

# **Metodologie pro Informační studia a knihovnictví 2**

## **Modul 7: Třídění druhého stupně. Kontingenční tabulky**

### **Co se dozvíte v tomto modulu?**

- Co je třídění druhého stupně
- Jak vytvořit a interpretovat kontingenční tabulku

## Třídění druhého stupně

Doposud jsme se zabývali jen **popisem jednotlivých proměnných** – prováděli jsme tzv. třídění prvního stupně. Často jsou pro nás ale mnohem zajímavější data, která vzniknou tzv. **tříděním druhého stupně**, ve kterém se porovnávají dvě proměnné.

Třídění druhého stupně se používá například:

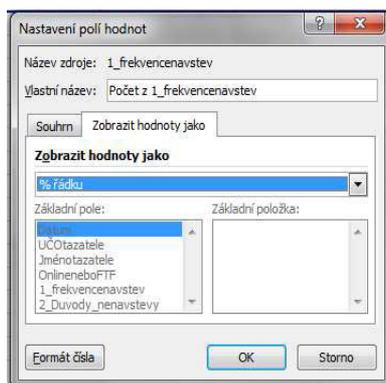
- chceme-li zjistit, zda odpovídali různě muži a ženy,
- chceme-li zjistit, zda jsou rozdíly v odpovědích respondentů dle věku,
- chceme-li zjistit, zda jsou rozdíly v odpovědích respondentů dle vzdělání,
- chceme-li zjistit, zda jsou rozdíly v odpovědích respondentů dle postojů k jinému problému.

Pro třídění druhého stupně se používá speciální tabulka četností – tzv. **kontingenční tabulka** (v Excelu funkce pivot table, v SPSS Crosstabs).

*Příklad: Chceme zjistit, zda v našem souboru existují rozdíly v tom, jak často chodí do knihoven muži a ženy.*

Počet z 1_frekvencenavstev	1_fré	1	2	3	4	5	6	7	Celkový součet
13_pohlavi	1	18	13	31	25	19	33	96	235
	2	38	25	103	38	24	22	82	332
Celkový součet		56	38	134	63	43	55	178	567

*Mužů a žen bylo ale v souboru rozdílné množství!! **Abychom mohli odpovědi porovnat, potřebujeme znát relativní četnosti!!!** Klikneme na data pravým tlačítkem myši a vybereme v nabídce Zobrazit hodnoty jako – řádková/sloupcová procenta.*



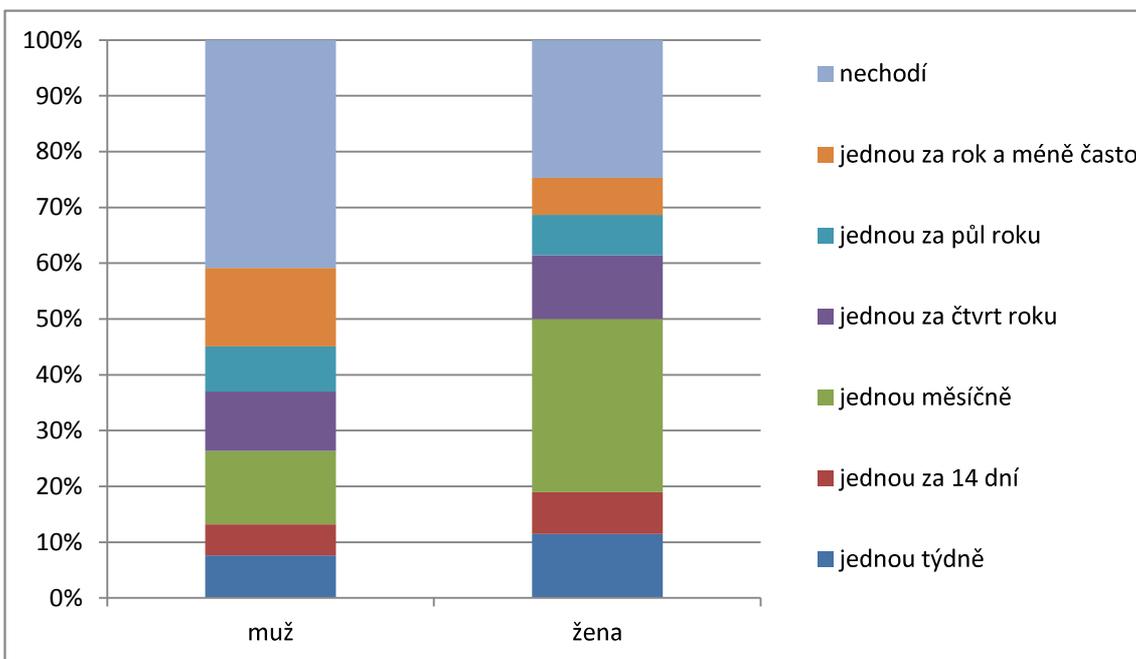
Nyní dostaneme tabulku, kde již hodnoty u mužů a žen můžeme porovnávat – nejsou totiž závislé na tom, že mužů a žen odpovídá různý počet.

Počet z 1_frekvencenavstev	1_frekve	1	2	3	4	5	6	7	Celkový součet
13_pohlavi	1	7,66%	5,53%	13,19%	10,64%	8,09%	14,04%	40,85%	100,00%
	2	11,45%	7,53%	31,02%	11,45%	7,23%	6,63%	24,70%	100,00%
Celkový součet		9,88%	6,70%	23,63%	11,11%	7,58%	9,70%	31,39%	100,00%

Tuto tabulku můžeme číst následujícím způsobem:

- 7,66 % mužů ve vzorku chod do knihovny jednou týdně. 5,53 % mužů chodí do knihovny jednou za 14 dní.
- 11,45 % žen chodí do knihovny jednou týdně. 7,53 % žen chodí do knihovny jednou za 14 dní.
- Atd...

Abychom mohli lépe rychle porovnat rozdíly mezi muži a ženami, necháme si zobrazit skládaný graf:



Zde vidíme patrné rozdíly. Ještě markantněji by byly patrné, kdybychom si překódovali proměnnou frekvence návštěv knihovny např. na pravidelné návštěvníky (hodnoty 1,2,3 – tedy chodí alespoň jednou za měsíc), nepravidelné návštěvníky (hodnoty 4, 5, 6) a nenávštěvním (hodnota 7). Tady se rozdíly pohybují již v desítkách procent.

## Tvorba a interpretace kontingenční tabulky

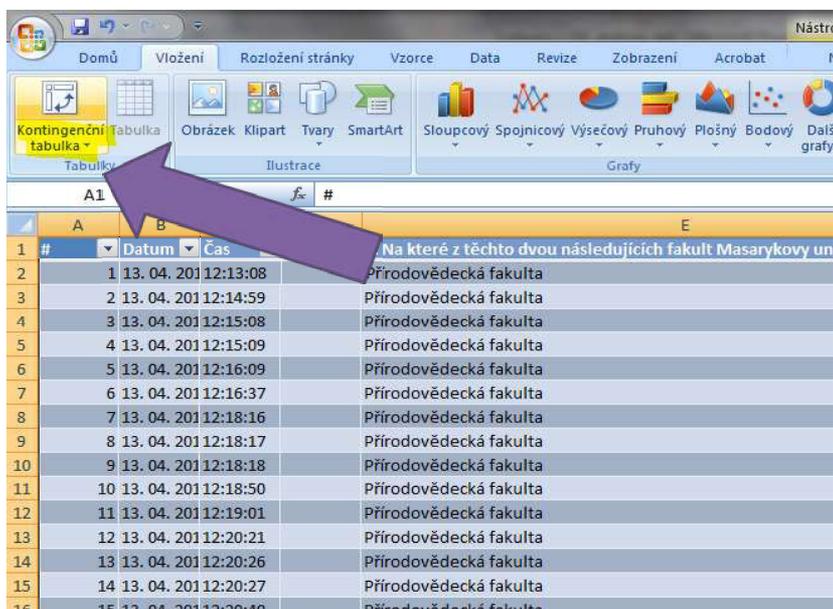
Vytvoření kontingenční tabulky je velmi jednoduché, je však třeba mít na paměti několik základních pravidel:

1. Důležité jsou pro nás **relativní četnosti**. Absolutní hodnoty jsou závislé na zastoupení jednotlivých skupin respondentů ve výběrovém vzorku.
2. Musíme určit, kterou proměnnou považujeme za **závislou** a kterou za **nezávislou**.
3. **Je-li nezávislá proměnná ve sloupcích, porovnáváme sloupcová procenta. Je-li nezávislá proměnná v řádcích, porovnáváme řádková procenta.**



## Kontingenční tabulka v Excelu

V Excelu budeme opět používat funkci Pivot tables (Kontingenční tabulka).



#	Datum	Čas	Na které z těchto dvou následujících fakult Masarykovy un
1	13. 04. 201	12:13:08	Přírodovědecká fakulta
2	13. 04. 201	12:14:59	Přírodovědecká fakulta
3	13. 04. 201	12:15:08	Přírodovědecká fakulta
4	13. 04. 201	12:15:09	Přírodovědecká fakulta
5	13. 04. 201	12:16:09	Přírodovědecká fakulta
6	13. 04. 201	12:16:37	Přírodovědecká fakulta
7	13. 04. 201	12:18:16	Přírodovědecká fakulta
8	13. 04. 201	12:18:17	Přírodovědecká fakulta
9	13. 04. 201	12:18:18	Přírodovědecká fakulta
10	13. 04. 201	12:18:50	Přírodovědecká fakulta
11	13. 04. 201	12:19:01	Přírodovědecká fakulta
12	13. 04. 201	12:20:21	Přírodovědecká fakulta
13	13. 04. 201	12:20:26	Přírodovědecká fakulta
14	13. 04. 201	12:20:27	Přírodovědecká fakulta
15	13. 04. 201	12:20:40	Přírodovědecká fakulta

Nejprve si musíme vybrat, jaké proměnné se budou zobrazovat v řádcích a jaké ve sloupcích. Neexistuje jednoznačný úzus (např. nezávislé proměnné v řádcích). Pokud má jedna z proměnných mnoho variant, je vhodné ji umístit do řádků (umístění do sloupců by znesnadňovalo orientaci v tabulce, případně by se tabulka musela umístit do listu s horizontální orientací).

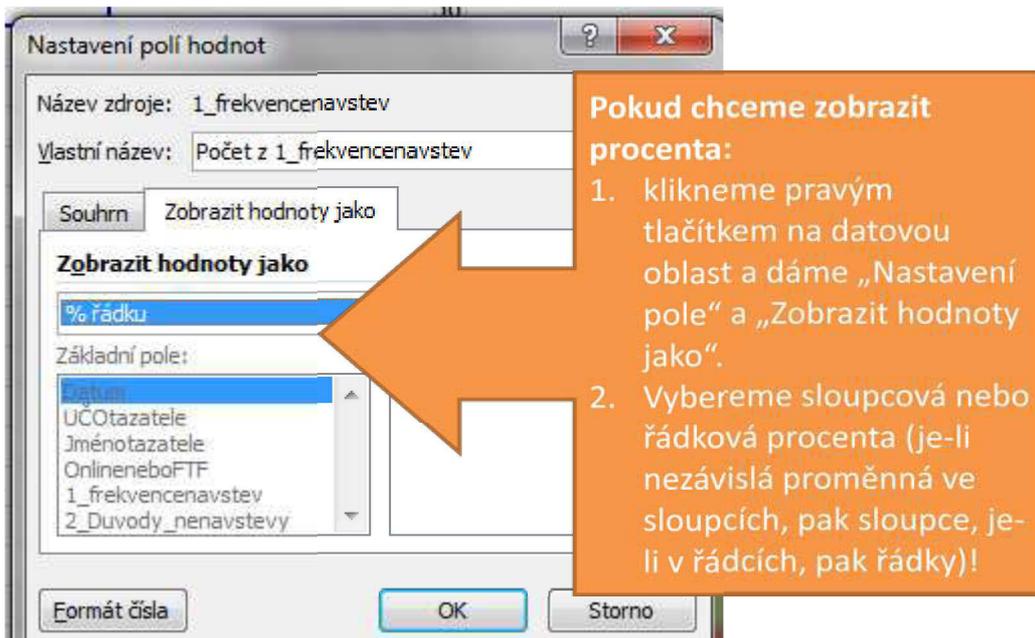
Zde vybíráme proměnné a přetahujeme je do políček „popisky řádků“ a „popisky sloupců“. Nakonec vybereme, co se bude zobrazovat v políčku „hodnoty“.

Filtrem si můžete „vyfiltrovat“ odpovědi – např. pokud chcete zobrazit jen odpovědi u žen, pak zadáte do filtru pohlaví a v tabulce nastavíte

Pokud v datech zůstaly nevalidní hodnoty (missing values), je lepší je pro analýzu vyřadit.

Počet z 1 frekvencenavstev	1_frekvencenavstev	
13_pohlavi	jednou týdně	jedn
muž		18
žena		38
Celkový součet		56

V tuto chvíli máme tabulku s absolutními četnostmi. Potřebujeme však tabulku, kde budou uvedeny i **četnosti relativní**. Kliknete pravým tlačítkem na datovou oblast a nastavíme si zobrazení polí hodnot. Podle toho, kde máme nezávislou proměnnou, vybereme řádková či sloupcová procenta.



## ***Kontingenční tabulka v SPSS***

V SPSS jsou kontingenční tabulky v záložce Analyze → Descriptive Statistics → Crosstabs. Zobrazení řádkových a sloupcových procent se nastavuje v nabídce „Cells“.

**SPSS umí generovat tabulku obsahující dohromady jak absolutní, tak relativní četnosti.**