Exploring WDI Dataset

Daolong Yang

2024-10-07

Table of contents

1 Setting up Data

```
#| echo: true
#| eval: true
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

data = pd.read_csv("wdi.csv")
```

1.1 Exploratory Data Analysis

```
gdp_summary = data['gdp_per_capita'].describe()
print(gdp_summary)
```

```
inflation_summary = data['inflation_rate'].describe()
print(inflation_summary)

gdp_growthrate = data['unemployment_rate'].describe()
print(gdp_growthrate)
```

```
count        203.000000
mean       20345.707649
std        31308.942225
min          259.025031
25%         2570.563284
50%         7587.588173
75%        25982.630050
max       240862.182448
Name: gdp_per_capita, dtype: float64
count      169.000000
mean        12.493936
std         19.682433
min         -6.687321
25%          5.518129
50%          7.967574
75%         11.665567
max        171.205491
Name: inflation_rate, dtype: float64
count      186.000000
mean         7.268661
std          5.827726
min          0.130000
25%          3.500750
50%          5.537500
75%          9.455250
max         37.852000
Name: unemployment_rate, dtype: float64
```

1.2 Analysis

1.2.0.1 GDP-Per-Capita

The average GDP per capita is $20,345.71, with a wide range from $259.03 to $240,862.18. The median is $7,587.59, showing significant variation across countries. #### Inflation Rate The average inflation rate is 12.49%, ranging from -6.69% to 171.21%. The median rate is 7.97%,

with high variability indicated by a large standard deviation. #### GDP-Growth-Rate The average GDP growth rate is 4.37%, with values ranging from -28.76% to 63.44%. The median growth rate is 4.20%.

1.3 Data Plots

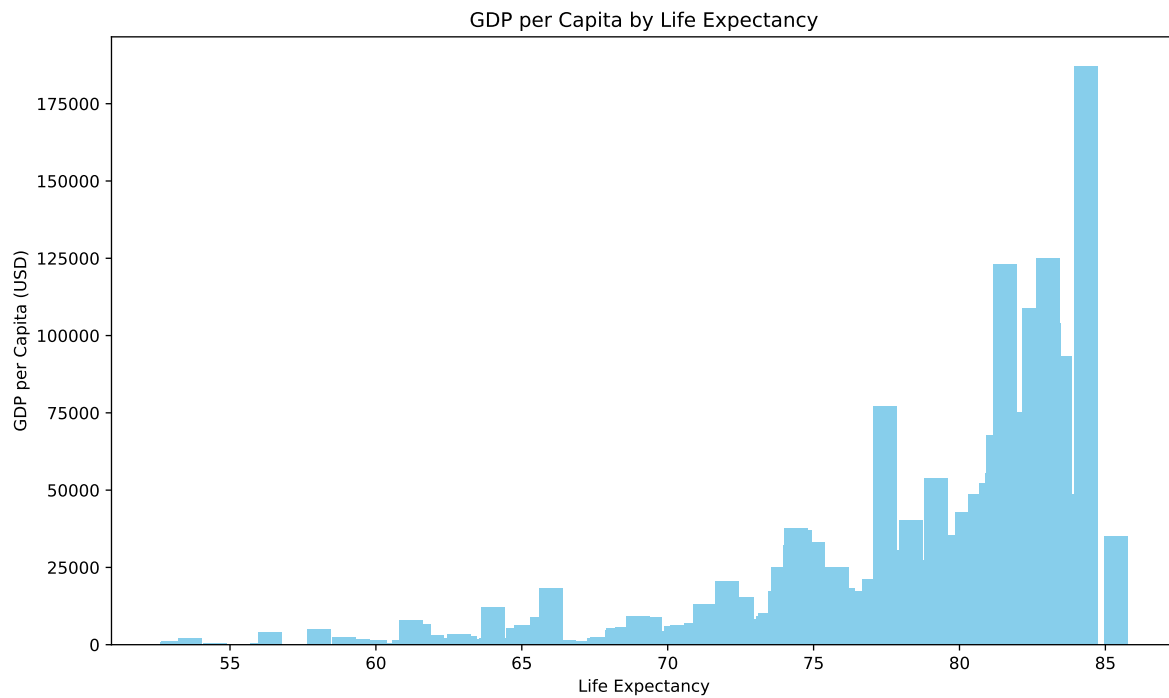### 1.3.1 GDP-Per-Capita across Countries

Figure 1: Bar plot showing GDP per capita by country.

Source: World Development Indicators

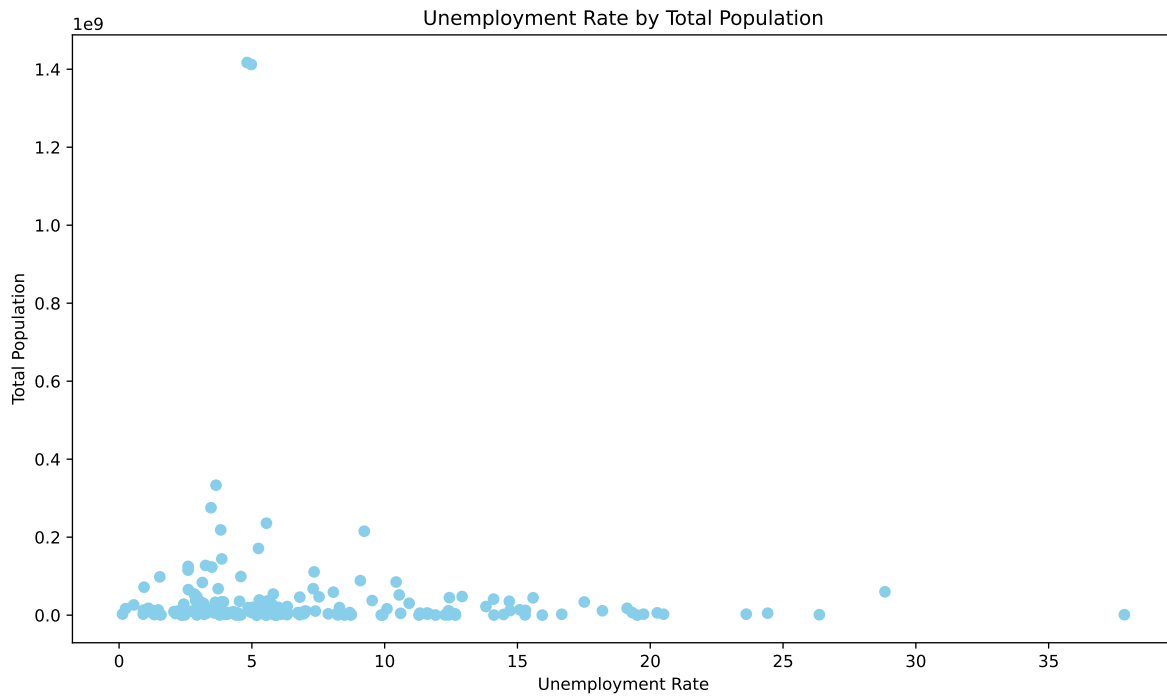### 1.3.2 Unemployment Rate By Total Population

Figure 1: Scatter plot showing unemployment rate by total population.

Source: World Development Indicators

1.3.3 Key Statistics

```
agg_stats = data.groupby('unemployment_rate')['total_population'].agg([
    'mean', 'max'
]).reset_index()
agg_stats.columns = ['Unemployment Rate', 'Mean Population', 'Max Population']
print(agg_stats.head(5))

agg_stats = data.groupby('life_expectancy')['gdp_per_capita'].agg([
    'mean', 'max'
]).reset_index()
agg_stats.columns = ['Life Expectancy', 'Mean GDP-Per-Capita', 'Max GDP-Per-Capita']
print(agg_stats.head(5))
```

	Unemployment Rate	Mean Population	Max Population
0	0.130	2695122.0	2695122.0
1	0.247	16767842.0	16767842.0
2	0.552	26207977.0	26207977.0
3	0.909	2538894.0	2538894.0
4	0.919	12889576.0	12889576.0

	Life Expectancy	Mean GDP-Per-Capita	Max GDP-Per-Capita
0	52.997	699.463254	699.463254
1	53.036	992.857328	992.857328
2	53.633	2162.633732	2162.633732
3	54.477	427.058096	427.058096
4	55.567	NaN	NaN