GDP Per Capita and Income Per Country

Hosea E. Wah

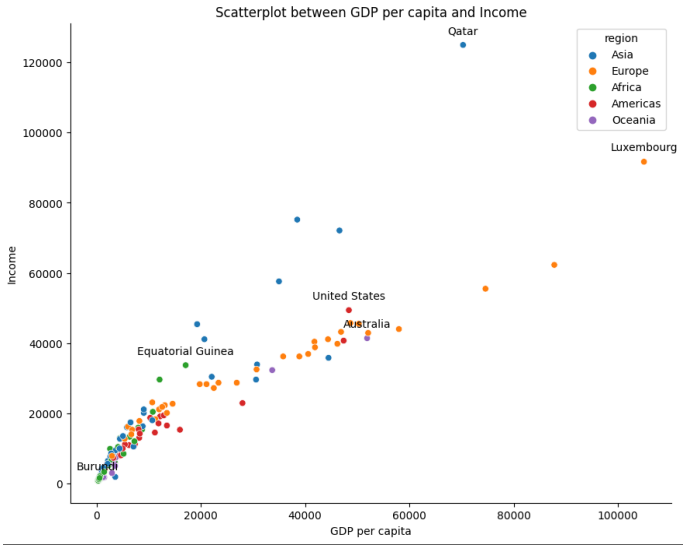
INFSCI 1520

Lingfei Wu

April 12th, 2023

[Link to Github Page]

ii. Main Figure Produced in Python



iii. Legend & Explanation

The legend that accompanies the graph provides valuable insight into the diverse geographical regions that are featured in the data visualization. The graph highlights the countries that are located in five major regions of the world, namely Asia, Europe, Africa, Americas, and Oceania. Each region is represented on the graph by a distinct color, which helps to visually distinguish between the various countries. For instance, the countries in Asia are represented by a cluster of blue dots, while the European countries are depicted as an orange dot. Similarly, the African countries are shown as green dots, while the countries located in the Americas are represented by a striking red dot. Lastly, the countries in the Oceania region are marked with purple dots. This comprehensive legend not only adds aesthetic appeal to the graph but also makes it easier for the viewer to interpret and understand the data presented.

iv. Findings from Visualization

- The scatter plot depicts GDP per capita and income and provides insights into global economic performance.

- African nations tend to have the lowest GDP per capita and income compared to other regions in the world due to historical and systemic factors.
- Equatorial Guinea has the highest GDP per capita of approximately $7,000 and an income of $25,000 among all African nations, while Burundi has the lowest GDP per capita and income.
- European countries have much higher GDP per capita and income values compared to African nations, with Luxembourg having the highest GDP per capita exceeding $100,000.
- The United States leads the Americas region with the highest GDP per capita of around $55,000 and an average income also around $55,000.
- Asian countries have more dispersion in their GDP per capita and income values, with Qatar standing out as the wealthiest nation in Asia.
- The scatterplot highlights significant disparities between regions and countries, underscoring the need for continued efforts towards promoting inclusive and sustainable economic growth, reducing poverty, and addressing the root causes of economic inequality.

v. Description of Data & Methods

The analysis used data from the "Unsupervised Learning on Country Data" dataset on Kaggle, which provides information on 167 countries' GDP per capita, income, and region. The dataset aims to enable unsupervised learning tasks to uncover underlying patterns and relationships among countries.

The analysis started by importing libraries such as pandas, numpy, matplotlib, and seaborn to manipulate, compute, and visualize data. A scatter plot was created using seaborn's scatterplot function to visualize the relationship between GDP per capita and income for countries in the dataset, with color-coding based on region.

To provide more context, the poorest and richest countries in each region were identified using the groupby method to sort the dataset based on GDP per capita. For each country, GDP per capita and income values were extracted, and annotations added to the scatter plot to label corresponding data points. Appropriate labels were also added to the x and y axes, and seaborn's despine function was used to remove top and right spines for presentation.

The analysis provided insights into wealth and income distribution across regions and economic disparities between the richest and poorest countries within each region. The scatter plot served as a starting point for further exploration, such as clustering countries based on GDP per capita and income, or incorporating additional variables like population, education, or health indicators to gain a more comprehensive understanding of the factors driving economic differences

vi. Significance Statement

A scatter plot between GDP per capita and income can reveal the relationship between a country's economic output and the average earnings of its citizens. This visualization can provide insights into how economic growth is distributed and whether it leads to an improvement in living standards.

When comparing African and European countries' economic disparities, especially in the context of colonization and slavery, the scatter plot can be particularly revealing. European countries that colonized African nations and benefited from exploitation may exhibit higher GDP per capita and income levels, while African countries that suffered from colonization and slavery might display lower levels.

The scatter plot can also shed light on income inequality within and between countries. A strong positive correlation between GDP per capita and income would suggest that economic growth benefits the average person. However, a weak or non-existent correlation could indicate that certain segments of the population are benefiting disproportionately, while others are left behind.

In the case of African countries, weak correlations between GDP per capita and income could be attributed to lingering structural issues caused by colonization and slavery, such as

unequal distribution of resources, lack of infrastructure, and limited access to quality education and healthcare. The scatter plot is a powerful tool for illustrating the long-lasting effects of historical injustices and the challenges that African nations continue to face in their pursuit of development and prosperity.