

## Ćwiczenia 1 & 2 (Wizualizacja i przetwarzanie danych)

1. Wygeneruj liczby całkowite od 10 do 20 i zachowaj je w wektorze  $x$ .
2. Wygeneruj 4 powtórzenia sekwencji liczb (3, 5, 7).
3. Wygeneruj sekwencję zawierającą osiem czwórek, następnie siedem szóstek i na koniec dziewięć trójek. Zapisz ją w macierzy  $M$  o sześciu wierszach i czterech kolumnach.
4. Wygeneruj wszystkie liczby nieparzyste z przedziału  $[1, 100]$ .
5. Utwórz wektor zawierający jedną jedynkę, dwie dwójki, ..., dziewięć dziewiątek.
6. Skonstruuj wektor  $x$  używając poniższego kodu:

```
x <- c(NA, 3, 14, NA, 33, 17, NA, 41)
```

- Zlicz liczbę braków.
- Wyznacz średnią arytmetyczną nie biorąc braków pod uwagę.
- Usuń braki w danych.
- Zastąp braki liczbą 11.

7. Zbiór danych `cfb` z pakietu `UsingR` zawiera dane finansowe 1000 klientów. Utwórz ramkę danych zawierającą jedynie klientów o dodatnich dochodach (`INCOME`) i ujemnej całkowitej wartości netto (`NETWORTH`). Ilu jest takich klientów?
8. Sprawdź, które zmienne w zbiorze danych `Cars93` (sprzedaż samochodów w USA w roku 1993) z pakietu `MASS` są czynnikami. Wyznacz w postaci tabeli liczbę samochodów dla miejsca pochodzenia (zmienna `Origin`) oraz rodzaju samochodu (zmienna `Type`).
9. **(S)** Dla zbioru danych `Cars93` wyświetl informacje funkcją `summary` dla każdego poziomu zmiennej `Type`.
10. Dla zbioru danych `Insurance` z pakietu `MASS`, zawierającego informacje na temat roszczeń ubezpieczeniowych samochodów w 1973 roku, skonstruuj tabelę zawierającą liczbę roszczeń (zmienna `Claims`) z podziałem na wiek (zmienna `Age`) oraz typ samochodu (zmienna `Group`).
11. **(S)** Ze zbioru `mtcars` utwórz ramkę `mtcars6`, która zawiera informacje jedynie o samochodach z 6 cylindrami.
12. Ze zbioru `Cars93` z pakietu `MASS`, utwórz ramkę `samochody`, która zawiera informacje jedynie o małych oraz sportowych samochodach.
13. Zastosuj funkcję `sapply` do funkcji `is.factor` na każdej kolumnie zbioru `tinting` (wpływ przyciemnienia szyb na jakość widzenia) z pakietu `DAAG`. Dla każdej kolumny będącej czynnikiem określ poziom. Które czynniki są uporządkowane (`is.ordered`)?
14. **(S)** Wykonaj poniższe polecenie tworzące listę `list1`:

```
list1 <- list(observationA = c(1:5, 7:3), observationB = matrix(1:6, nrow = 2))
```

Korzystając z grupy funkcji `apply()` wyznacz liczbę unikatowych wartości w każdym elemencie listy.

15. Wprowadź dane do R poleceniem

```
x <- c(1, 8, 2, 6, 3, 8, 5, 5, 5, 5)
```

Oblicz:

- sumę wszystkich elementów,

- logarytm dziesiętny wszystkich elementów,
- różnicę pomiędzy największym i najmniejszym elementem wektora  $x$ .

16. **(S)** Skonstruuj macierz diagonalną o wymiarze 4 o elementach 4, 1, 2, 3 na przekątnej.

17. Utwórz następującą macierz

$$D = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 2 & 1 \\ 2 & 3 & 0 \end{pmatrix}$$

oraz oblicz jej transpozycję i odwrotność. Pomnóż macierz przez jej odwrotność.

18. Walec o wysokości  $h$  oraz promieniu podstawy  $r$  ma objętość  $V = \pi r^2 h$  oraz pole powierzchni  $P = 2\pi r(r + h)$ . Dla długości promienia 1:5 oraz wysokości 4:8 oblicz odpowiednie objętości i pola powierzchni. Skonstruuj ramkę danych o kolumnach:  $r$ ,  $h$ ,  $V$ ,  $P$ .

19. Znajdź 8 największych lądów (kontynenty, wyspy) świata (zbiór danych `islands`).

20. W pakiecie `schoolmath` znajduje się zbiór danych `primlist`, który zawiera liczby pierwsze pomiędzy 1 a 9 999 999. Znajdź największą liczbę pierwszą mniejszą od 1000. Ile jest liczb pierwszych większych od 100 a mniejszych od 500?

21. Utwórz wektor liczb naturalnych od 1 do 1000, a następnie zamień liczby parzyste na ich odwrotności.

22. **(S)** Jak znaleźć nazwę (indeks) poszczególnej wartości wektora? Na przykład, gdzie jest trasa wspinaczkowa o długości 2100 stóp (zbiór danych `hills` z pakietu `MASS`, który zawiera najlepsze wyniki w wyścigach wspinaczkowych w Szkocji (1984 rok))?

23. Zmienna `ftv` w zbiorze danych `birthwt`, z pakietu `MASS`, zawiera liczbę wizyt matek u lekarza w pierwszym trymestrze ciąży. Przekształć ją do czynnika o trzech poziomach 0, 1 oraz 2 lub więcej (użyj funkcji `factor` oraz `levels`).

24. W pewnych sytuacjach przydatna może się okazać tzw. kategoryzacja zmiennych, czyli inny podział na kategorie niżby wynikał z danych. Wygeneruj 100 obserwacji, które są odpowiedziami na pytania ankiety, każda odpowiedź może przyjąć jedną z wartości: 'a', 'b', 'c', 'd', 'e'. Dokonaj kategoryzacji w taki sposób, aby kategoria 1 obejmowała odpowiedzi 'a' i 'b', 2 odpowiedzi 'c' i 'd' oraz 3 odpowiedzi 'e'.

**Wskazówka:** Wykorzystaj funkcję `recode` z pakietu `car`.

25. **(S)** Zbiór danych `Pima.tr2` z pakietu `MASS`, zawiera informacje na temat indiańskich kobiet (powyżej 21 lat) z okolic Phoenix chorych na cukrzycę. Napisz funkcję, która wyświetli liczbę brakujących danych w każdej ze zmiennych.

26. **(S)** Utwórz zmienną `tekst` o zawartości: „The current year is 2019”.

- Sprawdź czy w tym wyrażeniu znajdują się cyfry.
- Znajdź miejsca, w których znajdują się cyfry.
- Znajdź liczbę czterocyfrową znajdującą się na końcu wyrażenia i wyświetl ją.

27. **(S)** Utwórz wektor adresów `adresy` o zawartości: `www.dogman.com`, `http://rotterdam.com`, `https://facebook.com`, `httpx://sims.com`, `fungame.http`. Wyszukaj i wyświetl jedynie adresy zaczynające się od `http` lub `https`.

28. Wczytaj dane z pliku `Table1.txt`.

- Zmień nazwy kolumn na następujące: `Name`, `Age`, `Height`, `Weight` oraz `Sex`.
- Zmień nazwy wierszy na takie jak wartości zmiennej `Name`. Następnie usuń zmienną `Name`.

29. Wczytaj dane z pliku `Table2.txt`.