Disclaimer:

The content of this report and the product made is only meant for the learning process as part of the course. This is not for use of making publication or making commercialization without the mentor's consent. My contribution won't demand any claim and is only meant for the future for further progress of Mentor's development and innovation along the direction unless there is a special continuous involvement.

Group 11

Non-Linear Forecasting

Group Members

Aditya Raj (170123004)
Aryan Raj (170123010)
Aayush Bansal (170123001)
Mayank Saharan (170123033)
Sumedh Jouras (170123050)

Supervised By -
Dr. Arabin Dey

Nonlinear time series modeling

What is a nonlinear time series?

A nonlinear process is any stochastic process that is not linear. To this aim, a linear process must be defined. Realizations of time-series processes are called time series but the word is also often applied to the generating processes.

Nonlinear time series are generated by nonlinear dynamic equations. They display features that cannot be modeled by linear processes: time-changing variance, asymmetric cycles, higher-moment structures, thresholds, and breaks.

ARCH and GARCH models

ARCH models were introduced by Robert F. Engle in 1982 to model time-changing volatility (variance) in a time-homogeneous model. The model, at first introduced for monthly inflation, proved to be extremely successful for daily finance data. The model is rarely used outside economics.

Definition of the ARCH model

An *autoregressive conditionally heteroskedastic* (ARCH) model with integer-order p ≥ 1 is defined by

$$X_t = \sigma_t \varepsilon_t \, , \; \sigma_t^2 = c_0 + b_1 X_{t-1}^2 + \ldots + b_p X_{p-1}^2 \, ,$$

with $c_0 > 0$, $b_j \geq 0$, (ε_t) *iid*(0,1), and (ε_t) independent of X_s ,s < t. A process (X_t) obeying these equations is then an ARCH(p) process.

Note. There is a mean and variance equation. The variance equation does not have any additional error and exactly determines the unobserved local variance σ_t^2 .

Stability of the ARCH model

Theorem:
The ARCH model has a strictly and covariance stationary solution iff $\sum_{j=1}^{p} b_j < 1$ *. Then,* $EX_t = 0$ *and*

$$EX_t^2 = \frac{c_0}{1 - \sum_{j=1}^{p} b_j}$$

Note. This condition is not necessary for a strictly stationary solution with infinite variance. For Gaussian ε_t, the sum may even be as large as 3.

Properties of ARCH processes

Proposition:

Assume (X_t) is a strictly stationary ARCH(p) process with $c_0 > 0$ and $\sum_{j=1}^{p} b_j < 1$. Then,

1. *(X_t) is white noise with variance $\dfrac{c_0}{1 - \sum_{j=1}^{p} b_j}$;*
2. *if fourth moments are finite, (X_t^2) is a (causal) AR(p) process with non-negative ACF;*

GARCH processes

In 1986, Tim Bollerslev found that large ARCH orders are needed to fit observed financial time series. He suggested a more parsimonious representation with geometric decay of ARCH coefficients. These GARCH processes, even as GARCH(1,1) were very successful. Steve Taylor made the same discovery but overlooked that the correspondence to the ARMA model is not perfect and that his MACH models do not work.

Definition of the GARCH model

Definition:
A generalized autoregressive conditional heteroskedastic (GARCH) model with orders $p \geq 1$ and $q \geq 0$ is defined by

$$X_t = \sigma_t \varepsilon_t \,, \quad \sigma_t^2 = c_0 + \sum_{j=1}^{p} b_j X_{t-1}^2 + \sum_{j=1}^{q} a_j \sigma_{t-j}^2 \,,$$

with $c_0 > 0$, $b_j \geq 0$, $a_j \geq 0$, (ε_t) *iid*(0,1), and (ε_t) independent of X_s ,s < t. A thus defined stochastic process is called a GARCH(p, q) process.

Remark. GARCH(0, q) for q > 0 does not work, as it would define a non-stochastic linear difference equation for σ_t^2 . Some coefficients might be permitted to be negative, but conditions are restrictive for this case.

Stability of the GARCH model

Theorem:
The GARCH(p, q) model has a unique strictly and covariance stationary solution iff

$$\sum_{j=1}^{p} b_j + \sum_{j=1}^{q} a_j < 1$$

Then, $EX_t = 0$, (X_t) is white noise, and

$$\mathit{var}\, X_t = \frac{c_0}{1 - \sum_{j=1}^{p} b_j - \sum_{j=1}^{q} a_j}$$

$EX_t^4 < \infty$ if $E\varepsilon_t^4 < \infty$ and

$$\sqrt{E\varepsilon_t^4}\,\frac{\sum_{j=1}^{p} b_j}{1 - \sum_{j=1}^{q} a_j} < 1$$

<u>The ARMA representation of the GARCH model</u>

Proposition:

If (X_t) is a strictly and covariance stationary GARCH(p, q) process with finite fourth moments, (X_t^2) will be a causal and invertible ARMA(max(p, q), q) process. Its kurtosis exceeds the kurtosis of ε_t.

In detail,

$$X_t^2 = c_0 + \sum_{j=1}^{max(p,q)} (b_j + a_j) X_{t-1}^2 + e_t - \sum_{j=1}^{q} a_j e_{t-j}$$

For $e_t = X_t^2 - \sigma_t^2$.

Estimation of GARCH models

1. Least squares (for ARCH only)

Engle was interested in the regression with ARCH errors

$$X_t^2 = \mu + \sum_{j=1}^{P} \phi_j X_{t-1}^2 + u_t ,$$

$$u_t = \sigma_t \varepsilon_t \ , \ \sigma_t^2 = c_0 + b_1 u_{t-1}^2 + \ldots \ + b_p u_{t-p}^2 ,$$

He considered estimating the mean equation by OLS (consistent but inefficient), then the variance equation for squared OLS residuals by OLS (inefficient), then the mean equation by weighted LS, etc. The procedure is unreliable and mainly used for obtaining starting values for ML–based estimation

2. Conditional ML for ARCH models

It is relatively straightforward to show that the log-likelihood $\ell(X_1, \ . \ . \ . \ , X_t \mid c_0, \ b_1, \ . \ . \ . \ , \ b_p)$ can be represented as

$$\ell(X_{p+1}, \ . \ . \ . \ ,X_T \ |X_1, \ . \ . \ . \ ,X_t \ ,c_0, \ b_1, \ . \ . \ . \ , \ b_p)$$

$$\alpha \ - \ \sum_{t=p+1}^{T} (log\sigma_t^2 \ + \ \frac{X_t^2}{\sigma_t^2}), \ \ \sigma_t^2 = c_0 + \sum_{j=1}^{p} b_j X_{t-1}^2$$

Most estimation algorithms maximize this likelihood by brute force, possibly restricting the area of admissible parameter values (non-negativity, stability constraints).

ARIMA-GARCH model

The hybrid ARIMA-GARCH model is a non linear time series model which combines a linear ARIMA model with the conditional variance of a GARCH model. The estimation procedure of ARIMA and GARCH models are based on maximum likelihood method.

The logarithmic likelihood function has the following equation:

$$lnL[(y_t), \theta] = \sum_{t=1}^{T} \{ln[D(z_t(\theta)), \nu] - \tfrac{1}{2} ln[\sigma_t^2(\theta)]$$

Where θ is the vector of the parameters that have to be estimated for the conditional mean, conditional variance and density function, z_t denoting their density function, $D(z_t(\theta), \nu)$, is the log-likelihood function of $[(y_t), \theta]$, for a sample of T observation. The maximum likelihood estimator $\widehat{\theta}$ for the true parameter vector is found by maximizing the equation above.

Diagnostic Checking of ARIMA-GARCH Model

The diagnostic tests of ARIMA-GARCH models are based on residuals. Residuals' normality test is employed with Jarque and Bera (1980) test. Ljung and Box (1978) (Q-statistics) statistic for all time lags of autocorrelation is used for the serial correlation test. Also, for the conditional heteroscedasticity test, we use the squared residuals of the autocorrelation function.

Forecast Evaluation

On ARIMA-GARCH models we use both the static and dynamic forecast. The dynamic forecast, also known as n-step ahead forecast, uses the actual lagged value of Y variable in order to compute the first forecasted value. The static forecast (one-step-ahead forecast) of Y_{t+1} based on ARIMA-GARCH model is defined as:

$$\widehat{Y}_t(1) = E(Y_{t+1} | Y_t, Y_{t+1}, \ldots) = \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j}$$

Where the ε_s follow the stated GARCH model.

To evaluate the forecast efficiency, we use two statistical measures, mean squared error (MSE) and mean absolute error (MAE). MSE it computes the squared difference between every forecasted value and every realised value of the quantity being estimated, and finds the mean of them afterwards.

MAE it computes the mean of all the absolute, instead of squared, forecast errors.

Dickey-Fuller Test

A Dickey-Fuller test is a unit root test that tests the null hypothesis that $\alpha = 1$ in the following model equation. alpha is the coefficient of the first lag on Y.

Null Hypothesis (H0): alpha=1

$$y_t = c + \beta t + \alpha y_{t-1} + \phi \Delta Y_{t-1} + e_t$$

where,

- y_{t-1} = lag 1 of time series
- ΔY_{t-1} = first difference of the series at time (t-1)

Fundamentally, it has a similar null hypothesis as the unit root test. That is, the coefficient of Y_{t-1} is 1, implying the presence of a unit root. If not rejected, the series is taken to be non-stationary.

Augmented Dickey-Fuller (ADF) Test

As the name suggests, the ADF test is an 'augmented' version of the Dickey-Fuller test.

The ADF test expands the Dickey-Fuller test equation to include high order regressive process in the model.

$$y_t = c + \beta t + \alpha y_{t-1} + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + ... + \phi_p \Delta Y_{t-p} + e_t$$

If you notice, we have only added more differencing terms, while the rest of the equation remains the same. This adds more thoroughness to the test.

The null hypothesis however is still the same as the Dickey-Fuller test.

A key point to remember here is: Since the null hypothesis assumes the presence of unit root, that is $\alpha=1$, the p-value obtained should be less than the significance level (say 0.05) in order to reject the null hypothesis. Thereby, inferring that the series is stationary.

However, this is a very common mistake analysts commit with this test. That is, if the p-value is less than the significance level, people mistakenly take the series to be non-stationary.

Proposed Shiny App (our work):

The Shiny App we have developed makes use of stock prices of top 25 stocks from Yahoo Finance. We use the *GetSymbols()* method in R to crawl these prices. The dashboard also lets the user select the start date and end date to select the window for which stock prices are to be considered and imports the data from Yahoo Finance.
We have used the proposed ARIMA+GARCH method to do the forecasting of the stock prices.

We have also included a slider that lets the user choose the duration(number of days to forecast) .The slider allows us to choose days between 1 and 365 and default value of the slider is set at 15.

There is also a drop down menu where we can choose to forecast using any of these :- log returns, absolute returns and closing price. By default this value is set to log returns.

Further, Prediction Interval values can be selected from a drop down menu with values '99', '95', '90', '80', '70','60', and the default value of this drop down menu is set at 95. Prediction intervals basically tells us where we can expect to see the next data point sampled i.e. a **prediction interval** is an estimate of a value (or rather, the range of likely values) that isn't yet known but is going to be observed at some point in the future.

Similarly, from the drop down menus, we can select values of AR(p), MA(q) and G(p). We can also see that for which of these values our forecast is best fitted.

Fitting of the ARIMA-GARCH model is done in multiple steps. First , we use the ARIMA model ,use the residuals from the latter and plug-them as data for your GARCH model.

Calculations are done on 3 various types of data transformation -

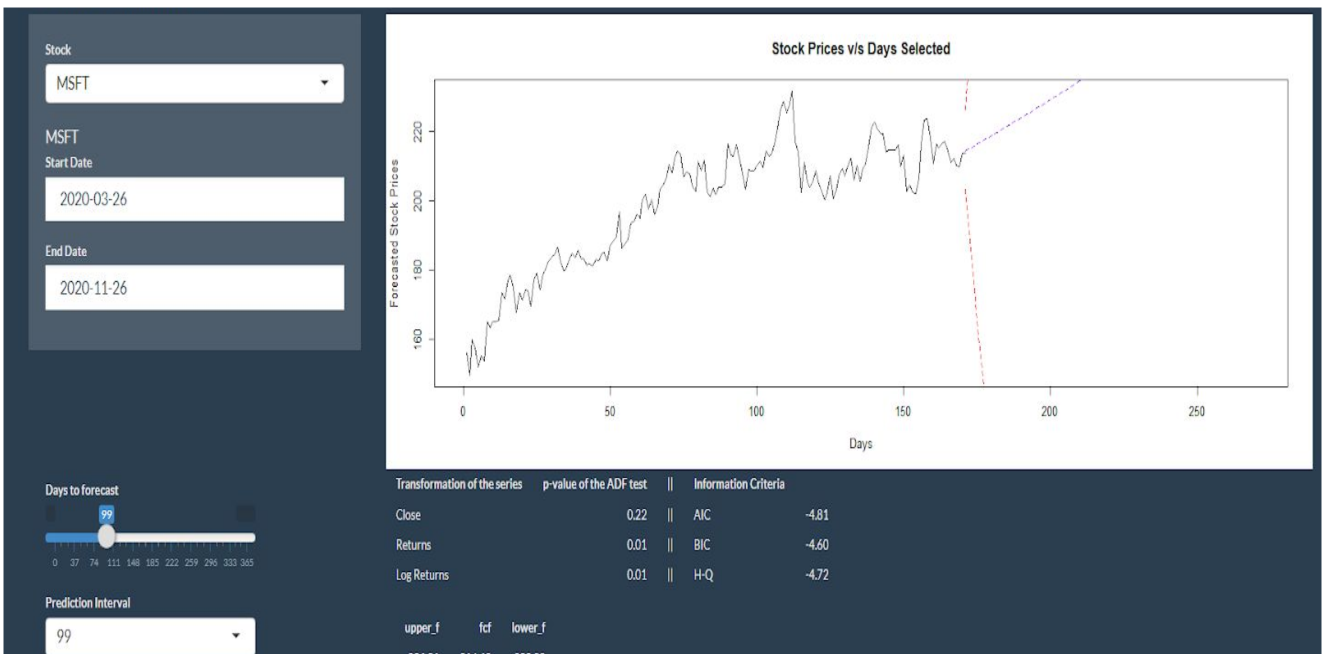
1.) closed prices
2.) returns
3.) log returns

ADF(Augmented Dickey-Fuller) test is performed .(positive values are printed out) .

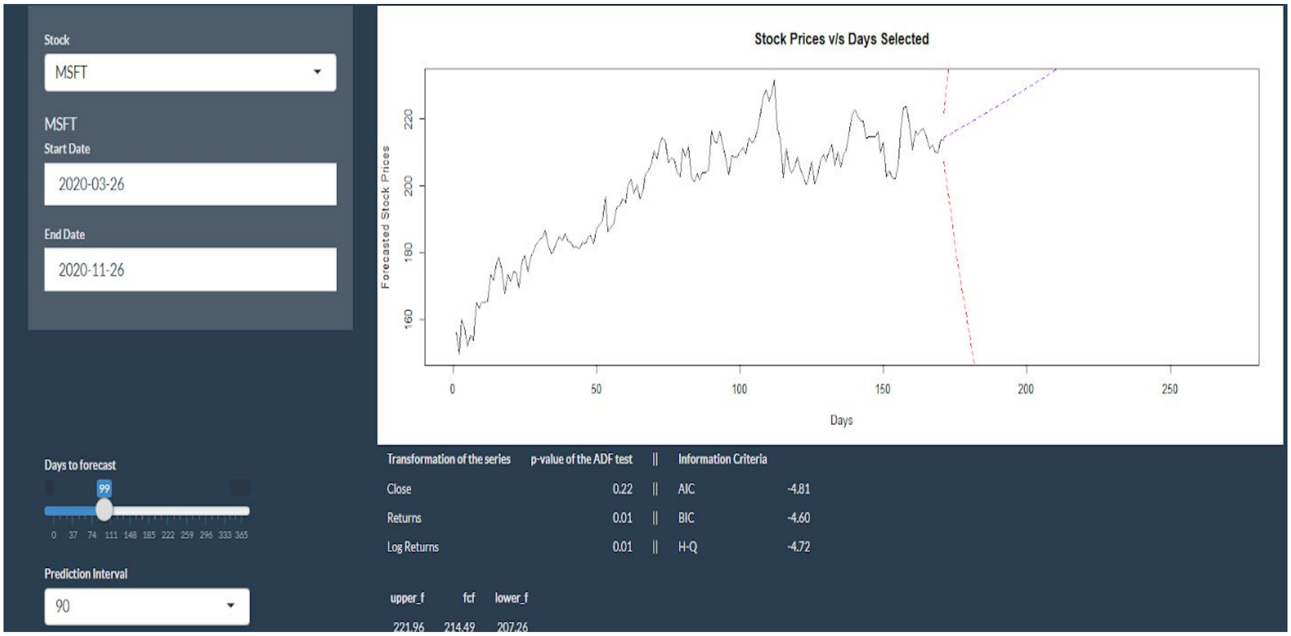
Results -

a.) differences in forecasting on changing the Prediction Interval (we keep AR(p) = 1, MA(q) = 1, G(p) = 1 , and stock considered is ‘MSFT’) -

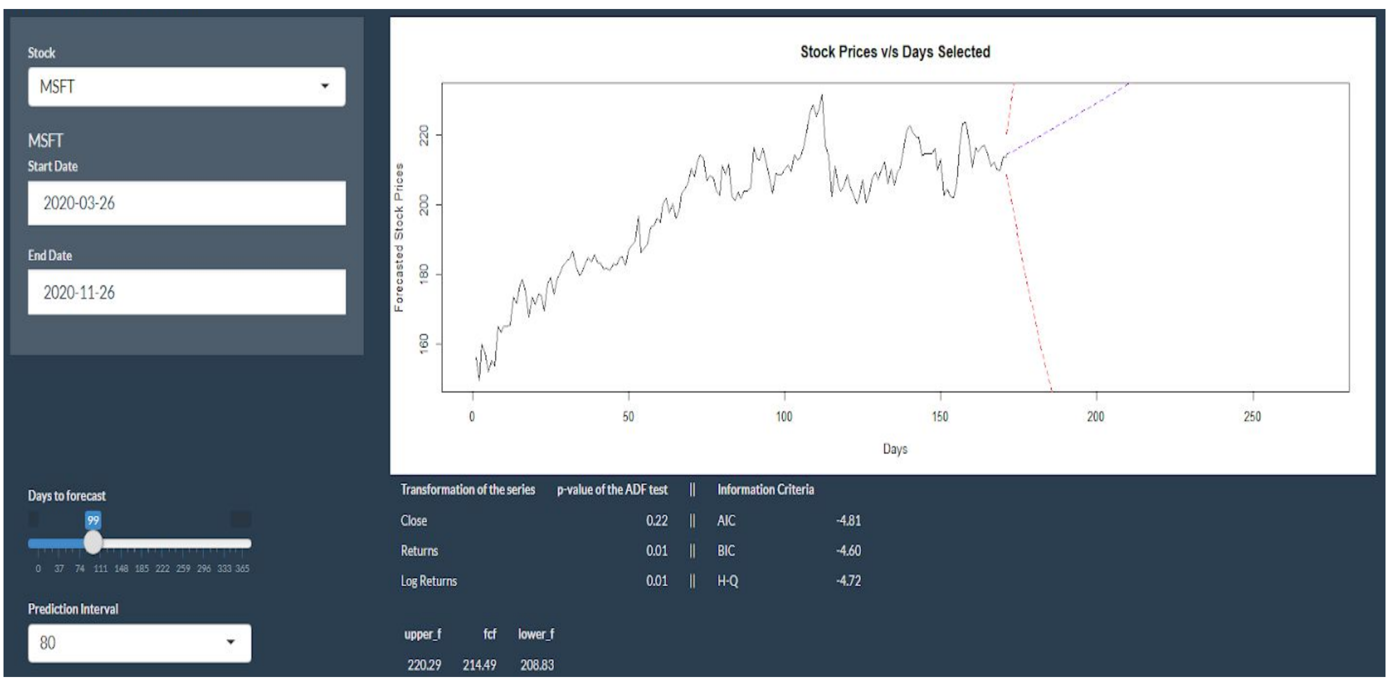
1.) 99 % Prediction Interval -



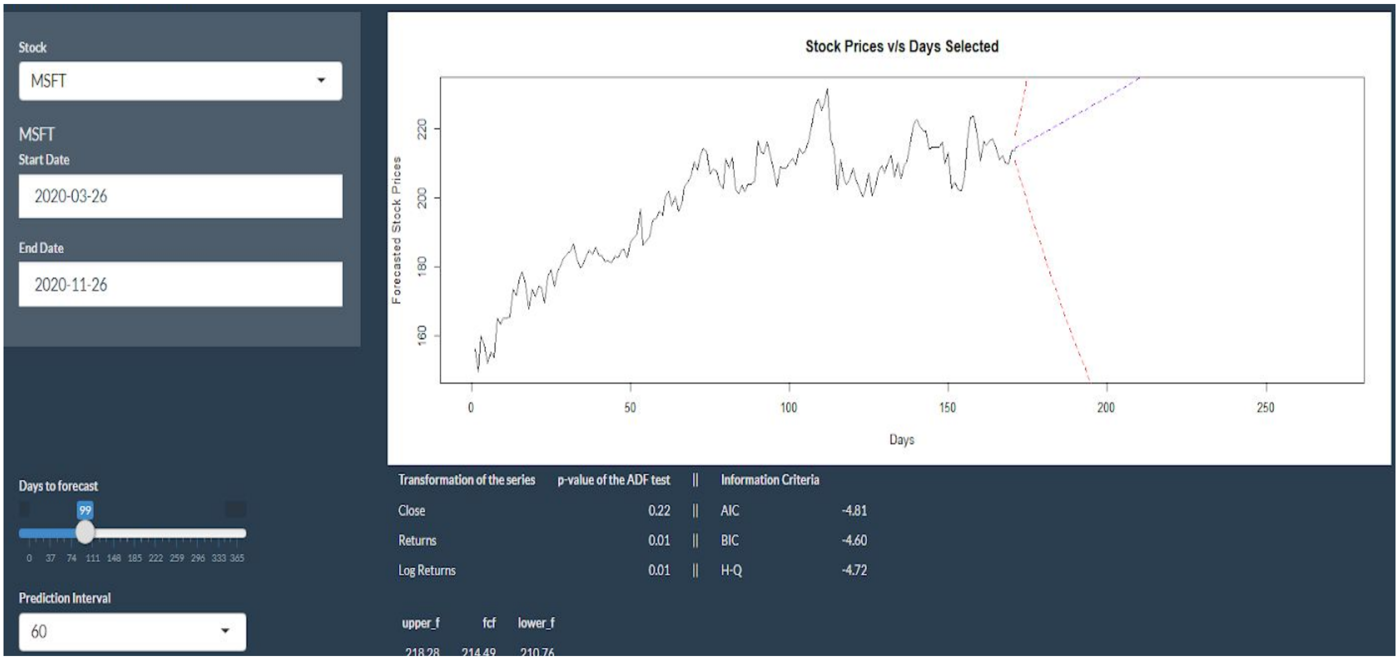
2) 90 % Prediction Interval -



3) 80 % Prediction Interval -



4) 60 % Prediction Interval -



Now, for 95% confidence level and for data from 2018-11-26 to 2020-11-26 for stock Amazon (AMZN) and no. of days to forecast =101, plots for different values of p(i.e. AR) , q(i.e. MA) and g(i.e. G) :-

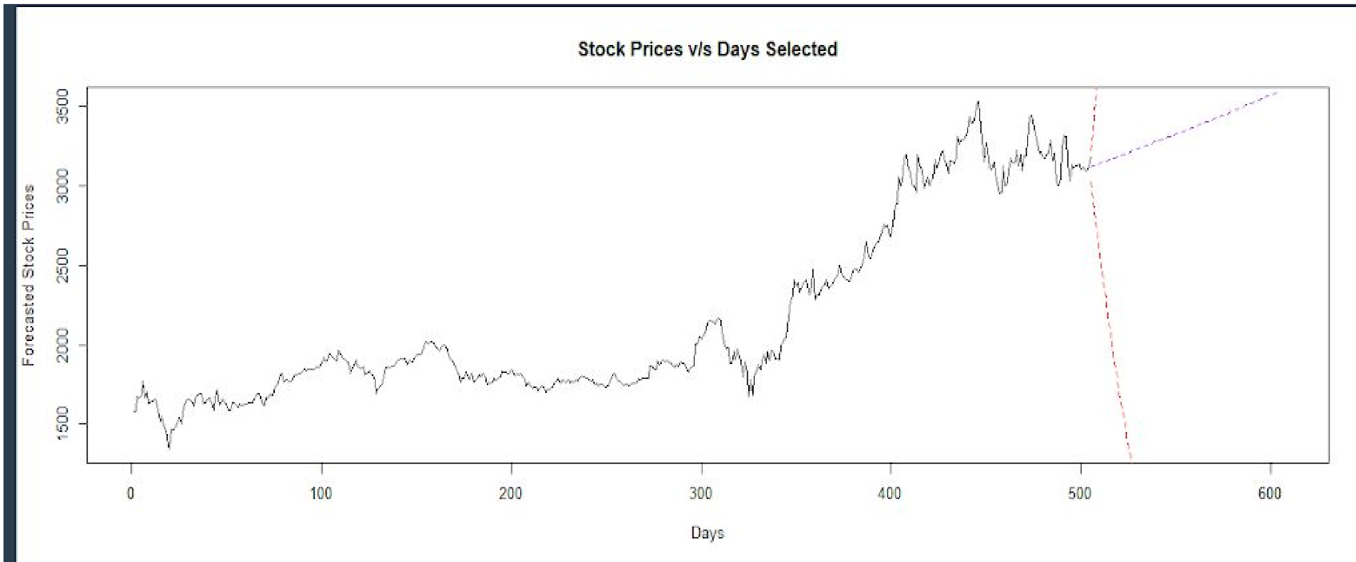
1. p = 1, q=1, g=1:

Transformation of the series	p-value of the ADF test	||	Information Criteria	
Close	0.64	||	AIC	-5.15
Returns	0.01	||	BIC	-5.04
Log Returns	0.01	||	H-Q	-5.11

2. p=2,q=1,g=1:

Transformation of the series	p-value of the ADF test	||	Information Criteria	
Close	0.64	||	AIC	-5.11
Returns	0.01	||	BIC	-5.00
Log Returns	0.01	||	H-Q	-5.07

3. p=2,q=2,g=1:



Transformation of the series	p-value of the ADF test	||	Information Criteria	
Close	0.64	||	AIC	-5.12
Returns	0.01	||	BIC	-5.00
Log Returns	0.01	||	H-Q	-5.07