STAT443 Group Project

Colton Cheung, Darren Ku, Yoonha Kim

31/03/2020

Introduction

As we entered late Janurary to early Feburary, the concern surrounding the novel Coronavirus dubbed Covid-19 soon grew beyond the limited geographic scope of its origins and entered the minds of the world. It was these growing global concerns that motivated our study of the closing price of the NASDAQ stock exchange and how the world's second largest stock exchange would react in the face of a potential pandemic. Under the premise that China serves as a major economic power and important step in global supply chains, our hypothesis would be that many of the companies listed on the NASDAQ stack exchange would have their businesses impacted negatively by the spread of Covid-19, and thus consequently we would see that reflected in the overall closing price of NASDAQ.

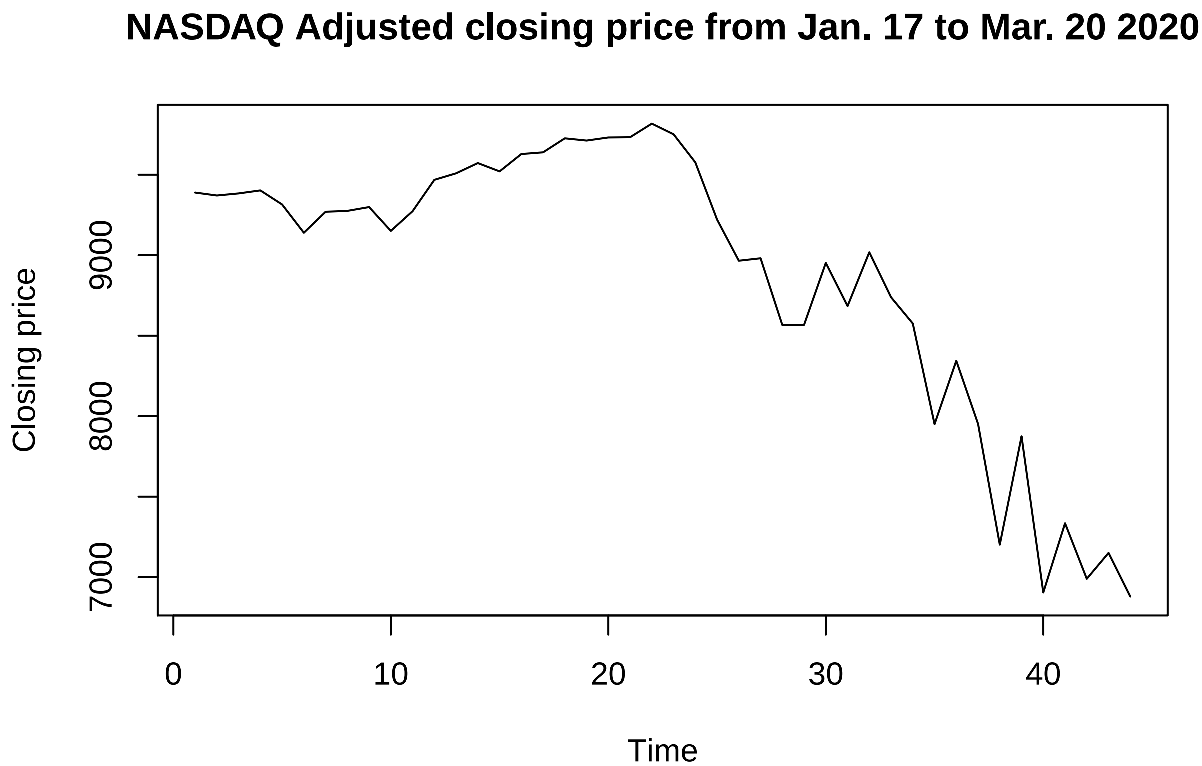
Eventually Covid-19 would be dubbed a pandemic and we would see NASDAQ react as we expected.

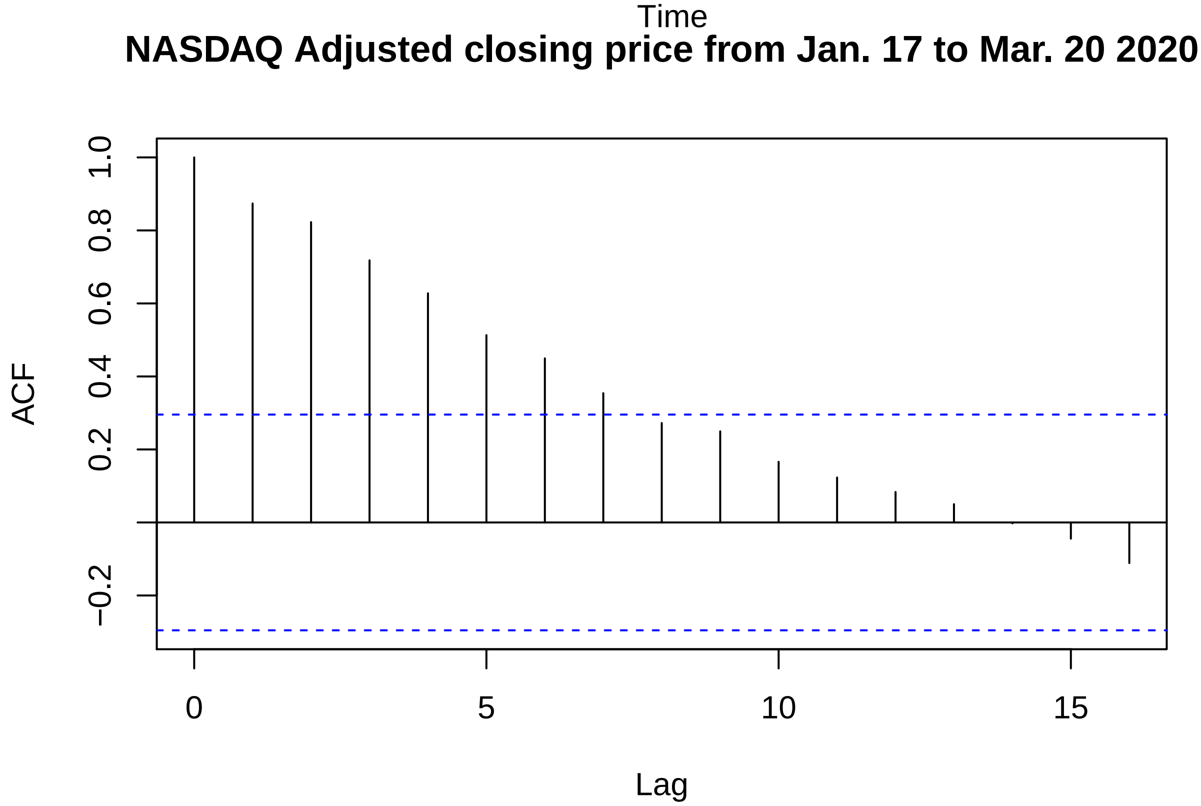
We collected our Data by looking at the summary of the Open, High, Low, Close, Adjusted Close, and Volume on Google Finance after the end of the trading day, which ends at 4:00pm EST each business day. Our data starts from January 17th to March 20th. For comparison's sake, we also took a look at the NASDAQ closing prices for the 46 day period between October 17th and December 20th as well as the 44 day period between January 17th to March 20th of the past year. The historical data was obtained via Google Finance's historical data archives.

Our main variable of interest is the adjusted closing price with Open, High, Low and Volume as supplementary variables.

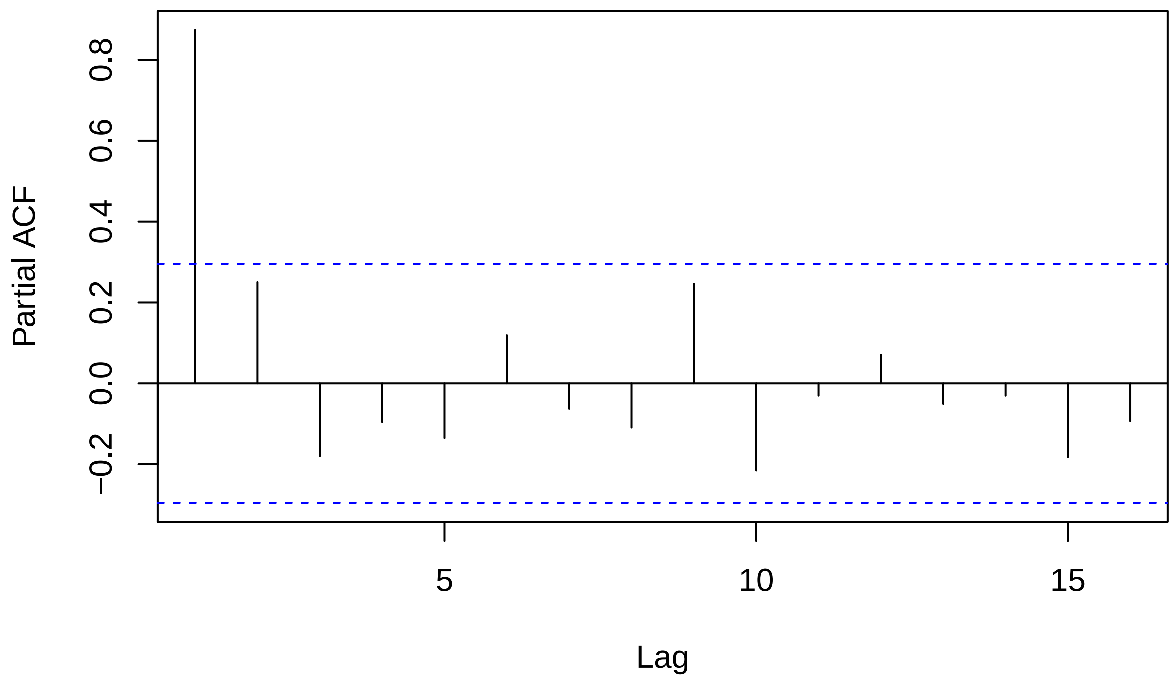
```r
OctToDec<-read.table("NASDAQ_OCT17_DEC20.csv", header=TRUE, sep=",")
JanToMar<-read.table("NASDAQ_JAN17_MAR20.csv", header=TRUE, sep=",")
JanToMar2019<-read.table("NASDAQ_JAN17_MAR20_2019.csv", header=TRUE, sep=",")
```

Original Time Series Plots

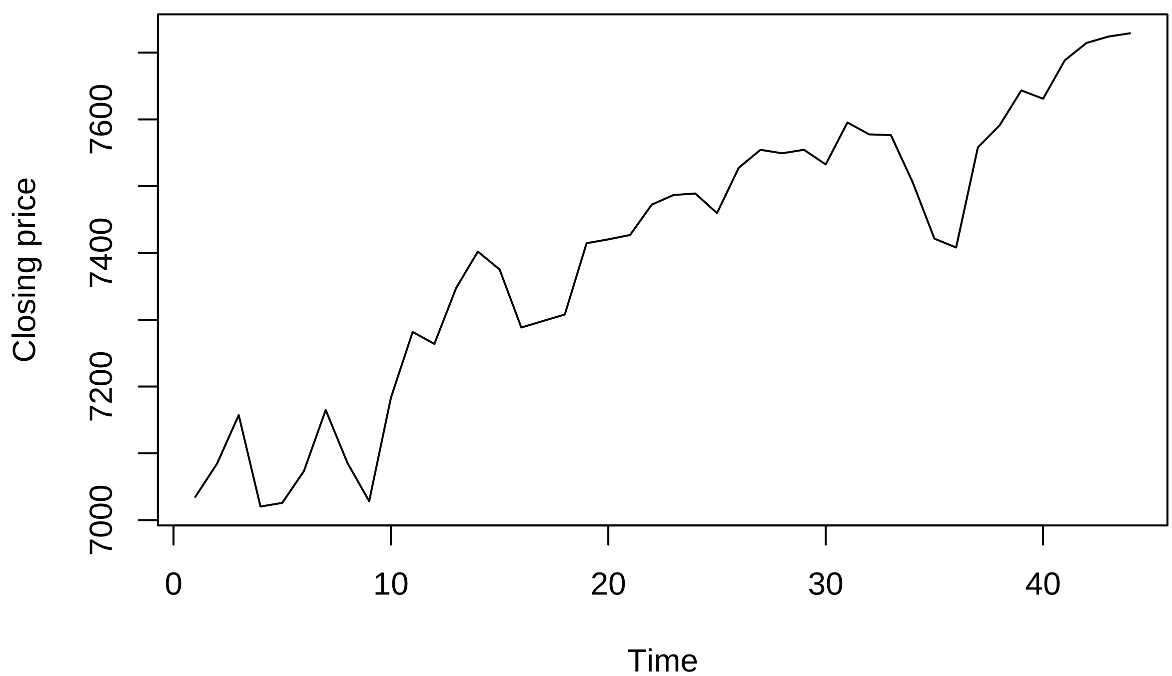


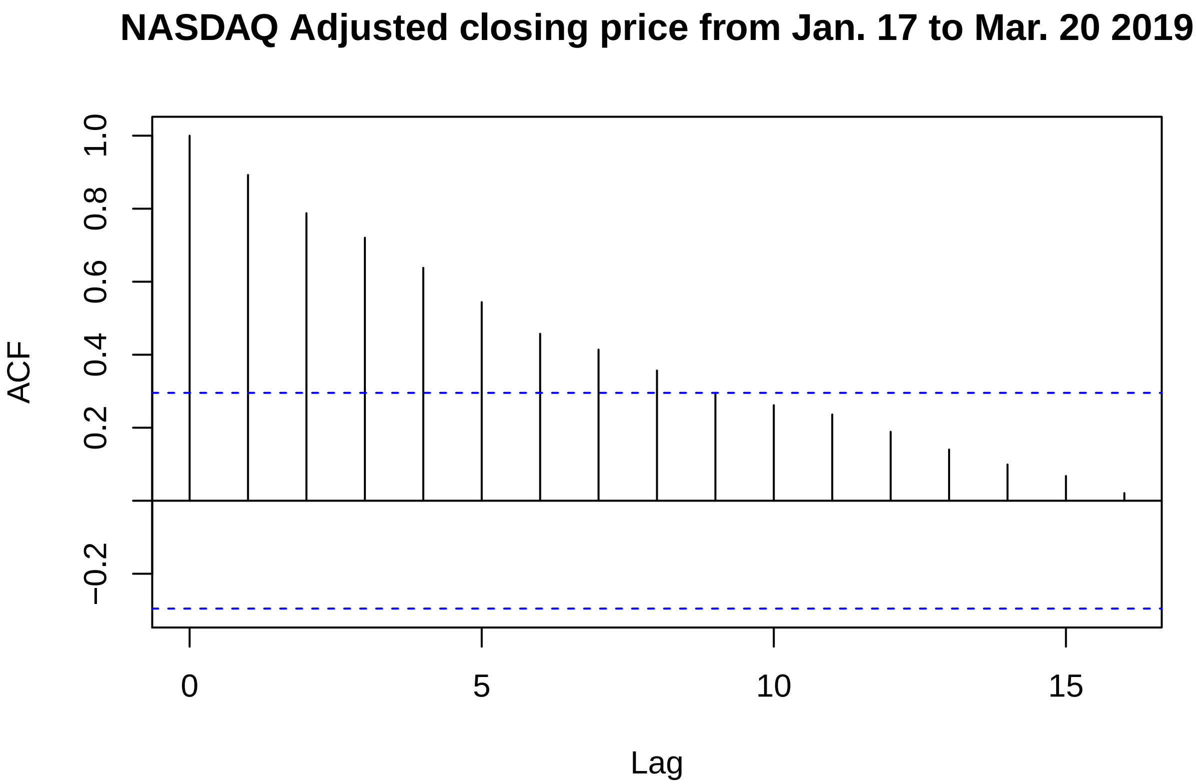


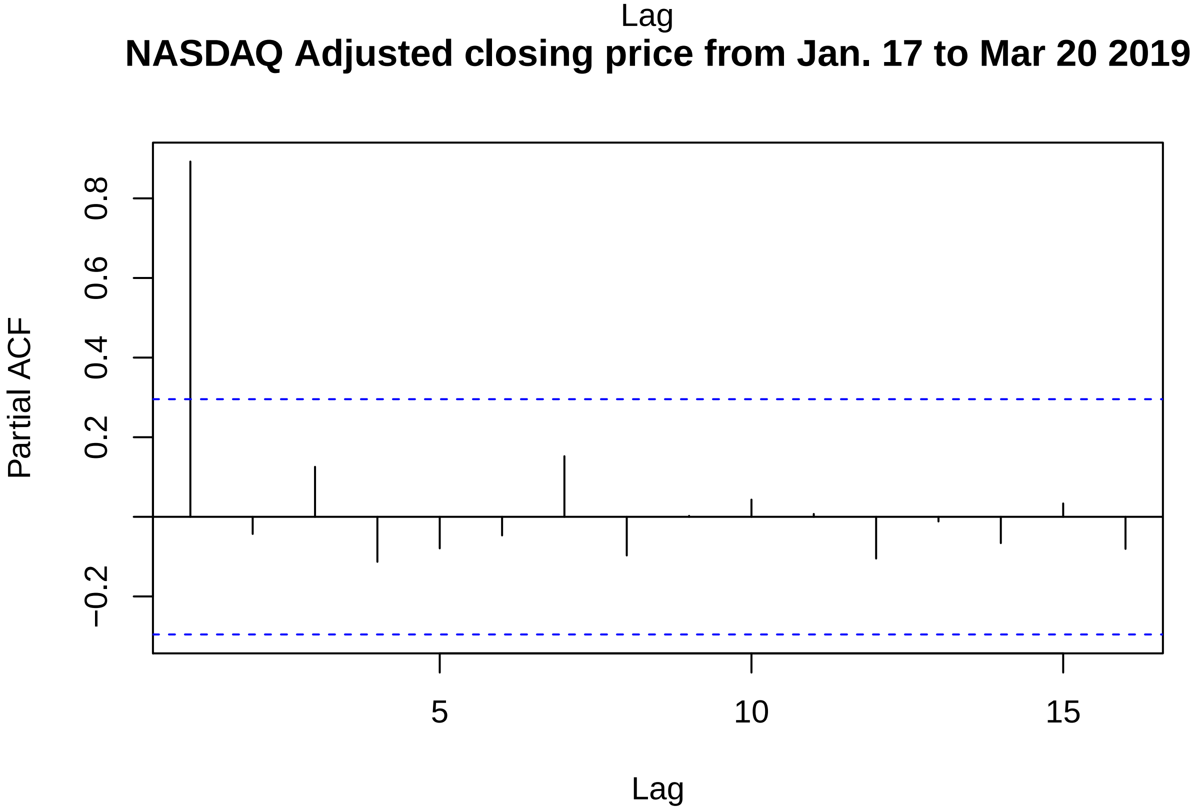
NASDAQ Adjusted closing price from Jan. 17 to Mar. 20 2020

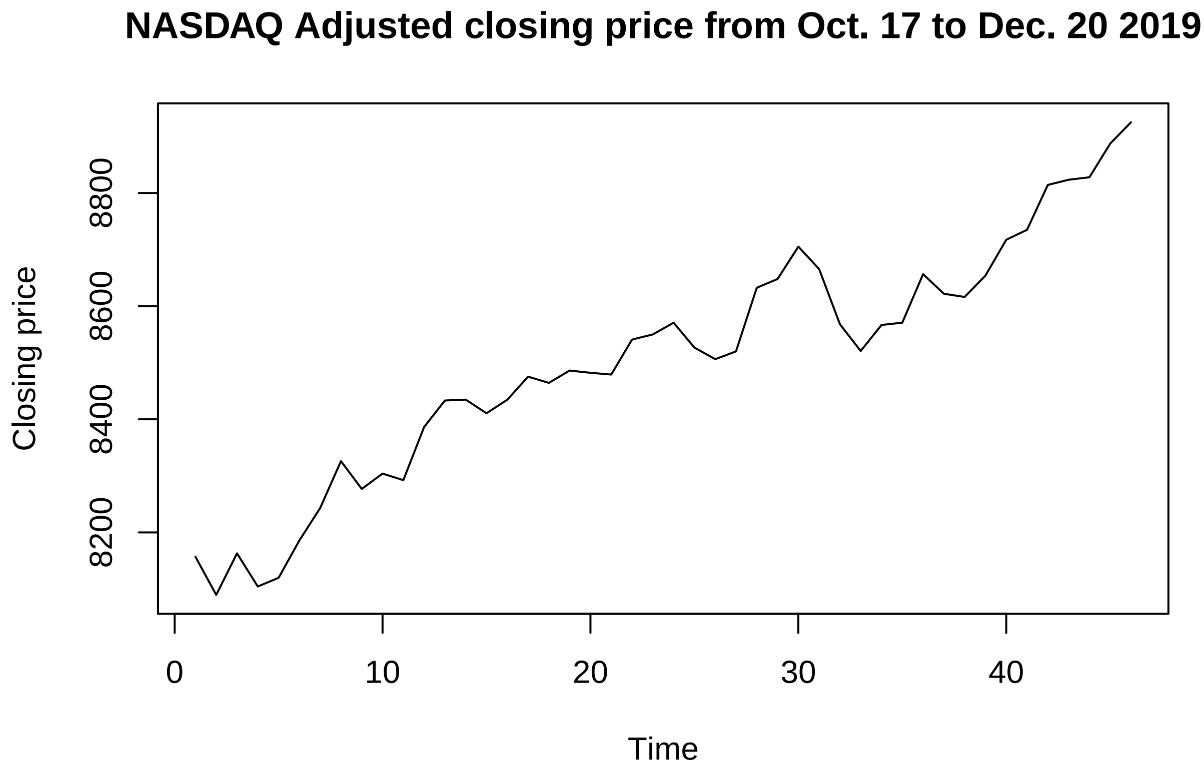


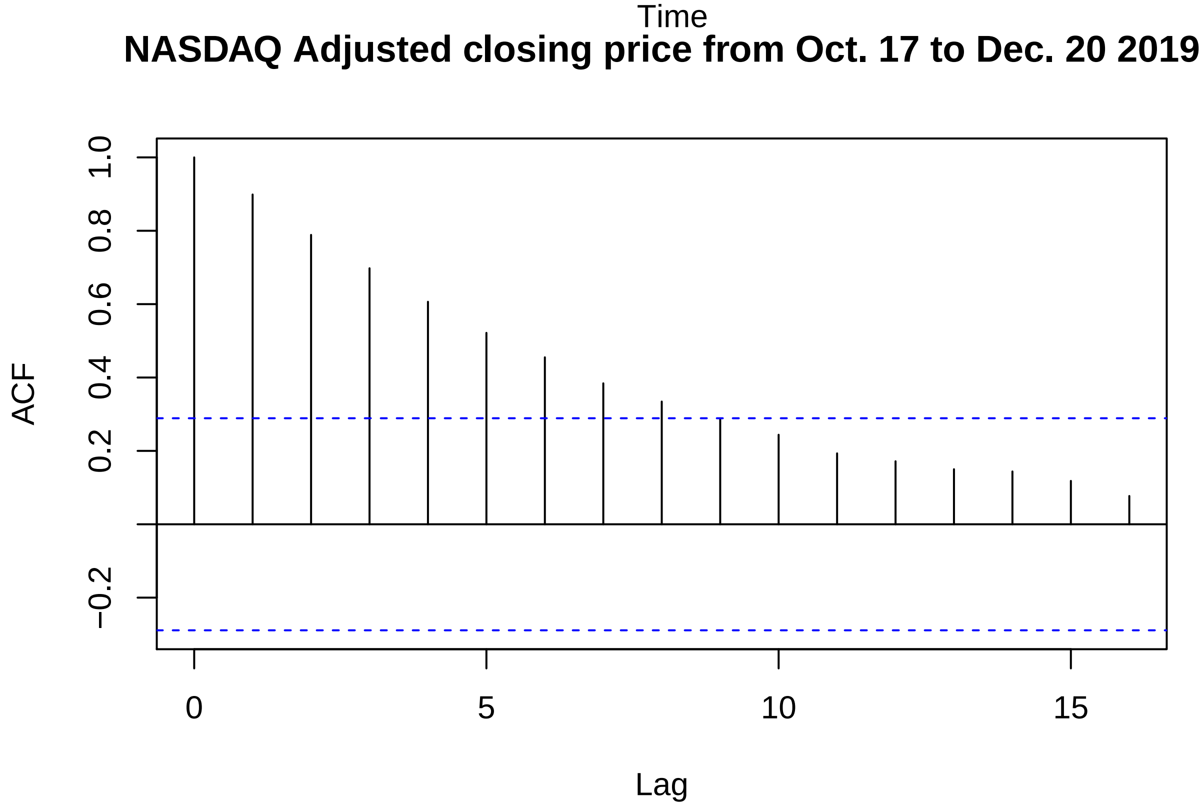
NASDAQ Adjusted closing price from Jan. 17 to Mar. 20 2019

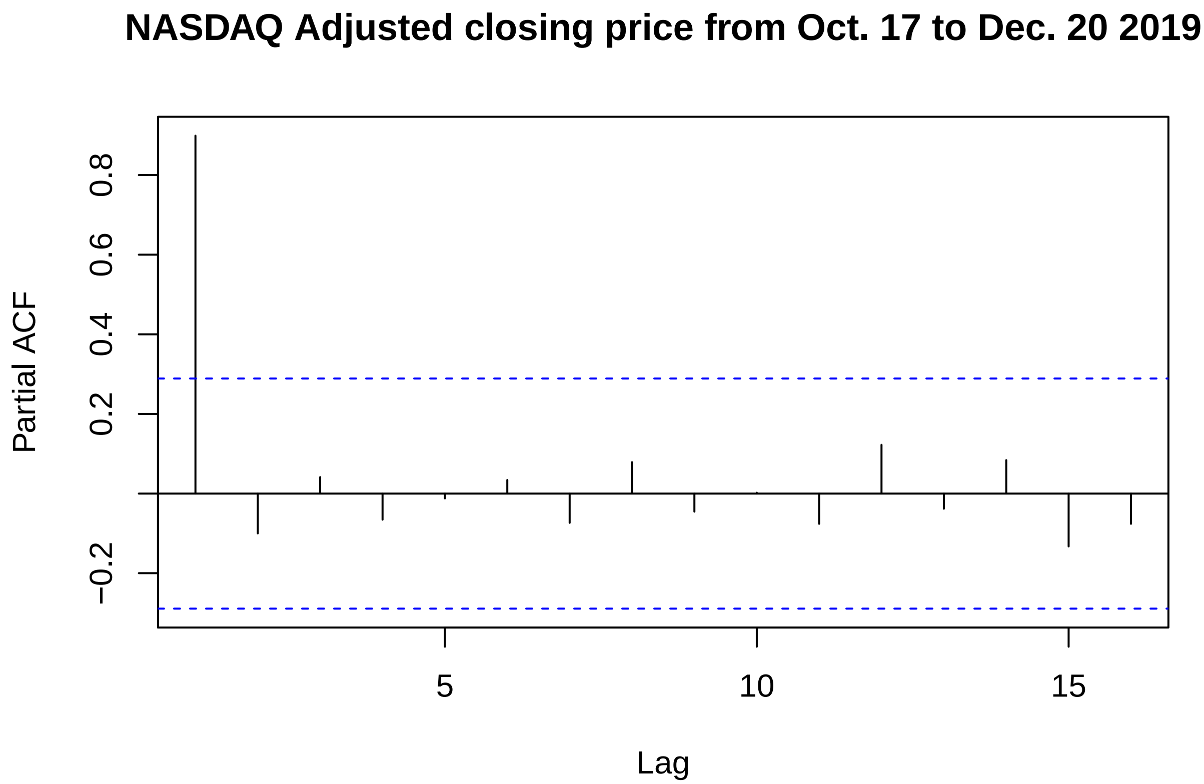



NASDAQ Adjusted closing price from Jan. 17 to Mar. 20 2019
ACF
1.0
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
Lag


NASDAQ Adjusted closing price from Jan. 17 to Mar 20 2019
Partial ACF
0.8
0.6
0.4
0.2
−0.2
5
10
15
Lag


NASDAQ Adjusted closing price from Oct. 17 to Dec. 20 2019
Closing price
8200
8400
8600
8800
0
10
20
30
40
Time


NASDAQ Adjusted closing price from Oct. 17 to Dec. 20 2019
ACF
−0.2
0.2
0.4
0.6
0.8
1.0
0
5
10
15
Lag



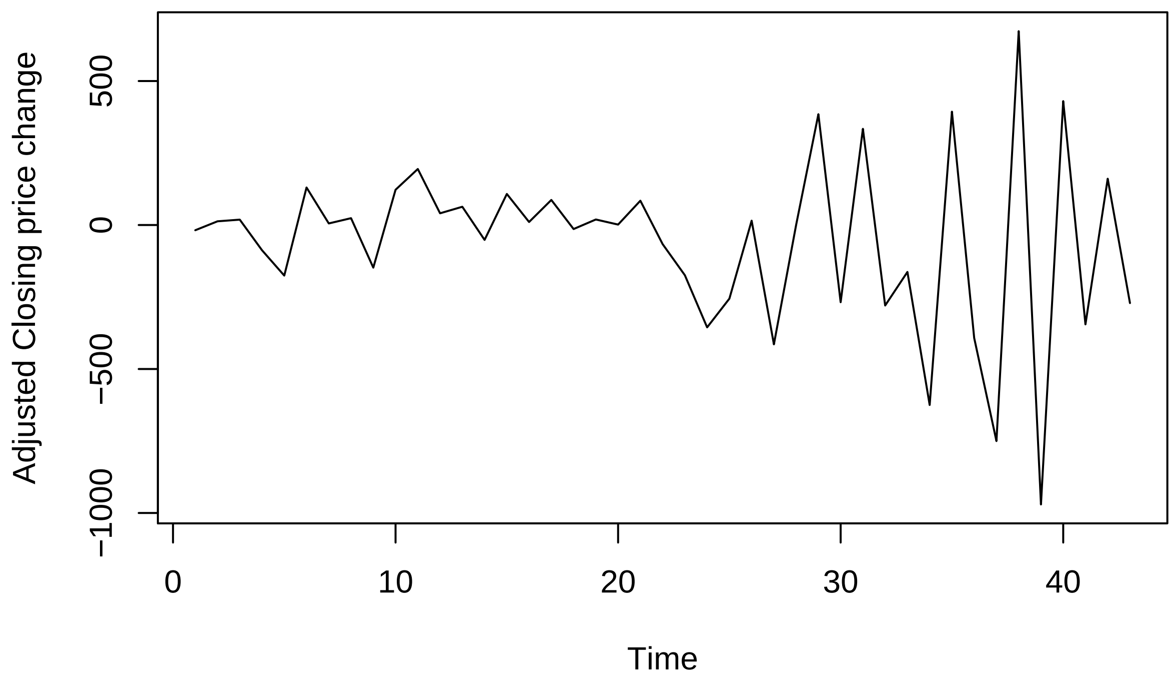
Preliminary Plots Analysis

As evident from the plots of the 3 time series, our January to March 2020 data has an overall negative trend, while both our 2019 datasets have positive trends. In fact look at the historical data for NASDAQ, it had a positive change of 35.23% over all of 2019, the highest yearly increase since 2013.

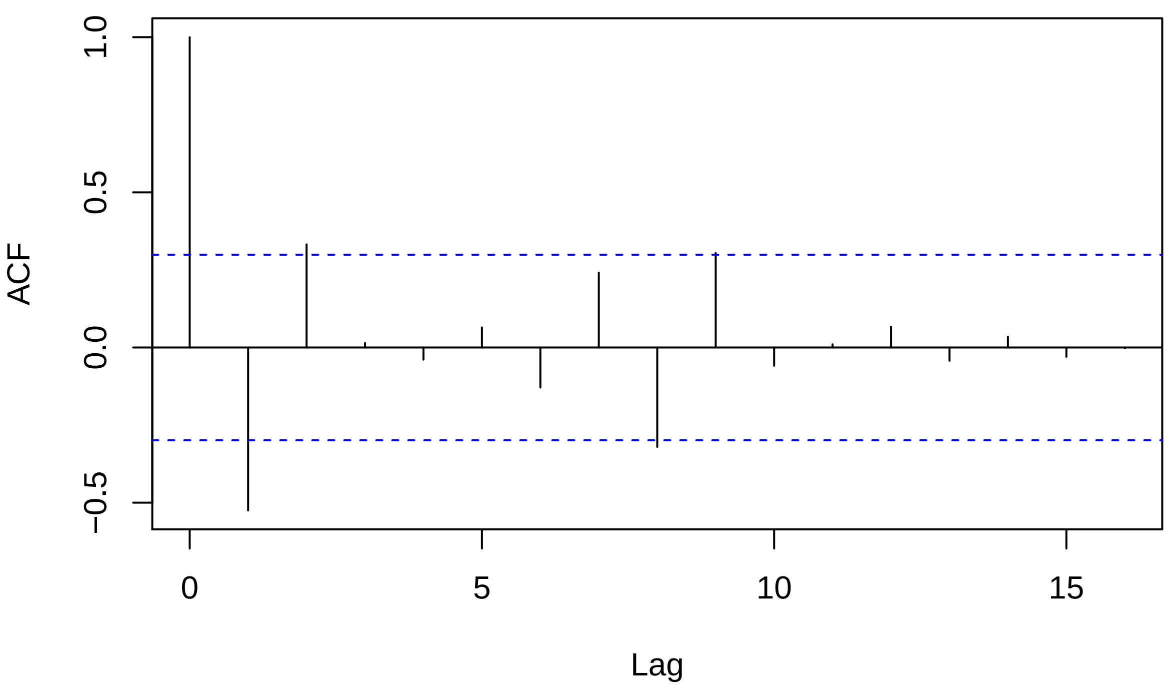
Meanwhile, the ACFs of all three datasets are all very similar, each simply decaying with increasing lag. Looking to the PACF, we see that there are no significant PACF values outside of lag 1. This inidicates that potentially an AR(1) model may be the appropriate thing to fit to this data.

First Differences Plots

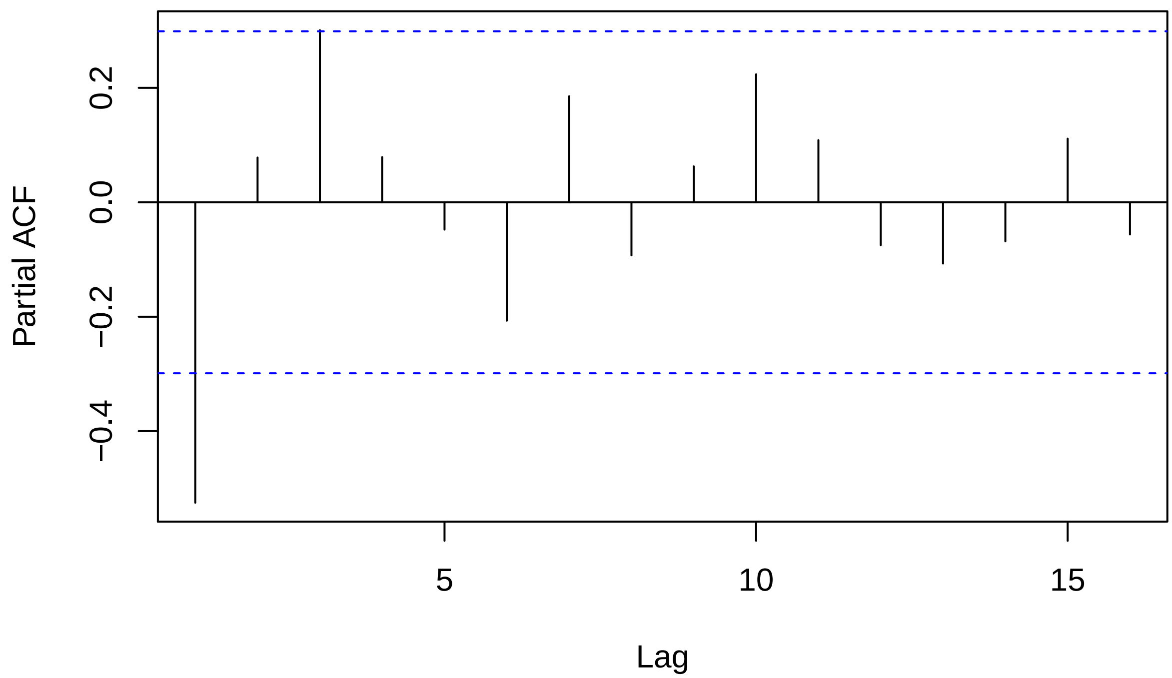
NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20 2020



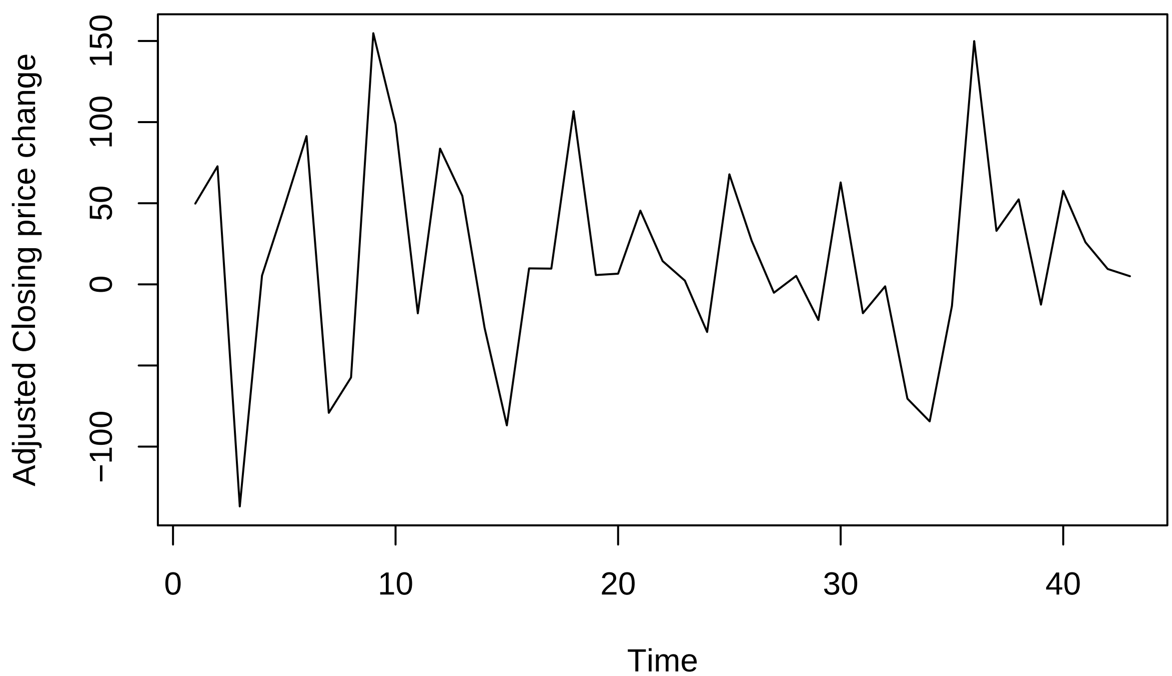
NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20 2020

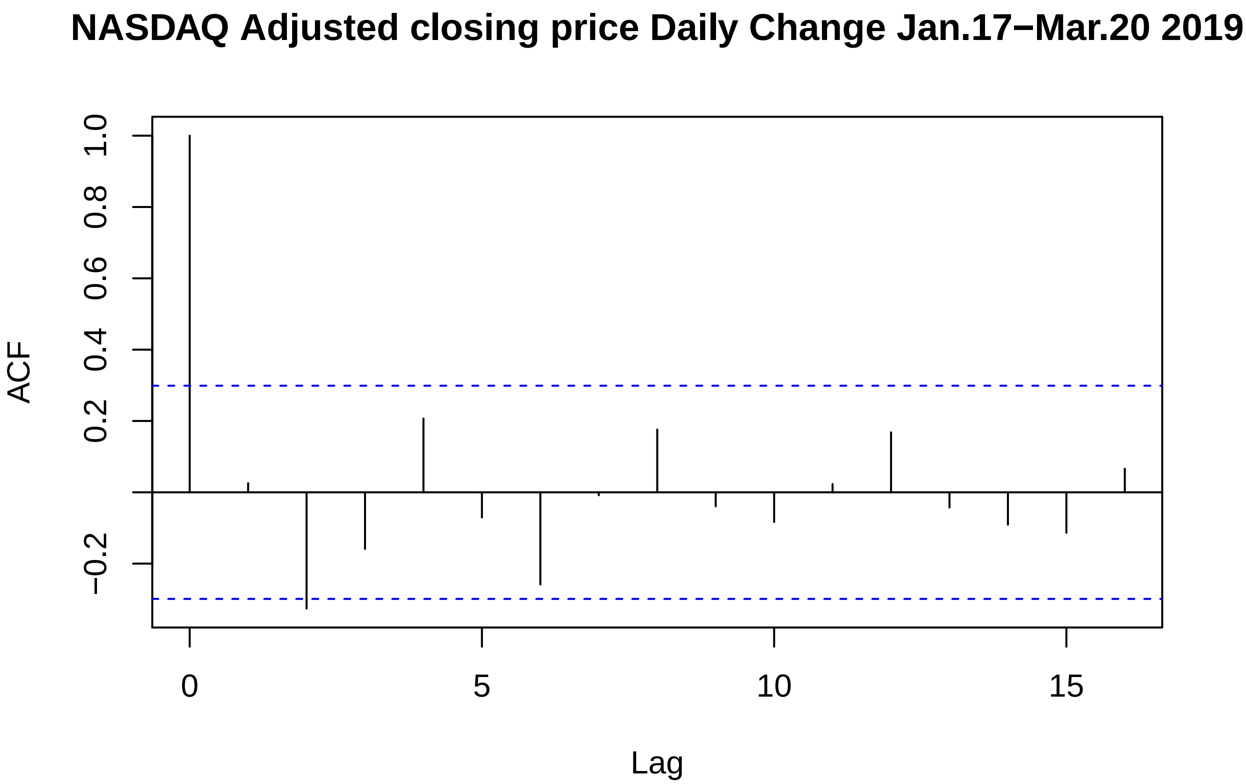


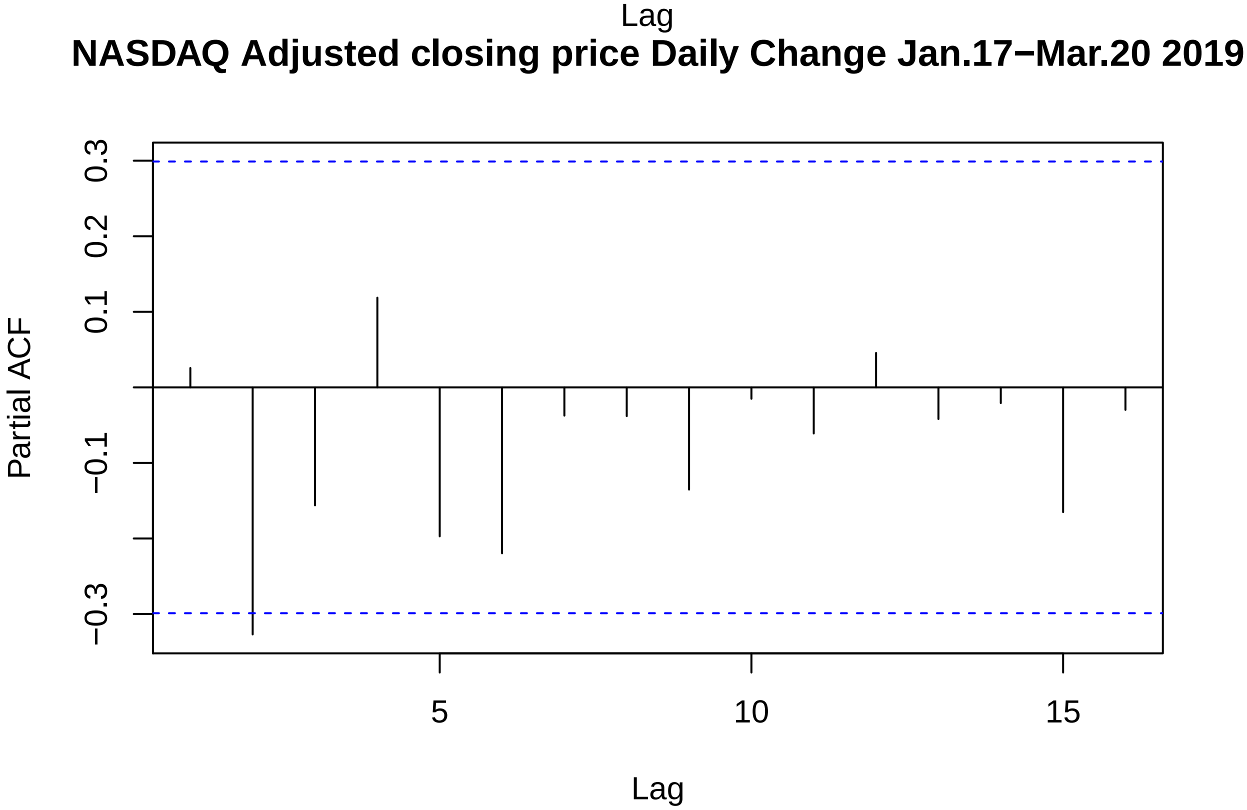
NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20 2020

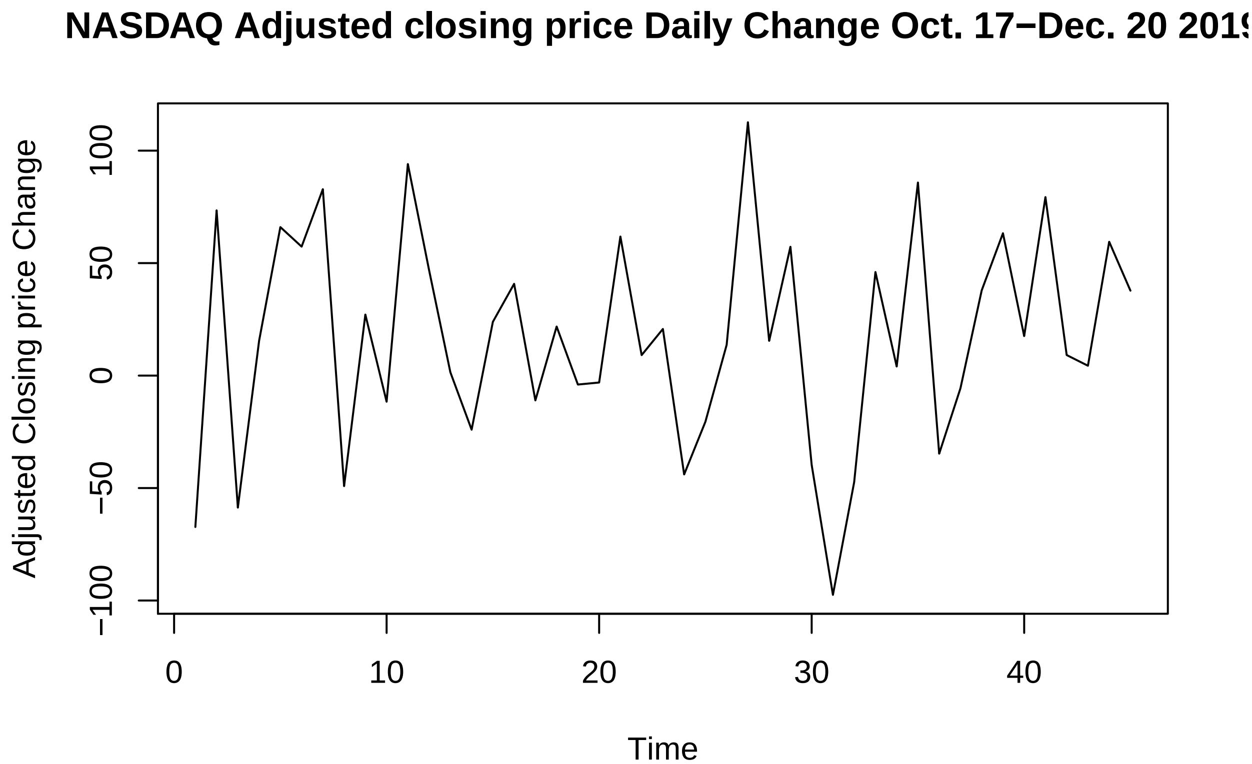


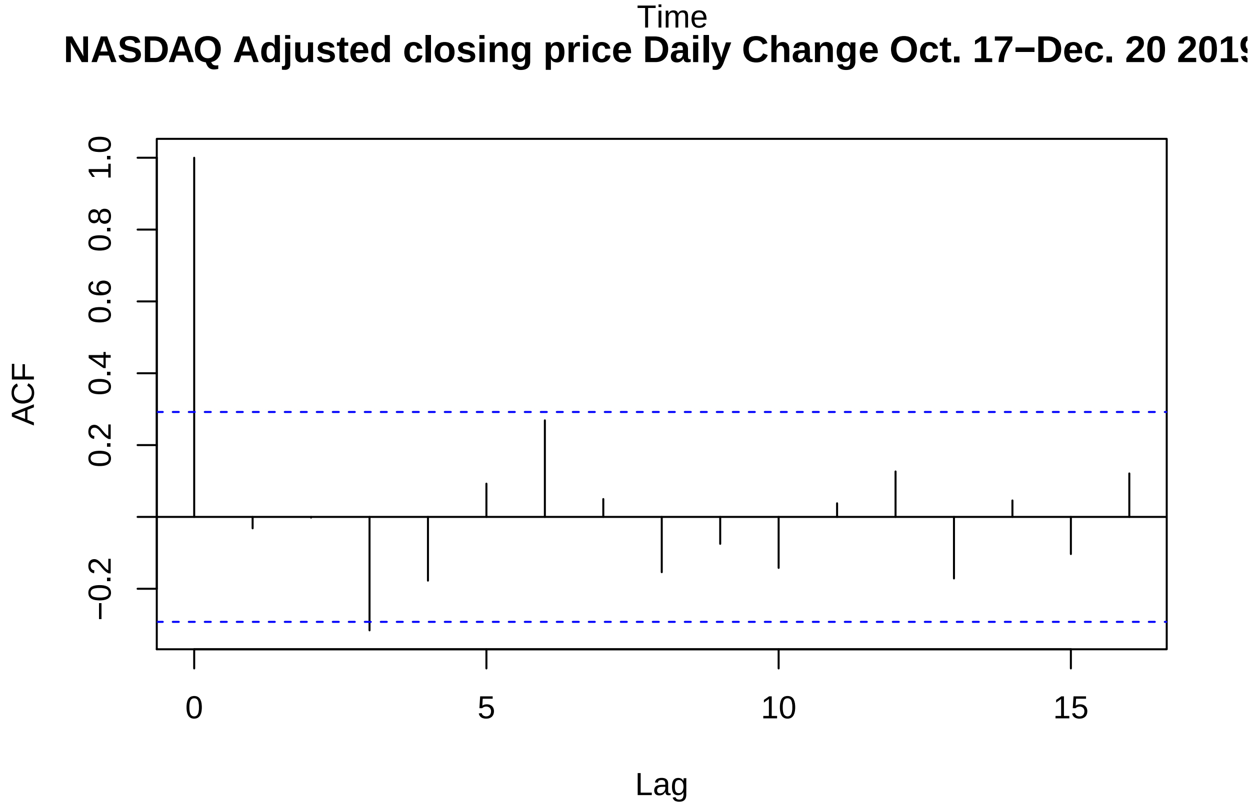
NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20 2019

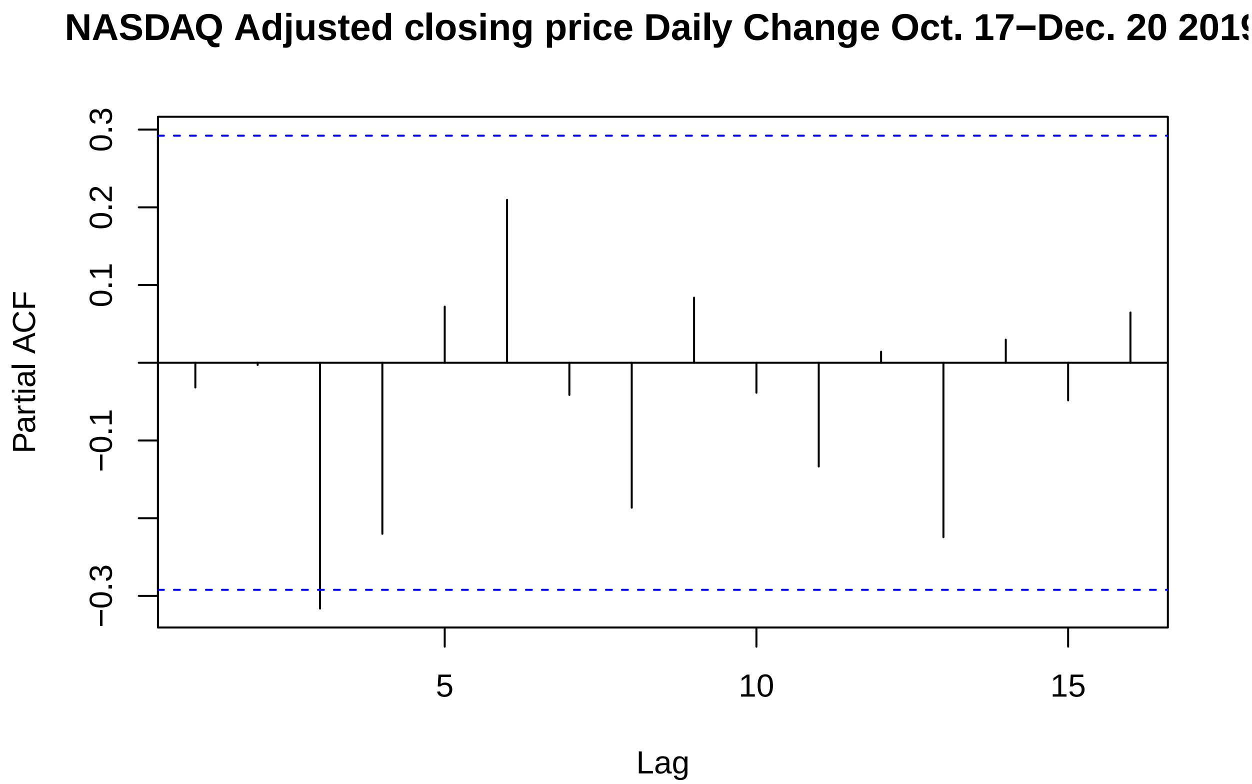



NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20 2019
ACF
1.0
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
Lag


NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20 2019
Partial ACF
0.3
0.2
0.1
−0.1
−0.3
5
10
15
Lag


NASDAQ Adjusted closing price Daily Change Oct. 17−Dec. 20 2019
Adjusted Closing price Change
100
50
0
−50
−100
0
10
20
30
40
Time


NASDAQ Adjusted closing price Daily Change Oct. 17−Dec. 20 2019
ACF
1.0
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
Lag



First Differences Anaylsis

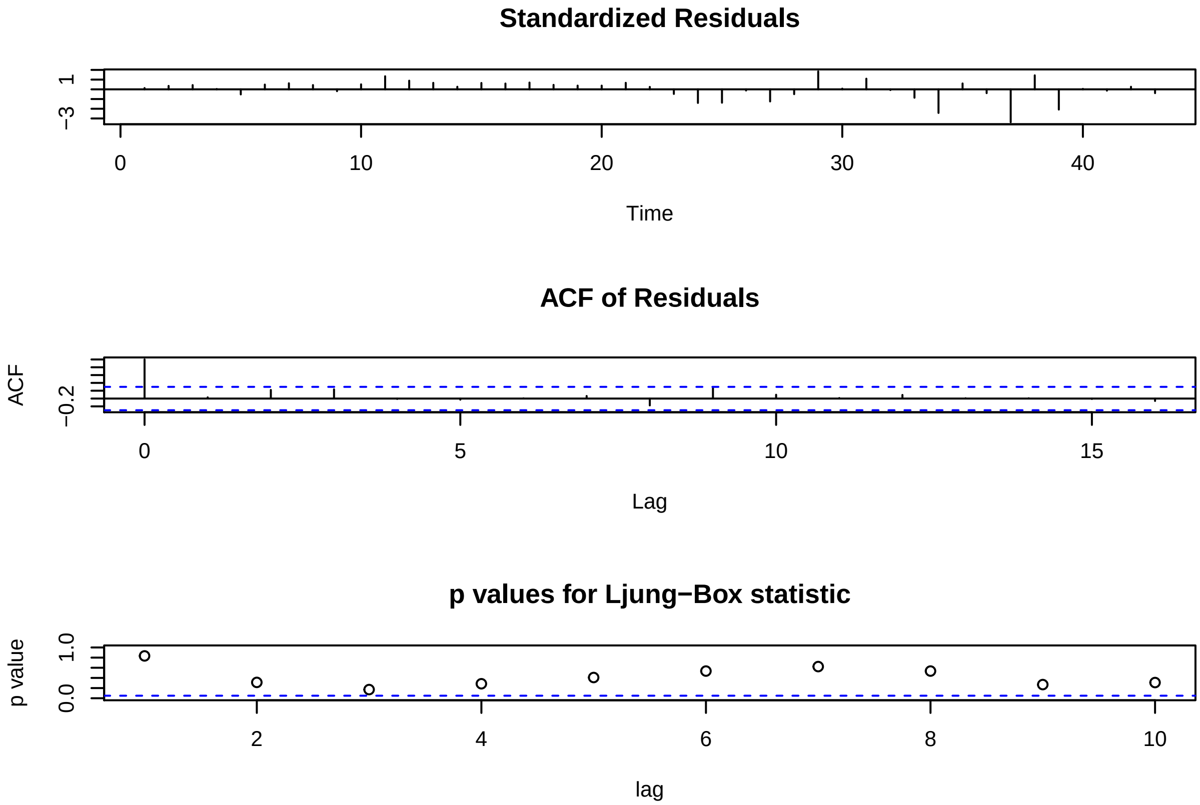
Applying first differences to the 3 time series, we see a departure of the January to March 2020 series from the other two. While the 2019 time series exhibit similar fluctuation behaviour, the 2020 time series shows a persisting increase in fluctuation magnitude starting from late February / early March 2020.

These sudden changes in the 2020 time series are expected in light of increasing reports of COVID-19 related cases and deaths. It was during early March (the 11th) that the World Health Organization declared COVID-19 a pandemic. Social distancing was more frequently practiced and businesses categorized as "non-essential" had to cease operations during this period as well.

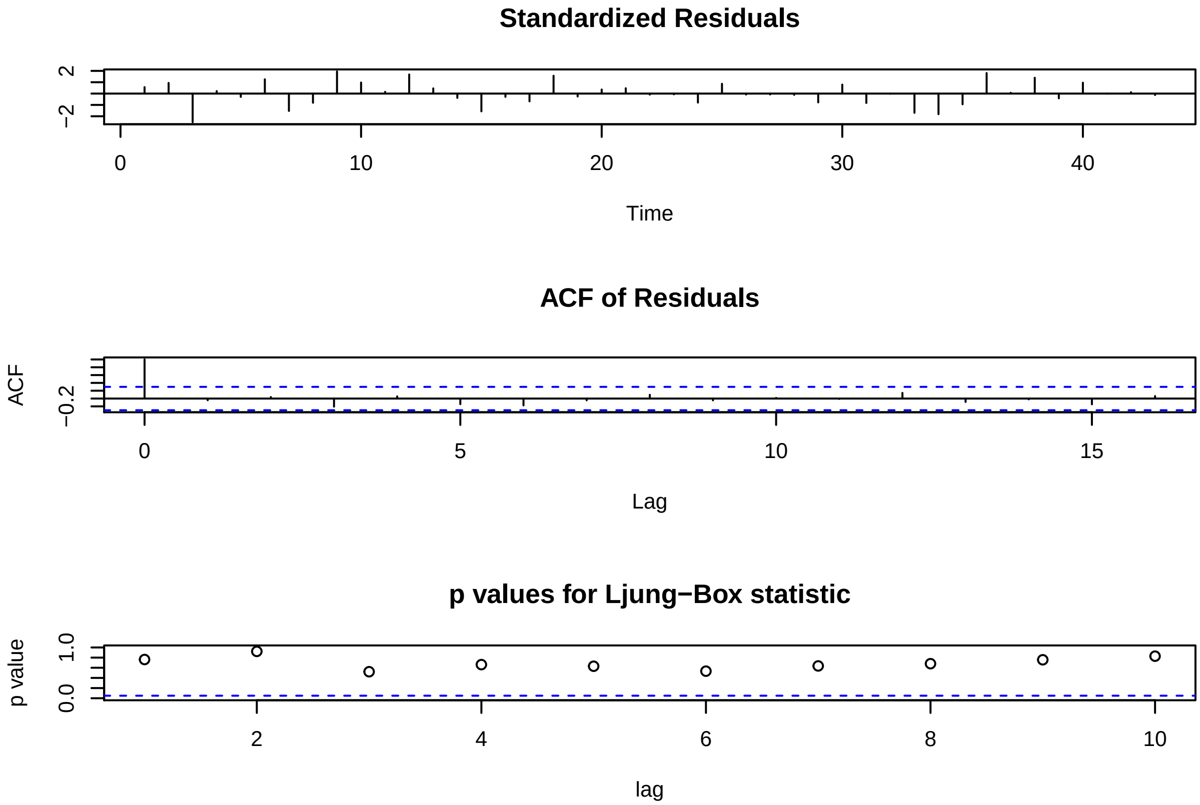
Meanwhile, the ACFs tails off after lag 1. Looking to the PACF, we see that the PACF of the data from Jan.17-Mar.20 2020 cuts off at lag 1. The PACF of 2019 Jan-Mar data cuts off at lag 2 and the PACF of 2019 Oct-Dec data cuts off at lag 3.This inidicates that potentially AR(1), AR(2), AR(3) models may be the appropriate thing to fit to this data respectively.

Manually-Fitted Model

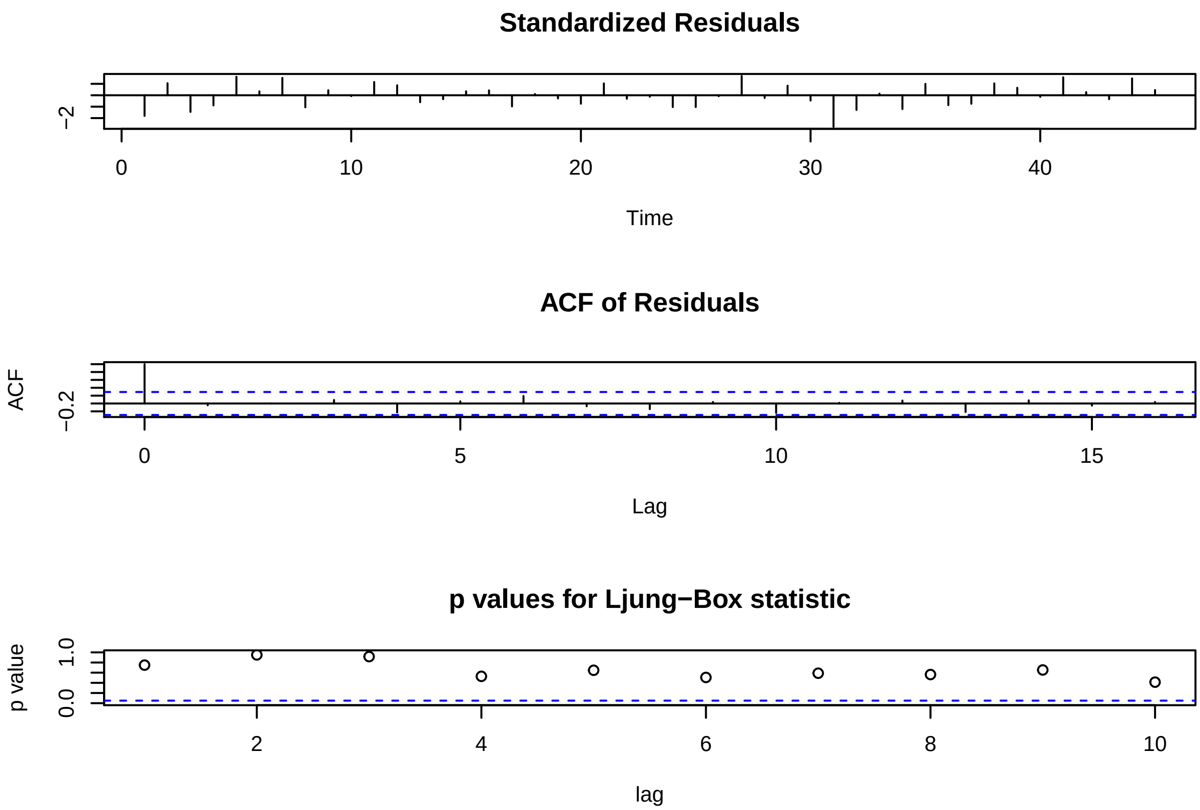
```
##
## Call:
## arima(x = JanMarDiff, order = c(1, 0, 0))
##
## Coefficients:
##           ar1  intercept
##       -0.5195   -56.9652
## s.e.   0.1278    25.9234
##
## sigma^2 estimated as 65650:  log likelihood = -299.65,  aic = 605.3
```




```
##
## Call:
## arima(x = JanMar2019Diff, order = c(2, 0, 0))
##
## Coefficients:
##          ar1      ar2  intercept
##       0.0296  -0.3212    15.7299
## s.e.  0.1436   0.1420     6.7748
##
## sigma^2 estimated as 3218:  log likelihood = -234.77,  aic = 477.54
```

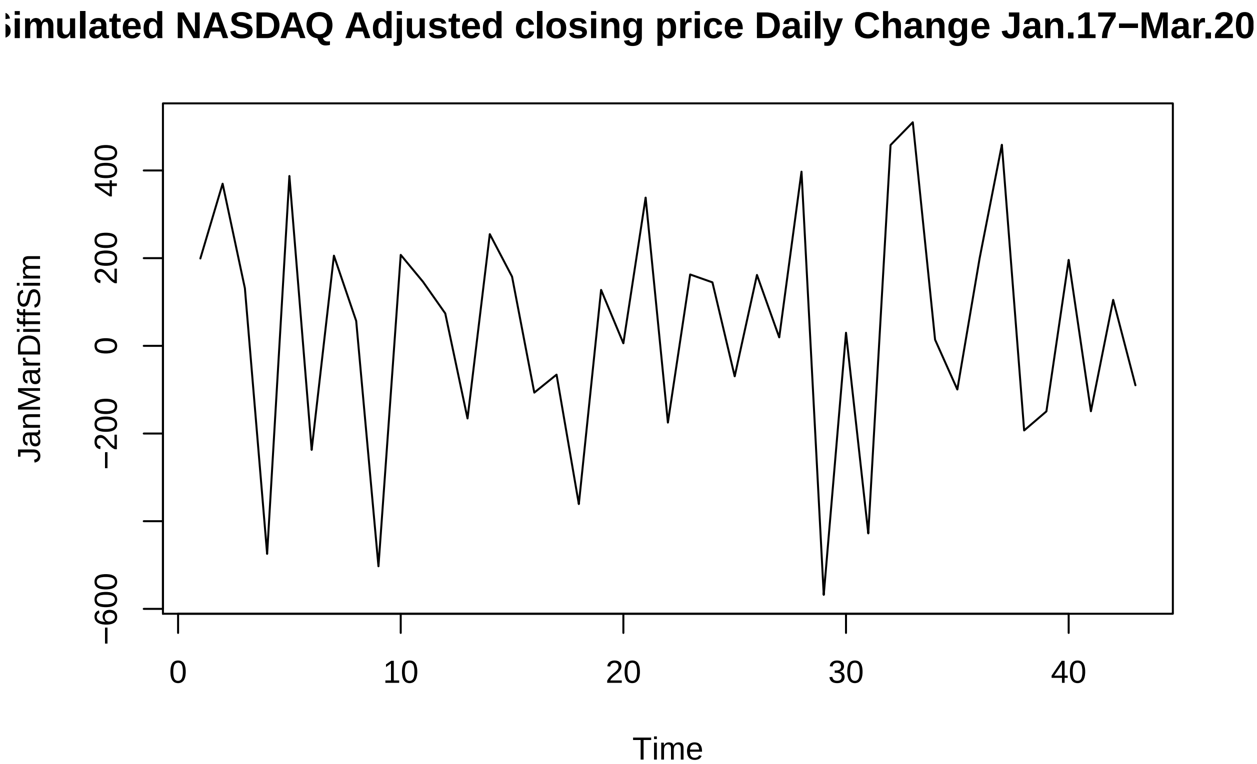


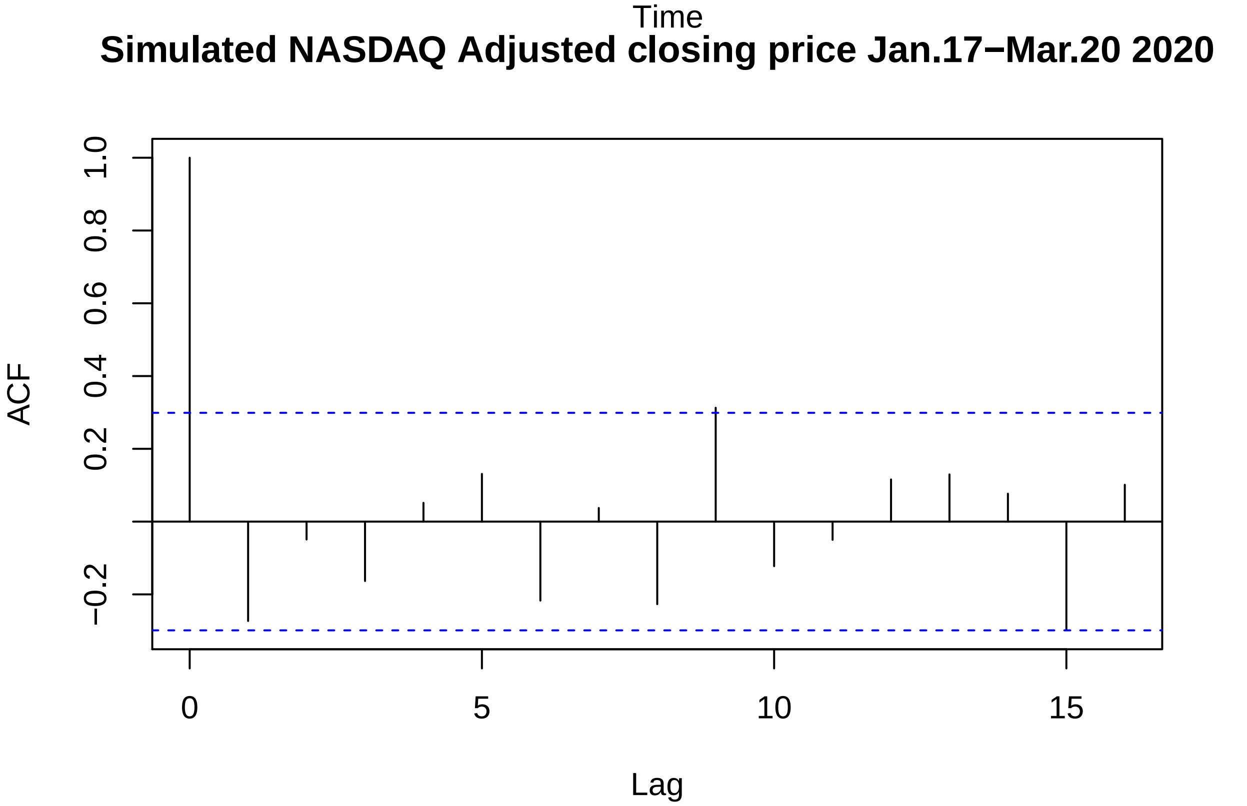

```
##
## Call:
## arima(x = OctDecDiff, order = c(3, 0, 0))
##
## Coefficients:
##           ar1     ar2      ar3  intercept
##       -0.0580  0.0195  -0.3655    17.4599
## s.e.   0.1425  0.1447   0.1476     4.6945
##
## sigma^2 estimated as 1883:  log likelihood = -233.73,  aic = 477.47
```

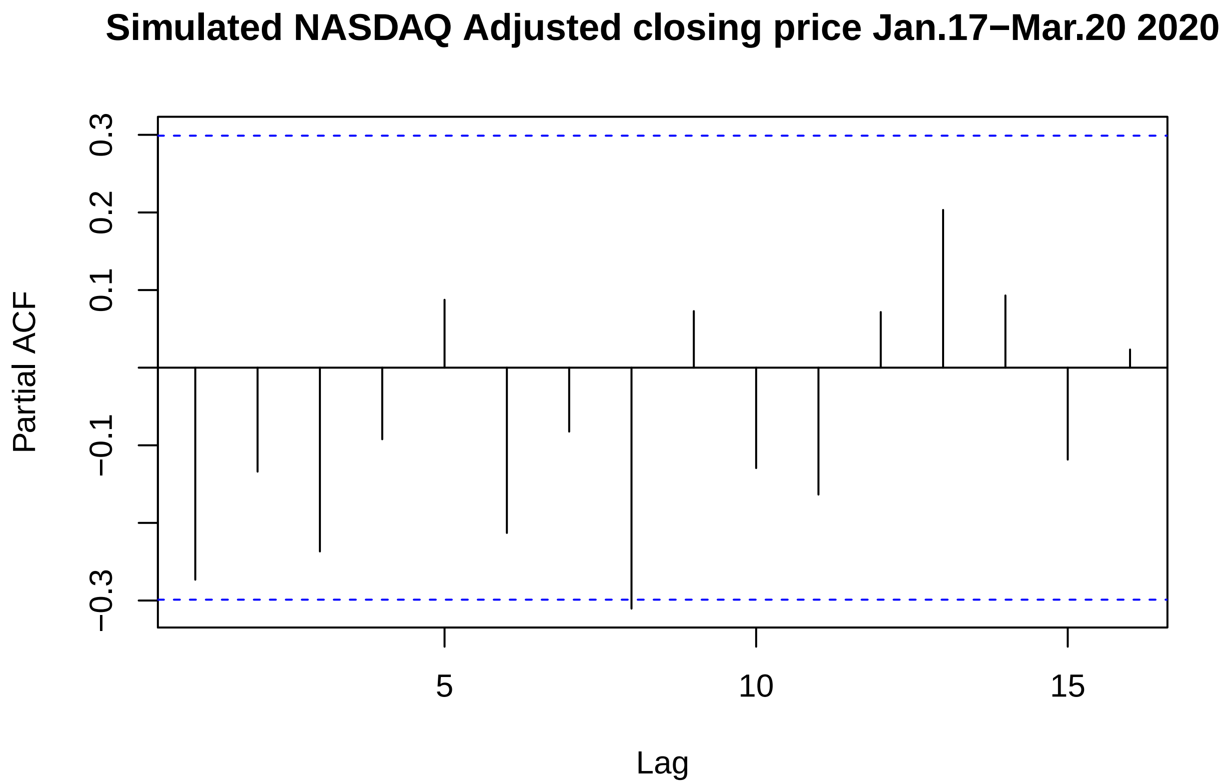


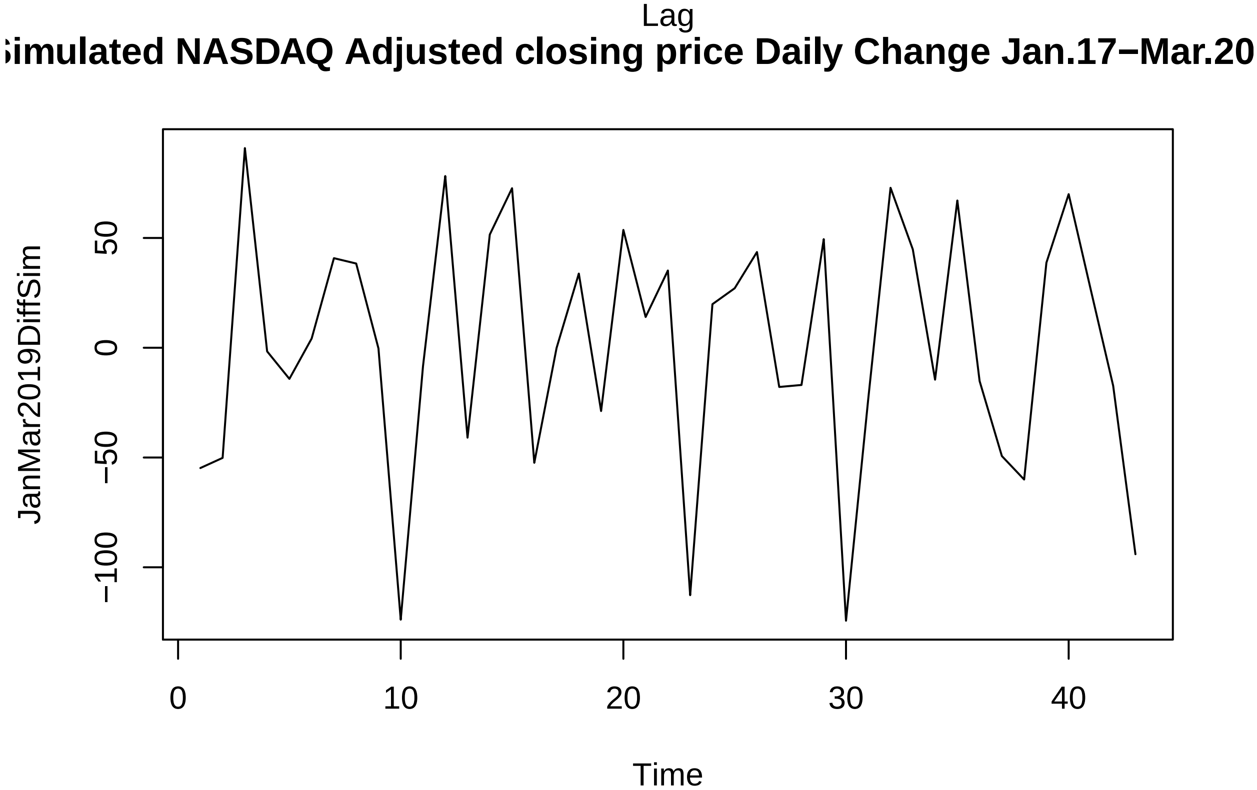
Model diagnostics of Manually-fitted Model

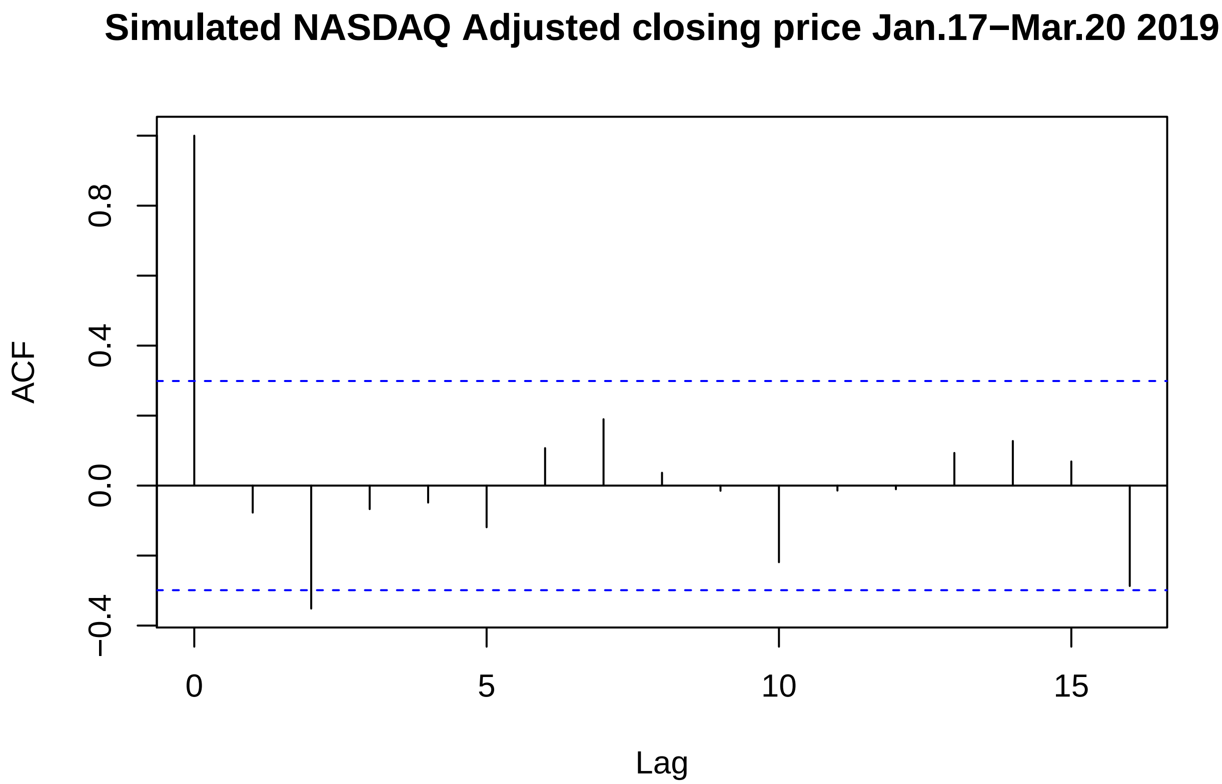
Fitting AR(1), AR(2), AR(3) models to 2020 January-March, 2019 January-March, 2019 October-December data is reasonable. The model diagonisics of all three fitted models via tsdiag command tells that the Residual plots are random, independent and the p-values for Ljung-Box is higher than 0.05. Thus, we do not reject our null-hypothesis which is the model fits well. We can also simulate AR(1), AR(2), AR(3) models and see its behaviours.

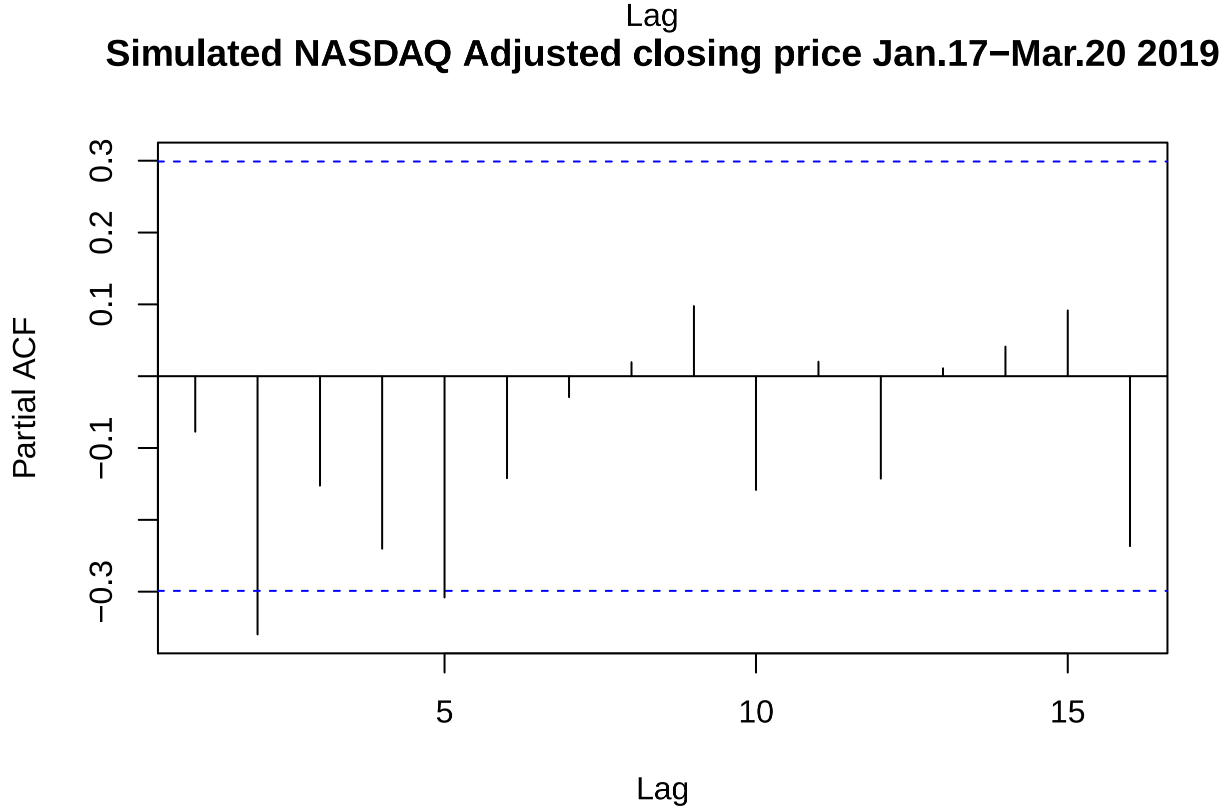

Simulated NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20
JanMarDiffSim
400
200
0
−200
−600
0
10
20
30
40
Time


Simulated NASDAQ Adjusted closing price Jan.17−Mar.20 2020
ACF
1.0
0.8
0.6
0.4
0.2
−0.2
0
5
10
15
Lag

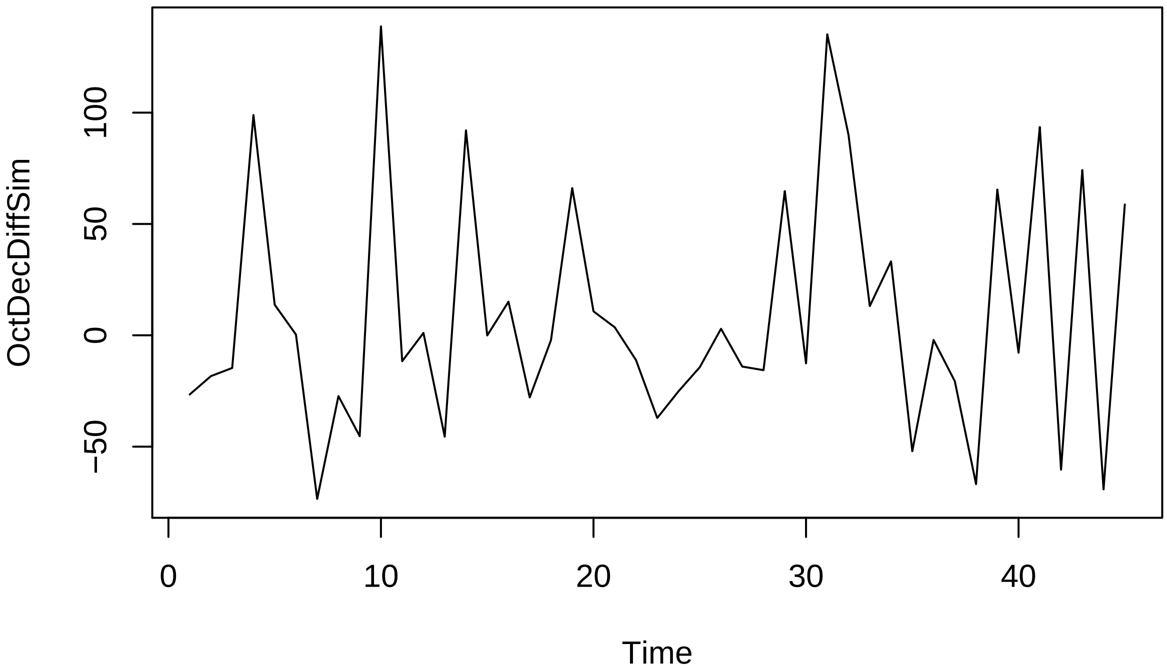

Simulated NASDAQ Adjusted closing price Jan.17−Mar.20 2020
Partial ACF
0.3
0.2
0.1
−0.1
−0.3
5
10
15
Lag


imulated NASDAQ Adjusted closing price Daily Change Jan.17−Mar.20
JanMar2019DiffSim
50
0
−50
−100
0
10
20
30
40
Time

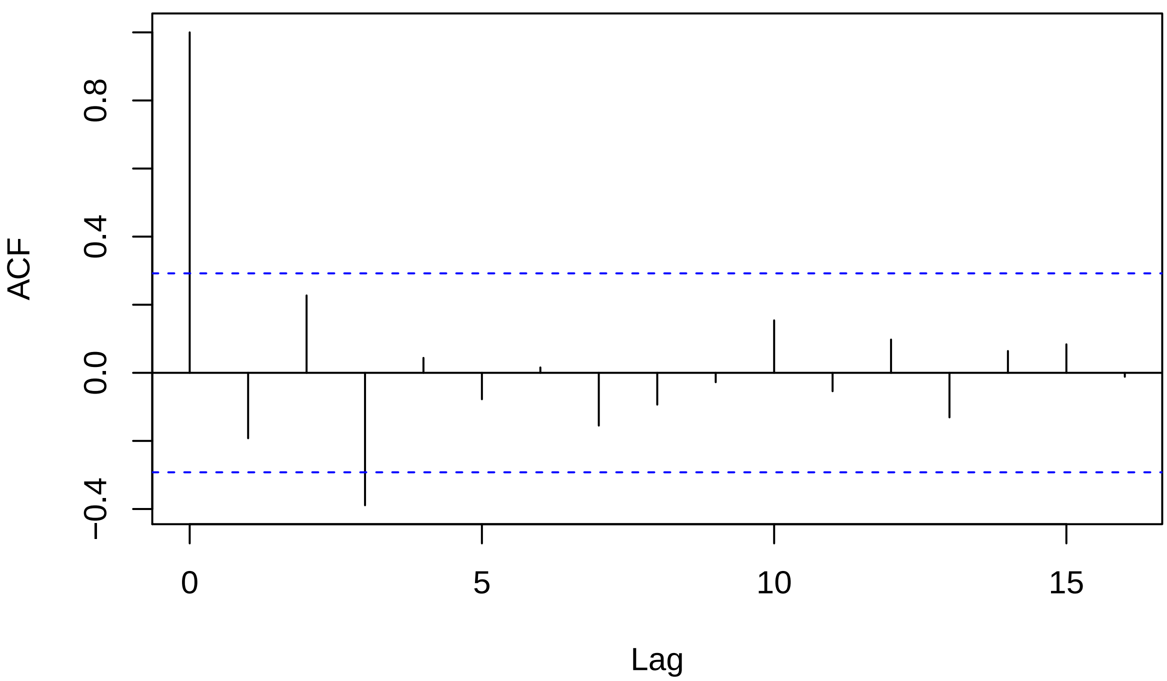

Simulated NASDAQ Adjusted closing price Jan.17−Mar.20 2019
ACF
−0.4
0.0
0.4
0.8
0
5
10
15
Lag

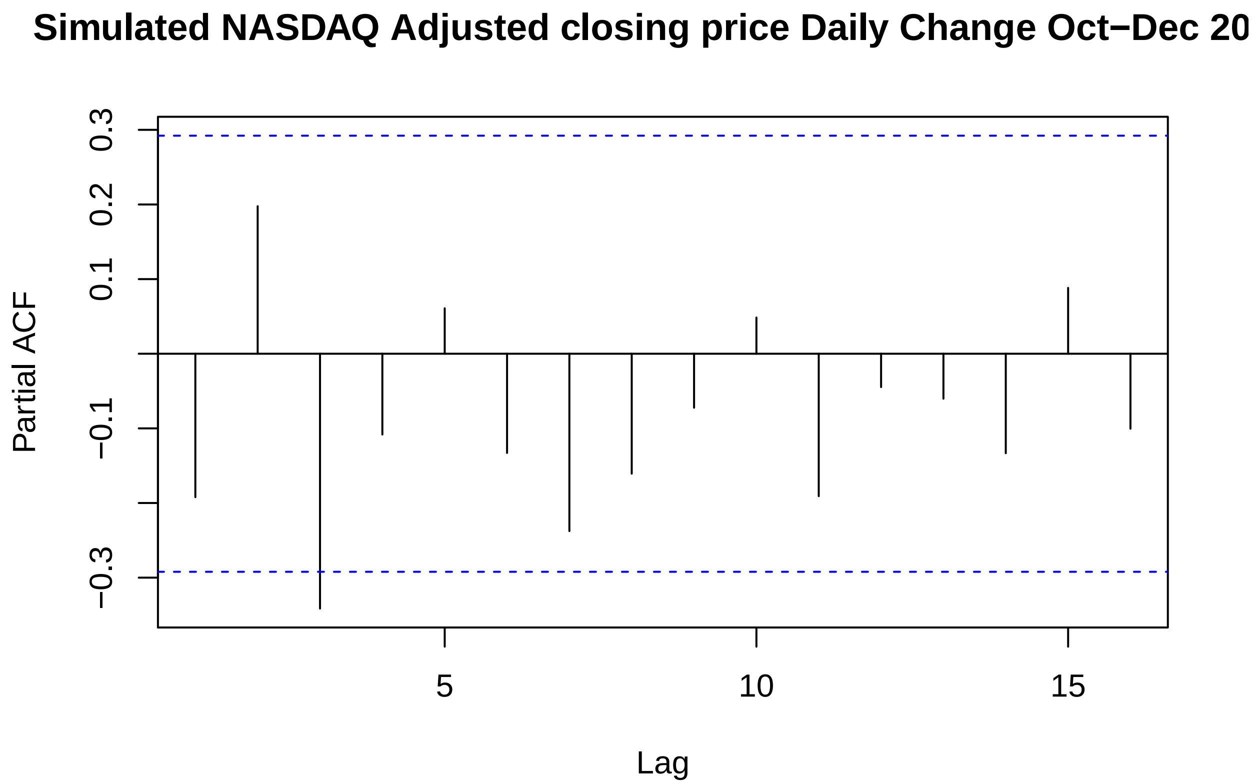

Simulated NASDAQ Adjusted closing price Jan.17−Mar.20 2019
Partial ACF
−0.3
−0.1
0.1
0.2
0.3
5
10
15
Lag

Simulated NASDAQ Adjusted closing price Daily Change Oct−Dec 20



Simulated NASDAQ Adjusted closing price Daily Change Oct−Dec 20



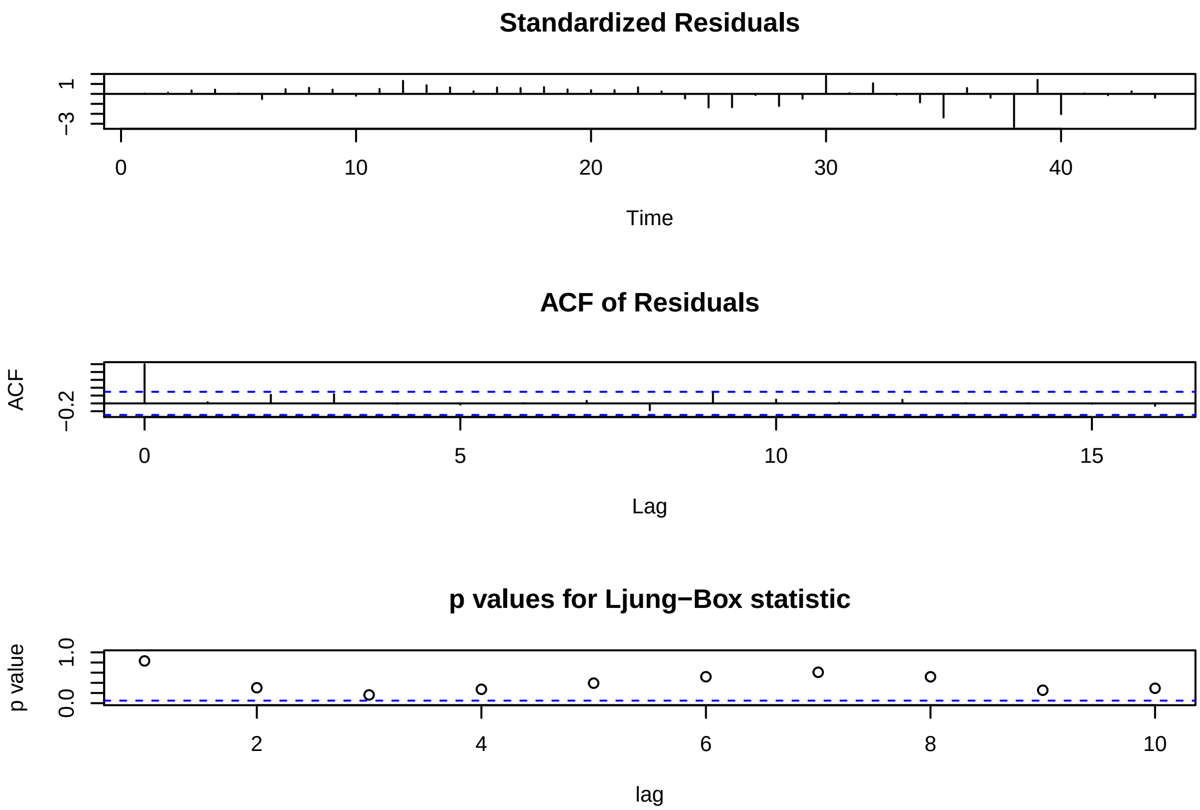


Manually-fitted Model Analysis

The simulated data behaviours as we expected. The simulated models are not perfectly look like our differenced plots, but it is still gives the information the behaviours are similar. The ACF plots tails off, and PACF plots cuts off at lag p (AR(p)) for all three models. Model diagnoistics and the simulated models using our fitted models tell it matches expectation from First Differences Anaylsis.

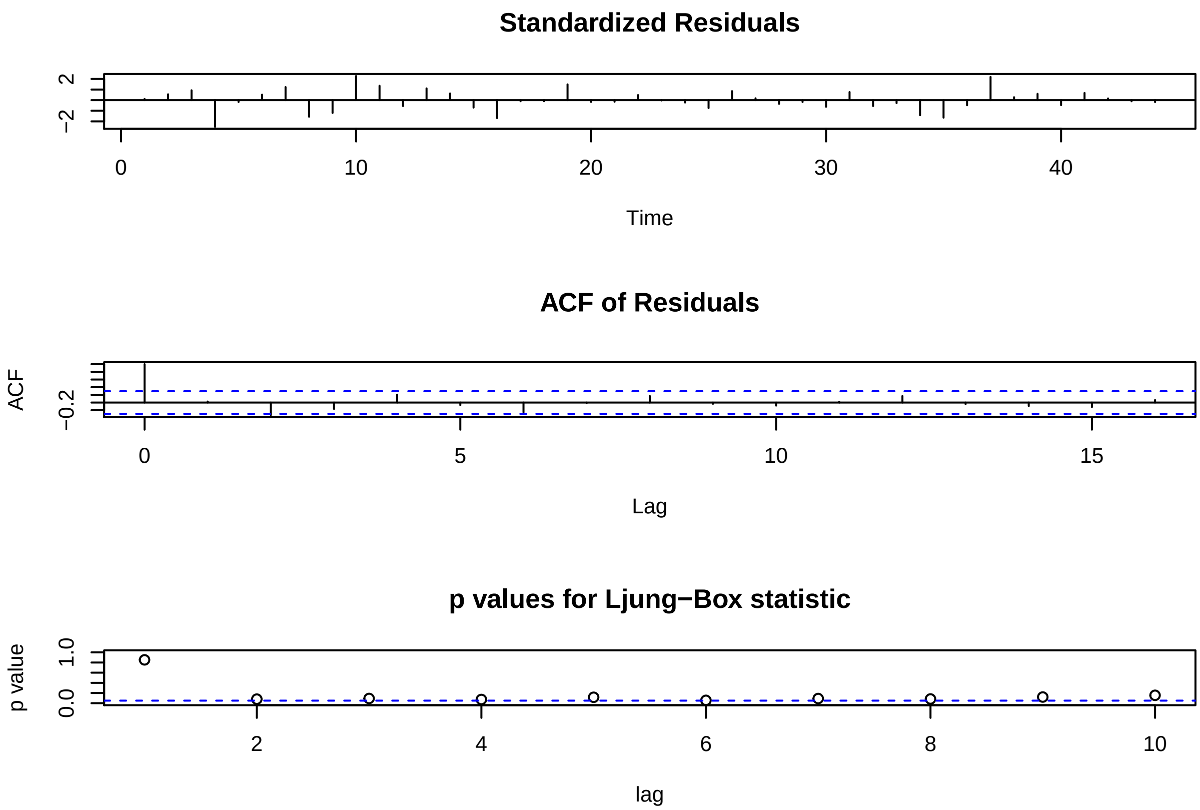
Automatically-Fitted Models

```
JanToMarAuto = auto.arima(JanToMar$Adj.Close)
tsdiag(JanToMarAuto)
```



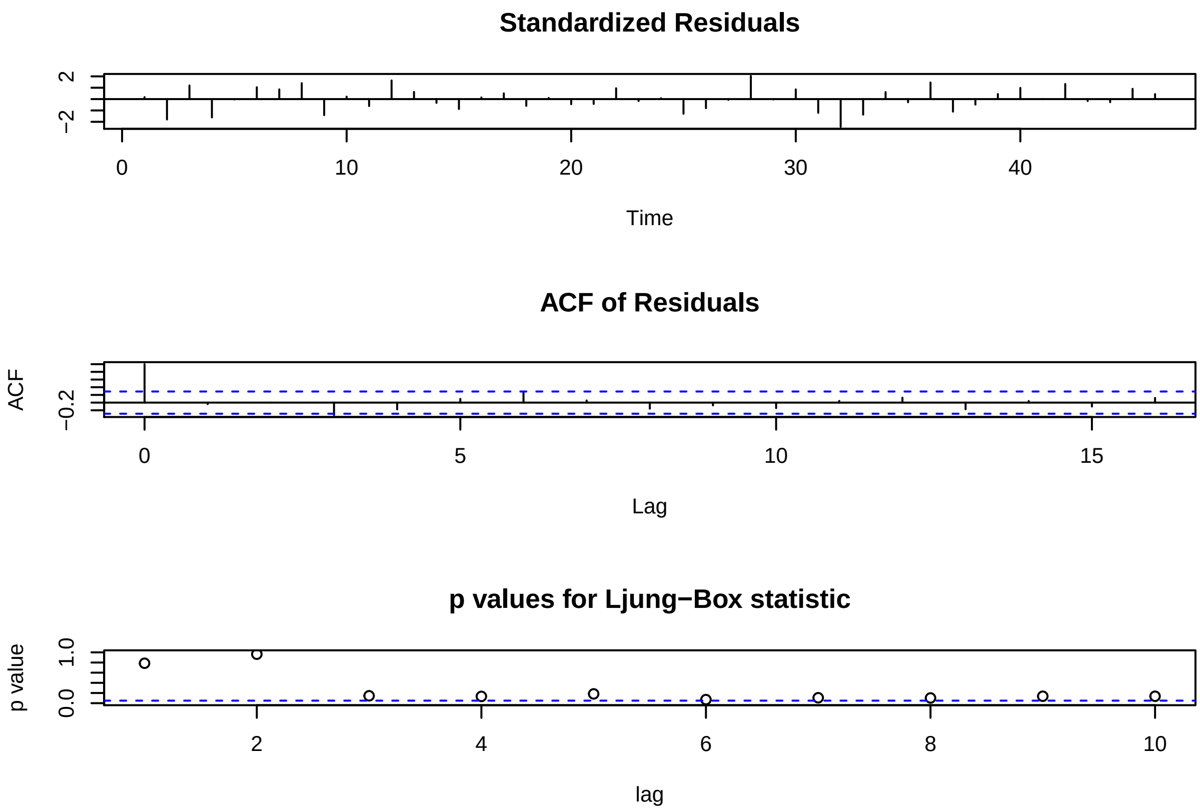

```
#ARIMA(1,1,0)

JanToMar2019Auto = auto.arima(JanToMar2019$Adj.Close)
tsdiag(JanToMar2019Auto)
```




```
#ARIMA(0,1,0)

OctToDecAuto = auto.arima(OctToDec$Adj.Close)
tsdiag(OctToDecAuto)
```




```
#ARIMA(0,1,0)
```

Global pandemics are not common variables to include within models, and so, as we are presented with the opportunity to see how an actual pandemic can affect model fitting, we decided to use the auto.arima() function from the package "forecast". The best ARIMA models were automatically fit by the function based on AIC. First looking at both the 2019 series, auto.arima() returns an ARIMA(0,1,0), while an ARIMA(1,1,0) is returned for the January to March 2020 series.

As expected, the time series affected by the pandemic deviates from the other two. Intuitively, one would think a model based on a stock market index should exhibit AR(P) coefficients, where $P > 0$. Thus, by removing stationary effects such as a trend on what we suspect is an AR(1) model, sequential observations in a differenced series should only be dictated by white noise. This is confirmed by the ARIMA(0,1,0) models returned by auto.arima() for both the 2019 series.

```
Since fitted models do not capture reality perfectly and exhaustively, we can see the realization of thi
```

Forecasting Model

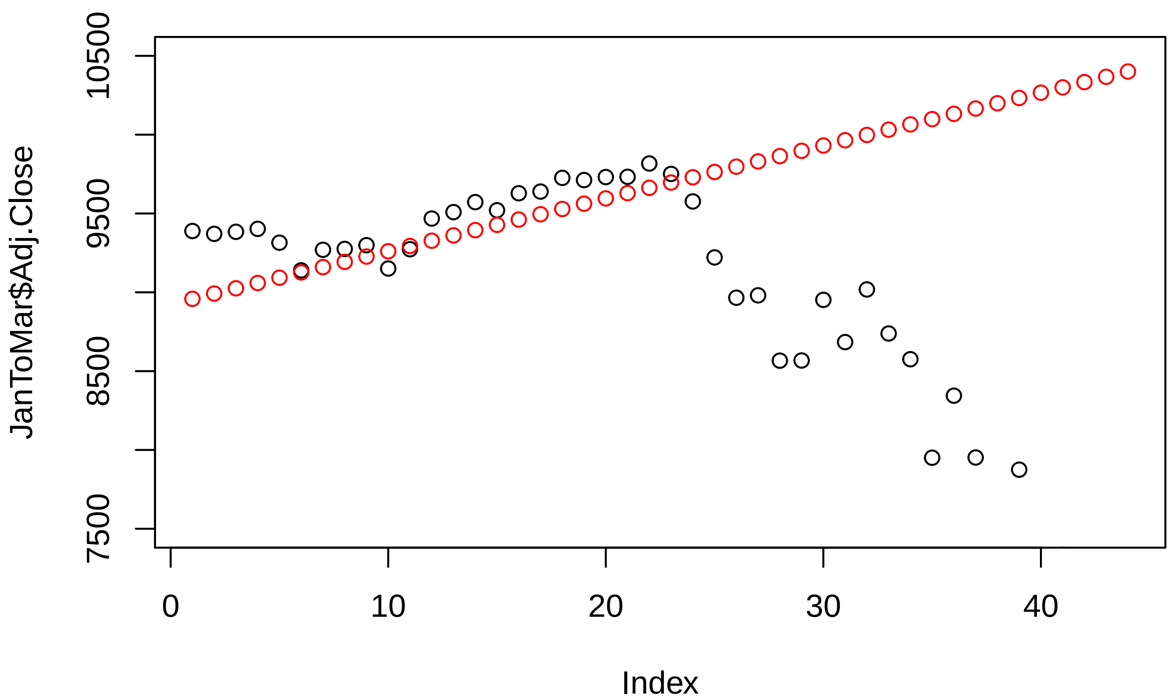
To see how much the real data for January to March 2020 diverged, we can compare it to the forecasted values of a Holt-Winters model trained on the data for October to December 2019, where the NASDAQ index still seemed to be behaving "normally". While there is a grain of salt to be taken here with such a forecasted model since we're essentially forecasting as many data points as there are in the training set, We think it'll still be an interesting comparison to make.

```
hw <- HoltWinters(as.ts(OctToDec$Adj.Close,frequency=7), alpha = NULL, beta = NULL,
gamma = FALSE, seasonal = c("additive", "multiplicative"), l.start = NULL, b.start = NULL, s.start = NU
hw
```

```
## Holt-Winters exponential smoothing with trend and without seasonal component.
##
```

```
## Call:
## HoltWinters(x = as.ts(OctToDec$Adj.Close, frequency = 7), alpha = NULL,     beta = NULL, gamma = FALS
##
## Smoothing parameters:
##  alpha: 1
##  beta : 0.2407177
##  gamma: FALSE
##
## Coefficients:
##          [,1]
## a 8924.95996
## b   33.53791
```

```
JanToMar.predict <- predict(hw, n.ahead = 44)
plot(JanToMar$Adj.Close, ylim=c(7500,10500))
points(1:44, JanToMar.predict,col="2")
```



###

Analysis of Forecast As we can see, with the dearth of data the Holt-Winters method resulted in a prediction that appears almost linear. However, the predictions for a look-ahead of 44 actually appears to fit the data quite well for the data points leading up to before news and concerns of coronavirus really spread. The data only really diverges from this model starting in the time point corresponding to early Feburary where the markets started reacting.

It's certainly hard to say for sure with how our forecast model was fitted, and how the models don't account for anything other potential shift in trends, the NASDAQ index maybe have continued experiencing a positive trend throughout Febuary and into March.

Conclusion

In conclusion, our models show that Covid-19 affected the NASDAQ adjusted closing price in a few facets. It greatly increased the day to day volatility of the daily change in the price as evidenced by our analysis of the time series for the first differences. It has also certianly shifted the course of the trend to be a steep negative trend as opposed to the gentle positive trend that the data from the preceeding 3 months would indicate.

All in all, our expectations that the growing concerns of the Coronavirus would greatly impact the adjusted closing price of NASDAQ were met fully. Through our analysis, we have come to see how we should not blindly follow models and believe they can predict major events such as this one.