Post-Pandemic Analysis of Leading Investment Banks

Team Member: Takaaki Morita (ID:7075405771), Ming Shan Lee (ID: 1895899214)

Introduction:

This report provides an analysis of the post-pandemic performance of leading investment banks. It examines their financial health, risk management, liquidity status, and strategic changes made in response to the pandemic. We aim to understand the recovery rates and overall resilience through comprehensive data analysis. Using predictive models, we forecast the stock prices for the upcoming month, contributing valuable insights into the expected market behavior. The COVID-19 pandemic has unprecedentedly impacted global financial markets, requiring significant strategic shifts among investment banks. This project seeks to assess these institutions' post-pandemic financial health and performance. By analyzing stock prices and financial reports from the past five years, we evaluate liquidity, risk management, and the effectiveness of strategic changes. The research problem is to ascertain investment banks' economic resilience and strategic positioning in a post-pandemic world.

1. Data Collection

1.1 Data Sources

As the introduction states, we will analyze the stock prices and financial reports of the past five years. It is not practical to analyze every investment bank, so we chose to analyze the top 5 investment banks and try to infer the whole market from their analysis.

We obtained the monthly closing prices of stocks for the past five years of Goldman Sachs Group Inc. (GS), Morgan Stanley (MS), JP Morgan Chase & Co. (JPM), Bank of America Corporation (BAC), and Citigroup Inc. (C), which allowed us to track the performance of each investment bank over a significant period, encompassing pre-pandemic, during-pandemic, and post-pandemic phases from the Yahoo finance webpage: https://finance.yahoo.com/.

The financial reports from the investment companies are also crucial to present the whole picture of the finance market. As it is not scrapable from the Yahoo finance webpage, we obtained the investment banks' income statements, balance sheets, and cash flow statements on a quarterly and annual basis from the Python package 'yfinance' for the past four years. These documents provide insights into the financial health of the banks, including their revenue streams, profitability, liquidity positions, and capital structures.

1.2 Collection Methodology

We utilized Python 'requests' and 'BeautifulSoup' packages to scrap historical stock prices from the Yahoo finance webpage. By sending HTTP requests to Yahoo Finance's server, we can obtain the relevant financial data we need, and then parse the valid data using BeautifulSoup. At last, we extracted the necessary data from the HTML elements, such as tables and div tags, where financial data is typically housed.

In addition to the historical stock prices data, we also collected financial statements from the Python package 'yfinance,' which we noticed that the Yahoo finance webpage had prevented us from scrapping financial information from its webpage.

2. Data Cleaning

2.1 Data Cleaning

Since we utilized the JSON and Beautifulsoup packages, it is straightforward that the data we extracted is in string type. We first convert every data in the dataframe to its proper type, converting the 'Date' column to the DateTime type and the rest columns into integer type by removing the digit grouping commas and then converting them into integers.

Then, we noticed some missing values in the financial data. After manual checking, we are sure there isn't any missing value in the stock prices data, but we still drop any instance with no value in the 'Volume' column, which happens if the market is not open. Although there are many ways to deal with missing values, the only proper way to deal with them in financial data is to drop them.

We can obtain clean data after converting data types and dropping missing values. In addition, we sort the data with the 'Date' column to maintain the data is still in time series and reset the index after sorting.

3. Analysis

In this section, we will present different analyses of the two data obtained. For the financial data, we will be giving statistical analysis. As for the stock price data, we will present a correlation heatmap between features to give insight into the data.

3.1 Statistical Analysis

We began with a descriptive analysis to understand our financial data's distribution, central tendency, and dispersion. This included calculating means, medians, standard deviations, and percentiles using the pandas and NumPy libraries.
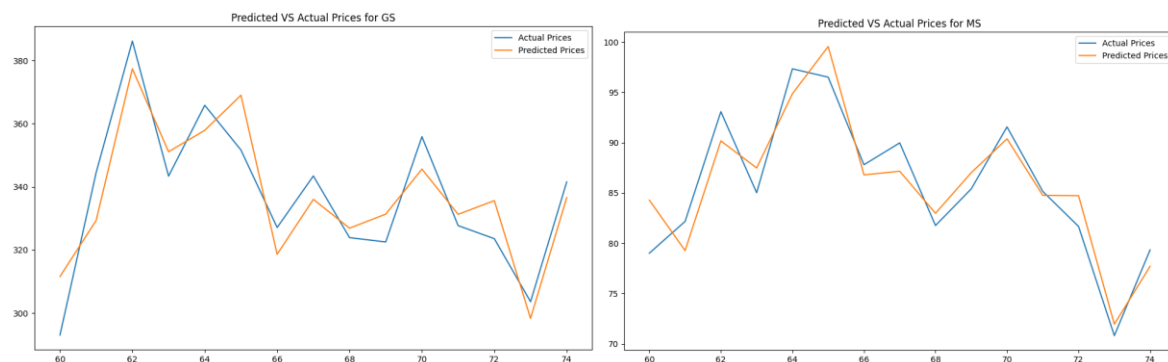
3.2 Correlation Heatmap

To give further insight into our stock price data, we present the correlation heatmap between columns of the data.
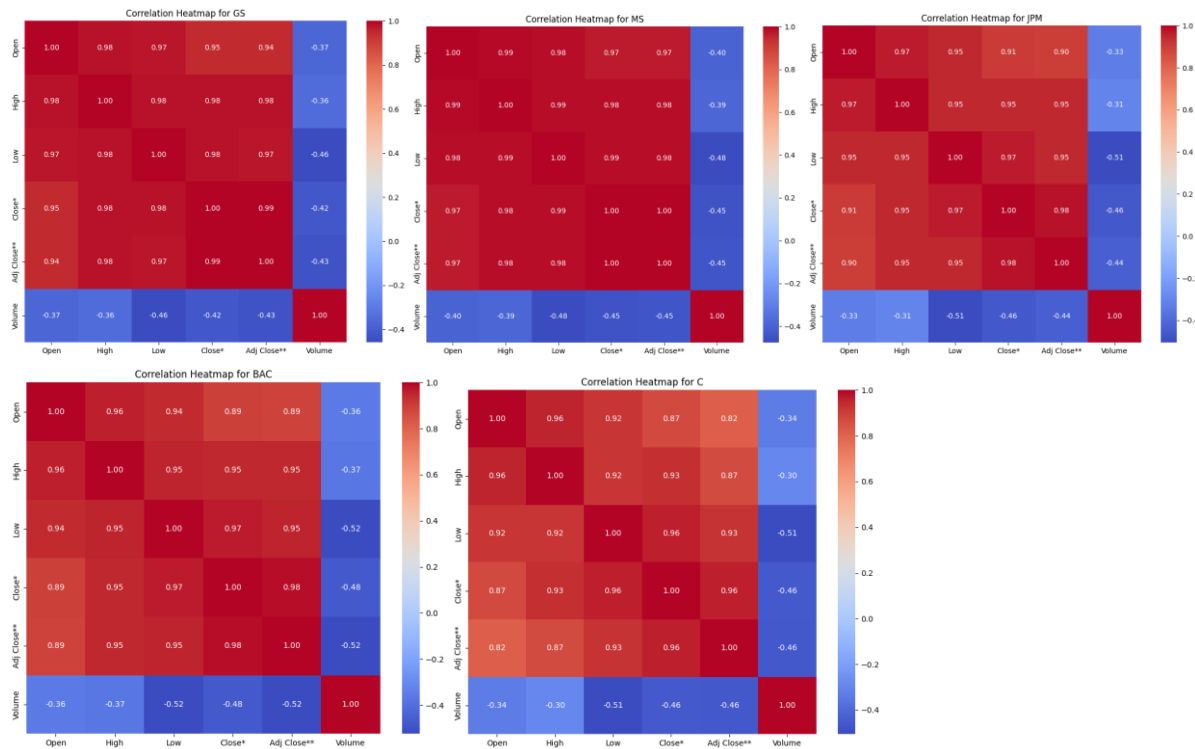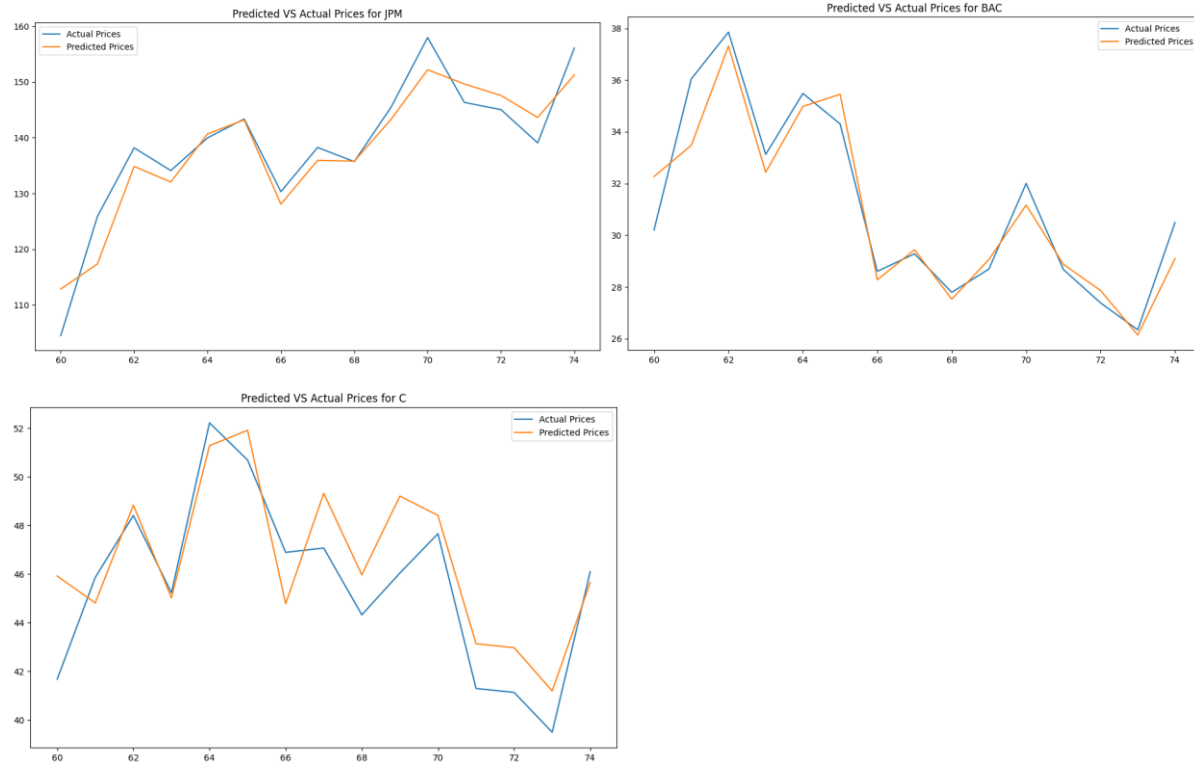
As shown in the graphs, we can observe that other features have a high positive correlation except for the 'Volume' column. From our domain knowledge of stock prices, the correlations make sense since a higher value in one of the features often means higher values in other features. At the same time, higher values in the stock prices often mean lower volume since high stock prices are less attractive to stock buyers.

3.3 Predictions (Machine Learning)

After observing the correlations between features, we are curious about how accurately we can predict future stock prices from these high correlation features. Knowing potential future price movements helps in managing risk. Investors and companies use predictions to hedge against downturns and to calculate the VaR (value at risk for their portfolios.

Predicted VS Actual Prices for JPM

Predicted VS Actual Prices for BAC

Predicted VS Actual Prices for C

	Machine Learning Model Metrics			Moving Average Model (Baseline Model)		
Company	RMSE	MAE	R-SQUARED	RMSE	MAE	R-SQUARED
GS	10.3359	9.28517	0.793687	18.6163	14.9098	0.330712
MS	2.50337	2.20785	0.868495	4.61539	3.86933	0.552997
JPM	4.22662	3.39448	0.882208	7.29578	6.00889	0.649027
BAC	1.04851	0.78317	0.905328	2.13426	1.76333	0.607742
C	1.90588	1.59086	0.696150	2.72039	2.28467	0.380942

From the metrics, we can see that the model we presented has outperformed the moving average model, which is the baseline model, significantly in every metrics.

3.4 Extra visualizations

We had included the most relevant visualizations in this report. Extra visualizations can be found on GitHub.

4. Discussion

This section of our report synthesizes the data analysis, highlighting the resilience of investment banks to post-pandemic market conditions. We discuss the implications of our financial ratio analysis, which suggests that banks with robust digital transformation strategies and diversified asset portfolios demonstrated superior resilience. The time-series analysis revealed that despite initial volatility, stock prices for most banks stabilized, indicating market confidence in their long-term stability. Furthermore, our predictive models suggest a cautiously optimistic outlook for stock prices in the next 30 days, albeit with a recommendation for investors to monitor market dynamics closely due to potential external shocks.

5. Conclusions and Future Work

The comprehensive analysis conducted in this study provides a clear picture of the investment banking sector's health in a post-pandemic environment. Our findings indicate that: the leading investment banks have shown commendable adaptability in the face of pandemic-induced economic challenges. Strategic shifts towards digitalization and asset diversification have effectively buffered against the pandemic's disruptive impact. Predictive models suggest a gradual improvement in stock prices, signifying a recovery trajectory. These insights underscore the importance of strategic agility and risk management in navigating unprecedented market disruptions. The conclusion drawn from this study is that while the investment banking sector was not immune to the pandemic's impacts, its core institutions have remained resilient, and there is a positive outlook for recovery in the short to medium term. To build on the findings of this study, we propose several avenues for future research: Longitudinal Study: Conducting a long-term study that tracks the performance of these banks over the next month to validate our predictions and observe long-term trends. Broader Industry Analysis: Expanding the scope of analysis to include a more comprehensive range of financial institutions, including regional banks and non-traditional financial services companies, to gain a more comprehensive view of the sector. Impact of Regulatory Changes: Examining the impact of post-pandemic regulatory changes on investment banking strategies and performance. Technological Innovations: Investigating the role of emerging technologies such as blockchain and artificial intelligence in shaping the future of investment banking.