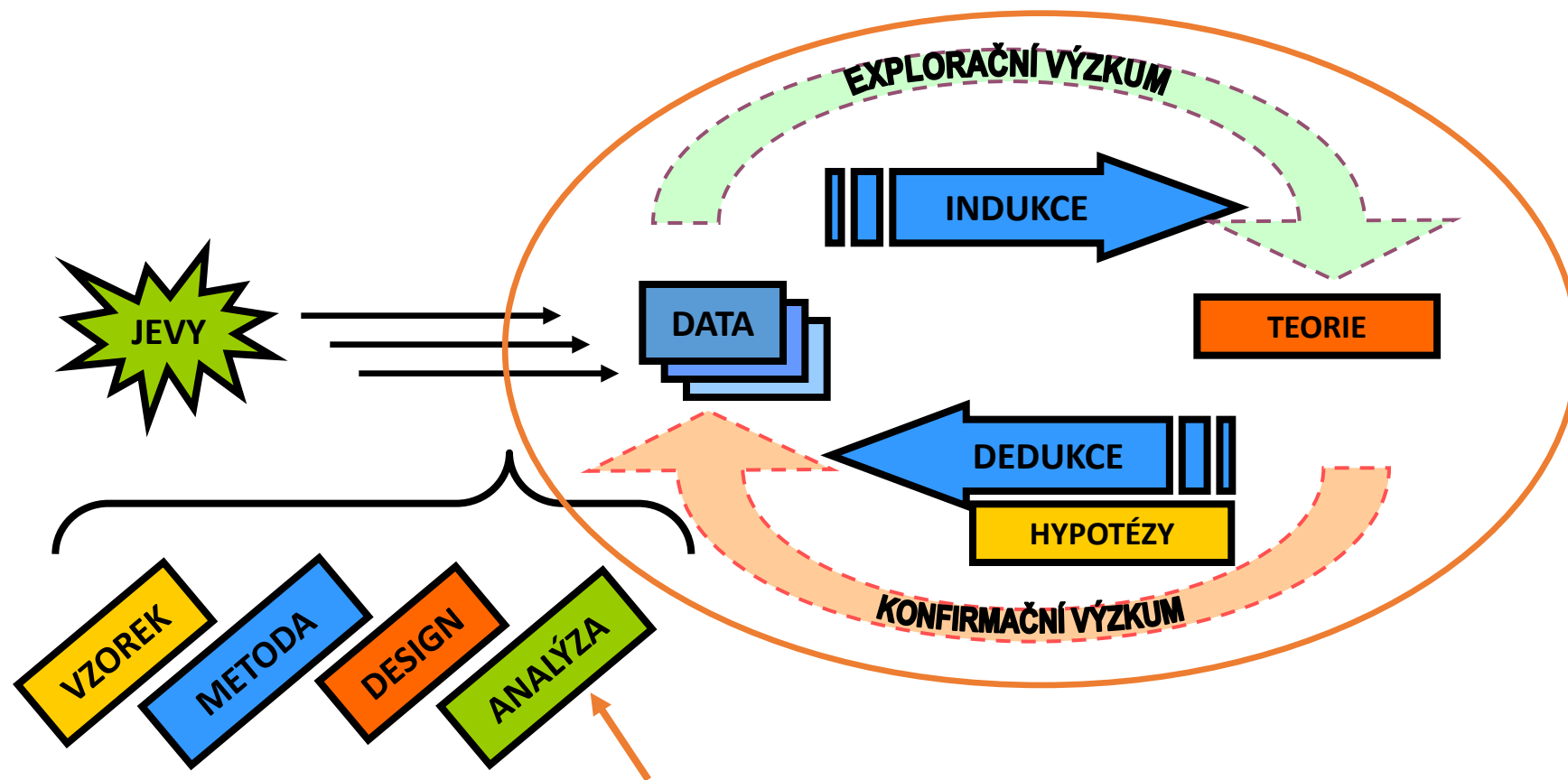


# Zpracování dat

S využitím podkladů K. Vlčkové a Jiřího Šafra

# Když se řekne... věda



AJ: phenomena, (empirical) data, facts, induction, theory, deduction, hypotheses, sample, method, design, analysis, exploratory and confirmatory research, statistics

# Zpracování dat jako jedna z fází výzkumu

- Následuje nejčastěji po všech přípravných fázích
  - formulace problému,
  - studium literatury,
  - formulování hypotéz,
  - vymezení základních pojmů,
  - tvorba výzkumného nástroje,
  - volba výzkumného vzorku,
  - ověřování konstrukce a vlastností výzkumného nástroje aj. v předvýzkumu,
  - sběr dat.
  
- Následuje fáze zpracování dat – odlišná fáze, dosti mechanická

# Rozdíl zpracování dat v kvalitativním a kvantitativním výzkumu

- O tom, jak bude vypadat zpracování dat rozhoduje to, zda jsme dělali kvalitativní či kvantitativní výzkum:
  - **u kvalitativního**
    - budeme získaná data třídit, kategorizovat,
    - kvalitativní analýza (typické, reprezentativní příklady X atypické),
    - interpretovat, vysvětlovat,
  - **u kvantitativního výzkumu**
    - bylo o způsobu zpracování dat bylo už rozhodnuto předem, než se začala sbírat data!!!,
    - způsob zpracování byl také ozkoušen v předvýzkumu
    - četnosti, statistika.
  - **u smíšeného designu**
    - kombinace obojího

# Kvalitně a adekvátně statistice sbíraná data

## Problémy

- Jednoznačné přiřazení do kategorií.
- Dostatečně naplněné kategorie.
- Dostatečný počet respondentů.
- Otevřené položky a jejich kategorizace.
- Moc dlouhé otázky a náročné nebo nesmyslné.
- Nečitelné odpovědi.
- Zavádějící odpovědi respondentů.
- Plán kódování dat.

# Předzpracování dat

- Nejprve se data předzpracovávají
  - připravují pro zpracování,
  - zvažují se možné kategorizace dat,
  - zadávají se proměnné a jejich hodnoty do hlavičky datové matice,
  - pak se zadávají data převedená většinou do čísel.
- vše už rovnou psát do počítače
  - do (nejlépe) statistického programu nebo alespoň do Excelu,
  - z něj se to dá převést do SPSS, Statistica, PSPP (free ware), R (free ware), aj.

# Excel

- Není statistický program
- Mnohé jde zpracovat i v Excelu
- Problém, jak pracovat s chybějícími daty
  - ručně

Microsoft Excel - database\_disertacka

Úpravy Zobrazit Vložit Formát Nástroje Data Okno Nápověda

Nápověda - zadejte dotaz

C616

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Cislo	škola	trida	hlavi	doxa-Aj	znamka-Aj	znalost-Aj	mluveni-Aj	cteni-Aj	psani Aj	poslech-Aj	doxa-Nj	znamka-Nj	znalost-Nj	mluveni NJ	cteni-Nj	psani-Nj	poslec
2	1	Taborska	3.A	z	3	2	3	2	1	2	3	8	3	4	4	3	4	
3	2	Taborska	3.A	m	11	3	3	2	2	2	3	3	4	4	4	4	4	
4	3	Taborska	3.A	m	10	2	3	2	2	2	2	3	2	4	3	3	3	
5	4	Taborska	3.A	m	10	4	4	4	4	3	4	3	4	4	4	4	4	
6	5	Taborska	3.A	z	10	3	4	4	4	3	4	3	3	4	4	4	4	
7	6	Taborska	3.A	z	3	2	4	3	3	3	4	10	2	3	2	3	2	
8	7			m	3	3	3	3	3	3	3	7	3	3	2	3	2	
9	8	Taborska	3.A	m	7	3	2	2	1	3	1	3	2	3	2	3	3	
10	9	Taborska	3.A	m	8	3	3	2	2	2	2	3	2	4	3	3	3	
11	10	Taborska	3.A	m	10	2	2	2	2	1	2	6	2	4	2	3	1	
12	11	Taborska	3.A	z	8	3	4	4	3	4	2	6	1	3	2	2	3	
13	12	Taborska	3.A	m	3	2	3					9	3	2	2	2	4	
14	13	Taborska	3.A	m	3	3	4	3	2	3	3	9	2	3	2	1	3	
15	14	Taborska	3.A	z	8	3	3	3	3	4	4	3	2	4	4	4	4	
16	15	Taborska	3.A	z	3	3	4	4	3	3	3	11	2	3	3	2	1	
17	16	Taborska	3.A	m	3	3	4	4	3	4	3	10	2	3	3	3	2	
18	17	Taborska	3.B	z	10	2	3	3	2	2	3	3	2	4	4	3	3	
19	18	Taborska	3.B	z	10	4	3	3	2	2	3	3	3	4	4	4	4	
20	19	Taborska	3.B	z	10	3	3	3	2	3	4	3	2	4	4	3	3	
21	20	Taborska	3.B	z	8	3	3	3	2	4	3	3	3	4	4	4	4	
22	21	Taborska	3.B	m	10	2	2	3	1	2	3	3	3	3	4	3	3	
23	22	Taborska	3.B	m	12	1	2	2	1	1	1	7	2	3	3	1	3	
24	23	Taborska	3.B	z	10	1		3	2	3	3	3	1	1	1	1	1	
25	24	Taborska	3.B	m	8	2	2	2	1	1	2	3	2	3	3	3	3	
26	25	Taborska	3.B	z	10	2	4	4	2	4	2	5	2				4	
27	26	Taborska	3.B	z	10	2	2	3	3	3	3	3	1				4	
28	27	Taborska	3.B	z	9	1	3	3	2	3	2	4	2				4	
29	28	Taborska	3.B	z	10	3	4	3	3	3	4	3	2				3	
30	29	Taborska	3.B	m	6	3	2	3	3	3	2	3	3				4	
31	30	Taborska	3.B	m	11	2	2	2	2	3	1	3	2	4	4	4	4	
32	31	Taborska	3.B	z	10	2	3	3	2	3	3	3	3	4	4	3	3	

database List2 List3

Kreslení Automatické tvary

Připraven

**datová  
matice**



Microsoft Excel - databse\_disertacka

Nápověda – zadejte dotaz

Seřadit...  
Filtř

Formulář...  
Souhrny...  
Ověření...  
Tabulka...  
Text do sloupců...  
Sloučit...  
Skupina a přehled  
Kontingenční tabulka a graf...  
Importovat externí data  
Aktualizovat data

	CD	CE	CF	CG	CH	CO	CP	CQ	CR	CS	CT	CU	CV	CW	CX	CY	CZ	DA	DB	DC	DD	DE
1	otiv-kultura	otiv-zajem	partn. skola	otiv-rodice	otiv-DA																	
2	0	0	0	0																		
3	0	1	0	1																		
4	0	0	0	0																		
5	0	0	0	1																		
6	0	0	0	1																		
7	0	0	0	0																		
8	0	0	0	0																		
9	0	0	0	0																		
10	0	0	0	0																		
11	0	0	0	0																		
12	0	0	0	0																		
13	0	0	0	0																		
14	1	1	0	0	0			2	2	2	2	3	3	3	2							
15	0	0	0	1	0			2	3	1	1	5	4	2	2							
16	0	1	0	0	0			2	3	3	1	4	4	1	2							
17	0	0	0	0	0			3	2	3	3	1	2	2	2							
18	0	0	0	0	0			2	1	1	3	4	3	1	3							
19	0	1	0	0	0			1	3	2	3	4	2	1	3							
20	1	1	0	0	0			1	2	3	3	4	2	2	2							
21	0	1	0	0	1	cteni knize		1	1	1	2	3	4	1	1	1	1	2	2	1	1	1
22	0	0	0	0	1	potrebuji to		1	4	3	1	3	3	1	2	1	1	2	2	1	1	1
23	1	0	0	1	0			4	2	4	3	5	5	3	5	3	3	4	2	1	3	1
24	1	1	0	0	0			1	2	4	4	2	1	1	1	2	2	5	3	3	1	1
25	0	0	0	1	0			2	4	2	1	4	3	2	1	1	1	5	5	5	1	1
26	0	0	0	1	0			2	1	3	2	3	1	1	2	1	1	4	1	2	1	1
27	0	0	0	0	1	pribuzni v d		4	3	2	1	4	1	3	3	1	3	2	2	2	1	1
28	1	1	0	0	0			1	1	1	2	4	1	1	1	1	1	4	1	1	1	5
29	0	0	0	0	0			4	1	5	4	4	3	1	1	1	2	5	2	5	1	1
30	0	0	0	0	0			1	3	2	4	3	1	1	2	1	3	5	3	3	3	1
31	1	0	0	0	0			1	1	4	4	4	1	3	2	1	2	3	3	4	1	5
32	0	0	0	0	0			1	1	3	1	4	3	2	1	2	3	5	5	5	1	5

**Data lze filtrovat – vybrat si pro analýzy jen např. dívky**

database / List2 / List3

Kreslení Automatické tvary

Připraven

Microsoft Excel - database\_disertacka

Soubor Úpravy Zobrazit Vložit Formát Nástroje Data Okno Nápověda

Nápověda - zadejte dotaz

CB16 fx

A B C H I J K L M N O P Q R

600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631

- Buňky...
- Řádek
- Sloupec
- List
- Graf...
- Symbol...
- Konec stránky
- fx** Funkce...
- Název
- Komentář
- Obrázek
- Diagram
- Objekt...
- Hypertextový odkaz... Ctrl+K

Kreslení Automatické tvary

Připraven

Start

stare database zaklady\_statistiky -... stat\_zprac\_dat - Mi... Microsoft PowerPoi... Microsoft Excel - ...

9:32

**Pomocí Excelu lze počítat mnohé statistické funkce**

Vyhledat funkci:

Zadejte stručný popis požadované činnosti a potom klepněte na tlačítko Přejít.

Přejít

Vybrat kategorii: Statistické

Vybrat funkci:

MAX  
MAXA  
MEDIAN  
MIN  
MINA  
MODE  
NEGBINOMDIST

Použití funkcí v  
Excelu

### **AVERAGEA(hodnota1;hodnota2;...)**

Vrátí průměrnou hodnotu (aritmetický průměr) argumentů. Text a logická hodnota NEPRAVDA mají hodnotu 0, logická hodnota PRAVDA má hodnotu 1. Argumenty mohou být čísla, názvy, matice nebo odkazy.

[Nápověda k této funkci](#)

OK

Storno

Microsoft Excel - Kopie - PISA strategie

Nápověda - zadejte dotaz

S3

**Tabulka výsledků**

Table 4.5  
**Index of control strategies and performance on the combined reading literacy scale, by national quarters of the index**  
*Results based on students' self-reports*

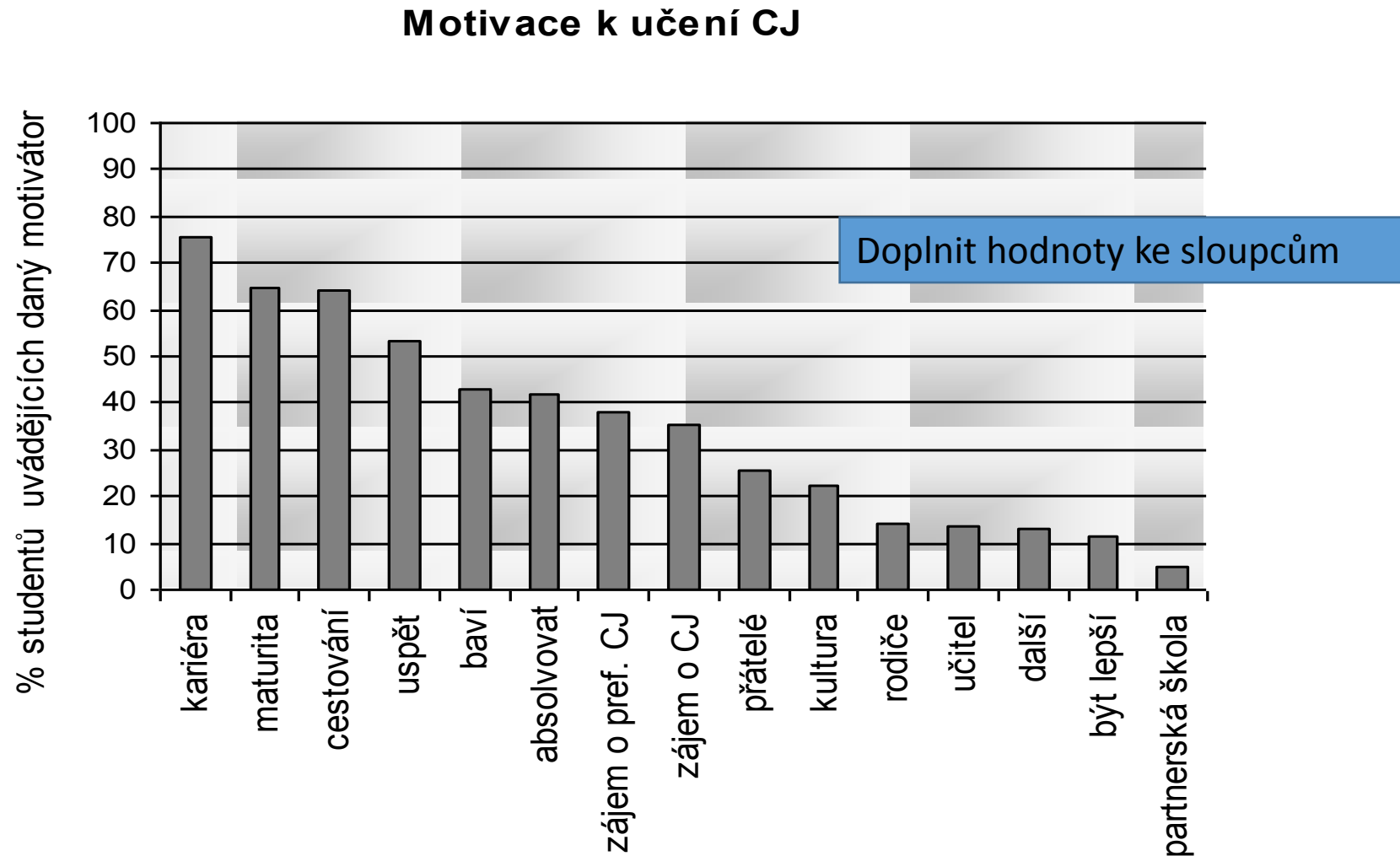
Country	Index of control strategies <sup>1</sup>													
	All students		Males		Females		Bottom quarter		Second quarter		Third quarter		Top quarter	
	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.	Mean index	S.E.
<b>OECD Countries</b>														
Australia	0,02	(0,02)	-0,05	(0,02)	0,10	(0,03)	-1,20	(0,02)	-0,26	(0,01)	0,30	(0,01)	1,24	(0,02)
Austria	0,40	(0,02)	0,31	(0,03)	0,48	(0,02)	-0,72	(0,02)	0,12	(0,01)	0,64	(0,01)	1,54	(0,02)
Belgium (F)	0,14	(0,02)	0,07	(0,02)	0,21	(0,03)	-0,96	(0,02)	-0,11	(0,01)	0,37	(0,01)	1,25	(0,02)
Czech Rep	0,27	(0,02)	0,11	(0,03)	0,42	(0,02)	-0,84	(0,01)	-0,02	(0,00)	0,50	(0,01)	1,45	(0,02)
Denmark	-0,23	(0,01)	-0,24	(0,02)	-0,22	(0,02)	-1,25	(0,02)	-0,48	(0,01)	0,00	(0,01)	0,81	(0,02)
Finland	-0,47	(0,02)	-0,52	(0,02)	-0,42	(0,02)	-1,54	(0,02)	-0,71	(0,01)	-0,21	(0,01)	0,58	(0,02)
Germany	0,24	(0,02)	0,14	(0,04)	0,33	(0,02)	-0,94	(0,02)	-0,05	(0,01)	0,48	(0,00)	1,45	(0,03)
Hungary	0,21	(0,02)	0,09	(0,03)	0,33	(0,03)	-0,91	(0,02)	-0,05	(0,01)	0,45	(0,01)	1,34	(0,02)
Iceland	-0,35	(0,02)	-0,36	(0,03)	-0,34	(0,02)	-1,53	(0,02)	-0,60	(0,01)	-0,08	(0,01)	0,79	(0,02)
Ireland	0,07	(0,02)	-0,10	(0,04)	0,23	(0,03)	-1,28	(0,03)	-0,22	(0,01)	0,37	(0,01)	1,39	(0,02)
Italy	0,23	(0,02)	0,05	(0,04)	0,41	(0,02)	-0,94	(0,03)	-0,04	(0,01)	0,48	(0,00)	1,41	(0,01)
Korea	-0,44	(0,02)	-0,47	(0,03)	-0,41	(0,03)	-1,74	(0,02)	-0,70	(0,01)	-0,12	(0,01)	0,79	(0,02)
Luxembou	0,05	(0,02)	-0,10	(0,03)	0,19	(0,03)	-1,26	(0,03)	-0,23	(0,01)	0,33	(0,01)	1,34	(0,03)
Mexico	0,16	(0,02)	0,06	(0,03)	0,25	(0,02)	-0,98	(0,01)	-0,18	(0,01)	0,39	(0,01)	1,41	(0,02)
New Zeala	0,07	(0,03)	-0,03	(0,03)	0,17	(0,03)	-1,12	(0,02)	-0,23	(0,01)	0,30	(0,01)	1,32	(0,02)
Norway	-0,58	(0,02)	-0,50	(0,03)	-0,66	(0,02)	-1,76	(0,02)	-0,81	(0,01)	-0,28	(0,01)	0,54	(0,02)
Portugal	0,19	(0,02)	0,03	(0,02)	0,34	(0,02)	-0,90	(0,02)	-0,13	(0,01)	0,39	(0,01)	1,41	(0,02)
Sweden	0,03	(0,02)	0,04	(0,03)	0,02	(0,02)	-1,09	(0,02)	-0,22	(0,01)	0,29	(0,01)	1,17	(0,02)
Switzerlan	0,11	(0,02)	0,00	(0,03)	0,22	(0,03)	-1,00	(0,02)	-0,15	(0,00)	0,35	(0,01)	1,26	(0,03)
United Sta	-0,08	(0,03)	-0,26	(0,04)	0,09	(0,04)	-1,44	(0,03)	-0,40	(0,01)	0,24	(0,01)	1,30	(0,03)
OECD	0,01	(0,02)	-0,12	(0,02)	0,14	(0,02)	-1,24	(0,02)	-0,29	(0,01)	0,29	(0,00)	1,28	(0,01)
OECD	0,00	(0,01)	-0,09	(0,01)	0,09	(0,01)	-1,17	(0,01)	-0,28	(0,00)	0,26	(0,00)	1,19	(0,01)
<b>Non-OECD countries</b>														

List1 / List2 / List3 /

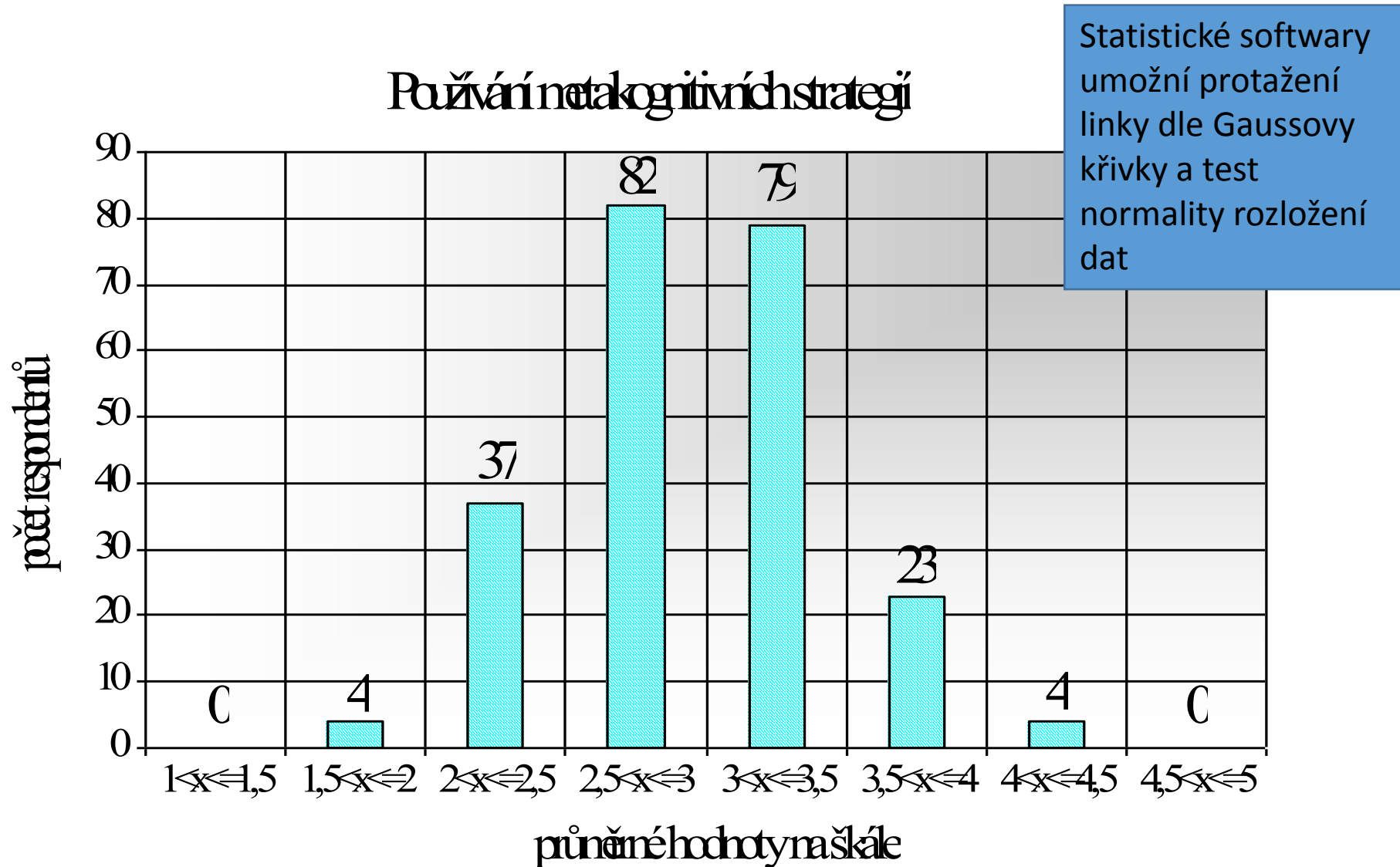
Kreslení Automatické tvary

Připraven

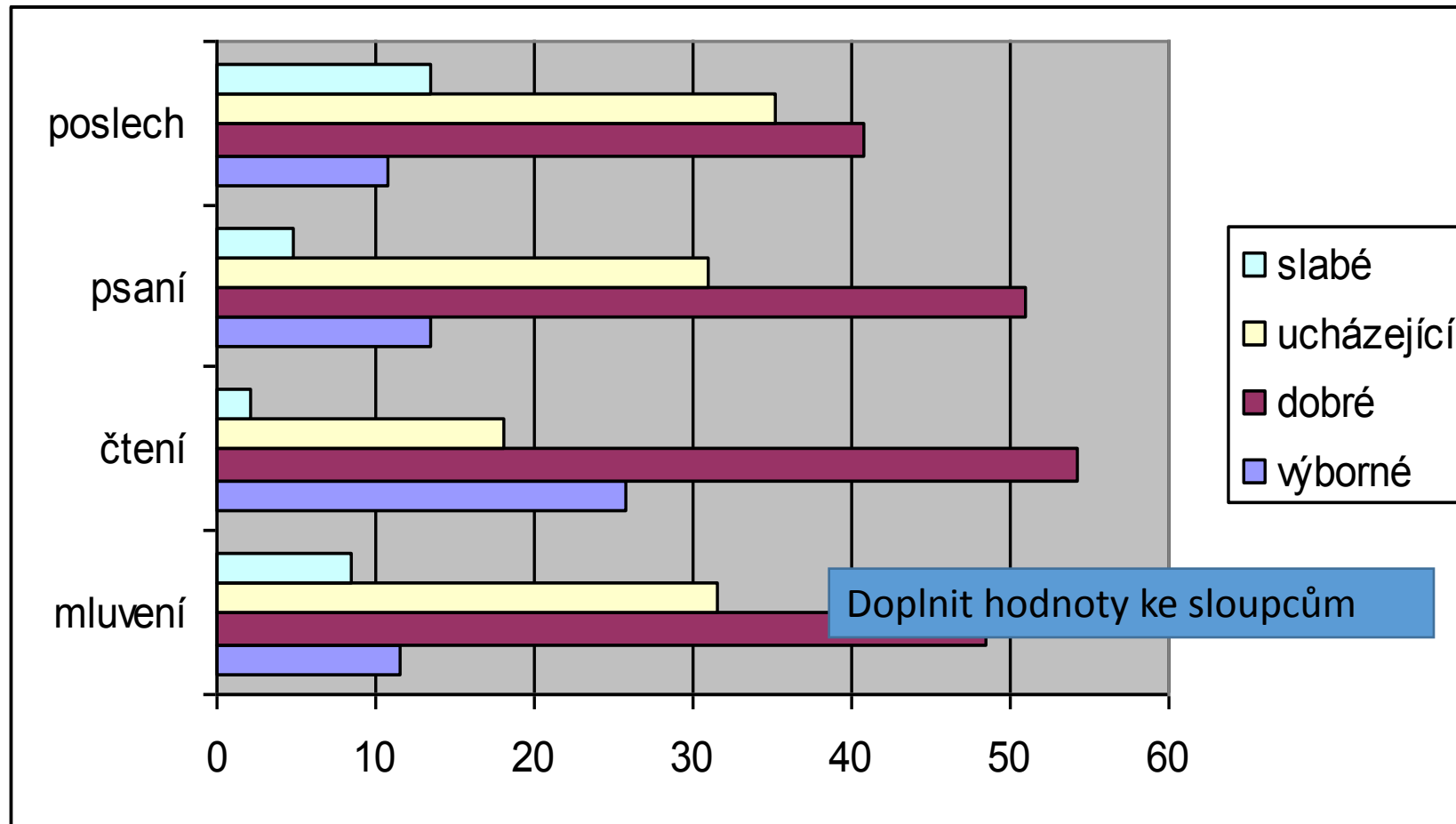
# Ukázka grafického znázornění dat (Excel, sloupcový graf)



# Ukázka grafického znázornění dat (Excel, histogram)

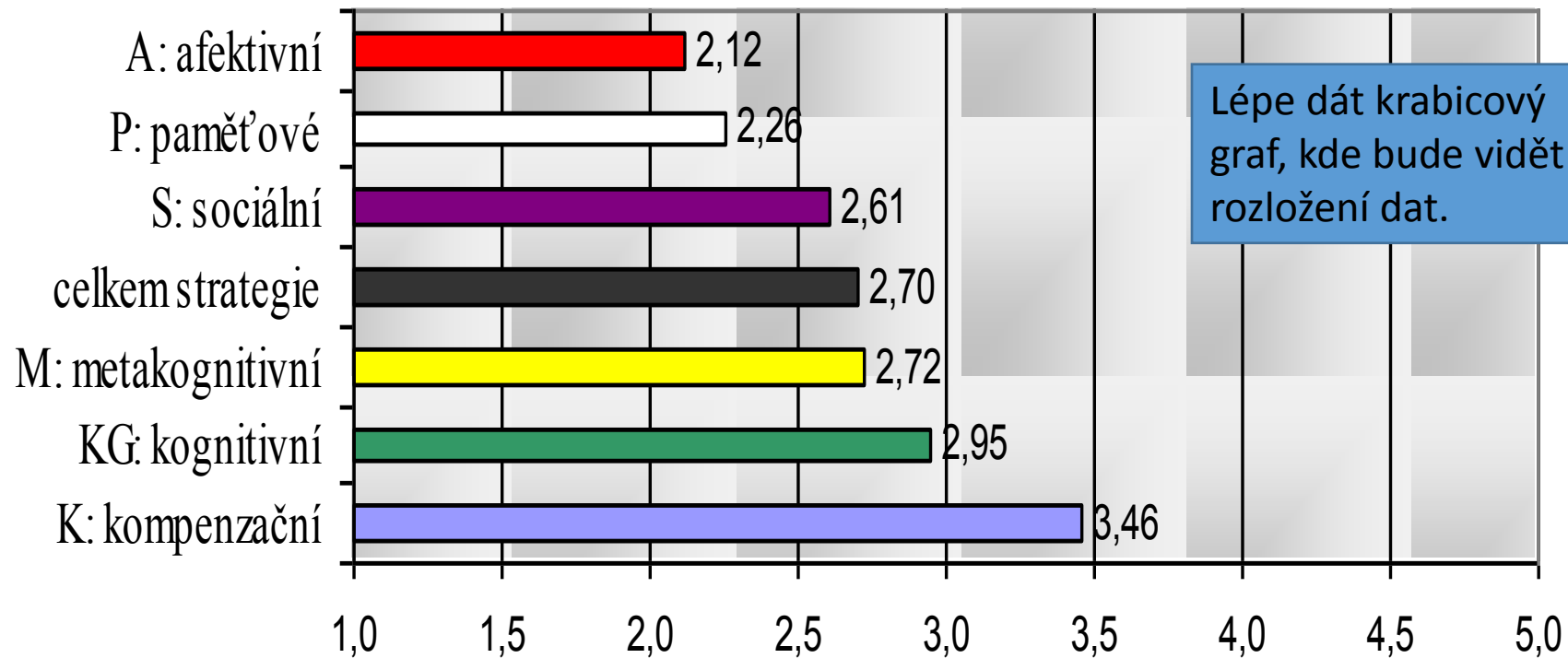


# Ukázka grafického znázornění dat (Excel, pruhový graf)



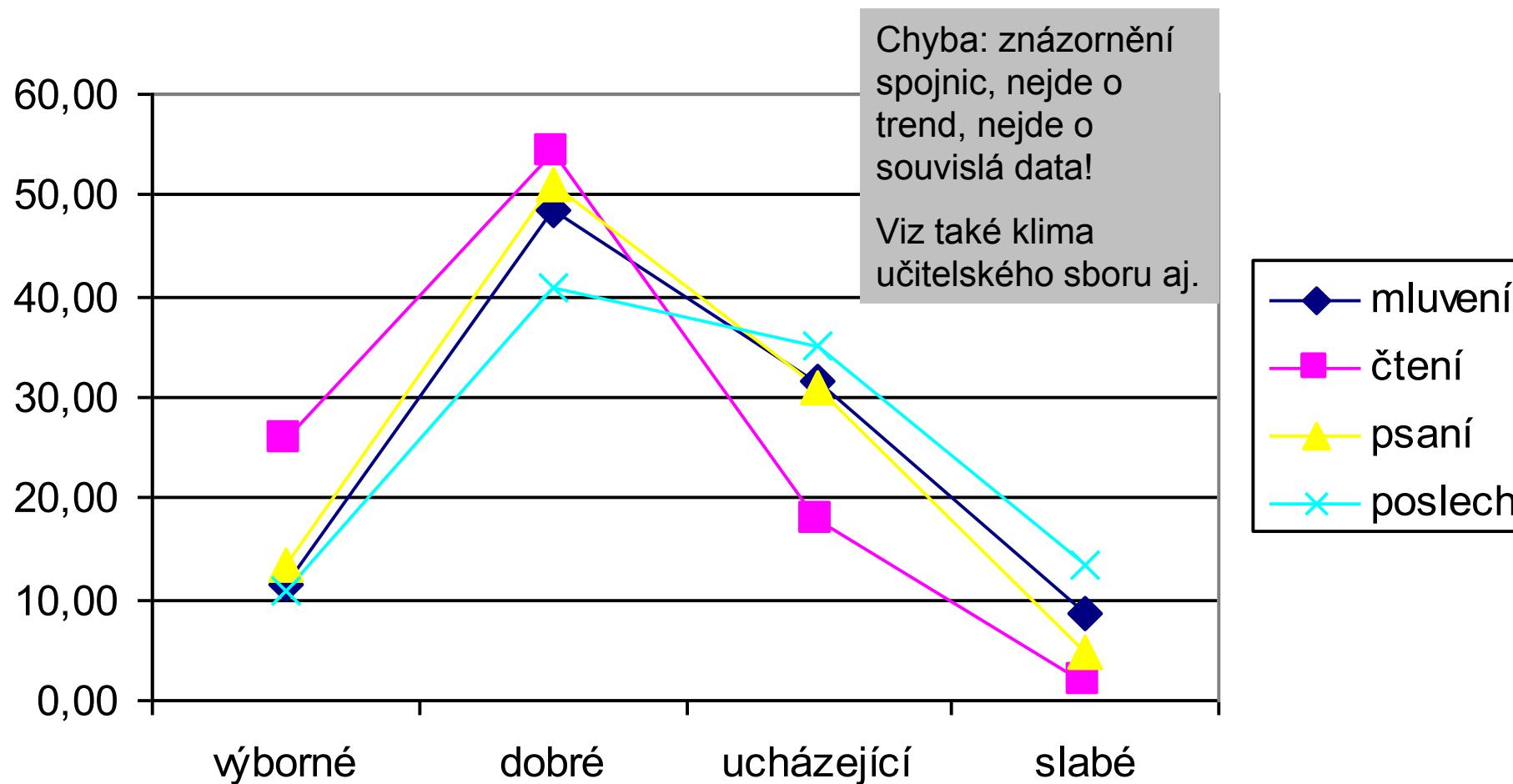
# Ukázka grafického znázornění dat (Excel, pruhový graf)

## Používání skupin strategií





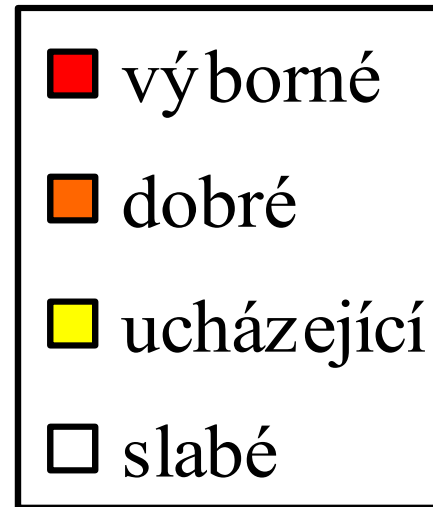
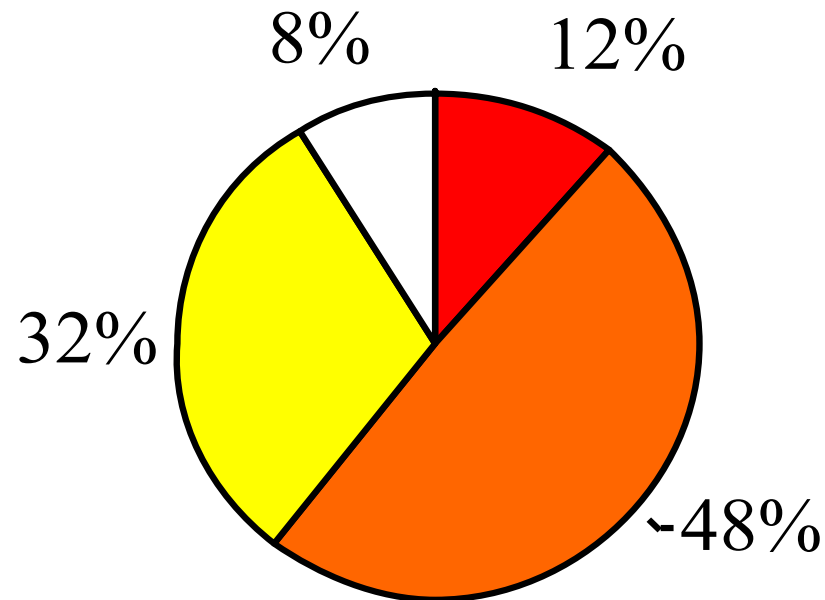
# Ukázka grafického znázornění dat (Excel)



# Ukázka grafického znázornění dat (Excel, koláčový graf)

mluvení

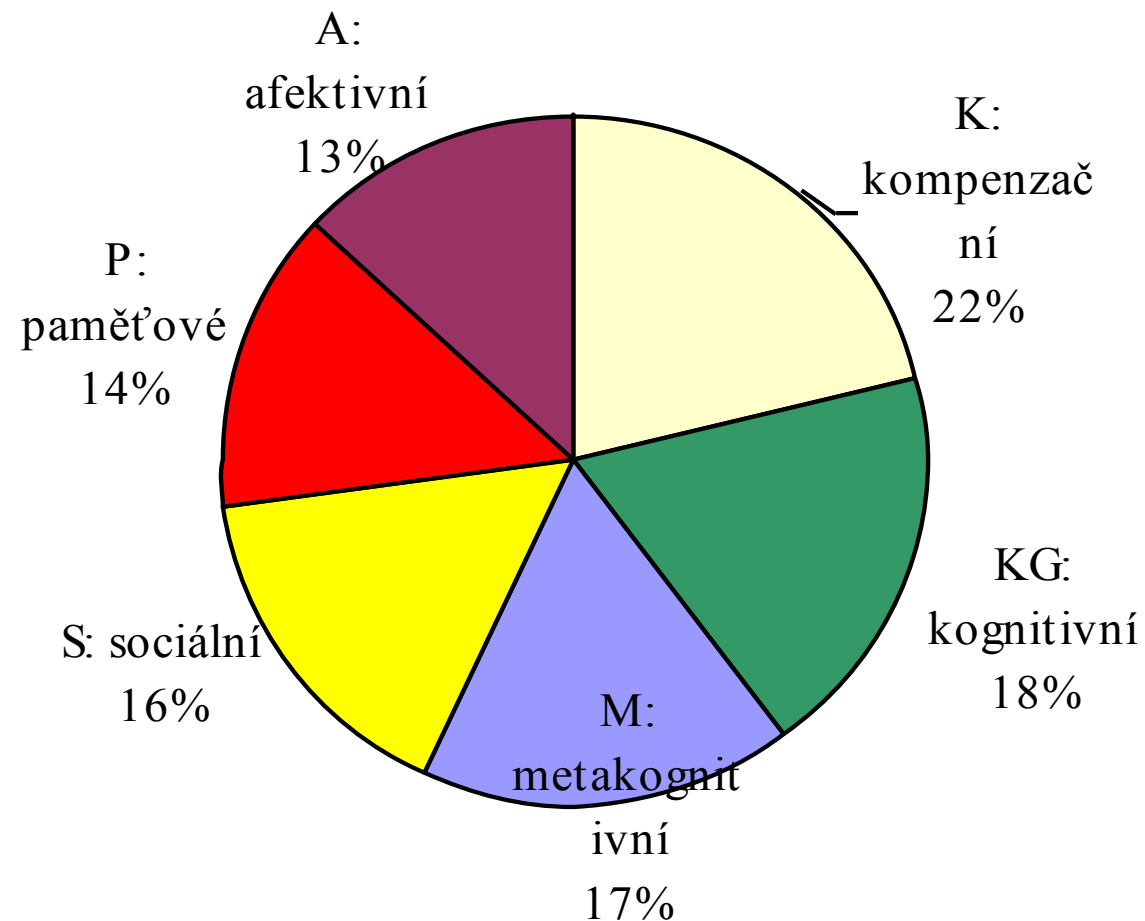
Pozor na zkreslení koláči,  
jde dát i poměrový  
pruhový graf



# Ukázka grafického znázornění dat (Excel, koláčový graf)

Nepostrádá  
tento graf  
smysl?

## Poměr strategií v procesu učení



# Zpracování dat ve statistickém softwaru

Co by měl umět:

- vytvoření kódovacího rámce,
- zakódování dat,
- transformace dat,
- statistická analýza dat.

# Hromadný sběr dat

- Dotazníky (ev. jiné nástroje generující množství dat)
  - Tužka –papír:
    - Sebereme a očíslovíme (typicky levý horní roh – např. 001 – 999)
    - Vytvoříme datovou matici („přepis dotazníku do PC“)
  - Google form či jiný / podobný elektronický sběr dat (předchozí krok odpadá)
    - Možno v Excelu (či LibreOffice Calculator) ev. rovnou ve statistickém programu
      - Kde vzít Excel - <https://it.muni.cz/sluzby/microsoft-office-365>
      - Kde vzít statistický software – viz návod doc. Vlčkové... nebo zkuste <https://www.gnu.org/software/pspp/>
    - První sloupec ID (číslo dotazníku/ů)
    - Další postupně jednotlivé odpovědi (např. „ot\_1“ až „ot\_99“)
      - Jednotlivé typy odpovědí kódujeme
      - Chybějící údaje kódujeme vysokým číslem, které nemůže být v intervalu použitého pro kódování (např. „99“ v případě likertovské škály míry souhlasu s výrokem v dotazníku „1-5“)

# SPSS / PSPP

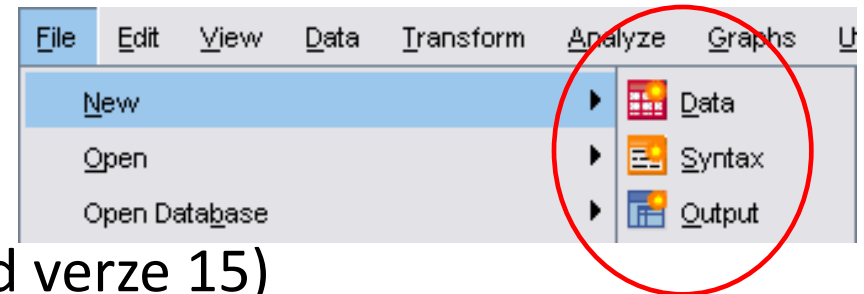
Stat. software

# 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**

# Jak to vypadá

The screenshot displays the SPSS interface with three main windows:

- Output – PSPPIRE Output Viewer:** Shows a list of menu items (GET, SORT CASES, SPLIT FILE, etc.) and a table of statistics for 'lowup ethnicit quiz1 final'.

Variable	N	Mean	Std. Deviation
lowup	3	1.00	0.816
ethnicit	3	4.00	1.414
quiz1	3	5.00	1.414
final	3	59.33	10.423
- \*students.sav [DataSet1] – PSPPIRE Data Editor:** Shows a list of variables (id, gender, ethnicit, etc.) and a data grid with columns for 'year' (1, 2) and quiz scores (quiz1, quiz2, quiz3, quiz4, quiz5, final, total).

Case	id	lastname	firstnam	year	quiz1	quiz2	quiz3	quiz4	quiz5	final	total			
1	908754	MARQUEZ	CHYRELLE	1	2	2	1	2	3.70	3.60	1	2	3	5
2	192627	MISCHKE	ELAINE	2	1	2	1	2	3.84	2.98	1	1	4	6
3	822485	VALENZUELA	KATHRYN	1	4	2	1	2	3.90	3.15	1	1	7	3
4	106484	VILLARRUZ	ALFRED	1	4	2	1	3	3.90	3.65	1	2	8	4
5	725987	BATILLER	FRED	1	3	2	1	3	3.90	3.95	1	2	8	2
- Descriptives:** A dialog box for selecting variables and statistics. The 'Variables' list includes 'quiz1' and 'quiz2'. The 'Statistics' list includes Mean, Standard deviation, Minimum, Maximum, Variance, and Kurtosis. The 'Include user-missing data in analysis' checkbox is checked.

At the bottom of the Data Editor window, the 'Data View' and 'Variable View' tabs are visible, with 'Data View' selected.



# Proměnné - úpravy

The screenshot shows the SPSS Data Editor window for a file named \*students.sav. The main window displays a list of variables with their properties. The variable 'ethnicit' is selected, and the 'psppire' dialog box is open, showing the 'Value Labels' section. The dialog box contains a list of value labels: 1 = 'AMERICAN INDIAN', 2 = 'ASIAN', 3 = 'AFRO-AMERICAN', 4 = 'CAUCASIAN', and 5 = 'HISPANIC'. The 'Value' field is set to 2, and the 'Value Label' field is set to 'ASIAN'. The 'Data View' and 'Variable View' tabs are visible at the bottom, with 'Variable View' selected and circled in orange.

Variab	Name	Type	Width	Decimal	Label	Value Labels	Missing Values	Columns	Align	Measure
1	id	Numeric	6	0		None	None	8	Right	Scale
2	lastname	String	10			None	None	10	Left	Nominal
3	firstnam	String	10			None	None	10	Left	Nominal
4	gender	Numeric	1	0		{1, FEMALE}...	None	8	Right	Ordinal
5	ethnicit	Numeric	1	0						Ordinal
6	year	Numeric	1	0	YEAR IN SCHOOL					Ordinal
7	lowup	Numeric	1	0	LOWER OR UPPER DIVIS					Ordinal
8	section	Numeric	1	0						Ordinal
9	hsgpa	Numeric	4	2	High School GPA					Scale
10	colgpa	Numeric	8	2	College GPA					Scale
11	extrcred	Numeric	1	0	DID EXTRA CREDIT PRO					Ordinal
12	review	Numeric	1	0	ATTENDED REVIEW SES					Ordinal
13	quiz1	Numeric	2	0						Ordinal
14	quiz2	Numeric	2	0						Ordinal
15	quiz3	Numeric	2	0						Ordinal
16	quiz4	Numeric	2	0		None	None	8	Right	Ordinal
17	quiz5	Numeric	2	0		None	None	8	Right	Ordinal
18	final	Numeric	2	0		None	None	8	Right	Scale
19	total	Numeric	3	0		None	None	8	Right	Scale
20	percent	Numeric	4	1		None	None	8	Right	Scale
21	grade	String	8			None	None	8	Left	Nominal
22	passfail	String	1			None	None	1	Left	Nominal

psppire

Value Labels

Value: 2

Value Label: ASIAN

+ Add

Apply

Remove

1 = 'AMERICAN INDIAN'

2 = 'ASIAN'

3 = 'AFRO-AMERICAN'

4 = 'CAUCASIAN'

5 = 'HISPANIC'

OK

Cancel

Help

Data View Variable View

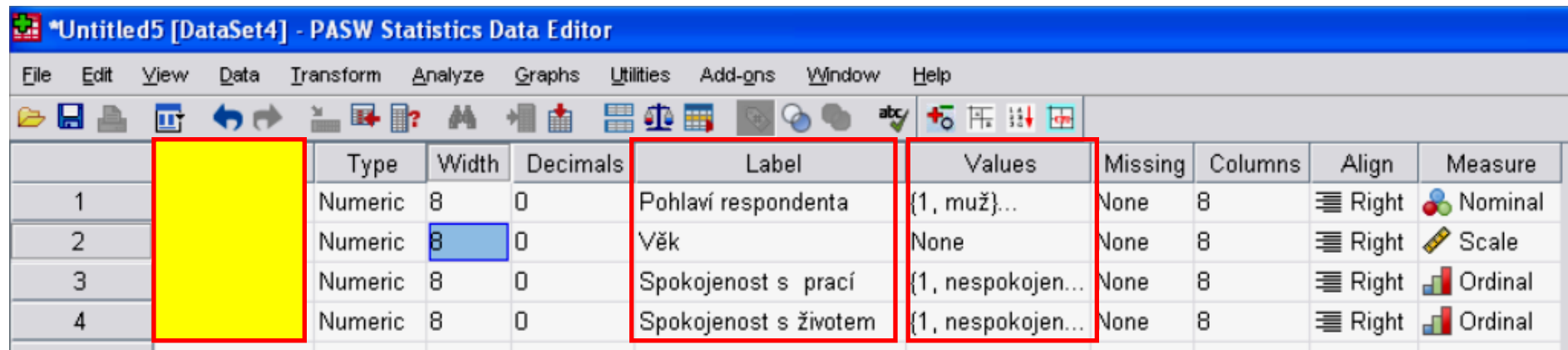
Filter off Weights off Split by year

# Tři základní informace o proměnných:

- **jméno proměnné** **VARIABLE NAME**  
max. 8 znaků bez diakritiky, mezer; musí začínat písmenem  
(např. *pohlavi*) → **bezpodmínečně v datech musí být**

pro uživatele může být dále užitečná nadstavba:

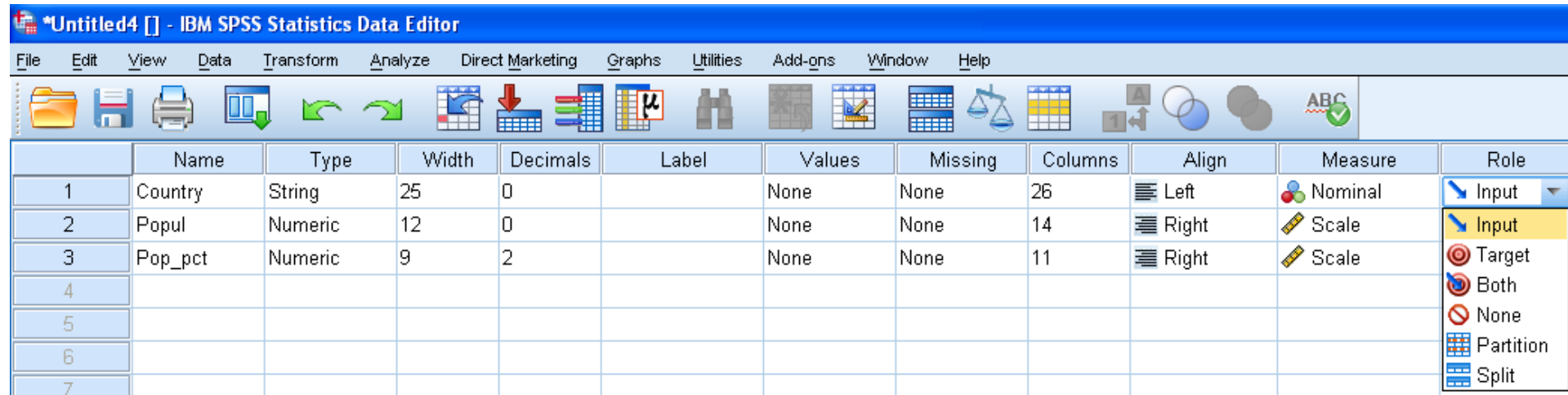
- **popiska proměnné** **VARIABLE LABEL**  
Delší popis názvu znaku, lze s diakritikou, např. (zkrácené)  
znění otázky v dotazníku, který se bude objevovat ve  
výstupech (např. *Pohlaví respondenta*)
- **popisky kategorií (hodnot)** proměnné **VALUE LABELS**  
Popis kategorií znaku, které se budou objevovat ve výstupech  
(např. *1 = muž, 2 = žena*)



The screenshot shows the 'Variable View' tab in PASW Statistics Data Editor. A yellow box highlights the variable names in the first column. Red boxes highlight the 'Label' and 'Values' columns for the first four variables.

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

# Ve Variable View je toho ale více



The screenshot shows the Variable View in IBM SPSS Statistics. The main table lists variables with their properties. A dropdown menu is open for the 'Role' column, showing options like Input, Target, Both, None, Partition, and Split.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Country	String	25	0		None	None	26	Left	Nominal	Input
2	Popul	Numeric	12	0		None	None	14	Right	Scale	Input
3	Pop_pct	Numeric	9	2		None	None	11	Right	Scale	Target
4											Both
5											None
6											Partition
7											Split

**Nejdůležitější** jsou:

- **Type:** druh záznamu dat (v zásadě buď jako číslo = Numeric nebo jako slova = String). Preferujeme Numeric
- **Decimals:** počet desetinných míst, co se bude zobrazovat ve výstupech (samotný záznam dat ale zachovává více desetinných míst)
- **Measure:** typ proměnné (dříve to nevadilo, novější verze vyžadují pro určitý typ analýz)
- **Missings:** uživatelsky definované hodnoty (o tom více jinde)
- v novějších verzích (SPSS 22) ještě přibyla **role proměnné v analýze** (závislá, nezávislá, třídící atd.)

# Překódování a transformace

The screenshot shows the SPSS Data Editor interface with the 'Compute Variable' dialog box open. The dialog is configured to create a new variable named 'log\_colgpa' using the expression 'LN(colgpa)'. The 'Target Variable' field contains 'log\_colgpa' and the 'Numeric Expressions' field contains 'LN(colgpa)'. The 'Functions' list includes 'LN(number)'. The 'Data View' tab is selected at the bottom of the window.

Case	id	lastname	firstnam	gender	ethnicit	year	lowup	section	hsgpa	colgpa	extrcred	review	quiz1	quiz2
1	302400	JONES												
2	106484	VILLARRUZ												
3	664653	KHAN												
4	595177	WILLIAMS												
5	506467	SCARBROUG												
6	681855	GRISWOLD												
7	721311	SONG												
8	237983	LEE												
9	725987	BATILLER												
10	615115	VASENIUS												
11	979028	NEUHARTH												
12	140219	GUADIZ												
13	908754	MARQUEZ												
14	417003	EVANGELIST												
15	818528	CARRINGTON												
16	938666	SUAREZ-TAN												
17	354601	CARPIO	MARY	1	2	2	1	1	2.03	2.40	1	2	10	1
18	307894	TORRENCE	GWEN	1	3	2	1	2	2.09	2.21	2	2	6	6
19	983522	SLOAT	AARON	2	3	3	2	3	2.11	2.45	1	1	4	6
20	108642	VALAZQUEZ	SCOTT	2	4	3	2	2	2.19	3.50	2	1	10	1
21	287617	CUMMINGS	DAVENA	1	5	3	2	3	2.21	3.82	1	2	9	1

# 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 (pro spojitě-kardinální znaky např. dle percentilů)

# 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**

# Typy proměnných

	Kategorické	Ordinální	Kardinální	
Stevensova typologie	Nominální	Ordinální	Intervalová	Poměrová
Matematické operace	$=, \neq$	$>, <$	$+, -$	$\times, \div$
Statistiky	Frekvence, Modus	Medián, percentily	Průměr, směrodatná odchylka	Koeficient variace
Příklady	<i>Město bydliště, Vystudovaný obor</i>	<i>dosažené vzdělání, hodnocení (výroku) na Likertově škále (např. 1-7)</i>	<i>Hodiny, IQ skór, Stupně Celsia</i>	<i>Věk, Stupně Celsia</i>
Další dělení	dichotomické vs. polytomické		spojité vs. nespojité	

kombinace úrovní proměnných - test	
1 s 1	Kontingenční tabulka s testem chí-kvadrát
1 s 2	Kontingenční tabulka s testem chí-kvadrát
1 s 3	Kontingenční tabulka s testem chí-kvadrát H-test podle Kruskala a Wallise
1 s 4	H-test podle Kruskala a Wallise
1 s 5	Jednofaktorová analýza rozptylu
2 s 2	Kontingenční tabulka s testem chí-kvadrát Přesný test podle Fischera a Yatese Korelace Chí-kvadrát-test podle McNemara*
2 s 3	Kontingenční tabulka s testem chí-kvadrát U-test podle Manna a Whitneyho Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla
2 s 4	U-test podle Manna a Whitneyho Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla
2 s 5	t-test podle Studenta bodová biseriální korelace
3 s 3	Kontingenční tabulka s testem chí-kvadrát H-test podle Kruskala a Wallise Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla Wilcoxonův test*

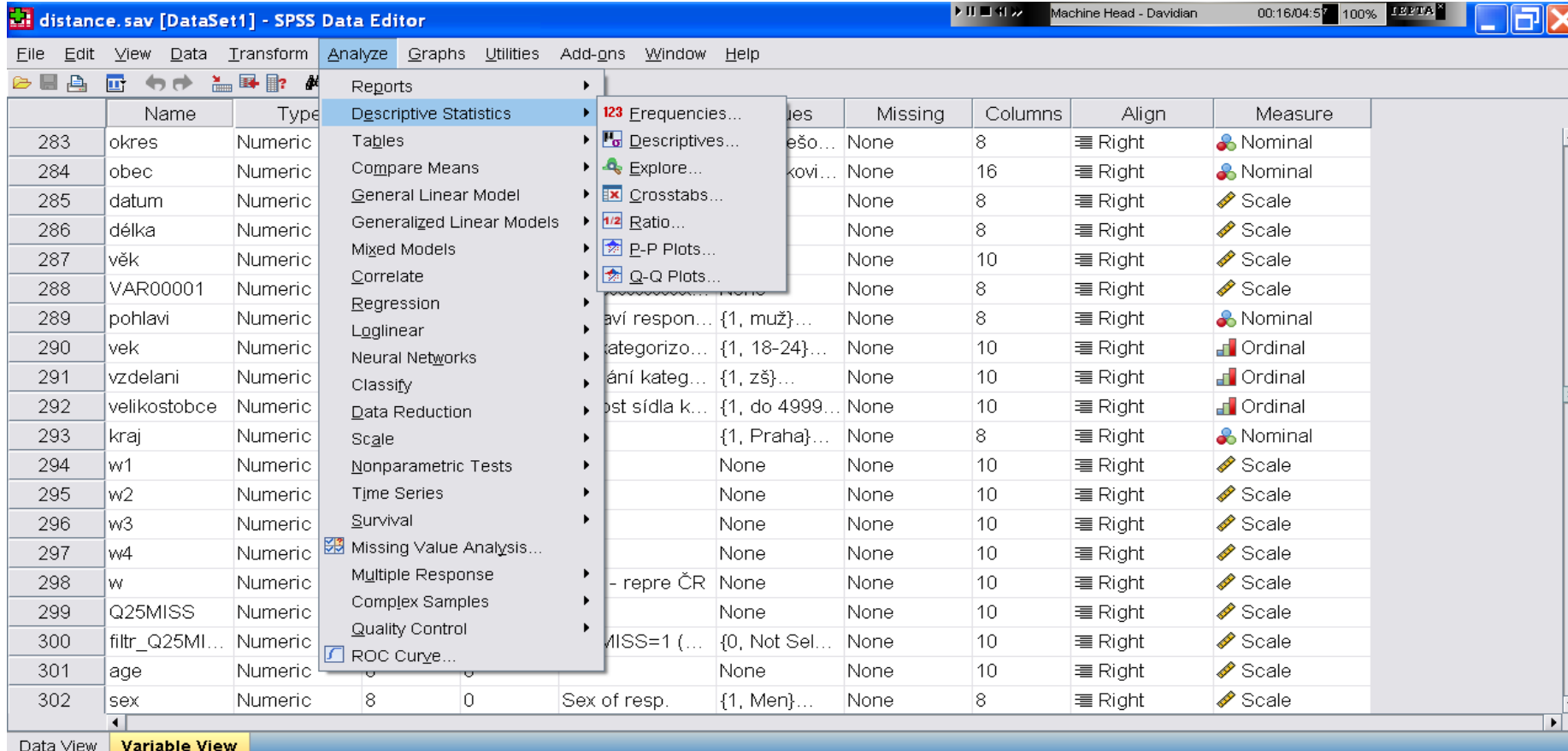
\*nelze je provést ve všech situacích, resp. nejsou smysluplné



## kombinace úrovní proměnných - test (pokračování)

3 s 4	H-test podle Kruskala a Wallise Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla
3 s 5	Jednofaktorová analýza rozptylu Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla
4 s 4	Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla Wilcoxonův test*
4 s 5	Pořadová korelace podle Spearmana Pořadová korelace podle Kendalla (Kendallův koeficient pořadové korelace) Wilcoxonův test*
5 s 5	<b>Produkt-moment korelace</b> Parciální korelace t-test pro nezávislé výběry

# Analýzy



The screenshot shows the SPSS Data Editor interface with the 'Analyze' menu open. The 'Descriptive Statistics' option is selected, and a sub-menu is visible showing options like 'Frequencies...', 'Descriptives...', 'Explore...', 'Crosstabs...', 'Ratio...', 'P-P Plots...', and 'Q-Q Plots...'. The background shows a variable list with columns for Name, Type, Missing, Columns, Align, and Measure.

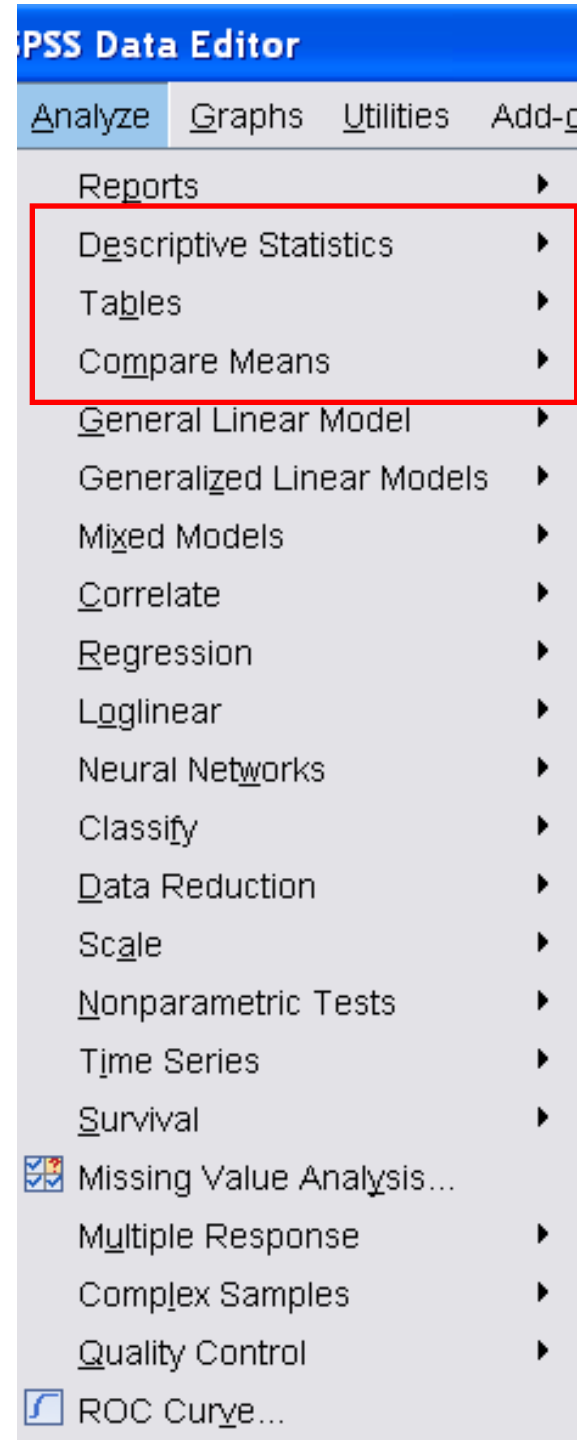
Name	Type	Missing	Columns	Align	Measure
283 okres	Numeric		8	Right	Nominal
284 obec	Numeric		16	Right	Nominal
285 datum	Numeric		8	Right	Scale
286 délka	Numeric		8	Right	Scale
287 věk	Numeric		10	Right	Scale
288 VAR00001	Numeric		8	Right	Scale
289 pohlavi	Numeric		8	Right	Nominal
290 vek	Numeric		10	Right	Ordinal
291 vzdelani	Numeric		10	Right	Ordinal
292 velikostobce	Numeric		10	Right	Ordinal
293 kraj	Numeric		8	Right	Nominal
294 w1	Numeric		10	Right	Scale
295 w2	Numeric		10	Right	Scale
296 w3	Numeric		10	Right	Scale
297 w4	Numeric		10	Right	Scale
298 w	Numeric		10	Right	Scale
299 Q25MISS	Numeric		10	Right	Scale
300 filtr_Q25MI...	Numeric		10	Right	Scale
301 age	Numeric		10	Right	Scale
302 sex	Numeric	8	0	Right	Scale

→ příkazy v sekci hlavního menu **Analyze**

Popisné statistiky: **Descriptive statistics** → **Frequencies / Descriptives / Explore**

# Analýzy → *Analyze*

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



# Grafy → Graphs

The screenshot shows the PASW Statistics Viewer interface. The main window displays a bar chart titled "Graph" with the y-axis labeled "Mean" ranging from 0 to 4. The x-axis categories are "Práce", "Rodina", "Přátelé a známi", "Volný čas", "Politika", and "Náboženství". The bars represent mean values for each category. A red circle highlights the "Graphs" menu, which is open, showing various chart types like Bar, 3-D Bar, Line, Area, Pie, High-Low, Boxplot, Error Bar, Population Pyramid, Scatter/Dot, and Histogram. The "Interactive" submenu is also visible, showing more chart options.

Output  
Log  
Graph  
Title  
Notes  
Bar of mean(q30\_a, q30\_b,

GRAPH  
/BAR (SIMPLE) = I  
Legacy Dialogs

Graph

Mean

Práce Rodina Přátelé a známi Volný čas Politika Náboženství

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...
  - Drog-Line...
  - Area...
  - Pie...
  - Boxplot...
  - Error Bar...
  - Histogram...
  - Scatterplot...