CHITH SABESH
chithsabesh@gmail.com

ANALYSING AND FORECASTING CLOSE PRICES OF 4 INDIAN BANKS
Milestone report-1

OVERVIEW

The problem here deals with forecasting the close prices of 4 major Indian Banks namely ICICI,HDFC,Axis and SBI.Project aims at coming up with a forecasting model for these respective banks.

THE DATA

The data comes from https://eodhistoricaldata.com/ and hence no scraping is required.Contains stock prices of all companies present in the Bombay Stock Exchange(BSE).Out of which i have chosen the above 4 banks respectively.

DATA WRANGLING

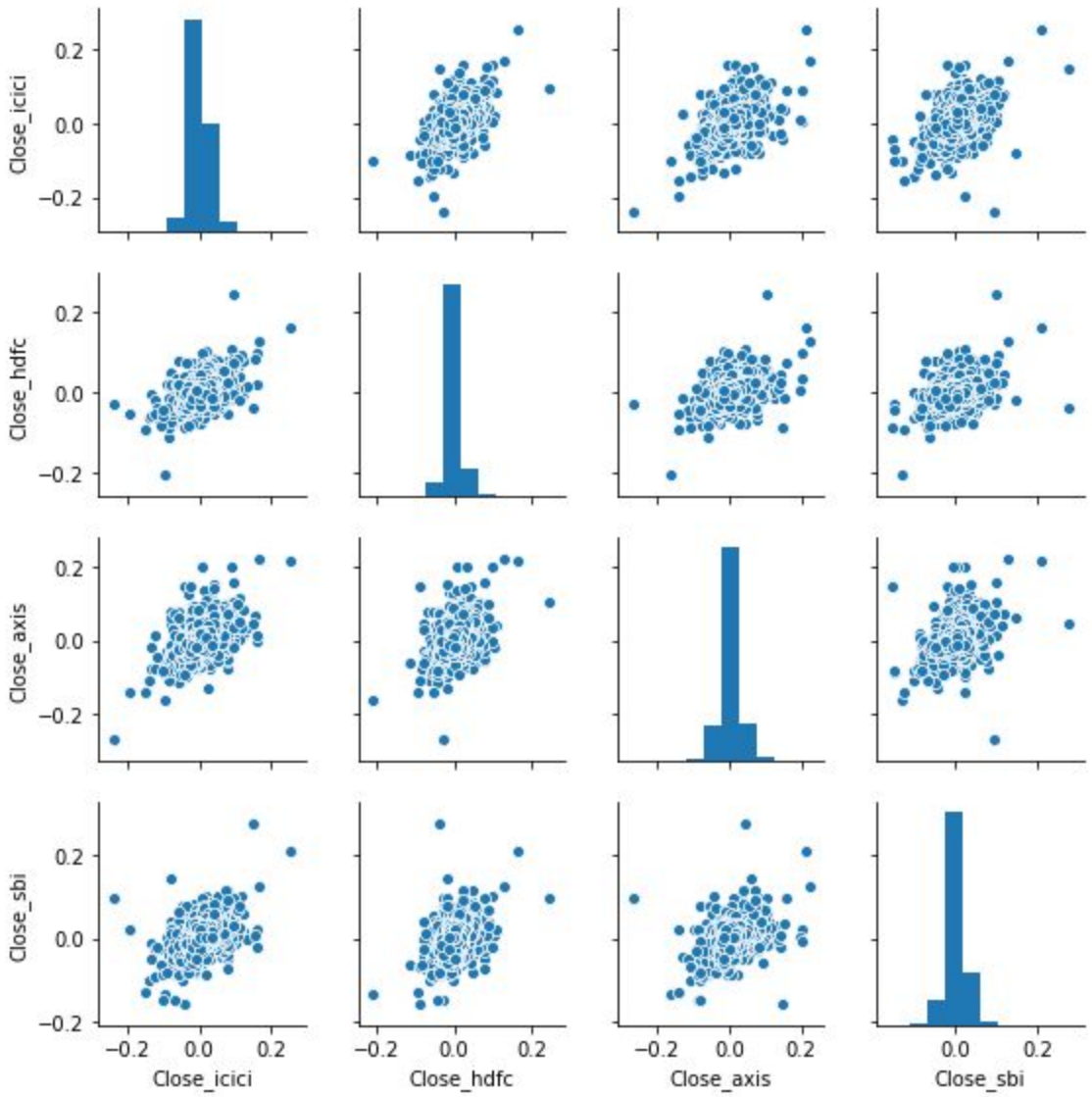
Not much data wrangling was required here. Only the process of getting the right company codes from https://www.moneycontrol.com/ was done. Since the initial data sets only had company codes.

EXPLORATORY DATA ANALYSIS

Plotting the timeline of closed prices for all 4 banks

Daily returns of all with respect to each other

Daily returns is the percentage change of prices from one day to the next



Shows very large correlation of returns between all 4 banks. This will be further explained by plotting the correlation matrix.

WORST SINGLE DAY DROP

```
Close_icici   2001-07-02
Close_hdfc    2004-05-17
Close_axis    2001-07-02
Close_sbi     2000-07-24
dtype: datetime64[ns]
```

ICICI,Axis and SBI all had their worst drops in the same month. In fact axis and ICICI had it in the same day.

BEST SINGLE DAY GAIN

```
Close_icici   2009-05-18
Close_hdfc    2004-05-18
Close_axis    2008-10-13
Close_sbi     2017-10-25
dtype: datetime64[ns]
```

Weirdly HDFC had its biggest gain the next day after it had its biggest drop. While SBI grew only recently
[SBI grows Rs 61,000 crore bigger in a day]

WHICH BANK IS THE RISKIEST

```
Close_icici    0.027665
Close_hdfc     0.019688
Close_axis     0.028203
Close_sbi      0.023709
dtype: float64
```

Axis bank is the most riskiest while HDFC is the least

LETS CHECK RISKS FOR 2019

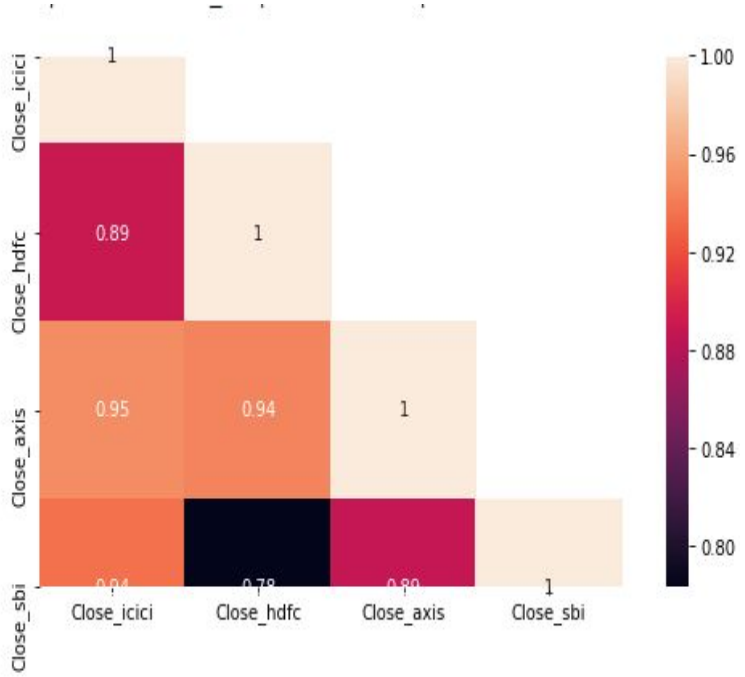
```
Close_icici    0.016087
Close_hdfc     0.009590
Close_axis     0.015328
Close_sbi      0.016895
dtype: float64
```

Risks has reduced for all 4 banks

INFERENTIAL STATISTICS

CORRELATION

	Close_icici	Close_hdfc	Close_axis	Close_sbi
Close_icici	1.000000	0.890280	0.948446	0.935630
Close_hdfc	0.890280	1.000000	0.944846	0.783369
Close_axis	0.948446	0.944846	1.000000	0.888123
Close_sbi	0.935630	0.783369	0.888123	1.000000

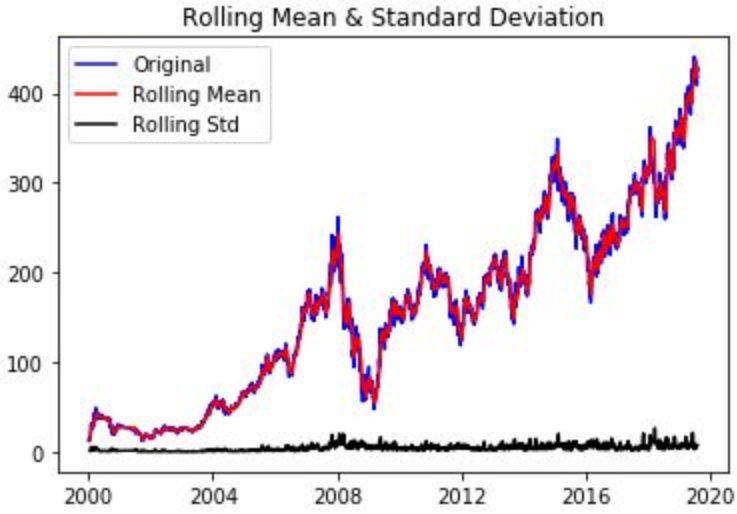


All 4 banks have large correlations between their closing prices.

CHECKING STATIONARITY

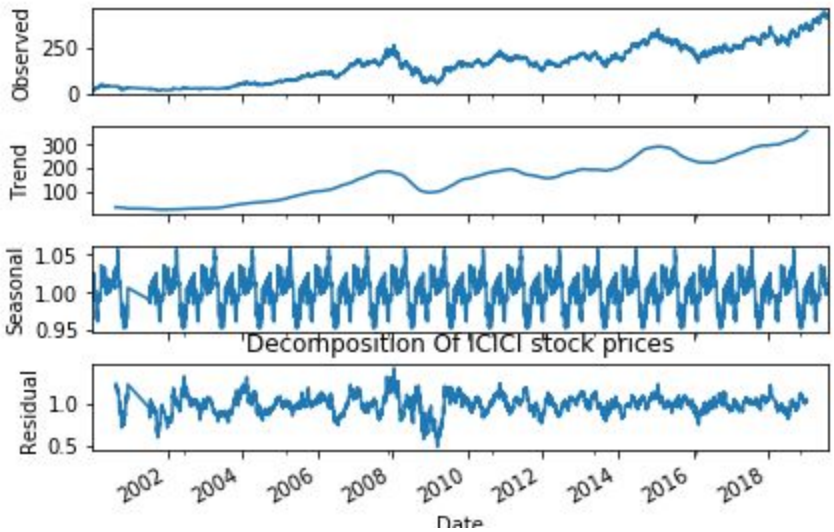
I am only showing the stationarity of ICICI. The stationarity of other 3 banks is available in the jupyter notebook.

The Graph in the next page and along with the Augmented Dickey Fuller Test (which tests whether a particular Time series is stationary) is given and from the values we can tell that the time series is not stationary since the calculated test statistic is less then its critical value.



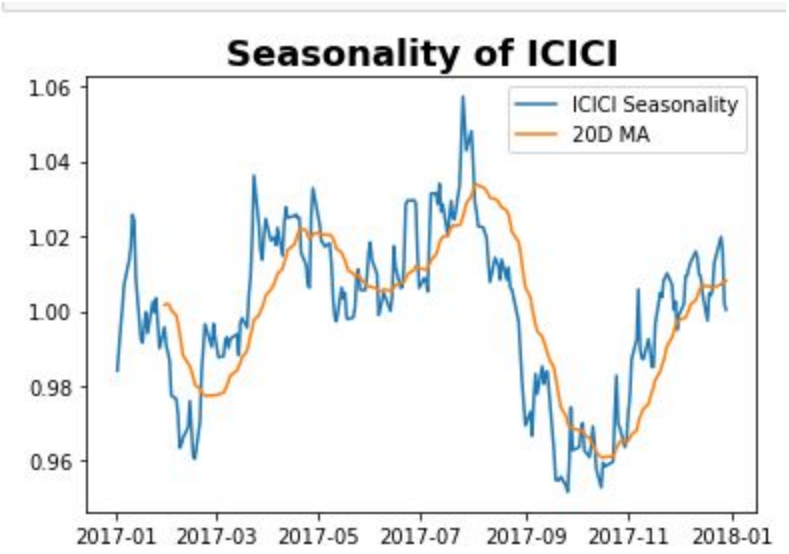

```
Results of Dickey-Fuller Test:
Test Statistic                  -0.301125
p-value                          0.925345
#Lags Used                       6.000000
Number of Observations Used    4626.000000
Critical Value (1%)             -3.431764
Critical Value (5%)             -2.862165
Critical Value (10%)            -2.567103
dtype: float64
```

DECOMPOSING THE SAME ICICI STOCK PRICES



General upward trend can be spotted and a seasonal pattern which is repeating every year with a wave like pattern is also found.

DRILLING DOWN THE SEASONALITY



When plotting seasonality year wise we can find out that there are major lows during October and November. While the highs are during August.

CONCLUSION

This report highlights the Wrangling, Exploratory data analysis and inferential statistics done on the stock price dataset. With these insights it is now possible to move to the steps Machine learning