CHITH SABESH
chithsabesh@gmail.com

ANALYSING AND FORECASTING STOCK PRICES OF 4 INDIAN BANKS
Milestone report-2

The Milestone report 1 looked at the overview,Dataset, Exploratory Data Analysis and Inferential Statistics for 4 major Banks in the Indian Banking sector.

This report will look into the Machine Learning part of the project.Wherein we try different forecasting methods for each bank in question.

STATIONARISING EACH SERIES

ICICI: Differencing of order 1 was required to make it stationary

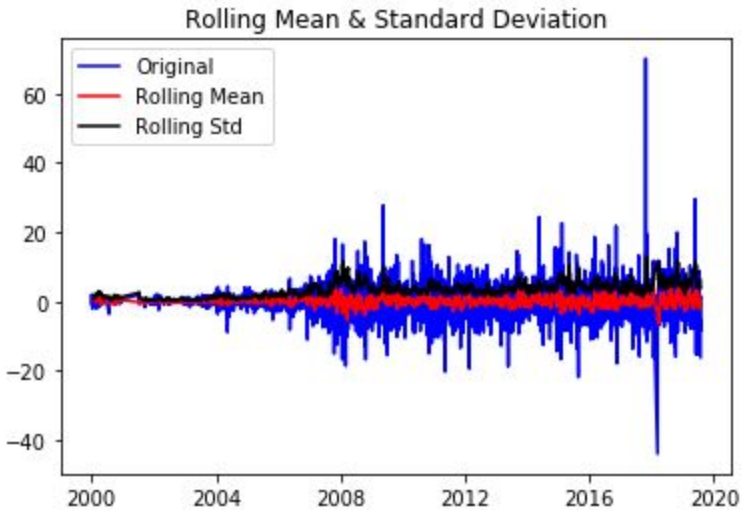
HDFC: Differencing of order 10 was required to make it stationary

AXIS: Differencing of order 20 was required to make it stationary

SBI: Differencing of order 1 was required to make it stationary

NOTE:Differencing of 10 and 20 is wrong and should not be done.This will be addressed towards the end.

The Augumented Dickey Fuller Test was done for the four banks the test statistic was found to be lesser then the critical value.




```
Results of Dickey-Fuller Test:
Test Statistic                  -63.301655
p-value                           0.000000
#Lags Used                        0.000000
Number of Observations Used    4631.000000
Critical Value (1%)              -3.431763
Critical Value (5%)              -2.862164
Critical Value (10%)             -2.567102
dtype: float64
```

MODELS BUILT

1) Moving Average
2) ARIMA(Auto Regressive Integrated Moving Average)
3) Linear Regression

MODEL PERFORMANCE:

	Moving Average	Arima	Linear Regression
icici	5.111	5.131	5.115
hdfc	44.207	44.719	44.467
axis	36.912	37.187	36.922
sbi	5.184	5.262	5.182

From the model performance table it is easy to tell that all models gave similar performance for each individual bank.

But the differencing for HDFC and AXIS were 10 and 20 respectively. Such high levels of differencing is not right and hence we should try another method to make it stationary.

Decomposing is another such method where we model only on the residual component by separating the trend and seasonality.

This increased model accuracy for both HDFC and Axis.With Auto ARIMA giving an RMSE of 0.48 and 0.42 respectively.

BEST VALUES:
Best values for p,q,r for ARIMA in each bank was
ICICI: (1,1,1)
HDFC:(0,1,0)
AXIS::(0,1,0)
SBI:(1,1,1)

FEATURE ENGINEERING:
Only for Linear Regression feature Engineering was done. With emphasis being given on the date. Preference given to beginning of the week,month and year. Since Stock prices depend on starting and ending of these three.

CONCLUSION:
Since Stock Price forecasting is a tough and vast domain.The further work that can be done in this project is to get News data about all the 4 banks.