

PX-Edit 3.3

Structural tables

Principles

Principles

- The structural table is the way to move information from other applications to PX-Edit (and further to px format)
- They include only the **necessary** information:
 - table title
 - variable names
 - value names for each variable
 - data matrix
- Most statistical tables can easily be converted to structural tables
- The table may be a spreadsheet or a well structured text file

Good to know

- PX-Edit also accepts
 - asymmetrical tables
(i.e. tables with missing information or tables with disorder)
 - *Fill Item* is used for the missing information
 - redundant tables
 - the user will be warned about duplicated values
 - the **last value** will be used
- Possible problems
 - there are notes below the table (especially in Excel sheets)
 - title cells that should be empty are not
 - the tabulation recognition algorithm is based on empty areas

Short description

- The first line shows the **table title**
 - the **language code** may be given beside the title
- The next lines contain the **column variable names** in the first column
- The next line contains the **row variable names** side by side
 - column or even row variables may be missing
- **Row** variable **value names** are in their own columns
- **Column** variable **value names** are in their own rows
- Data are represented as a table, every figure positioned by its headers

Layouts

Example 1

Married couples with children		table title					
Year		1992	... column variables				
Age		-19	20-24	25-29	30-34	35-39 ...	
Area	Sex						
Akaa	Males	1	20	108	238	379	...
	Females	6	33	191	305	406	
Alajärvi	Males	0	21	107	237	291	
	Females	3	45	188	272	304	
Alavieska	Males	0	4	27	45	75	
	Females	1	7	45	66	69	
Alavus	Males	0	14	102	227	345	
	Females	2	51	170	273	347	
Asikkala	Males	0	8	42	111	174	
	Females	0	13	74	150	229	
Askola	Males	0	2	26	94	99	
	Females	0	7	53	106	131	
Aura	Males	0	3	32	52	91	
...					
row variables		data part					

Shortcuts

- Variable **codes** may be given in their own rows or columns (i.e. the general structural table)
 - they share the **same** variable name
 - the code must be given first
 - table indexing is based on value-code combinations
- There may be empty columns (not in the text files) and single empty rows
 - when reading Excel tables or tables from clipboard two successive empty rows stop the table reading, because they can be used to separate the table from its metadata part
- The second cell after the title may contain the **language code**
 - tables are always monolingual

Example 2

Married couples with children			table title			
Year			1991	...	column variables	
Year			1991	...		
Age			-19	20-24	25-29	30-34 ...
Area	Area	Sex				
020	Akaa	Males	1	20	108	238 ...
		Females	6	33	191	305
005	Alajärvi	Males	0	21	107	237
		Females	3	45	188	272
009	Alavieska	Males	0	4	27	45
		Females	1	7	45	66
010	Alavus	Males	0	14	102	227
		Females	2	51	170	273
016	Asikkala	Males	0	8	42	111
		Females	0	13	74	150
018	Askola	Males	0	2	26	94
...			
row variables			data part			

Example 2 in Excel with language code

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Married	en														
2	Year			1992												
3	Year			1992												
4	Age			- 19	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64	65 - 69	70 - 74	75 -
5	Area	Area	Sex													
6	020	Akaa	Males	1	20	108	238	379	452	329	186	110	73	53	27	15
7			Females	6	33	191	305	406	429	266	146	95	58	29	21	6
8	005	Alajärvi	Males	0	21	107	237	291	363	242	160	128	92	60	34	25
9			Females	3	45	188	272	304	332	229	140	91	79	51	18	8
10	009	Alavieska	Males	0	4	27	45	75	76	50	32	30	27	16	6	5
11			Females	1	7	45	66	69	68	49	24	30	23	5	7	3
12	010	Alavus	Males	0	14	102	227	345	374	271	215	130	119	51	30	31
13			Females	2	51	170	273	347	339	272	172	117	96	42	14	14
14	016	Asikkala	Males	0	8	42	111	174	276	209	145	65	58	17	16	12
15			Females	0	13	74	150	229	241	208	97	61	26	13	14	7
16	018	Askola	Males	0	2	26	94	99	136	108	58	40	28	13	10	5
17			Females	0	7	53	106	131	120	84	53	29	16	11	6	3
18	019	Aura	Males	0	3	32	52	91	93	86	30	18	11	11	5	4
19			Females	0	10	50	75	95	94	47	29	14	12	2	5	3
20	035	Brändö	Males	0	0	0	8	8	3	12	6	9	3	2	1	2
21			Females	0	0	7	4	5	13	9	7	3	3	1	1	1
22	043	Eckerö	Males	0	0	10	7	19	19	13	9	7	5	4	2	3

Layout: one variable table

Population by Region	
Region	
regions	data column

Layouts: two variable tables

Population by Region and Year		
Region	Year	
regions	years	data column

Population by Region and Year	
Year	years
Region	
regions	data table (matrix)

Layouts: three variable tables

Population by Region, Year and Age group			
Region	Year	Age group	
regions	years	age groups	data column

Population by Region, Year and Age group		
Age group		age groups
Region	Year	
regions	years	data table

Population by Region, Year and Age group	
Year	years
Age group	age groups
Region	
regions	data table

Special cases

Keyword block

- The Excel tables may have a separate keyword setting block under the table separated by at least two empty rows
- The format of the keyword table is similar to the [control csv](#) for PX-Job

	A	B	C	D	E	F	G	H
1	Establishment and personnel							
2	Industry			C		D		
3	Industry			C Mining and quarrying		D Manufacturing		
4	Data			Establishment	Personnel	Establishment	Personnel	
5	Year	Region	Region					
6	1995	0	Whole country	1271	3528	26009	380791	
7		4	Alahärmä	2	0	42	458	
8		5	Alajärvi	6	6	84	548	
9	1996	0	Whole country	1409	3961	27442	386001	
10		4	Alahärmä	2	0	41	406	
11		5	Alajärvi	6	11	88	655	
12								
13								
14	UNITS	persons						
15	NOTE	table footnote	variable footnote for Industry					
16	variablename		Industry	Year				
17	VALUENOTE			value footnote for Year 1996				
18	valuetext			1996				
19								
20	CELLNOTE				cell note for year 1995, Alajärvi, establishment and all industries			
21	Industry			*				
22	Data			Establishment				
23	Year			1995				
24	Region			Alajärvi				
25								
26								

Sequential files

- **Mask control** can be used to transfer sequential files containing fixed-size records obtained from e.g. some legacy systems
 - all the information is given in columns (i.e. sequential records)
 - the record mask is given in the **second** row
 - the record mask uses the same character for one record
 - mask spaces will skip the corresponding columns
 - the data part is the last column, it needs only one marker (because the data part length usually varies)
 - the variable names and codes have to be changed afterwards

Column mask example

Statistical table	table title
aaabbbbc d	column mask
00019901;4847	<i>code columns, data column at the end</i>
00019902;583676	
00019903;488	
00019904;113653	
00019905;93	
00019906;1443	
00019907;279	
00019908;0	
00019911;4823	
00019912;586586	
00019913;489	
00019914;123399	
00019915;93	
00019916;1427	
00019917;279	
...	

Cell-specific information

***DATANOTECELL* input**

- It is possible to attach the *DATANOTECELL* text strings with the corresponding figures or dot codes in the structured tables
- The possible strings must be given in the main settings file:
[System]
datanotecells=
 - comma separated list, may have leading or trailing blanks, may be enclosed with double quotes
- Strings will be recognised without any spaces, but they will be added to the table metadata with the same formatting as in the settings file

***DATANOTECELL* example**

- If there is the following line in the settings file:

```
datanotecells=*, **, "    ***"
```

then the input table may contain such figures or cells as:

123** 123 with '**' (quotes are for visibility only)

3.456 * 3.456 with '*'

..*** dot code with '***'

Cell specific footnotes

- The *Separator characters* for different types of footnotes is an alternative to *datanotecells*
- If one uses separator characters that are not the defaults, they **must** be configured in the main settings file:

```
[System]
separator_attribute=      (default |)
separator_cellnote=      (default !)
separator_datanote=      (default #)
```

- Footnote strings may now be attached to the cell figure with the separator character

Separator examples

- Separators enable multiple footnotes for a single cell, even with dot codes
 - strings containing special characters (such as spaces, commas or separators) should be enclosed in quotes
- For example the following input will be recognised (the cell separator is ; and the other separators are coloured just for clarity)
`1;2|A,B;3.14!"Just a comment"#a);4;..|A,C; ...`
- NB: PX-Edit will copy the Excel cell comments to the corresponding PC-Axis keywords