WDI Data Analysis

Alix Morales

2024-10-08

Loading Packages and Data for Analysis

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

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

Running Analysis on the Code

```
# exploratory analysis
print(wdi.describe())
# focusing on adult_literacy_rate
# income_inequality
# gdp_per_capita
```

```
       inflation_rate  exports_gdp_share  gdp_growth_rate  gdp_per_capita  \
count          169.00             169.00           202.00          203.00
mean           12.49               46.17             4.37        20345.71
std            19.68               34.00             6.63        31308.94
min            -6.69                1.57           -28.76          259.03
25%             5.52               24.53             2.44         2570.56
50%             7.97               40.22             4.20         7587.59
75%            11.67               55.46             6.20        25982.63
```

```
max            171.21             211.28            63.44       240862.18

       adult_literacy_rate  primary_school_enrolment_rate  \
count                49.00                         114.00
mean                 79.57                         100.87
std                  19.38                          12.04
min                  27.28                          64.40
25%                  72.40                          94.19
50%                  83.78                         100.02
75%                  95.50                         105.04
max                 100.00                         138.19

       education_expenditure_gdp_share  measles_immunisation_rate  \
count                           105.00                     193.00
mean                              4.23                      83.85
std                               2.07                      16.00
min                               1.03                       0.00
25%                               2.90                      76.00
50%                               3.89                      90.00
75%                               5.16                      96.00
max                              16.58                      99.00

       health_expenditure_gdp_share  income_inequality  unemployment_rate  \
count                         20.00              28.00             186.00
mean                           9.04              38.33               7.27
std                            2.70               7.72               5.83
min                            5.10              26.40               0.13
25%                            7.26              32.90               3.50
50%                            8.93              38.10               5.54
75%                           10.63              43.12               9.46
max                           16.57              54.80              37.85

       life_expectancy  total_population
count           209.00            217.00
mean             72.42       36536447.72
std               7.71      141058261.99
min              53.00          11312.00
25%              66.78         808726.00
50%              73.51        6465097.00
75%              78.47       26069416.00
max              85.38     1417173173.00
```

The GDP per capita variable is very interesting as it shows the massive disparities in wealth

between countries, providing context for a lot of the other variables (min: 259, max: 240862). Adult literacy rate is interesting since there are only 20 entries and a lot of missing data. Likewise, income inequality is missing a significant amount of data.

Plotting the Data

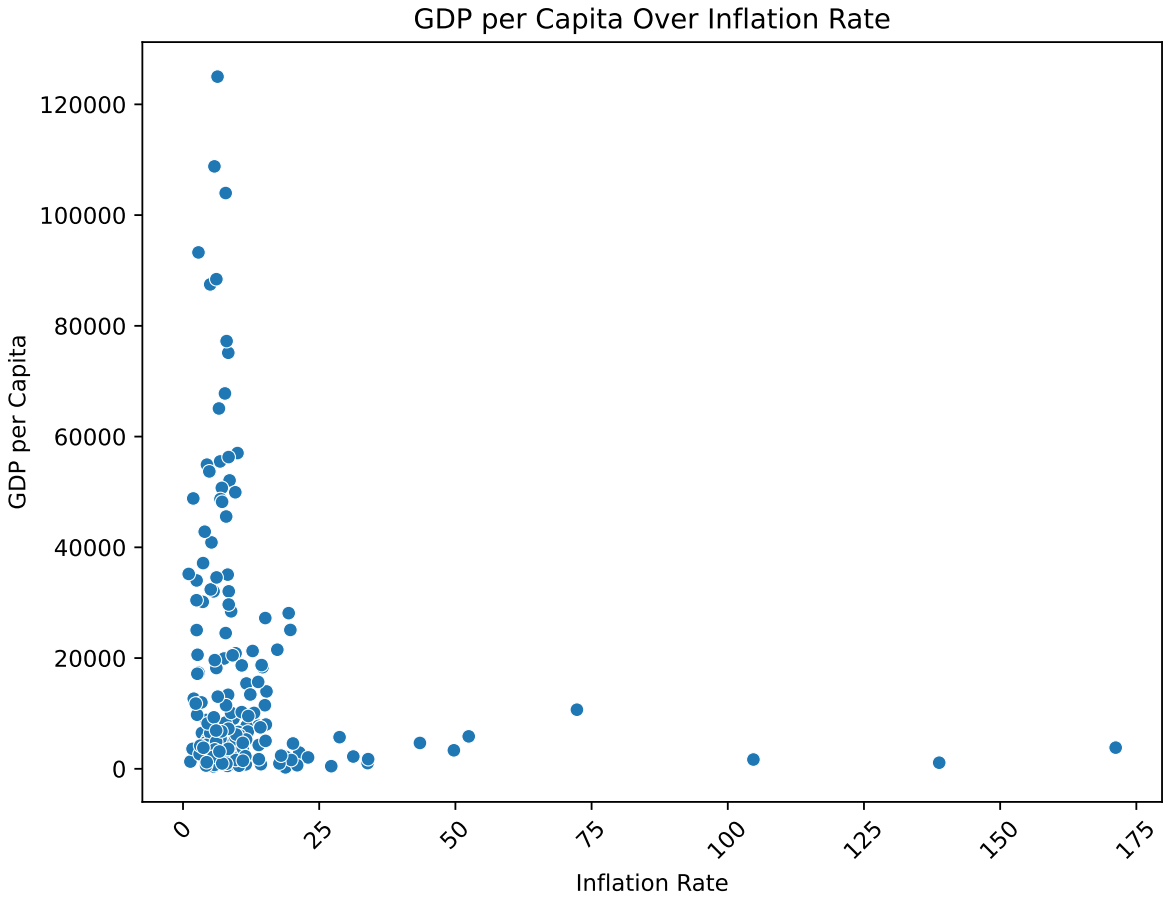

Figure 1

Figure 1: Scatter plot showing the relationship between GDP per capita and total population.

Source: World Development Indicators dataset.

According to Figure 1 countries with high inflation rates generally seem to have low GDP per capita values. Further analysis is needed to derive a correlation, especially regarding the lower inflation countries.

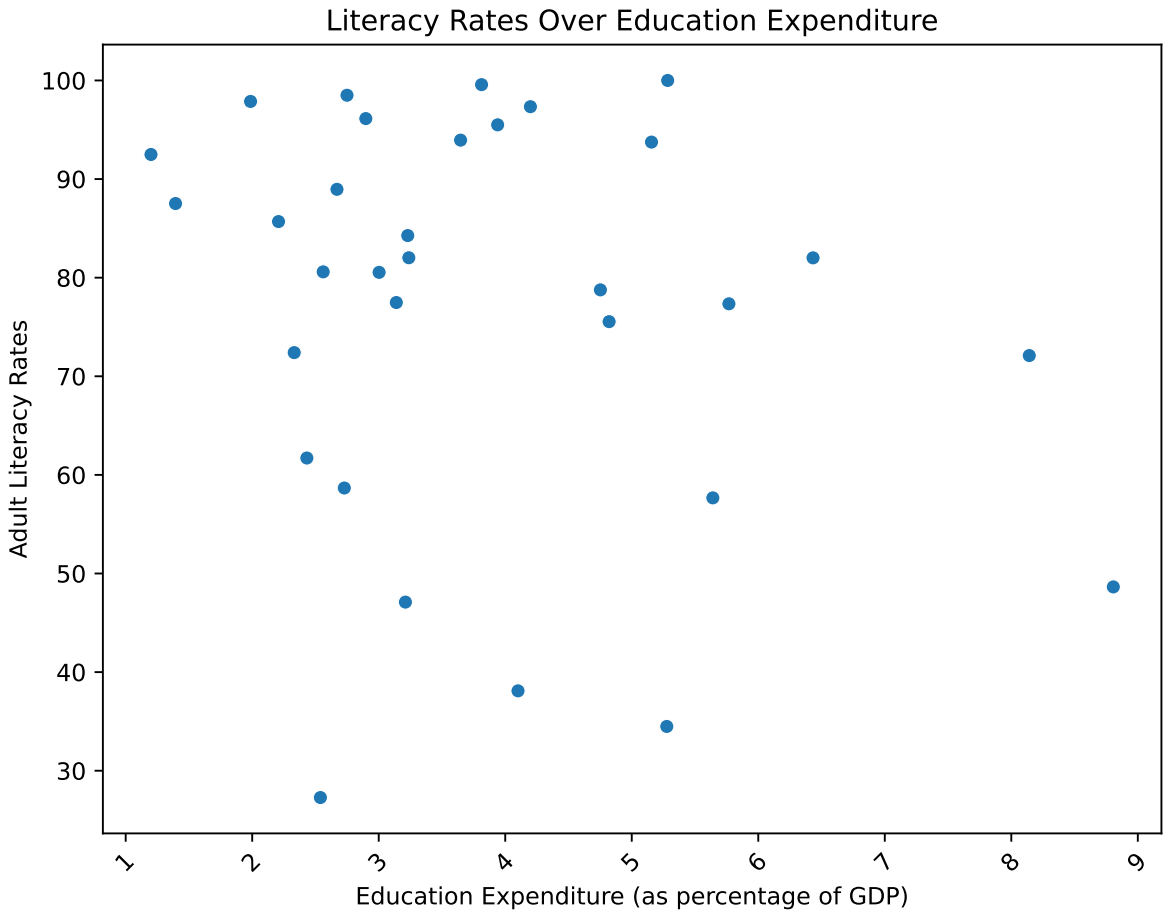
Literacy Rates Over Education Expenditure
Adult Literacy Rates
Education Expenditure (as percentage of GDP)
100
90
80
70
60
50
40
30
1
2
3
4
5
6
7
8
9

Figure 2

Figure 2: Scatter plot showing the relationship between Literacy Rates and education expenditure.

Source: World Development Indicators dataset.

Figure 2 is very sparse and has very few data points. This is an example of a plot from which conclusions cannot be drawn. We should be very cautious of plots like these as they are missing lots of data!

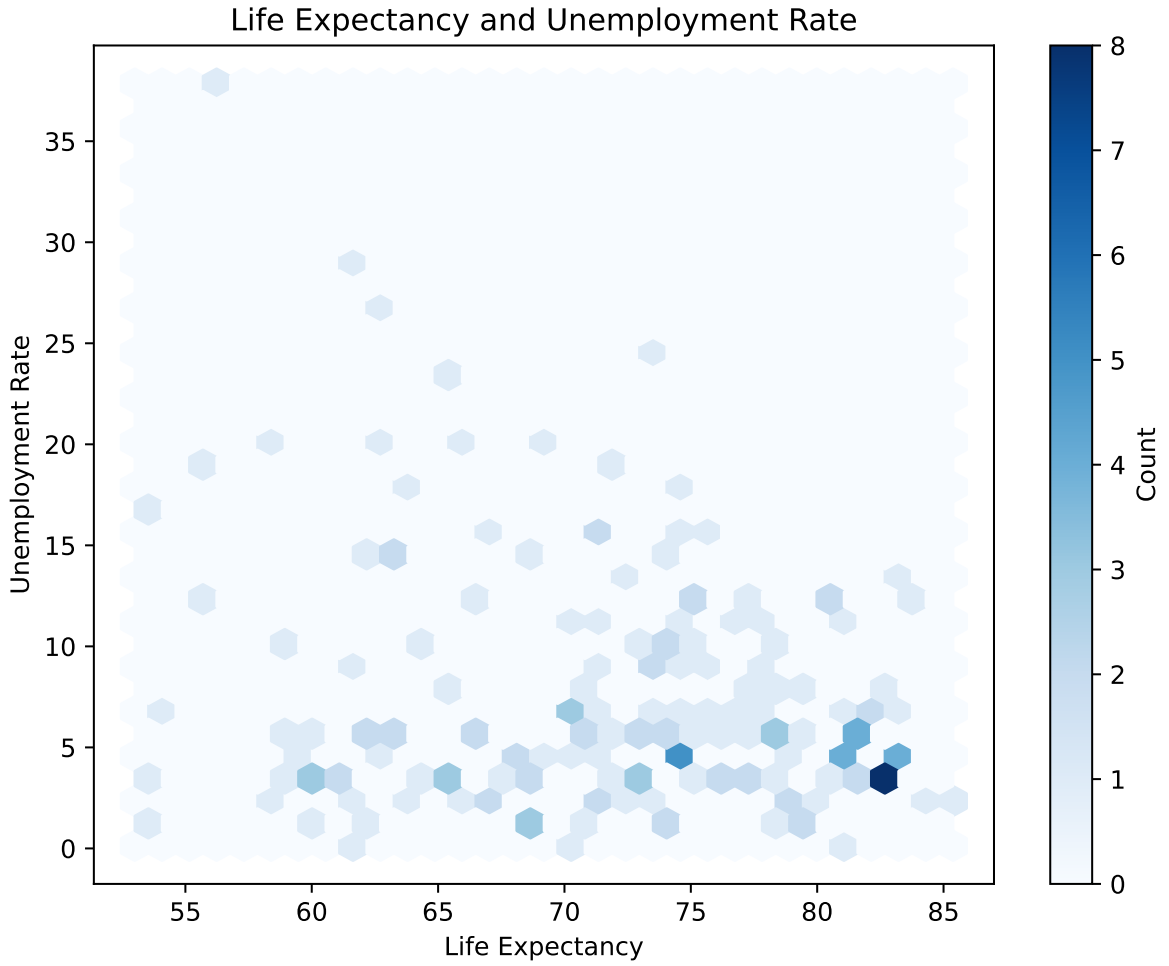

Figure 3

Figure 3: Scatter plot showing the relationship between Life Expectancy and Unemployment Rate.

Source: World Development Indicators dataset.

According to Figure 3 countries with high life expectancies generally seem to have low unemployment rates. This is likely the result of powerful countries having stronger economies and healthcare systems.

Key Statistics

The table below highlights the variables used

	mean	min	max
gdp_per_capita	20345.71	259.03	240862.18
inflation_rate	12.49	-6.69	171.21
adult_literacy_rate	79.57	27.28	100.00
education_expenditure_gdp_share	4.23	1.03	16.58
life_expectancy	72.42	53.00	85.38
unemployment_rate	7.27	0.13	37.85