of the SME tabular database, provided on the diskettes:

Geographical coverage	The whole of each country
Analytical unit	Enterprise
Sectoral classification	NACE 70, three digit level
Size class breakdown (number of employees)	0, 1-9, 10-19, 20-49, 50-99, 100-199, 200-249, 250-499, 500 +
Economic variables	Total employment
	Salaried employment
	Turnover
	Value added
	Labour costs

A. General Methodology

1. Statistics on SMEs: relevant concepts and practical difficulties

The whole problem with policies for supporting small firms rests on the theory that a small unit has certain features which distinguish it from a large one and, above all, that it has specific economic difficulties. The job of the statistician is to put together a set of information which will give policy-makers clear guidelines, while ensuring that the indicators are feasible, easy to interpret and internationally comparable. Three basic questions, not always easy to answer with the information currently available, have to be asked.

Defining a statistical unit

In most cases it is clear that policy-makers are interested in units with strategic power, particularly over investment decisions. Statisticians find this elementary decision-making unit in the enterprise concept set out in the European regulation on statistical units.

The definition of an enterprise requires a coherent economic entity to be constructed from legal entities. This does not, however, guarantee international comparability which is entirely adequate for analysis by size class. Deciding how to aggregate legal entities is often a complicated matter.

In every Member State, the organization and size of businesses is influenced by institutional conditions (taxation systems, social contributions, subsidies available) but also by the nature of the economic fabric and managerial culture (in some countries it is standard practice to call in an external firm as a subcontractor or provider of services; in other countries this is unknown). This may affect the breakdown not only by size class, but also by sector of activity.

It is therefore useful when making comparisons between countries or over time to have an idea of the effects of national legislative and economic contexts on the structure of units.

The use of the 'group of enterprises' unit would make the figures more comparable by systematically covering all units belonging to a group above a specific participation percentage, and would also be better attuned to the problems of SME policy-making, since belonging to a powerful group brings strategic opportunities not available to small independent firms. This unit is, however, impossible to construct at present in most Member States of the European Union.

Choosing the criteria for breakdown by size class

Once the unit is selected, the criteria (employment, turnover, other financial ratios, etc.) and the classes likely to bring out the specific features sought in the target population have to be chosen.

Recently, the Commission has adopted a communication harmonizing the definitions in current use at the Union level[1]. This text gives thresholds based on employment and turnover (or total assets) plus a criterion of not belonging to a group.

At present, most available statistics do not enable all these criteria to be cross-referenced; the most commonly-available breakdown is employment.

A single definition put forward for political or administrative reasons may seem restrictive for in-depth analyses. The response of the statistician will be to give analysts a measure of flexibility by suggesting narrower size classes, making it possible to work on a larger number of sub-populations. Eurostat's tabular database on SMEs is tending in this direction, thanks to the cross-referencing of detailed breakdowns by size class and sector of activity.

How to interpret trends

Interpretations of trends over time by size class may lead to misunderstandings. An interpretation of the growth of economic variables corresponding to a given size class should take into account the effect of firms entering and leaving the size class, as well as the actual growth of firms remaining in it.

The difference between net entries and growth proper is extremely difficult to calculate, since individual data have to be followed up over time. This is not usually the case with the aggregated data available, which are based on cross-sections of the population at a given moment. The best way of obtaining a detailed analysis of the dynamics of certain categories of firms would be to conduct longitudinal studies (linkage over time of individual data to firms). Such approaches are increasingly common and very promising.

The possible scale of the effects of changes in size classes should be constantly borne in mind when using aggregated data. Growth rates may be misleading: the population in a class containing very few individuals may easily be doubled by the arrival of a single firm. Since data broken down into a number of sub-populations often include classes with small numbers, the small numbers effect should be taken into account by calculating, for example, the contributions to total change.

---

[1] Commission Recommendation No 96/280/EC of 3 April 1996 concerning the definition of small and medium-sized enterprises (text with EEA relevance).

2. Main sources of the SME tabular database

For most of the countries, data have been produced by public institutions.

These are primarily:

* the national statistical institutes: in many cases they integrate different official data sources such as enterprise censuses, the VAT register, the statistical business register, results of surveys, etc.;

* the social security authority, which is a major source of employment data.

    In some cases, data come from private institutions.

In general, the SME tabular data are created from several sources, either at the national level, or from Eurostat.

This is necessary when:

* one source cannot provide all the variables requested (example: the social security authority generally cannot provide data on turnover or value added);

* one source does not cover all economic sectors (in several countries surveys are carried out for some sectors as VAT registers do not contain exempted sectors);

* one source cannot provide information on all enterprise sizes (example: industrial censuses generally do not cover the smaller firms; social security registers do not include enterprises without employees).

3. Classification of economic activities

Data on the diskette are broken down according to the NACE 70 nomenclature, established by Eurostat in 1970. From 1993 onwards, data are classified by the more recent NACE Rev. 1 nomenclature.

The NACE 70 divides economic activities into 10 broad divisions (one digit levels, for example '4' is 'Other manufacturing industries') which are subdivided further into more detailed classes (two digit levels, e.g. '43' is 'Textile industry') and further into groups (three digit levels, e.g. '433' is 'Silk industry') and even more detailed sub-groups. The exhaustive list of all NACE 70 divisions and classes is given in an annex.

Most countries provide the data directly in the NACE 70 classification. However, several countries have applied a different nomenclature, either a national one, or the ISIC (International Standard Industrial Classification from the UN).

In these cases, Eurostat had to convert the data to the NACE 70 classification.

The more detailed the data provided by the countries, the greater the success of the conversion procedure. However, in most cases, sectoral aggregations and exceptions cannot be avoided.

Data cover all non-agricultural market activities, a definition which mainly excludes the following economic sectors:

* agriculture, hunting, forestry and fishing (NACE 0);
* central banking authorities (NACE 811);
* public administration, national defence and compulsory social security, (NACE 91);
* other services provided to the general public, e.g. social work, trade unions, tourist offices, etc. (NACE 96);
* domestic services (NACE 99);
* diplomatic representations, international organizations and allied armed forces (NACE 00).

The separation of market and non-market services, which affects mainly 'other services' (NACE 9), is not always possible in a satisfactory way, especially if data are not provided in NACE 70 and have to be converted. Comparisons of country results for this economic activity should therefore be carried out cautiously, especially when a high level of disaggregation is applied.

4. Units

The enterprise is the unit principally used in the national SME data, but several countries cannot provide enterprise data and thus send data on local units, kind-of-activity units or others. Most of the statistical units used are defined in the Council Regulation (EEC) on the statistical units for the observation and analysis of the production system in the Community.[1]

The enterprise

is the smallest combination of legal units that is an organizational unit producing goods or services, which benefits from a certain degree of autonomy in decision-making, especially for the allocation of its current resources. An enterprise carries out one or more activities in one or more locations. An enterprise may be a sole legal unit.

The local unit

is an enterprise or part thereof (e.g. a workshop, factory, warehouse, office, mine or depot) situated in a geographically-identified place. At or from this place economic activity is carried out for which (save for certain exceptions) one or more persons work (even only part-time) for one and the same enterprise.

The kind-of-activity unit (KAU)

groups all the parts of an enterprise contributing to the performance of an activity at a detailed sectoral level (four digit level of NACE Rev. 1) and corresponds to one or more operational subdivisions of the enterprise.

The local kind-of-activity unit (local KAU)

is the part of a KAU which corresponds to a local unit. Several countries send data on establishments. The definition of an establishment comes closest to the definition of a local KAU.

The enterprise group

is an association of enterprises bound together by legal and/or financial links. A group of enterprises can have more than one decision-making centre, especially for policy or production, sales and profits. It may centralize certain aspects of financial management and taxation. It constitutes an economic entity which is empowered to make choices, particularly concerning the units which it comprises.

The other units used by some Member States are legal units: the employer or the VAT unit; they do not correspond directly to one of the definitions above and are defined in national administrative registers.

Due to such specific definitions, data for different statistical units are not fully comparable. The size class breakdown should be read with caution as the definition of a unit strongly influences its size.

---

[1] Regulation (EEC) No 696/93 of 15 March 1993.

42

5. Definition of the size classes

The main characteristic of the SME data is its breakdown by employment size classes. The size class boundaries are expressed in terms of employees; this allows us to separate sole proprietorships (size class zero) from the rest of the enterprise population, as they employ by definition only non-salaried staff, consisting of the entrepreneur and often family members or persons working under a very small contract.

In accordance with Eurostat's data request, most countries send the national data using the following detailed employment size classes:

- 0
- 1-9
- 10-19
- 20-49
- 50-99
- 100-199
- 200-249
- 250-499
- 500 +

6. Economic variables

As well as the number of enterprises by size class, the SME data contain several economic variables:

employment: number of persons employed and number of employees;
total turnover;
value added at factor costs;
labour costs.

Employment

*   ***Number of persons employed*** (total employment)

    The number of persons employed is defined as the total number of persons who work in the enquiry unit (inclusive of working proprietors, partners working regularly in the enterprise and unpaid family workers), as well as persons who work outside the unit but belong to it and are paid by it (e.g. commercial representatives, delivery men, repair and maintenance groups). Included are persons absent for a short period (e.g. on sick leave, paid leave or special leave), and also those on strike, but not those absent for an indefinite period.

    Also included are part-time workers who are regarded as such under the laws of the country concerned and who are on the payroll, as well as seasonal workers, apprentices and homeworkers on the payroll.

    Excluded are labour forces made available to the unit by other enterprises and charged for, persons carrying out repair and maintenance work in the enquiry unit on behalf of other enterprises, as well as those doing their compulsory military service.

*   ***Number of employees*** (salaried employment)

    Persons who work for an employer and who have a contract of employment and receive compensation in the form of wages, salaries, fees, gratuities, payments by result or payments in kind. Homeworkers, regardless of whether they are on the payroll or not, are excluded.

Monetary variables

Monetary data such as turnover, value added and labour costs are provided by the countries in the respective national currency.

All data published are converted into ecu by Eurostat using the yearly average official exchange rate. Data are in current prices; no deflator has been applied.

● **Turnover**

The turnover is the totals invoiced by the unit during the reference period, that is all market sales of goods or services supplied to third parties. Turnover includes all duties and taxes on the goods or services invoiced by the unit with the exception of value added tax (VAT) invoiced by the unit *vis-à-vis* its customers. It includes all other charges (transport, packaging, etc.) ascribed to the customer even if these charges are listed separately in the invoice. Reduction in prices, rebates and discounts, as well as the value of returned packaging must be deducted, but not cash discounts.

Turnover does not include sales of fixed assets. Operating subsidies received from public authorities or the EU are also excluded.

In the banking sector, turnover is defined as interest and commission received. In the insurance sector it corresponds to gross premiums written and pension contributions.

● **Value added at factor costs**

Value added is a measure of output used to assess the wealth generation by an enterprise through its transformation of bought-in goods and services into marketable goods and services. It is measured in the following way.

Value added at factor cost =
production value
- intermediate consumption
- consumption taxes invoiced by the enterprise to its customers
- other indirect taxes
+ operating subsidies.

Production value =
total turnover
- cost of products for resale (excluding deductible VAT)
+ change in stocks of finished and half-finished products
+ change in stocks of goods for resale
+ value of investment goods manufactured or constructed by own personnel for own use.

Intermediate consumption =
purchases of raw materials, intermediate goods, etc. (excluding deductible VAT)
+ cost of industrial services received (excluding deductible VAT)
- changes in stocks or raw and ancillary materials, intermediate goods and energy (excluding deductible VAT)
+ cost of non-industrial services received (excluding deductible VAT).

- *Labour costs*

They measure the expenditure borne by employers in order to employ workers. They can be subdivided into direct and indirect costs.

Direct labour costs

- direct remuneration (including bonuses and gratuities paid regularly);
- other bonuses and gratuities;
- payments for days not worked;
- payments in kind.

Indirect labour costs

- statutory social security contributions and family allowances paid by the employer;
- non-statutory payments made by employers under agreement or contracts or voluntarily;
- vocational training costs;
- other social expenditure.

7. Treatment of confidential data

Due to confidentiality, most countries do not allow the disclosure of data concerning a very small number of enterprises, usually two or three units. As in the SME tabular data the enterprise population is broken down by two dimensions, size and economic sector, there are, especially in the smaller countries, quite a lot of cells which have to be hidden.

Countries apply their own rules for confidentiality treatment. Some of them do not consider any data as confidential, for most of them confidentiality only applies to the economic variables, and there are exceptions where even the number of units has to be treated as confidential.

Two different methods of confidentiality treatment are applied.

- Some countries hide confidential data at the national level and do not provide such data to Eurostat.

- The majority of the countries, on the other hand, provide confidential data which has to be hidden by Eurostat. This allows a better harmonization of results because the same aggregates can be calculated for each country.

8. Estimations

For the purpose of having data for the whole of the EU, Eurostat has carried out a set of estimations by country. These estimations were necessary because the original data sets provided by the Member States were not always complete or fully harmonized.

The estimated data sets for EUR 12 (1990) and EUR 15 (1992) contain three variables which are broken down by the detailed size classes and by NACE 70, two digit level:

- number of enterprises
- total employment
- turnover.

Depending on the country, the following main estimation steps were necessary:

- estimation of missing economic sectors
- estimation of a missing variable
- estimation of a missing size class (especially the class with no salaried employment)
- splitting aggregated sectors
- splitting aggregated size classes
- harmonization of the statistical units.

The user should bear in mind that estimations have been made by simple methods using national accounts data to fill in the major gaps.

The figures for EUR 15 are not necessarily equivalent to the sum of the country raw data published here. Compared to the EUR 12 estimations for 1990, estimation procedures for EUR 15 have changed considerably, and comparisons on a disaggregated level should therefore be avoided.

9. Data comparability

Comparability between years

Countries sometimes change their method of data collection: a new source may be added, the underlying register may be improved, the analytical unit may change, or results of a periodic census may be available one year but not the next one. Such factors severely reduce comparability.

Comparability with other sources

To compile the SME tabular data, countries generally use or combine sources. These may be different from the sources used for the national accounts data.

The SME data cover all non-agricultural market activities which means that an important part of the economy and of employment is excluded.

Differences in the underlying sources and in the sectoral coverage lead to diverging values when comparing the SME data with other statistical sources.

B. Methodology by Country

Belgium - I.N.S.

Sources

The Institut National de Statistique (INS) provided data using the VAT register. The VAT register covers all units liable for value added tax. Statistics are produced monthly, quarterly and annually.

Reference year(s)

1992.

Reference unit

The **VAT-declaring unit**: all units regularly engaged in activities providing goods or services that are subject to VAT payment. All units active for any part of the year are included.

No threshold for VAT payment in terms of minimum turnover exists. However, some VAT units can be exonerated from periodical VAT statements:

• small retailers in food, footwear, clothing, books, newspapers, etc. when their purchases do not exceed BFR 4.5 million (food sector) and BFR 2.5 million (the others); they can pay instead what is called the 'taxe d'égalisation';

• some other sectors can also skip VAT payment: farmers, second-hand dealers, waste collectors, artists, etc.

Only the first category has a relevant turnover.

This reference unit has an important overlap with the employer which is the reference unit used by Organisation Nationale de la Sécurité Sociale (ONSS). All employers who are included in the Social Security Register, being active in any of the sectors liable for VAT, are also counted as a VAT-declaring unit.

Size class 0

Size class 0 is included in the VAT register. The coverage is not 100% because some small units are exempt from periodical VAT statements.

For 1992, a descriptive register ('fichier signalétique') with all the units liable for VAT payment (but not necessarily actually paying) was used to count the number of units in size class 0.

Variables

1. *Number of units:* the number of VAT units.

2 a. *Total employment:* only for size class zero.

2 b. *Salaried employment:* not available. See: Belgium - O.N.S.S.

3. *Turnover:* as declared to the VAT register.

4. *Value added:* not available.

5. *Labour costs:* not available.

Classification

Economic sector: the nomenclature used is the VAT-NACE, which corresponds at three digits exactly to NACE 70. The main activity is determined by the largest turnover.

Sectoral coverage: some sectors which are not liable to VAT payment are missing.

*   certain services such as notaries, lawyers, hospitals, resting houses, medical services and some paramedical services;

*   social and cultural institutions such as libraries, museums, education centres;

*   some banking and financial activities, certain counselling activities;

*   public administration.

Employment size class: expressed in terms of salaried employment.

Size classes 6 (200 to 249 employees) and 7 (250 to 499 employees) are joined together in one size class (200 to 499 employees).

Monetary unit

Data supplied in Belgian francs (BFR), the exchange rate being:

ECU 1 = BFR 41.5932 (1992).

Confidentiality

Data considered confidential by INS have been treated by Eurostat in accordance with INS specifications to prevent disclosure (three or fewer units).

Belgium - O.N.S.S.

Sources

The Organisation Nationale de la Sécurité Sociale (ONSS) provided data based on their own registration of employers. Employers' declarations were used to compose annual statistics. Every employer with one or more employees is obliged to contribute to the social security system and is consequently represented in the register.

Reference year(s)

1990, 1991 and 1992.

Reference unit

The **employer**: every employer who engaged one or more workers in the second trimester, who were fully or partially submitted to the social security system, is included.

This reference unit has an important overlap with the VAT-declaring unit of the Belgian Institut National de Statistique (INS). All employers included in the Social Security Register, active in any of the sectors liable for VAT are also counted as VAT-declaring units. All sectors are covered in the Social Security Register, but public administration was excluded for this exercise.

Size class 0

Not available.

Variables

1.   *Number of units:* the number of employers.

2 a.   *Total employment:* not available. See: Belgium - I.N.S.

2 b.   *Salaried employment:* number of full and part-time employees in the Social Security Register at the end of the second trimester. Workers with more than one job working for different employers are counted more than once.

3.   *Turnover:* not available.

4.   *Value added:* not available.

5.   *Labour costs:* data only available for 1992. Data supplied in millions of Belgian francs and converted into thousands of ecu.

Classification

Economic sector: all sectors from NACE 70 divisions 1 to 9 are included.

A firm's main activity is determined primarily on the basis of turnover. If these figures were not available, the number of workers engaged in a certain activity was taken.

Sectoral coverage: all employers regardless of the legal status. Data available for both private and public sectors, that are kept separately.

Employment size class: expressed in terms of salaried employment.

Monetary unit

Data supplied in Belgian francs (BFR), the exchange rates being:

ECU 1 = BFR 42.4252 (1990)

ECU 1 = BFR 42.2232 (1991)

ECU 1 = BFR 41.5932 (1992).

Confidentiality

No data were considered confidential; all data are published.

Denmark

Sources

Danmarks Statistik provided data, the result of a special ad hoc exercise, integrating four official sources.

1. The Business Register: based on two units – legal and local – using the Directorate of Tax and Customs registers.

2. The VAT Register: based on the legal administrative VAT unit. This accounts for all units in all sectors except those exempt from VAT payment – personal transport, banking, insurance and financial services, doctors and dentists and education. The threshold for registration is ECU 15 000.

3. The register-based business employment statistics: using the employer's annual salary declarations submitted to the Tax Directorate. From 1990 onwards, this source was supplemented with information concerning the self-employed.

4. The register-based account statistics: these aim for an automatic procedure to select those enterprises whose tax declarations require special attention. The reporting unit is the legal, taxable enterprise. The register represents approximately 50% of total turnover realized by all Danish enterprises.

Reference year(s)

1990, 1991 and 1992.

Reference unit

The **legal unit**: a unit close, though not identical, to the enterprise, which resulted from combining the four sources above. It comprises all units existing at the end of November of the reference year with a turnover threshold of ECU 15 000 per year. All market sectors liable to VAT are included – industry, construction, services. However, in order to produce a completely compatible set for the five variables, some sectors were excluded from the tables before transmission to Eurostat.

Size class 0

This represents active legal units without employees and an annual turnover of at least ECU 15 000.

Variables

1. *Number of units:* number of legal units.

2 a. *Total employment:* number of persons employed, both full and part-time and the self-employed, at the end of November of the reference year.

2 b. *Salaried employment:* not available.

3. *Turnover:* does not cover activities which are not liable for VAT payment.

4. *Value added at factor cost:* The Danish definition of value added at factor costs is slightly different from that of Eurostat. Some items are under-estimated, others are overestimated, the differences are thus compensated resulting in negligible overall differences.

5. *Labour costs.*

Classification

Economic sector: the main activity is the one with the largest turnover and is expressed in NACE 70 divisions 1-9 at the three digit level.

Employment size class: expressed in terms of salaried persons (full-time equivalent) at the end of November of the reference year.

Monetary unit

Data supplied in Danish kroner (DKR), the exchange rates being:

ECU 1 = DKR 7.85645 (1990)

ECU 1 = DKR 7.90852 (1991)

ECU 1 = DKR 7.80925 (1992).

Confidentiality

Data considered confidential were eliminated by Danmarks Statistik before transmission to Eurostat.

Remark

Since employment figures are expressed in terms of persons employed (heads), while intervals of employment size classes are in terms of full-time employees, there are cases in which average employment exceeds the limits of the interval.

Germany

Sources

The Institut für Mittelstandsforschung (IFM), Bonn, provided data based on the following official sources, submitted to extensive estimates:

1. For the number of enterprises.

* VAT Statistics 1990 and 1992 (Federal Statistical Office).

* Census of workplaces 1987 (Federal Statistical Office);

* Statistics of the Social Security 1992 (Federal Labour Office);

* Exception: for NACE 37 and NACE 5 the Statistics of the Manufacturing Industries has been partly used.

2. For persons employed.

* Census of workplaces 1987 (Federal Statistical Office);

* Statistics of the Social Security (Federal Labour Office);

* Exception: for NACE 37 and NACE 5 there has been used partly the Statistics of the Manufacturing Industries.

3. For the value added and turnover.

* VAT Statistics 1990 and 1992 (Federal Statistical Office)

All tables (i.e. enterprises, employment, turnover and value added) refer to the same population.

Reference year(s)

1992.

Reference unit

The **enterprise**: the smallest unit that has to keep books for commercial and/or tax law reasons and has a commercial/profit orientation.

Size class 0

It represents the self-employed, that is units staffed only by self-employed and non-salaried persons. Size class 0 is estimated for the year 1992. The number of enterprises in this size class is estimated on the basis of the number of enterprises in 1987 with a turnover below DM 20 000 per year.

Variables

1. *Number of units:* number of enterprises.

2 a. *Total employment:* salaried and non-salaried, full-time and part-time employed persons. In a partnership the owner is included. In a corporation the owner is included only if employed officially.
The reference date is 30th June 1992.

2 b. *Salaried employment:* not available.

3. *Turnover:* enterprises with a turnover below DM 25 000 are excluded.

4. *Value added:* enterprises with a turnover below DM 25 000 are excluded..

5. *Labour costs:* not available.

Classification

Economic sector: the firm's activity with the highest value in the tax declaration. It is expressed in NACE 70 divisions 1-9 at the two digit level.

Sectoral coverage: all non-agricultural market sectors are covered except the following NACE codes: NACE 15 (nuclear fuels industry), 71 (railways), 79 (communications), 95 (medical and other health services; veterinary services).

Employment size class: expressed in the total number of employees.

Size classes 6 (200-249) and 7 (250-499) are joined together in one size class: 200-499.

Monetary unit

Data supplied in thousands of Deutschmark (DM), the exchange rate being:

ECU 1 = DM 2.02031 (1992).

Confidentiality

No data were considered as confidential.

Territorial coverage

West Germany and West Berlin before reunification.

Greece

Sources

The National Statistical Service of Greece (NSSG) provided data using the results of the Annual Survey of Industry. The sampled population comes from the Census of Enterprises and Establishments, and the sample is renewed every five years, the latest update being in 1989.

The annual survey is based on a sample of units employing 10 to 19 persons per year and is exhaustive for units with an average annual employment of 20 persons or more. Units with less than 10 employees are not included.

Reference year(s)

1992.

Reference unit

The **establishment**: unit in which one or more persons are engaged permanently, under single management control, or in one of the following – extraction, production, repair, assembly, sale of goods or any auxiliary activity.

NB: The activity of such a unit is relatively homogeneous and therefore closer to a Local Kind-of-Activity Unit (LKAU) than to a local unit.

Size class 0

Not available. Data were only provided for units with at least 10 employees.

Variables

1.   *Number of units:* annual average number of establishments.

2 a. *Total employment:* average number of persons employed per year.

2 b. *Salaried employment:* average number of salaried persons employed per year.

3.   *Turnover:* annual total based on the accountancy year, which coincides with the calendar year or, if not, ends in the reference year. Taxes, such as consumption tax and stamp duty, are excluded.

4.   *Value added at factor cost:* annual total of value added (accountancy year: see turnover).

5.   *Labour costs:* not available.

Classification

Economic sector: data have been provided in the NACE 70 nomenclature up to the three digit level. A firm's most important activity is determined by a combination of employment, turnover and profit figures.

Sectoral coverage: NACE sectors 1 to 4 and 67 up to the three digit level. Construction and services are therefore not covered.

Employment size class: based on salaried employment, excluding units with less than 10 employees, is formed as follows:

10-19, 20-49, 50-99, 100-199, 200-249, 250-499, 500 +.

The average number of persons employed (salaried plus non-salaried) determines to which size class an establishment belongs. A unit is assigned to a size class at the time of the five-yearly census and, regardless of changes in its average employment, remains in that size class until the next census (see: Remark).

Monetary unit

Data supplied in drachmas (DR) and converted into ECU, the exchange rates being:

ECU 1 = DR 201.412 (1990)

ECU 1 = DR 225.215 (1991)

ECU 1 = DR 247.026 (1992).

Confidentiality

Confidential data – regarding the data broken down by size classes – transmitted by NSSG, have been treated by Eurostat to prevent their disclosure in this publication.

Remark

The sampling basis is updated once every five years (latest 1989).

Establishments are allocated to a certain employment size class and remain in it for five years until the next updating during a new census round. An annual survey does not take into account potential shifts of an establishment between strata. New establishments and those that have ceased all activity are integrated in the survey, based on an administrative register.

Spain

Sources

The Instituto Nacional de Estadistica (INE) provided the data, using two sources:

1. 'Directorio Central de empresas del INE' (DIRCE): a central register of enterprises based on administrative sources. This covers all the non-agricultural sectors (except public administration, health and education).

The sources for updating DIRCE are:

- tax on economic activity (Impuesto sobre Actividades Económicas);

- the 'Retenedores a Cuenta' (another fiscal source);

- the Social Security Register (Cuentas de Cótisation de Seguridad Social).

2. Census of Local Units (1990). This covers all local units (and enterprises as a derived unit) engaged in activities carried out in buildings. It is expressed in NACE Rev. 1 and covers the whole country.

Reference year(s)

1990 and 1991.

Reference unit

The **enterprise**: all existing enterprises on 31st December of the reference year, including those recently created.

Size class 0

This represents all enterprises which are active but without employees engaged in the Social Security Register. It is thought that this class may be somewhat overestimated in terms of both numbers and employment because of difficulties and delays in detecting 'deaths' or transitional periods of inactivity.

Variables

1. *Number of units:* number of enterprises.

2 a. *Salaried employment.*

2 b. *Total employment.*

   Data refer to 31st December of the reference year.

3. *Turnover:* not available.

4. *Value added:* not available.

5. *Labour costs:* not available.

Classification

Economic sector: data were supplied in the 'Clasificatión Nacional de Actividades Económicas' (CNAE) at three digits, with three exceptions at two digits (CNAE 71, 73, 81). A conversion to NACE 70 covering NACE divisions 1-9 was carried out by Eurostat resulting in a classification at three digits and in some cases at two digits.

Enterprises active in more than one sector were classed according to the activity with the highest number of employees.

Sectoral coverage: all non-agricultural sectors are included, with the exception of official post offices, libraries, public activities and museums, sports facilities, health and veterinary services.

Employment size class: intervals are expressed by the total number of employees.

Monetary unit

No monetary variables have been provided.

Confidentiality

No data were considered as confidential.

Territorial coverage

The Peninsular territory, the Islands and Ceuta and Melilla.

Remark

Due to the conversion from the CNAE nomenclature to NACE, some NACE codes had to be aggregated.

France

Sources

The Institut National de la Statistique et des Etudes Economiques (INSEE) provided the data. They were extracted from two sources, the Unified System of Business Statistics (SUSE) and the Coordinated Survey of Small Enterprises (ECPE).

1. The 'Système Unifié de Statistiques d'Entreprises' is composed of:

    a. The Fiscal Registers.

    - Bénéfices Non Commerciaux (BNC): non-trade profits, such as those in the professions, health sector, etc.
    - Bénéfices Industriels et Commerciaux (BIC): industrial and trade profits.

    These can be divided into three types:

    (i) the normal one (enterprises with three or more employees);
    (ii) the simplified one (fewer than three);
    (iii) the 'forfait' (enterprises below an annual turnover of FF 700 000, i.e. mainly independent crafts).

    b. Annual Survey of Enterprises. It is exhaustive for enterprises employing a minimum of 20 persons and is based on samples for the smaller enterprises. It is also used to control and complete the fiscal sources.

2. The 'Enquête Coordonnée des Petites Entreprises' (ECPE) was used to estimate total employment on the base of salaried employment. This survey is carried out every five years – the latest one was in 1988, with more recent information therefore not being available yet.

Reference year(s)

1990, 1991 and 1992.

Reference unit

The **enterprise**: all enterprises that have existed for at least six months are registered regardless of their legal form or employment. All sectors are included except certain non-profit making activities.

Size class 0

This represents enterprises with only non-salaried employment and enterprises with a turnover below FF 500 000 regardless of the number of employees.

Variables

1. *Number of units:* number of enterprises.

2 a. *Total employment:* only available for 1990.

2 b. *Salaried employment:* full and part-time salaried employment on 31st December of the reference year.

3. *Turnover:* as defined within the ESAI (Enquête Structurelle sur l'Activité de l'Industrie), expressed in millions of francs.

4. *Value added:* as defined within the ESAI. Changes in stock are not considered for very small units (régime forfait).

5. *Labour costs:* only available for 1992.

Classification

Economic sector: expressed in NACE 70 divisions 1-9 at the three digit level. A firm's sector is determined by the activity with the largest share of employment and, in services, the one with the largest share of turnover.

Employment size class: expressed in terms of salaried employment.

Monetary unit

Data were supplied in French francs (FF), the exchange rates being:

ECU 1 = FF 6.91416 (1990)
ECU 1 = FF 6.97334 (1991)
ECU 1 = FF 6.84839 (1992).

Comparability

It is not possible to compare absolute numbers or totals for any of the four variables for size classes 0 and 1-9, only ratios. For size classes 10 + absolute numbers may be compared cautiously.

Confidentiality

Confidential data were transmitted by INSEE and treated by Eurostat.

Territorial coverage

Metropolitan France and overseas territories.

Italy

Sources

The Istituto Nazionale di Statistica (ISTAT) provided the data. They are deduced from the Census for Industry, Trade, Transport and Services (1991) from which the results are now available.

Reference year(s)

1991.

Reference unit

The **enterprise**.

Size class 0

Size class 0 (self-employed) is not registered as a separate size class but mostly included in class 1-9.

Variables

1.   *Number of units:* the number of enterprises existing on 21st October 1991
2 a. *Total employment:* total number of employees on 21st October 1991.
2 b. *Salaried employment:* salaried employment on 21st October 1991.
3.   *Turnover:* not available.
4.   *Value added:* not available.
5.   *Labour costs:* not available.

Classification

Economic sector: Data have been provided in NACE Rev. 1 classification up to the four digit level. Data have been converted to NACE 70. Where there is more than one activity, the main one is based on the largest share of turnover.

Employment size class: size classes are in terms of total persons employed. The size classes are as follows:

> 1 - 9, 10 - 19, 20 - 99, 100 - 199, 200 - 249, 250 - 499, 500 +.

Monetary unit

Data supplied in Italian lire (LIT), the exchange rate being:
ECU 1 = LIT 1533.23 (1991).

Confidentiality

The data provided had no confidentiality constraints.

Luxembourg

Sources

The Service Central de la Statistique et des Etudes Economiques (STATEC) provided the data, compiled from two sources.

1. The Social Security Register.

2. The VAT register.

Reference year(s)

1992.

Reference unit

The **enterprise**: all units are registered bearing the same identification number and activity code.

All enterprises active in the reference year are considered. New enterprises are included as long as they declare some turnover during the year. There is an optional threshold of LFR 400 000 (from 1990) on the VAT Register; however, even below this threshold, most companies choose to register.

Size class 0

All active enterprises which in October of the reference year do not count with any employee in the Social Security Register. Due to some delays in the accounting of the cessations, this size class might be somewhat inflated.

Variables

1. *Number of units:* number of enterprises.

2 a. *Total employment:* not available

2 b. *Salaried employment:* number of salaried employees.

   The reference date is 30th October 1992.

3. *Turnover:* total sales excluding VAT. Data are supplied in millions of LUF.

4. *Value added:* not available.

5. *Labour costs:* not available.

Luxembourg (continued)

Classification

Economic sector: NACE 70, divisions 1-9 at three digits. Firms active in several sectors are assigned to the one generating the largest value added.

Employment size class: expressed in terms of salaried employment.

The size classes are:

0 (for the units), 1-9, 10-19, 20-99, 100-199, 200-499, 500 +.

Monetary unit

Data provided in Luxembourgish francs, the exchange rate being:

ECU 1 = LFR 41.5932 (1992).

Confidentiality

The number of enterprises and the number of employees are not confidential. Turnover is confidential if it applies to less than four enterprises or if one enterprise accounts for more than 75% of the total turnover. The treatment has been done by Eurostat.

The Netherlands

Sources

The Centraal Bureau voor de Statistiek (CBS) provided the data, mainly from an annual multi-sectoral structural survey whose population and sampling frame is provided by the Central Business Register.

The sectors which are distinguished are the manufacturing industry, public utilities, construction, transport and other services. The activity is expressed in the national nomenclature – SBI – at the three digit level for all sectors (conversion to NACE was carried out subsequently), both for integral and sample statistics. With the exception of manufacturing, sample statistics are sometimes expressed in two digits rather than three.

Reference year(s)

1991 and 1992.

Reference unit

This is the **Kind-of-Activity Unit**. It can be an enterprise, part of an enterprise or a combination of different enterprises which predominantly engages in one kind of economic activity, and is not restricted to a geographical area. The KAU coincides in the majority of the cases with the enterprise.

Reference population: the KAUs, including takeovers and newly-created enterprises, that receive the annual enquiry form. In November of the reference year they must have been part of the monthly observation of that year or observed before that date.

Thresholds: industry and construction – a minimum of one employee. The whole economy – at least one person regularly working a minimum of 15 hours per week (i.e. approximately half a man-year).

Size class 0

This size class is registered for the service sector. The unit is included if it is active for at least 15 hours per week on a regular basis. There is no size class 0 for industry and construction.

Three sources were used for demographic data on size class 0: the Business Register, the Survey on Services and the Annual Census of Firms.

The Netherlands (continued)

Variables

1. *The number of units:* the number of Kind-of-Activity Units.

2 a. *Total employment:* includes all full-time, part-time and temporary employees as well as working proprietors, family workers, managers, shareholders, etc.

2 b. *Salaried employment:* it includes wage-earners working 15 or more hours per week and insured under the 'Sickness Insurance Act', homeworkers under the same insurance and wage and salary earners aged 65 and over.

3. *Turnover:*

*   Industry and construction:

    sector's turnover = industrial sales + trade activities + other activities

*   Services:

    turnover = total revenues (excluding : interest yield, operating subsidies, reduction of provisions, VAT, value of own production and exceptional receipts)

4. *Value added:*

    gross value added = production value - intermediary consumption.

5. *Labour costs:* the total cost of employees.

Classification

Economic sector: data were provided in NACE 70 divisions 1-9 at the three digit level. The main activity was assigned on the basis of the highest value added. For industry, construction and transport, in practice, it comes to an assignment by the largest number of employees.

Sectoral coverage.

*   Industry and construction: all sectors are covered, whether public or private, and regardless of legal status.

*   Services: only private sectors are included. In NACE 70, the following are excluded: 63, 643, 663, 666, 672-5, 81, 82, 831-4, 850, 91, 93, 94, 95, 96, 97 (but 972 is included), 983, 984, 99.

The Netherlands (continued)

Employment size class: For industry and construction: number of paid man-days per year registered by the professional federations. These data are converted into man-years. For the transport sector: the size is based on the number of paid employees at the end of the year. Other services: average size.

The size classes are:

0, 1-2, 0-4, 5-9, 10-19, 20-49, 50-99, 100-199, 200-499, 500 +.

Monetary unit

Data supplied in Dutch guilders (HFL), the exchange rates being:

ECU 1 = HFL 2.31093 (1991)

ECU 1 = HFL 2.27482 (1992).

Confidentiality

In some cases, size classes were aggregated by the CBS in order to maintain confidentiality. This results in non-harmonized size classes. NACE classifications at the three digit level were sometimes aggregated to the one and two digit level for confidentiality reasons.

Remark

1. Size coverage differs between industry/construction and services; industry/construction cover enterprises with at least one employee, services also include size class 0. Due to some methodological differences in the collection of data, care should be taken when comparing the different sectors.

2. The employment registration differs between industry – part of construction and the remaining part of construction – (salaried employment) and services (total employment).

3. Reference date: the total is at the end of September for all sectors except for the transport sector where the total is at the end of the year.

Austria

Sources

The Österreichisches Statistisches Zentralamt provided the SME data compiled from the 1991 and 1992 annual industrial surveys.

In Austria, establishments carrying out economic activities defined by ISIC Rev. 3 codes 10 to 45 are differentiated by their classification into the sections 'Industrie' and 'Gewerbe' of the Federal Chamber of Commerce. The section 'Gewerbe' itself is divided into 'Großgewerbe' (20 employees or more) and 'Klein- und Dienstleistungsgewerbe' (0 to 19 employees).

The data collection is exhaustive for all establishments belonging to the section 'Industrie' of the Austrian Federal Economic Chamber and for the establishments with 20 or more employees in the 'Gewerbe' section. Establishments in the 'Klein- und Dienstleistungsgewerbe' class (less than 20 employees) are surveyed by sampling.

Reference year(s)

1991 and 1992.

Reference unit

The **establishment**: a local self-contained unit capable of reporting production costs. No inclusion thresholds or limitations regarding certain legal forms exists.

Size class 0

The size class 0 is included: establishments with no salaried employment belonging to the 'Industrie' section. Enterprises without salaried employment belonging to the 'Gewerbe' section are not included.

Variables

1.   *Number of units:* the number of establishments.

2 a.   *Total employment:* salaried and unsalaried employment (self-employed, family workers), both full and part-time.

2 b.   *Salaried employment:* wage and salary earners, both full and part-time.

3.   *Turnover:* total sales.

4.   *Gross value added:* gross production value minus intermediate consumption.

5.   *Labour costs:* all costs related to the salaried employment, that is wages and salaries, contributions of the employer to social security, and other related costs.

Classification

Economic sector: expressed in ISIC Rev. 3 at the three digit level. The ISIC codes are the result of a reclassification according to the Austrian 'Betriebssystematik 1968'; this reclassification has been done by the Austrian Central Statistical Office. The conversion to NACE 70 has been done by Eurostat. The main activity is determined by the activity which accounts for the largest part of net production value.

Sectoral coverage: the manufacturing industry and construction sectors, from ISIC Rev. 3, codes 10 to 45. No services are included.

(See also: Sources).

Employment size classes: expressed in terms of number of salaried employees.

A common size class breakdown for both the exhaustive and the sampled data is not possible (see Sources). Data for 1991 and 1992 therefore include only the results of the exhaustive survey, so size classes 0, 1-9 and 10-19 are not complete.

Monetary unit

Economic variables are expressed in thousands of Austrian schillings (ÖS), the exchange rates being:

ECU 1 = ÖS 14.4309 (1991)

ECU 1 = ÖS 14.2169 (1992).

Confidentiality

Data considered confidential (i.e. information on four establishments or less) have been eliminated from the set of data transmitted to Eurostat.

Portugal

Sources

The Instituto Nacional de Estatistica (INE) provided data compiled from the Register of Statistical Units (FUE). This is updated from several sources.

1. The INE Annual Survey on Enterprises.

2. Employment tables from the Employment Ministry. These only cover the medium-sized and large enterprises with at least 100 employees. The rest are only partially covered.

3. The National Register of Enterprises from the Ministry of Justice. All new units have to register to start an activity.

4. The Tax Register on Business Income from the Finance Ministry.

5. A special survey for updating the FUE.

6. The Register of Company Accounts.

Reference year(s)

1991 and 1992.

Reference unit

The **enterprise**: active for at least one year. The reference date is 31st December of the reference year.

Size class 0

Most units in this group are enterprises without employees (i.e. a real class 0). Enterprises for which there is no employment information are also included. However, this number is negligible.

Variables

1. *Number of units:* number of enterprises active for at least one year.

2 a. *Total employment:* the number of (full and part-time) active persons expressed in number of heads. The national accounts registers employment in full-time equivalents, thus resulting in a lower employment figure.

2 b. *Salaried employment:* the number of salaried employees.

3. *Turnover.*

4. *Value added at market prices.*

5. *Labour costs.*

Classification

Economic sector: is expressed in NACE 70 at the three digit level. Firms with more activities are classified according to the largest share of turnover.

Sectoral coverage: industry, construction and services, NACE 1-9.

Employment size class: expressed in terms of salaried employment.

Monetary unit

Data supplied in ECU according to the following exchange rates:

ECU 1 = ESC 181.107 (1990)

ECU 1 = ESC 178.614 (1991)

ECU 1 = ESC 174.614 (1992).

Confidentiality

Confidential data have been treated by Eurostat according to INE criteria (data relating to one or two units).

Remark

For the four years before 1992 the number of enterprises has been estimated using the 1992 data. This was hampered by insufficient information on cessations from 1988-92. Consequently, having only information on active and created enterprises, the total number of enterprises in the past is underestimated.

Territorial coverage

Continental Portugal plus autonomous regions.

Finland

Sources

Statistics Finland provided data from the Enterprise and Establishment Register, which is based on the tax authorities database.

All new businesses receive an enquiry form from the Enterprise and Establishment Register. At the start or closure of activities, enterprises must be registered with the tax authorities either as an employer or as a personal enterprise liable for turnover tax. This updating is carried out quarterly.

The business register is updated by quarterly enquiries into all new enterprises, annual enquiries into all enterprises with two or more establishments and additional enquiries into all enterprises which have become active in new localities.

Reference year(s)

1992.

Reference unit

This is the **enterprise**, whether made up of single units or physical persons. The register contains information on all enterprises with paid employees, plus corporate bodies and self-employed people liable to turnover taxes.

Enterprises have to fulfil two criteria. They must be active for at least six months in the reference year and either:

* employ at least 0.5 persons in the reference year (i.e. the full-time equivalent of half a year), or

* have a turnover in the reference year 1992 of at least MKF 46 000 (approximately ECU 9 000).

Size class 0

This represents enterprises liable to turnover taxes with less than 0.5 full-time employees. The coverage is restricted by the following conditions:

* a turnover above MKF 46 000 (1992);

* the sector/the enterprise should be liable to turnover tax payments.

NB: the positive labour costs for size class 0 is due to the definition of this class: between zero and 0.5 full-time employment equivalents. This is the same for salaried employment.

Variables

1. *Number of units:* number of enterprises.

2 a. *Total employment:* persons employed, both full and part-time, converted into man-years.

2 b. *Salaried employment:* the number of salaried persons, both full-time and part-time, converted into man-years.

3. *Turnover:* as reported in the business financial statement. For NACE 81, 82 and 83 turnover is not available.

4. *Value added:* not available.

5. *Labour costs:* wages and salaries for salaried employment.

Classification

Economic sector: following the Standard Industrial Classification 1988 (SIC Finland) at the three digit level. The conversion into NACE 70 divisions 1 to 9 has been carried out by Eurostat. Where there is more than one activity, the one with highest share of value added is included.

Sectoral coverage: the sectors which are excluded are extraction of petroleum and natural gas (NACE 13), letting of real estate by the owner (NACE 15).

Employment size class: in terms of number of salaried persons.

Monetary unit

Data were given in Finnish marks (MKF), the exchange rate being:

ECU 1 = MKF 5.80703 (1992).

Confidentiality

Confidential data have been hidden by Eurostat; monetary variables on less than three enterprises are considered confidential.

Sweden

Sources

Statistics Sweden provided the data based on the Financial Accounts Survey (EFA), compiled from the financial accounts of enterprises. The sampling frame of EFA is the Central Register of Enterprises and Establishments (CFAR), which is updated every fortnight.

The enquiry is exhaustive for enterprises with at least 50 employees. Data for size class 0-49 are either collected exhaustively or sampled from external sources, namely the National Board of Patents and Registration of Trademarks (PRV).

For the extraction of the variables, two different sources are used:

- for the number of enterprises and the number of employees: the business register;
- for turnover, value added and labour costs: the VAT register.

Reference year(s)

1992.

Reference unit

The **enterprise**: the legal or balance sheet unit, i.e. the smallest unit for which most balance sheets and profit and loss data can be obtained. The threshold for turnover is SKR 100 000. All units in existence in October of the reference year – a snapshot – are included. Enterprise creations after this date are included only if they have more than 50 employees.

Size class 0

This represents enterprises without salaried employment. In principle they are all considered as no threshold is applied. Estimates for size class 0 of the number of enterprises, turnover and value added were carried out by Statistics Sweden.

Size class 0 (and even size classes 1-9 and 10-19) is sometimes difficult to identify because:

- it is often based on very small samples and a very detailed kind-of-activity code;

- in Sweden there are companies operating within a group of companies where the staff is employed by one enterprise but is actually working in another. Therefore, there are enterprises with very large incomes but no, or very few, employees.

Thus, size classes 0, 1-9 and 10-19 are not published separately but in an aggregated band: 0-19.

Variables

1. *Number of units:* number of enterprises.
2 a. *Total employment:* not available.
2 b. *Salaried employment:* average number of employees during the financial year. Both full and part-time employees are included, calculated in man-years. Non-salaried employment is not included.
3. *Turnover:* total turnover in gross accounts including rents and fees. It excludes VAT and other indirect taxes, internal deliveries, returns and discounts.
4. *Value added:* adjusted gross operating margin which corresponds in the national accounts to the gross national product at factor cost adjusted by price charges (i.e. capital gains or loss) on stocks.
5. *Labour costs.*

Government grants (e.g. to theatres or recreational services) are included in the value added figure, but not in the turnover figure.

Classification

Economic sector: expressed in the Swedish nomenclature SNI (equivalent to ISIC 1968 at four digits) at the four digit level with some aggregated codes. The conversion into NACE 70 was carried out by Eurostat, attaining three digits of the NACE, but ISIC codes which were already aggregated in SNI led to the aggregation of NACEs at three and two digits. The activity of the enterprise accounting for the largest share of employment is considered the main activity.

Sectoral coverage: all non-financial enterprises in the corporate sector. Both private and public enterprises are included, regardless of their legal status or size. Certain ISIC codes are absent from the tables:

*   some activities only exist at the local unit level, not at the enterprise level as main activities: ISIC 2902 (chemical and fertilizer mineral mining), 2903 (salt mining) and 7115 (pipeline transport);

*   the financial sector and related activities – not covered by the EFA survey (ISIC 8101, 935, 939, 953 and 96);

*   Real Estate activities (ISIC 831) are not covered either.

Employment size class: expressed in average number of salaried employees: 0 - 19, 20 - 99, 100 - 199, 200-499, 500 +.

Users should be aware that employment figures only refer to enterprises with at least one employee (size class 1-19), while the rest include class 0.

Monetary unit

Data were supplied in thousands of Swedish kroner (SKR). Conversion into ECU was carried out by Eurostat, the exchange rate being:

ECU 1 = SKR 7.53295 (1992).

Confidentiality

Data considered confidential transmitted by Statistics Sweden have been treated by Eurostat to prevent their disclosure.

United Kingdom

Sources

Consultants Graham Bannock & Partners provided the data, having submitted official statistics to extensive estimates.

Sources were:

1. The VAT register from the Department of Trade and Industry (DTI) is the main source. This covers all VAT registered units, excluding central and local government but including public corporations, sole proprietorships, partnerships and incorporated businesses. The optional threshold for VAT registration was UKL 25 400 for the financial year 1990-91 and UKL 35 000 for 1991-92. Below this threshold estimates were carried out by Graham Bannock.

2. The *Annual Census of Production* (ACOP) which covers all establishments with 100 + employees and is based on samples for the rest. The census is used firstly to estimate average employment (in order to transform the turnover size classes into employment size classes) and then to add second digit data (class level) to divisions 2-4.

3. The *Census of Employment* which records all active units in April plus those newly created. The reference unit is the employer; size class 0 is excluded.

4. The annual *Labour Force Survey* which provides for quarterly estimates of the self-employed, at the one digit level of the NACE.

5. The Inland Revenue who supply information on the average number of partners per partnership; this is necessary for class 0 estimates.

6. Others, such as The Times 1000 database on the largest UK firms, the Leyshon Report and employment figures from the Central Statistical Office (CSO). These data identify the private sector and public corporations from central and local government.

Reference year(s)

1990 and 1991.

Reference unit

An **estimate derived from the VAT unit,** but not always directly comparable with the enterprise: a snapshot of the register (VAT units) plus the self-employed not registered for VAT.

Size class 0

This represents the self-employed with no employees whether registered for VAT or not (estimated).

Size class 0 cannot be given separately from size class 1-9, hence 0-9.

Variables

1. *Number of units:* enterprises as derived from the VAT unit.

2 a. *Total employment:* average number of persons employed during the year.

2 b. *Salaried employment:* not available.

3. *Turnover:* totals invoiced by the enterprise during the reference year. Available for industry and construction (NACE 1- 5) but not for services.

4. *Value added:* not available.

5. *Labour costs:* not available.

Classification

Economic sector: expressed at the one or two digit level (NACE 70), result of the conversion from the VAT codes and SIC nomenclature: divisions 1, 2-4 (at two digit level), 5, 6, 7 (class 71/79, rest of 7), 8 (class 83, class 85, rest of 8), 9 (class 91-96, class 97-99). The main activity is determined by the assessment of the entrepreneur himself.

Sectoral coverage: all market services, both private and public, are covered with the exception of NACE class 85 (letting of real estate by the owner).

Employment size class: expressed in total number of employees. Size class 0 and 1 are taken together, resulting in the following size class: 0-9 employees.

The following size breakdown is available:

0-9, 10-19, 20-49, 50-99, 100-199, 200-249, 250-499, 500 +.

Monetary unit

Data supplied in pounds sterling, the exchange rates being:

ECU 1 = UKL 0.71386 (1990)
ECU 1 = UKL 0.70101 (1991).

Confidentiality

No data are confidential.

Territorial coverage

The United Kingdom (Great Britain and Northern Ireland).

Iceland

Sources

The Statistical Bureau of Iceland provided data using the tax register, which contains the tax returns of all enterprises, including their annual accounts as well as special wages/employment returns. The data transmitted to Eurostat are based on wage and salary statements from the tax authorities.

Reference year(s)

1992.

Reference unit

The **enterprise**.

Size class 0

There are two size classes 0: **0** and **0A**.

Size class 0: individual firms without employees (no wage payments other than the owner's wages), i.e. self-employed with no employees.

Size class 0A: sole proprietorships with registered activity of owner (not necessarily amounting to one man-year) and one or more employees (who may work less than full-time, or only part of the year). The owner is counted as one of the persons employed only when there are working weeks registered in his name on the salary statement.

Variables

1. *Number of units:* number of enterprises.

2 a. *Total employment:* number of persons employed measured in man-years (52 working weeks per year). Employers are to report annually the number of weeks worked by the employees, or by the self-employed themselves. It covers both the full-time and part-time positions according to the length of their employment.

2 b. *Salaried employment:* not available.

3. *Turnover:* not available.

4. *Value added:* not available.

5. *Labour costs:* defined as wages and salaries plus calculated owner's income. As employment refers to the number of man-years it may include labour costs for size class zero.

Classification

Economic sector: expressed in ISIC at the five digit level. A conversion into NACE 70 divisions 1 to 9 at the two digit level was carried out by Eurostat.

When the enterprise is involved in more than one activity, the main activity is determined by the highest sum of wages and salaries. In the wage and salary statements, employers report the number of weeks worked by employees in any branch of activity. If the same person is also employed in another activity, he/she will be credited with extra weeks for that one.

Sectoral coverage: the ISIC (Rev. 2). There is no information based on the enterprise for the following: whale processing (30312), manufacture of tobacco products (31220), construction and repair of harbours and lighthouses not contracted away (50432). Otherwise all non-agricultural sectors are covered.

Employment size class: expressed in rounded off man-years of total employment.

The size classes are:

    0:    self-employed persons without employees

    0A: self-employed persons with one or more employees

    and 1-9, 10-19, 20-49, 50-99, 100 +.

Confidentiality

No data transmitted by the Statistical Bureau of Iceland are confidential.

Remark

Additional comment concerning the employment figures (full and part-time workers), counted as man-years: a labour force sample survey carried out in Iceland showed that 71% of workers work full-time (90% of the active men and 48% of the active women are in this category). The rest are employed part-time (23%), have occasional work (2%) or have unspecified work-hours (4%).

Liechtenstein

Sources

The Economic Office provided the data for Liechtenstein. The source is the Census of Enterprises conducted by the 'Bundesamt für Statistik' (Swiss Federal Statistical Office) in co-operation with the Economic Office.

Reference year(s)

1991.

Reference unit

This is the **local unit**. The reference date is 30th September 1991.

Size class 0

Not available.

Variables

1.   *Number of units:* active local units.
2 a. *Total employment:* not available.
2 b. *Salaried employment:* number of employees working in the local unit on 30th September 1991.
3.   *Turnover:* not available.
4.   *Value added:* not available.
5.   *Labour costs:* not available.

Classification

Economic sector: as declared by the **enterprise,** expressed in the **Swiss** activity nomenclature (ASWZ 85) at the four digit level. Conversion to NACE was carried out by Eurostat at the two digit level.
Sectoral coverage: all industries and services.

Employment size class: expressed in terms of salaried employment. Assigned according to the number of employees on 30th September of the reference year. Size class 0 is not available and size classes 6 and 7 are joined resulting in the following size class: 200-499 employees.

Available size classes are:

1-9, 10-19, 20-49, 50-99, 100-199, 200-499, 500 +.

Monetary unit

ECU 1 = SFR 1.77245 (1991).

Norway

Sources

The Norwegian Central Bureau of Statistics provided the data using the Central Register of Establishments and Enterprises. This register is based on and updated with surveys and information held in two other registers: the Value Added Tax Register and the Register of Employers.

1. Surveys: most enterprises are subject to some kind of survey each year. Apart from surveys, e.g. on industrial statistics, simple questionnaires are sent to the rest of the registered units.

2. The Value Added Tax Register, in the Directorate of Taxes, lists all new activities subject to registration. With few exceptions, VAT registration is compulsory for all enterprises with establishments which have an expected annual turnover of at least NKR 12 000.

3. The Register of Employers, in the National Insurance Office, contains information on all new employers who have to register their activity.

Reference year(s)

1991 and 1992.

Reference unit

The **enterprise.** The legal institutional unit comprising all the economic activities engaged in by one and the same owner, and covering one or more productive units (establishments).

No minimum period of activity is required for an enterprise to register, in that respect they are all included in the population. However, a minimum annual turnover of NKR 12 000 (approximately ECU 1 500) is required.

Size class 0

This covers enterprises where the equivalent of less than half a man-year is worked. Class 0 is not fully covered for all activities. Enterprises with less than one man-year are not included in the table for employment.

Norway (continued)

Variables

1. *The number of units:* the number of enterprises registered at the end of the reference year.

2 a. *Total employment:* number of salaried and non-salaried persons employed, calculated in man-years.

2 b. *Salaried employment:* not available.

3. *Turnover:* total annual amount of Gross Production Value.

4. *Value added:* not available.

5. *Labour costs:* not available.

No information on total employment and turnover is provided for the following activities: electricity, gas and water supply, railway transport, provision of tramway and subway transport, harbour administration, lighterage and pilotage services, air transport, communications, financial institutions, public administration and defence, sewage disposal, social and related community services, recreation and cultural services.

Classification

Economic sector: data were supplied in the ISIC Rev. 2 nomenclature up to the five digit level. Conversion to the NACE 70 nomenclature was carried out by Eurostat, obtaining a three and sometimes two digit level of classification. If there is more than one activity, the criterion used to assign the main one is the number of man-years.

Sectoral coverage: all non-agricultural market sectors are included except ocean transport vessels, letting of residential and other buildings, some cultural services, domestic services, international and other extra-territorial bodies.

Employment size class: an additional size class option is included, namely 'not reported'. This category mainly contains enterprises registered as new in the reference year, for which no information is available on employment/turnover.

Monetary unit

Data were supplied in thousands of Norwegian kroner (NKR), the exchange rate being:

ECU 1 = NKR 8.01701 (1991)
ECU 1 = NKR 8.04177 (1992).

Confidentiality

Confidential figures have been blanked out by the Norwegian Bureau of Statistics.

Switzerland

Sources

The Bundesamt für Statistik (Swiss Federal Statistical Office) provided the data for Switzerland.

Its source was the *Census of Enterprises*. This census is conducted every 10 years in years ending with a 5 (1985, 1995, etc.). Intermediate censuses are carried out two or three times during the 10-year period. These include less variables and are exhaustive for units with at least one employee and sampled for class 0. The census used here is an intermediate one, dated 30th September 1991. As the census is mandatory the response rate is almost 100%.

Reference year(s)

1991.

Reference unit

The **enterprise**.

Size class 0

Size class 0 represents all enterprises without employees. This size class was estimated from a sample and is joined with size class 1 (1-9 employees) resulting in size class 0-9.

Variables

1. *Number of units:* all active enterprises on 30th September 1991.

2 a. *Total employment:* not available.

2 b. *Salaried employment:* number of employees working in the enterprise on 30th September 1991.

3. *Turnover:* not available.

4. *Value added:* not available.

5. *Labour costs:* not available.

..

Switzerland (continued)

Classification

Economic sector: as declared by the enterprise, expressed in the Swiss activity nomenclature (ASWZ 85) at the four digit level. Conversion to the NACE divisions 1 to 9 was carried out by Eurostat attaining the three digit level.

Sectoral coverage: all industries and services.

Employment size class: expressed in terms of salaried employment. Assigned according to the number of employees on 30th September of the reference year.

Employment size classes 0, 1 and 6, 7 are joined resulting in the following size classes:

0-9, 10-19, 20-49, 50-99, 100-199, 200-499, 500 +.

Monetary unit

Data supplied in Swiss francs, the exchange rate being:

ECU 1 = SFR 1.77245 (1991).

Remark

As a result of the conversion from the ASWZ nomenclature to NACE 70, some sectors may not always match precisely the NACE three/two digit levels.

Installation procedure for CUB.X and the SME database
(CUB.X version 1.0 – SME diskette version 1.0)

1. About CUB.X

1.1. CUB.X is the property of the European Commission. It has been produced by Eurostat. This software may be used by you free of charge, but cannot be resold. It may not be used for commercial purposes or distributed without prior authorisation by Eurostat.

1.2. For more information, questions, remarks and to receive the official release please contact:

EUROSTAT DATA SHOP
2 rue J. Engling
L-1466 Dommeldange
Tel (+352) 43 35 22 51
Fax (+352) 43 35 22 221

For questions concerning CUB.X. please contact Mr. Tayenne,
Tel (+352) 4301 37267.

2. Content of the diskettes

One diskette containing the CUB.X software and one containing the SME database.

3. What you need to run CUB.X for SME data

- PC 486 (CUB.X for Win runs on a PC 386 but slowly ...)

- WINDOWS version 3.1 (or later)

- At least 15 Mbytes on your hard disk

N.B. CUB.X has not been tested on OS/2

4. How to install CUB.X and the SME database

4.1. Installation of CUB.X on the hard disk.

- load WINDOWS, start the file manager

- insert the CUB.X diskette in the disk drive

- go to disk a: (or whichever drive you placed the CUB.X diskette in)

- run the SETUP.EXE file.

You will be guided through the installation procedure. During the procedure files will be decompressed.

4.2 Installation of the SME database on the hard disk.

- load WINDOWS, start the file manager

- insert the SME disk in the disk drive

- go to disk a: (or whichever drive you placed the SME diskette in)

- run the SETUP.EXE file.

You will be guided through the installation procedure. During the procedure files will be decompressed.

5. Running SME

To run SME database simply double click on the New Cronos Products group icon in the WINDOWS screen. Then double click on the SME /Eurostat icon. You are then in the SME database using the CUB.X Browser.

Enterprises in Europe, Fourth Report

◊ What is the role of Small and Medium-sized Enterprises in the EU?
◊ How many people are employed in EU enterprises and how much turnover do they generate?
◊ In which countries are enterprises particularly big or small?

The importance of SME's for growth and employment has been stressed for years. SME's represent over 90% of all enterprises in the EU. Therefore a European policy for SME's needs reliable and comparable enterprise data for all Member States.

The publication *Enterprises in Europe, Fourth Report* provides not only the general methodology of the SME data, but also a set of specific analyses on different subjects:

◊ a general overview of enterprises in the EU;
◊ specific analyses such as the creation of new firms, the innovation behaviour and labour costs of SME's, regional analyses, etc.;
◊ a snapshot of the business structure of 16 European countries and 27 economic sectors.

This package contains:

One diskette containing the CUB.X software and one diskette containing SME data. Also included is a CUB.X quick reference manual and table notes by country.

Export formats:

Table: DIF,EXCEL / TAB,CSV. Stream: DBASE,TEXT,CSV.

What you need to run Enterprises in Europe data with CUB.X

- PC 486 (CUB.X for Win runs on a PC 386 but slowly ...),
- WINDOWS version 3.1 (or later),
- At least 15 Mbytes on your hard disk.

Please note - CUB.X has not been tested with OS/2

Price (excluding VAT) in Luxembourg
ECU 50