

Stock price forecasting

Murali Mandayam
July 31st, 2019



"Stocks fell sharply today on predictions that
the sun will rise again and tomorrow
will be another day..."

Are stock prices predictable?

Mark Twain's quotes

OCTOBER: This is one of the peculiarly dangerous months to speculate in stocks in. The others are July, January, September, April, November, May, March, June, December, August, and February.

- *Pudd'nhead Wilson's Calendar*

There are two times in a man's life when he should not speculate: when he can't afford it and when he can.

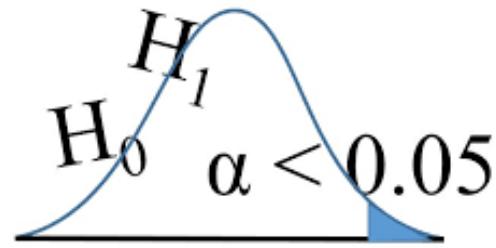
- *Following the Equator, Pudd'nhead Wilson's New Calendar*

Problem Overview

Many businesses and government agencies use forecasting techniques

- Companies forecast their revenues for their shareholders
- Meteorologists forecast the weather for citizens
- Economists forecasts GDP growths for countries

The stock market is no different. As people throng to the stock market to become monetarily richer and as companies looking to raise money get listed in the stock exchange there is so much data and potential to use machine learning techniques in the stock market.



Can machine learning really be used to
predict stock prices?

How accurate are such predictions?



Solution Approach

- Exploratory Data Analysis:
 - ACF
 - PACF
 - Decomp plots
 - Smoothing data with moving average
- Run these models:
 - ARIMA
 - SARIMA
 - LSTM
 - Linear Regression and
 - Random Forest Regressor
- Explore the error metrics such as RMSE, MAE, MPE etc.
- Analyze which model performs better and why
- Conclusion and Next steps

Data Sets

- We pull the data from here:
- <https://www.kaggle.com/camnugent/sandp500/downloads/sandp500.zip/4>
- 5 years worth data for 505 S&P 500 stocks

	date	open	high	low	close	volume	Name
0	2013-02-08	15.07	15.12	14.63	14.75	8407500	AAL
1	2013-02-11	14.89	15.01	14.26	14.46	8882000	AAL
2	2013-02-12	14.45	14.51	14.10	14.27	8126000	AAL
3	2013-02-13	14.30	14.94	14.25	14.66	10259500	AAL
4	2013-02-14	14.94	14.96	13.16	13.99	31879900	AAL



Exploratory Data Analysis

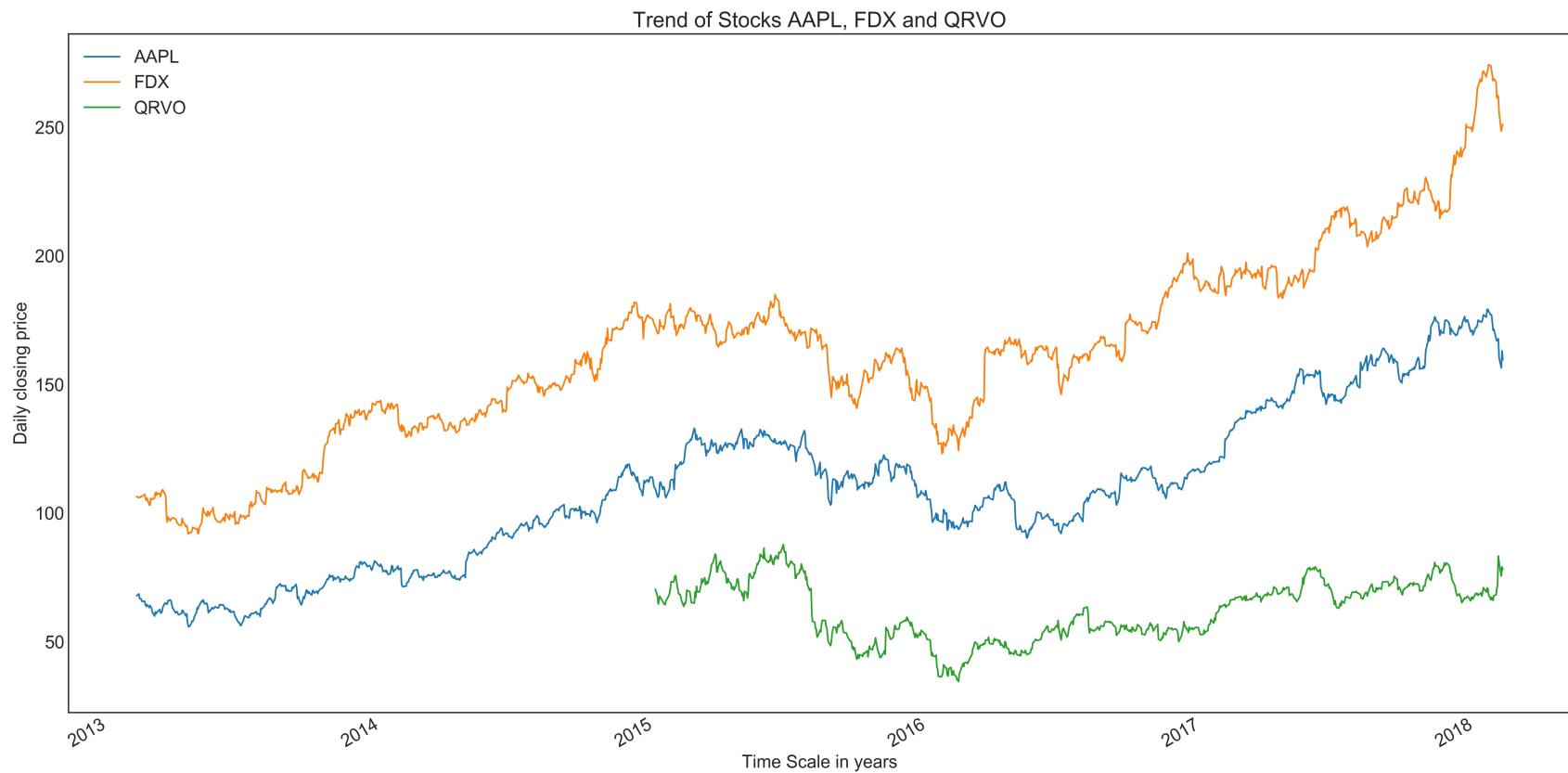


Figure 1: Stock prices for Apple, FedEx and Qorvo Inc.

Exploratory Data Analysis

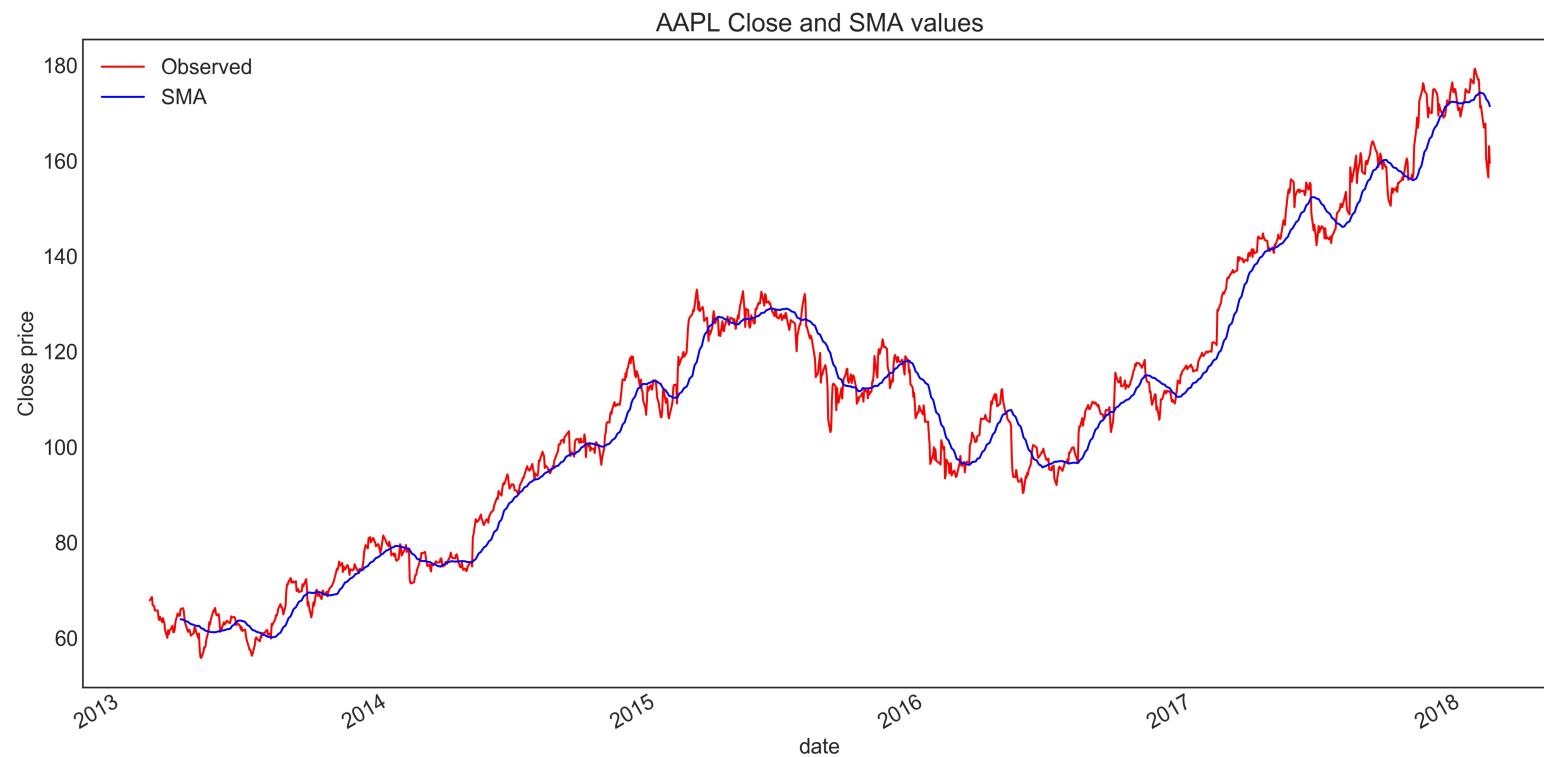


Figure 2: Actual data vs. SMA for Apple stock



Exploratory Data Analysis

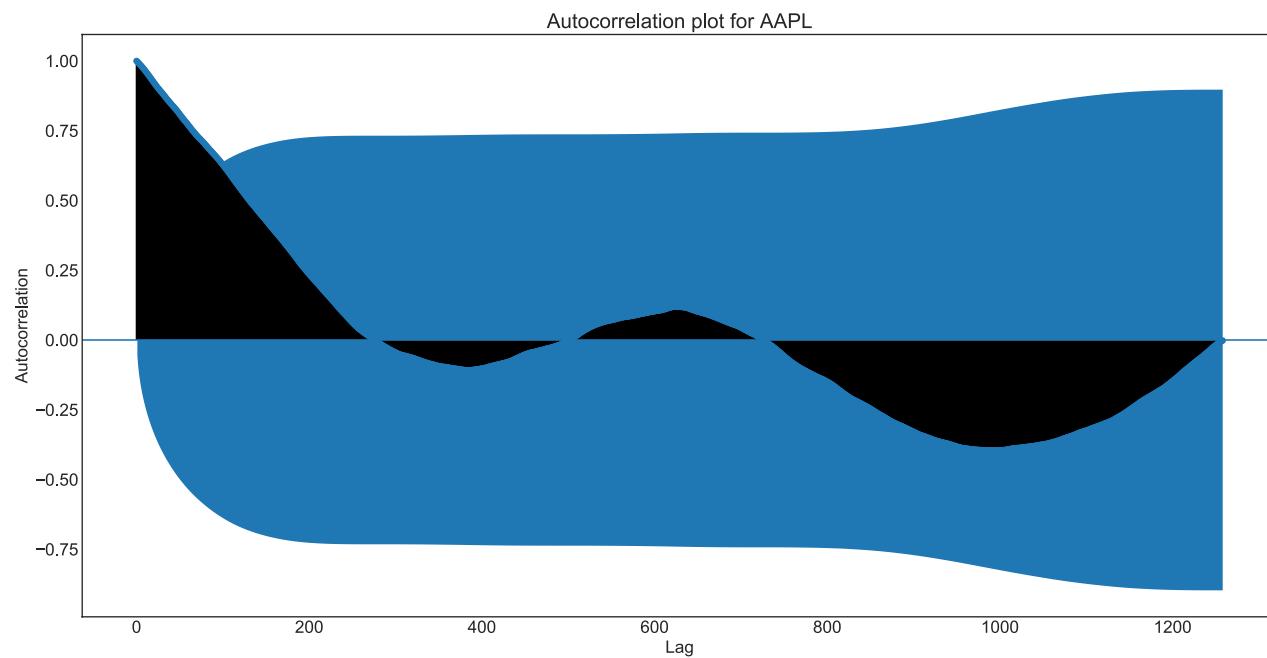


Figure 3: ACF Plot for Apple stock

Exploratory Data Analysis

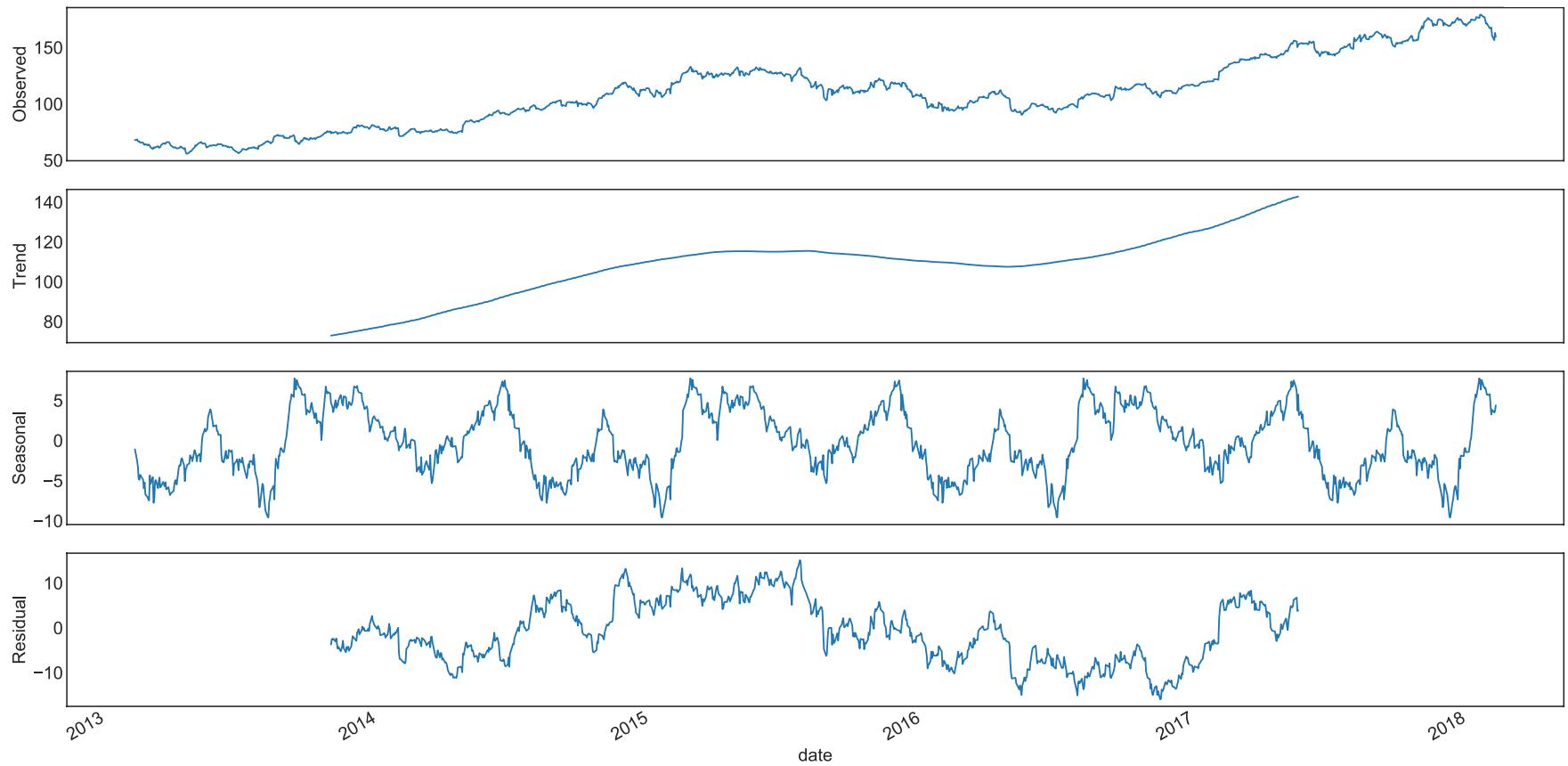


Figure 4: Decomposition Plot for AAPL

The ‘How’ of modeling data

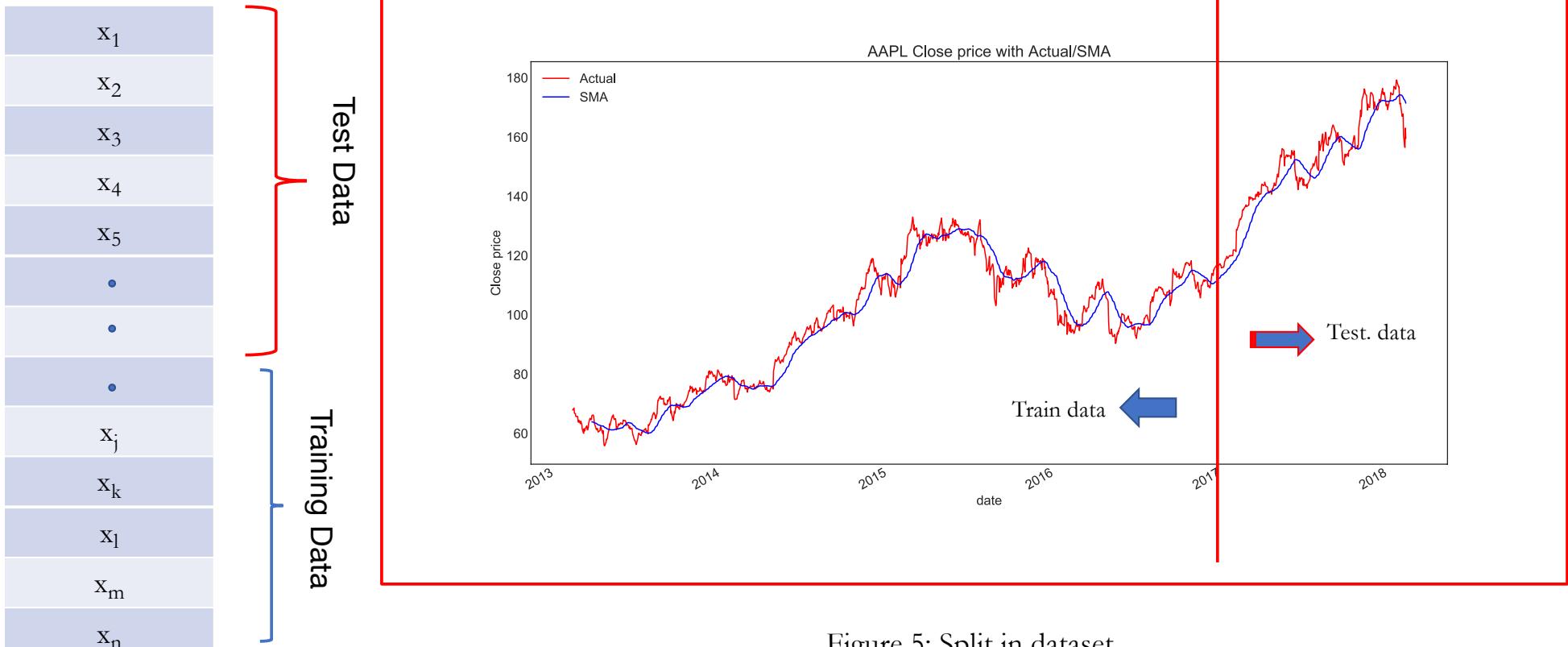


Figure 5: Split in dataset

Forecasting model with ARIMA

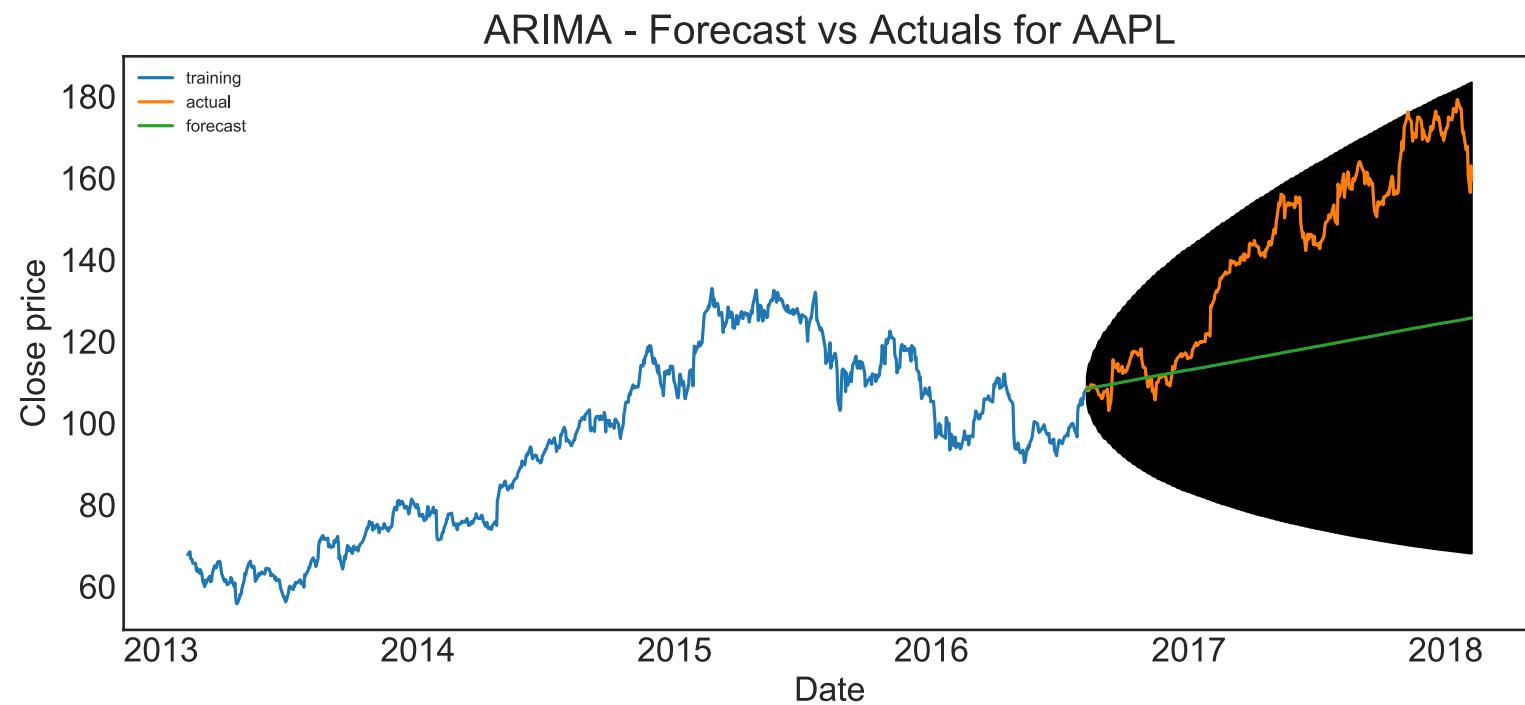


Figure 6: ARIMA Forecasting with data for Apple stock



Sample output from ARIMA

```
ARIMA params for AAPL = (0, 0, 0) - AIC: 12186.408478862091
ARIMA params for AAPL = (0, 0, 1) - AIC: 10526.589425187827
ARIMA params for AAPL = (0, 1, 0) - AIC: 4716.111652325981
ARIMA params for AAPL = (0, 1, 1) - AIC: 4717.799319165133
ARIMA params for AAPL = (0, 1, 2) - AIC: 4718.540181220545
ARIMA params for AAPL = (0, 2, 0) - AIC: 5562.699826380353
ARIMA params for AAPL = (0, 2, 1) - AIC: 4722.420589683508
ARIMA params for AAPL = (0, 2, 2) - AIC: 4724.079583482782
ARIMA params for AAPL = (1, 0, 0) - AIC: 4730.643045240879
ARIMA params for AAPL = (1, 0, 1) - AIC: 4732.208275747267
ARIMA params for AAPL = (1, 0, 2) - AIC: 4733.1566003511625
ARIMA params for AAPL = (1, 1, 0) - AIC: 4717.819594814397
ARIMA params for AAPL = (1, 2, 0) - AIC: 5243.601187272494
ARIMA params for AAPL = (1, 2, 1) - AIC: 4724.100659477893
ARIMA params for AAPL = (1, 2, 2) - AIC: 4724.344488071991
ARIMA params for AAPL = (2, 0, 0) - AIC: 4732.234227436988
ARIMA params for AAPL = (2, 0, 1) - AIC: 4732.512342291195
ARIMA params for AAPL = (2, 0, 2) - AIC: 4735.128522049959
ARIMA params for AAPL = (2, 1, 0) - AIC: 4718.454914364371
ARIMA params for AAPL = (2, 1, 1) - AIC: 4720.405935267261
ARIMA params for AAPL = (2, 2, 0) - AIC: 5093.838420056897
ARIMA params for AAPL = (2, 2, 1) - AIC: 4724.798194049899
ARIMA params for AAPL = (2, 2, 2) - AIC: 4726.761465921321
```

Figure 7: ARIMA Forecasting with data for Apple stock

Forecasting with SMA for AAPL with ARIMA

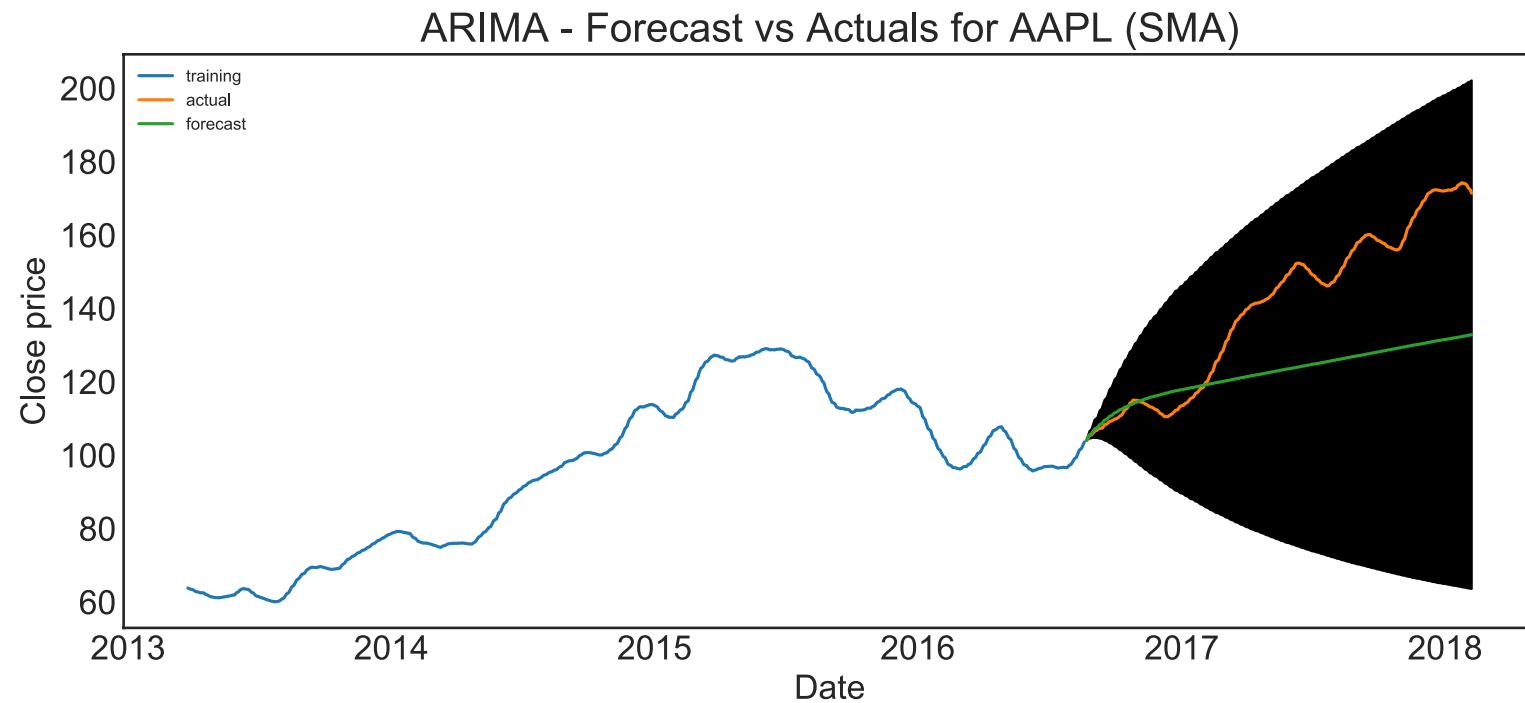


Figure 8: ARIMA Forecasted data of SMA for Apple stock

Forecasting model with SARIMA

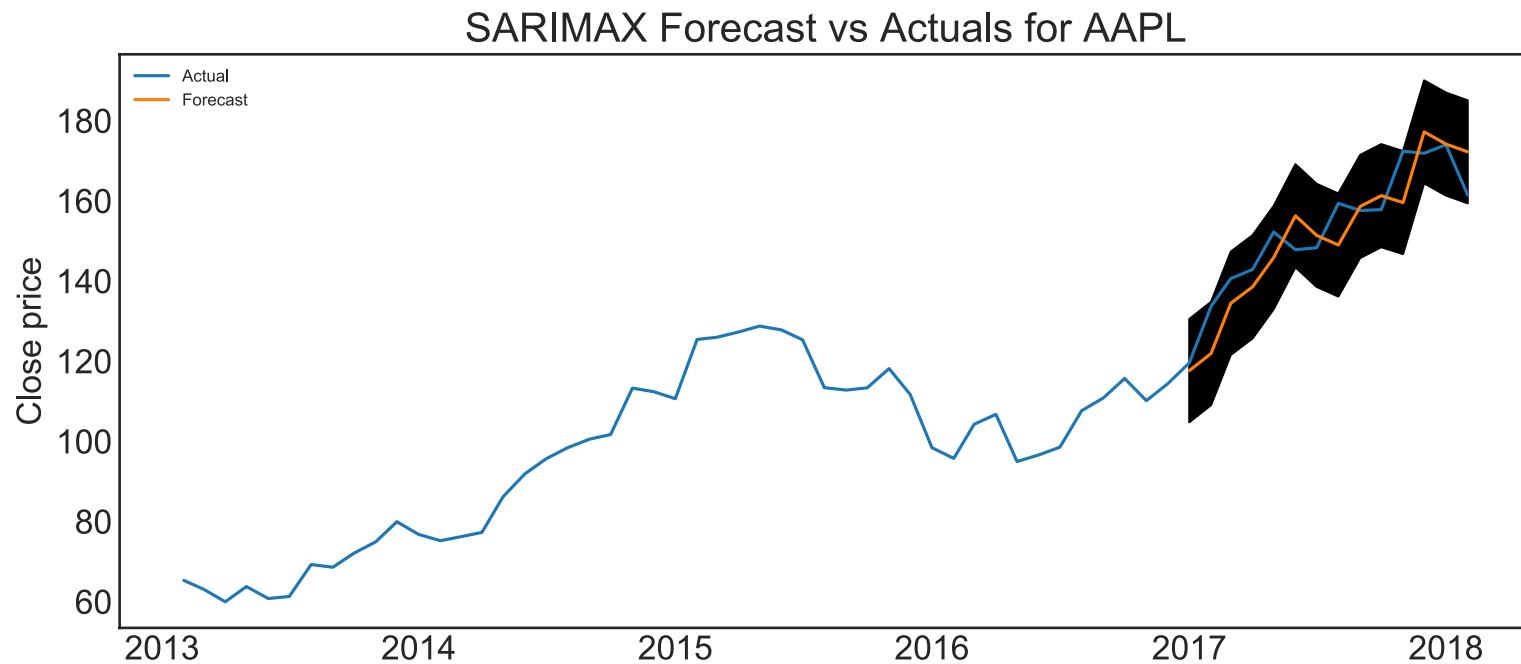


Figure 9: SARIMA Forecasted data for Apple stock

Forecasting model with Linear Regression

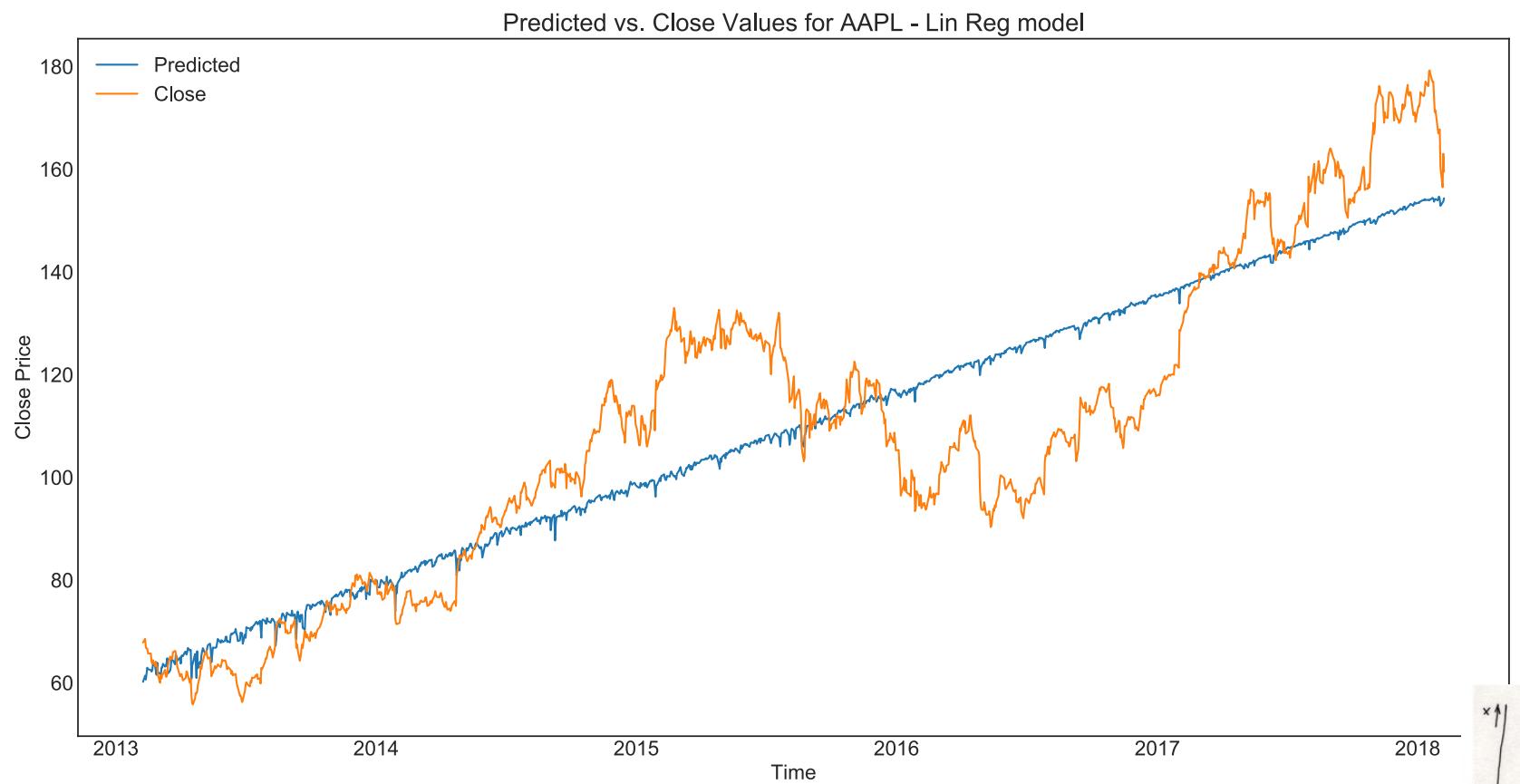
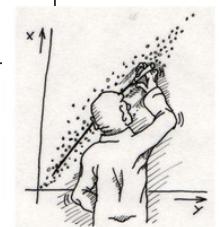


Figure 10: Linear Regression forecast for Apple stock



Forecasting model with LSTM

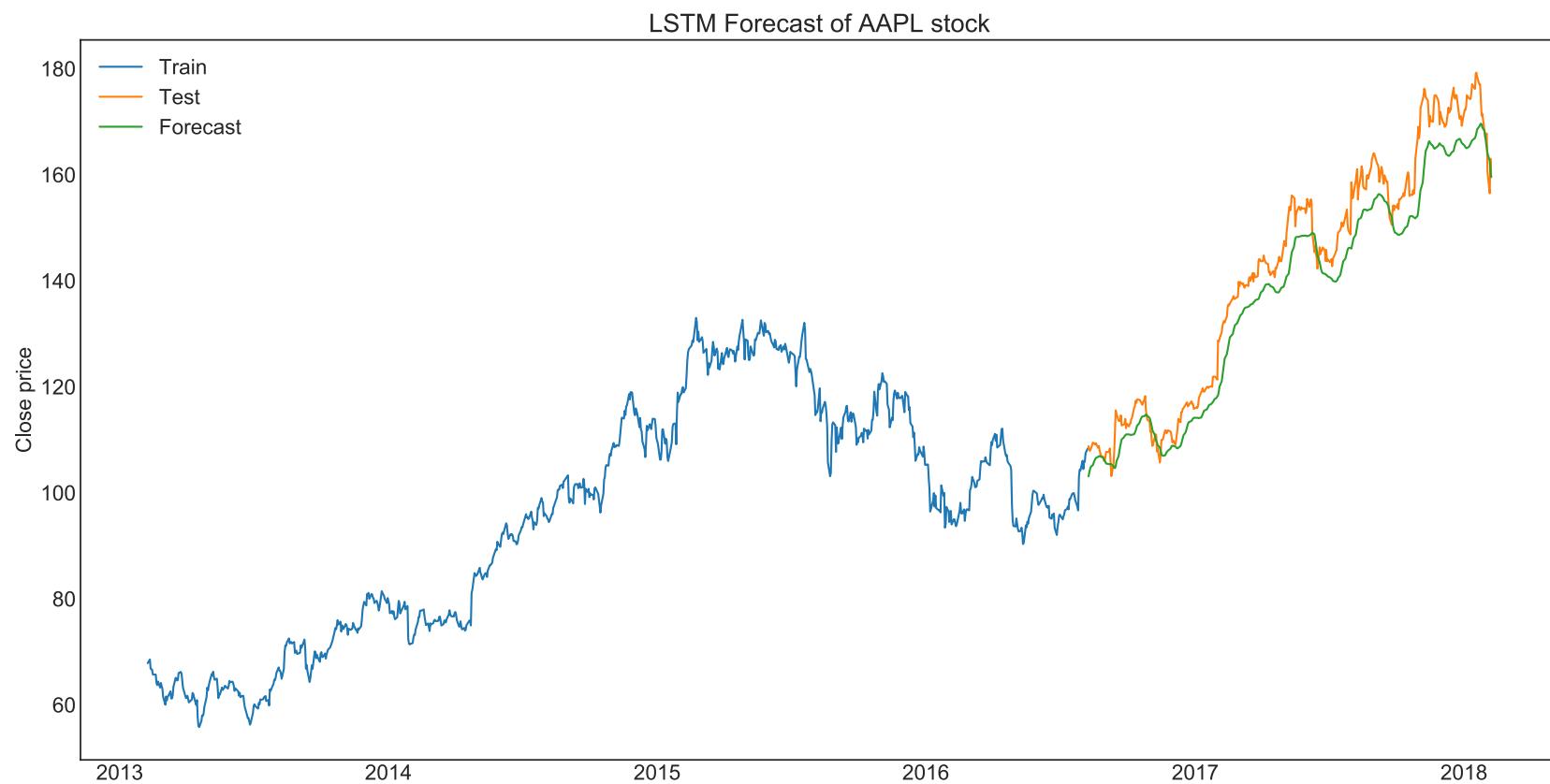


Figure 11: LSTM forecasted data for Apple stock

K Keras

Forecasting model with Random Forest

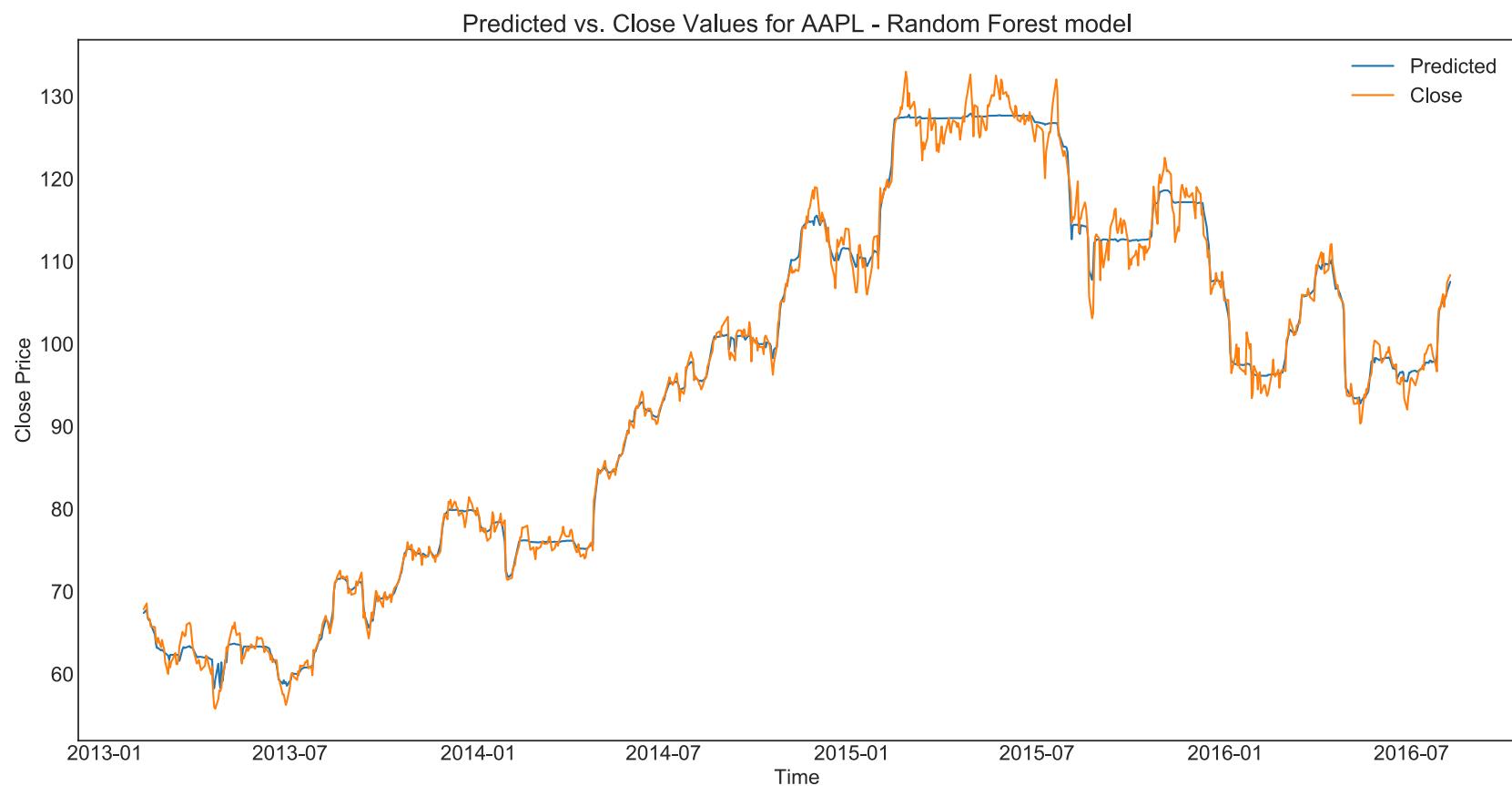


Figure 12: Random Forest forecasted data for Apple stock

RMSE values for various models on AAPL

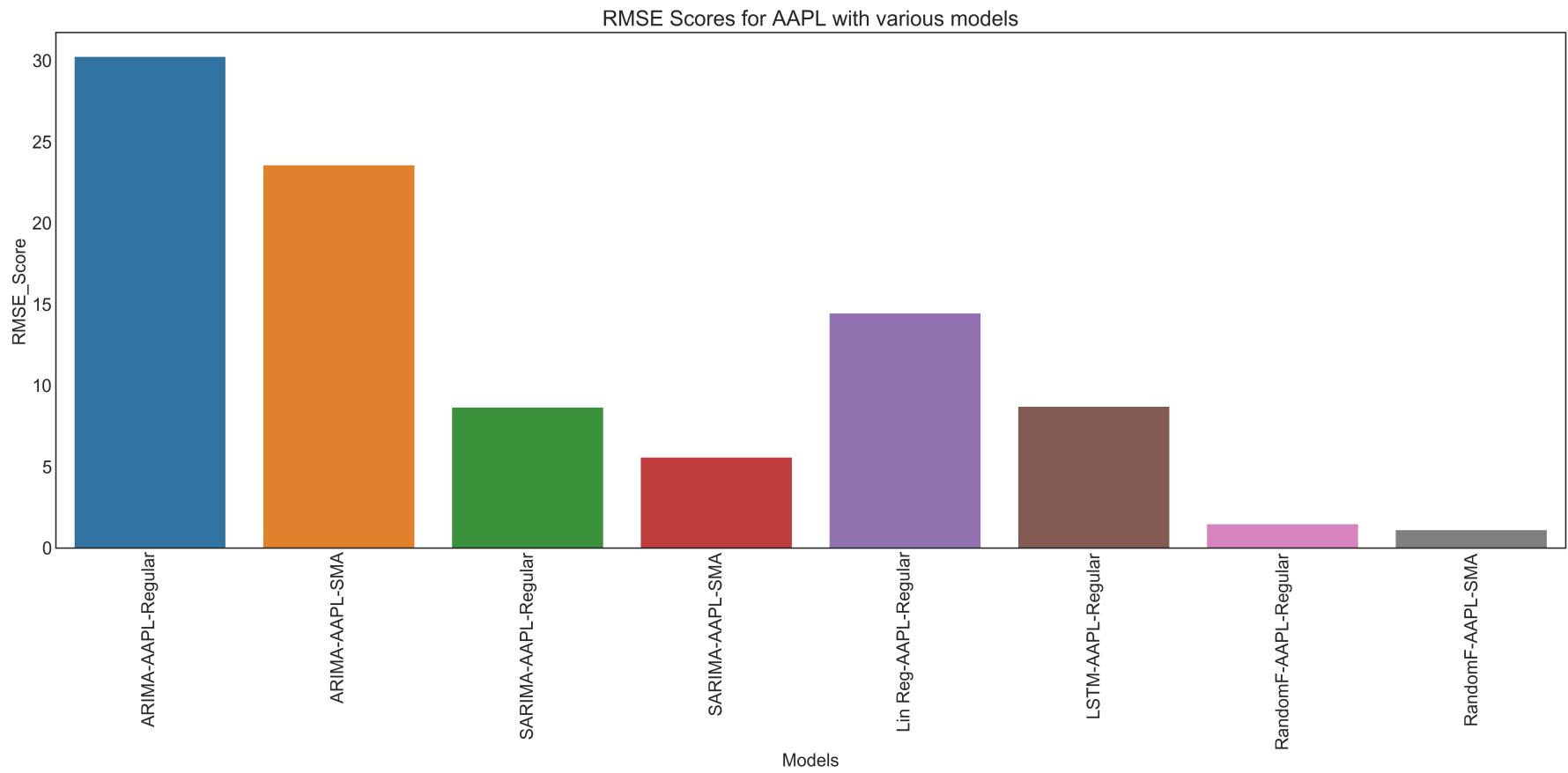


Figure 13: RMSE Scores by model for AAPL

RMSE values for various models on FDX

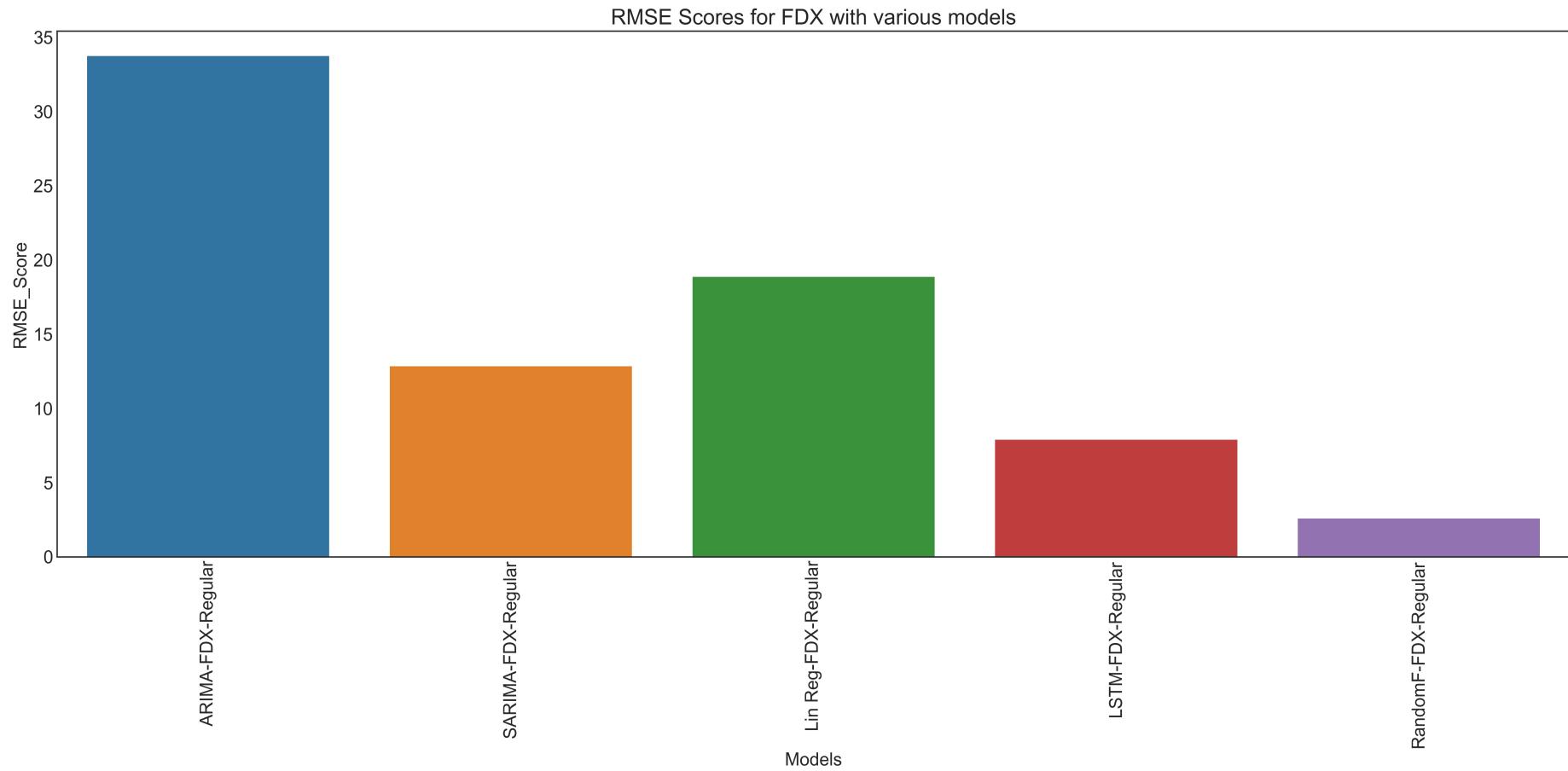


Figure 14: RMSE Scores by model for FDX

RMSE values for various models on QRVO

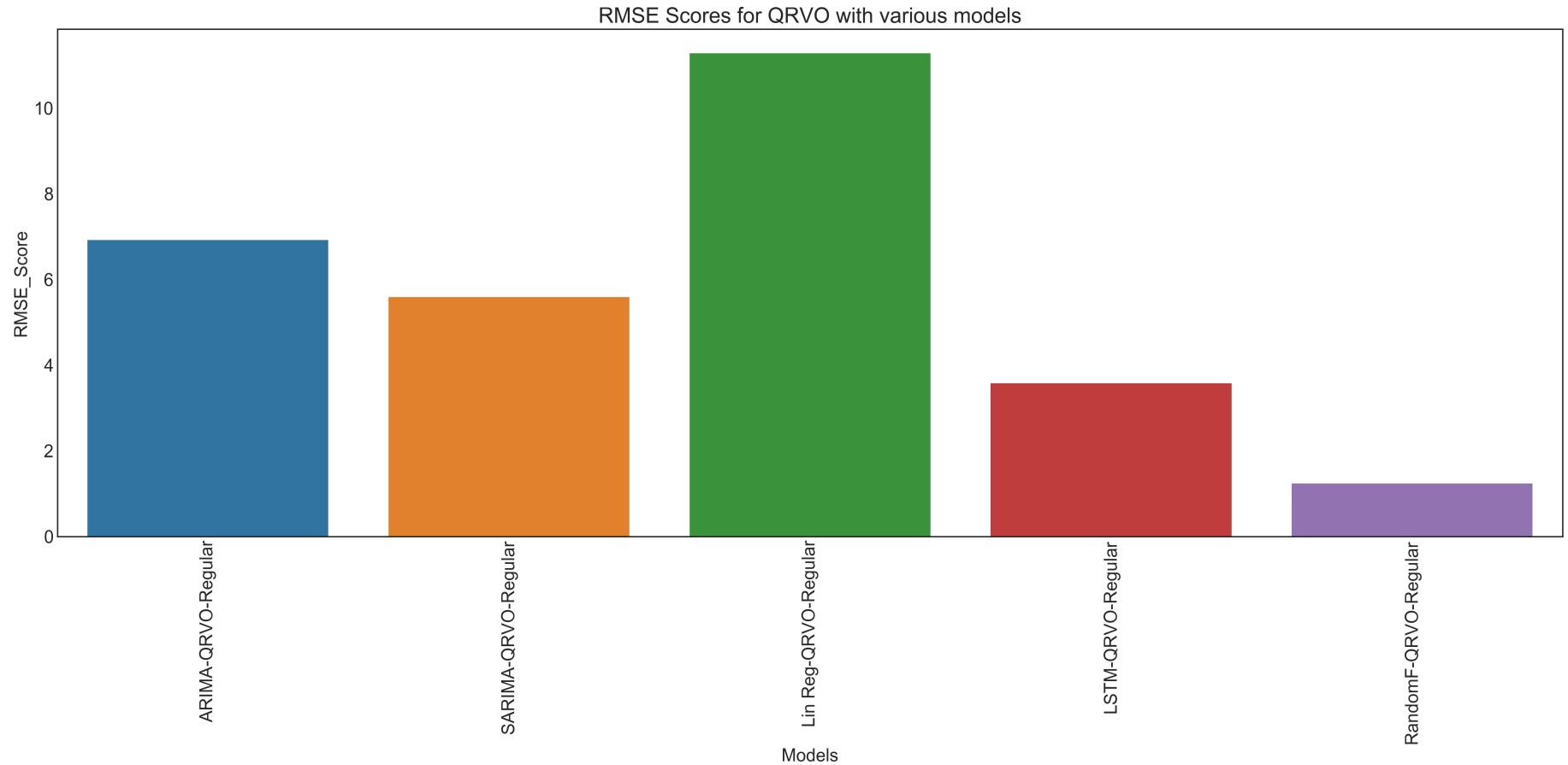


Figure 15: RMSE Scores by model for QRVO

Takeaways

- Great! We know the secret sauce to become rich!!! 😊
- Stock price fluctuations depend on a lot of factors and not just volume of stocks traded and close price
- We could:
 - Do sentiment analysis about a company using NLP by scraping the web
 - Add holidays and other seasonal factors as a feature
- We need to combine domain knowledge with specialist technical skills to deliver high-impact statistical insights, predictive models, and data-driven business advice.
- The possibilities are endless