GDP Correlation Analysis

Lisa Blackmer-Raynolds

2025-06-04

Table of contents

The goal of this project is to examine the correlation between GDP per capita and three different variables: adult literacy rates, life expectancy, and measles immunisation rate. First I will run the correlation analysis for each variable. The data comes from the World Bank's World Development Indicators (**worldbank2022?**). Code is created with the help of GitCopilot (**copilot?**).

```
setwd("/Users/Lisa/Desktop/QTM/assignment5")
library(readr)
library(ggplot2)
```

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```
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
```

```r
wdi <- read_csv("wdi.csv")
```

```
Rows: 217 Columns: 14
```

```
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (1): country
dbl (13): inflation_rate, exports_gdp_share, gdp_growth_rate, gdp_per_capita...

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.
```

```r
# Filter for rows with both gdp_per_capita and adult_literacy_rate
wdi_filtered <- subset(wdi, !is.na(gdp_per_capita) & !is.na(adult_literacy_rate))

# GDP and Adult Literacy Correlation
cor_literacy <- cor.test(
  wdi$gdp_per_capita, wdi$adult_literacy_rate,
  use = "complete.obs", method = "pearson"
)

# Filter for rows with both gdp_per_capita and Life Expectancy
wdi_filtered <- subset(wdi, !is.na(gdp_per_capita) & !is.na(life_expectancy))

# GDP and Life Expectancy
cor_life <- cor.test(
  wdi$gdp_per_capita, wdi$life_expectancy,
  use = "complete.obs", method = "pearson"
)

# Remove rows with missing GDP per capita or  measles immunisation rate data
wdi_filtered <- subset(wdi, !is.na(gdp_per_capita) & !is.na(measles_immunisation_rate))

# GDP and Measles Immunisation Rate
cor_immunisation <- cor.test(
  wdi$gdp_per_capita, wdi$measles_immunisation_rate,
  use = "complete.obs", method = "pearson"
)

# Create summary table
cor_table <- data.frame(
```

```
  Comparison = c(
    "GDP per Capita vs. Adult Literacy Rate",
    "GDP per Capita vs. Life Expectancy",
    "GDP per Capita vs. Measles Immunisation Rate"
  ),
  Correlation = c(
    round(cor_literacy$estimate, 3),
    round(cor_life$estimate, 3),
    round(cor_immunisation$estimate, 3)
  ),
  P_Value = c(
    signif(cor_literacy$p.value, 3),
    signif(cor_life$p.value, 3),
    signif(cor_immunisation$p.value, 3)
  )
)

# Display the table
knitr::kable(cor_table, caption = "Summary of Correlation Analyses")
```

Table 1: Summary of Correlation Analyses

Comparison	Correlation	P_Value
GDP per Capita vs. Adult Literacy Rate	0.425	0.002870
GDP per Capita vs. Life Expectancy	0.639	0.000000
GDP per Capita vs. Measles Immunisation Rate	0.314	0.000013

0.0.1 Interpretation

All three variables are positivly correlated with GDP. Of all the variables, life expectancy has the strongest and most significant correlation with GDP (Figure 1). Adult literacy has the least significant, but second strongest correlation (Figure 2). Measles immunisation rate has the weakest correlation with GDP (Figure 3).

0.0.2 Figure 1

```
Warning: Removed 3 rows containing missing values or values outside the scale range
(`geom_point()`).
```

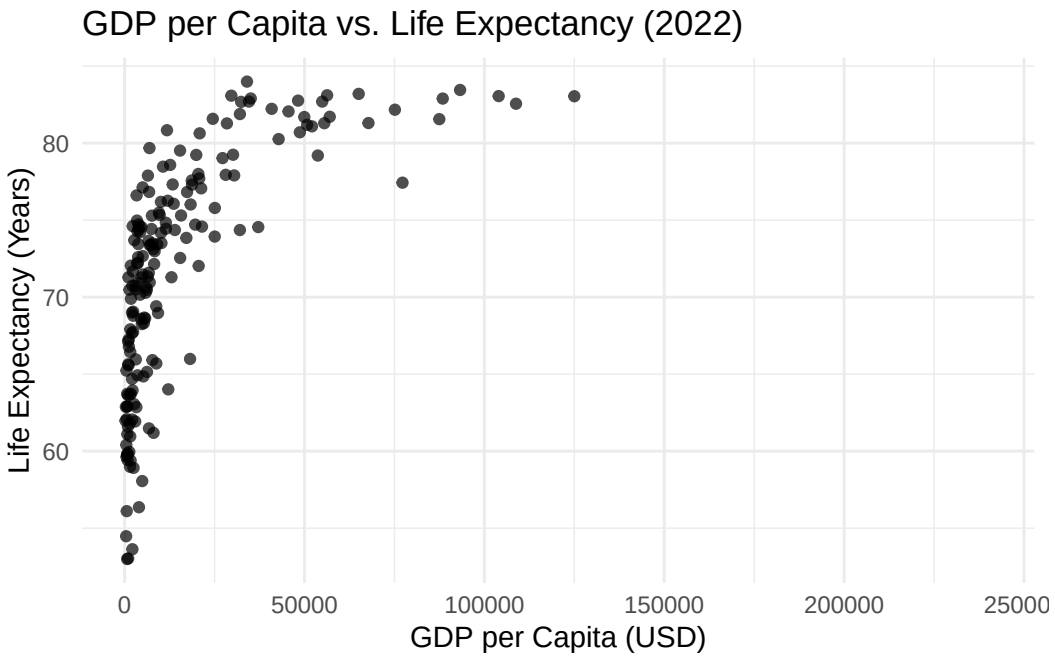

Figure 1: Scatter plot of GDP per Capita vs. Life Expectancy

0.0.3 Figure 2

```
Warning: Removed 139 rows containing missing values or values outside the scale range
(`geom_point()`).
```

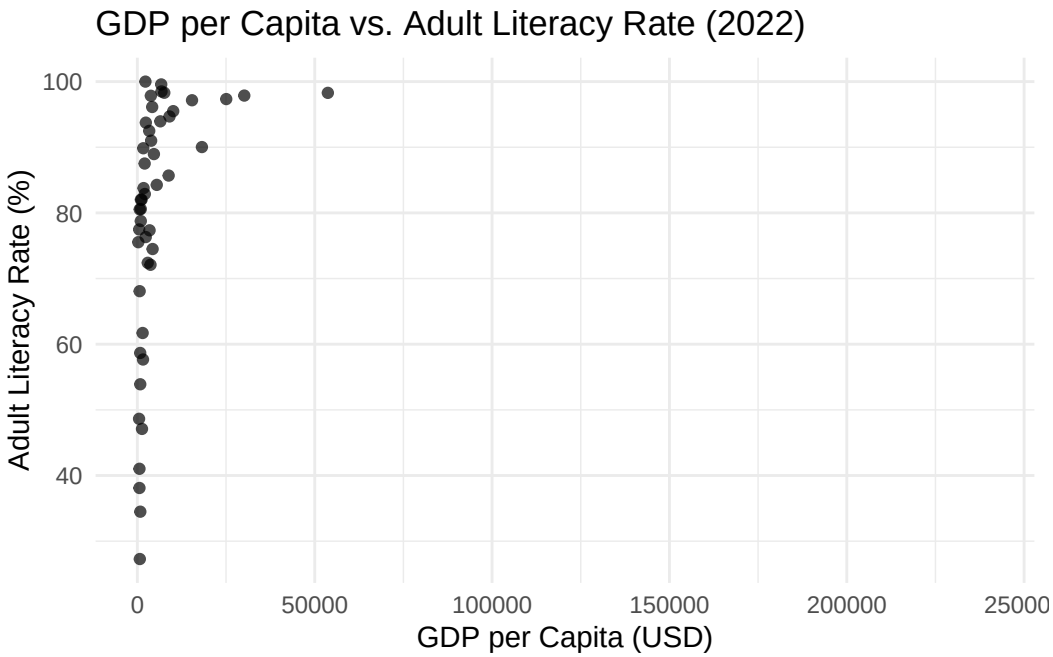

Figure 2: Scatter plot of GDP per Capita vs. Adult Literacy Rate

0.0.4 Figure 3

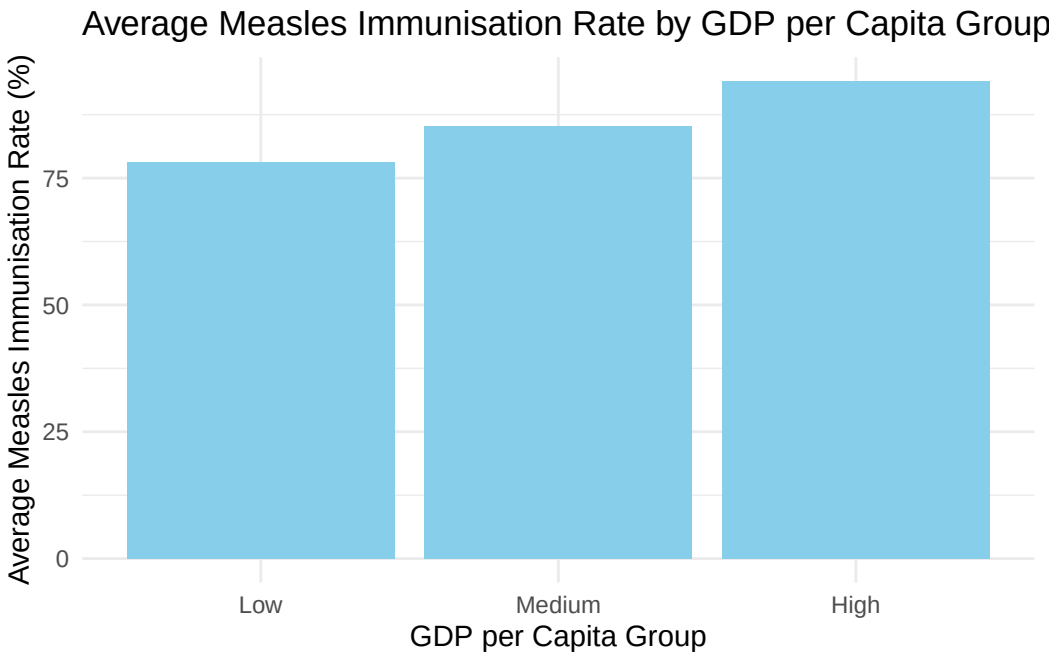

Figure 3: Bar plot of GDP per Capita and Measles Immunisation Rate