

GDP per capita and development help received

Haraldur Bjarni Davíðsson

10/24/2019

Inngangur

Gögnin sem þetta verkefni byggir á eru fengin frá opinberri gagnasiðu Alþjóðabankans, [datacatalog.worldbank.org](https://data.worldbank.org/indicator/ny.gdp.mktp.cd). Fyrra gagnasettið, sótt af <https://data.worldbank.org/indicator/ny.gdp.mktp.cd> sýnir verga landsframleiðslu allra landa veraldar, ásamt ákveðinna landsvæða (t.d. í Suður-Aasíu eða Mið-Austurlöndum og Norður-Afríku) miðað við höfðatölu í bandaríkjadöllum. Einnig má þar finna verga landsframleiðslu m.v. höfðatölu landa sem falla undir aðra flokka, s.s. og hátekju- og lágtækjulönd, þróunarlönd (skv. skilgreiningu Efnahags- og framfarastofnunarinnar, OECD) o.fl. Seinna gagnasettið, sótt af <https://data.worldbank.org/indicator/DT.ODA.ODAT.CD>, er sett upp á sama máta og það fyrra, en það byggir á upplýsingum um hve mikla þróunaraðstoð hvert land hefur notið á hverju ári, í bandaríkjadöllum talið. Bæði gagnasettin verða síðan síuð þannig að auðveldara verður að skoða gögnin á myndrænan máta, en einungis verður byggt á gögnum einstakra landa, en ekki annarra flokkana. Markmið þessa verkefnis er að kanna verga landsframleiðslu m.v. höfðatölu hinna ýmissu landa og setja gögnin fram á myndrænan hátt. Þá verða tengslin milli vergrar landsframleiðslu og þróunaraðstoðar könnuð, þ.e. hvort stuðningur sem ríki njóta leiði til aukinnar vergrar landsframleiðslu eða hvort aukin landsframleiðsla hafi áhrif á umfang þróunaraðstoðar.

Innlestur gagna

Gögnin eru lesin inn og í kjölfarið eru þau hreinsuð. Ýmsir dálkar, ásamt fyrstu fjórum línunum eru teknar í burtu svo gögnin verði læsilegri og auðveldari í meðförum. Fyrir neðan kóðann eru fyrstu 9 dálkar og 6 línur í gdp gagnasettinu.

```
gdp<-read.csv(file="gdp/API_NY.GDP.PCAP.PP.CD_DS2_en_csv_v2_382630.csv",sep=";",skip=4,header=T)
devhelp<-read.csv(file="devhelp/API_DT.ODA.ODAT.CD_DS2_en_csv_v2_385593.csv",sep=";",skip=4,header=T)
colnames(gdp)<-gsub('X','"',colnames(gdp),fixed=TRUE)
colnames (devhelp)<-gsub('X','"',colnames(devhelp),fixed=TRUE)
```

```
head(gdp %>% select(1:9))
```

```
##      Country.Name Country.Code Indicator.Name
## 1      Aruba      ABW GDP per capita, PPP (current international $)
## 2 Afghanistan      AFG GDP per capita, PPP (current international $)
## 3      Angola      AGO GDP per capita, PPP (current international $)
## 4      Albania      ALB GDP per capita, PPP (current international $)
## 5      Andorra      AND GDP per capita, PPP (current international $)
## 6 Arab World      ARB GDP per capita, PPP (current international $)
##      Indicator.Code 1960 1961 1962 1963 1964
## 1 NY.GDP.PCAP.PP.CD NA NA NA NA NA
## 2 NY.GDP.PCAP.PP.CD NA NA NA NA NA
## 3 NY.GDP.PCAP.PP.CD NA NA NA NA NA
## 4 NY.GDP.PCAP.PP.CD NA NA NA NA NA
## 5 NY.GDP.PCAP.PP.CD NA NA NA NA NA
## 6 NY.GDP.PCAP.PP.CD NA NA NA NA NA
```

Gagnaummyndun

Gögnunum er breytt, tvær nýjar töflur eru lesnar inn fyrir sitt hvora töfluna, þ.e. gdp og devhelp, með því að nota melt skipunina. Fyrir töfluna sem geymir verga landsframleiðslu eru töflurnar gdp_by_country_name og gdp_by_country_code búnar til. Báðar töflurnar innihalda þrjá dálka, bæði hafa þau Year, sem geymir árið, GDP_per_capita sem geymir verga landsframleiðslun m.v. höfðatölu. Auk þess inniheldur önnur þeirra Country_name sem geymir nafn landsins, en hin Country_id sem geymir auðkennisbókstafi landsins. Ástæðan fyrir þessari skiptingu er vegna þess að þegar Country.name er borið saman við nafn á sama landi í öðru landagagnasetti gætu nöfnin verið á mismunandi formi. T.d. er country name fyrir Rússland "Russian Federation" en í einhverju öðru gæti það verið "Russia". Country.id-ið myndi hins vegar mjög sennilega vera það sama. Sama skiptingin var síðan gerð fyrir devhelp töfluna nema að Development_help dálkurinn kom í staðinn fyrir GDP_per_capita. Eftir þessa síu eru gögnin komin á betra form fyrir frekari úrvinnslu gagnanna. Fyrir neðan kóðann eru allir dálkar og 6 fyrstu línur í öllum gagnasettunum.

```
gdp_by_country_code<-gdp
gdp_by_country_name <- gdp[ -c(2:34,64) ]
gdp_by_country_code <- gdp_by_country_code[ -c(1, 3:34) ]
gdp_by_country_name<-melt(gdp_by_country_name,"Country.Name")
gdp_by_country_code<-melt(gdp_by_country_code,"Country.Code")
colnames(gdp_by_country_name)<-c("Country_name", "Year", "GDP_per_capita")
gdp_by_country_name$Year<-as.integer(as.character(gdp_by_country_name$Year))
colnames(gdp_by_country_code)<-c("Country_id", "Year", "GDP_per_capita")
gdp_by_country_code$Year<-as.integer(as.character(gdp_by_country_code$Year))
```

```
devhelp_by_country_code<-devhelp
devhelp_by_country_name <- devhelp[ -c(2:34,64) ]
devhelp_by_country_code <- devhelp_by_country_code[ -c(1, 3:34) ]
devhelp_by_country_name<-melt(devhelp_by_country_name,"Country.Name")
devhelp_by_country_code<-melt(devhelp_by_country_code,"Country.Code")
colnames(devhelp_by_country_name)<-c("Country_name", "Year", "Development_help")
devhelp_by_country_name$Year<-as.integer(as.character(devhelp_by_country_name$Year))
colnames(devhelp_by_country_code)<-c("Country_id", "Year", "Development_help")
devhelp_by_country_code$Year<-as.integer(as.character(devhelp_by_country_code$Year))
```

```
head(gdp_by_country_name)
```

```
##      Country_name Year GDP_per_capita
## 1      Aruba 1990    24101.109
## 2 Afghanistan 1990         NA
## 3      Angola 1990    3089.683
## 4      Albania 1990    2549.473
## 5      Andorra 1990         NA
## 6 Arab World 1990    6806.042
```

```
head(gdp_by_country_code)
```

```
##      Country_id Year GDP_per_capita
## 1      ABW 1990    24101.109
## 2      AFG 1990         NA
## 3      AGO 1990    3089.683
## 4      ALB 1990    2549.473
## 5      AND 1990         NA
## 6      ARB 1990    6806.042
```

```
head(devhelp_by_country_name)
```

```
##      Country_name Year Development_help
## 1      Aruba 1990    29980000
## 2 Afghanistan 1990    121720000
## 3      Angola 1990    265770000
## 4      Albania 1990    11050000
## 5      Andorra 1990         NA
## 6 Arab World 1990    12762750000
```

```
head(devhelp_by_country_code)
```

```
##      Country_id Year Development_help
## 1      ABW 1990    29980000
## 2      AFG 1990    121720000
## 3      AGO 1990    265770000
## 4      ALB 1990    11050000
## 5      AND 1990         NA
## 6      ARB 1990    12762750000
```