

M4130 cvičení 9

Práce s faktory, datovými tabulkami a datovými soubory

12.11.2019

1. Vytvořte faktor `f` s hodnotami 0, 1, 2, 0, 2, 1, 1, 2, 2, 0, 1.

- Upravte `f` tak, aby byly jednotlivé úrovně faktoru (“levels”) uspořádány.
- Zjistěte, na kterých pozicích jsou hodnoty vyšší než 0 (zkuste, co by ste stalo, kdybychom “levels” neuspořádali).

```
f <- factor(c(0, 1, 2, 0, 2, 1, 1, 2, 2, 0, 1))
f <- ordered(f)
which(f >= 0)
```

2. Vytvořte nový faktor `fact` použitím `f` tak, aby se místo 0 vypisovalo “kategorie”, místo 1 “1. kategorie” a místo 2 “2. kategorie”.

```
fact <- f
levels(fact) <- c("kategorie", "1. kategorie", "2.kategorie")
```

3. Vytvořte datovou tabulku `osoby` se sloupci “jmeno”, “pohlavi”, “bydliste”, “vyska”, která bude obsahovat údaje o 6 osobách:

1. Petr, M, Brno, 178
2. Tereza, F, Praha, 173
3. Karolina, F, Brno, 158
4. Filip, M, Plzen, 190
5. Jiri, M, Ostrava, 167
6. Jana, F, Ostrava, 173

- Zjistěte rozměry tabulky.
- Zjistěte strukturu a datové typy sloupců tabulky.
- Zjistěte názvy sloupců tabulky.
- Vypište `summary` tabulky.

```
jmeno <- c("Petr", "Tereze", "Karolina", "Filip", "Jiri", "Jana")
pohlavi <- c("M", "F", "F", "M", "M", "F")
bydliste <- c("Brno", "Praha", "Brno", "Plzen", "Ostrava", "Ostrava")
vyska <- c(178, 173, 158, 190, 167, 173)
osoby <- data.frame(jmeno, pohlavi, bydliste, vyska)

dim(osoby)
str(osoby)
names(osoby)
summary(osoby)
```

4. Pracujte nadále s tabulkou `osoby`:

- Vypište výšky všech osob (pomocí indexů a/nebo názvu sloupců).
- Vypište jména a bydliště všech osob (pomocí indexů a/nebo názvu sloupců).
- Změňte úrovně sloupce `pohlavi`: M -> male, F -> female.
- Vypište všechny údaje o Karolíně.

```
levels(osoby$pohlavi) <- c("female", "male")
```

5. Tabulce `osoby` přidejte názvy řádků, stejné jako jména osob.

- Opět vypište všechny údaje o Karolíně, využijte pojmenování řádků.
- Vypište jméno, výšky a bydliště Karolíny a Filipa.
- Do tabulky `osoby` přidejte sloupec `skola` s hodnotami SS, VS, VS, VS, SS, SS.
- Prozkoumejte strukturu a datové typy sloupců tabulky.

```
row.names(osoby) <- osoby$jmeno
```

6. Vytvořte tabulku `skola` s jedním sloupcem `skola` s hodnotami SS, VS, VS, VS, SS, SS.

- Spojte tabulky `osoby` a `skola` vedle sebe do nové tabulky `osoby1`.
- Prozkoumejte strukturu a datové typy sloupců tabulky `osoby1`.
- V čem se liší/neliší poslední dva sloupce tabulky `osoby1`?

7. Pracujte s tabulkou `osoby1`:

- Smažte předposlední sloupec `skola`.
- Zbýlý sloupec `skola` přejmenujte na `vzdelani` a sloupec `bydliste` přejmenujte na `mesto`.

8. Pokračujte v práci s tabulkou `osoby1`:

- Vypište tabulku `osoby1` seřazenou podle pohlaví.
- Vypište tabulku `osoby1` seřazenou podle pohlaví opačně.
- Vypište tabulku `osoby1` seřazenou podle pohlaví a dále podle názvu města (abecedně vzestupně).
- Vypište tabulku `osoby1` seřazenou podle pohlaví a dále podle názvu města (abecedně sestupně).

9. Vypište tabulku `osoby1` seřazenou podle názvu města, když na ní je definováno řazení `Brno < Praha < Plzen < Ostrava`.

```
osoby1$mesto <- ordered(osoby1$mesto, levels = c("Brno", "Praha", "Plzen", "Ostrava"))
```

10. Vytvořte tabulku `osoby2` se sloupci `jmeno`, `pohlavi`, `bydliste`, `vyska`, `skola` s údaji:

1. Adam, male, Brno, 173, ZS
2. Eva, female, Praha, 169, SS

11. Spojte tabulky `osoby1` a `osoby2` pod sebe do tabulky `osoby3`. (Nápověda: funkce `merge`).

- Je tabulka `osoby3` správně?
- Prozkoumejte strukturu a datové typy sloupců tabulky `osoby3`.

- Opravte tabulku osoby2 (příp. i osoby1) tak, aby spojení do osoby3 proběhlo správně.

12. Pracujte s tabulkou osoby3:

- Spočítejte průměrnou výšku osob.
- Spočítejte průměrnou výšku mužů.
- Spočítejte průměrnou výšku žen.
- Vypište frekvenční tabulku měst. (Nápověda: funkce `table`).

13. Práce s reálnými daty

Načtěte data ze souboru `rajcata.dat` udávající hmotnosti rajčat. Poté spočítejte:

- Jejich průměrnou hmotnost a medián hmotností rajčat.
- Kolik z nich je těžších než 65 g? Zjistěte jejich pořadí.
- Jaký je procentuální podíl rajčat, jejichž hmotnost je menší než průměrná hmotnost?
- Potřebujeme odvážit nejvýše 1 kg rajčat. Rajčata na váhu přidáváme postupně po jednom dle tabulky. Kolik nejvýše jich položíme, aby hmotnost nepřesáhla 1 kg?
- Potřebujeme odvážit nejvýše 1 kg rajčat. Rajčata můžeme přidávat v libovolném pořadí. Kolik nejvýše jich položíme, aby hmotnost nepřesáhla 1 kg? (Nápověda: funkce `which` a `cumsum`).