ASSIGNMENT: DATA VISUALIZATION ANALYSIS - G7 ECONOMIC AND SOCIAL INDICATORS

Dhairyasheel Patil *CS-GY 6313 B*

Introduction

This report presents an analysis of two visualizations regarding key economic and social indicators in G7 countries from 2000 to 2022. The first visualization is a correlation heatmap that explores the relationships among Fertility Rate, Female Labor Force Participation Rate, and GDP per Capita. The second visualization is a combined line chart that illustrates the trend of the average log-transformed unemployment rate and GDP per Capita over time. The following sections describe the question each visualization aims to answer, the design rationale, and the key findings.

Clarification of Dataset and Bias

The dataset used for this analysis comes from trustworthy sources such as the World Bank and national statistical databases for G7 countries. It is important to note, however, that potential bias exists due to differences in data collection methods across countries. Additionally, focusing only on G7 countries may introduce economic bias, limiting the global applicability of these trends.

1 Visualization 1: Correlation Heatmap

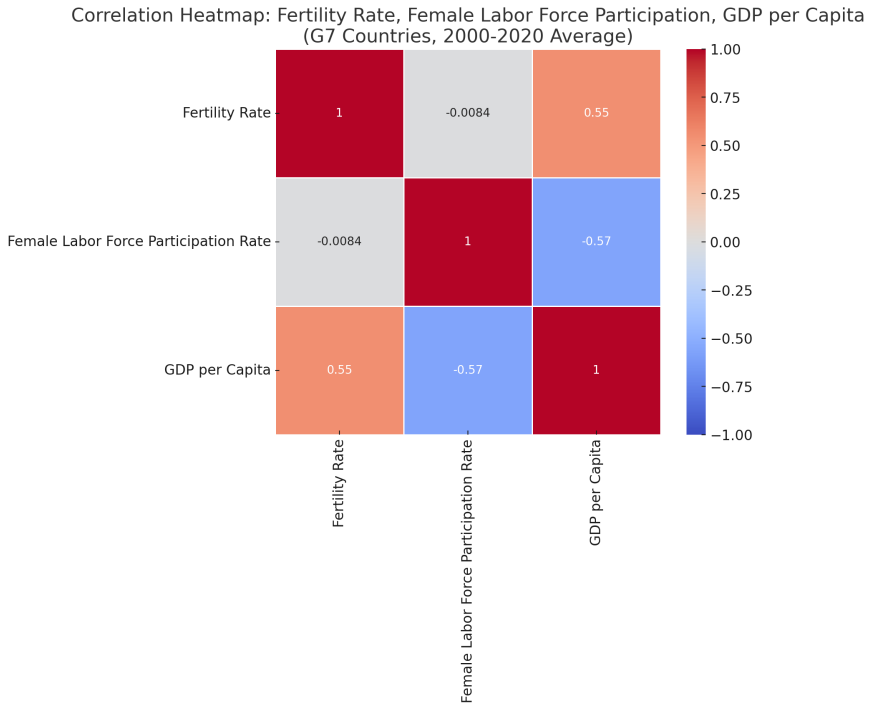

Figure 1: Correlation Heatmap of Fertility Rate, Female Labor Force Participation Rate, and GDP per Capita (G7, 2000-2020 Average)

Question

"What is the correlation among Fertility Rate, Female Labor Force Participation Rate, and GDP per Capita in G7 countries from 2000 to 2020?"

Interpretation

The heatmap reveals a moderate positive correlation (0.55) between Fertility Rate and GDP per Capita, suggesting that countries with higher fertility rates also tend to have higher GDP per capita within this dataset. A moderate negative correlation (-0.57) exists between Female Labor Force Participation Rate and GDP per Capita, indicating an inverse relationship between these two factors. Minimal correlation (-0.0084) is observed between Fertility Rate and Female Labor Force Participation Rate, implying independence in their variations.

Design Rationale and Considerations

The heatmap uses a gradient color scheme from blue to red, where red indicates a positive correlation and blue indicates a negative correlation. This color scheme provides a visual cue to identify the strength and direction of the relationships between variables. Labels along both axes and annotated correlation coefficients within each cell make the information directly accessible to the viewer. The focus on three variables allows for a simple grid format, which keeps the visualization uncluttered.

2 Visualization 2: Unemployment Rate and GDP Trend Line Chart

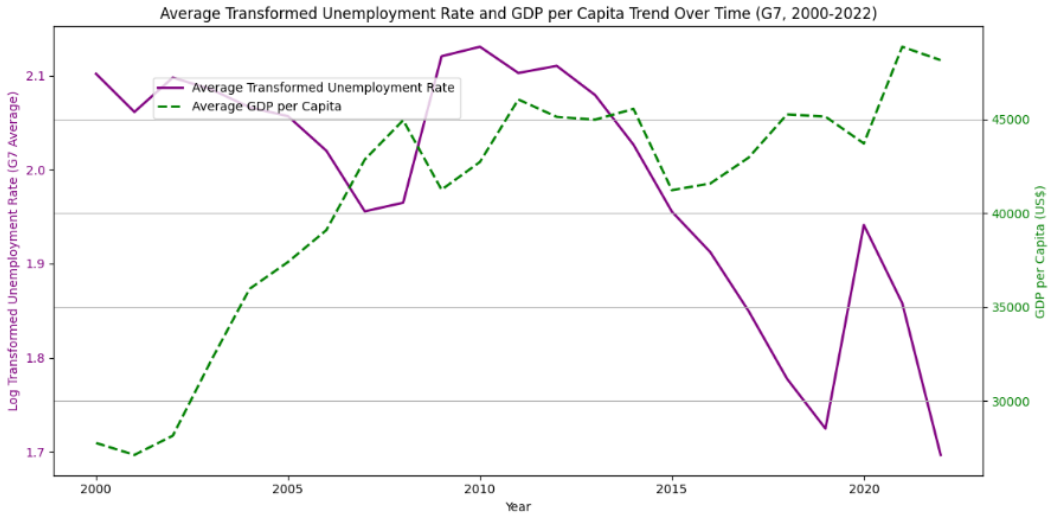

Figure 2: Average Transformed Unemployment Rate and GDP per Capita Trend Over Time (G7, 2000-2022)

Question

"What is the trend of the average log-transformed unemployment rate and GDP per Capita in G7 countries from 2000 to 2022?"

Interpretation

The graph shows an inverse correlation between GDP per capita and the unemployment rate for G7 countries from 2000 to 2022. As GDP per capita (green) generally rises, the unemployment rate (purple) tends to decrease, especially after 2010. This suggests that economic growth is associated with lower unemployment. Notably, fluctuations around 2020 reflect economic impacts likely related to global events.

Design Rationale and Considerations

This visualization uses two separate y-axes to represent the two variables. The left y-axis shows the log-transformed unemployment rate, while the right y-axis shows GDP per Capita in USD. The solid magenta line represents the

unemployment rate, while the dashed green line represents GDP per Capita, with each line designed to stand out in its own color. The use of dual axes allows both trends to be observed concurrently, providing insights into how economic health, indicated by GDP per capita, relates to unemployment.

Potential for Misleading Interpretations

Visualization 2 can be deceptive for a few reasons:

1. **Different Y-Axis Scales**: The unemployment rate and GDP per capita are on separate y-axes, making it hard to directly compare the magnitude of changes. This separation may exaggerate or downplay the apparent relationship between the two variables, depending on the viewer's interpretation of the line positions.
2. **Transformed Data**: The unemployment rate is log-transformed, which may obscure the actual rate of change. This transformation can create a smoother or more dramatic trend than exists in the raw data, potentially leading viewers to misinterpret the actual dynamics of unemployment. It was applied to normalize the large variance in unemployment rates, but viewers may not easily understand the rationale without this explanation.
3. **Non-Intuitive Color and Style Choices**: The use of bold colors and different line styles can draw attention to trends that aren't as significant. This visual emphasis may amplify the perceived inverse relationship, even if the correlation isn't strong or consistent over time.
4. **Single-Axis Overlay**: Presenting both trends on the same chart may suggest a causal relationship between GDP per capita and unemployment, though correlation does not imply causation. Without additional context or analysis, viewers might misinterpret this relationship as directly causal rather than merely correlative.

Clarification of Bias and Context

It is important to highlight that the use of G7 countries as the dataset for this analysis introduces economic bias, as these countries are advanced economies with unique economic structures. Additionally, the selection of only two variables (GDP per Capita and unemployment) may not provide a complete picture of the socio-economic dynamics at play.