**WORLD BANK DATA - Analysis of Unemployment and Gross Domestic Product**

**INTRODUCTION**

Gross Domestic Product (GDP) is a key economic indicator that measures the total value of all goods and services produced within a country's borders over a specific period. It serves as a yardstick for a country's economic performance and size.

While a higher GDP is generally associated with better development and lower unemployment, it's essential to consider other factors. The distribution of wealth, the quality of jobs created, and the effectiveness of government policies also play crucial roles in determining the overall impact on development and unemployment. Additionally, economic growth should be sustainable and inclusive to ensure long-term benefits for the entire population.

Unemployment among individuals with advanced education can negatively impact economic growth and GDP. It leads to underutilization of skilled talent, reduced productivity, and potential economic drag. Skill mismatches and long-term consequences, such as a "brain drain," are also concerns. On the positive side, addressing this unemployment can enhance innovation, productivity, and human capital development, contributing to economic growth and competitiveness. The relationship is complex, requiring careful consideration of factors influencing both the workforce and economic development.

The objective of this analysis is to explore the relationship between Gross Domestic Product (GDP) and Unemployment across various countries. Two primary datasets from the World Bank, focusing on GDP and Unemployment, were meticulously curated from diverse sources, providing a comprehensive overview of economic indicators. The objective was clear — identify a dataset that harmonizes seamlessly with GDP, providing a holistic perspective on economic landscapes.

In our exploration, we're curious about something important. We often think that countries with more money (higher Gross Domestic Product or GDP) are better off, with a higher quality of life. But here's the big question: how do these richer countries handle an essential issue like unemployment?

Beyond the conventional scope of GDP-centric analyses, our quest extends to unravel the relationship between economic prosperity and a tangible societal metric—unemployment. To navigate this exploration, we deliberately choose an unconventional approach. Instead of examining the overall unemployment rate, we narrow our focus to a subset of the labor force—those with advanced levels of education. The rationale behind this decision is grounded in a simple yet profound question: Does having a degree translate into a tangible impact on unemployment rates?

**DATASET OVERVIEW**

***Dataset 1: World Bank's GDP Ranking Table***

Our analysis revolves on the GDP ranking table—an intricate compilation of economic data sourced from global statistical offices. This dataset undergoes meticulous annual updates, ensuring the provision of current and reliable economic indicators. Utilizing diverse methodologies, including surveys and administrative record scrutiny, this dataset represents a comprehensive aggregation process.

The conversion of data to current US dollars is facilitated by the World Bank's Atlas method, employing a weighted average of exchange rates. This process addresses variations in currency purchasing power, ensuring a nuanced and accurate portrayal of economic trends.

*Data Cleaning*

1. Columns 'Unnamed 2' and 'Unnamed 5' were identified as non-essential and removed.

As we navigate the sea of data, we spotted a couple of columns—'Unnamed 2' and 'Unnamed 5'—that weren't playing a part in our analytical adventure. To keep things focused, we decided to let go of these columns and streamline our dataset.

1. Missing values in the 'Country' column were addressed to prepare for merging.

We have identified missing values in the 'Country' column, and as part of the upcoming merging process, we plan to remove all NaN values from this column. This will facilitate a smoother merge operation, as we aim to merge based on the 'Country' column.

***Dataset 2: World Development Indicators (WDI) - Unemployment with Advanced Education:***

In this dataset we are uncovering insights from the World Development Indicators (WDI) dataset, specifically honing in on Unemployment with Advanced Education. This dataset, considered the World Bank's masterpiece in development indicators, brings together the most up-to-date and accurate global development information.

The key metric we're keenly observing is the percentage of highly educated individuals who find themselves without a job. This information is sourced from the Education and Mismatch Indicators database (EMI), a creation of the International Labour Organization (ILO). Updated annually using a weighted average aggregation method, this data provides a unique lens for analysis. It zooms in on aspects like social protection and labor dynamics, giving us a closer look at the challenges faced by individuals with advanced education.

So, when we talk about the percentage of the labor force with advanced levels of education facing unemployment, dataset focuses on the percentage of highly educated individuals facing unemployment.

**THE BIG MERGE**

We're merging Unemployment and GDP datasets to understand economic trends better. We're using 'Country Code' as the common link between Unemployment's 'Country Code' and GDP's 'Country' columns. Opting for 'Country Code' minimizes errors due to potential formatting differences in 'Country Name.' The goal is to correlate unemployment rates with GDP data for more accurate insights. This choice ensures a smooth merge, enhancing the reliability of our analysis on the interplay between unemployment and economic performance.

*Initial Findings:*

In our analysis of unemployment and GDP datasets, major economies like India and China exhibit missing unemployment figures, posing a challenge in assessing their employment landscapes comprehensively. Rather than discarding these entries, we've chosen to retain them, recognizing that missing values don't necessarily imply an absence of unemployment. Various factors contribute to reporting challenges, such as differing reporting practices, issues in data collection, substantial informal sectors, and methodological variations. India and China, with large populations and significant informal economies, face difficulties in accurately reporting unemployment figures, a challenge shared by many countries globally.

*A white background with numbers and symbols

Description automatically generated***ANALYSIS**

**Descriptive Statistics:**

Descriptive statistics were calculated for 'Ranking,' 'US dollars (millions),' and '2017 [YR2017].' These provide insights into central tendencies, variability, and distribution within the dataset.

In 2017, data on unemployment rates is available for 120 countries. On average, the global unemployment rate was around 7.44%, with a moderate level of variability (standard deviation of approximately 6.13%). The middle point of the dataset, or the median, is 5.13%. Notably, 25% of countries have unemployment rates below 3.23%, while 75% have rates below 9.88%. The reported rates range from a minimum of 0.37% for Qatar to a maximum of 33.29% for West Bank and Gaza

Qatar, with a minimum unemployment rate of 0.37%, reflects a situation characterized by economic stability and robust employment opportunities. This low rate aligns with Qatar's economic profile, which is largely driven by its petroleum and natural gas industries, coupled with significant investments in infrastructure and diversification efforts. The nation's strategic economic planning and wealth from natural resources contribute to a favourable employment landscape.

Conversely, West Bank and Gaza present a starkly different scenario, with the highest reported unemployment rate of 33.29%. This high rate is indicative of substantial economic challenges in the region, influenced by factors such as political instability, limited access to resources, and the impact of longstanding conflicts. Economic conditions in West Bank and Gaza are often intertwined with geopolitical events, making the pursuit of sustained employment opportunities more challenging.

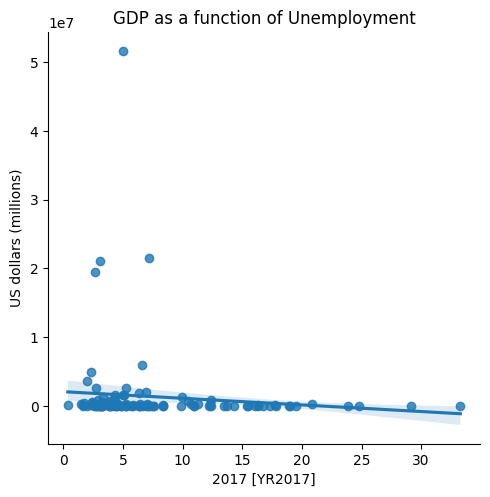
In summary, the disparity in unemployment rates between Qatar and West Bank and Gaza underscores the diverse economic contexts and challenges faced by different regions, shaped by factors ranging from resource abundance to geopolitical complexities.

**DATA VISUALIZATION**

A chart of a heatmap

Description automatically generatedThe correlation heatmap reveals interesting relationships between GDP and unemployment rate across different countries. A weak negative correlation exists between the two, indicating that as a country's GDP increases, its unemployment rate tends to decrease. This suggests a potential link between economic prosperity and lower unemployment levels.

However, it's crucial to remember that correlation doesn't imply causation. Other factors, like government policies and social safety nets, likely play a role in shaping unemployment rates as well.

**LINEAR MODELLING:**

The results from the linear regression analysis show that when the unemployment rate ('2017 [YR2017]') goes up by one unit, the GDP ('US dollars (millions)') tends to go down by approximately $96,319.55. The starting point, called the intercept, is around $2,076,709.53.

The negative coefficient indicates that there's a negative connection between GDP and unemployment, suggesting that as GDP increases, unemployment tends to decrease. However, it's important to note that the correlation is weak, and the observed relationship might not be statistically strong, according to the P-Value and the R-Square. So, while there's a hint of a connection, a strong conclusion cannot be drawn from this analysis which would make intuitive sense as there are many indicators that go into the analysis.

**Non-Linear Relationship**

The graph illustrates a negative correlation between unemployment and GDP, implying that as unemployment rises, GDP tends to decrease, and conversely, as unemployment falls, GDP tends to increase. Additionally, the non-linear nature of the relationship suggests that changes in GDP are not proportional to changes in the unemployment rate. A one-percentage-point increase in unemployment may cause a more substantial GDP decrease than a corresponding decrease would lead to a GDP increase.

**CONCLUSION:**

In conclusion, our analysis of the relationship between Gross Domestic Product (GDP) and unemployment, with a focus on individuals with advanced education, offers valuable insights into the economic dynamics of diverse countries. While GDP provides a broad view of economic performance, our unique approach narrows the focus to the subset of the labor force with advanced education.

The merger of GDP and Unemployment datasets, based on 'Country Code,' allowed for a nuanced analysis. Retaining entries with missing unemployment figures for major economies like India and China acknowledges reporting challenges, demonstrating our commitment to a thorough exploration.

Backed by descriptive statistics and visualizations, we observe a weak negative correlation between GDP and unemployment rates, suggesting higher GDP associates with lower unemployment. However, caution is warranted, as correlation doesn't imply causation. Government policies and social safety nets likely contribute to the complex unemployment landscape.

Linear regression supports a negative connection between GDP and unemployment, but the weak correlation and non-significant p-values caution against definitive conclusions. This aligns with the recognition that economic growth should be sustainable and inclusive, considering wealth distribution and policy effectiveness.

Examining cases like Qatar's stable employment landscape and challenges in regions like West Bank and Gaza underscores the need for context-specific analyses. In summary, our analysis provides a stepping stone for a nuanced understanding of the GDP-unemployment relationship, emphasizing careful interpretation and the multifaceted nature of economic indicators. Future research could delve into regional dynamics and policy impacts for a more comprehensive view.