

Návod pro práci s SPSS

Návody pro práci s programem SPSS pro kurz Metodologie pro Informační studia a knihovnictví 2 (jaro 2013)

Ladislava Zbiejczuk Suchá

Instalace programu

SPSS najdete v INETu. Po přihlášení se se svým UČO a sekundárním heslem najdete programy v sekci Provozní služby – Software – Nabídka softwaru.

The screenshot shows the 'Nabídka softwaru' (Software Offer) page on the INET portal. The page includes a search bar, navigation tabs (Navinky, Osobní, Personalistika, Ekonomika, Provoz), and a sidebar with 'Hledat aplikaci' and 'Provoz' sections. The main content area displays a table of software offers with columns for 'Název softwaru', 'Lokalizace', 'Popis', 'Platnost od', 'Platnost do', and 'Získat'. The 'IBM SPSS Statistics 21' offer is highlighted in yellow.

Název softwaru	Lokalizace	Popis	Platnost od	Platnost do	Získat
ACREA CR, spol. s r.o.					
IBM SPSS Data Access Pack 6.1	EN - Anglická verze	Akademická multilicence pro MU 2012			Získat
IBM SPSS Data Access Pack 6.1 with sp3	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013			Získat
IBM SPSS Modeler 14.2	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	05.01.2012	01.02.2014	Získat
IBM SPSS Modeler 15	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	23.11.2012	01.02.2014	Získat
IBM SPSS Statistics 18	EN - Anglická verze	Akademická multilicence pro MU 2009 - 2013	09.12.2009	01.02.2014	Získat
IBM SPSS Statistics 19	EN - Anglická verze	Akademická multilicence pro MU 2011 - 2013	22.12.2010	01.02.2014	Získat
IBM SPSS Statistics 20	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	05.01.2012	01.02.2014	Získat
IBM SPSS Statistics 20 Fix Pack 1 32b	EN - Anglická verze	Fix Pack 1 32b			Získat
IBM SPSS Statistics 20 Fix Pack 1 64b	EN - Anglická verze	Fix Pack 1 64b			Získat
IBM SPSS Statistics 21	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2013	23.11.2012	01.02.2014	Získat
ALTAIP, Ltd.					
Altap Salamander 2.5	NS - Nespecifikováno	Celouniverzitní licence	11.01.2008		Získat
MathWorks					
Matlab 7.13	EN - Anglická verze	Matlab 7.13 (2011b)			Získat
Matlab 8.0	EN - Anglická verze	Matlab 8.0 (2012b)			Získat
SAS Institute					
SAS 9.3	EN - Anglická verze	Akademická multilicence pro MU 2012 - 2015	15.09.2012	31.05.2015	Získat
SAS 9.3 SID files 2013	EN - Anglická verze	Licenční soubory pro SAS 9.3 pro MU 2012 - 2013	31.10.2012	31.12.2013	Získat
StatSoft					
Statistics 10 MR1	CZ - Česká verze	Jednouzivatelská verze	05.09.2012	31.12.2013	Získat
Statistics 10 MR1	EN - Anglická verze	Jednouzivatelská verze	05.09.2012	31.12.2013	Získat

Program si můžete stáhnout ve formátu ISO. Pro spuštění je tedy nutné jej vypálit na DVD nebo vytvořit virtuální disk. Při registraci nezapomeňte uvést registrační kód dostupný v INETu.

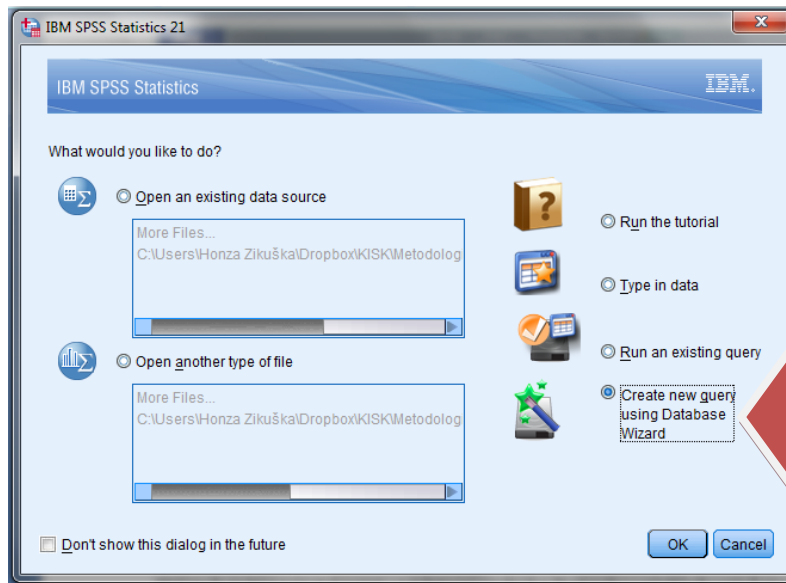
Dostupných je hned několik druhů licencí – doporučuji vybrat licenci **IBM SPSS Statistics 21** (nejnovější verze programu).

Otevření souborů s daty

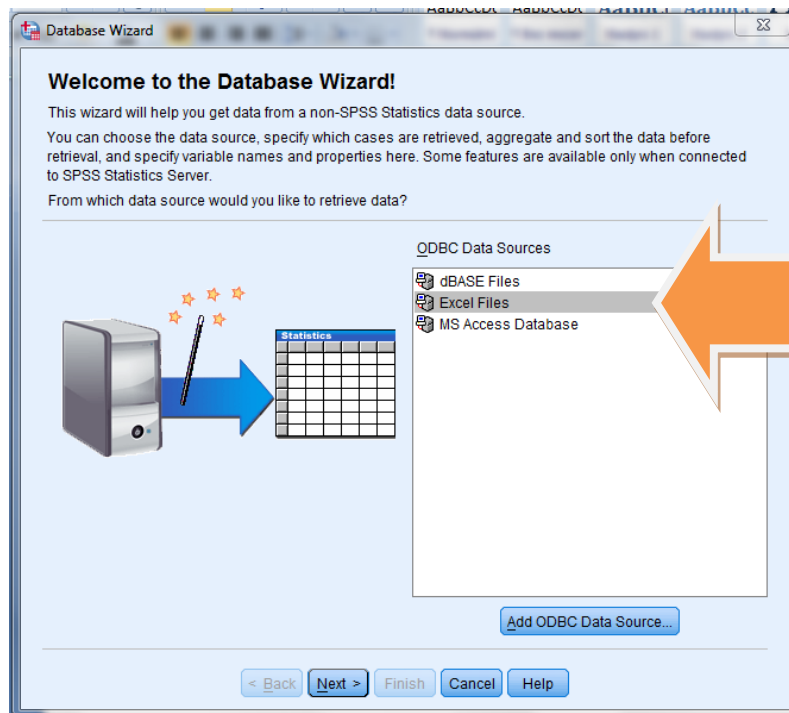
SPSS tedy máme nainstalované – najdete jej v nabídce Start nebo v přehledu vašich programů. Do SPSS můžete data dostat několika způsoby – ten nezákladnější je přímé tvoření datasetu v SPSS. My ale budeme potřebovat pracovat s daty, která již máme ve formátu .xls.

Postupovat budeme následovně:

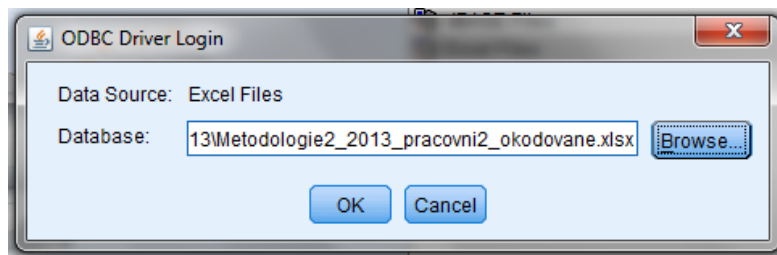
1. Uložíme si na své PC datový soubor ve formátu pro Excel (najdeme jej v ISu).
2. Pro převedení excelového souboru do souboru typu .sav spustíme „Database Wizzard“:



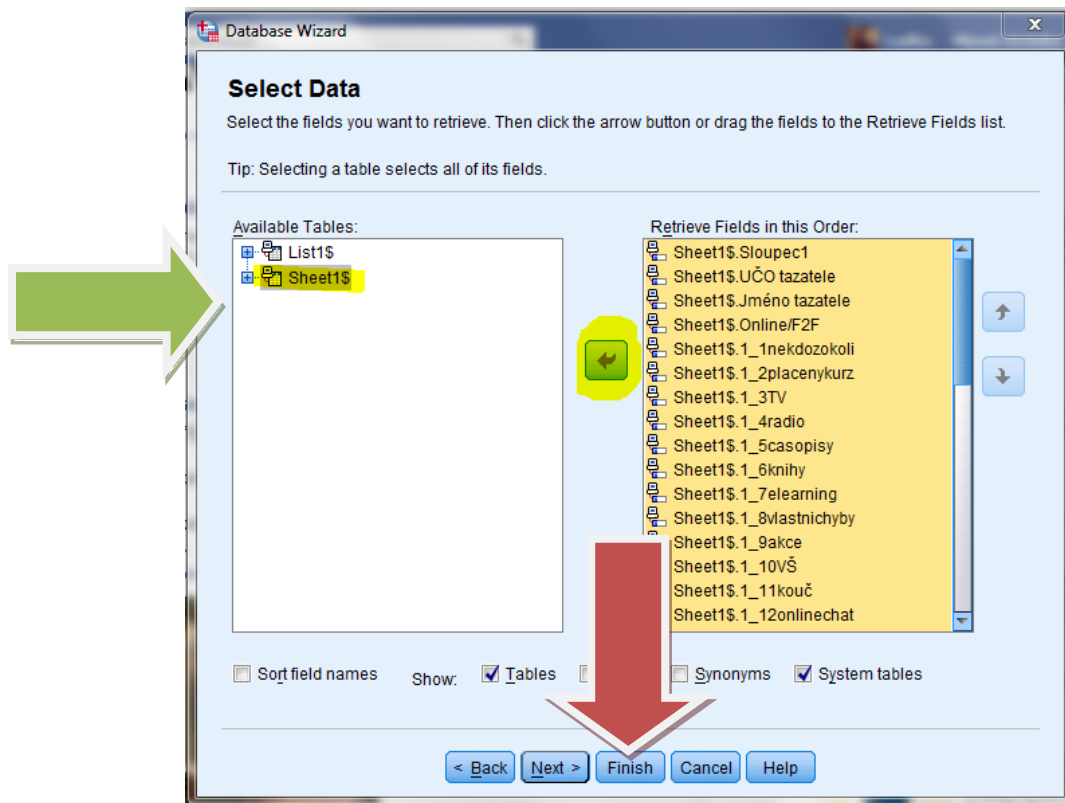
3. Z nabízených možností v dalším okně si vyberte „Excel files“:



4. Vyberte soubor ze svého PC:



5. Vyberte si oblast, kterou chcete převést a poté potvrďte stiskem „**Finish**“



6. V počítači se vám otevřou dvě nová okna. Jedno přímo s datasetem a druhé je tzv. „Output“ – okno, kam se zapisují procesy a výsledky operací SPSS.

Práce s datasetem

Dataset je neprve potřeba upravit a popsat. Všimněte si, že v SPSS lze přepínat mezi dvěma druhy zobrazení:

- pohled na data,
- pohled na proměnné.

The screenshot shows the IBM SPSS Statistics Data Editor interface. The main window displays a dataset in 'Data View' mode, which is a grid where each row represents a respondent and each column represents a variable. The variables shown are 'Sloupec1', 'UČO_tazat...', 'Jméno_tazatele', and 'OnlineF2F'. The data rows contain dates, UČO numbers, names, and online/offline status. At the bottom of the window, there are two buttons: 'Data View' (highlighted in yellow) and 'Variable View'. A large blue arrow points from the 'Data View' button towards the text below.

Pohled na data je velmi podobný tomu, co znáte z Excelu – co řádek, to respondent, co sloupec, to proměnná. **Pohled na proměnné** upřesňuje parametry jednotlivých proměnných.

Ukažme si to na příkladu této otázky:

- 2. Považujete obor Informační studia a knihovnictví za perspektivní?

- velmi perspektivní
- spíše perspektivní
- spíše neperspektivní
- zcela neperspektivní
- nevím, nemohu odpovědět
- neodpověděl/a

- 1
- 2
- 3
- 4
- 1
- 2

Hodnoty proměnné okódované

Chybějící hodnoty (missing values)

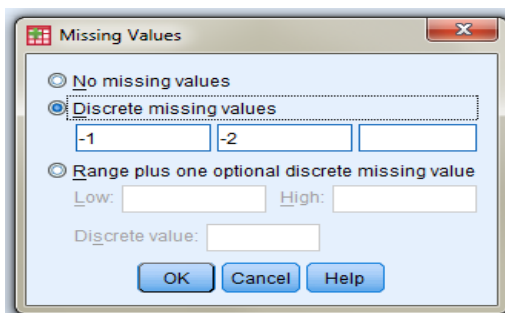
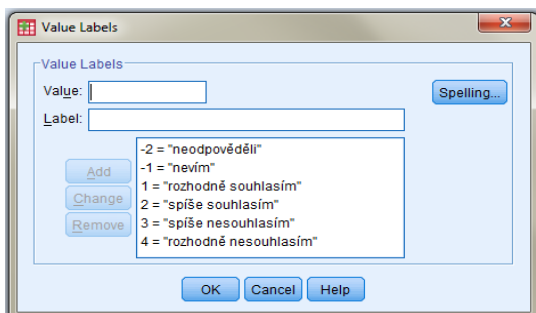
Takto bude vypadat matice dat:

Q1_prinos
Studium na
KISK hodnotim
jako:

	q1_prinos	q2_perspektiva	q3_doplnění	q4_znovustud	q5_eal	q8_1_osl	q8_2_pr...	q8_3_tema	q8_4_pocetkurzu	q8_5_pra	q9_1_navhkurzu	q9_2_na	q9_3_na	q10
1	1													
2	1													
3	2													
4	1	2	1	1	1									
5	1	2	2	1	1									
6	2	1	1	1	2									
7	1	1	2	1	1									
8	1	1	1	1	1									
9	2	2	1	1	1									
10	2	2	4	2	2									
11	2	2	4	2	2									
12	2	1	1	1	2									
13	1	2	2	1	1									
14	1	2	2	1	1									
15	2	1	1	1	1									
16	1	2	2	2	2									
17	2	2	2	1	2									
18	2	2	2	2	1									
19	2	2	2	1	2									
20	2	2	2	1	1									
21	2	2	2	1	2									
22	2	2	2	1	1									
23	2	2	1	1	1									

Zároveň je potřeba popsat jednotlivé proměnné na kartě **Variable view**:

- **Name:** zkrácené označení proměnné.
- **Typ:** číselné/slovní (SPSS potřebuje vědět, jaké operace může provádět s jednotlivými proměnnými)
- **Decimal:** desetinná místa (pouze kardinální proměnné) – automaticky jsou nastavena dvě desetinná místa, snižte si jejich počet na 0.
- **Label:** většinou kopírujeme znění otázky.
- **Value labels:** hodnoty proměnné – popíšeme všechny hodnoty proměnné včetně „missing values“
- **Missing values:** které hodnoty nezahrnujeme do dané analýzy – SPSS s nimi v konkrétních operacích nebude počítat.
- **Measure:** typ proměnné (nominální/ordinální/kardinální)

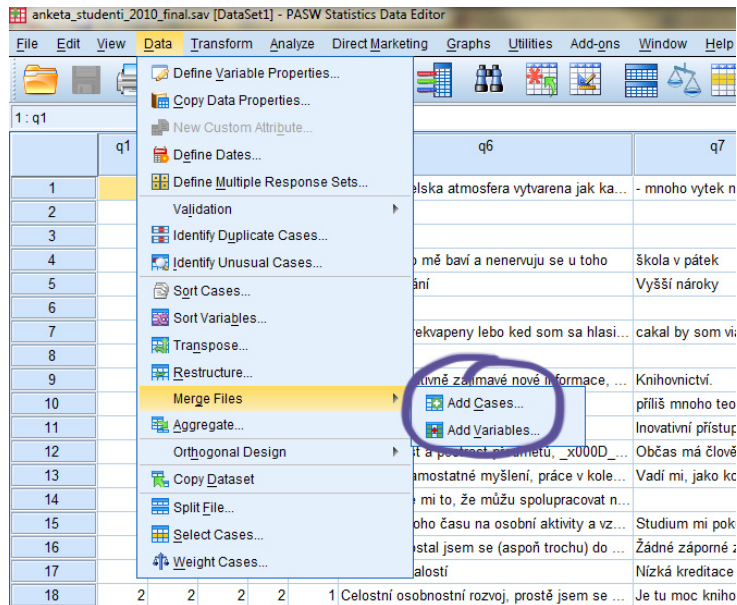


Ve studijních materiálech v ISu máte již datasety s popsanými proměnnými.

Slučování datových souborů

Někdy potřebujeme sloučit více datových souborů. Máme na výběr dvě varianty:

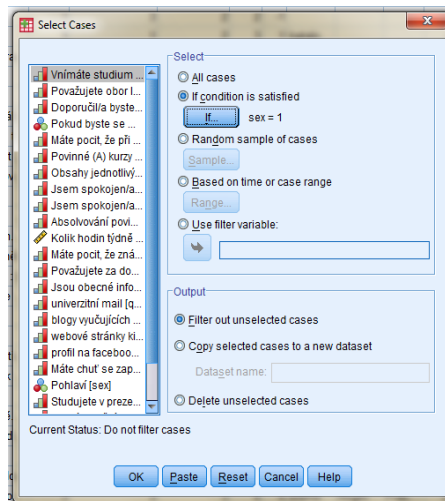
- Chceme sloučit více dat o stejných případech: Merge Files → Add variables
- Chceme sloučit soubory s různými jednotkami a stejnými proměnnými Merge Files → Add Cases



Výběr případů

Někdy naopak potřebujeme pracovat jen s některými případy (například se ženami):

- **Data → Select Cases**
- Lze vybírat náhodně nebo dle kritéria – pokud např. chceme pracovat jen s muži, pak musíme použít proceduru IF



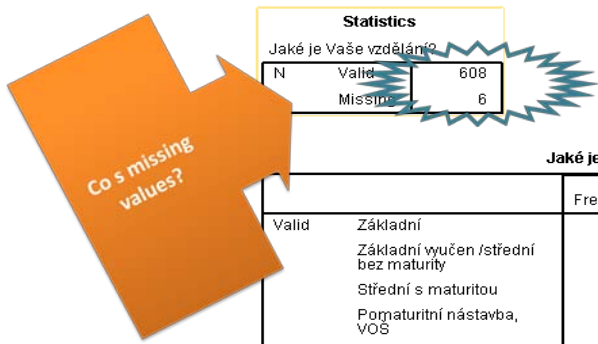
Kontrola dat

V SPSS probíhá kontrola dat se stejnou logikou jako v jakémkoliv jiném programu. Její provedení je jen jednodušší, protože SPSS je přizpůsobeno na provádění statistických operací. SPSS má také tu výhodu, že nám v Outputu dává tabulky již v té podobě, v jaké by se měly objevit v odborné práci – tedy kompletní tabulky četností s nevalidními validními absolutními i relativními hodnotami.

Pro použití v odborné práci je pouze třeba **přeložit popisky tabulek**.

Kontrola kategorizovaných dat

SPSS nám prostřednictvím jednoduchého příkazu **Analyze → Descriptive Statistics → Frequencies** (zde si vyberete konkrétní proměnnou) vrátí počet validních a nevalidních hodnot proměnných. Výsledky najdeme v okně Output:



Statistics
Jaké je Vaše vzdělání?
N Valid 608
Missing 6

Co s missing values?

Jaké je Vaše vzdělání?

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Základní	46	7,5	7,6	7,6
	Základní vyučen /střední bez maturity	62	10,1	10,2	17,8
	Střední s maturitou	307	50,0	50,5	68,3
	Pomaturitní nástavba, VOS	40	6,5	6,6	74,8
	Vysokoškolské	153	24,9	25,2	100,0
Total		608	99,0	100,0	
Missing	System	6	1,0		
	Total	614	100,0		

Stejně jako v případě SPSS nás bude zajímat výpis četností jednotlivých výskytů hodnot proměnné. Zde máme příklad chybného zápisu jména studentky či chybného zápisu v proměnné „pohlaví“:



Chybný zápis jména

Jméno výzkumníka

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Andrea Szászová	10	1,6	1,6	1,6
	Babara Ondrušová	1	,2	,2	1,8
	Barbora Ondrušová	9	1,5	1,5	3,3
	Barbora Pitašová	10	1,6	1,6	4,9
	Blanka Justová	10	1,6	1,6	6,5
	Blanka Svobodová	10	1,6	1,6	8,1
	Dagmar Chládková	11	1,8	1,8	9,9
	Dagmar Šíková	10	1,6	1,6	11,6
	Dalibor Bláha	10	1,6	1,6	13,2
	Daniela Králová	10	1,6	1,6	14,8

Statistics

Pohlaví		
N	Valid	613
	Missing	1

Pohlaví

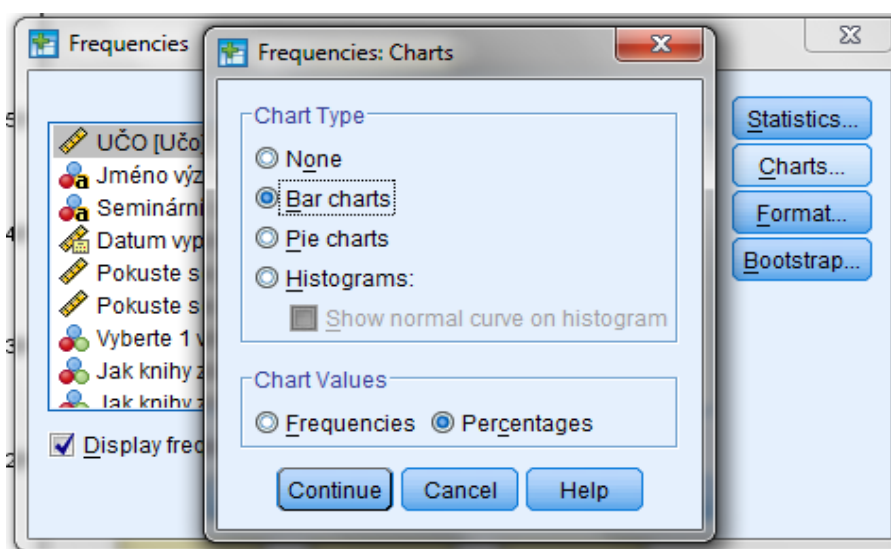
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Muž	279	45,4	45,5	45,5
	Žena	333	54,2	54,3	99,8
	7	1	,2	,2	100,0
	Total	613	99,8	100,0	
Missing	System	1	,2		
Total		614	100,0		

Chyba:
proměnná
„Pohlaví“ by
neměla
nabývat
hodnoty 7

Poté co naleznete chybná data, můžete je v datasetu vyhledat pomocí příkazu CTRL+F stejně jako v Excelu.

Tabulky četností a grafy v SPSS

Tabulky četností v SPSS získáme příkazem Analyze → Descriptive Statistics → Frequencies . Grafy vytvoříme cestou Analyze → Descriptive Statistics → Frequencies → **Charts**.



Modus a medián v SPSS

Modus, medián a aritmetický průměr jednoduše získáte v SPSS touto cestou:

Analyze → Descriptive Statistics → Frequencies → **Statistics** → **Mean, Median, Mode**