Analyzing the Effect of U.S. Public Debt on Economic Growth

1. Introduction

1.1 Motivation for Analyzing the U.S. Economy

The U.S. economy is a cornerstone of global financial stability and has undergone significant changes over time. Understanding the dynamics between key indicators, such as GDP, public debt, inflation, and unemployment, provides valuable insights for policymakers, economists, and researchers. This project investigates how the growth of the U.S. economy is related to the increase in government borrowing, aiming to uncover trends and relationships that influence economic health.

1.2 Main Research Question

The central question this project seeks to answer is:
   • "Is the growth of the U.S. economy related to the increase in government borrowing over time?"

2. Datasets

For this project, we have identified two primary datasets to analyze the relationship between economic growth and government borrowing over time:

2.1 U.S. Economic Indicators (1974–2024)

Datasets

The U.S. Economic Indicators dataset provides detailed data on key economic trends in the United States from 1974 to 2024. Key attributes include monthly measurements of GDP, inflation (CPI), and unemployment rates. These indicators are crucial for understanding the performance of the U.S. economy and assessing its stability over time.

   • **MetaData URL:** U.S. Economic Indicators Dataset
   • **Data URL:** Available for download via Kaggle API (data from 1974 to 2020 is used in this project).

- **Usability : 10.00**
- **Description:** This dataset is complete, well-structured, and requires minimal preprocessing.
- **License:** Open Data Commons Public Domain Dedication and License (PDDL) v1.0
- CPIAUCSL: Consumer Price Index.
- UNRATE: Unemployment rate.
- GDP: Gross Domestic Product in millions.

Data spans from 1974 to 2024, offering a comprehensive view of economic trends.

2.2 U.S. Public Debt vs. GDP (1947–2020)

Dataset: Provides quarterly data on GDP and public debt in the U.S. Source: Kaggle Dataset License: Standard Kaggle license. Description:
- Quarter: Date of the quarter.
- Gross Domestic Product ($mil): GDP in millions.
- Total Public Debt ($mil): Public debt in millions.

Data spans from 1947 to 2020.

3. Data Pipeline

The data pipeline processes datasets on U.S. Economic Indicators and Public Debt using the Kaggle API. The pipeline involves steps such as data extraction, preprocessing, quarterly aggregation, and merging. Figure 1 illustrates the pipeline. Figure 1: Data Pipeline

Figure 1: Data Pipeline
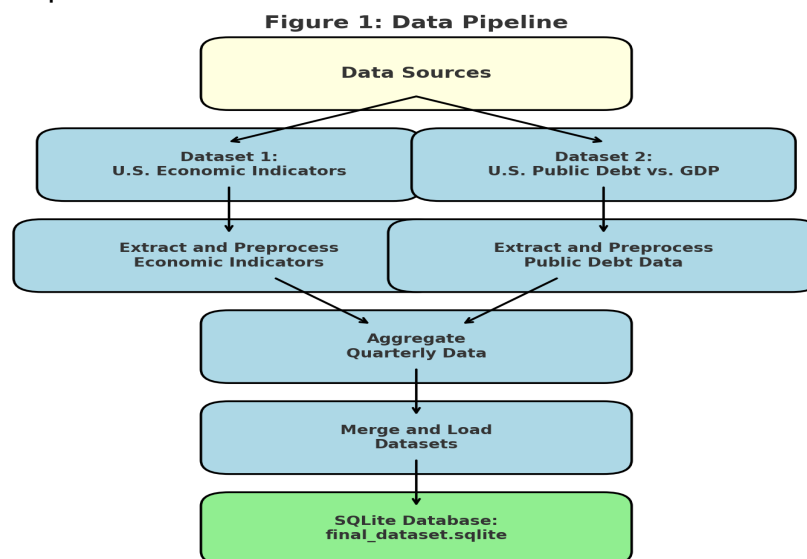
Transformations and Cleaning Steps:
  • Monthly data on CPI, GDP, and Unemployment aggregated to quarterly intervals.
  • Public Debt data merged with quarterly aggregated economic indicators.
  • Calculated metrics: Debt-to-GDP Ratio, GDP Growth Rate, and Debt Growth Rate.
  • Missing values handled using forward and backward filling.

4. Results and Limitations

4.1 Output Data Description The final dataset includes:
  • Quarterly data on CPI, GDP, Unemployment, Public Debt, and Debt-to-GDP Ratios.
  • Derived metrics: GDP Growth Rate and Debt Growth Rate.
  • The dataset is stored in an SQLite database (`final_dataset.sqlite`).

4.2 Data Quality and Limitations The dataset is well-prepared and normalized, but some limitations exist:
  • Temporal mismatch: Economic Indicators extend to 2024, while Public Debt data ends in 2020.
  • Derived metrics (e.g., Growth Rates) may introduce minor inaccuracies due to aggregation.