

# **SPRAWOZDANIE**

Zajęcia: Analiza Procesów Ucznienia

Prowadzący: prof. dr hab. Vasyl Martsenyuk

## **Laboratorium 1**

30.09.2018

**Temat:** "Podstawy języka R"

**Wariant 1**

Maciej Wojcieszek  
Informatyka II stopień,  
stacjonarne (zaoczne),  
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.

Sprawozdania w postaci:

1. Sprawozdanie (plik .pdf)
  2. plik R 3. wyniki z konsoli (dowolny plik tekstowy) zachować w zdalnym repozytorium (np Github) link na który wysłać w mailu z tematem APU\_Gr\_numer\_grupy na adres mailowy vmartsenyuk@ath.bielsko.pl
- 
- 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. Preprint submitted to Elsevier October 12, 2018
  - 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 połączaniem 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-letnim procesorem. Potem stwórz wektory wyświetlacz, pamięć RAM i pamięć wbudowana, aparat\_foto, cena, liczba\_opinii zawierające kolejno dane 10 smartfonów. Następnie stwórz ramkę danych smartfony z 10 kolumnami 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 cenę 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 2 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 konkatenację 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:

- a)

```
a <- 2 ^ 5
b <- a * 2
max(a, b)
```
- b)

```
help(sum)
??sum
```
- c)

```
a<- c(10, 11, 12, 13, 14, 15)
sum(a)
```
- d)

```
apropos("sum", mode = "function")
```
- e)

```
setwd("C:/Users/noof94/OneDrive - Akademia Techniczno-Humanistyczna/R")
a <- "smartfony Samsung"
save(a, file = "ala.RData")
remove(a)
a
load("ala.RData")
a
```
- f)

```
install.packages("gridExtra")
library(gridExtra)
help(package = "gridExtra")
grid.table(mtcars[1:10, ])
```
- g)

```
seq(100, 20, -4)
```
- h)

```
a <- 9:5
b <- 11:16
d <- c(b, a)
d
```
- i)

```
model <- c("Xiaomi Mi Mix 3",
            "Samsung Galaxy A9",
            "Xiaomi MI 8 Lite",
            "Honor 8X",
            "Huawei P20",
            "Huawei P20 Lite",
            "Samsung Galaxy A7",
            "Honor 9 Lite",
            "Huawei P20 Pro",
            "Xiaomi Redmi 6")
pamiec_RAM <- c("6GB",
                "6GB",
                "4GB",
                "4GB",
                "4GB",
                "4GB",
                "4GB",
                "4GB",
                "3GB",
                "6GB",
                "3GB")
```

```

pamiec_wbudowana <- c("128GB",
                      "128GB",
                      "64GB",
                      "128GB",
                      "64GB",
                      "64GB",
                      "64GB",
                      "64GB",
                      "32GB",
                      "128GB",
                      "32GB")

aparat_foto <- c(
  "12 Mpix + 12 Mpix",
  "24 Mpix + 8 Mpix + 10 Mpix + 5 Mpix",
  "12 Mpix + 5 Mpix",
  "20 Mpix + 2 Mpix",
  "20 Mpix + 12 Mpix",
  "16 Mpix + 2 Mpix",
  "24 Mpix + 5 Mpix + 8 Mpix",
  "13 Mpix + 2 Mpix",
  "40 Mpix + 20 Mpix + 8 Mpix",
  "12 Mpix + 5 Mpix")

cena <- c(2699,
          1990,
          1099,
          1299,
          1799,
          1099,
          1399,
          749,
          2699,
          599)

liczba_opinii <- c(0,
                  7,
                  29,
                  396,
                  11,
                  179,
                  18,
                  53,
                  72,
                  25)

smartfony <- data.frame(model, pamiec_RAM, pamiec_wbudowana, apa-
  rat_foto, cena, liczba_opinii, stringsAsFactors = FALSE)
smartfony

j)
nowy_smartfon <- data.frame(model = "Motorola Moto G6 Play",
  pamiec_RAM = "3GB",
  pamiec_wbudowana = "32GB",
  aparat_foto = "13 Mpix",
  cena = 669,
  liczba_opinii = 304)

smartfony <- rbind(smartfony, nowy_smartfon)
mean(smartfony$cena)

```

k)

```
ocena <- c(0, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5)
smartfony <- cbind(smartfony, ocena)
smartfony
tapply(smartfony$ocena, smartfony$ocena, mean)
```

l)

```
model <- c("Huawei Y7 Prime 2018",
           "Huawei Mate 20 Lite",
           "Honor 10",
           "Samsung Galaxy A6+")
pamiec_RAM <- c("3GB",
                "4GB",
                "4GB",
                "3GB")
pamiec_wbudowana <- c("32GB",
                      "64GB",
                      "64GB",
                      "32GB")
aparat_foto <- c("13 Mpix + 2 Mpix",
                 "20 Mpix + 2 Mpix",
                 "24 Mpix + 16 Mpix",
                 "16 Mpix + 5 Mpix")
cena <- c(699,
          1399,
          1599,
          999)
liczba_opinii <- c(23,
                  17,
                  34,
                  28)

ocena <- c(5, 5, 5, 5)

nowe_smartfony <- data.frame(model, pamiec_RAM, pamiec_wbudowana, aparat_foto, cena,
                             liczba_opinii, ocena)

smartfony <- rbind(smartfony, nowe_smartfony)
smartfony
liczebnosc <- table(smartfony$ocena)
barplot(liczebnosc)
```

m)

```
procenty <- liczebnosc / sum(liczebnosc)
pie(procenty)
library(plotrix)
fan.plot(liczebnosc, labels = names(liczebnosc))
```

n)

```
smartfony[, "status_opini"] <- ifelse(smartfony$liczba_opinii < 1, "Nie ma",
                                     ifelse(smartfony$liczba_opinii < 50, "Mniej niż 50
                                     opini",
                                     ifelse(smartfony$liczba_opinii < 100, "Miedzy 50 a 100",
                                     "Wiecej niz 100"))))

smartfony$status_opini <- factor(smartfony$status_opini)
pie(table(smartfony$status_opini))
```

o)

```
paste(smartfony$model, "ma ocene klientow", smartfony$ocena, "bo ma liczbę opini",
      smartfony$liczba_opinii)
```

p)

```
setwd("C:\\Users\\noof94\\OneDrive - Akademia Techniczno-Humanistyczna\\R\\Zadanie 1")  
a <- smartfony  
save(a, file = "smartfony.csv")
```

## 1. Wykorzystane komendy:

a) R LAB 1

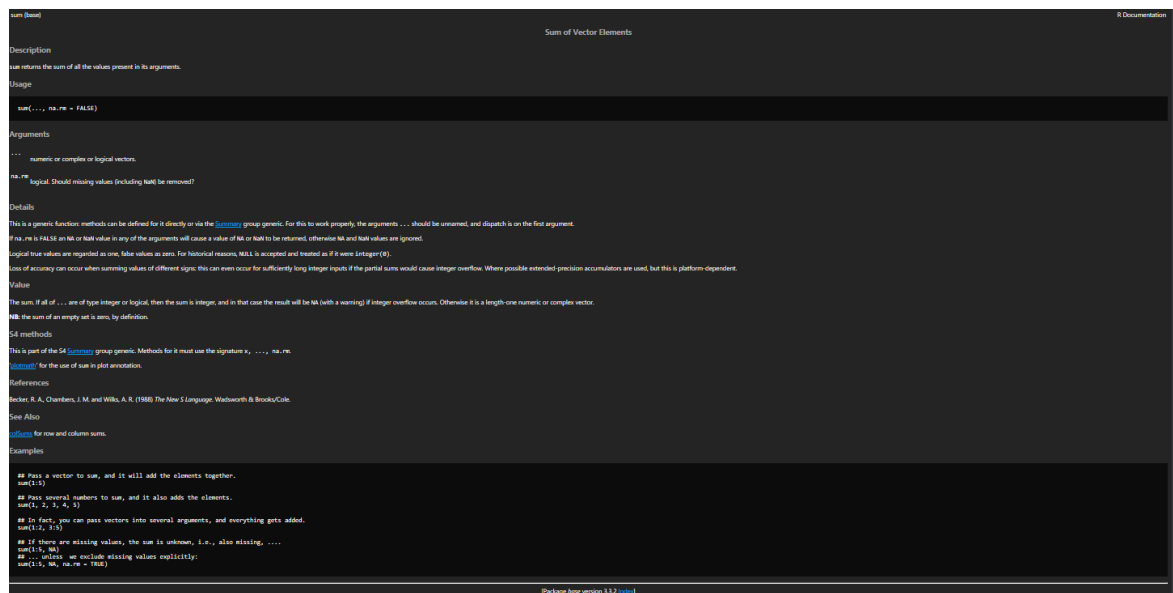
b) smartfony.csv

## 2. Wyniki działań:

a)

```
> a <- 2 ^ 5  
  
> b <- a * 2  
|  
> max(a, b)  
[1] 64
```

b)



The screenshot shows the R Documentation page for the `sum` function. The page title is "Sum (base)". The description states: "sum returns the sum of all the values present in its arguments." The usage is: `sum(..., na.rm = FALSE)`. The arguments section indicates that `...` are numeric or complex or logical vectors, and `na.rm` is a logical value indicating whether missing values should be removed. The details section explains that `sum` is a generic function and that `na.rm = TRUE` will cause a value of NA to be returned if any NA values are present. The value section states that the sum of an empty set is zero by definition. The methods section lists 54 methods for `sum`. The references section cites Becker, R. A., Chambers, J. M., and Wilks, A. R. (1988) *The New S Language*. The examples section shows several uses of `sum`, including summing a vector, summing multiple numbers, summing elements of a matrix, and handling missing values.

c)

```
> a <- c(10, 11, 12, 13, 14, 15)  
  
> sum(a)  
[1] 75
```

d)

```
> apropos("sum", mode = "function")
[1] ".colSums"          ".rowSums"          "colSums"
[4] "contr.sum"         "cumsum"            "format.summaryDefault"
[7] "print.summary.table" "print.summaryDefault" "resume"
[10] "rowsum"            "rowsum.data.frame" "rowsum.default"
[13] "rowSums"           "rxSummary"         "sum"
[16] "summary"           "Summary"           "summary.aov"
[19] "summary.connection" "summary.data.frame" "Summary.data.frame"
[22] "summary.Date"      "Summary.Date"      "summary.default"
[25] "Summary.difftime"  "summary.factor"    "Summary.factor"
[28] "summary.glm"       "summary.lm"        "summary.manova"
[31] "summary.matrix"    "Summary.numeric_version" "Summary.ordered"
[34] "summary.POSIXct"   "Summary.POSIXct"   "summary.POSIXlt"
[37] "Summary.POSIXlt"   "summary.proc_time" "summary.srcfile"
[40] "summary.srcref"    "summary.stepfun"   "summary.table"
[43] "summaryRprof"
```

e)

```
> setwd("C:\\Users\\noof94\\OneDrive - Akademia Techniczno-Humanistyczna\\R")

> a <- "smartfony Samsung"

> save(a, file = "ala.RData")

> remove(a)

> a
Error in eval(expr, envir, enclos) : object 'a' not found
> load("ala.RData")
> a
[1] "smartfony Samsung"
```

f)

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
<i>Mazda RX4</i>	21	6	160	110	3.9	2.62	16.46	0	1	4	4
<i>Mazda RX4 Wag</i>	21	6	160	110	3.9	2.875	17.02	0	1	4	4
<i>Datsun 710</i>	22.8	4	108	93	3.85	2.32	18.61	1	1	4	1
<i>Hornet 4 Drive</i>	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
<i>Hornet Sportabout</i>	18.7	8	360	175	3.15	3.44	17.02	0	0	3	2
<i>Valiant</i>	18.1	6	225	105	2.76	3.46	20.22	1	0	3	1
<i>Duster 360</i>	14.3	8	360	245	3.21	3.57	15.84	0	0	3	4
<i>Merc 240D</i>	24.4	4	146.7	62	3.69	3.19	20	1	0	4	2
<i>Merc 230</i>	22.8	4	140.8	95	3.92	3.15	22.9	1	0	4	2
<i>Merc 280</i>	19.2	6	167.6	123	3.92	3.44	18.3	1	0	4	4

g)

```
> seq(100, 20, -4)
[1] 100 96 92 88 84 80 76 72 68 64 60 56 52 48 44 40 36 32 28 24 20
```

h)

```
> a <- 9:5

> b <- 11:16

> d <- c(b, a)

> d
[1] 11 12 13 14 15 16 9 8 7 6 5
```

i)

```
> smartfony
```

	model	pamiec_RAM	pamiec_wbudowana	aparat_foto	cena	liczba_opinii
1	Xiaomi Mi Mix 3	6GB	128GB	12 Mpix + 12 Mpix	2699	0
2	Samsung Galaxy A9	6GB	128GB	24 Mpix + 8 Mpix + 10 Mpix + 5 Mpix	1990	7
3	Xiaomi MI 8 Lite	4GB	64GB	12 Mpix + 5 Mpix	1099	29
4	Honor 8X	4GB	128GB	20 Mpix + 2 Mpix	1299	396
5	Huawei P20	4GB	64GB	20 Mpix + 12 Mpix	1799	11
6	Huawei P20 Lite	4GB	64GB	16 Mpix + 2 Mpix	1099	179
7	Samsung Galaxy A7	4GB	64GB	24 Mpix + 5 Mpix + 8 Mpix	1399	18
8	Honor 9 Lite	3GB	32GB	13 Mpix + 2 Mpix	749	53
9	Huawei P20 Pro	6GB	128GB	40 Mpix + 20 Mpix + 8 Mpix	2699	72
10	Xiaomi Redmi 6	3GB	32GB	12 Mpix + 5 Mpix	599	25

j)

```
> nowy_smartfon <- data.frame(model = "Motorola Moto G6 Play",
+                             pamiec_RAM = "3GB",
+                             pamiec_ .... [TRUNCATED]

> smartfony <- rbind(smartfony, nowy_smartfon)

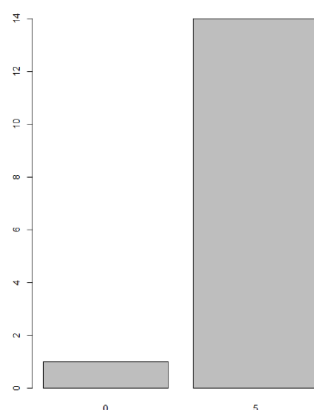
> mean(smartfony$cena)
[1] 1463.636
```

k)

```
model pamiec_RAM pamiec_wbudowana aparat_foto cena liczba_opinii ocena
1 Xiaomi Mi Mix 3 6GB 128GB 12 Mpix + 12 Mpix 2699 0 0
2 Samsung Galaxy A9 6GB 128GB 24 Mpix + 8 Mpix + 10 Mpix + 5 Mpix 1990 7 5
3 Xiaomi MI 8 Lite 4GB 64GB 12 Mpix + 5 Mpix 1099 29 5
4 Honor 8X 4GB 128GB 20 Mpix + 2 Mpix 1299 396 5
5 Huawei P20 4GB 64GB 20 Mpix + 12 Mpix 1799 11 5
6 Huawei P20 Lite 4GB 64GB 16 Mpix + 2 Mpix 1099 179 5
7 Samsung Galaxy A7 4GB 64GB 24 Mpix + 5 Mpix + 8 Mpix 1399 18 5
8 Honor 9 Lite 3GB 32GB 13 Mpix + 2 Mpix 749 53 5
9 Huawei P20 Pro 6GB 128GB 40 Mpix + 20 Mpix + 8 Mpix 2699 72 5
10 Xiaomi Redmi 6 3GB 32GB 12 Mpix + 5 Mpix 599 25 5
11 Motorola Moto G6 Play 3GB 32GB 13 Mpix 669 304 5

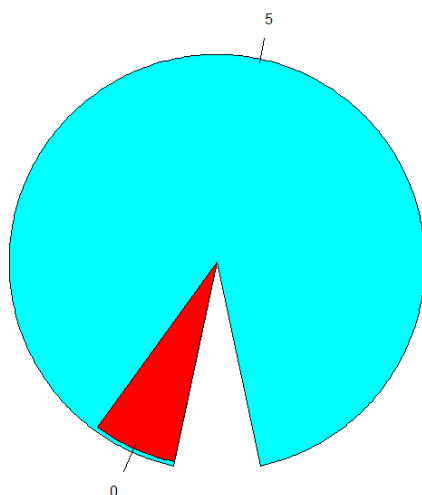
> tapply(smartfony$cena, smartfony$ocena, mean)
0 5
2699.0 1340.1
>
```

l)

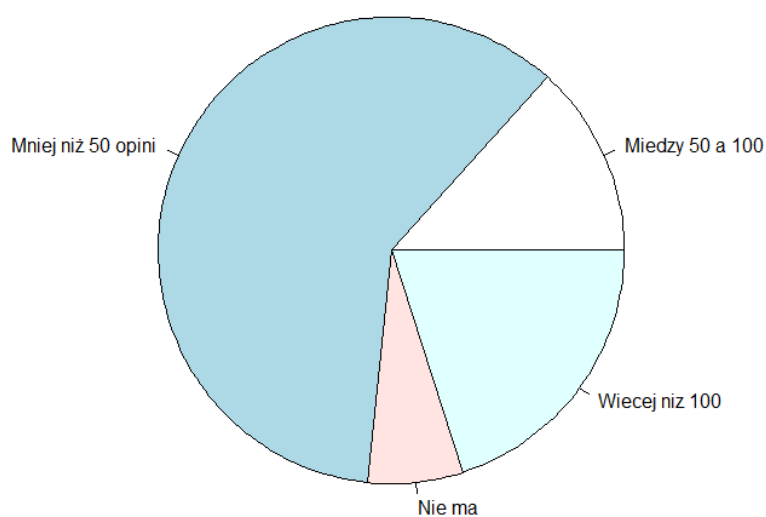


m)





n)



o)

```
[1] "Xiaomi Mi Mix 3 ma ocene klientow 0 bo ma liczbe opini 0"
[2] "Samsung Galaxy A9 ma ocene klientow 5 bo ma liczbe opini 7"
[3] "Xiaomi MI 8 Lite ma ocene klientow 5 bo ma liczbe opini 29"
[4] "Honor 8X ma ocene klientow 5 bo ma liczbe opini 396"
[5] "Huawei P20 ma ocene klientow 5 bo ma liczbe opini 11"
[6] "Huawei P20 Lite ma ocene klientow 5 bo ma liczbe opini 179"
[7] "Samsung Galaxy A7 ma ocene klientow 5 bo ma liczbe opini 18"
[8] "Honor 9 Lite ma ocene klientow 5 bo ma liczbe opini 53"
[9] "Huawei P20 Pro ma ocene klientow 5 bo ma liczbe opini 72"
[10] "Xiaomi Redmi 6 ma ocene klientow 5 bo ma liczbe opini 25"
[11] "Motorola Moto G6 Play ma ocene klientow 5 bo ma liczbe opini 304"
[12] "Huawei Y7 Prime 2018 ma ocene klientow 5 bo ma liczbe opini 23"
[13] "Huawei Mate 20 Lite ma ocene klientow 5 bo ma liczbe opini 17"
[14] "Honor 10 ma ocene klientow 5 bo ma liczbe opini 34"
[15] "Samsung Galaxy A6+ ma ocene klientow 5 bo ma liczbe opini 28"
```

### **3. Wnioski:**

Po przerobieniu wszystkich tych zadań udało się zdobyć wiedzę, która umożliwi wykonywanie innych bardziej skomplikowanych zadań w przyszłości.