

Projekt SAP

Tema 2 - Uloga izvoza i uvoza u gospodarstvu

Pavo Matanović, Karla Baričević, Slavko Boldin

Učitavanje podataka i deskriptivna analiza

Na početku učitavamo podatke i analiziramo kako izgledaju podaci.

```
export.data = read.csv("Export_data.csv", fileEncoding="UTF-8-BOM")
# head(export.data)

import.data = read.csv("Import_data.csv", fileEncoding="UTF-8-BOM")
# head(import.data)

gdp.data = read.csv("GDP_data.csv", fileEncoding="UTF-8-BOM")
# head(gdp.data)

gdp.pc.data = read.csv("GDPpercapita_data.csv", fileEncoding="UTF-8-BOM")
# head(gdp.pc.data)
```

Sljedeći blok koda generira dataframe sa brojem upisanih podataka te brojem procjena među upisanim podacima.

```
export.loc.cnt = export.data %>% group_by(LOCATION) %>%
  summarise(exp_n = n(), exp_est = sum(Flag.Codes == 'E')) %>%
  arrange(desc(exp_n), exp_est)
import.loc.cnt = import.data %>% group_by(LOCATION) %>%
  summarise(imp_n = n(), imp_est = sum(Flag.Codes == 'E')) %>%
  arrange(desc(imp_n), imp_est)
gdp.loc.cnt = gdp.data %>% group_by(LOCATION) %>%
  summarise(gdp_n = n(), gdp_est = sum(Flag.Codes == 'E')) %>%
  arrange(desc(gdp_n), gdp_est)
gdp.pc.loc.cnt = gdp.pc.data %>% group_by(LOCATION) %>%
  summarise(gdp_pc_n = n(), gdp_pc_est = sum(Flag.Codes == 'E')) %>%
  arrange(desc(gdp_pc_n), gdp_pc_est)
loc.cnt = merge(merge(export.loc.cnt, import.loc.cnt), merge(gdp.loc.cnt, gdp.pc.loc.cnt))
knitr::kable(
  head(arrange(loc.cnt,
    desc(loc.cnt[,2]), desc(loc.cnt[,4]), desc(loc.cnt[,6]), desc(loc.cnt[,8]),
    loc.cnt[,3], loc.cnt[,5], loc.cnt[,7], loc.cnt[,9]), 20),
  caption = "Broj podataka za pojedinu državu"
)
```

Table 1: Broj podataka za pojedinu državu

LOCATION	exp_n	exp_est	imp_n	imp_est	gdp_n	gdp_est	gdp_pc_n	gdp_pc_est
CAN	41	0	41	0	41	0	41	0

LOCATION	exp_n	exp_est	imp_n	imp_est	gdp_n	gdp_est	gdp_pc_n	gdp_pc_est
DNK	41	0	41	0	41	0	41	0
FRA	41	0	41	0	41	0	41	0
CHE	41	1	41	1	41	1	41	1
FIN	41	1	41	1	41	1	41	1
DEU	41	12	41	12	41	12	41	12
SWE	41	14	41	14	41	14	41	14
GBR	41	16	41	16	41	0	41	0
AUT	41	16	41	16	41	16	41	16
BEL	41	16	41	16	41	16	41	16
ESP	41	16	41	16	41	16	41	16
GRC	41	16	41	16	41	16	41	16
IRL	41	16	41	16	41	16	41	16
ITA	41	16	41	16	41	16	41	16
NLD	41	16	41	16	41	16	41	16
ISL	41	16	41	16	41	16	41	17
PRT	41	17	41	17	41	17	41	17
AUS	40	0	40	0	40	0	40	0
NOR	40	0	40	0	40	0	40	0
USA	40	0	40	0	40	0	40	0

```
{r, include=FALSE} remove(export.loc.cnt, import.loc.cnt, gdp.loc.cnt, gdp.pc.loc.cnt,
loc.cnt)
```

Odabrane države

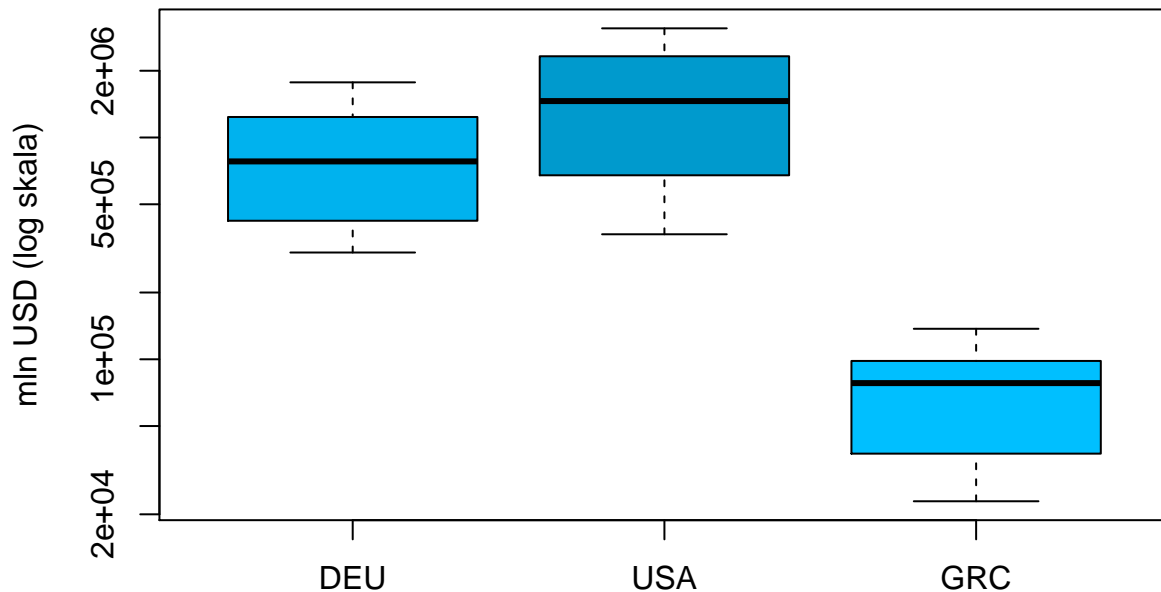
Odabrali smo USA, Njemačku(DEU) i Grčku(GRC) za analizu.

```
usa = data.frame(export.mln_usd = export.data$Value[export.data$LOCATION == "USA"],
  import.mln_usd = import.data$Value[import.data$LOCATION == "USA"],
  gdp.mln_usd = gdp.data$Value[gdp.data$LOCATION == "USA"],
  gdp.pc.usd_cap = gdp.pc.data$Value[gdp.pc.data$LOCATION == "USA"])
usa$net.trade = usa$export.mln_usd - usa$import.mln_usd
deu = data.frame(export.mln_usd = export.data$Value[export.data$LOCATION == "DEU"],
  import.mln_usd = import.data$Value[import.data$LOCATION == "DEU"],
  gdp.mln_usd = gdp.data$Value[gdp.data$LOCATION == "DEU"],
  gdp.pc.usd_cap = gdp.pc.data$Value[gdp.pc.data$LOCATION == "DEU"])
deu$net.trade = deu$export.mln_usd - deu$import.mln_usd
grc = data.frame(export.mln_usd = export.data$Value[export.data$LOCATION == "GRC"],
  import.mln_usd = import.data$Value[import.data$LOCATION == "GRC"],
  gdp.mln_usd = gdp.data$Value[gdp.data$LOCATION == "GRC"],
  gdp.pc.usd_cap = gdp.pc.data$Value[gdp.pc.data$LOCATION == "GRC"])
grc$net.trade = grc$export.mln_usd - grc$import.mln_usd
```

Deskriptivna statistika

```
boxplot(deu$import.mln_usd,
  usa$import.mln_usd,
  grc$import.mln_usd,
  names = c("DEU", "USA", "GRC"), main = "Sredine uvoza",
  col = c("deepskyblue2", "deepskyblue3", "deepskyblue"),
  ylab = "mln USD (log skala)",
  log = "y")
```

Sredine uvoza



Vidimo da im se čisti uvoz u mil. USD razlikuje jako čak i na logaritamskoj skali.

To možemo potvrditi i statističkim t testom.

```
t.test(log(usa$import.mln_usd), log(deu$import.mln_usd),
       alternative = "g", var.equal = FALSE)
```

```
##
## Welch Two Sample t-test
##
## data: log(usa$import.mln_usd) and log(deu$import.mln_usd)
## t = 3.607, df = 76.264, p-value = 0.0002755
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.2795736      Inf
## sample estimates:
## mean of x mean of y
## 14.01251 13.49321
```

```
t.test(log(deu$import.mln_usd), log(grc$import.mln_usd),
       alternative = "g", var.equal = FALSE)
```

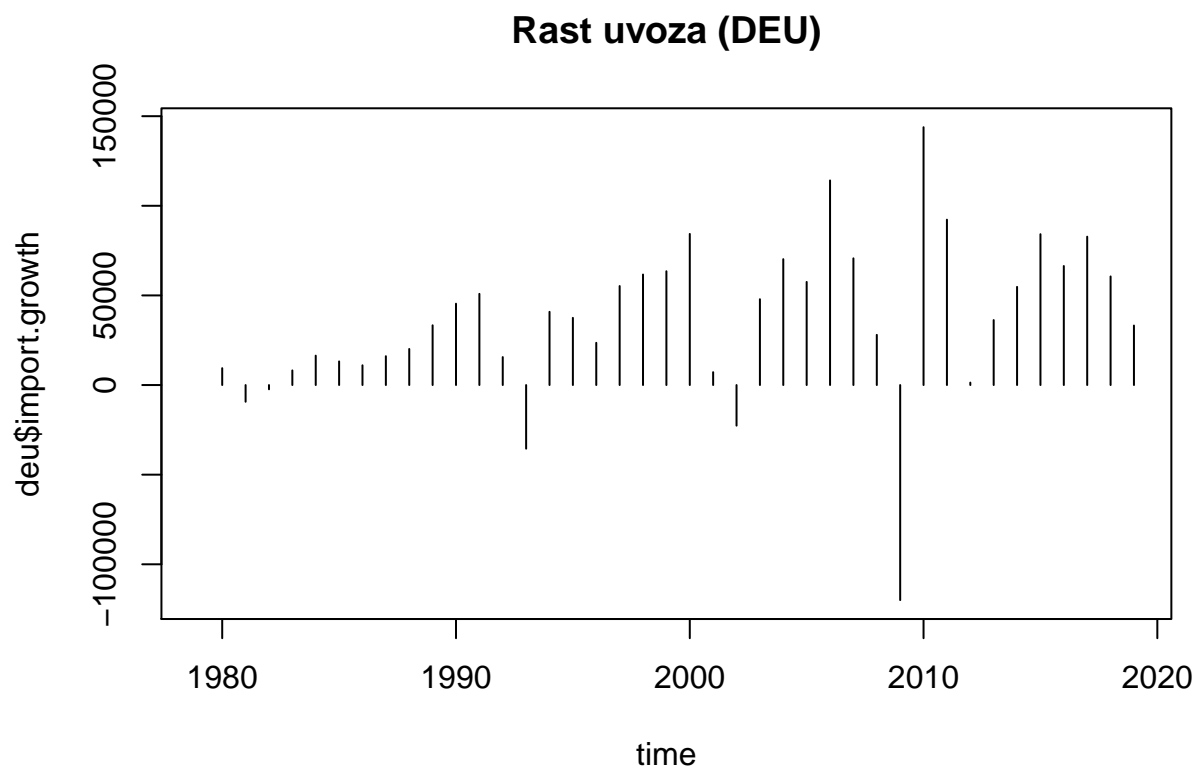
```
##
## Welch Two Sample t-test
##
## data: log(deu$import.mln_usd) and log(grc$import.mln_usd)
## t = 19.365, df = 79.86, p-value < 2.2e-16
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  2.265246      Inf
## sample estimates:
## mean of x mean of y
## 13.49321 11.01499
```

```

usa = usa %>% mutate(import.mln_usd, import.growth = import.mln_usd - lag(import.mln_usd))
deu = deu %>% mutate(import.mln_usd, import.growth = import.mln_usd - lag(import.mln_usd))
grc = grc %>% mutate(import.mln_usd, import.growth = import.mln_usd - lag(import.mln_usd))

time = 1979:2019
plot(time, deu$import.growth, type = "h", main = "Rast uvoza (DEU)")

```



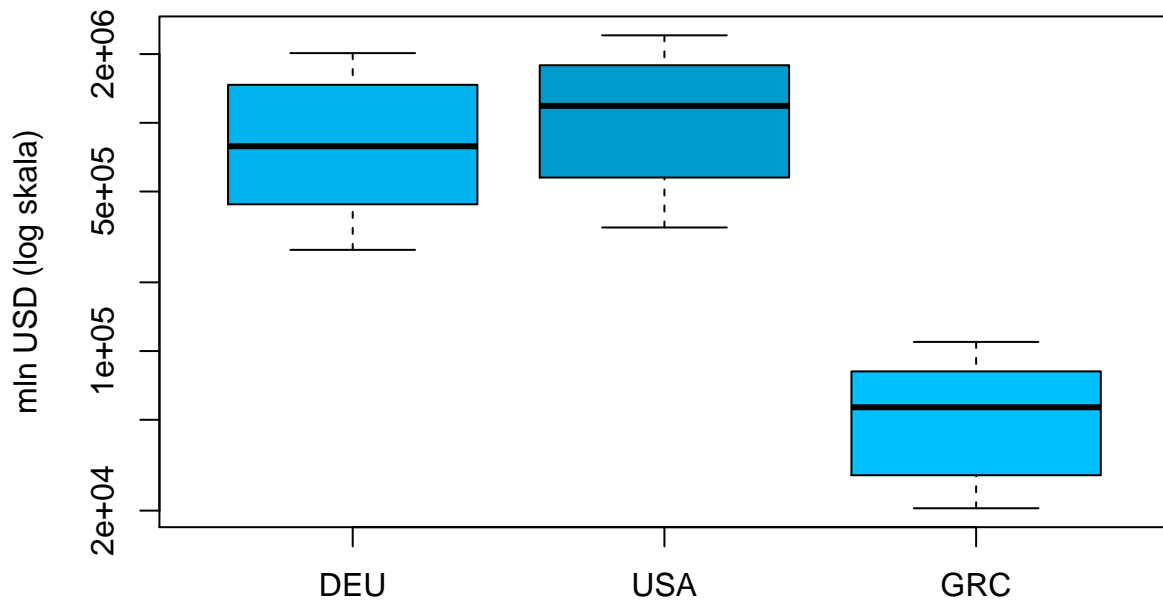
Izvoz

```

boxplot(deu$export.mln_usd,
        usa$export.mln_usd,
        grc$export.mln_usd,
        names = c("DEU", "USA", "GRC"), main = "Sredine izvoza",
        col = c("deepskyblue2", "deepskyblue3", "deepskyblue"),
        ylab = "mln USD (log skala)",
        log = "y")

```

Sredine izvoza



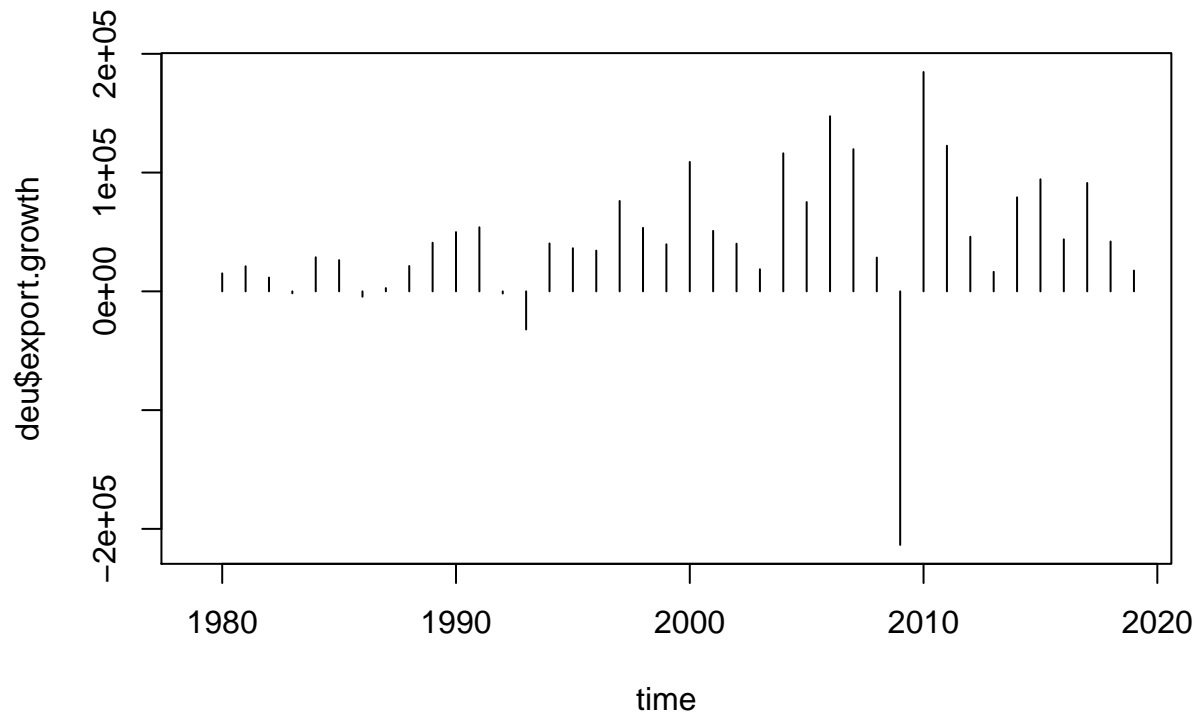
Podaci su slični kao i kod uvoza, SAD prednjači i u izvozu u odnosu na Njemačku i Grčku.

```
usa = usa %>% mutate(export.mln_usd, export.growth = export.mln_usd - lag(export.mln_usd))
deu = deu %>% mutate(export.mln_usd, export.growth = export.mln_usd - lag(export.mln_usd))
grc = grc %>% mutate(export.mln_usd, export.growth = export.mln_usd - lag(export.mln_usd))
```

```
time = 1979:2019
```

```
plot(time, deu$export.growth, type = "h", main = "Rast izvoza (DEU)")
```

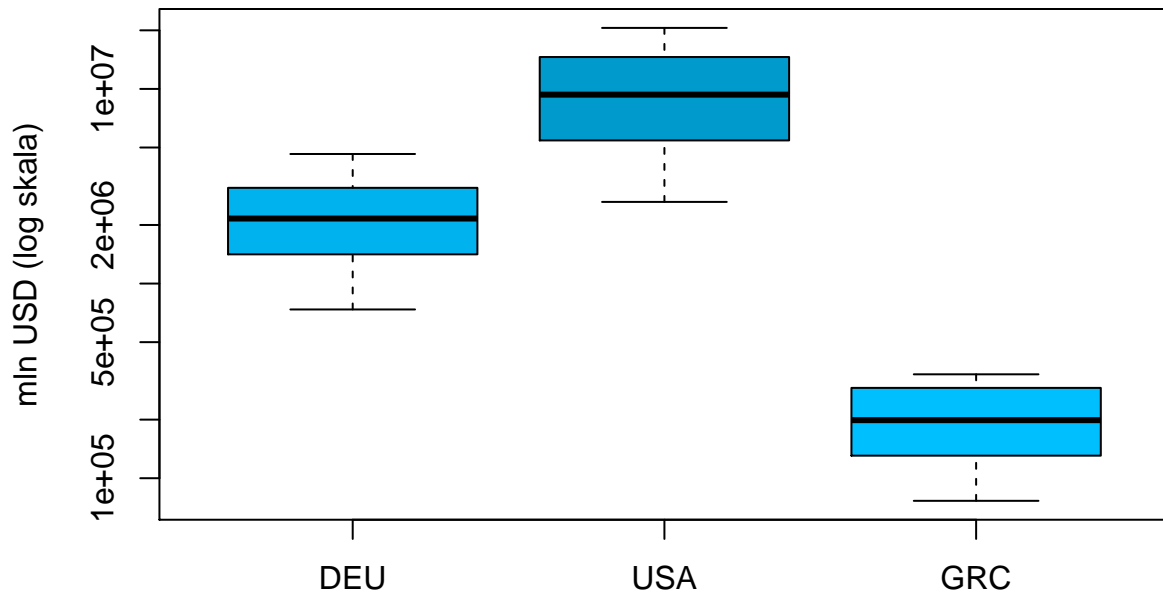
Rast izvoza (DEU)



Za razliku od uvoza koji linearno raste, izvoz više “osjeća” promjene na tržištu (veće fluktuacije), npr. značajan pad izvoza 2009. godine zbog tadašnje svjetske gospodarske krize.

```
boxplot(deu$gdp.mln_usd,
        usa$gdp.mln_usd,
        grc$gdp.mln_usd,
        names = c("DEU", "USA", "GRC"), main = "Sredine BDP-a",
        col = c("deepskyblue2", "deepskyblue3", "deepskyblue"),
        ylab = "mln USD (log skala)",
        log = "y")
```

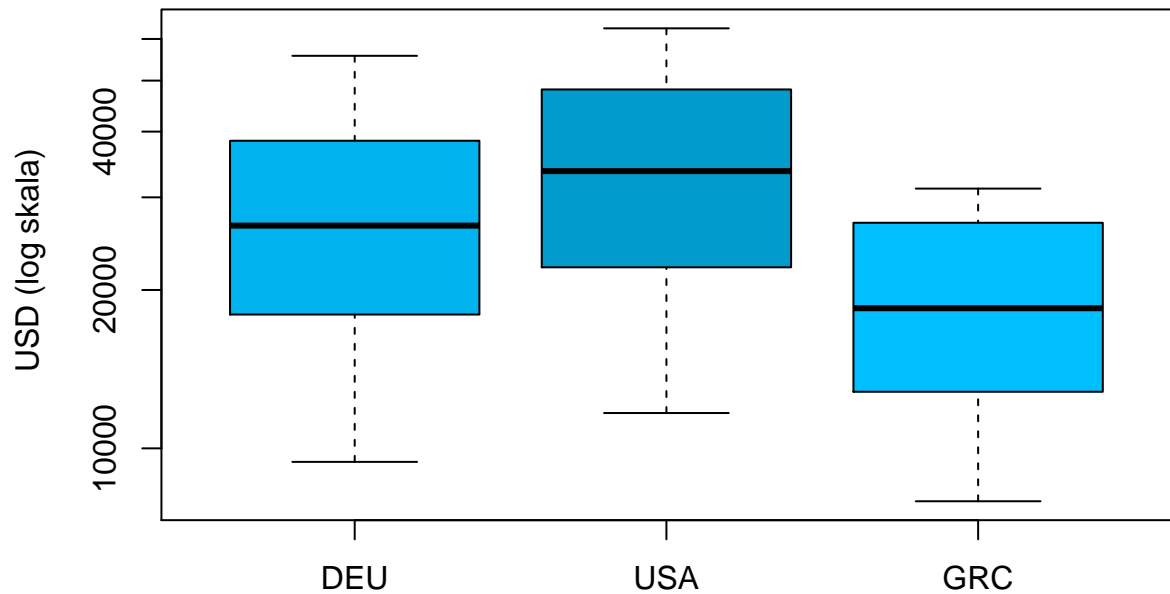
Sredine BDP-a



Kao i kod uvoza i izvoza, po čistom BDP-u SAD značajno prednjači, dok je razlika između Njemačke i Grčke veća od one između SAD-a i Njemačke. No, ovaj prikaz možda nije mjerodavan što se tiče razvijenosti. Treba pogledati BDP po stanovniku:

```
boxplot(deu$gdp.pc.usd_cap,
        usa$gdp.pc.usd_cap,
        grc$gdp.pc.usd_cap,
        names = c("DEU", "USA", "GRC"), main = "Sredine BDP-a po stanovniku",
        col = c("deepskyblue2", "deepskyblue3", "deepskyblue"),
        ylab = " USD (log skala)",
        log = "y")
```

Sredine BDP-a po stanovniku

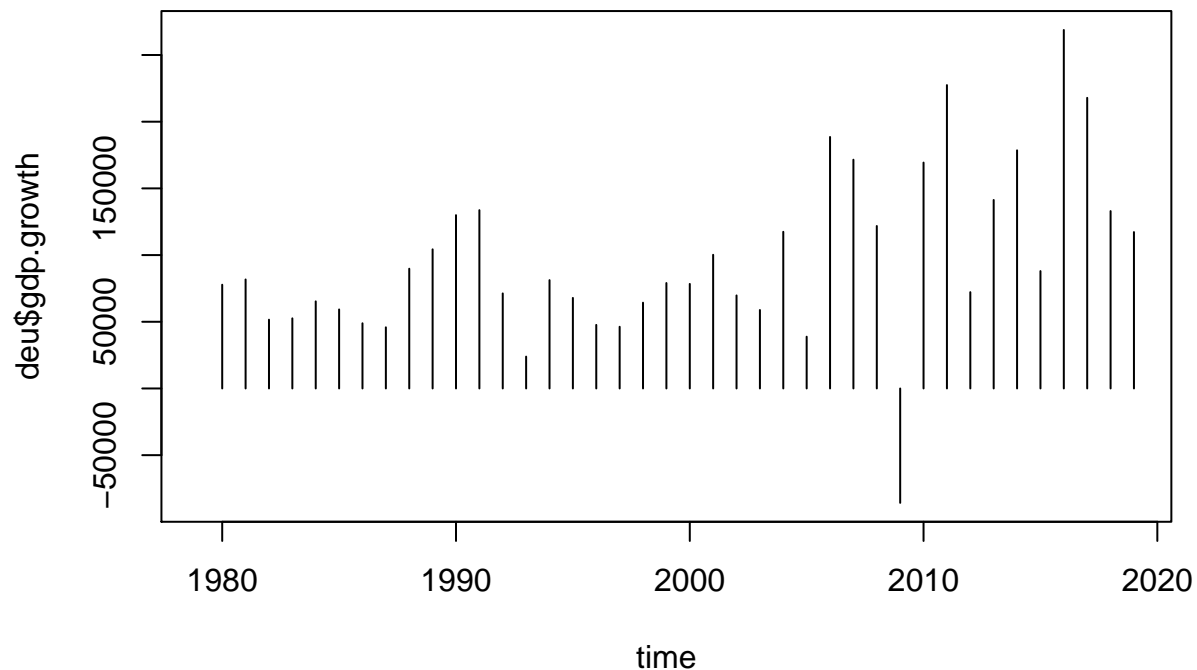


Na prikazu BDP-a po stanovniku podaci su normalizirani brojem stanovnika, razlike nisu toliko značajne, no SAD i dalje prednjači.

```
usa = usa %>% mutate(gdp.mln_usd, gdp.growth = gdp.mln_usd - lag(gdp.mln_usd))
deu = deu %>% mutate(gdp.mln_usd, gdp.growth = gdp.mln_usd - lag(gdp.mln_usd))
grc = grc %>% mutate(gdp.mln_usd, gdp.growth = gdp.mln_usd - lag(gdp.mln_usd))
```

```
time = 1979:2019
plot(time, deu$gdp.growth, type = "h", main = "Rast BDP-a (DEU)")
```

Rast BDP-a (DEU)



BDP

Njemačke je u stalnom porastu uz fluktuacije, a jedini pad BDP-a koji primjećujemo vezan je uz gospodarsku krizu 2009. godine, kada primjećujemo i značajne padove u uvozu i izvozu. Rast BDP-a po stanovniku bit će proporcionalan.