Overview of "Bank Safety Score"

Objective:

The main goal of this project is to analyze historical stock price data for major banks and GDP data to calculate a safety score for each bank. The safety score provides an indication of the bank's financial stability in the coming years based on the given data.

Data Sources:

- Stock price data for major banks ('JPM', 'BAC', 'WFC', 'C', 'GS') is obtained from Yahoo Finance using the yfinance library.
- GDP data is obtained from the Federal Reserve Economic Data (FRED) using the fredapi library.
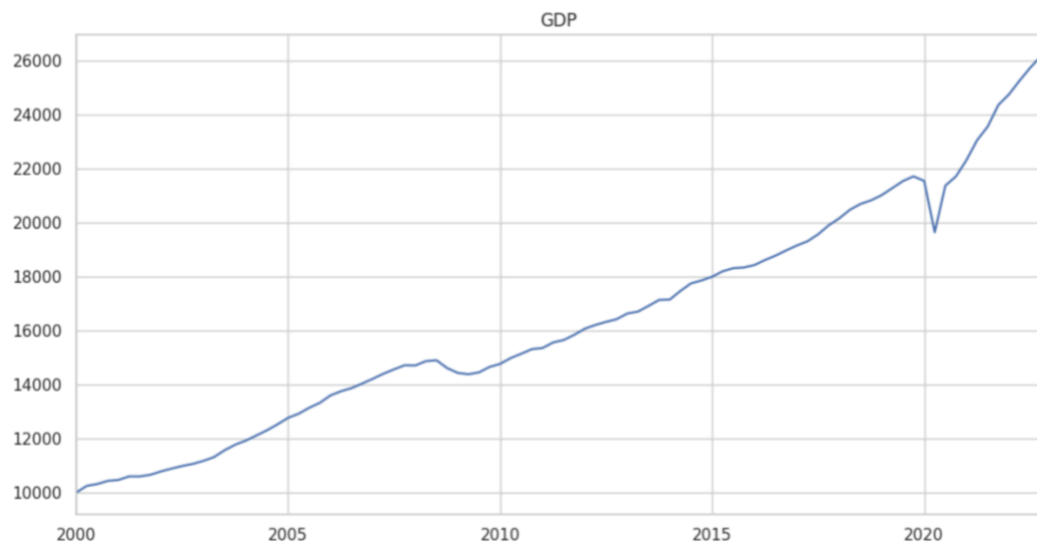
Data Preprocessing:

- Merged the bank stock data and GDP data.
- Filled missing values using forward fill and backward fill methods.
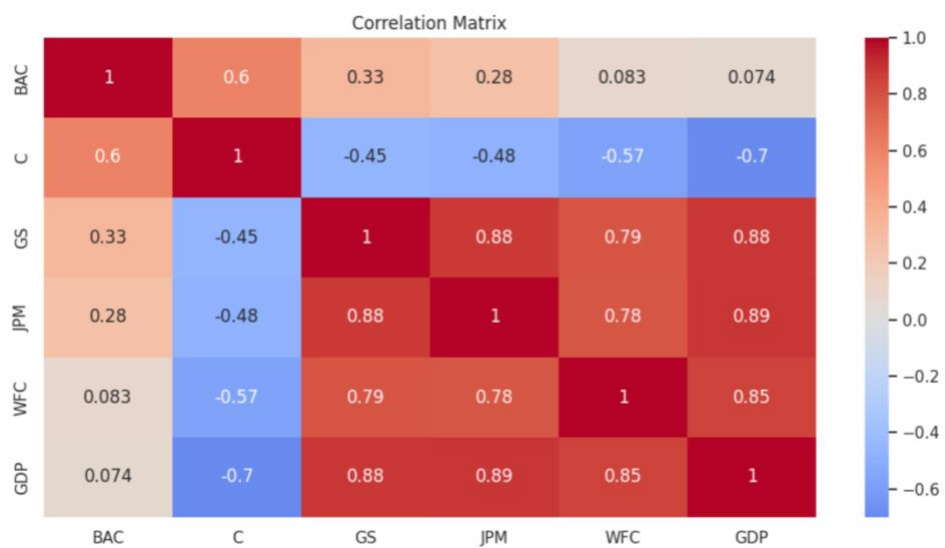- Removed outliers using the interquartile range (IQR) method.

Exploratory Data Analysis (EDA):

- Plotted the time series data for major banks and GDP.

Major Banks Stock Prices

GDP

- Calculated and visualized the correlation matrix using a heatmap.

Correlation Matrix

Time Series Modeling:

- Calculated monthly returns for the dataset.
- Split the dataset into training (80%) and testing (20%) sets.
- Normalized the data using mean and standard deviation.
- Created and fit a VAR model on the training data.
- Forecasted future returns using the fitted model.

Model Evaluation:

- Calculated the safety scores for each bank based on the forecasted data.

- Calculated the Mean Absolute Error (MAE) for each bank's actual vs. predicted stock prices.

Bank	Mean Absolute Error scores
GS - Goldman Sachs	0.01478337291046335
C - Citigroup Inc.	0.016861853937068915
JPM - JPMorgan Chase & Co.	0.014050040881805661
BAC - Bank of America Co.	0.01587713801676777
WFC - Wells Fargo & Co.	0.01589671869004082

Results:

Safety scores were calculated for each bank, with higher scores indicating better financial stability relative to other banks.

In this project, the safety score for each bank is calculated based on the standard deviation of their forecasted stock returns. The idea behind using standard deviation is that it measures the volatility or the amount of variation in the stock prices. Higher volatility generally implies higher risk.

The safety score calculation steps are as follows:
1. After fitting the VAR model, forecasted future returns are generated for the test data period.
2. The standard deviation of the forecasted returns for each bank is calculated.
3. The safety score for each bank is computed as the inverse of the standard deviation (i.e., 1 / standard deviation).

The rationale behind using the inverse of the standard deviation is that a lower standard deviation implies lower volatility and, therefore, lower risk. By inverting the standard deviation, banks with lower volatility will have higher safety scores, indicating better financial stability relative to other banks.

Safety scores:

Bank	Safety Score
GS - Goldman Sachs	97.866694
C - Citigroup Inc.	86.230598
JPM - JPMorgan Chase & Co.	83.076303
BAC - Bank of America Co.	75.385437
WFC - Wells Fargo & Co.	69.397305

Conclusion:

This analysis provides insights into the financial stability of major banks based on historical stock price data and GDP.

This data shows that the safety score is maximum for Goldman Sachs followed by Citigroup Inc.