News Sentiment Analysis on Stock Prices

Meghana Kasula

Pruthvi Nagaraju

Shaloma Ghosh

Vinit Gupta

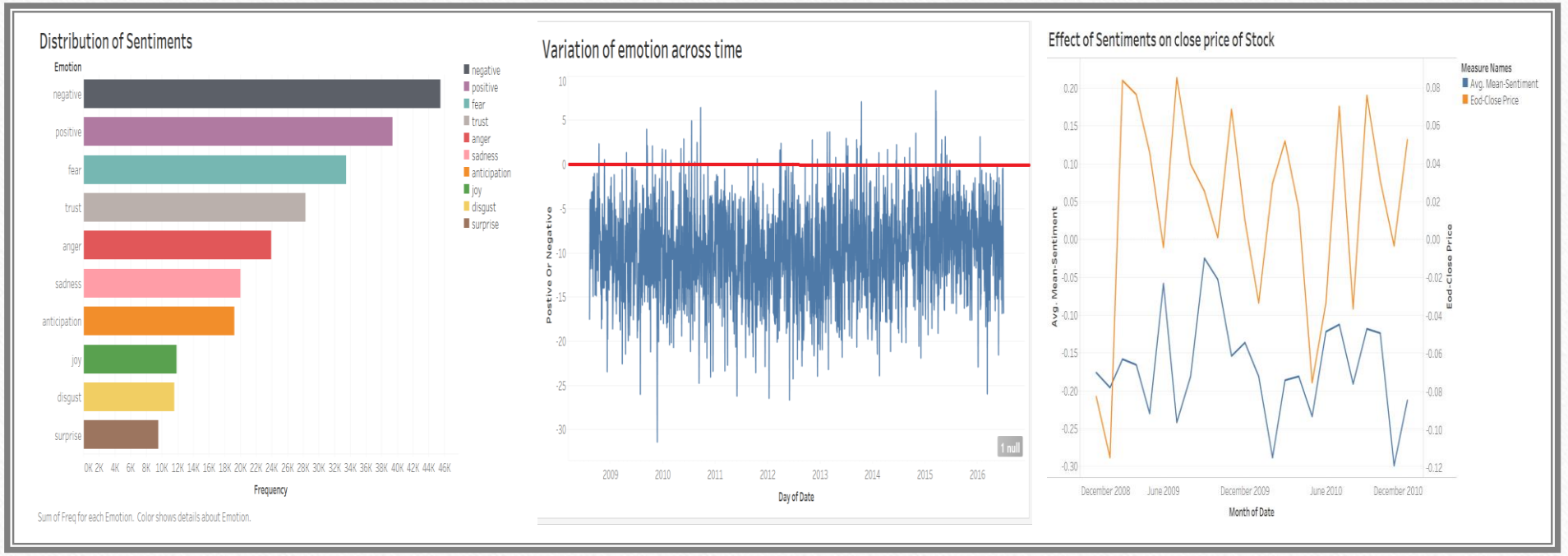
Weining Zheng

Objective and Data Set Used

- AIM : We aim to analyze the impact of public sentiments on stock market movements. We wanted to combine the conventional time series analysis with news from Reddit to predict weekly changes in stock price.
- Stock Price Data Set
- News Data Set

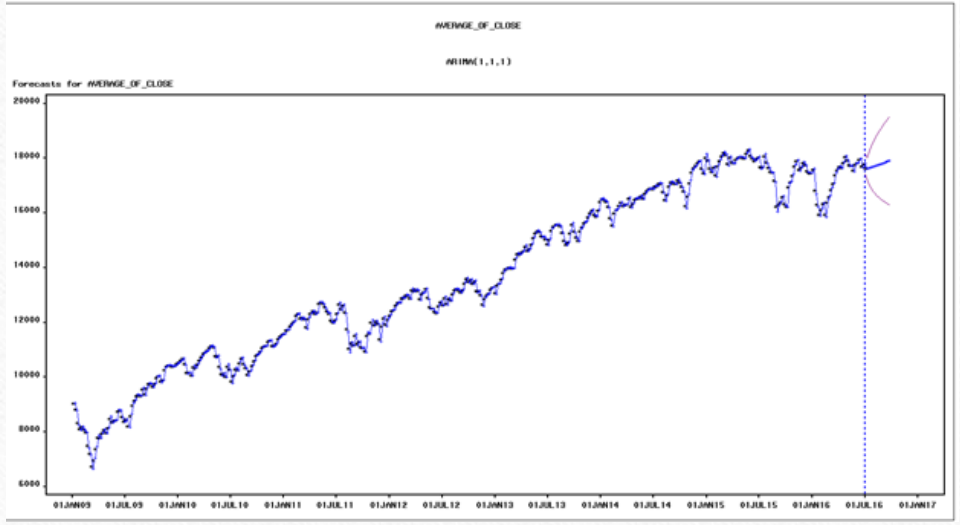
Sentiment Analysis(News Data)

- Used a package called “Syuzhet” on R to do an sentiment analysis of the combined news
- We have now 11 sentiments namely - Positive or Negative (sign indicated both), Anger, Anticipation, Disgust, Fear, Joy, Sadness, Surprise, Trust, Negative and Positive.
- We will be only using Positive or Negative Dataset for forecasting since it can store information of all kinds of news.

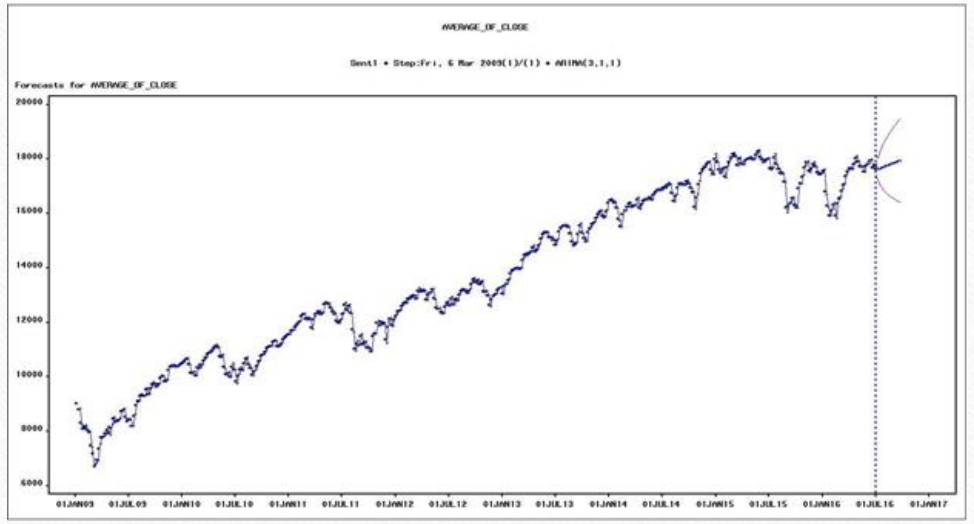


Visualization

Modeling - Forecasting Without & With Sentiments

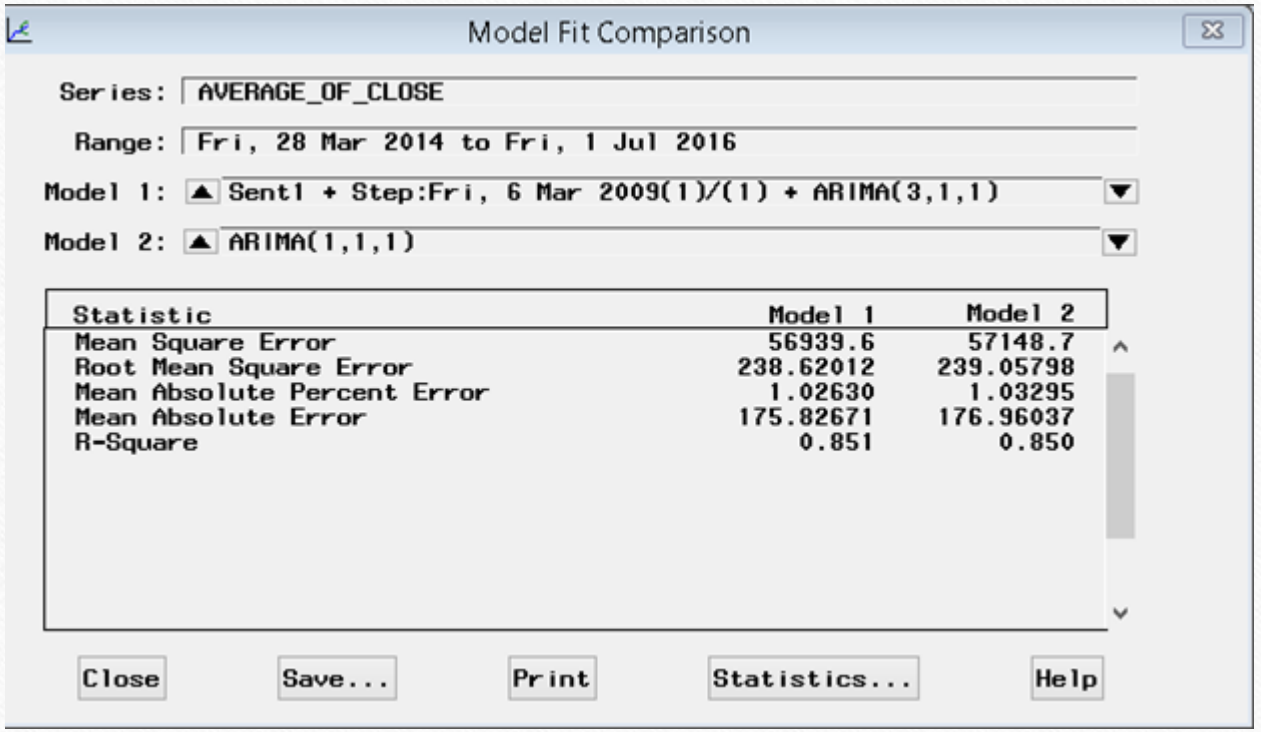


Statistic of Fit	Value
Mean Square Error	57148.7
Root Mean Square Error	239.05798
Mean Absolute Percent Error	1.03295
Mean Absolute Error	176.96037
R-Square	0.850

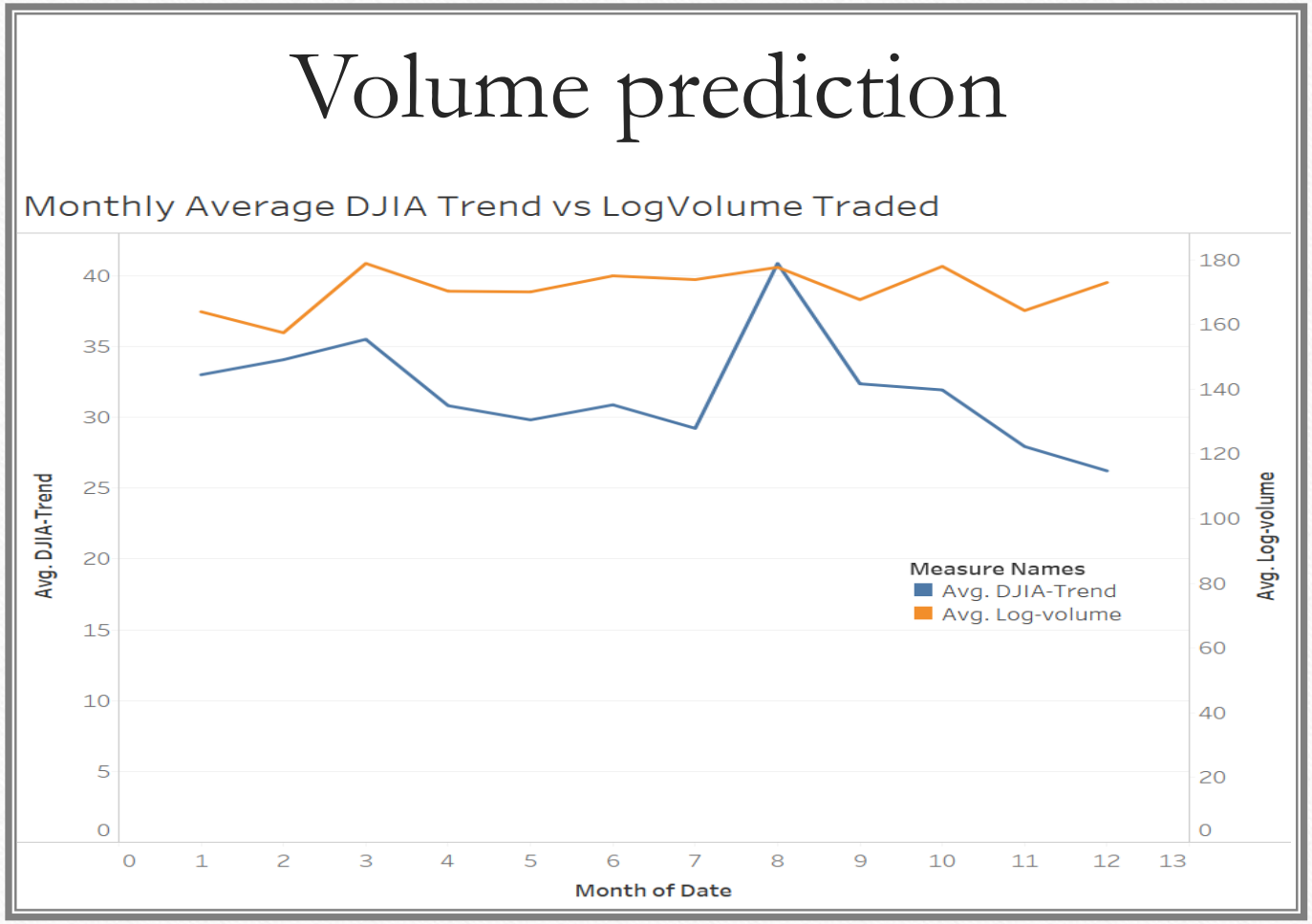


Statistic of Fit	Value
Mean Square Error	56939.6
Root Mean Square Error	238.62012
Mean Absolute Percent Error	1.02630
Mean Absolute Error	175.82671
R-Square	0.851

Comparison between Models in Part 1 and Part 2



Volume prediction



Objective - Our objective is to try and predict the volume of stock that will be traded based on the Google Trends obtained during the relevant time period.

Conclusion - we can conclude that, up to a certain extent we can predict the volume of stock to be traded during a given month based on the google trends.

Business Applications

- **Trading Institutions -** Trading Institutions can prepare the trend changes of the stock price in advance
- **Individual Investors -** Common sentiments, such as happy and anger, of certain news can allow the individual investors to understand what the price trend will be in the future and allow them to change their own investment

Conclusion

- Public mood can indeed be captured from the large-scale news feeds by means of simple natural language processing techniques.
- However we have obtained MAPE of about 1% using SAS Time Series Forecasting System.
- News is also highly biased since the stock news is predominantly negative.
- Our dataset only considers the reddit using, english speaking people. It's possible to obtain a higher correlation if the actual mood is studied.
- It may be hypothesized that people's mood indeed affect their investment decisions, hence the correlation.
- But in that case, there's no direct correlation between the people who invest in stocks and who use twitter more frequently, though there certainly is an indirect correlation - investment decisions of people may be affected by the moods of people around them, ie. the general public sentiment. All these remain as areas of future research.

Thank You