World Bank Data Analysis

Eric Zou

2025-02-21

Load Dataset

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np

# Load the dataset
df = pd.read_csv("wdi.csv")

# Display the first few rows
print(df.head())
```

```
          country  inflation_rate  exports_gdp_share  gdp_growth_rate  \
0     Afghanistan             NaN          18.380042        -6.240172
1         Albania        6.725203          37.197085         4.826688
2         Algeria        9.265516          30.808979         3.600000
3  American Samoa             NaN          46.957520         1.735016
4         Andorra             NaN                NaN         9.564612

   gdp_per_capita  adult_literacy_rate  primary_school_enrolment_rate  \
0      357.261153                  NaN                            NaN
1     6846.426143                 98.5                      96.371231
2     4961.552577                  NaN                     108.343933
3    18017.458938                  NaN                            NaN
4    42414.059009                  NaN                      90.147346

   education_expenditure_gdp_share  measles_immunisation_rate  \
0                              NaN                       56.0
```

```
1                          2.744330                       86.0
2                          4.749247                       79.0
3                               NaN                        NaN
4                          2.647290                       98.0

   health_expenditure_gdp_share  income_inequality  unemployment_rate  \
0                           NaN                NaN             14.100
1                           NaN                NaN             10.137
2                           NaN                NaN             12.346
3                           NaN                NaN                NaN
4                           NaN                NaN                NaN

   life_expectancy  total_population
0           62.879        40578842.0
1           76.833         2777689.0
2           77.129        45477389.0
3              NaN           48342.0
4              NaN           79705.0
```

Exploratory Data Analysis

GDP per Capita Distribution

Figure shows the distribution of GDP per capita across different countries.

```
plt.hist(df["gdp_per_capita"].dropna(), bins=30, color="blue", alpha=0.7)
plt.title("GDP per Capita Distribution")
plt.xlabel("GDP per Capita")
plt.ylabel("Count")
plt.show()
```

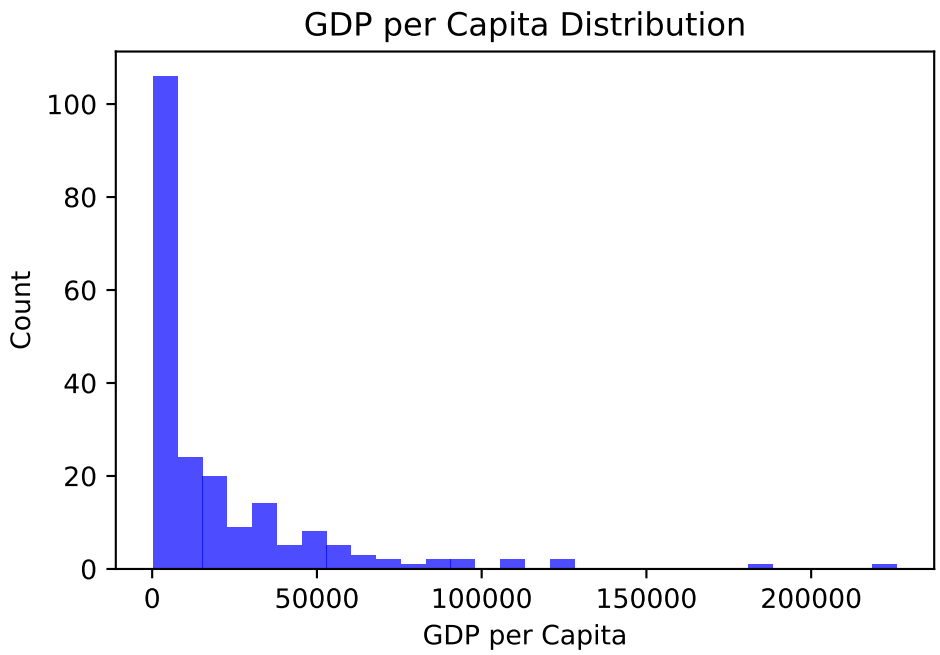

Figure 1: A histogram showing the distribution of GDP per capita across countries. The data highlights significant disparities between nations.

Finding: most of the countries have less than 100000 GDP per captia.

Inflation Rate vs GDP Growth Rate

Figure illustrates the relationship between inflation rates and GDP growth rates.

```
plt.scatter(df["inflation_rate"], df["gdp_growth_rate"], color="red", alpha=0.6)
plt.title("Inflation Rate vs GDP Growth Rate")
plt.xlabel("Inflation Rate (%)")
plt.ylabel("GDP Growth Rate (%)")
plt.show()
```

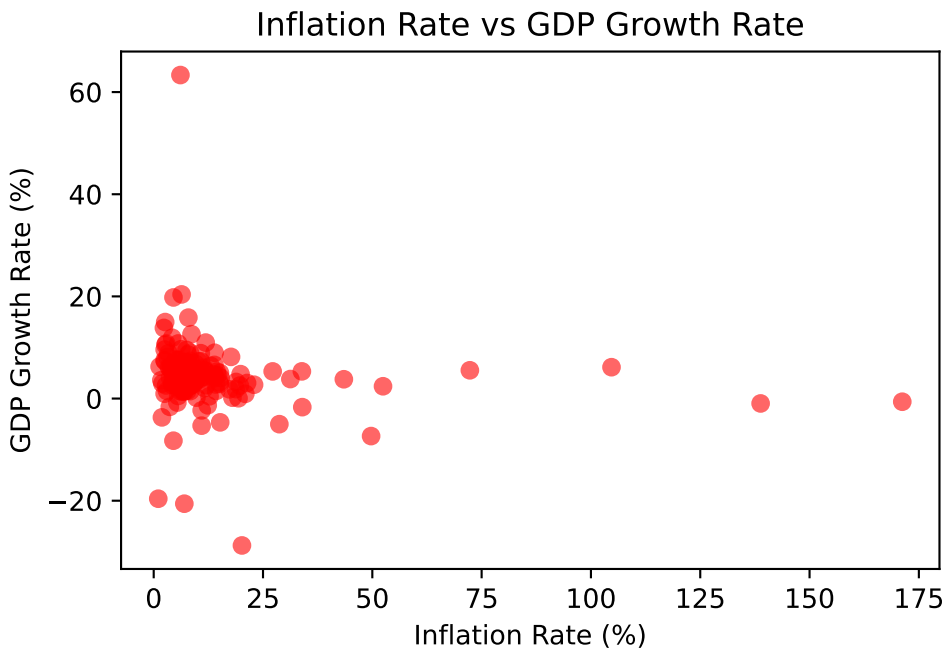

Figure 2: A scatter plot displaying the relationship between inflation rates and GDP growth rates.

Finding: Some countries experience high inflation with negative GDP growth, while other countries have high inflation rate with positive GDP growth together.

Health Expenditure vs Life Expectancy

Figure presents a scatter plot of health expenditure as a percentage of GDP and its correlation with life expectancy.

```
plt.scatter(df["health_expenditure_gdp_share"], df["life_expectancy"], color="green", alpha=
plt.title("Health Expenditure vs Life Expectancy")
plt.xlabel("Health Expenditure (% of GDP)")
plt.ylabel("Life Expectancy (Years)")
plt.show()
```

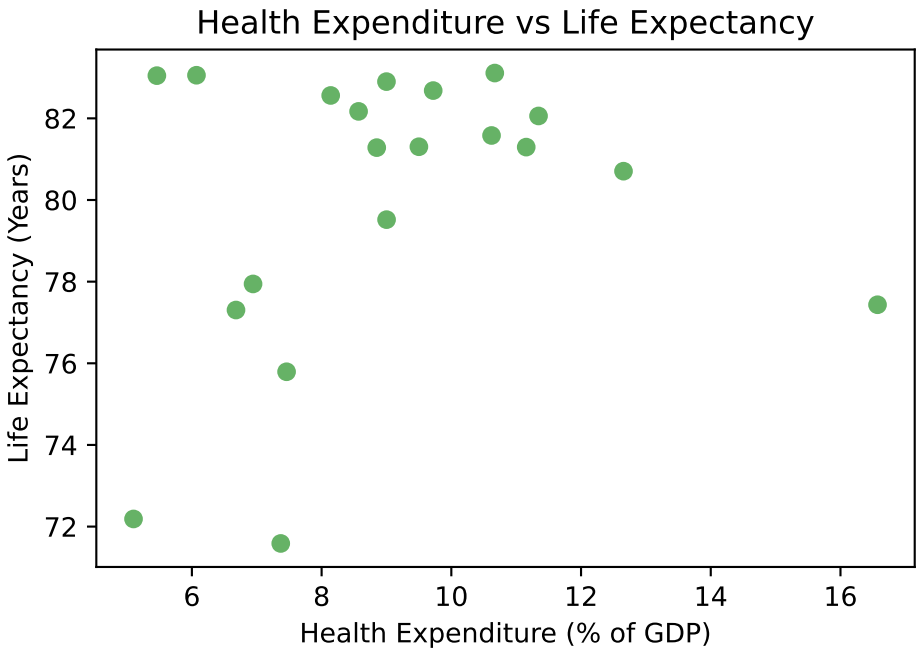

Figure 3: A scatter chart showing how life expectancy varies with health expenditure as a percentage of GDP.

Finding: More health expediture, more life expectency people have.

Summary of Findings

This part include the cross-reference to my figures above.

1. **GDP per Capita Distribution:** The distribution of GDP per capita varies widely among countries, with some extreme outliers, as shown in Figure Figure 1.
2. **Inflation Rate vs GDP Growth Rate:** There appears to be a mixed relationship between inflation and GDP growth, with no clear trend across all countries, as depicted in Figure Figure 2.
3. **Health Expenditure vs Life Expectancy:** Countries that spend more on healthcare as a percentage of GDP tend to have higher life expectancies, indicating a positive correlation (Figure Figure 3).

Bar Chart of Health Expenditure vs Life Expectancy

Figure highlights the distribution of health expenditure across randomly selected five countries.

```
sample_countries = df.dropna(subset=["health_expenditure_gdp_share"]).sample(5)

plt.figure(figsize=(12, 6))
sns.barplot(x=sample_countries["country"], y=sample_countries["health_expenditure_gdp_share"]
plt.title("Health Expenditure vs Country (Random 5)")
plt.xlabel("Country")
plt.ylabel("Health Expenditure (% of GDP)")
plt.xticks(rotation=45)
plt.show()
```

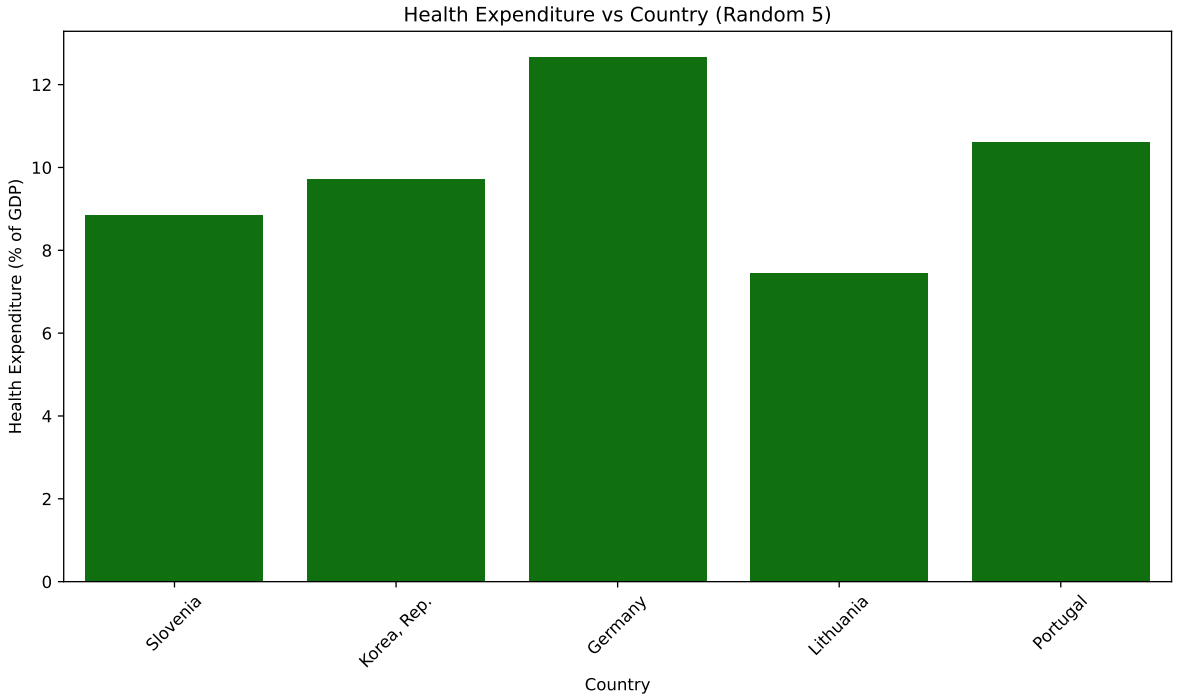

Figure 4: Bar chart representing health expenditure by country.

Source: Data retrieved from the World Bank Indicators dataset (Bank 2025).

Unemployment Rate vs GDP per Capita

Figure illustrates how unemployment rates relate to GDP per capita.

This graph should hide the code and only show the graph.

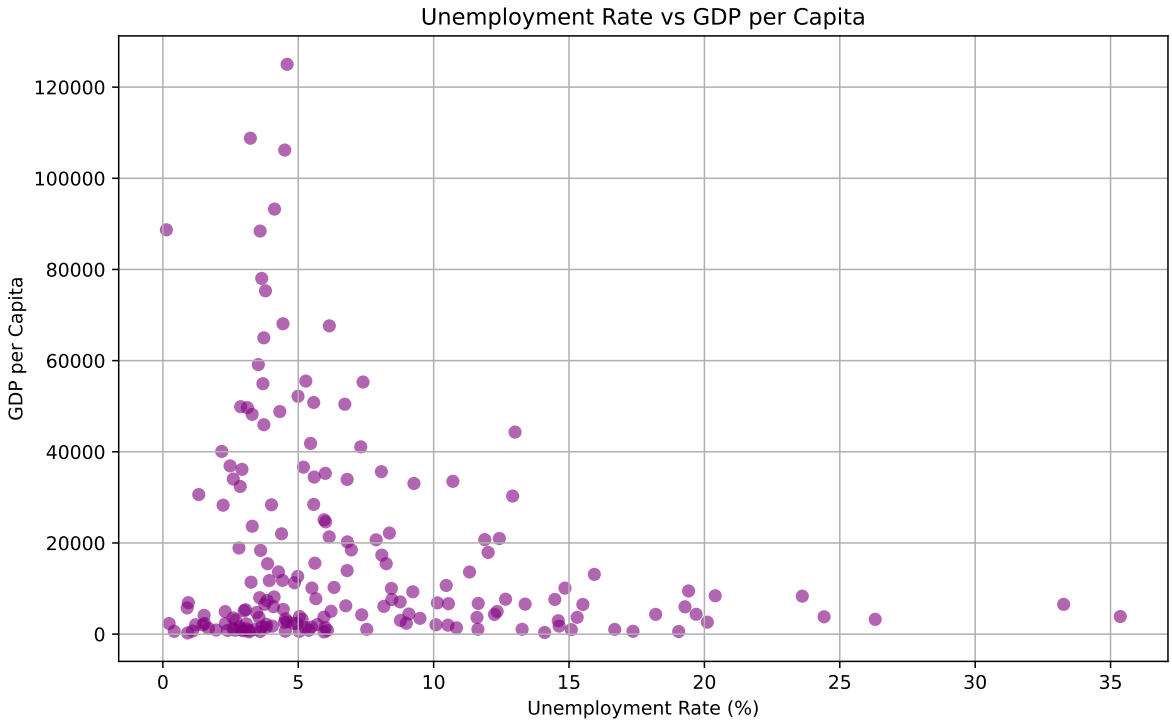

Figure 5: A scatter plot illustrating the relationship between unemployment rate and GDP per capita. Countries with lower unemployment tend to have higher GDP per capita.

Key Statistics Table

Data (Bank 2025) provides key insight for key economic indicators analyzed in this report.

Statistic	GDP per Capita	Inflation Rate	GDP Growth Rate	Unempl
Mean	20520.34	12.40	4.39	7.23
Median	7606.24	7.93	4.21	5.33
Standard Deviation	30640.74	19.47	6.71	5.84
Minimum	250.63	-6.69	-28.76	0.13
Maximum	226052.00	171.21	63.33	35.36

Source: Data retrieved from the World Bank Indicators dataset (Bank 2025; Smith 2020; Johnson 2018).

Bank, World. 2025. *World Development Indicators*. The World Bank. https://databank.worldbank.org/source/world-development-indicators.

Johnson, Emily. 2018. *Healthcare Spending and Longevity*. Oxford, UK: Oxford University Press.

Smith, John. 2020. "The Impact of Inflation on Economic Growth." *Journal of Economic Studies* 45 (3): 123–45. https://doi.org/10.1000/jec.2020.0045.