

R analysis for BRI dataset

Steve

2025-03-16

```
# -----  
# Step 1: Setup Environment  
# -----
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.4.3
```

```
## Warning: package 'ggplot2' was built under R version 4.4.3
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
```

```
## v dplyr      1.1.4      v readr      2.1.5
```

```
## v forcats   1.0.0      v stringr   1.5.1
```

```
## v ggplot2    3.5.1      v tibble    3.2.1
```

```
## v lubridate  1.9.3      v tidyr     1.3.1
```

```
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(plm)
```

```
## Warning: package 'plm' was built under R version 4.4.3
```

```
##
```

```
## Attaching package: 'plm'
```

```
##
```

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

```
##
```

```
##      between, lag, lead
```

```
library(stargazer)
```

```
##
```

```
## Please cite as:
```

```
##
```

```
## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
library(ggplot2)
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 4.4.3
```

```
## corrplot 0.95 loaded
```

```
# -----
# Step 2: Load and Inspect Data
# -----
library(tidyverse)
bri_data <- read_csv("C:/Users/HomePC/Downloads/bri_dataset.csv")
```

```
## Rows: 70 Columns: 11
## -- Column specification -----
## Delimiter: ","
## chr (1): Country
## dbl (10): Year, GDP_Growth, Total_FDI, Renewable_FDI, Renewable_Share, Solar...
##
## 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.
```

```
# Basic inspection
glimpse(bri_data)
```

```
## Rows: 70
## Columns: 11
## $ Country      <chr> "Ethiopia", "Ethiopia", "Ethiopia", "Ethiopia~
## $ Year         <dbl> 2010, 2011, 2012, 2013, 2014, 2015, 2016, 201~
## $ GDP_Growth   <dbl> 5.0, 4.0, 4.2, 5.0, 4.7, 4.0, 4.0, 4.9, 5.6, ~
## $ Total_FDI    <dbl> 541.2, 504.2, 529.4, 605.9, 722.8, 652.8, 760~
## $ Renewable_FDI <dbl> 131.9, 174.4, 202.5, 171.2, 258.7, 252.4, 230~
## $ Renewable_Share <dbl> 24.4, 34.6, 38.3, 28.3, 35.8, 38.7, 30.3, 35.~
## $ Solar_Capacity <dbl> 16.5, 21.5, 25.0, 28.9, 37.0, 49.2, 55.9, 75.~
## $ Wind_Capacity <dbl> 46.8, 52.5, 63.2, 76.0, 88.6, 103.3, 120.6, 1~
## $ Total_Renewable_Capacity <dbl> 63.3, 74.0, 88.2, 104.9, 125.6, 152.5, 176.5,~
## $ Gov_Score    <dbl> 0.42, 0.42, 0.42, 0.42, 0.42, 0.42, 0.42, 0.4~
## $ Regulatory_Quality <dbl> 0.39, 0.39, 0.39, 0.39, 0.39, 0.39, 0.39, 0.3~
```

```
summary(bri_data)
```

```
##      Country      Year      GDP_Growth      Total_FDI
## Length:70      Min.   :2010      Min.   :4.000      Min.   : 400.4
## Class :character 1st Qu.:2013      1st Qu.:4.800      1st Qu.: 552.4
## Mode  :character Median :2016      Median :5.200      Median : 705.5
##                      Mean   :2016      Mean   :5.393      Mean   : 728.0
##                      3rd Qu.:2020      3rd Qu.:5.900      3rd Qu.: 881.0
##                      Max.   :2023      Max.   :7.700      Max.   :1246.3
## Renewable_FDI Renewable_Share Solar_Capacity Wind_Capacity
## Min.   : 81.8      Min.   :20.40      Min.   : 8.80      Min.   : 5.00
## 1st Qu.:152.8      1st Qu.:23.88      1st Qu.: 34.52      1st Qu.: 24.05
```

```
## Median :207.0   Median :29.55   Median : 76.10   Median : 47.45
## Mean    :213.7   Mean    :29.41   Mean    :111.00   Mean    : 76.99
## 3rd Qu.:267.6   3rd Qu.:34.50   3rd Qu.:149.62   3rd Qu.: 95.67
## Max.    :406.1   Max.    :39.60   Max.    :455.90   Max.    :388.50
## Total_Renewable_Capacity   Gov_Score   Regulatory_Quality
## Min.      : 13.80           Min.      :0.420   Min.      :0.390
## 1st Qu.: 65.47           1st Qu.:0.470   1st Qu.:0.430
## Median :127.80           Median :0.530   Median :0.480
## Mean    :187.99           Mean    :0.524   Mean    :0.476
## 3rd Qu.:257.40           3rd Qu.:0.580   3rd Qu.:0.510
## Max.    :666.00           Max.    :0.620   Max.    :0.570
```

```
sum(is.na(bri_data)) # Check for missing values
```

```
## [1] 0
```

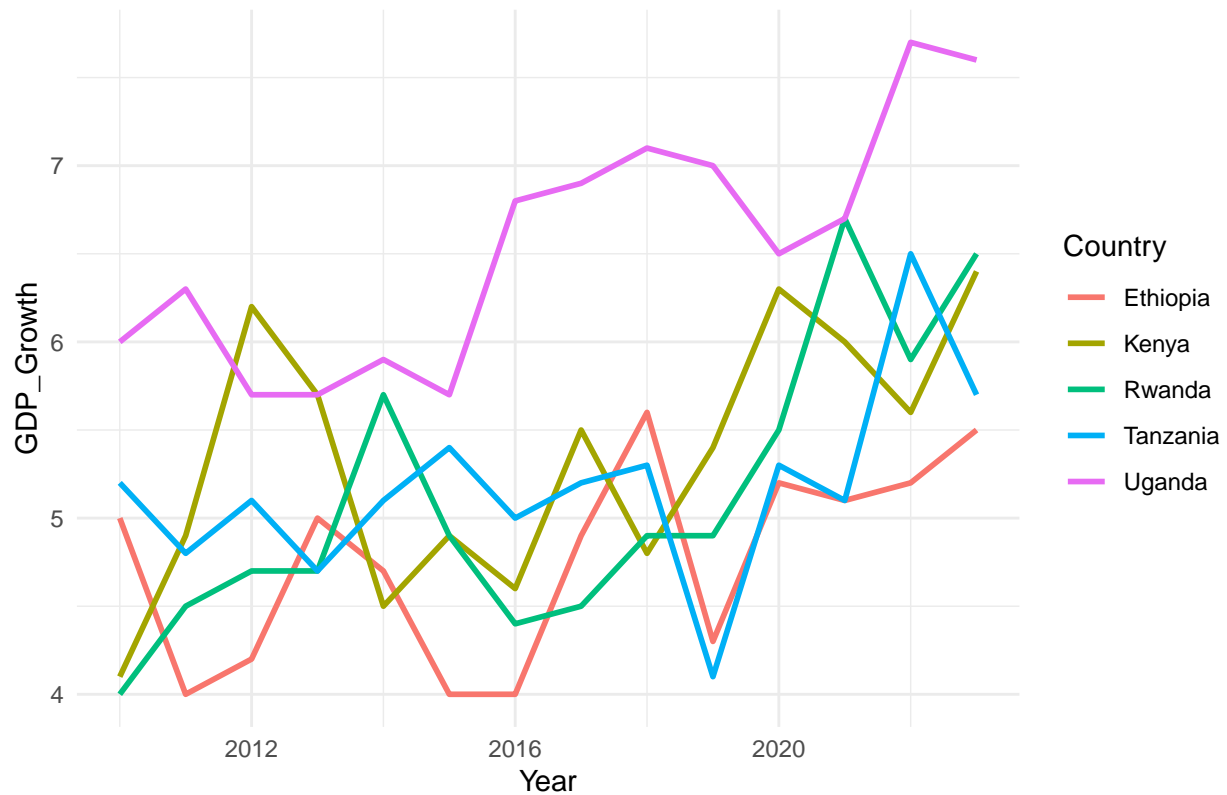
```
# -----
# Step 3: Descriptive Analysis
# -----
# Summary statistics by country
bri_data %>%
  group_by(Country) %>%
  summarise(
    Avg_GDP = mean(GDP_Growth),
    Avg_Renew_FDI = mean(Renewable_FDI),
    Avg_Capacity = mean(Total_Renewable_Capacity)
  )
```

```
## # A tibble: 5 x 4
##   Country Avg_GDP Avg_Renew_FDI Avg_Capacity
##   <chr>    <dbl>    <dbl>    <dbl>
## 1 Ethiopia  4.76      248.     259.
## 2 Kenya   5.35      226.     236.
## 3 Rwanda    5.13      191.     82.4
## 4 Tanzania  5.18      198.     213.
## 5 Uganda    6.54      207.     150.
```

```
# Time trends visualization
ggplot(bri_data, aes(x = Year, y = GDP_Growth, color = Country)) +
  geom_line(size = 1) +
  labs(title = "GDP Growth Trends by Country") +
  theme_minimal()
```

```
## Warning: Using 'size' aesthetic for lines was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.
```

GDP Growth Trends by Country

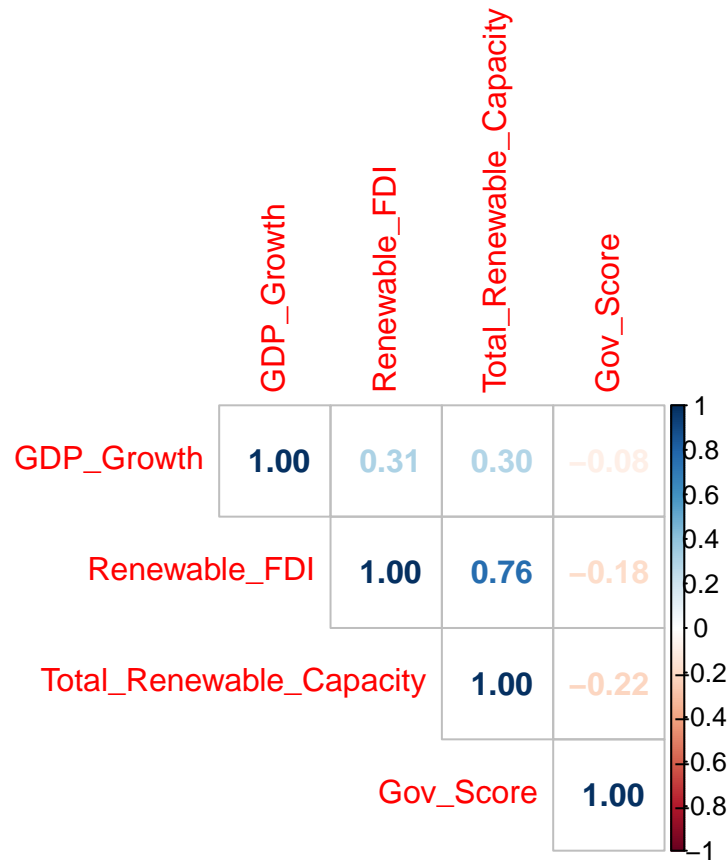


```
# -----
# Step 4: Correlation Analysis
# -----
# 1. Install the package (if not already installed)
install.packages("corrplot")

## Warning: package 'corrplot' is in use and will not be installed

# 2. Load the package
library(corrplot)

# 3. Now run your correlation plot code
cor_matrix <- cor(bri_data %>% select(GDP_Growth, Renewable_FDI, Total_Renewable_Capacity, Gov_Score))
corrplot(cor_matrix, method = "number", type = "upper")
```



```
# -----
# Step 5: Panel Data Regression
# -----
library(plm)

# Convert to panel data format
pdata <- pdata.frame(bri_data, index = c("Country", "Year"))

# Fixed Effects Model
fe_model <- plm(GDP_Growth ~ Renewable_FDI + Total_Renewable_Capacity + Gov_Score,
  data = pdata,
  model = "within")

# Random Effects Model
re_model <- plm(GDP_Growth ~ Renewable_FDI + Total_Renewable_Capacity + Gov_Score,
  data = pdata,
  model = "random")

# Hausman Test
h_test <- phtest(fe_model, re_model)
print(h_test)

##
## Hausman Test
##
## data: GDP_Growth ~ Renewable_FDI + Total_Renewable_Capacity + Gov_Score
```

```
## chisq = 0.083727, df = 2, p-value = 0.959
## alternative hypothesis: one model is inconsistent
```

```
# -----
# Step 6: Model Diagnostics (Fixed)
# -----
# 1. Load required packages
library(tidyverse) # For %>% operator and arrange()
library(lmtest)    # For bptest/bgtest (NOT "Intest")
```

```
## Warning: package 'lmtest' was built under R version 4.4.3
```

```
## Loading required package: zoo
```

```
##
```

```
## Attaching package: 'zoo'
```

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

```
##
```

```
##      as.Date, as.Date.numeric
```

```
# 2. Fix dataframe name and sort data
```

```
bri_data <- bri_data %>% # Original name was "bri_data"
  arrange(Country, Year) # Case-sensitive column names
```

```
# 3. Fix variable names (case-sensitive)
```

```
fe_model_lm <- lm(
  GDP_Growth ~ Renewable_FDI + Total_Renewable_Capacity + Gov_Score + factor(Country),
  data = bri_data
)
```

```
# 4. Correct function names
```

```
bptest(fe_model_lm) # Not "bytes"
```

```
##
```

```
## studentized Breusch-Pagan test
```

```
##
```

```
## data: fe_model_lm
```

```
## BP = 4.9614, df = 6, p-value = 0.5488
```

```
bgtest(fe_model_lm) # Not "bytes"
```

```
##
```

```
## Breusch-Godfrey test for serial correlation of order up to 1
```

```
##
```

```
## data: fe_model_lm
```

```
## LM test = 1.8506, df = 1, p-value = 0.1737
```

```
# -----
```

```
# Step 7: Results Interpretation (Fixed)
```

```
# -----
```

```
# 1. Install and load package
```

```
install.packages("stargazer") # Run once
```

```
## Warning: package 'stargazer' is in use and will not be installed
```

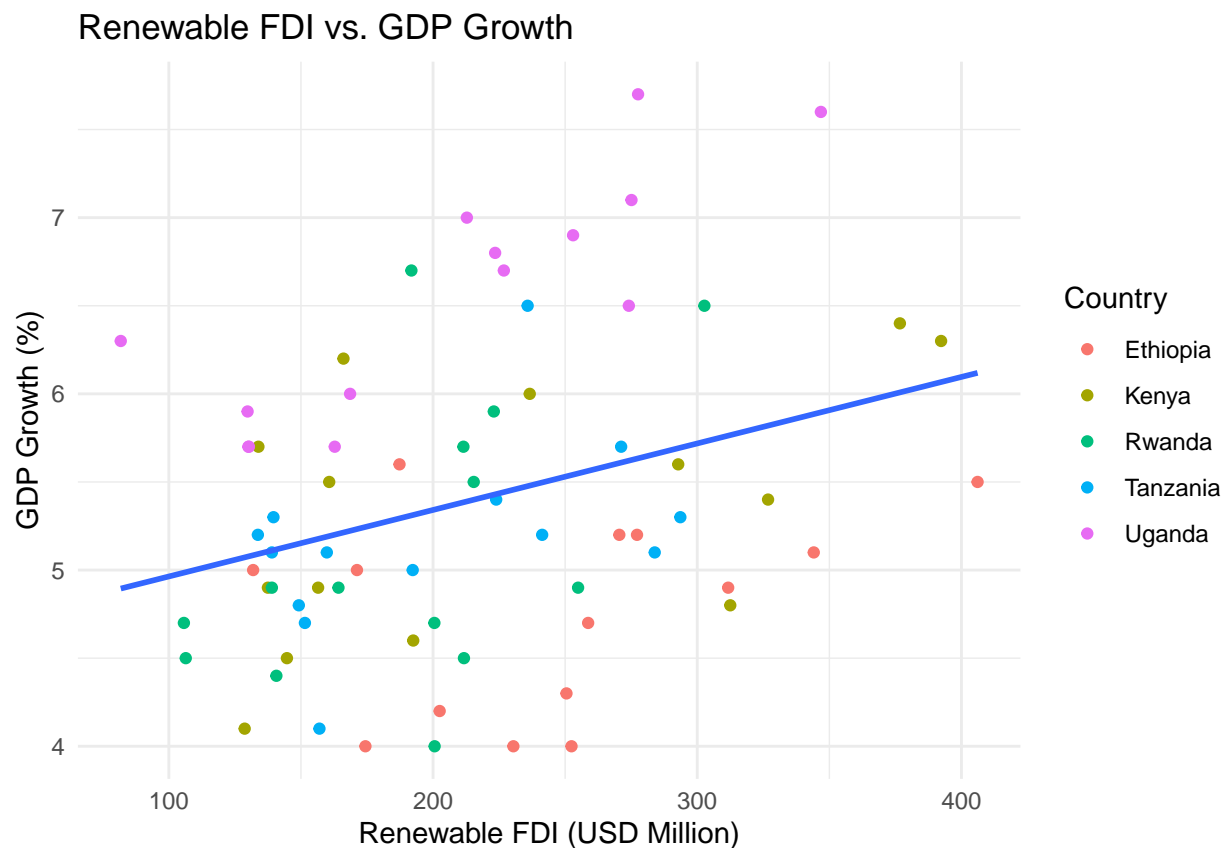
```
library(stargazer)           # Required for stargazer()

# 2. Use correct model objects
stargazer(fe_model_lm,       # Use the lm model for compatibility
           type = "text",
           title = "Fixed Effects Regression Results",
           dep.var.labels = "GDP Growth",
           covariate.labels = c("Renewable FDI",
                                "Renewable Capacity",
                                "Governance Score"))
```

```
##
## Fixed Effects Regression Results
## =====
##                               Dependent variable:
##                               -----
##                               GDP Growth
## -----
## Renewable FDI                0.002
##                               (0.001)
##
## Renewable Capacity           0.002***
##                               (0.001)
##
## Governance Score             41.272***
##                               (4.226)
##
## factor(Country)Kenya        -5.933***
##                               (0.601)
##
## factor(Country)Rwanda       -7.446***
##                               (0.751)
##
## factor(Country)Tanzania     -3.943***
##                               (0.404)
##
## factor(Country)Uganda
##
## Constant                    -13.535***
##                               (1.921)
## -----
## Observations                70
## R2                          0.661
## Adjusted R2                 0.629
## Residual Std. Error         0.544 (df = 63)
## F Statistic                 20.474*** (df = 6; 63)
## =====
## Note:                        *p<0.1; **p<0.05; ***p<0.01
```

```
# -----
# Step 8: Visualization of Key Findings
# -----
# Marginal effects plot
ggplot(bri_data, aes(x = Renewable_FDI, y = GDP_Growth)) +
  geom_point(aes(color = Country)) +
  geom_smooth(method = "lm", se = FALSE) +
  labs(title = "Renewable FDI vs. GDP Growth",
       x = "Renewable FDI (USD Million)",
       y = "GDP Growth (%)") +
  theme_minimal()
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```



```
# Capacity vs GDP growth
ggplot(bri_data, aes(x = Total_Renewable_Capacity, y = GDP_Growth)) +
  geom_point(aes(size = Renewable_FDI, color = Country)) +
  geom_smooth(method = "loess") +
  labs(title = "Renewable Capacity vs. GDP Growth",
       x = "Total Renewable Capacity (MW)",
       y = "GDP Growth (%)") +
  theme_minimal()
```

```
## 'geom_smooth()' using formula = 'y ~ x'
```


Renewable Capacity vs. GDP Growth

