

Analiza e Indikatorëve të Ekonomisë Globale

Era Dervishi

Hyrje

Analiza e treguesve ekonomikë globale është një fushë e studimit që përfshin përpunimin dhe interpretimin e të dhënave për të kuptuar performancën ekonomike të vendeve në nivel global. Kjo përfshin analizën e shifrave si: prodhimi i brendshëm bruto (PBB), inflacioni, tregti ndërkombëtare, shkëmbimi valutor, dhe tregues të tjerë që reflektojnë shëndetin e ekonomisë së një vendi. Analizat mund të përdorin mjete statistikore dhe matematikore për të zbuluar tendencat, lidhjet dhe shkaqet e ndryshimeve në treguesit ekonomikë. Këto analiza ndihmojnë ekonomistët, politikanët dhe investitorët të marrin vendime të informuara në lidhje me politikat ekonomike, investimet dhe zhvillimin ekonomik global.

Eksplorimi fillestar i të dhënave

```
# Importing libraries
```

```
library(readr)
```

```
## Warning: package 'readr' was built under R version 4.4.1
```

```
library(dplyr)
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
## intersect, setdiff, setequal, union
```

```
library(ggplot2)
```

```
library(janitor)
```

```
## Warning: package 'janitor' was built under R version 4.4.1
```

```
##
```

```
## Attaching package: 'janitor'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
## chisq.test, fisher.test
```

```
# Importing the dataset
```

```
dataset <- read_csv("C:/Users/User/Desktop/Global Economy Indicators.csv")
```

```
## Rows: 10512 Columns: 26
```

```
## — Column specification —————
```

```
## Delimiter: ","
```

```
## chr (2): Country, Currency
```

```
## dbl (24): CountryID, Year, AMA exchange rate, IMF based exchange rate, Popul...
```

```
##
```

```
## i Use `spec()` to retrieve the full column specification for this data.
```

```
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# Initial exploration of the data
glimpse(dataset)
```

```
## Rows: 10,512
## Columns: 26
## $ CountryID <dbl> ...
## $ Country <chr> ...
## $ Year <dbl> ...
## $ `AMA exchange rate` <dbl> ...
## $ `IMF based exchange rate` <dbl> ...
## $ Population <dbl> ...
## $ Currency <chr> ...
## $ `Per capita GNI` <dbl> ...
## $ `Agriculture, hunting, forestry, fishing (ISIC A-B)` <dbl> ...
## $ `Changes in inventories` <dbl> ...
## $ `Construction (ISIC F)` <dbl> ...
## $ `Exports of goods and services` <dbl> ...
## $ `Final consumption expenditure` <dbl> ...
## $ `General government final consumption expenditure` <dbl> ...
## $ `Gross capital formation` <dbl> ...
## $ `Gross fixed capital formation (including Acquisitions less disposals of valuables)` <dbl> ...
## $ `Household consumption expenditure (including Non-profit institutions serving households)` <dbl> ...
## $ `Imports of goods and services` <dbl> ...
## $ `Manufacturing (ISIC D)` <dbl> ...
## $ `Mining, Manufacturing, Utilities (ISIC C-E)` <dbl> ...
## $ `Other Activities (ISIC J-P)` <dbl> ...
## $ `Total Value Added` <dbl> ...
## $ `Transport, storage and communication (ISIC I)` <dbl> ...
## $ `Wholesale, retail trade, restaurants and hotels (ISIC G-H)` <dbl> ...
## $ `Gross National Income(GNI) in USD` <dbl> ...
## $ `Gross Domestic Product (GDP)` <dbl> ...
```

```
summary(dataset)
```

```
## CountryID Country Year AMA exchange rate
## Min. : 4.0 Length:10512 Min. :1970 Min. : 0.00
## 1st Qu.:212.0 Class :character 1st Qu.:1984 1st Qu.: 1.00
## Median :430.0 Mode :character Median :1997 Median : 2.81
## Mean :431.1 Mean :1996 Mean : 357.40
## 3rd Qu.:643.0 3rd Qu.:2009 3rd Qu.: 51.34
## Max. :894.0 Max. :2021 Max. :111636.63
##
## IMF based exchange rate Population Currency
## Min. : 0.00 Min. :4.359e+03 Length:10512
## 1st Qu.: 1.00 1st Qu.:6.331e+05 Class :character
## Median : 2.76 Median :5.052e+06 Mode :character
## Mean : 341.98 Mean :2.852e+07
## 3rd Qu.: 48.07 3rd Qu.:1.679e+07
## Max. :42000.00 Max. :1.426e+09
##
## Per capita GNI Agriculture, hunting, forestry, fishing (ISIC A-B)
## Min. : 34 Min. :2.814e+04
## 1st Qu.: 730 1st Qu.:1.305e+08
## Median : 2316 Median :9.314e+08
## Mean : 8966 Mean :7.793e+09
## 3rd Qu.: 8966 3rd Qu.:4.023e+09
## Max. :234317 Max. :1.350e+12
## NA's :121
## Changes in inventories Construction (ISIC F) Exports of goods and services
## Min. : -1.510e+11 Min. :1.577e+05 Min. : -1.670e+09
## 1st Qu.: 4.674e+05 1st Qu.:6.816e+07 1st Qu.: 4.027e+08
## Median : 6.830e+07 Median :3.718e+08 Median : 2.408e+09
## Mean : 1.627e+09 Mean :1.002e+10 Mean : 4.671e+10
## 3rd Qu.: 6.911e+08 3rd Qu.:2.816e+09 3rd Qu.: 1.542e+10
## Max. : 2.110e+11 Max. :1.240e+12 Max. : 3.530e+12
## NA's :1841 NA's :21
```

Pastrimi dhe Transformimi i të Dhënave

Pas eksplorimit fillestar, u krye pastrimi dhe transformimi i të dhënave për të siguruar që të dhënat janë të rregullta dhe të përshtatshme për analizë. Këtu u fokusua në largimin e rreshtave me vlera të munguara dhe përzgjedhjen e kolonave të rëndësishme për analizën tonë.

```
# Cleaning and Transforming the Data
cleaned_dataset <- dataset %>%
  clean_names() %>% # Cleaning column names using janitor
  filter(!is.na(population), !is.na(gross_national_income_gni_in_usd), !is.na(ama_exchange_rate)) %>%
  select(population, gross_national_income_gni_in_usd, ama_exchange_rate) %>%
  arrange(population)

# Viewing the cleaned data
glimpse(cleaned_dataset)
```

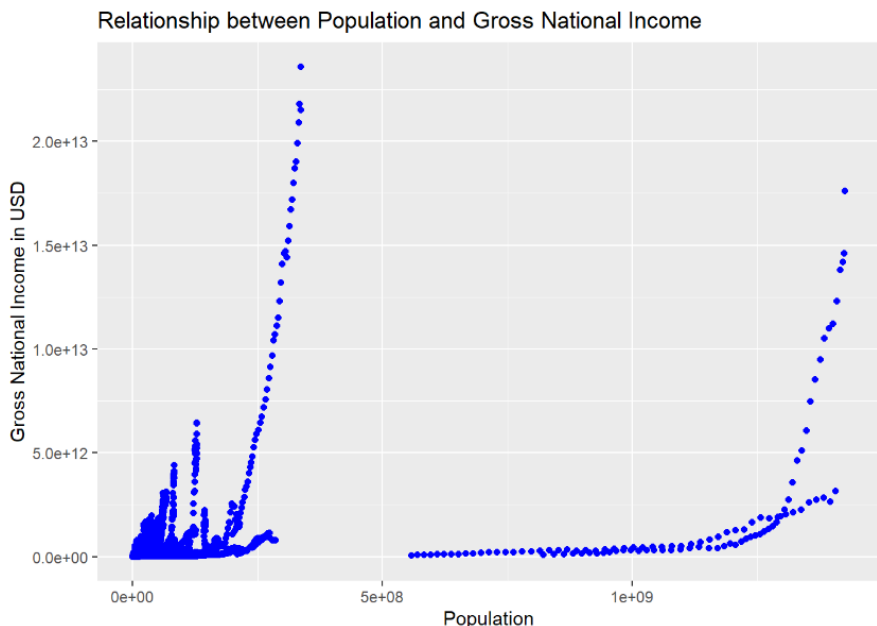
```
## Rows: 10,512
## Columns: 3
## $ population          <dbl> 4359, 4417, 4500, 4505, 4601, 4627, 4...
## $ gross_national_income_gni_in_usd <dbl> 65070370, 71572379, 67822222, 7010000...
## $ ama_exchange_rate    <dbl> 2.700000, 2.700000, 2.700000, 2.700000...
```

```
summary(cleaned_dataset)
```

##	population	gross_national_income_gni_in_usd	ama_exchange_rate
## Min.	:4.359e+03	Min. :3.565e+06	Min. : 0.00
## 1st Qu.	:6.331e+05	1st Qu.:1.411e+09	1st Qu.: 1.00
## Median	:5.052e+06	Median :7.887e+09	Median : 2.81
## Mean	:2.852e+07	Mean :1.825e+11	Mean : 357.40
## 3rd Qu.	:1.679e+07	3rd Qu.:5.006e+10	3rd Qu.: 51.34
## Max.	:1.426e+09	Max. :2.360e+13	Max. :111636.63

Vizualizimi i të Dhënave

```
# Scatter plot to show the relationship between population and gross national income
population_income_plot <- ggplot(cleaned_dataset, aes(x = population, y = gross_national_income_gni_in_usd)) +
  geom_point(color = "blue") +
  labs(title = "Relationship between Population and Gross National Income", x = "Population", y = "Gross National
Income in USD")
population_income_plot
```



Nga grafiku “Marrëdhënia midis Popullsisë dhe të Ardhurave Kombëtare Bruto”, mund të bëjmë disa vëzhgime dhe analiza:

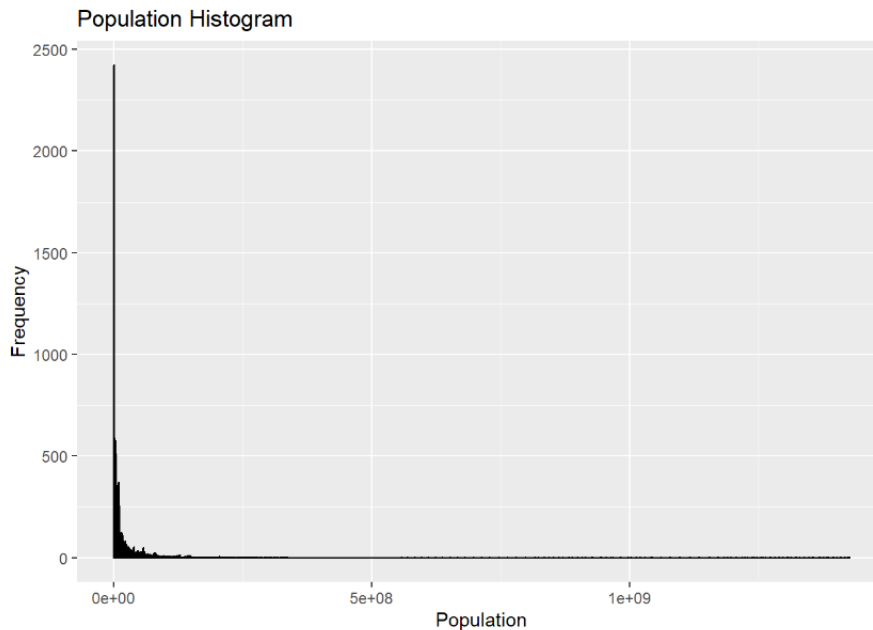
Trendi i Përgjithshëm: Ka një tendencë që të Ardhurat Kombëtare Bruto të rriten me rritjen e madhësisë së popullsisë. Kjo mund të tregojë se vendet me popullsi më të madhe kanë potencialin të prodhojnë më shumë të ardhura ekonomike.

Konsentrasi i Pikave të Të Dhënave: Shumica e pikave janë të grumbulluara afër origjinës, që do të thotë se shumë vende kanë popullsi të vogla dhe GNI të ulët. Kjo mund të tregojë se vendet me popullsi të vogla përballen me sfida në arritjen e niveleve të larta të të ardhurave.

Rritja Eksponenciale: Duke u larguar nga origjina dhe duke u lëvizur djathtas në boshtin horizontal (që tregon rritje në popullsi), vërehet se GNI rritet në një mënyrë që duket eksponenciale, jo lineare. Kjo mund të tregojë se pas kalimit të një prag të caktuar popullsie, kapaciteti për të gjeneruar të ardhura kombëtare bruto rritet në mënyrë të shpejtë.

Variabiliteti: Edhe pse ka një trend të përgjithshëm rritës, ka një variabilitet të madh midis vendeve me popullsi të ngjashme. Kjo mund të tregojë se faktorë të tjerë përveç madhësisë së popullsisë ndikojnë në GNI.

```
# Histogram to show the distribution of population
population_histogram <- ggplot(cleaned_dataset, aes(x = population)) +
  geom_histogram(binwidth = 1000000, fill = "skyblue", color = "black") +
  labs(title = "Population Histogram", x = "Population", y = "Frequency")
population_histogram
```



Grafiku i histogramit të popullsisë paraqet shpërndarjen e madhësive të popullsisë në një dataset.

Shumica e shiritave janë të përqendruar afër boshtit y, që tregon se shumica e vlerave të popullsisë në këtë dataset janë të ulëta. Ka një shirit shumë të lartë në fillim që pastaj zvogëlohet shpejt; kjo sugjeron se ka shumë raste të popullsisë të vogla dhe pak raste të popullsisë të mëdha brenda këtij dataseti.

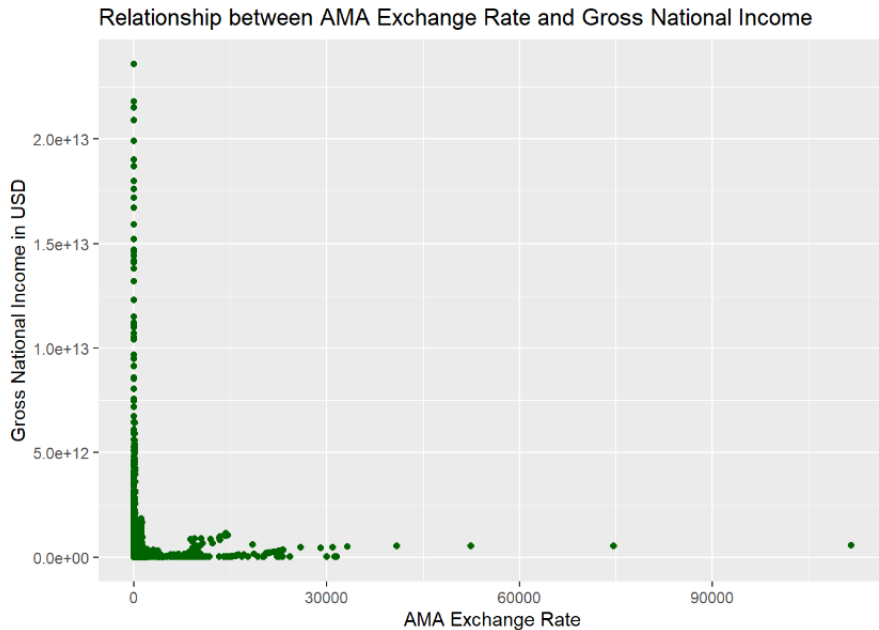
Disa pikat kyçe të grafikut:

Shiriti më i lartë: Shiriti i parë dhe më i lartë tregon një frekuencë prej rreth **2500**, që do të thotë se ka një numër të madh të vlerave të popullsisë që janë shumë të ulëta, ndoshta nën **5 milion**.

Zvogëlimi i shpejtë: Pas shiritit të parë, ka një zvogëlim të shpejtë në lartësinë e shiritave, që tregon se ka shumë më pak raste të popullsisë të mëdha.

Popullsia maksimale: Nuk ka shumë shiritave që përfaqësojnë popullsi mbi **500 milion**, që sugjeron se rastet e popullsisë të mëdha janë të rralla në këtë dataset.

```
# Scatter plot to show the relationship between AMA exchange rate and gross national income
exchange_gni_plot <- ggplot(cleaned_dataset, aes(x = ama_exchange_rate, y = gross_national_income_gni_in_usd)) +
  geom_point(color = "darkgreen") +
  labs(title = "Relationship between AMA Exchange Rate and Gross National Income", x = "AMA Exchange Rate", y = "Gross National Income in USD")
exchange_gni_plot
```



Nga grafiku “Marrëdhënia mes Kursit të Këmbimit të AMA dhe të Ardhurave Kombëtare Bruto”, mund të bëjmë disa vëzhgime dhe analiza:

Dendësia e Pikave të Të Dhënave: Shumica e pikave janë të grumbulluara pranë origjinës, që tregon se për vlerat më të ulëta të kursit të këmbimit të AMA, të ardhurat kombëtare bruto janë gjithashtu relativisht të ulëta. Kjo mund të sugjerojë se ekonomitë me një kurs të këmbimit më të ulët të AMA-së kanë tendencë të kenë një PBB më të ulët.

Shpërndarja e Pikave: Ka një numër të vogël të pikave që janë të shpërndara drejt vlerave më të larta të kursit të këmbimit dhe të PBB-së. Kjo mund të tregojë se vetëm disa ekonomi arrijnë nivele të larta të të ardhurave kombëtare bruto, ndërsa shumica mbeten me të ardhura më të ulëta.

Mungesa e një Trendi të Qartë: Nuk duket të ketë një trend të qartë ose një marrëdhënie lineare midis kursit të këmbimit të AMA dhe të ardhurave kombëtare bruto. Kjo mund të tregojë se faktorë të tjerë ndikojnë në PBB përveç kursit të këmbimit të AMA.

Vlerat Anomale: Nëse ka ndonjë pikë që shfaqet larg grupit kryesor të të dhënave, ato mund të konsiderohen si vlera anomale. Këto vlera anomale mund të përfaqësojnë ekonomi që kanë performancë të jashtëzakonshme ose rrethana unike që i bëjnë ato të dallohen nga të tjerat.

Interpretimi i Vlerave të Larta: Për vlerat më të larta të kursit të këmbimit dhe të PBB-së, do të ishte e dobishme të kuptonim kontekstin më të gjerë ekonomik dhe politik që mund të ndikojë në këto vlera.

Analiza Statistike

Pas vizualizimit të të dhënave, u krye një analizë statistikore për të zbuluar korrelacionin midis popullsisë dhe të ardhurave kombëtare bruto, si dhe për të kryer një model regresioni linear për të parashikuar të ardhurat kombëtare bruto bazuar në popullsi.

```
# Correlation between population and gross national income
correlation <- cor(cleaned_dataset$population, cleaned_dataset$gross_national_income_gni_in_usd)
print(correlation)
```

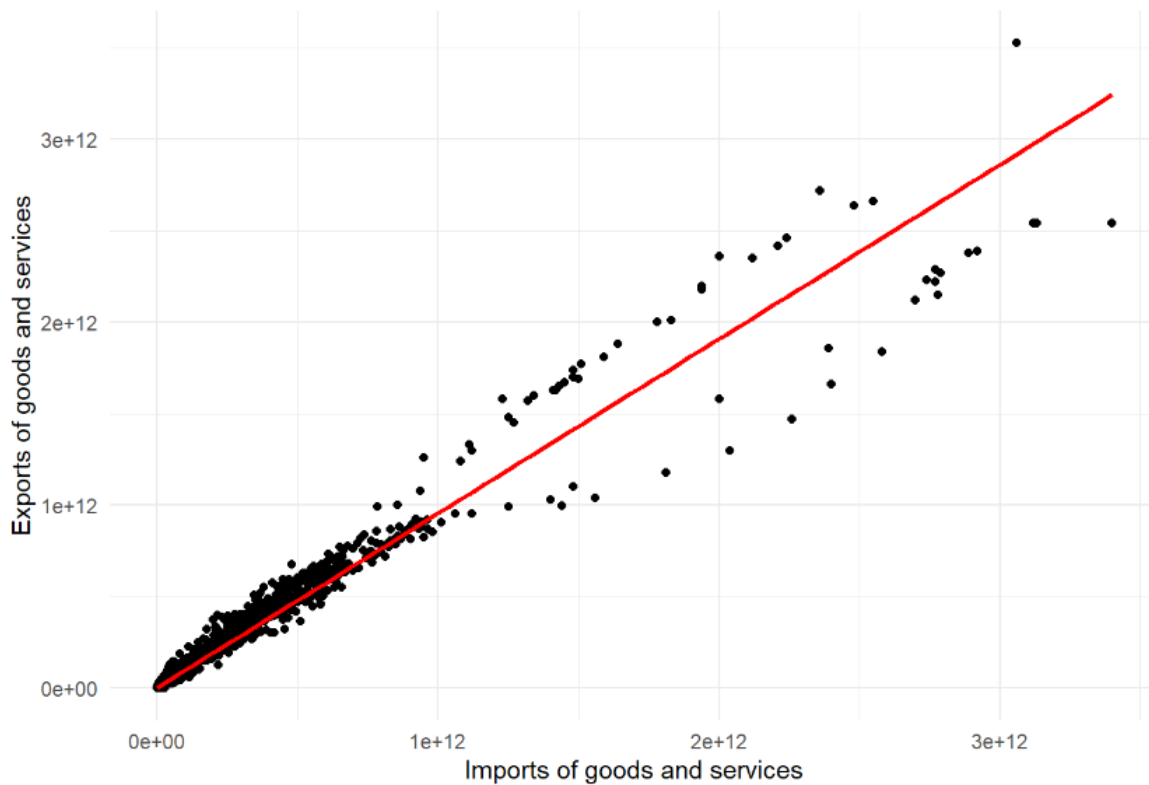
```
## [1] 0.3945702
```

Vlera e korrelacionit është 0.3945702, e cila është pozitive dhe afër 0.4. Kjo tregon një lidhje të dobët, por pozitive midis popullsisë dhe të ardhurave kombëtare bruto. Pra, kur popullsia e një vendi rritet, ka tendencë që të ardhurat kombëtare bruto të rriten gjithashtu, megjithatë, lidhja nuk është shumë e fortë.

```
column1 <- "Imports of goods and services"
column2 <- "Exports of goods and services"
correlation_df <- data.frame(x = dataset[[column1]], y = dataset[[column2]])

# Krijë një grafik të shfaqjes së korelacionit
ggplot(correlation_df, aes(x = x, y = y)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  labs(title = paste("Korelacioni midis", column1, "dhe", column2),
       x = column1, y = column2) +
  theme_minimal()
```

Korelacioni midis Imports of goods and services dhe Exports of goods and services



```

column1 <- "Gross Domestic Product (GDP)"
column2 <- "Population"

# Llogaritja e korelacionit midis kolonave të përzgjedhura
correlation_matrix <- cor(dataset[[column1]], dataset[[column2]])

# Konverto matricën e korelacionit në një data frame
correlation_df <- data.frame(col1 = column1, col2 = column2, corr = correlation_matrix)

# Krijo një heatmap për të paraqitur korelacionin
ggplot(correlation_df, aes(x = col1, y = col2, fill = corr)) +
  geom_tile(color = "white") +
  scale_fill_gradient(low = "green", high = "darkblue") +
  labs(title = paste("Heatmap i Korelacionit midis", column1, "dhe", column2),
       x = column1, y = column2) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

```

Heatmap i Korelacionit midis Gross Domestic Product (GDP) dhe Population



```

# Linear Regression
linear_model <- lm(gross_national_income_gni_in_usd ~ population, data = cleaned_dataset)
summary(linear_model)

```

```

##
## Call:
## lm(formula = gross_national_income_gni_in_usd ~ population, data = cleaned_dataset)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.694e+12 -1.043e+11 -8.502e+10 -7.892e+10  2.236e+13
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  8.451e+10  9.185e+09   9.202  <2e-16 ***
## population   3.437e+03  7.808e+01  44.022  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.136e+11 on 10510 degrees of freedom
## Multiple R-squared:  0.1557, Adjusted R-squared:  0.1556
## F-statistic: 1938 on 1 and 10510 DF, p-value: < 2.2e-16

```

Analiza e rezultateve të modelit regresionit linear tregon disa informacione të rëndësishme lidhur me marrëdhënien midis popullsisë dhe të ardhurave kombëtare bruto (Gross National Income - GNI):

Intercept dhe Koeficienti i Popullsisë: - Intercept tregon vlerën e pritshme të GNI kur popullsia është zero. Në këtë rast, intercept është 8.451e+10, që tregon vlerën fillestare të pritshme të GNI. -

Koeficienti për popullsinë është $3.437e+03$, që do të thotë se për çdo njësi rritje në popullsi, pritshmëria e GNI rritet me $3.437e+03$.

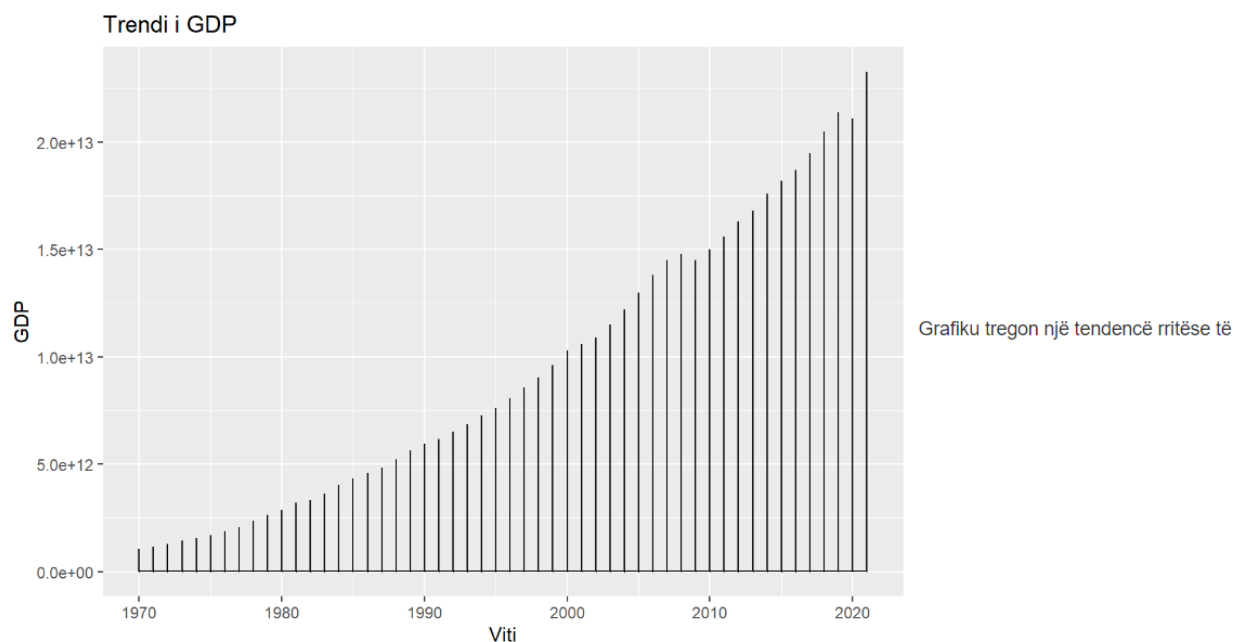
Humbjet e Mbetura (Residuals): - Humbjet e mbetura përshkruajnë ndryshimet mes vlerës reale të GNI dhe vlerës së parashikuar nga modeli. Në rastin tonë, humbjet e mbetura variojnë nga $-3.694e+12$ deri në $2.236e+13$, tregon që modeli mund të ketë humbje të mëdha për disa vlera të dhënash.

Gabimi Standard i Mbetur (Residual Standard Error): - Një gabim standard i ulët tregon që modeli ka performuar mirë në përdorimin e të dhënave për të bërë parashikime. Në këtë rast, gabimi standard i mbetur është $9.136e+11$, që tregon një shkallë të konsiderueshme të variacionit në vlerat e pritshme të GNI.

R-katrori i Shumëfishtë dhe R-katrori i Rregulluar: - R-katrori i shumëfishtë dhe i rregulluar tregon se sa mirë modeli ka përputhje me të dhënat. Në rastin tonë, R-katrori i shumëfishtë është 0.1557 dhe R-katrori i rregulluar është 0.1556. Kjo tregon se vetëm rreth 15.57% e variabilitetit të GNI është shpjeguar nga variabla e popullsisë në model.

F-Statistika: - F-Statistika jep informacion në lidhje me rëndësinë e modelit në përgjithësi. Një p-vlerë shumë e ulët ($p\text{-value} < 0.05$) tregon se modeli është i rëndësishëm statistikisht. Në rastin tonë, p-vlera është shumë e vogël ($p\text{-value} < 2.2e-16$), që tregon se modeli është i rëndësishëm.

```
# Linear Regression
library(ggplot2)
ggplot(data = dataset, aes(x = Year, y = `Gross Domestic Product (GDP)`)) +
  geom_line() +
  labs(title = "Trendi i GDP", x = "Viti", y = "GDP")
```

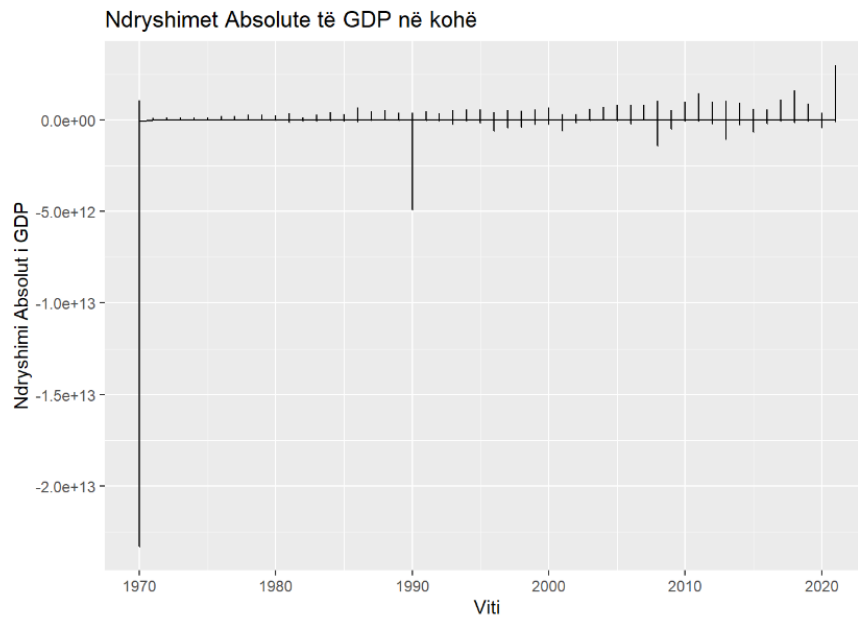


GDP-së me disa fluktuacione. Ky grafik është interesant sepse vizualisht paraqet zhvillimin ekonomik gjatë pesë dekadave.

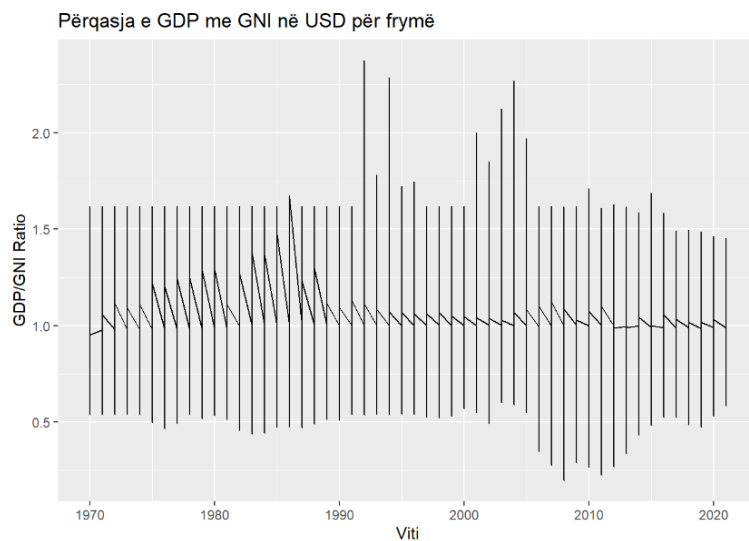
```
dataset$GDP_Absolute_Change <- c(NA,diff(dataset$`Gross Domestic Product (GDP)`))

# Përqasja e GDP me GNI në USD për frymë
dataset$GDP_to_GNI_Ratio <- dataset$`Gross Domestic Product (GDP)` / dataset$`Gross National Income(GNI) in USD`

# Vizualizimi i trendit të ndryshimeve absolute të GDP
ggplot(data = dataset, aes(x = Year, y = GDP_Absolute_Change)) +
  geom_line() +
  labs(title = "Ndryshimet Absolute të GDP në kohë", x = "Viti", y = "Ndryshimi Absolut i GDP")
```



```
# Vizualizimi i përqasjes së GDP me GNI në USD për frymë
ggplot(data = dataset, aes(x = Year, y = GDP_to_GNI_Ratio)) +
  geom_line() +
  labs(title = "Përqasja e GDP me GNI në USD për frymë", x = "Viti", y = "GDP/GNI Ratio")
```



Shifrat e GDP shkojnë nga 0e+00 deri në 2.0e+13. Grafiku tregon një tendencë rritëse të GDP-së me disa fluktuacione. Ky grafik është interesant sepse vizualisht paraqet zhvillimin ekonomik gjatë pesë dekadave.