«Казахстанско-Британский технический университет»

Report

Abylay Toremuratuly

Sanzhar Aubakirov

Introduction

The project explores the economic performance of five developing and emerging countries — Poland, Turkey, India, Brazil, and South Africa — by comparing their Gross Domestic Product (GDP) and GDP per capita.

The goal was to investigate how the total size of an economy relates to the standard of living of its citizens using real data collected from two different sources:

- World Bank API (for total GDP data, 2013–2023)
- Wikipedia (for nominal GDP per capita values)

This comparison provides insights into economic inequality and development efficiency among these countries.

Data Collection & Preparation

API Data (World Bank)

Using the World Bank Open Data API (NY.GDP.MKTP.CD), I collected yearly GDP values (in current US dollars) for the years 2013–2023. The API returned structured JSON data that was converted into a pandas DataFrame.

Web Scraping (Wikipedia)

GDP per capita values were scraped from the Wikipedia page "List of countries by GDP (nominal) per capita" using requests and BeautifulSoup.

The relevant HTML table was identified and read via pandas.read_html().

Data cleaning included removing commas, converting strings to numeric values, and standardizing country names (e.g., "türkiye" → "turkey").

After cleaning, both datasets were merged by the country column, resulting in a unified dataset containing five countries and their most recent GDP statistics for 2023.

Analysis & Visualization

Three visualizations were produced to support the analysis:

- Figure 1. *Relationship between Total GDP and GDP per Capita (2023)* — a scatter plot showing how larger economies tend to have higher GDP per capita.

- Figure 2. *GDP per Capita Comparison by Country* — a bar chart illustrating differences in average income per person.
- Figure 3. *Correlation Heatmap* — confirming a strong positive relationship between total GDP and GDP per capita.

Descriptive statistics revealed wide variation across countries.

- India has the highest total GDP (~3.6 trillion USD) but the lowest GDP per capita (~2,900 USD).
- Poland has a smaller total GDP (~812 billion USD) but a much higher per capita GDP (~28,000 USD).
- Turkey and Brazil occupy middle positions, while South Africa shows moderate figures in both metrics.

The correlation coefficient between total GDP and GDP per capita was ≈ 0.84, indicating a strong positive association between economic size and average income.

Key Insights

- Economic size does not always translate to higher living standards — India's example shows that population scale can reduce per-person wealth.
- Poland demonstrates efficient economic growth, achieving high per-capita income with a moderate total GDP.
- Turkey and Brazil illustrate balanced mid-range economies, while South Africa remains behind in both indicators.
- The overall positive correlation suggests that larger economies tend to have higher individual prosperity, but demographic factors significantly shape this relationship.

Conclusion

This project successfully applied the complete data collection and preparation pipeline:

1. API extraction (World Bank),
2. Web scraping (Wikipedia),
3. Cleaning, merging, and EDA in pandas,
4. Visualization with Seaborn and Matplotlib.

The analysis highlighted clear differences between economic scale and individual wealth levels, reinforcing that GDP alone is not enough to measure prosperity.

Future work could include analyzing additional indicators such as education, unemployment, or CO_2 emissions to better understand the broader development context.