Analysing the Relationship Between Unemployment Rate and GDP Growth in the United States

1. Introduction

Understanding how unemployment rates and GDP growth are interconnected is crucial for grasping the broader dynamics of the U.S. economy. This project aims to explore this relationship over both long-term and short-term periods, focusing on annual data from 1960 to 2023 and quarterly data from 2014 to 2024.

By analysing trends and patterns, the project seeks to uncover whether these indicators consistently move together, diverge during certain periods, or exhibit significant correlations. The results could help provide valuable insights into labour market dynamics, policy implications, and economic resilience.

2. Used Data

Source and Description

The data for this project is sourced from **Federal Reserve Economic Data (FRED)**, a trusted database managed by the Federal Reserve Bank of St. Louis. It includes two datasets:

- **Annual Data** (1960–2023): Provides yearly unemployment rates and inflation-adjusted GDP.
- **Quarterly Data** (2014–2024): Captures seasonal trends in unemployment and GDP on a quarterly basis.

These datasets were processed using a data pipeline that cleaned, merged, and stored them in an SQLite database for efficient analysis.

Structure and Features

The cleaned data has the following structure:

1. **Columns**:
    - **DATE**: Timestamps for data points.
    - **UnemploymentRate**: Seasonally adjusted percentage of the unemployed labor force.
    - **GDPRate**: Inflation-adjusted GDP in billions of dollars.
2. **Tables**:
    - **annual_data**: Annual averages of unemployment and GDP.
    - **quarterly_data**: Seasonal trends for unemployment and GDP.

```
=== Schema of Table: annual_data ===          === Schema of Table: quarterly_data ===
   cid            name  type  notnull dflt_value  pk      cid            name  type  notnull dflt_value  pk
0    0            date  TEXT        0       None   0    0    0            date  TEXT        0       None   0
1    1 UnemploymentRate  REAL        0       None   0    1    1 UnemploymentRate  REAL        0       None   0
2    2         GDPRate  REAL        0       None   0    2    2         GDPRate  REAL        0       None   0

=== First 5 Rows of Table: annual_data ===    === First 5 Rows of Table: quarterly_data ===
         date  UnemploymentRate   GDPRate              date  UnemploymentRate   GDPRate
0  1960-01-01               5.5  14001.090      0  2014-07-01               6.1  18406.941
1  1961-01-01               6.7  14360.265      1  2014-10-01               5.7  18500.031
2  1962-01-01               5.6  15240.493      2  2015-01-01               5.5  18666.621
3  1963-01-01               5.6  15904.567      3  2015-04-01               5.4  18782.243
4  1964-01-01               5.2  16821.109      4  2015-07-01               5.1  18857.418
```

*Fig.1 Schema of Table: **annual_data** and **quarterly_data***

License Compliance

The data complies with FRED's **Terms of Use**:

1. Attribution: Results and visualizations include proper acknowledgment of FRED as the source.
2. Non-commercial Use: The project is strictly academic, with no commercial application.
3. Copyright Retention: All relevant copyright notices are preserved.

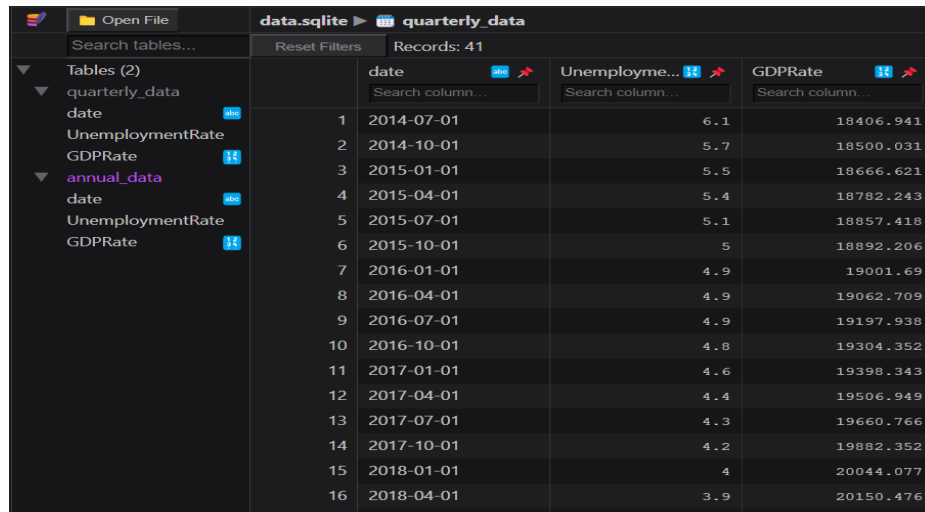## 3. Pipeline

Pipeline Overview

To process the raw datasets, an automated pipeline was developed in Python. The pipeline's primary functions include:

1. **Downloading Data**: Fetching annual and quarterly data for unemployment and GDP from FRED in CSV format.
2. **Data Cleaning and Transformation**:
    - Renaming columns for clarity (e.g., converting UNRATE to UnemploymentRate and GDPC1 to GDPRate).
    - Merging datasets on the date column for both annual and quarterly data.
3. **Database Integration**: Storing the merged datasets in an SQLite database for efficient querying.

Pipeline Output:

The pipeline outputs the following SQLite tables:

1. **quarterly_data**: Contains date, quarterly unemployment rates and GDP.
2. **annual_data**: Contains date, annual unemployment rates and GDP.

Fig.2 Output Data from the Pipeline

4. Analysis

Methodology

The analysis involved examining both annual and quarterly datasets to assess trends, distributions, and correlations between unemployment and GDP. Key techniques included:

1. **Descriptive Statistics**: Summarized data properties like mean, median, and standard deviation.

```
=== EDA for Table: quarterly_data ===                      === EDA for Table: annual_data ===

Summary statistics:                                        Summary statistics:
              date  unemploymentrate          gdprate                    date  unemploymentrate         gdprate
count           41         41.000000        41.000000      count           64         64.000000       64.000000
unique          41               NaN              NaN      unique          64               NaN             NaN
top    2014-07-01 00:00:00        NaN              NaN      top    1960-01-01 00:00:00        NaN             NaN
freq             1               NaN              NaN      freq             1               NaN             NaN
mean           NaN          4.748780     20641.702537      mean           NaN          5.923438    45736.325500
std            NaN          1.713056      1484.353886      std            NaN          1.619400    22634.390824
min            NaN          3.500000     18406.941000      min            NaN          3.500000    14001.090000
25%            NaN          3.800000     19304.352000      25%            NaN          4.850000    25268.680000
50%            NaN          4.200000     20548.793000      50%            NaN          5.600000    40906.349500
75%            NaN          5.100000     21919.222000      75%            NaN          7.025000    66061.888000
max            NaN         13.000000     23400.294000      max            NaN          9.700000    90684.386000
```

*Fig. 3 Descriptive statistics of **annual_data** and **quarterly_data***

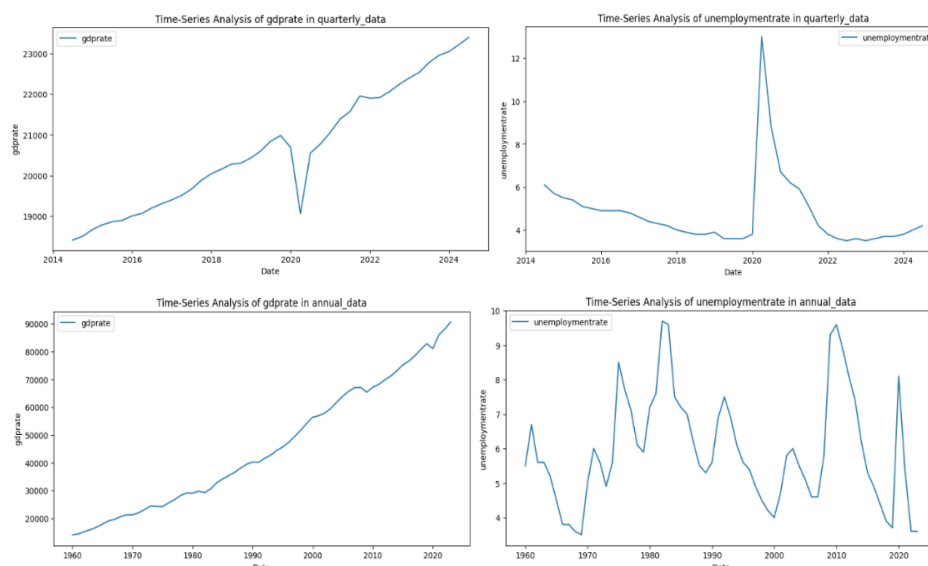2. **Time-Series Analysis**: Explored how unemployment and GDP have evolved over time.

*Fig. 4 Time-Series Analysis **annual_data (gdprate and unemploymentrate) and quarterly_data (gdprate and unemploymentrate)***

3. **Correlation Analysis**: Quantified the strength of the relationship between unemployment and GDP.
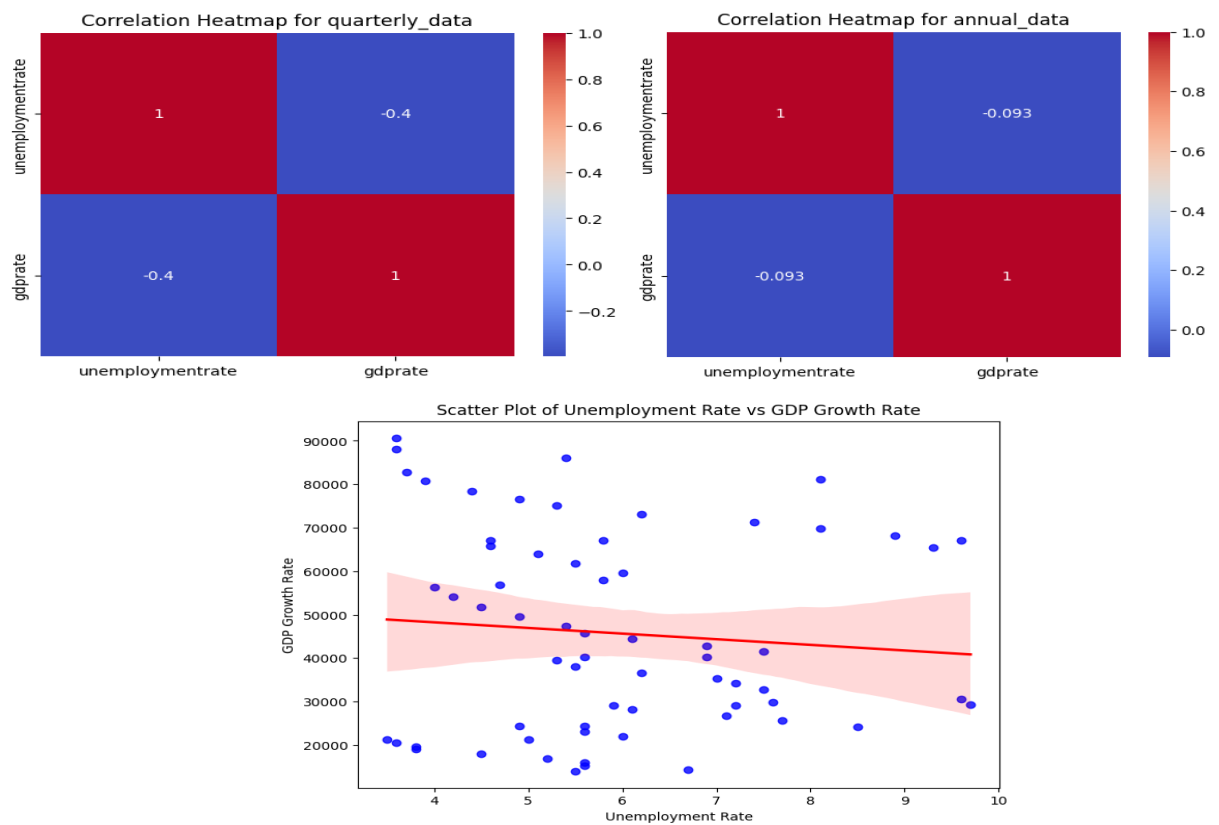

Fig. 5 *Correlation Analysis **annual_data and quarterly_data and Pearson Correlation***

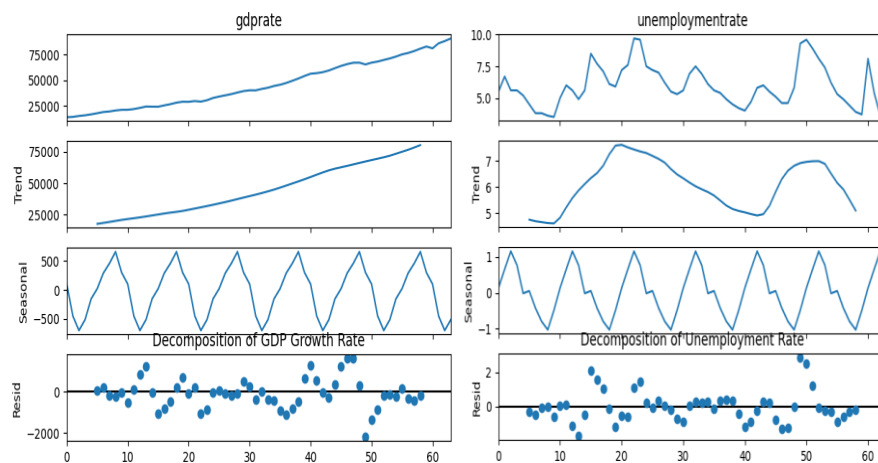**4. Seasonal Decomposition:** Identified seasonal patterns and long-term trends.

Fig.6 *Seasonal Decomposition of gdprate and unemploymentrate*

Findings

1. **Trends Over Time**:
   o **Annual GDP Growth**: Showed a consistent upward trend over decades, reflecting long-term economic expansion.
   o **Unemployment Rates**: Exhibited cyclical patterns, with sharp increases during major economic downturns such as the 2008 financial crisis and the COVID-19 pandemic in 2020, followed by declines during periods of recovery.

2. **Correlation Results**:
   o **Annual Data**:
     ▪ A weak negative correlation of **-0.093** suggests that over the long term, GDP growth has a limited direct influence on unemployment rates.
     ▪ This weaker relationship may be influenced by structural shifts, technological advancements, and labour market adaptations that dilute GDP's direct impact on employment.
   o **Quarterly Data**:

- A stronger negative correlation of **-0.4** highlights a clearer and more immediate relationship between GDP growth and unemployment.
- This suggests that short-term economic activity has a noticeable impact on employment levels, aligning with expectations that GDP growth often fosters job creation.

  o The Pearson Correlation Coefficient of **-0.09** indicates a very weak negative relationship between GDP growth and unemployment rates in the annual dataset. This suggests that GDP growth alone is not a strong predictor of unemployment trends over the long term. External factors such as policy decisions, technological shifts, and labour market dynamics likely play significant roles in shaping unemployment trends.

3. **Seasonal Decomposition**:
Seasonal decomposition for both GDP growth and unemployment provided deeper insights into underlying patterns:
   o **GDP Growth Rate**:
     - The trend component indicates steady economic growth over time.
     - The seasonal component captures regular fluctuations, reflecting predictable economic cycles.
     - The residuals reveal deviations caused by unexpected events or anomalies, such as economic shocks.

   o **Unemployment Rate**:
     - The trend highlights long-term cyclical changes, such as rising or falling unemployment tied to economic phases.
     - The seasonal component captures industry-specific hiring and employment cycles.
     - The residuals reflect variations due to external shocks or policy shifts, including events like COVID-19.

Conclusions

The analysis supports the expected dynamics between GDP growth and unemployment rates, with nuanced observations:
1. **Long-Term Perspective**:
   o The annual data suggests that while GDP growth generally aligns with lower unemployment, the relationship is influenced by external factors such as policy interventions, technological evolution, and global economic events.
2. **Short-Term Perspective**:
   o Quarterly data reveals a stronger and more immediate link between GDP growth and employment, demonstrating that short-term economic activities significantly influence unemployment levels.

Despite these insights, the analysis has limitations. For example, it does not explicitly model external shocks such as COVID-19, and the assumption of linear relationships may overlook the complexities of economic interactions.

Future Work

To build on these findings, the following directions are recommended:
1. **Incorporate Additional Variables**:
   o Analyse other factors such as inflation rates, labour force participation, and industry-specific GDP contributions to better understand their roles in shaping unemployment trends.

* Pipeline section have been discussed in the data report

* Grammar corrections and structural suggestions in this document were assisted by ChatGPT.