



FINANCE & RETAIL ANALYTICS

By Kanchan Narsinghani

Contents

Problem definition	5
Data Dictionary	5
Top 5 Rows	8
Bottom 5 rows.....	8
Check shape, Data types, and statistical summary	8
Statistical Summary.....	10
Univariate analysis	12
Multivariate analysis	35
Key meaningful observations on individual variables and the relationship between variables	38
Missing Values Detection	39
Missing Value Treatment.....	40
Outlier Detection	41
Logistic Regression	47
Linear Discriminant Analysis	84
Random Forest Classifier.....	86
Check the most important features in the final model and draw inferences	88
Metrics of Choice (Justify the evaluation metrics).....	89
Actionable insights and recommendations.....	89
PART B	90
Top 5 Rows	90
Shape of data	90
Column Type	90
Statistical Summary.....	91
Draw a Stock Price Graph (Stock Price vs Time) for the given stocks - Write observations.....	92
Calculate Returns for all stocks	97
Calculate the Mean and Standard Deviation for the returns of all stocks	98
Draw a plot of Mean vs Standard Deviation for all stock returns	99
Write observations and inferences	99
 Figure 1 - Top 5 Rows	 8
Figure 2- Bottom 5 rows.....	8

Figure 3 Column Types	10
Figure 4 Statistical Summary	11
Figure 5 - Change in Stock & Total Expenses.....	14
Figure 6 Profit after tax and PBDITA.....	15
Figure 7 PBT and Cash Profit	16
Figure 8 PDITA to total income and PBT to total income	17
Figure 9 PAT to total income and Cash Profit to total income	18
Figure 10 PAT to net worth and Sales.....	19
Figure 11 Income from financial Services and Other income	20
Figure 12 Total capital and Reserves & Funds.....	21
Figure 13 Borrowings & Current Liabilities and Provisions	22
Figure 14 Deferred tax liability.....	22
Figure 15 Shareholders funds and Cumulative retained profits	23
Figure 16 Capital employed and TOL/TNW.....	24
Figure 17 Total term liabilities / tangible net worth	24
Figure 18 Contingent liabilities / Net worth and Contingent liabilities.....	25
Figure 19 Net fixed assets	25
Figure 20 Investments and Current assets.....	26
Figure 21 Net working capital and Quick Ratio	27
Figure 22 Current ratio and Debt to equity ratio	28
Figure 23 Cash to current liabilities and Cash to average cost of sales per day.....	29
Figure 24 Creditors Turnover and Debtors turnover.....	30
Figure 25 Finished goods turnover and WIP turnover	31
Figure 26 Raw material turnover and Shares outstanding	32
Figure 27 Equity face value and EPS.....	33
Figure 28 Adjusted EPS and Total liabilities.....	34
Figure 29 PE on BSE.....	34
Figure 30 Heatmap.....	35
Figure 31 Missing Value Detection.....	39
Figure 32 Missing Value Treatment.....	40
Figure 33 Column Outliers	41
Figure 34 Outlier Treatment.....	43
Figure 35 - Missing Values Heatmap	44
Figure 36 Missing Value RowWise	44
Figure 37 Missing Value (Percentage)	45
Figure 38 VIF values	48
Figure 39 Post Total Assets Removal.....	49
Figure 40 Post Sales Removal.....	49
Figure 41 Post Total Income Removal	50
Figure 42 Post Net_worth Removal	50
Figure 43 Post Total_liabilities Removal.....	51
Figure 44 Post PBT Removal.....	51
Figure 45 Post Cash_profit removal	52
Figure 46 Post Capital_employed Removal.....	52

Figure 47 Post PBT_to_total_income Removal	53
Figure 48 Post Current_assets Removal.....	53
Figure 49 Post PBDITA Removal	54
Figure 50 Post Shareholders_funds Removal.....	54
Figure 51 Cash_profit_to_total_income Removal	55
Figure 52 Post Debt_to_equity_ratio Removal	55
Figure 53 Post Reserves_and_funds removal	56
Figure 54 Post Total_expenses Removal	56
Figure 55 Model 2 Summary	58
Figure 56 Model 3 Summary	59
Figure 57 Model 4 Summary	60
Figure 58 Model 5 Summary	61
Figure 59 Model 6 Summary	62
Figure 60 Model 7 Summary	63
Figure 61 Model 8 Summary	64
Figure 62 Model 9 Summary	65
Figure 63 Model 10 Summary	66
Figure 64 Model 11 Summary	67
Figure 65 Model 12 Summary	68
Figure 66 Model 13 Summary	69
Figure 67 Model 14 Summary	70
Figure 68 Model 15 Summary	71
Figure 69 Model 16 Summary	72
Figure 70 Model 17 Summary	73
Figure 71 Model 18 Summary	74
Figure 72 Model 19 Summary	75
Figure 73 Model 20 Summary	76
Figure 74 Model 21 Summary	77
Figure 75 Model 22 Summary	78
Figure 76 Model 23 Summary	79
Figure 77 Model 24 Summary	80
Figure 78 Confusion Matrix.....	81
Figure 79 Confusion Matrix after selecting Optimum threshold	82
Figure 80 Classification Report on train data	82
Figure 81 Confusion Matrix on test data.....	83
Figure 82 Classification Report on test data	83
Figure 83 LDA Classification Report on train data	84
Figure 84 LDA Classification Report on test data	84
Figure 85 LDA CONfusion matrix on train data.....	85
Figure 86 LDA Classification report on train data	85
Figure 87 Confusion Matrix on test data.....	86
Figure 88 LDA classification report on test data	86
Figure 89 Random Forest Classification Report on train data.....	87
Figure 90 Random Forest Test data.....	87

Figure 91 Top 5 Rows	90
Figure 92 Column Types	90
Figure 93 Statistical Summary	91
Figure 94 ITC_Limited.....	92
Figure 95 Bharti_Airtel	93
Figure 96 TATA_Motors	94
Figure 97 DLF_Limited.....	95
Figure 98 YES_Bank.....	96
Figure 99 Stock Returns	97
Figure 100 Means of Stocks	98
Figure 101 Standard deviation of stocks	98
Figure 102 Mean vs Standard deviation of stocks	99

Problem definition

In the realm of modern finance, businesses encounter the perpetual challenge of managing debt obligations effectively to maintain a favorable credit standing and foster sustainable growth. Investors keenly scrutinize companies capable of navigating financial complexities while ensuring stability and profitability. A pivotal instrument in this evaluation process is the balance sheet, which provides a comprehensive overview of a company's assets, liabilities, and shareholder equity, offering insights into its financial health and operational efficiency. In this context, leveraging available financial data, particularly from preceding fiscal periods, becomes imperative for informed decision-making and strategic planning.

Objective

A group of venture capitalists want to develop a Financial Health Assessment Tool. With the help of the tool, it endeavors to empower businesses and investors with a robust mechanism for evaluating the financial well-being and creditworthiness of companies. By harnessing machine learning techniques, they aim to analyze historical financial statements and extract pertinent insights to facilitate informed decision-making via the tool. Specifically, they foresee facilitating the following with the help of the tool:

Debt Management Analysis: Identify patterns and trends in debt management practices to assess the ability of businesses to fulfill financial obligations promptly and efficiently, and identify potential cases of default.

Credit Risk Evaluation: Evaluate credit risk exposure by analyzing liquidity ratios, debt-to-equity ratios, and other key financial indicators to ascertain the likelihood of default and inform investment decisions.

They have hired you as a data scientist and provided you with the financial metrics of different companies. The task is to analyze the data provided and develop a predictive model leveraging machine learning techniques to identify whether a given company will be tagged as a defaulter in terms of net worth next year. The predictive model will help the organization anticipate potential challenges with the financial performance of the companies and enable proactive risk mitigation strategies.

Data Dictionary

The data consists of financial metrics from the balance sheets of different companies. The detailed data dictionary is given below.

Networth Next Year: Net worth of the customer in the next year

Total assets: Total assets of customer

Net worth: Net worth of the customer of the present year

Total income: Total income of the customer

Change in stock: Difference between the current value of the stock and the value of stock in the last trading day

Total expenses: Total expenses done by the customer

Profit after tax: Profit after tax deduction

PBDITA: Profit before depreciation, income tax, and amortization

PBT: Profit before tax deduction

Cash profit: Total Cash profit

PBDITA as % of total income: $\text{PBDITA} / \text{Total income}$

PBT as % of total income: $\text{PBT} / \text{Total income}$

PAT as % of total income: $\text{PAT} / \text{Total income}$

Cash profit as % of total income: $\text{Cash Profit} / \text{Total income}$

PAT as % of net worth: $\text{PAT} / \text{Net worth}$

Sales: Sales done by the customer

Income from financial services: Income from financial services

Other income: Income from other sources

Total capital: Total capital of the customer

Reserves and funds: Total reserves and funds of the customer

Borrowings: Total amount borrowed by the customer

Current liabilities & provisions: current liabilities of the customer

Deferred tax liability: Future income tax customer will pay because of the current transaction

Shareholders funds: Amount of equity in a company which belongs to shareholders

Cumulative retained profits: Total cumulative profit retained by customer

Capital employed: Current asset minus current liabilities

TOL/TNW: Total liabilities of the customer divided by Total net worth

Total term liabilities / tangible net worth: Short + long term liabilities divided by tangible net worth

Contingent liabilities / Net worth (%): $\text{Contingent liabilities} / \text{Net worth}$

Contingent liabilities: Liabilities because of uncertain events

Net fixed assets: The purchase price of all fixed assets

Investments: Total invested amount

Current assets: Assets that are expected to be converted to cash within a year

Net working capital: Difference between the current liabilities and current assets

Quick ratio (times): Total cash divided by current liabilities

Current ratio (times): Current assets divided by current liabilities

Debt to equity ratio (times): Total liabilities divided by its shareholder equity

Cash to current liabilities (times): Total liquid cash divided by current liabilities

Cash to average cost of sales per day: Total cash divided by the average cost of the sales

Creditors turnover: Net credit purchase divided by average trade creditors

Debtors turnover: Net credit sales divided by average accounts receivable

Finished goods turnover: Annual sales divided by average inventory

WIP turnover: The cost of goods sold for a period divided by the average inventory for that period

Raw material turnover: Cost of goods sold is divided by the average inventory for the same period

Shares outstanding: Number of issued shares minus the number of shares held in the company

Equity face value: cost of the equity at the time of issuing

EPS: Net income divided by the total number of outstanding share

Adjusted EPS: Adjusted net earnings divided by the weighted average number of common shares outstanding on a diluted basis during the plan year

Total liabilities: Sum of all types of liabilities

PE on BSE: Company's current stock price divided by its earnings per share

Note: A company will not be tagged as a defaulter if its net worth next year is positive, or else, it'll be tagged as a defaulter.

Top 5 Rows

	Num	Networth Next Year	Total assets	Net worth	Total income	Change in stock	Total expenses	Profit after tax	PBDITA	PBT	...	Debtors turnover	Finished goods turnover	WIP turnover	Raw material turnover	Shares outstanding	Equity face value	EPS
0	1	395.3	827.6	336.5	534.1	13.5	508.7	38.9	124.4	64.6	...	5.65	3.99	3.37	14.87	8760056.0	10.0	4.44
1	2	36.2	67.7	24.3	137.9	-3.7	131.0	3.2	5.5	1.0	...	NaN	NaN	NaN	NaN	NaN	NaN	0.00
2	3	84.0	238.4	78.9	331.2	-18.1	309.2	3.9	25.8	10.5	...	2.51	17.67	8.76	8.35	NaN	NaN	0.00
3	4	2041.4	6883.5	1443.3	8448.5	212.2	8482.4	178.3	418.4	185.1	...	1.91	18.14	18.62	11.11	10000000.0	10.0	17.60
4	5	41.8	90.9	47.0	388.6	3.4	392.7	-0.7	7.2	-0.6	...	68.00	45.87	28.67	19.93	107315.0	100.0	-6.52

Figure 1 - Top 5 Rows

Bottom 5 rows

	Num	Networth Next Year	Total assets	Net worth	Total income	Change in stock	Total expenses	Profit after tax	PBDITA	PBT	...	Debtors turnover	Finished goods turnover	WIP turnover	Raw material turnover	Shares outstanding	Equity face value	EPS
4251	4252	0.2	0.4	0.2	NaN	NaN	NaN	NaN	NaN	NaN	...	0.00	NaN	NaN	0.00	NaN	NaN	0.00
4252	4253	93.3	159.6	86.7	172.9	0.1	169.7	3.3	18.4	3.7	...	1.80	11.00	8.28	9.88	8162700.0	10.0	0.42
4253	4254	932.2	833.8	664.6	2314.7	32.1	2151.6	195.2	348.4	303.0	...	6.08	59.28	31.14	9.87	7479762.0	10.0	26.58
4254	4255	64.6	95.0	48.5	110.5	4.6	113.5	1.6	9.7	2.6	...	3.71	78.99	11.51	14.95	NaN	NaN	0.00
4255	4256	0.0	384.6	111.3	345.8	11.3	341.7	15.4	57.6	20.7	...	4.71	53.37	8.33	3.74	960000.0	10.0	15.63

Figure 2- Bottom 5 rows

Check shape, Data types, and statistical summary

The number of rows (observations) is 4256

The number of columns (variables) is 51.

Details of column types in dataframe are as below:

0	Num	4256 non-null	int64
1	Networth Next Year	4256 non-null	float64
2	Total assets	4256 non-null	float64
3	Net worth	4256 non-null	float64
4	Total income	4025 non-null	float64
5	Change in stock	3706 non-null	float64
6	Total expenses	4091 non-null	float64
7	Profit after tax	4102 non-null	float64
8	PBDITA	4102 non-null	float64
9	PBT	4102 non-null	float64
10	Cash profit	4102 non-null	float64
11	PBDITA to total income	4177 non-null	float64
12	PBT to total income	4177 non-null	float64
13	PAT to total income	4177 non-null	float64
14	Cash profit to total income	4177 non-null	float64
15	PAT to net worth	4256 non-null	float64
16	Sales	3951 non-null	float64
17	Income from fincial services	3145 non-null	float64
18	Other income	2700 non-null	float64
19	Total capital	4251 non-null	float64
20	Reserves and funds	4158 non-null	float64
21	Borrowings	3825 non-null	float64
22	Current liabilities & provisions	4146 non-null	float64
23	Deferred tax liability	2887 non-null	float64
24	Shareholders funds	4256 non-null	float64
25	Cumulative retained profits	4211 non-null	float64
26	Capital employed	4256 non-null	float64
27	TOL/TNW	4256 non-null	float64
28	Total term liabilities / tangible net worth	4256 non-null	float64
29	Contingent liabilities / Net worth (%)	4256 non-null	float64
30	Contingent liabilities	2854 non-null	float64
31	Net fixed assets	4124 non-null	float64
32	Investments	2541 non-null	float64
33	Current assets	4176 non-null	float64
34	Net working capital	4219 non-null	float64
35	Quick ratio (times)	4151 non-null	float64
36	Current ratio (times)	4151 non-null	float64
37	Debt to equity ratio (times)	4256 non-null	float64
38	Cash to current liabilities (times)	4151 non-null	float64
39	Cash to average cost of sales per day	4156 non-null	float64
40	Creditors turnover	3865 non-null	float64
41	Debtors turnover	3871 non-null	float64

42	Finished goods turnover	3382 non-null	float64
43	WIP turnover	3492 non-null	float64
44	Raw material turnover	3828 non-null	float64
45	Shares outstanding	3446 non-null	float64
46	Equity face value	3446 non-null	float64
47	EPS	4256 non-null	float64
48	Adjusted EPS	4256 non-null	float64
49	Total liabilities	4256 non-null	float64
50	PE on BSE	1629 non-null	float64

Figure 3 Column Types

Statistical Summary

	count	mean	std	min	25%	50%	75%	max
Num	4256.0	2.128500e+03	1.228746e+03	1.000000e+00	1064.750	2128.500	3.192250e+03	4.256000e+03
Networth Next Year	4256.0	1.344741e+03	1.593674e+04	-7.426560e+04	3.975	72.100	3.308250e+02	8.057734e+05
Total assets	4256.0	3.573617e+03	3.007444e+04	1.000000e-01	91.300	315.500	1.120800e+03	1.176509e+06
Net worth	4256.0	1.351950e+03	1.296131e+04	0.000000e+00	31.475	104.800	3.898500e+02	6.131516e+05
Total income	4025.0	4.688190e+03	5.391895e+04	0.000000e+00	107.100	455.100	1.485000e+03	2.442828e+06
Change in stock	3706.0	4.370248e+01	4.369150e+02	-3.029400e+03	-1.800	1.600	1.840000e+01	1.418550e+04
Total expenses	4091.0	4.356301e+03	5.139809e+04	-1.000000e-01	96.800	426.800	1.395700e+03	2.366035e+06
Profit after tax	4102.0	2.950506e+02	3.079902e+03	-3.908300e+03	0.500	9.000	5.330000e+01	1.194391e+05
PBDITA	4102.0	6.059406e+02	5.646231e+03	-4.407000e+02	6.925	36.900	1.587000e+02	2.085765e+05
PBT	4102.0	4.102590e+02	4.217415e+03	-3.894800e+03	0.800	12.600	7.417500e+01	1.452926e+05
Cash profit	4102.0	4.082675e+02	4.143926e+03	-2.245700e+03	2.900	19.400	9.625000e+01	1.769118e+05
PBDITA to total income	4177.0	3.179892e+00	1.722566e+02	-6.400000e+03	4.970	9.680	1.647000e+01	1.000000e+02
PBT to total income	4177.0	-1.819683e+01	4.199111e+02	-2.134000e+04	0.560	3.340	8.940000e+00	1.000000e+02
PAT to total income	4177.0	-2.003367e+01	4.235762e+02	-2.134000e+04	0.350	2.370	6.420000e+00	1.500000e+02
Cash profit to total income	4177.0	-9.021278e+00	2.999574e+02	-1.502000e+04	2.000	5.660	1.073000e+01	1.000000e+02
PAT to net worth	4256.0	1.016786e+01	6.153240e+01	-7.487200e+02	0.000	8.040	2.020250e+01	2.466670e+03
Sales	3951.0	4.645685e+03	5.308090e+04	1.000000e-01	113.350	468.600	1.481200e+03	2.384984e+06
Income from fincial services	3145.0	8.136006e+01	1.042759e+03	0.000000e+00	0.500	1.900	9.800000e+00	5.193820e+04
Other income	2700.0	5.595289e+01	1.178415e+03	0.000000e+00	0.400	1.500	6.200000e+00	4.285670e+04
Total capital	4251.0	2.245577e+02	1.684951e+03	1.000000e-01	13.200	42.600	1.031500e+02	7.827320e+04
Reserves and funds	4158.0	1.210562e+03	1.281623e+04	-6.525900e+03	5.300	55.150	2.825250e+02	6.251378e+05
Borrowings	3825.0	1.176248e+03	8.581249e+03	1.000000e-01	24.400	99.800	3.583000e+02	2.782573e+05
Current liabilities & provisions	4146.0	9.606314e+02	9.140536e+03	1.000000e-01	17.500	70.300	2.659250e+02	3.522403e+05
Deferred tax liability	2887.0	2.344951e+02	2.106253e+03	1.000000e-01	3.200	13.500	5.130000e+01	7.279660e+04
Shareholders funds	4256.0	1.376487e+03	1.301069e+04	0.000000e+00	32.300	107.600	4.089000e+02	6.131516e+05

Cumulative retained profits	4211.0	9.371820e+02	9.853096e+03	-6.534300e+03	1.100	37.400	2.062000e+02	3.901338e+05
Capital employed	4256.0	2.433618e+03	2.049640e+04	0.000000e+00	61.300	221.200	7.903000e+02	8.914089e+05
TOL/TNW	4256.0	4.025343e+00	2.087909e+01	-3.504800e+02	0.600	1.420	2.830000e+00	4.730000e+02
Total term liabilities / tangible net worth	4256.0	1.854288e+00	1.587506e+01	-3.256000e+02	0.050	0.345	1.000000e+00	4.560000e+02
Contingent liabilities / Net worth (%)	4256.0	5.570750e+01	3.691657e+02	0.000000e+00	0.000	5.360	3.101250e+01	1.470427e+04
Contingent liabilities	2854.0	9.485522e+02	1.205674e+04	1.000000e-01	6.000	37.850	1.953250e+02	5.595068e+05
Net fixed assets	4124.0	1.209487e+03	1.250240e+04	0.000000e+00	26.200	93.850	3.528250e+02	6.366046e+05
Investments	2541.0	7.218659e+02	6.793860e+03	0.000000e+00	1.000	8.200	6.380000e+01	1.999786e+05
Current assets	4176.0	1.350360e+03	1.015557e+04	1.000000e-01	36.600	148.350	5.150000e+02	3.548152e+05
Net working capital	4219.0	1.628742e+02	3.182030e+03	-6.383900e+04	-1.100	16.700	8.650000e+01	8.578280e+04
Quick ratio (times)	4151.0	1.497355e+00	9.327519e+00	0.000000e+00	0.410	0.670	1.030000e+00	3.410000e+02
Current ratio (times)	4151.0	2.257398e+00	1.247829e+01	0.000000e+00	0.930	1.230	1.720000e+00	5.050000e+02
Debt to equity ratio (times)	4256.0	2.871563e+00	1.559997e+01	0.000000e+00	0.220	0.790	1.750000e+00	4.560000e+02
Cash to current liabilities (times)	4151.0	5.284197e-01	4.796342e+00	0.000000e+00	0.020	0.070	1.900000e-01	1.650000e+02
Cash to average cost of sales per day	4156.0	1.451579e+02	2.521992e+03	0.000000e+00	2.880	8.040	2.197000e+01	1.280408e+05
Creditors turnover	3865.0	1.681226e+01	7.567492e+01	0.000000e+00	3.720	6.170	1.169000e+01	2.401000e+03
Debtors turnover	3871.0	1.792903e+01	9.016443e+01	0.000000e+00	3.810	6.470	1.185000e+01	3.135200e+03
Finished goods turnover	3382.0	8.436999e+01	5.626374e+02	-9.000000e-02	8.190	17.320	4.001250e+01	1.794760e+04
WIP turnover	3492.0	2.868451e+01	1.696509e+02	-1.800000e-01	5.100	9.860	2.024000e+01	5.651400e+03
Raw material turnover	3828.0	1.773393e+01	3.431259e+02	-2.000000e+00	3.020	6.410	1.182250e+01	2.109200e+04
Shares outstanding	3446.0	2.376491e+07	1.709790e+08	-2.147484e+09	1308382.500	4750000.000	1.090602e+07	4.130401e+09
Equity face value	3446.0	-1.094829e+03	3.410136e+04	-9.999989e+05	10.000	10.000	1.000000e+01	1.000000e+05
EPS	4256.0	-1.962175e+02	1.306195e+04	-8.431818e+05	0.000	1.490	1.000000e+01	3.452253e+04
Adjusted EPS	4256.0	-1.975276e+02	1.306193e+04	-8.431818e+05	0.000	1.240	7.615000e+00	3.452253e+04
Total liabilities	4256.0	3.573617e+03	3.007444e+04	1.000000e-01	91.300	315.500	1.120800e+03	1.176509e+06
PE on BSE	1629.0	5.546229e+01	1.304445e+03	-1.116640e+03	2.970	8.690	1.700000e+01	5.100274e+04

Figure 4 Statistical Summary

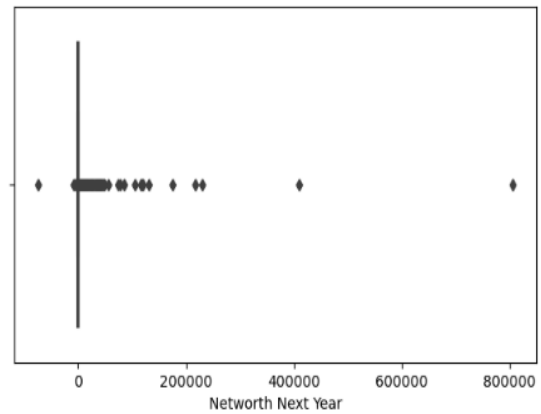
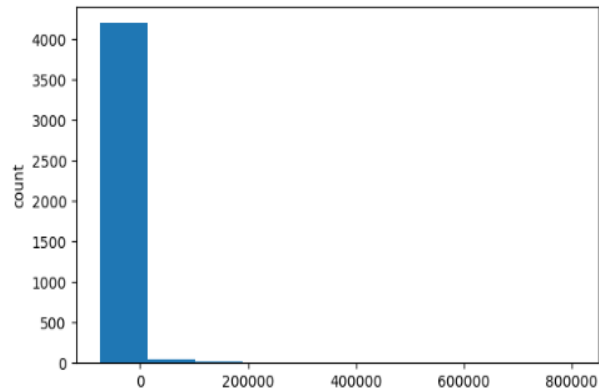
There are few missing values in Total income, change in stock, total expenses, profit after tax, PBDITA, PBT, Cash Profit, PBDITA to total income, PBT to total income, PAT to total income, Cash profit to total income, Sales, Income from financial services, Other income, Total capital, Reserves and funds, Borrowings, Current liabilities and provisions, Deferred tax liability, Cumulative retained profits, contingent liabilities, net fixed assets, investments, current assets, net working capital, quick ratio, current ratio, cash to current liabilities, cash to average cost of sales per day, creditors turnover, debtors turnover, finished goods turnover, WIP turnover, Raw materials turnover, Shares outstanding, Equity face value, PE on BSE.

We checked and there are no duplicate values.

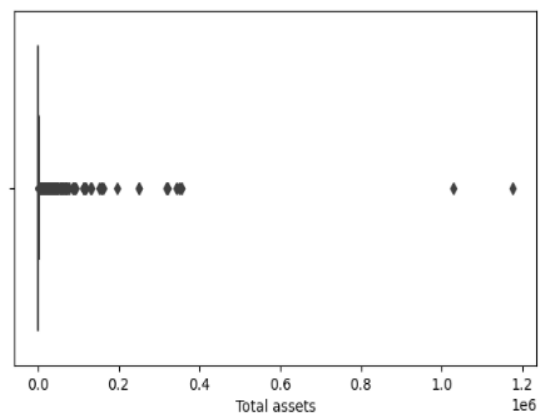
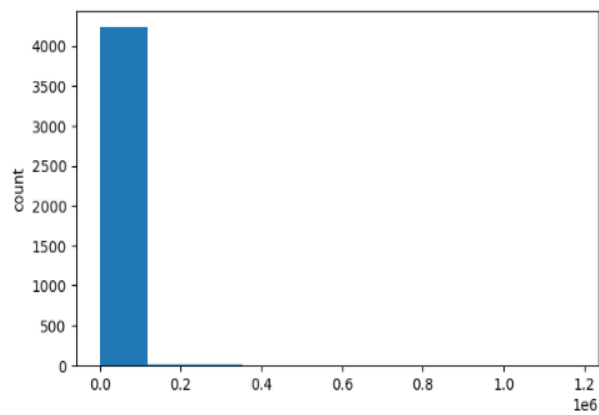
We remove special characters from columns to standardize column names. We removed following characters from column names : 'as % of', '/', ' (%) ', ' & ', ' (times)', and spaces .

Univariate analysis

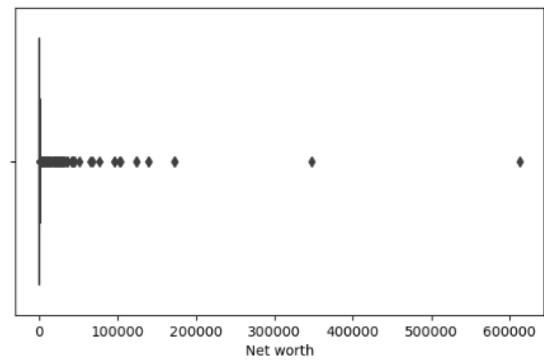
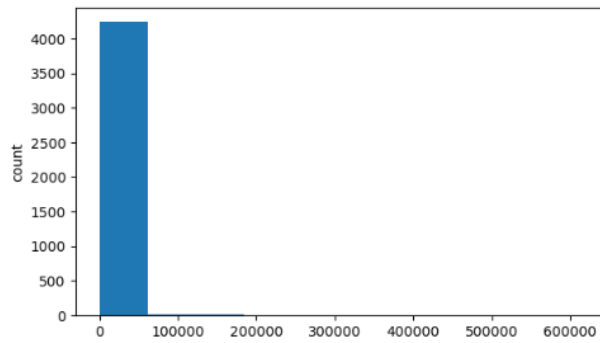
Networth Next Year
Count : 4256
Number of outliers: 624



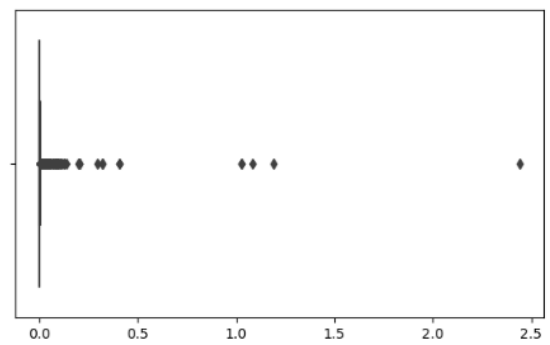
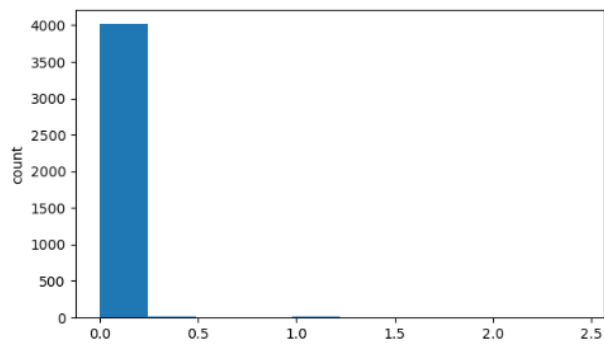
Total assets
Count : 4256
Number of outliers: 585



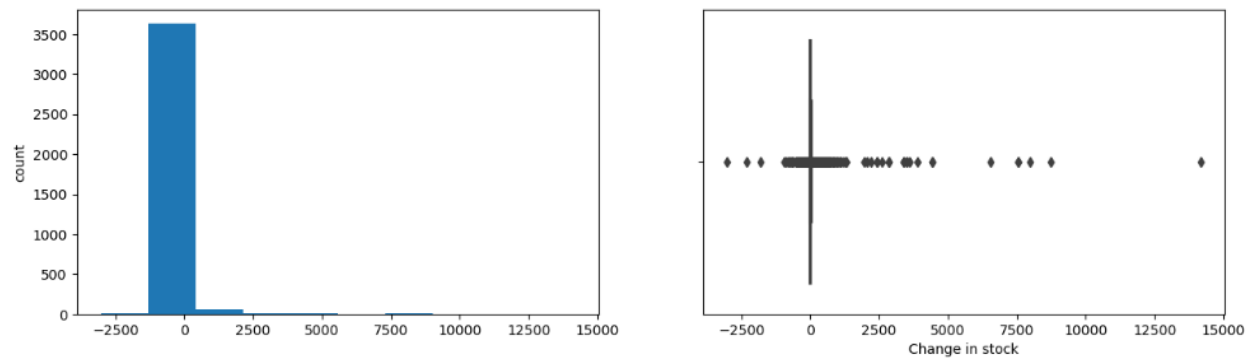
Net worth
Count : 4256
Number of outliers: 595



Total income
Count : 4025
Number of outliers: 0



Change in stock
Count : 3706
Number of outliers: 0



Total expenses
Count : 4091
Number of outliers: 0

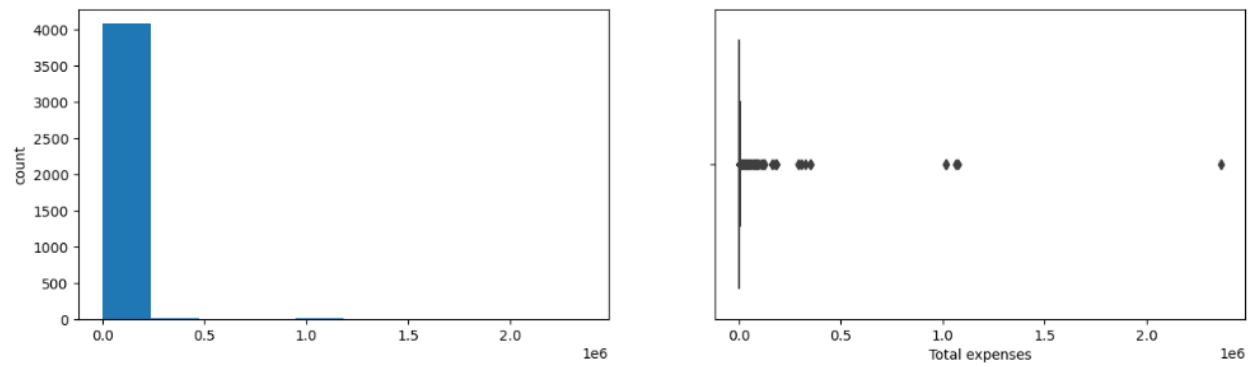
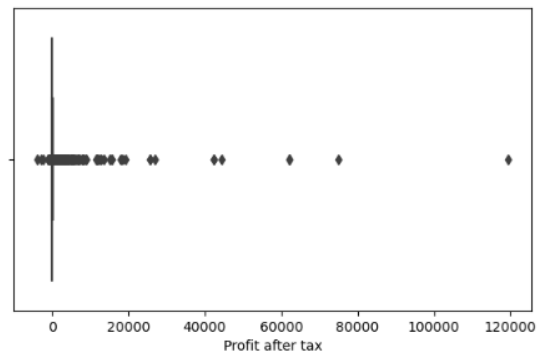
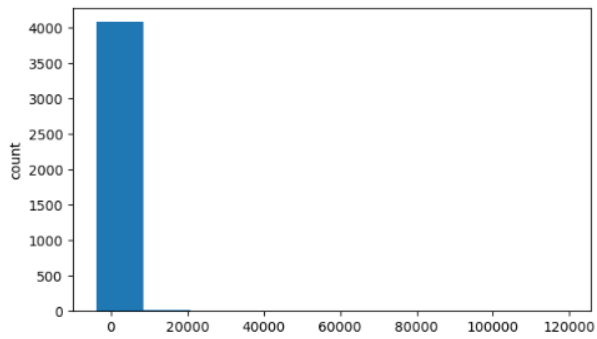


Figure 5 - Change in Stock & Total Expenses

Profit after tax
Count : 4102
Number of outliers: 0



PBDITA
Count : 4102
Number of outliers: 0

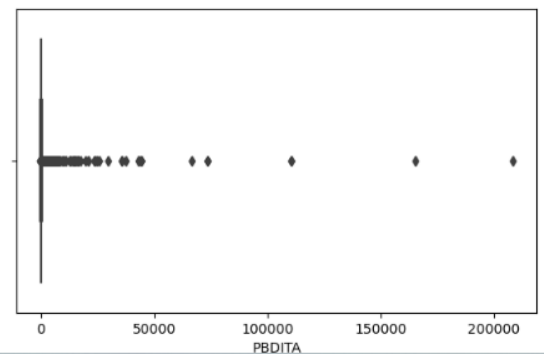
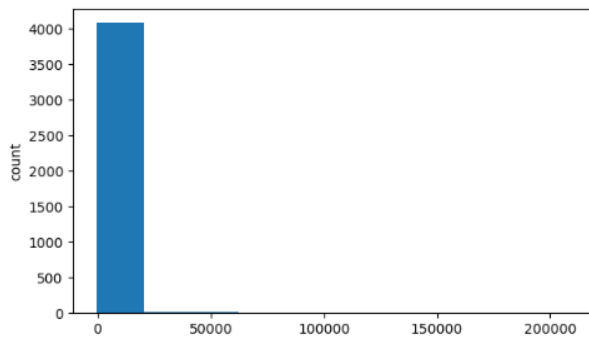
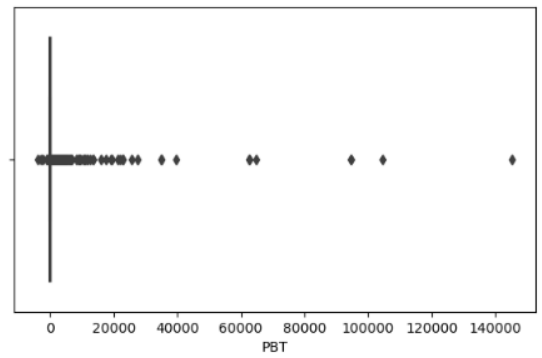
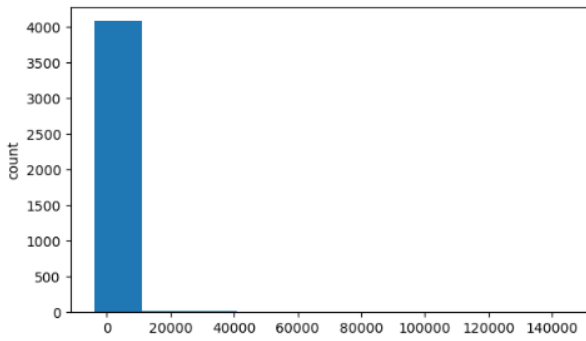


Figure 6 Profit after tax and PBDITA

PBT
Count : 4102
Number of outliers: 0



Cash profit
Count : 4102
Number of outliers: 0

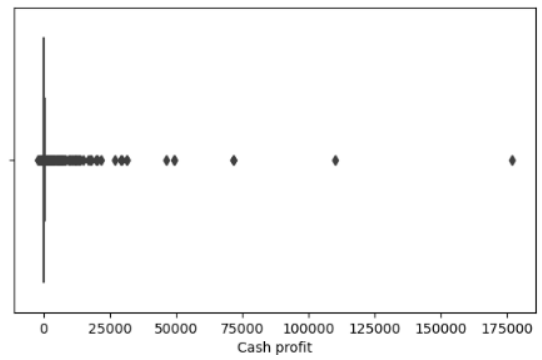
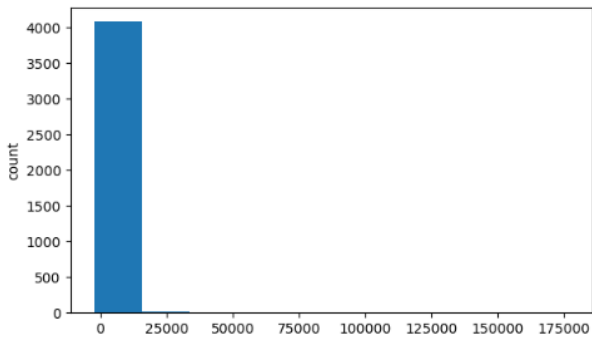
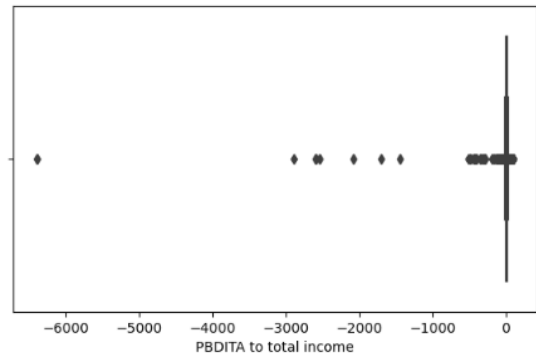
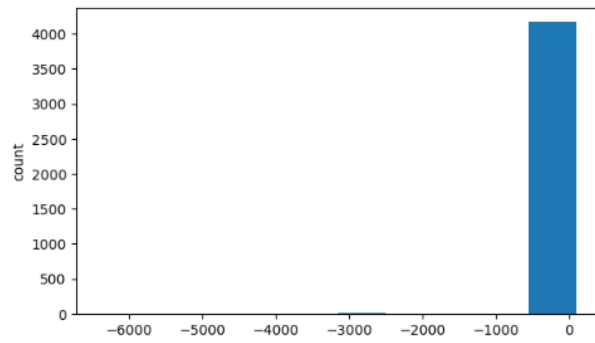


Figure 7 PBT and Cash Profit

PBDITA to total income
Count : 4177
Number of outliers: 0



PBT to total income
Count : 4177
Number of outliers: 0

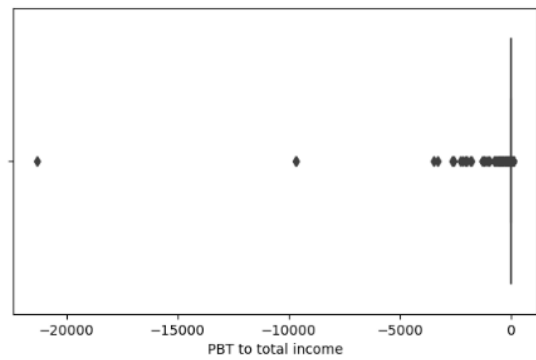
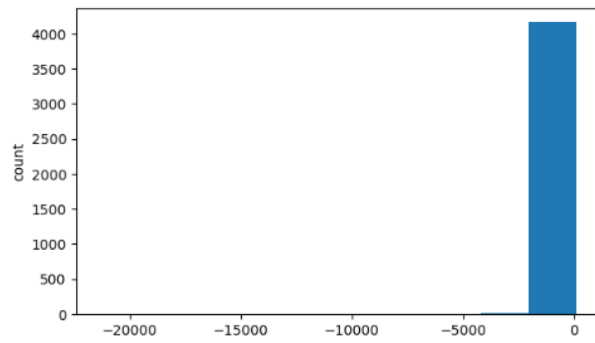
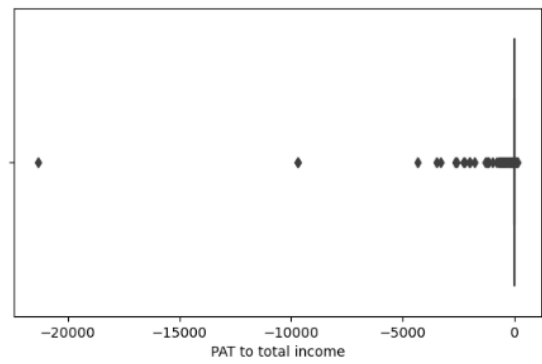
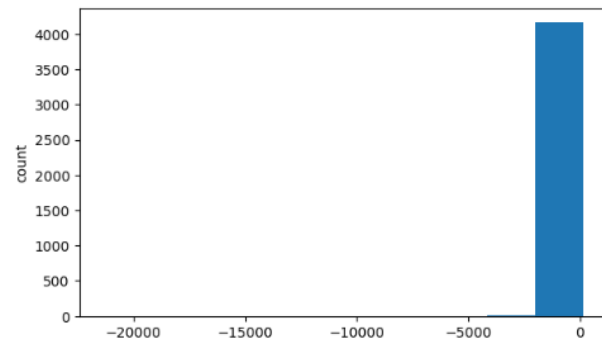


Figure 8 PDITA to total income and PBT to total income

PAT to total income
 Count : 4177
 Number of outliers: 0



Cash profit to total income
 Count : 4177
 Number of outliers: 0

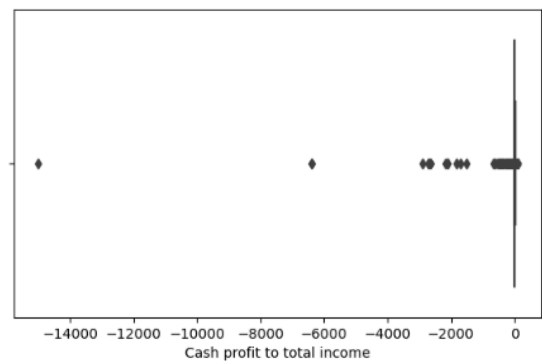
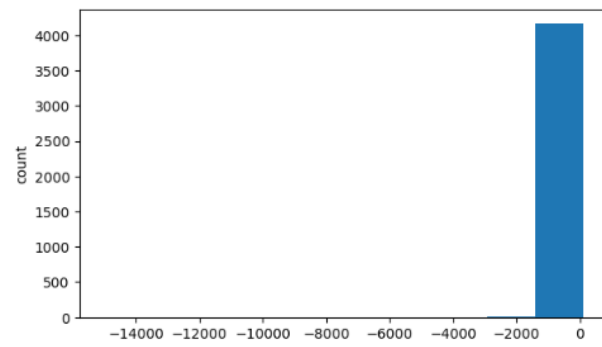
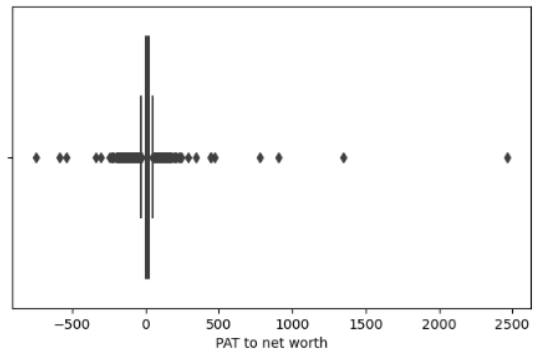
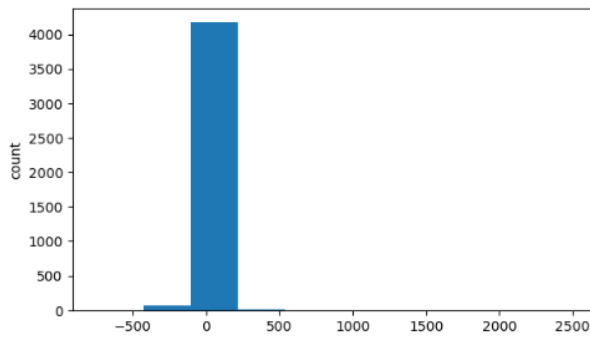


Figure 9 PAT to total income and Cash Profit to total income

PAT to net worth
Count : 4256
Number of outliers: 427



Sales
Count : 3951
Number of outliers: 0

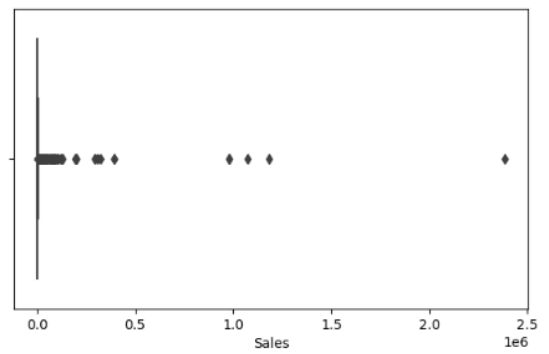
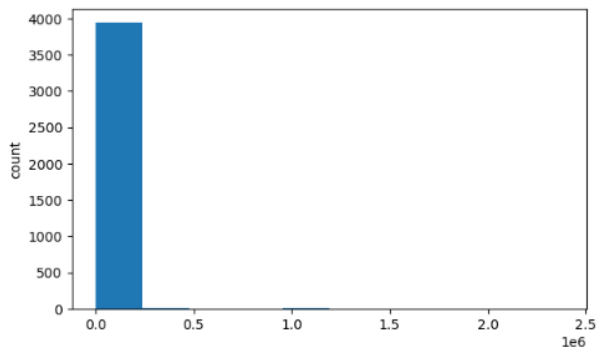
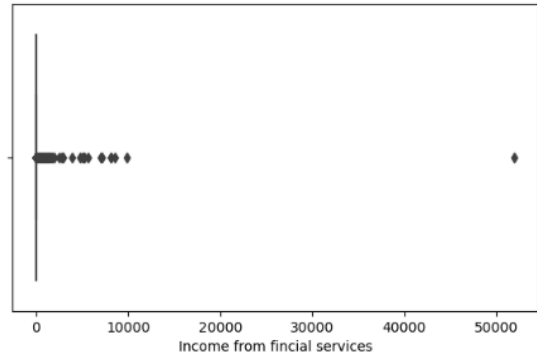
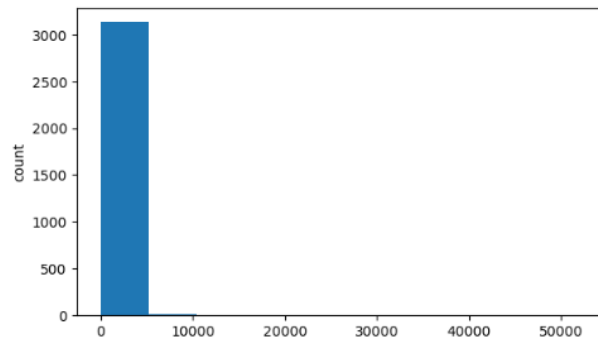


Figure 10 PAT to net worth and Sales

Income from financial services
Count : 3145
Number of outliers: 0



Other income
Count : 2700
Number of outliers: 0

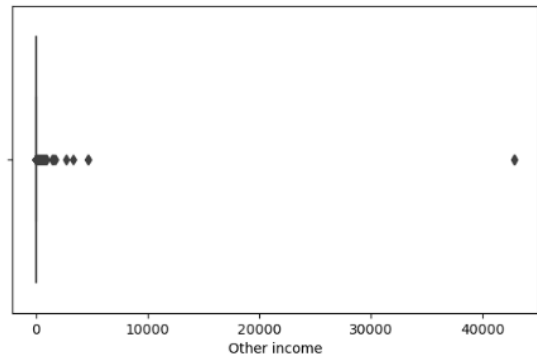
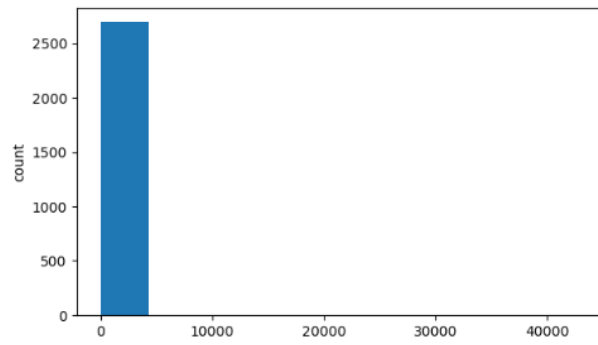
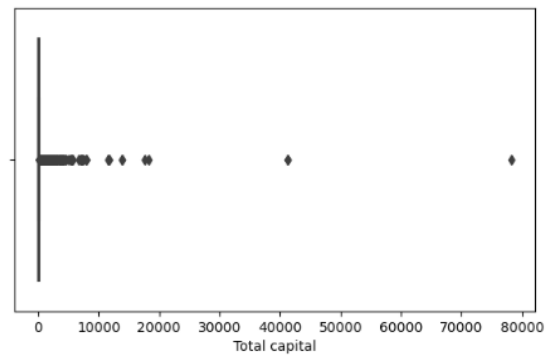
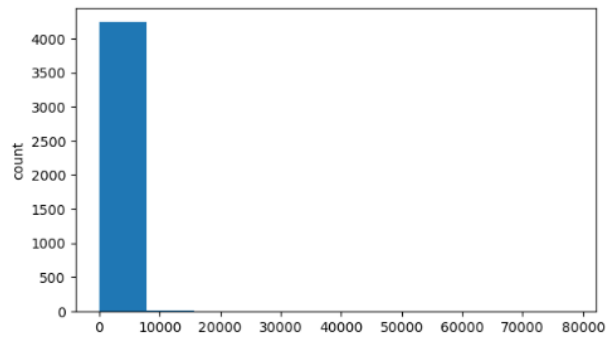


Figure 11 Income from financial Services and Other income

Total capital
Count : 4251
Number of outliers: 0



Reserves and funds
Count : 4158
Number of outliers: 0

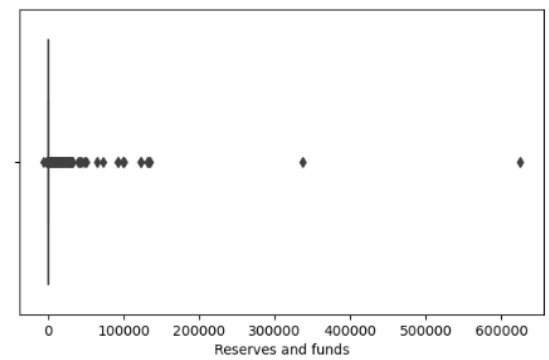
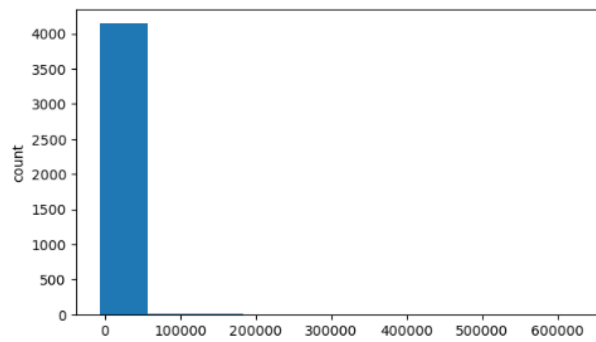
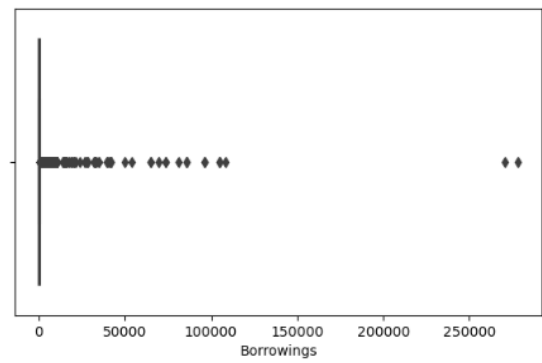
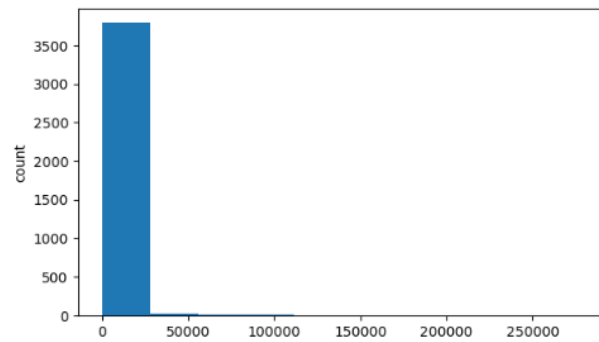


Figure 12 Total capital and Reserves & Funds

Borrowings
 Count : 3825
 Number of outliers: 0



Current liabilities & provisions
 Count : 4146
 Number of outliers: 0

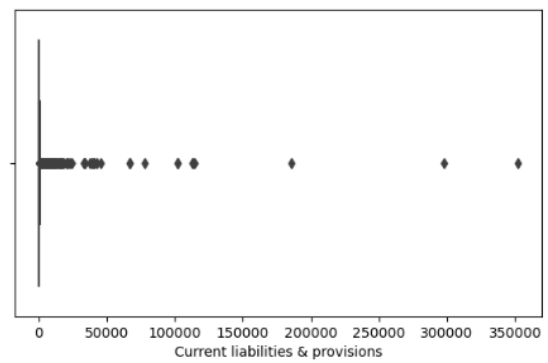
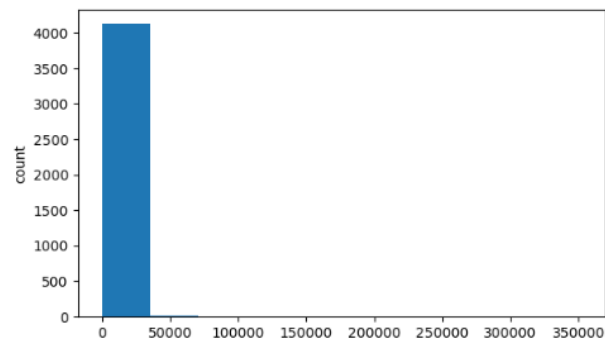


Figure 13 Borrowings & Current Liabilities and Provisions

Deferred tax liability
 Count : 2887
 Number of outliers: 0

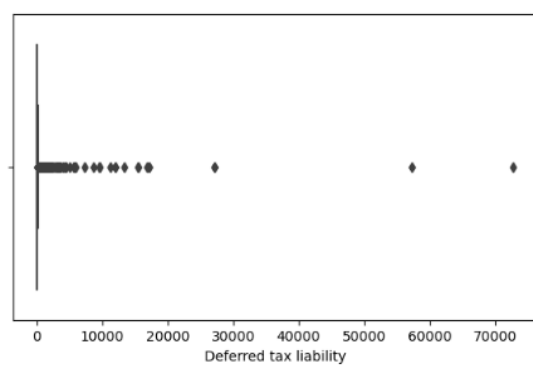
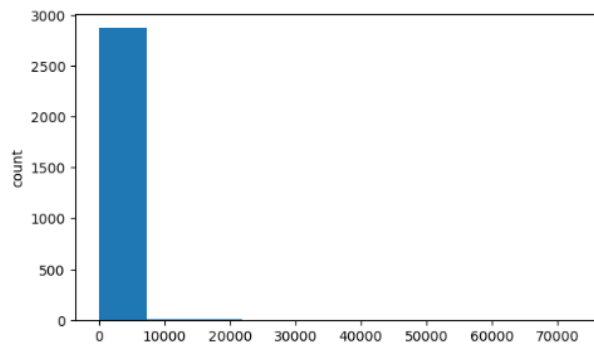
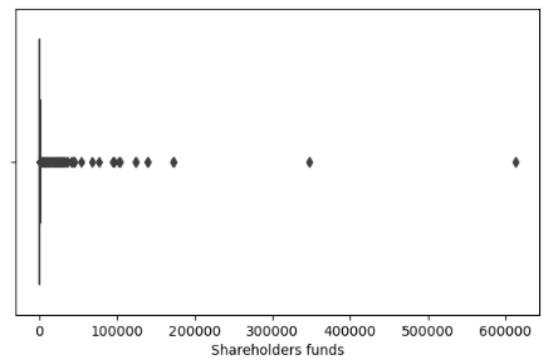
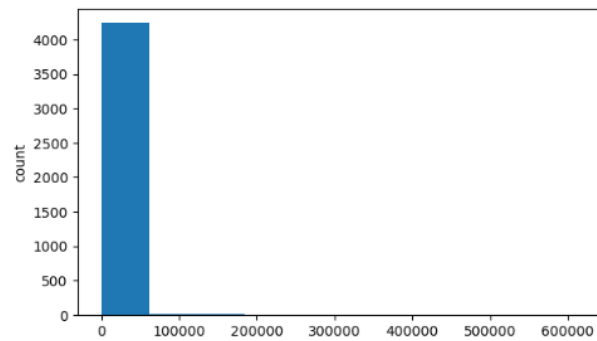


Figure 14 Deferred tax liability

Shareholders funds
Count : 4256
Number of outliers: 588



Cumulative retained profits
Count : 4211
Number of outliers: 0

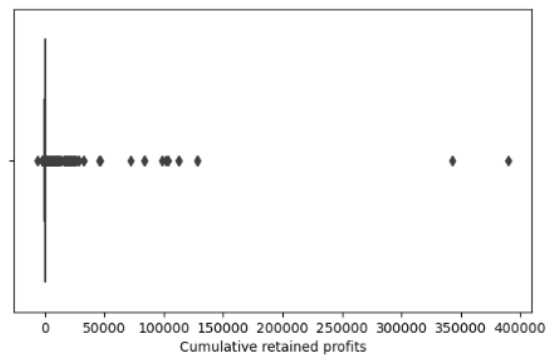
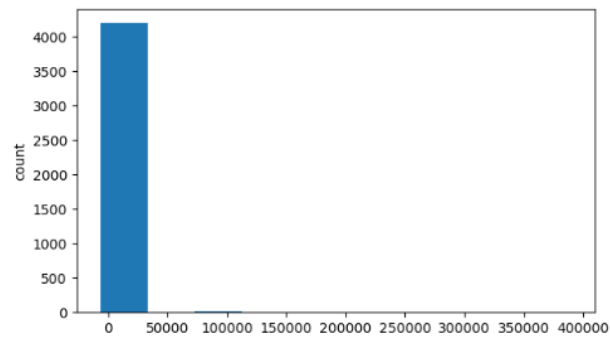
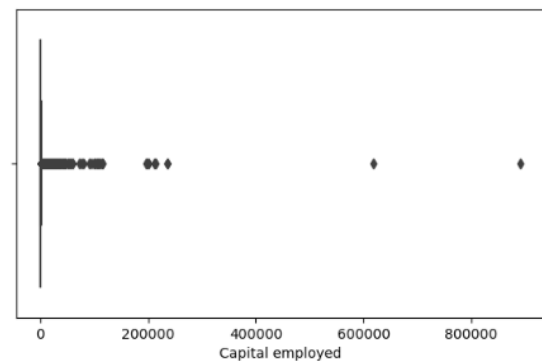
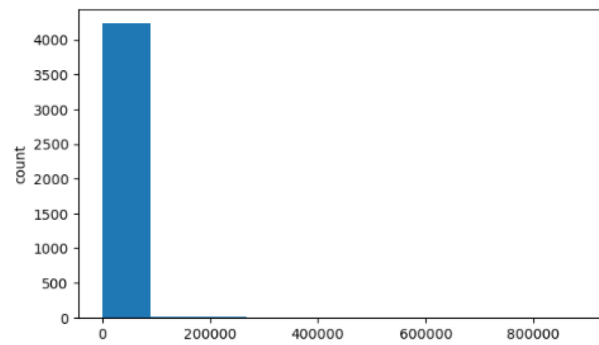


Figure 15 Shareholders funds and Cumulative retained profits

Capital employed
 Count : 4256
 Number of outliers: 572



TOL/TNW
 Count : 4256
 Number of outliers: 414

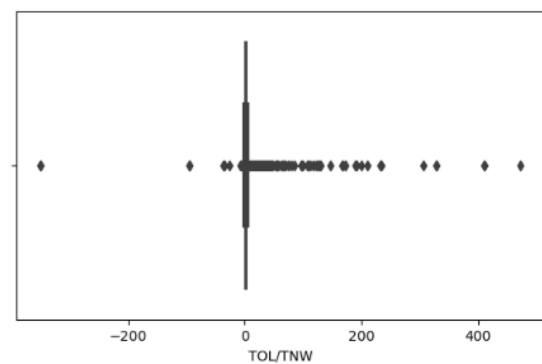
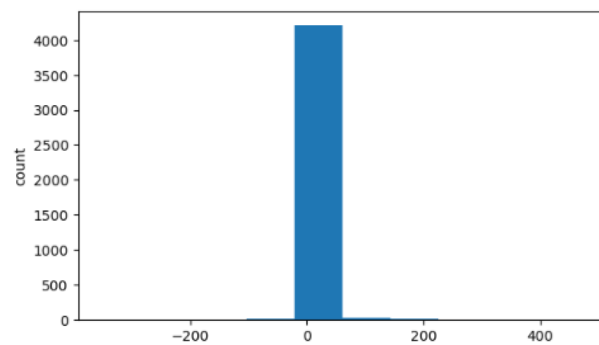


Figure 16 Capital employed and TOL/TNW

Total term liabilities / tangible net worth
 Count : 4256
 Number of outliers: 406

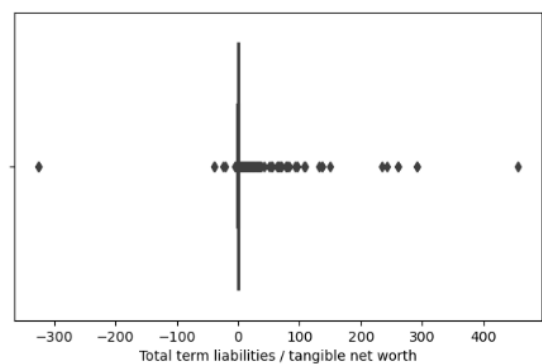
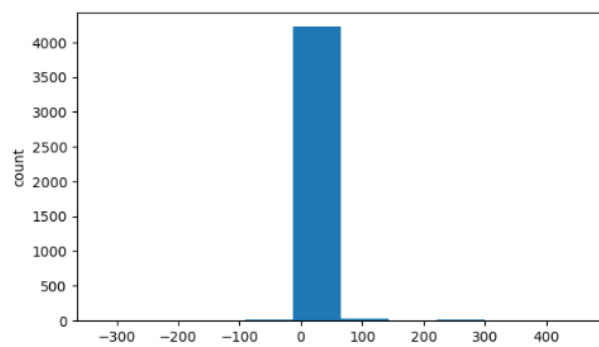
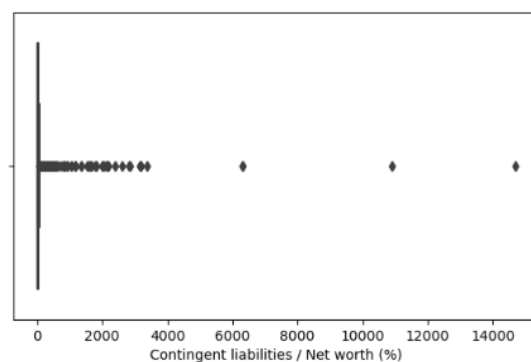
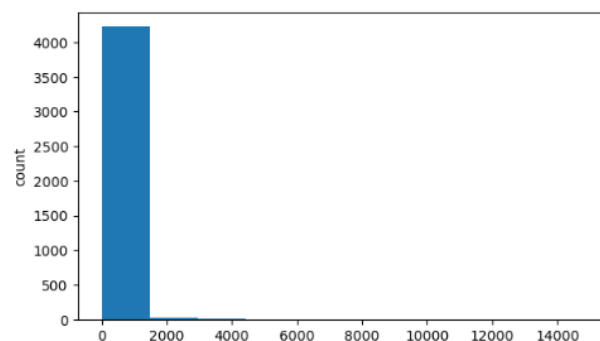


Figure 17 Total term liabilities / tangible net worth

Contingent liabilities / Net worth (%)

Count : 4256

Number of outliers: 478



Contingent liabilities

Count : 2854

Number of outliers: 0

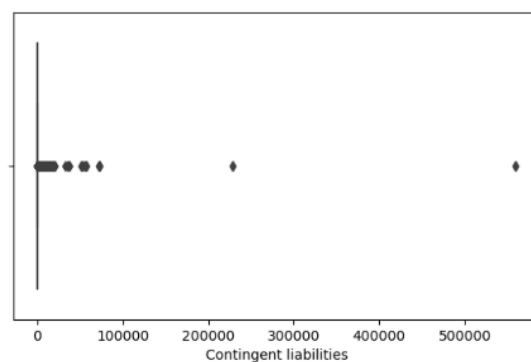
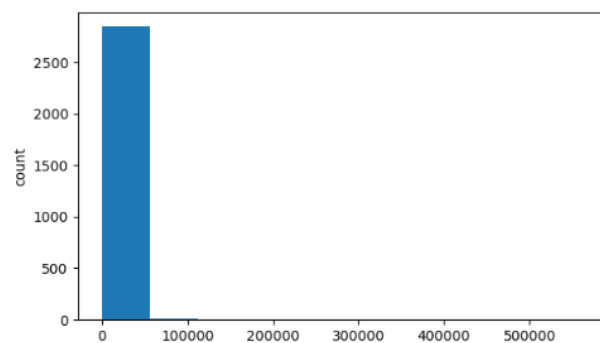


Figure 18 Contingent liabilities / Net worth and Contingent liabilities

Net fixed assets

Count : 4124

Number of outliers: 0

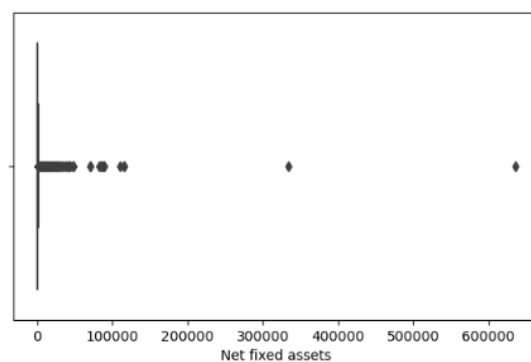
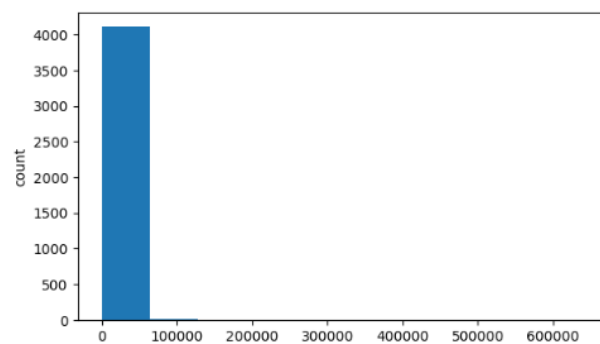
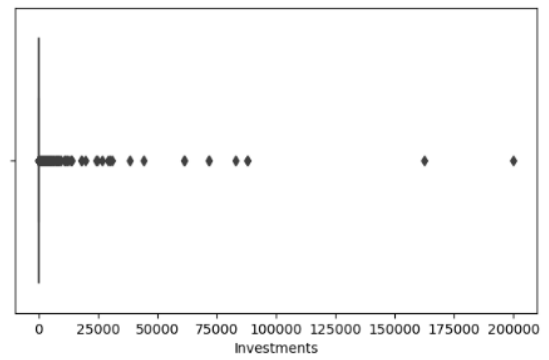
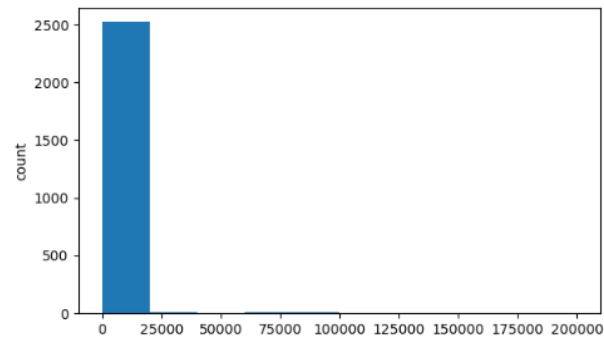


Figure 19 Net fixed assets

Investments
Count : 2541
Number of outliers: 0



Current assets
Count : 4176
Number of outliers: 0

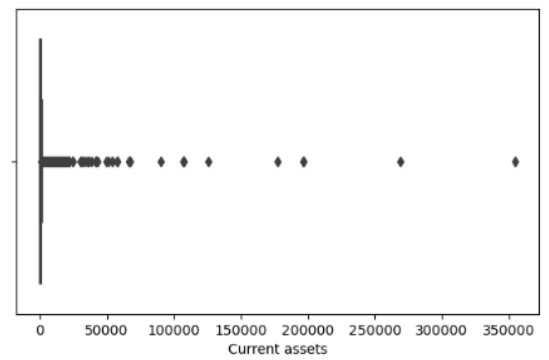
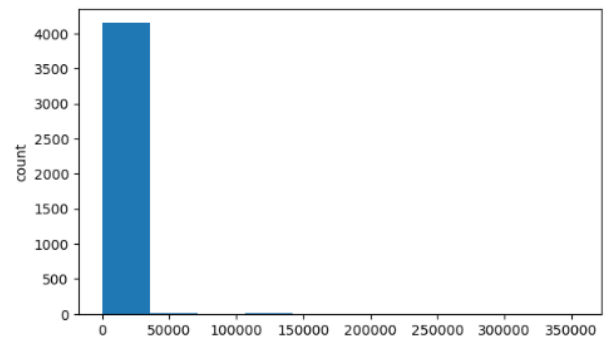
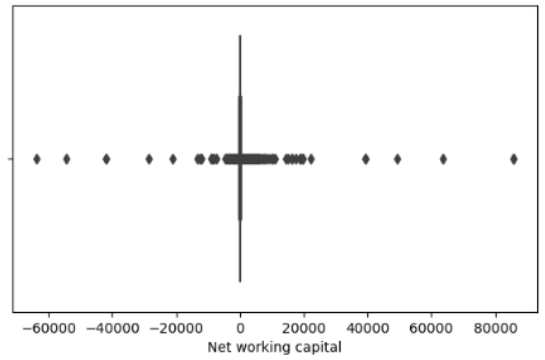
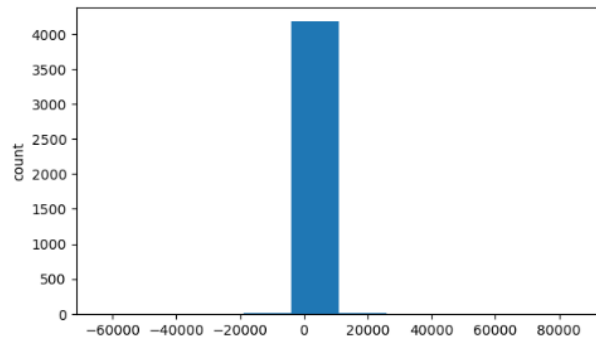


Figure 20 Investments and Current assets

Net working capital
 Count : 4219
 Number of outliers: 0



Quick ratio (times)
 Count : 4151
 Number of outliers: 0

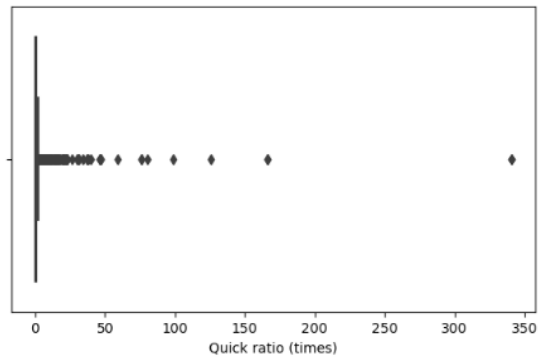
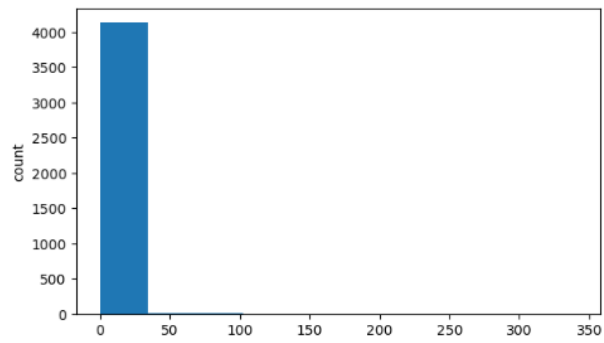
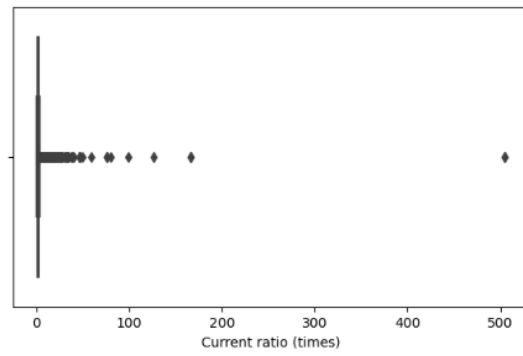
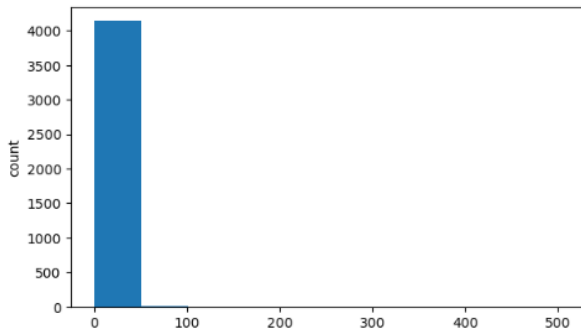


Figure 21 Net wroking capital and Quick Ratio

Current ratio (times)
Count : 4151
Number of outliers: 0



Debt to equity ratio (times)
Count : 4256
Number of outliers: 381

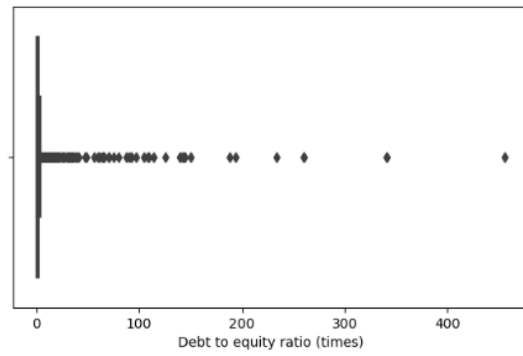
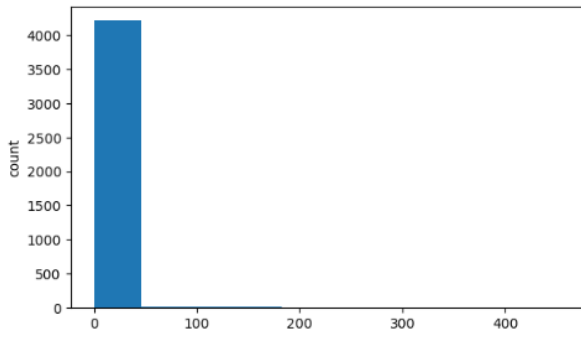
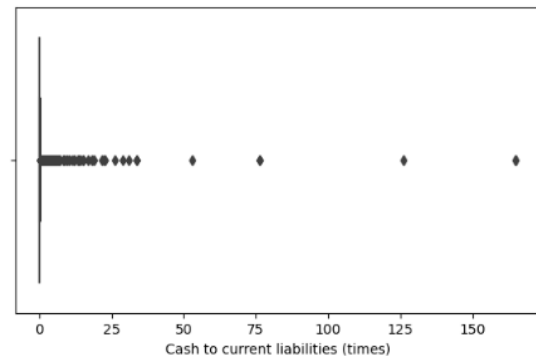
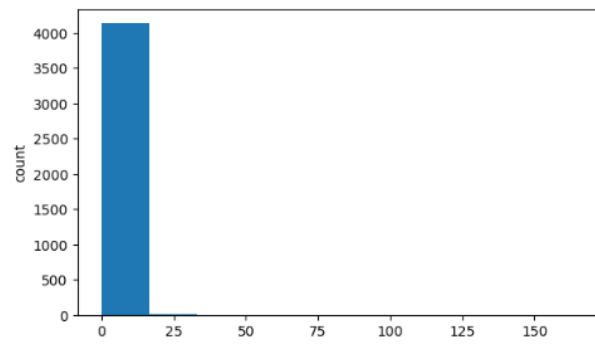


Figure 22 Current ratio and Debt to equity ratio

Cash to current liabilities (times)
 Count : 4151
 Number of outliers: 0



Cash to average cost of sales per day
 Count : 4156
 Number of outliers: 0

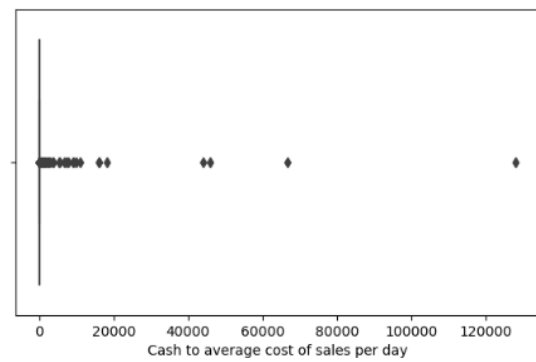
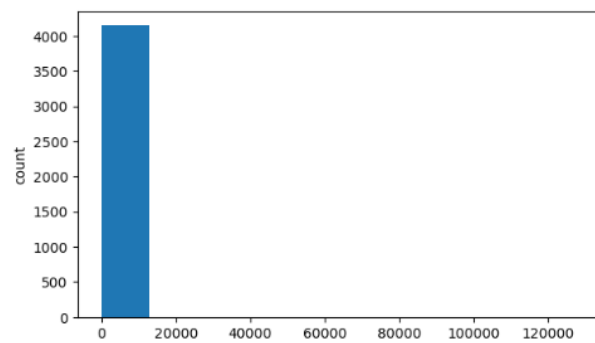
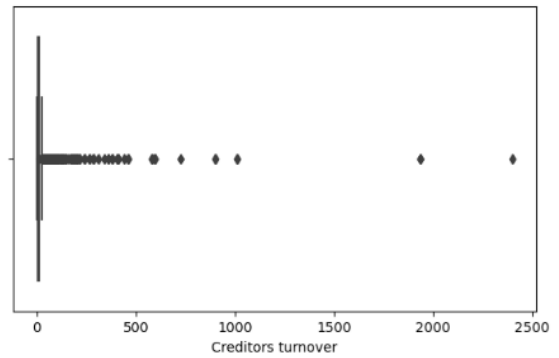
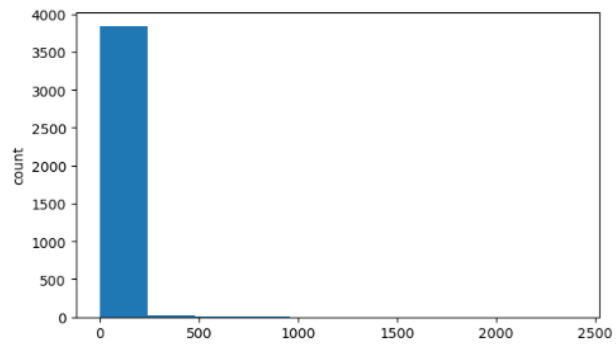


Figure 23 Cash to current liabilities and Cash to average cost of sales per day

Creditors turnover
 Count : 3865
 Number of outliers: 0



Debtors turnover
 Count : 3871
 Number of outliers: 0

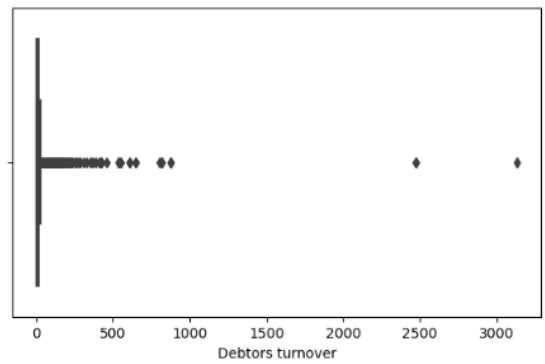
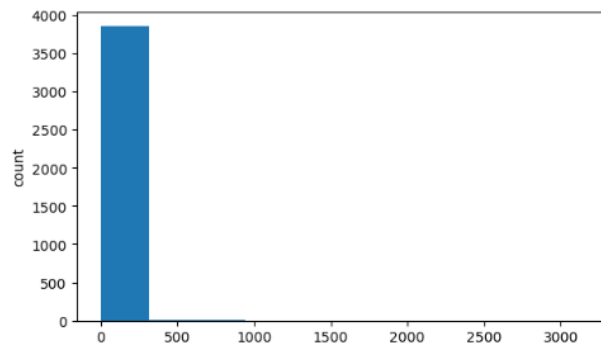
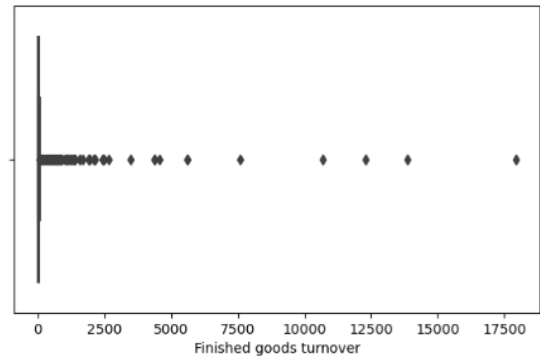
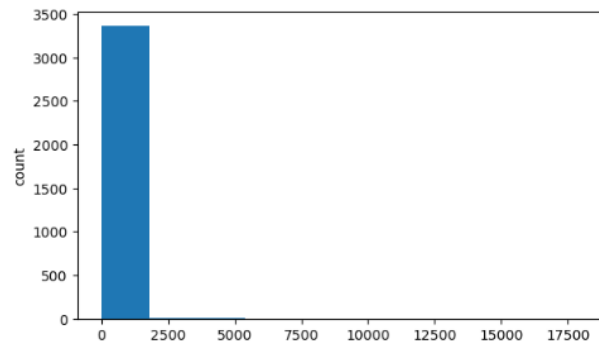


Figure 24 Creditors Turnover and Debtors turnover

Finished goods turnover
Count : 3382
Number of outliers: 0



WIP turnover
Count : 3492
Number of outliers: 0

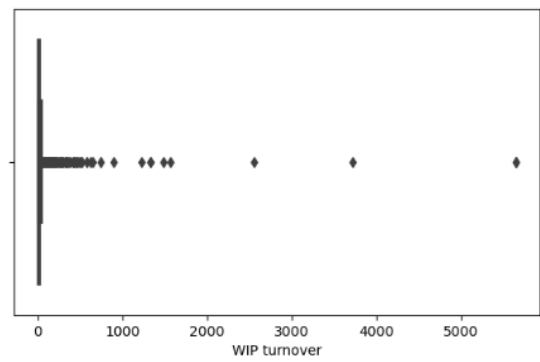
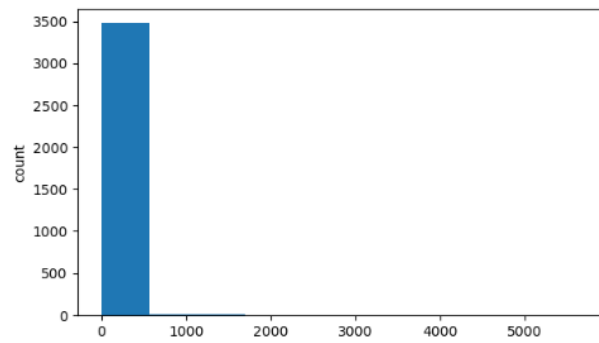
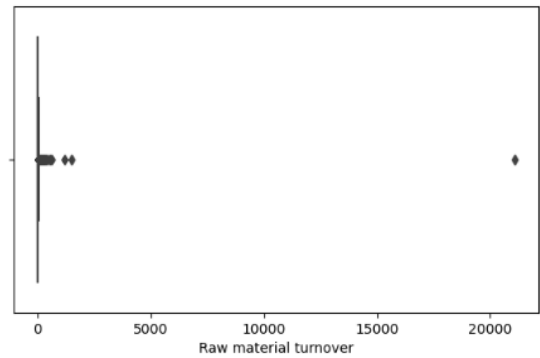
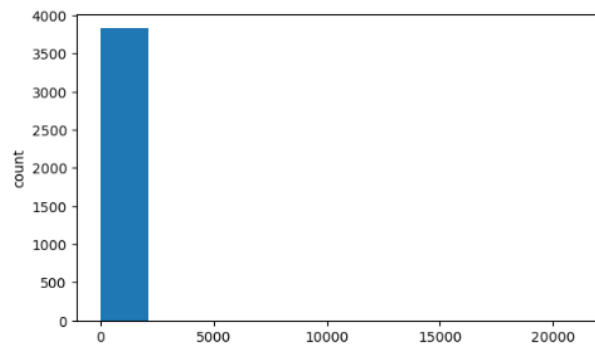


Figure 25 Finished goods turnover and WIP turnover

Raw material turnover
 Count : 3828
 Number of outliers: 0



Shares outstanding
 Count : 3446
 Number of outliers: 0

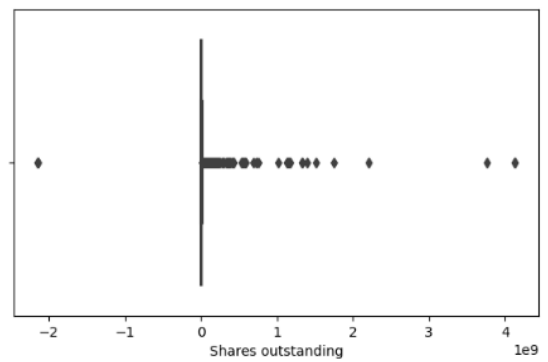
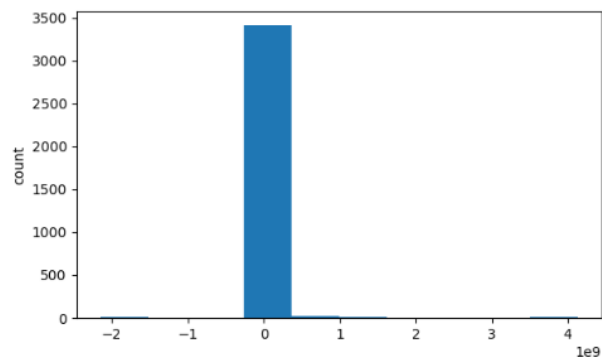
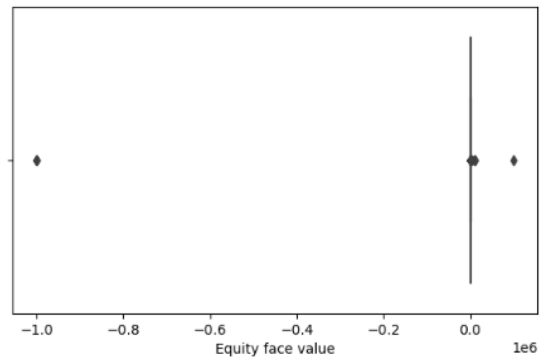
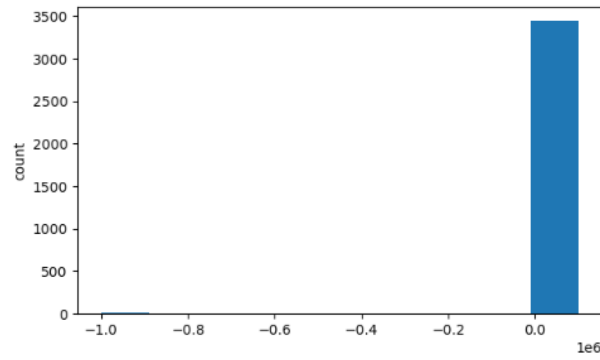


Figure 26 Raw material turnover and Shares outstanding

Equity face value
Count : 3446
Number of outliers: 0



EPS
Count : 4256
Number of outliers: 638

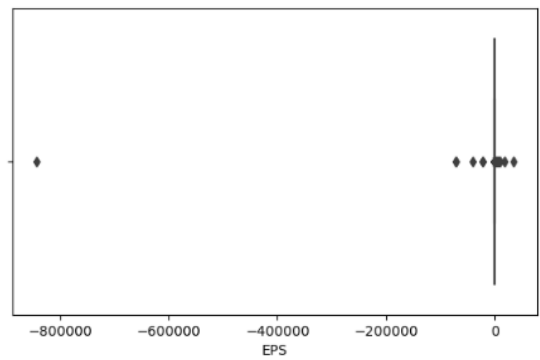
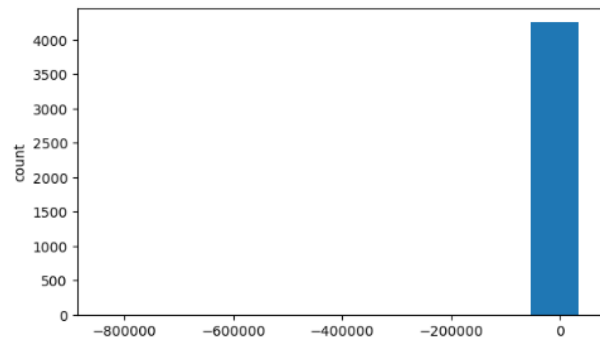
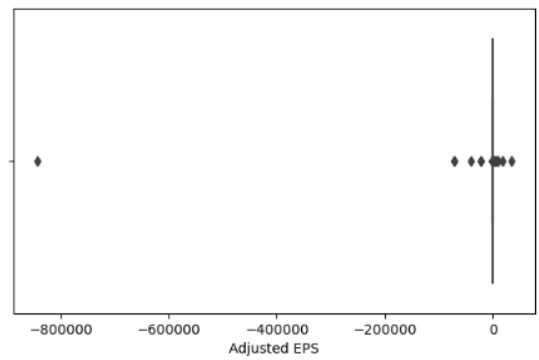
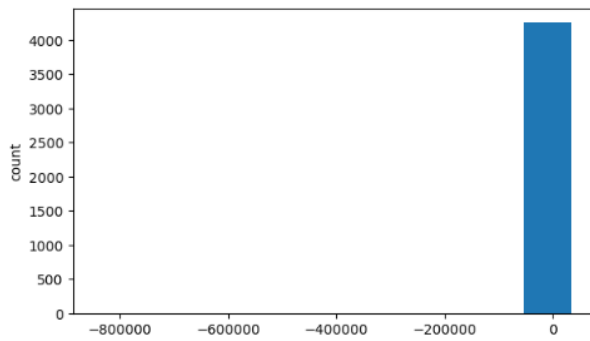


Figure 27 Equity face value and EPS

Adjusted EPS
Count : 4256
Number of outliers: 694



Total liabilities
Count : 4256
Number of outliers: 585

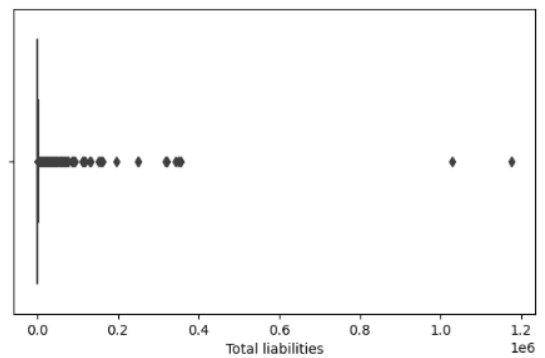
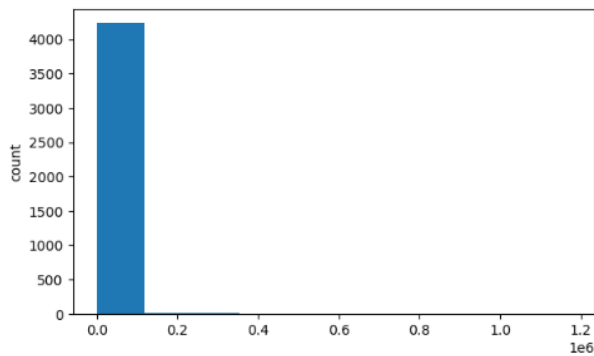


Figure 28 Adjusted EPS and Total liabilities

PE on BSE
Count : 1629
Number of outliers: 0

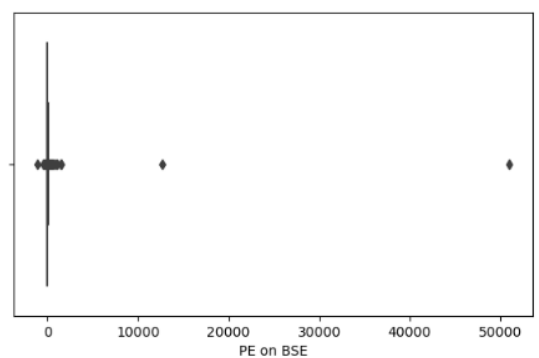
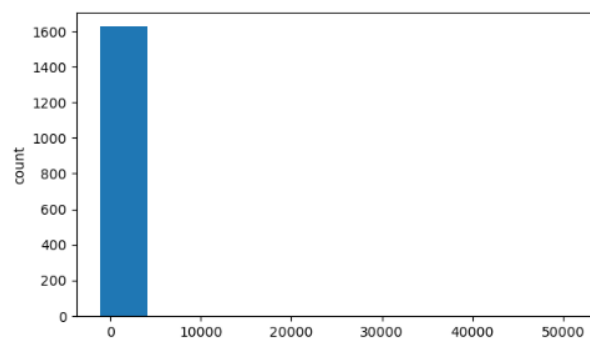


Figure 29 PE on BSE

Multivariate analysis

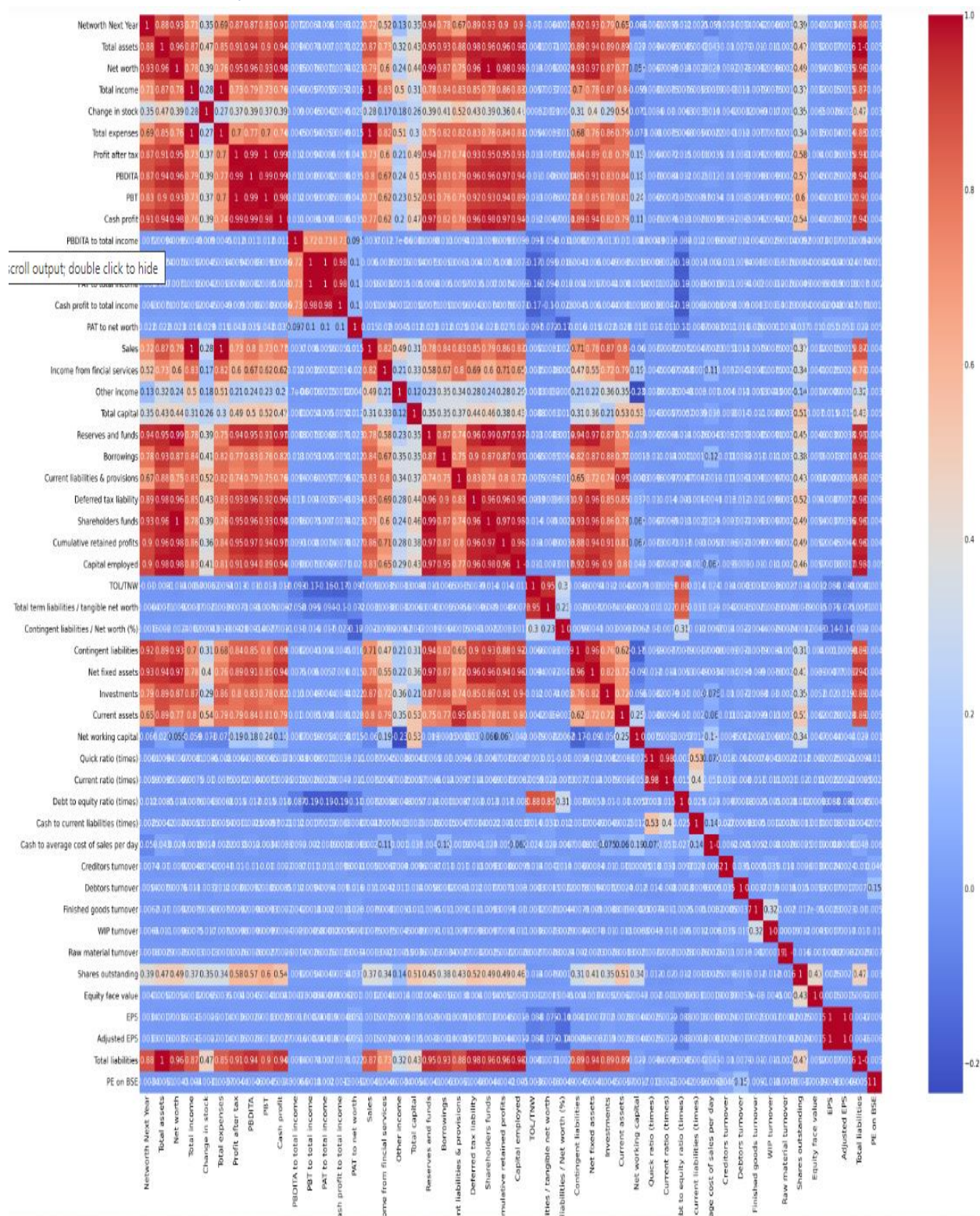


Figure 30 Heatmap

	Networth Next Year	Total assets	Net worth	Total income	Change in stock	Total expenses	Profit after tax	PBDITA	PBT	Cash profit	...	Debtors turnover	Finished goods turnover	WIP turnover
Networth Next Year	1.000000	0.877803	0.930135	0.710953	0.345199	0.690526	0.867992	0.872395	0.833797	0.906610	...	0.005430	-0.006176	-0.006573
Total assets	0.877803	1.000000	0.959404	0.868607	0.470735	0.852863	0.907560	0.943393	0.895117	0.939904	...	0.007887	-0.010292	-0.010790
Net worth	0.930135	0.959404	1.000000	0.783831	0.393760	0.761549	0.954399	0.962925	0.932463	0.978439	...	0.007610	-0.009197	-0.009574
Total income	0.710953	0.868607	0.783831	1.000000	0.276395	0.999203	0.727438	0.793210	0.727190	0.762634	...	0.011494	-0.007918	-0.007481
Change in stock	0.345199	0.470735	0.393760	0.276395	1.000000	0.273717	0.366994	0.389530	0.369046	0.391782	...	-0.003189	-0.006912	-0.016666
Total expenses	0.690526	0.852863	0.761549	0.999203	0.273717	1.000000	0.699991	0.769182	0.700041	0.737395	...	0.011745	-0.007705	-0.007221
Profit after tax	0.867992	0.907560	0.954399	0.727438	0.366994	0.699991	1.000000	0.990417	0.995142	0.990290	...	0.008058	-0.009191	-0.009841
PBDITA	0.872395	0.943393	0.962925	0.793210	0.389530	0.769182	0.990417	1.000000	0.988875	0.991634	...	0.009173	-0.009781	-0.009929
PBT	0.833797	0.895117	0.932463	0.727190	0.369046	0.700041	0.995142	0.988875	1.000000	0.977955	...	0.008490	-0.009342	-0.009857
Cash profit	0.906610	0.939904	0.978439	0.762634	0.391782	0.737395	0.990290	0.991634	0.977955	1.000000	...	0.008489	-0.009195	-0.009369
PBDITA to total income	0.007173	0.009445	0.009477	0.004892	0.009038	0.004474	0.011708	0.011380	0.011640	0.010631	...	0.011906	0.004232	0.002864
PBT to total income	0.006385	0.007374	0.007559	0.005655	0.004543	0.005387	0.009412	0.008920	0.009290	0.008648	...	0.009445	0.001795	-0.000580
PAT to total income	0.005965	0.006957	0.007052	0.005486	0.004208	0.005254	0.008648	0.008226	0.008456	0.007980	...	0.009446	0.002019	-0.000118
Cash profit to total income	0.006330	0.007124	0.007391	0.005192	0.004492	0.004933	0.009015	0.008593	0.008863	0.008561	...	0.008965	0.001303	-0.000542
PAT to net worth	0.021671	0.022346	0.023274	0.016113	0.029495	0.015005	0.043264	0.034824	0.041568	0.035205	...	0.016442	0.026036	0.000107
Sales	0.716262	0.870477	0.787302	0.999872	0.276323	0.998926	0.729793	0.795414	0.729000	0.765923	...	0.011127	-0.007911	-0.007475
Income from fincial services	0.520435	0.725773	0.601567	0.825077	0.169750	0.823314	0.603779	0.673578	0.619435	0.618667	...	0.004237	-0.008132	-0.007481
Other income	0.134670	0.321466	0.243282	0.504255	0.175114	0.513296	0.213823	0.244751	0.228176	0.201111	...	0.010557	-0.005328	-0.004451
Total capital	0.349069	0.434464	0.437481	0.310340	0.256250	0.295956	0.485572	0.500089	0.518158	0.466547	...	0.013523	-0.011168	-0.008892
Reserves and funds	0.938546	0.951959	0.993691	0.775874	0.391050	0.754176	0.939000	0.945006	0.910684	0.968124	...	0.005809	-0.008469	-0.009082

Borrowings	0.783406	0.932630	0.873876	0.835287	0.410316	0.823565	0.771153	0.825291	0.755756	0.816983	...	0.008171	-0.010633	-0.011071
Current liabilities & provisions	0.666964	0.883252	0.745515	0.831367	0.524679	0.824869	0.741344	0.793694	0.749743	0.763801	...	0.006059	-0.009110	-0.009697
Deferred tax liability	0.893295	0.975395	0.963976	0.848337	0.430406	0.830606	0.928427	0.959768	0.916677	0.961215	...	0.011804	-0.011157	-0.009818
Shareholders funds	0.929903	0.958777	0.999338	0.782088	0.392806	0.759785	0.953456	0.962275	0.931991	0.977435	...	0.007677	-0.009323	-0.009712
Cumulative retained profits	0.897528	0.961634	0.976813	0.855363	0.358181	0.836042	0.949638	0.968331	0.937503	0.967477	...	0.007258	-0.009884	-0.009832
Capital employed	0.901214	0.978867	0.980948	0.827770	0.409940	0.809373	0.911376	0.938500	0.891610	0.944844	...	0.008030	-0.010028	-0.010637
TOL/TNW	-0.010461	-0.008082	-0.013888	-0.005692	0.000518	-0.005425	-0.013077	-0.010454	-0.013060	-0.011802	...	-0.000299	-0.003218	-0.003607
Total term liabilities / tangible net worth	-0.006422	-0.007065	-0.009203	-0.003727	-0.002064	-0.003941	-0.007320	-0.005964	-0.007594	-0.006716	...	-0.001508	-0.002065	-0.002336
Contingent liabilities / Net worth (%)	-0.001563	0.001999	-0.002447	0.001966	0.000429	0.001790	-0.002765	-0.000140	-0.002686	-0.001109	...	0.002245	-0.004443	-0.002948
Contingent liabilities	0.918809	0.885966	0.932598	0.701370	0.306289	0.681963	0.839393	0.847374	0.795825	0.892474	...	0.007785	-0.007947	-0.008425
Net fixed assets	0.929506	0.939418	0.966310	0.779827	0.402203	0.761523	0.887541	0.910122	0.854963	0.938912	...	0.009397	-0.008969	-0.007756
Investments	0.787223	0.889105	0.865117	0.870066	0.293957	0.859897	0.795688	0.828215	0.779294	0.819465	...	0.007220	-0.008802	-0.010316
Current assets	0.654580	0.889097	0.772956	0.804262	0.544442	0.793950	0.787846	0.836138	0.808023	0.792468	...	0.002433	-0.009924	-0.012911
Net working capital	-0.065879	0.026920	0.055997	-0.058992	-0.071436	-0.073125	0.187539	0.177657	0.243233	0.113852	...	-0.012262	0.000225	-0.006807
Quick ratio (times)	-0.006067	-0.009413	-0.006705	-0.008053	-0.008621	-0.008011	-0.006434	-0.007793	-0.006486	-0.007120	...	-0.014001	0.000740	0.004296
Current ratio (times)	-0.005920	-0.009482	-0.006906	-0.007533	-0.010498	-0.007476	-0.007161	-0.008362	-0.007302	-0.007608	...	-0.008046	-0.010851	-0.010739
Debt to equity ratio (times)	-0.011834	-0.008463	-0.013857	-0.007580	-0.006282	-0.006811	-0.015163	-0.011714	-0.015014	-0.013218	...	-0.001780	-0.002541	-0.005019
Cash to current liabilities (times)	-0.002498	-0.004250	-0.002363	-0.005259	-0.003878	-0.005352	-0.001119	-0.002479	-0.000970	-0.002103	...	-0.000929	-0.005034	-0.001249
Cash to average cost of sales per day	-0.058829	0.042757	0.023962	-0.001865	0.014323	-0.002210	-0.003542	0.011960	-0.003440	-0.003752	...	-0.004992	-0.005166	-0.008046
Creditors turnover	-0.007441	-0.010484	-0.009213	-0.004803	-0.004242	-0.004144	-0.010084	-0.010459	-0.010146	-0.009735	...	0.035325	0.000502	0.035394
Debtors turnover	0.005430	0.007887	0.007610	0.011494	-0.003189	0.011745	0.008058	0.009173	0.008490	0.008489	...	1.000000	0.003740	0.018596
Finished goods turnover	-0.006176	-0.010292	-0.009197	-0.007918	-0.006912	-0.007705	-0.009191	-0.009781	-0.009342	-0.009195	...	0.003740	1.000000	0.323246
WIP turnover	-0.006573	-0.010790	-0.009574	-0.007481	-0.016666	-0.007221	-0.009841	-0.009929	-0.009857	-0.009369	...	0.018596	0.323246	1.000000
Raw material turnover	-0.001818	-0.002903	-0.002498	-0.001259	-0.001335	-0.001084	-0.002624	-0.002814	-0.002675	-0.002563	...	0.001631	-0.001986	-0.000187
Shares outstanding	0.385854	0.474330	0.494503	0.369170	0.351013	0.343341	0.578448	0.571384	0.604438	0.542068	...	0.014801	-0.012413	-0.012258
Equity face value	0.004116	0.005216	0.005377	0.003178	0.006547	0.003479	0.004040	0.004489	0.004138	0.004145	...	0.005308	0.000070	-0.004491
EPS	0.003431	0.001654	0.003622	0.001539	-0.002635	0.001411	0.003558	0.002872	0.003338	0.002812	...	0.001684	0.002251	0.001686
Adjusted EPS	0.003344	0.001567	0.003533	0.001489	-0.002671	0.001364	0.003455	0.002771	0.003239	0.002708	...	0.001690	0.002254	0.001696
Total liabilities	0.877803	1.000000	0.959404	0.868607	0.470735	0.852863	0.907560	0.943393	0.895117	0.939904	...	0.007887	-0.010292	-0.010790
PE on BSE	-0.003430	-0.005067	-0.004323	-0.004044	-0.003053	-0.003688	-0.004433	-0.004650	-0.004502	-0.004310	...	0.146838	-0.005089	-0.017937

Key meaningful observations on individual variables and the relationship between variables

Networth next year is positively correlated to Total Assets, Net worth, Total income, total expenses, profit after tax, PBDITA, PBT, Cash Profit, sales, Reserves and funds, borrowings, Current liabilities and provision, deferred tax and liability, shareholders fund, cumulative retained funds, capital employed, contingent liabilities, net fixed assets, investments, current assets, and total liabilities.

Networth next year is very less but positively correlated to change in stocks, PBDITA to total income, PBT to total income, PAT to total income, cash profit to total income, PAT to net worth, income from financial service, other income, total capital.

Networth next year is negatively correlated to TOL/TNW, Total term liabilities / tangible net worth, Contingent liabilities / Net worth (%), Net working capital, Quick Ratio, Current ratio, Debt to equity ratio, Cash to current liabilities, cash to average cost of sales per day, creditors turnover, finished goods turnover, WIP turnover, Raw materials turnover, PE on BSE.

Missing Values Detection

Networth Next Year	0
Total assets	0
Net worth	0
Total income	231
Change in stock	550
Total expenses	165
Profit after tax	154
PBDITA	154
PBT	154
Cash profit	154
PBDITA as % of total income	79
PBT as % of total income	79
PAT as % of total income	79
Cash profit as % of total income	79
PAT as % of net worth	0
Sales	305
Income from fincial services	1111
Other income	1556
Total capital	5
Reserves and funds	98
Borrowings	431
Current liabilities & provisions	110
Deferred tax liability	1369
Shareholders funds	0
Cumulative retained profits	45
Capital employed	0
TOL/TNW	0
Total term liabilities / tangible net worth	0
Contingent liabilities / Net worth (%)	0
Contingent liabilities	1402
Net fixed assets	132
Investments	1715
Current assets	80
Net working capital	37
Quick ratio (times)	105
Current ratio (times)	105
Debt to equity ratio (times)	0
Cash to current liabilities (times)	105
Cash to average cost of sales per day	100
Creditors turnover	391
Debtors turnover	385
Finished goods turnover	874
WIP turnover	764
Raw material turnover	428
Shares outstanding	810
Equity face value	810
EPS	0

Figure 31 Missing Value Detection

Missing Value Treatment

We address the missing values in columns with mean value of the column. This resulted in zero missing values.

Networth Next Year	0
Total assets	0
Net worth	0
Total income	0
Change in stock	0
Total expenses	0
Profit after tax	0
PBDITA	0
PBT	0
Cash profit	0
PBDITA as % of total income	0
PBT as % of total income	0
PAT as % of total income	0
Cash profit as % of total income	0
PAT as % of net worth	0
Sales	0
Income from fincial services	0
Other income	0
Total capital	0
Reserves and funds	0
Borrowings	0
Current liabilities & provisions	0
Deferred tax liability	0
Shareholders funds	0
Cumulative retained profits	0
Capital employed	0
TOL/TNW	0
Total term liabilities / tangible net worth	0
Contingent liabilities / Net worth (%)	0
Contingent liabilities	0
Net fixed assets	0
Investments	0
Current assets	0
Net working capital	0
Quick ratio (times)	0
Current ratio (times)	0
Debt to equity ratio (times)	0
Cash to current liabilities (times)	0
Cash to average cost of sales per day	0
Creditors turnover	0
Debtors turnover	0
Finished goods turnover	0
WIP turnover	0
Raw material turnover	0
Shares outstanding	0
Equity face value	0

Figure 32 Missing Value Treatment

Outlier Detection

Number of outliers in each column:

Networth Next Year	624
Total assets	585
Net worth	595
Total income	647
Change in stock	396
Total expenses	604
Profit after tax	770
PBDITA	663
PBT	763
Cash profit	693
PBDITA as % of total income	346
PBT as % of total income	628
PAT as % of total income	690
Cash profit as % of total income	423
PAT as % of net worth	427
Sales	361
Income from fincial services	125
Other income	92
Total capital	545
Reserves and funds	661
Borrowings	316
Current liabilities & provisions	634
Deferred tax liability	130
Shareholders funds	588
Cumulative retained profits	711
Capital employed	572
TOL/TNW	414
Total term liabilities / tangible net worth	406
Contingent liabilities / Net worth (%)	478
Contingent liabilities	125
Net fixed assets	629
Investments	109
Current assets	574
Net working capital	776
Quick ratio (times)	333
Current ratio (times)	369
Debt to equity ratio (times)	381
Cash to current liabilities (times)	605
Cash to average cost of sales per day	631
Creditors turnover	272
Debtors turnover	281
Finished goods turnover	172
WIP turnover	245
Raw material turnover	177
Shares outstanding	253
Equity face value	1343
EPS	638
Adjusted EPS	694
Total liabilities	585
PE on BSE	97

Figure 33 Column Outliers

We treat the outliers in the column with interquartile range method and replaced the outliers above upper range with upper range value and replace outliers below lower range with lower range value.

This resulted in removal of all outliers.

Networth Next Year	0
Total assets	0
Net worth	0
Total income	0
Change in stock	0
Total expenses	0
Profit after tax	0
PBDITA	0
PBT	0
Cash profit	0
PBDITA as % of total income	0
PBT as % of total income	0
PAT as % of total income	0
Cash profit as % of total income	0
PAT as % of net worth	0
Sales	0
Income from fincial services	0
Other income	0
Total capital	0
Reserves and funds	0
Borrowings	0
Current liabilities & provisions	0
Deferred tax liability	0
Shareholders funds	0
Cumulative retained profits	0
Capital employed	0
TOL/TNW	0
Total term liabilities / tangible net worth	0
Contingent liabilities / Net worth (%)	0
Contingent liabilities	0
Net fixed assets	0
Investments	0
Current assets	0
Net working capital	0
Quick ratio (times)	0
Current ratio (times)	0
Debt to equity ratio (times)	0
Cash to current liabilities (times)	0
Cash to average cost of sales per day	0
Creditors turnover	0
Debtors turnover	0
Finished goods turnover	0
WIP turnover	0
Raw material turnover	0
Shares outstanding	0
Equity face value	0
EPS	0

Figure 34 Outlier Treatment

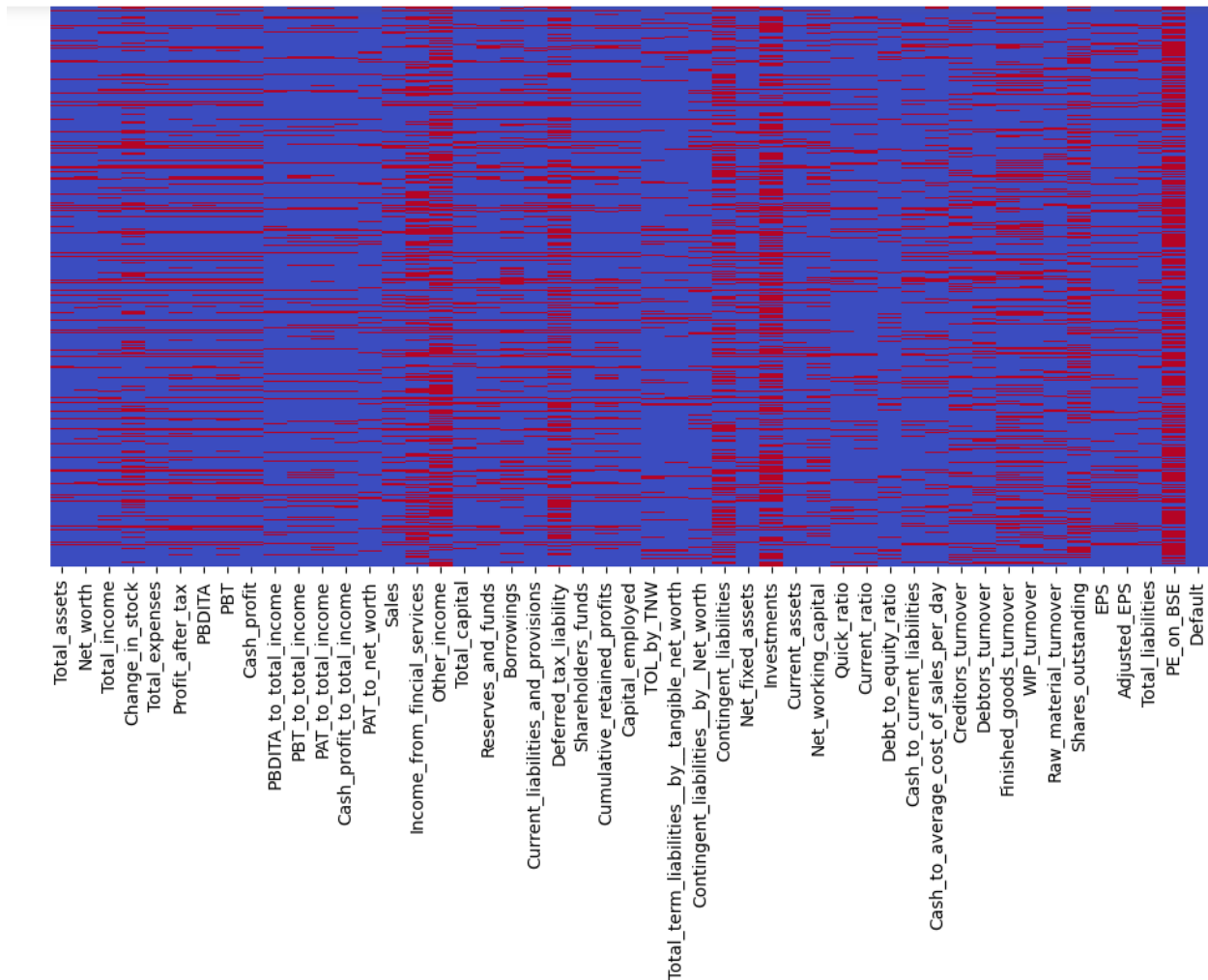


Figure 35 - Missing Values Heatmap

We should inspect total missing values by each row.

```
0      4
1      8
2      3
3     21
4      4
..
4251   24
4252    5
4253    8
4254    5
4255    2
Length: 4256, dtype: int64
```

Figure 36 Missing Value RowWise

If we consider availability of features for deciding the observations to be considered, we will end up losing more than 90% of the actual defaulters.

PE_on_BSE	0.672932
Investments	0.508929
Other_income	0.457002
Contingent_liabilities	0.421758
Deferred_tax_liability	0.417058
Income_from_fincial_services	0.382519
Change_in_stock	0.305451
Shares_outstanding	0.302162
Finished_goods_turnover	0.299107
WIP_turnover	0.268327
Borrowings	0.226269
Profit_after_tax	0.203477
PBT	0.201598
Net_working_capital	0.198073
Creditors_turnover	0.195724
Sales	0.189145
Debtors_turnover	0.186325
Cash_profit	0.183506
Cumulative_retained_profits	0.174812
Reserves_and_funds	0.174107
Total_income	0.173637
PBDITA	0.173402
Raw_material_turnover	0.170113
Net_fixed_assets	0.164709
Adjusted_EPS	0.163064
Current_liabilities_and_provisions	0.162359
PAT_to_total_income	0.161889
Cash_to_average_cost_of_sales_per_day	0.160479
Total_expenses	0.160479
Cash_to_current_liabilities	0.151316
EPS	0.149906
PBT_to_total_income	0.146852
Current_assets	0.143797
Net_worth	0.139803
Shareholders_funds	0.138158
Total_liabilities	0.137453
Total_assets	0.137453
Capital_employed	0.134398
Total_capital	0.130639
Cash_profit_to_total_income	0.118656
Current_ratio	0.117951
Contingent_liabilities__by__Net_worth	0.112312
Quick_ratio	0.111842
PAT_to_net_worth	0.100329
PBDITA_to_total_income	0.099859
TOL_by_TNW	0.097274
.....	
Total_term_liabilities__by__tangible_net_worth	0.095395
Debt_to_equity_ratio	0.089521
Default	0.000000
dtype: float64	

Figure 37 Missing Value (Percentage)

Dropping columns with more than 30% missing values

As a result, we dropped these columns:

'PE_on_BSE', 'Investments', 'Other_income', 'Contingent_liabilities', 'Deferred_tax_liability', 'Income_from_fincial_services'.

Scale the predictors

We use StandardScaler to scale and transform the variables

Imputing the remaining missing values

We use KNN imputer with n_neighbor = 10 to impute missing values and end up with 0 missing values in columns.

Logistic Regression

Model Building using Logistic Regression for 'Probability at default'

The equation of the Logistic Regression by which we predict the corresponding probabilities and then go on predict a discrete target variable is

$$y = \frac{1}{1 + e^{-z}}$$

Note: $z = \beta_0 + \sum_{i=1}^n (\beta_i X_i)$

Now, Importing statsmodels modules

Creating logistic regression equation & storing it in f_1

```
model = SM.logit(formula='Dependent Variable ~  $\Sigma$ Independent Variables (k)' data = 'Data Frame containing the required values').fit()
```


Checking the Variance Inflation Factor

	variables	VIF
0	Total_assets	inf
41	Total_liabilities	inf
14	Sales	131.649782
2	Total_income	108.809103
4	Total_expenses	54.141945
1	Net_worth	43.466172
19	Shareholders_funds	41.124173
21	Capital_employed	26.394973
7	PBT	21.950059
5	Profit_after_tax	20.968554
8	Cash_profit	14.246202
6	PBDITA	12.900187
10	PBT_to_total_income	12.776815
11	PAT_to_total_income	11.170363
26	Current_assets	9.892854
16	Reserves_and_funds	7.429286
12	Cash_profit_to_total_income	7.029608
18	Current_liabilities_and_provisions	6.206138
30	Debt_to_equity_ratio	5.794471
9	PBDITA_to_total_income	5.733987
17	Borrowings	5.353832
20	Cumulative_retained_profits	5.004437
39	EPS	4.963555
25	Net_fixed_assets	4.904620
40	Adjusted_EPS	4.288094
22	TOL_by_TNW	4.118524
15	Total_capital	3.912192
23	Total_term_liabilities__by__tangible_net_worth	3.532824
38	Shares_outstanding	3.394687
28	Quick_ratio	2.994345

Figure 38 VIF values

We dropped 'Total assets' and VIF scores are as below:

	variables	VIF
13	Sales	131.649782
1	Total_income	108.809103
3	Total_expenses	54.141945
0	Net_worth	43.466172
18	Shareholders_funds	41.124173
40	Total_liabilities	31.325368
20	Capital_employed	26.394973
6	PBT	21.950059
4	Profit_after_tax	20.968554
7	Cash_profit	14.246202
5	PBDITA	12.900187

Figure 39 Post Total Assets Removal

We dropped 'Sales' and VIF scores are as below:

	variables	VIF
1	Total_income	43.539484
0	Net_worth	43.294984
3	Total_expenses	43.103651
17	Shareholders_funds	41.084918
39	Total_liabilities	31.255299
19	Capital_employed	26.375296
6	PBT	21.940392
4	Profit_after_tax	20.937985
7	Cash_profit	14.136906
5	PBDITA	12.875341
9	PBT_to_total_income	12.776753

Figure 40 Post Sales Removal

We dropped 'Total_income' and VIF scores are as below:

	variables	VIF
0	Net_worth	43.294710
16	Shareholders_funds	41.082961
38	Total_liabilities	31.174769
18	Capital_employed	26.236527
5	PBT	21.929146
3	Profit_after_tax	20.921349
6	Cash_profit	14.118911
8	PBT_to_total_income	12.776751
4	PBDITA	12.480905
9	PAT_to_total_income	11.151488
23	Current_assets	9.887915

Figure 41 Post Total Income Removal

We dropped 'Net_worth' and VIF score are as below

	variables	VIF
37	Total_liabilities	31.169340
17	Capital_employed	25.959404
4	PBT	21.871176
2	Profit_after_tax	20.921348
5	Cash_profit	14.106940
7	PBT_to_total_income	12.762832
3	PBDITA	12.424094
8	PAT_to_total_income	11.131740
15	Shareholders_funds	10.822254
22	Current_assets	9.886009
9	Cash_profit_to_total_income	7.021444

Figure 42 Post Net_worth Removal

Dropping 'Total_liabilities' and updated VIF scores are:

	variables	VIF
4	PBT	21.780170
2	Profit_after_tax	20.912312
5	Cash_profit	14.096848
17	Capital_employed	13.586146
7	PBT_to_total_income	12.762481
3	PBDITA	12.373568
8	PAT_to_total_income	11.126614
15	Shareholders_funds	10.759140
22	Current_assets	9.688740
9	Cash_profit_to_total_income	7.019572
12	Reserves_and_funds	6.960163

Figure 43 Post Total_liabilities Removal

Dropping 'PBT' and updated VIF scores are:

	variables	VIF
4	Cash_profit	13.892897
16	Capital_employed	13.547460
3	PBDITA	12.170908
6	PBT_to_total_income	12.064827
14	Shareholders_funds	10.750856
7	PAT_to_total_income	10.727936
21	Current_assets	9.685193
2	Profit_after_tax	7.869473
8	Cash_profit_to_total_income	7.017592
11	Reserves_and_funds	6.960113
1	Total_expenses	6.775318

Figure 44 Post PBT Removal

Dropping Cash_profit and updated VIF scores are:

	variables	VIF
15	Capital_employed	13.542594
5	PBT_to_total_income	12.058696
13	Shareholders_funds	10.670712
6	PAT_to_total_income	10.660854
20	Current_assets	9.673750
3	PBDITA	8.047602
10	Reserves_and_funds	6.913829
1	Total_expenses	6.747742
7	Cash_profit_to_total_income	6.384480
2	Profit_after_tax	6.145296
24	Debt_to_equity_ratio	5.750171

Figure 45 Post Cash_profit removal

Dropping 'Capital_employed' and updated VIF scores are:

	variables	VIF
5	PBT_to_total_income	12.056419
6	PAT_to_total_income	10.657554
19	Current_assets	9.253055
3	PBDITA	8.012468
13	Shareholders_funds	7.701626
10	Reserves_and_funds	6.903012
1	Total_expenses	6.738717
7	Cash_profit_to_total_income	6.367282
2	Profit_after_tax	6.131936
23	Debt_to_equity_ratio	5.742449
4	PBDITA_to_total_income	5.543391

Figure 46 Post Capital_employed Removal

Dropping 'PBT_to_total_income' and updated VIF scores are:

	variables	VIF
18	Current_assets	9.251724
3	PBDITA	7.998028
12	Shareholders_funds	7.700954
9	Reserves_and_funds	6.898907
1	Total_expenses	6.736205
6	Cash_profit_to_total_income	6.248586
2	Profit_after_tax	6.129941
22	Debt_to_equity_ratio	5.698268
11	Current_liabilities_and_provisions	5.458636
4	PBDITA_to_total_income	5.407204
13	Cumulative_retained_profits	4.924534

Figure 47 Post PBT_to_total_income Removal

Dropping 'Current_assets' and updated VIF scores are:

	variables	VIF
3	PBDITA	7.983288
12	Shareholders_funds	7.205917
9	Reserves_and_funds	6.895570
6	Cash_profit_to_total_income	6.220607
2	Profit_after_tax	6.129569
1	Total_expenses	5.882669
21	Debt_to_equity_ratio	5.688713
4	PBDITA_to_total_income	5.382932
13	Cumulative_retained_profits	4.919685
30	EPS	4.899187
5	PAT_to_total_income	4.608514

Figure 48 Post Current_assets Removal

Dropping 'PBDITA' and updated VIF scores are:

	variables	VIF
11	Shareholders_funds	7.198628
8	Reserves_and_funds	6.817984
5	Cash_profit_to_total_income	6.216095
20	Debt_to_equity_ratio	5.669430
1	Total_expenses	5.350676
3	PBDITA_to_total_income	5.271123
12	Cumulative_retained_profits	4.919104
29	EPS	4.886325
2	Profit_after_tax	4.770087
4	PAT_to_total_income	4.558801
10	Current_liabilities_and_provisions	4.444305

Figure 49 Post PBDITA Removal

Dropping 'Shareholders_funds and updated VIF scores are:

	variables	VIF
5	Cash_profit_to_total_income	6.207692
19	Debt_to_equity_ratio	5.668709
8	Reserves_and_funds	5.383456
1	Total_expenses	5.327894
3	PBDITA_to_total_income	5.269708
11	Cumulative_retained_profits	4.915499
28	EPS	4.877467
2	Profit_after_tax	4.604105
4	PAT_to_total_income	4.558595
29	Adjusted_EPS	4.259592
10	Current_liabilities_and_provisions	4.226985

Figure 50 Post Shareholders_funds Removal

Dropping 'Cash_profit_to_total_income' and updated VIF scores are:

	variables	VIF
18	Debt_to_equity_ratio	5.660805
7	Reserves_and_funds	5.383429
1	Total_expenses	5.326076
10	Cumulative_retained_profits	4.914155
27	EPS	4.877238
2	Profit_after_tax	4.598769
28	Adjusted_EPS	4.254777
9	Current_liabilities_and_provisions	4.204259
4	PAT_to_total_income	4.053324
11	TOL_by_TNW	4.007892
14	Net_fixed_assets	3.873218

Figure 51 Cash_profit_to_total_income Removal

Dropping Debt_to_equity_ratio' and updated VIF scores are:

	variables	VIF
7	Reserves_and_funds	5.383425
1	Total_expenses	5.316590
10	Cumulative_retained_profits	4.913792
26	EPS	4.874680
2	Profit_after_tax	4.597908
27	Adjusted_EPS	4.254288
9	Current_liabilities_and_provisions	4.116111
4	PAT_to_total_income	4.048094
14	Net_fixed_assets	3.858244
6	Total_capital	3.785985
8	Borrowings	3.611351

Figure 52 Post Debt_to_equity_ratio Removal

Dropping 'Reserves_and_funds' and updated VIF scores are:

	variables	VIF
1	Total_expenses	5.236616
25	EPS	4.857865
2	Profit_after_tax	4.521754
26	Adjusted_EPS	4.253969
8	Current_liabilities_and_provisions	4.113171
4	PAT_to_total_income	4.046197
13	Net_fixed_assets	3.809390
6	Total_capital	3.779194
7	Borrowings	3.607526
24	Shares_outstanding	3.325041
15	Quick_ratio	2.968943

Figure 53 Post Reserves_and_funds removal

Dropping Total_expenses' and updated VIF scores are:

	variables	VIF
24	EPS	4.856870
1	Profit_after_tax	4.332428
25	Adjusted_EPS	4.253725
3	PAT_to_total_income	4.035344
5	Total_capital	3.779134
12	Net_fixed_assets	3.722924
6	Borrowings	3.471118
23	Shares_outstanding	3.304115
7	Current_liabilities_and_provisions	3.106442
14	Quick_ratio	2.967292
15	Current_ratio	2.946776

Figure 54 Post Total_expenses Removal

We use all remaining variables as X and 'Default' as y . We do 70:30 split of data for train and test data

We create first model with formula where default is derived as summation of rest of variables. We use statsmodel library's logit function to create it and check the summary.

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2952
Method:	MLE	Df Model:	26
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01761
Time:	13:05:31	Log-Likelihood:	-1508.5
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	0.0009911

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3297	0.050	-26.835	0.000	-1.427	-1.233
Change_in_stock	0.0150	0.057	0.263	0.793	-0.097	0.127
Profit_after_tax	0.0132	0.081	0.162	0.872	-0.147	0.173
PBDITA_to_total_income	-0.1179	0.075	-1.573	0.116	-0.265	0.029
PAT_to_total_income	-0.0347	0.089	-0.390	0.697	-0.209	0.140
PAT_to_net_worth	-0.0311	0.074	-0.420	0.674	-0.176	0.114
Total_capital	0.0814	0.088	0.926	0.354	-0.091	0.254
Borrowings	0.0695	0.080	0.866	0.386	-0.088	0.227
Current_liabilities_and_provisions	-0.0015	0.074	-0.021	0.984	-0.147	0.144
Cumulative_retained_profits	-0.1970	0.069	-2.847	0.004	-0.333	-0.061
TOL_by_TNW	-0.0313	0.072	-0.434	0.664	-0.172	0.110
Total_term_liabilities__by__tangible_net_worth	0.0288	0.069	0.418	0.676	-0.106	0.164
Contingent_liabilities__by__Net_worth	-0.0556	0.055	-1.018	0.309	-0.163	0.051
Net_fixed_assets	0.0245	0.081	0.303	0.762	-0.134	0.183
Net_working_capital	-0.0191	0.062	-0.309	0.757	-0.140	0.102
Quick_ratio	0.0773	0.081	0.954	0.340	-0.081	0.236
Current_ratio	-0.0817	0.080	-1.016	0.310	-0.239	0.076
Cash_to_current_liabilities	-0.1172	0.073	-1.601	0.109	-0.261	0.026
Cash_to_average_cost_of_sales_per_day	0.0647	0.071	0.911	0.362	-0.074	0.204
Creditors_turnover	-0.0379	0.062	-0.609	0.543	-0.160	0.084
Debtors_turnover	0.0780	0.061	1.277	0.202	-0.042	0.198

Figure 55 Model 2 Summary

Current_liabilities_and_provisions is an insignificant variable, therefore, we will eliminate it.

We remove the column and create new model.

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2953
Method:	MLE	Df Model:	25
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01761
Time:	13:05:32	Log-Likelihood:	-1508.5
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	0.0006483

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3297	0.050	-26.847	0.000	-1.427	-1.233
Change_in_stock	0.0149	0.057	0.262	0.793	-0.097	0.127
Profit_after_tax	0.0126	0.077	0.163	0.870	-0.139	0.164
PBDITA_to_total_income	-0.1179	0.075	-1.574	0.115	-0.265	0.029
PAT_to_total_income	-0.0346	0.089	-0.389	0.697	-0.209	0.139
PAT_to_net_worth	-0.0311	0.074	-0.420	0.674	-0.176	0.114
Total_capital	0.0812	0.087	0.929	0.353	-0.090	0.253
Borrowings	0.0694	0.080	0.867	0.386	-0.087	0.226
Cumulative_retained_profits	-0.1972	0.068	-2.882	0.004	-0.331	-0.063
TOL_by_TNW	-0.0316	0.070	-0.454	0.650	-0.168	0.105
Total_term_liabilities__by__tangible_net_worth	0.0291	0.068	0.429	0.668	-0.104	0.162
Contingent_liabilities__by__Net_worth	-0.0557	0.054	-1.026	0.305	-0.162	0.051
Net_fixed_assets	0.0241	0.078	0.307	0.759	-0.130	0.178
Net_working_capital	-0.0191	0.062	-0.310	0.757	-0.140	0.102
Quick_ratio	0.0773	0.081	0.954	0.340	-0.081	0.236
Current_ratio	-0.0817	0.080	-1.016	0.310	-0.239	0.076
Cash_to_current_liabilities	-0.1172	0.073	-1.601	0.109	-0.261	0.026
Cash_to_average_cost_of_sales_per_day	0.0646	0.071	0.913	0.361	-0.074	0.203
Creditors_turnover	-0.0377	0.061	-0.615	0.538	-0.158	0.082

Figure 56 Model 3 Summary

Shares_outstanding is still insignificant, therefore, we will eliminate it

We create new model removing Shares_outstanding.

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2954
Method:	MLE	Df Model:	24
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01761
Time:	13:09:19	Log-Likelihood:	-1508.5
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	0.0004158

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3293	0.048	-27.425	0.000	-1.424	-1.234
Change_in_stock	0.0149	0.057	0.262	0.794	-0.097	0.127
Profit_after_tax	0.0124	0.077	0.161	0.872	-0.139	0.164
PBDITA_to_total_income	-0.1179	0.075	-1.574	0.115	-0.265	0.029
PAT_to_total_income	-0.0346	0.089	-0.390	0.696	-0.209	0.139
PAT_to_net_worth	-0.0311	0.074	-0.420	0.674	-0.176	0.114
Total_capital	0.0787	0.058	1.353	0.176	-0.035	0.193
Borrowings	0.0692	0.080	0.866	0.386	-0.087	0.226
Cumulative_retained_profits	-0.1972	0.068	-2.882	0.004	-0.331	-0.063
TOL_by_TNW	-0.0315	0.070	-0.452	0.651	-0.168	0.105
Total_term_liabilities__by__tangible_net_worth	0.0290	0.068	0.428	0.669	-0.104	0.162
Contingent_liabilities__by__Net_worth	-0.0557	0.054	-1.026	0.305	-0.162	0.051
Net_fixed_assets	0.0239	0.078	0.306	0.760	-0.130	0.177
Net_working_capital	-0.0192	0.062	-0.312	0.755	-0.140	0.102
Quick_ratio	0.0774	0.081	0.956	0.339	-0.081	0.236
Current_ratio	-0.0818	0.080	-1.017	0.309	-0.239	0.076
Cash_to_current_liabilities	-0.1172	0.073	-1.601	0.109	-0.261	0.026
Cash_to_average_cost_of_sales_per_day	0.0646	0.071	0.913	0.361	-0.074	0.203
Creditors_turnover	-0.0376	0.061	-0.614	0.539	-0.157	0.082

Figure 57 Model 4 Summary

Profit_after_tax is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2955
Method:	MLE	Df Model:	23
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01760
Time:	13:15:06	Log-Likelihood:	-1508.6
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	0.0002631

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3277	0.048	-27.942	0.000	-1.421	-1.235
Change_in_stock	0.0154	0.057	0.272	0.786	-0.096	0.127
PBDITA_to_total_income	-0.1187	0.075	-1.589	0.112	-0.265	0.028
PAT_to_total_income	-0.0313	0.086	-0.363	0.717	-0.200	0.138
PAT_to_net_worth	-0.0282	0.072	-0.393	0.694	-0.169	0.113
Total_capital	0.0803	0.057	1.401	0.161	-0.032	0.193
Borrowings	0.0707	0.079	0.889	0.374	-0.085	0.226
Cumulative_retained_profits	-0.1936	0.065	-2.991	0.003	-0.321	-0.067
TOL_by_TNW	-0.0316	0.070	-0.454	0.650	-0.168	0.105
Total_term_liabilities__by__tangible_net_worth	0.0285	0.068	0.420	0.674	-0.104	0.161
Contingent_liabilities__by__Net_worth	-0.0556	0.054	-1.024	0.306	-0.162	0.051
Net_fixed_assets	0.0258	0.077	0.333	0.739	-0.126	0.178
Net_working_capital	-0.0186	0.062	-0.302	0.763	-0.139	0.102
Quick_ratio	0.0776	0.081	0.958	0.338	-0.081	0.236
Current_ratio	-0.0820	0.080	-1.019	0.308	-0.240	0.076
Cash_to_current_liabilities	-0.1167	0.073	-1.596	0.111	-0.260	0.027
Cash_to_average_cost_of_sales_per_day	0.0643	0.071	0.909	0.363	-0.074	0.203
Creditors_turnover	-0.0380	0.061	-0.623	0.534	-0.158	0.082
Debtors_turnover	0.0777	0.061	1.273	0.203	-0.042	0.197

Figure 58 Model 5 Summary

Change_in_stock is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2956
Method:	MLE	Df Model:	22
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01758
Time:	13:16:30	Log-Likelihood:	-1508.6
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	0.0001655

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3280	0.048	-27.950	0.000	-1.421	-1.235
PAT_to_total_income	-0.1193	0.075	-1.598	0.110	-0.266	0.027
PAT_to_total_income	-0.0304	0.086	-0.353	0.724	-0.199	0.139
PAT_to_net_worth	-0.0274	0.072	-0.382	0.702	-0.168	0.113
Total_capital	0.0812	0.057	1.418	0.156	-0.031	0.193
Borrowings	0.0715	0.079	0.900	0.368	-0.084	0.227
Cumulative_retained_profits	-0.1924	0.065	-2.979	0.003	-0.319	-0.066
TOL_by_TNW	-0.0308	0.070	-0.442	0.658	-0.167	0.106
Total_term_liabilities_by_tangible_net_worth	0.0283	0.068	0.418	0.676	-0.104	0.161
Contingent_liabilities_by_Net_worth	-0.0550	0.054	-1.014	0.310	-0.161	0.051
Net_fixed_assets	0.0264	0.077	0.341	0.733	-0.125	0.178
Net_working_capital	-0.0177	0.061	-0.288	0.774	-0.138	0.103
Quick_ratio	0.0758	0.081	0.939	0.348	-0.082	0.234
Current_ratio	-0.0807	0.080	-1.005	0.315	-0.238	0.077
Cash_to_current_liabilities	-0.1163	0.073	-1.590	0.112	-0.260	0.027
Cash_to_average_cost_of_sales_per_day	0.0642	0.071	0.908	0.364	-0.074	0.203
Creditors_turnover	-0.0380	0.061	-0.622	0.534	-0.158	0.082
Debtors_turnover	0.0772	0.061	1.266	0.206	-0.042	0.197
Finished_goods_turnover	-0.1171	0.073	-1.603	0.109	-0.260	0.026

Figure 59 Model 6 Summary

PAT_to_total_income is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2957
Method:	MLE	Df Model:	21
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01754
Time:	13:19:19	Log-Likelihood:	-1508.7
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	0.0001036

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3281	0.048	-27.956	0.000	-1.421	-1.235
PBDITA_to_total_income	-0.1351	0.060	-2.252	0.024	-0.253	-0.018
PAT_to_net_worth	-0.0390	0.064	-0.613	0.540	-0.164	0.086
Total_capital	0.0800	0.057	1.399	0.162	-0.032	0.192
Borrowings	0.0731	0.079	0.923	0.356	-0.082	0.228
Cumulative_retained_profits	-0.1944	0.064	-3.020	0.003	-0.321	-0.068
TOL_by_TNW	-0.0280	0.069	-0.406	0.685	-0.163	0.107
Total_term_liabilities__by__tangible_net_worth	0.0302	0.067	0.447	0.655	-0.102	0.162
Contingent_liabilities__by__Net_worth	-0.0552	0.054	-1.018	0.309	-0.161	0.051
Net_fixed_assets	0.0272	0.077	0.352	0.725	-0.124	0.179
Net_working_capital	-0.0175	0.061	-0.285	0.776	-0.138	0.103
Quick_ratio	0.0755	0.081	0.936	0.349	-0.083	0.233
Current_ratio	-0.0807	0.080	-1.006	0.314	-0.238	0.077
Cash_to_current_liabilities	-0.1173	0.073	-1.605	0.108	-0.261	0.026
Cash_to_average_cost_of_sales_per_day	0.0635	0.071	0.897	0.370	-0.075	0.202
Creditors_turnover	-0.0374	0.061	-0.613	0.540	-0.157	0.082
Debtors_turnover	0.0776	0.061	1.273	0.203	-0.042	0.197
Finished_goods_turnover	-0.1174	0.073	-1.607	0.108	-0.261	0.026
WIP_turnover	0.0792	0.070	1.133	0.257	-0.058	0.216

Figure 60 Model 7 Summary

TOL_by_TNW is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2958
Method:	MLE	Df Model:	20
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01748
Time:	13:22:19	Log-Likelihood:	-1508.7
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	6.427e-05

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3289	0.047	-27.993	0.000	-1.422	-1.236
PBDITA_to_total_income	-0.1308	0.059	-2.216	0.027	-0.246	-0.015
PAT_to_net_worth	-0.0418	0.063	-0.661	0.509	-0.166	0.082
Total_capital	0.0819	0.057	1.438	0.151	-0.030	0.193
Borrowings	0.0668	0.078	0.860	0.390	-0.085	0.219
Cumulative_retained_profits	-0.1928	0.064	-3.001	0.003	-0.319	-0.067
Total_term_liabilities__by__tangible_net_worth	0.0128	0.052	0.246	0.806	-0.090	0.115
Contingent_liabilities__by__Net_worth	-0.0568	0.054	-1.052	0.293	-0.163	0.049
Net_fixed_assets	0.0310	0.077	0.404	0.686	-0.119	0.181
Net_working_capital	-0.0165	0.061	-0.269	0.788	-0.137	0.104
Quick_ratio	0.0754	0.081	0.935	0.350	-0.083	0.233
Current_ratio	-0.0794	0.080	-0.990	0.322	-0.237	0.078
Cash_to_current_liabilities	-0.1122	0.072	-1.560	0.119	-0.253	0.029
Cash_to_average_cost_of_sales_per_day	0.0595	0.070	0.849	0.396	-0.078	0.197
Creditors_turnover	-0.0354	0.061	-0.583	0.560	-0.155	0.084
Debtors_turnover	0.0766	0.061	1.257	0.209	-0.043	0.196
Finished_goods_turnover	-0.1181	0.073	-1.617	0.106	-0.261	0.025
WIP_turnover	0.0795	0.070	1.137	0.255	-0.057	0.216
Raw_material_turnover	-0.0650	0.057	-1.147	0.251	-0.176	0.046

Figure 61 Model 8 Summary

Total_term_liabilities__by__tangible_net_worth is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2959
Method:	MLE	Df Model:	19
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01746
Time:	13:26:37	Log-Likelihood:	-1508.8
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	3.758e-05

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3280	0.047	-28.063	0.000	-1.421	-1.235
PBDITA_to_total_income	-0.1299	0.059	-2.205	0.027	-0.245	-0.014
PAT_to_net_worth	-0.0415	0.063	-0.656	0.512	-0.166	0.083
Total_capital	0.0804	0.057	1.420	0.156	-0.031	0.191
Borrowings	0.0732	0.073	1.002	0.316	-0.070	0.216
Cumulative_retained_profits	-0.1968	0.062	-3.167	0.002	-0.319	-0.075
Contingent_liabilities__by__Net_worth	-0.0564	0.054	-1.045	0.296	-0.162	0.049
Net_fixed_assets	0.0311	0.077	0.406	0.685	-0.119	0.181
Net_working_capital	-0.0167	0.061	-0.272	0.786	-0.137	0.104
Quick_ratio	0.0744	0.081	0.924	0.355	-0.083	0.232
Current_ratio	-0.0781	0.080	-0.975	0.329	-0.235	0.079
Cash_to_current_liabilities	-0.1128	0.072	-1.569	0.117	-0.254	0.028
Cash_to_average_cost_of_sales_per_day	0.0595	0.070	0.849	0.396	-0.078	0.197
Creditors_turnover	-0.0351	0.061	-0.578	0.563	-0.154	0.084
Debtors_turnover	0.0772	0.061	1.268	0.205	-0.042	0.196
Finished_goods_turnover	-0.1185	0.073	-1.623	0.105	-0.262	0.025
WIP_turnover	0.0804	0.070	1.152	0.249	-0.056	0.217
Raw_material_turnover	-0.0646	0.057	-1.142	0.253	-0.176	0.046
EPS	0.1528	0.098	1.553	0.120	-0.040	0.346

Figure 62 Model 9 Summary

Net_working_capital is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2960
Method:	MLE	Df Model:	18
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01744
Time:	13:27:32	Log-Likelihood:	-1508.8
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	2.151e-05

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3283	0.047	-28.072	0.000	-1.421	-1.236
PBDITA_to_total_income	-0.1292	0.059	-2.195	0.028	-0.244	-0.014
PAT_to_net_worth	-0.0416	0.063	-0.657	0.511	-0.166	0.083
Total_capital	0.0787	0.056	1.398	0.162	-0.032	0.189
Borrowings	0.0702	0.072	0.972	0.331	-0.071	0.212
Cumulative_retained_profits	-0.2008	0.060	-3.327	0.001	-0.319	-0.083
Contingent_liabilities__by__Net_worth	-0.0570	0.054	-1.056	0.291	-0.163	0.049
Net_fixed_assets	0.0329	0.076	0.431	0.666	-0.117	0.182
Quick_ratio	0.0740	0.081	0.919	0.358	-0.084	0.232
Current_ratio	-0.0852	0.076	-1.127	0.260	-0.233	0.063
Cash_to_current_liabilities	-0.1136	0.072	-1.582	0.114	-0.254	0.027
Cash_to_average_cost_of_sales_per_day	0.0599	0.070	0.855	0.393	-0.077	0.197
Creditors_turnover	-0.0357	0.061	-0.587	0.557	-0.155	0.083
Debtors_turnover	0.0785	0.061	1.294	0.196	-0.040	0.197
Finished_goods_turnover	-0.1180	0.073	-1.616	0.106	-0.261	0.025
WIP_turnover	0.0807	0.070	1.157	0.247	-0.056	0.217
Raw_material_turnover	-0.0645	0.057	-1.140	0.254	-0.175	0.046
EPS	0.1527	0.098	1.551	0.121	-0.040	0.346
Adjusted_EPS	-0.1548	0.094	-1.640	0.101	-0.340	0.030

Figure 63 Model 10 Summary

Net_fixed_assets is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2961
Method:	MLE	Df Model:	17
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01738
Time:	13:28:48	Log-Likelihood:	-1508.9
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	1.247e-05

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3268	0.047	-28.127	0.000	-1.419	-1.234
PBDITA_to_total_income	-0.1246	0.058	-2.156	0.031	-0.238	-0.011
PAT_to_net_worth	-0.0443	0.063	-0.703	0.482	-0.168	0.079
Total_capital	0.0854	0.054	1.581	0.114	-0.020	0.191
Borrowings	0.0884	0.058	1.517	0.129	-0.026	0.203
Cumulative_retained_profits	-0.1952	0.059	-3.314	0.001	-0.311	-0.080
Contingent_liabilities__by__Net_worth	-0.0553	0.054	-1.028	0.304	-0.161	0.050
Quick_ratio	0.0753	0.080	0.936	0.349	-0.082	0.233
Current_ratio	-0.0879	0.075	-1.167	0.243	-0.236	0.060
Cash_to_current_liabilities	-0.1134	0.072	-1.580	0.114	-0.254	0.027
Cash_to_average_cost_of_sales_per_day	0.0591	0.070	0.844	0.399	-0.078	0.196
Creditors_turnover	-0.0369	0.061	-0.608	0.543	-0.156	0.082
Debtors_turnover	0.0822	0.060	1.371	0.170	-0.035	0.200
Finished_goods_turnover	-0.1173	0.073	-1.608	0.108	-0.260	0.026
WIP_turnover	0.0816	0.070	1.171	0.242	-0.055	0.218
Raw_material_turnover	-0.0645	0.057	-1.141	0.254	-0.175	0.046
EPS	0.1563	0.098	1.594	0.111	-0.036	0.348
Adjusted_EPS	-0.1577	0.094	-1.675	0.094	-0.342	0.027

Figure 64 Model 11 Summary

Creditors_turnover is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2962
Method:	MLE	Df Model:	16
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01726
Time:	13:29:56	Log-Likelihood:	-1509.1
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	7.547e-06

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3270	0.047	-28.133	0.000	-1.419	-1.235
PBDITA_to_total_income	-0.1249	0.058	-2.160	0.031	-0.238	-0.012
PAT_to_net_worth	-0.0434	0.063	-0.689	0.491	-0.167	0.080
Total_capital	0.0856	0.054	1.585	0.113	-0.020	0.191
Borrowings	0.0868	0.058	1.490	0.136	-0.027	0.201
Cumulative_retained_profits	-0.1979	0.059	-3.368	0.001	-0.313	-0.083
Contingent_liabilities_by__Net_worth	-0.0538	0.054	-1.003	0.316	-0.159	0.051
Quick_ratio	0.0758	0.080	0.942	0.346	-0.082	0.234
Current_ratio	-0.0975	0.074	-1.321	0.186	-0.242	0.047
Cash_to_current_liabilities	-0.1153	0.072	-1.606	0.108	-0.256	0.025
Cash_to_average_cost_of_sales_per_day	0.0621	0.070	0.888	0.375	-0.075	0.199
Debtors_turnover	0.0694	0.056	1.233	0.217	-0.041	0.180
Finished_goods_turnover	-0.1179	0.073	-1.615	0.106	-0.261	0.025
WIP_turnover	0.0770	0.069	1.110	0.267	-0.059	0.213
Raw_material_turnover	-0.0719	0.055	-1.299	0.194	-0.180	0.037
EPS	0.1588	0.098	1.622	0.105	-0.033	0.351
Adjusted_EPS	-0.1577	0.094	-1.676	0.094	-0.342	0.027

Figure 65 Model 12 Summary

PAT_to_net_worth is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2963
Method:	MLE	Df Model:	15
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01710
Time:	13:31:08	Log-Likelihood:	-1509.3
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	4.628e-06

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3247	0.047	-28.180	0.000	-1.417	-1.233
PBDITA_to_total_income	-0.1385	0.054	-2.544	0.011	-0.245	-0.032
Total_capital	0.0858	0.054	1.589	0.112	-0.020	0.192
Borrowings	0.0866	0.058	1.487	0.137	-0.028	0.201
Cumulative_retained_profits	-0.1999	0.059	-3.400	0.001	-0.315	-0.085
Contingent_liabilities_by_Net_worth	-0.0564	0.054	-1.052	0.293	-0.161	0.049
Quick_ratio	0.0738	0.080	0.917	0.359	-0.084	0.232
Current_ratio	-0.0960	0.074	-1.301	0.193	-0.241	0.049
Cash_to_current_liabilities	-0.1154	0.072	-1.607	0.108	-0.256	0.025
Cash_to_average_cost_of_sales_per_day	0.0640	0.070	0.916	0.360	-0.073	0.201
Debtors_turnover	0.0673	0.056	1.197	0.231	-0.043	0.178
Finished_goods_turnover	-0.1196	0.073	-1.639	0.101	-0.263	0.023
WIP_turnover	0.0742	0.069	1.071	0.284	-0.062	0.210
Raw_material_turnover	-0.0771	0.055	-1.408	0.159	-0.185	0.030
EPS	0.1485	0.097	1.534	0.125	-0.041	0.338
Adjusted_EPS	-0.1643	0.094	-1.753	0.080	-0.348	0.019

Figure 66 Model 13 Summary

Cash_to_average_cost_of_sales_per_day is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2964
Method:	MLE	Df Model:	14
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01683
Time:	13:32:24	Log-Likelihood:	-1509.7
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	3.173e-06

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3263	0.047	-28.227	0.000	-1.418	-1.234
PBDITA_to_total_income	-0.1277	0.053	-2.406	0.016	-0.232	-0.024
Total_capital	0.0874	0.054	1.619	0.105	-0.018	0.193
Borrowings	0.0880	0.058	1.511	0.131	-0.026	0.202
Cumulative_retained_profits	-0.2005	0.059	-3.407	0.001	-0.316	-0.085
Contingent_liabilities__by__Net_worth	-0.0547	0.054	-1.021	0.307	-0.160	0.050
Quick_ratio	0.0745	0.080	0.926	0.354	-0.083	0.232
Current_ratio	-0.1005	0.074	-1.366	0.172	-0.245	0.044
Cash_to_current_liabilities	-0.0717	0.053	-1.357	0.175	-0.175	0.032
Debtors_turnover	0.0637	0.056	1.134	0.257	-0.046	0.174
Finished_goods_turnover	-0.1206	0.073	-1.652	0.099	-0.264	0.022
WIP_turnover	0.0720	0.069	1.039	0.299	-0.064	0.208
Raw_material_turnover	-0.0779	0.055	-1.421	0.155	-0.185	0.030
EPS	0.1469	0.097	1.517	0.129	-0.043	0.337
Adjusted_EPS	-0.1619	0.094	-1.729	0.084	-0.346	0.022

Figure 67 Model 14 Summary

Quick_ratio is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2965
Method:	MLE	Df Model:	13
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01655
Time:	13:33:35	Log-Likelihood:	-1510.2
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	2.146e-06

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3251	0.047	-28.229	0.000	-1.417	-1.233
PBDITA_to_total_income	-0.1257	0.053	-2.373	0.018	-0.230	-0.022
Total_capital	0.0882	0.054	1.635	0.102	-0.018	0.194
Borrowings	0.0828	0.058	1.426	0.154	-0.031	0.196
Cumulative_retained_profits	-0.1990	0.059	-3.383	0.001	-0.314	-0.084
Contingent_liabilities__by__Net_worth	-0.0546	0.054	-1.021	0.307	-0.160	0.050
Current_ratio	-0.0511	0.050	-1.020	0.308	-0.149	0.047
Cash_to_current_liabilities	-0.0572	0.050	-1.135	0.256	-0.156	0.042
Debtors_turnover	0.0442	0.052	0.847	0.397	-0.058	0.147
Finished_goods_turnover	-0.1157	0.073	-1.593	0.111	-0.258	0.027
WIP_turnover	0.0789	0.069	1.148	0.251	-0.056	0.213
Raw_material_turnover	-0.0706	0.054	-1.302	0.193	-0.177	0.036
EPS	0.1474	0.097	1.523	0.128	-0.042	0.337
Adjusted_EPS	-0.1609	0.094	-1.717	0.086	-0.345	0.023

Figure 68 Model 15 Summary

Debtors_turnover is still insignificant, therefore, we will eliminate it

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2966
Method:	MLE	Df Model:	12
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01632
Time:	13:34:28	Log-Likelihood:	-1510.5
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	1.331e-06

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3265	0.047	-28.274	0.000	-1.418	-1.235
PBDITA_to_total_income	-0.1224	0.053	-2.318	0.020	-0.226	-0.019
Total_capital	0.0870	0.054	1.613	0.107	-0.019	0.193
Borrowings	0.0880	0.058	1.525	0.127	-0.025	0.201
Cumulative_retained_profits	-0.1980	0.059	-3.367	0.001	-0.313	-0.083
Contingent_liabilities__by__Net_worth	-0.0520	0.053	-0.975	0.330	-0.157	0.053
Current_ratio	-0.0521	0.050	-1.039	0.299	-0.150	0.046
Cash_to_current_liabilities	-0.0538	0.050	-1.072	0.284	-0.152	0.045
Finished_goods_turnover	-0.1163	0.073	-1.601	0.109	-0.259	0.026
WIP_turnover	0.0869	0.068	1.279	0.201	-0.046	0.220
Raw_material_turnover	-0.0598	0.053	-1.137	0.255	-0.163	0.043
EPS	0.1483	0.097	1.531	0.126	-0.042	0.338
Adjusted_EPS	-0.1585	0.094	-1.690	0.091	-0.342	0.025

Figure 69 Model 16 Summary

Contingent_liabilities__by__Net_worth is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2967
Method:	MLE	Df Model:	11
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01601
Time:	13:35:22	Log-Likelihood:	-1511.0
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	8.872e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3248	0.047	-28.275	0.000	-1.417	-1.233
PBDITA_to_total_income	-0.1241	0.053	-2.350	0.019	-0.228	-0.021
Total_capital	0.0814	0.054	1.514	0.130	-0.024	0.187
Borrowings	0.0745	0.056	1.328	0.184	-0.035	0.184
Cumulative_retained_profits	-0.1988	0.059	-3.381	0.001	-0.314	-0.084
Current_ratio	-0.0482	0.050	-0.965	0.334	-0.146	0.050
Cash_to_current_liabilities	-0.0556	0.050	-1.109	0.268	-0.154	0.043
Finished_goods_turnover	-0.1173	0.073	-1.615	0.106	-0.260	0.025
WIP_turnover	0.0866	0.068	1.274	0.203	-0.047	0.220
Raw_material_turnover	-0.0627	0.053	-1.193	0.233	-0.166	0.040
EPS	0.1460	0.097	1.507	0.132	-0.044	0.336
Adjusted_EPS	-0.1617	0.094	-1.726	0.084	-0.345	0.022

Figure 70 Model 17 Summary

Current_ratio is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2968
Method:	MLE	Df Model:	10
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01570
Time:	13:36:12	Log-Likelihood:	-1511.5
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	5.660e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3225	0.047	-28.284	0.000	-1.414	-1.231
PBDITA_to_total_income	-0.1295	0.053	-2.460	0.014	-0.233	-0.026
Total_capital	0.0782	0.054	1.458	0.145	-0.027	0.183
Borrowings	0.0802	0.056	1.436	0.151	-0.029	0.190
Cumulative_retained_profits	-0.2064	0.058	-3.539	0.000	-0.321	-0.092
Cash_to_current_liabilities	-0.0679	0.049	-1.396	0.163	-0.163	0.027
Finished_goods_turnover	-0.1173	0.073	-1.616	0.106	-0.259	0.025
WIP_turnover	0.0854	0.068	1.258	0.208	-0.048	0.219
Raw_material_turnover	-0.0700	0.052	-1.345	0.178	-0.172	0.032
EPS	0.1508	0.097	1.557	0.119	-0.039	0.341
Adjusted_EPS	-0.1655	0.094	-1.766	0.077	-0.349	0.018

Figure 71 Model 18 Summary

WIP_turnover is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2969
Method:	MLE	Df Model:	9
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01519
Time:	13:37:13	Log-Likelihood:	-1512.3
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	4.572e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3236	0.047	-28.324	0.000	-1.415	-1.232
PBDITA_to_total_income	-0.1303	0.053	-2.473	0.013	-0.234	-0.027
Total_capital	0.0762	0.054	1.420	0.156	-0.029	0.181
Borrowings	0.0840	0.056	1.509	0.131	-0.025	0.193
Cumulative_retained_profits	-0.2061	0.058	-3.532	0.000	-0.320	-0.092
Cash_to_current_liabilities	-0.0635	0.048	-1.311	0.190	-0.159	0.031
Finished_goods_turnover	-0.0606	0.055	-1.097	0.273	-0.169	0.048
Raw_material_turnover	-0.0574	0.051	-1.128	0.260	-0.157	0.042
EPS	0.1468	0.097	1.517	0.129	-0.043	0.336
Adjusted_EPS	-0.1595	0.093	-1.706	0.088	-0.343	0.024

Figure 72 Model 19 Summary

Finished_goods_turnover is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2970
Method:	MLE	Df Model:	8
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01479
Time:	13:38:15	Log-Likelihood:	-1512.9
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	3.064e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3192	0.046	-28.379	0.000	-1.410	-1.228
PBDITA_to_total_income	-0.1293	0.053	-2.454	0.014	-0.233	-0.026
Total_capital	0.0751	0.054	1.400	0.161	-0.030	0.180
Borrowings	0.0862	0.056	1.548	0.122	-0.023	0.195
Cumulative_retained_profits	-0.2069	0.058	-3.545	0.000	-0.321	-0.093
Cash_to_current_liabilities	-0.0643	0.048	-1.327	0.185	-0.159	0.031
Raw_material_turnover	-0.0700	0.050	-1.408	0.159	-0.168	0.027
EPS	0.1490	0.097	1.541	0.123	-0.040	0.338
Adjusted_EPS	-0.1670	0.093	-1.791	0.073	-0.350	0.016

Figure 73 Model 20 Summary

Cash_to_current_liabilities is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2971
Method:	MLE	Df Model:	7
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01421
Time:	13:39:48	Log-Likelihood:	-1513.8
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	2.524e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3229	0.046	-28.494	0.000	-1.414	-1.232
PBDITA_to_total_income	-0.1387	0.052	-2.656	0.008	-0.241	-0.036
Total_capital	0.0733	0.054	1.368	0.171	-0.032	0.178
Borrowings	0.0964	0.055	1.753	0.080	-0.011	0.204
Cumulative_retained_profits	-0.2130	0.058	-3.671	0.000	-0.327	-0.099
Raw_material_turnover	-0.0710	0.050	-1.429	0.153	-0.168	0.026
EPS	0.1512	0.097	1.565	0.118	-0.038	0.340
Adjusted_EPS	-0.1697	0.093	-1.821	0.069	-0.352	0.013

Figure 74 Model 21 Summary

Total_capital is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2972
Method:	MLE	Df Model:	6
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01360
Time:	13:40:50	Log-Likelihood:	-1514.7
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	2.035e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3171	0.046	-28.545	0.000	-1.408	-1.227
PBDITA_to_total_income	-0.1281	0.052	-2.485	0.013	-0.229	-0.027
Borrowings	0.1350	0.047	2.869	0.004	0.043	0.227
Cumulative_retained_profits	-0.2177	0.058	-3.740	0.000	-0.332	-0.104
Raw_material_turnover	-0.0715	0.050	-1.440	0.150	-0.169	0.026
EPS	0.1535	0.097	1.585	0.113	-0.036	0.343
Adjusted_EPS	-0.1754	0.093	-1.880	0.060	-0.358	0.007

Figure 75 Model 22 Summary

Raw_material_turnover is still insignificant, therefore, we will eliminate it

Logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2973
Method:	MLE	Df Model:	5
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01292
Time:	13:41:35	Log-Likelihood:	-1515.7
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	1.738e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3111	0.046	-28.592	0.000	-1.401	-1.221
PBDITA_to_total_income	-0.1193	0.051	-2.326	0.020	-0.220	-0.019
Borrowings	0.1259	0.047	2.698	0.007	0.034	0.217
Cumulative_retained_profits	-0.2211	0.058	-3.796	0.000	-0.335	-0.107
EPS	0.1494	0.097	1.541	0.123	-0.041	0.340
Adjusted_EPS	-0.1846	0.093	-1.978	0.048	-0.367	-0.002

Figure 76 Model 23 Summary

EPS is still insignificant, therefore, we will eliminate it

.logit Regression Results

Dep. Variable:	Default	No. Observations:	2979
Model:	Logit	Df Residuals:	2974
Method:	MLE	Df Model:	4
Date:	Wed, 16 Oct 2024	Pseudo R-squ.:	0.01217
Time:	13:42:18	Log-Likelihood:	-1516.9
converged:	True	LL-Null:	-1535.6
Covariance Type:	nonrobust	LLR p-value:	1.501e-07

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-1.3116	0.046	-28.616	0.000	-1.401	-1.222
PBDITA_to_total_income	-0.1112	0.051	-2.183	0.029	-0.211	-0.011
Borrowings	0.1275	0.047	2.741	0.006	0.036	0.219
Cumulative_retained_profits	-0.1982	0.056	-3.533	0.000	-0.308	-0.088
Adjusted_EPS	-0.0698	0.056	-1.250	0.211	-0.179	0.040

Figure 77 Model 24 Summary

We created predictions using the last model.

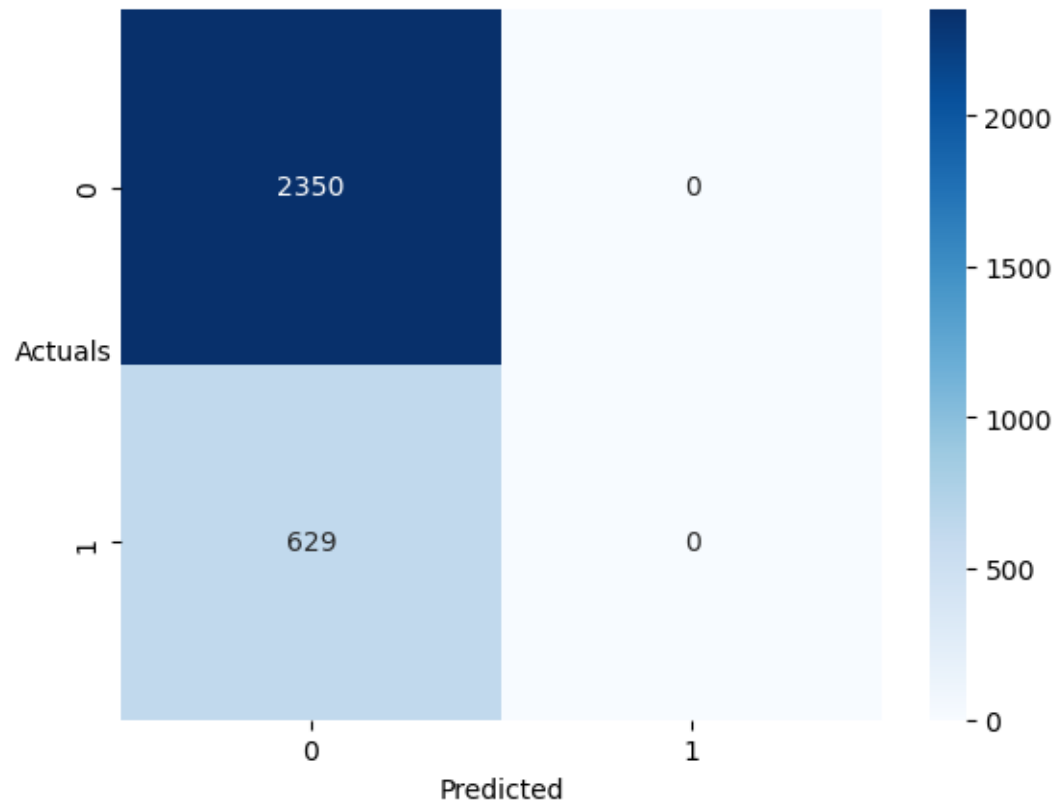


Figure 78 Confusion Matrix

Choosing the optimal threshold

We find tpr, fpr, thresholds based on ROC Curve. Optimal threshold is calculated as $tpr - fpr$.

Validating on the train set with revised threshold

Our new confusion matrix looks like:

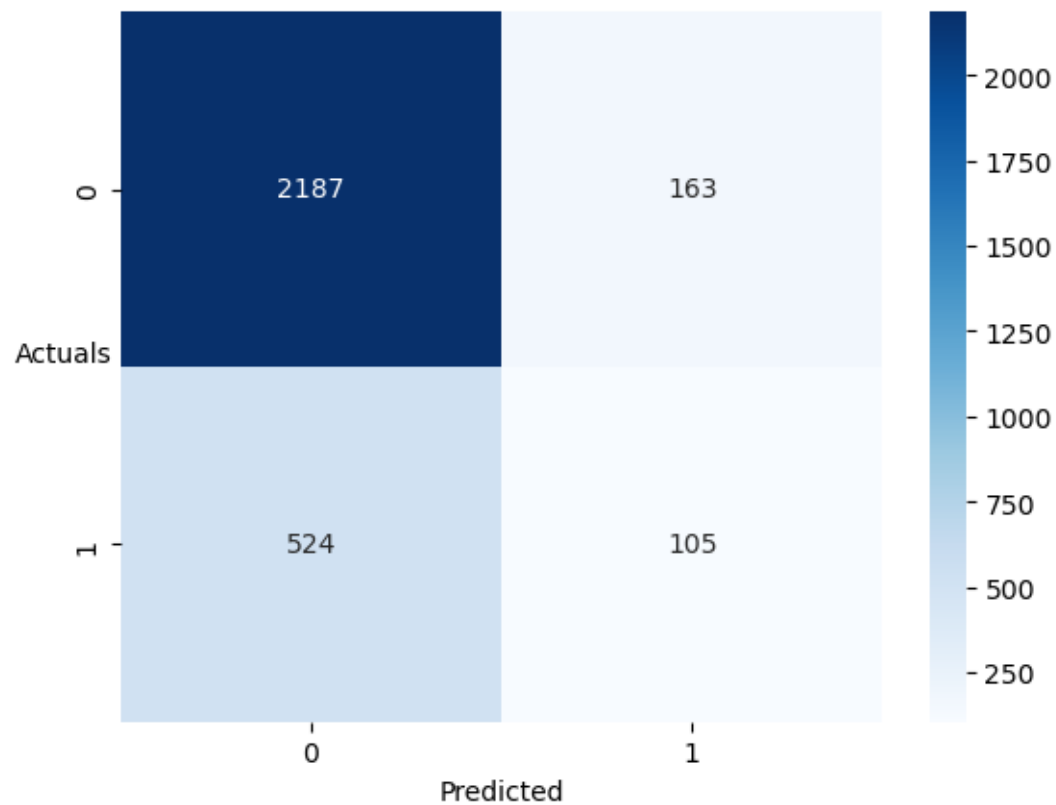


Figure 79 Confusion Matrix after selecting Optimum threshold

Let us now go ahead and print the classification report to check the various other parameters.

	precision	recall	f1-score	support
0.0	0.807	0.931	0.864	2350
1.0	0.392	0.167	0.234	629
accuracy			0.769	2979
macro avg	0.599	0.549	0.549	2979
weighted avg	0.719	0.769	0.731	2979

Figure 80 Classification Report on train data

Validating on the test set

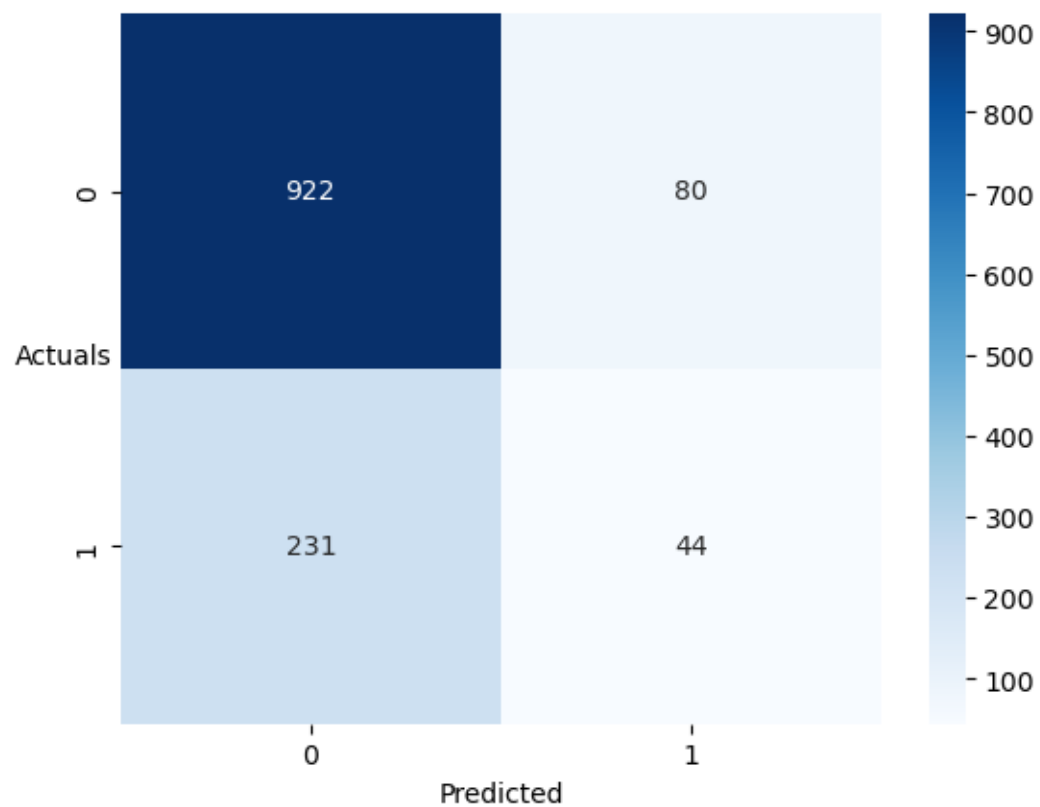


Figure 81 Confusion Matrix on test data

Our classification report looks like:

	precision	recall	f1-score	support
0.0	0.800	0.920	0.856	1002
1.0	0.355	0.160	0.221	275
accuracy			0.756	1277
macro avg	0.577	0.540	0.538	1277
weighted avg	0.704	0.756	0.719	1277

Figure 82 Classification Report on test data

Linear Discriminant Analysis

We create LDA model on train and test data

Classification report on train data :

	precision	recall	f1-score	support
0.0	0.79	1.00	0.88	2350
1.0	1.00	0.00	0.00	629
accuracy			0.79	2979
macro avg	0.89	0.50	0.44	2979
weighted avg	0.83	0.79	0.70	2979

Figure 83 LDA Classification Report on train data

Classification report on test data:

	precision	recall	f1-score	support
0.0	0.78	1.00	0.88	1002
1.0	0.00	0.00	0.00	275
accuracy			0.78	1277
macro avg	0.39	0.50	0.44	1277
weighted avg	0.62	0.78	0.69	1277

Figure 84 LDA Classification Report on test data

Choosing the optimal threshold

We find tpr, fpr, thresholds based on ROC Curve. Optimal threshold is calculated as $tpr - fpr$.

Validating on the train set with revised threshold

Our new confusion matrix on train data looks like:

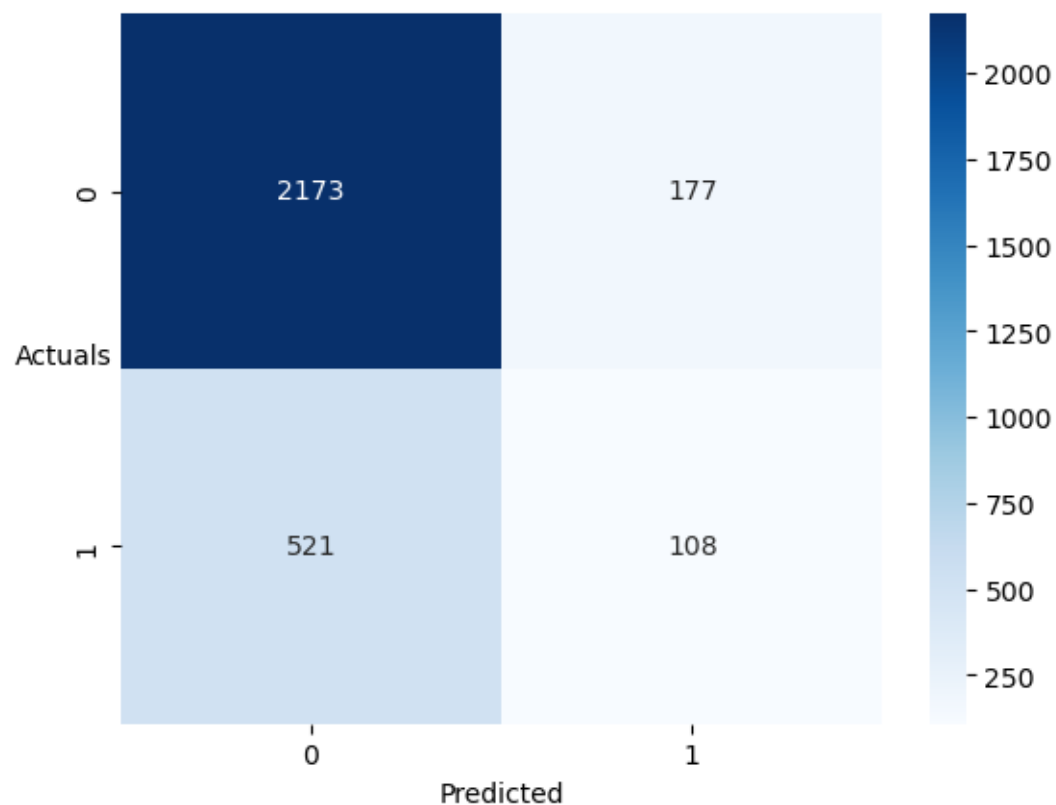


Figure 85 LDA Confusion matrix on train data

	precision	recall	f1-score	support
0.0	0.807	0.925	0.862	2350
1.0	0.379	0.172	0.236	629
accuracy			0.766	2979
macro avg	0.593	0.548	0.549	2979
weighted avg	0.716	0.766	0.730	2979

Figure 86 LDA Classification report on train data

Confusion Matrix on test data:

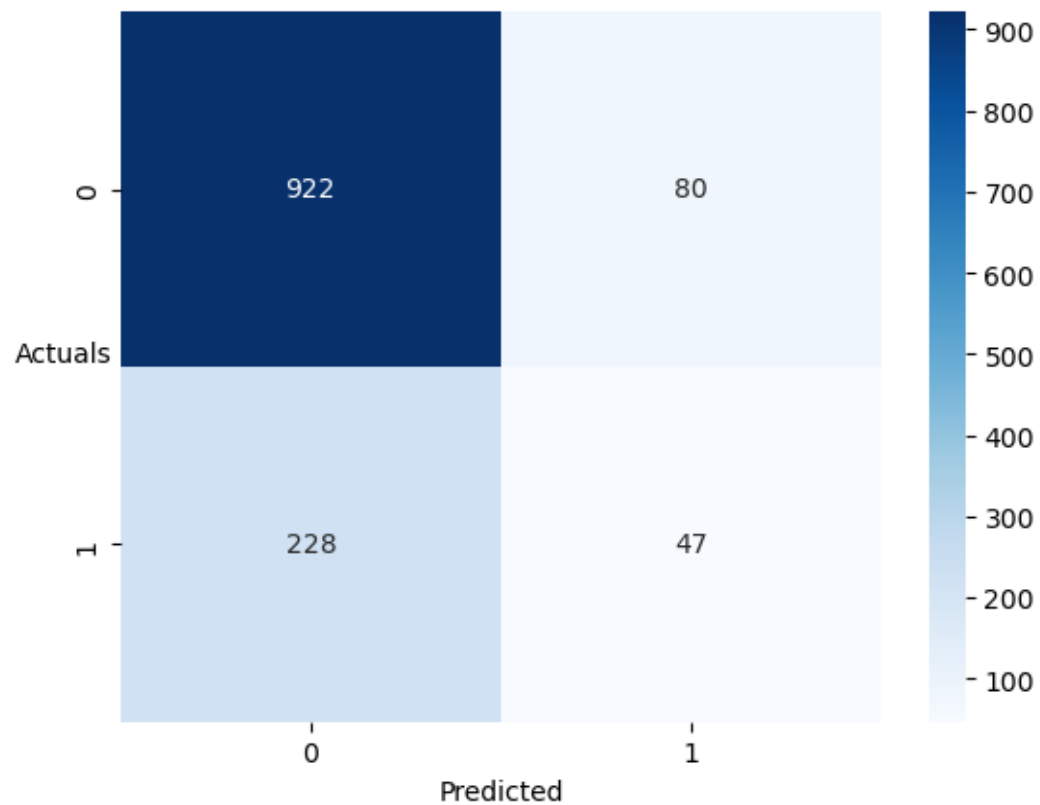


Figure 87 Confusion Matrix on test data

	precision	recall	f1-score	support
0.0	0.802	0.920	0.857	1002
1.0	0.370	0.171	0.234	275
accuracy			0.759	1277
macro avg	0.586	0.546	0.545	1277
weighted avg	0.709	0.759	0.723	1277

Figure 88 LDA classification report on test data

Random Forest Classifier

We create Random forest classifier on train and test data

We use GridsearchCV with following parameters: param_grid = {

'max_depth': [3, 5, 7],

'min_samples_leaf': [5, 10, 15],

```
'min_samples_split': [15, 30, 45],
'n_estimators': [25, 50]
}
```

Results on train data

	precision	recall	f1-score	support
0.0	0.80	1.00	0.89	2350
1.0	0.84	0.05	0.09	629
accuracy			0.80	2979
macro avg	0.82	0.52	0.49	2979
weighted avg	0.81	0.80	0.72	2979

Figure 89 Random Forest Classification Report on train data

Results of test data

	precision	recall	f1-score	support
0.0	0.79	0.99	0.88	1002
1.0	0.40	0.01	0.03	275
accuracy			0.78	1277
macro avg	0.59	0.50	0.45	1277
weighted avg	0.70	0.78	0.69	1277

Figure 90 Random Forest Test data

Check the most important features in the final model and draw inferences

$$\text{default} = -1.311618710470406 * (\text{PBDITA_to_total_income}) + -0.11119945913392916 * (\text{Borrowings}) + 0.12752993242608837 * (\text{Cumulative_retained_profits}) + -0.1982275820294953 * (\text{Adjusted_EPS})$$

-- 1 unit increase in the PBDITA_to_total_income lead to a 1.3 times decrease in the default

-- 1 unit increase in the borrowings lead to a 0.11 times decrease in the default

--1 unit increase in the Cumulative_retained_profits lead to a 0.12 times increase in the default

--1 unit increase in the Adjusted_EPS lead to a 0.19 times decrease in the default

Metrics of Choice (Justify the evaluation metrics)

We choose the evaluation metrics to ensure reliability and effectiveness of the classification models. We optimize the models to increase accuracy of predictions.

Let's start with accuracy, which calculates the ratio of correctly predicted outcomes compared to the total number of predictions. Another metric is precision, which evaluates the proportion of true positive predicted outcomes out of all predicted positive outcomes. Next, recall calculates the proportion of true positive predictions out of all actual positive predictions.

Model performance check across different metrics

We can see overfitting on train data and test data for Random Forest Classifier model.

We can see overfitting on train data for Linear Discriminant Analysis model.

Bagging and KNN show significant difference between performance metrics of train and test data.

Accuracy, Precision, Recall and F-1 Score for LDA and Random Forest models are almost similar

precision and the F1 score are most important. As a result, Logistic Regression therefore is probably a better choice as it provides the better accuracy than other models. Logistic Regression has a better F1 score than that of others, the difference is minimal.

Actionable insights and recommendations

- We should work on increasing PBDITA_to_total_income to decrease default
- We should work on increasing borrowings to decrease default
- We should work on decreasing Cumulative_retained_profits to decrease default
- We should work on increasing Adjusted_EPS to decrease default

PART B

Top 5 Rows

	Date	ITC_Limited	Bharti_Airtel	Tata_Motors	DLF_Limited	Yes_Bank
0	28-03-2016	217	316	386	114	173
1	04-04-2016	218	302	386	121	171
2	11-04-2016	215	308	374	120	171
3	18-04-2016	223	320	408	122	172
4	25-04-2016	214	319	418	122	175

Figure 91 Top 5 Rows

Shape of data

The number of rows (observations) is 418

The number of columns (variables) is 6

Column Type

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 418 entries, 0 to 417
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Date            418 non-null   object
1   ITC_Limited     418 non-null   int64
2   Bharti_Airtel  418 non-null   int64
3   Tata_Motors    418 non-null   int64
4   DLF_Limited     418 non-null   int64
5   Yes_Bank       418 non-null   int64
```

Figure 92 Column Types

We have 6 variables. 1 Date variable and 5 stock variables of integer type.

Statistical Summary

	ITC_Limited	Bharti_Airtel	Tata_Motors	DLF_Limited	Yes_Bank
count	418.000000	418.000000	418.000000	418.000000	418.000000
mean	278.964115	528.260766	368.617225	276.827751	124.442584
std	75.114405	226.507879	182.024419	156.280781	130.090884
min	156.000000	261.000000	65.000000	110.000000	11.000000
25%	224.250000	334.000000	186.000000	166.250000	16.000000
50%	265.500000	478.000000	399.500000	213.000000	30.000000
75%	304.000000	706.750000	466.000000	360.500000	249.750000
max	493.000000	1236.000000	1035.000000	928.000000	397.000000

Figure 93 Statistical Summary

We can see that we have 418 records.

Average of Bharti_Airtel is highest and Yes_bank has lowest average.

Standard deviation is lowest for ITC_Limited and highest for Bharti_Airtel.

Minimum value for YES_Bank is smallest and Bharti_Airtel is highest.

Maximum value for YES_Bank is smallest and Bharti_Airtel is highest.

Draw a Stock Price Graph (Stock Price vs Time) for the given stocks-
Write observations

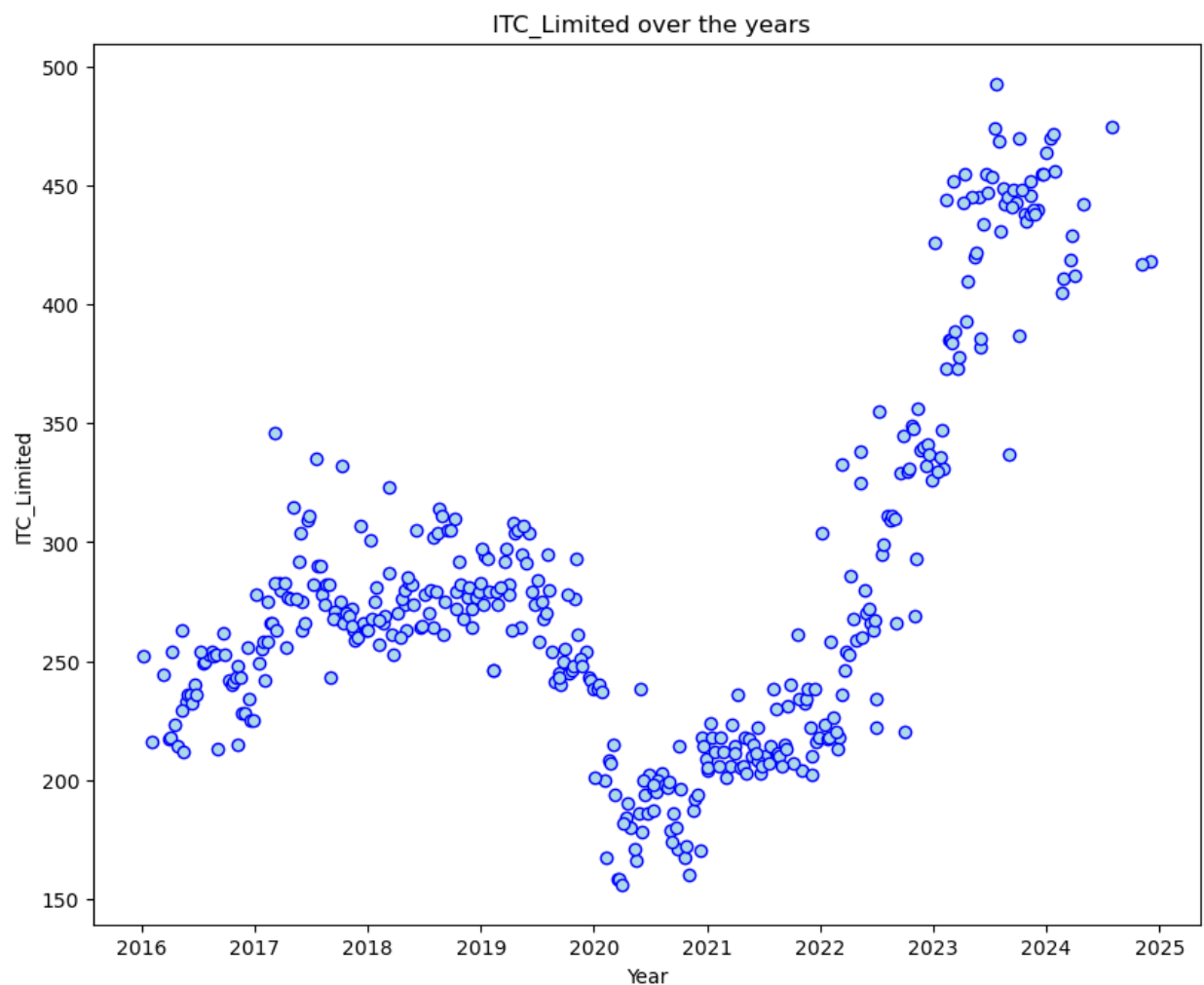


Figure 94 ITC_Limited

We are analyzing data for ITC_Limited.

We have data for years 2016 to 2025.

In initial years, from 2016 – 2019. Price started from 200 and reached 350 in 2019.

In 2020, it came down to 150 and there was exponential increase till 2025.

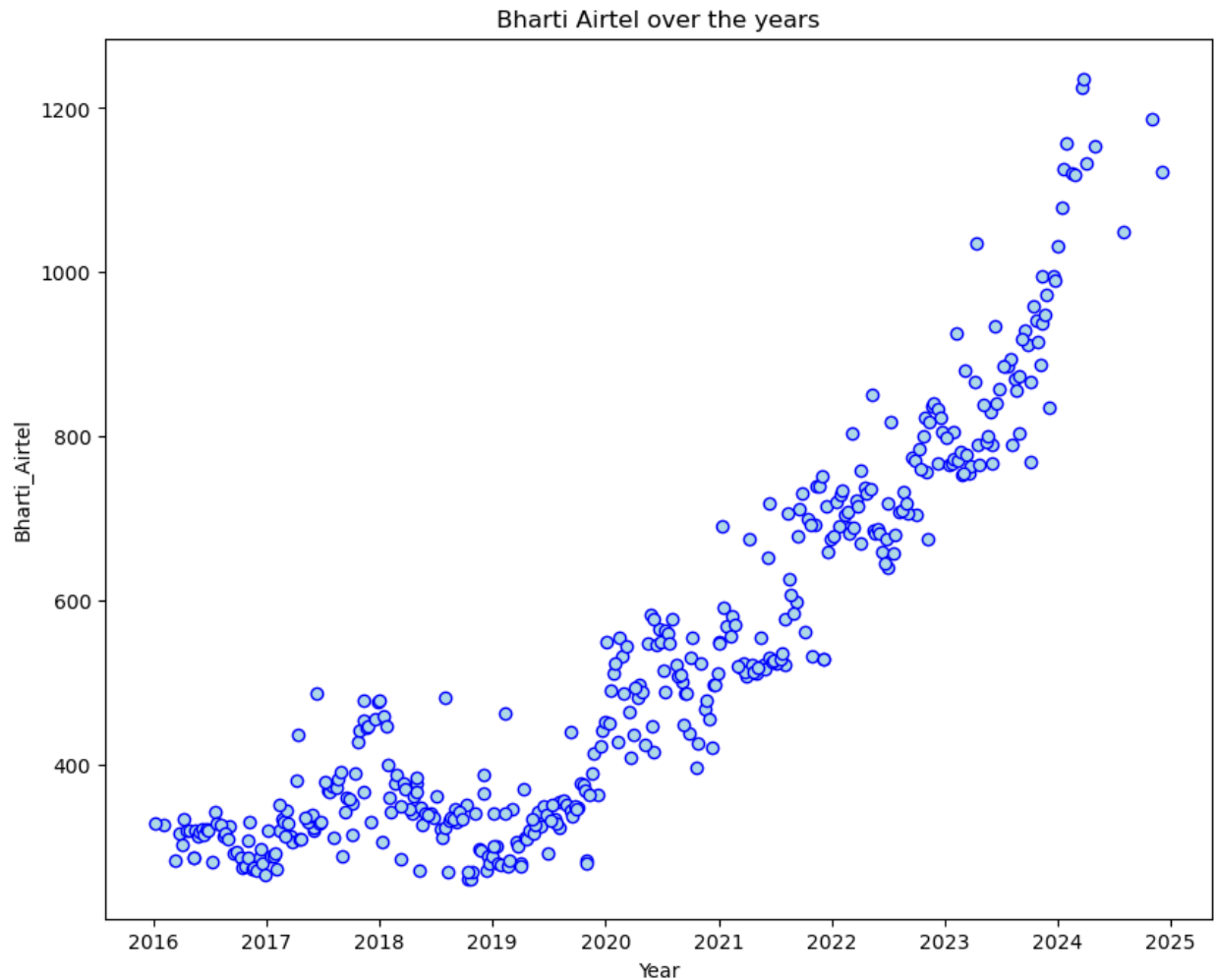


Figure 95 Bharti_Airtel

We are analyzing data for Bharti_Airtel.

We have data for years 2016 to 2025.

In initial years, from 2016 – 2019. Price stayed in lower range mostly upto 400 and reached upto 550 in 2018 .

In 2020, there was exponential increase till 2025.

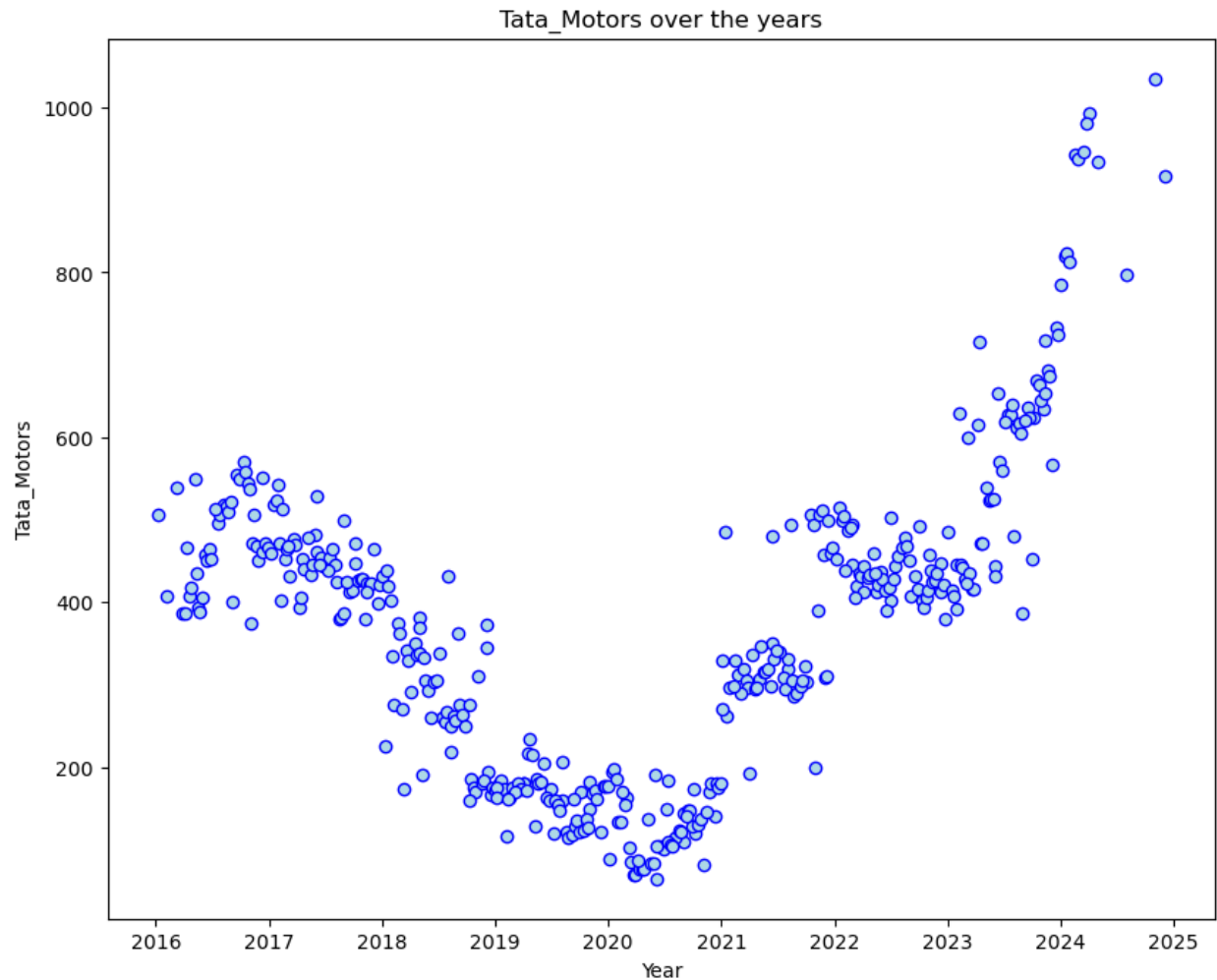


Figure 96 TATA_Motors

We are analyzing data for TATA_Motors.

We have data for years 2016 to 2025.

In initial years, from 2016 – 2019. Price started from 400 and reached below 200 in 2019.

In 2020, it came down to 150 and there was exponential increase till 2025.

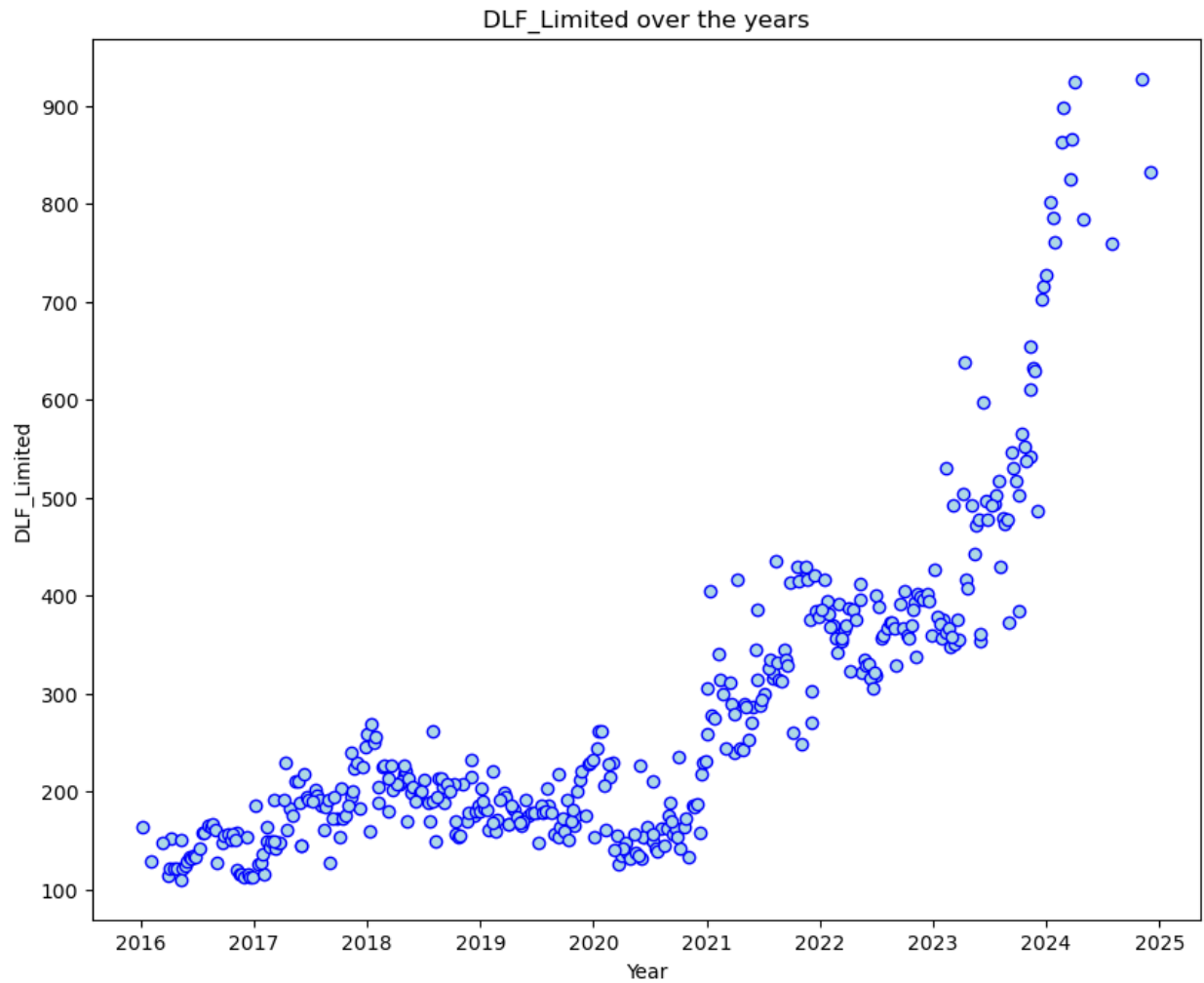


Figure 97 DLF_Limited

We are analyzing data for DLF_Limited.

We have data for years 2016 to 2025.

In initial years, from 2016 – 2021. Price started from 100 and ranged till 300

In 2021, it came down to 100 and there was exponential increase till 2025.

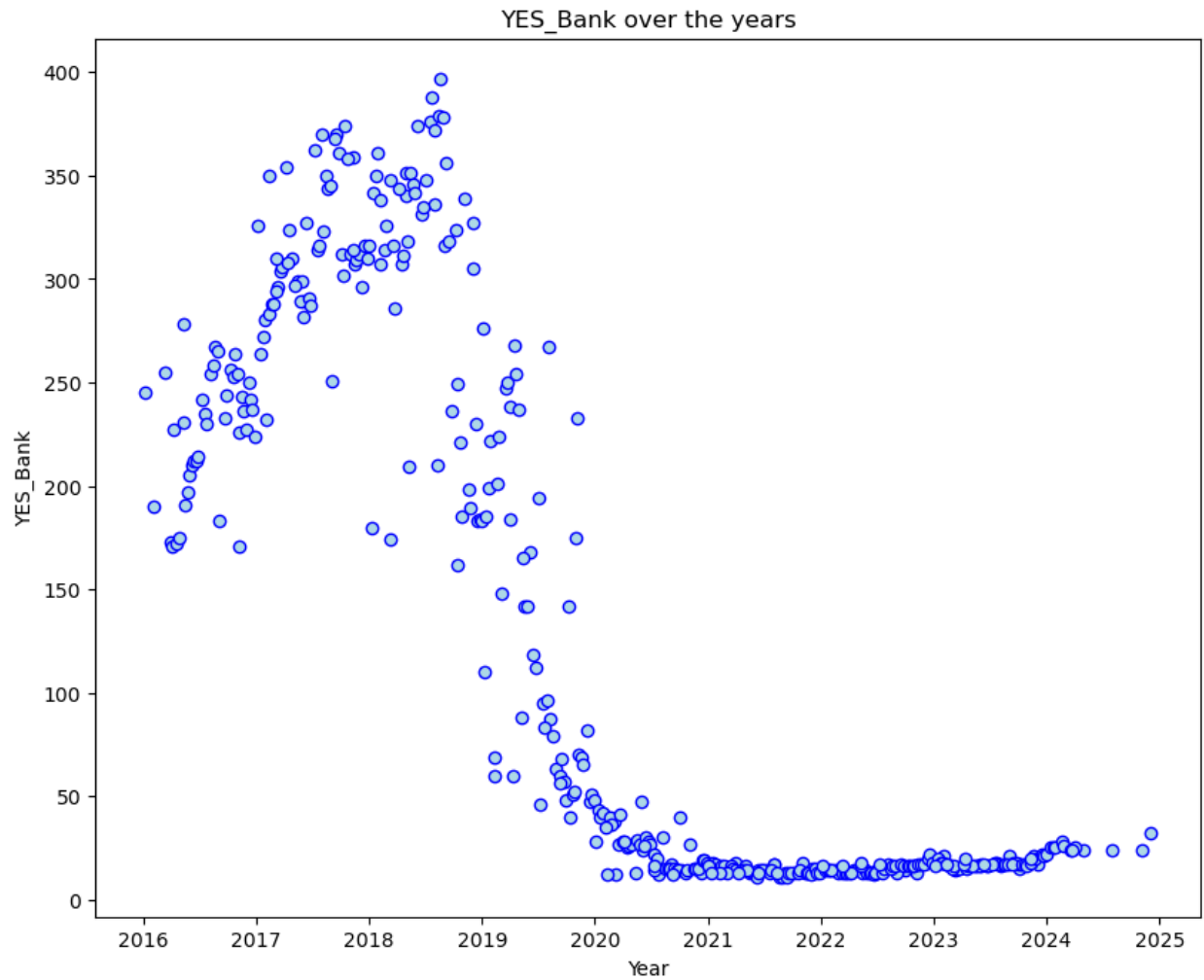


Figure 98 YES_Bank

We are analyzing data for YES_Bank.

We have data for years 2016 to 2025.

In initial years, from 2016 – 2019. Price started from 150 and reached 400 in 2019.

In 2020, it came down to below 50 and it stayed in the range till 2025.

Calculate Returns for all stocks

We take logarithmic value of data and find difference between records.

	ITC_Limited	Bharti_Airtel	Tata_Motors	DLF_Limited	Yes_Bank
0	NaN	NaN	NaN	NaN	NaN
1	0.004598	-0.045315	0.000000	0.059592	-0.011628
2	-0.013857	0.019673	-0.031582	-0.008299	0.000000
3	0.036534	0.038221	0.087011	0.016529	0.005831
4	-0.041196	-0.003130	0.024214	0.000000	0.017291

Figure 99 Stock Returns

Looking at top 5 rows

We have weekly data for stocks and top 5 weeks analysis tells us:

For ITC_Limited, We see there is a positive change and then negative change

For Bharti_Airtel, We see a negative change initially and positive change twice then negative again

For TATA_Motors, we have no change then negative change then all positive change.

For DLF_Limited, we have no change then negative change then all positive change.

For YES_Bank, we have no change then negative change then all positive change.

Calculate the Mean and Standard Deviation for the returns of all stocks

Means of Stock

```
ITC_Limited      0.001634
Bharti_Airtel    0.003271
Tata_Motors      0.002234
DLF_Limited      0.004863
Yes_Bank         -0.004737
dtype: float64
```

Figure 100 Means of Stocks

We see that average of stock returns is positive for all stocks except YES_Bank. YES_Bank has lowest average in negative numbers and DLF_Limited has highest average.

Standard Deviation of Stocks

```
ITC_Limited      0.035904
Bharti_Airtel    0.038728
Tata_Motors      0.060484
DLF_Limited      0.057785
Yes_Bank         0.093879
dtype: float64
```

Figure 101 Standard deviation of stocks

We see that standard deviation for all stocks is positive. YES_Bank has highest standard deviation and ITC_limited has lowest standard deviation.

Draw a plot of Mean vs Standard Deviation for all stock returns

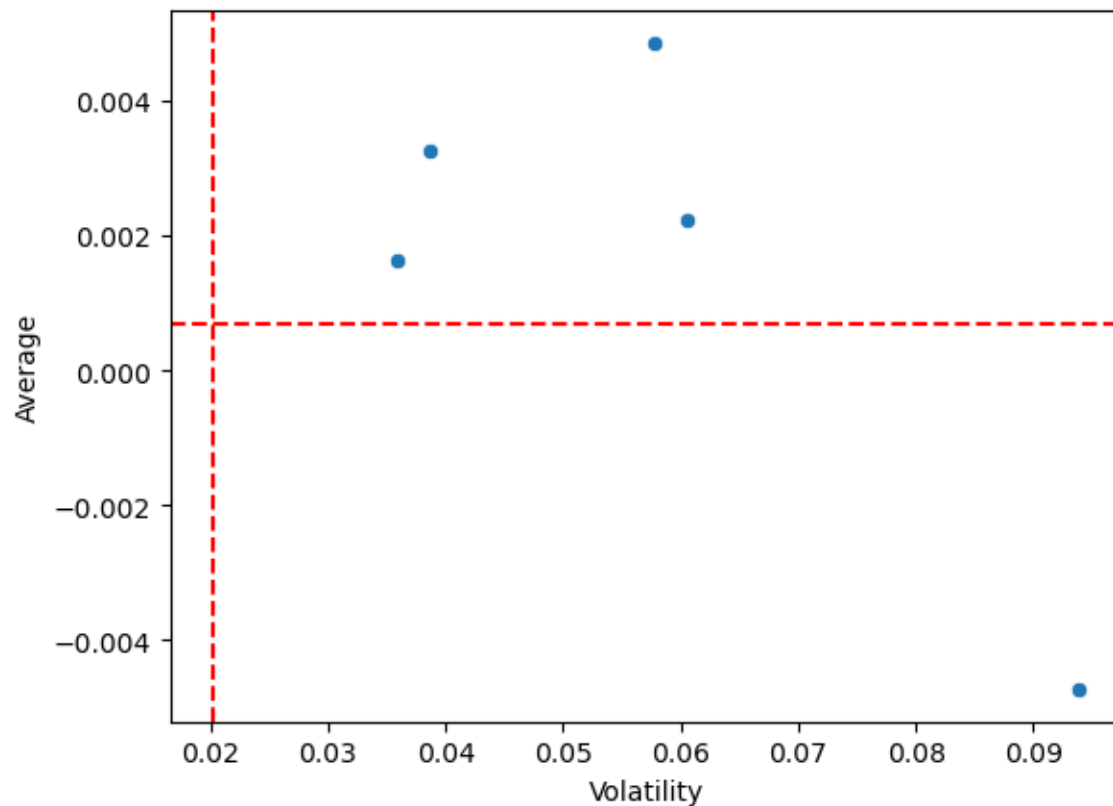


Figure 102 Mean vs Standard deviation of stocks

Write observations and inferences

Standard deviation measures market volatility by calculating how stock prices are dispersed from average price.

Standard deviation helps in calculating risk of the stock. Higher the standard deviation, higher the risk.

We can see that average of stock returns for 4 stocks is positive and volatility is lower.

We can see the one of stocks has very high volatility and low average. Hence, it doesn't play important role in portfolio.