

PX-Edit 3.3

Introduction

About PX-Edit

- Developed at Statistics Finland to efficiently **edit and create** statistical tables
- Based on the PC-Axis file format, but not restricted to it
- Main parts:
 - main window (+ system tray)
 - table spreadsheets
 - metadata editor
 - other special purpose windows
- Multilingual, easy to translate

Design philosophy

- PC-Axis tables
 - must be able to handle all **normal** px files
 - should be able to handle incomplete or even damaged tables
 - is able to repair numerous errors
 - is able to point out numerous other errors
 - changes are reflected immediately in the spreadsheet
 - no artificial table size limitations
- Other source file formats
 - imports data or metadata from different sources including text files, Excel tables and clipboard data
- **Feasible** new features are being added regularly

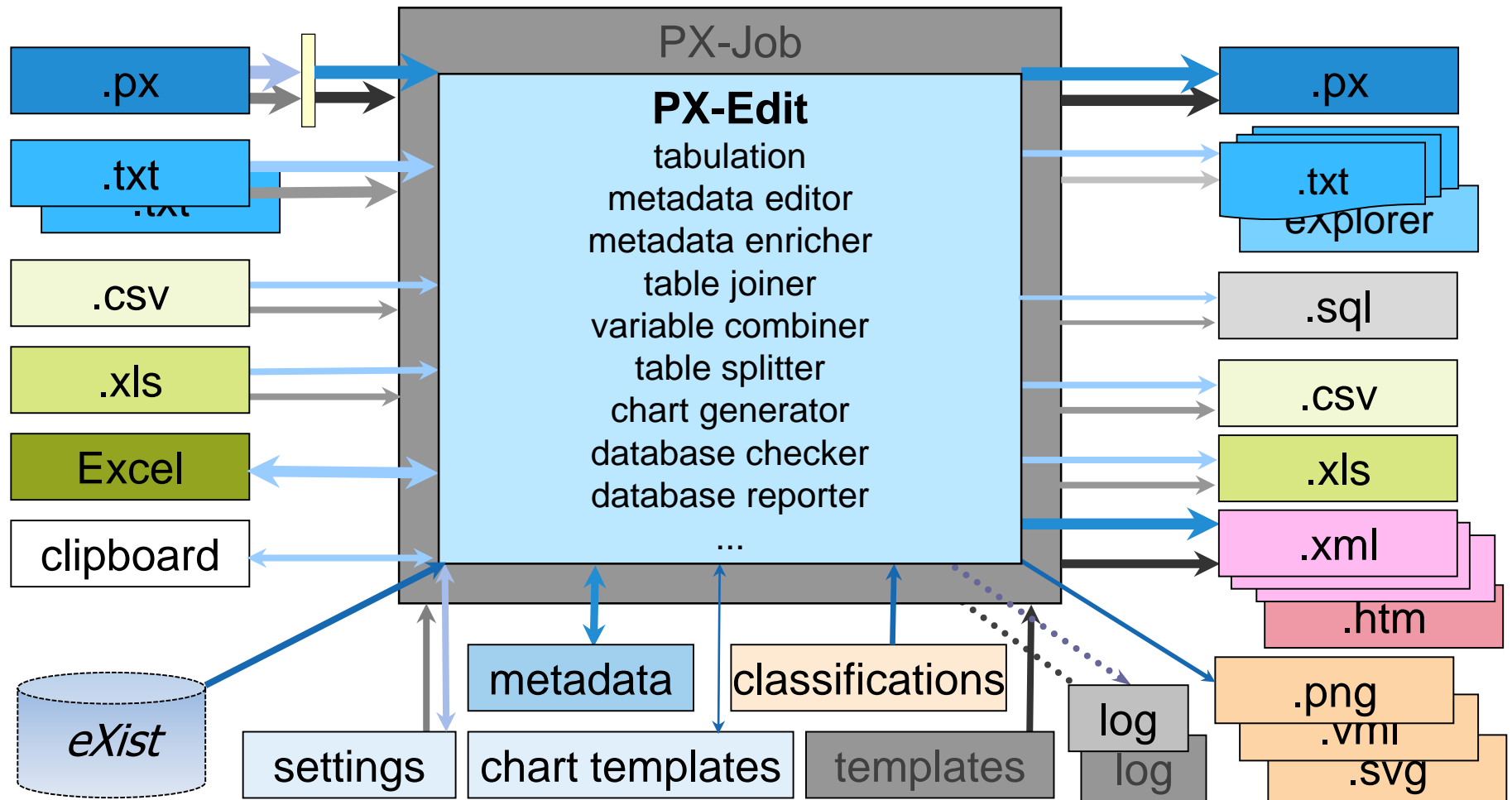
Terminology

- Variable
 - datacube edge (~ classification)
 - e.g. Country, Gender, Year
- Value
 - a single variable item (~ class)
 - either value texts (compulsory) or value codes (may be missing)
 - e.g. Finland, Sweden; man, woman; 1995, 1996, 1997, ...
- Data
 - numeric data
 - dot codes for missing, concealed, etc. data: " ." – " "

Settings

- Working environment: workstation or LAN
- Organisational settings: `px-edit_main_33.ini`, `xdf_33.ini`
 - mandatory keywords
 - environmental settings
 - user and error logs
 - own keywords
- User settings: `px-edit_33.ini`
 - personal settings

PX-Edit in a nutshell



Main window

- Open files
 - *File/Open* [**Ctrl+o**], most recently used list
 - drag and drop support (files or directory folders)
 - read *structural* tables from text files, Excel, xls, csv, ...
- Save in different file formats
- Change user settings
 - language, fonts, *Fill item*, user level ...
- Special operations
 - table joining, table cloning, database checking, reporting and standardizing, ...
- Help always accessible [**F1**]

Table window

- Displays the data **visually**
 - range colouring
 - table, variable, value or cell notes
 - data range, variable names, ...
- *File, Edit, Language* and *Window* menus
- Clipboard support
- Multiple tables may be open simultaneously
 - only one table is **active** at a time
- All data is **not** loaded into the spreadsheet!

Free tabulation

- Change the variable order
 - aggregation is possible, too
- Select variable values
- Sort variable values or codes
 - alphabetically
 - according to the data values
 - free sorting
- All settings can be reset to their original state

Input data

- PC-Axis files (px) [**Ctrl+o**]
- Metadata files (pxk)
- Structural text files (txt, dat, csv, ...)
 - tabulator, semicolon, comma or space delimited
 - arbitrarily large
- Structural Excel files (xls, xlsx, csv)
- Structural tables from Excel worksheet [**Ctrl+e**, **Ctrl+Shift+e**]
- Archives (zip)
- Structural tables from clipboard
- *eXist* database (Statistics Finland's)
- New empty tables [**Ctrl+n**]

Opening a px file: Unicode or *CODEPAGE*

- If the file contains the *Byte Order Mark* (BOM) header → **Unicode**
- Selecting *PC-Axis Unicode files* in the *File Open* window
 - the Unicode check is passed → **Unicode**
- *Character conversion/Ignore coding* is not set, and *CODEPAGE* is
 - `utf-8` and the Unicode check is passed → **Unicode**
 - a valid iso-8859 coding → **ISO-8859**
 - a valid WinANSI coding → **WinANSI**
 - a valid DOS codepage → **DOS**
- *Character conversion/Ignore coding* is set, or *CODEPAGE* is not recognised
 - *Settings/Check for Unicode* is set, and the Unicode check is passed → **Unicode**

Opening a px file: *LANGUAGE*

- The language dependent codes are based on the main table language
 - the system language will be used if *Character conversion/System language coding* is set or there is no *LANGUAGE* setting in the px file
 - if the code cannot be deducted, the Western code will be used instead
- *Character conversion/ISO-8859* is set
 - a valid iso-8859 coding (or `iso-8859-15`) → **ISO-8859**
- *Character conversion/DOS coding* is set and *CHARSET* is not ANSI
 - a valid DOS codepage (or `codepage-858`) → **DOS**
- Otherwise a valid WinANSI coding (or `Windows-1252`) → **WinANSI**

Structural tables

- A structural table contains the **necessary** data
 - table title
 - language code (optional)
 - variable names
 - variable values and/or codes
 - data part
 - cell level information possible
- Two different types
 - general table format
 - record mask for fixed-length sequential files

Syntax checking

- Severe errors
 - mandatory keywords: table will not be opened
 - others: errors will be bypassed
- All found errors will be shown
 - message contains the table and file names
- Some errors may be **repaired**
 - e.g. missing *UNITS*, timestamps and *DECIMALS* problems
 - multiple table opening includes the **Repair all** option

Saving a px file: Unicode or *CODEPAGE*

- *Settings/Save in Unicode* is set → **Unicode**
- Selecting *PC-Axis Unicode files* in the *File Save* window → **Unicode**
- If *CODEPAGE* exists and no *Character conversion* is set
 - `utf-8` → **Unicode**
 - a valid iso-8859 coding → **ISO-8859**
 - a valid WinANSI coding → **WinANSI**

Saving a px file: *LANGUAGE*

- The language dependent codes are based on the main table language
 - the system language will be used if *Character conversion/System language coding* is set or there is no *LANGUAGE* setting in the px file
 - if the code cannot be deducted, the Western code will be used instead
- *Character conversion* is set to *iso-8859*
 - a valid iso-8859 coding (or `iso-8859-15`) → **ISO-8859**
- Otherwise a valid WinANSI coding (or `Windows-1252`) → **WinANSI**
- The Unicode files will be saved with the Byte Order Mark (BOM) header
- *CODEPAGE* will be set as the used conversion code
- *CHARSET* will be set as `ANSI`
- DOS conversion is **not** supported in saving

Output formats

- PC-Axis file (px) [**Ctrl+s**]
- Pure metadata file (pxk)
- Excel files (xls, csv)
- Structural text files (txt, ...) [**Ctrl+t**]
 - freely adjustable
- CoSSI/XML table formats (xml)
- *simplePXsql* files (sql)
- *eXplorer* tables (txt)
- Export to Excel
- Copy to clipboard [**Ctrl+c**]

Values and codes editing

- Manual editing
- Search and replace
- *Value* Tab menus
 - character case changes
 - remove leading characters or dots
 - handle leading codes
 - clipboard support [**Ctrl+c/v**]
 - ...
- Adjust classifications with external classification file (pxc)
- Code creation

Metadata editing

- Multilingual keywords
 - single line keywords
 - multiline keywords
- Menu selections
- Date fields
- Special input fields (e.g. *TIMEVAL*)
- Fill empty fields with default keywords [**Ctrl+k**]
- **Import** keywords from another file or open table [**Ctrl+i**]
- *Meta part* Tab for checking the file format

Data editing

- Cell by cell editing (may also be prohibited)
- Paste from clipboard [**Ctrl+v**]
- Search and replace [**Ctrl+f**]
- Aggregate variables
- Mask or delete values
- Data anonymizing
- *Calculator*
 - make new values
 - change existing values

Special functions

- Merge two or more tables into a new table [**Ctrl+j**]
- Add a new variable or new values
- Combine variables
- Remove values
- Check database
- Standardize database
- Convert database
- Database report
- Compare variables from two tables
- Clone (duplicate) tables [**Ctrl+d**, **Shift+Ctrl+d**]
- Split the table into partial tables along one variable

Multilingualism

- The multilingual tables will be manipulated with the keywords *LANGUAGE* and *LANGUAGES*
- *LANGUAGE* is the main language, the others are extra languages
 - almost all keywords have language based variants in the px file
 - the language code of the extra language must be given after the keyword in brackets (*KEYWORD*[fi])
- The current language can be changed with the table or metadata editor language menus
- PX-Edit will fill in the **missing** keyword contents for each extra language from the corresponding main language keyword

There are many ways to add new languages

- Add a new language to a monolingual table with the **Add** function in the *LANGUAGES* input field
 - other metadata input manually
- Join two or more tables with different languages
 - the tables should be uniform
- Import metadata from another table with different languages
 - the import window will then have a separate language menu

PX-Job: PX-Edit command line interface (batch)

PX-Job {job} [**in**] {out} {err} {copy} {meta} {set} {path} {log} {-} {!}

job	job type (default: px)
in	source directory or file (mandatory)
out	output directory or file (default: source files)
err	directory for erroneous files
copy	directory for source files
meta	metadata file or directory
set	settings file
path	common directory path
log	log file
-	options
!	switches

PX-Edit history

- 0.9 1999 Basic structure (Kimmo *Linna*)
- 1.0 2003 GUI, memory handling, variable compare
- 1.1 2003 Data anonymizing, structured files, sparse matrices
- 2.0 2003 Batch, StatXDF, keyword fetching, cell notes
- 2.1 2004 Huge tables, *TIMEVAL*, New table, font charsets
- 2.2 2004 Batch templates, classification adjusting, PX-Tool
- 2.3 2005 System Tray, aggregations, formulae, table mask
- 2.4 2006 CoSSI/XML, *DATASYMBOL**, Excel comments
- 3.0 2011 Multilingualism, Unicode, database standardization
- 3.1 2013 PX-Job, *simplePXsql*, database conversions
- 3.2 2015 User levels, metadata injection, reporting, variable combining
- 3.3 2017 Data Report, Join/Merge, PX-Job enhancements

New versions are available at:

www.stat.fi/tup/pcaxis/lataus_tyokalut_en.html