Time Series Forecasting

Homework -6

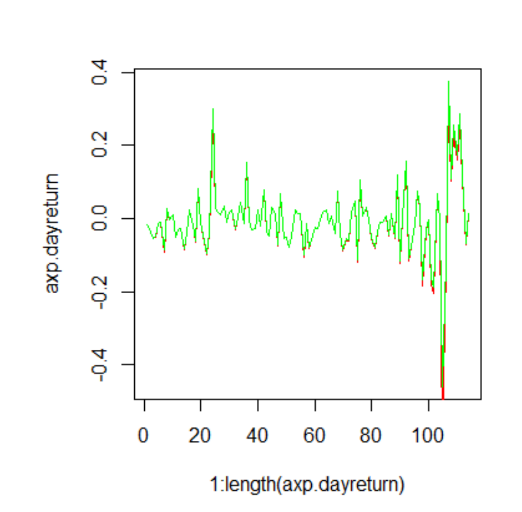
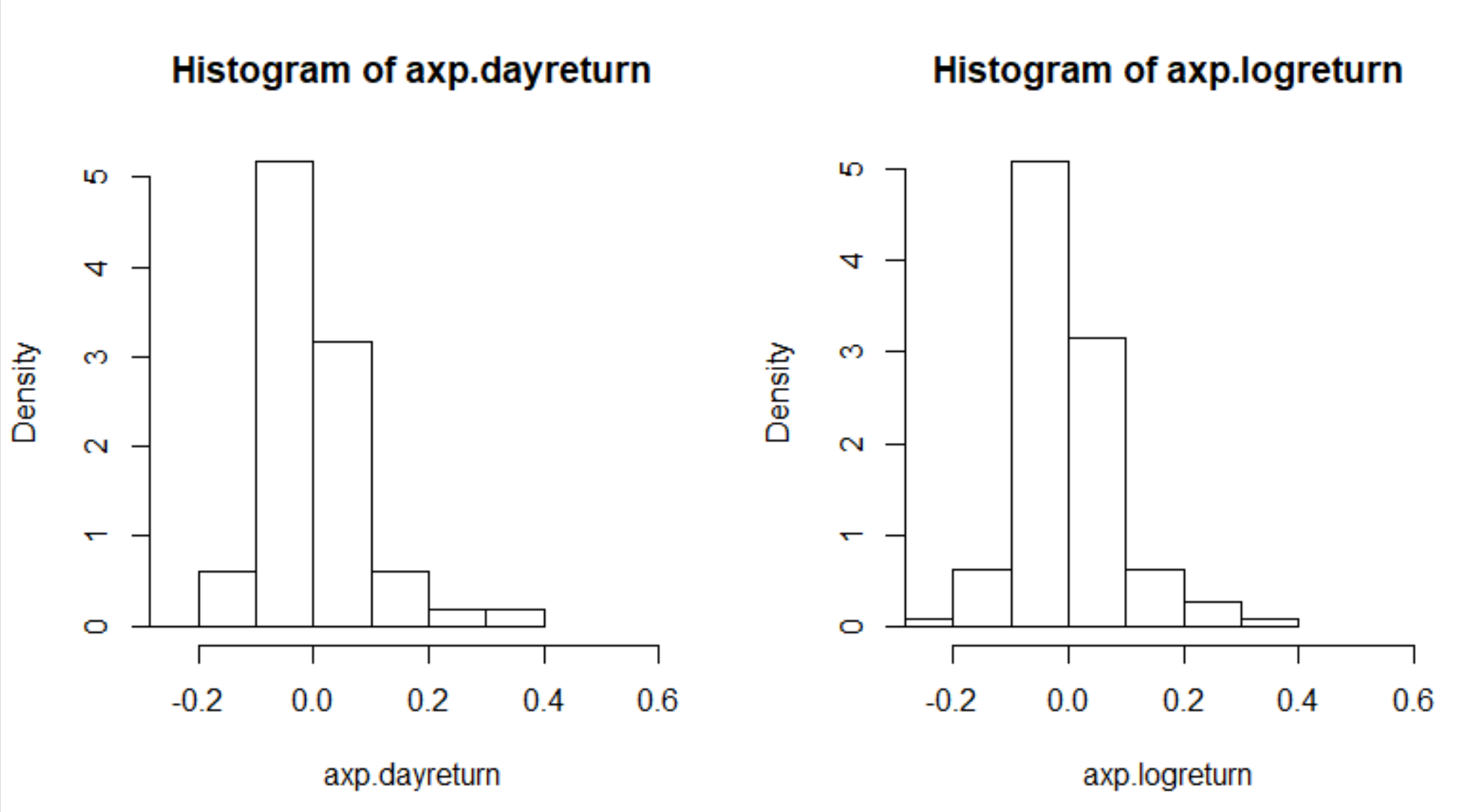
American Express Company Stock Analysis and Forecasting

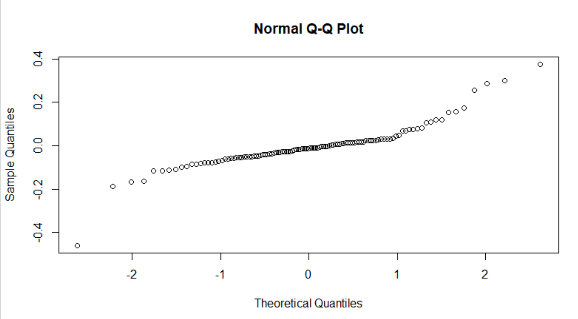
By

Jaswanth Das Gunturu

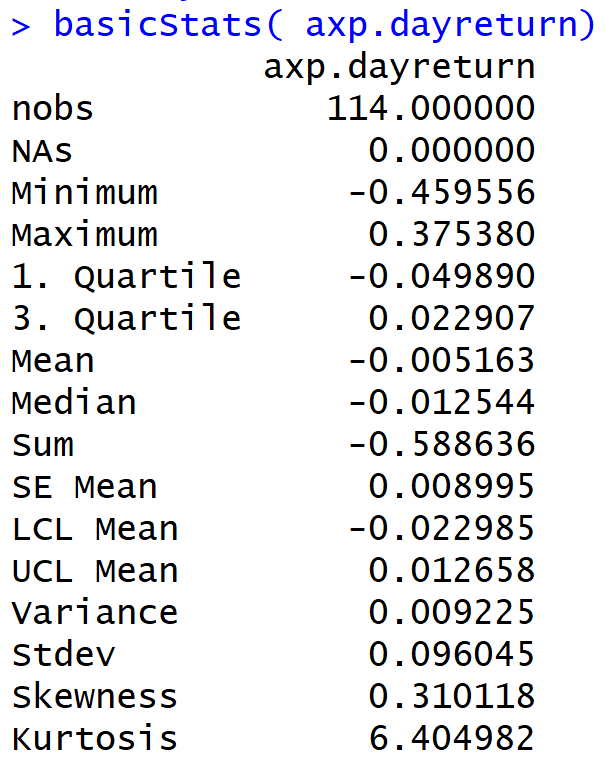
Aditya Gidda

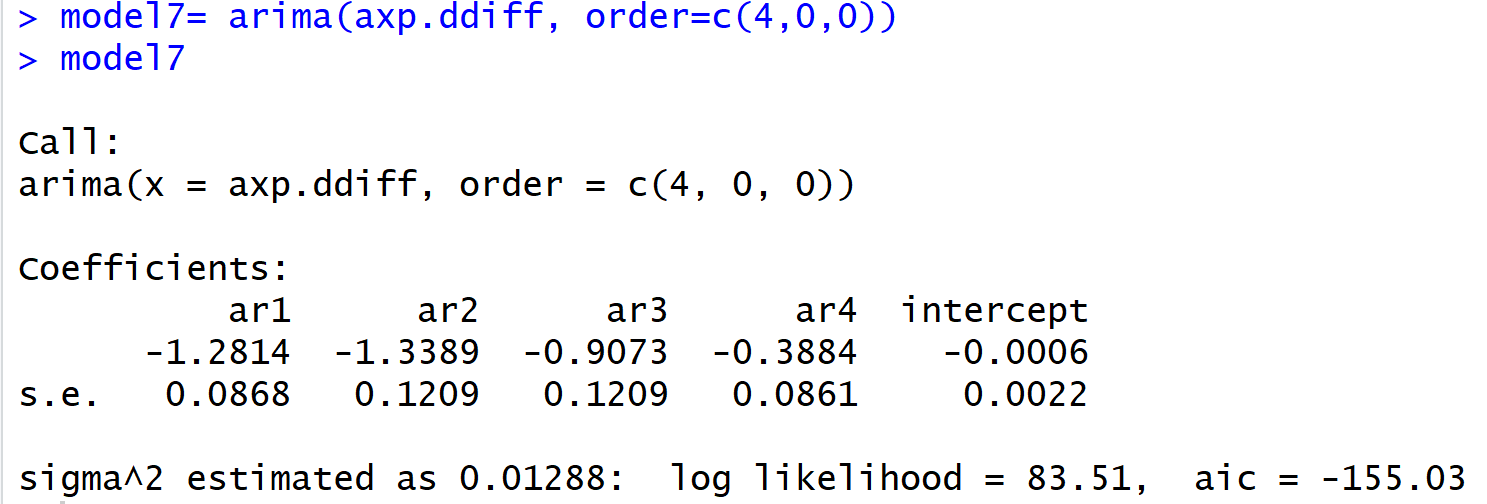
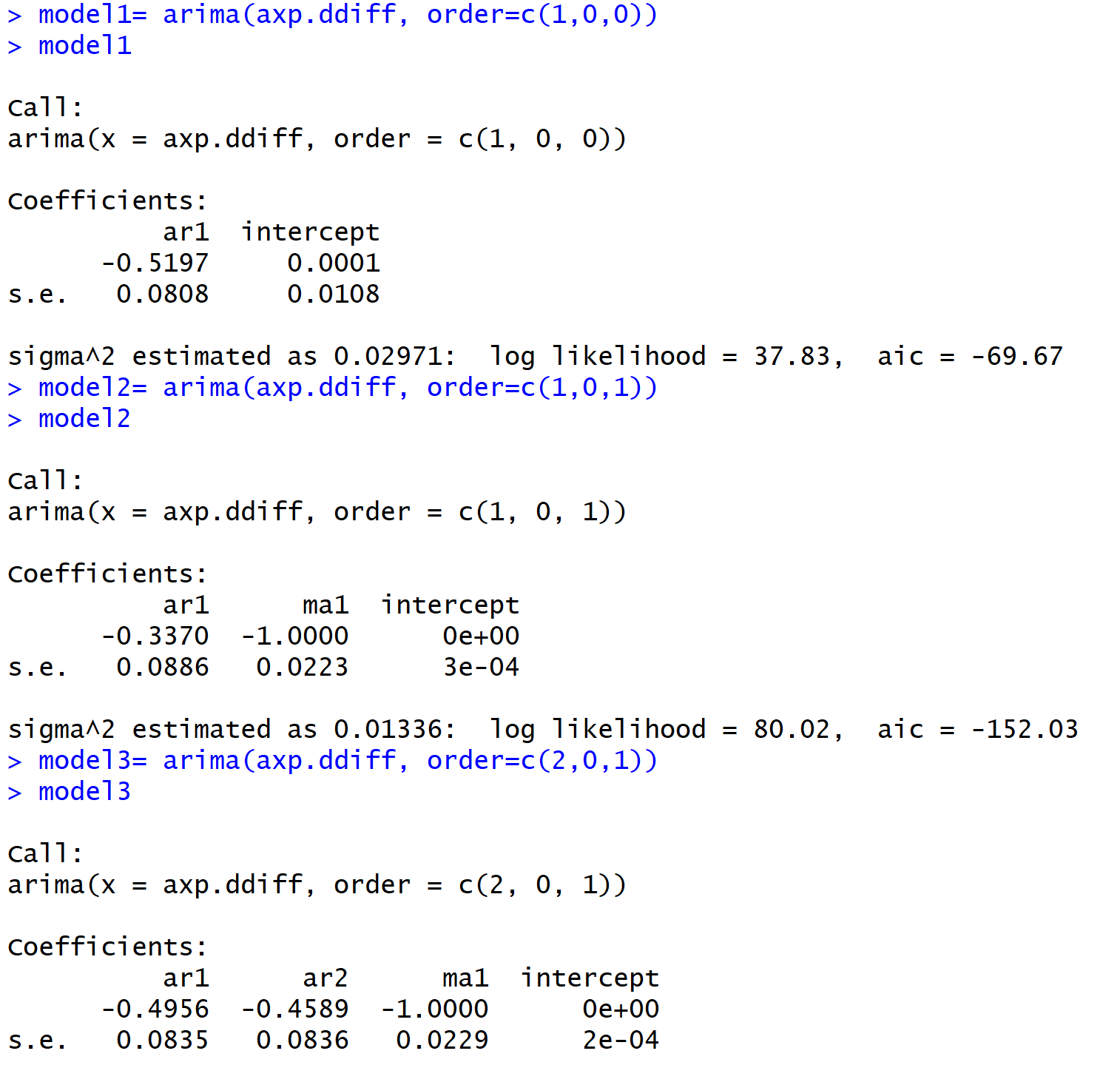
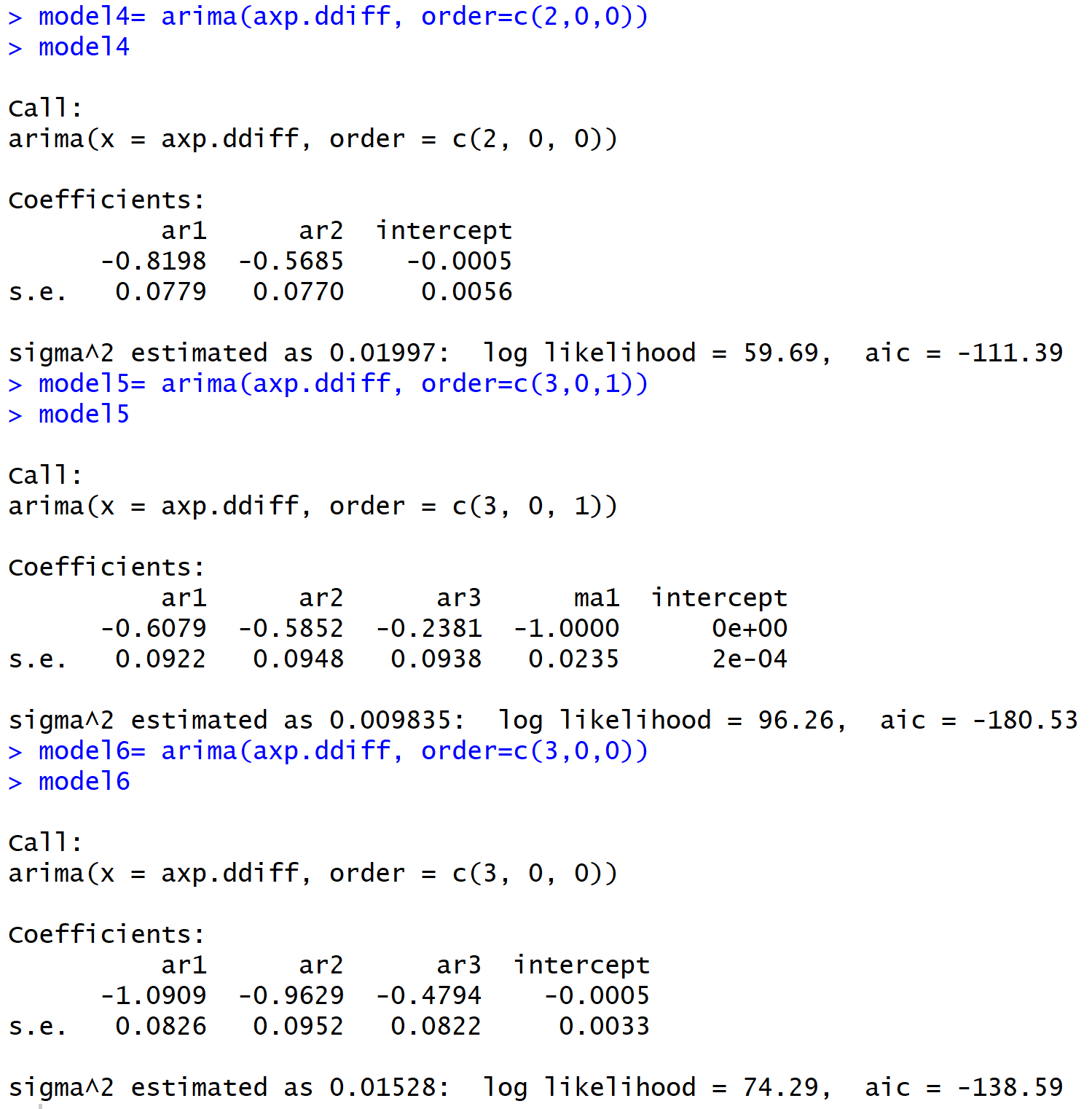
Introduction

The data set contains the stock prices of the company shares in real time for monthly data. The dataset consists of stock opening value, closing value, price and adj.close value. We calculated simple returns and named it as day return and we even did the logreturns too. The TS plots of both logreturn and day return are almost the same except for the peak towards the negative side. And when we plotted the histograms we found that both of them are right skewed. 

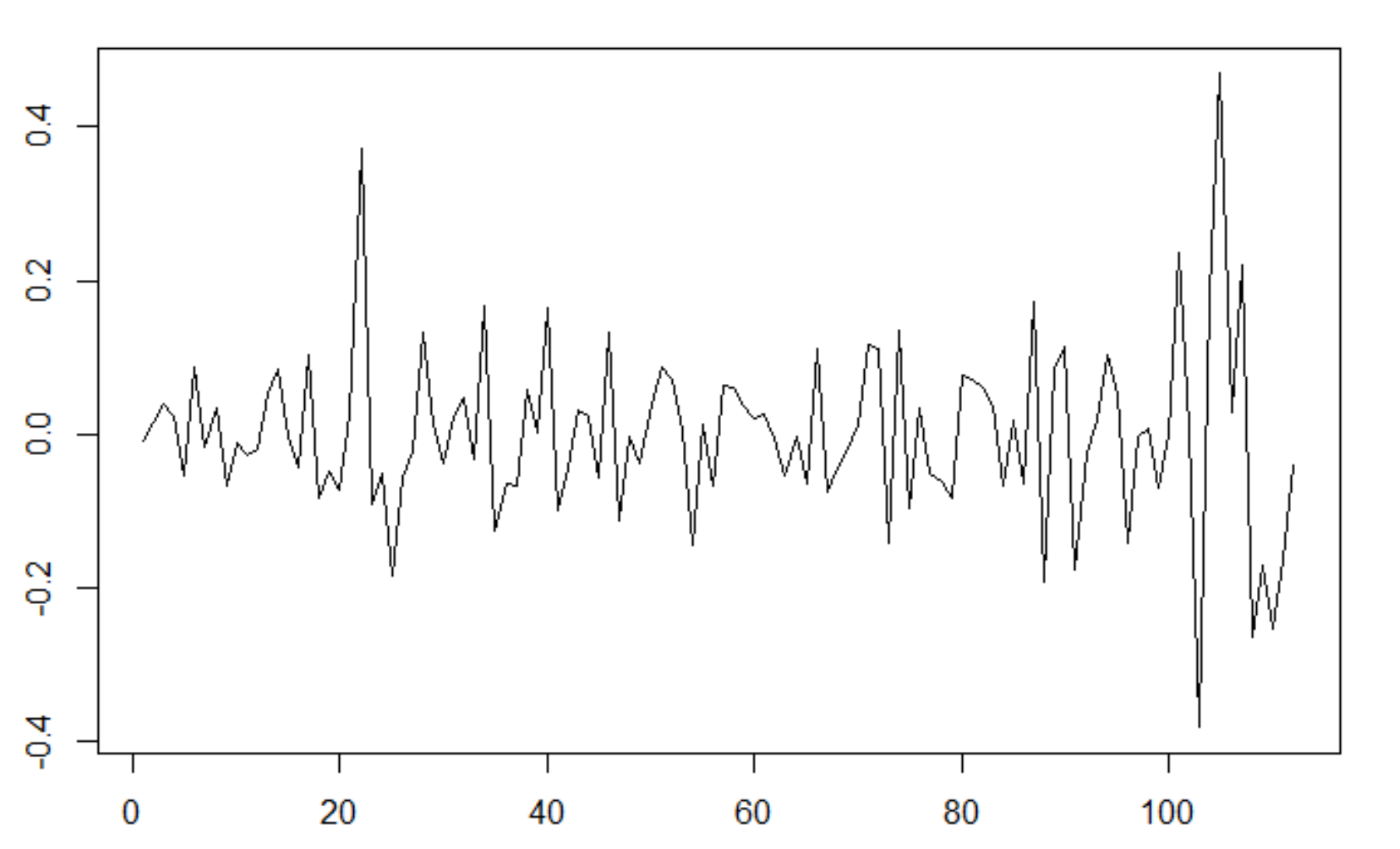
And when looked into the Normal Q-Q Plot we can know that the data is almost normal.

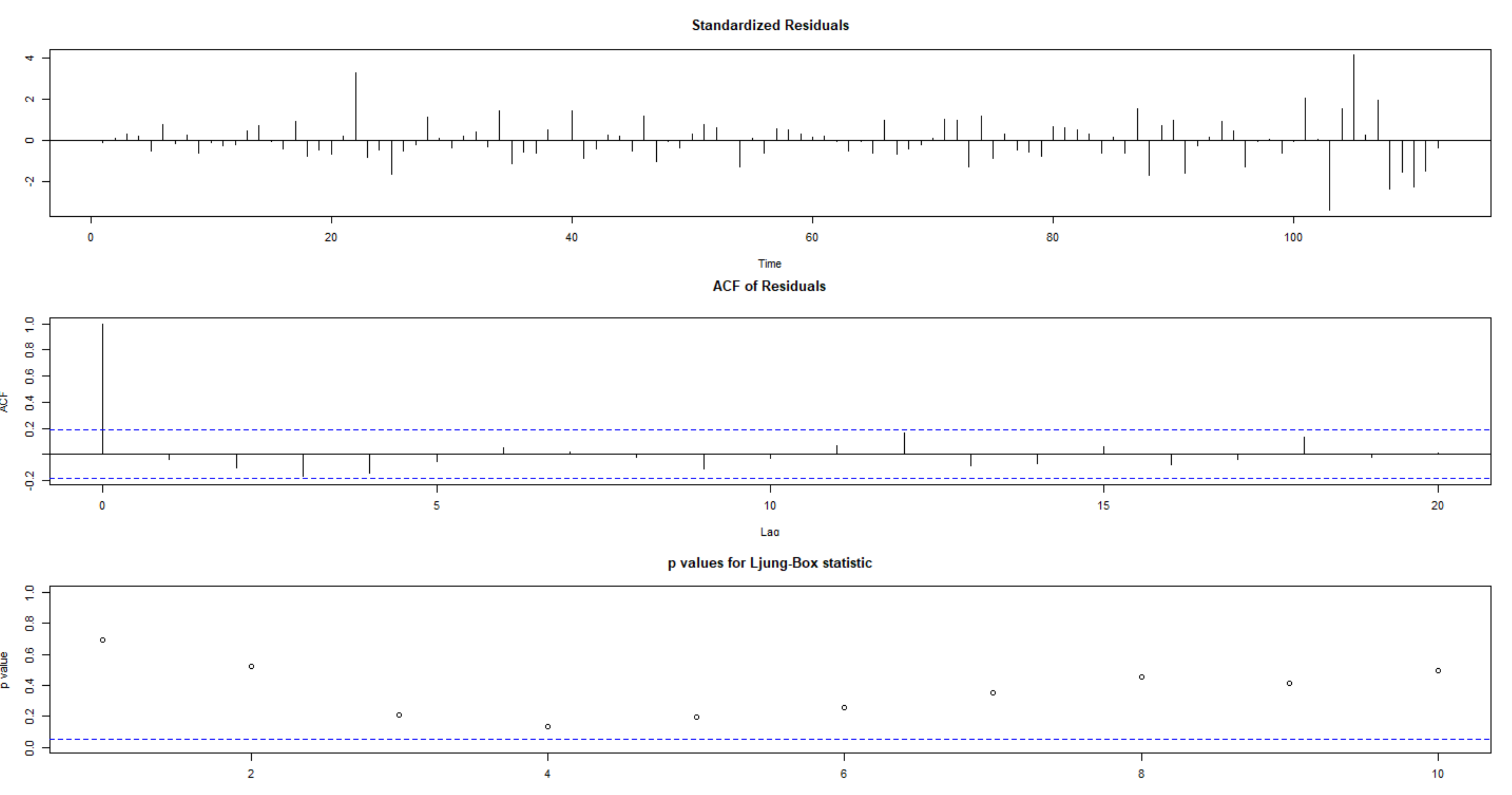
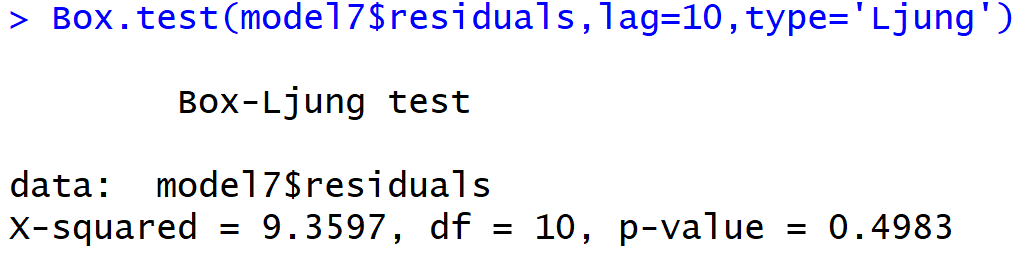
So, as there is not much difference between the dayreturn and the logreturn we take the dayreturn for further analysis. And we obtain the basic stats of the dayreturn. Further analysis we do the t-test and obtain the conclusion that the true mean is not equal to zero.

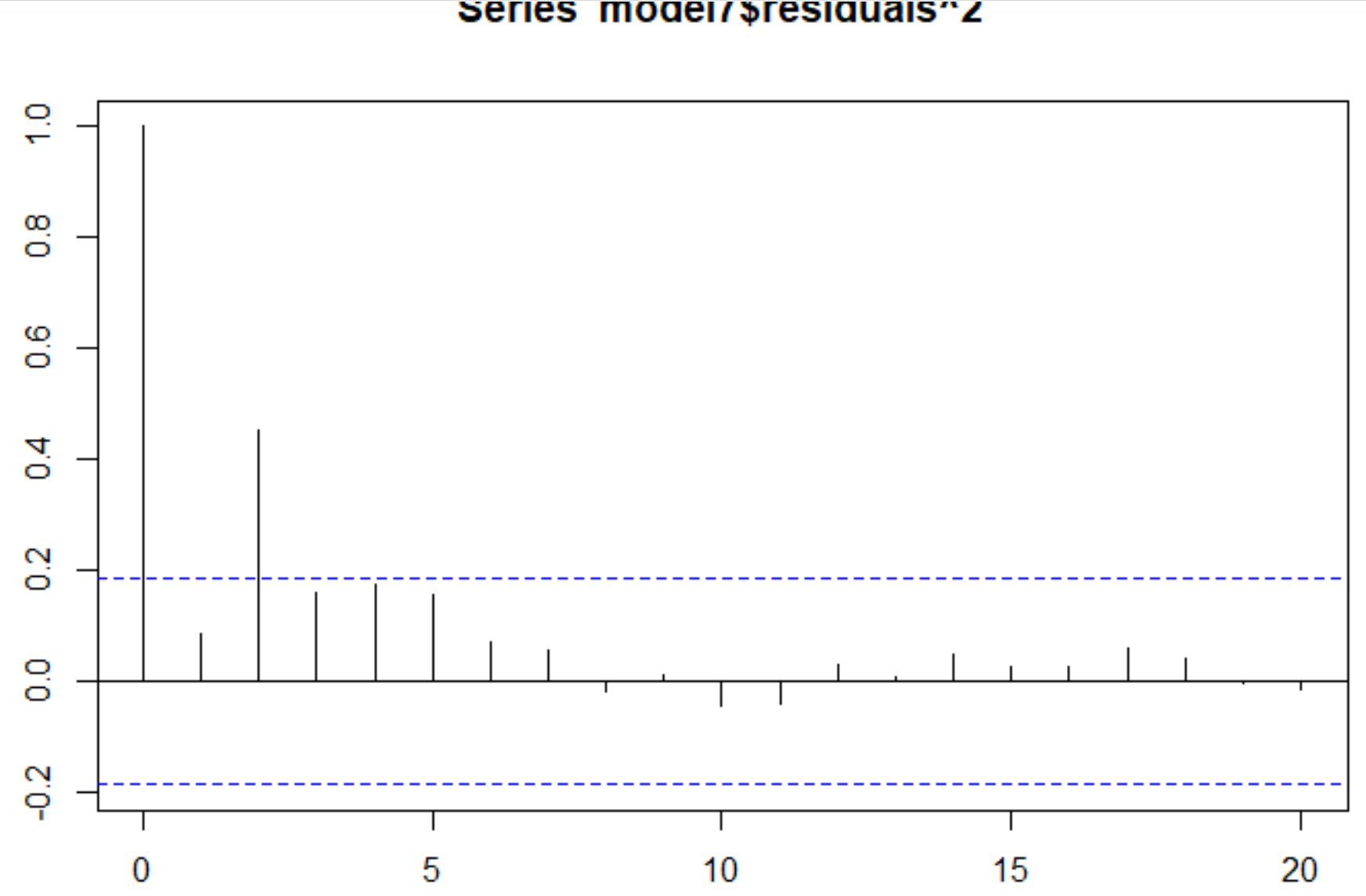
And we do the Jarque-Bera Normality test and we obtain a p value near to zero which states that it is less than 0.05 and we can reject the null hypothesis and it is normal. From the Box-pierce and Box-Ljung test we found that both the p-values are less than 0.05, to reject the null hypothesis.

So, we started with the ARIMA models, we checked the ARIMA for about 7 models and found (4,0,0) as the best fit with the aic value of about -155.03.

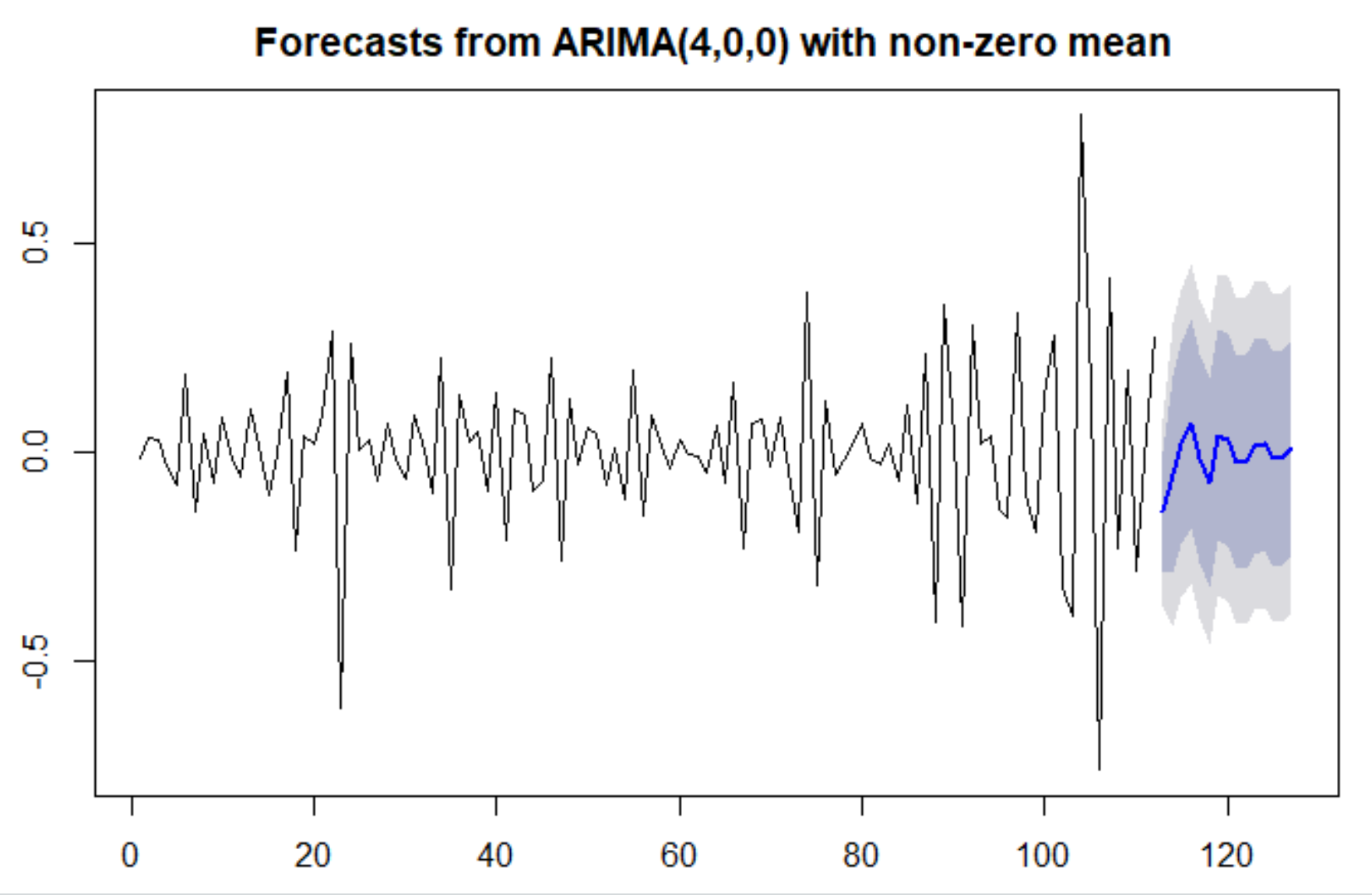
Though there were some models with the aic values better than the model7, we found that the p values are’nt significant and for some other models they were not good because of the z-stat value.

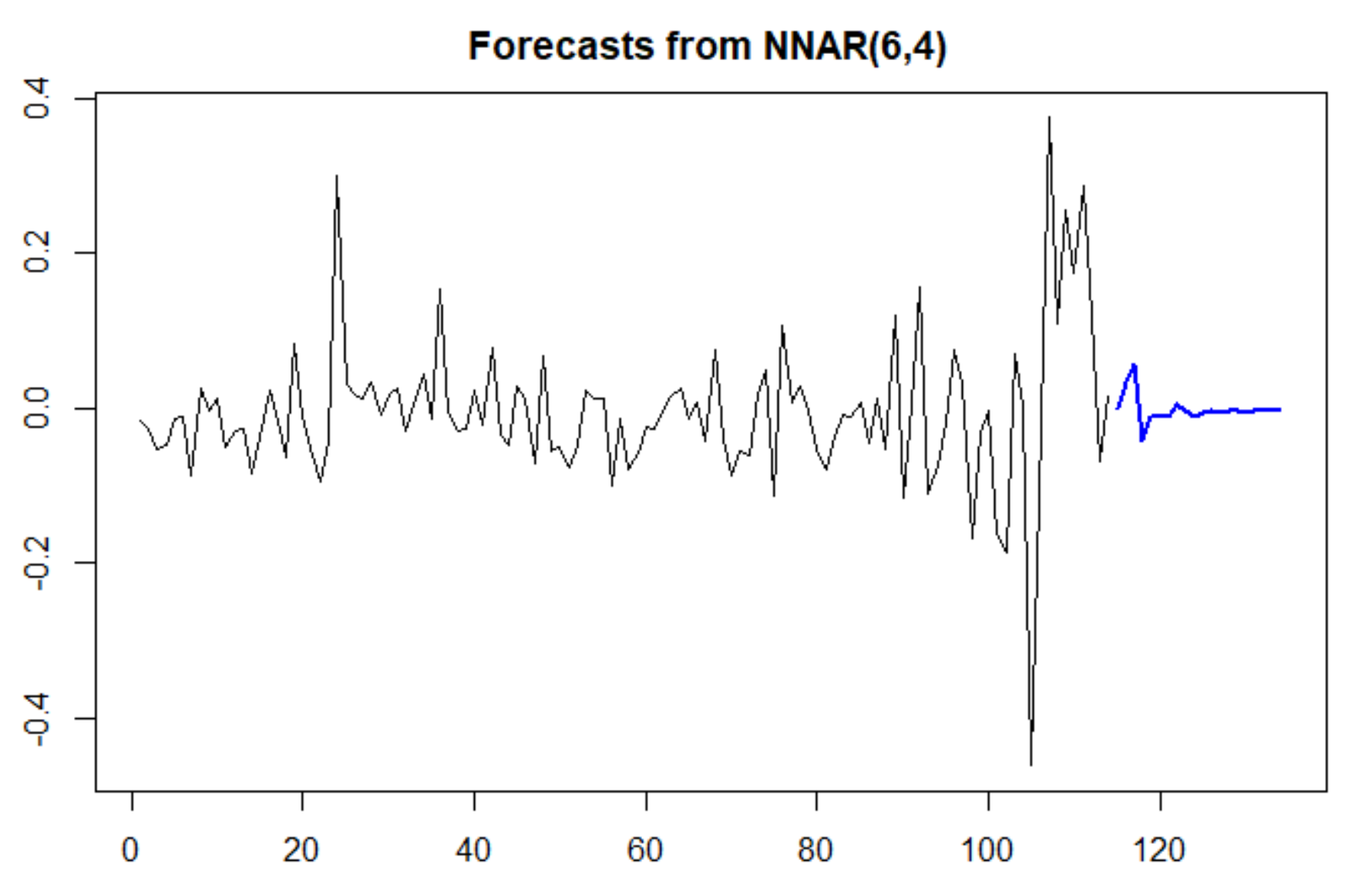
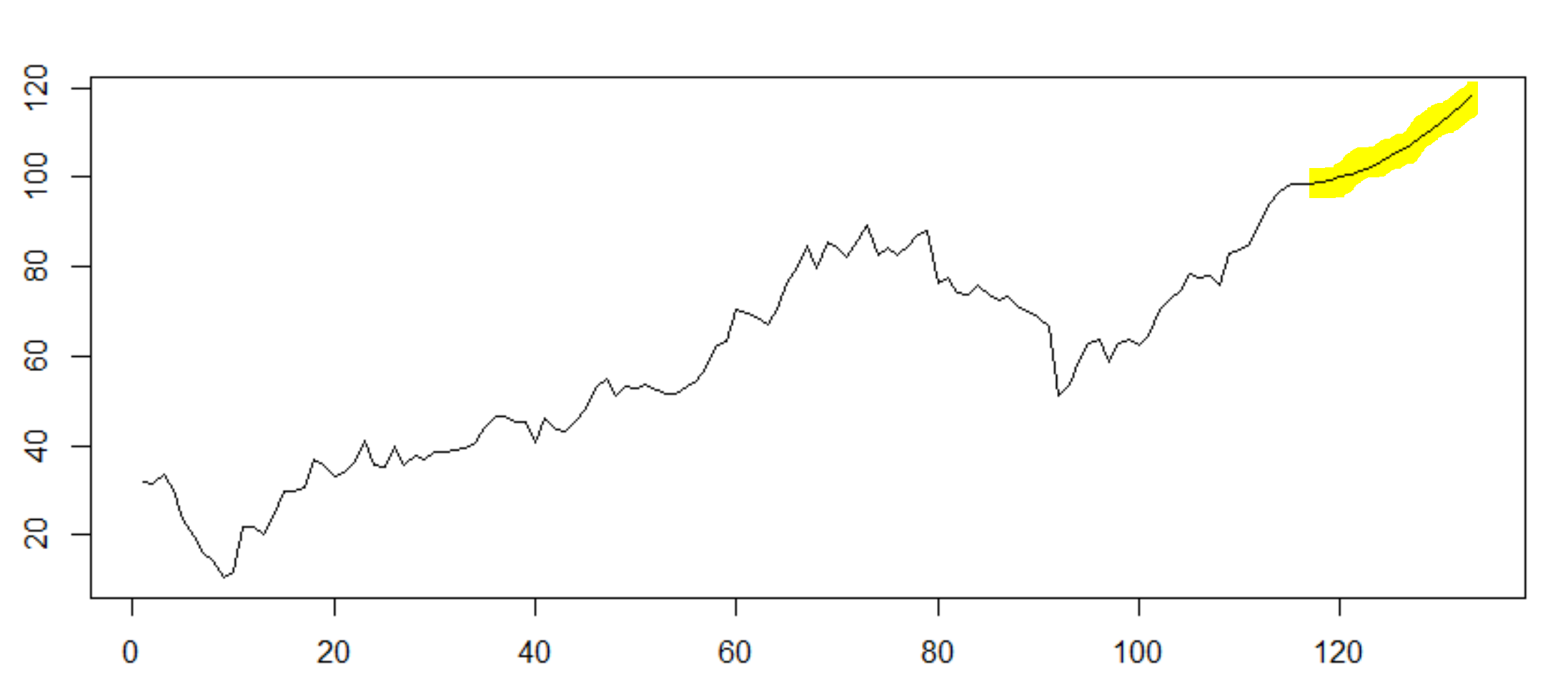
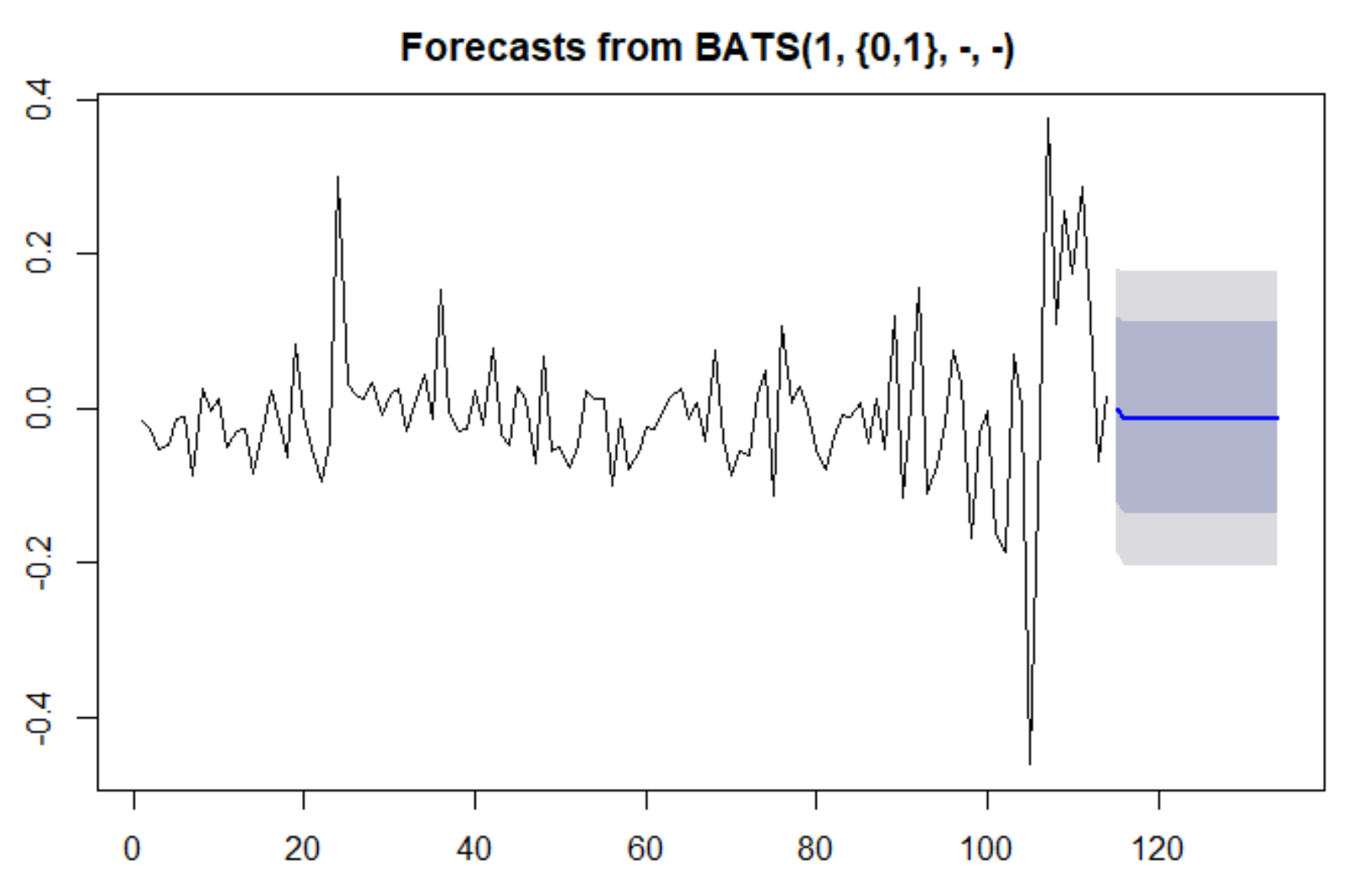
But for the model7 the p values are significant and the acf plot also doesn’t have any autocorrelation as shown.



When the box test is done on the residuals of model7 we found that the p-value is more than 0.05 and the residuals are independent.

When the acf of the residuals^2 is plotted there is a little bit of autocorrelation in the starting lags and there isn’t any autocorrelation after that and this suggests we cannot go for the ARCH and GARCH models here in this situation.

The data is taken only till the May of 2017, so that we can predict the stock values for the next year until May 1, 2018.

And we predicted the stocks using three methods, one was the ARIMA method, BATS method and the other was the neural network(NNAR) method. We found that the best suited model was the first forecast method from ARIMA(4,0,0), since it is the model which goes with the flow of the rest of the residuals and even the predicted values were very near to the real values we obtained from the data source.

The graph below shows the predicted values highlighted which is the best one we’ve obtained from the ARIMA(4,0,0) forecast.