

Stock forecast for banks



Table of Contents

[Stock price forecast of bank 3](#_Toc23197173)

[Stock market index 4](#_Toc23197174)

[Literature Review 4](#_Toc23197175)

[Time series forecast for Banks 4](#_Toc23197176)

[Data Manipulation 5](#_Toc23197177)

[Descriptive analysis 5](#_Toc23197178)

[Descriptive statistics 7](#_Toc23197179)

[T-test b/w open and close stock 7](#_Toc23197180)

[Exponential smoothing forecasts Model: 14](#_Toc23197181)

[Holt's method 17](#_Toc23197182)

[Holt winter’s Model for Trend & Seasonality 20](#_Toc23197183)

[Automating model building using ETS () 23](#_Toc23197184)

[Forecast by ETS 24](#_Toc23197185)

[ARIMA Model 25](#_Toc23197186)

[Validating of Arima model 27](#_Toc23197188)

[Model accuracy 32](#_Toc23197191)

[Garch model 34](#_Toc23197192)

[GARCH Forecasting 37](#_Toc23197193)

[Conclusion 38](#_Toc23197194)

[References: 38](#_Toc23197195)

Introduction:

Here we aimed to forecast the stock price of all banks by using time series analysis on given data set. In given data set 5711 observations and 12 variables has given to us. We have provided 4 companies stock prices. In stock price an open price, high price, low price and price close, volume prices chart (also OHLC) is type of stock value that typically used for illustrate movements in the price of financial tool over time periods. By using different time series model we forecast the stock price for all banks. Before start the time series analysis, we have to know about stock. So in below we can see the stock trends behaviors:

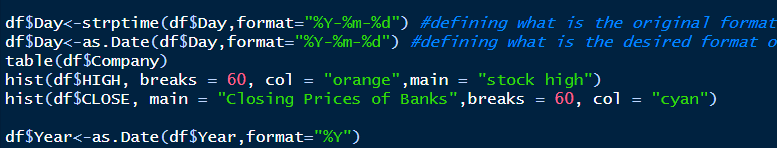
In stock:

* Open value- price when stock opened day
* High value- High price of the stock on day
* Low value- Low price of the stock on day
* Close value- The price when stock closed and last price of stock.
* Volume: Stock price

# Time series forecast for Banks

We have installed the required library to perform time series analysis. We have converted Year and day columns in Data format. Then we have forecast the close price of stock for banks by using different time series model like Arima, garch, ets etc.

## Data Manipulation: We have change the data type of year and day into date format



## Descriptive analysis

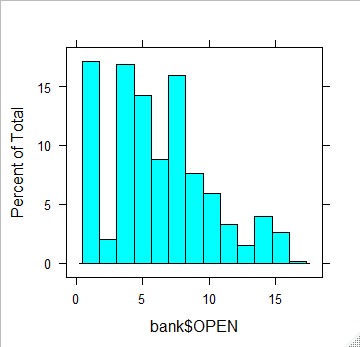


Fig: Open price distribution

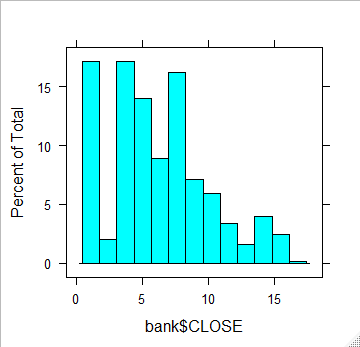
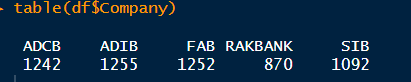


Fig: Close price of stock distribution

All banks details in the data:



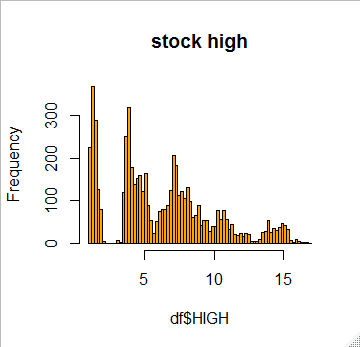


Fig: High Price distribution

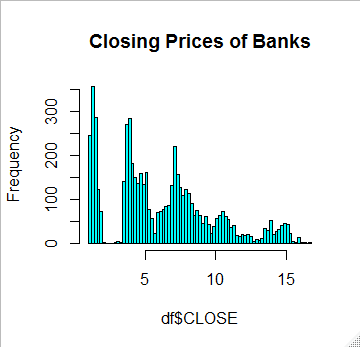
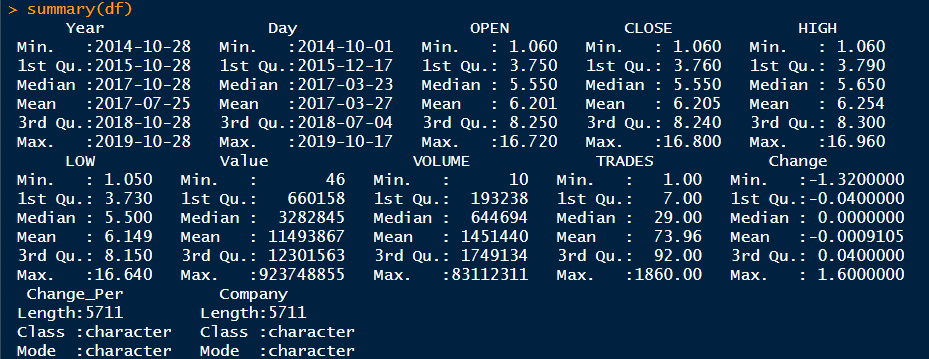
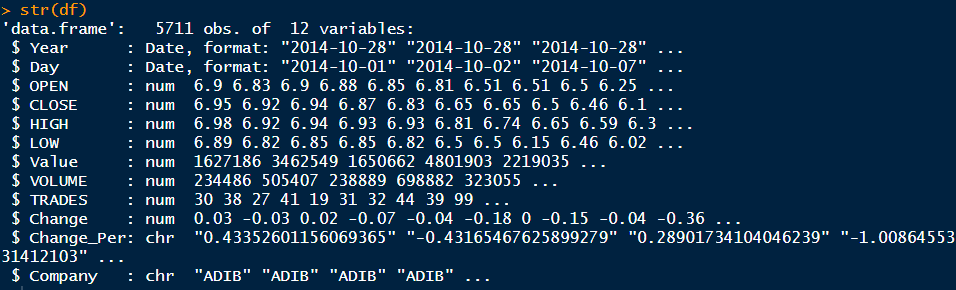


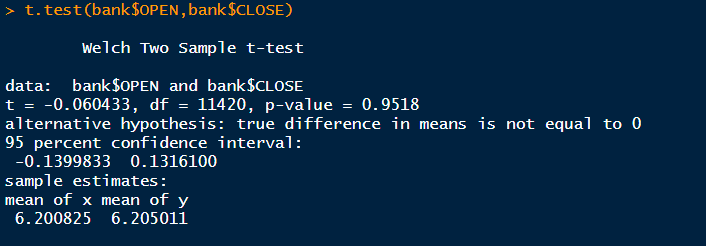
Fig: Closing price of banks stock

## Descriptive statistics





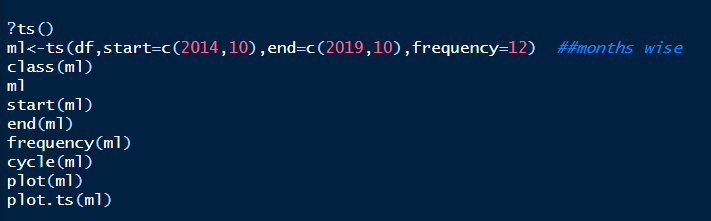
## T-test b/w open and close stock

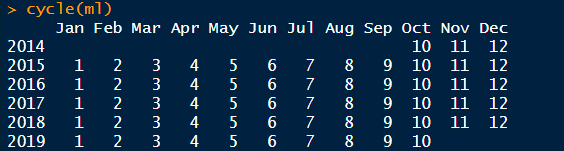


Then we have Converted the data frame into a time series object, by using ts() function.

ts() function belongs to forecast package:

Converted in time series object





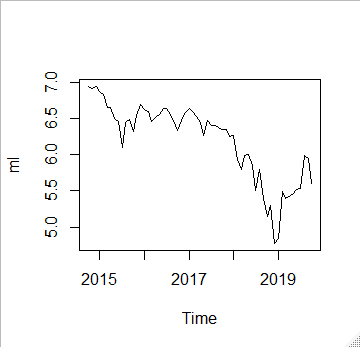
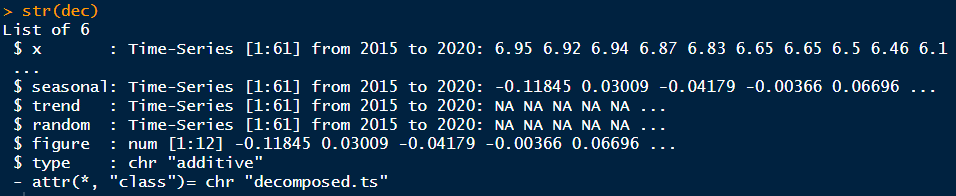


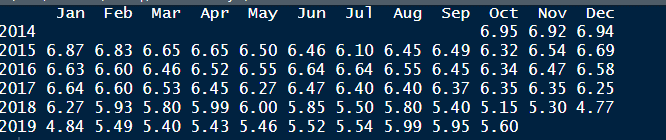
Fig: By ts function plot has been created

We can see Data has both trend and seasonality

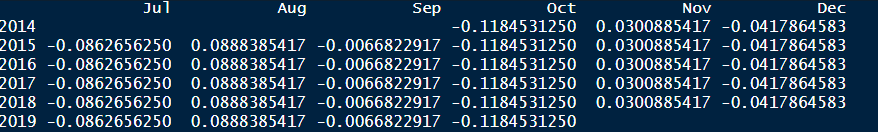
Decomposing



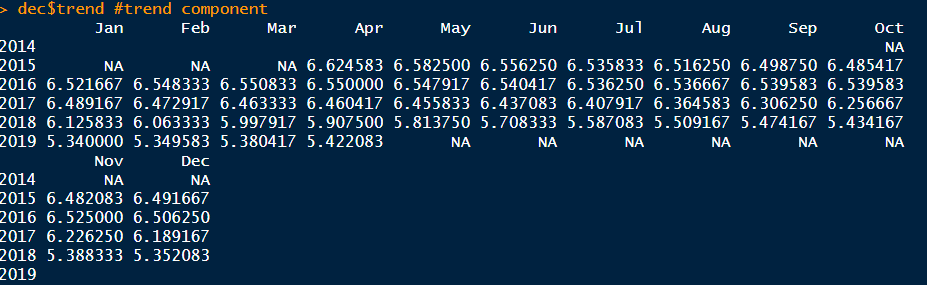
Ordinal series for close price:

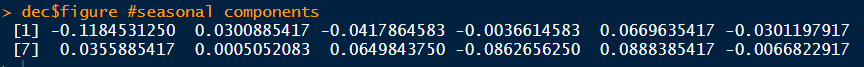


Seasonal indices (guessed based on frequency of ts() object):



Trend component:





We have used additive seasonality:



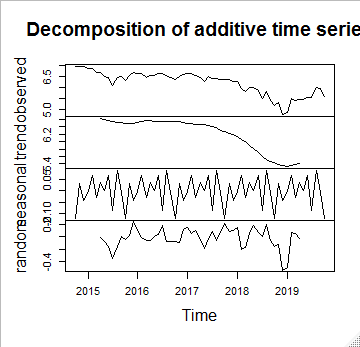


Fig: Des additive time series

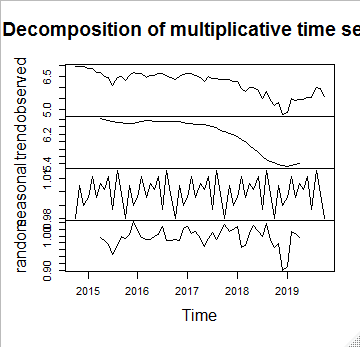
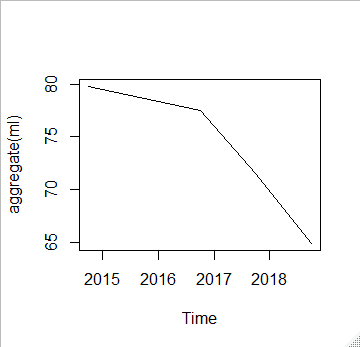


Fig: Dec of multiplicative time series

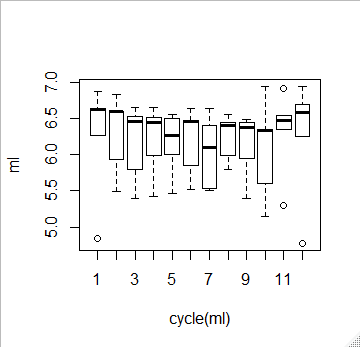
In this plot

* Top panel contains the original time series,
* Second panel contains the trend,
* Third panel contains the seasonality component,
* Last panel contains random fluctuations

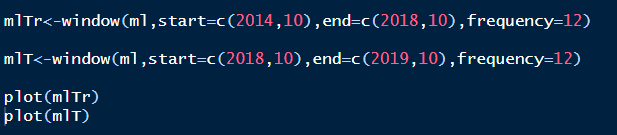
Then we aggregating the time series



Boxplot creates one boxplot for each month



Then we have created Subset the time series by using window function. We should to create two windows (2014 to 2018) and (2018 to 2019). We have Created Training and Testing Window datasets as mltr(train) and mlt(test). Frequency is the number of time periods in a year



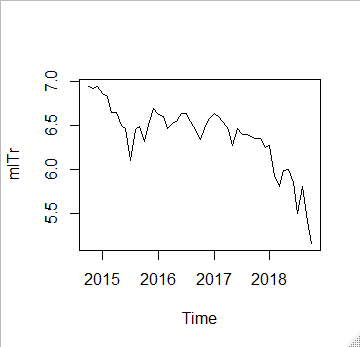


Fig: Train time series visualization

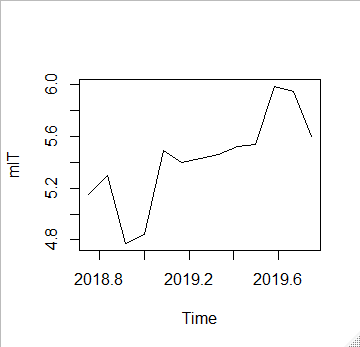
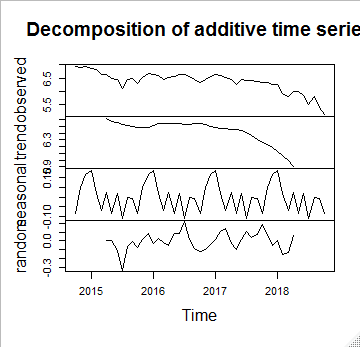


Fig: test time series visualization

We shall create a time series model on (2014 to 2017) and forecast for the period (2017 to 2019) and more then 1 years. We shall then compare the forecasts with the actual time series data

* Data has both trend and seasonality
* Decomposing the Time Series into seasonal, trend and irregular
* Components with additive seasonality



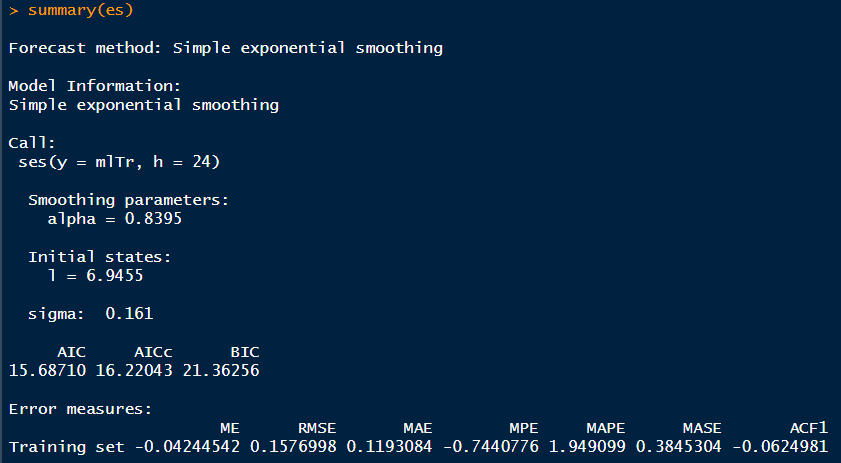
In this plot:

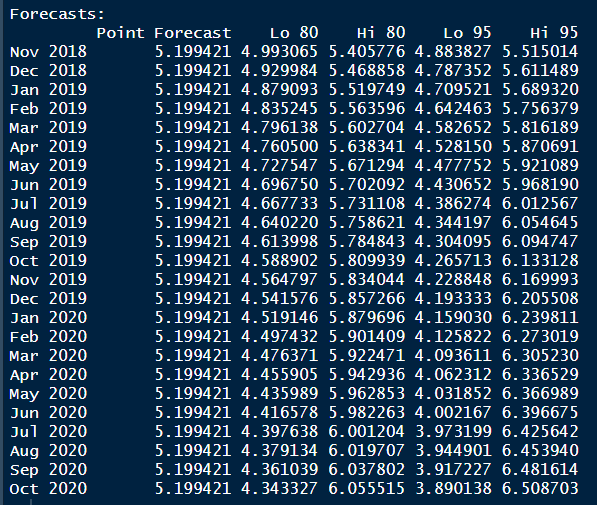
* Top panel contains the original time series
* Second panel contains the trend
* Third panel contains the seasonality component
* Last panel contains random fluctuations

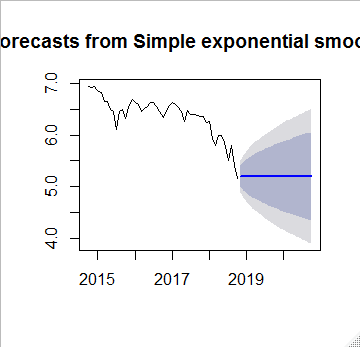
# Exponential smoothing forecasts Model:

Exponential smoothing on training window dataset:

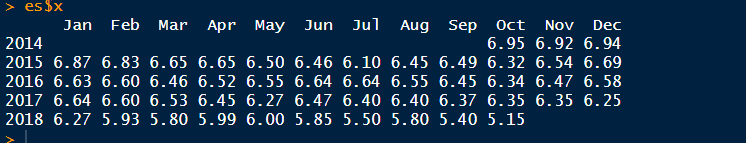
To forecast by Exponential smoothing forecasts share for the next 12 months

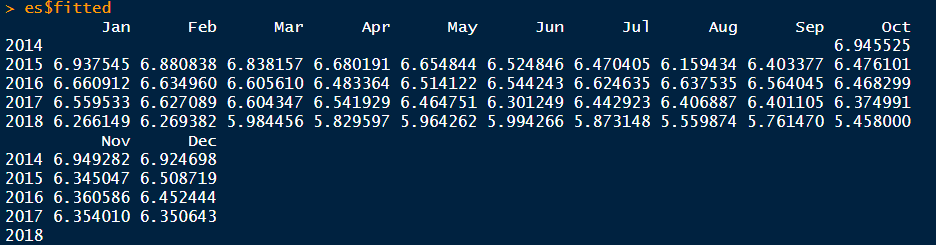




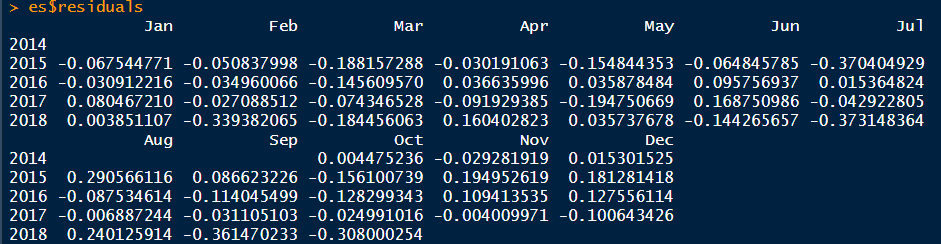


Blue line is slightly elevated, meaning that this forecast



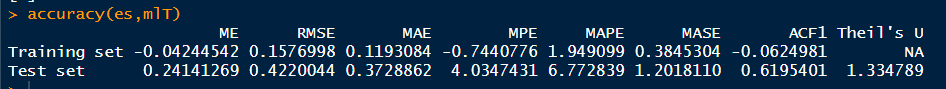


Residuals: Actual-predicted

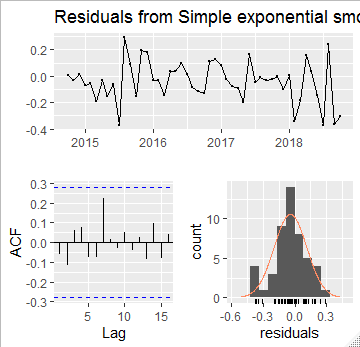




Finding the accuracy of exponential smoothing model:



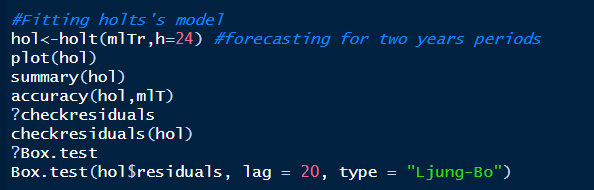
Checking residuals for ES model:



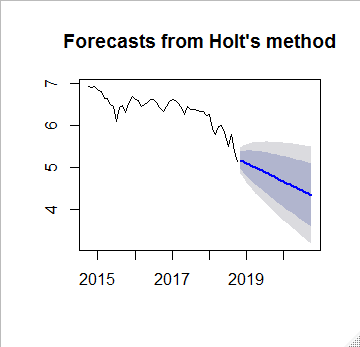
There is a pattern in the residuals. Therefore this forecast is not sufficiently an accurate forecast.

Hence, we use the Holt's method.

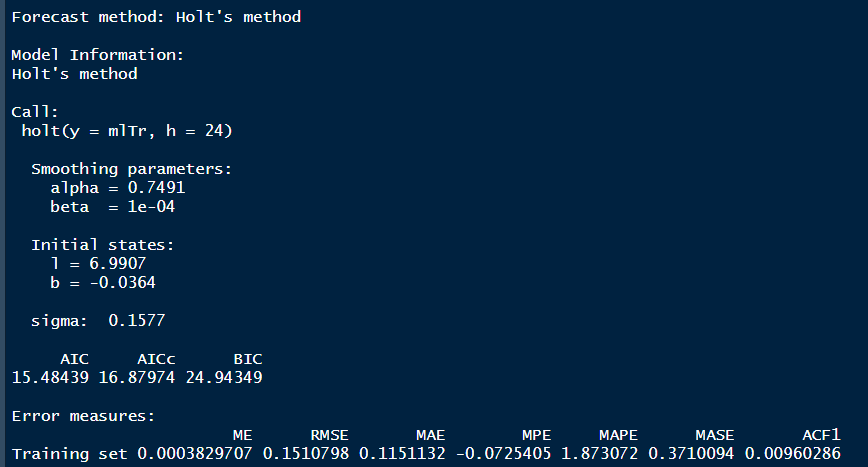
# Holt's method

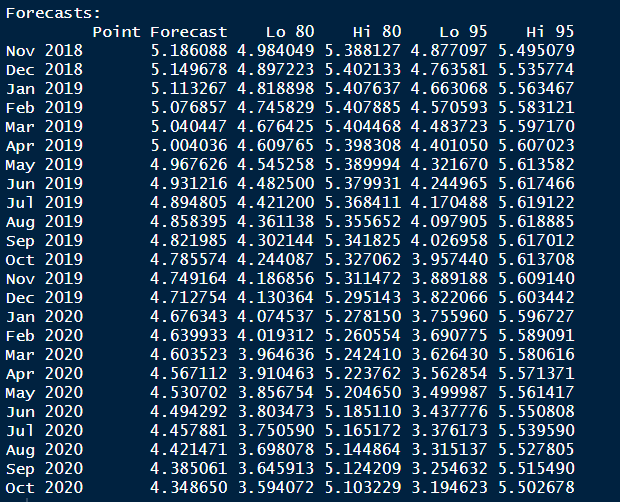


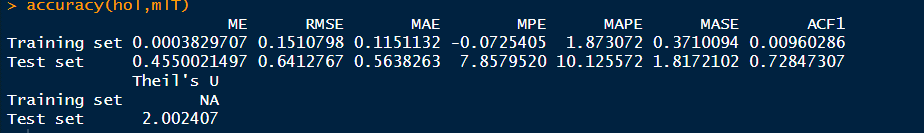
Fitting holt's model:



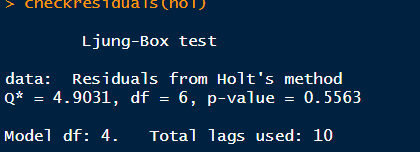
Blue line is slightly elevated, meaning that this forecast

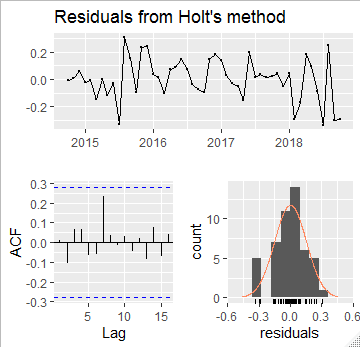


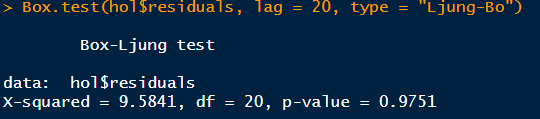




Check that residuals by checkresuduals () function from a ts model:





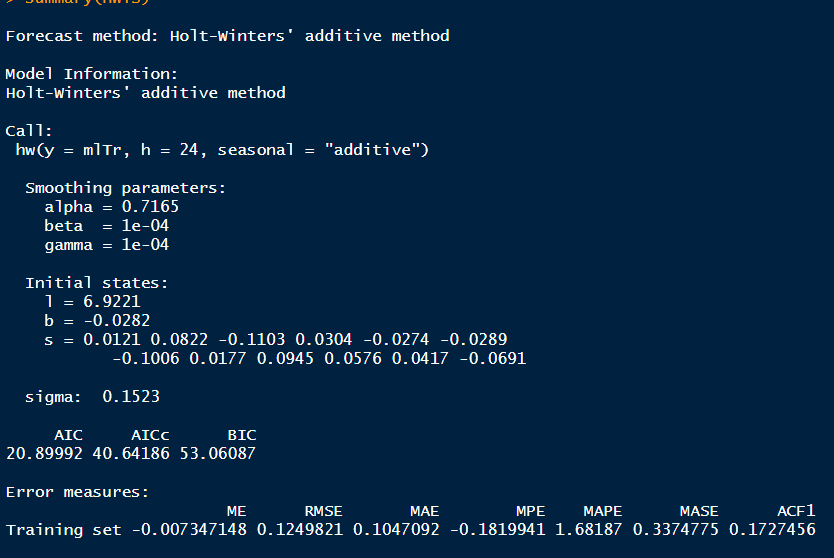


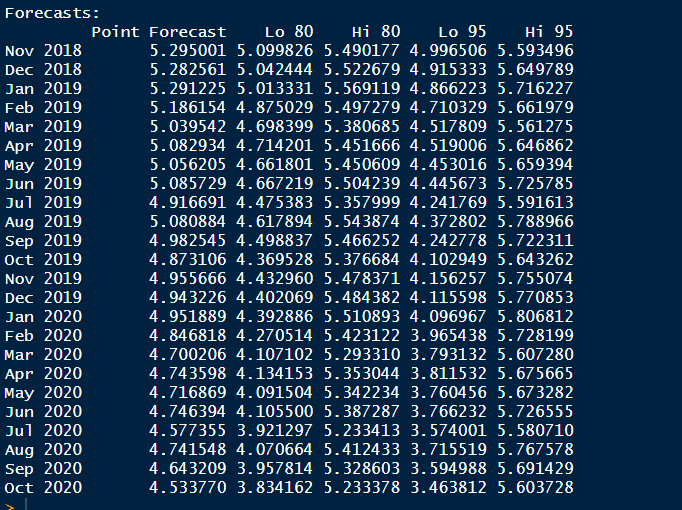
Here in this plot:

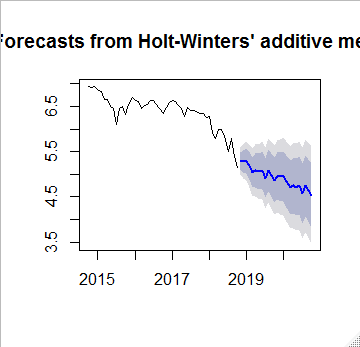
* Blue line is slightly elevated, meaning that this forecast
* This is the better than the previous one.
* The Ljung-Box test reveals that p-value is low.
* Data is identically and independently distributed (white noise)
* Data is not identically and independently
* Distributed
* Therefore we reject Ho and conclude that data exhibits serial correlation.
* clearly the series is seasonal so seasonal component is required

Finally trying the Holt winter's method

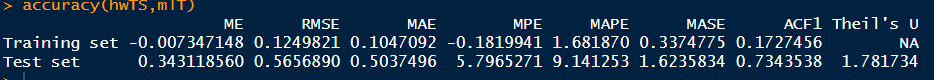
# Holt winter’s Model for Trend & Seasonality

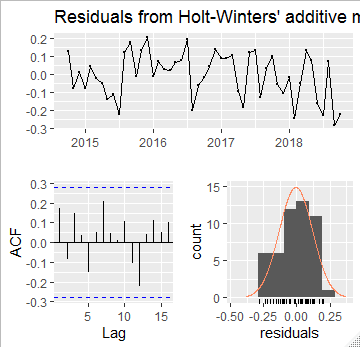


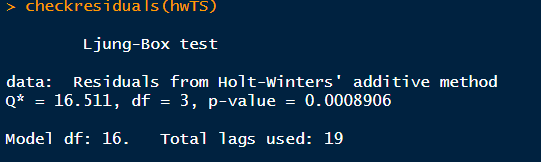




Blue line is slightly elevated, meaning that this forecast



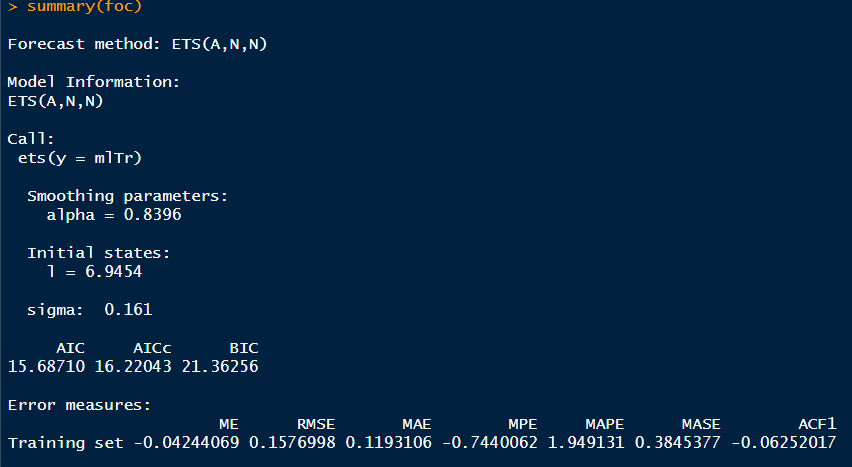




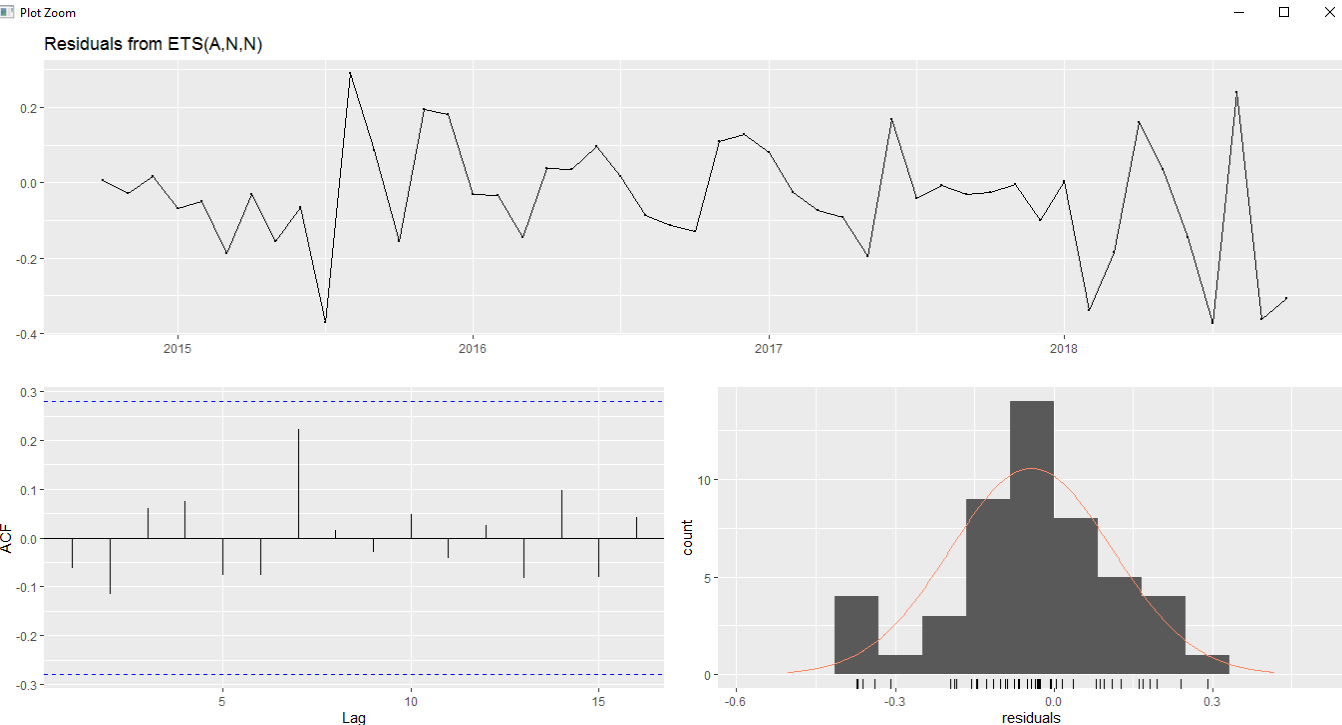
Here

* p-value is less, indicating that data is identically and
* Independently distributed
* All spikes are below blue line.
* There's no pattern in the residuals
* Histogram is also closer to bell-shaped

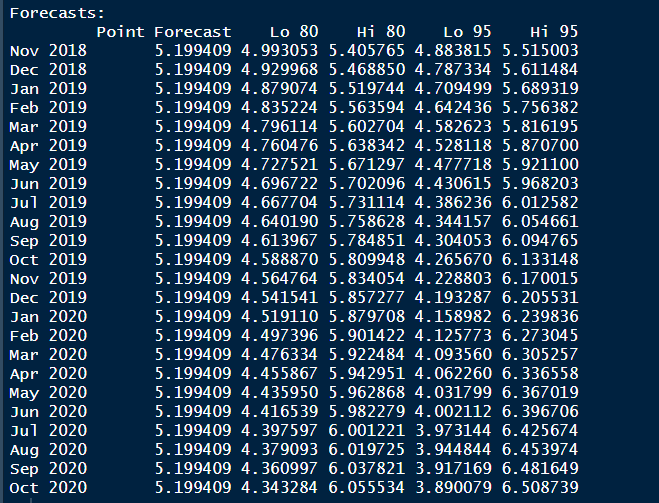
# Automating model building using ETS ()

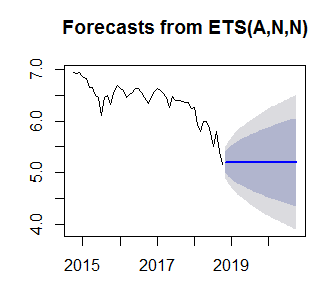


The MAPE is less than 7 and it seems like a good MAPE.



# Forecast by ETS

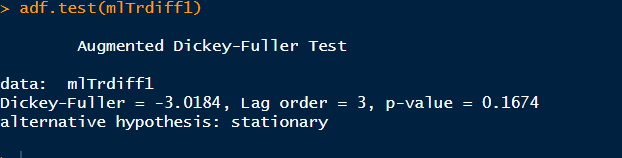


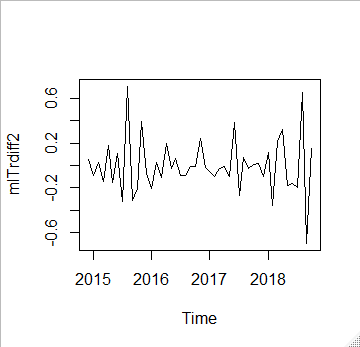


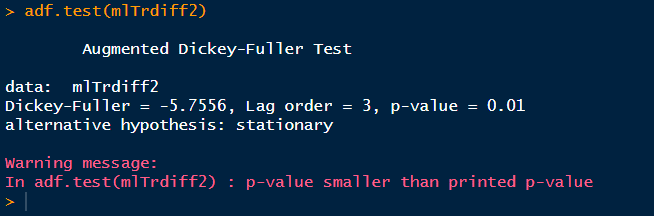
Blue line is slightly elevated, meaning that this forecast

# ARIMA Model

# Integration Part; statistical way of checking stationarity:



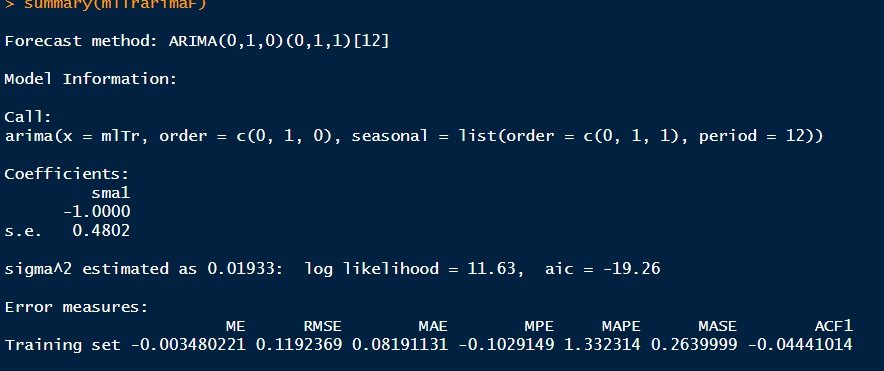


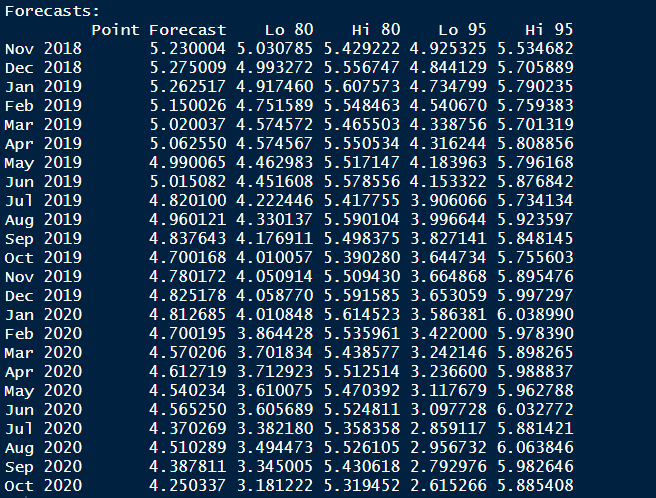


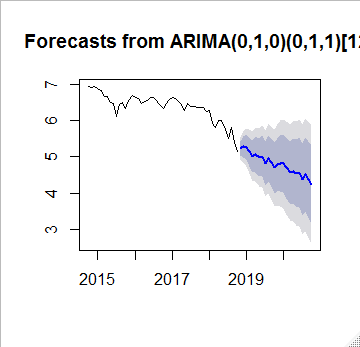
Sometimes we might require taking the log of the time series

We can use auto.arima() to find optimal values of p.d and q.

# Auto Arima model

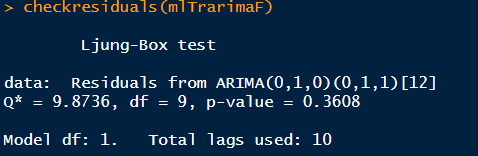


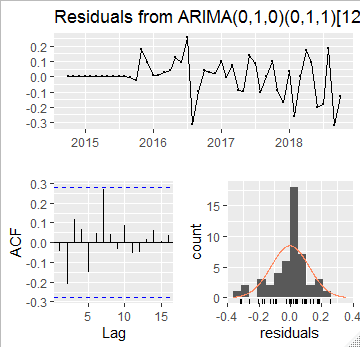




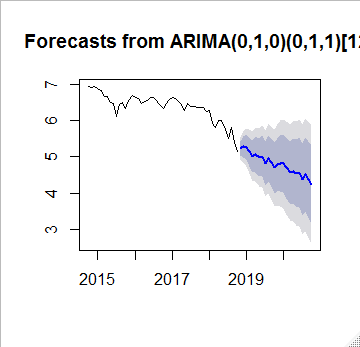
Blue line is slightly elevated, meaning that this forecast

# Validating of Arima model





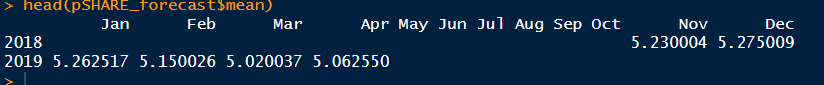
Dataset STOCK forecasting for bank the next 2 years:



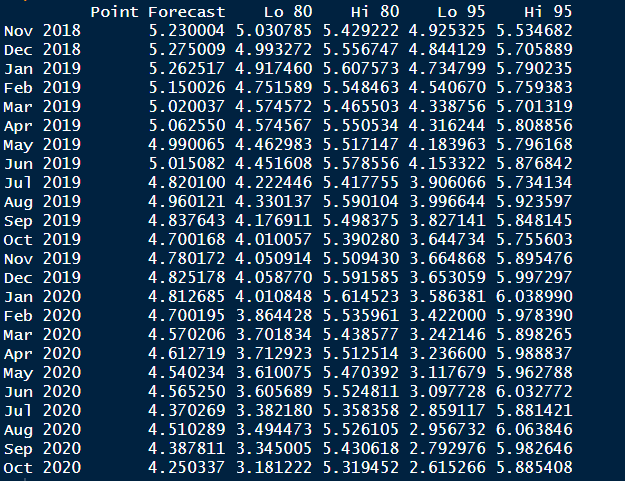
Blue line is slightly elevated, meaning that this forecast

As we can see, we have a blue line that represents the mean of our prediction:

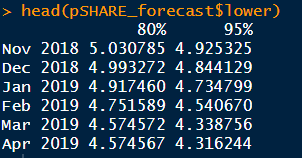
Dataset forecast mean first 5 values



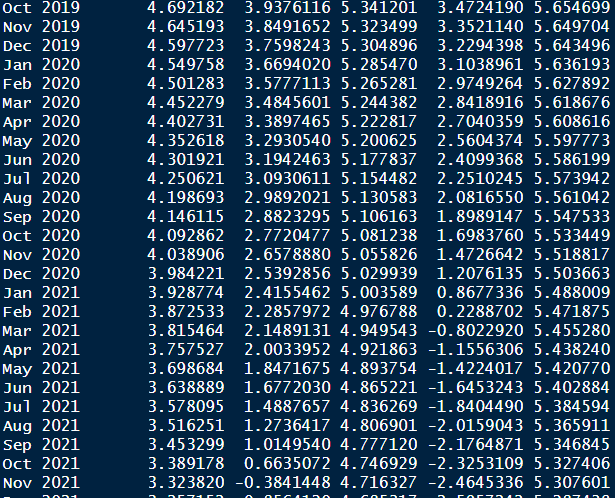
Stock forecast by Arima Model

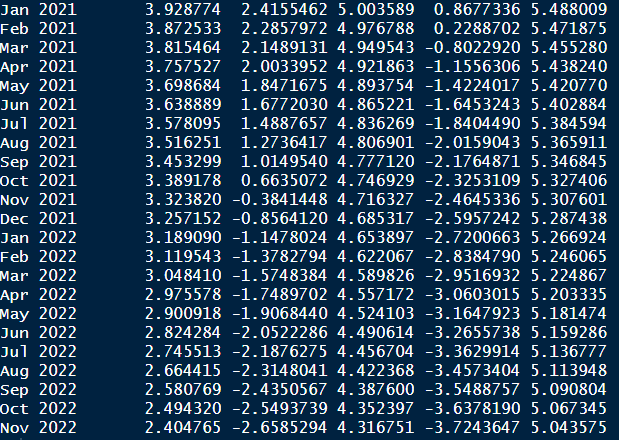


Dataset forecast lower first 5 values

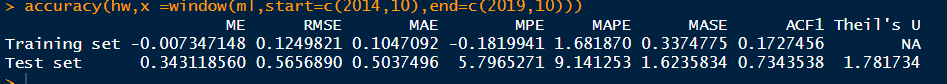


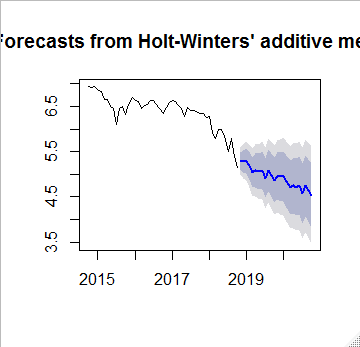
Auto Arima model



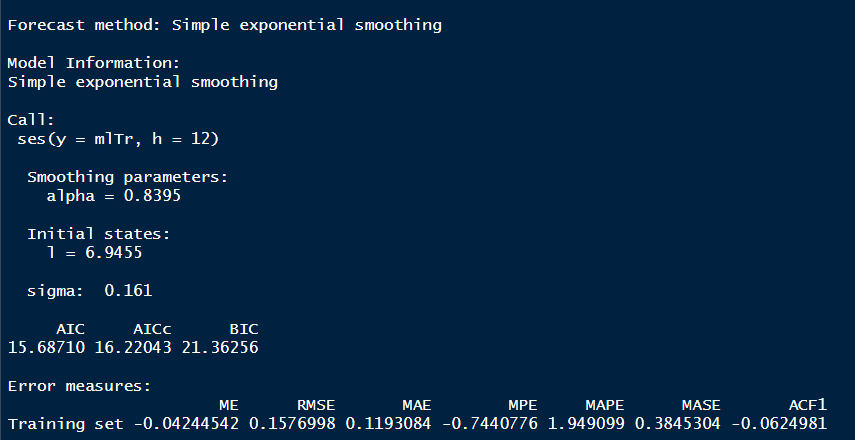


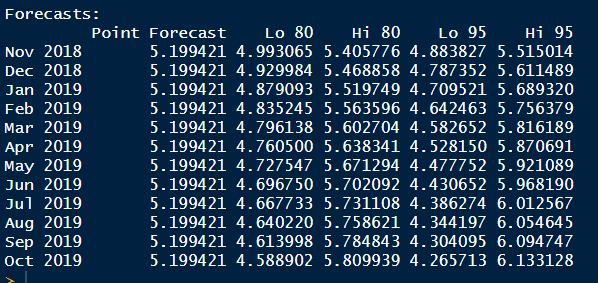
Clearly the series is seasonal so seasonal component is required:



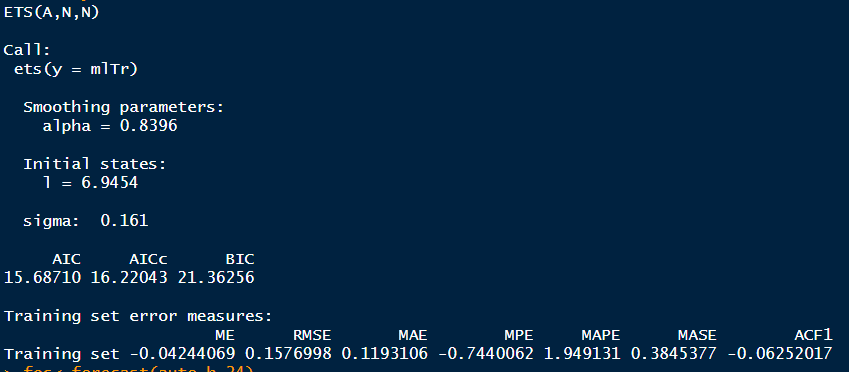


Blue line is slightly elevated, meaning that this forecast

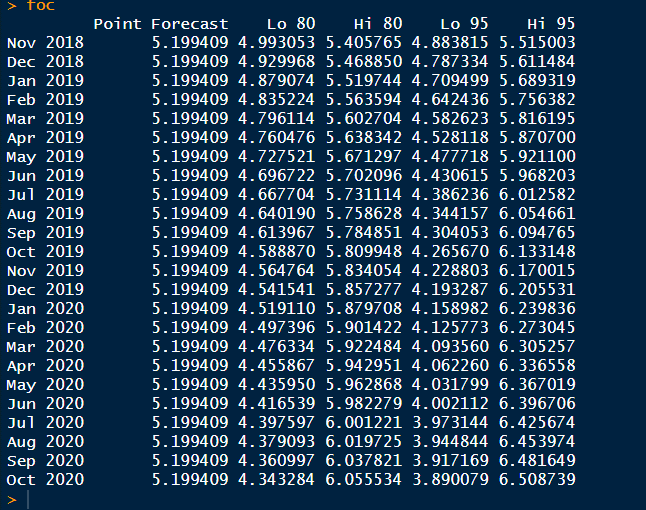




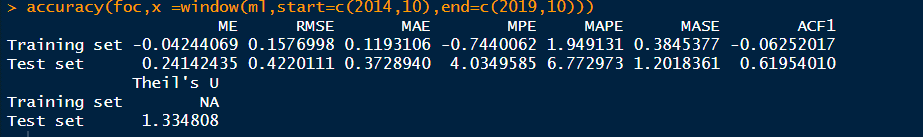
# Automating model building using ETS ()

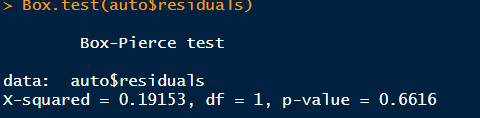


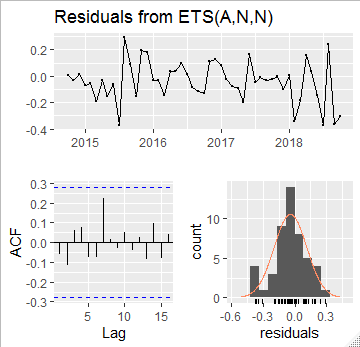
# Forecast by ETS:

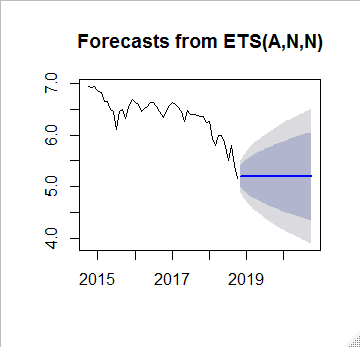


# Model accuracy









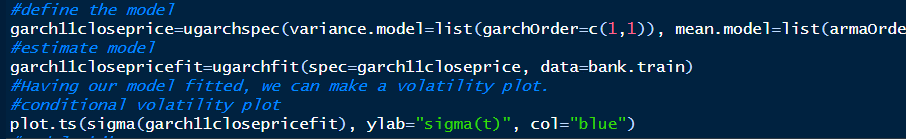
Blue line is slightly elevated, meaning that this forecast

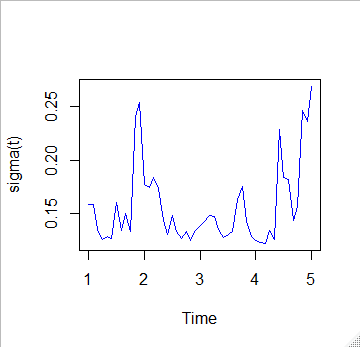
# Garch model:

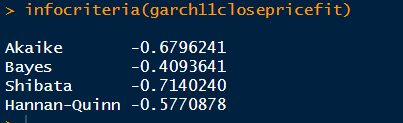
Once we have our package for garch loaded we proceed to apply the model previously defined to the close price dataset. For finding ARFIMA parameters we run auto Arima function.

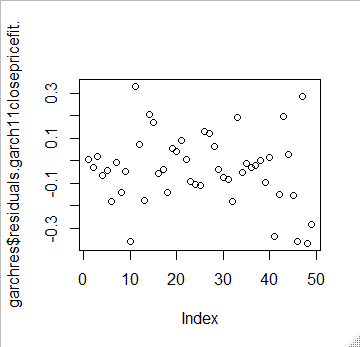
##Dataset forecast upper first 5 values

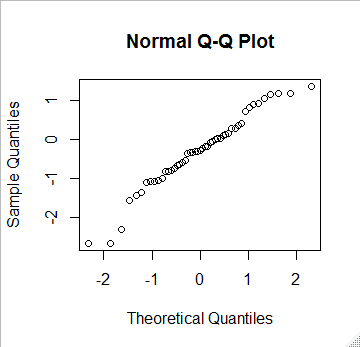
With the parameters collected we choose ARFIMA (1, 0, and 2) and incorporate the parameters to a garch model. Then we define the Garch model

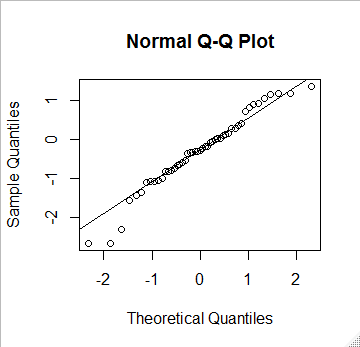




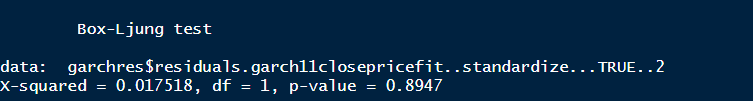




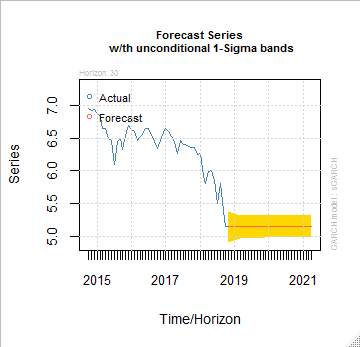


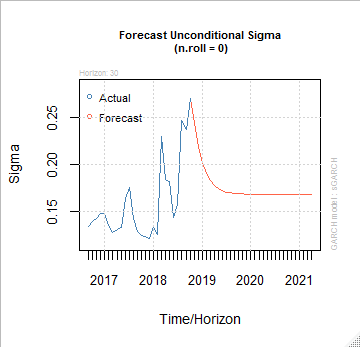


Squared standardized residuals Ljung Box



GARCH Forecasting





# Conclusion:

Time series analysis result designate that:

* Different time series algorithms has performed differently on different stock of banks.
* The model which get trained in a smooth period and predict well in shock period may be considered for future predictions of banks.
* For trader it can be better to watch stock prices those stock prices (previous day), as well, which will influence the movement of the bank stock.
* Here we have applied 4+ time series model to check the accuracy of each model, then we have found Arima and Garch model forecast better than other forecast model.

# References:

* Liu, H., Huang, D., & Wang, Y. (2011, July). Chaotic dynamics analysis and forecast of stock time series. In *2011 International Symposium on Computer Science and Society* (pp. 75-78). IEEE.
* Narain, L. S., & Narain, R. L. (2002). Stock market prediction: A comparitive study of multi vairate statistical and artificial neural network models. *Journal of Accounting and Finance Research*, *10*(2), 85.
* SONG, L., & CHEN, J. X. (2004). An Empirical Investigation on Effect of Listed Companies' Annual Report Announcements on Stock Prices [J]. *Commercial Research*, *19*.
* Kute, S., & Tamhankar, S. (2013). A survey on stock market prediction techniques. *International Journal of Science and Research*
* Monahan, S. J. (2006). Discussion of Why Do Managers Voluntarily Issue Cash Flow Forecasts?. *Journal of Accounting Research*, 431-436.
* EMELE, C. R., OKPALEKE, F. C., & UMEH, O. L. (2014). Commercial real estate market forecasts: complexities, methodologies and opportunities in the Lagos mega city. *Journal January/February*, *51*, 67.