

Mobilnost turista i poslovanje hrvatskog turističkog sektora

Pacici

2022-12-06

Ucitavanje i obrada podataka

Ucitavanje svih potrebnih podataka i obrada istih, uklonjene su strsece vrijednosti iz podataka o prometu za domaci i strani promet na granicnim prijelazima kako bi dobili jasniju sliku podataka.

Ucitavanje cijena i fundamenata cetiri velika poduzeca u Hrvatskoj i cijene indeksa CROBEXturist Zagrebacke burze.

Ucitavanje podataka o dnevnom ulasku i izlasku turista iz Hrvatske, podijeljeno po vrsti granicnog prijelaza.

```
domaci_ulaz <- read_promet("podaci/domaci_promet/domaci_ulaz_putnici.csv")
domaci_izlaz <- read_promet("podaci/domaci_promet/domaci_izlaz_putnici.csv")

strani_ulaz <- read_promet("podaci/strani_promet/strani_ulaz_putnici.csv")
strani_izlaz <- read_promet("podaci/strani_promet/strani_izlaz_putnici.csv")

domaci_ulaz$Dan <- factor(domaci_ulaz$Dan, levels = c("Mon", "Tue", "Wed", "Thu", "Fri", "Sat", "Sun"))
domaci_izlaz$Dan <- factor(domaci_izlaz$Dan, levels = c("Mon", "Tue", "Wed", "Thu", "Fri", "Sat", "Sun"))
strani_ulaz$Dan <- factor(strani_ulaz$Dan, levels = c("Mon", "Tue", "Wed", "Thu", "Fri", "Sat", "Sun"))
strani_izlaz$Dan <- factor(strani_izlaz$Dan, levels = c("Mon", "Tue", "Wed", "Thu", "Fri", "Sat", "Sun"))

#Uklanjanje Outliera u prometu
domaci_ulaz <- subset(domaci_ulaz, domaci_ulaz$UKUPNO < 1e6)
domaci_izlaz <- subset(domaci_izlaz, domaci_izlaz$UKUPNO < 1e6)

strani_ulaz <- subset(strani_ulaz, strani_ulaz$UKUPNO < 1e6)
strani_izlaz <- subset(strani_izlaz, strani_izlaz$UKUPNO < 1e6)

#za riječni promet uocavamo jedan drastični outlier > 2000, pogledajmo datum kad je to bilo (oko 3000 l)
domaci_ulaz <- subset(domaci_ulaz, domaci_ulaz$Riječni < 2000)
domaci_izlaz <- subset(domaci_izlaz, domaci_izlaz$Riječni < 2000)

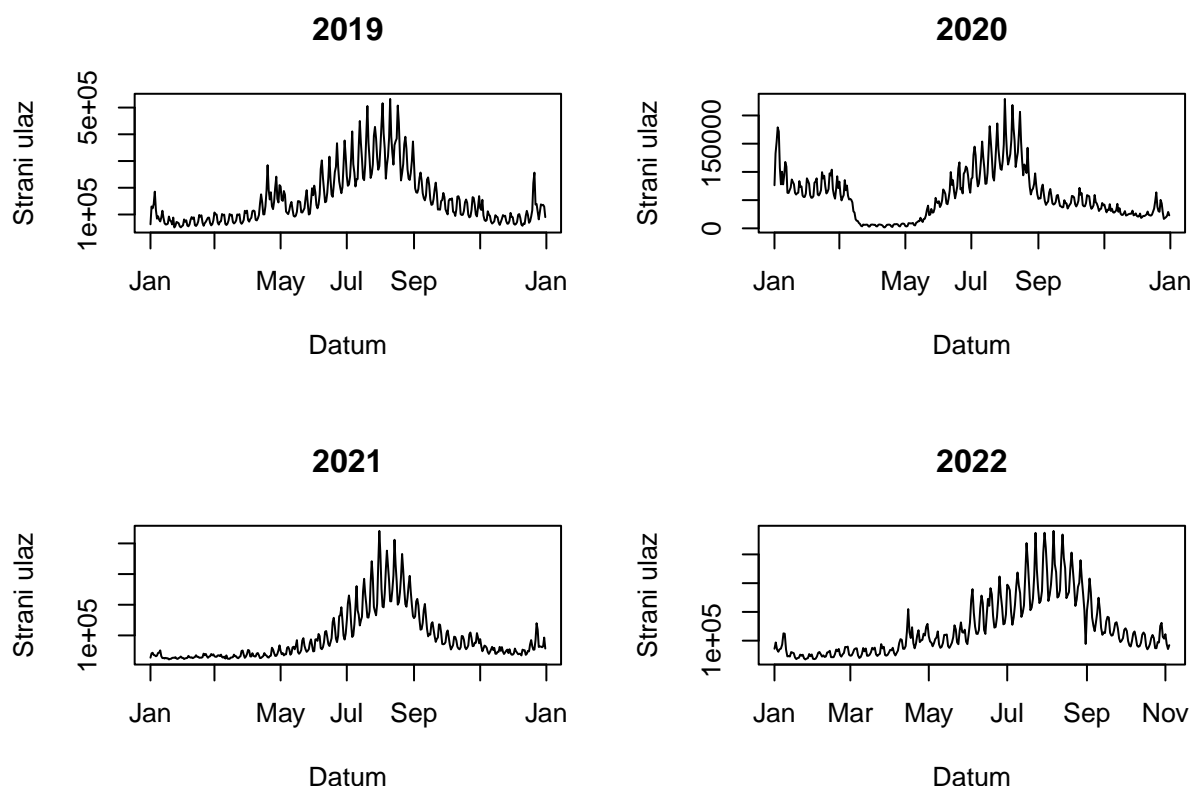
strani_ulaz <- subset(strani_ulaz, strani_ulaz$Riječni < 2000)
strani_izlaz <- subset(strani_izlaz, strani_izlaz$Riječni < 2000)

#za pomorski nekoliko podataka koji odstupaju, realni su ali ih micemo da jasnije vidimo kompletnu sliku
domaci_ulaz <- subset(domaci_ulaz, domaci_ulaz$Pomorski < 1000)
domaci_izlaz <- subset(domaci_izlaz, domaci_izlaz$Pomorski < 1000)
```

Analiza podataka o dnevnom prometu preko granicnih prijelaza

Uspoređujemo strani ulaz po godinama u periodu od zadnje 4 godine (2019. - 2022.)

```
promet <- strani_ulaz
par(mfrow = c(2, 2))
plot(UKUPNO ~ X, data = subset(promet, promet[, "Godina"] == 2019), type = "l", xlab = "Datum", ylab = "Strani ulaz")
plot(UKUPNO ~ X, data = subset(promet, promet[, "Godina"] == 2020), type = "l", xlab = "Datum", ylab = "Strani ulaz")
plot(UKUPNO ~ X, data = subset(promet, promet[, "Godina"] == 2021), type = "l", xlab = "Datum", ylab = "Strani ulaz")
plot(UKUPNO ~ X, data = subset(promet, promet[, "Godina"] == 2022), type = "l", xlab = "Datum", ylab = "Strani ulaz")
```



```
print(max(subset(promet, promet[, "Godina"] == 2019)$UKUPNO))
```

```
## [1] 532290
```

```
print(max(subset(promet, promet[, "Godina"] == 2020)$UKUPNO))
```

```
## [1] 229293
```

```
print(max(subset(promet, promet[, "Godina"] == 2021)$UKUPNO))
```

```
## [1] 440641
```

```
print(max(subset(promet, promet[, "Godina"] == 2022)$UKUPNO))
```

```
## [1] 481596
```

Vidimo nagli pad uzrokovan zatvaranjem granica pocetkom covid pandemije u 2020. godini. Takoder u

vrhuncu sezone 2020. godine imamo 3 puta manje turista nego 2019. godine. U 2021. i 2022. vidimo lagani oporavak, međutim vrhunci u 2021. i 2022. su redom manji za 17% i 10% nego u 2019. godini.

Postoji li razlika među vrstama granicnih prijelaza te među različitim danima u tjednu ili mjesecima

Plotamo domaci i strani ulaz i izlaz po mjesecima i danima za ukupan promet i posebno za svaku vrstu prometa.

Ukupni promet

```
par(mfrow = c(1, 2))
```

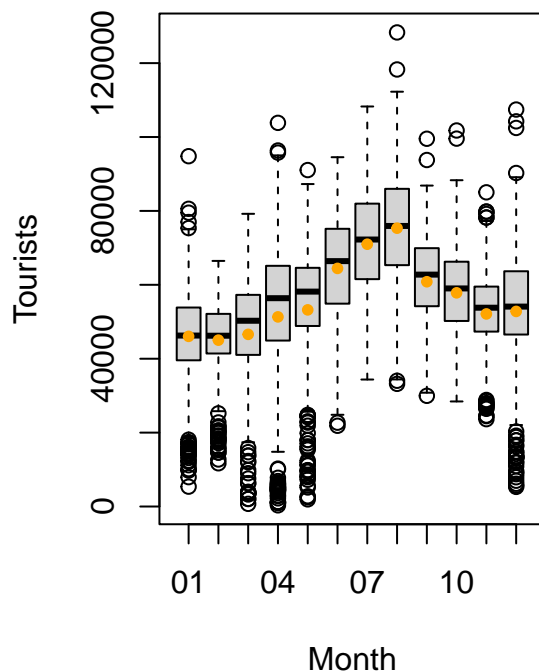
```
#Domaci ulaz turista po mjesecu
```

```
boxplot(domaci_ulaz$UKUPNO ~ domaci_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci ulaz t  
means <- tapply(domaci_ulaz$UKUPNO, domaci_ulaz$Mjesec, mean)  
points(means, pch = 20, col = "orange")
```

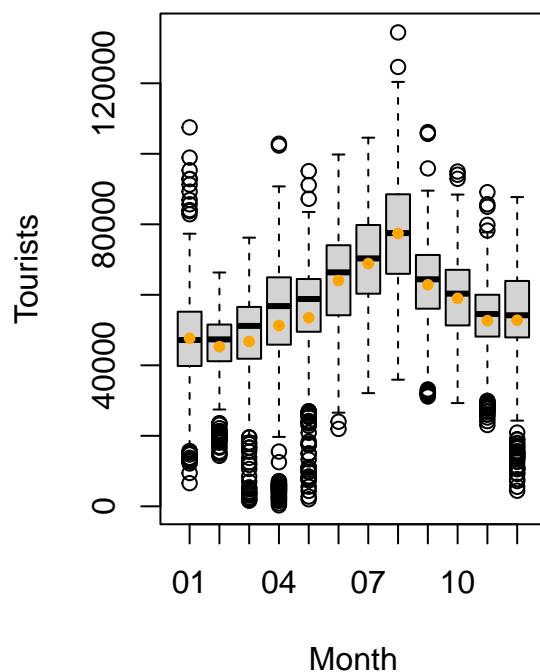
```
#Domaci izlaz turista po mjesecu
```

```
boxplot(domaci_izlaz$UKUPNO ~ domaci_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci izlaz  
means <- tapply(domaci_izlaz$UKUPNO, domaci_izlaz$Mjesec, mean)  
points(means, pch = 20, col = "orange")
```

Domaci ulaz turista po mjesecu



Domaci izlaz turista po mjesecu



```
par(mfrow = c(1, 2))
```

```
#Strani ulaz turista po mjesecu
```

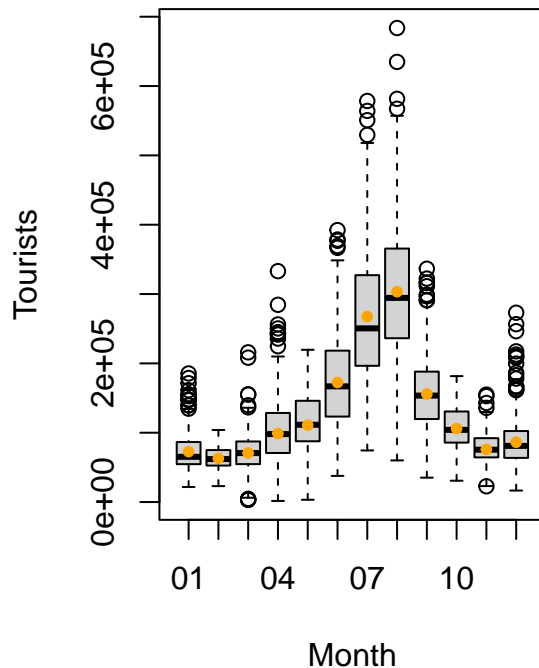
```

boxplot(strani_ulaz$UKUPNO ~ strani_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani ulaz t
means <- tapply(strani_ulaz$UKUPNO, strani_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")

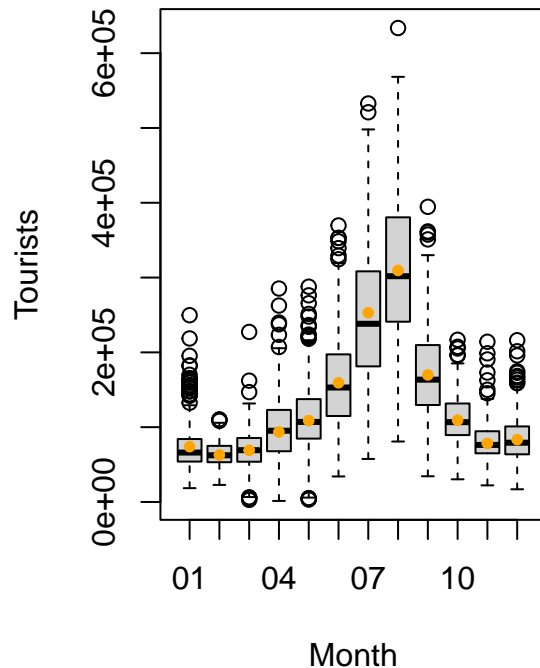
#Strani izlaz turista po mjesecu
boxplot(strani_izlaz$UKUPNO ~ strani_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani izlaz
means <- tapply(strani_izlaz$UKUPNO, strani_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")

```

Strani ulaz turista po mjesecu



Strani izlaz turista po mjesecu



```

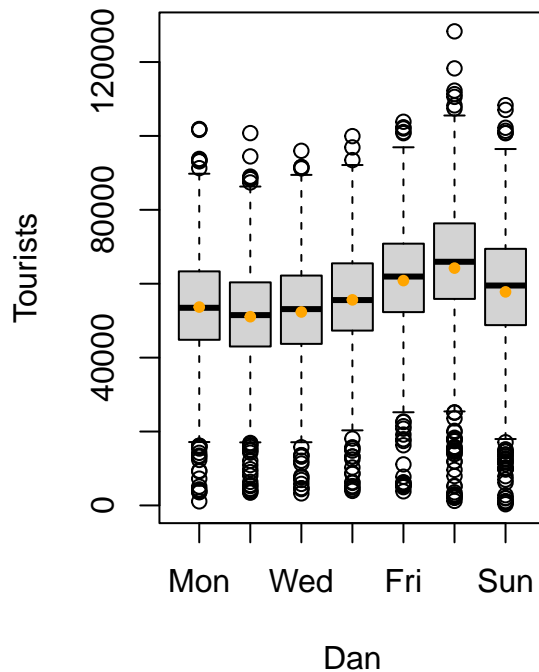
par(mfrow = c(1, 2))

#Domaci ulaz turista po danima
boxplot(domaci_ulaz$UKUPNO ~ domaci_ulaz$Dan, xlab = "Dan", ylab = "Tourists", main="Domaci ulaz turist
means <- tapply(domaci_ulaz$UKUPNO, domaci_ulaz$Dan, mean)
points(means, pch = 20, col = "orange")

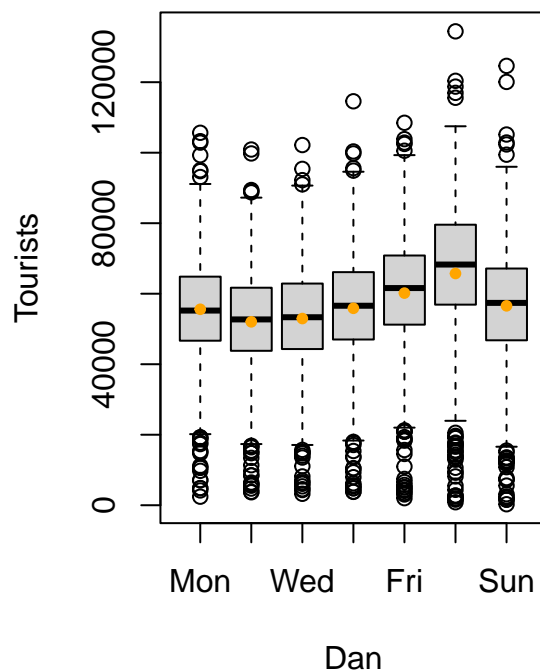
#Domaci izlaz turista po danima
boxplot(domaci_izlaz$UKUPNO ~ domaci_izlaz$Dan, xlab = "Dan", ylab = "Tourists", main="Domaci izlaz tur
means <- tapply(domaci_izlaz$UKUPNO, domaci_izlaz$Dan, mean)
points(means, pch = 20, col = "orange")

```

Domaci ulaz turista po danima



Domaci izlaz turista po danima



Iz podataka jasno vidimo najveće prosjeke u 7. i 8. mjesecu u domaćem i stranom prometu te da su ulaz i izlaz turista najveći subotom.

7. i 8. mjesec smo uzeli kao sezonu i plotali promet po danu tijekom ta dva mjeseca.

Ukupni promet u sezoni

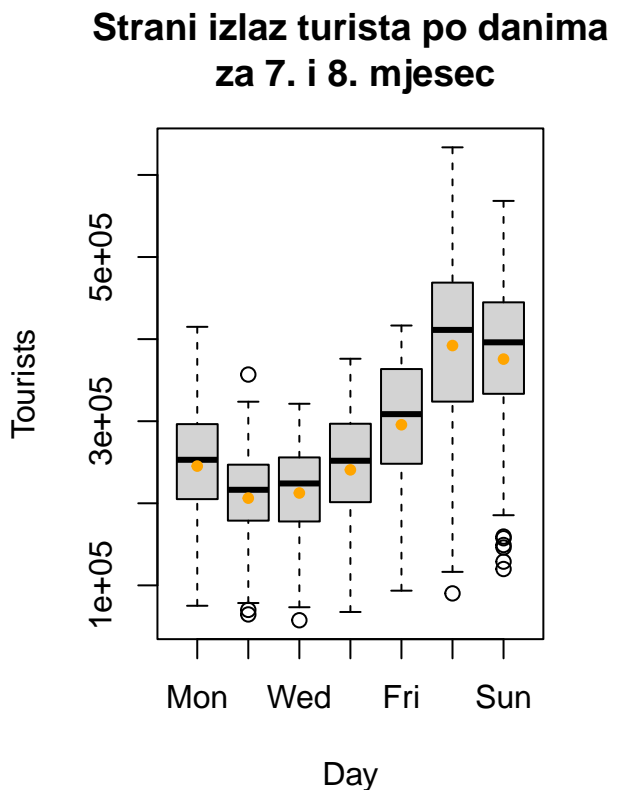
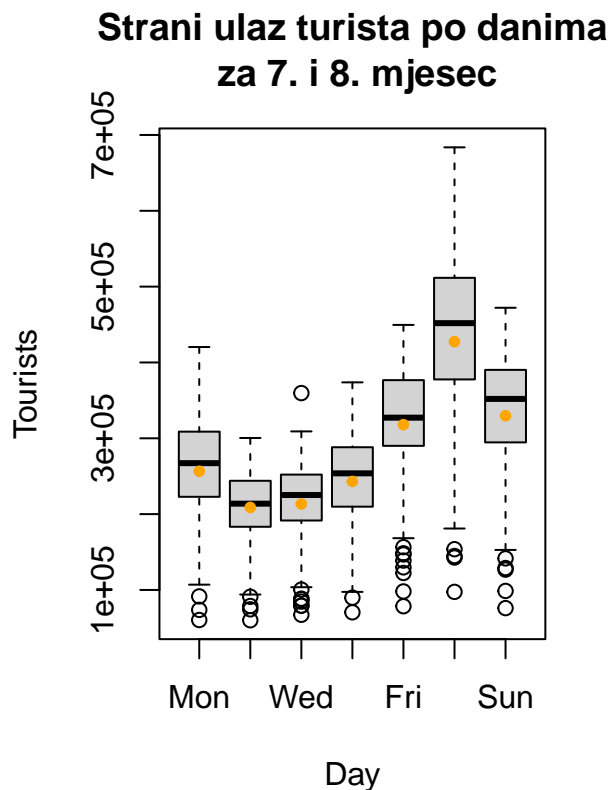
```
sezona = c("07", "08")

#strani sezona
strani_ulaz_sezona <- subset(strani_ulaz, strani_ulaz$Mjesec %in% sezona)
strani_izlaz_sezona <- subset(strani_izlaz, strani_izlaz$Mjesec %in% sezona)

par(mfrow = c(1, 2))

#Strani ulaz turista po danima za 7. i 8. mjesec
boxplot(strani_ulaz_sezona$UKUPNO ~ strani_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani ulaz")
means <- tapply(strani_ulaz_sezona$UKUPNO, strani_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

#Strani izlaz turista po danima za 7. i 8. mjesec
boxplot(strani_izlaz_sezona$UKUPNO ~ strani_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani izlaz")
means <- tapply(strani_izlaz_sezona$UKUPNO, strani_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")
```



*#vidimo po danima da je ulaz najveći u petak, subotu i nedjelju (malo veći brojevi za sub i ned)
#izlaz daleko najveći u subotu i nedjelju*

Vidimo po danima da je ulaz turista najveći u petak, subotu i nedjelju (malo veći brojevi subotom i nedjeljom), dok je izlaz daleko najveći u subotu i nedjelju.

željeznički promet

```
par(mfrow = c(2, 2))
```

#Domaci ulaz turista po mjesecima

```
boxplot(domaci_ulaz$Željeznički ~ domaci_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci u  
means <- tapply(domaci_ulaz$Željeznički, domaci_ulaz$Mjesec, mean)  
points(means, pch = 20, col = "orange")
```

#Domaci izlaz turista po mjesecima

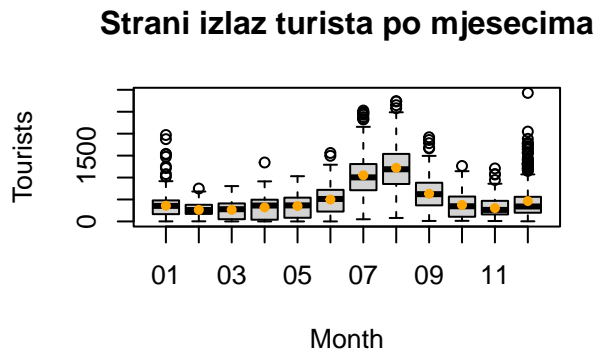
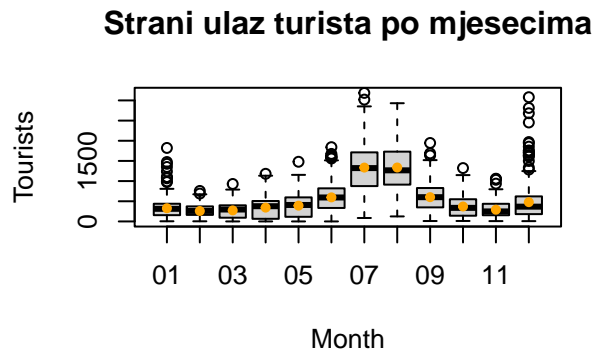
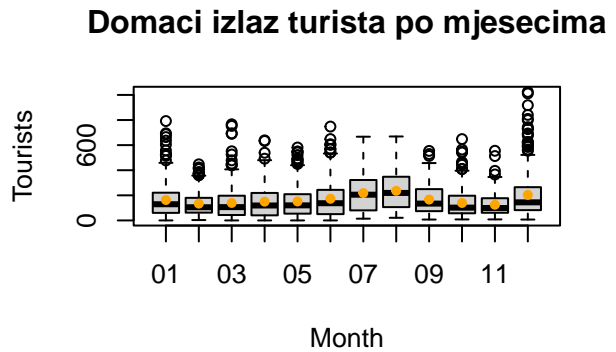
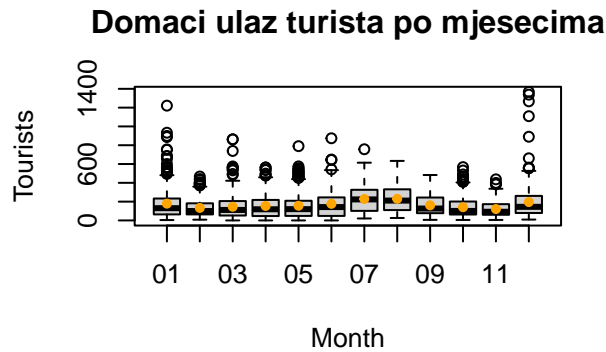
```
boxplot(domaci_izlaz$Željeznički ~ domaci_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci i  
means <- tapply(domaci_izlaz$Željeznički, domaci_izlaz$Mjesec, mean)  
points(means, pch = 20, col = "orange")
```

#Strani ulaz turista po mjesecima

```
boxplot(strani_ulaz$Željeznički ~ strani_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani u  
means <- tapply(strani_ulaz$Željeznički, strani_ulaz$Mjesec, mean)  
points(means, pch = 20, col = "orange")
```

#Strani izlaz turista po mjesecima

```
boxplot(strani_izlaz$Željeznički ~ strani_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani izlaz",
means <- tapply(strani_izlaz$Željeznički, strani_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

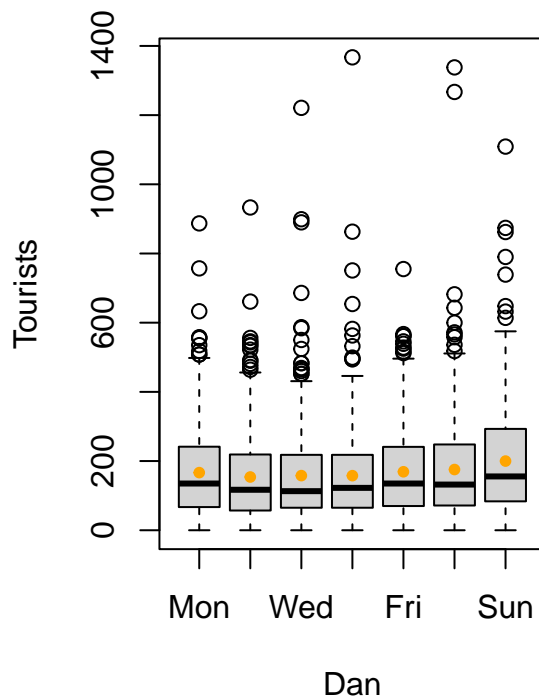


```
par(mfrow = c(1, 2))

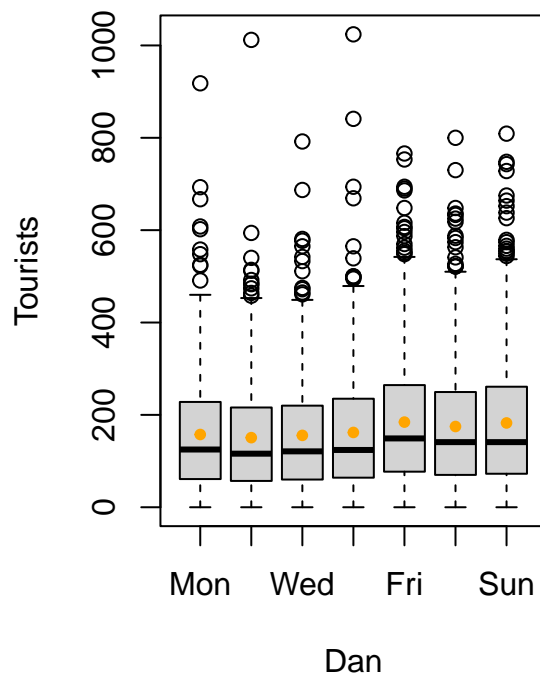
#Domaci ulaz turista po danima
boxplot(domaci_ulaz$Željeznički ~ domaci_ulaz$Dan, xlab = "Dan", ylab = "Tourists", main="Domaci ulaz t",
means <- tapply(domaci_ulaz$Željeznički, domaci_ulaz$Dan, mean)
points(means, pch = 20, col = "orange")

#Domaci izlaz turista po danima
boxplot(domaci_izlaz$Željeznički ~ domaci_izlaz$Dan, xlab = "Dan", ylab = "Tourists", main="Domaci izlaz",
means <- tapply(domaci_izlaz$Željeznički, domaci_izlaz$Dan, mean)
points(means, pch = 20, col = "orange")
```

Domaci ulaz turista po danima



Domaci izlaz turista po danima



Najveći prosjeci su u 7. i 8. mjesecu i u domaćem i u stranom prometu, izraženije u stranom. Također, prosinac i siječanj su malo bolji od listopada, studenog, veljače i ožujka. Nema značajne razlike po danima u domaćem prometu.

željeznički promet u sezoni

```
sezona = c("07", "08")

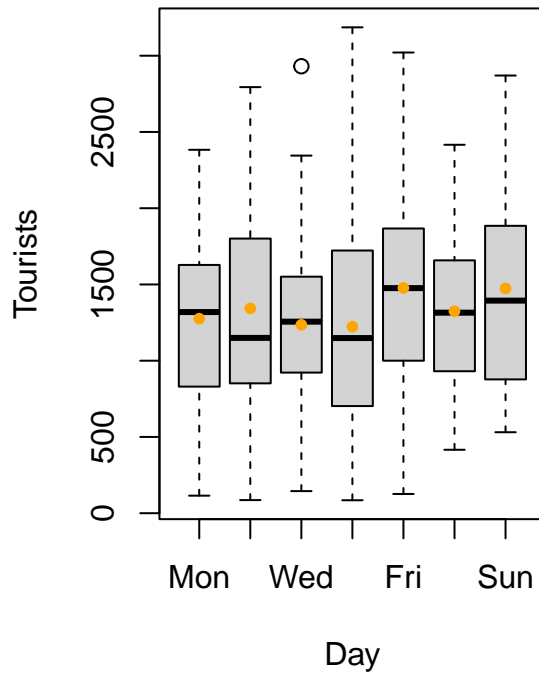
#strani sezona
strani_ulaz_sezona <- subset(strani_ulaz, strani_ulaz$Mjesec %in% sezona)
strani_izlaz_sezona <- subset(strani_izlaz, strani_izlaz$Mjesec %in% sezona)

par(mfrow = c(1, 2))

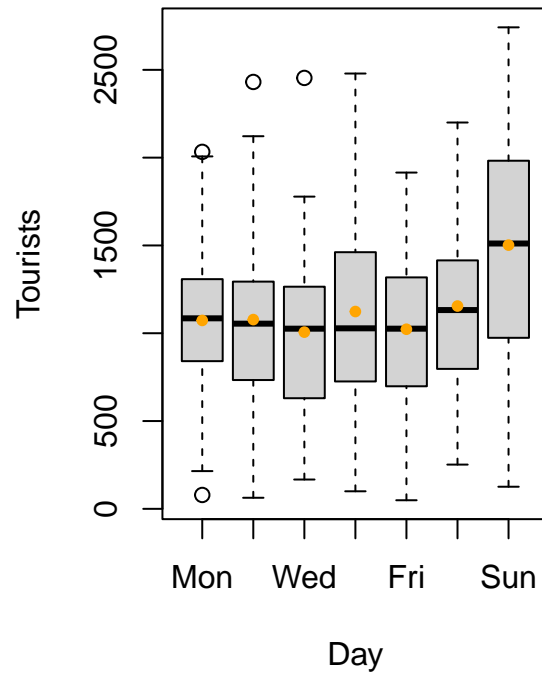
#Strani ulaz turista po danima za 7. i 8. mjesec
boxplot(strani_ulaz_sezona$Željeznički ~ strani_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani ulaz")
means <- tapply(strani_ulaz_sezona$Željeznički, strani_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

#Strani izlaz turista po danima za 7. i 8. mjesec
boxplot(strani_izlaz_sezona$Željeznički ~ strani_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani izlaz")
means <- tapply(strani_izlaz_sezona$Željeznički, strani_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")
```


**Strani ulaz turista po danima
za 7. i 8. mjesec**



**Strani izlaz turista po danima
za 7. i 8. mjesec**



Najveći izlaz nedjeljom.

Rijecni promet

```
par(mfrow = c(2, 2))
```

```
#Domaci ulaz turista po mjesecima
```

```
boxplot(domaci_ulaz$Riječni ~ domaci_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci ulaz")
means <- tapply(domaci_ulaz$Riječni, domaci_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

```
#Domaci izlaz turista po mjesecima
```

```
boxplot(domaci_izlaz$Riječni ~ domaci_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci izlaz")
means <- tapply(domaci_izlaz$Riječni, domaci_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

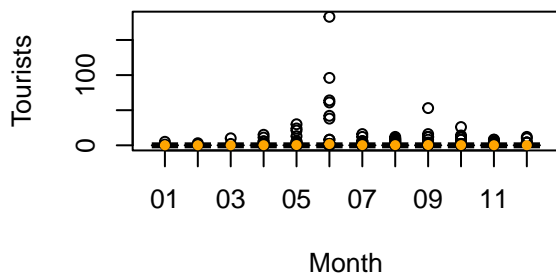
```
#Strani ulaz turista po mjesecima
```

```
boxplot(strani_ulaz$Riječni ~ strani_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani ulaz")
means <- tapply(strani_ulaz$Riječni, strani_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

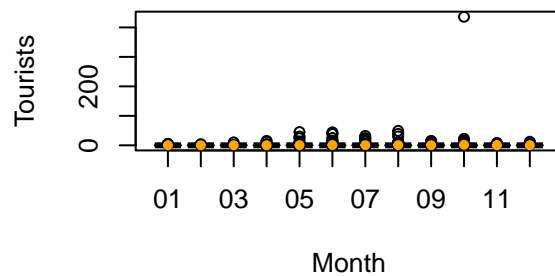
```
#Strani izlaz turista po mjesecima
```

```
boxplot(strani_izlaz$Riječni ~ strani_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani izlaz")
means <- tapply(strani_izlaz$Riječni, strani_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

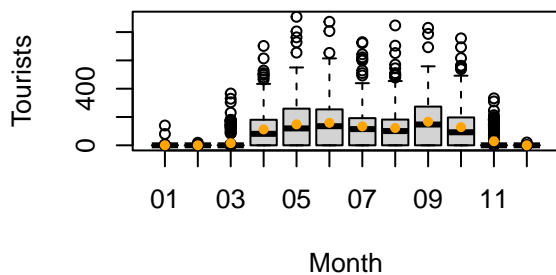
Domaci ulaz turista po mjesecima



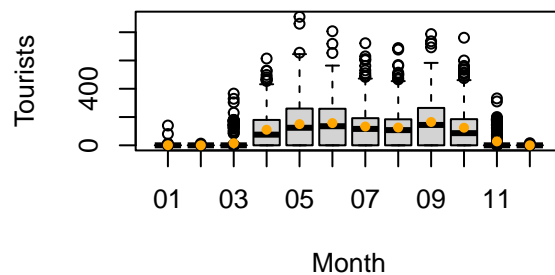
Domaci izlaz turista po mjesecima



Strani ulaz turista po mjesecima



Strani izlaz turista po mjesecima



Za domaci promet vidimo da je medijan blizu nule sto znaci da prometa vecinom nema, vidimo nesto vecu brojke od 5. do 8. mjeseca, za strani promet vidimo jednolik promet od 4. do 10. mjeseca.

Sezonu smo odredili od travnja do listopada.

Rijecni promet u sezoni

```
sezona <- c("04", "05", "06", "07", "08", "09", "10")

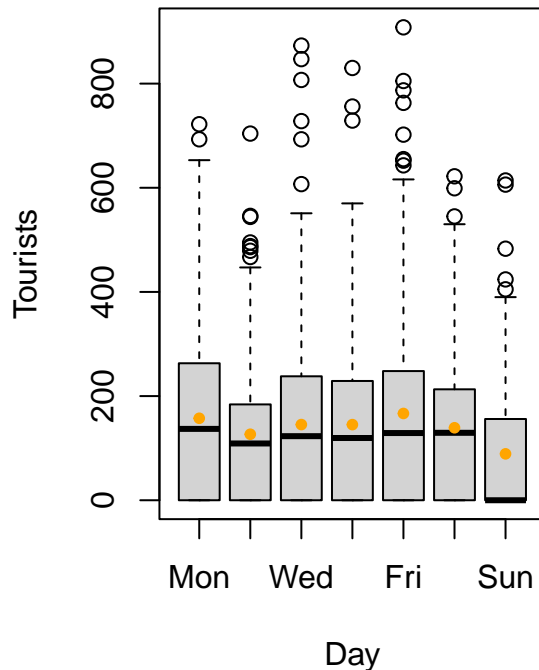
#strani sezona
strani_ulaz_sezona <- subset(strani_ulaz, strani_ulaz$Mjesec %in% sezona)
strani_izlaz_sezona <- subset(strani_izlaz, strani_izlaz$Mjesec %in% sezona)

par(mfrow = c(1, 2))

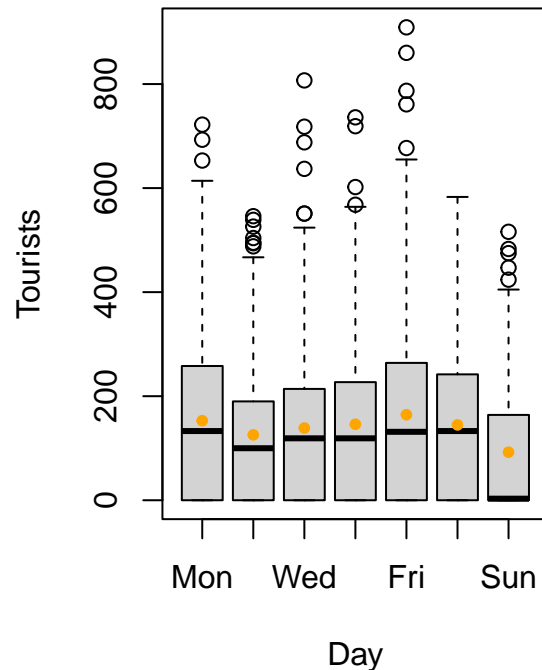
#Strani ulaz turista po danima u sezoni
boxplot(strani_ulaz_sezona$Riječni ~ strani_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani ulaz")
means <- tapply(strani_ulaz_sezona$Riječni, strani_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

#Strani izlaz turista po danima u sezoni
boxplot(strani_izlaz_sezona$Riječni ~ strani_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani izlaz")
means <- tapply(strani_izlaz_sezona$Riječni, strani_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")
```

Strani ulaz turista po danima u sezoni



Strani izlaz turista po danima u sezoni



nedjelju komentirat

Zanimljiv slucaj je nedjelja, tijekom koje vecinom nema prometa. Najvjerojatnije zbog toga sto su nedjeljom rijecni granicni prijelazi zatvoreni za promet.

Cestovni promet

```
par(mfrow = c(2, 2))
```

#Domaci ulaz turista po mjesecima

```
boxplot(domaci_ulaz$Cestovni ~ domaci_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci ulaz")
means <- tapply(domaci_ulaz$Cestovni, domaci_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

#Domaci izlaz turista po mjesecima

```
boxplot(domaci_izlaz$Cestovni ~ domaci_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci izlaz")
means <- tapply(domaci_izlaz$Cestovni, domaci_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

#Strani ulaz turista po mjesecima

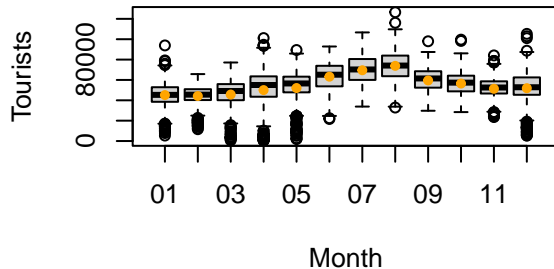
```
boxplot(strani_ulaz$Cestovni ~ strani_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani ulaz")
means <- tapply(strani_ulaz$Cestovni, strani_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

#Strani izlaz turista po mjesecima

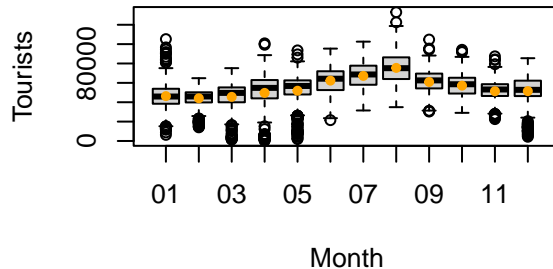
```
boxplot(strani_izlaz$Cestovni ~ strani_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani izlaz")
```

```
means <- tapply(strani_izlaz$Cestovni, strani_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

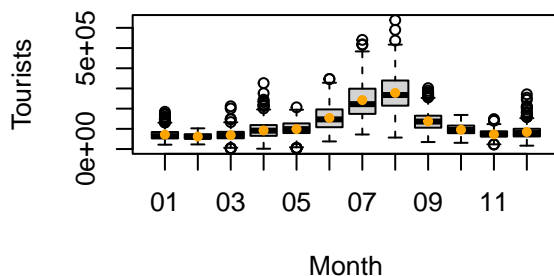
Domaci ulaz turista po mjesecima



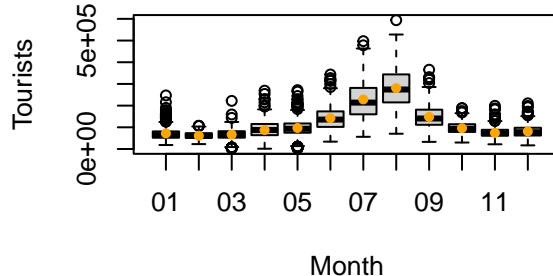
Domaci izlaz turista po mjesecima



Strani ulaz turista po mjesecima



Strani izlaz turista po mjesecima



Iako je i za domaci i za strani vrhunac ulaza i izlaza u 7. i 8. mjesecu, u stranom prometu ti mjeseci puno vise iskacu od ostalih, dok je u domacem prometu prijelaz postupan.

Cestovni promet u sezoni

```
sezona = c("07", "08")
```

```
domaci_ulaz_sezona <- subset(domaci_ulaz, domaci_ulaz$Mjesec %in% sezona)
domaci_izlaz_sezona <- subset(domaci_izlaz, domaci_izlaz$Mjesec %in% sezona)
```

```
strani_ulaz_sezona <- subset(strani_ulaz, strani_ulaz$Mjesec %in% sezona)
strani_izlaz_sezona <- subset(strani_izlaz, strani_izlaz$Mjesec %in% sezona)
```

```
par(mfrow = c(1, 2))
```

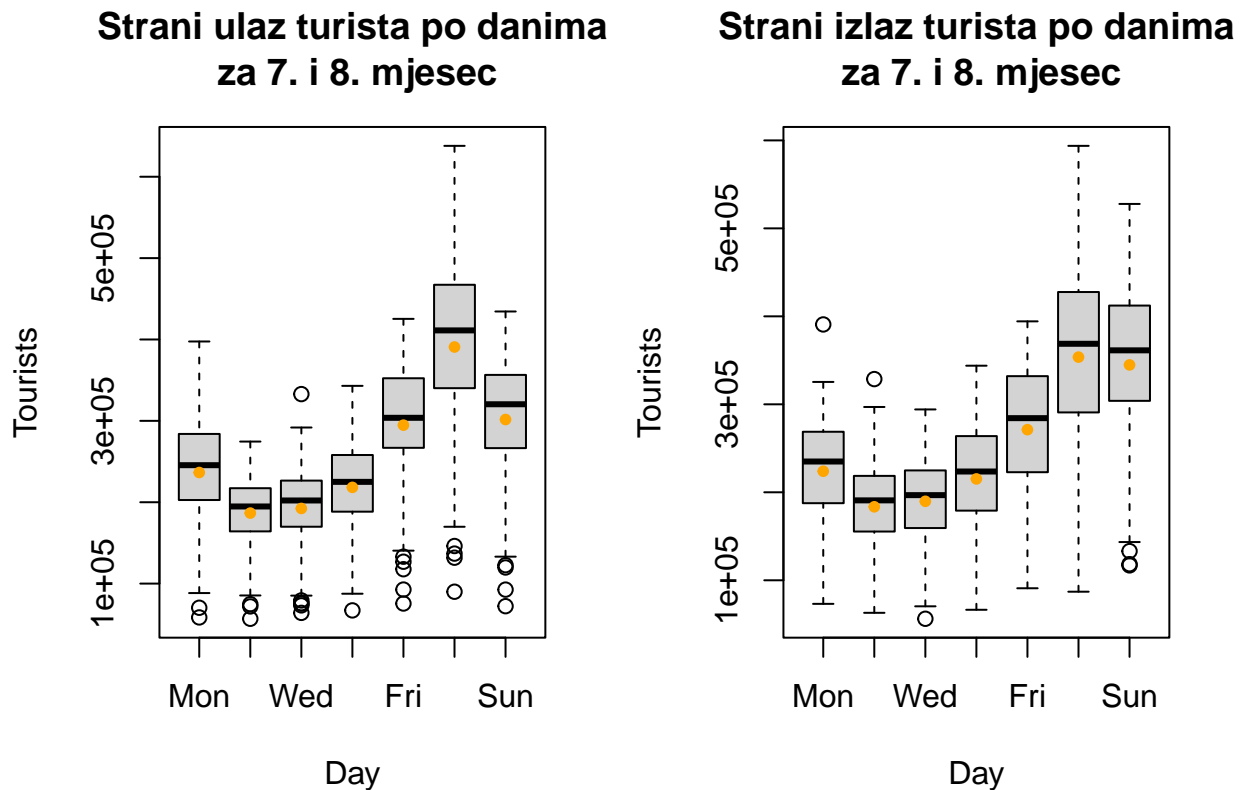
```
#Strani ulaz turista po danima za 7. i 8. mjesec
```

```
boxplot(strani_ulaz_sezona$Cestovni ~ strani_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani ulaz turista po danima za 7. i 8. mjesec")
means <- tapply(strani_ulaz_sezona$Cestovni, strani_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")
```

```
#Strani izlaz turista po danima za 7. i 8. mjesec
```

```
boxplot(strani_izlaz_sezona$Cestovni ~ strani_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani izlaz turista po danima za 7. i 8. mjesec")
means <- tapply(strani_izlaz_sezona$Cestovni, strani_izlaz_sezona$Dan, mean)
```

```
points(means, pch = 20, col = "orange")
```



Tijekom sezone strani ulaz ima izrazeni vrhunac subotom, a izlaz vikendom. Također možemo uočiti trend rasta od utorka prema vikendu.

Pomorski promet

```
par(mfrow = c(2, 2))
```

```
#Domaci ulaz turista po mjesecima
```

```
boxplot(domaci_ulaz$Pomorski ~ domaci_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci ulaz
```

```
means <- tapply(domaci_ulaz$Pomorski, domaci_ulaz$Mjesec, mean)
```

```
points(means, pch = 20, col = "orange")
```

```
#Domaci izlaz turista po mjesecima
```

```
boxplot(domaci_izlaz$Pomorski ~ domaci_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci iz
```

```
means <- tapply(domaci_izlaz$Pomorski, domaci_izlaz$Mjesec, mean)
```

```
points(means, pch = 20, col = "orange")
```

```
#Strani ulaz turista po mjesecima
```

```
boxplot(strani_ulaz$Pomorski ~ strani_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani ulaz
```

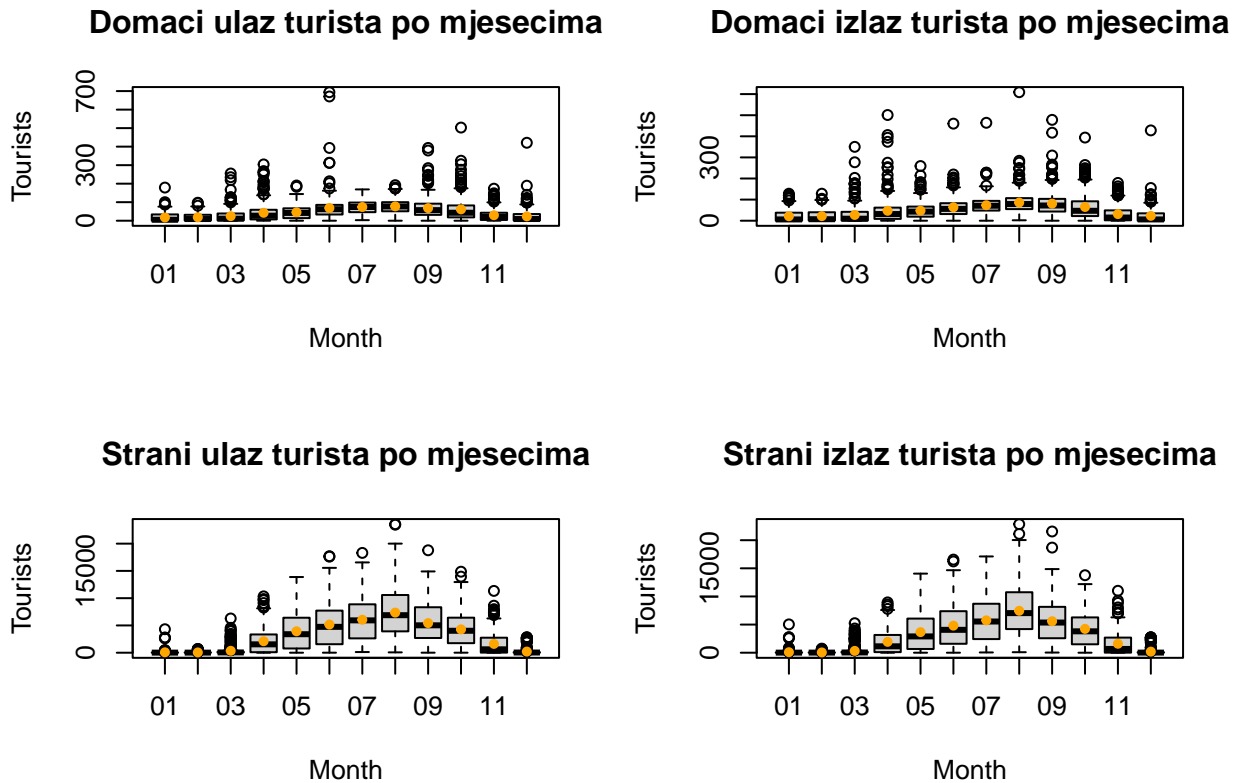
```
means <- tapply(strani_ulaz$Pomorski, strani_ulaz$Mjesec, mean)
```

```
points(means, pch = 20, col = "orange")
```

```
#Strani izlaz turista po mjesecima
```

```
boxplot(strani_izlaz$Pomorski ~ strani_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani iz
```

```
means <- tapply(strani_izlaz$Pomorski, strani_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```



Kod domaceg i stranog prometa vidimo normalnu razdiobu s vrhuncem u 8. mjesecu.

Pomorski promet u sezoni

```
sezona = c("05", "06", "07", "08", "09", "10")

domaci_ulaz_sezona <- subset(domaci_ulaz, domaci_ulaz$Mjesec %in% sezona)
domaci_izlaz_sezona <- subset(domaci_izlaz, domaci_izlaz$Mjesec %in% sezona)

par(mfrow = c(2, 2))

#Domaci ulaz turista po danima za 7. i 8. mjesec
boxplot(domaci_ulaz_sezona$Pomorski ~ domaci_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Domaci ulaz")
means <- tapply(domaci_ulaz_sezona$Pomorski, domaci_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

#Domaci izlaz turista po danima za 7. i 8. mjesec
boxplot(domaci_izlaz_sezona$Pomorski ~ domaci_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Domaci izlaz")
means <- tapply(domaci_izlaz_sezona$Pomorski, domaci_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

strani_ulaz_sezona <- subset(strani_ulaz, strani_ulaz$Mjesec %in% sezona)
```

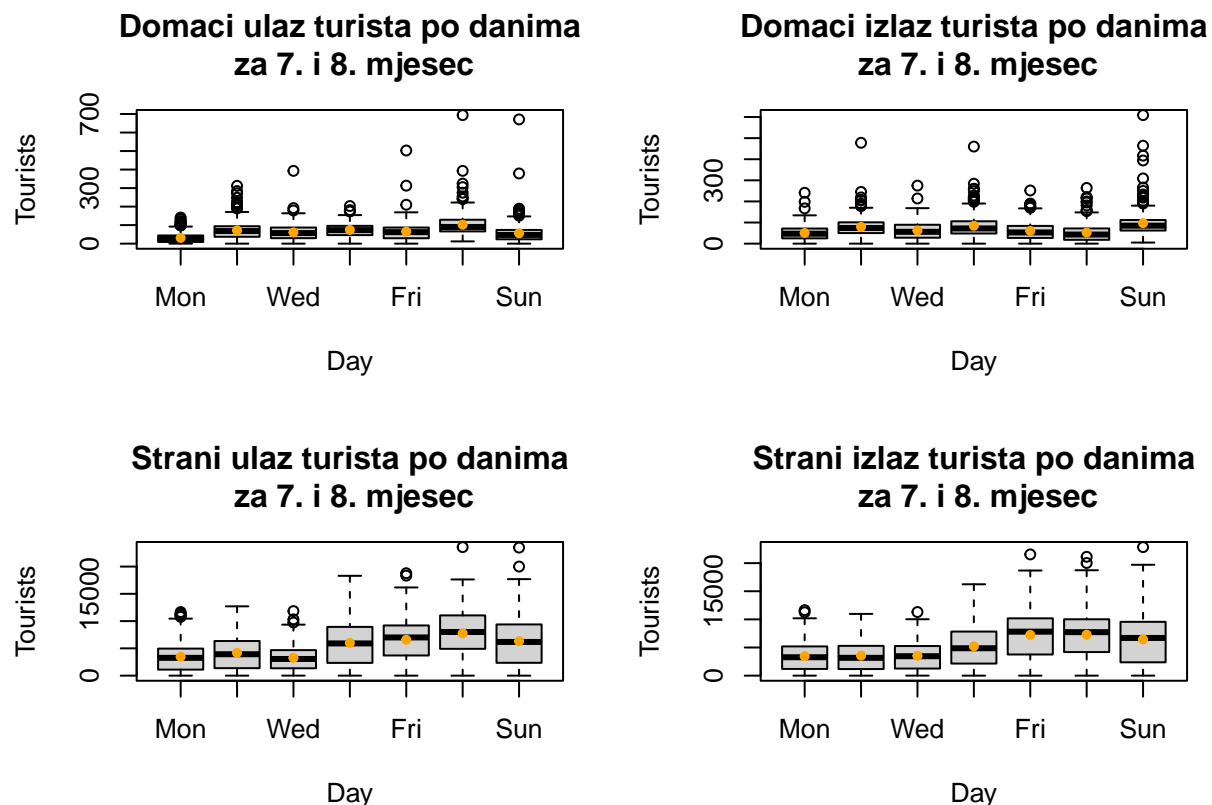
```
strani_izlaz_sezona <- subset(strani_izlaz, strani_izlaz$Mjesec %in% sezona)
```

```
#Strani ulaz turista po danima za 7. i 8. mjesec
```

```
boxplot(strani_ulaz_sezona$Pomorski ~ strani_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani ulaz turista po danima za 7. i 8. mjesec")
means <- tapply(strani_ulaz_sezona$Pomorski, strani_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")
```

```
#Strani izlaz turista po danima za 7. i 8. mjesec
```

```
boxplot(strani_izlaz_sezona$Pomorski ~ strani_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani izlaz turista po danima za 7. i 8. mjesec")
means <- tapply(strani_izlaz_sezona$Pomorski, strani_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")
```



Za domaci ulaz najmanji ponedjeljak i nedjelja, a izlaz najveći nedjeljom. Strani promet je nizak početkom tjedna, znatno se povećava od četvrtka i ostaje jednolik do kraja tjedna.

Zracni promet

```
par(mfrow = c(2, 2))
```

```
#Domaci ulaz turista po mjesecima
```

```
boxplot(domaci_ulaz$Zračni ~ domaci_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci ulaz turista po mjesecima")
means <- tapply(domaci_ulaz$Zračni, domaci_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

```
#Domaci izlaz turista po mjesecima
```

```
boxplot(domaci_izlaz$Zračni ~ domaci_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Domaci izlaz turista po mjesecima")
means <- tapply(domaci_izlaz$Zračni, domaci_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")
```

```

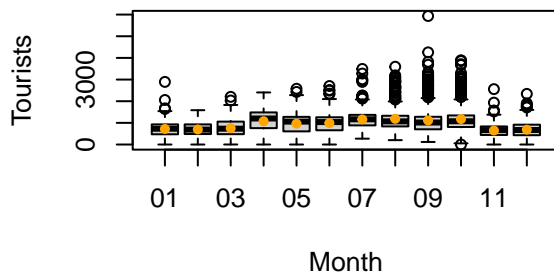
means <- tapply(domaci_izlaz$Zračni, domaci_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")

#Strani ulaz turista po mjesecima
boxplot(strani_ulaz$Zračni ~ strani_ulaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani ulaz t
means <- tapply(strani_ulaz$Zračni, strani_ulaz$Mjesec, mean)
points(means, pch = 20, col = "orange")

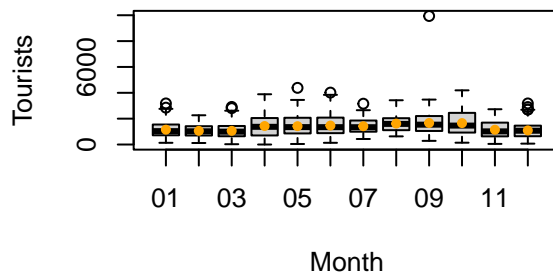
#Strani izlaz turista po mjesecima
boxplot(strani_izlaz$Zračni ~ strani_izlaz$Mjesec, xlab = "Month", ylab = "Tourists", main="Strani izlaz
means <- tapply(strani_izlaz$Zračni, strani_izlaz$Mjesec, mean)
points(means, pch = 20, col = "orange")

```

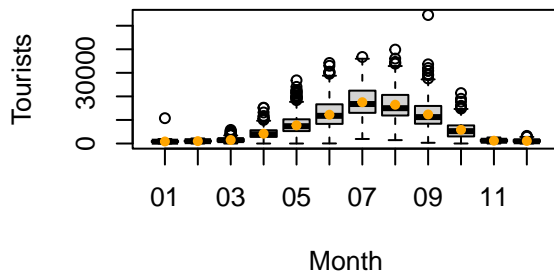
Domaci ulaz turista po mjesecima



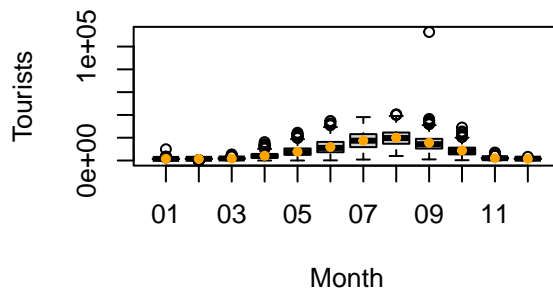
Domaci izlaz turista po mjesecima



Strani ulaz turista po mjesecima



Strani izlaz turista po mjesecima



Zračni promet je pojačan od 4. do 10. mjeseca pa uzimamo taj period za sezonu. Primijetimo da u tom periodu domaci ulaz/izlaz prati uniformnu distribuciju, a strani ulaz/izlaz normalnu distribuciju

Zračni promet po sezoni

```

sezona = c("04", "05", "06", "07", "08", "09", "10")

domaci_ulaz_sezona <- subset(domaci_ulaz, domaci_ulaz$Mjesec %in% sezona)
domaci_izlaz_sezona <- subset(domaci_izlaz, domaci_izlaz$Mjesec %in% sezona)

par(mfrow = c(1, 2))

#Domaci ulaz turista po danima u sezoni

```

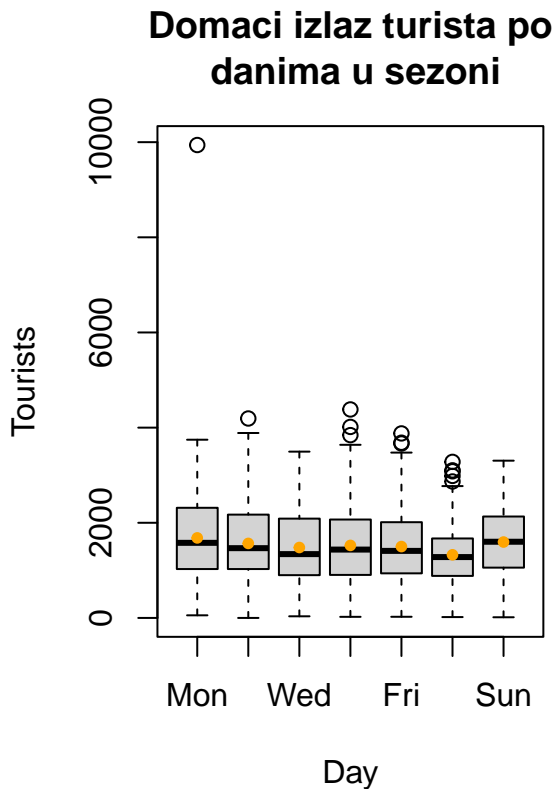
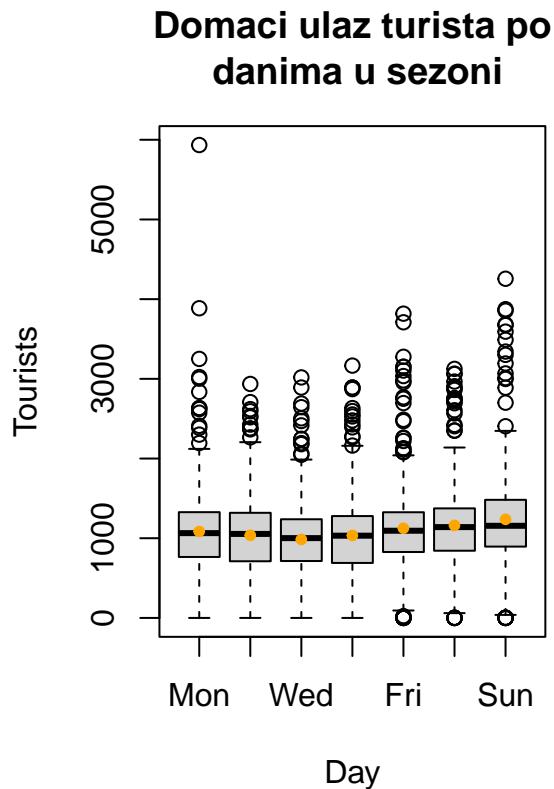


```

boxplot(domaci_ulaz_sezona$Zračni ~ domaci_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Domaci ulaz turista po danima u sezoni")
means <- tapply(domaci_ulaz_sezona$Zračni, domaci_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

#Domaci izlaz turista po danima u sezoni
boxplot(domaci_izlaz_sezona$Zračni ~ domaci_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Domaci izlaz turista po danima u sezoni")
means <- tapply(domaci_izlaz_sezona$Zračni, domaci_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

```



```

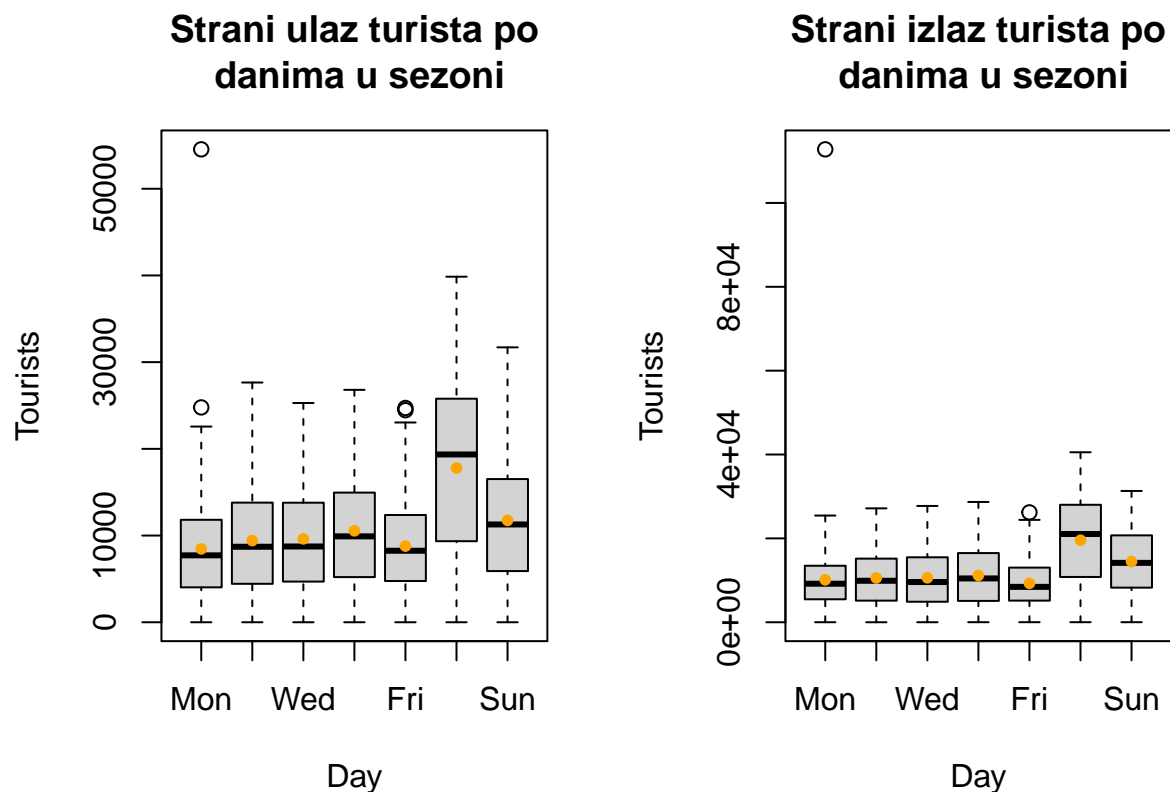
strani_ulaz_sezona <- subset(strani_ulaz, strani_ulaz$Mjesec %in% sezona)
strani_izlaz_sezona <- subset(strani_izlaz, strani_izlaz$Mjesec %in% sezona)

par(mfrow = c(1, 2))

#Strani ulaz turista po danima u sezoni
boxplot(strani_ulaz_sezona$Zračni ~ strani_ulaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani ulaz turista po danima u sezoni")
means <- tapply(strani_ulaz_sezona$Zračni, strani_ulaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

#Strani izlaz turista po danima u sezoni
boxplot(strani_izlaz_sezona$Zračni ~ strani_izlaz_sezona$Dan, xlab = "Day", ylab = "Tourists", main="Strani izlaz turista po danima u sezoni")
means <- tapply(strani_izlaz_sezona$Zračni, strani_izlaz_sezona$Dan, mean)
points(means, pch = 20, col = "orange")

```



U zračnom prometu nema puno varijacije po danima u tjednu, izuzev stranih ulazaka i izlazaka kojih ima najviše subotom.

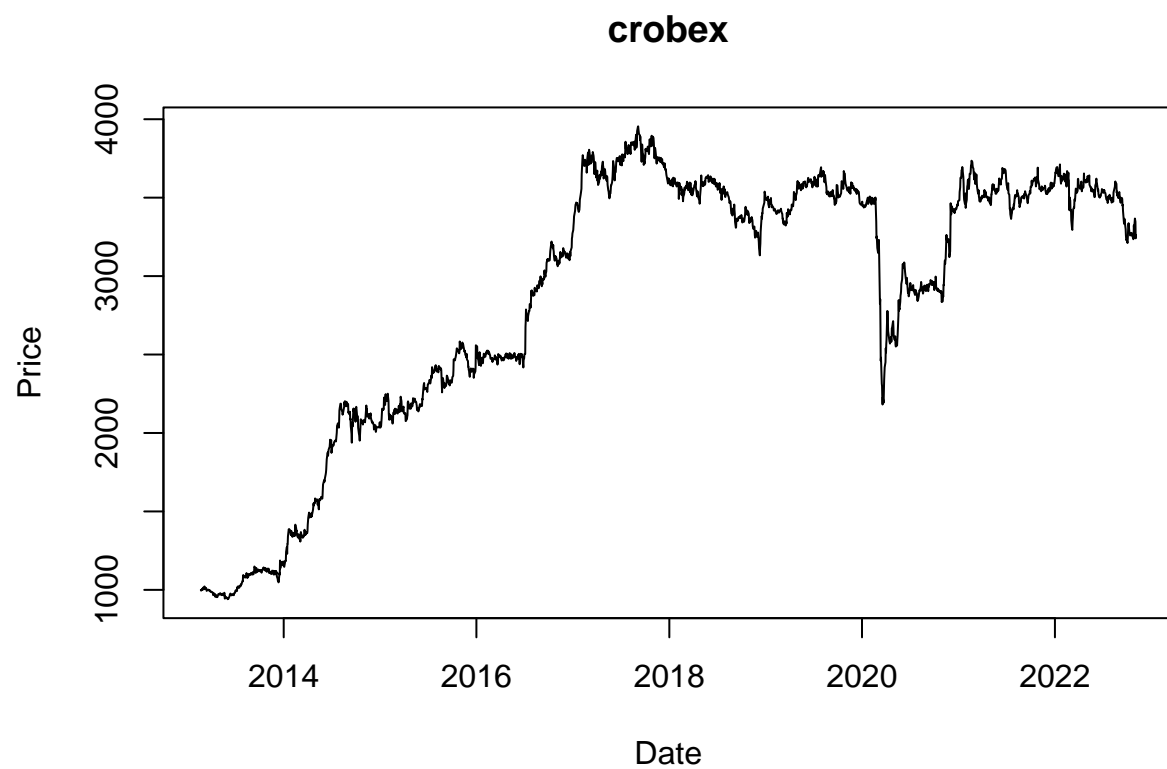
Analiza fundamentalnih podataka turističkih tvrtki

Analizirajte fundamentalne podatke turističkih tvrtki - postoje li sezonalnosti i u njima?

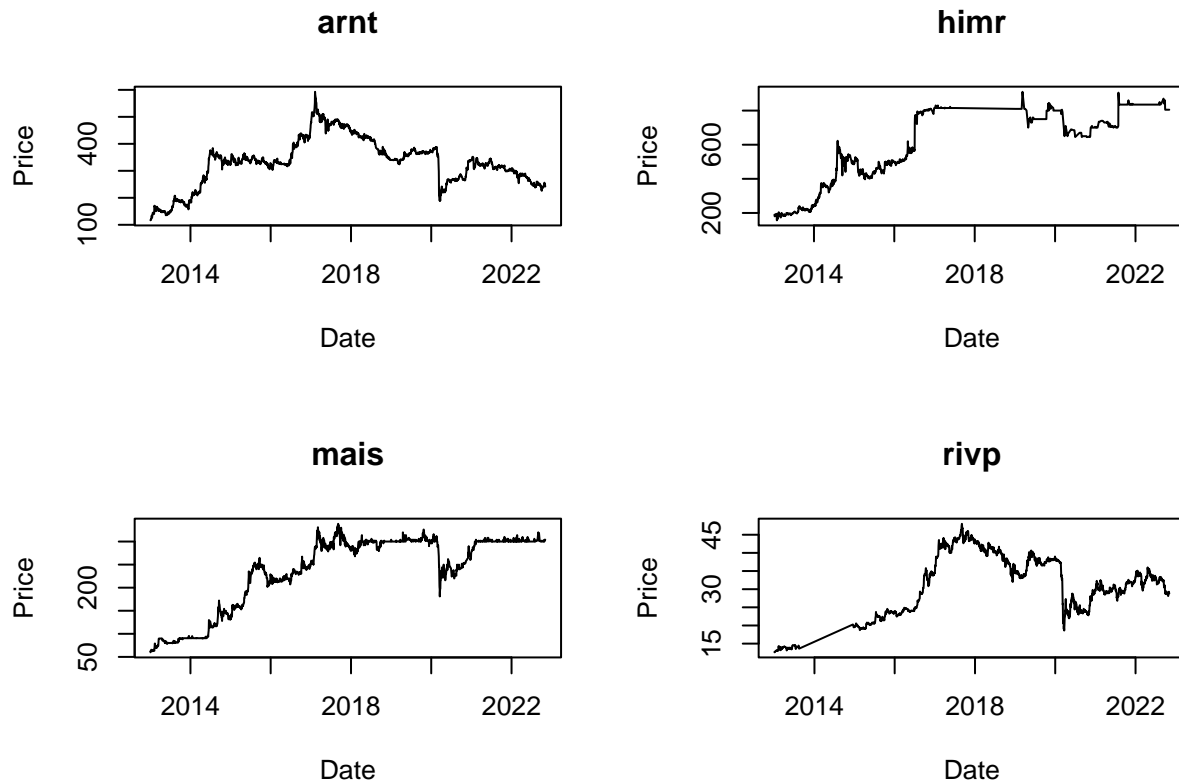
Vizualizacija podataka

Plotamo crobex indeks i cijene dionica za arnt, himr, mais i rivp. Jedino na himr plotu ne vidimo jasno pad 2020. za vrijeme covid krize. Primjećujemo vizualno da se himr razlikuje u ponašanju od ostale 3 dionice.

```
plot(crobex, type = "l", main = "crobex")
```



```
layout(matrix(c(1, 2, 3, 4), nrow = 2, ncol = 2, byrow = TRUE))
plot(arnt, type = "l", main = "arnt")
plot(himr, type = "l", main = "himr")
plot(mais, type = "l", main = "mais")
plot(rivp, type = "l", main = "rivp")
```



Pogledajmo barplot prodaje i profita za ARNT od 2013. do 2022. godine

```
library(ggplot2)
library(reshape2)

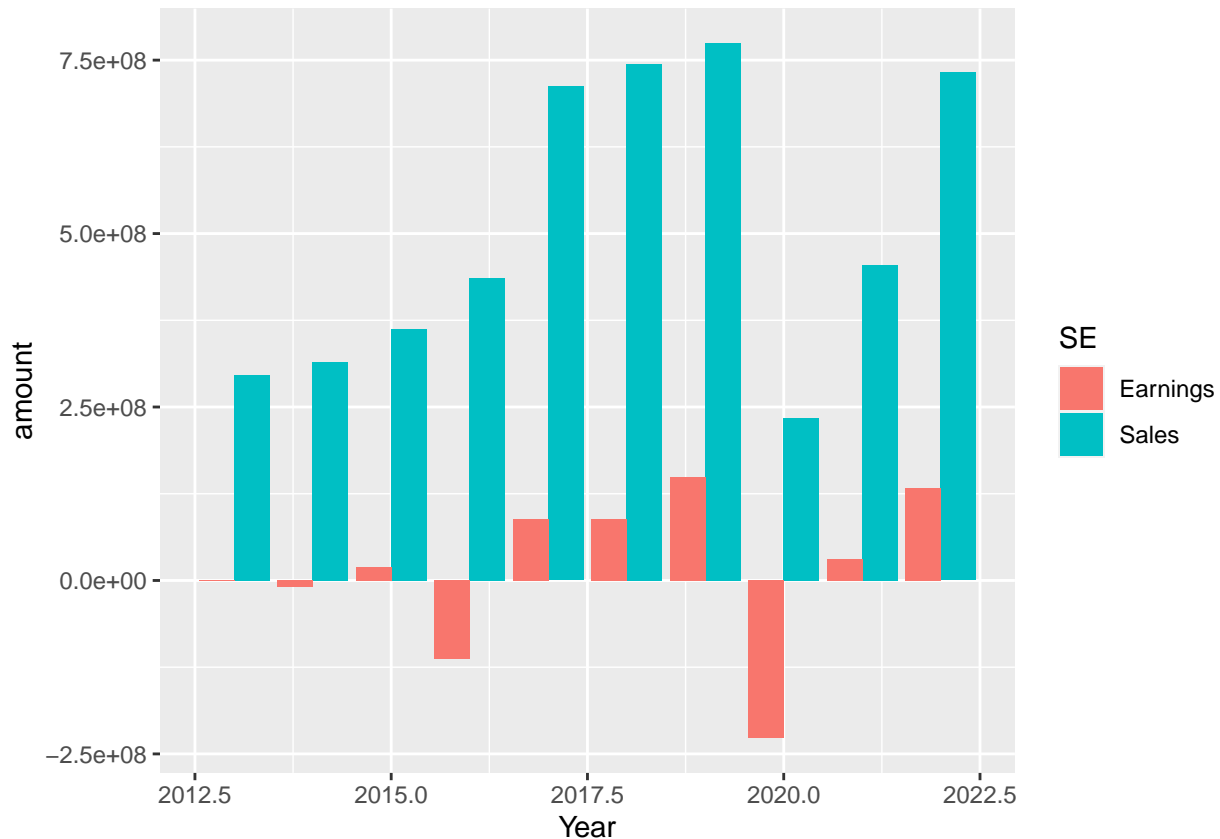
df_sum <- arnt_fund %>% group_by(Year) %>% summarize(Sales = sum(Sales), Earnings = sum(Earnings))

df <- as.data.frame(matrix(nrow = 2 * nrow(df_sum), ncol=3))
colnames(df) <- c("Year", "SE", "amount")

for(i in 1:nrow(df_sum)){
  df$Year[2 * (i - 1) + 1] <- df_sum$Year[i]
  df$SE[2 * (i - 1) + 1] <- "Sales"
  df$amount[2 * (i - 1) + 1] <- df_sum$Sales[i]

  df$Year[2 * i] <- df_sum$Year[i]
  df$SE[2 * i] <- "Earnings"
  df$amount[2 * i] <- df_sum$Earnings[i]
}

#crtamo barplot
ggplot(df, aes(fill=SE, y=amount, x=Year)) +
  geom_bar(position='dodge', stat='identity')
```



Vidimo trend rasta prodaje, medutim pad 2020. zbog covida, a 2021. i 2022. oporavak. Kompanija nije bila znacajno profitabilna do 2017. godine.

Pogledajmo vizualno s boxplotovima prodaje u kvartalima.

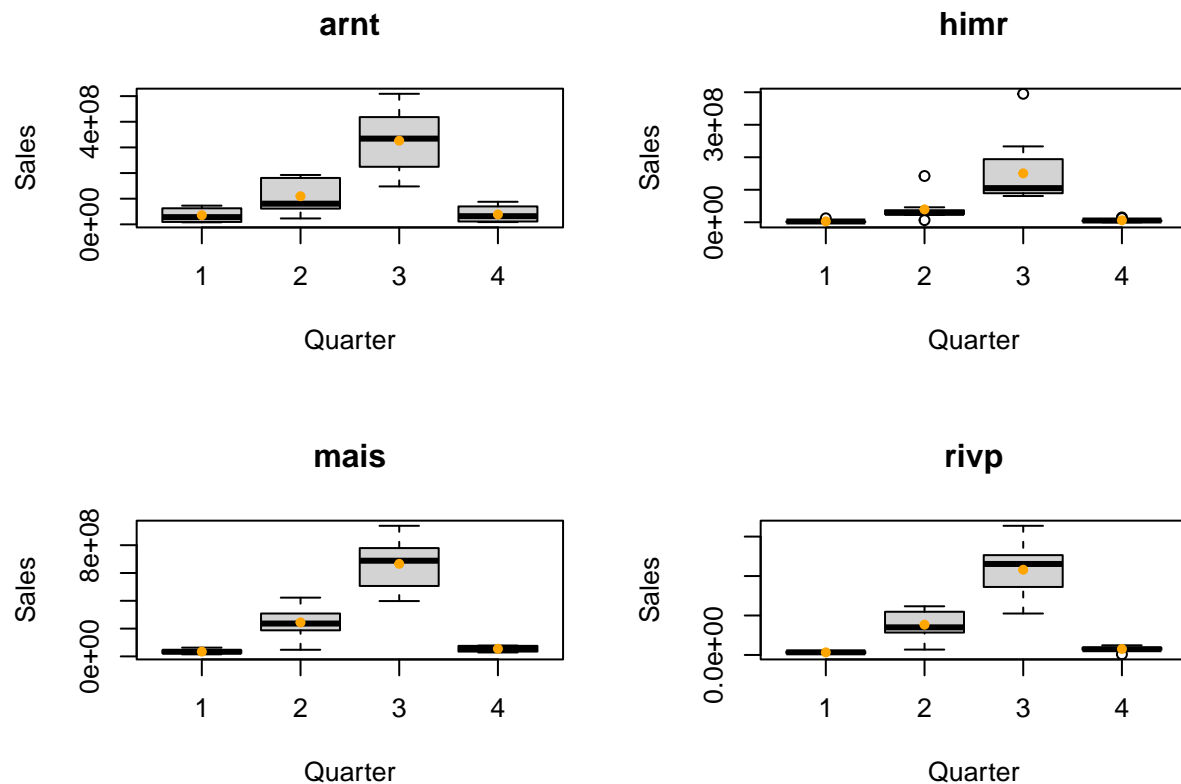
```
par(mfrow = c(2, 2))

boxplot(arnt_fund$Sales ~ arnt_fund$Quarter, xlab = "Quarter", ylab = "Sales", main = "arnt")
means <- tapply(arnt_fund$Sales, arnt_fund$Quarter, mean)
points(means, pch = 20, col = "orange")

boxplot(himr_fund$Sales ~ himr_fund$Quarter, xlab = "Quarter", ylab = "Sales", main = "himr")
means <- tapply(himr_fund$Sales, himr_fund$Quarter, mean)
points(means, pch = 20, col = "orange")

boxplot(mais_fund$Sales ~ mais_fund$Quarter, xlab = "Quarter", ylab = "Sales", main = "mais")
means <- tapply(mais_fund$Sales, mais_fund$Quarter, mean)
points(means, pch = 20, col = "orange")

boxplot(rivp_fund$Sales ~ rivp_fund$Quarter, xlab = "Quarter", ylab = "Sales", main = "rivp")
means <- tapply(rivp_fund$Sales, rivp_fund$Quarter, mean)
points(means, pch = 20, col = "orange")
```



Ocekivano prodaje su najveće u trećem kvartalu koji odgovara periodu sezone.

Analiza razlikovanja kvartala prema ostvarenim prihodima od prodaje

Kao istraživačko pitanje smo testirali postoji li razlika između kvartala prema ostvarenim prihodima prodaje, kao podatke koristimo prošireni skup podataka za poduzeće Arena Hospitality Group d.d. koji uključuje podatke od 4Q 2002.

Vizualizacija prihoda po kvartalima za ARNT:

```
summary(arnt_sales_ext$Sales)
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.
## 6536000 10318172 43776719 87334982 118178500 508990943
```

```
summary(arnt_sales_ext$Sales[arnt_sales_ext$Quarter==1])
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.
## 7013443 8019191 9952000 22363812 20943925 72847427
```

```
summary(arnt_sales_ext$Sales[arnt_sales_ext$Quarter==2])
```

```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.
## 22931057 34748869 50636942 74762652 79383600 192415297
```

```
summary(arnt_sales_ext$Sales[arnt_sales_ext$Quarter==3])
```

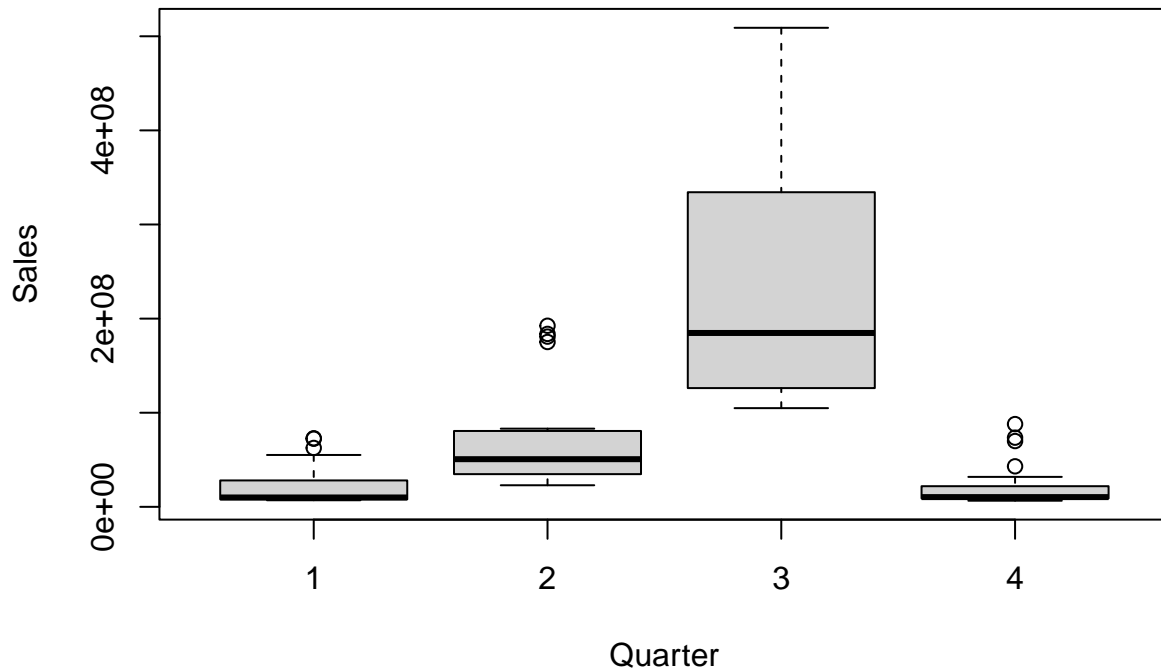
```
##      Min.   1st Qu.   Median     Mean  3rd Qu.    Max.
## 104887000 126723039 184754074 229697401 323177521 508990943
```

```
summary(arnt_sales_ext$Sales[arnt_sales_ext$Quarter==4])
```

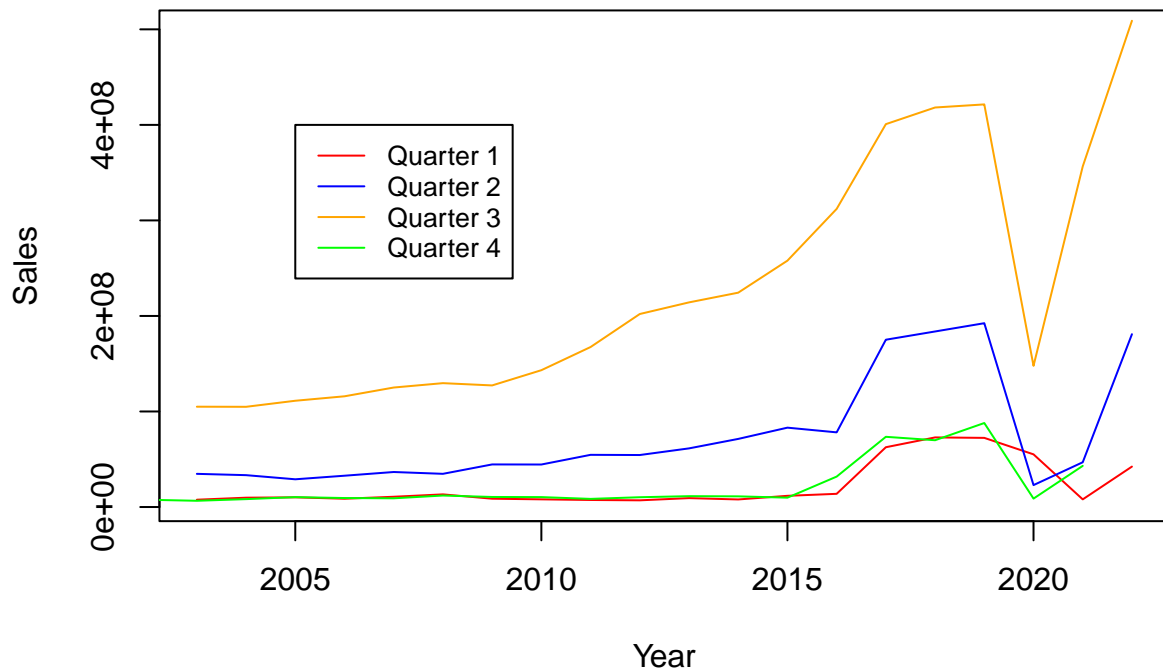
```
##      Min.   1st Qu.   Median     Mean 3rd Qu.     Max.
## 6536000  9056528 10360793 22516062 16984682 87929736
```

```
boxplot(arnt_sales_ext$Sales ~ arnt_sales_ext$Quarter, xlab="Quarter", ylab="Sales", main="Prihodi prema kvartalima")
```

Prihodi prema kvartalima



```
plot(arnt_sales_ext$Year[arnt_sales_ext$Quarter==1], arnt_sales_ext$Sales[arnt_sales_ext$Quarter==1], xlab="Year", ylab="Sales", main="Prihodi prema godinama", col="red", lty="n")
lines(arnt_sales_ext$Year[arnt_sales_ext$Quarter==2], arnt_sales_ext$Sales[arnt_sales_ext$Quarter==2], col="blue", lty="n")
lines(arnt_sales_ext$Year[arnt_sales_ext$Quarter==3], arnt_sales_ext$Sales[arnt_sales_ext$Quarter==3], col="green", lty="n")
lines(arnt_sales_ext$Year[arnt_sales_ext$Quarter==4], arnt_sales_ext$Sales[arnt_sales_ext$Quarter==4], col="black", lty="n")
legend(2005, 4e+08, legend = c("Quarter 1", "Quarter 2", "Quarter 3", "Quarter 4"), col = c("red", "blue", "green", "black"), bty="n")
```



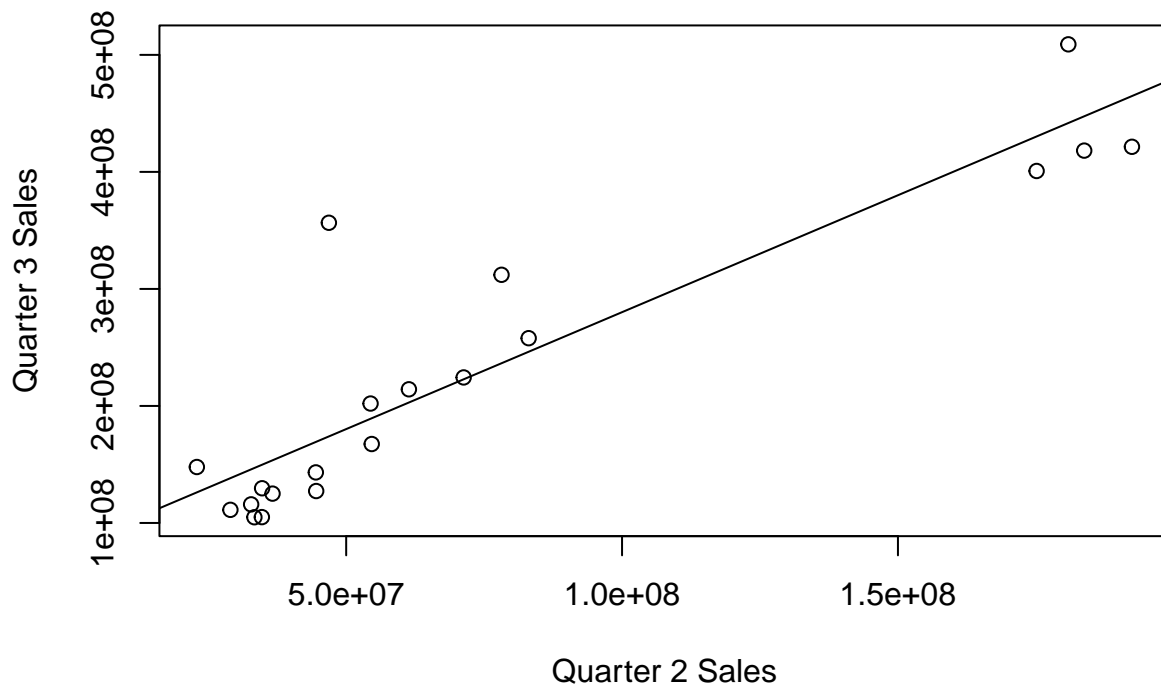
Uocava se ocekivano najveći prihodi u 3. kvartalu, zatim 2., a 1. i 4. su otprilike jednaki na dnu. Prihodi u 3. kvartalu jednoliko rastu kroz godine, a u ostalima su držali prihode na otprilike istoj razini do 2017. godine kada se dogodio nagli skok. Također se jasno vidi period covid pandemije 2020. godine, kada su prihodi naglo pali.

ANOVA

Nezavisnost

Testiramo postoji li veza između prodaje u drugom i trećem kvartalu.

```
model <- lm(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3] ~ arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2])
plot(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2], arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3],
     abline(model))
```

```
summary(model)
```

```
##
## Call:
## lm(formula = arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3] ~
##     arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2])
##
## Residuals:
```

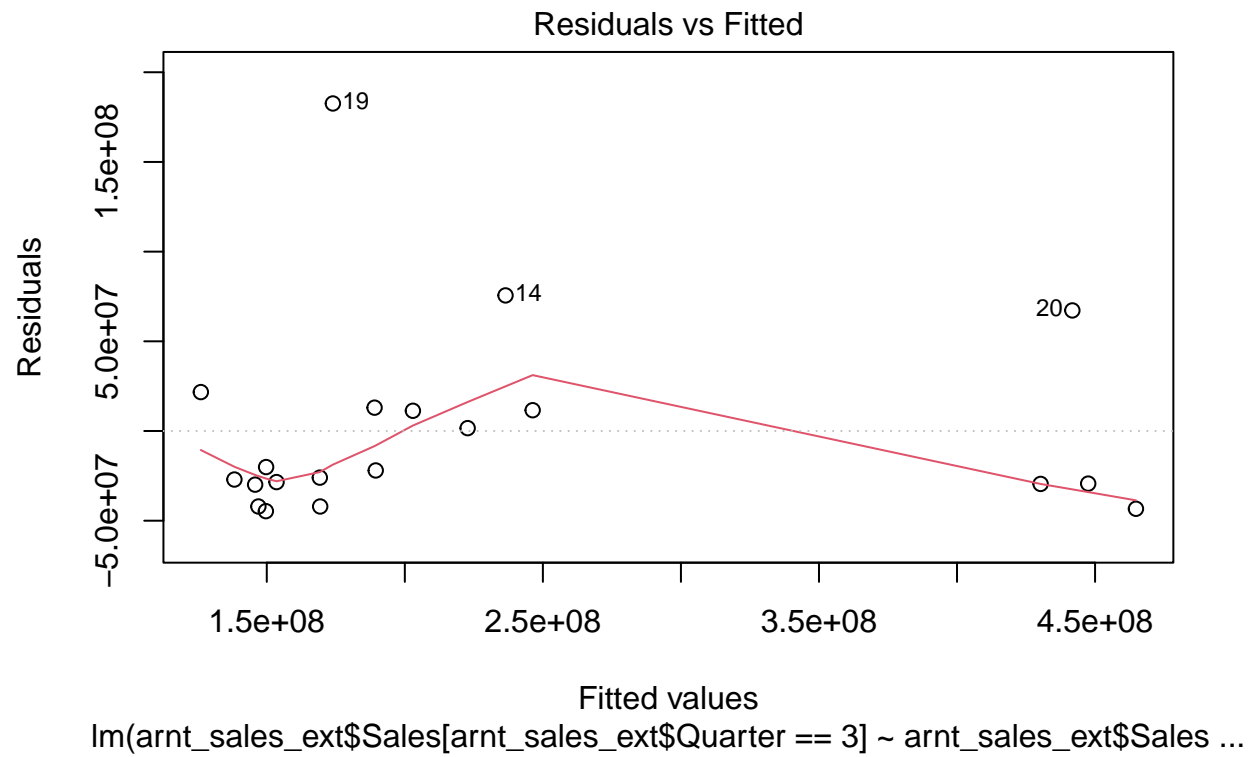
	Min	1Q	Median	3Q	Max
	-44676642	-29603367	-23998413	11945280	182589629

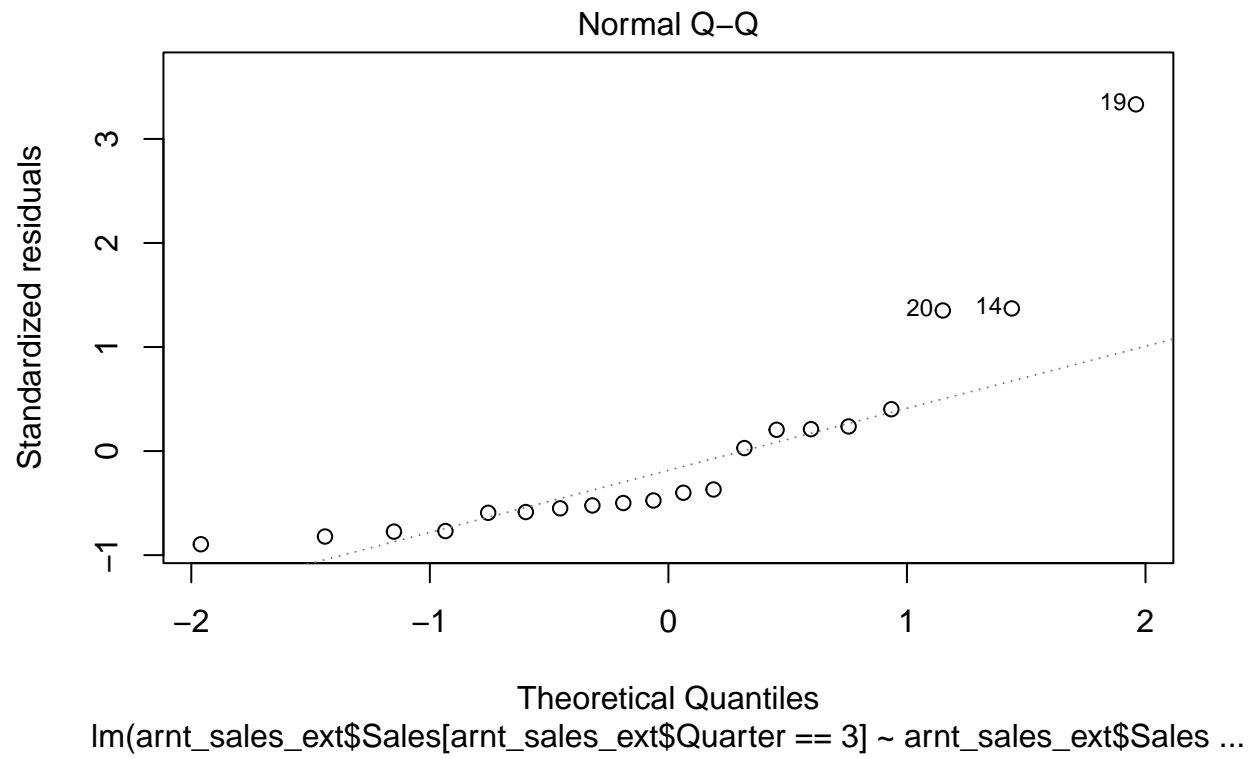
```
##
## Coefficients:
```

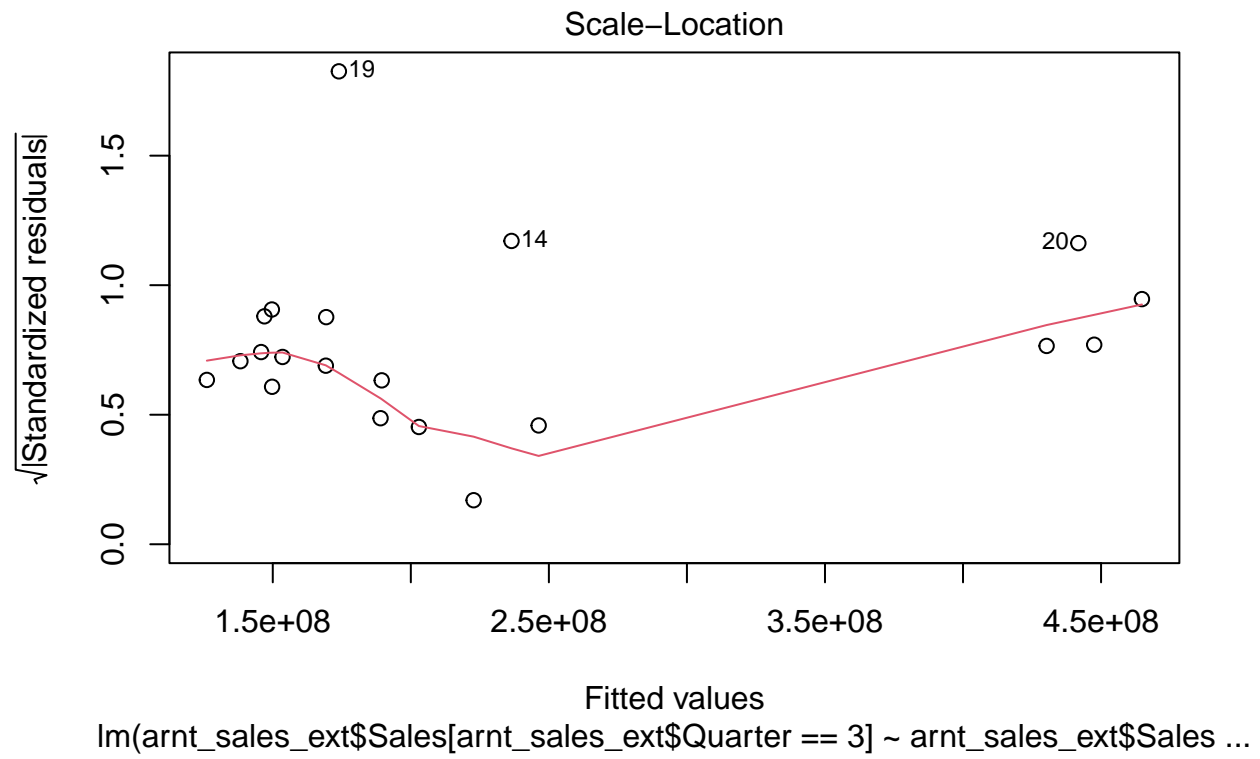
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	8.030e+07	2.100e+07	3.823	0.00125 **
arnt_sales_ext\$Sales[arnt_sales_ext\$Quarter == 2]	1.998e+00	2.242e-01	8.912	5.1e-08 ***

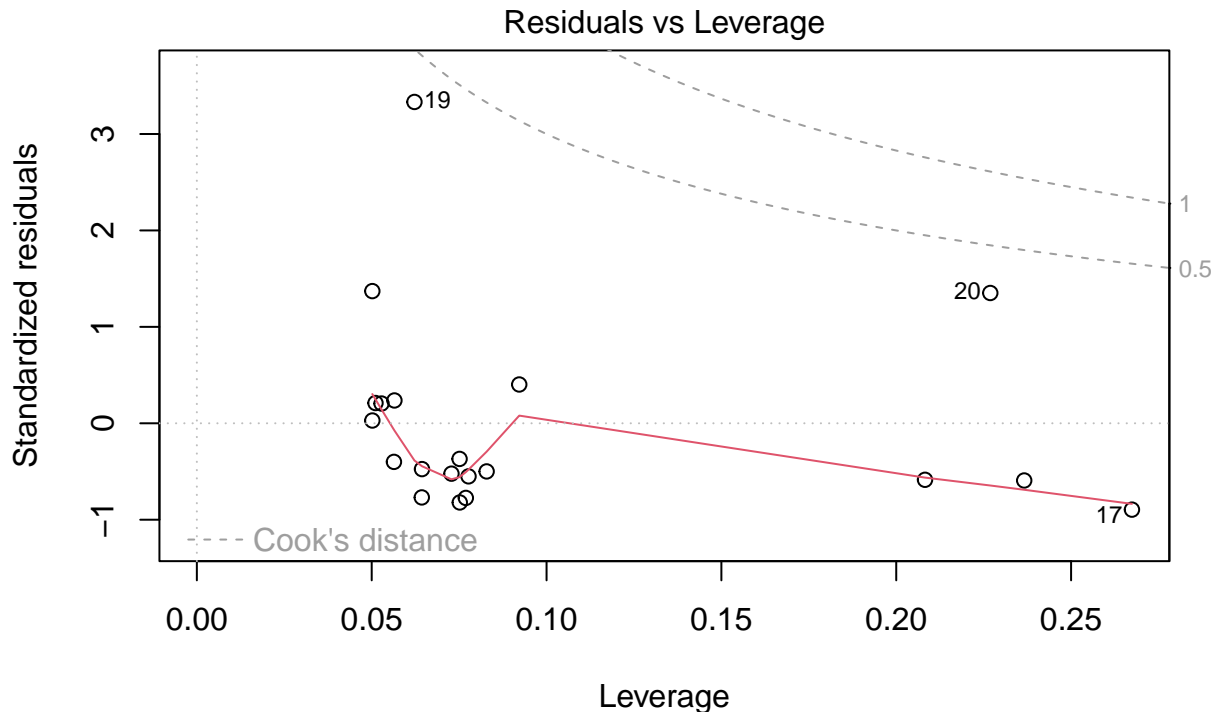
```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 56580000 on 18 degrees of freedom
## Multiple R-squared:  0.8152, Adjusted R-squared:  0.805
## F-statistic: 79.42 on 1 and 18 DF,  p-value: 5.097e-08
```

```
plot(model)
```









`lm(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3] ~ arnt_sales_ext$Sales ...`

Vidimo korelaciju između prodaje u drugom i trećem kvartalu, tj. nemamo nezavisnost za ANOVU.

Koristit ćemo logaritam za transformaciju podataka

Normalnost

Provodimo test normalnosti podataka.

```
require(nortest)
```

```
## Loading required package: nortest
```

```
lillie.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 1]))
```

```
##
```

```
## Lilliefors (Kolmogorov-Smirnov) normality test
```

```
##
```

```
## data: log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 1])
```

```
## D = 0.28112, p-value = 0.0002192
```

```
lillie.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]))
```

```
##
```

```
## Lilliefors (Kolmogorov-Smirnov) normality test
```

```
##
```

```
## data: log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2])
```

```
## D = 0.14832, p-value = 0.2965
```

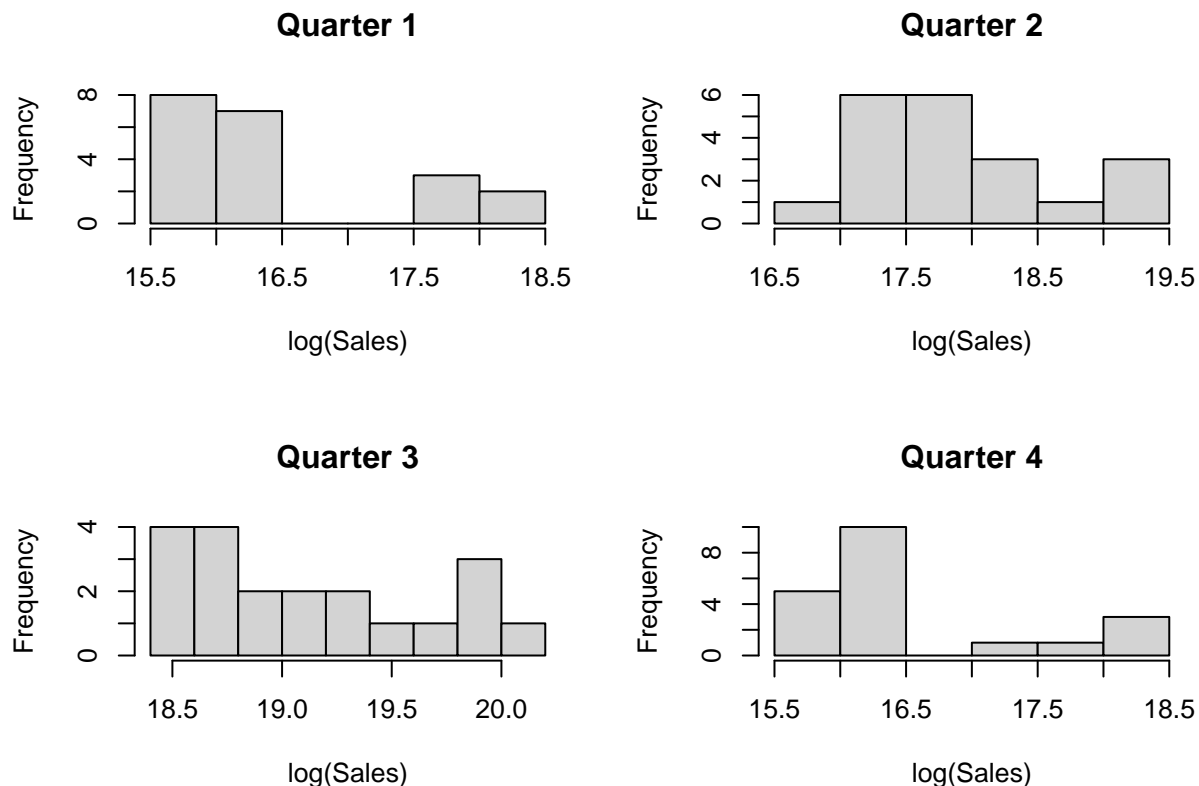
```
lillie.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]))
```

```
##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data: log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3])
## D = 0.16462, p-value = 0.1664
```

```
lillie.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 4]))
```

```
##
## Lilliefors (Kolmogorov-Smirnov) normality test
##
## data: log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 4])
## D = 0.35043, p-value = 5.427e-07
```

```
par(mfrow = c(2, 2))
hist(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 1]), xlab = "log(Sales)", main = "Quarter 1")
hist(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]), xlab = "log(Sales)", main = "Quarter 2")
hist(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]), xlab = "log(Sales)", main = "Quarter 3")
hist(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 4]), xlab = "log(Sales)", main = "Quarter 4")
```

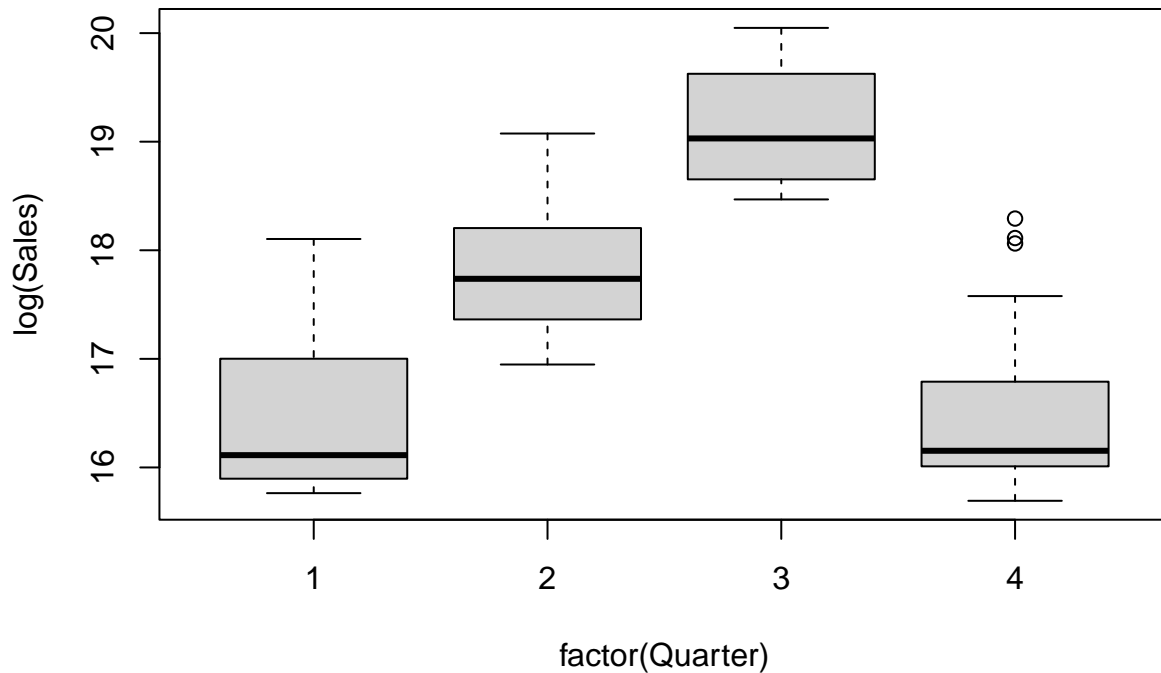


```
anova <- aov(Sales ~ factor(Quarter), data = arnt_sales_ext)
```

Nemamo dovoljno malu p-vrijednost da bi negirali normalnost podataka za drugi i treci kvartal, medutim u 1. i 4. kvartalu odbacujemo normalnost.

Jednakost varijanci

```
boxplot(log(Sales) ~ factor(Quarter), data = arnt_sales_ext)
```



```
bartlett.test(log(arnt_sales_ext$Sales) ~ arnt_sales_ext$Quarter)
```

```
##  
## Bartlett test of homogeneity of variances  
##  
## data: log(arnt_sales_ext$Sales) by arnt_sales_ext$Quarter  
## Bartlett's K-squared = 5.0608, df = 3, p-value = 0.1674
```

```
var(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 1]))
```

```
## [1] 0.7302634
```

```
var(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]))
```

```
## [1] 0.4441857
```

```
var(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]))
```

```
## [1] 0.2837433
```

```
var(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 4]))
```

```
## [1] 0.6954283
```

Nemamo dovoljno malu p-vrijednost da bi odbacili jednakost varijanci.

Provodimo ANOVA-u

```
anova <- aov(log(Sales) ~ factor(Quarter), data = arnt_sales_ext)
summary(anova)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## factor(Quarter) 3  94.25   31.417   58.35 <2e-16 ***
## Residuals      76  40.92    0.538
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

p-value vrlo mali, na visokoj razini znacajnosti odbacujemo hipotezu da su sve ocekivane vrijednosti jednake.

Post-hoc test

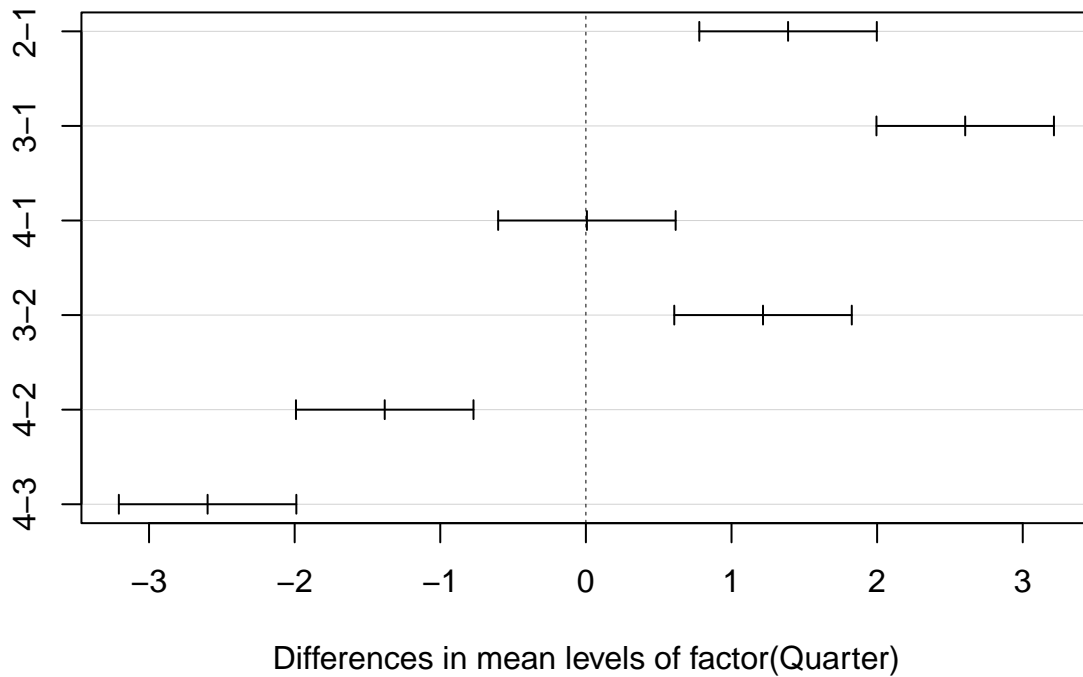
Uspoređujemo u parovima kvartale, vidimo da se treci znacajno razlikuje od svih ostalih.

```
post_test <- TukeyHSD(anova, conf.level=.95)
print(post_test)
```

```
##    Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = log(Sales) ~ factor(Quarter), data = arnt_sales_ext)
##
## $`factor(Quarter)`
##              diff              lwr              upr              p adj
## 2-1  1.388445826  0.7789355  1.9979562 0.0000004
## 3-1  2.604866960  1.9953566  3.2143773 0.0000000
## 4-1  0.006974809 -0.6025355  0.6164851 0.9999902
## 3-2  1.216421134  0.6069108  1.8259315 0.0000081
## 4-2 -1.381471017 -1.9909813 -0.7719607 0.0000005
## 4-3 -2.597892151 -3.2074025 -1.9883818 0.0000000
```

```
plot(post_test)
```


95% family-wise confidence level



T test o jednakosti srednjih vrijednosti za drugi i treći kvartal

Vidimo da se drugi i treći kvartal razlikuju, pa provodimo t-test s ciljem dokazivanja da je prodaja u trećem kvartalu barem dvostruko veća od prodaje u drugom kvartalu.

Normalnost

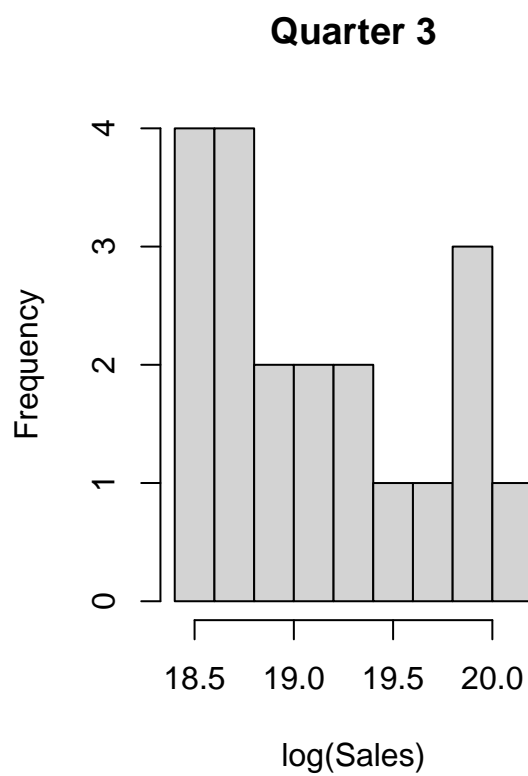
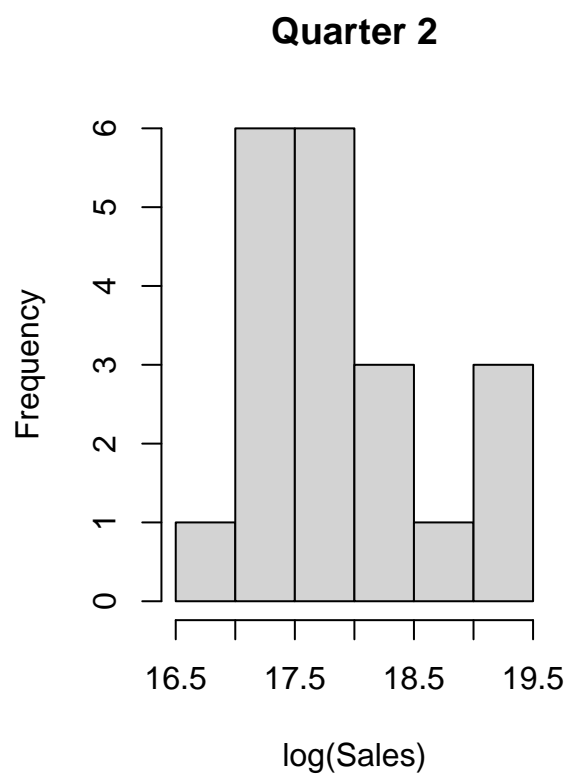
```
lillie.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2])
## D = 0.14832, p-value = 0.2965
```

```
lillie.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]))
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3])
## D = 0.16462, p-value = 0.1664
```

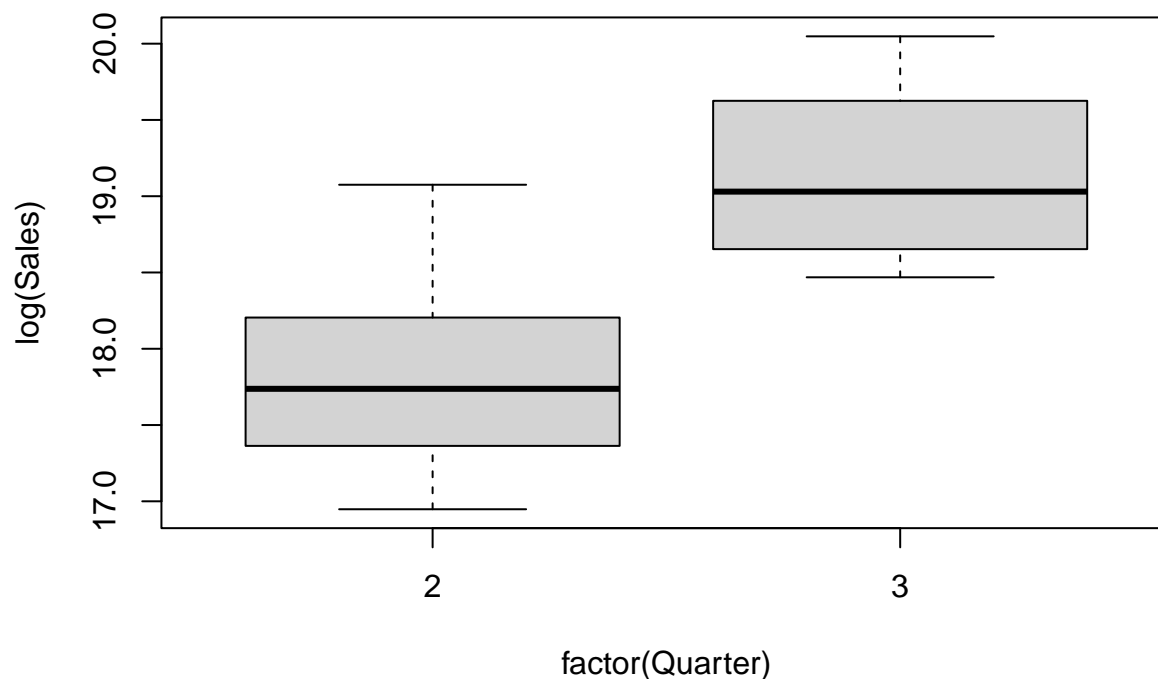
```
par(mfrow = c(1, 2))
hist(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]), xlab = "log(Sales)", main = "Quarter 2")
hist(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]), xlab = "log(Sales)", main = "Quarter 3")
```



Ne odbacujemo normalnost.

Jednakost varijanci

```
boxplot(log(Sales) ~ factor(Quarter), data = subset(arnt_sales_ext, arnt_sales_ext$Quarter == 2 | arnt_s
```



```
bartlett.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2 | arnt_sales_ext$Quarter == 3]) ~ arnt_sales_ext$Quarter)
```

```
##
## Bartlett test of homogeneity of variances
##
## data: log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2 | arnt_sales_ext$Quarter == 3]) by arnt_sales_ext$Quarter
## Bartlett's K-squared = 0.92194, df = 1, p-value = 0.337
```

```
var(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]))
```

```
## [1] 0.4441857
```

```
var(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]))
```

```
## [1] 0.2837433
```

Ne odbacujemo jednakost varijanci.

Provodimo t-test

```
t.test(log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]), 1.049 * log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2]))
```

```
##
## Welch Two Sample t-test
##
## data: log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 3]) and 1.049 * log(arnt_sales_ext$Sales[arnt_sales_ext$Quarter == 2])
## t = 1.7273, df = 35.499, p-value = 0.04641
## alternative hypothesis: true difference in means is greater than 0
```

```
## 95 percent confidence interval:
##  0.007533595          Inf
## sample estimates:
## mean of x mean of y
##  19.11342  18.77396
```

Na razini znacajnosti 5% zakljucujemo da je srednja vrijednost prodaje u 3. kvartalu barem $e^{1.049} = 2.85$ veca nego u 2. kvartalu.

Veza prometa po razlicitim granicnim prijelazima s fundamentima turistickih poduzeca

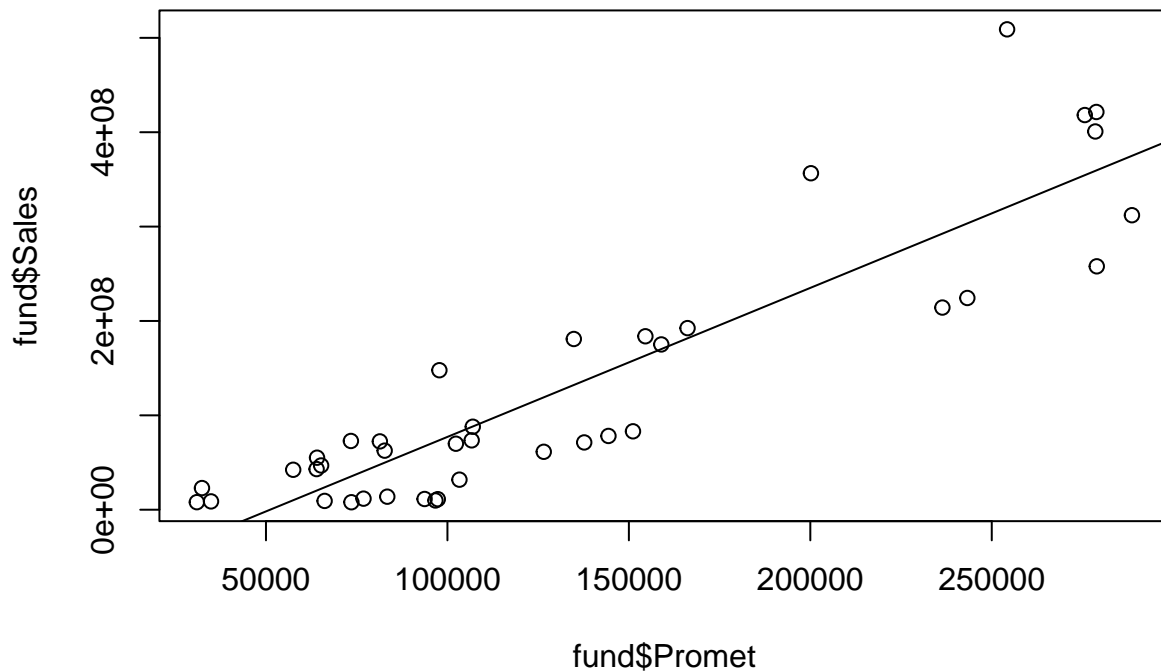
Provodimo linearnu regresiju kako bi vidjeli postoji li veza izmedu prometa na granicnim prijelazima i fundamenata ARNT.

```
x = list()
promet = strani_ulaz
fund <- arnt_fund
for(i in 1:nrow(fund)){
  p <- get_promet(promet, fund$Year[i], fund$Quarter[i])
  x <- append(x, p)
}

fund$Promet <- unlist(x)
model <- lm(Sales ~ Promet, data = fund)

plot(fund$Promet, fund$Sales)

abline(model)
```



```
summary(model)
```

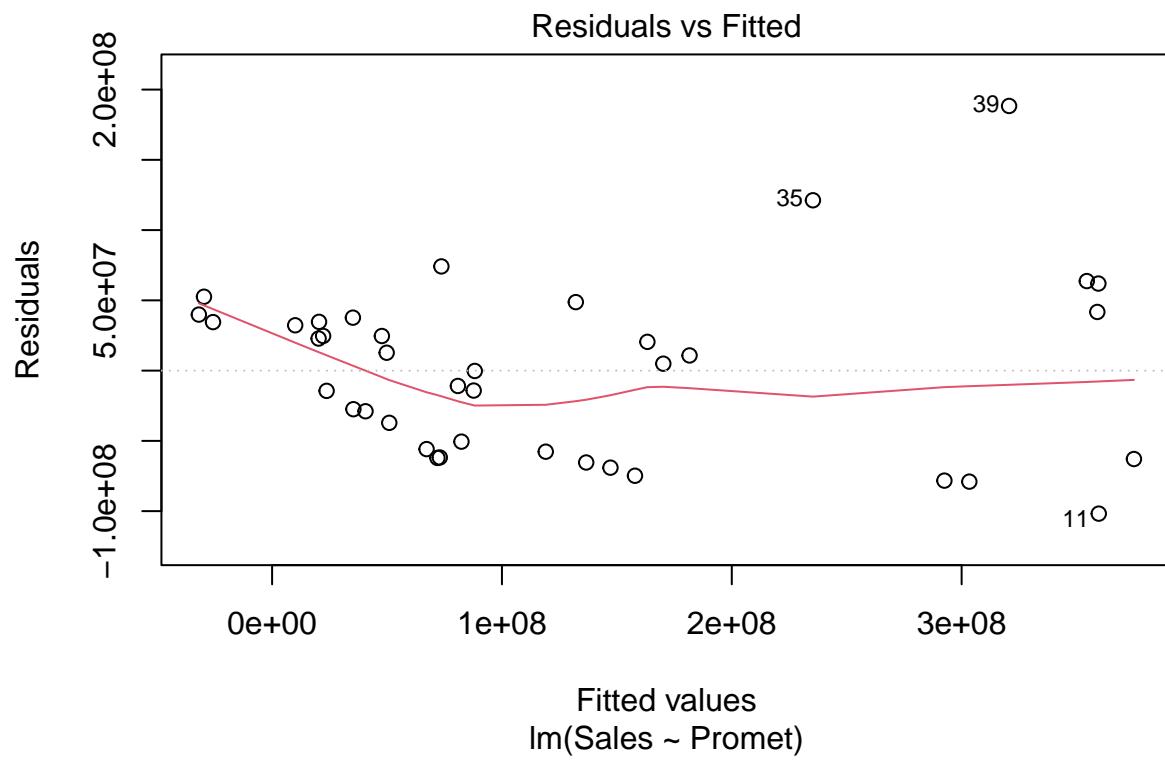
```
##
## Call:
## lm(formula = Sales ~ Promet, data = fund)
##
## Residuals:
```

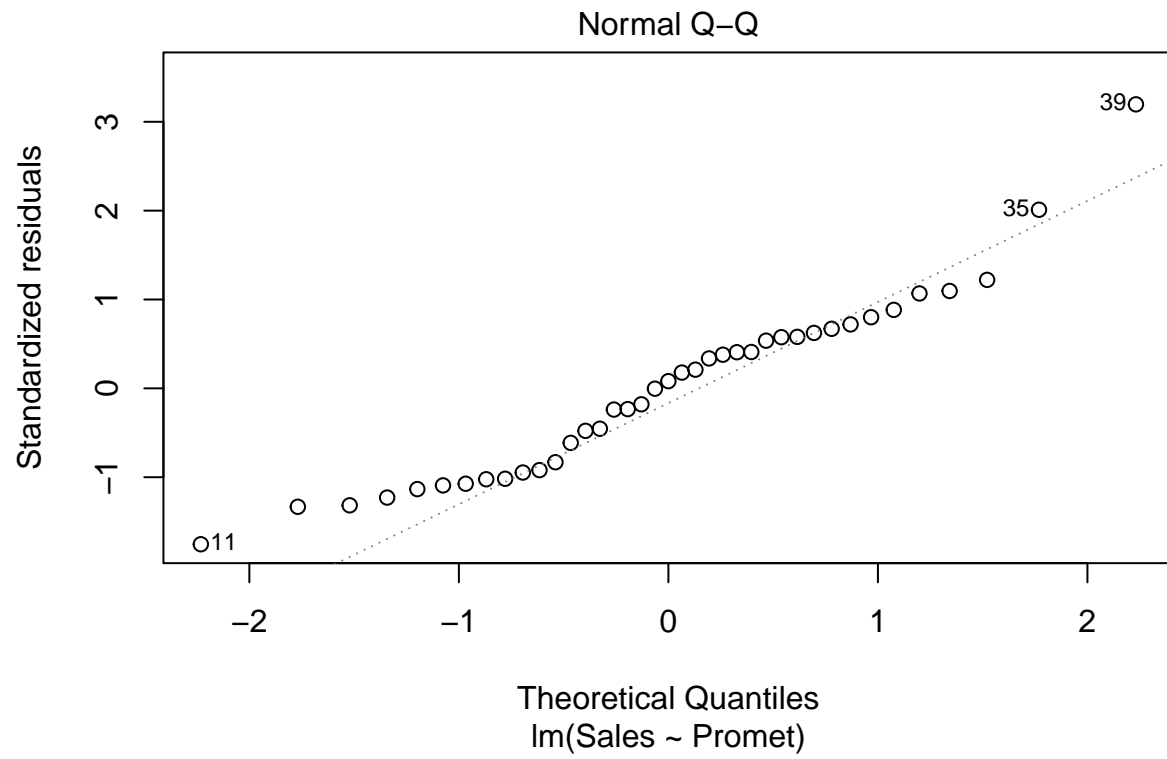
	Min	1Q	Median	3Q	Max
	-101772499	-56779501	4928669	36194795	188339129

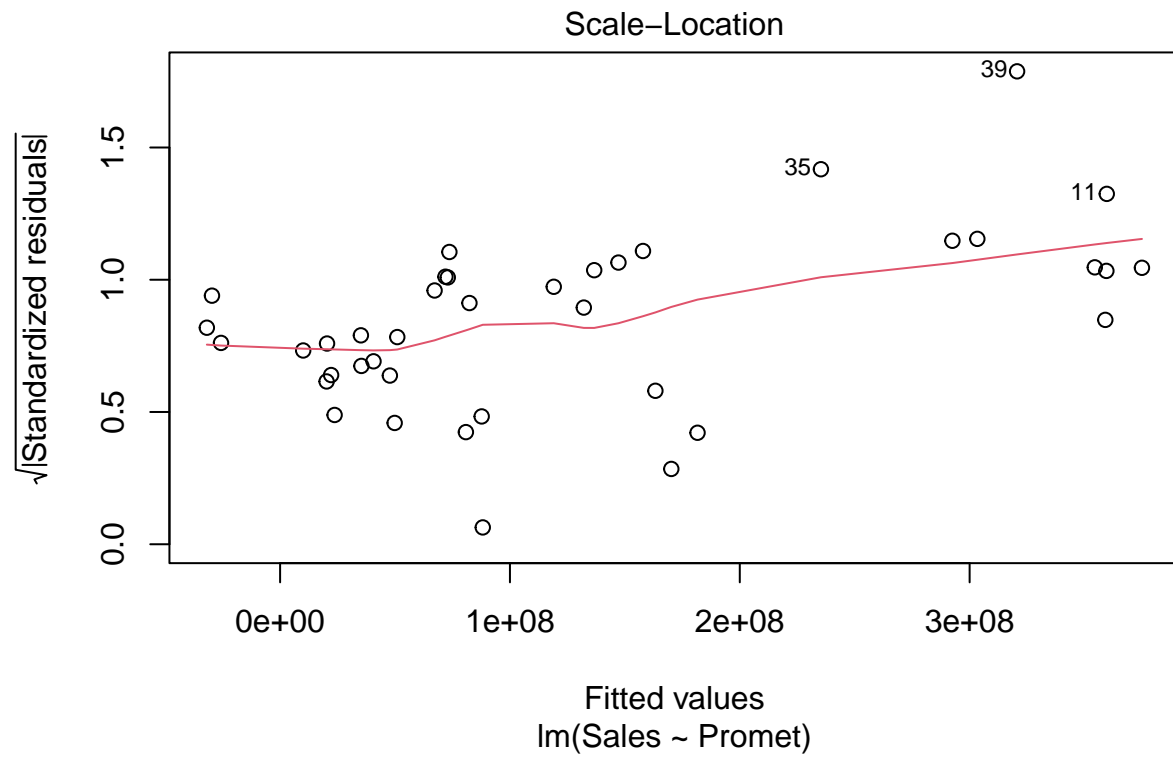
```
##
## Coefficients:
```

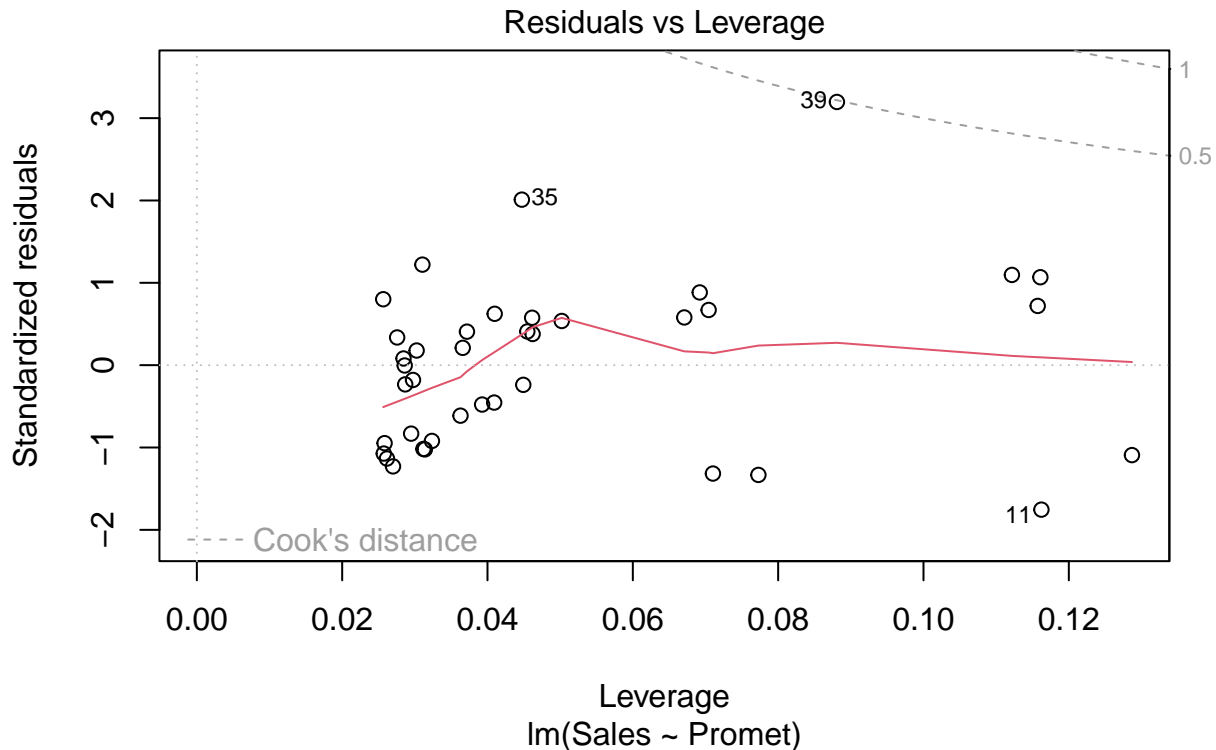
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-8.074e+07	1.967e+07	-4.105	0.000214 ***
Promet	1.579e+03	1.275e+02	12.379	1.01e-14 ***

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 61700000 on 37 degrees of freedom
## Multiple R-squared:  0.8055, Adjusted R-squared:  0.8003
## F-statistic: 153.2 on 1 and 37 DF,  p-value: 1.008e-14
plot(model)
```









$R^2 = 0.8$, $p_value = 1e-14$ nam govori da imamo znacajnu vezu sto mozemo i vidjeti na grafu.

Veza dnevnog kretanja cijene indeksa CROBEXturist i kretanja turista preko hrvatskih granica

Provodimo linearnu regresiju kako bi vidjeli postoji li veza izmedu dnevnog kretanja turista preko hrvatskih granica i cijene CROBEXturist indeksa.

```
strani_ulaz$Date <- as.Date(strani_ulaz$X)
strani_izlaz$Date <- as.Date(strani_izlaz$X)

cijena_promet <- merge(crobex, strani_ulaz, by="Date")
cijena_promet <- merge(cijena_promet, strani_izlaz, by="Date")

summary(cijena_promet)
```

```
##      Date      Price      X.x      Cestovni.x
## Min.   :2013-02-21 Min.   : 940.3 Min.   :2013-02-21 Min.   : 5593
## 1st Qu.:2015-08-03 1st Qu.:2321.2 1st Qu.:2015-08-03 1st Qu.: 62982
## Median :2017-12-28 Median :3322.7 Median :2017-12-28 Median : 88287
## Mean   :2017-12-31 Mean   :2885.5 Mean   :2017-12-31 Mean   :110289
## 3rd Qu.:2020-06-05 3rd Qu.:3558.0 3rd Qu.:2020-06-05 3rd Qu.:137878
## Max.   :2022-11-04 Max.   :3954.5 Max.   :2022-11-04 Max.   :425337
##      NA's      :1
## Željeznički.x  Riječni.x  Pomorski.x  Zračni.x
## Min.   : 0.0   Min.   : 0   Min.   : 0   Min.   : 0
```

```

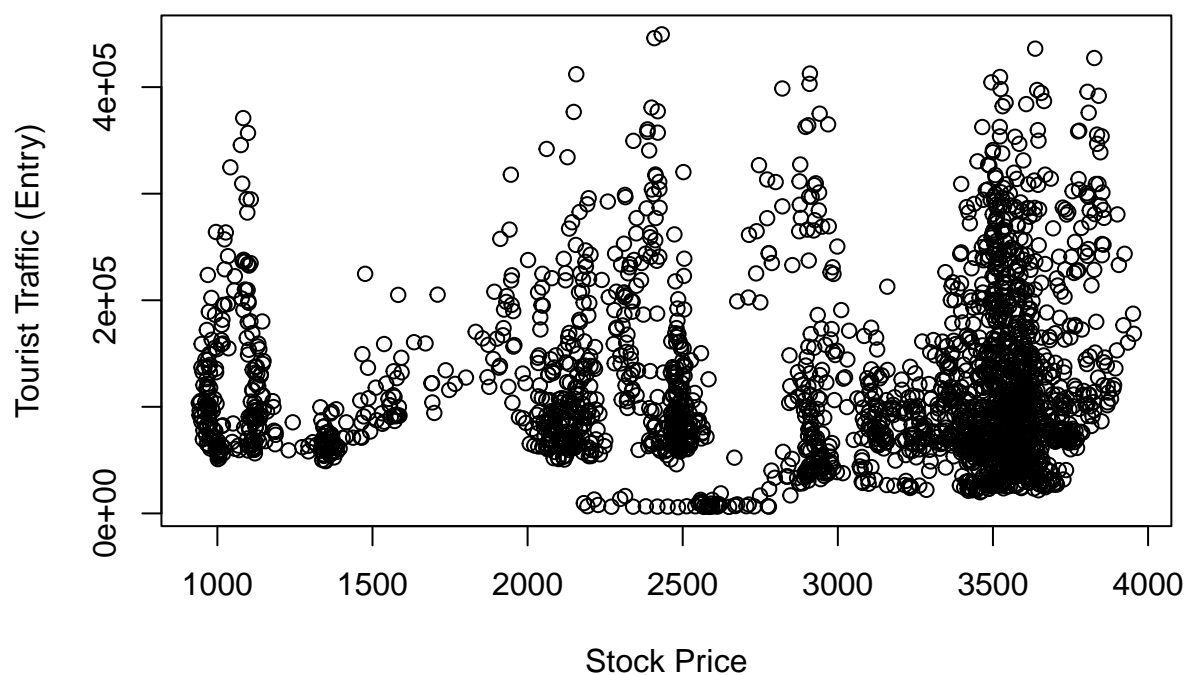
## 1st Qu.: 192.0    1st Qu.: 0    1st Qu.: 36    1st Qu.: 1064
## Median : 402.0    Median : 0    Median : 1040    Median : 3446
## Mean   : 533.9    Mean   : 93    Mean   : 2733    Mean   : 6118
## 3rd Qu.: 694.0    3rd Qu.:155    3rd Qu.: 4666    3rd Qu.:10119
## Max.   :3186.0    Max.   :908    Max.   :18782    Max.   :54546
##
##      UKUPNO.x      Dan.x      Mjesec.x      Godina.x
## Min.   : 5778    Mon:477    Length:2413    Length:2413
## 1st Qu.: 65181    Tue:489    Class :character    Class :character
## Median : 95564    Wed:485    Mode  :character    Mode  :character
## Mean   :119768    Thu:479
## 3rd Qu.:151740    Fri:483
## Max.   :449501    Sat: 0
##                      Sun: 0
##      Quarter.x      X.y      Cestovni.y      Željeznički.y
## Min.   :1.000    Min.   :2013-02-21    Min.   : 5885    Min.   : 0.0
## 1st Qu.:2.000    1st Qu.:2015-08-03    1st Qu.: 62201    1st Qu.: 187.0
## Median :3.000    Median :2017-12-28    Median : 84683    Median : 390.0
## Mean   :2.499    Mean   :2017-12-31    Mean   :105518    Mean   : 476.5
## 3rd Qu.:3.000    3rd Qu.:2020-06-05    3rd Qu.:127539    3rd Qu.: 638.0
## Max.   :4.000    Max.   :2022-11-04    Max.   :394291    Max.   :2479.0
##
##      Riječni.y      Pomorski.y      Zračni.y      UKUPNO.y      Dan.y
## Min.   : 0.0    Min.   : 0    Min.   : 0    Min.   : 5902    Mon:477
## 1st Qu.: 0.0    1st Qu.: 30    1st Qu.: 1597    1st Qu.: 64460    Tue:489
## Median : 0.0    Median : 1003    Median : 3705    Median : 92308    Wed:485
## Mean   : 91.2    Mean   : 2669    Mean   : 6895    Mean   :115650    Thu:479
## 3rd Qu.:153.0    3rd Qu.: 4526    3rd Qu.: 11021    3rd Qu.:144886    Fri:483
## Max.   :909.0    Max.   :21549    Max.   :112784    Max.   :416560    Sat: 0
##                      Sun: 0
##      Mjesec.y      Godina.y      Quarter.y
## Length:2413    Length:2413    Min.   :1.000
## Class :character    Class :character    1st Qu.:2.000
## Mode  :character    Mode  :character    Median :3.000
##                      Mean   :2.499
##                      3rd Qu.:3.000
##                      Max.   :4.000
##

```

```
#Veza između cijene i ulaza turista
```

```
plot(cijena_promet$Price, cijena_promet$UKUPNO.x, xlab="Stock Price", ylab="Tourist Traffic (Entry)", ma
```

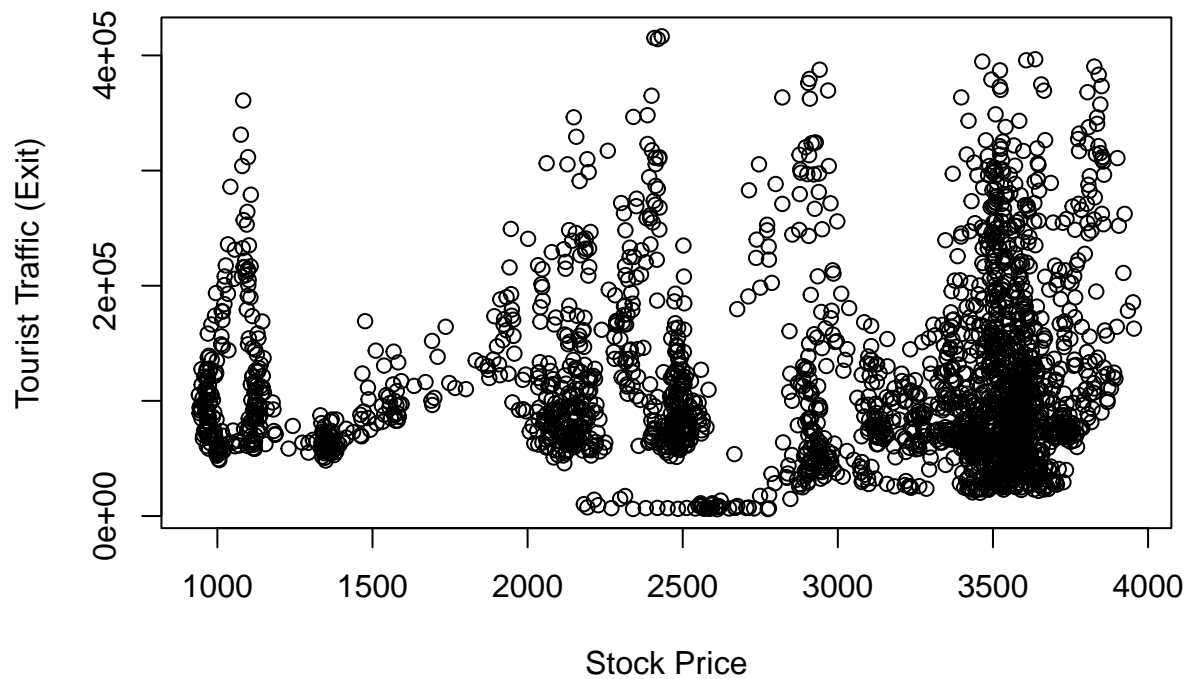
Veza između cijene i ulaza turista



```
#Veza između cijene i izlaza turista
```

```
plot(cijena_promet$Price, cijena_promet$UKUPNO.y, xlab="Stock Price", ylab="Tourist Traffic (Exit)", ma
```

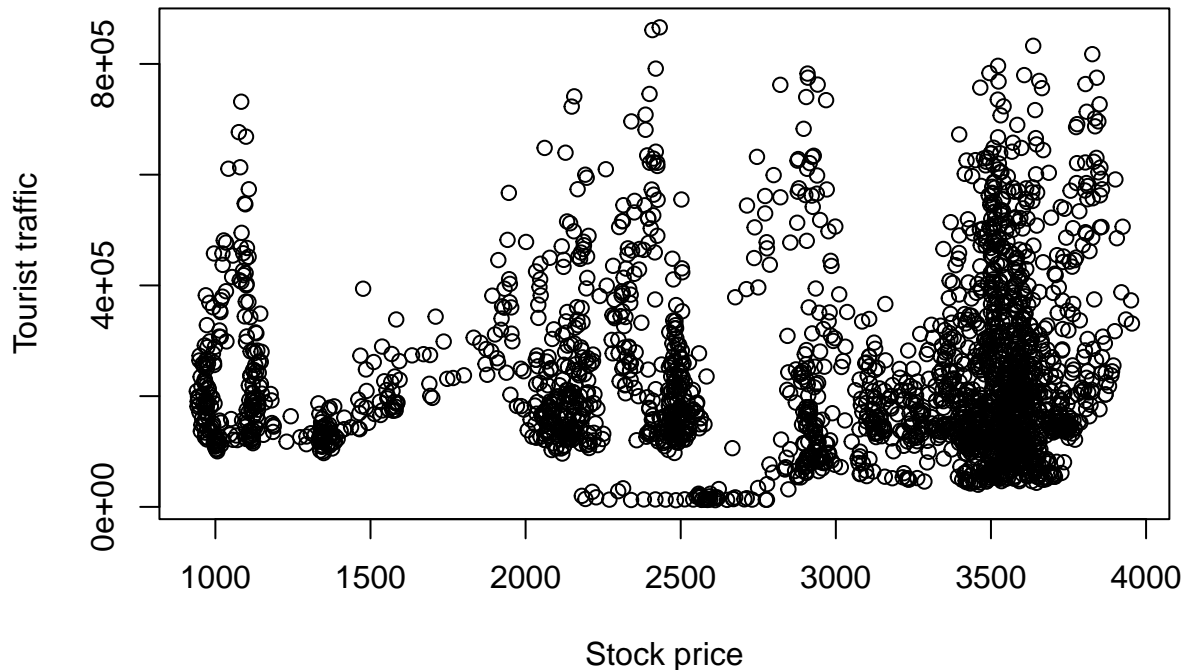
Veza između cijene i izlaza turista



```
#Veza između cijene i ukupnog prometa turista
```

```
plot(cijena_promet$Price, cijena_promet$UKUPNO.x + cijena_promet$UKUPNO.y, xlab = "Stock price", ylab =
```

Veza između cijene i ukupnog prometa turista



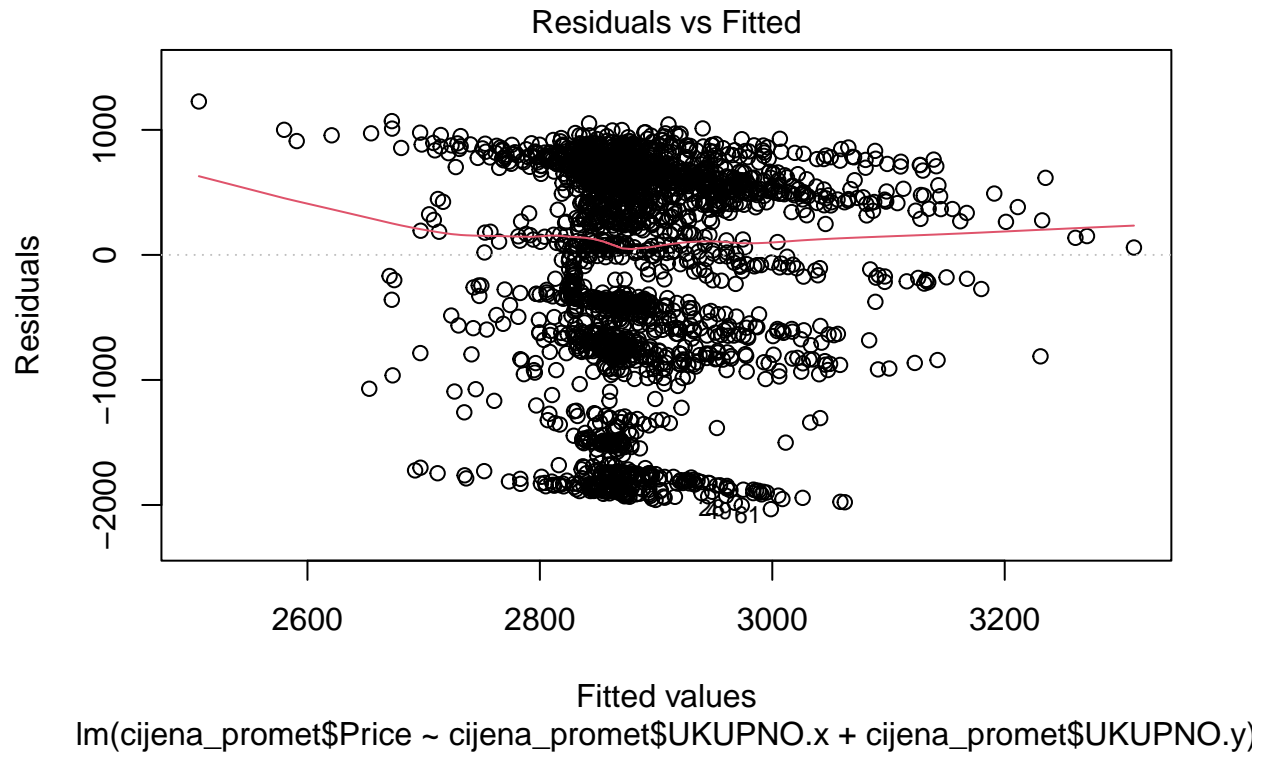
Iz vizualizacije ne vidimo moguću korelaciju.

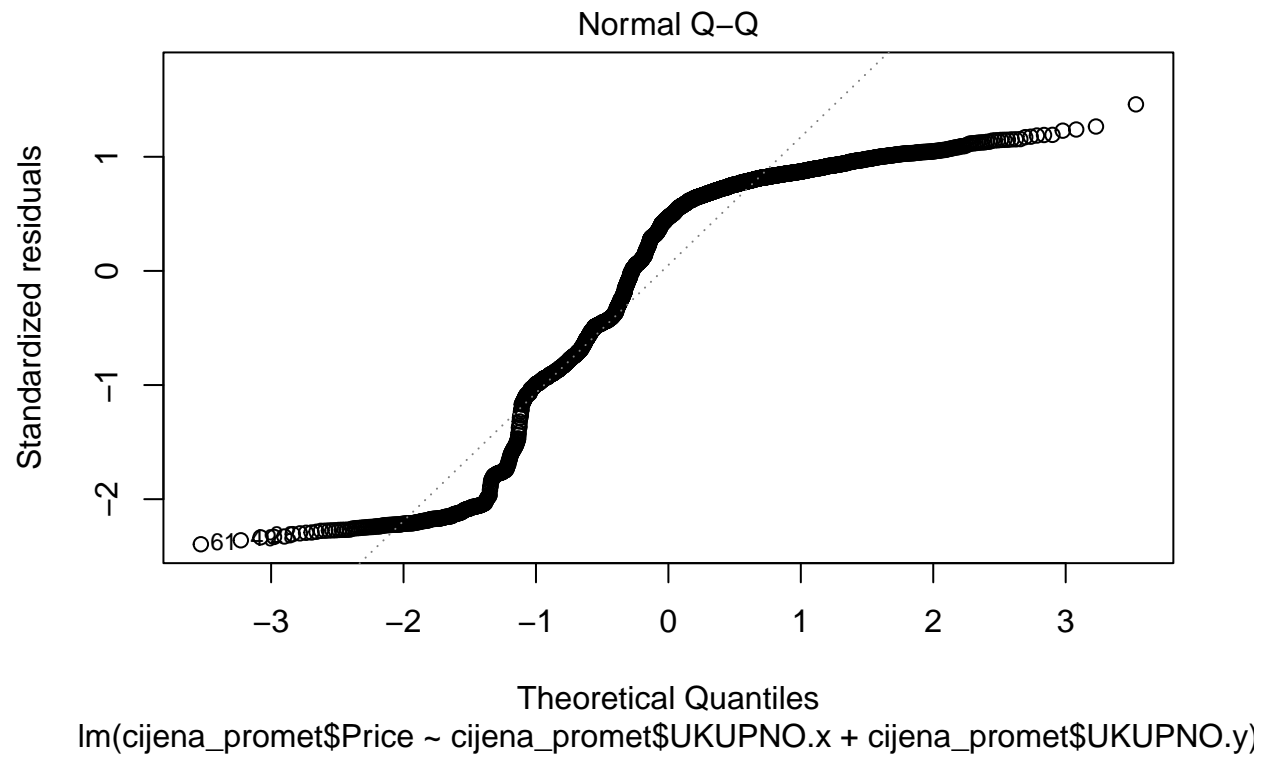
Implementacija linearne regresije sa razinom značajnosti od 5%

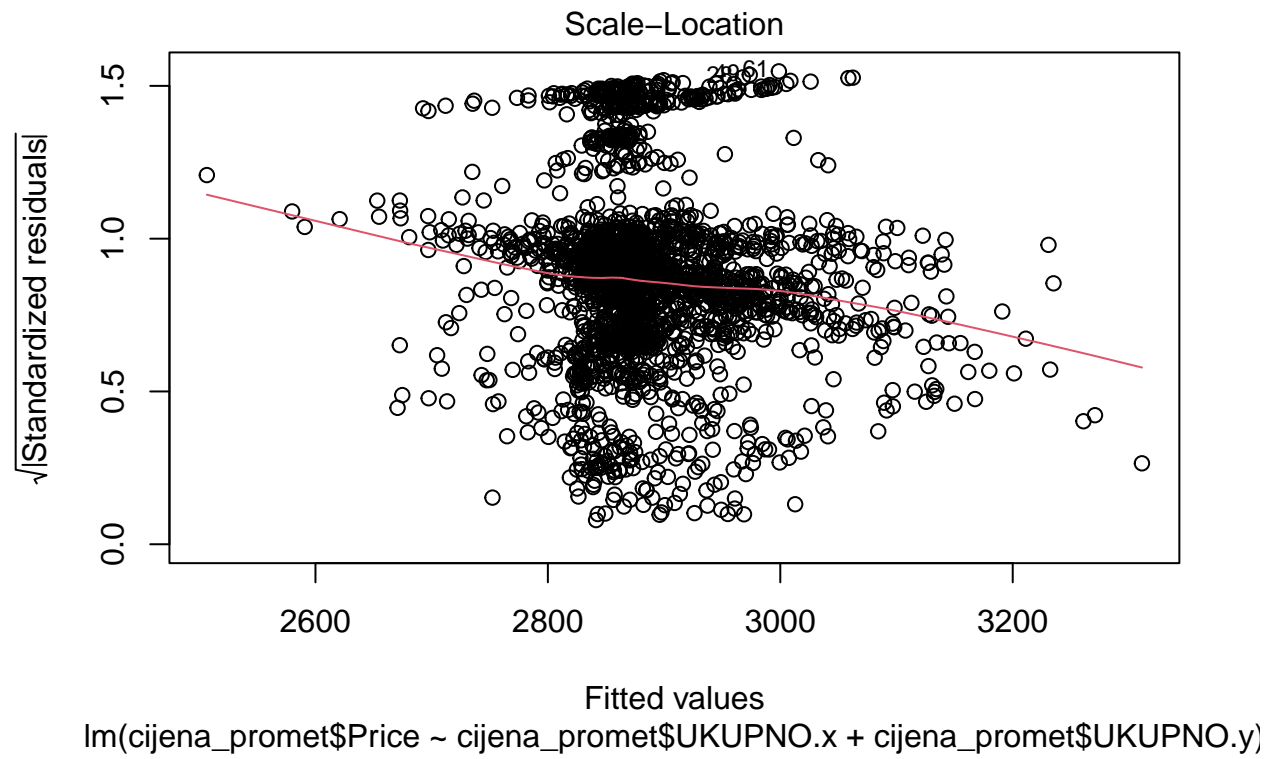
```
model <- lm(cijena_promet$Price ~ cijena_promet$UKUPNO.x + cijena_promet$UKUPNO.y)
summary(model)
```

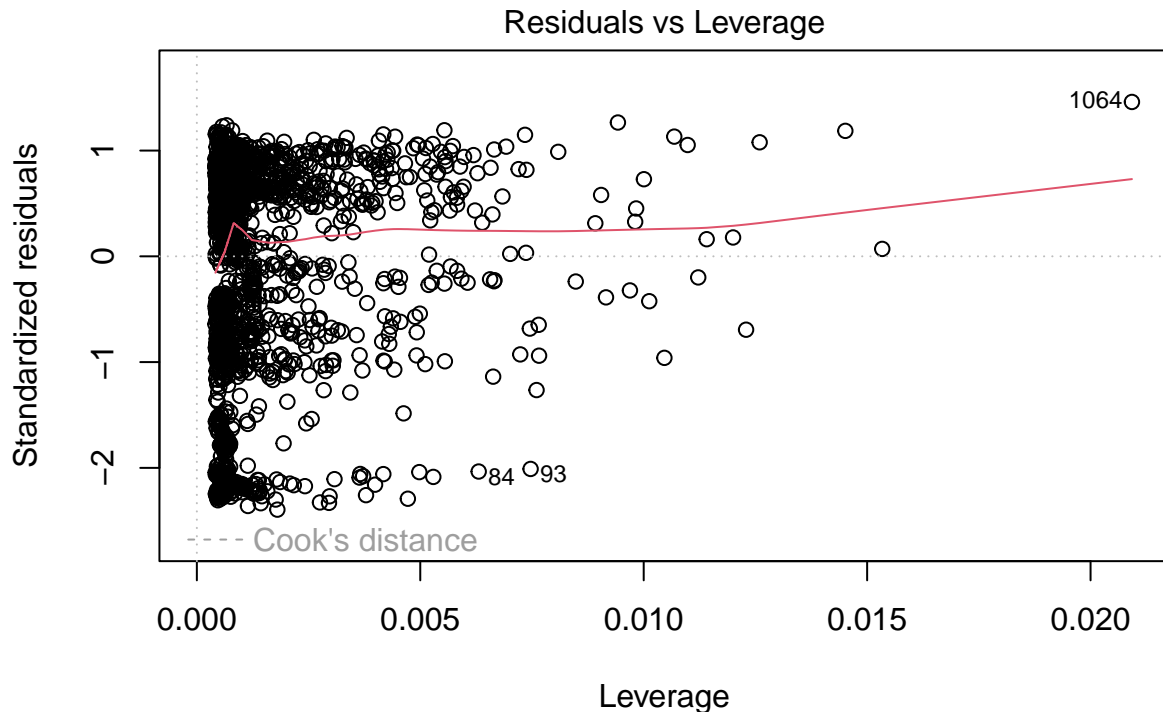
```
##
## Call:
## lm(formula = cijena_promet$Price ~ cijena_promet$UKUPNO.x + cijena_promet$UKUPNO.y)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2033.2  -598.4   398.9   686.6  1227.1
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.824e+03  3.152e+01  89.597 < 2e-16 ***
## cijena_promet$UKUPNO.x -3.615e-03  1.072e-03  -3.372 0.000757 ***
## cijena_promet$UKUPNO.y  4.276e-03  1.121e-03   3.816 0.000139 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 849.7 on 2409 degrees of freedom
## (1 observation deleted due to missingness)
## Multiple R-squared:  0.007325, Adjusted R-squared:  0.006501
## F-statistic: 8.889 on 2 and 2409 DF, p-value: 0.0001425
```

```
plot(model)
```









$\text{lm}(\text{cijena_promet}\$Price \sim \text{cijena_promet}\$UKUPNO.x + \text{cijena_promet}\$UKUPNO.y)$

Mozemo zakljuciti da nema linearne korelacije, sto smo mogli i prepostaviti.

Takoder smo testirali postoji li veza izmedu day to day cijene indeksa i kolicine ulaza stranih turista. Pretpostavka je da ce veci broj ulaza dovesti do vece spekulacije, te do veceg trgovanja indeksom, sto ce dovesti do vecih zamaha u cijeni.

Analiza veze izmedu ulaza i day to day promjene cijene

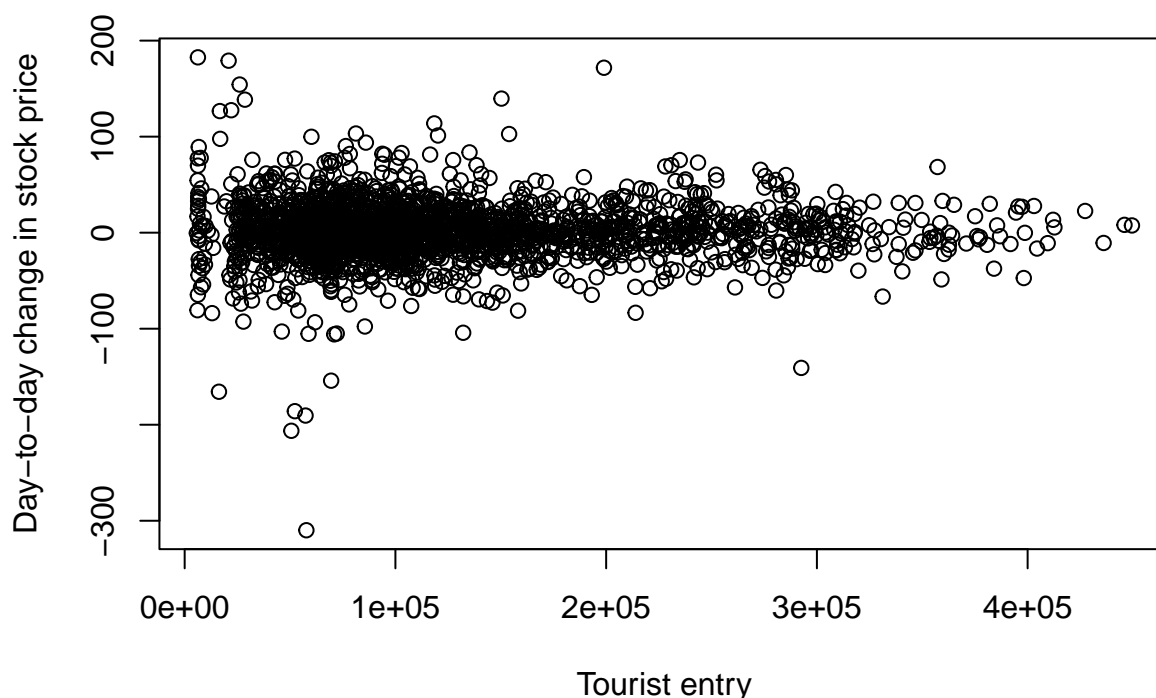
```
cijena_promet_ulaz <- merge(crobex, strani_ulaz, by="Date")
change <- rep(NA, nrow(cijena_promet_ulaz))

for (i in 2:nrow(cijena_promet_ulaz)) {
  change[i] <- cijena_promet_ulaz$Price[i] - cijena_promet_ulaz$Price[i - 1]
}

cijena_promet_ulaz$change <- change

plot(cijena_promet_ulaz$UKUPNO, cijena_promet_ulaz$change, xlab = "Tourist entry", ylab = "Day-to-day change in price")
```

Veza između promjene cijene i ulaza turista



Fit a linear model to the data and summarize the results

```
model <- lm(change ~ UKUPNO, data = cijena_promet_ulaz)
summary(model)
```

##

Call:

```
## lm(formula = change ~ UKUPNO, data = cijena_promet_ulaz)
```

##

Residuals:

	Min	1Q	Median	3Q	Max
##	-310.553	-14.139	-0.624	13.496	182.119

##

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	4.242e-01	1.073e+00	0.395	0.693
## UKUPNO	4.312e-06	7.462e-06	0.578	0.563

##

Residual standard error: 29.14 on 2409 degrees of freedom

(3 observations deleted due to missingness)

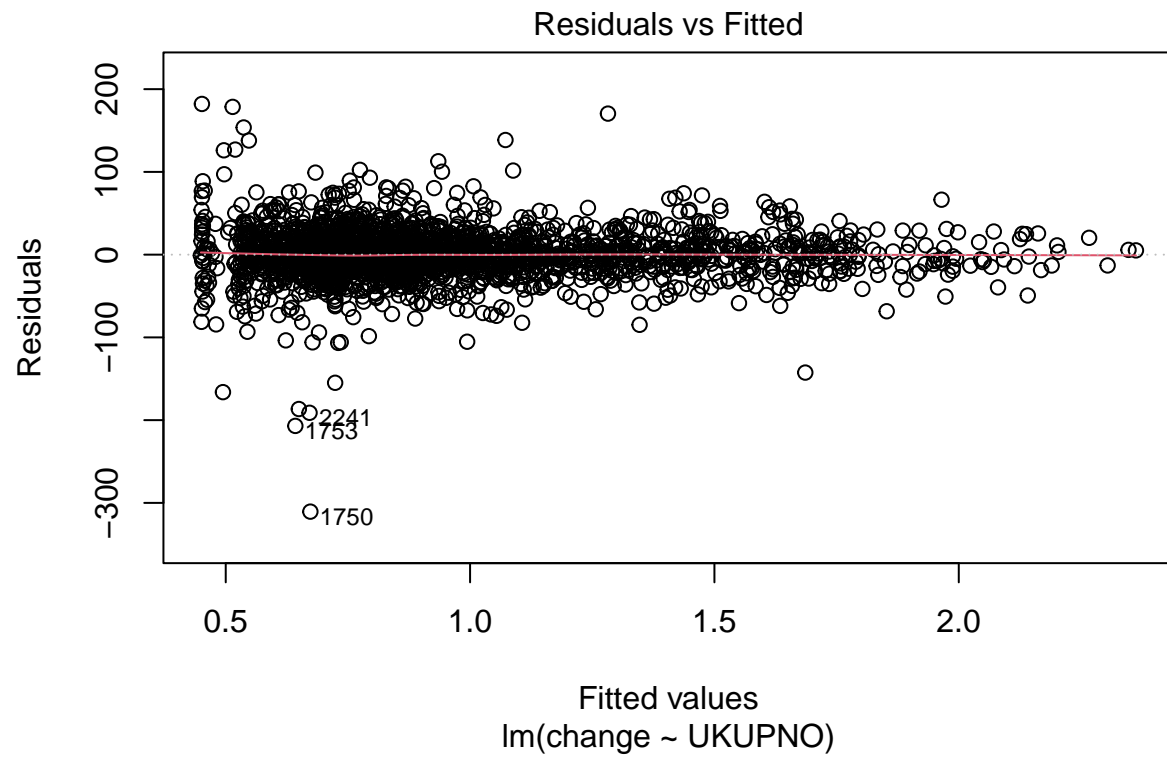
Multiple R-squared: 0.0001386, Adjusted R-squared: -0.0002764

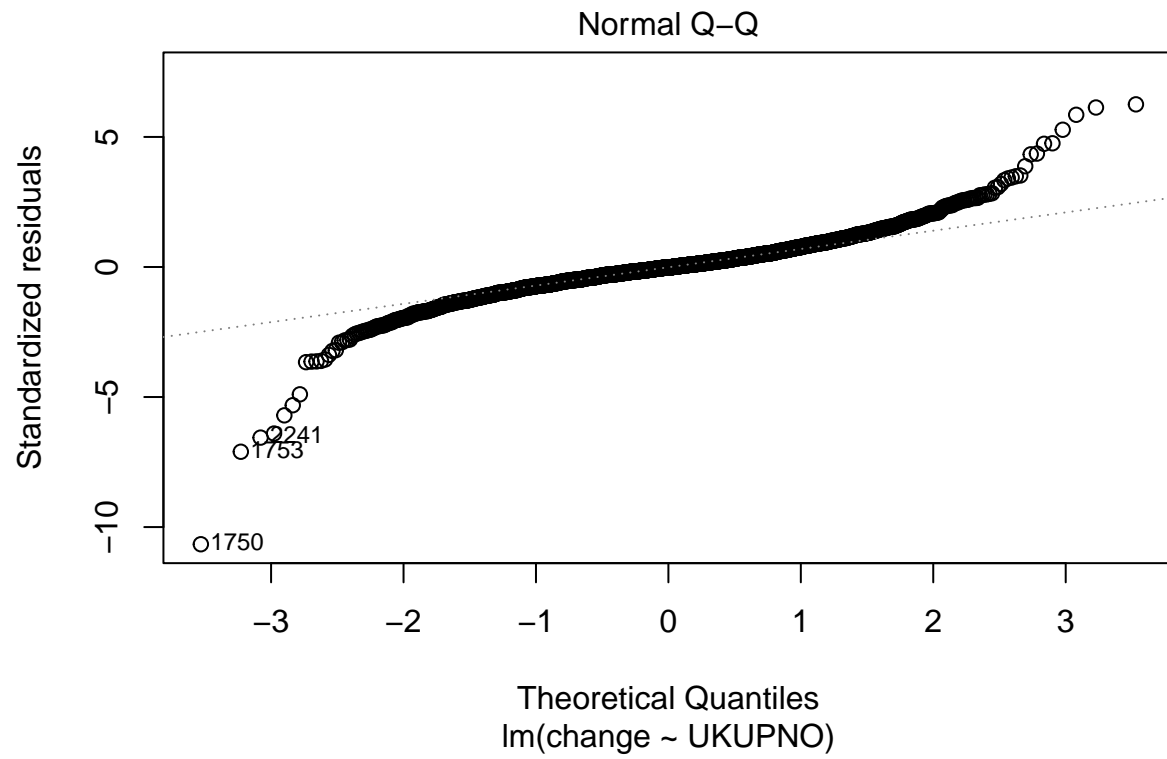
F-statistic: 0.334 on 1 and 2409 DF, p-value: 0.5634

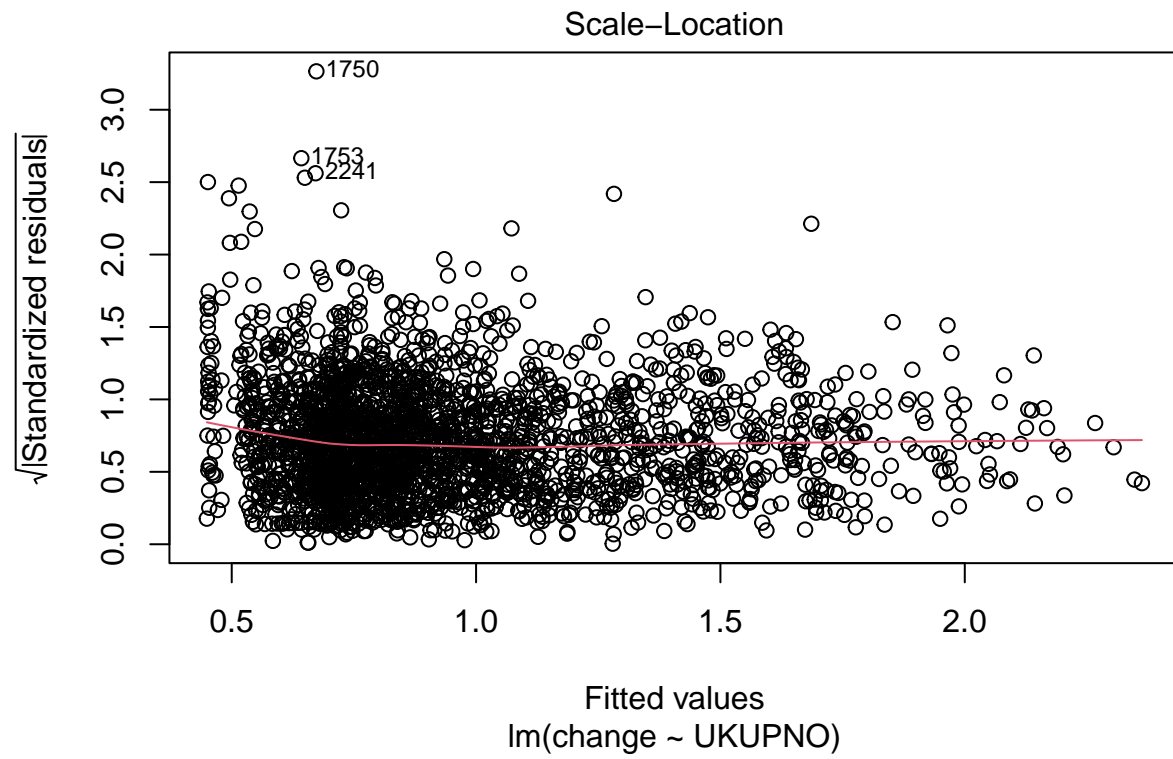
```
cor(cijena_promet_ulaz$UKUPNO, cijena_promet_ulaz$change)
```

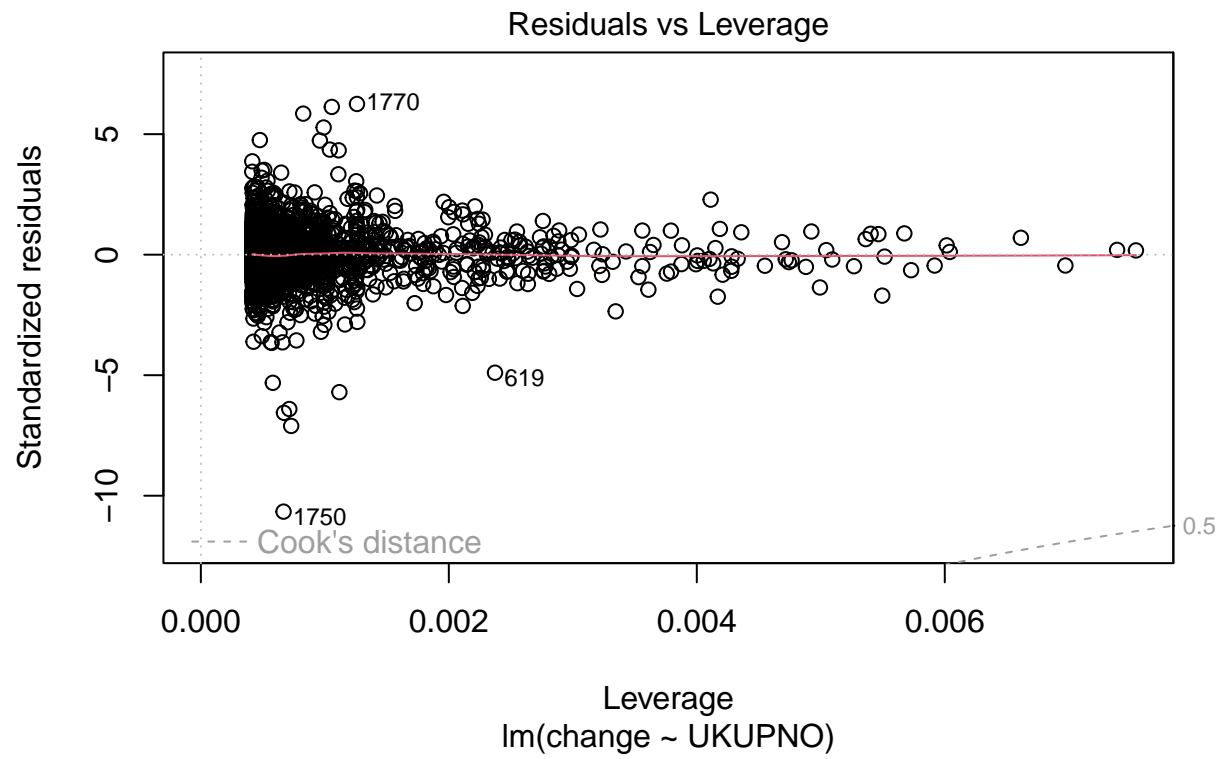
[1] NA

```
plot(model)
```









I tu smo dokazali da nema korelacije.