

SPRAWOZDANIE

Zajęcia: Analiza Procesów Uczenia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium 1

Data 3.03.2023

Temat: Podstawy języka R

Wariant 1

Rafał Klinowski
Informatyka II stopień,
Stacjonarne,
1 semestr,
Gr. a

1. Polecenie: Wariant 1

- (a) Do zmiennej `a` podstaw wartość wyrażenia $2 * \exp(5)$. Do zmiennej `b` podstaw podwojoną wartość zmiennej `a`. Wywołaj funkcję sprawdzającą, która z wartości zmiennych jest większa.
- (b) Uruchom i poczytaj dokumentację dla funkcji `sum()`.
- (c) Stwórz wektor `a` zawierający liczby od 15 do 25. Policz sumę liczb zawartych w wektorze.
- (d) Wyświetl wszystkie funkcje zawierające frazę `sum` w swojej nazwie.
- (e) Ustaw dowolny katalog roboczy. Następnie stwórz zmienną `a` zawierającą łańcuch znaków "smartfony Samsung". Zapisz zmienną `a` z obszaru roboczego do pliku w katalogu roboczym. Następnie usuń zmienną `a`. Sprawdź wartość zmiennej `a` (powinno jej brakować). Na końcu wczytaj plik ze zmienną `a` i sprawdź jej wartość.
- (f) Zainstaluj i załaduj pakiet `gridExtra`, który umożliwia m.in. ładną wizualizację danych tabelarycznych. Następnie przy pomocy dokumentacji pakietu znajdź funkcję do wizualizacji danych tabelarycznych. Użyj jej na pierwszych 10 wierszach zbioru danych `mtcars`.
- (g) Stwórz wektor zawierający ciąg liczb 100, 96, 92, ... 20.
- (h) Stwórz wektora `a` z liczbami od 9 do 5 oraz wektor `b` z liczbami od 11 do 16. Utwórz nowy wektor `d` będący połączeniem wektora `b` i `a` (w takiej kolejności). Wyświetl go.
- (i) Stwórz wektor `nazwa` zawierający nazwy 10 smartfonów Samsung z systemem Android 8 i o 8-miordziennej mocy procesora. Potem stwórz wektory `wyświetlacz`, `pamięć_RAM` i `pamięć_wbudowana`, `aparat_foto`, `cena`, `liczba_opinii` zawierające kolejno dane 10 smartfonów. Następnie stwórz ramkę danych `smartfony` złożoną z wektorów `nazwa`, `wyświetlacz`, `pamięć_RAM`, `pamięć_wbudowana`, `aparat_foto`, `cena` oraz `liczba_opinii`. Wylicz średnią cenę smartfonów.
- (j) Do stworzonej w poprzednim zadaniu ramki danych smartfonów dodaj wpis zawierający dane nowego smartfonu. Wylicz średnią ceny ponownie.
- (k) Korzystając z ramki danych `smartfony` dodaj nową kolumnę określającą ocenę klientów. Wpisz do kolumny odpowiednio oceny w skali od 0 do 5 krok 0.5. Dodana kolumna powinna się automatycznie przekonwertować do cech jakościowych (tzw. `factors`). Wylicz średnią cenę każdej oceny.
- (l) Do ramki danych `smartfony` dodaj kolejne 4 smartfony. Narysuj na wykresie słupkowym liczebność reprezentantów każdej z ocen

klientów.

- (m) Wykorzystując ramkę danych smartfony pokaż procentowy udział każdej oceny przy pomocy wykresu kołowego oraz wachlarzowego.
- (n) Do ramki danych smartfony dodaj nową kolumnę status_opinii z wartościami: "nie ma", "mniej 50 opinii", "50-100 opinii", "więcej 100 opinii" w zależności od liczby opinii. Zamień dodaną kolumnę na cechy jakościowe. Następnie przy pomocy wykresu kołowego wyrysuj procentowy udział smartfonów o konkretnym statusie opinii.
- (o) Wykorzystując ramkę danych smartfony stwórz zdanie o każdym z smartfonów postaci: nazwa + " ma ocenę klientów " + ocena_klientów + " bo ma liczbę opinii" + liczba_opinii. Plus oznacza konkatencję łańcuchów i wartości.
- (p) Zachować ramkę danych w pliku .csv. Załadować ramkę danych z pliku .csv
Dane (15 smartfonów) pobrać ze strony <http://www.euro.com.pl>

2. Wprowadzane dane:

Dane wykorzystane w programie dotyczące parametrów smartfonów zostały pobrane ze strony Euro RTV AGD: https://www.euro.com.pl/telefony-komorkowe,system-operacyjny_2!android_1_1.bhtml

3. Wykorzystane komendy:

Poniżej można znaleźć wszystkie wykorzystane komendy:

```
# Autor: Rafal Klinowski, wariant 1.
```

```
# a
a <- 2*exp(5)
b <- 2*a
a>b
```

```
# b
?sum()
```

```
# c
a <- c(15,16,17,18,19,20,21,22,23,24,25)
suma <- sum(a)
suma
```

```
# d
# ?
```

```

# e
a <- "smartfony Samsung"
writeLines(a,"zmienna_a")
a <- readLines("zmienna_a")
a

# f
install.packages("gridExtra")
library(gridExtra)
library(grid)
grid.table(mtcars)

# g
number <- 96
x <- c(100)
for (i in 1:20) {
  x <- c(x, number)
  number <- number - 4
}

# h
a <- c(9,8,7,6,5)
b <- c(11,12,13,14,15,16)
d <- c(b,a)
d

# i
nazwy <- c("Samsung Galaxy S23 Ultra", "Xiaomi 13", "Motorola moto G22", "Samsung
Galaxy M33 5G",
           "Motorola edge 30 neo", "OPPO Find N2 Flip", "Motorola moto G82 5G",
           "Infinix SMART 6 HD", "Xiaomi Redmi Note 11s", "Samsung Galaxy S20 FE
5G")

pamiec_ram <- c(8, 8, 4, 6, 8, 8, 6, 2, 6, 6)

pamiec_wbudowana <- c(256, 256, 64, 128, 128, 256, 128, 32, 64, 128)

aparat_foto <- c(200, 50, 50, 50, 64, 50, 50, 8, 108, 12)

cena <- c(6799, 4799, 649, 1299, 1699, 4999, 1299, 399, 999, 1999)

liczba_opinii <- c(3, 0, 41, 6, 16, 1, 44, 0, 28, 313)

smartfony <- data.frame(nazwy, pamiec_ram, pamiec_wbudowana, aparat_foto, cena,
liczba_opinii)
smartfony

srednia_cena <- mean(smartfony$cena)
srednia_cena

# j
nowy_smartfon <- data.frame(nazwy="realme 9
Pro+",pamiec_ram=6,pamiec_wbudowana=128,
                           aparat_foto=50,cena=1444,liczba_opinii=7)
smartfony <- rbind(smartfony, nowy_smartfon)
smartfony

# k

```

```

oceny_klientow <- c(5,0,5,5,5,1,5,0,5,5,5)
smartfony$oceny_klientow <- oceny_klientow
smartfony

# l
nowy_smartfon <- data.frame(nazwy="Motorola edge 20 pro
5G",pamiec_ram=12,pamiec_wbudowana=256,

aparat_foto=108,cena=2599,liczba_opinii=26,oceny_klientow=5)
smartfony <- rbind(smartfony, nowy_smartfon)
nowy_smartfon <- data.frame(nazwy="Samsung Galaxy M23
5G",pamiec_ram=4,pamiec_wbudowana=128,

aparat_foto=50,cena=999,liczba_opinii=52,oceny_klientow=5)
smartfony <- rbind(smartfony, nowy_smartfon)
nowy_smartfon <- data.frame(nazwy="OnePlus 11
5G",pamiec_ram=16,pamiec_wbudowana=256,

aparat_foto=50,cena=4599,liczba_opinii=7,oceny_klientow=5)
smartfony <- rbind(smartfony, nowy_smartfon)
nowy_smartfon <- data.frame(nazwy="Samsung Galaxy
S22+",pamiec_ram=8,pamiec_wbudowana=128,

aparat_foto=50,cena=4499,liczba_opinii=20,oceny_klientow=5)
smartfony <- rbind(smartfony, nowy_smartfon)

smartfony

ilosci_ocen <- table(smartfony$oceny_klientow)
barplot(ilosci_ocen, xlab="Ocena", ylab="Ilosc")

# m
pie(ilosci_ocen)

library(fanplot)
plot.new()
fan(ilosci_ocen)

# n
status_opinii <- c("mniej niz 50 opinii")

for (i in 2:15) {
  if (smartfony$liczba_opinii[i] == 0) {
    status_opinii <- c(status_opinii, "nie ma")
  }
  else if (smartfony$liczba_opinii[i] < 50) {
    status_opinii <- c(status_opinii, "mniej niz 50 opinii")
  }
  else if (smartfony$liczba_opinii[i] < 100) {
    status_opinii <- c(status_opinii, "50-100 opinii")
  }
  else {
    status_opinii <- c(status_opinii, "wiecej niz 100 opinii")
  }
}

smartfony$status_opinii <- status_opinii

smartfony

```

```

ilosci_ocen <- table(smartfony$status_opinii)
pie(ilosci_ocen)

# o
for (i in 1:15) {
  zdanie <- paste(smartfony$nazwy[i], " ma ocene klientow ",
smartfony$oceny_klientow[i],
                  " bo ma liczbe opinii", smartfony$liczba_opinii[i])
  print(zdanie)
}

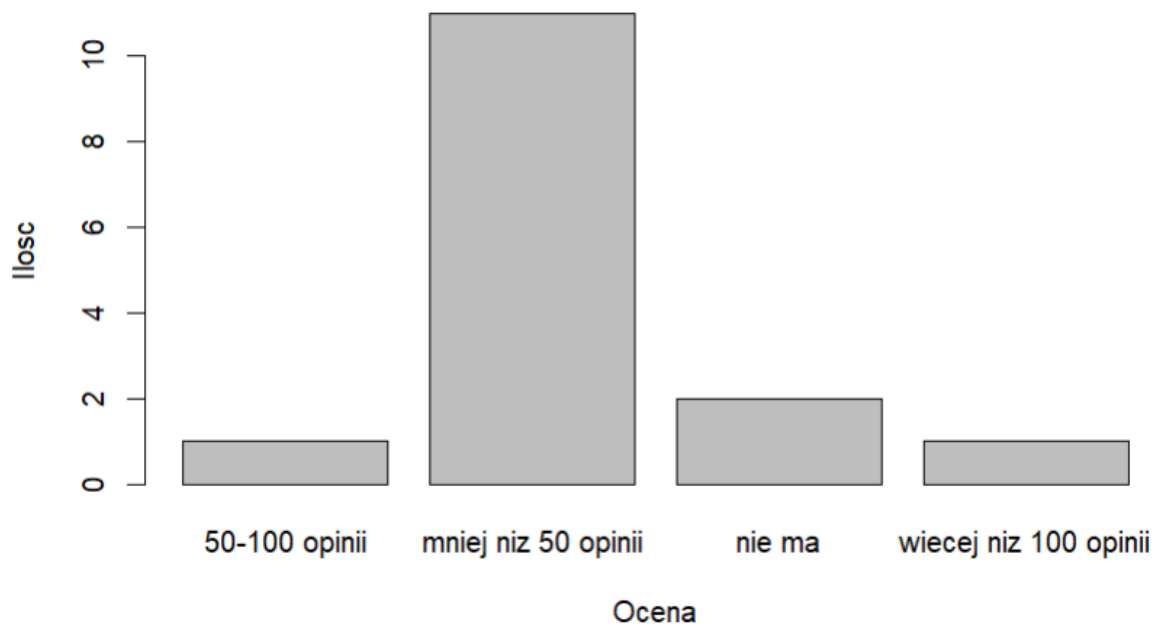
# p
write.csv(smartfony, "smartfony.csv", row.names=FALSE)

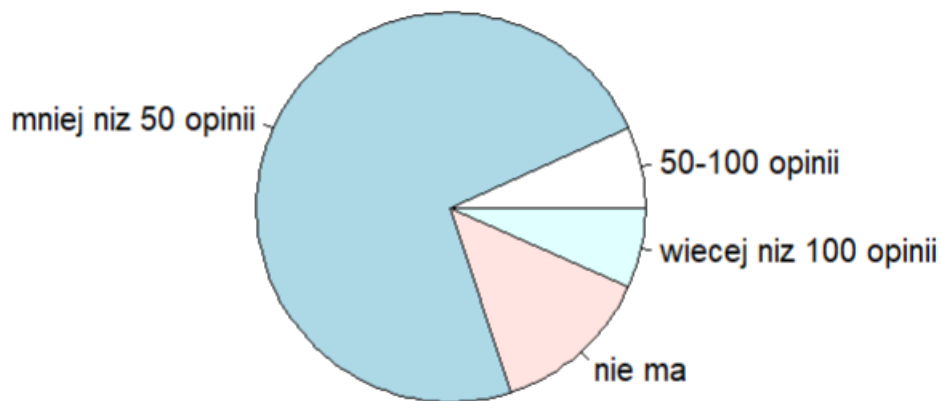
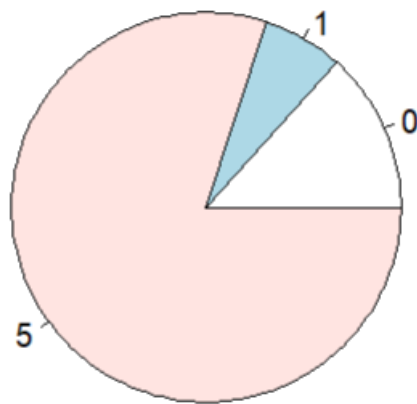
```

4. Wynik działania:

Wyniki poleceń w konsoli można znaleźć w pliku „console.txt”, link do repozytorium poniżej.

Zrzuty ekranu przedstawiające niektóre z wykresów uzyskanych w środowisku:





Zgodnie z poleceniem, link do repozytorium GitHub zawierający niezbędne pliki znajduje się tutaj: https://github.com/Stukeley/APU_Lab1

5. Wnioski:

Na podstawie otrzymanego wyniku można stwierdzić, że praca z ramkami danych (dataframe) w języku R jest bardzo prosta i dobrze zintegrowana. Dodawanie zarówno nowych rzędów jak i kolumn do ramki danych jest trywialne. R zawiera cały szereg funkcji do pracy z danymi, m.in. funkcje odpowiedzialne za rysowanie wykresów, obliczenie średniej, zliczania wartości czy dynamicznie dodawania danych (np. w pętli).