FINANCIAL RISK ASSESSMENT

MILESTONE 2

BUSINESS REPORT

ABSTRACTS OF CREDIT AND MARKET RISK

Dataset available includes information from the financial statement of the companies for the previous year (2015). Also, information about the Net worth of the company in the following year (2016). The dataset contains 6 years of information (weekly stock information) on the stock prices of 10 different Indian Stocks. Calculate the mean and standard deviation on the stock returns and share insights.

Yashveer Kothari. A
Post Graduate Programme in Data Science & Business Analytics

LIST OF CONTENTS

ABOUT FINANCIAL RISK ASSESSMENT

- A financial risk assessment is an evaluation of the probability that a financial asset or investment could lose value.
- A financial risk assessment can help you identify the risk level specific to your business, prioritize those risks, develop ways to avoid them and outline steps to manage them should they happen. A financial risk assessment also helps you understand your appetite for risk.
- Some common financial risks are credit, operational, foreign investment, legal, equity, and liquidity risks.

TYPES OF FINACIAL RISKS

FINANCIAL RISK

- **CREDIT RISK**
- **LIQUIDITY RISK**
- **SYSTEMATIC RISK**
- **MARKET RISK**
- **OPERATIONAL RISK**

1. **Credit Risk:** This type of risk arises when one fails to fulfil their obligations towards their counterparties. Credit risk can be classified into Sovereign Risk and Settlement Risk. Sovereign risk usually arises due to difficult foreign exchange policies. Settlement risk, on the other hand, arises when one party makes the payment while the other party fails to fulfil the obligations

2. **Liquidity Risk:** This type of risk arises out of an inability to execute transactions. Liquidity risk can be classified into Asset Liquidity Risk and Funding Liquidity Risk. Asset Liquidity risk arises either due to insufficient buyers or insufficient sellers against sell orders and buys orders respectively.

3. **Systematic Risk:** It refers to the risk inherent to the entire market or market segment. Systematic risk, also known as "undiversifiable risk," "volatility" or "market risk," affects the overall market, not just a particular stock or industry.

4. **Market Risk:** This type of risk arises due to the movement in prices of financial instrument. Market risk can be classified as Directional Risk and Non-Directional Risk. Directional risk is caused due to movement in stock price, interest rates and more. Non-Directional risk, on the other hand, can be volatility risks.

5. **Operational Risk:** This type of risk arises out of operational failures such as mismanagement or technical failures. Operational risk can be classified into Fraud Risk and Model Risk. Fraud risk arises due to the lack of controls and Model risk arises due to incorrect model application.

PROBLEM STATEMENTS

- **CREDIT RISK:** Businesses or companies can fall prey to default if they are not able to keep up their debt obligations. Defaults will lead to a lower credit rating for the company which in turn reduces its chances of getting credit in the future and may have to pay higher interests on existing debts as well as any new obligations. From an investor's point of view, he would want to invest in a company if it is capable of handling its financial obligations, can grow quickly, and is able to manage the growth scale. A balance sheet is a financial statement of a company that provides a snapshot of what a company owns, owes, and the amount invested by the shareholders. Thus, it is an important tool that helps evaluate the performance of a business. Data that is available includes information from the financial statement of the companies for the previous year (2015). Also, information about

the Net worth of the company in the following year (2016) is provided which can be used to drive the labeled field. Explanation of data fields available in Data Dictionary, 'Credit Default Data Dictionary.xlsx'

DATA DICTIONARY – CREDIT RISK

#	FIELD NAME	DESCRIPTION
1	Co_Code	Company Code
2	Co_Name	Company Name
3	Net worth Next Year	Value of a company as on 2016 - Next Year (difference between the value of total assets and total liabilities)
4	Equity Paid Up	Amount that has been received by the company through the issue of shares to the shareholders
5	Net worth	Value of a company as on 2015 - Current Year
6	Capital Employed	Total amount of capital used for the acquisition of profits by a company
7	Total Debt	The sum of money borrowed by the company and is due to be paid
8	Gross Block	Total value of all of the assets that a company owns
9	Net Working Capital	The difference between a company's current assets (cash, accounts receivable, inventories of raw materials and finished goods) and its current liabilities (accounts payable).
10	Current Assets	All the assets of a company that are expected to be sold or used as a result of standard business operations over the next year.
11	Current Liabilities and Provisions	Short-term financial obligations that are due within one year (includes amount that is set aside cover a future liability)
12	Total Assets/Liabilities	Ratio of total assets to liabilities of the company
13	Gross Sales	The grand total of sale transactions within the accounting period

14	Net Sales	Gross sales minus returns, allowances, and discounts
15	Other Income	Income realized from non-business activities (e.g., sale of long-term asset)
16	Value Of Output	Product of physical output of goods and services produced by company and its market price
17	Cost of Production	Costs incurred by a business from manufacturing a product or providing a service
18	Selling Cost	Costs which are made to create the demand for the product (advertising expenditures, packaging and styling, salaries, commissions and travelling expenses of sales personnel, and the cost of shops and showrooms)
19	PBIDT	Profit Before Interest, Depreciation & Taxes
20	PBDT	Profit Before Depreciation and Tax
21	PBIT	Profit before interest and taxes
22	PBT	Profit before tax
23	PAT	Profit After Tax
24	Adjusted PAT	Adjusted profit is the best estimate of the true profit
26	CP	Commercial paper, a short-term debt instrument to meet short-term liabilities.
27	Revenue earnings in forex	Revenue earned in foreign currency
28	Revenue expenses in forex	Expenses due to foreign currency transactions
29	Capital expenses in forex	Long term investment in forex
30	Book Value (Unit Curr)	Net asset value
31	Book Value (Adj.) (Unit Curr)	Book value adjusted to reflect asset's true fair market value
32	Market Capitalisation	Product of the total number of a company's outstanding shares and the current market price of one share
33	CEPS (annualised) (Unit Curr)	Cash Earnings per Share, profitability ratio that measures the financial performance of a company by calculating cash flows on a per share basis

34	Cash Flow from Operating Activities	Use of cash from ongoing regular business activities
35	Cash Flow from Investing Activities	Cash used in the purchase of non-current assets–or long-term assets– that will deliver value in the future
36	Cash Flow from Financing Activities	Net flows of cash that are used to fund the company (transactions involving debt, equity, and dividends)
37	ROG-Net Worth (%)	Rate of Growth - Net worth
38	ROG-Capital Employed (%)	Rate of Growth - Capital Employed
39	ROG-Gross Block (%)	Rate of Growth - Gross Block
40	ROG-Gross Sales (%)	Rate of Growth - Gross Sales
41	ROG-Net Sales (%)	Rate of Growth - Net Sales
42	ROG-Cost of Production (%)	Rate of Growth - Cost of Production
43	ROG-Total Assets (%)	Rate of Growth - Total Assets
44	ROG-PBIDT (%)	Rate of Growth- PBIDT
45	ROG-PBDT (%)	Rate of Growth- PBDT
46	ROG-PBIT (%)	Rate of Growth- PBIT
47	ROG-PBT (%)	Rate of Growth- PBT
48	ROG-PAT (%)	Rate of Growth- PAT
49	ROG-CP (%)	Rate of Growth- CP
50	ROG-Revenue earnings in forex (%)	Rate of Growth - Revenue earnings in forex
51	ROG-Revenue expenses in forex (%)	Rate of Growth - Revenue expenses in forex
52	ROG-Market Capitalisation (%)	Rate of Growth - Market Capitalisation
53	Current Ratio [Latest]	Liquidity ratio, company's ability to pay short-term obligations or those due within one year
54	Fixed Assets Ratio [Latest]	Solvency ratio, the capacity of a company to discharge its obligations towards long-term lenders indicating

55	Inventory Ratio [Latest]	Activity ratio, specifies the number of times the stock or inventory has been replaced and sold by the company
56	Debtors Ratio [Latest]	Measures how quickly cash debtors are paying back to the company
57	Total Asset Turnover Ratio [Latest]	The value of a company's revenues relative to the value of its assets
58	Interest Cover Ratio [Latest]	Determines how easily a company can pay interest on its outstanding debt
59	PBIDTM (%) [Latest]	Profit before Interest Depreciation and Tax Margin
60	PBITM (%) [Latest]	Profit Before Interest Tax Margin
61	PBDTM (%) [Latest]	Profit Before Depreciation Tax Margin
62	CPM (%) [Latest]	Cost per thousand (advertising cost)
63	APATM (%) [Latest]	After tax profit margin
64	Debtors Velocity (Days)	Average days required for receiving the payments
65	Creditors Velocity (Days)	Average number of days company takes to pay suppliers
66	Inventory Velocity (Days)	Average number of days the company needs to turn its inventory into sales
67	Value of Output/Total Assets	Ratio of Value of Output (market value) to Total Assets
68	Value of Output/Gross Block	Ratio of Value of Output (market value) to Gross Block

Q.1.8. Build a Random Forest Model on Train Dataset. Also showcase your model building approach.

Using RandomisedSearchCV method, we can define a grid of hyperparameter ranges and randomly sample from the grid, performing K-Fold CV with each combination of values.

1.8.1 Showcase your model building approach

We now fit our model to the GridSearchCV for Random Forest model by training the model with our independent variable and dependent variables

- n_estimators = number of trees in the forest
- max_features = max number of features considered for splitting a node
- max_depth = max number of levels in each decision tree
- min_samples_split = min number of data points placed in a node before the node is split
- min_samples_leaf = min number of data points allowed in a leaf node

TABLE 1: PROBABILITY OF TRAINING AND TEST DATA

	0	1
0	0.982523	0.017477
1	0.994532	0.005468
2	0.932613	0.067387
3	0.945453	0.054547
4	1.000000	0.000000

	0	1
0	0.992703	0.007297
1	0.977112	0.022888
2	0.880947	0.119053
3	0.196207	0.803793
4	0.936337	0.063663

```
                                      IMP
Networth                         0.232167
Book_Value_Unit_Curr             0.216041
Book_Value_Adj_Unit_Curr         0.183907
Current_Ratio_Latest             0.054498
Capital_Employed                 0.049386
...                                   ...
ROG_PBIDT_perc                   0.000314
ROG_PAT_perc                     0.000305
Debtors_Ratio_Latest             0.000293
Cash_Flow_From_Financing_Activities 0.000233
Capital_expenses_in_forex        0.000000
```

- Top three are most important features, Book_Value_Unit_Curr, Networth, Book_Value_Adj_Unit_Curr.

1.9 Validate the Random Forest Model on test Dataset and state the performance matrices. Also state interpretation from the model.

## FIGURE 1: CONFUSION MATRIX TRAIN & TEST DATA

TABLE 3: CLASSIFICATION REPORT TRAIN & TEST DATA

	precision	recall	f1-score	support
0.0	0.97	0.99	0.98	2157
1.0	0.91	0.77	0.83	245
accuracy			0.97	2402
macro avg	0.94	0.88	0.91	2402
weighted avg	0.97	0.97	0.97	2402

	precision	recall	f1-score	support
0.0	0.98	0.99	0.99	1041
1.0	0.93	0.85	0.89	143
accuracy			0.97	1184
macro avg	0.95	0.92	0.94	1184
weighted avg	0.97	0.97	0.97	1184

FIGURE 2: AUC & ROC CURVE -TRAIN & TEST DATA

AUC for the Training Data: 0.991
AUC for the Test Data: 0.988
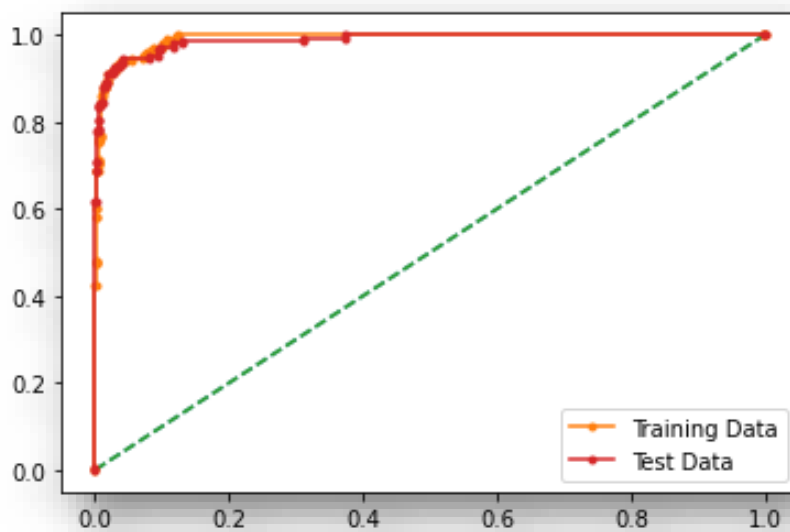
- **Here, recall has increased to 84% from 76% in test data even F1 Score is also increased to 89% with precision of 94%. It is comparatively good model.**

Q.1.10. Build a LDA Model on Train Dataset. Also showcase your model building approach

- Linear Discriminant Analysis (LDA) is a dimensionality reduction technique which is commonly used for the supervised classification problems.
- It is used for modelling differences in groups i.e., separating two or more classes. It is used to project the features in higher dimension space into a lower dimension space.
- Library used in LDA is Sklearn
- Using GridsearchCV, we input various parameters like 'max_iter', 'penalty', solver', 'tol' which will help us to find best grid for prediction of the better model.

TABLE 4: PROBABILITY OF TRAINING AND TEST DATA

	0	1
0	0.958705	0.041295
1	0.996202	0.003798
2	0.546485	0.453515
3	0.999866	0.000134
4	0.999335	0.000665

	0	1
0	0.983966	0.016034
1	0.950279	0.049721
2	0.859898	0.140102
3	0.919085	0.080915
4	0.996063	0.003937

1.11 Validate the LDA Model on test Dataset and state the performance matrices. Also state interpretation from the model.

FIGURE 3: CONFUSION MATRIX TRAIN & TEST DATA

TABLE 5: CLASSIFICATION REPORT TRAIN & TEST DATA

	precision	recall	f1-score	support
0.0	0.94	0.99	0.97	2157
1.0	0.84	0.48	0.61	245
accuracy			0.94	2402
macro avg	0.89	0.74	0.79	2402
weighted avg	0.93	0.94	0.93	2402

	precision	recall	f1-score	support
0.0	0.93	0.98	0.96	1041
1.0	0.79	0.45	0.58	143
accuracy			0.92	1184
macro avg	0.86	0.72	0.77	1184
weighted avg	0.91	0.92	0.91	1184

FIGURE 4: AUC & ROC CURVE -TRAIN & TEST DATA

```
AUC for the Training Data: 0.950
AUC for the Test Data: 0.935
```
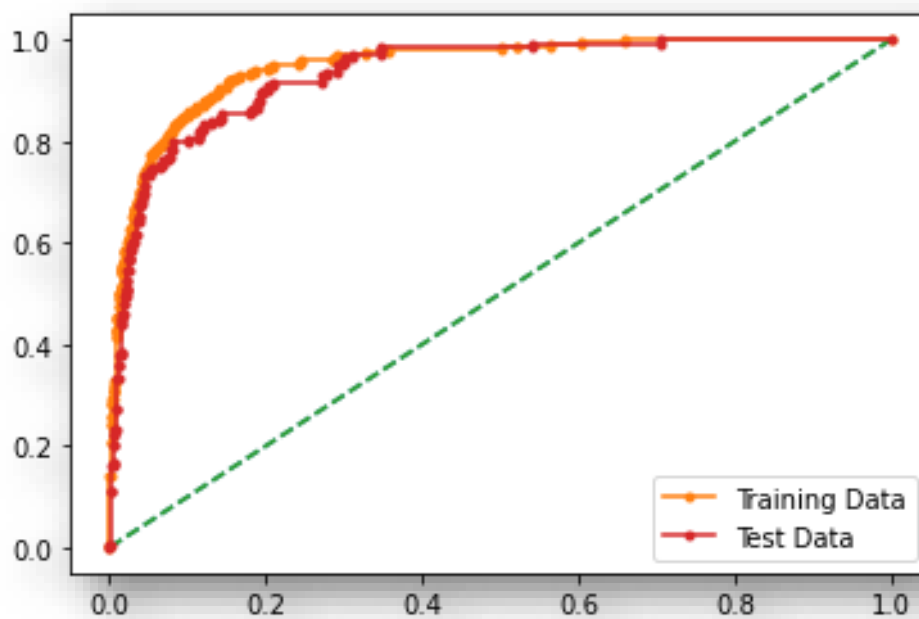

- **In this model recall is very low with 48%. It is not a good model.**

1.12 Compare the performances of Logistics, Radom Forest and LDA models (include ROC Curve).

TABLE 6: MODEL COMPARISION

MODELS	DATA	ACCURACY	PRECISION	RECALL	F1-SCORE	AUC
RANDOM FOREST	TRAIN	97	93	76	84	99
	TEST	97	94	84	89	98
LINEAR DISCRIMINANT ANALYSIS	TRAIN	94	84	48	61	95
	TEST	92	79	45	58	93
LOGISTIC REGRESSION	TRAIN	95	86	64	95	96
	TEST	95	82	69	75	96

- **Random forest with grid search performed well with highest recall and good f1 score. Roc Curve shows it's not unfitting or overfitting. While comparing other models, it is observed that Random Forest is best model for credit risk analysis with accuracy of 97%.**

CONCLUSION

- Credit report analysis provides information on the credit worthiness of a potential customer The model with selected features will predict a relatively high probability of default. Next step is to integrate with classification model where defaulters further classified into "very high risk", "high risk", "medium risk", "low risk", etc.

MARKET RISK ANALYSIS – PROBLEM STATEMENT

- The dataset contains 6 years of information (weekly stock information) on the stock prices of 10 different Indian Stocks. Calculate the mean and standard deviation on the stock returns and share insights.

PROBLEM OBJECTIVE:

- The Objective of the report is to explore the Market risk dataset in Python (JUPYTER NOTEBOOK) and generate insights about the dataset. This exploration report will consist of the following:

1. Importing the dataset in Jupyter notebook.

2. Understanding the structure of dataset.

3. Exploratory Data analysis

4. Graphical exploration

5. Calculate the mean and standard deviation on the stock returns

6. Insights from the dataset

TABLE 7: DATASET SAMPLE

	Date	Infosys	Indian_Hotel	Mahindra_&_Mahindra	Axis_Bank	SAIL	Shree_Cement	Sun_Pharma	Jindal_Steel	Idea_Vodafone	Jet_Airways
0	31-03-2014	264	69	455	263	68	5543	555	298	83	278
1	07-04-2014	257	68	458	276	70	5728	610	279	84	303
2	14-04-2014	254	68	454	270	68	5649	607	279	83	280
3	21-04-2014	253	68	488	283	68	5692	604	274	83	282
4	28-04-2014	256	65	482	282	63	5582	611	238	79	243

TABLE 8: SHAPE OF THE DATASET

```
The number of rows 314
The number of columns 11
```

TABLE 9: INFORMATION OF THE DATASET

```
RangeIndex: 314 entries, 0 to 313
Data columns (total 11 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Date                314 non-null    object
 1   Infosys             314 non-null    int64
 2   Indian_Hotel        314 non-null    int64
 3   Mahindra_&_Mahindra 314 non-null    int64
 4   Axis_Bank           314 non-null    int64
 5   SAIL                314 non-null    int64
 6   Shree_Cement        314 non-null    int64
 7   Sun_Pharma          314 non-null    int64
 8   Jindal_Steel        314 non-null    int64
 9   Idea_Vodafone       314 non-null    int64
 10  Jet_Airways         314 non-null    int64
dtypes: int64(10), object(1)
```

TABLE 10: DESCRIPTION OF THE DATASET

	Infosys	Indian_Hotel	Mahindra_&_Mahindra	Axis_Bank	SAIL	Shree_Cement	Sun_Pharma	Jindal_Steel	Idea_Vodafone	Jet_Airways
count	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000	314.000000
mean	511.340764	114.560510	636.678344	540.742038	59.095541	14806.410828	633.468153	147.627389	53.713376	372.659236
std	135.952051	22.509732	102.879975	115.835569	15.810493	4288.275085	171.855893	65.879195	31.248985	202.262668
min	234.000000	64.000000	284.000000	263.000000	21.000000	5543.000000	338.000000	53.000000	3.000000	14.000000
25%	424.000000	96.000000	572.000000	470.500000	47.000000	10952.250000	478.500000	88.250000	25.250000	243.250000
50%	466.500000	115.000000	625.000000	528.000000	57.000000	16018.500000	614.000000	142.500000	53.000000	376.000000
75%	630.750000	134.000000	678.000000	605.250000	71.750000	17773.250000	785.000000	182.750000	82.000000	534.000000
max	810.000000	157.000000	956.000000	808.000000	104.000000	24806.000000	1089.000000	338.000000	117.000000	871.000000

- From the above we can infer the following;
1. There are 314 rows and 11 columns in the dataset.
2. There no missing values in the dataset.
3. Shree Cements have the highest stock value.
4. SAIL company have the low stock price
5. There are 10 integer datatypes and object data type.

Q.2.1 Draw Stock Price Graph (Stock Price vs Time) for any 2 given stocks with inference.

Year vs Infosys
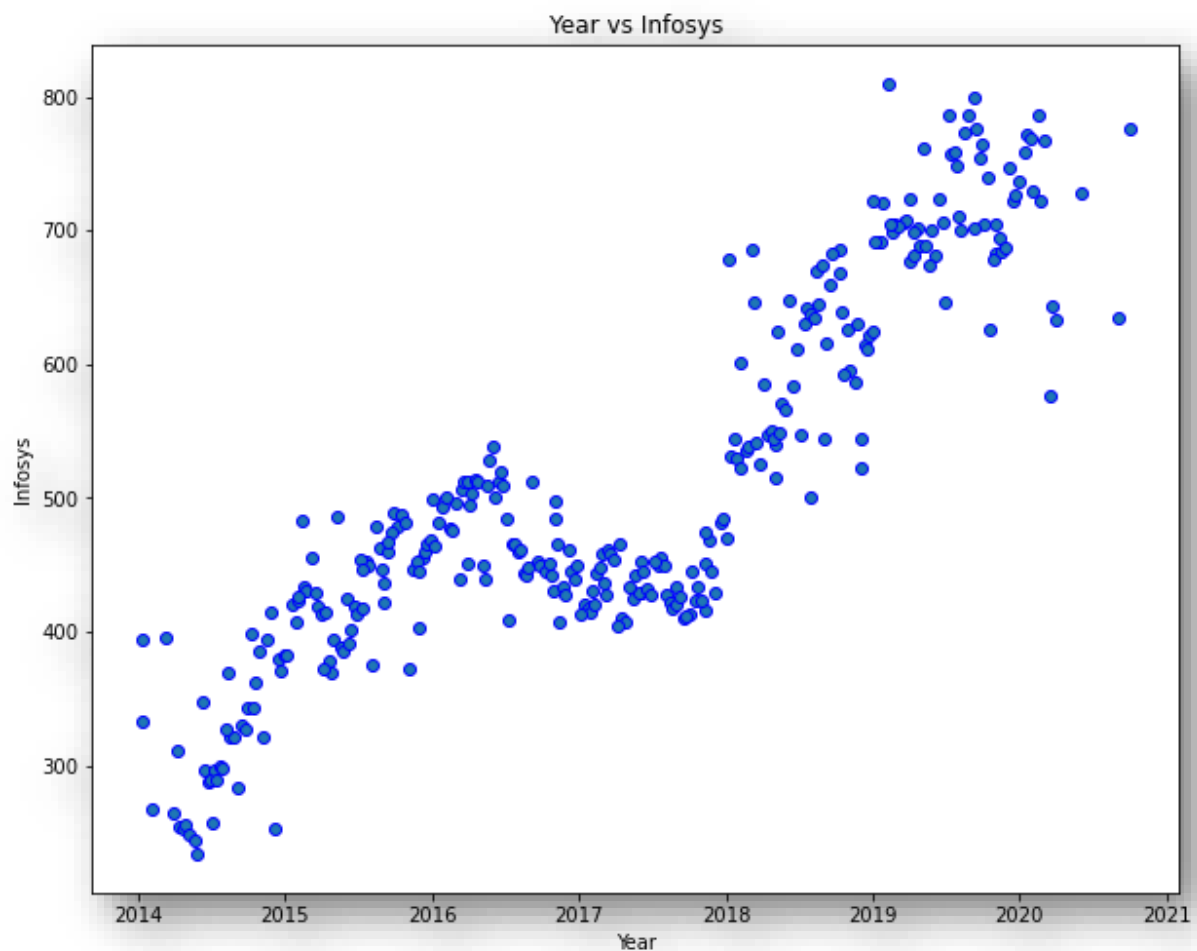
FIGURE 6: SCATTER PLOT - PRICE GRAPH – MAHINDRA & MAHINDRA
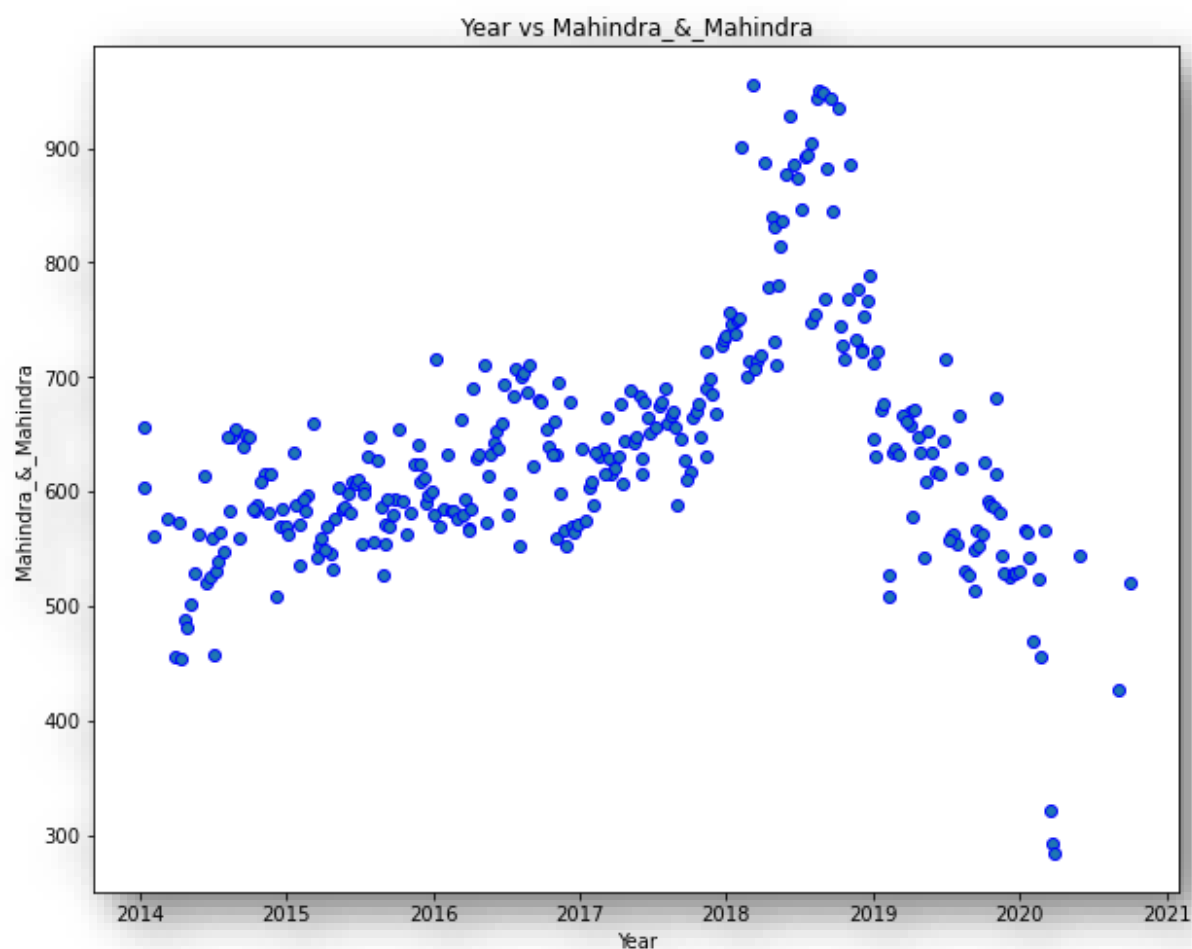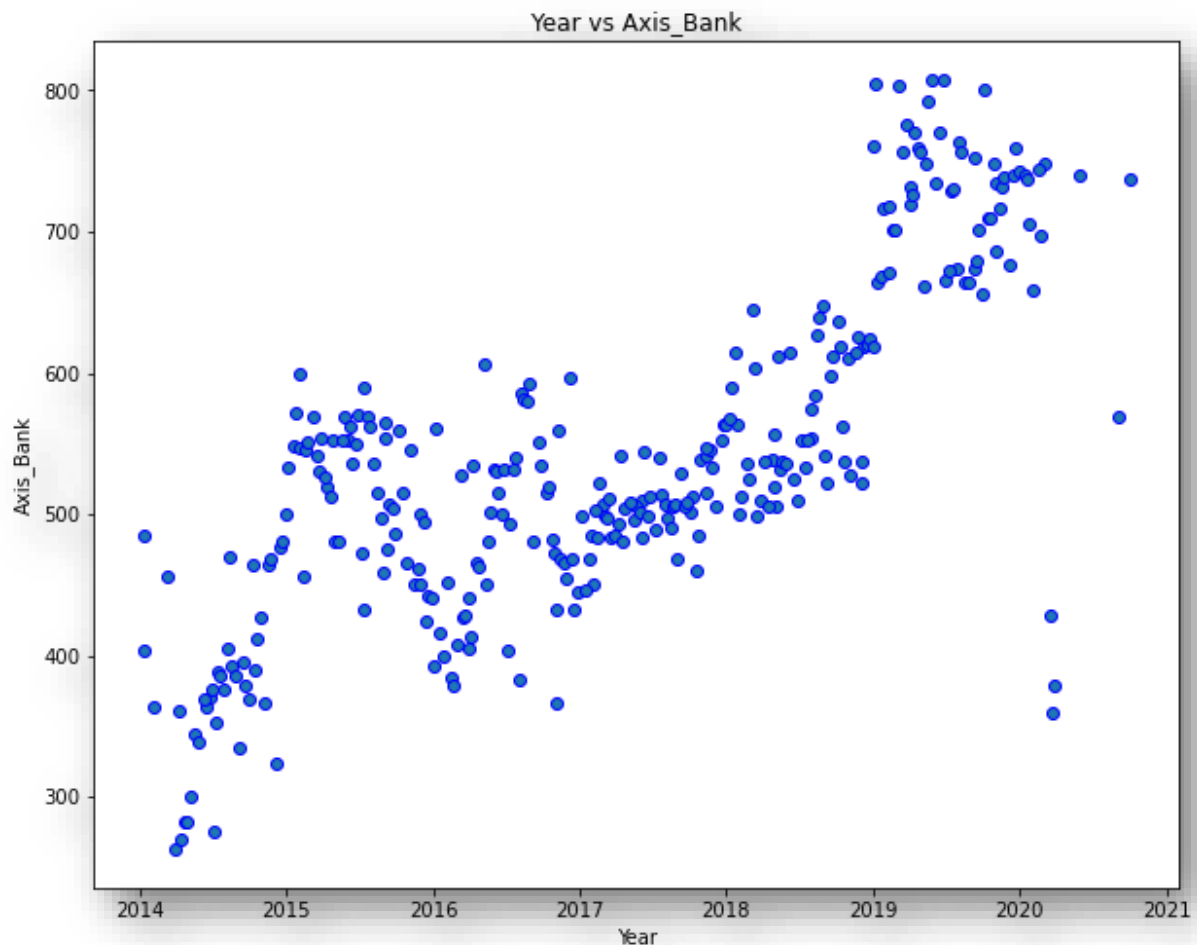

Year vs Mahindra_&_Mahindra

FIGURE 7: SCATTER PLOT - PRICE GRAPH – AXIS BANK

Year vs Axis_Bank

INFERENCE FROM THE PRICE GRAPHS

1. For the first analysis of the price and time chosen stock is Infosys, shows increasing trend for 2014 to 2016, then started consolidating from 2016 to 2018 and then again showed upward trend from 2018 till 2021.

2. For M&M we see the gradual increase from the start of year till its peak (more than 900) of price in the year 2018 to 2019, afterwards started down fall till 2021.

3. For Axis bank there is up and down in every year but there is steady growth in the price from start of year till the last year 2021.

Q.2.2. Calculate Returns for all stocks with inference.

TABLE 11: LOGARITHMIC DIFFERENCE

	Infosys	Indian_Hotel	Mahindra_&_Mahindra	Axis_Bank	SAIL	Shree_Cement	Sun_Pharma	Jindal_Steel	Idea_Vodafone	Jet_Airways
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	-0.026873	-0.014599	0.006572	0.048247	0.028988	0.032831	0.094491	-0.065882	0.011976	0.086112
2	-0.011742	0.000000	-0.008772	-0.021979	-0.028988	-0.013888	-0.004930	0.000000	-0.011976	-0.078943
3	-0.003945	0.000000	0.072218	0.047025	0.000000	0.007583	-0.004955	-0.018084	0.000000	0.007117
4	0.011788	-0.045120	-0.012371	-0.003540	-0.076373	-0.019515	0.011523	-0.140857	-0.049393	-0.148846

Q.2.3 Calculate Stock Means and Standard Deviation for all stocks with inference.

TABLE 12: STOCK MEANS

```
Shree_Cement          0.003681
Infosys               0.002794
Axis_Bank             0.001167
Indian_Hotel          0.000266
Sun_Pharma           -0.001455
Mahindra_&_Mahindra  -0.001506
SAIL                 -0.003463
Jindal_Steel         -0.004123
Jet_Airways          -0.009548
Idea_Vodafone        -0.010608
dtype: float64
```

```
Idea_Vodafone            0.104315
Jet_Airways              0.097972
Jindal_Steel             0.075108
SAIL                     0.062188
Indian_Hotel             0.047131
Axis_Bank                0.045828
Sun_Pharma               0.045033
Mahindra_&_Mahindra      0.040169
Shree_Cement             0.039917
Infosys                  0.035070
dtype: float64
```

- **The highest risky stock is idea Vodafone and on the other hand Infosys has the least risk factors, so best in the stock list to invest.**

Q.2.4 Draw a plot of Stock Means vs Standard Deviation and state your inference.

FIGURE 8: SCATTER PLOT- GRAPH PLOT- STOCK MEANS Vs STANDARD DEVIATION
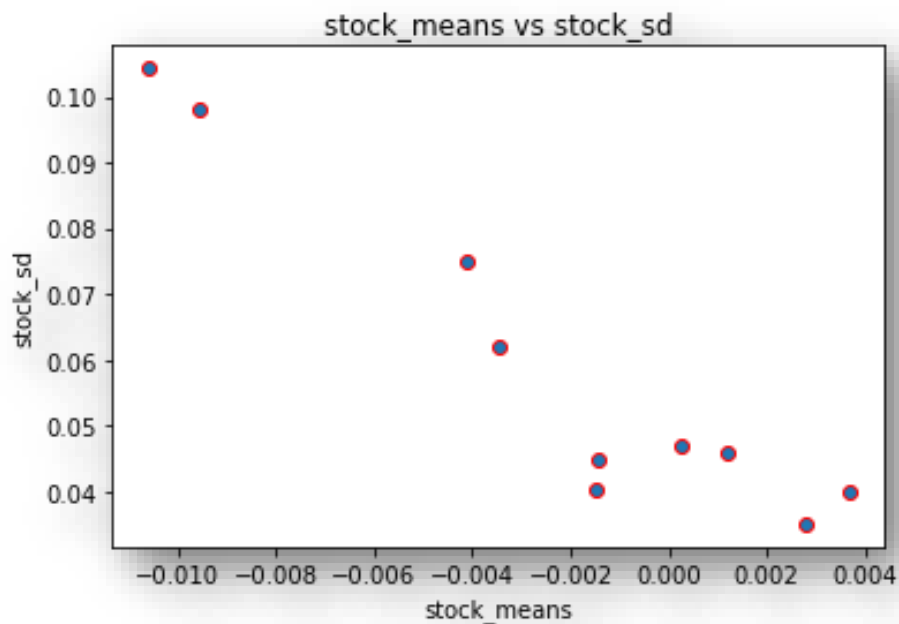
- **Stocks higher up but on the far left indicate high volatility and low returns, while the stocks on the bottom right indicate low volatility and high returns. During the investment, this graph is very useful in analysing the risk from different companies.**

Q.2.5. CONCLUSION AND RECOMMENDATIONS.

1. Stocks with lower mean and higher SD do not play all the roles in selecting the stocks for portfolio that has competing stock with more return and less risk. Therefore, from pure returns perspective, Shree cement followed by Infosys and axis bank looks good in this dataset.

2. From pure risk perspective as measured by SD, Infosys followed by Shree cement and Mahindra & Mahindra looks good in this dataset.

3. We recommend using the stock means vs standard deviation plot to assess the risk to reward ratio. More volatile stock might give short term gains but might not be a good stock for investment for long term. Whereas low volatile stock might not be a good investment in short term, but might be good return in long term.

4. Hence based on the type of investment that one is looking for an inference should be made from the above-mentioned plot.

5. Strongly recommended not to invest in shares like Idea Vodafone, Sun Pharma and Jet Airways, as these stocks have been on steady descent over the years due to various losses. They have low stock price mean and high risks value.