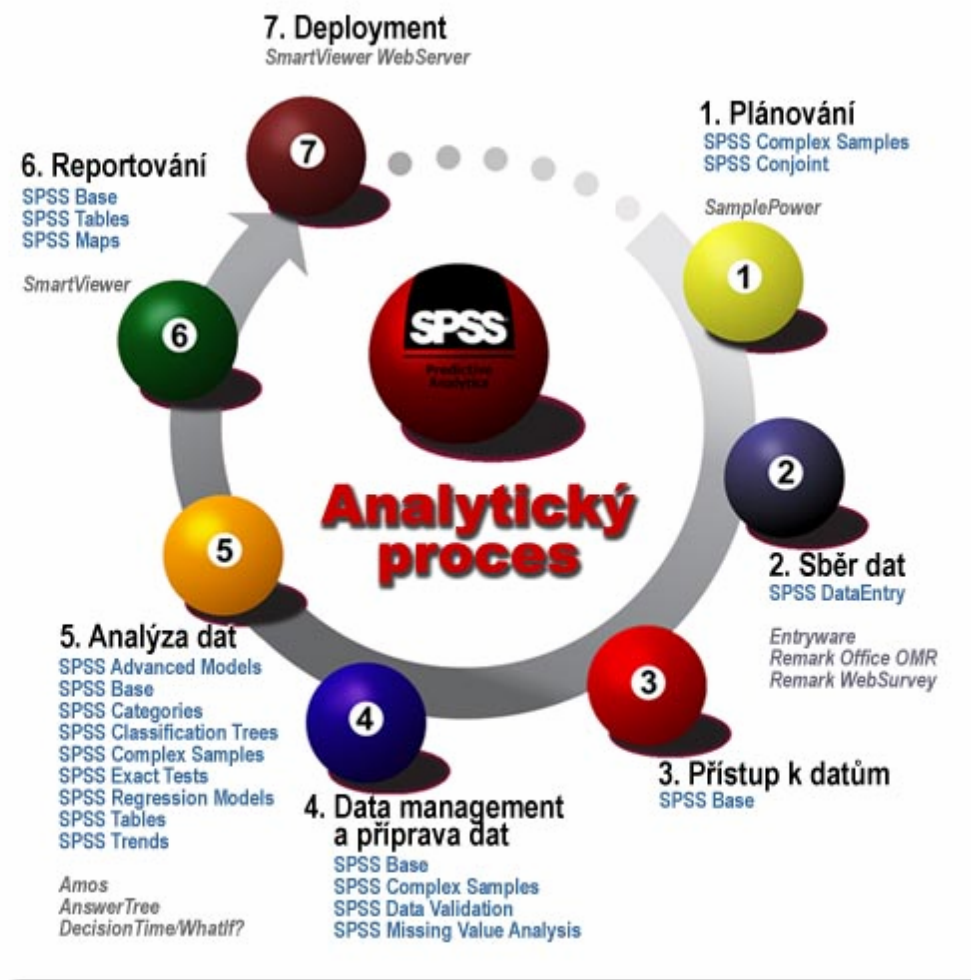


Uživatelské prostředí SPSS

Uživatelské prostředí SPSS

- SPSS je univerzální analytický systém určený pro osobní počítače
- Široké spektrum využití: věda a výzkum, marketing, personalistika, statistické analýzy, finanční analýzy, rozhodovací modely, analýzy časových řad, ...
- Software SPSS pokrývá celý analytický proces, proto je možné kombinovat jednotlivé moduly podle individuálních potřeb

Struktura SW SPSS



Přenos dat do SPSS

- Data lze do prostředí SPSS transportovat z mnoha různých zdrojů (jako např. Excel, Access, Lotus, dBase apod.)
- Základní proceduru pro přenos dat do SPSS spustíme:
 - File
 - Open Database
 - New Query
- Po počáteční volbě výchozího zdroje dat dokončíme podle instrukcí průvodce Database Wizard přenos dat, nakonec uložíme soubor ve formátu SPSS

Untitled1 [DataSet0] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

New
 Open
Open Database
 Read Text Data...

Close Ctrl+F4
 Save Ctrl+S
 Save As...
 Save All Data
 Export to Database...
 Mark File Read Only

Rename Dataset...
 Display Data File Information
 Cache Data...
 Stop Processor Ctrl+.
 Switch Server...
 Connect To Repository...

Print Preview
 Print... Ctrl+P

Recently Used Data
 Recently Used Files

Exit

New Query ...
 Edit Query ...
 Run Query ...

	var	var	var	var
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				

Otevření souboru v SPSS

- Otevření již existujícího souboru:

File

Open

Data

- Vytvoření nového souboru:

File

New

Data

Data Editor

- Data Editor umožňuje pohodlně vytvářet a editovat soubory v SPSS, otevírá se automaticky po spuštění programu
- Je koncipován do dvou separátních částí, tím poskytuje komplexní pohled na data:
 - Data View – lze prohlížet označení proměnných a jejich aktuální hodnoty
 - Variable View – obsahuje definice proměnných a jejich vlastnosti
- V obou oknech lze průběžně přidávat, měnit a mazat data

Variable View

- Pro všechny proměnné v souboru je třeba předem definovat následující atributy:
 - Name – jméno proměnné, musí být originální, v rámci souboru se nesmí opakovat
 - Type – typ proměnné (číslo, datum a čas, měnové jednotky, vlastní proměnná apod.)
 - Width – zobrazovaná šířka sloupce
 - Decimals – počet desetinných míst
 - Label - popisky

Variable view - pokračování

- Values – popisky kódů, číselníky; definujeme tehdy, když pro popis nenumerických kategorií chceme použít numerický kód (např. žena x muž, 1 x 2)
- Missing Values – chybějící hodnoty
- Align – zarovnání (standardně nastaveno vpravo pro numerické proměnné a vlevo pro proměnné typu string)
- Measure – typ proměnné (nominální, ordinální, škála)

Variable view - pokračování

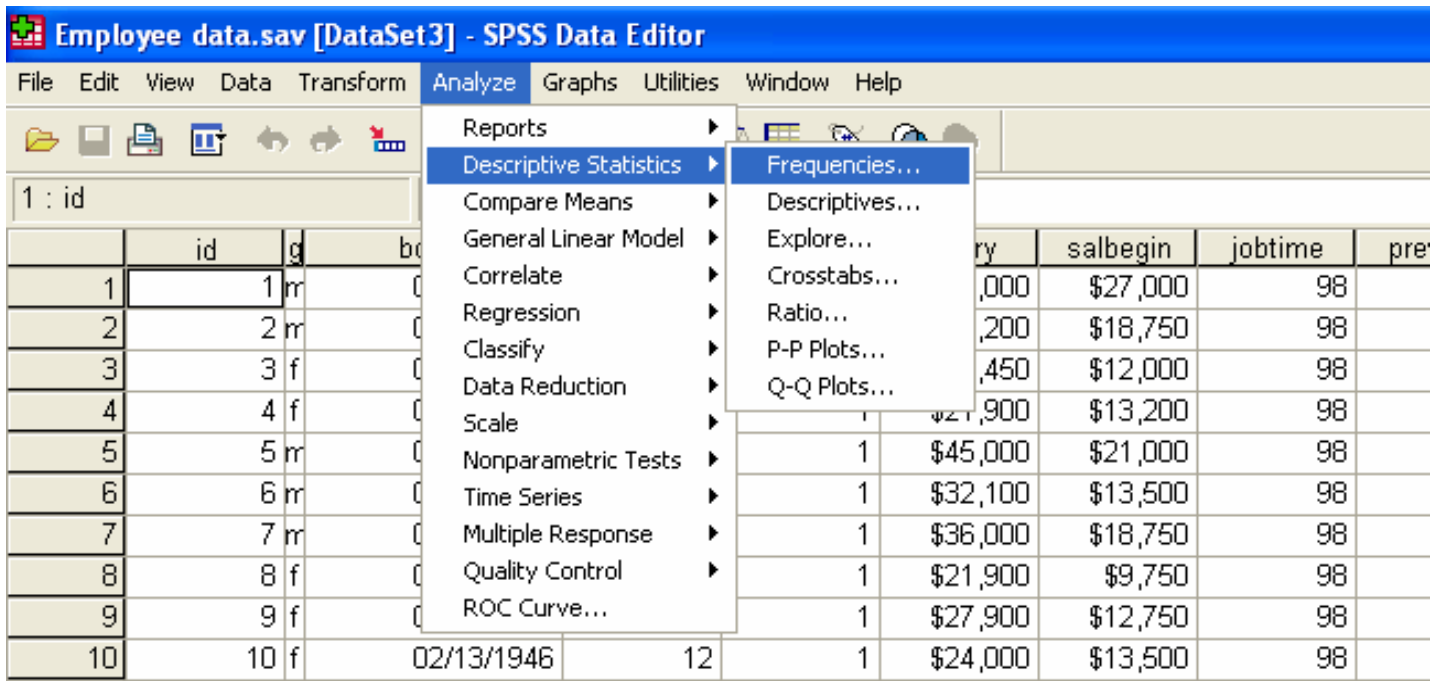
- Definování jednotlivých proměnných je třeba věnovat značnou pozornost, jsou východiskem pro následující analýzy
- V prostředí SPSS se nedoporučuje pořizovat vlastní data, je lepší transportovat již upravená data z některého programu, který je pro tento úkol speciálně určen

Analýzy

- Poté, co jsou definovány proměnné a datový list je naplněn, je možno začít soubor analyzovat
- Výsledky každé analytické procedury se zobrazují do zvláštního souboru s obecným názvem „Output“, řadí se automaticky za sebou ve sledu, v jakém jsou prováděny
- Výsledky analýz nejsou ukládány automaticky, aby nedošlo ke ztrátě dat, je nutné je průběžně ukládat
- V nižších verzích SPSS není možné mít otevřeno více souborů najednou, lze pracovat vždy pouze s jedním

Základní statistické procedury – Descriptive statistics

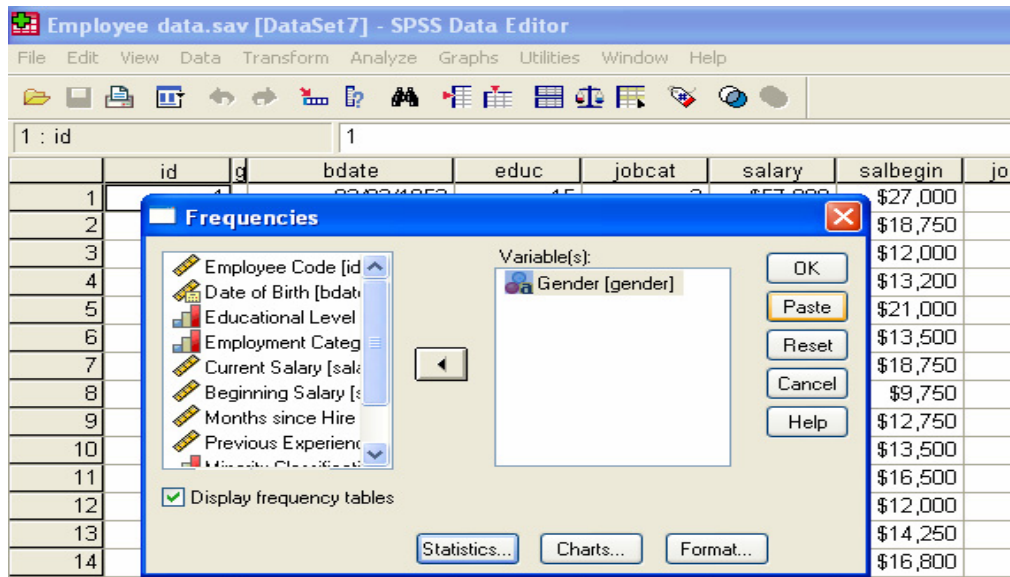
- **Frequencies** – tabulky četností
 - používáme pouze pro kategorizované proměnné



The screenshot shows the SPSS Data Editor interface for a file named 'Employee data.sav [DataSet3]'. The 'Analyze' menu is open, and the 'Descriptive Statistics' submenu is also open, with 'Frequencies...' selected. The main data grid is visible in the background, showing columns for 'id', 'g', 'b', 'ry', 'salbegin', 'jobtime', and 'pre'.

	id	g	b	ry	salbegin	jobtime	pre	
1	1	m	0	,000	\$27,000	98		
2	2	m	0	,200	\$18,750	98		
3	3	f	0	,450	\$12,000	98		
4	4	f	0	,900	\$13,200	98		
5	5	m	0	1	\$45,000	\$21,000	98	
6	6	m	0	1	\$32,100	\$13,500	98	
7	7	m	0	1	\$36,000	\$18,750	98	
8	8	f	0	1	\$21,900	\$9,750	98	
9	9	f	0	1	\$27,900	\$12,750	98	
10	10	f	02/13/1946	12	1	\$24,000	\$13,500	98

Frequencies



- Do okna Variables zadáme jednu či více proměnných, které chceme analyzovat

- Analýzu lze doplnit také o následující statistiky:

The screenshot shows the SPSS Data Editor interface with the 'Frequencies: Statistics' dialog box open. The dialog box is configured with the following options:

- Percentile Values:**
 - Quartiles
 - Cut points for: 10 equal groups
 - Percentile(s):
- Central Tendency:**
 - Mean
 - Median
 - Mode
 - Sum
- Dispersion:**
 - Std. deviation
 - Variance
 - Range
 - Minimum
 - Maximum
 - S.E. mean
- Distribution:**
 - Values are group midpoints
 - Skewness
 - Kurtosis

The 'Continue' button is highlighted in the dialog box. The background data table shows columns: id, g, bdate, educ, jobcat, salary, salbegin, job.

- Součástí nabídky jsou také grafy:

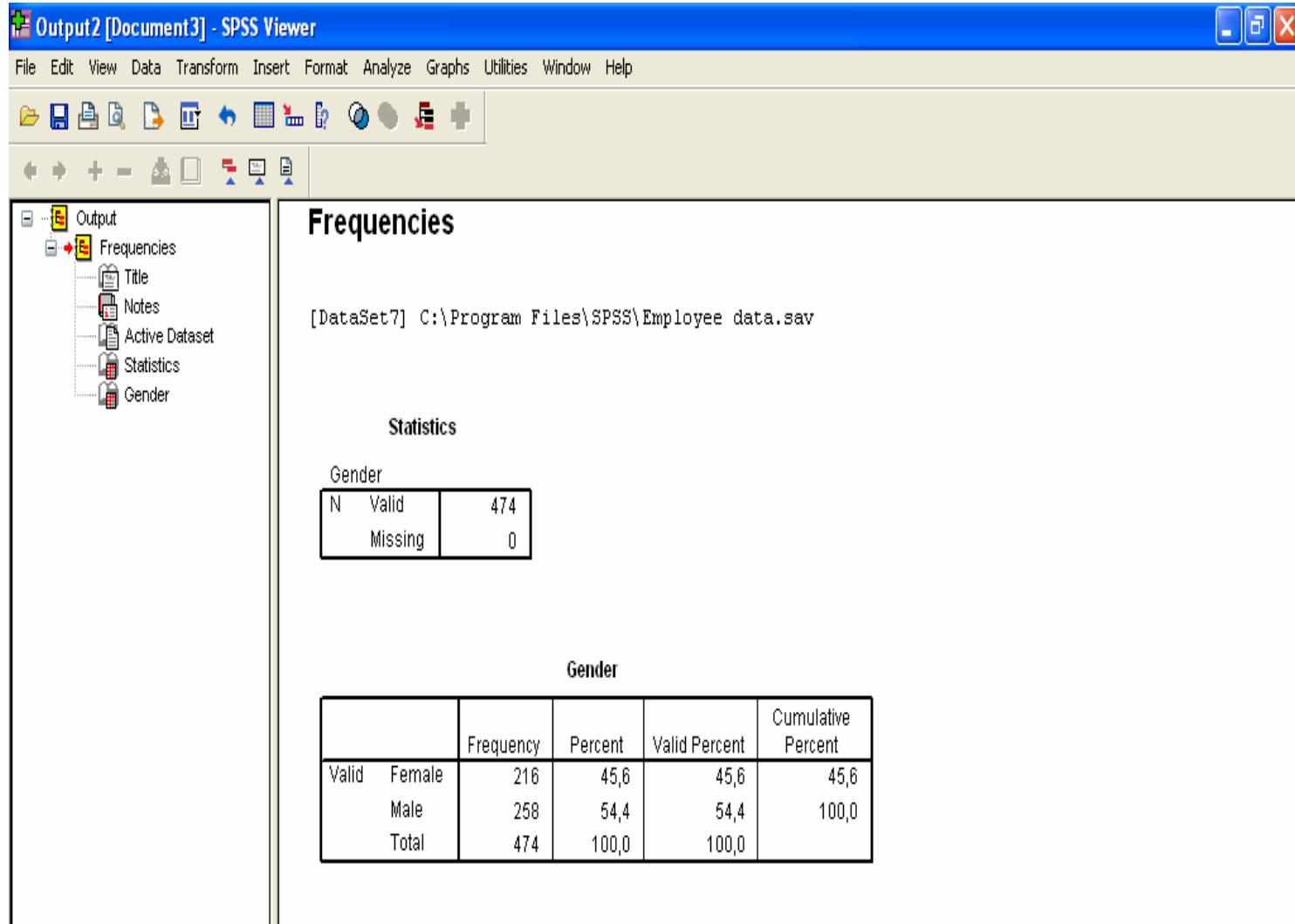
The screenshot shows the SPSS Data Editor interface with the 'Frequencies' dialog box open. The 'Charts' sub-dialog is also open, showing the following settings:

- Chart Type:** Bar charts (selected), None, Pie charts, Histograms: (unchecked)
 - With normal curve (unchecked)
- Chart Values:** Frequencies (selected), Percentages (unchecked)

The background data table is partially visible, showing columns for id, gender, bdate, educ, jobcat, salary, salbegin, jobtime, and previous salary.

	id	g	bdate	educ	jobcat	salary	salbegin	jobtime	pre
1							\$27,000	98	
2							\$18,750	98	
3							\$12,000	98	
4							\$13,200	98	
5							\$21,000	98	
6							\$13,500	98	
7							\$18,750	98	
8							\$9,750	98	
9							\$12,750	98	
10							\$13,500	98	
11							\$16,500	98	
12							\$12,000	98	
13							\$14,250	98	
14							\$16,800	98	
15	15	m				\$27,300	\$13,500	97	
16	16	m				\$40,800	\$15,000	97	
17	17	m				\$46,000	\$14,250	97	
18	18	m	03/20/1956	16	3	\$103,750	\$27,510	97	

- Výsledek nalezneme ve výstupním souboru:



The screenshot shows the SPSS Viewer interface. The title bar reads "Output2 [Document3] - SPSS Viewer". The menu bar includes File, Edit, View, Data, Transform, Insert, Format, Analyze, Graphs, Utilities, Window, and Help. The toolbar contains various icons for file operations and analysis. The left pane shows a tree view of the output, with "Output" expanded to show "Frequencies", "Title", "Notes", "Active Dataset", "Statistics", and "Gender". The main pane displays the "Frequencies" output for the variable "Gender".

Frequencies

[DataSet7] C:\Program Files\SPSS\Employee data.sav

Statistics

Gender

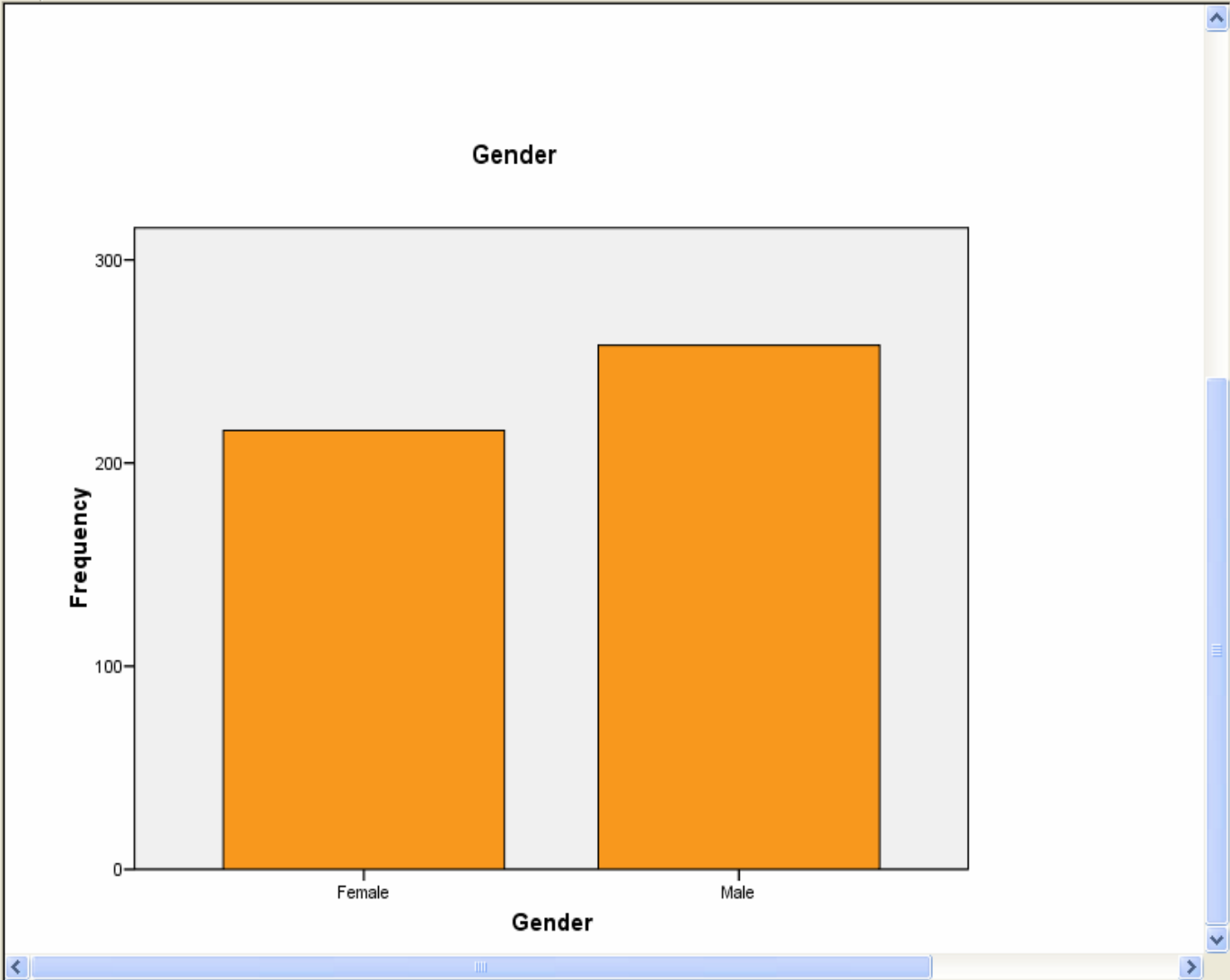
N	Valid	474
	Missing	0

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	216	45,6	45,6	45,6
	Male	258	54,4	54,4	100,0
	Total	474	100,0	100,0	



- Output
 - Frequencies
 - Title
 - Notes
 - Active Dataset
 - Statistics
 - Gender
 - Bar Chart



Descriptives

- Popisné statistiky, které se využívají při popisu základních vlastností číselných (spojitých) proměnných
- Proceduru lze také využít při ukládání bezrozměrných standardizovaných (aritmetický průměr je roven 1 a rozptyl 0) proměnných (z-skóry)

- Po vyvolání dialogu vložíme analyzovanou proměnnou do okna Variables

Employee data.sav [DataSet7] - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : id 1

Descriptives

Variable(s):
Current Salary [salary]

Save standardized values as variables

OK
Paste
Reset
Cancel
Help
Options...

						salbegin	jobtime	pre
1						\$27,000	98	
2						\$18,750	98	
3						\$12,000	98	
4						\$13,200	98	
5						\$21,000	98	
6						\$13,500	98	
7						\$18,750	98	
8						\$9,750	98	
9						\$12,750	98	
10						\$13,500	98	
11	11	f	02/07/1950	16	1	\$30,300	\$16,500	98
12	12	m	01/11/1966	8	1	\$28,350	\$12,000	98
13	13	m	07/17/1960	15	1	\$27,750	\$14,250	98
14	14	f	02/26/1949	15	1	\$35,100	\$16,800	98

- V nabídce Options vybíráme požadované statistiky:

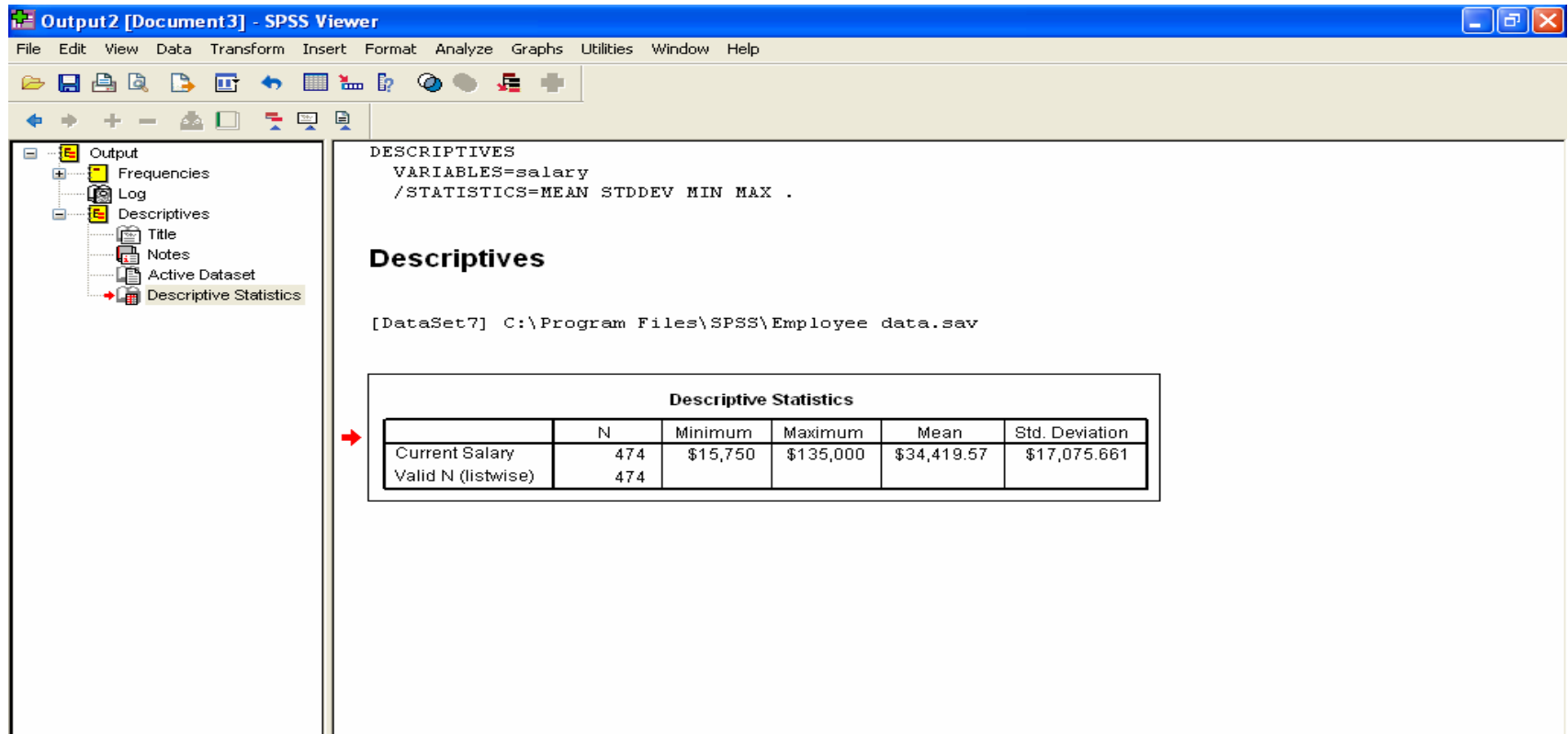
The screenshot shows the SPSS Data Editor window titled "Employee data.sav [DataSet7] - SPSS Data Editor". The main window displays a data table with columns "id", "salbegin", "jobtime", and "pre". A "Descriptives" dialog box is open, and its "Options" sub-dialog box is also open, showing the following settings:

- Mean
- Sum
- Dispersion**
 - Std. deviation
 - Minimum
 - Variance
 - Maximum
 - Range
 - S.E. mean
- Distribution**
 - Kurtosis
 - Skewness
- Display Order**
 - Variable list
 - Alphabetic
 - Ascending means
 - Descending means

The background data table is as follows:

id	salbegin	jobtime	pre
1	\$27,000	98	
2	\$18,750	98	
3	\$12,000	98	
4	\$13,200	98	
5	\$21,000	98	
6	\$13,500	98	
7	\$18,750	98	
8	\$9,750	98	
9	\$12,750	98	
10	\$13,500	98	
11	\$30,300	98	
12	\$28,350	98	
13	\$27,750	98	
14	\$35,100	98	
15	\$27,300	97	
16	\$40,800	97	
17	\$46,000	97	
18	\$103,750	97	

- Výsledek nalezneme ve výstupním souboru:



The screenshot shows the SPSS Viewer interface. The title bar reads "Output2 [Document3] - SPSS Viewer". The menu bar includes File, Edit, View, Data, Transform, Insert, Format, Analyze, Graphs, Utilities, Window, and Help. The toolbar contains various icons for file operations and analysis. The left sidebar shows a tree view with "Output" expanded to "Descriptives", which is further expanded to "Descriptive Statistics". The main window displays the following text:

```
DESCRIPTIVES  
VARIABLES=salary  
/STATISTICS=MEAN STDDEV MIN MAX .
```

Descriptives

[DataSet7] C:\Program Files\SPSS\Employee data.sav

A table titled "Descriptive Statistics" is shown, with a red arrow pointing to it from the left. The table has the following data:

	N	Minimum	Maximum	Mean	Std. Deviation
Current Salary	474	\$15,750	\$135,000	\$34,419.57	\$17,075.661
Valid N (listwise)	474				

Crosstabs – kontingenční tabulky

- Využívají se při zkoumání vztahu mezi dvěma kategorizovanými proměnnými
- Pro grafické znázornění lze doplnit sloupcový graf četností
- K posouzení síly vztahu mezi proměnnými software disponuje škálou nejrůznějších statistických testů

- Analyzované proměnné vložíme do polí Rows a Columns
- Pokud chceme vložit do některého z polí více proměnných, výsledky se zobrazí v samostatných tabulkách
- Pro lepší orientaci požadujeme sloupcový graf označením možnosti Display clustered bar charts
- V nabídce Statistics vybereme statistický test, pomocí něhož budeme posuzovat významnost zjištěné závislosti (obvykle chí kvadrát)

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
------	------	-------	----------	-------	--------	---------	---------	-------	---------

Crosstabs

- Employee Code [id]
- Date of Birth [bdate]
- Educational Level [year]
- Current Salary [salary]
- Beginning Salary [salb]
- Months since Hire [job]
- Previous Experience [pre]
- Minority Classification [mnc]

Row(s):

Gender [gender]

Column(s):

Employment Category

OK

Paste

Reset

Cancel

Help

Layer 1 of 1

Previous Next

Display clustered bar charts

Suppress tables

Statistics...

Cells...

Format...

	None	8	Right	Scale
ale}...	None	1	Left	Nominal
	None	13	Right	Scale
issing))0	0	8	Right	Ordinal
issing))0	0	8	Right	Ordinal
sing}...\$0	\$0	8	Right	Scale
sing}...\$0	\$0	8	Right	Scale
ing}...0	0	8	Right	Scale
ing}...None	None	8	Right	Scale

Crosstabs: Statistics

Chi-square

Nominal

Contingency coefficient

Phi and Cramér's V

Lambda

Uncertainty coefficient

Nominal by Interval

Eta

Cochran's and Mantel-Haenszel statistics

Test common odds ratio equals:

Correlations

Ordinal

Gamma

Somers' d

Kendall's tau-b

Kendall's tau-c

Kappa

Risk

McNemar

Continue

Cancel

Help

- Chceme-li ve výstupní tabulce zobrazit pozorované a očekávané četnosti, jejich rozdíly nebo řádková či sloupcová procenta, výběr je k dispozici v nabídce Cells:

The image shows two dialog boxes from the SPSS software interface. The background is a data table with columns: Name, Type, Width, Decimals, Label, Values, Missing, Columns, and A. The **Crosstabs** dialog box is open, showing 'Gender [gender]' in the Row(s) field and 'Employment Category' in the Column(s) field. The **Crosstabs: Cell Display** sub-dialog box is also open, showing options for Counts (Observed checked), Percentages (Row checked), Residuals (Unstandardized, Standardized, Adjusted standardized), and Noninteger Weights (Round cell counts selected).

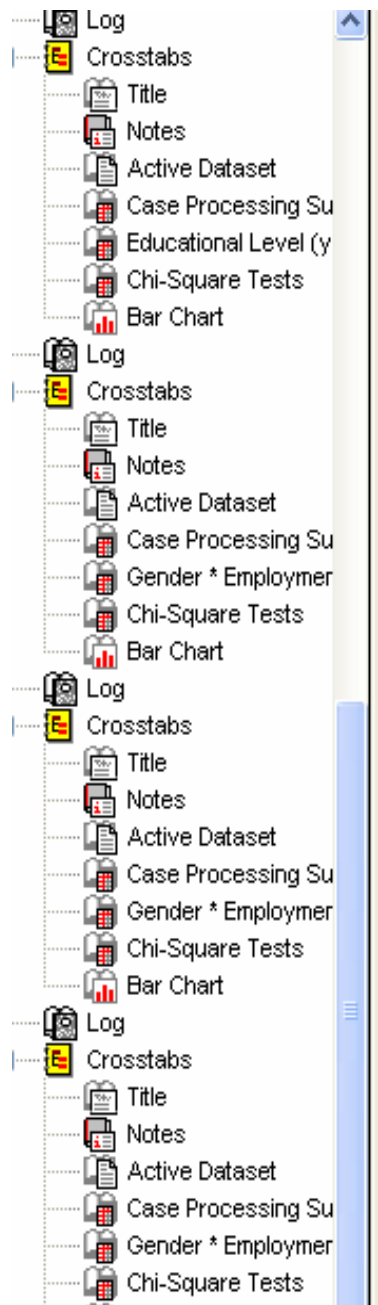
Crosstabs Dialog:

- Row(s): Gender [gender]
- Column(s): Employment Category
- Display clustered bar charts:
- Suppress tables:

Crosstabs: Cell Display Dialog:

- Counts: Observed, Expected
- Percentages: Row, Column, Total
- Residuals: Unstandardized, Standardized, Adjusted standardized
- Noninteger Weights: Round cell counts, Truncate cell counts, No adjustments; Round case weights, Truncate case weights

- Ve výstupním okně je znázorněna tabulka s počtem platných a chybějících hodnot (Case Processing Summary)
- Kontingenční tabulka (Crosstabulation) obsahující v řádcích legendy varianty, kterých může nabývat jedna proměnná, ve sloupcích legendy možné varianty druhé proměnné a ve vnitřních políčkách sdružené četnosti, popř. je doplněna o požadovaná procenta (viz. výše)
- Požadovaný statistický test (např. Chi-Square Tests)
- Sloupcový graf



Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Employment Category	474	100,0%	0	,0%	474	100,0%

Gender * Employment Category Crosstabulation

			Employment Category			Total
			Clerical	Custodial	Manager	
Gender	Female	Count	206	0	10	216
		% within Gender	95,4%	,0%	4,6%	100,0%
	Male	Count	157	27	74	258
		% within Gender	60,9%	10,5%	28,7%	100,0%
Total		Count	363	27	84	474
		% within Gender	76,6%	5,7%	17,7%	100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	79,277 ^a	2	,000
Likelihood Ratio	95,463	2	,000
N of Valid Cases	474		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 12,30.

Bar Chart

