Stat380FinalReport

Dhruv Nasit, Felicia Vijayarangam, Joseph Doan, and Eric Feng

1. **Introduction**

   Our project focuses on the United States' market. We know that financial markets are sensitive to not only the contemporary economy state, but also how new information compares to what investors expected. The macroeconomic announcements like the Consumer Price Index (CPI), the unemployment rate, and Federal Reserve interest rate decisions are among the most closely monitored. These events are important because they can change expectations about inflation, economic growth, and future interest rates. When these pivotal changes are apparent, investors may adjust their portfolios, which can cause the stock market volatility to go either up or down.

   In this project, we want to study the affects of these macroeconomic announcements towards the short-term volatility in the U.S stock market. We would be using the SPDR S&P 500 ETF (SPY) as our overall representation of the stock market. Instead of just looking at the average returns, we focus on the realized volatility around the macroeconomic announcement dates. We compute a 5-day rolling volatility measure using daily SPY logs returns and then examine how this volatility behaves before and after CPI, unemployment, and Federal funds rate releases.

   Our main research question is: "How do month-to-month changes in key U.S. macroeconomic indicators (CPI, unemployment, and the Federal funds rate) relate to the short-term changes in SPY volatility around announcement dates?"

   As a group, we were motivated to work on this topic because many of us are interested in learning about investing and follow macroeconomic news. We regularly see headlines in articles and television about CPI, unemployment, and Fed decisions "moving the market", but we want to closely see the effect of each of these events carefully by using data to prove its importance to the volatility of the market. This topic also lets us apply several methods we learned from STAT 380. For example, we can use regression and expand into time series models to measure these effects. In this way, the project is both personally interesting to us and a good practice to use what we learn in class to real-world situations.

2. **Data Description** – Describe your dataset, including its source, key variables. Include EDA and any preprocessing, cleaning, or feature engineering steps, if applicable.

3. **Methodology** – Explain the statistical learning methods you used in detail, and briefly discuss why you chose them. At least one method should be something covered in class, and at least one should be new or extended beyond class material.
4. **Data Analysis Results** – Present and interpret your main findings. Include relevant figures and tables that support your discussion.
5. **Conclusion** – Summarize the key takeaways, note any limitations, and suggest possible extensions or future directions.
6. **Author Contribution Statement** – Briefly describe each team member's contribution.