

Příklady k prezentaci na 11. cvičení

Budete-li chtít prezentovat i příkazy v R-studiu, přineste si prosím příslušný kód na flashce, nebo mi ho s předstihem pošlete emailem. Pokud byste v příkladu narazili na nějakou nejasnost, neváhejte se mi ozvat.

Dataset `Muzi50.RData` byl opět součástí dlouhodobého výzkumu provedeného studentkou 2. lékařské fakulty UK na mužích středního věku. Tentokrát je zaměřen na muže ve věku 50 let. Byly zkoumány následující veličiny:

- ID = identifikace osoby
- vyska = výška postavy
- vaha = hmotnost v kg
- syst1 = měření systolického tlaku v mmHg
- syst2 = opakované měření systolického tlaku v mmHg
- chlst = hladina cholesterolu v mg/100 ml
- Vino = zda pije víno
- cukr = denní spotřeba cukru (počet kostek)
- bmi = body mass index v kg/m^2
- vek = věk
- KOURrisk = kouření jako rizikový faktor (více než 15 cigaret denně)
- Skupina = rozlišení čtyř skupin (faktor): RSI (riziková skupina určená k intervenci/školení), RSK (riziková skupina bez intervence, tedy kontrolní), PS (muži s vážnějším onemocněním, tj. patologická skupina), NS&NSS (normální, podle výzkumníka bez problémů). Do rizikové skupiny byly zařazováni muži, u kterých se objevil aspoň jeden z pěti rizikových faktorů (pozitivní rodinná anamnéza, obezita, kouření, hypertenze, cholesterol vyšší než 260 mg/100 ml).

Tabulku si do R můžete načíst příkazem

```
load("Muzi50.RData")
```

Příklad 1

U mužů ve věku 50 let prozkoumejte hladinu cholesterolu (veličina `chlst`) ve čtyřech skupinách podle proměnné (faktoru) `Skupina`.

- Pomocí vhodného obrázku (boxplot či graf průměrů) znázorněte základní popisné statistiky pro hladinu cholesterolu v závislosti na skupinách.
- Pomocí analýzy rozptylu otestujte, zda se střední hodnota hladiny cholesterolu v jednotlivých skupinách liší. (Formulujte nulovou a alternativní hypotézu, zapište a ověřte předpoklady analýzy rozptylu).
- Pokud vám vyjde, že mezi skupinami je signifikantní rozdíl, pomocí Tukeyho metody zjistěte, které skupiny se od sebe významně liší.