

Datová matice

Cílem je převést tu kupu do podoby jedné datové matice proměnné X objekty (lidé).

- 1 proměnná = 1 sloupec; 1 člověk(objekt) = 1 řádek

Převod by měl být tak kompletní, abychom mohli dotazníky zahodit (tedy přesněji archivovat).

1. Pojmenovat (identifikovat) proměnné.

- vytvořit si systém, kterého se budeme držet
- max 8 znaků, začít písmenem, používat řádové nuly (p01), vyhýbat se diakritice a mezerám
- vždy začít identifikační proměnnou
- příslušnost ke skupině je také proměnnou

2. Kódování – jak budeme zapisovat jednotlivé odpovědi – hodnoty proměnné

- Vytvoříme si systém
 - snažíme se udržet maximum informace (redukce je možná kdykoli později)
 - veškerá data by měla mít numerickou podobu
 - volné výpovědi opišeme a budeme kódovat později
 - je-li proměnná typu „zaškrtněte vše, co platí“, každá možnost je jednou proměnnou
 - vytvoříme kódy i pro chybějící či špatné odpovědi¹ (čísla mimo škálu, např. 99); v datové matici by neměly být prázdné buňky
 - vybrat si systém přidělování arbitrárních hodnot: vyšší číslo/vyšší souhlas vs. stoupáme zleva doprava
 - problémy bývají s časovými údaji – je praktické je rozdělit či jinak převést na obyčejná čísla (den, měsíc, rok; hodiny, minuty)
- Zapišeme si to všechno do vzorového dotazníku, ideálně způsobem, kterému porozumí i druhý.
- Jména proměnných vložíme do (pouze) 1. řádku. Můžeme použít komentář na delší popis proměnné a kódování hodnot. Pro stejný účel lze využít i jiný list. V prvním sloupci je identifikační proměnná. Počítat s omezením 256 sloupců na list.
- Pokud je dat přibližně více než se vejde na 1 obrazovku, pak zvážíme využití „data – ověření“.
- Aplikujeme systém na všechny dotazníky – tj. vložíme data.

¹ missing values

3. Třetím krokem je kontrola dat

- třídění (automatický filtr)
- počítání různých kontrolních statistik

4. Čtvrtým krokem je podle potřeby překódování a vytváření nových/počítaných proměnných.

- Zde je třeba umět pracovat s funkcemi v Excelu, popř. v statistickém programu své volby.
- Kromě samozřejmých matematických a statistických je třeba znát i funkce pro práci s textem (v Excelu CONCATENATE, HODNOTA, ČÁST) a funkce logické (KDYŽ, A, NE NEBO, PRAVDA, NEPRAVDA).