

# Kviz 1

Miha Prajs

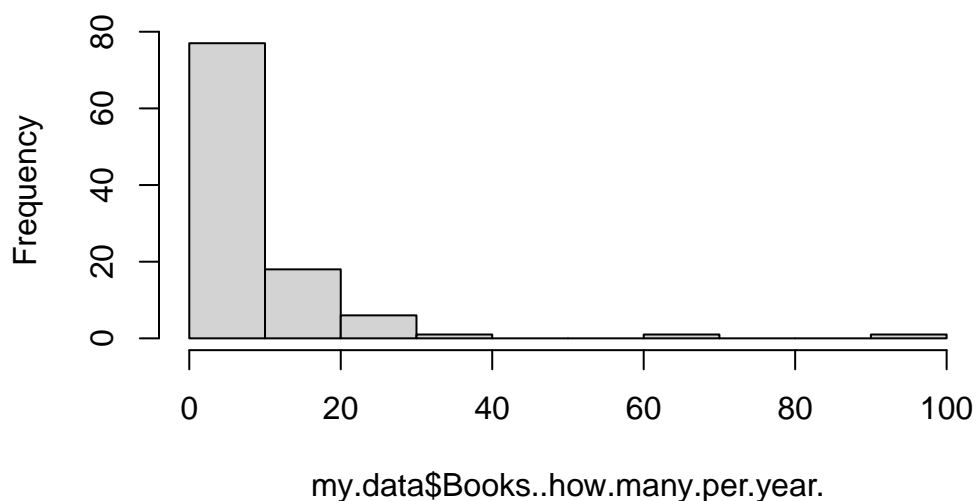
## Priprava podatkov

```
my.data <-  
  read.table("C:/Users/mihap/Code/Faks/famnit24-statistika/HW/data.txt",  
    header=TRUE, stringsAsFactors=TRUE, sep=" ", na.strings="NA", dec=".",  
    strip.white=TRUE)  
x <- 6  
z <- 4  
my.data <- my.data[unique(c(seq(x,  
                                nrow(my.data)  
                                ,by=10),  
                            seq(z,  
                                nrow(my.data)  
                                ,by=10))),]
```

## 1. naloga

```
hist(my.data$Books..how.many.per.year.)
```

### Histogram of my.data\$Books..how.many.per.year.



```
median(my.data$Books..how.many.per.year.,na.rm=T)
```

```
[1] 9
```

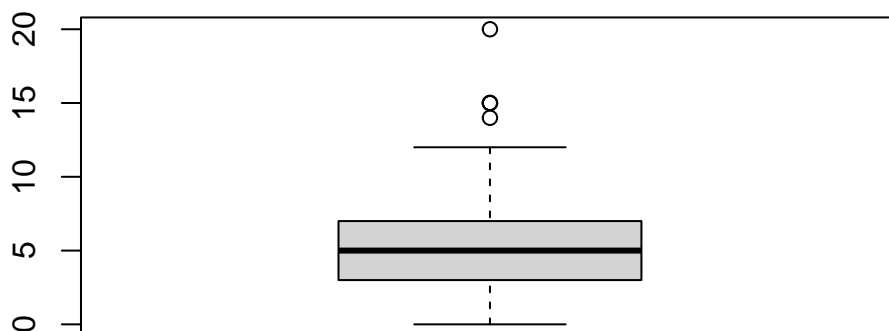
```
IQR(my.data$Books..how.many.per.year.,na.rm=T)
```

```
[1] 8
```

Histogram je desno asimetričen. Mediana je ustrezna mera središčnosti, saj prikaže srednjo vrednost med danimi podatki. Interkvartilni razpon pa primerna mera variabilnosti, saj nam pove koliko vzorcev predstavlja 50 % dane populacije.

## 2. naloga

```
boxplot(my.data$Sport..hours.per.week.)
```



```
osamelci <- boxplot.stats(my.data$Sport..hours.per.week.)$out
osamelci
```

```
[1] 20 15 14 15 15
```

Iz grafa je razvidno, da je osamelcev več. Z uporabo funkcije `boxplot.stats([data])$out` ugotovimo koliko in kateri so, v našem primeru jih je 5 in so 14, 15, 15, 15 in 20.

### 3. naloga

```
krog <- function(r) {
  obseg <- pi*r^2
  return(obseg)
}
krog(x)
```

```
[1] 113.0973
```

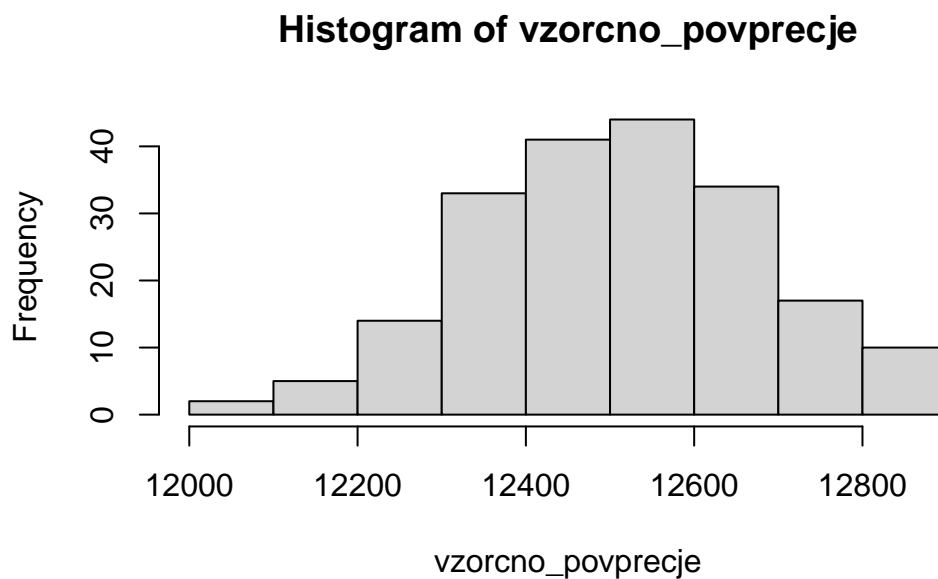
## 4. naloga

### Priprava podatkov

```
podatki <- read.csv("./podatki.csv",sep=";")
```

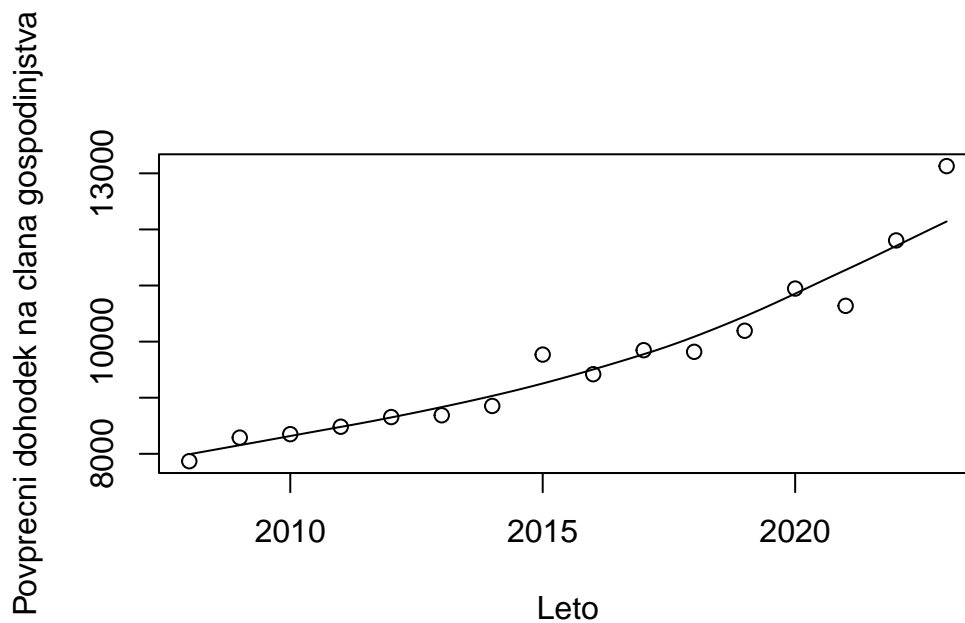
a)

```
populacija <- podatki$X2023.Dohodek.po.socialnih.transferjih  
  
mean_pop <- mean(populacija)  
  
vzorcnopovprecje <- c(1:200)  
for (i in 1:200) {  
  vzorec <- sample(populacija,size=6)  
  vzorcnopovprecje[i] <- mean(vzorec)  
}  
hist(vzorcnopovprecje)
```



b)

```
regija <- subset(podatki, STATISTICNA.REGIJA=="Obalno-kraska")
regija <- regija[-1]
rownames(regija) <- regija[[1]]
regija <- regija[-1]
regija <- unlist(regija)
leto <- c(2008:2023)
names(regija) <- leto
scatter.smooth(x=leto,
               y=regija,
               xlab="Leto",
               ylab="Povprečni dohodek na člana gospodinjstva")
```



c)

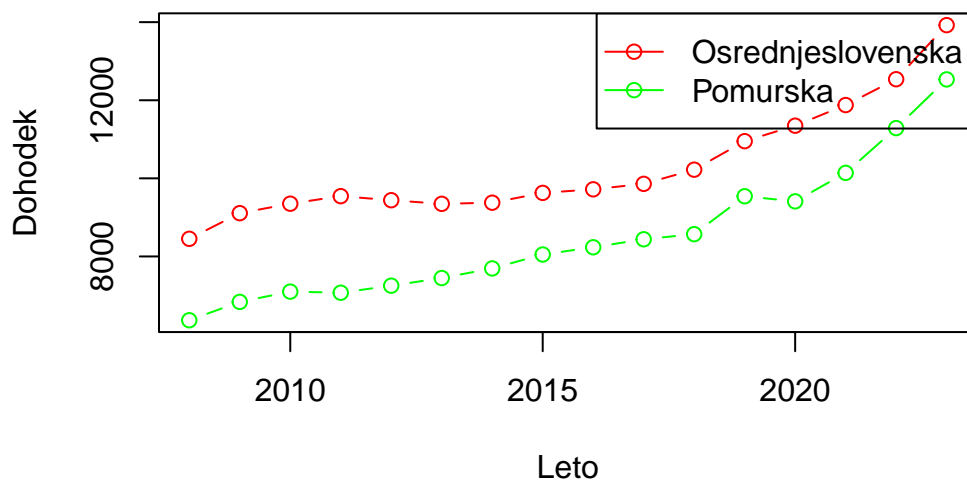
```
osrednjeslovenska <- subset(podatki, STATISTICNA.REGIJA == "Osrednjeslovenska")
pomurska <- subset(podatki, STATISTICNA.REGIJA == "Pomurska")
dohodek_osrednjeslovenska <- as.numeric(osrednjeslovenska[3:18])
```

```
dohodek_pomurska <- as.numeric(pomurska[3:18])

plot(leto, dohodek_osrednjeslovenska, type = "b", col = "red",
     ylim = range(c(dohodek_osrednjeslovenska, dohodek_pomurska)),
     main = "Primerjava dohodkov - Osrednjeslovenska in Pomurska",
     xlab = "Leto", ylab = "Dohodek")
lines(leto, dohodek_pomurska, type = "b", col = "green")

legend("topright", legend = c("Osrednjeslovenska", "Pomurska"),
      col = c("red", "green"), lty = 1, pch = 1)
```

## Primerjava dohodkov – Osrednjeslovenska in Pomurska



```
povprecna_razlika <- mean(dohodek_osrednjeslovenska - dohodek_pomurska)
povprecna_razlika
```

```
[1] 1794.688
```

Povprečna razlika med teh vrednosti je 1794,68.