

SPRAWOZDANIE

Zajęcia: Analiza Procesów Uczenia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

Laboratorium 1

03.03.2023

Temat: Podstawy języka R

Wariant 5

Jarosław Waliczek
Informatyka II stopień,
stacjonarne,
1 semestr,
Gr.1

1. Polecenie: Zadanie dotyczy tworzenia danych, które będą wykorzystywane na kolejnych zajęciach w celu podejmowania decyzji przy kupowaniu urządzeń RTV AGD. Sprawozdanie sporządzić zgodnie ze wzorem i odesłać przez system e-uczelnia. Pliki w postaci: plik .R, wyniki z konsoli (dowolny plik tekstowy), plik .csv zachować w zdalnym repozytorium (np Github) link na który wskazać w sprawozdaniu

2. Wprowadzane dane:

5. (a) Do zmiennej `a` podstaw wartość wyrażenia $14 \cos(\pi)$. Do zmiennej `b` podstaw podwójną wartość zmiennej `a`. Wywołaj funkcję, sprawdzającą, która z wartości zmiennych jest mniejsza.
- (b) Uruchom i poczytaj dokumentację dla funkcji `log10()`.
- (c) Stwórz wektor `a` zawierający liczby od 900 do 1150. Policz średnią kwadratów liczb zawartych w wektorze.

- (d) Wyświetl wszystkie funkcje zawierające frazę log w swojej nazwie.
- (e) Ustaw dowolny katalog roboczy. Następnie stwórz zmienną a zawierającą łańcuch znaków "łódówka turystyczna". 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 ładną wizualizację danych tabelarycznych. Następnie przy pomocy dokumentacji pakietu znajdź funkcję do wizualizacji danych tabelarycznych. Użyj jej na pierwszych 10 wierszach zbioru danych sleep.
- (g) Stwórz wektor zawierający ciąg liczb 1500, 1495, 1490, . . . 1400.
- (h) Stwórz wektora a z liczbami od 25 do 5 oraz wektor b z liczbami od 50 do 70. Utwórz nowy wektor d będący połączeniem wektora b i a (w takiej kolejności). Wyświetl go.
- (i) Stwórz wektor nazwa zawierający nazwy 15 łódówek turystycznych typu elektrycznego. Potem stwórz wektory pojemność, waga, cena, liczba_opinii zawierające kolejno dane 15 łódówek. Następnie stwórz ramkę danych łódówki złożoną z wektorów pojemność, waga, cena, liczba_opinii. Wylicz średnią cenę łódówek.
- (j) Do stworzonej w poprzednim zadaniu ramki danych łódówek dodaj wpis zawierający dane nowej łódówki. Wylicz średnią cenę ponownie.
- (k) Korzystając z ramki danych łódówki dodaj nową kolumnę określając 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 aparaty dodaj kolejne 4 łódówki. Narysuj na wykresie słupkowym liczebność reprezentantów każdej z ocen klientów.

- (m) Wykorzystując ramkę danych lodówki pokaż procentowy udział każdej oceny przy pomocy wykresu kołowego oraz wachlarzowego.
- (n) Do ramki danych lodówki 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 lodówek o konkretnym statusie opinii.
- (o) Wykorzystując ramkę danych lodówki stwórz zdanie o każdej z lodówek postaci: nazwa + " ma ocenę klientów " + ocena_klientów + " bo ma liczbę opinii" + liczba_opinii. Plus oznacza konkatenację łańcuchów i wartości.
- (p) Zachowaj ramkę danych w pliku .csv. Załaduj ramkę danych z pliku .csv
Dane (15 lodówek turystycznych) pobrać ze strony <http://www.euro.com.pl>

3. Wykorzystane komendy:

```

a <- 14
a <- 14 * cos(3.14)
b <- a * 2
min(a, b)
a <- 900:1150
help(log)
getwd()
setwd("C:/Users/jaro9/OneDrive/Desktop/apu")
getwd()
a <- "lodówka turystyczna"
save(a, file = "workspace.RData")
rm(a)
a
load("workspace.RData")
a
install.packages("gridExtra")
x1 <- seq(1400, 1500, by = 5)
a <- 25:5
b <- 50:70
d <- b + a
d
nazwa <- c("HB", "Curver", "Ravanson", "QVANT", "Yeticool", "G21", "Camry", "Dometic", "Sencor",
"Aisberg", "Mobicool", "Polarbox", "Fresh", "Newestcool", "Fantastic")
waga <- c(4.3, 4.4, 4.5, 4.6, 4.7, 4.7, 4.2, 4.3, 4.4, 4.9, 4.9, 4.7, 4.7, 4.8, 5)
cena <- c(300, 800, 1500, 2000, 3000, 450, 600, 900, 1200, 1300,

```

```

100, 150, 200, 300, 500)
liczba_opinii <- c(0, 40, 80, 120, 130, 10, 70, 80, 40, 20, 90, 150, 25, 55, 5)
df <- data.frame(nazwa, waga, cena, liczba_opinii)
mean(df[[3]])
df[nrow(df) + 1,] <- list("Polarbox2.0", 5.5, 550, 15)
mean(df[[3]])
ocena_klientow1 <- seq(0, 5, by=0.5)
ocena_klientow2 <- seq(0, 2, by=0.5)
ocena_klientow <- append(ocena_klientow1, ocena_klientow2)
df["ocena_klientow"] = ocena_klientow
lodowki <- df
lodowki[nrow(lodowki) + 1,] <- list("CV60", 5, 2000, 12, 2.5)
lodowki[nrow(lodowki) + 1,] <- list("CV70", 4.5, 2400, 11, 3)
lodowki[nrow(lodowki) + 1,] <- list("CV80", 2, 1000, 5, 3.5)
lodowki[nrow(lodowki) + 1,] <- list("CV90", 1, 800, 2, 4)
count <- table(lodowki$ocena_klientow)
barplot(count,
  main = "liczebność reprezentantów każdej z ocen klientów",
  ylim = c(0, 5),
  xlab = "Ocena klienta",
  ylab = "Ilość")
percentage <- table(lodowki$ocena_klientow) /
  length(lodowki$ocena_klientow)
pie(percentage)

percentage2 <- table(lodowki$ocena_klientow) /
  length(lodowki$ocena_klientow)
fan.plot(percentage2, labels = names(percentage2))

lodowki$status_opinii <- with(lodowki, ifelse(liczba_opinii > 100, 'więcej niż 100 opinii',
  ifelse(liczba_opinii > 50, '50-100 opinii',
  ifelse(liczba_opinii == 0, 'nie ma', 'mniej niż 50 opinii'))))

for(i in 1:nrow(lodowki)) {
  cat(lodowki[i, 1], "ma ocenę klientów", lodowki[i, 5], "bo ma liczbę opinii", lodowki[i, 4], '\n')
}

write.csv(lodowki, "C:/Users/jaro9/OneDrive/Desktop/apu/lodowki.csv", row.names=FALSE)
read.csv("C:/Users/jaro9/OneDrive/Desktop/apu/lodowki.csv")

```

4. Wynik działania:

Wyniki z konsoli:

```

a <- 14
> a <- 14 * cos(3.14)
> b <- a * 2
> min(a, b)
[1] -27.99996
> a <- 900:1150
> help(log)
> getwd()
[1] "C:/Users/jaro9/OneDrive/Dokumenty"
> setwd("C:/Users/jaro9/OneDrive/Desktop/apu")
> getwd()
[1] "C:/Users/jaro9/OneDrive/Desktop/apu"
> a <- "lódówka turystyczna"
> save(a, file = "workspace.RData")
> rm(a)
> a

```

```
Error: object 'a' not found
> load("workspace.RData")
> a
[1] "łódówka turystyczna"
> install.packages("gridExtra")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and install
the appropriate version of Rtools before proceeding:
```

```
https://cran.rstudio.com/bin/windows/Rtools/
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/gridExtra_2.3.zip'
Content type 'application/zip' length 1109604 bytes (1.1 MB)
downloaded 1.1 MB
```

pakiet 'gridExtra' został pomyślnie rozpakowany oraz sumy MD5 zostały sprawdzone

Pobrane pakiety binarne są w

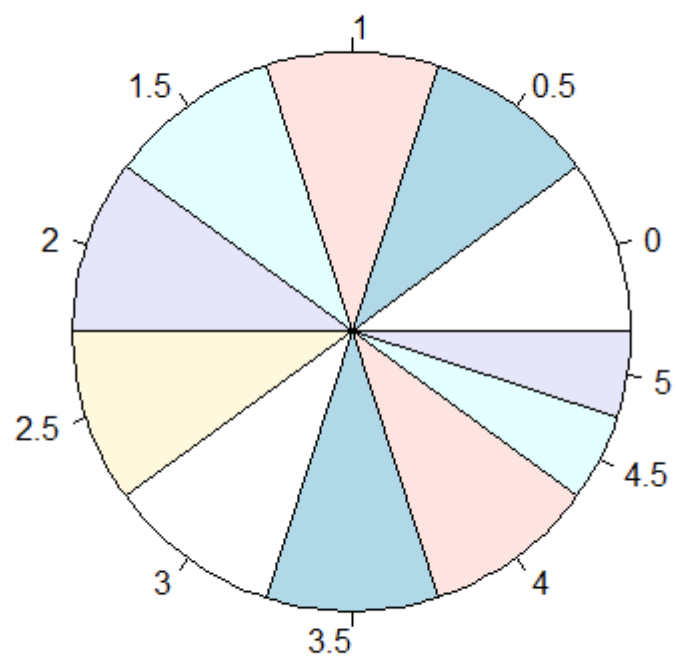
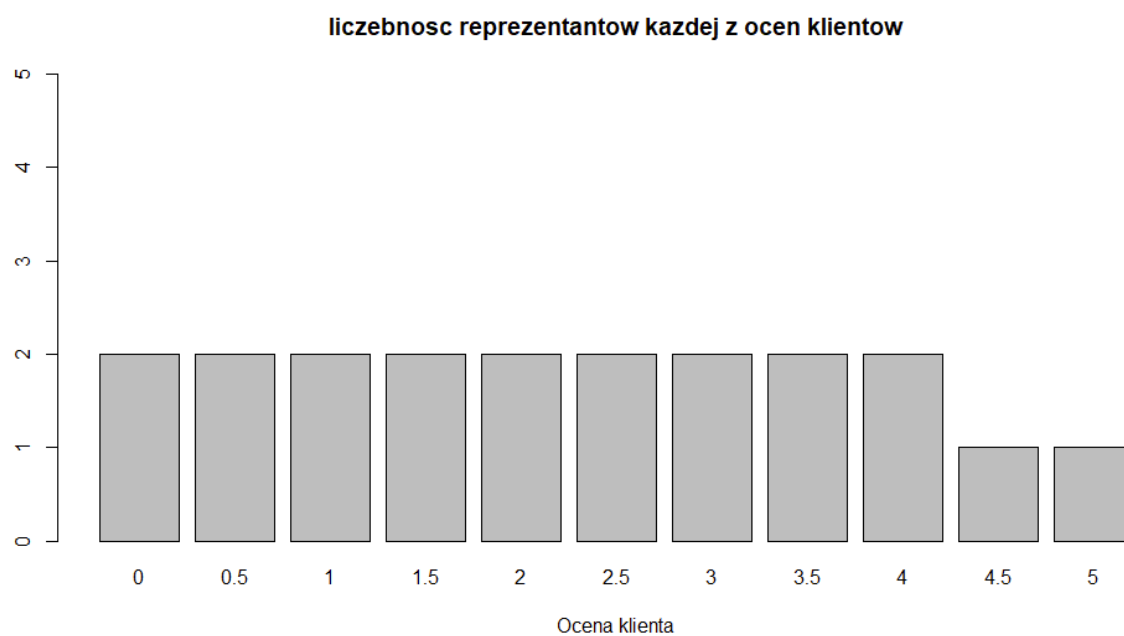
```
C:\Users\jaro9\AppData\Local\Temp\RtmpCeENKs\downloaded_packages
> x1<-seq(1400,1500,by=5)
> a <- 25:5
> b <- 50:70
> d <- b + a
> d
[1] 75 75 75 75 75 75 75 75 75 75 75 75 75 75 75 75 75 75 75 75
> nazwa <- c("HB", "Curver", "Ravanson", "QVANT", "Yeticool", "G21", "Camry", "Dometic", "Sencor",
+           "Aisberg", "Mobicoool", "Polarbox", "Fresh", "Newestcool", "Fantastic")
> waga <- c(4.3, 4.4, 4.5, 4.6, 4.7, 4.7, 4.2, 4.3, 4.4, 4.9, 4.9, 4.7, 4.7, 4.8, 5)
> cena <- c(300, 800, 1500, 2000, 3000, 450, 600, 900, 1200, 1300,
+          100, 150, 200, 300, 500)
> liczba_opinii <- c(0, 40, 80, 120, 130, 10, 70, 80, 40, 20, 90, 150, 25, 55, 5)
> df <- data.frame(nazwa, waga, cena, liczba_opinii)
> mean(df[[3]])
[1] 886.6667
> df[nrow(df) + 1,] <- list("Polarbox2.0", 5.5, 550, 15)
> mean(df[[3]])
[1] 865.625
> ocena_klientow1 <- seq(0, 5, by=0.5)
> ocena_klientow2 <- seq(0, 2, by=0.5)
> ocena_klientow <- append(ocena_klientow1, ocena_klientow2)
> df["ocena_klientow"] = ocena_klientow
> lodowki <- df
> lodowki[nrow(lodowki) + 1,] <- list("CV60", 5, 2000, 12, 2.5)
> lodowki[nrow(lodowki) + 1,] <- list("CV70", 4.5, 2400, 11, 3)
> lodowki[nrow(lodowki) + 1,] <- list("CV80", 2, 1000, 5, 3.5)
> lodowki[nrow(lodowki) + 1,] <- list("CV90", 1, 800, 2, 4)
> count <- table(lodowki$ocena_klientow)
> barplot(count,
+         main = "liczebność reprezentantów każdej z ocen klientów",
+         ylim = c(0, 5),
+         xlab = "Ocena klienta",
+         ylab = "Ilość")
> percentage <- table(lodowki$ocena_klientow) /
+   length(lodowki$ocena_klientow)
> pie(ppercentage)
> percentage2 <- table(lodowki$ocena_klientow) /
+   length(lodowki$ocena_klientow)
> fan.plot(ppercentage2, labels = names(ppercentage2))
Error in fan.plot(ppercentage2, labels = names(ppercentage2)) :
could not find function "fan.plot"
> lodowki$status_opinii <- with(lodowki, ifelse(liczba_opinii > 100, 'więcej niż 100 opinii',
+                                             ifelse(liczba_opinii > 50, '50-100 opinii',
```

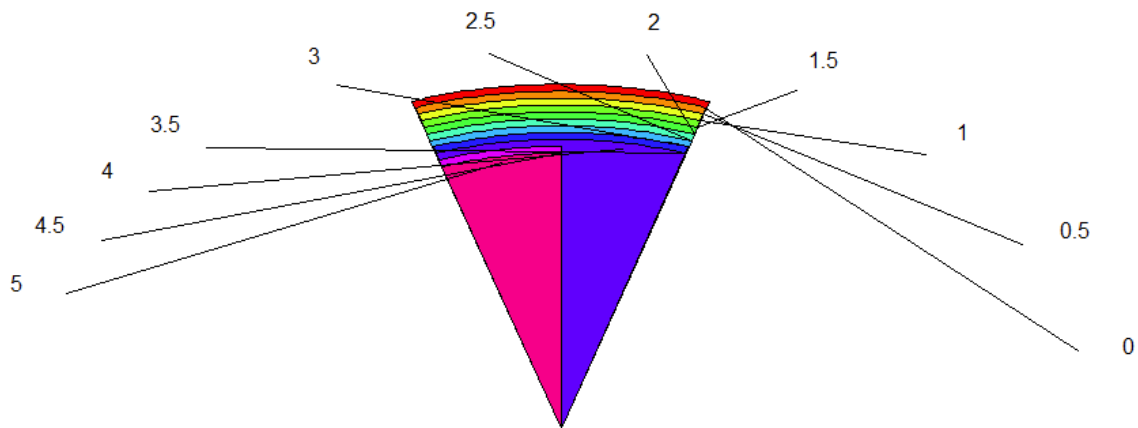
```

+               ifelse(liczba_opinii == 0, 'nie ma', 'mniej niz 50 opinii'))))
> for(i in 1:nrow(lodowki)) {
+   cat(lodowki[i, 1], "ma ocenę klientów", lodowki[i, 5], "bo ma liczbę opinii", lodowki[i, 4], '\n')
+ }
HB ma ocenę klientów 0 bo ma liczbę opinii 0
Curver ma ocenę klientów 0.5 bo ma liczbę opinii 40
Ravanson ma ocenę klientów 1 bo ma liczbę opinii 80
QVANT ma ocenę klientów 1.5 bo ma liczbę opinii 120
Yeticool ma ocenę klientów 2 bo ma liczbę opinii 130
G21 ma ocenę klientów 2.5 bo ma liczbę opinii 10
Camry ma ocenę klientów 3 bo ma liczbę opinii 70
Dometic ma ocenę klientów 3.5 bo ma liczbę opinii 80
Sencor ma ocenę klientów 4 bo ma liczbę opinii 40
Aisberg ma ocenę klientów 4.5 bo ma liczbę opinii 20
Mobicool ma ocenę klientów 5 bo ma liczbę opinii 90
Polarbox ma ocenę klientów 0 bo ma liczbę opinii 150
Fresh ma ocenę klientów 0.5 bo ma liczbę opinii 25
Newestcool ma ocenę klientów 1 bo ma liczbę opinii 55
Fantastic ma ocenę klientów 1.5 bo ma liczbę opinii 5
Polarbox2.0 ma ocenę klientów 2 bo ma liczbę opinii 15
CV60 ma ocenę klientów 2.5 bo ma liczbę opinii 12
CV70 ma ocenę klientów 3 bo ma liczbę opinii 11
CV80 ma ocenę klientów 3.5 bo ma liczbę opinii 5
CV90 ma ocenę klientów 4 bo ma liczbę opinii 2
> write.csv(lodowki, "C:/Users/jaro9/OneDrive/Desktop/apu/lodowki.csv", row.names=FALSE)
> read.csv("C:/Users/jaro9/OneDrive/Desktop/apu/lodowki.csv")
  nazwa waga cena liczba_opinii ocena_klientow status_opinii
1    HB 4.3 300      0      0.0      nie ma
2  Curver 4.4 800     40     0.5  mniej niz 50 opinii
3 Ravanson 4.5 1500    80     1.0  50-100 opinii
4   QVANT 4.6 2000   120     1.5  wiecej niz 100 opinii
5 Yeticool 4.7 3000   130     2.0  wiecej niz 100 opinii
6    G21 4.7 450    10     2.5  mniej niz 50 opinii
7   Camry 4.2 600    70     3.0   50-100 opinii
8  Dometic 4.3 900    80     3.5   50-100 opinii
9   Sencor 4.4 1200   40     4.0  mniej niz 50 opinii
10  Aisberg 4.9 1300   20     4.5  mniej niz 50 opinii
11 Mobicool 4.9 100    90     5.0   50-100 opinii
12 Polarbox 4.7 150   150     0.0  wiecej niz 100 opinii
13  Fresh 4.7 200    25     0.5  mniej niz 50 opinii
14 Newestcool 4.8 300   55     1.0   50-100 opinii
15 Fantastic 5.0 500    5     1.5  mniej niz 50 opinii
16 Polarbox2.0 5.5 550   15     2.0  mniej niz 50 opinii
17   CV60 5.0 2000   12     2.5  mniej niz 50 opinii
18   CV70 4.5 2400   11     3.0  mniej niz 50 opinii
19   CV80 2.0 1000    5     3.5  mniej niz 50 opinii
20   CV90 1.0 800    2     4.0  mniej niz 50 opinii

```

rzuty ekranu:





5. Wnioski: Na podstawie otrzymanego wyniku można stwierdzić, że Język R jest bardzo przydatny do obliczeń statystycznych i wizualizacji wyników

Repo: https://github.com/Jaro233/Zadanie1_APU