

Procvičování 9

1. Importujte do **R** tabulku jmen s daty narození českých občanů *jmena.txt*. Nazvěte ji *jm.data*. Tabulka obsahuje v řádcích všechna jednoslovná křestní jména českých obyvatel, ve sloupcích pak jejich frekvenci v jednotlivých ročnících od r. 1896 do 2013. Tabulka je docela velká, nepokoušejte se ji proto celou zobrazit v **R**, nechejte si raději vypsát jen prvních pár řádků a sloupců. Najdete ji ve studijních materiálech, nebo ji můžete načíst přímo z <http://www.sci.muni.cz/~syrovat/jmena.txt>, to ale bude chvilku trvat.
2. Do vektoru *jm.freq* vypočítejte frekvenci výskytu jednotlivých jmen v české populaci.
3. Do vektoru *jm* si vytáhněte křestní jména z dataframu *jm.data*.
4. Seřaďte křestní jména podle abecedy a nechejte si zobrazit prvních a posledních 20.
5. Zjistěte nejdelší křestní jméno. Postup: vytvořte si vektor délek jmen, seřaďte jména podle jejich délek od nejvyšších po nejnižší a vyberte první. Délkou rozumíme počet znaků (characters). Funkci, která zjistí délku textového řetězce musíte vyhledat.
6. Nechejte si vypsát 100 nejkratších jmen.
7. Nechejte si vypsát 50 nejběžnějších jmen.
8. Zjistěte, zda se Vaše jméno nachází mezi 50 nejběžnějšími.
9. Zjistěte kolikáté nejběžnější je Vaše jméno.
10. Naimportujte do **R** dataframu *spe* a *env* a vektor korelací abundancí jednotlivých druhů s Froudeho číslem *korelace*. Vše najdete v *dat09.xls*.
11. Nechejte si vypsát zaokrouhlené hodnoty Froudeho čísla (na 2 desetinná místa) seřazené od nejmenší po největší, a pak od největší po nejmenší.
12. Zjistěte 5 nejnižších a 5 nejvyšších hodnot Froudeho čísla.
13. Nechejte si vypsát jména 5 lokalit s nejvyšším Froudeho číslem.
14. Do dataframu *env* přidejte proměnné *ind* a *spec* obsahující celkové počty jedinců a druhů pakomárů na lokalitách.
15. Nechejte si vypsát počty druhů seřazené podle Froudeho čísla od nejnižšího po nejvyšší.
16. Nechejte si vypsát dataframe *env* seřazený podle Froudeho čísla od nejvyššího po nejnižší.
17. Zjistěte, kolik druhů a kolik jedinců pakomárů bylo na 5 lokalitách s nejvyšším Froudeho číslem (vyberte jen příslušné elementy z proměnných *ind* a *spec* dataframu *env*).
18. Zjistěte, jaké Froudeho číslo bylo naměřeno na 3 lokalitách s nejvyšším počtem jedinců pakomárů.
19. Nechejte si vypsát dataframe *env* s lokalitami seřazenými podle Froudeho čísla.
20. Nechejte si vypsát dataframe *env* s lokalitami seřazenými podle typu habitatu (*gr*) a uvnitř typů podle hloubky vody.
21. Nechejte si vypsát jména pěti nejfrekventovanějších druhů (tedy těch s výskytem na nejvyšším počtu lokalit).
22. Nechejte si vypsát jména druhů seřazená podle jejich korelací s Froudeho číslem.

23. Zjistěte, kterých 10 druhů nejlépe koreluje s Froudeho číslem. Pozor na to, že korelace může být kladná nebo záporná, zajímají nás druhy, které nejlépe na hydraulické podmínky reagují, tedy ty s nejnižší a nejvyšší korelací (při řazení tedy použijeme absolutní hodnoty korelací).
24. Vytvořte kopii dataframu *spe*, v níž bude 7 druhů které nejlépe reagovaly na hydraulické podmínky.
25. Seřadte sloupce tohoto dataframu podle typu odezvy (kladná nebo záporná korelace), a uvnitř každého typu podle síly korelace (první budou ty s nejvyšší odezvou). Řádky pak seřadte podle Froudeho čísla.
26. Exportujte do excelu výsledný dataframe. Použijte funkci `write.table()`.