

Kviz 1

Statistika 2024/25

Jan Panjan

2024-11-03

```
library(Rcmdr)
library(tidyverse)
```

```
my.data <- read.delim("data.txt")
```

```
x <- 2
z <- 8
```

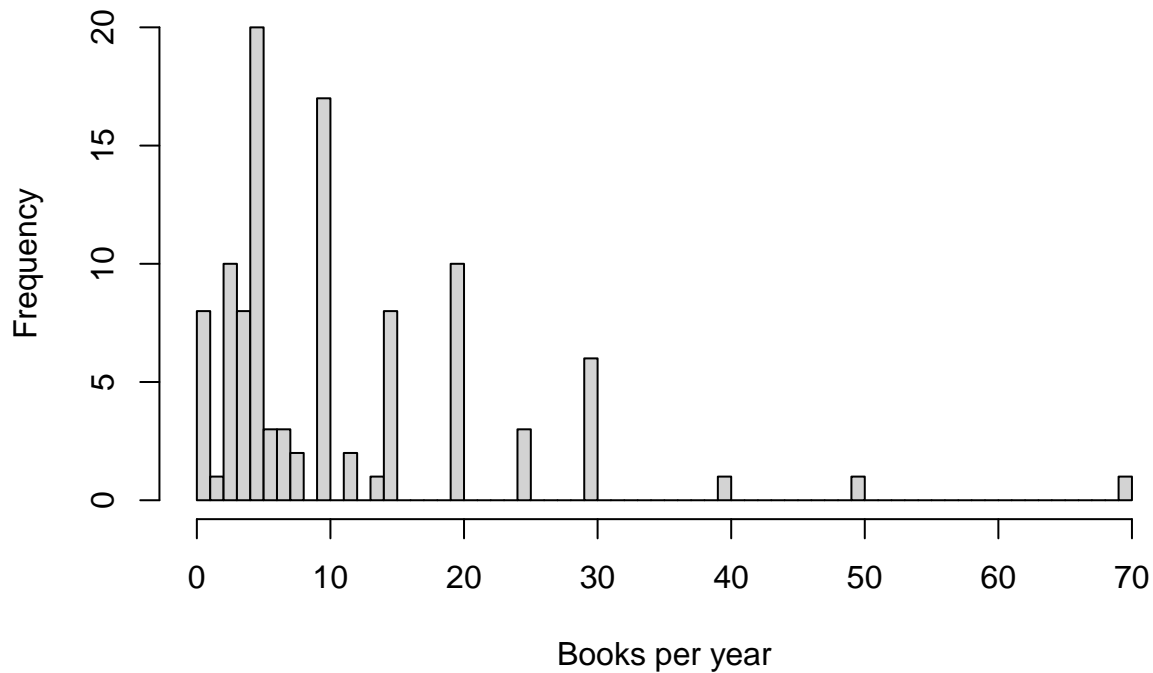
```
my.data <- my.data[unique(c(seq(x, nrow(my.data), 10), seq(z, nrow(my.data), 10))), ] %>%
  as.data.frame()
```

Books..how.many.per.year

Narišite histogram spremenljivke Books..how.many.per.year. Kakšne oblike je porazdelitev? Poročajte ustrezno mero središčnosti in mero variabilnosti glede na obliko porazdelitve. Utemeljite vašo odločitev.

```
col <- my.data$Books..how.many.per.year
hist(col,
      xlab = "Books per year",
      breaks = 75,
      main = "Histogram od Books..how.many.per.year")
```

Histogram od Books..how.many.per.year



```
Me <- median(col)
avg <- mean(col)
Me;avg
```

```
## [1] 7
```

```
## [1] 11.09524
```

Porazdelitev je asimetrična (v desno).

Mera središčnosti - Mediana, saj imajo podatki osamelce, ki kvarijo vrednost aritmetične sredine in ker je porazdelitev asimetrična.

Mera variabilnosti - Interkvartilni razpon, zaradi osamelcev.

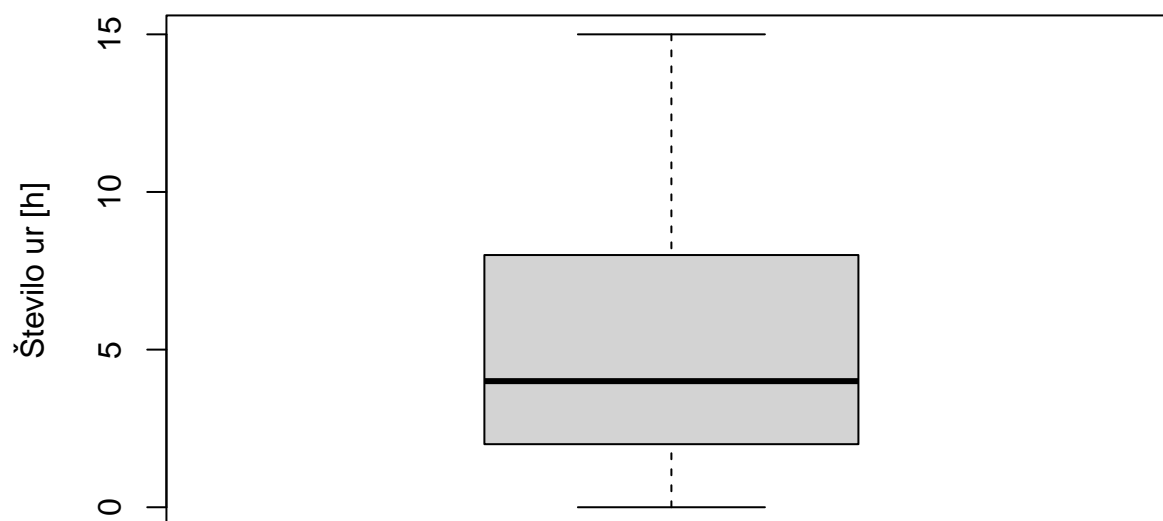
Sport..hours.per.week..

Narišite okvir z ročaji spremenljivke Sport..hours.per.week.. Poročajte vse pomembne vrednosti iz grafa in jih interpretirajte. Ali vidite kakšen osamelec? Obrazložite izraz.

```
col <- my.data$Sport..hours.per.week.
```

```
boxplot(col,
  main = "Okvir z ročaji za ure športanja na teden",
  ylab = "Število ur [h]")
```

Okvir z rocaji za ure športanja na teden



Max vrednost je 15, min vrednost je 0. Mediana je malo manj kot 5. Večina vrednosti se pojavi med cca 2.5 in 8. Mediana je manjša od aritmetične sredine, torej je porazdelitev asimetrična (v desno).

Osamelci so vrednosti, ki se veliko razlikujejo od večine ostalih vrednosti.

Obseg kroga