

Data poskytl Masarykův onkologický ústav v Brně. Byly provedeny 2 sběry dat od různých pacientů. První sběr proběhl na jaře (březen, duben) – soubor `bd.csv` a druhý na podzim (září, říjen) – soubor `zr.csv`. Každý datový soubor obsahuje 4 sloupce s těmito názvy:

vek	věk pacienta
pohlavi	pohlaví pacienta, M – muž, F – žena
vitD	množství vitamínu D v krvi
Ca	množství vápníku v krvi

Řešte následující úkoly:

- 1) Načtěte data z obou souborů a odstraňte řádky s chybějícími proměnnými. Zjistěte, kolik řádků se odstraní.
- 2) Pro oba soubory testujte normalitu rozložení pro spojité proměnné, tj. pro vek, vitD, Ca. V případě, že je normalita výrazně porušena, zamyslete se nad vhodnou transformací.
- 3) Zodpovězte následující výzkumné otázky:
 - a) Je korelace mezi hladinou vitamínu D a vápníku?
 - b) Je rozdíl mezi hladinou vápníku v jednotlivých obdobích?
 - c) Je rozdíl mezi hladinou vitamínu D resp. Ca mezi ženami a muži?
 - d) Je závislost vitamínu D resp. Ca na věku?
 - e) Je sezónní efekt na hladiny vitamínu D resp. Ca ovlivněn pohlavím resp. věkem?Všechny svoje odpovědi statisticky zdůvodněte.