Final Project: Milestone 2

Mourad Askar, Gerardo Palacios, Joe Paszkiewicz, Evelina Ramoskaite[1,2]

[1] DSC 425 - Time Series Analysis and Forecasting

[2] DePaul University

Final Project: Milestone 2

Crypto-Currency and Stock Data

```
##      Index                   data
##  Min.   :2016-10-01   Min.   :   6.822
##  1st Qu.:2017-12-31   1st Qu.: 158.418
##  Median :2019-04-01   Median : 255.537
##  Mean   :2019-04-01   Mean   : 589.463
##  3rd Qu.:2020-06-30   3rd Qu.: 557.063
##  Max.   :2021-09-30   Max.   :4168.701
##                       NA's   :4
```

Graphing the Time Series

Figure 1 shows the time series for Ethereum over the past 5 years. It appears to be a multiplicative, non-stationary time series with an exponential positive trend that has exploded most recently in 2021.

```
autoplot(data)
```

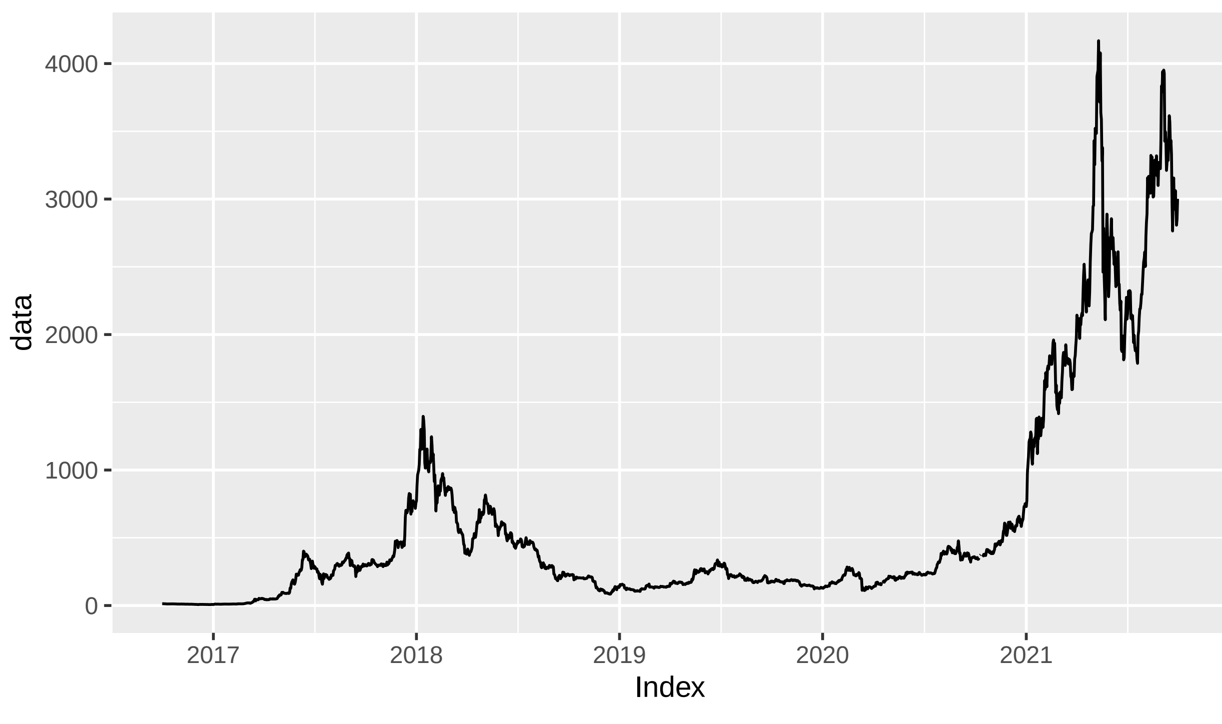

Figure 1

Figure 2, we can see the time series with a log transformation. It has transformed the exponential behavior into something more linear. There still remains a general increasing trend, and appears to be more additive.

```
autoplot(log(data))
```

Figure 2

Figure 3, we can see the log returns. The plot shows general white noise with a few outliers in 2017 and 2020.

```
autoplot(diff(log(data)))
```

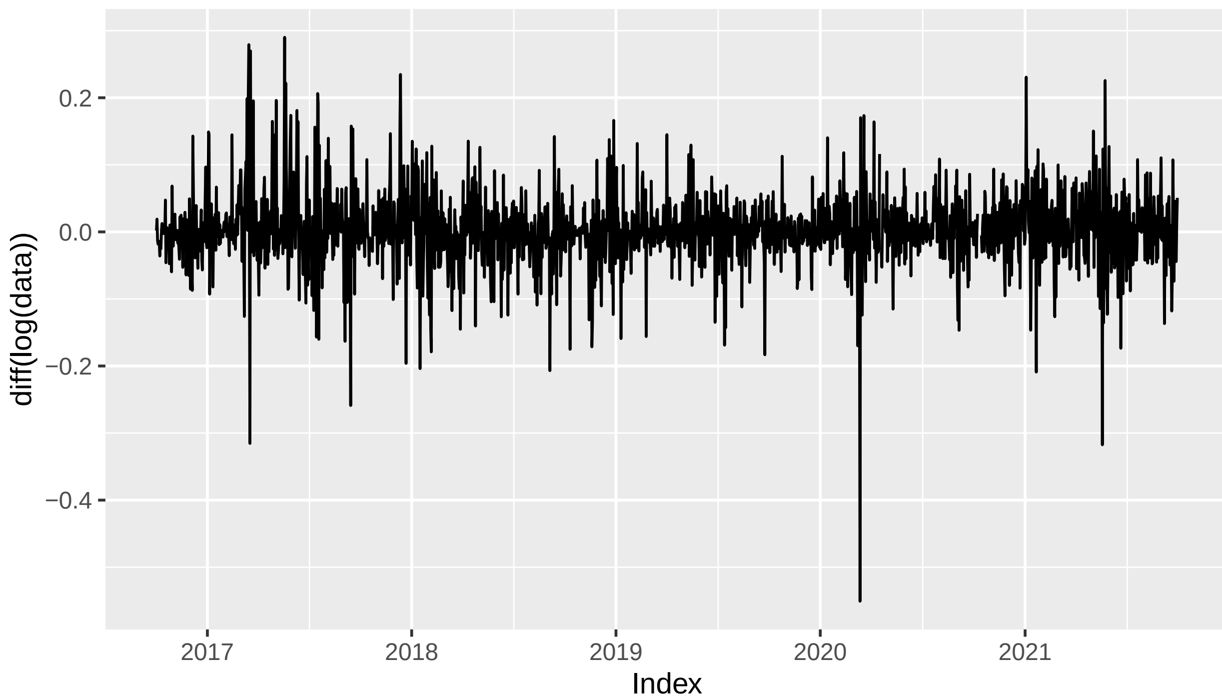

Figure 3

Auto-correlation

ACF of the time series

Figure 4 is the ACF plot. Auto-correlation has a strong presence in this time series. The ACF gradually decreases indicating a non-stationary series.

```
acf((data), na.action = na.pass)
```

ACF of the Log Returns

```
acf(diff(log(data)), na.action = na.pass)
```

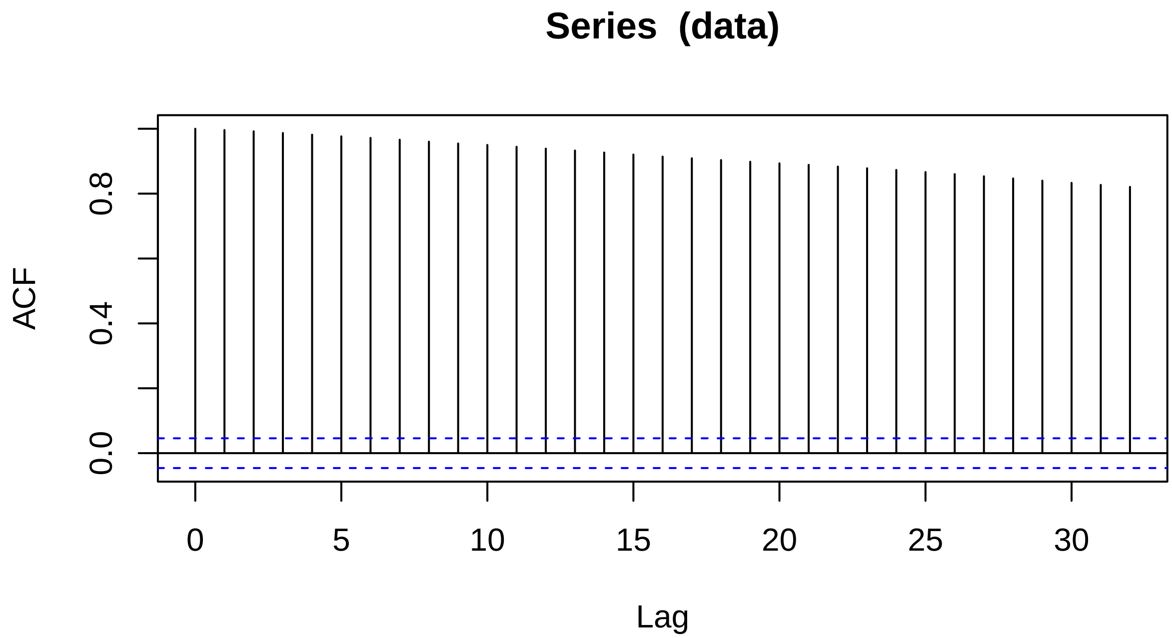

Figure 4

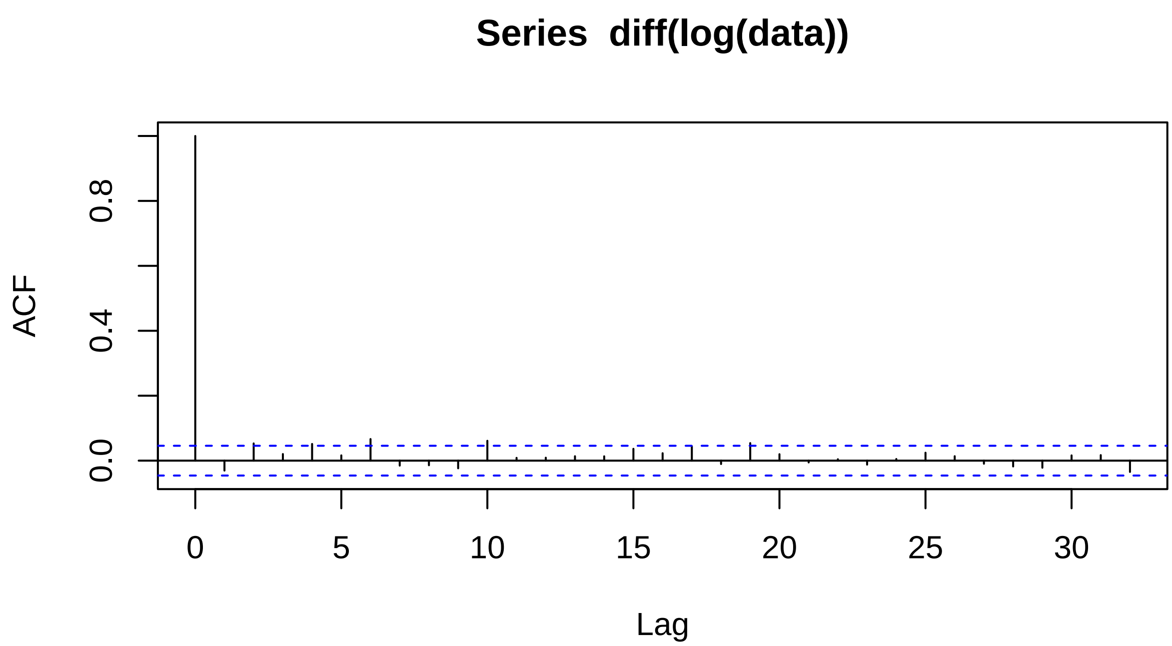

Figure 5

Ljung Box Test

This can be further confirmed by performing the Ljung Box test. At lag 1, the Ljung-Box p-value is close to zero. This indicates that at the 99% confidence, the null hypothesis is rejected and one can conclude that the series is not independently distributed and exhibit serial correlation.

```
Box.test(log(data), lag = 1, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  log(data)
## X-squared = 1819.9, df = 1, p-value < 2.2e-16
```

Coinbase and Cardano

Coinbase Summary Statistics

```
##      Index              stock.COIN
##  Min.   :2021-04-14   Min.   :220.6
##  1st Qu.:2021-05-25   1st Qu.:232.2
##  Median :2021-07-07   Median :244.8
##  Mean   :2021-07-06   Mean   :252.6
##  3rd Qu.:2021-08-17   3rd Qu.:260.8
##  Max.   :2021-09-29   Max.   :342.0
```

Cardano Summary Statistics

```
##      Index              stock.ADA
##  Min.   :2017-10-01   Min.   :0.01854
##  1st Qu.:2018-10-01   1st Qu.:0.04740
##  Median :2019-10-01   Median :0.09291
##  Mean   :2019-10-01   Mean   :0.35803
##  3rd Qu.:2020-09-30   3rd Qu.:0.28847
##  Max.   :2021-09-30   Max.   :2.96824
##                       NA's   :4
```

Graphing the Time Series

The figures below show the time series for Coinbase and Cardano (ADA-USD) since their inception in 2021 and 2017, respectively. While Cardano appears to be a multiplicative, non-stationary time series, Coinbase seem to be an additive time series.

In the figures below we can see both time series with a log transformation. The log transformation did not affect Coinbase much, confirming the additive nature. On the other hand, it has transformed Cardano's into a more stable form.

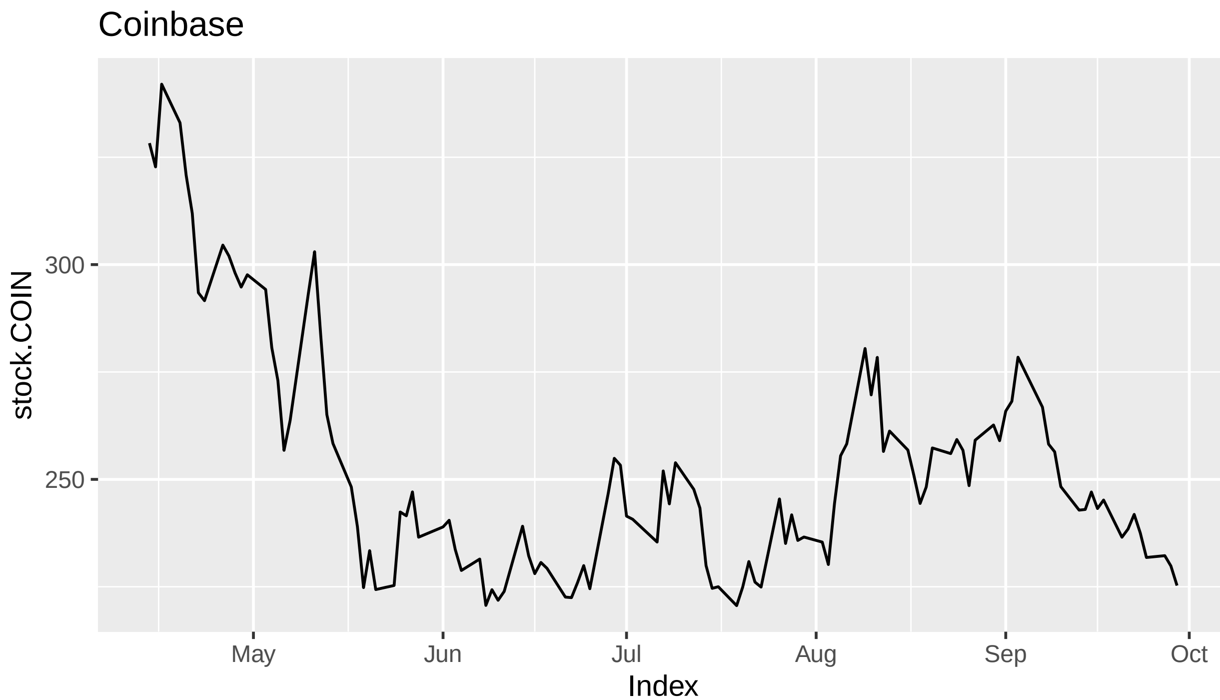

Figure 6

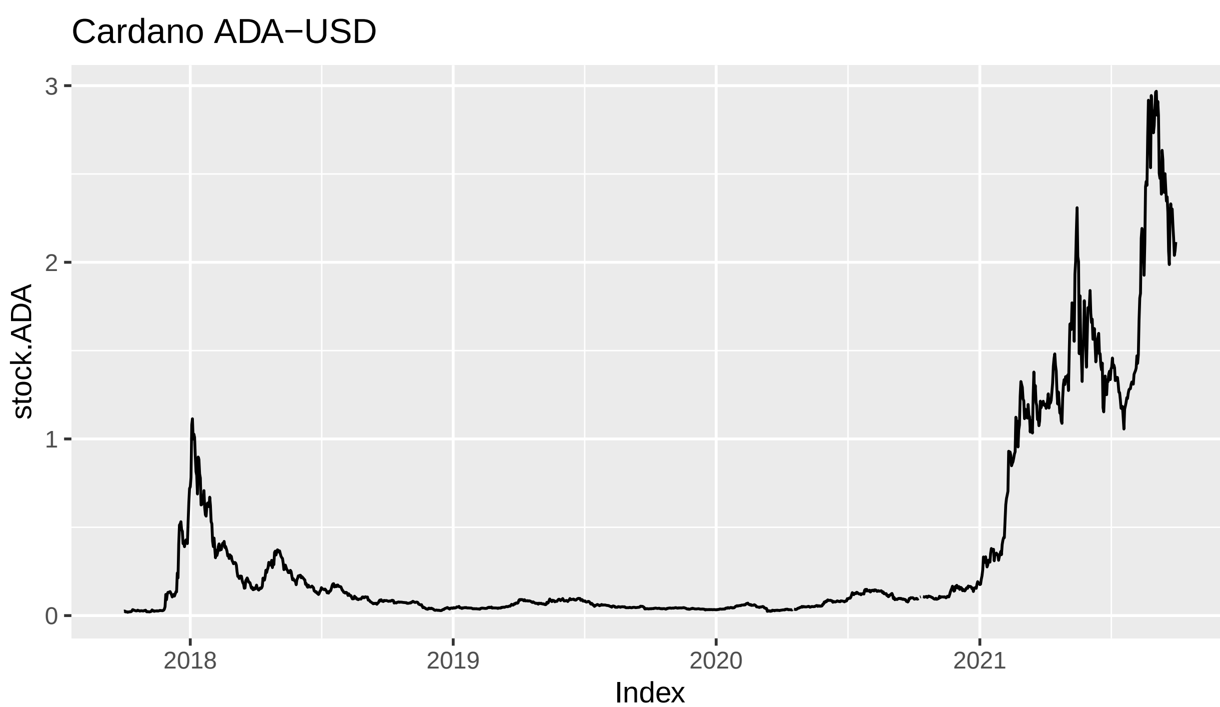

Figure 7

```
autoplot(log(stock.COIN)) +
  ggtitle('Coinbase "Log"')
```

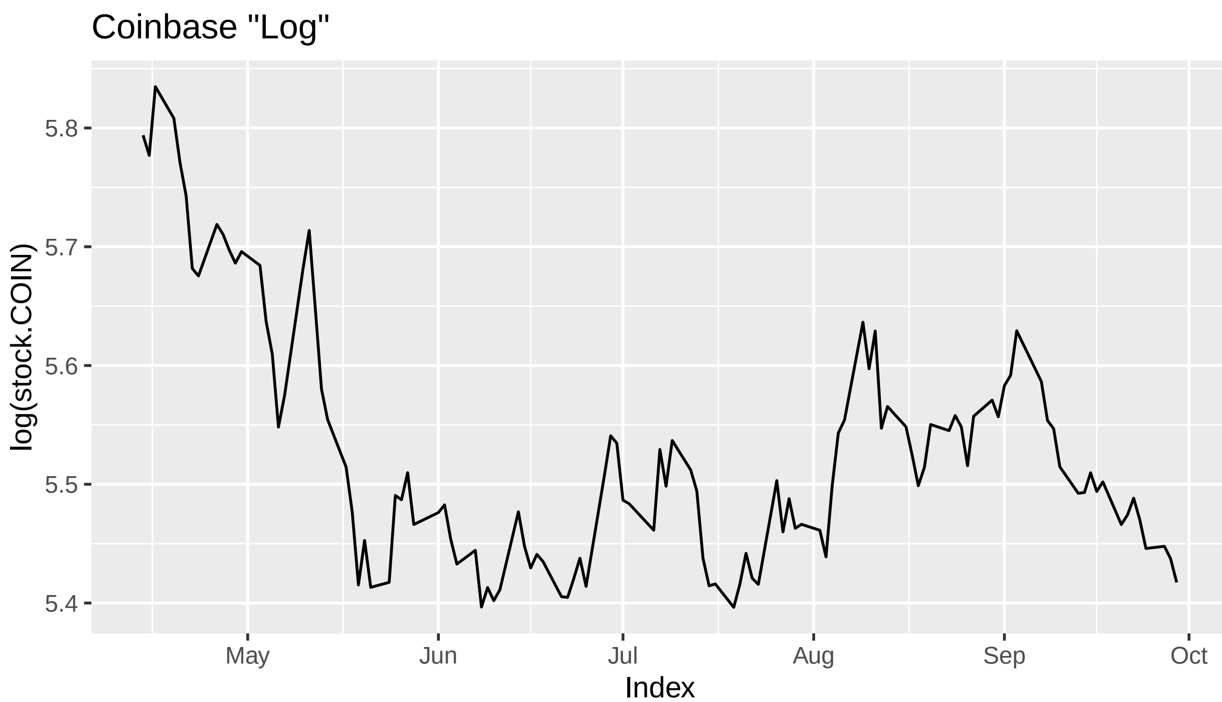

Figure 8

```
autoplot(log(stock.ADA)) +
  ggtitle('Cardano ADA-USD "Log"')
```

The figures below, we can see the log returns. The plot shows general white noise in both tickers with a few outliers fro Cardano in late 2017 and early 2020.

```
autoplot(diff(log(stock.COIN))) +
  ggtitle('Coinbase "Log Returns"')
```

```
autoplot(diff(log(stock.ADA))) +
  ggtitle('Cardano ADA-USD "Log Returns"')
```

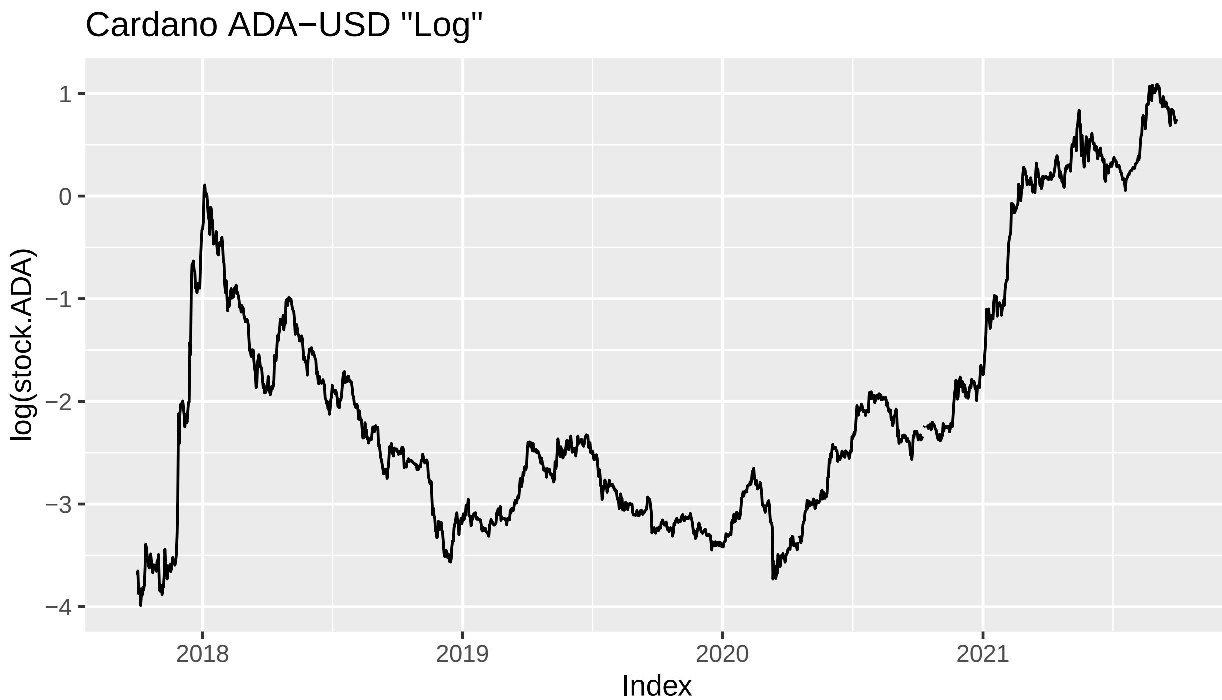

Figure 9

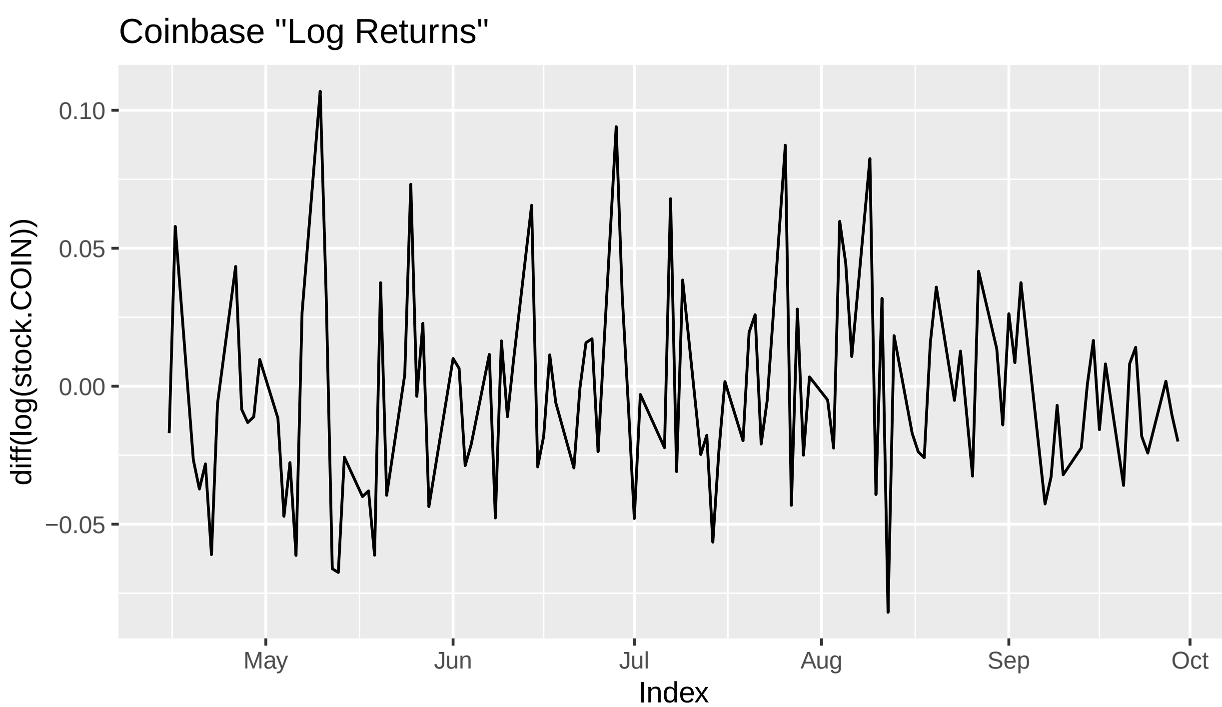

Figure 10

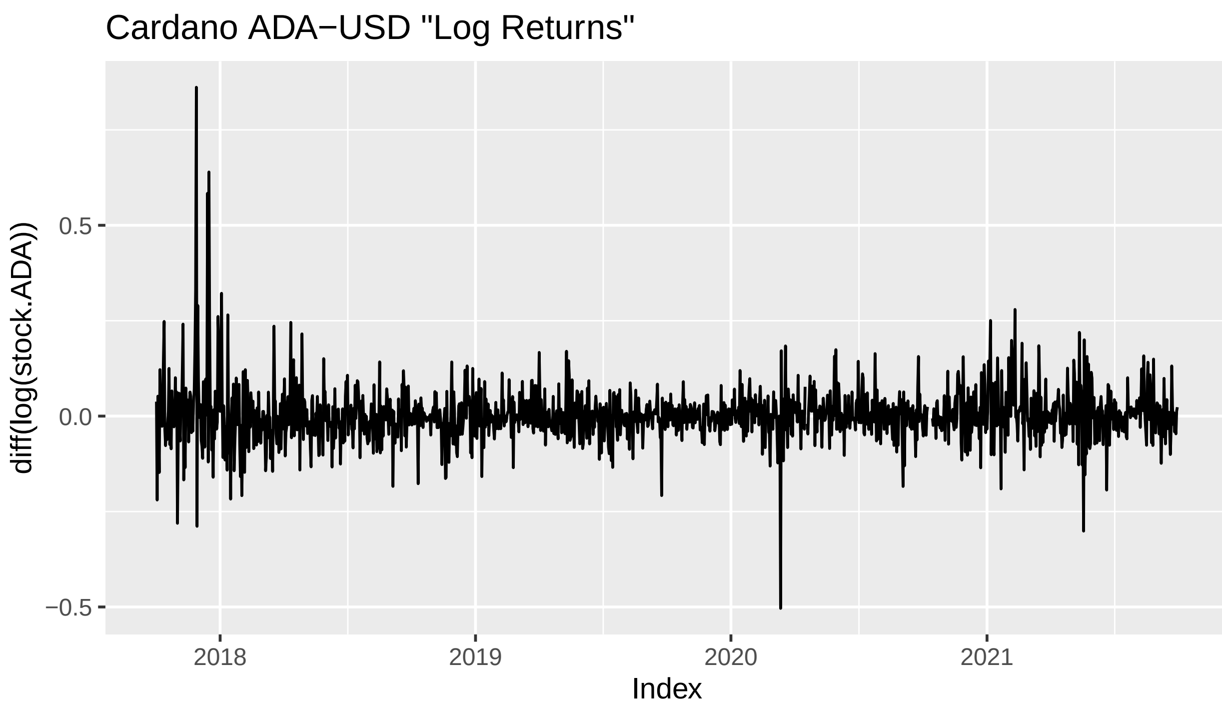

Figure 11

Auto-correlation for the series

The ACF plots shows that Cardano's Auto-correlation has a strong presence in this time series. The ACF gradually decreases indicating a non-stationary series. Coinbase's ACF quickly decays indicating also indicating non-stationary series.

```
acf((stock.COIN), na.action = na.pass)
```

```
acf((stock.ADA), na.action = na.pass)
```

Auto-correlation for the log returns

```
acf(diff(log(stock.COIN)), na.action = na.pass)
```

```
acf(diff(log(stock.ADA)), na.action = na.pass)
```

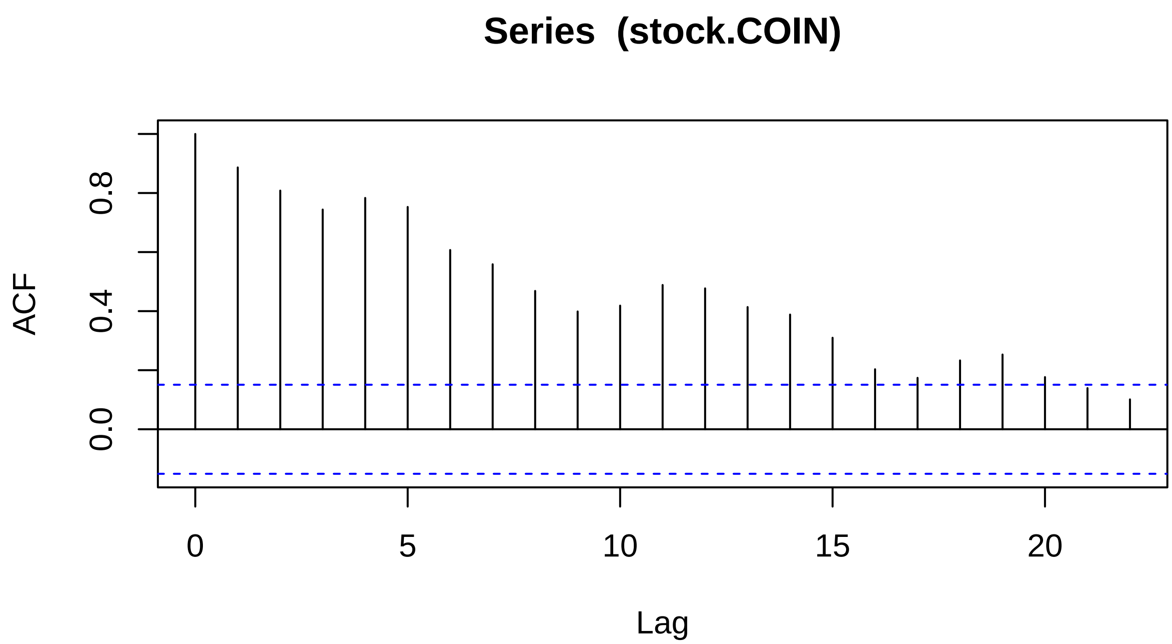

Figure 12

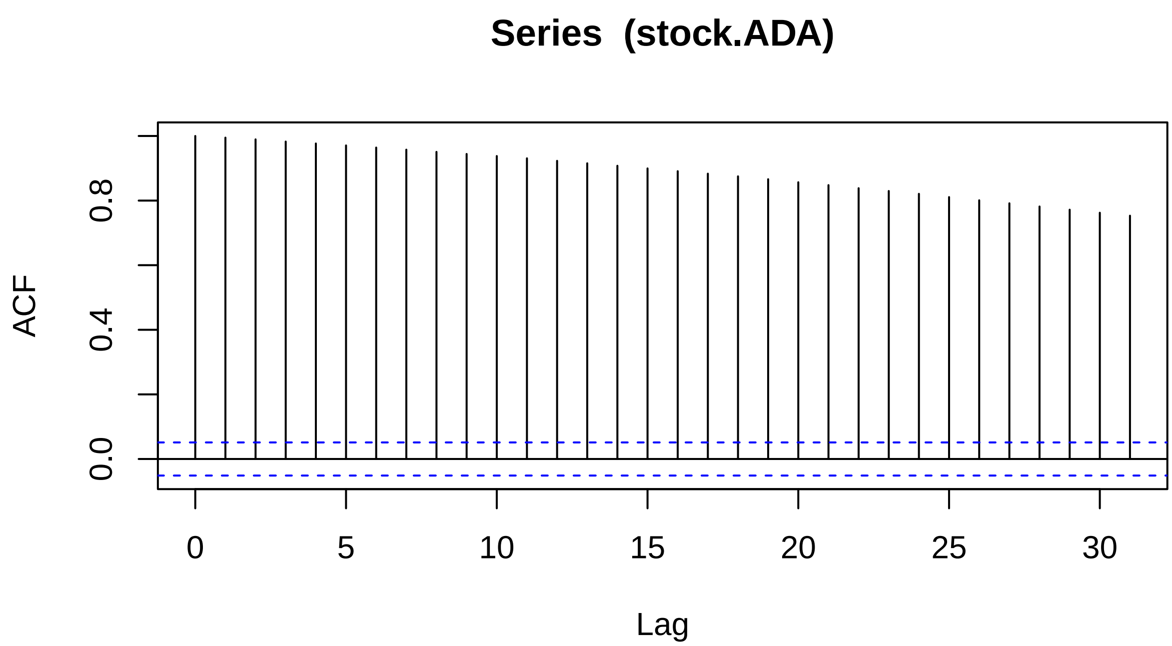

Figure 13

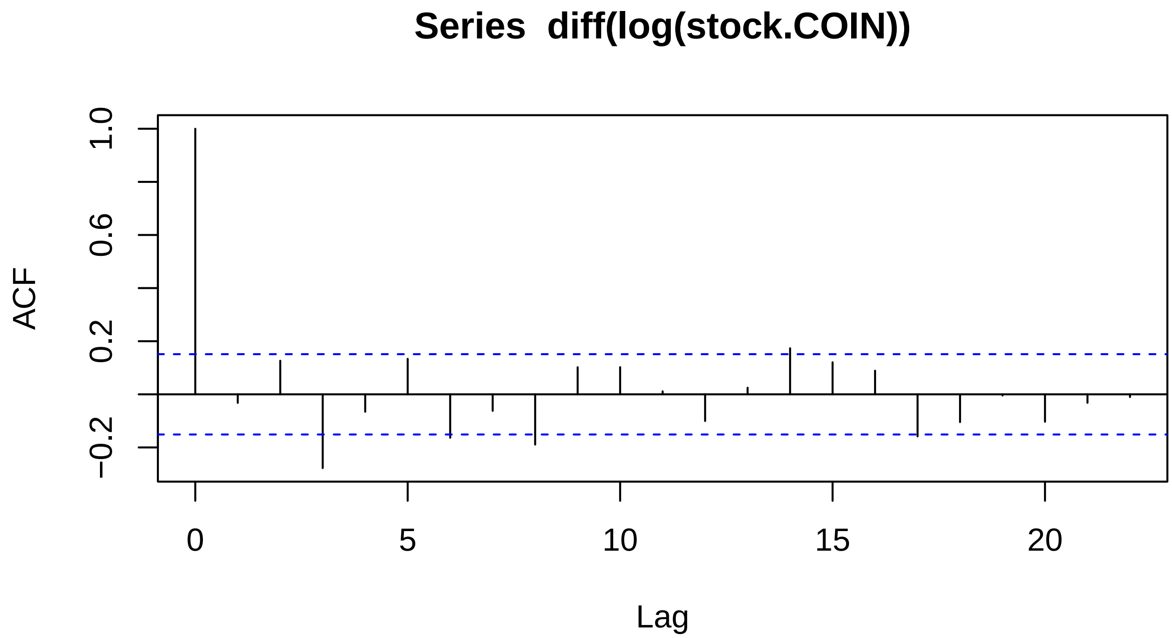

Figure 14

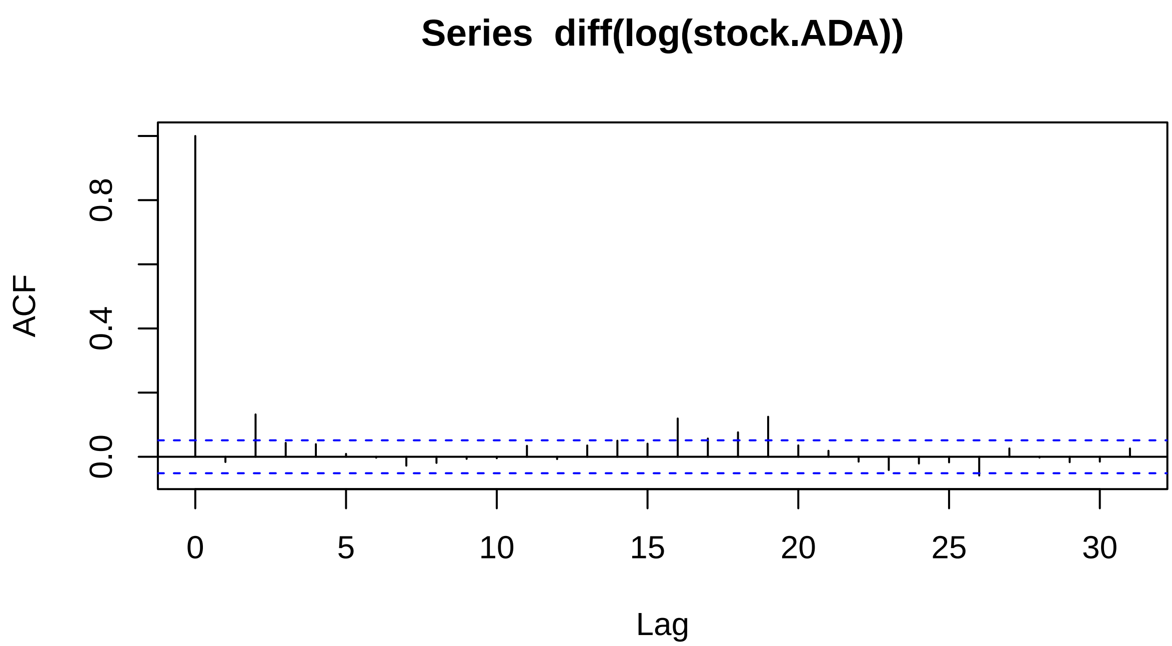

Figure 15

Ljung Box Test

This can be further confirmed by performing the Ljung Box test. At lag 100, the Ljung-Box p-value is close to zero. This indicates that at the 99% confidence, the null hypothesis is rejected and one can conclude that the series is not independently distributed and exhibit serial correlation.

```
Box.test(log(stock.COIN), lag = 100, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  log(stock.COIN)
## X-squared = 1029.8, df = 100, p-value < 2.2e-16
```

```
Box.test(log(stock.ADA), lag = 100, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  log(stock.ADA)
## X-squared = 85981, df = 100, p-value < 2.2e-16
```

Bitcoin

Bitcoin Summary Statistics

##From 2016 to 2021, Bitcoin had a minimum value of $610, and a maximum value of $63,503.

```
##      Index              stock.BTC
##  Min.   :2016-10-01   Min.   :  610.2
##  1st Qu.:2017-12-31   1st Qu.: 4110.7
##  Median :2019-04-01   Median : 8040.3
##  Mean   :2019-04-01   Mean   :12728.8
##  3rd Qu.:2020-06-30   3rd Qu.:11346.4
##  Max.   :2021-09-30   Max.   :63503.5
##                       NA's   :4
```

Graphing the Bitcoin Time Series

The figures below show the time series for bitcoin(BTC-USD)from 2016 to 2021. Bitcoin appears to be a multiplicative, non-stationary time series.

```
autoplot(stock.BTC) +
  ggtitle('Bitcoin Prices in USD')
```

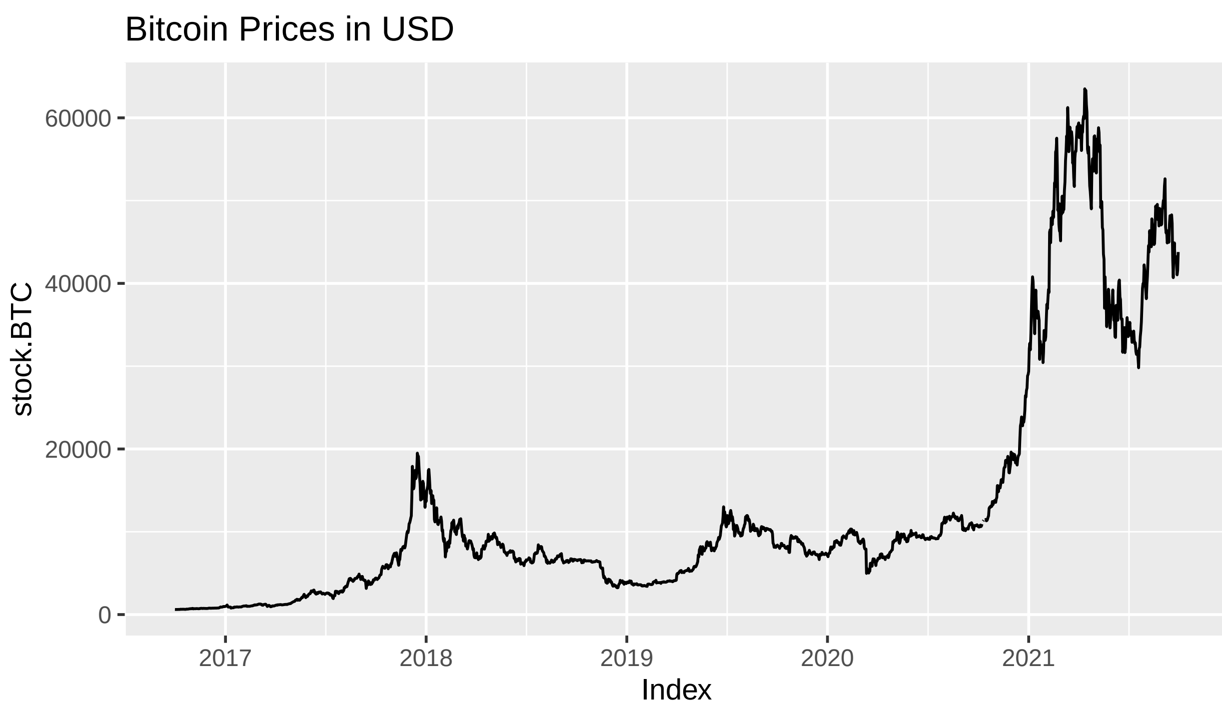

Figure 16

Log Transformed Series

```
autoplot(log(stock.BTC)) +
  ggtitle('Bitcoin Price in USD (Log)')
```

Figure 17

Log Returns

The log return plot for Bitcoin shows a stark contrast to the original plot. It is more uniform in comparison, confirming that the time series shows strong multiplicative behavior.

```
autoplot(diff(log(stock.BTC))) +
  ggtitle('Bitcoin Log Returns')
```

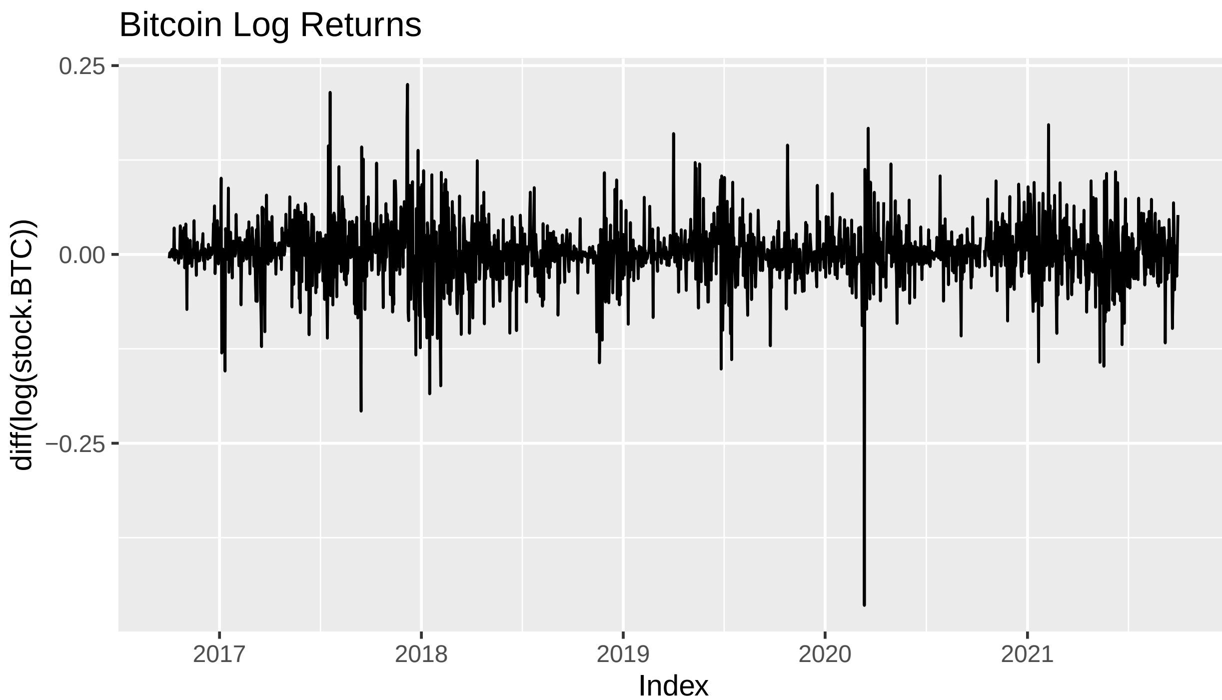

Figure 18

Acf for the Bitcoin Series

The acf plot of series has a very slow decay, indicating that the series is not stationary.

```
acf(((stock.BTC)), na.action = na.pass)
```

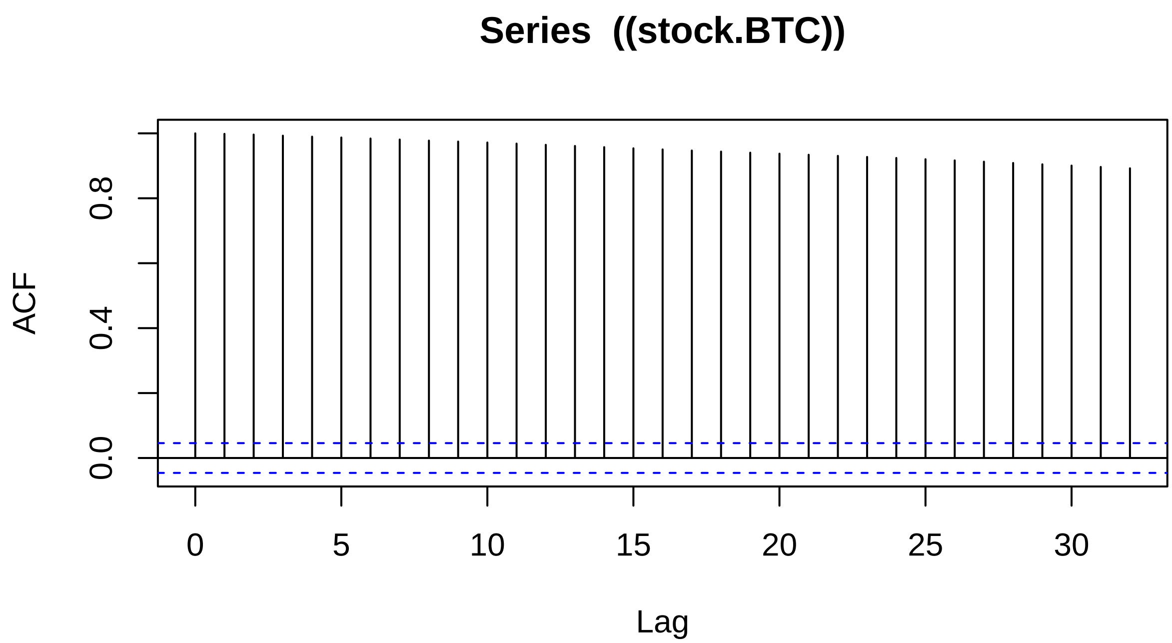

Figure 19

Acf of the log returns

The acf of the log returns is shows little activity above the confidence interval, showing behavior akin to white noise.

```
acf(diff(log(stock.BTC)), na.action = na.pass)
```

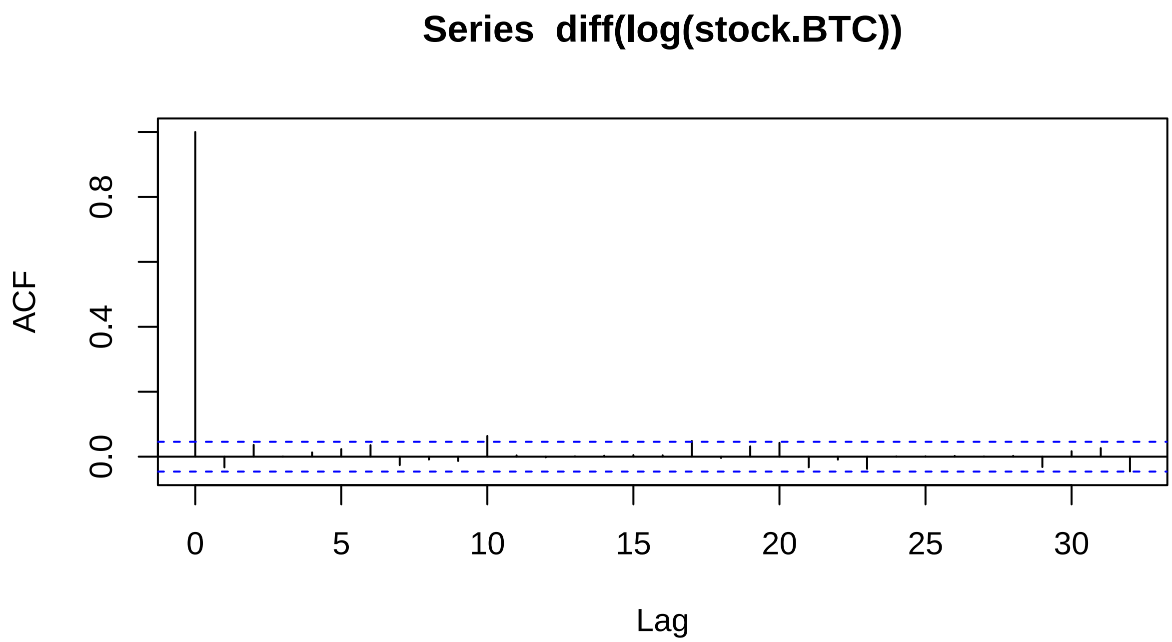

Figure 20

Ljung-Box test

The Ljung box test rejects the null hypothesis that the log(Bitcoin)series is stationary. However, when the test is run on the log returns, the null hypothesis is not rejected, showing some promise for forecasting.

```
# log (Bitcoin)
Box.test((log(stock.BTC)), lag = 100, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  (log(stock.BTC))
## X-squared = 134390, df = 100, p-value < 2.2e-16
```

```
# log returns
Box.test(diff(log(stock.BTC)), lag = 100, type = "Ljung-Box")
```

```
##
##  Box-Ljung test
##
## data:  diff(log(stock.BTC))
## X-squared = 112.16, df = 100, p-value = 0.191
```