Daniel Lesh

Zhenhao Gong

ECON 3133 - Elementary Economic Forecasting

9 May 2021

Economic Forecasting: Apple (*AAPL*) Stock Price Prediction

Introduction

Since the founding of the New York Stock Exchange (NYSE) in May of 1792, investors have worked tirelessly to try and accurately predict the price of stocks using a variety of different statistical and mathematical tools. In the age of technology, programming languages such as R and Python are being utilized to try and make better price predictions and forecasts of different stocks daily. With harboring the power of complex statistical and mathematical formulas within these programming languages, investors both small and large are able to make better investment decisions with different forecasting models.

In this project, I chose to build a forecasting model to try and predict the price of Apple (AAPL) stock from *2021.01* to *2021.12*. Since we are already almost halfway through the year of 2021, I plan to compare the price of AAPL from January 2021 to May 2021 with the forecasting predictions the models I have developed provide.

I decided to use Apple's stock for two reasons: I have a large position in Apple's stock currently, and Apple is also one of my all-time favorite companies. They have revolutionized humanity with their products, and I believe their future developments are key to the success of the human race.

In Q1 2020, the onset of the COVID-19 Pandemic caused the stock market to collapse. On March 9, 2020, the Dow Jones Industrial Average fell by over 2,000 points - the highest single day point drop in U.S. stock market history. During this same time, the price of AAPL stock fell to $57.09, its lowest price since 2019. With these volatile price changes, the model selection process was key in accounting for attributing external factors to both increases and decreases in the price of Apple stock. With this in mind, I will outline the model selection process to propose the best model that fits the AAPL stock data.

Data Description

The data set I used is from yahoofinance.com. Yahoo Finance provides real-time live data of all current stocks listed on the NYSE, along with publishing financial news and articles. The data set has 132 observations. These observations are the closing price of AAPL on the first day of every month from 2010.01 to 2020.12. I used only 10 years of pricing data for AAPL due to the relatively small market capitalization of AAPL stock pre-2010 (compared to its over 2 trillion dollar market capitalization today).

There is one listed variable in this data set: Closing_Price. To be able to attribute time to the 132 closing prices listed from 1:132 in the data set, I created a subset of data designated as 'tsaapl'. I was able to do this using the *ts()* function in R, using a start date of 2010.01 and a frequency of 12. When executing the code 'tsaapl' within the R console, an output of 132 observations with each month from 2010.01 to 2020.12 is provided.

As can be examined with the summary stats output of the 'tsaapl' data set in (***Figure 1***), there is a minimum value of 6.86, a maximum value of 132.69 and a mean value of 34.08. There is also considerable variation within the data set - providing evidence as to some of the volatility within AAPL price over the past 10 years (***Figure 2***).

```
> summary(tsaapl) # Summary stats for time series AAPL data
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  6.859  17.307  27.124  34.088  42.272 132.690
```

Figure 1

```
> var(tsaapl) # 636.1989
[1] 636.1989
> sd(tsaapl)  # 25.22298
[1] 25.22298
```

Figure 2

The time series plot of the AAPL stock's historical price data provides further visualization of the volatility within the data set (***Figure 3***). Although slight volatility appears from 2012 to 2018, large increases to the stock price are produced from the beginning of 2018 to the end of 2020. This attributes to the large price variation in the data set . The AAPL stock time series does not appear to be time series correlated, as no seasonal factors and patterns are attributed to the data set since the fluctuation in the price of the stock is random and independent of any seasonal changes.

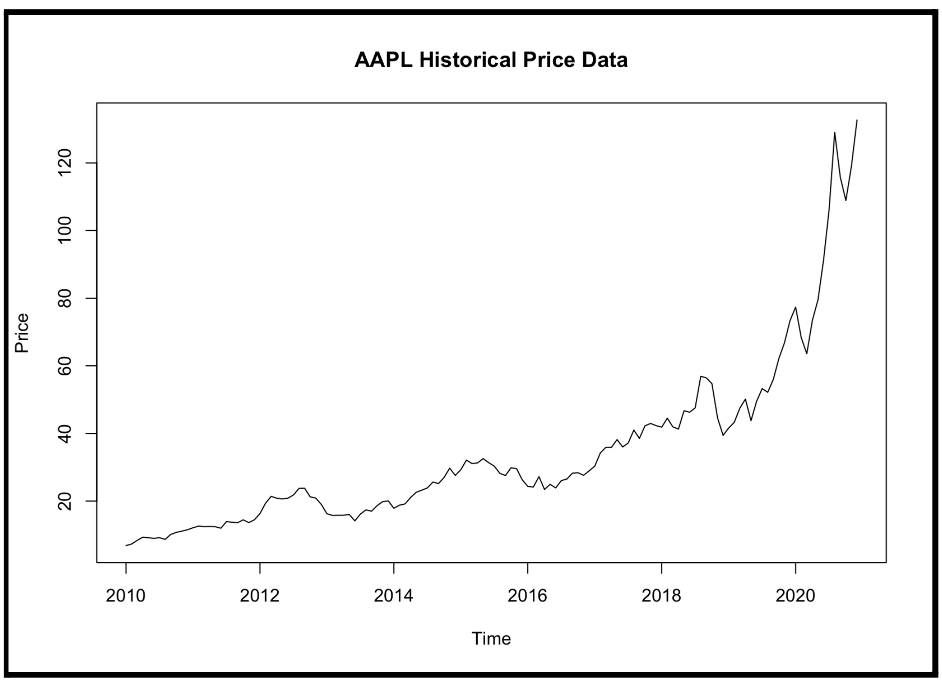


Figure 3

Model Selection

To forecast the price of AAPL stock from 2021.01 - 2021.12, I created 7 forecasting models within RStudio[1]. These models were assessed using a variety of different model selection criteria and diagnostic statistics. The 3 most commonly used model selection techniques used in my R Script were the *Akaike Information Criterion*, the *Bayesian Information Criterion*, and the *Durbin-Watson Test Statistic*. Below in **Table 1** are the results of these model selection criterion used on the forecasting models I produced:

Table 1: AIC, BIC, and DW Test Model Selection Criterion

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AIC	874.16	757.37	794.14	795.86	728.46	-58.22	-172.99
BIC	891.46	754.49	797.01	801.58	742.8	-49.58	-164.34
DW Test						0.17	0.15

Before diving into each model, it is important to note why I selected each ARIMA model I wanted to use for the forecasting process of AAPL price predictions. I utilized the ACF and PACF autocorrelation functions to derive an output with regard to what models would best fit the data set. With a slightly descending lag from 0.0 - 2.0 in the ACF plot (***Figure 4***) and a sharp cut off from lags 0.1 - 0.2 in the PACF plot (***Figure 5***), I knew that the best model to fit this data set would be an Autoregressive (AR) or Autoregressive Moving Average (ARMA) model. A visualization of the ACF and PACF plots can be found below:

[1] Model (1) is a MA(4) model. Model (2) is an AR(1) model. Model (3) is an AR(2) model. Model (4) is an ARMA(3,1) model. Model (5) is an Auto Arima model. Model (6) is a Log-Linear trend model. Model (7) is a Cycle Trend model.

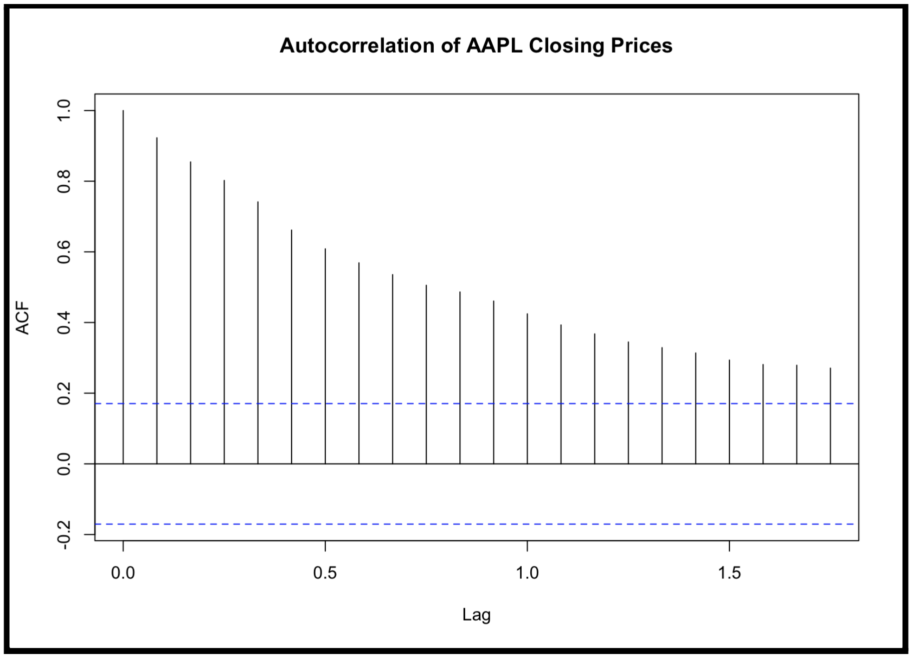


Figure 4

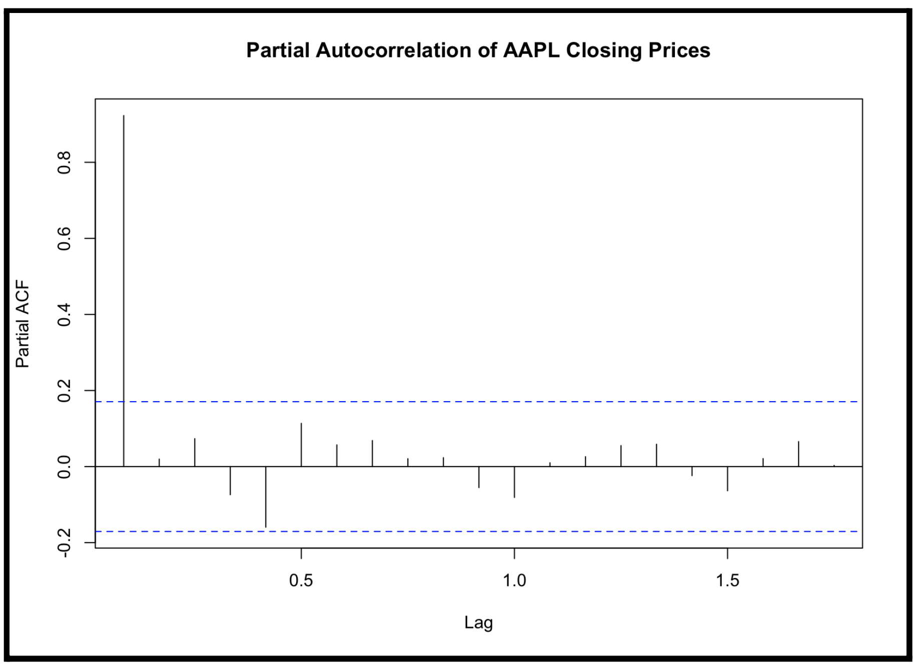


Figure 5

Forecasting Models

The first ARIMA model that I programmed in R was a Moving Average (MA) 4 model (***Figure 6***). Although the ACF and PACF plots provided indication that the best ARIMA model to fit this time series data set would be an AR model, I wanted to include an MA(4) model to see if it could potentially provide a good fit for the forecasting model. With the highest AIC/BIC scores and the volatile residual values, the MA(4) model proved not to be a good fit for the forecasting model. Because of this, I decided not to utilize the MA(4) model in the AAPL price prediction.

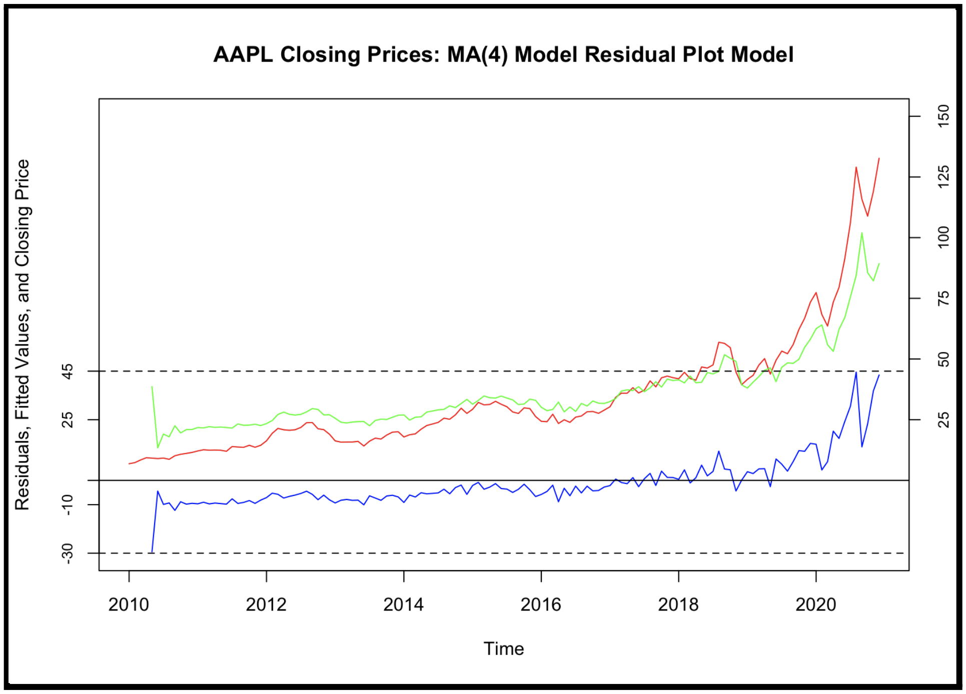


Figure 6

The second ARIMA model that I programmed was an AR(1) model (***Figure 7***). As was shown earlier, the ACF and PACF plots provided indication that the best ARIMA model to fit this time series data set would be an AR model. With the lowest AIC/BIC scores and least volatile residual values out of all the ARIMA models tested, the AR(1) model proved to be a good fit for the forecasting model. Because of this, I decided to utilize the AR(1) model in the AAPL price prediction for months 2021.01 - 2021.02 (2-month ahead forecast).

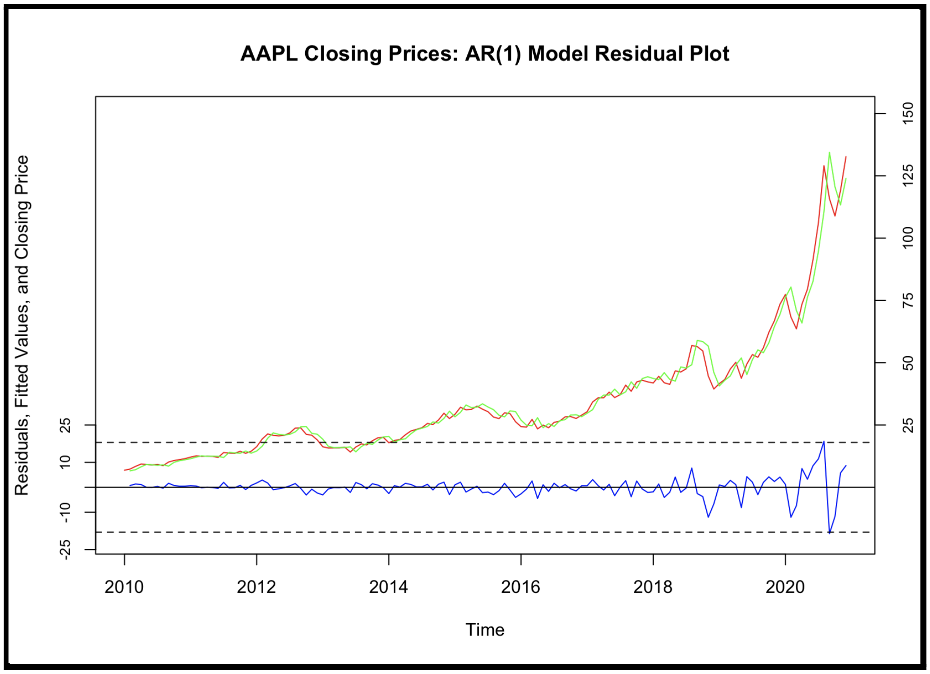


Figure 7

The third and fourth ARIMA models that I programmed were the AR(2) and ARMA(3,1) models (***Figure 8*** & ***Figure 9***). The AR(2) and ARMA(3,1) models had relatively the same AIC/BIC scores. The plotted residual values for both of the models were practically the same as well. This is due to the close relationship both models shared after the 0.1 cut-off value in the PACF plot. Since the lag was brought to zero almost immediately after 0.1, the models share a lot of the same characteristics with each other. With this in mind, I decided to utilize the AR(2) model in the AAPL price prediction for the 6-month and 12-month ahead forecasts.

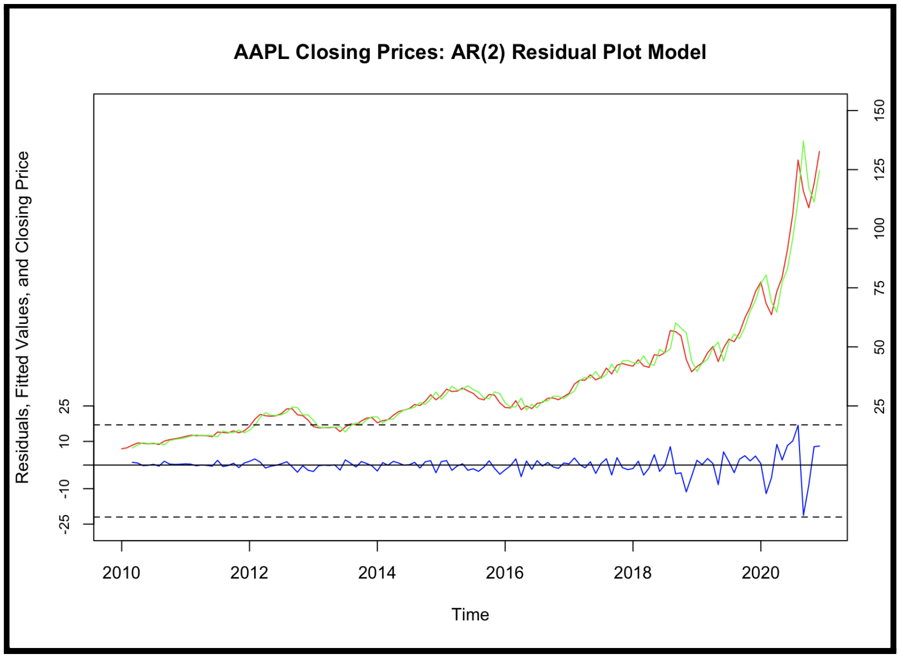


Figure 8

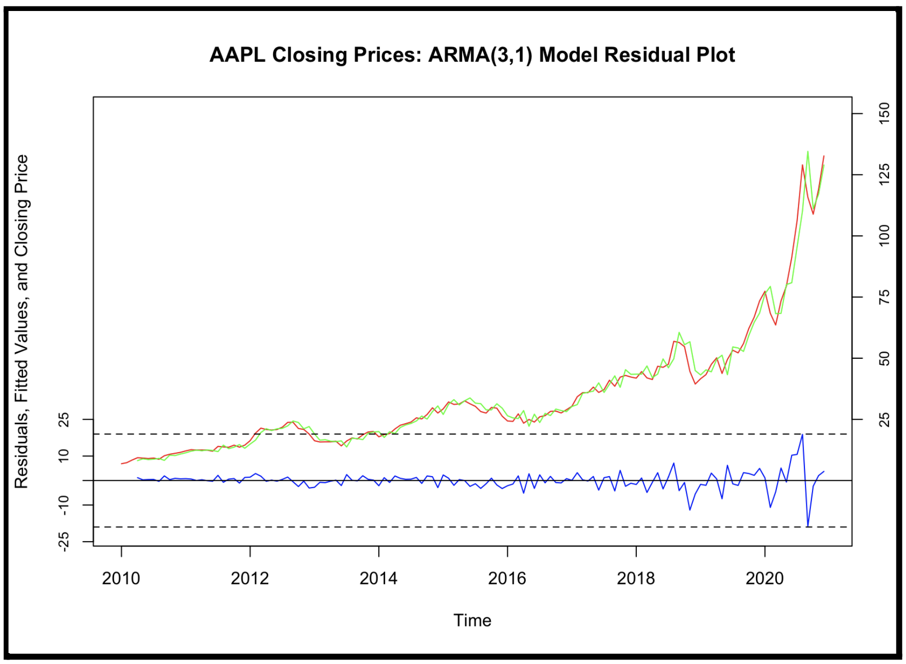


Figure 9

The fifth ARIMA model that I programmed was an Auto ARIMA model (***Figure 10***). The Auto ARIMA model had the lowest AIC/BIC scores out of the the 5 ARIMA models. The purpose of creating an Auto ARIMA model was to highlight the residual values within the ARIMA(1, 1, 1) plot. Essentially, I utilized the Auto ARIMA model as a cross-check in the model selection process for the AR(1) and AR(2) models. Since it reestablished the AR models as the best ones to use for the forecasting along with the AIC/BIC model selection criterion, I used the AR(1) and AR(2) models in the price prediction forecast for AAPL stock.

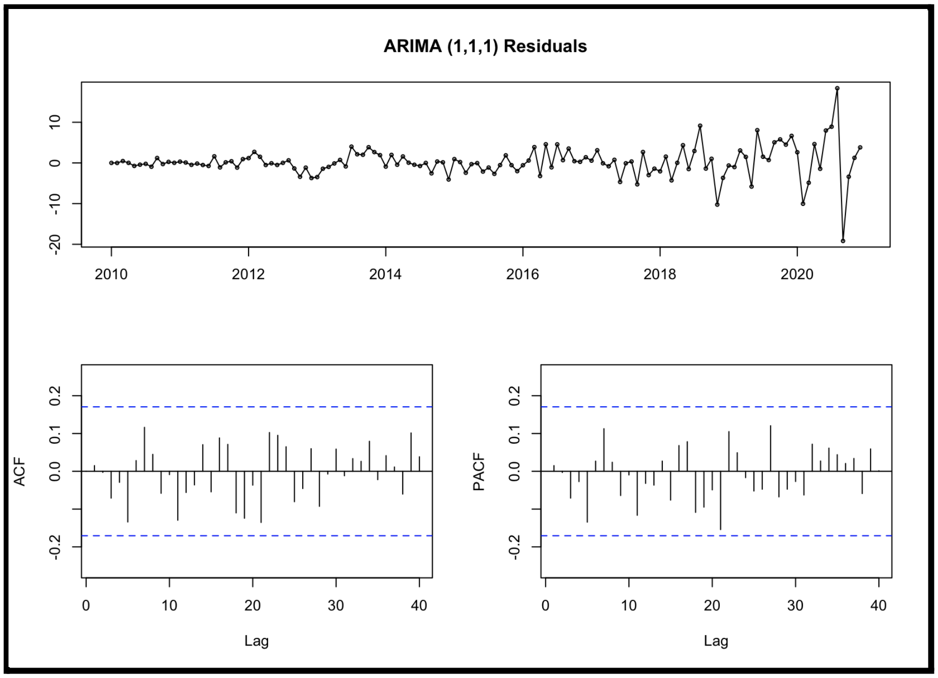


Figure 10

Results

The AR(1) and AR(2) models were used to predict the price of AAPL stock in three different h-step ahead forecasts: 2-month (2021.01 - 2021.02), 6-month (2021.01 - 2021.06), and 12-month 2021.01 - 2021.12).

The following output below is from the AR(1) 2-month ahead forecast (***Figure 11***). Since the AR(1) model only produces 1 lag, the forecast method cannot directly predict the price frontier ahead of the last month of record (2020.12), but it can use point forecasting that is useful with price prediction. For the month of February 2021, the point forecasting output of $132.69 was only less than $1.45 dollars off the closing price of AAPL stock on February 1st, 2021 ($134.14).

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2021	132.69	127.2091	138.1709	124.3077	141.0723
Feb 2021	132.69	124.9389	140.4411	120.8357	144.5443

Figure 11

The next output is from the AR(2) 6-month ahead forecast (***Figure 12***). Since the AR(2) model produces multiple lags, we are able to utilize the model for forecasting ahead of a short-term period. The point forecast values for the months of January 2021 - May 2021 were very high. These high forecast values can be attributed to the immense increase in value the AAPL stock price saw from 2019 through the midway point of 2021. In comparison to the actual values of AAPL stock from January 2021- May 2021, the point forecasting price targets missed. What managed to be closer to predicting the actual values of the price of AAPL stock was the 'Lo 95' forecast values. In between the months of April 2021 and May 2021, the values were close to meeting the interval forecasting targets of $132.58 and $126.86. This shows that although the initial price target may have missed, the model still has encapsulated some of the current price prediction of AAPL stock during the months of January 2021 - May 2021 through its 'Lo 95' interval forecast.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2021	146.33	139.8037	152.8563	136.3489	156.3111
Feb 2021	159.97	145.3767	174.5633	137.6515	182.2885
Mar 2021	173.61	149.1908	198.0292	136.2641	210.9559
Apr 2021	187.25	151.5039	222.9960	132.5811	241.9189
May 2021	200.89	152.4896	249.2904	126.8680	274.9120
Jun 2021	214.53	152.2730	276.7870	119.3161	309.7438

Figure 12

The next output is from the AR(2) 12-month forecast (***Figure 13***). The point forecast values for the months of January 2021 - December 2021 were once again, very high. Even though the forecasted values could be attained in future years for AAPL stock, these values are in no way feasible for AAPL stock to hit in the fiscal year of 2021. What is interesting to see though is that as similar to the 6-month ahead forecast, the ‘Lo 80’ and 'Lo 95’ interval forecast values show more accurate price predictions for what could come of AAPL stock for the rest of this year. Both the ‘Lo 80’ and ‘Lo 95’ forecasts tend to show a declining trend in price of AAPL stock from May 2021 - Dec 2021. These ‘Lo’ interval forecasts may be a better predictor of the price target of AAPL stock in the coming months.

Forecasts:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Jan 2021	146.33	139.8037	152.8563	136.34888	156.3111
Feb 2021	159.97	145.3767	174.5633	137.65153	182.2885
Mar 2021	173.61	149.1908	198.0292	136.26406	210.9559
Apr 2021	187.25	151.5039	222.9960	132.58113	241.9189
May 2021	200.89	152.4896	249.2904	126.86801	274.9120
Jun 2021	214.53	152.2730	276.7870	119.31615	309.7438
Jul 2021	228.17	150.9497	305.3903	110.07175	346.2682
Aug 2021	241.81	148.5957	335.0243	99.25104	384.3689
Sep 2021	255.45	145.2733	365.6267	86.94924	423.9507
Oct 2021	269.09	141.0346	397.1454	73.24621	464.9338
Nov 2021	282.73	135.9244	429.5356	58.21018	507.2498
Dec 2021	296.37	129.9812	462.7588	41.90028	550.8397

Figure 13

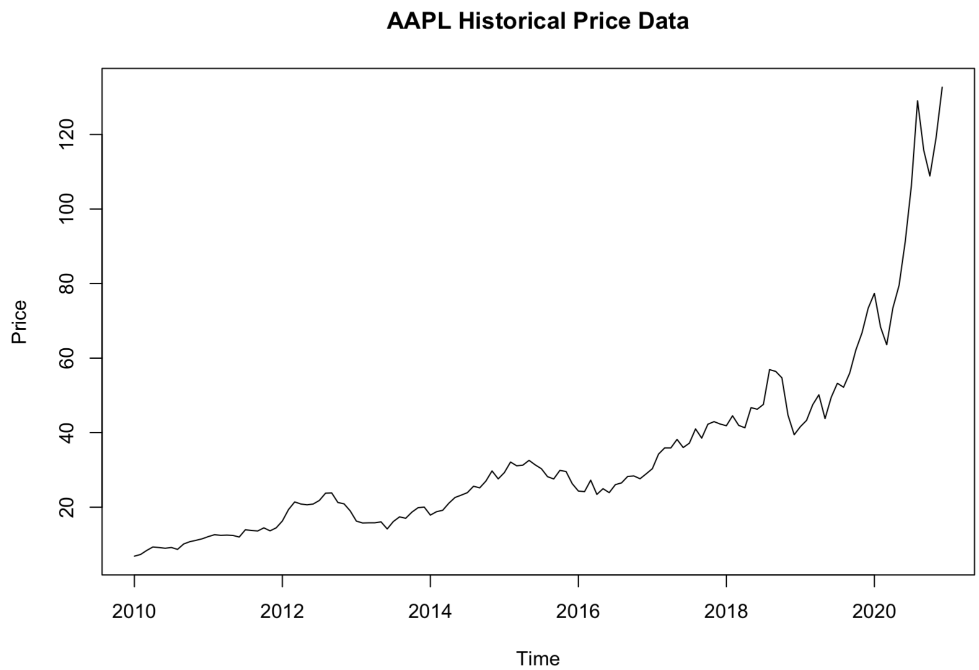
Conclusions

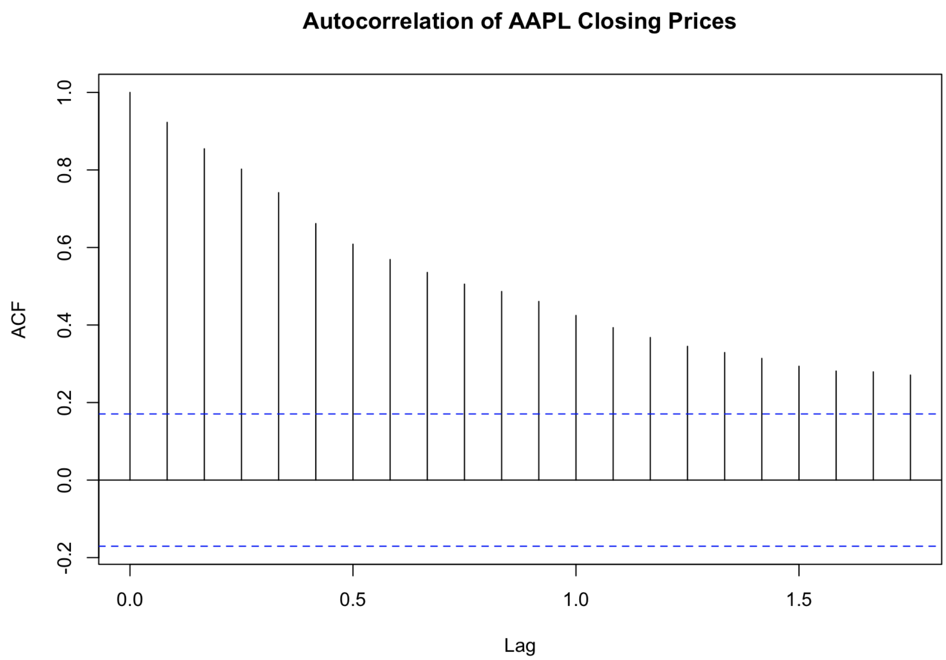
Predicting the price of a stock is never easy. When it came to predicting the price of AAPL’s stock, there are numerous considerations that can be made for why the ‘Lo’ interval forecast produced a more accurate price prediction than the actual point forecasting price target. In the past 2 years, the value of AAPL stock has grown immensely and has consistently broken it all time stock price highs. With this and the high AIC and BIC scores, model selection proved to be more difficult as compared to seasonal or less volatile time series data. Because of the high amounts of variation in the stock price and attributing external economic factors such as the market crash caused by the COVID-19 Pandemic, trying to predict the price of AAPL’s stock proved to be even more difficult. Some of the other models I created when logging the stock price of AAPL provided lower AIC and BIC scores, but were not feasible for forecasting due the same increasing residual trends found in the ARIMA models. Training a model (i.e. a machine learning model) to better predict the price of AAPL would be the most efficient way to try and predict the price of the stock. This model would be much more accurate in evaluating and reacting to trends found within the stock's data. Although our analysis was able to somewhat predict the price of AAPL stock correctly through May 2021, there are other potential models that would be able to provide more accurate price predictions in the coming months.

Sources

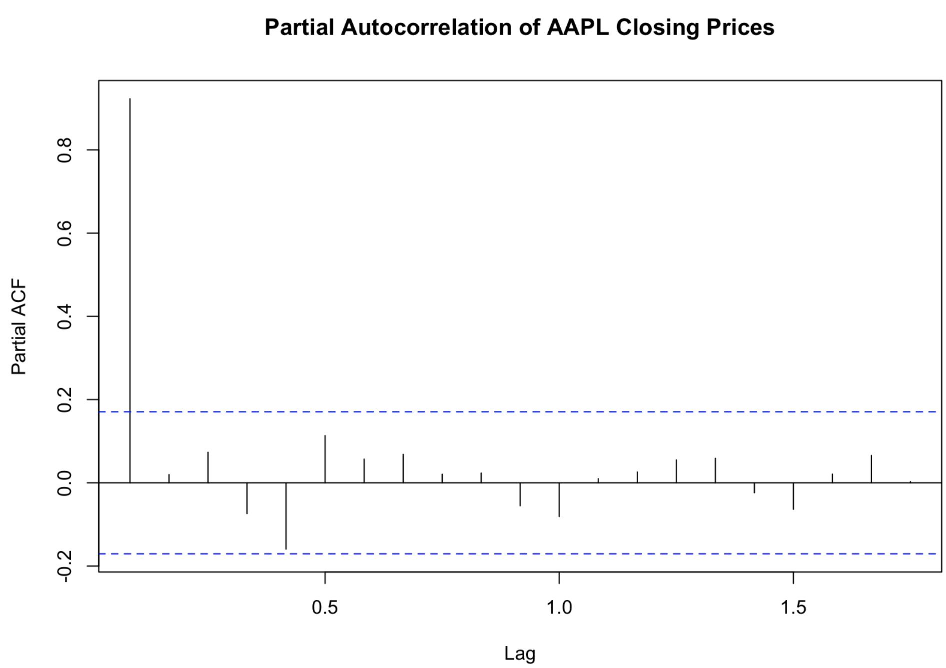
1. "The Birth of Stock Exchanges." *Investopedia*, www.investopedia.com/articles/07/stock-exchange-history.asp. Accessed 8 May 2021.
2. "Apple (AAPL) Stock Price Data" *Https://Finance.Yahoo.Com/Quote/AAPL/*, finance.yahoo.com/quote/AAPL. Accessed 8 May 2021.
3. Valk, Serge de. "Nowcasting: An R Package for Predicting Economic Variables Using Dynamic Factor Models." *R-Project.Org*, R-Project.Org, May 2019, journal.r-project.org/archive/2019/RJ-2019-020/RJ-2019-020.pdf.
4. Otani, Akane. "Covid-19 Fueled Stock Market Selloff Last Year. Here Are Some Lessons Learned." *WSJ*, 7 Mar. 2021, www.wsj.com/articles/covid-19-crashed-the-stock-market-a-year-ago-here-are-some-lessons-learned-11615113001.

Appendix

- ```
  > summary(tsaapl) # Summary stats for time series AAPL data
     Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
    6.859  17.307  27.124  34.088  42.272 132.690 
  ```
- ```
  > var(tsaapl) # 636.1989
  [1] 636.1989
  > sd(tsaapl)  # 25.22298
  [1] 25.22298
  ```
- 

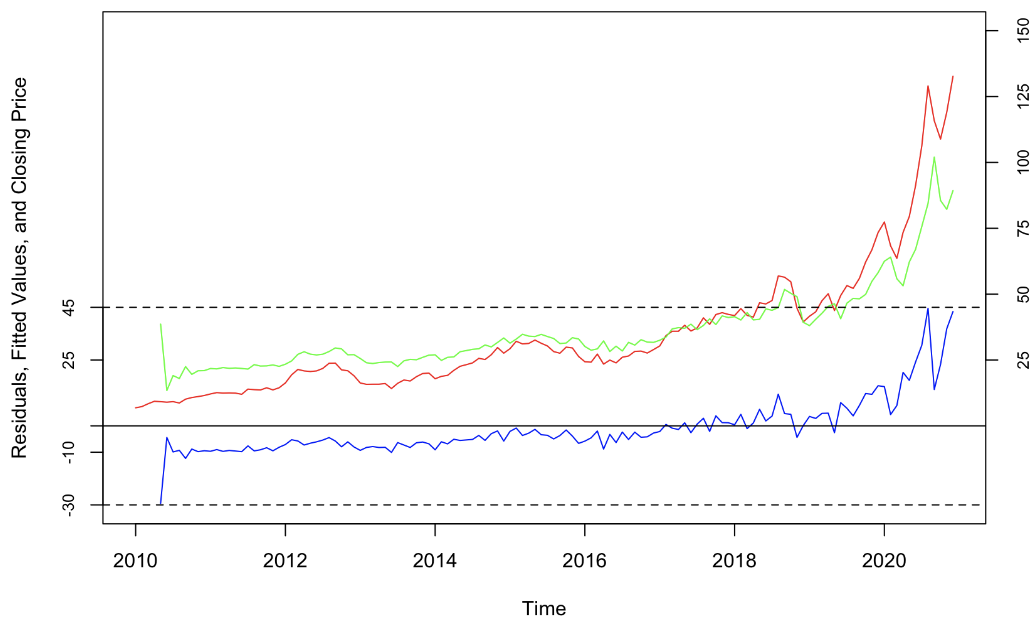

Autocorrelation of AAPL Closing Prices
ACF
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
0.0
0.5
1.0
1.5
Lag

-


Partial Autocorrelation of AAPL Closing Prices
Partial ACF
0.8
0.6
0.4
0.2
0.0
-0.2
0.5
1.0
1.5
Lag

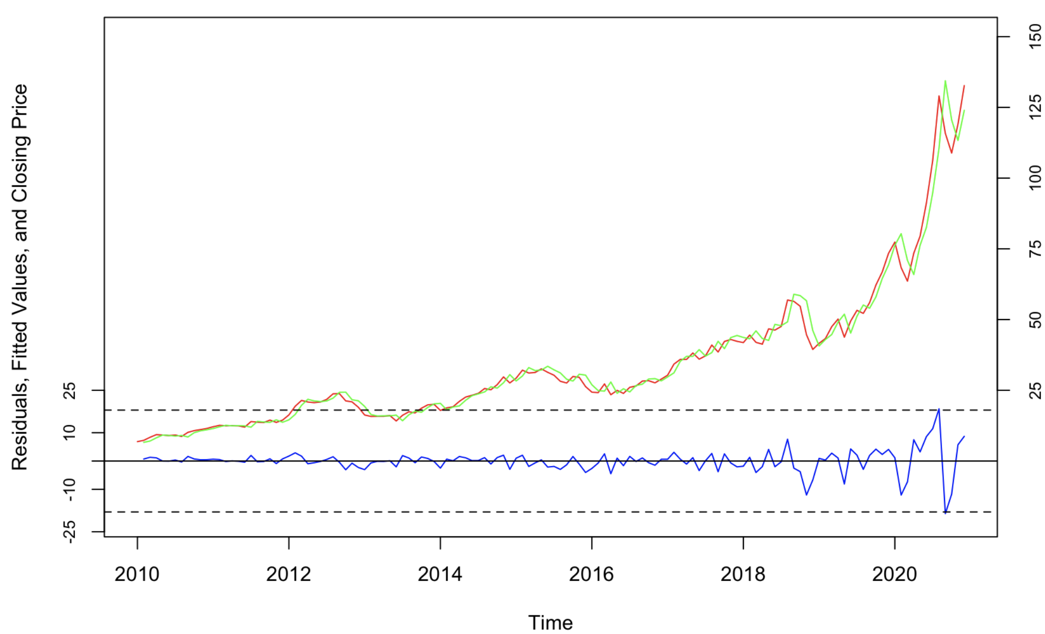
-

AAPL Closing Prices: MA(4) Model Residual Plot Model

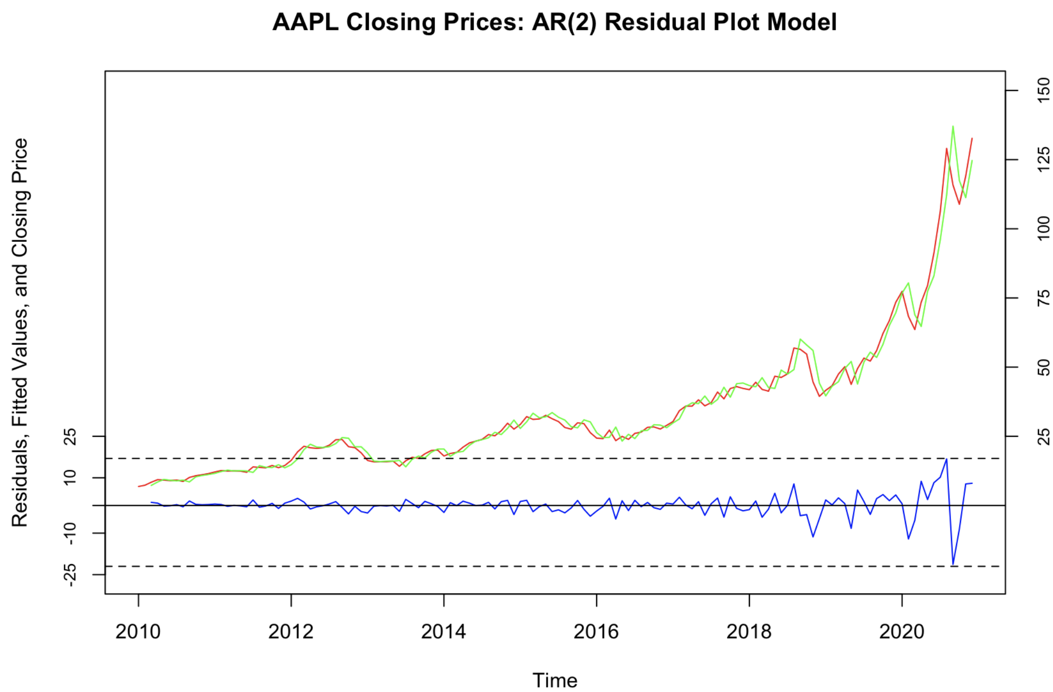


-

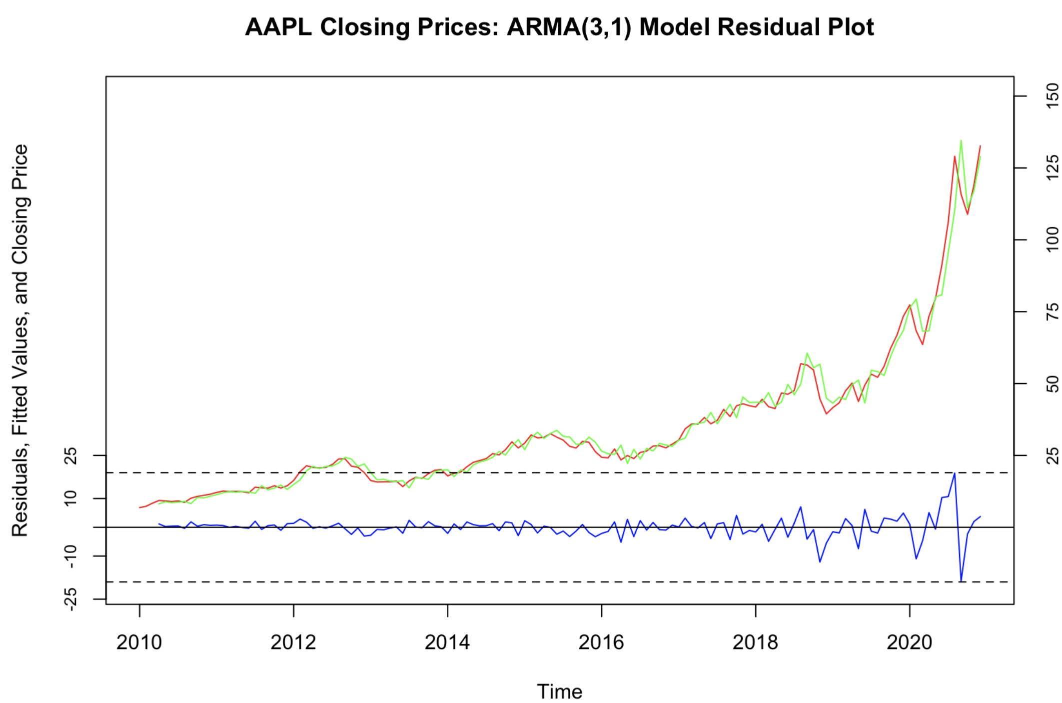
AAPL Closing Prices: AR(1) Model Residual Plot



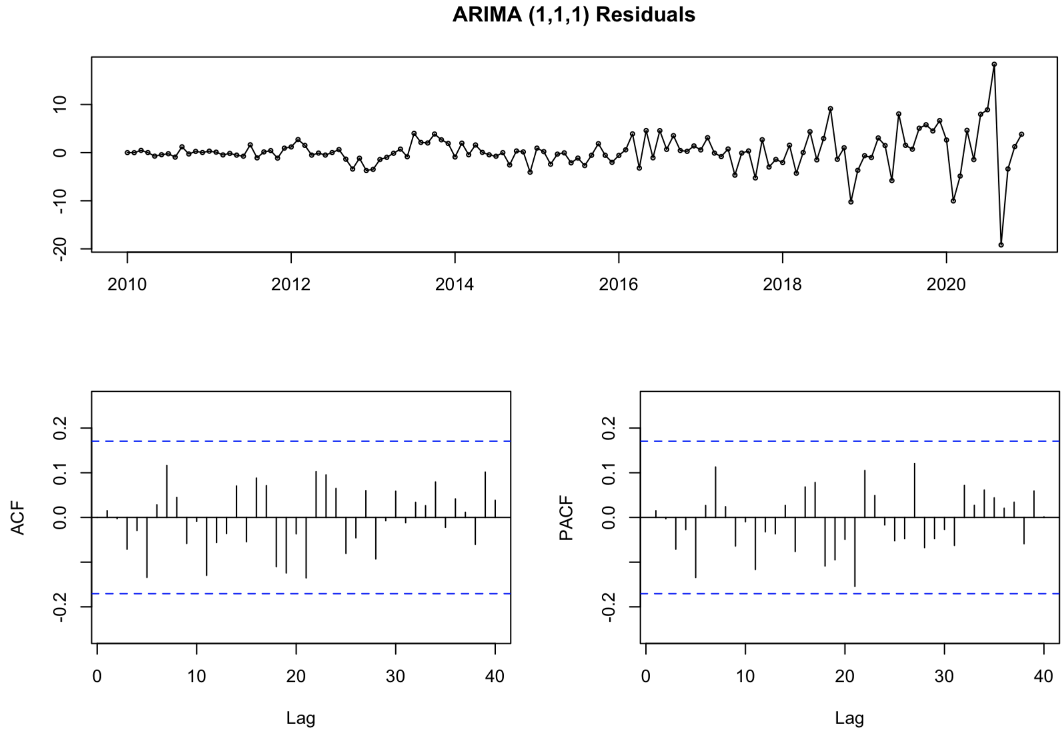
-


AAPL Closing Prices: AR(2) Residual Plot Model
Residuals, Fitted Values, and Closing Price
Time
2010
2012
2014
2016
2018
2020
25
10
-10
-25
150
125
100
75
50
25

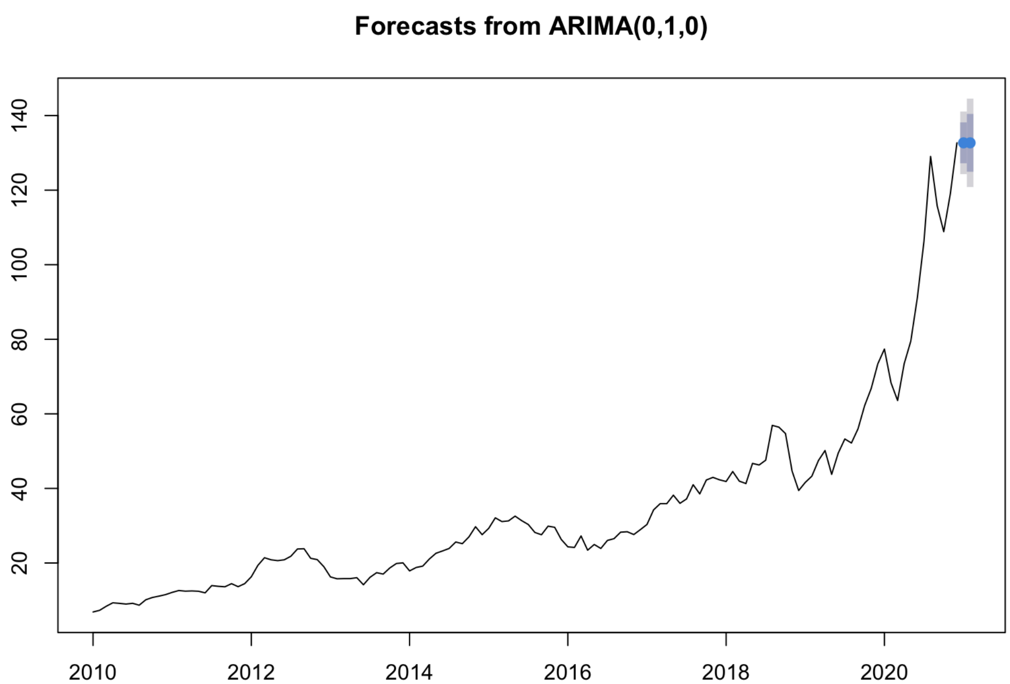
-


AAPL Closing Prices: ARMA(3,1) Model Residual Plot
Residuals, Fitted Values, and Closing Price
Time
2010
2012
2014
2016
2018
2020
25
10
-10
-25
150
125
100
75
50
25

-

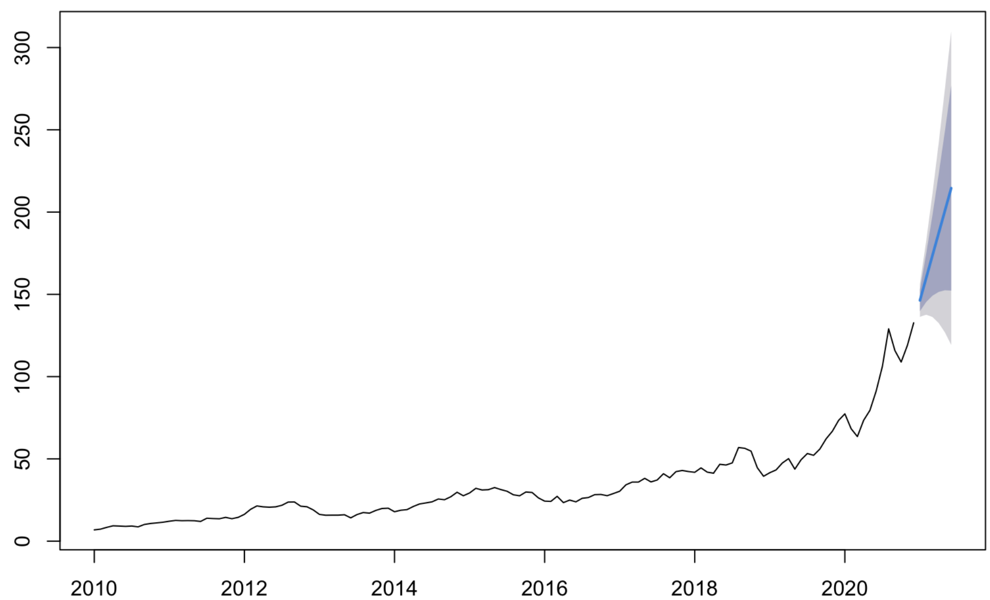

ARIMA (1,1,1) Residuals
10
0
-10
-20
2010
2012
2014
2016
2018
2020
ACF
PACF
0.2
0.1
0.0
-0.2
0
10
20
30
40
Lag

•


Forecasts from ARIMA(0,1,0)
140
120
100
80
60
40
20
2010
2012
2014
2016
2018
2020

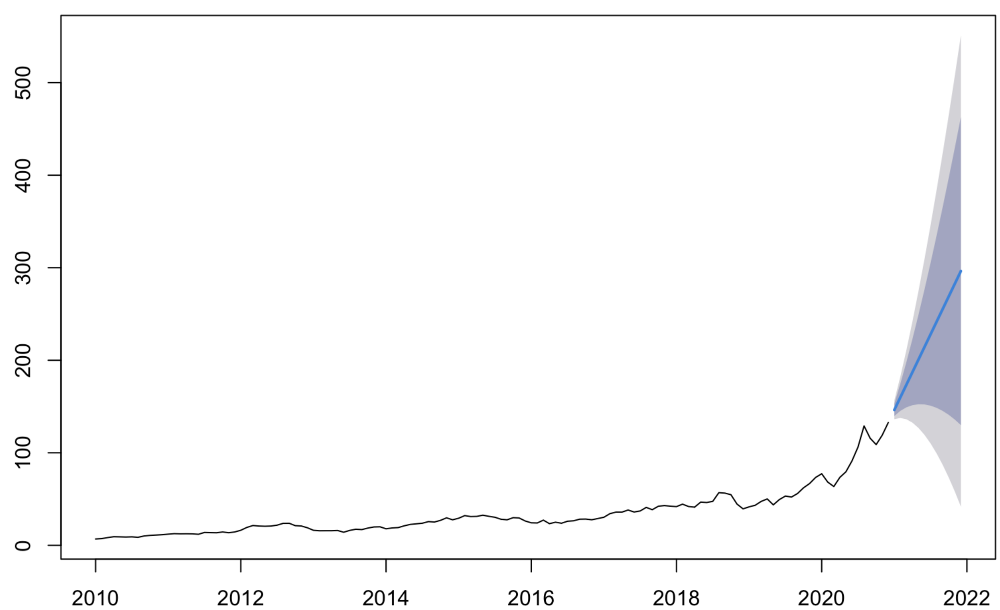
•

Forecasts from ARIMA(0,2,0)

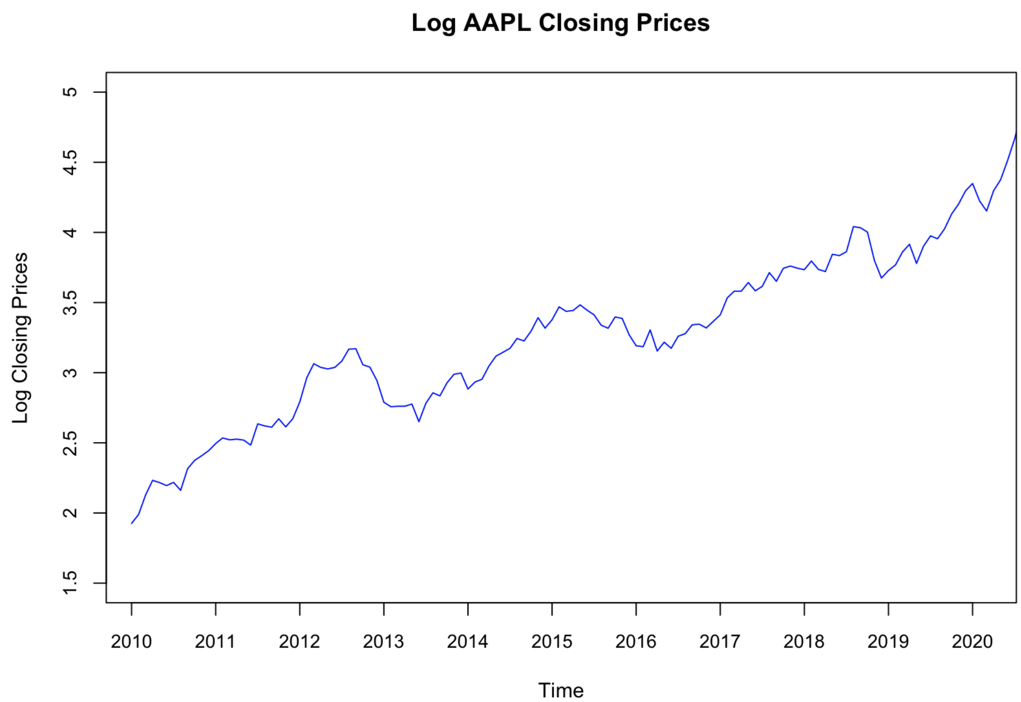


•

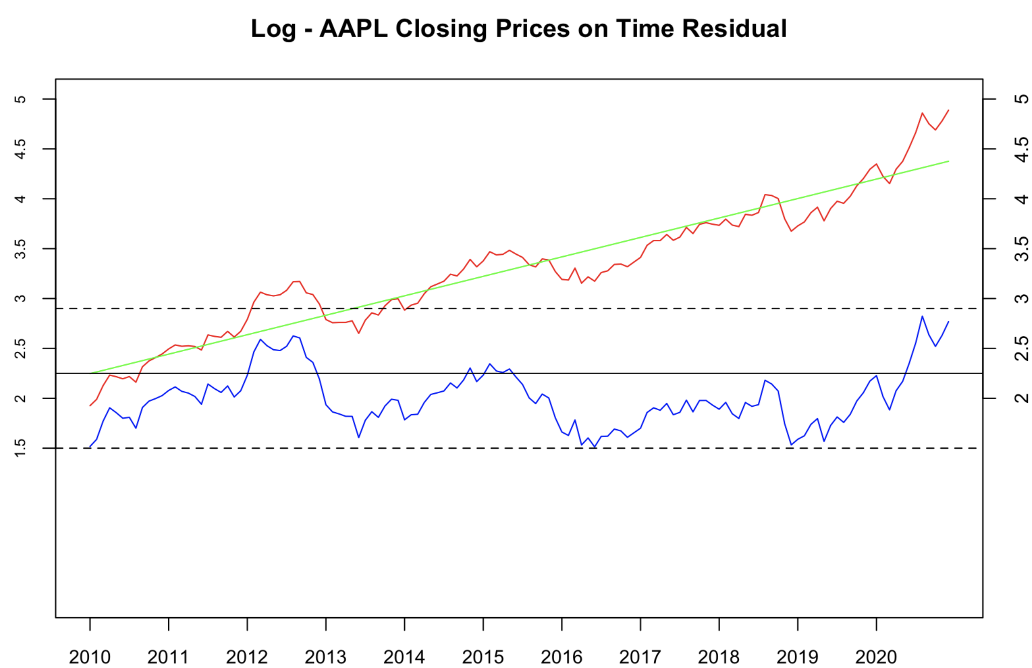
Forecasts from ARIMA(0,2,0)



•

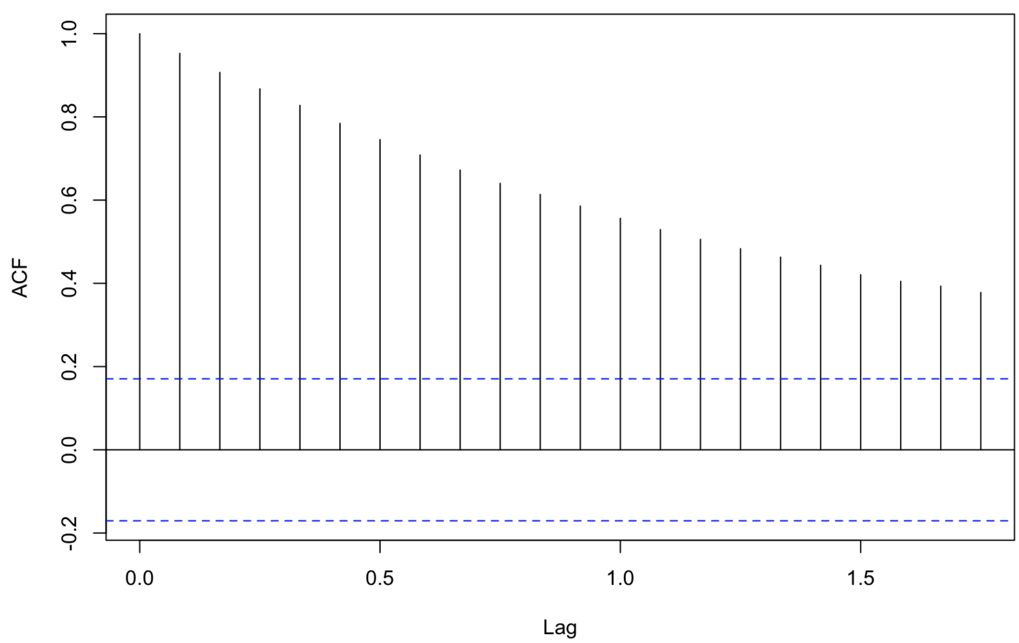

Log AAPL Closing Prices
Log Closing Prices
1.5
2
2.5
3
3.5
4
4.5
5
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
Time

-


Log - AAPL Closing Prices on Time Residual
1.5
2
2.5
3
3.5
4
4.5
5
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020

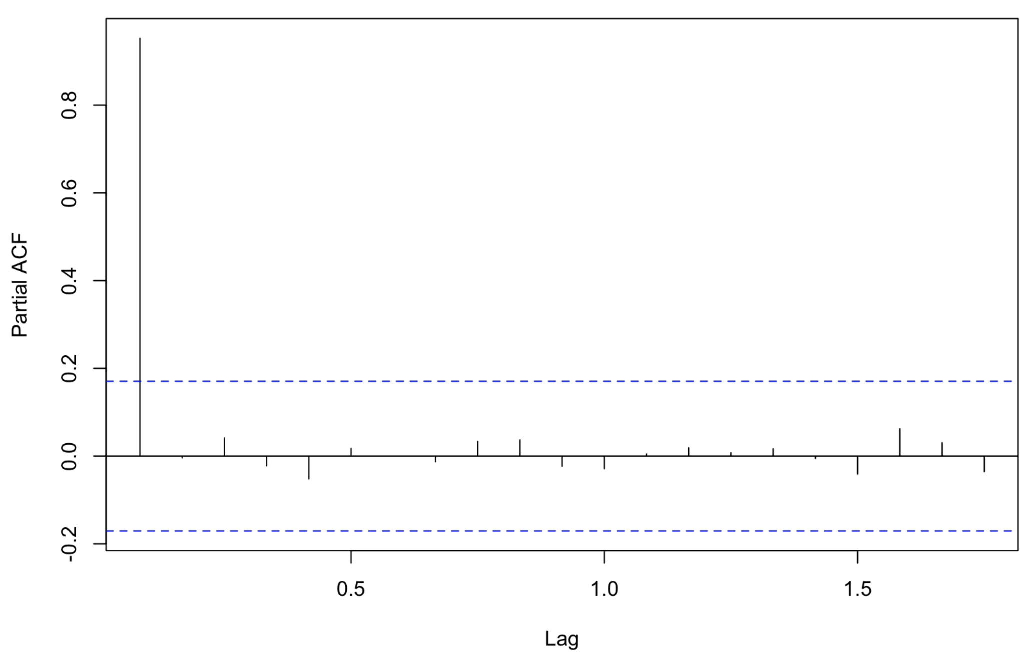
-

Autocorrelation of Log AAPL Closing Prices



-

Partial Autocorrelation of Log AAPL Closing Prices



-