

UK FHS
Historická sociologie, Řízení a supervize
(LS 2011)

Analýza kvantitativních dat I.

Úvod do SPSS (PSPP)

Jiří Šafr

[jiri.safr\(AT\)seznam.cz](mailto:jiri.safr(AT)seznam.cz)

Poslední aktualizace 10/3/2011

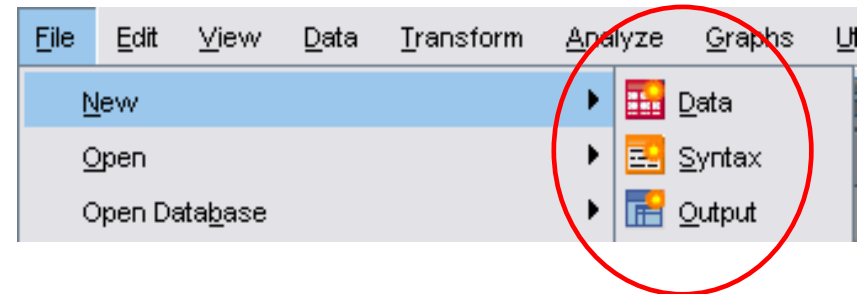
Vše je podrobně popsáno v souboru

Návod na statistický software PSPP, část 1. – úvod

http://metodykv.wz.cz/PSPP_navod_1_uvod.pdf

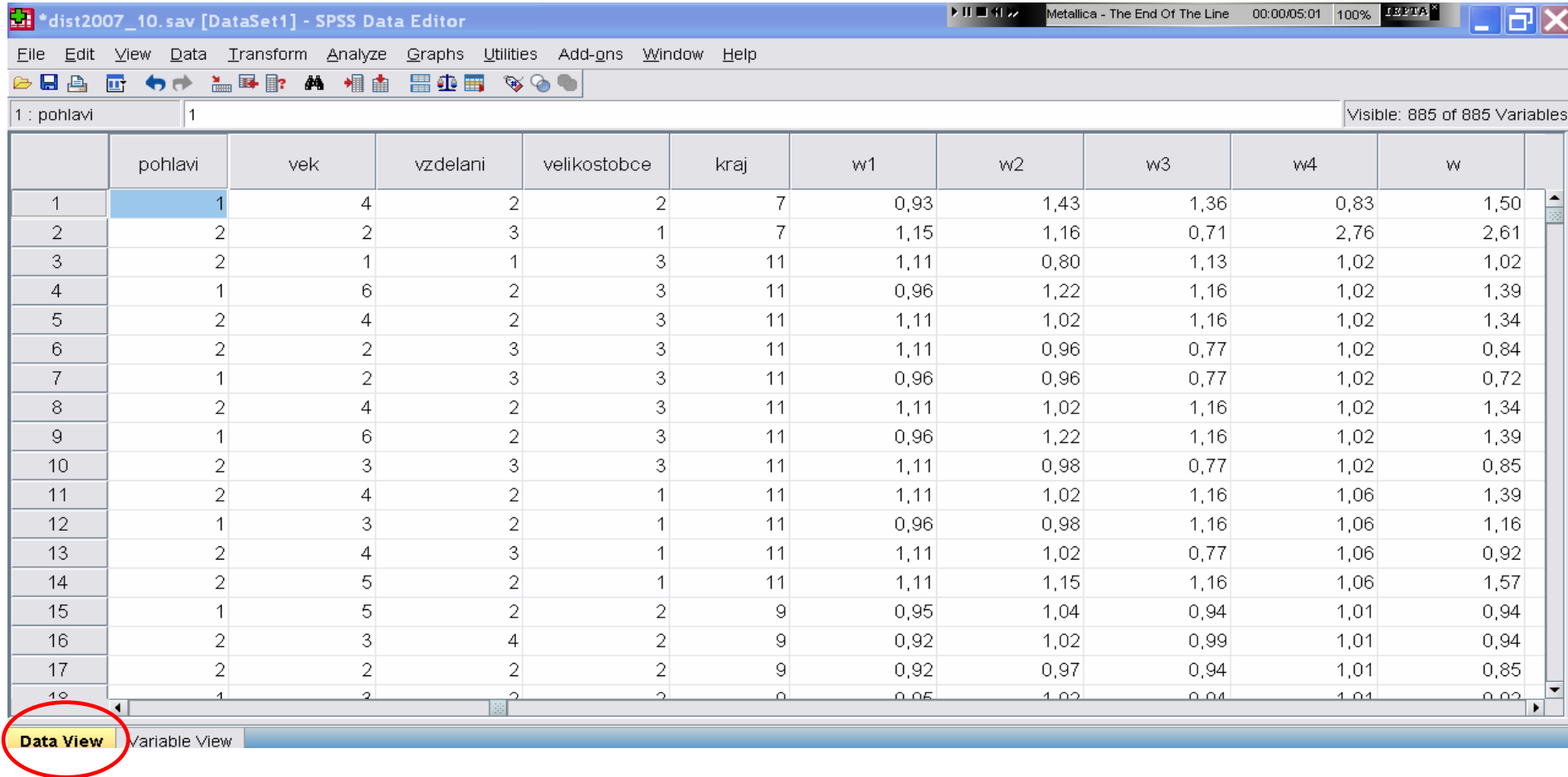
Tři typy oken v SPSS / PSPP

- **Data editor** (záložky **Variable View** a **Data View**)
 - správa datového souboru
 - soubory s koncovkou ***.sav**



- **Output** → výstupy analýz
 - soubory s koncovkou ***.spv** (od verze 15)
- **Syntax** → zadávání pomocí příkazového řádku
 - soubory s koncovkou ***.sps**

Data: Data Editor



The screenshot shows the SPSS Data Editor window with the following data table:

	pohlavi	vek	vzdelani	velikostobce	kraj	w1	w2	w3	w4	w	
1	1	4	2	2	7	0,93	1,43	1,36	0,83	1,50	
2	2	2	3	1	7	1,15	1,16	0,71	2,76	2,61	
3	2	1	1	3	11	1,11	0,80	1,13	1,02	1,02	
4	1	6	2	3	11	0,96	1,22	1,16	1,02	1,39	
5	2	4	2	3	11	1,11	1,02	1,16	1,02	1,34	
6	2	2	3	3	11	1,11	0,96	0,77	1,02	0,84	
7	1	2	3	3	11	0,96	0,96	0,77	1,02	0,72	
8	2	4	2	3	11	1,11	1,02	1,16	1,02	1,34	
9	1	6	2	3	11	0,96	1,22	1,16	1,02	1,39	
10	2	3	3	3	11	1,11	0,98	0,77	1,02	0,85	
11	2	4	2	1	11	1,11	1,02	1,16	1,06	1,39	
12	1	3	2	1	11	0,96	0,98	1,16	1,06	1,16	
13	2	4	3	1	11	1,11	1,02	0,77	1,06	0,92	
14	2	5	2	1	11	1,11	1,15	1,16	1,06	1,57	
15	1	5	2	2	9	0,95	1,04	0,94	1,01	0,94	
16	2	3	4	2	9	0,92	1,02	0,99	1,01	0,94	
17	2	2	2	2	9	0,92	0,97	0,94	1,01	0,85	
18	1	2	2	2	9	0,95	1,02	0,94	1,01	0,92	

The 'Data View' tab is highlighted with a red circle.

V řádcích jsou **případy** (respondenti), ve sloupcích **proměnné** (otázky)

→ v okně **Data Editor** záložka **Data View**

Správa proměnných

(labely, chybějící hodnoty - missingy, formát, ...)

ISSP2007_v2_1.sav [DataSet1] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
197	q29	Numeric	2	0	Považujete se za věřícího člověka?	{1, ano, jsem praktikující věřící...}	98, 99	3	Right	Ordinal
198	q30_a	Numeric	2	0	Jak je ve vašem životě důležitá/é: Práce	{1, velmi důležitá}...	98, 99	5	Right	Ordinal
199	q30_b	Numeric	2	0	Jak je ve vašem životě důležitá/é: Rodina	{1, velmi důležitá}...	98, 99	5	Right	Ordinal
200	q30_c	Numeric	2	0	Jak je ve vašem životě důležitá/é: Přátelé a známí	{1, velmi důležitá}...	98, 99	5	Right	Ordinal
201	q30_d	Numeric	2	0	Jak je ve vašem životě důležitá/é: Volný čas	{1, velmi důležitá}...	98, 99	5	Right	Ordinal
202	q30_e	Numeric	2	0	Jak je ve vašem životě důležitá/é: Politika	{1, velmi důležitá}...	98, 99	5	Right	Ordinal
203	q30_f	Numeric	2	0	Jak je ve vašem životě důležitá/é: Náboženství	{1, velmi důležitá}...	98, 99	5	Right	Ordinal
204	q31_a	Numeric	2	0	q31 - Cizincům zakázat politická činnost	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
205	q31_b	Numeric	2	0	q31 - Problémy by měl řešit stát	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
206	q31_c	Numeric	2	0	q31 - Jistota a blahobyt je důležitější než svoboda	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
207	q31_d	Numeric	2	0	q31 - Nejlepší je nevybočovat z řady	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
208	q31_e	Numeric	2	0	q31 - Rozdíly v příjmech by se měly zmenšit	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
209	q31_f	Numeric	2	0	q31 - Bylo by dobré, aby většina lidí věřila v Boha	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
210	q31_g	Numeric	2	0	q31 - Každý musí své problémy řešit sám	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
211	q31_h	Numeric	2	0	q31 - Život si musíme udělat příjemný, jak to jen jde	{1, rozhodně souhlasím}...	98, 99	5	Right	Ordinal
212	q32_a	Numeric	2	0	q32 - Člověk sotva co může změnit na průběhu vlastního života	{1, naprosto souhlasím}...	98, 99	5	Right	Ordinal

Data View **Variable View**

PASW Statistics Processor is ready

→ záložka **Variable View** v okně Data Editor

Output: výstupy (analýz, úpravy dat)

*Output_ukazka.spv [Document5] - PASW Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Add-ons Window Help

Output
Custom Tables
Title
Notes
Active Dataset
Table 1
Bar Chart from Pivot Table

Custom Tables

[DataSet1] D:\!FHS_HS\web\ISSP2007_v2_1.sav

<i>Jak je ve vašem životě důležitá:</i>	1 velmi důležitá	2 dost důležitá	3 ne příliš důležitá	4 vůbec není důležitá	Total	
q30_a Práce	42,6%	39,1%	9,3%	9,1%	100%	1149
q30_b Rodina	77,4%	18,9%	2,6%	1,0%	100%	1215
q30_c Přátelé a známí	39,1%	51,5%	8,7%	,7%	100%	1221
q30_d Volný čas	30,9%	52,8%	14,2%	2,1%	100%	1216
q30_e Politika	2,8%	11,8%	38,4%	46,9%	100%	1202
q30_f Náboženství	6,1%	11,1%	22,1%	60,7%	100%	1178

Jak je ve vašem životě důležitá (průměry):

Variable	Average Value
q30_b Rodina	~1.4
q30_c Přátelé a známí	~1.8
q30_a Práce	~1.9
q30_d Volný čas	~1.9
q30_e Politika	~3.3
q30_f Náboženství	~3.3

Úprava dat

(agregování, filtrování, rekódování, ...)

The screenshot shows the SPSS Data Editor interface. The 'Data' menu is open, displaying various options for data manipulation. The main window shows a list of variables with their properties.

Variable	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
okres	Nominal	8	0	Okres	{1, Benešo...	None	8	Right	Nominal
obec	Nominal	16	0	Obec	{1, Boskovi...	None	16	Right	Nominal
datum	Scale	8	0	datum rozhovoru	None	None	8	Right	Scale
délka	Scale	8	0	délka rozhovoru	None	None	8	Right	Scale
věk	Scale	10	0	věk respondenta	None	None	10	Right	Scale
VAR0	Scale	8	0	xxxxxxxxxxxxx...	None	None	8	Right	Scale
pohlaví	Nominal	8	0	Pohlaví respon...	{1, muž}...	None	8	Right	Nominal
vek	Ordinal	10	0	věk kategorizo...	{1, 18-24}...	None	10	Right	Ordinal
vzdela	Ordinal	10	0	vzdělání kateg...	{1, zš}...	None	10	Right	Ordinal
veliko	Ordinal	10	0	velikost sídla k...	{1, do 4999...	None	10	Right	Ordinal
kraj	Nominal	8	0	kraj	{1, Praha}...	None	8	Right	Nominal
w1	Scale	10	0		None	None	10	Right	Scale
w2	Scale	10	0		None	None	10	Right	Scale
w3	Scale	10	0		None	None	10	Right	Scale
w4	Scale	10	0		None	None	10	Right	Scale
w	Scale	10	0	váha - repre ČR	None	None	10	Right	Scale
Q25M	Scale	10	0		None	None	10	Right	Scale
filtr_Q25MI...	Numeric	10	0	Q25MISS=1 (...	{0, Not Sel...	None	10	Right	Scale
age	Numeric	8	0		None	None	10	Right	Scale
sex	Numeric	8	0	Sex of resp.	{1, Men}...	None	8	Right	Scale

→ příkazy v sekcích hlavního menu **Data** a **Transform**
(nebo přímé zadání pomocí příkazového řádku v Syntaxu)

Analýzy

The screenshot shows the SPSS Data Editor interface. The 'Analyze' menu is open, displaying a list of statistical procedures. The 'Descriptive Statistics' submenu is also open, showing options like 'Frequencies...', 'Descriptives...', 'Explore...', 'Crosstabs...', 'Ratio...', 'P-P Plots...', and 'Q-Q Plots...'. The main window displays a list of variables with their names, types, and measurement scales.

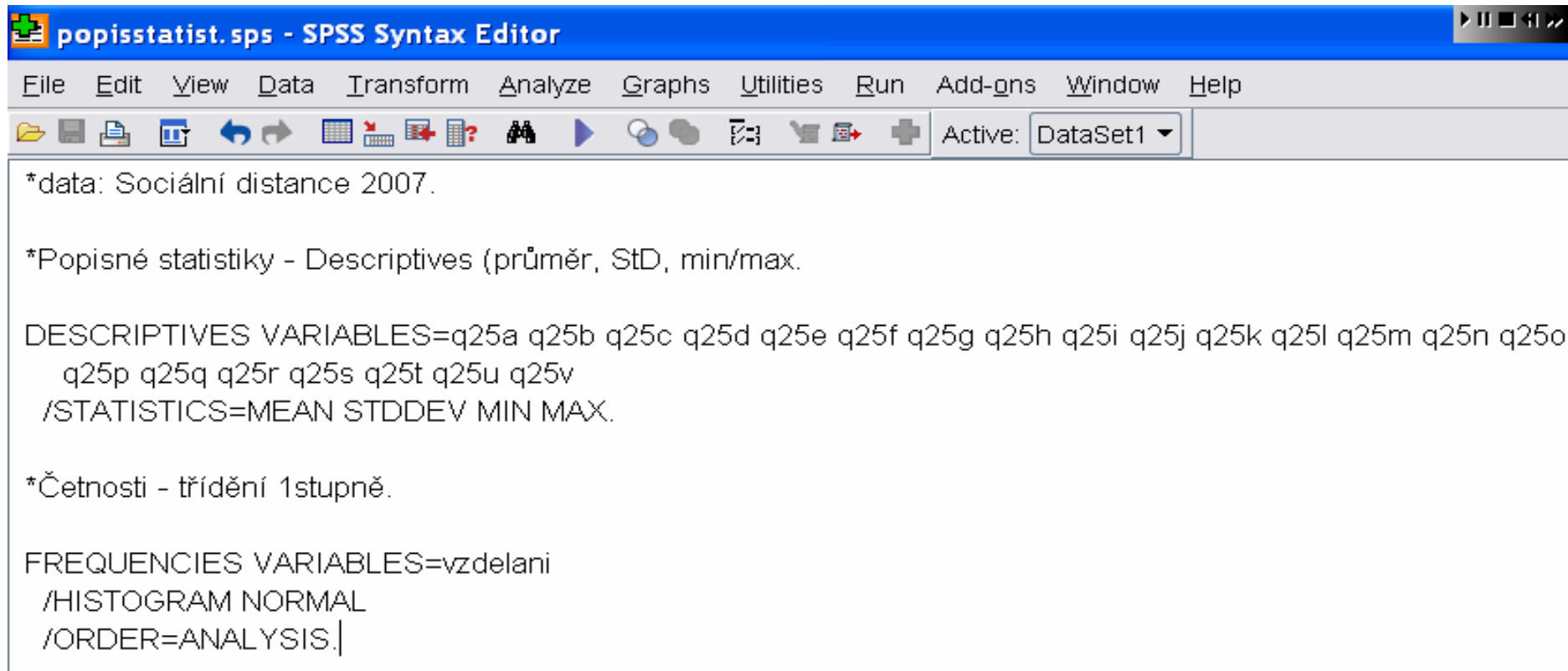
Name	Type	Measure
283 okres	Numeric	Nominal
284 obec	Numeric	Nominal
285 datum	Numeric	Scale
286 délka	Numeric	Scale
287 věk	Numeric	Scale
288 VAR00001	Numeric	Scale
289 pohlavi	Numeric	Nominal
290 vek	Numeric	Ordinal
291 vzdelani	Numeric	Ordinal
292 velikostobce	Numeric	Ordinal
293 kraj	Numeric	Nominal
294 w1	Numeric	Scale
295 w2	Numeric	Scale
296 w3	Numeric	Scale
297 w4	Numeric	Scale
298 w	Numeric	Scale
299 Q25MISS	Numeric	Scale
300 filtr_Q25MI...	Numeric	Scale
301 age	Numeric	Scale
302 sex	Numeric	Scale

→ příkazy v sekci hlavního menu **Analyze**

Popisné statistiky: **Descriptive statistics** → **Frequencies** / **Descriptives** / **Explore**
(nebo přímé zadání pomocí příkazového řádku v Syntaxu)

Syntax:

přímé zadání pomocí příkazového řádku



The screenshot shows the SPSS Syntax Editor window titled "popisstatist.sps - SPSS Syntax Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Run, Add-ons, Window, and Help. The toolbar contains various icons for file operations, navigation, and execution. The active dataset is "DataSet1". The script content is as follows:

```
*data: Sociální distance 2007.  
  
*Popisné statistiky - Descriptives (průměr, StD, min/max.  
  
DESCRIPTIVES VARIABLES=q25a q25b q25c q25d q25e q25f q25g q25h q25i q25j q25k q25l q25m q25n q25o  
q25p q25q q25r q25s q25t q25u q25v  
/STATISTICS=MEAN STDDEV MIN MAX.  
  
*Četnosti - třídění 1stupně.  
  
FREQUENCIES VARIABLES=vzdelani  
/HISTOGRAM NORMAL  
/ORDER=ANALYSIS.
```

Výhodou je, že sekvenci příkazů můžeme uložit a kdykoliv se k ní vrátit, případně měnit vstupní proměnné.

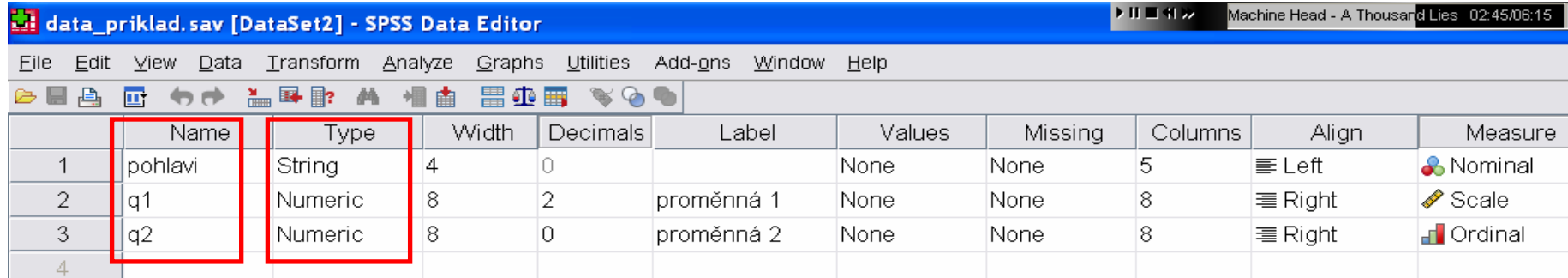
V případě úpravy dat poskytuje kontrolu operací a umožňuje úpravy sdílet.

Syntax lze u každé operace (analýzy či úpravy dat) vygenerovat při zadávání z menu pomocí tlačítka **Paste**

Vkládání dat v Data editoru

Variable View

Zde si nejprve připravíme datovou matici: **jména proměnných (Name)**, jejich **typ/formát (Type)** a případně labely

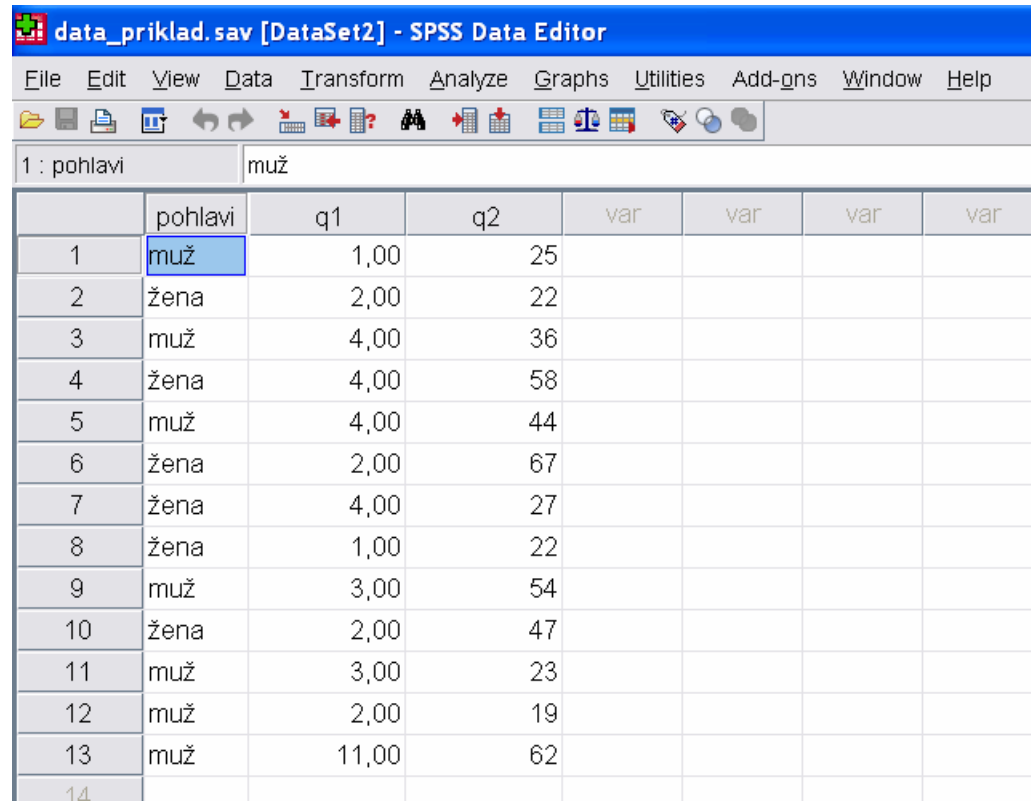


The screenshot shows the SPSS Data Editor interface in Variable View. The title bar reads 'data_priklad.sav [DataSet2] - SPSS Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations and data manipulation. The main table lists variables with columns for Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, and Measure. The 'Name' and 'Type' columns are highlighted with red boxes.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	pohlavi	String	4	0		None	None	5	Left	Nominal
2	q1	Numeric	8	2	proměnná 1	None	None	8	Right	Scale
3	q2	Numeric	8	0	proměnná 2	None	None	8	Right	Ordinal
4										

Data View

Zde pak ukládáme data – po řádcích (1 řádek = 1 respondent/ dotazník)

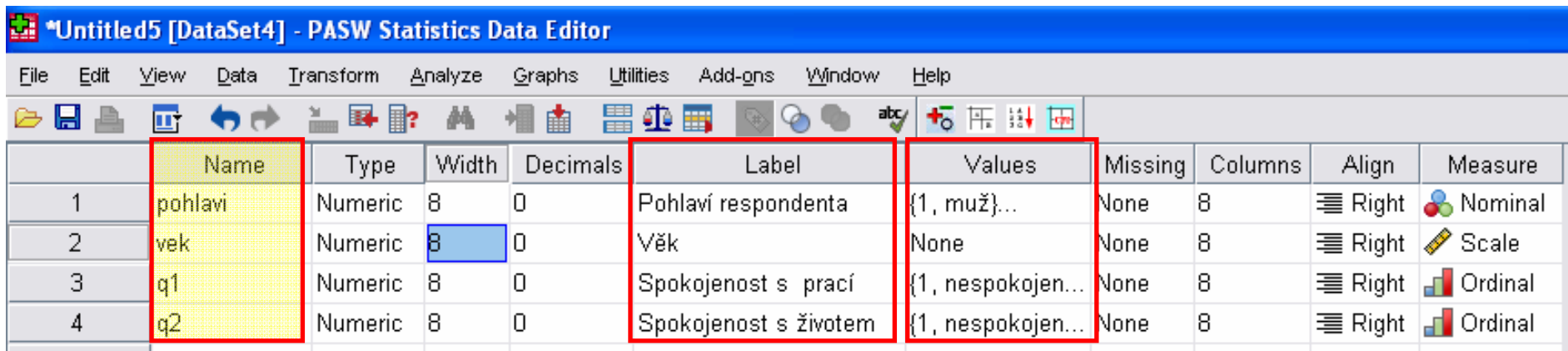


The screenshot shows the SPSS Data Editor interface in Data View. The title bar reads 'data_priklad.sav [DataSet2] - SPSS Data Editor'. The menu bar and toolbar are the same as in the Variable View. The main table shows data for 14 rows. The first row is highlighted, and the first column is labeled 'pohlavi'.

	pohlavi	q1	q2	var	var	var	var
1	muž	1,00	25				
2	žena	2,00	22				
3	muž	4,00	36				
4	žena	4,00	58				
5	muž	4,00	44				
6	žena	2,00	67				
7	žena	4,00	27				
8	žena	1,00	22				
9	muž	3,00	54				
10	žena	2,00	47				
11	muž	3,00	23				
12	muž	2,00	19				
13	muž	11,00	62				
14							

Tři informace o proměnných:

- **jméno proměnné** **VARIABLE NAME**
max. 8 znaků bez diakritiky, mezer; musí začínat písmenem (např. *pohlaví*)
- **popiska proměnné** **VARIABLE LABEL**
Delší popis názvu znaku, lze s diakritikou, např. (zkrácené) znění otázky v dotazníku, který se bude objevovat ve výstupech (např. *Pohlaví respondenta*)
- **popisky kategorií** **VALUE LABELS**
Popis kategorií znaku, které se budou objevovat ve výstupech (např. *1 = muž, 2 = žena*)



The screenshot shows the PASW Statistics Data Editor interface. The title bar reads '*Untitled5 [DataSet4] - PASW Statistics Data Editor'. The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Add-ons, Window, and Help. The toolbar contains various icons for file operations and data manipulation. Below the toolbar is a table with the following columns: Name, Type, Width, Decimals, Label, Values, Missing, Columns, Align, and Measure. The 'Name' column is highlighted in yellow, and the 'Label' and 'Values' columns are highlighted in red. The table contains four rows of variables:

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	pohlaví	Numeric	8	0	Pohlaví respondenta	{1, muž}...	None	8	Right	Nominal
2	vek	Numeric	8	0	Věk	None	None	8	Right	Scale
3	q1	Numeric	8	0	Spokojenost s prací	{1, nespokojen...}	None	8	Right	Ordinal
4	q2	Numeric	8	0	Spokojenost s životem	{1, nespokojen...}	None	8	Right	Ordinal

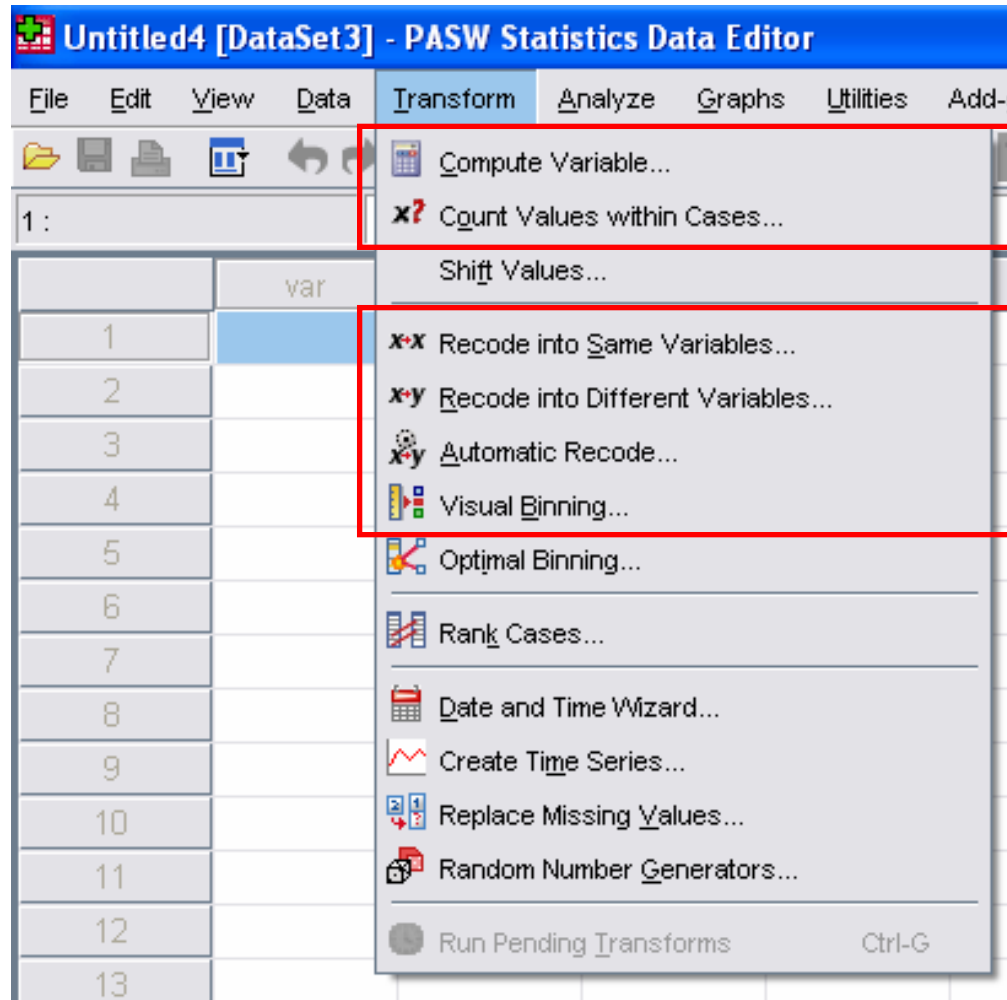
Transformace dat → *Transform*

- Výpočet/vytváření nových (syntetických) znaků
→ **COMPUTE** (nebo pro „načítání“ **COUNT**)
- Rekódování
→ **RECODE** (do stejné nebo nové proměnné)
- *Visual Binding* – nástroj pro snadné rekódování (dle percentilů apod.)

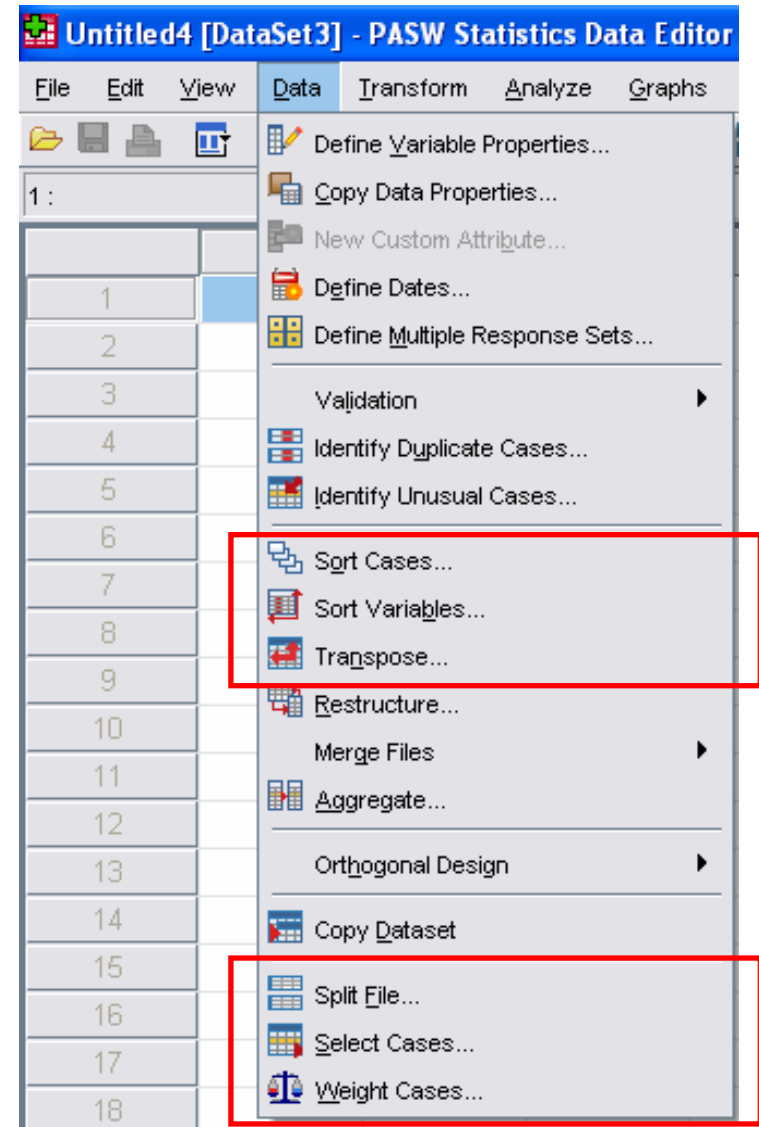
Úpravy dat a výstupů → *Data*

- Uspořádání případů → **SORT CASES**
- Rozdělení na podsoubory → **SPLIT FILE**
- Výběr případů (filtrování) → **SELECT CASES**
- Vážení → **WEIGHT CASES**
- Agregace (např. průměry pro skupiny)
→ **AGGREGATE**
- Spojování souborů dat / přidávání proměnných
→ **MERGE FILES**

Transformace dat → Transform

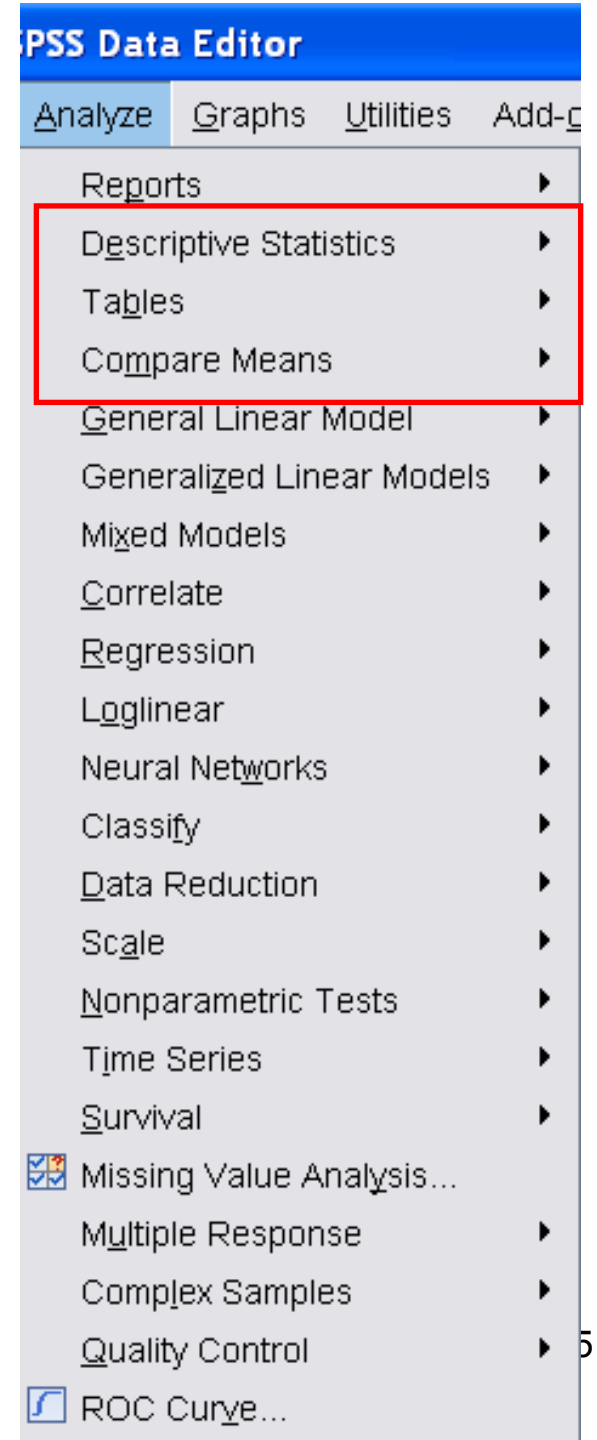


Úpravy dat → Data



Analýzy → *Analyze*

- **Descriptive statistics**
- **Tables**
- **Compare means**
- **Correlate**
- **Data Reduction**
- **Nonparametric Tests**
- **Missing Value Analysis**
- **Multiple Response**



Grafy → Graphs

***Output9 [Document9] - PASW Statistics Viewer**

File Edit View Data Transform Insert Format Analyze **Graphs** Utilities Add-ons Window Help

Chart Builder...
Graphboard Template Chooser...
Legacy Dialogs

Output
Log
Graph
Title
Notes
Bar of mean(q30_a, q30_b, ...)

GRAPH
/BAR (SIMPLE) = I

Graph

Category	Mean
Práce	1.9
Rodina	1.3
Přátelé a známí	1.7
Volný čas	1.9
Politika	3.3
Náboženství	1.3

MEAN(q30_e) MEA:

- Bar...
- 3-D Bar...
- Line...
- Area...
- Pie...
- High-Low...
- Boxplot...
- Error Bar...
- Population Pyramid...
- Scatter/Dot...
- Histogram...
- Interactive
 - Bar...
 - Dot...
 - Line...
 - Ribbon...
 - Drop-Line...
 - Area...
 - Pie...
 - Boxplot...
 - Error Bar...
 - Histogram...
 - Scatterplot...

Načtení dat ze syntaxu

(vhodné pro menší datové soubory nebo tabulární data)

*hypotetická data: hodnoty proměnných odděleny mezerou, stringové znaky v " ".

DATA LIST LIST

/ okres (A15) progr domduch vek obyv.

BEGIN DATA.

"Benešov" 1 3 52 40990

"Kladno" 0 4 61 67490

"M. Boleslav" 1 6 50 6099

"Příbram" 0 2 45 4816

"Dobříš" 1 1 49 7827

END DATA.

***Labely.**

VAR LAB progr "Program pro seniory".

VAR LAB domduch "Domovy důchodců – počet".

VAR LAB vek "Průměrný věk".

VAR LAB obyv "Počet obyvatel".

VAL LAB progr

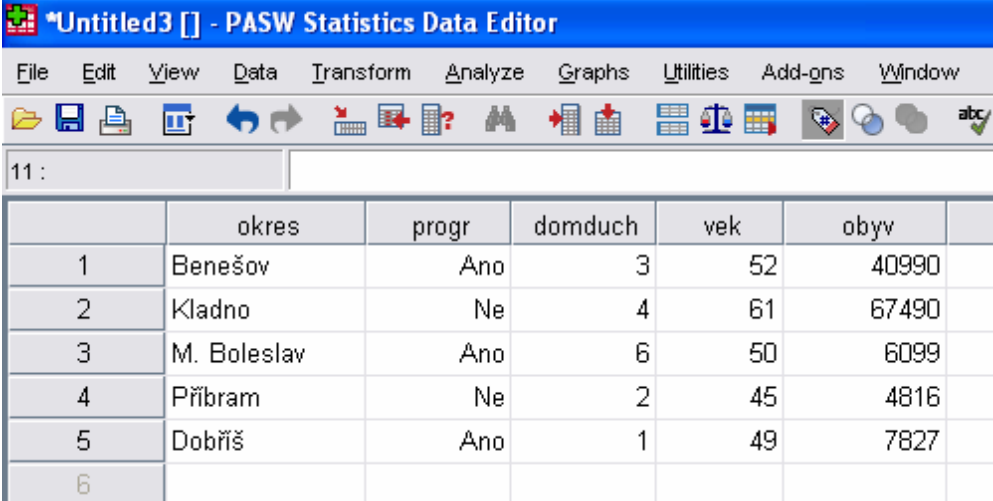
0 "Ne"

1 "Ano".

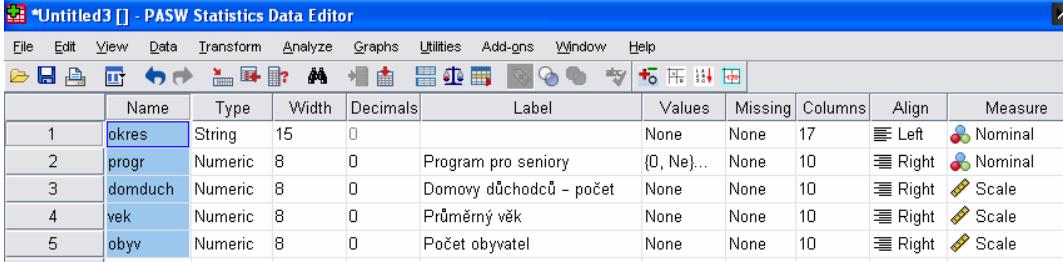
***formát čísel** (bez desetinného místa) a **typ proměnné.**

FORMATS progr domduch vek obyv (f8).

VARIABLE LEVEL domduch vek obyv (scale).



	okres	progr	domduch	vek	obyv
1	Benešov	Ano	3	52	40990
2	Kladno	Ne	4	61	67490
3	M. Boleslav	Ano	6	50	6099
4	Příbram	Ne	2	45	4816
5	Dobříš	Ano	1	49	7827
6					



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	okres	String	15	0		None	None	17	Left	Nominal
2	progr	Numeric	8	0	Program pro seniory	{0, Ne}..	None	10	Right	Nominal
3	domduch	Numeric	8	0	Domovy důchodců – počet	None	None	10	Right	Scale
4	vek	Numeric	8	0	Průměrný věk	None	None	10	Right	Scale
5	obyv	Numeric	8	0	Počet obyvatel	None	None	10	Right	Scale

Načtení dat z agregované existující kontingenční tabulky (→ vážení procenty)

věk	vzdělání	volil	
		1 nevolil	2 volil
1 <49	1 ZŠ+VY	138	92
	2 SŠ+VŠ	106	218
2 >50	1 ZŠ+VY	143	257
	2 SŠ+VŠ	56	175

Pozice pole v tabulce

věk	vzdělání	volil	
		0 nevolil	1 volil
1 <49	0 ZŠ+VY	111	112
	1 SŠ+VŠ	121	122
2 >50	0 ZŠ+VY	211	212
	1 SŠ+VŠ	221	222

****nacteni kontingencni tabulky aneb sekundarni analyza (ČR, ISSP 2007).
 DATA LIST LIST/vek vzdel volil freq.
 VAL LAB vzdel 1 "ZŠ+VY" 2 "SŠ+VŠ" /
 vek 1 "<49" 2 ">50" / volil 1
 "nevolil" 2 "volil".

BEGIN DATA

1 1 1 138

1 1 2 92

1 2 1 106

1 2 2 218

2 1 1 143

2 1 2 257

2 2 1 56

2 2 2 175

END DATA.

FORMATS vek vzdel volil freq (f8).

WEIGHT by freq.

CROSS vzdel by volil by vek.

CROSS vzdel by volil.

Syntax:

crosstab_data_input.sps