

# PC-Axis file format manual version 2006-5

## Technical description of the file format and application guidelines for Statistics Finland's metadata standard

<b>1. Foreword</b> .....	<b>2</b>
<b>2. Application guidelines at Statistics Finland</b> .....	<b>3</b>
<b>3. About the terminology used</b> .....	<b>4</b>
<b>4. About dot codes</b> .....	<b>5</b>
<b>5. About keywords</b> .....	<b>6</b>
<b>Variables</b> .....	6
<b>Classifications and additivity</b> .....	6
<b>Character set</b> .....	6
<b>Basic language</b> .....	6
<b>Topic information</b> .....	7
<b>Titling</b> .....	7
<b>Server management</b> .....	7
<b>Small table</b> .....	7
<b>Decimal numbers</b> .....	7
<b>Measurement units</b> .....	7
<b>Creation and updating</b> .....	7
<b>Footnotes</b> .....	7
<b>Source and contact information</b> .....	7
<b>Data</b> .....	7
<b>Content variables</b> .....	7
<b>Time series</b> .....	7
<b>Economic time series</b> .....	7
<b>Keywords to be introduced in near future</b> .....	7
<b>6. Table examples</b> .....	<b>9</b>
6.1 Only mandatory keywords.....	9
6.3 Mandatory and optional keywords.....	11
6.4 Sparse matrix saving with the KEYS keyword.....	12
6.5 Definition of content variables with the CONTVARIABLE keyword.....	13
<b>7. Keywords in alphabetical order</b> .....	<b>14</b>
<b>8. Typical sources of error</b> .....	<b>36</b>
<b>Appendices</b> .....	<b>37</b>
Appendix 1. Keywords in alphabetical order.....	37
Appendix 2. Keywords in presentation order .....	39
Appendix 3. About grouping files .....	41
Appendix 4. Statistics Finland's classification of statistical topics, abbreviations of statistics and naming of files .....	42
Appendix 5. About multi-language metadata .....	43
Appendix 6. References to Statistics Finland's other guidelines and standards .....	45
Appendix 7. References .....	45

## 1. Foreword

The PC-Axis file is the saving format of a multi-dimensional statistical table (cube). The extension of the file format is .px (e.g. *table.px*). The file is a structured document where so-called keywords are used as initial identifiers of content sections. The file is formed of two main blocks: metadata and data. Some of the keywords used for definition of metadata are mandatory and some optional, the data part is mandatory and **always** the last one in the file.

When all the mandatory keywords of the meta part and their contents and the data are included and technically correctly recorded, the file can be used in all main applications of the PC-Axis product family: PC-Axis, PX-Web, PX-Map and PX-Edit (in what follows these will be referred to as applications). PX-Edit, which is also a correction and revision tool for files, can be used to open files with deficient contents or slightly erroneous syntax, in which case PX-Edit indicates the errors it identifies; for some typical errors it can also suggest corrections. At its simplest, standardisation of a PC-Axis file is done so that the file is opened in PX-Edit and it is saved, when the order and saving format of keywords are standardised. Opening and saving can also be made as a PX-Edit batch function.

**This extensive version of the document is primarily intended as guidelines and a manual for designers and programmers of applications producing PC-Axis files.** The document is meant to be not only a comprehensive description of the PC-Axis file format and its application at Statistics Finland but also a description of those specified features of the file format which are not generally in use at Statistics Finland. The document is not complete because it was thought that there was no need to record some outmoded features and rare special cases that were considered unnecessary. Available is also a shorter version of this document describing only those features currently in general use at Statistics Finland.

The document will be updated and specified whenever necessary.

### 6.1 Acknowledgements

The following workgroup has taken part in making this document possible:

Riia *Arvela*, Hans *Baumgartner*, Veli-Matti *Jantunen*, Mia *Kilpiö* (translation), Jaakko *Laakso* and Tuula *Taivainen*.

Each of us can be reached via email at: *firstname.surname@stat.fi*.

## 2. Application guidelines at Statistics Finland

To improve the quality of metadata and usability of statistics **an internal standard has been specified at Statistics Finland** as application guidelines for keywords, which is more demanding and extensive than the mandatory amount of metadata required for mere technical functionality.

The guidelines lay considerable weight on the metadata needed in the control of product processes and archiving and cataloguing of tables. For this purpose, some new keywords related to the creation and updating of tables have been added to the file specification by Statistics Finland. Even if the present applications did not quite yet meet them in all respects, they should be fulfilled in line with the guidelines, so that archive search functions, for example, can be made to function well in future and cataloguing and classification of table content can be automated. In addition to data management, user interfaces of services will be increasingly based on metadata in the following. Modern web distribution systems, such as PX-Web, create service user interfaces automatically on the basis of metadata. Each content producer must bear in mind that poor or deficient metadata will be visible for customers as a poor user interface of the web service as well.

Each keyword is dealt with one at a time. For the sake of technical functionality, mandatory keywords are indicated in the keyword titles with the text *mandatory* and those mandatory at Statistics Finland with the text *mandatory (SF)*. Some keywords are recommended for use and then they are marked as *recommended (SF)*. Keywords given with the text *in special cases (SF)* may be important or useful only in certain services. Their mode of use should be discussed with the information services that standardise services and IT services that maintain the servers. To simplify and lighten the processes the general **recommendation** is that only those features be used that are marked as *mandatory* or *mandatory (SF)*.

No separate application guidelines have been compiled for time series keywords. They are mostly processing rules guiding applications that can be utilised especially when the px file format is the common source of various transfer formats, such as GESMES and Aremos, and the PC-Axis application is used as a transformer to produce these transfer modes.

A summary of the main information on all keywords is compiled in the table (**Virhe. Kirjanmerkkiä ei ole määritetty.**), where the PC-Axis column shows the mandatory specifications for technical functionality and the SF standard column the quality requirement of metadata demanded of all Statistics Finland's database services.

Separate instructions have been issued on naming of files, building of topic directories of the database service, application of standard classifications and PC-Axis groupings based on them, which are available on Statistics Finland's Intranet (StatFin2006-intra).

The technical correctness of files can be checked with PX-Edit.

### 3. About the terminology used

A **cube** refers to a multi-dimensional table (array) created as a result of cross-tabulation of several variables. Different dimensions of the statistical cube (classifications) are called **variables** in PC-Axis. Variables are described in PC-Axis with keywords `STUB` (row variables) `HEADING` (column variables). It is also possible to create a variable consisting of various content variables; the keyword `CONTVARIABLE` is used for this.

The variable is divided into **values**, which are expressed as texts and possible codes. The values of variables (classes, headings) are indicated by variable with the `VALUES` and codes with the `CODES` keywords.

No leading or trailing spaces must be used in variable names and value or code texts, and dots should be used in place of spaces especially when showing the hierarchy. Use of some special characters should also be avoided (tabulator, quotation marks, parentheses and brackets and semi-colons), because applications may interpret such notations undesiredly. The number sign (#) is used in PC-Axis to show applications the text line feeds, so it must not be used for any other purpose.

The numerical values of a statistical table are given in the numerical table following the `DATA` keyword. The decimal separator of the figures is a full stop, the minus sign is in front of a negative figure and thousand separators are not used. The figures are separated from one another with space or tabulator. The end of each data row must have a space before the line feed. In addition to numerical values, missing, masked or other hidden information can be expressed in PC-Axis data part by means of so-called **dot codes**.

The grouping and summing of variables in PC-Axis tables can be managed with separate **grouping files**. A variable in a file may be connected to a shared grouping file by means of a separate value group keyword. Further information is available in the description of the keyword `DOMAIN` and in [Appendix 3](#).

## 4. About dot codes

**Dot codes** describe missing, masked or other such data that cannot be presented. Dot codes are limited by quotation marks, and a code can have from one to six dots (" . " - " . . . . . ").

**NB** PC-Axis applications do not control for which purpose dot codes are used.

In addition to dot codes, a **dash code** is used as a special notation: "-". The dash code is interpreted in old application versions as zero and saved as zero. Then it cannot be known afterwards whether the original file had the value "-" or 0. Therefore the dash code should not be used for any other purpose, and a four-dot dot code should be used for 'magnitude nil' ". . . . ". The newest application versions treat a dash code as so-called sharp null (exact zero) and save it back to a dash code.

***Statistics Finland's application guideline:** The following notations in accordance with the Official Statistics of Finland standard accepted in 2005 are in use:*

0 . 0	Magnitude less than half of unit employed (not working properly in PC-Axis)
" . "	<b>Category not applicable</b>
" . . "	<b>Data not available or too uncertain for presentation</b>
" . . . "	<b>Data subject to secrecy</b>
" . . . . "	<b>Magnitude nil (do not use a dash!)</b>
" . . . . . "	
" . . . . . . "	
" - "	Magnitude nil (do not use in PC-Axis because it is interpreted as zero).

When using **mask codes** the code specifications should always be included in the table footnote. The used masking rule or the reason for masking must also be given for masked data, for example: data masked, because less than 3 observations.

## 5. About keywords

The PC-Axis file is built of keyword specifications, which consist of the keyword part, equality sign, keyword value and **semi-colon**, which always ends the specification. Because of their mutual dependency, keywords must be in a certain **order** shown in [Appendix 2](#). The last keyword of the file is always DATA. Own, non-standard keywords can be specified but it is recommended **not** to use them!

Keywords are either table, variable, value or cell-specific or special keywords. The keyword part has a keyword given in uppercase, followed by a possible extension in parentheses. In multi-language tables keywords may be followed by a language code in brackets before the extension (see [Appendix 5](#)).

- a table-specific keyword has the form KEYWORD=
- a variable-specific keyword has the form KEYWORD( "Variable" )=
- a value-specific keyword has the form KEYWORD( "Variable", "Value" )=
- the form of cell and special keywords is given in the specifications of each keyword (e.g. CELLNOTE).

The value of the keyword is given in quotation marks, such as KEYWORD="text". Quotation marks are not used for numerical or YES/NO type keywords. The content of some keywords (e.g. footnotes) may be shown in multiple lines in applications, then the # character is used as the line feed character.

**Semi-colons, quotation marks or other special characters (e.g. tabulator) should not be used in the value texts of keywords.**

If the keyword value is split into several lines, the line to be split is ended in a quotation mark (before line feed) and the next line is started with a quotation mark. If possible, word space should be selected as the splitting point. Lists (e.g. VALUES) are split so that the comma separating the elements of the list remains as the last character of the list to be split.

**NB** For historical reasons, treatment of some keywords differs from the general rules (mandatoriness, maximum lengths of value texts, etc.) and the use of some has clearly been outdated.

### Variables

Variables are described directly as row and column variables, so the combined content of the keywords STUB (row variables) and HEADING (column variables) is a list of all variables in the table. At the same time the **arrangement** (pivot) of the variables in the table default view is shown and it is determined in which order the numerical cells of the table should be located in the data part.

### Classifications and additivity

The classifications (variable values) are defined with the keywords VALUES and CODES. The **additivity** of variables is controlled with the keywords AGGREGALLOWED and ELIMINATION. The files related to the grouping of variables are defined with the keyword DOMAIN.

The map template for the area variable is defined with the keyword MAP.

Obsolete keywords DOUBLECOLUMN, PRETEXT and PARTITIONED are used for showing variable values and codes and for presenting area sub-hierarchies. The presentation of dot codes, dot code sum and dash code can be defined with the keywords DATASYMBOL1 - DATASYMBOL6, DATASYMBOLSUM and DATASYMBOLNIL.

### Character set

The keywords CHARSET and CODEPAGE are used for managing the character set.

### Basic language

The basic language is indicated with the keyword LANGUAGE and the languages contained in a multi-lingual table with the keyword LANGUAGES. PC-Axis compatibility is indicated with the AXIS-VERSION keyword.

## Topic information

The information on statistical topics is managed with the keywords SUBJECT-AREA, SUBJECT-CODE, DATABASE and MATRIX.

## Titling

Table titling is controlled with the keywords TITLE, DESCRIPTION, DESCRIPTIONDEFAULT and CONTENTS.

## Server management

The keywords PX-SERVER and DIRECTORY-PATH are suitable for server management.

## Small table

A small table can be made open direct in PC-Axis with the AUTOPEN keyword.

## Decimal numbers

Saving of decimal numbers and presentation of decimals is directed with three keywords. The number of saved decimals (DECIMALS) can be higher than the precision shown in the applications (SHOWDECIMALS). Part of the values can also be set a presentation precision differing from the general setting (PRECISION).

## Measurement units

Measurement units are given with the keyword UNITS.

## Creation and updating

Information about the creation and updating of the table is indicated with the keywords CREATION-DATE and LAST-UPDATED, updating times can also be given with the keywords NEXT-UPDATE and UPDATE-FREQUENCY.

In px files timestamps are marked as CCYYMMDD hh:mm.

## Footnotes

The keywords NOTE, NOTEX, VALUENOTE, VALUENOTEX, CELLNOTE, CELLNOTEX, DATANOTECELL and DATANOTESUM are used for management of different level footnote information in the table.

## Source and contact information

Source and contact information can be defined with the keywords SOURCE and CONTACT.

## Data

The data part is given with the keyword DATA. The data part can be presented in so-called sparse matrix format using the KEYS keyword. The rare keyword CONFIDENTIAL is used for management of the data mask rule, which is **not** recommended for use.

## Content variables

Handling of content variables is solved in PC-Axis with the keyword CONTVARIABLE. The keyword is used to define the variable that classifies the content of the table and whose values are content variables. The range of use for certain keywords changes at the same time from table into variable-based, whereby only the name of the content variable is written in the variable specification of the keyword (see example 6.5).

## Time series

The time series is defined with the keyword TIMEVAL.

## Economic time series

A set of special keywords are used for economic time series: BASEPERIOD, CFPRICES, DAYADJ, REFPERIOD, SEASADJ and STOCKFA. If the content variable ( CONTVARIABLE ) is defined, the keywords for economic time series are repeated for each content variable value.

## Keywords to be introduced in near future

The PC-Axis format will include a group of keywords which had not yet been tested when this document was completed, so their function may differ from the instructions given here. They include:

DATANOTECELL, DATANOTESUM, DATASYMBOL1, DATASYMBOL2, DATASYMBOL3, DATASYMBOL4, DATASYMBOL5, DATASYMBOL6, DATASYMBOLNIL, DATASYMBOLSUM, DIRECTORY-PATH, INFO, NEXT-UPDATE, PX-SERVER and UPDATE-FREQUENCY.

There is so far only little information about the following coming keywords, so guidelines will be issued on them later on.

DEFAULT-GRAPH,	to express the default graph type to be attached to the table in PX-Web;
LINK,	to attach an external web link to the table;
SURVEY,	to indicate the method using in acquiring basic information for the table; and
VARIABLE-TYPE,	meant to be used for indicating the type of variable (e.g. for time series and area variables).

In addition, there will be a keyword for individualising the table during its life cycle (TABLE-ID?) and a set of keywords used for expressing and processing the variable hierarchy.

This document will be supplemented when more detailed information is available on the keywords in question.



## 6. Table examples

### 6.1 Only mandatory keywords

To attain a technically functioning statistical cube, only 13 mandatory keywords are needed. Such a px file functions technically in applications, but its metadata are fairly scarce from the viewpoint of users. It is suitable as a preform whose metadata can be supplemented by means of PX-Edit, for example.

```
CHARSET="ANSI";
AXIS-VERSION="2000";
DECIMALS=0;
MATRIX="vaerak";
SUBJECT-CODE="VRM";
SUBJECT-AREA="Väestö";
TITLE="Väestö 31.12. muuttujina Sukupuoli, Kunta, Vuosi ja Siviilisääty";
CONTENTS="Väestö 31.12.";
UNITS="Henkilöä";
STUB="Sukupuoli","Kunta";
HEADING="Vuosi","Siviilisääty";
VALUES("Sukupuoli")="Sukupuoli yhteensä","Miehet","Naiset";
VALUES("Kunta")="Espoo","Helsinki","Vantaa";
VALUES("Vuosi")="2001","2002";
VALUES("Siviilisääty")="Siviilisääty yhteensä","Naimaton","Naimisissa";
DATA=
57516 43030 100546 57516 43030 100546
144564 85339 229903 144564 85339 229903
47131 33688 80819 47131 33688 80819
53821 43375 97196 53821 43375 97196
151536 86184 237720 151536 86184 237720
44071 33935 78006 44071 33935 78006
111337 86405 197742 111337 86405 197742
296100 171523 467623 296100 171523 467623
91202 67623 158825 91202 67623 158825;
```

## 6.2 All keywords required for Statistics Finland's databases

This table includes all the keywords mandatory in Statistics Finland's databases and some recommended keywords, such as NOTE. Such a file may contain more information for reading, and it can also be more easily further processed with a computer, since all classifications have classification codes, for instance.

```

CHARSET="ANSI";
AXIS-VERSION="2000";
LANGUAGE="fi";
CREATION-DATE="20001212 11:00";
DECIMALS=0;
SHOWDECIMALS=0;
MATRIX="vaerak";
COPYRIGHT=YES;
SUBJECT-CODE="VRM";
SUBJECT-AREA="Väestö";
DESCRIPTION="Väestö 31.12.2003";
TITLE="Väestö 31.12.2003 muuttujina Sukupuoli, Kunta ja Siviilisääty";
CONTENTS="Väestö 31.12.2003";
UNITS="Henkilöä";
STUB="Sukupuoli","Kunta";
HEADING="Siviilisääty";
VALUES("Sukupuoli")="Sukupuoli yhteensä","Miehet","Naiset";
VALUES("Kunta")="Espoo","Helsinki","Vantaa";
VALUES("Siviilisääty")="Siviilisääty yhteensä","Naimaton","Naimisissa";
CODES("Sukupuoli")="S","1","2";
CODES("Kunta")="049","091","092";
CODES("Siviilisääty")="S","1","2";
LAST-UPDATED="20040319 09:00";
CONTACT="Tilastokeskus, Väestötilastopalvelu#Postiosoite:"
"Väestötilastopalvelu, #PL 4A, 00022 Tilastokeskus#Puhelin: (09) 1734 3590"
"#Faksi: (09) 1734 #3251#Yhteyshenkilö: Nicola Brun#Sähköposti:"
"#vaestotilasto.palvelu@tilastokeskus.fi#<A HREF=http://tilastokeskus.fi/ "
"TARGET=_blank>Linkki tilastokeskuksen kotisivulle</A>";
SOURCE="Tilastokeskus";
INFOFILE="vaerak";
NOTE="Vuodesta 2002 lähtien on siviilisäätytietoihin lisätty rekisteröidyt"
"parisuhteet...";
DATA=
197742 111337 86405
467623 296100 171523
158825 91202 67623
100546 57516 43030
229903 144564 85339
80819 47131 33688
97196 53821 43375
237720 151536 86184
78006 44071 33935;

```

### 6.3 Mandatory and optional keywords

In the following example the file includes aggregation rules and processing rules of already calculated sums (ELIMINATION):

- if the user does not choose anything from the gender variable, the application calculates the sum of genders and shows it
- if the user does not choose anything from the area variable, the application shows the already calculated sum for the whole country.

The table also has an example of a variable-specific footnote (explanation of marital status) and of a value-specific footnote, which in this case is directed to all data on the area ESPOO.

```
CHARSET="ANSI";
AXIS-VERSION="2000";
LANGUAGE="fi";
CREATION-DATE="19930401 12:10";
DECIMALS=0;
MATRIX="vaerak";
COPYRIGHT=YES;
SUBJECT-CODE="VRM";
SUBJECT-AREA="Väestö";
DESCRIPTION="Väestö 1990-1992";
TITLE="Väestö muuttujina siviilisääty, sukupuoli, alue ja aika";
CONTENTS="Väestö";
UNITS="henkilöiden lukumäärä";
STUB="siviilisääty", "sukupuoli";
HEADING="alue", "aika";
VALUES("siviilisääty")="naimisissa", "naimaton", "eronnut", "leski";
VALUES("sukupuoli")="mies", "nainen";
VALUES("alue")="Koko maa", "Helsinki", "Espoo";
VALUES("aika")="1990", "1991", "1992";
TIMEVAL("aika")=TLIST(A1), "1990", "1991", "1992";
CODES("siviilisääty")="1", "2", "3", "4";
CODES("sukupuoli")="1", "2";
CODES("alue")="SSS", "091", "049";
ELIMINATION("alue")="Koko maa";
ELIMINATION("sukupuoli")=YES;
LAST-UPDATED="19950209 13:00";
SEASADJ=NO;
CONTACT="Tilastokeskus#PC-Axis-koulutus#fax 09 1734 1234#s-posti pcaxis@stat.fi";
DATABASE="TKDB";
SOURCE="Tilastokeskus";
REFPERIOD="Viiteajankohta on 31. joulukuuta joka vuosi";
INFOFILE="vaerak";
NOTE("siviilisääty")="Siviilisääty on riippuvainen väestölaskennan "
"rekistereistä. Naimisissa olevat yhdessä asuvat merkitään naimisissa "
"oleviksi. Muut yhdessä asuvat kuuluvat naimattomiin";
VALUENOTE("alue", "Espoo")="Tähän selitystä Espoon datasta#ja selitys "
"jaetaan näytössä usealle riville.#Tästä alkaakin jo kolmas selitysrivi";
DATA=
    (data are not listed here)
```

## 6.4 Sparse matrix saving with the KEYS keyword

A PC-Axis table can be saved in so-called sparse matrix format, where pure zero lines are not saved in order to save space. Sparse matrix saving is controlled with the keyword **KEYS**, which must be given to each row variable. The value of the keyword is either **VALUES** or **CODES**, and the key recorded in the rows of the data part corresponding to the variable is either the corresponding variable value text or code value with apostrophes.

```
CHARSET="ANSI";
AXIS-VERSION="2000";
LANGUAGE="fi";
DECIMALS=0;
SHOWDECIMALS=0;
MATRIX="tyti";
COPYRIGHT=YES;
SUBJECT-CODE="TYM";
SUBJECT-AREA="Työmarkkinat";
DESCRIPTION="T01E Väestö 31.12.2000";
TITLE="T01E Väestö 31.12.2000 muuttujina Osa-alue,"
"Ikä ja Pääasiallinen toiminta/ammattiasema.";
CONTENTS="T01E Väestö 31.12.2000";
UNITS="Henkilö";
STUB="Osa-alue","Ikä";
HEADING="Pääasiallinen toiminta/ammattiasema";
VALUES("Osa-alue")="049 Espoo","078 Hanko-Hangö","091 Helsinki",
"092 Vantaa","106 Hyvinkää";
VALUES("Ikä")="0-6","7-14","30-34";
VALUES("Pääasiallinen toiminta/ammattiasema")="Työvoima",
"Työlliset","Palkansaajat";
CODES("Osa-alue")="049","078","091","092","106";
DOMAIN("Osa-alue")="OSAL_01 2002";
INFOFILE="tyti";
KEYS("Osa-alue")=CODES;
KEYS("Ikä")=VALUES;
DATA=
"049","30-34",16064 15324 14660
"078","30-34",560 490 470
"091","30-34",42100 39027 37423
"092","30-34",13215 12358 11806
"106","30-34",2646 2423 2270;
```

## 6.5 Definition of content variables with the CONTVARIABLE keyword

The content classifying variable, whose values are content variables, is indicated with the CONTVARIABLE keyword. The range of use for certain keywords at the same time changed from table to **variable**-based, so a value must be given separately for each content variable value; these keywords are: BASEPERIOD, CFPRICES, CONTACT, DAYADJ, LAST-UPDATED, REFFERIOD, SEASADJ, STOCKFA and UNITS.

```

CHARSET="ANSI";
AXIS-VERSION="2000";
LANGUAGE="sv";
CREATION-DATE="20030425 18:21";
DECIMALS=0;
SHOWDECIMALS=0;
MATRIX="BE0101F1";
SUBJECT-CODE="BE";
SUBJECT-AREA="Befolkning";
DESCRIPTION="Flyttningar 2000";
TITLE="Flyttningar med region, ålder, kön, tid och typ";
CONTENTS="Flyttningar";
UNITS="antal";
STUB="region", "ålder", "kön";
HEADING="tid", "typ";
CONTVARIABLE="typ";
VALUES("region")="00 Riket";
VALUES("ålder")="2";
VALUES("kön")="män", "kvinnor";
VALUES("tid")="2000";
VALUES("typ")="Inflyttade", "Flyttningsöverskott", "in3";
TIMEVAL("tid")=TLIST(A1), "2000";
CODES("region")="00";
CODES("ålder")="2";
CODES("kön")="1", "2";
CODES("typ")="typ1", "typ2", "in3";
DOMAIN("ålder")="Ålder";
LAST-UPDATED("Inflyttade")="20030212 16:49";
LAST-UPDATED("Flyttningsöverskott")="20030212 16:49";
LAST-UPDATED("in3")="20030212 16:49";
UNITS("Inflyttade")="antal";
UNITS("Flyttningsöverskott")="antal";
UNITS("in3")="ton";
CONTACT("Inflyttade")="Ewa Eriksson, SCB#Tel: 019-17 67 43";
REFPERIOD("in3")="2000";
DATABASE="Sveriges Statistiska Databaser";
SOURCE="Statistiska centralbyrån (SCB)";
INFOFILE="BE0101";
NOTE="Det här är också bra att veta";
DATA=
0 0 0
0 0 0;

```

## 7. Keywords in alphabetical order

**AGGREGALLOWED** - in special cases (SF)

Metadata type: processing rule

If AGGREGALLOWED=NO, the user cannot aggregate or add values to the table in PC-Axis product family applications. If the keyword specification is missing or gets the value YES, aggregation and addition is allowed.

***Statistics Finland's application guideline:*** *The table content producer is responsible for the settings of aggregation prevention. Because aggregation prevention does not move to other file formats or applications when transferring, the table footnote should indicate non-aggregated data included in the table.*

*For example, if the table has empty cells on the municipality level because of masking and the ready counted sums for regions and the whole country, the setting prevents the user from calculating the region sums (because the result derived in this way would be lower than the correct result).*

**AUTOPEN** - not in use (SF)

Metadata type: processing rule

With the keyword value YES the table can be ordered to open in full in PC-Axis, and the table selection window is not shown. For the time being the keyword does not function in the PX-Web service.

AUTOPEN=YES ;

***Statistics Finland's application guideline:*** *Not used. The user instructions will be updated if PX-Web starts to support the keyword.*

**AXIS-VERSION** - mandatory

Metadata type: document metadata/processing rule

Maximum length 80 characters.

The PC-Axis version number that reports to the applications the PC-Axis compatibility level of the file.

AXIS-VERSION=" 2000 " ;

***Statistics Finland's application guideline:*** *In all currently produced tables the used setting is AXIS-VERSION=" 2000 " (similarly as CHARSET=" ANSI "). With the exception of PX-Web, the applications are able to read files of older versions as well. Before old PC-Axis files are imported to the present services they should always be converted, checked and supplemented with PX-Edit, for example.*

**BASEPERIOD**

Metadata type: statistical metadata (economic time series keyword)

The keyword specifies the base time period of the time series for indexed series, for instance. Shown in a footnote.

If the content variable is defined, the keyword is repeated for each value of the content variable.

```
BASEPERIOD="2001=100";
```

```
CONTVARIABLE="Export";
```

```
BASEPERIOD("Export")="2000";
```

***Statistics Finland's application guideline:*** *There is no guideline because Statistics Finland does not so far use PC-Axis production databases where metadata corresponding to time series keywords would be supported.*

**CELLNOTE** - in special cases (SF)

Metadata type: statistical metadata

A footnote (additional note) for an individual cell or group of cells. The cell note is given with the variable value and name. If the value is \*, the note refers to all the values of the variable. It is also possible to give only one value of the variable. The values are given in the variable order according to the STUB and HEADING keywords.

```
CELLNOTE("*", "*", "Finland", "1995")="footnote for Finland";
```

**CELLNOTEX** - not in use (SF)

Metadata type: statistical metadata

As CELLNOTE, but shown automatically when opening a table (similarly as NOTEX).

***Statistics Finland's application guideline:*** *Not used. In Statistics Finland's PX-Web servers the mandatory message window of the CELLNOTEX text is not activated.*

**CFPRICES**

Metadata type: statistical metadata (economic time series keyword)

The keyword is used to indicate whether the data are at current (C) or fixed (F) prices.

If the content variable is defined, the keyword is repeated for each value of the content variable.

```
CFPRICES="F";
```

```
CONTVARIABLE="Costs";
```

```
CFPRICES("Costs")="C";
```

***Statistics Finland's application guideline:*** *There is no guideline because Statistics Finland does not so far use PC-Axis production databases where metadata corresponding to time series keywords would be supported.*

**CHARSET** - mandatory

Metadata type: document metadata/processing rule

The keyword CHARSET indicates whether the file is saved as a Windows text file. If there is no keyword or its value is other than ANSI, the PC-Axis version 2000 or newer assumes the file to be in DOS format and converts the text automatically into Windows format (this has an effect on the presentation of national characters, for example).

```
CHARSET="ANSI" ;
```

**Statistics Finland's application guideline:** *The text saving format is always CHARSET="ANSI". PX-Web does not make the DOS-ANSI conversion automatically, so it always requires use of the ANSI character set. Before old PC-Axis files are imported to the present services they should be converted, checked and supplemented with PX-Edit, for example.*

**CODEPAGE** - not in use (SF)

Metadata type: processing rule

Maximum length 20 characters

The content of the keyword is used in XML conversions for the specification of the document text format and it will also be used when a new Excel conversion version using this setting is completed.

The default value is ISO-8859-1.

```
CODEPAGE="ISO-8859-1" ;
```

**Statistics Finland's application guideline:** *Will be introduced and instructed if necessary when the applications have a new Excel output version.*

**CODES** - mandatory (SF)

Metadata type: statistical metadata

Maximum length of one code text is 256 characters.

The codeword CODES is used when the variable classes are given codes in addition to headings (VALUES). The codes are in quotation marks and they are separated by commas. The codes have to be **unique** and the code list must be complete. If the list is divided into several lines, line feed is made after the comma separating the codes.

```
CODES("Municipality")="SSS","004","005","006","009","010","015","016",  
"017","018","019","035","040","043","044","045","046","047","049","050";
```

**Statistics Finland's application guideline:** *Official codes are used as classification codes. Uppercase S letters are always entered as the code for total sum according to the code length. For example, when the code has three characters, the sum code is SSS.*

*If it is absolutely necessary to release a variable for whose classes there are no codes, consecutive numbering is used as the code filled with leading zeros to uniform length. It is easy to produce such numbering in PX-Edit.*

**NB** *The PC-Axis file format sets only values (headings) as mandatory. All file internal footnotes, such as a footnote connected to the value, are coded with values. The code cannot then be used in the file for marking a class-specific footnote link or for indicating the ELIMINATION setting. The old PC-Axis custom, which allows empty values in the code list, must **not** be used!*



**CONFIDENTIAL** - not in use (SF)

Metadata type: processing rule

Dynamic masking rule for confidential information, e.g. 1 is replaced by 0 and 2 is replaced by 3. The masking function requires a service mode where the user cannot at any stage touch the original data. In the document-based PX-Web service this means that downloading of an entire table must be prevented.

***Statistics Finland's application guideline:*** Do not use! Data providers supply to the database only PC-Axis files from which all confidential information has already been deleted. The function is not supported or activated in Statistics Finland's PX-Web servers.

**CONTACT** - mandatory (SF)

Metadata type: document metadata

The keyword contains the contact information of the producer. The text can have line feeds, which are separated with the # sign. The content of the keyword is visible in the table selection window of the PX-Web service under *Information* and in the table footnote in the PC-Axis program. If several content variables are defined, the keyword is repeated for each content variable (see CONTVARIABLE).

An example of a well constructed keyword content which has all the mandatory contact information in Statistics Finland's databases and a link to the homepage of the statistics concerned:

```
CONTACT="Statistics Finland#National Accounts#"
Telephone: (09) 1734 3338#"
"e-mail: skt.95@stat.fi#<A HREF=http:// "
"stat.fi/til/vtp/index.html TARGET=_blank> "
"The home page for the statistics</A>"
```

***Statistics Finland's application guideline:*** The contact information should always contain the contact details of the organisation. The contact information should include at least the following: the name of the organisation unit in charge of the content, the general email address of the unit and the telephone number of the on-duty content support during office hours. In charged or restricted services contact details for persons can also be added. The contact information should also include a web link to the homepage of the statistics concerned. When Statistics Finland releases tables of other data producers, a link to the data producers' webpage can be added to the contact information when agreed with the producer in question. The PX-Edit controlled function can be used for making the link.

**CONTENTS** - mandatory

Metadata type: statistical metadata

Maximum length 256 characters

The **basic title** of the table, which the application uses as the prefix of the title in dynamically created titles. A dynamically formed title (TITLE) consists of a basic title (CONTENTS) and variable names chosen according to the selection limitation.

```
STUB="Kunta", "Vuosi";
HEADING="koulutusluokitus";
CONTENTS="Koululaitoksen oppilaat";
DESCRIPTION="Koululaitoksen oppilaat 2003-2005";
TITLE="Koululaitoksen oppilaat muuttujina kunta, vuosi ja "
"koulutusluokitus";
```

**NB** Please note how years are entered.

***Statistics Finland's application guideline:** The basic title of the table that defines the content more accurately than SUBJECT-AREA. Do not include any variable names in the CONTENTS keyword so that they will not appear twice in the keyword TITLE. For example, there may be a misleading title on a map printout if variable names or time series years are saved for the keyword that can be excluded from the table for the map printout.*

*If the table has information for only one year, CONTENTS may contain the year.*

**CONTVARIABLE** - not in use (SF)

Metadata type: statistical metadata

The keyword is used to define the variable classifying the content of the table in which the values of the variable are content variables. The use range of some keywords at the same time changes from table to variable-based, that is, the value should be given separately for each content variable value when using CONTVARIABLE with the following keywords: BASEPERIOD, CFPRICES, CONTACT, DAYADJ, LAST-UPDATED, REFPERIOD, SEASADJ, STOCKFA and UNITS.

**NB** The UNITS keyword must **always** be given table-based as well.

```
CONTVARIABLE="Toimipaikkatiedot";
```

***Statistics Finland's application guideline:** The use of the keyword should be avoided, but in some special cases its use can be considered (e.g. for presentation of different unit information).*

**COPYRIGHT** - mandatory (SF)

Metadata type: document metadata

If the keyword value is YES, copyright refers to the organisation given in the SOURCE keyword. The copyright information is shown in the footnote and in the PX-Web selection window under *Information*.

```
COPYRIGHT=YES;
```

***Statistics Finland's application guideline:** In Statistics Finland's databases the copyright setting must always be COPYRIGHT=YES, so the applications show in a uniform manner the data producer with copyright specified with the SOURCE keyword.*

**CREATION-DATE** - mandatory (SF)

Metadata type: document metadata

The creation time of the file or time of the first release of the database table. The timestamp (date and time) is always written as CCYYMMDD hh:mm.

```
CREATION-DATE="20040119 13:19";
```

**Statistics Finland's application guideline:** In PX-Web database services *CREATION-DATE* indicates the time of the first release of the database table or the time when the file was printed out for the first time from the production database. When the base year of the index changes, a table with the new index is added to the PX-Web service and *CREATION-DATE* is set there at the same time. When new observations are afterwards added to this table, the *LAST-UPDATED* keyword is updated to show the release time of the update.

**DATA** - mandatory

*DATA* specifies the starting point of the data part, it is **always** the last one of the keywords and the file ends in a semi-colon at the end of the data part. All data cells belonging to the table come after the keyword. The separator between the cells is a space or tabulator. A minus sign is in front of a number smaller than zero and full stop is the decimal separator. Other decoration characters (thousand separators, etc.) inside cells must not be used. Data are often written so that the number of data records is determined according to the number of row variable values. The length of data records is then dependent on the number of row variable values and the size of numbers. It is also possible to write all data cells into one unlimited long record, which ends in a semi-colon. Dot codes are written in quotation marks. Ends of data row must have a space before the line feed.

```
DATA=
1877721 133065 20472 36284 76309 7.0865
1917051 138585 24700 "." ". ."
1932752 147442 28835 38966 79641 7.6286;
```

**Statistics Finland's application guideline:** Saving of data to very small data records should be avoided. Some applications read much faster a file where data are in long records or as one record than a file with just one or a few cells per record. Use of a comma and semi-colon as a cell separator is **not** recommended anymore, although applications support them for compatibility reasons.

**DATABASE** - recommended (SF)

Metadata type: document metadata

Name of the source database. Used to indicate the production database from which the table is printed as a px file. If the table data derive from many different databases, their names are separated with the # sign.

```
DATABASE="The name of the production database";
```

**Statistics Finland's application guideline:** It is recommended that the keyword contains the name of the production database or other data source. The information should show from which database or file and from which version the px file is printed.

**DATANOTECELL** - not in use (SF)

Metadata type: statistical metadata

The keyword is used to define a text format additional note linked to the data cell, which intends to highlight when viewing the table in the application that a footnote is attached to the cell in question. The cell specification is given in the same way as with the CELLNOTE keyword. To distinguish the note text from the data value in the same cell, the text should be started with at least one space.

```
DATANOTECELL(" * ", "20", " * ", "Migration", " * ") = " a";
```

Footnotes must still be given a defining explanation in the specification text, e.g. **a=revised** or **\*=preliminary**.

**Statistics Finland's application guideline:** For the time being **not** taken into use in general database services. The use can be considered later for special services which need to attach character format information to the data cell to highlight the footnotes attached to them.

**DATANOTESUM** - not in use (SF)

Metadata type: statistical metadata

The keyword is used to specify the additional note that is shown when aggregating cells marked with the DATANOTECELL keyword. There can be only one attribute. The note given must still be explained in the specification text.

```
DATANOTESUM= " * * * " ;
```

**Statistics Finland's application guideline:** Will **not** be taken into use in general database services. The use can be considered later for special services which need to use the DATANOTECELL keyword.

**DATASYMBOL1** - not in use (SF)**DATASYMBOL2** - not in use (SF)**DATASYMBOL3** - not in use (SF)**DATASYMBOL4** - not in use (SF)**DATASYMBOL5** - not in use (SF)**DATASYMBOL6** - not in use (SF)

Metadata type: processing rule

These keywords can be used to inform applications how dot codes are shown in applications. The keyword does not have any effect on the saving format of the dot code.

```
DATASYMBOL1="epälooginen" ;
```

**Statistics Finland's application guideline:** In Statistics Finland's services only the dot codes of Official Statistics of Finland are used (see [Section 4](#)).

**DATASYMBOLNIL** - not in use (SF)

Metadata type: processing rule

This keyword can be used to inform applications how the dash code is shown in applications. The keyword does not have any effect on the saving format of the code.

```
DATASYMBOLNIL="sharp null";
```

***Statistics Finland's application guideline:*** *The keyword must not be used because the dash code is interpreted as zero in PC-Axis applications.*

**DATASYMBOLSUM** - not in use (SF)

Metadata type: processing rule

```
DATASYMBOLSUM="dot code sum";
```

This keyword can be used to inform applications how the dot code sum is shown. Calculation with dot codes produces a separate dot code (seven-dot code), which is used when the dot code is summed with some other dot code of the same type.

**DAYADJ**

Metadata type: statistical metadata (economic time series keyword)

If the value of the keyword is YES, the time series data are working-day adjusted. This information can be used as a processing rule to prevent adjustment of a series adjusted once for working days. If the content variable is defined, the keyword is repeated for each value of the content variable.

```
DAYADJ=YES;
```

```
CONTVARIABLE="Series";
DAYADJ("sarja")=YES;
```

***Statistics Finland's application guideline:*** *There is no guideline because Statistics Finland does not so far use PC-Axis production databases where metadata corresponding to time series keywords would be supported.*

**DECIMALS** - mandatory (SF)

Metadata type: statistical metadata

The keyword is used to set the saving precision to be used, which can be 0 to 15 decimals. The applications do not use more than six decimals unless the presentation precision has been set separately. The SHOWDECIMALS keyword can be used to set the presentation precision lower than the saving precision. The value of the keyword is written without quotation marks.

DECIMALS=0 ;

**Statistics Finland's application guideline:** *In materials critical in their rounding practices the data are imported to the px file always rounded to their official release precision. For example, it is thus necessary to set in index data only DECIMALS and possibly also SHOWDECIMALS, to which then the same value has been set as DECIMALS.*

*If a table including decimal numbers uses grouping controlling aggregation and the saving precision differs from the presentation precision, the table footnote should mention that the sum level data may contain differences due to rounding. If the table requires matching of sum level roundings, the sum level data must be given ready-made.*

**DESCRIPTION** - mandatory (SF)

Metadata type: statistical metadata

The keyword is used to define **only the title visible in applications**. DESCRIPTION changes into a title visible in tables and applications only in special cases with the DESCRIPTIONDEFAULT=YES setting.

```
STUB="kunta", "vuosi" ;
HEADING="koulutusluokitus" ;
CONTENTS="Koululaitoksen oppilaat" ;
DESCRIPTION="Koululaitoksen oppilaat 2003-2005" ;
TITLE="Koululaitoksen oppilaat muuttujina kunta, vuosi ja "
"koulutusluokitus" ;
```

**NB** Please note how years are entered.

**Statistics Finland's application guideline:** *DESCRIPTION is the recommended way to set the table title in the PX-Web service to get into view in the application interface a title describing better the whole table. Unless DESCRIPTION titling is used, the keyword content must be the same as that of the keyword CONTENTS.*

**DESCRIPTIONDEFAULT** - not recommended (SF)

Metadata type: processing rule

Activation of a fixed title: if the keyword value is YES, the application displays a fixed title set with the keyword DESCRIPTION. Otherwise the title shown is the content of the keyword TITLE (dynamic title).

DESCRIPTIONDEFAULT=YES ;

**Statistics Finland's application guideline:** *Used only in special services when the dynamic title is not suitable.*

**DIRECTORY-PATH** - mandatory in database distributions (SF)

Metadata type: document metadata/processing rule

The content of the keyword shows the file directory path in the server with which it is released. In an archived table the keyword documents the service module and subject directory with which the table is released.

```
DIRECTORY-PATH="altika\asuminen\asuntojen hinnat\vuositilasto";
```

**NB** The directory path can have spaces, but file names do not.

**Statistics Finland's application guideline:** *This is important control information for the updating processing of PX-Web databases. It is used for automatic direction of updates to the correct directory in the distribution server, which is defined with the keyword PX-SERVER. The directory path is indicated according to the Windows standard separated with backslashes.*

**DOMAIN**

Metadata type: statistical metadata/processing rule

Maximum length 80 characters

Defines which value groups the table has and which grouping files are connected to the table, that is, which classifications and their hierarchical aggregation rules can be used. The keyword can appear only once for each variable. If the value group of five-year age groups in a population table is described in the file *ikaryhmitys5v.vs* and the aggregation rule is in the file *Age\_grouping\_5y.agg*, the corresponding value group specification for the variable age can be as follows:

```
DOMAIN("Age")="Age_grouping_5y";
```

[Appendix 3](#) has an example of grouping files, on which separate detailed instructions are available.

**Statistics Finland's application guideline:** *The classifications used in the grouping files must be based on the classifications of Statistics Finland's Classification Database. The file names of grouping files are given in the Classification Database according to the standardised naming practice and the StatFin online service complies with the standard groupings produced from the Classification Database, which are instructed to be used in the StatFin online service.*

**DOUBLECOLUMN** - not in use (SF)

Metadata type: processing rule

If the keyword value is YES, the code and value of the variable shown by the keyword are given in the row separately in the Excel format and in PC-Axis when the table is shown in matrix format.

This is an obsolete keyword for which there is no use in the present applications.

```
DOUBLECOLUMN("Municipality")=YES;
```

**Statistics Finland's application guideline:** *Not used. In the present PX-Web version the user can select in the user interface whether to choose only headings, codes or both.*

**ELIMINATION** - recommended (SF)

Metadata type: processing rule

The keyword information is used for automatic selection when the user does not choose from the variable any value (class). The keyword is used to define automatic aggregation as allowed or it shows the aggregation class whose value will be automatically chosen when no class is selected from the variable.

Aggregation information calculated in advance should primarily be included in the table and order aggregation to use this information. For example, the specification `ELIMINATION("Municipality")="Koko maa"` means that the already calculated aggregation class `Koko maa` included in the table will always be selected automatically when the user does not select anything from the variable `Municipality`.

If the specification is written as `ELIMINATION("Municipality")=YES`, the selection counts and shows as default selection a sum it produces by dynamically aggregating all class values of the variable concerned.

**NB** Use of dynamic aggregation should be avoided.

***Statistics Finland's application guideline:** An aggregation allowing setting is suitable only for a variable whose information can be aggregated. If the table includes non-aggregated information or aggregates cannot be calculated correctly from basic level information because of masking, aggregation classes should be calculated to the table and show with the `ELIMINATION` setting these ready counted aggregates. In some cases it is also worth considering whether to make separate tables from mere aggregate level information.*

**HEADING** - mandatory

Metadata type: processing rule

Maximum length of individual variable name 75 characters

This keyword is used to specify the table column variables and table view (together with the keyword `STUB`). The column variables used are defined with the keyword, indicated with quotation marks and separated by comma. If the variable list is divided into several lines, line feed is made after the comma separating the variables. `STUB` and `HEADING` also determine the order of data. Transposition of the table **cannot** be made only by changing these keywords but data always have to follow the corresponding order.

**The name of each variable used in the file must be unique.**

```
HEADING="koulutusluokitus" , "sukupuoli" , "tarkastelukohde" ;
```

***Statistics Finland's application guideline:** When the table belongs to some service whole where there is a lot of content using the same classifications, the default view should be set so that the classifications are in the same order in the default views of all tables. The uniform default view is reflected in the `PX-Web` service as a logical selection interface. Needs of different users generally vary much and easy arrangement (pivoting) of the table gives the user the possibility to choose the best possible view, but in some cases with a smaller content or a service directed to special groups it can also be considered whether to set the default view according to the most frequently used view.*

**NB** A separate recommendation will be issued of the recommended default order of variables in `StatFin`.



**INFO** - not in use (SF)

A table specification whose content is placed in the same file as the table itself. The specification saved as the content of the keyword is always included when tables are moved from the table to one's application. The keyword enables the content of the INFOFILE text file when **archiving** the table file.

```
INFO="This text equals to the contents of the INFOFILE file";
```

***Statistics Finland's application guideline:** For the time being not in use at Statistics Finland.*

**INFOFILE** - mandatory (SF)

Metadata type: document metadata

The value of the keyword is the file name of the table specification attached to the table **without the extension**. The document file must be located in the same database as the corresponding .px file. The keyword can refer to the document types .txt, .pdf, .doc, .rtf and .html. PX-Web searches for the files of the allowed file formats corresponding to all keywords and shows the first of them, so the same file need not be imported to the service in any other file format than the one in which it is to be displayed to customers. If INFOFILE is specified, the information button on the tool row in PC-Axis is active, and a link *Table specification* is visible in the selection window of the PX-Web database.

```
INFOFILE="pop_info";
```

***Statistics Finland's application guideline:** The .txt file format is recommended for use in Statistics Finland's databases, which is sectioned according to the title levels used in StatFin. In public free services no such file format should be used whose use may be restricted in customers' technological environment. Use of other file formats can be considered in special cases, and then it must be made sure or agreed with service customers that customers are able to use the file format in question. The information file must be located in the database in the same directory as the corresponding .px file. The information files are named according to the name of the px file corresponding to them. The old practice where several tables can have a common specification should definitely be avoided. Separate instructions have been issued of the naming practices of files, see [Appendix 4](#).*

**KEYS** - in special cases (SF)

Metadata type: statistical metadata

The PC-Axis table can be saved in the so-called sparse matrix format, where pure zero lines are not saved in order to save space (dot and dash codes are **not** interpreted as zero). Sparse matrix saving is directed with the keyword KEYS, which must be given to each row variable in the STUB keyword. The value of the keyword is either VALUES or CODES, and either the corresponding variable value text or code value with apostrophes is recorded as the key in the rows of the data part corresponding to the variable.

```
KEYS("Age")=VALUES;
KEYS("Municipality")=CODES;
```

The sparse matrix data rows start with the text or code keys of row variables in the apostrophes separated by comma, which are followed by the data values of the corresponding column variables (see Example 6.4). The keyword is recommended for use only in **large tables with a very high zero content**.

***Statistics Finland's application guideline:** In order to save space, sparse matrix saving can be used in materials where at least 80 per cent of the data are zero. Not used in smallish (of under one megabyte) files if the zero length was large in them. The CODES setting should usually be used as the keyword value.*

**LANGUAGE** - mandatory (SF)

Metadata type: document metadata

The language setting whose value is a two-character language code according to the ISO-639-1 standard, which indicates the default language of the table.

```
LANGUAGE=" fi " ;
```

**Statistics Finland's application guideline:** During 2006 versions supporting multi-language files will start to be used in applications. One-language files should also be provided with language codes. The language code included in single language files ensures that in future controlled transferability of language versions will work in both directions between px and XML formats.

**LANGUAGES**

Metadata type: document metadata/processing rule

The keyword contains a list of languages used in the document, that is, in which languages the table metadata are in the file. The languages are given with two-character language codes according to the ISO-639-1 standard. The keyword LANGUAGE indicates the **default language** of the document.

```
LANGUAGES=" fi " , " sv " , " en " ;
```

**Statistics Finland's application guideline:** In the course of 2006 versions supporting multi-language files will start to be used in the applications. Instructions will be issued on multi-language files, and their production will start in phases as soon as new tabulation applications support multi-languages and multi-language metadata are available exhaustively to tabulation applications. The file structure of a multi-language table is described briefly in [Appendix 5](#).

**LAST-UPDATED** - mandatory (SF)

Metadata type: document metadata

The keyword contains the timestamp of the last updating time of the file, which always takes the form CCYYMMDD hh:mm. The timestamp is shown in the PX-Web service selection window under *Information*. The time of updating is needed in GESMES or Aremos format data transfer (in PC-Axis program with the selection *File/Save as/Convert*).

```
LAST-UPDATED=" 20050606 09:00 " ;
```

**Statistics Finland's application guideline:** Each PC-Axis file must have the time of the latest update of the file. This is important metadata for controlling the process.

**NB** LAST-UPDATED is the latest update from the viewpoint of the **customer** but from that of the **producer** it is the time of release or end of embargo.

**MAP** - in special cases (SF)

Metadata type: processing rule

The content of the keyword is the name of the map template file corresponding to the table. If a map template specified for the table with the MAP keyword can be used, the thematic map function is shown active in the applications. In the PX-Web server all map templates are located in the shared map template directory of the service.

```
MAP("Municipality")="Finland_municipality";
```

**Statistics Finland's application guideline:** Use is made of the official regional classifications and their common, centrally maintained map templates. If the map function is set active in free databases, this has to be agreed with the Database unit that co-ordinates the dissemination policy and chargeability of services. The map function must not be linked to just any area data table, but tables suitable for map output or those produced specifically for map use are placed in the service to their own directories in accordance with separate instructions.

**NB** The MAP keyword is also used to express an area variable, for which purpose the new VARIABLE-TYPE keyword should be used when the applications start to support it.

**MATRIX** - mandatory

Maximum length 8 characters.

The identifier referring to the origin of the table.

```
MATRIX="vaerak";
```

**Statistics Finland's application guideline:** The abbreviation of the statistics is written as the content of the keyword (see [Appendix 4](#)).

**NEXT-UPDATE** - recommended (SF)

Metadata type: document metadata

The timestamp of the next updating time of the table, which is always of the form CCYYMMDD hh:mm. Applicable for use as additional information in monthly statistics that have a pre-notified accurate release schedule.

```
NEXT-UPDATE="20060505 09:00";
```

**NOTE** - recommended (SF)

Metadata type: statistical metadata

This keyword is used for specifying both table and variable specific footnotes (comments). Footnotes also move on to several transfer files (e.g. Excel).

```
NOTE="Tämä opetus- ja demokäyttöön tarkoitettu esimerkkitaulukko, jota "
"ei säännöllisesti päivitetä";
NOTE("Kunta")="Selitystä Kunta-muuttujasta. . .#toisen rivin teksti"
"tähän #kolmannen rivin teksti#jne. . .";
```

**Statistics Finland's application guideline:** *The use is recommended. The content of the NOTE keyword also moves on to transfer files when transferring a table or its part to Excel. A note or information to be transferred to other file formats should not thus be included only in the table specification in the external file of the table (INFOFILE keyword). The footnote may also have a web link to external additional information pages. A link to the web page is indicated in the same way as with the keyword CONTACT.*

**NOTEX** - not in use (SF)

Metadata type: statistical metadata or document metadata

As NOTE, but functions as a so-called mandatory input comment, which is always visible when the table is opened in the PC-Axis program.

```
NOTEX="Table footnotes...";
NOTEX("Municipality")="Footnotes for the variable Municipality...";
```

**Statistics Finland's application guideline:** *Not used. In Statistics Finland's PX-Web servers the mandatory message window of the CELLNOTEX text is not activated. The content of the keyword is visible when loading the file to PC-Axis but it is not shown in PX-Web services. The footnote may also have a web link to external additional information pages. A link to the web page is indicated in the same way as with the keyword CONTACT.*

**PARTITIONED** - not in use (SF)

Metadata type: processing rule

This keyword is used to demonstrate at most three-level internal hierarchy (sub-field description) of the variable, when the variable codes are included direct in the heading texts. In this way the applications can be shown the field division of codes, when the hierarchical code is given in the value text. For example:

```
VALUES("Area")="0662","0662111","0662112","0662113";
PARTITIONED("Area")="municipality",1,3;
PARTITIONED("Area")="district",4;
```

Here the first three characters of the area value contain the municipality code and the sub-area code starts from position four.

**Statistics Finland's application guideline:** *Not used at Statistics Finland. The key should not be used for describing general hierarchical variables (a specific description mode will be taken into use).*

**PRECISION** - in special cases (SF)

Metadata type: processing rule

The keyword is used to define for an individual variable value the number of decimals differing from those specified in the SHOWDECIMALS keyword **and higher than that** but lower than or equal to the DECIMALS value.

For example, when the saving precision of the table is two decimals and the general presentation precision of the table is without decimals, but some of the data are to be shown with decimals:

```
DECIMALS=2;
SHOWDECIMALS=0;
PRECISION("Tunnusluku", "170 Tuloveroprosentti 2002")=2;
PRECISION("Tunnusluku", "280 Vuosikate%")=1;
PRECISION("Tunnusluku", "394 Suhteellinen velkaantuneisuus-%")=1;
```

***Statistics Finland's application guideline:*** The use has to be always considered case-specifically.

**PRESTEXT** - not in use (SF)

Metadata type: processing rule

This obsolete keyword is used to control the saving and display mode of variable texts and codes. With the keyword value 1 texts and codes are saved to their own keywords, value 2 is used to save texts as codes and vice versa (when PC-Axis knows to use correct headings in variable aggregation) and value 3 indicates that both texts and codes are shown in the table.

```
PRESTEXT("församling")=0;
```

***Statistics Finland's application guideline:*** Do not use! According to Statistics Finland's metadata standard the code is **always** mandatory. The codes are specified with the keyword CODES, whereby the corresponding heading texts are always specified with the keyword VALUES. When the codes are always included, there is no need to interchange texts and codes. In the PX-Web service customers can self choose whether to include the text, code or both in their output or transfer file.

**PX-SERVER** - mandatory in database distributions (SF)

Metadata type: document metadata/processing rule

The keyword content shows in which distribution server or servers the table will be released. In an archived table it indicates in which servers the table was released at the time. The server names are separated with the # sign.

```
PX-SERVER="pxweb2.stat.fi";
```

***Statistics Finland's application guideline:*** This is important control information for the updating process of PX-Web databases. It is used for automatic direction of updates to the correct distribution server.

**REFPERIOD**

Metadata type: statistical metadata (economic time series keyword)

The keyword indicates the statistical period specified for the time series. Shown in a footnote. If the content variable is defined, the keyword is repeated for each value of the content variable.

```
REFPERIOD="2002" ;
```

```
CONTVARIABLE="Import" ;
REFPERIOD("Import")="2000" ;
```

***Statistics Finland's application guideline:*** *There is no guideline because Statistics Finland does not so far use PC-Axis production databases where metadata corresponding to time series keywords would be supported.*

**SEASADJ**

Metadata type: statistical metadata (economic time series keyword)

If the value of the keyword is YES, the time series data are working-day adjusted. This information can be used as a processing rule to prevent re-adjustment of a series adjusted once for working days. If the content variable is defined, the keyword is repeated for each value of the content variable.

```
SEASADJ=YES ;
```

```
CONTVARIABLE="Series" ;
SEASADJ("Series")=YES ;
```

***Statistics Finland's application guideline:*** *There is no guideline because Statistics Finland does not so far use PC-Axis production databases where metadata corresponding to time series keywords would be supported.*

**SHOWDECIMALS** - recommended (SF)

Metadata type: processing rule

The presentation precision of table decimals, which can be 0 to 6 decimals. The value has to be the same or lower than the saving precision specified in the DECIMALS keyword. If no presentation precision is given, the applications assume it to be the same as the saving precision (DECIMALS), but not higher than 6.

```
SHOWDECIMALS=2 ;
```

***Statistics Finland's application guideline:*** *In materials critical in their rounding practices the data are imported to the px file always **rounded to their official release precision**. For example, it is necessary to set in index data only DECIMALS and possibly also SHOWDECIMALS, to which then the same value has to be set as DECIMALS.*

**SOURCE** - mandatory (SF)

Metadata type: document metadata

The producer organisation responsible for the content of the table is specified with this keyword. The source information is shown in the footnote and in the PX-Web selection window under *Information*.

```
SOURCE="Statistics Finland" ;
```

***Statistics Finland's application guideline:*** *Materials produced by Statistics Finland always have the mandatory source reference "Statistics Finland". When Statistics Finland publishes statistics produced by other producers, recorded as sources are what is agreed with the producer in question: the main principle is that external data producers supply the table with all metadata ready for publication.*

***NB*** *The setting COPYRIGHT=YES should always be used.*

**STOCKFA**

Metadata type: statistical metadata (economic time series keyword)

The keyword defines the information type of the data. The alternatives are S (stock), F (flow) or A (average). If the content variable is defined, the keyword is repeated for each value of the content variable.

```
STOCKFA="S" ;
```

```
CONTVARIABLE="Costs" ;
STOCKFA("Costs")="A" ;
```

***Statistics Finland's application guideline:*** *There is no guideline because Statistics Finland does not so far use PC-Axis production databases where metadata corresponding to time series keywords would be supported.*

**STUB** - mandatory

Metadata type: processing rule

Maximum length of individual variable name 75 characters

This keyword is used to specify the table row variables and table view (together with the keyword **HEADING**). The row variables used are defined with the keyword, indicated with quotation marks and separated by comma. If the variable list is divided into several lines, line feed is made after the comma separating the variables. **STUB** and **HEADING** also determine the order of data. Transposition of the table cannot be made only by changing these keywords but data always have to follow the corresponding order.

**The name of each variable used in the file must be unique.**

```
STUB="Municipality", "Year" ;
```

***Statistics Finland's application guideline:*** *When the table belongs to some service whole where there is a lot of content using the same classifications, the default view should be set so that the classifications are in the same order in the default views of all tables. The uniform default view is reflected in the PX-Web service as a logical selection interface. Needs of different users generally vary much and easy arrangement (pivoting) of the table gives the user the possibility to choose the best possible view, but in some cases with smaller content or a service directed to special groups it can also be considered whether to set the default view according to the most frequently used view.*

***NB*** *A separate recommendation will be issued of the recommended default order of variables in StatFin.*

**SUBJECT-AREA** - mandatory  
 Metadata type: statistical metadata  
 Maximum length 100 characters

Name of statistical topic

SUBJECT-AREA="Education" ;

***Statistics Finland's application guideline:*** When naming statistical topics, it should be noted that only by using the statistical topics according to the classification of Official Statistics of Finland (OSF) statistics can be ordered automatically in archiving, for instance, according to the topic. In the same way it is possible to build distribution servers whose directory structure is derived straight from topic names and the files are directed automatically to the correct distribution directories according to their topic names. PX-Edit has a selection list from which the official topic classification can be easily linked to the keyword.

The topic classification is given in [Appendix 4](#).

StatFin and other general databases follow as purely as possible the general topic classification (OSF). In general databases each table must contain data mainly from one statistical topic only. If a general database table includes individual comparison data from another topic, it is classified by its main topic and placed in the PX-Web service to the topic directory of its **primary** content.

In certain special cases deviations from the general classification can be allowed especially in sub-topics. For practical reasons, in indicator services and corresponding compilation services it may be necessary to build topic directories more on the terms of the service content and the viewpoint of its users than by deriving them directly from the general topic classification.

**SUBJECT-CODE** - mandatory  
 Metadata type: statistical metadata  
 Maximum length 5 characters.

The topic code which is also needed for data management needs, such as control information for updating processes, for table cataloguing and as a search key for archived tables.

SUBJECT-CODE="KOU" ;

***Statistics Finland's application guideline:*** The topic classification of Official Statistics of Finland (OSF) is used. PX-Edit has a selection list from which the official topic classification can be easily linked to the keyword.

The topic classification is given in [Appendix 4](#).



**TIMEVAL** - recommended (SF)

Metadata type: statistical metadata and processing rule

TIMEVAL is a variable-specific specification that can appear only once in a table. The keyword value is the TLIST specification expressing the time scale and period, which can be given either in time limit or list format.

For example, for an annual series (time limit format):

```
TIMEVAL("Time")=TLIST(A1,"1994"- "1996");
```

or (list format):

```
TIMEVAL("Time")=TLIST(A1),"1994","1995","1996";
```

for a monthly series:

```
TIMEVAL("Time")=TLIST(M1,"199605"- "199704");
```

or:

```
TIMEVAL("Time")=TLIST(M1),"199605","199606","199607","199608","199609",
"199610","199611","199612","199701","199702","199703","199704";
```

The following TLIST specifications are used:

- A1 in annual statistics in form CCYY (CC century, YY year)
- H1 quarterly in form CCYYH, where H is 1 or 2
- Q1 quarterly in form CCYYQ, where Q is 1-4.
- M1 in monthly statistics in form CCYYMM, where MM is 01-12
- W1 in weekly information in form CCYYWW, where WW is 01-52.

TIMEVAL is used in PC-Axis when converting tables into Aremos and Gesmes/Ecoser format.

***Statistics Finland's application guideline:** Even if no other time series features are used, it is recommended to utilise the keyword for expressing the time variable. Time is a special variable where the measures directed to it can be in advanced applications controlled when the application is able to distinguish time variables from other variables (until a separate VARIABLE-TYPE keyword can be used for this).*

**TITLE** - mandatory

Metadata type: statistical metadata

Maximum length of one text row is 98 characters

The keyword is used to specify the **changing title** of the table, which is formed of the CONTENTS keyword text (basic title) and the names of the variables selected to the table.

```
STUB="kunta","vuosi";
HEADING="koulutusluokitus";
CONTENTS="Koululaitoksen oppilaat";
DESCRIPTION="Koululaitoksen oppilaat 2003-2005";
TITLE="Koululaitoksen oppilaat muuttujina kunta, vuosi ja "
"koulutusluokitus";
```

**NB** Please note how years are entered.

***Statistics Finland's application guideline:** Always created to the file using the keyword CONTENTS and variable names.*

**UNITS** - mandatory

Metadata type: statistical metadata

Maximum length 80 characters

Describes the measurement unit, the index basic year, etc. (people, 1000 euro, 1995=100).

```
UNITS="Number of pupils";
```

If the table has several units, use the CONTVARIABLE keyword. If CONTVARIABLE is specified, the keyword UNITS is given to each content variable. Then it is also necessary (because the keyword is mandatory in PC-Axis) to give table-specific UNITS, whose value text can now be empty. For example:

```
UNITS=" ";
CONTVARIABLE="Information";
VALUES("Information")="population", "area";
UNITS("population")="people";
UNITS("area")="1000 ha";
```

**Statistics Finland's application guideline:** *In very small (two-dimensional) tables and in well-founded cases the unit information can also be written to the text of the row or column variable concerned (table heading), in which case the content of the UNITS keyword is the text "unit information in the variable name".*

**UPDATE-FREQUENCY** - recommended (SF)

Metadata type: document metadata

Repetition frequency of updating. Suitable for use as additional information in monthly statistics, for example, which do not have a pre-notified release schedule, but it can be reported that the data are updated regularly.

```
UPDATE-FREQUENCY=" 6 / A " ;
```

**Statistics Finland's application guideline:** *Used is language-independent presentation frequency/time period, where the time period is marked with letters A (annual), M (month), W (week) and D (day). Thus a table updated once a month is coded as 1/M and a table updated twice in five years as 2/5A.*

**VALUENOTE** - recommended (SF)

Metadata type: statistical metadata

A footnote for a specific value of the variable. Visible in PX-Web and PC-Axis in connection with the table footnote.

```
VALUENOTE("Municipality", "Espoo")="Footnote for Espoosta";
```

**Statistics Finland's application guideline:** *Used where necessary and in preparation for improving support for the feature in the applications. Visible in current application versions still only in the footnote, but in future in other applications than in PX-Edit as a link attached to table cells.*

**VALUENOTEX** - not in use (SF)

Metadata type: statistical metadata

As VALUENOTE, but displayed automatically when opening a table (similarly as NOTEX).

***Statistics Finland's application guideline:** Not used. In Statistics Finland's PX-Web servers the mandatory message window of the VALUENOTEX text is not activated. The content of the keyword is visible when downloading the file to PC-Axis but it is not shown in PX-Web services.*

**VALUES** - mandatory

Metadata type: statistical metadata

Maximum length of one text row is 256 characters

The keyword VALUES is specified for each variable. The values (classification headings) **must be unique**, and they also appear in the order they have in rows or columns. The heading texts are in quotation marks and they are separated by commas.

If the list is divided into several lines, line feed is made after the comma separating the values.

```
VALUES ("Vuosi")="1987", "1988", "1989";
VALUES ("Muuttotyyppi")="Maassamuuton tulomuutto",
"Maassamuuton lähtömuutto", "Maahanmuutto", "Maastamuutto";
```

***Statistics Finland's application guideline:** Standardised variable names from the Classification Database are always used, so the variable values are based on official classifications. The classifications and variables used in tailored services that are not entered in the Classification Database are documented in the product specifications of the product in question. Aggregate classes are always given before their sub-items. For example:*

```
VALUES ("Area")="Whole country", "Alahärmä", "Alajärvi", . . .
```

*If the classification has hierarchies, which are given with indentations, the indentations are filled with **leading dots** to ensure that applications do not remove leading blanks. For example:*

```
VALUES ("Area")="Whole country", ".Uusimaa", "..Espoo", "..Helsinki";
```

*The standard grouping produced from the Classification Database already has leading dots.*

## 8. Typical sources of error

Even if the same things come up in the relevant places in these instructions, we ought to go over some rules that have to be observed when programming applications producing PC-Axis files in order to avoid often repeated errors and frequently asked questions.

- spaces must **not** be used in file names
- the initial part of the file is always comprised of metadata, data are always in the final section of the file
- it should be noted for both metadata and data that words must not be divided in lines
- a long text is divided to **file** lines so that the line break ends in a quotation mark and the following line starts with a quotation mark (the # sign shows the line feed with which information is given in lines in **applications**)
- variable names, value texts or codes must not be broken
- the keywords must be in a certain order (see [Appendix 2](#))
- the end of text character of each section starting with a keyword is a semi-colon, directly followed by a line feed
- the last character of the text preceding the quotation mark must not be empty (= no tailing spaces)
- quotation marks ( " ) **must not** be used in texts
- special characters such as parentheses, brackets and semi-colons should be avoided in texts, because they are reserved for indicating the file structure
- use of some special characters (e.g. tabulator) may prevent opening of a file in applications
- transfer of data to Excel may become problematic if the texts have parentheses and arithmetic signs inside them
- variable names have to be unique, similarly as variable values and codes corresponding to them
- numerical cells of the data part are without quotation marks, similarly as YES/NO type settings and numerical values (e.g. decimal settings) in the meta part
- data part cells have to be set apart by a separator character (space or tabulator)
- the end of the data part must **not** have a space before a semi-colon
- the decimal part needs not be saved in integers (saves space)
- line feed **is not** a separator character

## Appendices

### Appendix 1. Keywords in alphabetical order

	SF standard	PC-Axis	PX-Web	PX-Edit	Length	Comment
AGGREGALLOWED		Supported	Supported	Supported		
AUTOPEN		Supported	Supported	Supported		
AXIS-VERSION	Mandatory	Supported	Supported	Supported	80	
BASEPERIOD		Supported	Supported	Supported	80/row	1)
CELLNOTE		Supported	Supported	Supported	80/row	
CELLNOTEX		Supported	Supported	Supported	80/row	
CFPRICES		Supported	Supported	Supported	1	1)
CHARSET	Mandatory (ANSI)	Supported	Mandatory (ANSI)	Mandatory		
CODEPAGE		Supported	Supported	Supported	20	
CODES	Mandatory	Supported	Supported	Supported	256	
CONFIDENTIAL	Forbidden	Supported	Supported	Supported	20	
CONTACT	Mandatory	Supported	Supported	Supported	80/row	1)
CONTENTS	Mandatory	Mandatory	Mandatory	Mandatory	256	
CONTVARIABLE		Supported	Supported	Supported	80	
COPYRIGHT	Mandatory	Supported	Supported	Supported		
CREATION-DATE	Mandatory	Supported	Supported	Supported	80	
DATA	Mandatory	Mandatory	Mandatory	Mandatory		
DATABASE	Recommended	Supported	Supported	Supported	80/row	
DATANOTECELL		Supported		Supported	20	5)
DATANOTESUM		Supported			20	5)
DATASYMBOL1		Supported		Supported	20	5)
DATASYMBOL2		Supported		Supported	20	5)
DATASYMBOL3		Supported		Supported	20	5)
DATASYMBOL4		Supported		Supported	20	5)
DATASYMBOL5		Supported		Supported	20	5)
DATASYMBOL6		Supported		Supported	20	5)
DATASYMBOLNIL		Supported		Supported	20	5)
DATASYMBOLSUM		Supported		Supported	20	5)
DAYADJ		Supported	Supported	Supported		1)
DECIMALS	Mandatory	Mandatory	Mandatory	Mandatory		
DEFAULT-GRAPH				Supported		5)
DESCRIPTION	Mandatory	Supported	Supported	Supported	98	
DESCRIPTIONDEFAULT		Supported	Supported	Supported		
DIRECTORY-PATH	Mandatory (those exported to PX-Web)			Supported		5)
DOMAIN		Supported	Supported	Supported	80	
DOUBLECOLUMN	Not recommended	Matrix format	Matrix format	Supported		
ELIMINATION	Recommended	Supported	Supported	Supported		
HEADING	Mandatory	Mandatory	Mandatory	Mandatory	80 (à)	
INFO				Supported		5)
INFOFILE	Mandatory	Supported	Supported	Supported	256/row	
KEYS		Supported	Supported	Supported		
LANGUAGE	Mandatory	Supported	Supported	Supported	2	
LANGUAGES		Supported			n x 2	5)
LAST-UPDATED	Mandatory	Supported	Supported	Supported	80	1)
LINK				Supported		5)
MAP		Supported	Supported	Supported	80	
MATRIX	Mandatory	Mandatory	Mandatory	Mandatory	8	

NEXT-UPDATE				Supported		5)
NOTE	Recommended	Supported	Supported	Supported	80/row	4)
NOTEX		Supported	Supported	Supported	80/row	4)
PARTITIONED		Supported	Supported	Supported		
PRECISION		Supported	Supported	Supported		
PRETEXT	Not recommended	Supported	Supported	Supported		
PX-SERVER	Mandatory (exported to PX- Web)			Supported	256	5)
REFPERIOD		Supported	Supported	Supported	80/row	1)
SEASADJ		Supported	Supported	Supported		1)
SHOWDECIMALS		Supported	Supported	Supported		
SOURCE	Mandatory	Supported	Supported	Supported	80/row	
STOCKFA		Supported	Supported	Supported	1	1)
STUB	Mandatory	Mandatory	Mandatory	Mandatory	80 (à)	
SUBJECT-AREA	Mandatory	Mandatory	Mandatory	Mandatory	100	
SUBJECT-CODE	Mandatory	Mandatory	Mandatory	Mandatory	5	
SURVEY				Supported		5)
TIMEVAL		Supported	Supported	Supported		
TITLE	Mandatory	Mandatory	Mandatory	Mandatory	98	
UNITS	Mandatory	Mandatory	Mandatory	Mandatory	80	1)
UPDATE-FREQUENCY				Supported		5)
VALUENOTE		Supported	Supported	Supported	80/row	
VALUENOTEX		Supported	No	Supported	80/row	2)
VALUES	Mandatory	Mandatory	Mandatory	Mandatory	256	
VARIABLE-TYPE				Supported		5)

- 1) Appears for each content variable if the CONTVARIABLE keyword is defined
- 2) Not in the user interface
- 3) A quality requirement recorded in the SF standard column will be adopted in Statistics Finland's database services during 2006.
- 4) The NOTE and NOTEX keywords can be defined both table and variable-specifically
- 5) A new keyword or one to be adopted during 2006.

## Appendix 2. Keywords in presentation order

	SF standard	PC-Axis	PX-Web	PX-Edit	Length	Comment
CHARSET	Mandatory (ANSI)	Supported	Mandatory (ANSI)	Mandatory		
AXIS-VERSION	Mandatory	Supported	Supported	Supported	80	
CODEPAGE		Supported	Supported	Supported	20	
LANGUAGE	Mandatory	Supported	Supported	Supported	2	
LANGUAGES					n x 2	
CREATION-DATE	Mandatory	Supported	Supported	Supported		
DECIMALS	Mandatory	Mandatory	Mandatory	Mandatory		
SHOWDECIMALS		Supported	Supported	Supported		
MATRIX	Mandatory	Mandatory	Mandatory	Mandatory	8	
AGGREGALLOWED		Supported	Supported	Supported		
AUTOPEN		Supported	Supported	Supported		
CONFIDENTIAL	Forbidden	Supported	Supported	Supported	20	
COPYRIGHT	Mandatory	Supported	Supported	Supported		
NEXT-UPDATE				Supported		5)
UPDATE-FREQUENCY				Supported		5)
PX-SERVER	Mandatory (exported to PX-Web)			Supported	256	5)
DIRECTORY-PATH	Mandatory (exported to PX-Web)			Supported	256	5)
DEFAULT-GRAPH				Supported		5)
SUBJECT-CODE	Mandatory	Mandatory	Mandatory	Mandatory	5	
SUBJECT-AREA	Mandatory	Mandatory	Mandatory	Mandatory	100	
DESCRIPTION	Mandatory	Supported	Supported	Supported	98	
TITLE	Mandatory	Mandatory	Mandatory	Mandatory	98	
DESCRIPTIONDEFAULT		Supported	Supported	Supported		
CONTENTS	Mandatory	Mandatory	Mandatory	Mandatory	256	
UNITS	Mandatory	Mandatory	Mandatory	Mandatory	80	1)
STUB	Mandatory	Mandatory	Mandatory	Mandatory	80 (à)	
HEADING	Mandatory	Mandatory	Mandatory	Mandatory	80 (à)	
CONTVARIABLE		Supported	Supported	Supported	80	
VALUES	Mandatory	Mandatory	Mandatory	Mandatory	256	
TIMEVAL		Supported	Supported	Supported		
CODES	Mandatory	Supported	Supported	Supported	256	
DOUBLECOLUMN	Not recommended	Matrix format	Matrix format	Supported		
PRETEXT	Not recommended	Supported	Supported	Supported		
DOMAIN		Supported	Supported	Supported	80	
VARIABLE-TYPE				Supported	256	5)
MAP		Supported	Supported	Supported	80	
PARTITIONED		Supported	Supported	Supported		
ELIMINATION	Recommended	Supported	Supported	Supported		
PRECISION		Supported	Supported	Supported		
LAST-UPDATED	Mandatory	Supported	Supported	Supported	80	1)
STOCKFA		Supported	Supported	Supported	1	1)
CFPRICES		Supported	Supported	Supported	1	1)
DAYADJ		Supported	Supported	Supported		1)
SEASADJ		Supported	Supported	Supported		1)
UNITS ( )		Supported	Supported	Supported	80	6)
CONTACT	Mandatory	Supported	Supported	Supported	80/row	1)
REFPERIOD		Supported	Supported	Supported	80/row	1)

BASEPERIOD		Supported	Supported	Supported	80/row	1)
DATABASE	Recommended	Supported	Supported	Supported	80/row	
SOURCE	Mandatory	Supported	Supported	Supported	80/row	
SURVEY				Supported	256	5)
INFOFILE	Mandatory	Supported	Supported	Supported	256/row	
INFO				Supported		5)
LINK				Supported	256	5)
NOTEX		Supported	Supported	Supported	80/row	4)
NOTE	Recommended	Supported	Supported	Supported	80/row	4)
VALUENOTEX		Supported	No	Supported	80/row	2)
VALUENOTE		Supported	Supported	Supported	80/row	
CELLNOTEX		Supported	Supported	Supported	80/row	
CELLNOTE		Supported	Supported	Supported	80/ row	
DATASYMBOL1		Supported		Supported	20	5)
DATASYMBOL2		Supported		Supported	20	5)
DATASYMBOL3		Supported		Supported	20	5)
DATASYMBOL4		Supported		Supported	20	5)
DATASYMBOL5		Supported		Supported	20	5)
DATASYMBOL6		Supported		Supported	20	5)
DATASYMBOLNIL		Supported		Supported	20	5)
DATASYMBOLSUM		Supported		Supported	20	5)
DATANOTECELL		Supported		Supported	20	5)
DATANOTESUM		Supported			20	5)
KEYS		Supported	Supported	Supported		
DATA	Mandatory	Mandatory	Mandatory	Mandatory		

- 1) Appears for each content variable if the CONTVARIABLE keyword is defined
- 2) Not in the user interface
- 3) A quality requirement recorded in the SF standard column will be adopted in Statistics Finland's database services during 2006.
- 4) The NOTE and NOTEX keywords can be defined both table and variable-specifically
- 5) A new keyword or one to be adopted during 2006
- 6) If CONTVARIABLE is in use, content variable-specific UNITS values are written here.



### Appendix 3. About grouping files

Grouping files are similar to Windows .ini files, where sections (blocks) are separated with keywords in brackets. The content of the sections is comprised of indexation, which combines class codes and names to aggregation rules. The index and classes related to it are separated by an equality sign. The grouping is described with two files, of which the .vs file defines the value group and the .agg file its aggregation rule. The grouping can be created with the PC-Axis function *File/Classifications*.

Separate detailed instructions are available on compilation of grouping files.

Model from the file ikaryhmitys5v.vs	Model from the file 25_vuotisloukat.agg
[Descr] Name=ikä5 Pretext=5-vuotisikäryhmät Type=V [Aggreg] 1=25_vuotisloukat.agg [Domain] 1=ikaryhmitys5v [Valuecode] 1=0-4 2=5-9 3=10-14 4=15-19 5=20-24 6=25-29 7=30-34 8=35-39 9=40-44 10=45-49 11=50-54 12=55-59 13=60-64 14=65-69 15=70-74 16=75+ [Valuetext] 1=0-4 vuotta 2=5-9 vuotta 3=10-14 vuotta 4=15-19 vuotta 5=20-24 vuotta 6=25-29 vuotta 7=30-34 vuotta 8=35-39 vuotta 9=40-44 vuotta 10=45-49 vuotta 11=50-54 vuotta 12=55-59 vuotta 13=60-64 vuotta 14=65-69 vuotta 15=70-74 vuotta 16=75+	[Aggreg] Name=25_vuotisloukat.agg Valueset=ikä25 Desc=25-vuotisikäryhmät 1=0-24 2=25-49 3=50-74 4=75+ [Aggttext] 1=0-24-vuotiaat 2=25-49-vuotiaat 3=50-74-vuotiaat 4=75+-vuotiaat [0-24] 1=0-4 2=5-9 3=10-14 4=15-19 5=20-24 [25-49] 1=25-29 2=30-34 3=35-39 4=40-44 5=45-49 [50-74] 1=50-54 2=55-59 3=60-64 4=65-69 5=70-74 [75+] 1=75+

## Appendix 4. Statistics Finland's classification of statistical topics, abbreviations of statistics and naming of files

The classification of statistical topics that is used in the keyword SUBJECT-CODE

ASU	Housing	PAV	Services
ELI	Living Conditions	RAH	Financing and Insurance
ENE	Energy	RAK	Construction
HIN	Prices and Costs	SOS	Social Protection
JUL	Government Finance	TEO	Manufacturing
KAN	National Accounts	TER	Health
KAU	Trade	TTT	Science, Technology and Information Society
KOU	Education	TUL	Income and Consumption
KLT	Culture and the Media	TYM	Labour Market
LII	Transport and Tourism	VAA	Elections
MAA	Agriculture, Forestry and Fishery	VRM	Population
OIK	Justice	YMP	Environment and Natural Resources
PAL	Wages, Salaries and Labour Costs	YRI	Enterprises

Abbreviations of statistics are given as a table on Statistics Finland's Intranet:

[http://intra.stat.fi/admin/tilrak/tilrak\\_tilastot\\_11\\_5\\_2004.xls](http://intra.stat.fi/admin/tilrak/tilrak_tilastot_11_5_2004.xls).

There are also instructions on the Intranet on naming of PC-Axis files according to the abbreviations of statistics

[http://intra.stat.fi/tilastokeskus/statfin2006/px\\_tiedostonimistandardi\\_2006.pdf](http://intra.stat.fi/tilastokeskus/statfin2006/px_tiedostonimistandardi_2006.pdf)

## Appendix 5. About multi-language metadata

The following is an example of a file where the metadata are in three languages. First the languages (LANGUAGES) in use are defined and of them the default language (LANGUAGE). Metadata according to the default language are presented as in a single-language file so that the file also functions (but only in its default language) in those old applications that do not have multi-language support. Texts differing from the default language are always given after the keyword with the language code in brackets.

```
CHARSET="ANSI";
AXIS-VERSION="2005";
LANGUAGE="en";
LANGUAGES="en", "sv", "fi";
CREATION-DATE="20050217 18:34";
DECIMALS=0;
SHOWDECIMALS=0;
MATRIX="BE0101F1";
SUBJECT-CODE="BE";
SUBJECT-AREA="Population";
SUBJECT-AREA[sv]="Befolkning";
SUBJECT-AREA[fi]="Väestö";
DESCRIPTION="Migration 2003";
DESCRIPTION[sv]="Flyttningar 2003";
DESCRIPTION[fi]="Väestönmuetokset 2003";
TITLE="Migration by region, age, period, type and sex";
TITLE[sv]="Flyttningar efter region, ålder, tid, typ och kön";
TITLE[fi]="Väestönmuetokset muuttujina kunta, ikä, aika, tyyppi ja sukupuoli";
CONTENTS="Migration";
CONTENTS[sv]="Flyttningar";
CONTENTS[fi]="Väestönmuetokset";
UNITS="number";
UNITS[sv]="antal";
UNITS[fi]="henkilöä";
STUB="region", "age";
STUB[sv]="region", "ålder";
STUB[fi]="kunta", "ikä";
HEADING="period", "type", "sex";
HEADING[sv]="tid", "typ", "kön";
HEADING[fi]="aika", "tyyppi", "sukupuoli";
CONTVARIABLE="type";
CONTVARIABLE[sv]="typ";
CONTVARIABLE[fi]="tyyppi";
VALUES("region")="Sweden", "Stockholm county", "Upplands Väsby", "Vallentuna";
VALUES[sv]("region")="Riket", "Stockholms län", "Upplands Väsby", "Vallentuna";
VALUES[fi]("kunta")="Ruotsi", "Tukholman lääni", "Upplands Väsby", "Vallentuna";
VALUES("age")="20", "21", "22";
VALUES[sv]("ålder")="20", "21", "22";
VALUES[fi]("ikä")="20", "21", "22";
VALUES("period")="2003";
VALUES[sv]("tid")="2003";
VALUES[fi]("aika")="2003";
VALUES("type")="Immigrated", "Outmigrated";
VALUES[sv]("typ")="Inflyttade", "Utflyttade";
VALUES[fi]("tyyppi")="Tulomuutto", "Lähtömuutto";
VALUES("sex")="men", "women";
VALUES[sv]("kön")="män", "kvinnor";
VALUES[fi]("sukupuoli")="miehet", "naiset";
CODES("region")="00", "01", "0114", "0115";
CODES[sv]("region")="00", "01", "0114", "0115";
CODES[fi]("kunta")="00", "01", "0114", "0115";
CODES("age")="20", "21", "22";
CODES[sv]("ålder")="20", "21", "22";
CODES[fi]("ikä")="20", "21", "22";
CODES("type")="BE0101F1", "BE0101F2";
CODES[sv]("typ")="BE0101F1", "BE0101F2";
CODES[fi]("tyyppi")="BE0101F1", "BE0101F2";
CODES("sex")="1", "2";
CODES[sv]("kön")="1", "2";
CODES[fi]("sukupuoli")="1", "2";
```

```

DOMAIN("age")="Age";
DOMAIN[sv]("ålder")="Ålder";
DOMAIN[fi]("ikä")="ikä";
ELIMINATION("region")="Sweden";
ELIMINATION[sv]("region")="Riket";
ELIMINATION[fi]("kunta")="Ruotsi";
LAST-UPDATED("Inmigrated")="20040212 16:28";
LAST-UPDATED[sv]("Inflyttade")="20040212 16:28";
LAST-UPDATED[fi]("Tulomuutto")="20040212 16:28";
UNITS("Inmigrated")="number";
UNITS[sv]("Inflyttade")="antal";
UNITS[fi]("Tulomuutto")="henkilö";
UNITS("Outmigrated")="number";
UNITS[sv]("Utflyttade")="antal";
UNITS[fi]("Lähtömuutto")="henkilö";
SOURCE="Statistics Sweden";
SOURCE[sv]="Statistiska centralbyrån (SCB)";
SOURCE[fi]="SCB";
CONTACT("Inmigrated")="Ewa Eriksson, SCB#Tel: +46 19-17 67 43";
CONTACT[sv]("Inflyttade")="Ewa Eriksson, SCB#Tel: 019-17 67 43";
CONTACT[fi]("Tulomuutto")="Ewa Eriksson, SCB#Tel: 019-17 67 43";
DATABASE="SDB";
DATABASE[sv]="SDB";
DATABASE[fi]="SDB";
INFOFILE="BE0101en";
INFOFILE[sv]="BE0101sv";
INFOFILE[fi]="BE0101fi";
NOTEX="Mandatory English note";
NOTEX[sv]="Obligatorisk svensk not";
NOTEX[fi]="Pakollinen suomalainen alaviite";
NOTEX("region")="Mandatory footnote for region";
NOTEX[sv]("region")="Obligatorisk fotnot för region";
NOTEX[fi]("kunta")="Pakollinen suomalainen alaviite muuttujalle kunta";
NOTE("age")="Age refers to age attained by the end of the year, "
"i.e. in principal, an account for the year of birth.";
NOTE[sv]("ålder")="Med ålder avses uppnådd ålder vid årets slut, "
"d.v.s. i princip en redovisning efter födelseår.";
NOTE[fi]("ikä")="Sama suomeksi"
VALUENOTEX("type","Inmigrated")="Migration cannot be added with regard to "
"regions. When, for instance, municipalities are added to metropolitan areas, "
"migration to the region is added to migrations within the region.#With "
"calculating of age-specific rates per 1 000 of the mean population you shall "
"use the mean population per year of birth. And some more";
VALUENOTEX[sv]("typ","Inflyttade")="Flyttningar kan inte summeras vad avser "
"region. Vid t ex summering av kommuner till storstadsområde, summeras "
"flyttningar till regionen med flyttningar inom regionen.#Vid beräkning av "
"åldersdifferentierade tal per 1 000 av medelfolkmängden skall medelfolkmängden "
"per födelseår användas.";
VALUENOTEX[fi]("tyyppi","Tulomuutto")="Väestönmuutoksia ei saa laskea yhteen";
VALUENOTE("type","Inmigrated")="Request value note for inmigrated";
VALUENOTE[sv]("typ","Inflyttade")="Frivillig fotnot för inflyttade.";
VALUENOTE[fi]("tyyppi","Tulomuutto")="Suomalainen alaviite, tulomuutto.";
CELLNOTE("","20","","","Inmigrated","")="Cellnote inmigrated 20 years";
CELLNOTE[sv]("","20","","","Inflyttade","")="Cellnot inflyttade 20 år.";
CELLNOTE[fi]("","20","","","Tulomuutto","")="Soluviite, tulomuutto, 20.";
DATA=
456 938 223 327
613 948 305 464
835 968 325 511
590 1017 480 741
771 1145 668 993
924 1273 672 906
14 38 22 43
21 36 34 57
26 36 34 34
17 20 22 30
13 21 29 42
20 23 33 35;

```

## **Appendix 6. References to Statistics Finland's other guidelines and standards**

### **Intranet > StatFin2006 > Guidelines and standards**

User instructions of classifications and PC-Axis groupings

- general classification guidelines (also including regulations on maintenance and division of responsibilities)
- recommendation for regional classifications (separate instructions on use of regional classifications)

Structural instructions of the table specification (changed from previous StatFin).

Instructions on naming of files (applied for databases from naming used in tilastokeskus.fi).

PX-iMap application instructions (use of map features).

Directory structure recommendation for PX-Web servers (subject folders).

PC-Axis technical support pages.

## **Appendix 7. References**

The original documents maintained by Statistics Sweden (SCB) can be found at:

<http://www.pc-axis.scb.se/>