

Jak pracovat s Data Explorerem v katalogu Dataverse ČSDA

Obsah

1. [Co je Data Explorer?](#)
2. [Možnosti stažení souboru](#)
3. [Proměnná a informace, které se zobrazují](#)
4. [Cross Tabulation a její funkce](#)

1. Co je Data Explorer?

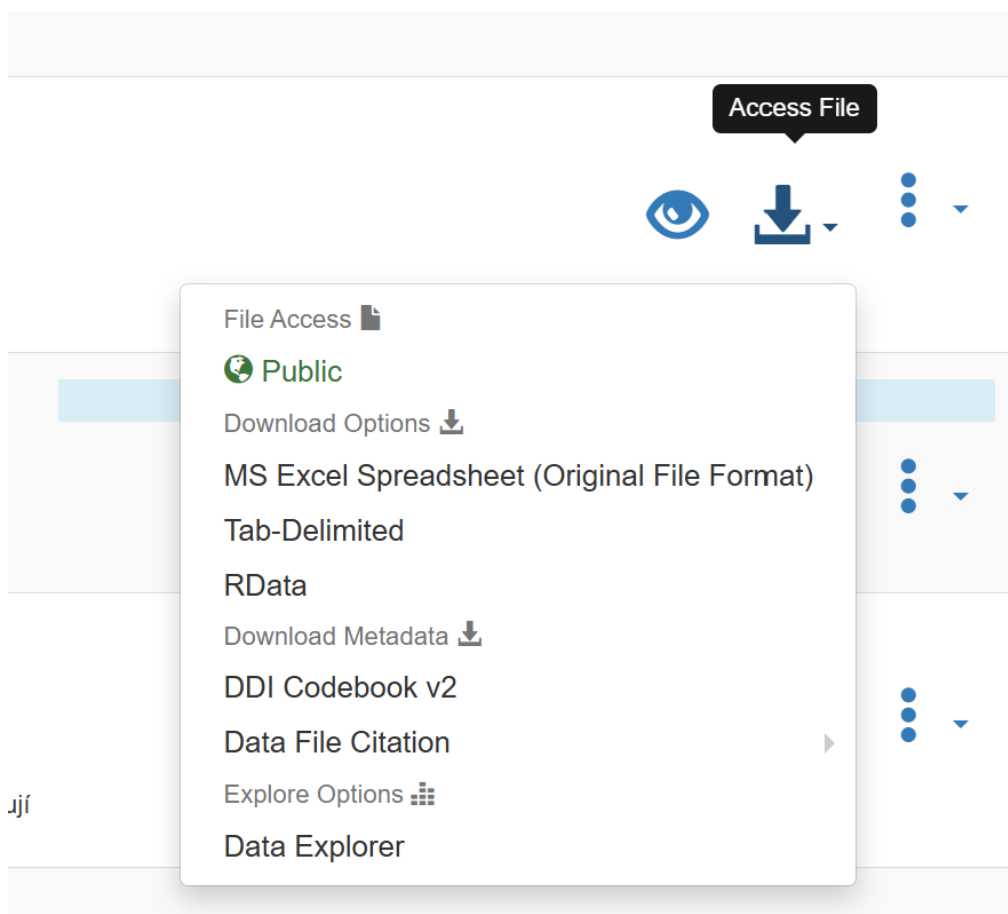
Data Explorer je webový nástroj pro rychlé prohlížení a základní exploraci datového souboru přímo v repozitáři. Uživatel nemusí nejprve stahovat data do statistického programu: může si prohlédnout seznam proměnných, jejich popisky, nastavení vah, vybrat proměnné a vytvořit kontingenční tabulku.

Explorer funguje pouze u tabulkových souborů, které Dataverse po nahrání převedl do svého tabulkového formátu (přípona **.tab**). Takto se zpracovávají soubory ve formátech SPSS, Stata, R, Excel a CSV. U ostatních souborů (dotazníky v PDF, technické zprávy, obrázky apod.) nabídku Data Explorer nevidíte.

Jako pracovní příklad pro tento návod jsme zvolili Průzkum správy výzkumných dat v České republice 2025 (<https://doi.org/10.14473/CSDA/OJTMJF>).

Jak Data Explorer spustím?

Nástroj spustíte volbou „Data Explorer“ z rozbalovací nabídky, která se otevře po kliknutí na ikonu „Access File“ (se šipkou). V případě Průzkumu správy výzkumných dat vidí možnost „Data Explorer“ všichni, protože jeho data jsou dostupná bez jakéhokoli omezení. Pokud jde o datový soubor s omezeným přístupem, musíte se nejprve přihlásit a mít k datům udělený přístup, abyste mohli s daty v Exploreru pracovat.



Hlavní obrazovka

Data Explorer

Průzkum správy výzkumných dat v České republice 2025

Vávra, Martin; Čížek, Tomáš; Klusáček, Jan; Sukhotskaya, Aliona; Salamounová, Petra, 2026, "Průzkum správy výzkumných dat v České republice 2025", <https://doi.org/10.14473/CSDA/OJTMJF>, CSDA, V4, UNF-6-8mvEZd8eysKIQqW206Hkw== [fileUNF]

Variables Cross Tabulation

6 Variables / Cross Tabulation

Groups

Download

Items per page 10 1 - 10 of 213 vari

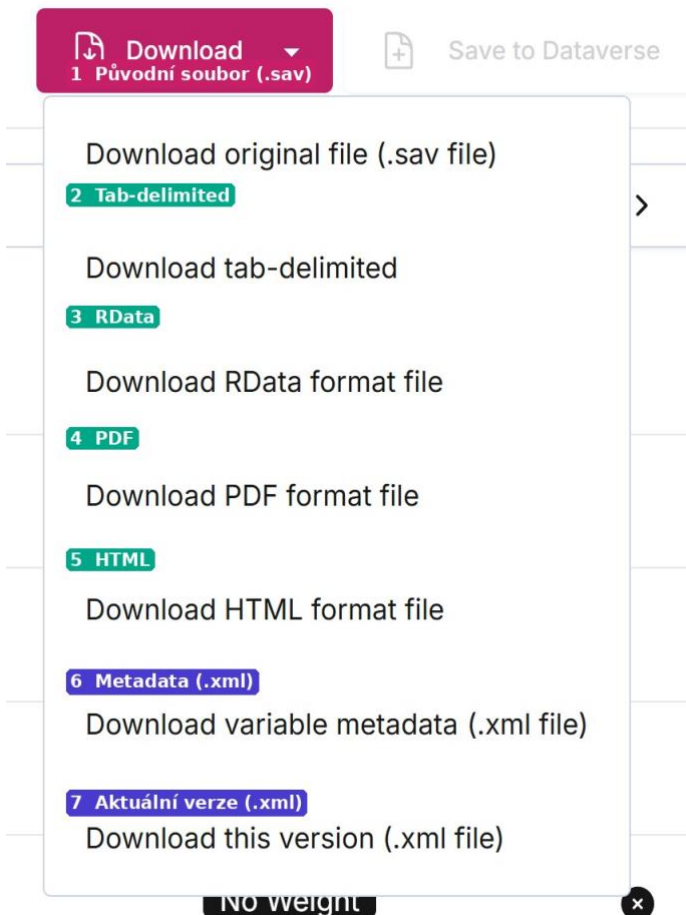
All Variables	1 Seznam proměnných	2 Stažení	Weight	Options
No Groups	3 ID / název / popis	4 Váha		
	☐ v239645 id id		No Weight	
	☐ v239604 sent Datum zaslání pozvánky		No Weight	
	☐ v239480 remindersent Datum odeslání poslední upomínky		No Weight	
	☐ v239484 remindercount Celkový počet zasláných připomínek		No Weight	
	☐ v239674 completed Hotovo		No Weight	
	☐ v239603 lastpage lastpage		No Weight	
	☐ v239599 startlanguage startlanguage		No Weight	
	☐ v239583 startdate startdate		No Weight	
	☐ v239504 datestamp datestamp		No Weight	
	☐ v239492 Q01 Jaka je velikost týmu uvedeného projektu podle počtu členů?		No Weight	

Na hlavní obrazovce je vidět zejména seznam proměnných se sloupci ID, Name, Label, Weight a Options. Nahoře jsou záložky Variables a Cross Tabulation a tlačítko Download.

Prakticky: Data Explorer je vhodný pro orientaci v datech, kontrolu kódování a rychlé vytvoření kontingenční tabulky. Pro pokročilé statistické analýzy je vhodné data stáhnout a zpracovat ve statistickém programu.

2. Možnosti stažení souboru

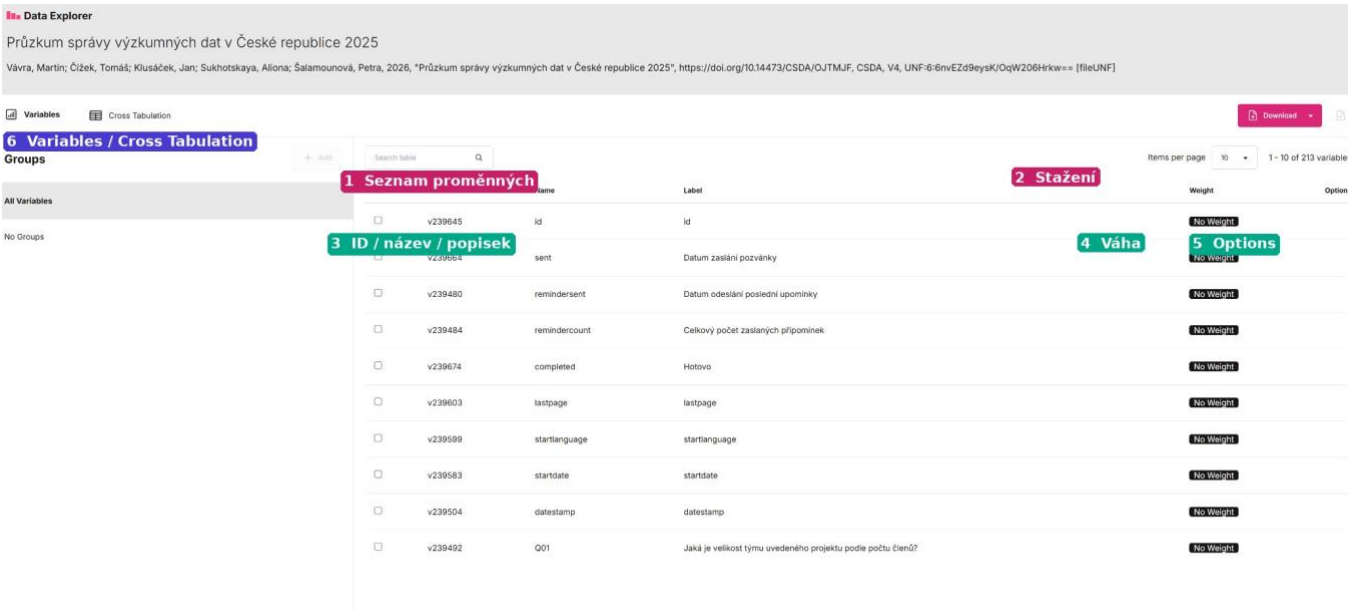
Nabídka Download v horní části Data Exploreru poskytuje několik možností, jak získat data nebo metadata v různých formátech.



Možnost	Co znamená
Download original file (.sav file)	Stážení původního datového souboru ve formátu SPSS (.sav).
Download tab-delimited	Stážení dat v textovém formátu s hodnotami oddělenými tabulátorem.
Download RData format file	Stážení dat ve formátu RData pro další práci v R.
Download PDF format file	Stážení codebooku (popisu proměnných) ve formátu PDF.
Download HTML format file	Stážení codebooku (popisu proměnných) ve formátu HTML.
Download variable metadata (.xml file)	Stážení metadat k proměnným ve formátu XML (standard DDI).

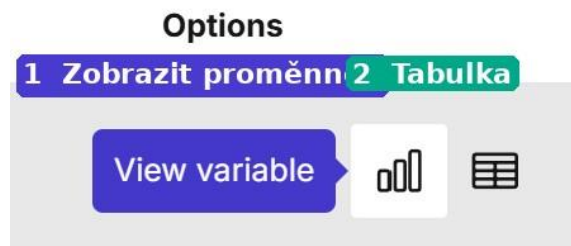
3. Proměnná a informace, které se zobrazují

V záložce Variables je každá proměnná zobrazena v samostatném řádku. Tato část slouží k orientaci v datovém souboru a k výběru proměnných pro další práci. Můžeme ji využít i k vyhledávání proměnných v datovém souboru.



Položka	Význam
ID	Jedinečný identifikátor proměnné v datovém souboru.
Name	Název proměnné používaný v datech, např. Q01 nebo datestamp.
Label	Popisek / slovní označení proměnné; v ukázce je u Q01 přímo text otázky.
Weight	Informace o použité váze. V ukázce je uvedeno No Weight.
Options	Akce dostupné pro danou proměnnou.

Funkce v nabídce Options



Ikona grafu (View variable) otevře detail konkrétní proměnné. Vedlejší ikona tabulky (Add to cross tabulation) přidá proměnnou do kontingenční tabulky. Přesný obsah detailu závisí na typu proměnné a dostupných metadatech.

Detail proměnné – View variable

Po kliknutí na View variable se otevře samostatné okno s podrobnějším přehledem vybrané proměnné. U kategoriálních proměnných, jako je Q01 na ukázce níže, obsahuje nejen základní identifikaci proměnné, ale také grafické a tabulkové zobrazení rozložení odpovědí.

View variable



 EDIT



ID: v239492

Name: Q01

Label: Jaká je velikost týmu uvedeného projektu podle počtu členů?

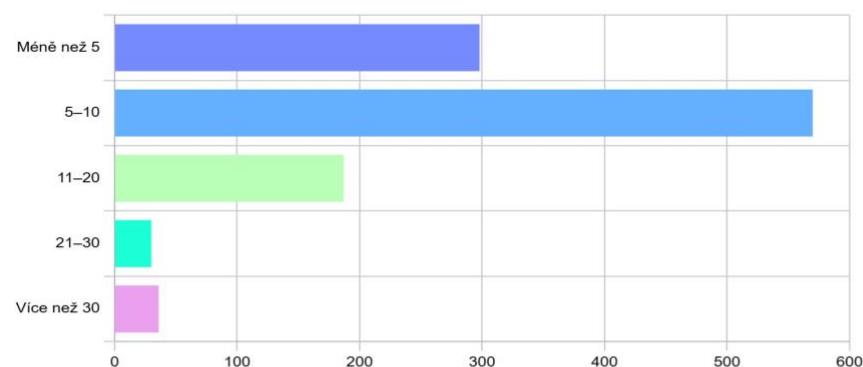


Chart data

Show in graph	Values	Categories	Count	Count (%)	Weighted count	Weighted (%)
☒	1	Méně než 5	298	26.58	None	-∞

Jaké informace a statistiky View variable poskytuje

Obsah detailu se odvíjí od typu proměnné a dostupných metadat. V našem příkladu získáme následující informace:

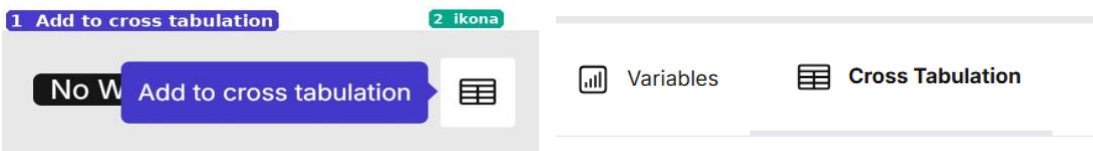
Prvek	Co uživatel vidí / k čemu slouží
ID	Jedinečný identifikátor proměnné, např. v239492.
Name	Název proměnné v datovém souboru, zde Q01.
Label	Slovní popisek proměnné; v ukázce je zde formulace otázky.
Graf rozložení	Graficky zobrazuje jednotlivé kategorie a jejich četnosti. V ukázce jde o horizontální sloupcový graf.
Show in graph	Zaškrtačací pole určující, zda má být daná kategorie zahrnuta do grafu.
Values	Kód / hodnota kategorie, např. 1.
Categories	Textový popis kategorie, např. „Méně než 5“.
Count	Absolutní počet pozorování v dané kategorii.
Count (%)	Procentuální vyjádření četnosti dané kategorie.
Weighted count	Vážený počet, pokud je pro analýzu použita váha. Při absenci váhy není tato hodnota věcně použitelná.
Weighted (%)	Vážený procentní podíl kategorie, pokud je použita váha.

Kromě toho získáme i popisné statistiky jednotlivých proměnných. U kategoriálních proměnných (jako je zde do kategorií převedený počet členů vědeckých týmů) jsou průměr a směrodatná odchylka vypočítány z číselných kódů a obvykle nemají smysluplný výklad; důležité jsou četnosti. U číselných proměnných (např. věk v letech, nekategorizovaný příjem) jsou naopak klíčové popisné statistiky.

4. Cross Tabulation – funkce a postup

Přidání proměnné do Cross Tabulation

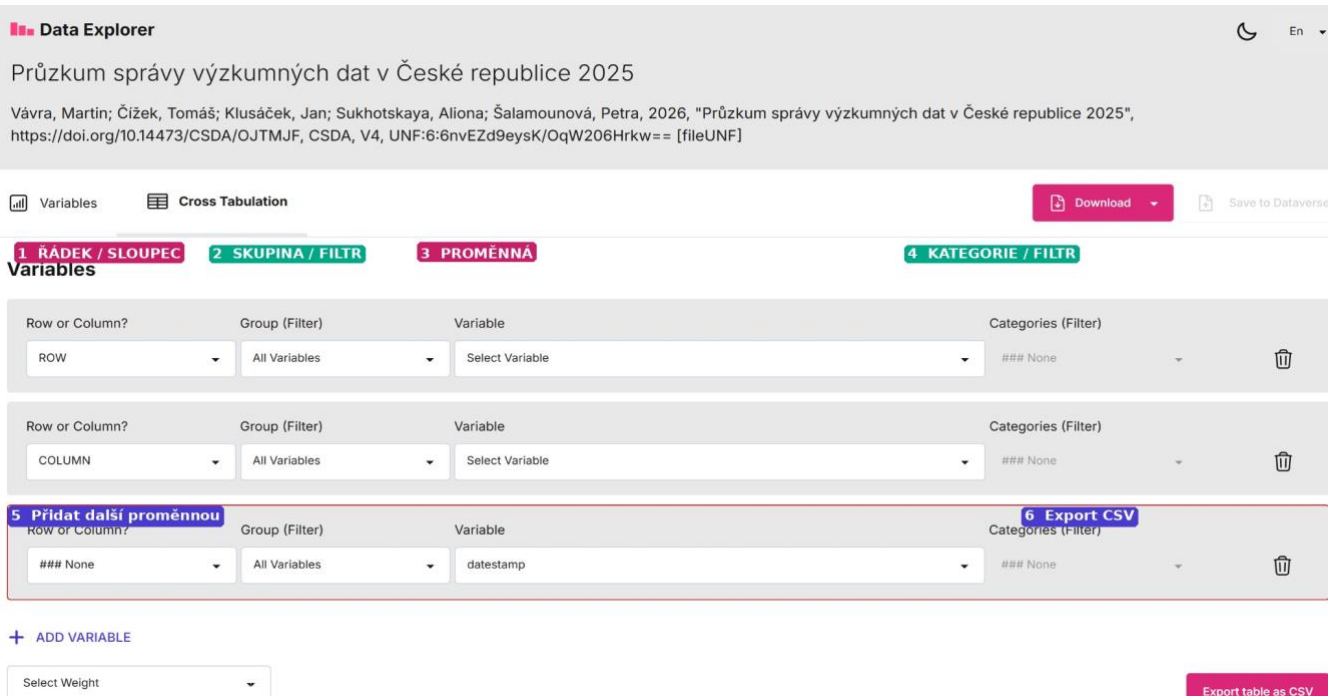
Akce Add to cross tabulation přenese vybranou proměnnou do části určené pro tvorbu kontingenční tabulky.



Poté, co proměnné do kontingenční tabulky přidáme, musíme se přepnout do jejího zobrazení, což lze udělat kliknutím na „Cross Tabulation“ na obrazovce vlevo pod názvem výzkumu.

Funkce	Popis
Row or Column?	Určuje, zda bude proměnná použita jako řádková nebo sloupcová proměnná.
Group (Filter)	Filtr skupiny pro zúžení nabídky proměnných.
Variable	Výběr konkrétní proměnné pro kontingenční tabulku.
Categories (Filter)	Výběr kategorií, které mají být ve výsledku zahrnuty.
Delete	Odstranění daného řádku s proměnnou.
+ ADD VARIABLE	Přidání další proměnné.
Select Weight	Výběr váhy, pokud je v datasetu dostupná.
Export table as CSV	Export výsledné kontingenční tabulky do CSV.

Cross Tabulation umožňuje sestavit kontingenční tabulku z vybraných proměnných. Pro každou přidanou proměnnou lze nastavit její roli a případné filtry.

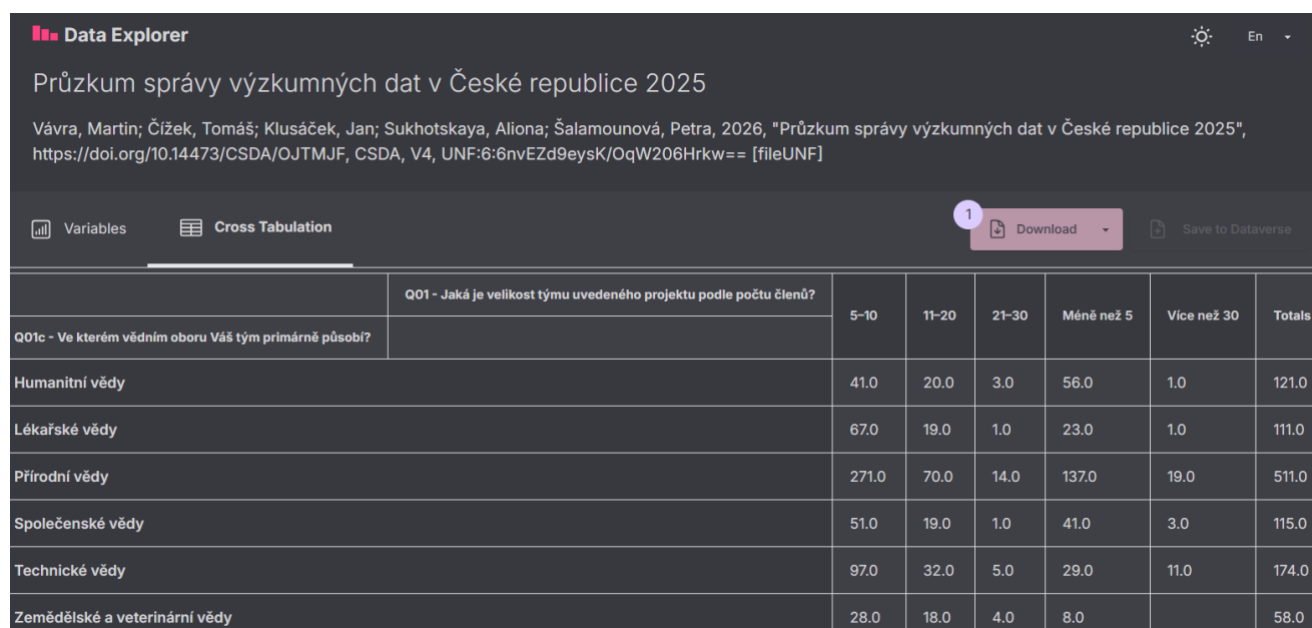


Jak vytvořit kontingenční tabulku

1. Otevřete Cross Tabulation nebo použijte Add to cross tabulation.
2. Nastavte první proměnnou jako ROW.
3. Nastavte druhou proměnnou jako COLUMN.
4. Podle potřeby použijte Categories (Filter) nebo váhu.
5. Další proměnné přidáte pomocí ADD VARIABLE.
6. Výsledek lze exportovat tlačítkem Export table as CSV do souboru CSV.

Při práci s kontingenční tabulkou je nutné zohlednit kódování proměnných, chybějící hodnoty a případné vážení dat.

Následující obrázek ukazuje výslednou kontingenční tabulku velikosti projektového týmu podle oboru, ve kterém respondent působí.



Q01 - Jaká je velikost týmu uvedeného projektu podle počtu členů?		5-10	11-20	21-30	Méně než 5	Více než 30	Totals
Q01c - Ve kterém vědním oboru Váš tým primárně působí?							
Humanitní vědy		41.0	20.0	3.0	56.0	1.0	121.0
Lékařské vědy		67.0	19.0	1.0	23.0	1.0	111.0
Přírodní vědy		271.0	70.0	14.0	137.0	19.0	511.0
Společenské vědy		51.0	19.0	1.0	41.0	3.0	115.0
Technické vědy		97.0	32.0	5.0	29.0	11.0	174.0
Zemědělské a veterinární vědy		28.0	18.0	4.0	8.0		58.0

Rychlá orientace – co kde najdete

Potřebuji...	Kde to najdu
stáhnout původní data	Download → Download original file (.sav)
stáhnout data pro R	Download → Download RData format file
získat metadata	Download → Download variable metadata (.xml file)
najít proměnnou	Záložka Variables a vyhledávání / seznam proměnných
prohlédnout proměnnou	Options → View variable
přidat proměnnou do kontingenční tabulky	Add to cross tabulation
vytvořit kontingenční tabulku	Záložka Cross Tabulation
nastavit řádek/sloupec	Pole Row or Column?
omezit kategorie	Pole Categories (Filter)
přidat další proměnnou	ADD VARIABLE
exportovat kontingenční tabulku	Export table as CSV