

Finance and Risk Analytics

Business Report

Athisya Nadar

February 2022

Post Graduate Program in Data Science and Business Analytics

Athisya@gmail.com

Problem Statement

Businesses or companies can fall prey to default if they are not able to keep up their debt obligations. Defaults will lead to a lower credit rating for the company which in turn reduces its chances of getting credit in the future and may have to pay higher interests on existing debts as well as any new obligations. From an investor's point of view, he would want to invest in a company if it is capable of handling its financial obligations, can grow quickly, and is able to manage the growth scale.

A balance sheet is a financial statement of a company that provides a snapshot of what a company owns, owes, and the amount invested by the shareholders. Thus, it is an important tool that helps evaluate the performance of a business.

Data that is available includes information from the financial statement of the companies for the previous year (2015). Also, information about the Networth of the company in the following year (2016) is provided which can be used to drive the labeled field.

Let us now go ahead and read the dataset and check the first five rows of the dataset.

```
Company.head()
```

	Co_Code	Co_Name	Networth_Next_Year	Equity_Paid_Up	Networth	Capital_Employed	Total_Debt	Gross_Block_	Net_Working_Capital_	Current_Asse
0	16974	Hind.Cables	-8021.60	419.36	-7027.48	-1007.24	5936.03	474.30	-1076.34	40
1	21214	Tata Tele. Mah.	-3986.19	1954.93	-2968.08	4458.20	7410.18	9070.86	-1098.88	486
2	14852	ABG Shipyard	-3192.58	53.84	506.86	7714.68	6944.54	1281.54	4496.25	9097
3	2439	GTL	-3054.51	157.30	-623.49	2353.88	2326.05	1033.69	-2612.42	1034
4	23505	Bharati Defence	-2967.36	50.30	-1070.83	4675.33	5740.90	1084.20	1836.23	4685

The number of rows (observations) is 3586

The number of columns (variables) is 67

Renaming the columns helps us to easily use variables for analysis by replacing “.”, “[.]”,etc with _

Next let us look at the variables, the data types, null values, missing values.

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 3586 entries, 0 to 3585
Data columns (total 67 columns):
 #   Column           Non-Null Count Dtype
 ---  ----
 0   Co_Code          3586 non-null  int64
 1   Co_Name          3586 non-null  object
 2   Networth_Next_Year 3586 non-null  float64
 3   Equity_Paid_Up   3586 non-null  float64
 4   Networth          3586 non-null  float64
 5   Capital_Employed 3586 non-null  float64
 6   Total_Debt        3586 non-null  float64
 7   Gross_Block_      3586 non-null  float64
 8   Net_Working_Capital_ 3586 non-null  float64
 9   Current_Assets_   3586 non-null  float64
 10  Current_Liabilities_and_Provisions_ 3586 non-null  float64
 11  Total_Assets_to_Liabilities_    3586 non-null  float64
 12  Gross_Sales       3586 non-null  float64
 13  Net_Sales         3586 non-null  float64
 14  Other_Income      3586 non-null  float64
 15  Value_Of_Output   3586 non-null  float64
 16  Cost_of_Production 3586 non-null  float64
 17  Selling_Cost      3586 non-null  float64
 18  PBIDT             3586 non-null  float64
 19  PBDT              3586 non-null  float64
 20  PBIT               3586 non-null  float64

 21  PBT               3586 non-null  float64
 22  PAT               3586 non-null  float64
 23  Adjusted_PAT      3586 non-null  float64
 24  CP                3586 non-null  float64
 25  Revenue_earnings_in_forex 3586 non-null  float64
 26  Revenue_expenses_in_forex 3586 non-null  float64
 27  Capital_expenses_in_forex 3586 non-null  float64
 28  Book_Value_Unit_Curr   3586 non-null  float64
 29  Book_Value_Adj_Unit_Curr 3582 non-null  float64
 30  Market_Capitalisation 3586 non-null  float64
 31  CEPS_annualised_Unit_Curr 3586 non-null  float64
 32  Cash_Flow_From_Operating_Activities 3586 non-null  float64
 33  Cash_Flow_From_Investing_Activities 3586 non-null  float64
 34  Cash_Flow_From_Financing_Activities 3586 non-null  float64
 35  ROG_Net_Worth_perc   3586 non-null  float64
 36  ROG_Capital_Employed_perc 3586 non-null  float64
 37  ROG_Gross_Block_perc 3586 non-null  float64
 38  ROG_Gross_Sales_perc 3586 non-null  float64
 39  ROG_Net_Sales_perc   3586 non-null  float64
 40  ROG_Cost_of_Production_perc 3586 non-null  float64
 41  ROG_Total_Assets_perc 3586 non-null  float64
 42  ROG_PBIDT_perc      3586 non-null  float64
 43  ROG_PBDT_perc      3586 non-null  float64
 44  ROG_PBIT_perc       3586 non-null  float64
 45  ROG_PBT_perc       3586 non-null  float64

```

```

46 ROG_PAT_perc                         3586 non-null float64
47 ROG_CP_perc                          3586 non-null float64
48 ROG_Revenue_earnings_in_forex_perc   3586 non-null float64
49 ROG_Revenue_expenses_in_forex_perc   3586 non-null float64
50 ROG_Market_Capitalisation_perc      3586 non-null float64
51 Current_Ratio_Latest_                3585 non-null float64
52 Fixed_Assets_Ratio_Latest_          3585 non-null float64
53 Inventory_Ratio_Latest_             3585 non-null float64
54 Debtors_Ratio_Latest_               3585 non-null float64
55 Total_Asset_Turnover_Ratio_Latest_  3585 non-null float64
56 Interest_Cover_Ratio_Latest_        3585 non-null float64
57 PBIDTM_perc_Latest_                 3585 non-null float64
58 PBITM_perc_Latest_                  3585 non-null float64
59 PBDTM_perc_Latest_                  3585 non-null float64
60 CPM_perc_Latest_                   3585 non-null float64
61 APATM_perc_Latest_                 3585 non-null float64
62 Debtors_Velocity_Days              3586 non-null int64
63 Creditors_Velocity_Days            3586 non-null int64
64 Inventory_Velocity_Days            3483 non-null float64
65 Value_of_Output_to_Total_Assets   3586 non-null float64
66 Value_of_Output_to_Gross_Block    3586 non-null float64
dtypes: float64(63), int64(3), object(1)
memory usage: 1.8+ MB

```

Company name is the only object datatype. Rest of the variables are either int or float datatype. Next let us look at the description of the data:

	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Co_Code	3586.0	NaN			16065.388734	19776.817379	4.0	3029.25	6077.5	24269.5	72493.0
Co_Name	3586	3586	AVT Natural Prod	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Networth_Next_Year	3586.0	NaN			725.045251	4769.681004	-8021.6	3.985	19.015	123.8025	111729.1
Equity_Paid_Up	3586.0	NaN			62.966584	778.761744	0.0	3.75	8.29	19.5175	42263.46
Networth	3586.0	NaN			649.746299	4091.988792	-7027.48	3.8925	18.58	117.2975	81657.35
Capital_Employed	3586.0	NaN			2799.611054	26975.135385	-1824.75	7.6025	39.09	226.605	714001.25
Total_Debt	3586.0	NaN			1994.823779	23652.842746	-0.72	0.03	7.49	72.35	652823.81
Gross_Block_	3586.0	NaN			594.178829	4871.547802	-41.19	0.57	15.87	131.895	128477.59
Net_Working_Capital_	3586.0	NaN			410.809665	6301.218546	-13162.42	0.9425	10.145	61.175	223257.56
Current_Assets_	3586.0	NaN			1960.349172	22577.570829	-0.91	4.0	24.54	135.2775	721166.0
Current_Liabilities_and_Provisions_	3586.0	NaN			391.992078	2675.001631	-0.23	0.7325	9.225	65.65	83232.98
Total_Assets_to_Liabilities_	3586.0	NaN			1778.453751	11437.57469	-4.51	10.555	52.01	310.54	254737.22
Gross_Sales	3586.0	NaN			1123.738985	10603.703837	-62.59	1.4425	31.21	242.25	474182.94
Net_Sales	3586.0	NaN			1079.702579	9996.574173	-62.59	1.44	30.44	234.44	443775.16
Other_Income	3586.0	NaN			48.729824	426.040665	-448.72	0.02	0.45	3.635	14143.4

	Value_Of_Output	3586.0	NaN	NaN	NaN	1077.187292	9843.880293	-119.1	1.4125	30.895	235.8375	435559.09			
	Cost_of_Production	3586.0	NaN	NaN	NaN	798.544621	9076.702982	-22.65	0.94	25.99	189.55	419913.5			
	Selling_Cost	3586.0	NaN	NaN	NaN	25.554997	194.244466	0.0	0.0	0.16	3.8825	5283.91			
	PBIDT	3586.0	NaN	NaN	NaN	248.175282	1949.59335	-4655.14	0.04	2.045	23.525	42059.26			
	PBDT	3586.0	NaN	NaN	NaN	116.268795	956.199566	-5874.53	0.0	0.795	12.945	23215.0			
	PBIT	3586.0	NaN	NaN	NaN	217.659395	1850.972782	-4812.95	0.0	1.15	16.6675	41402.96			
	PBT	3586.0	NaN	NaN	NaN	85.752909	799.925768	-6032.34	-0.06	0.31	7.4225	16798.0			
	PAT	3586.0	NaN	NaN	NaN	61.218313	620.298432	-6032.34	-0.06	0.255	5.54	13383.39			
	Adjusted_PAT	3586.0	NaN	NaN	NaN	60.058963	580.432912	-4418.72	-0.09	0.21	5.3425	13384.11			
	CP	3586.0	NaN	NaN	NaN	91.7342	780.790561	-5874.53	0.0	0.74	10.91	20760.2			
	Revenue_earnings_in_forex	3586.0	NaN	NaN	NaN	131.16527	1150.730209	0.0	0.0	0.0	7.2	46158.0			
	Revenue_expenses_in_forex	3586.0	NaN	NaN	NaN	256.327002	4132.339619	0.0	0.0	0.0	6.9875	193979.73			
	Capital_expenses_in_forex	3586.0	NaN	NaN	NaN	7.655689	111.43207	0.0	0.0	0.0	0.0	3722.1			
	Book_Value_Unit_Curr	3586.0	NaN	NaN	NaN	157.237836	1622.664105	-3371.57	7.9625	21.665	71.6675	75790.0			
	Book_Value_Adj__Unit_Curr	3582.0	NaN	NaN	NaN	2243.152917	128283.728186	-33715.7	7.06	18.925	60.01	7677600.29			
	Market_Capitalisation	3586.0	NaN	NaN	NaN	1664.092387	12805.173084	0.0	0.0	8.37	111.4575	260865.08			
	CEPS_annualised_Unit_Curr	3586.0	NaN	NaN	NaN	36.018709	828.420796	-1808.0	0.0	1.145	8.7725	45438.44			
	Cash_Flow_From_Operating_Activities	3586.0	NaN	NaN	NaN	65.77075	1455.048376	-25469.23	-0.3075	0.45	12.6475	44529.4			
	Cash_Flow_From_Investing_Activities	3586.0	NaN	NaN	NaN	-60.870365	701.974713	-23843.45	-5.1175	-0.12	0.12	3732.98			
	Cash_Flow_From_Financing_Activities	3586.0	NaN	NaN	NaN	11.436453	1272.257361	-38374.04	-5.8475	0.0	0.4575	28846.0			
	ROG_Net_Worth_perc	3586.0	NaN	NaN	NaN	1237.624576	41041.930017	-14485.71	-1.4875	1.84	11.3625	2144020.0			
	ROG_Capital_Employed_perc	3586.0	NaN	NaN	NaN	2988.884612	126472.870285	-8614.63	-3.835	1.375	12.5875	7412700.0			
	ROG_Gross_Block_perc	3586.0	NaN	NaN	NaN	37.554306	893.619402	-116.12	0.0	0.25	6.72	47400.0			
	ROG_Gross_Sales_perc	3586.0	NaN	NaN	NaN	242.672962	6103.527897	-5503.7	-8.0775	3.31	21.525	320200.0			
	ROG_Net_Sales_perc	3586.0	NaN	NaN	NaN	242.58853	6103.487655	-5503.7	-8.1175	3.205	21.5675	320200.0			
	ROG_Cost_of_Production_perc	3586.0	NaN	NaN	NaN	310.488405	5573.215095	-2130.23	-7.2425	4.415	23.1225	267150.0			
	ROG_Total_Assets_perc	3586.0	NaN	NaN	NaN	2793.282621	125941.653747	-136.13	-3.9725	1.475	12.5	7422120.0			
	ROG_PBIDT_perc	3586.0	NaN	NaN	NaN	375.852181	23278.396117	-52200.0	-23.3625	4.57	47.875	1386200.0			
	ROG_PBDT_perc	3586.0	NaN	NaN	NaN	336.379947	20353.39666	-52200.0	-30.5975	3.365	52.915	1208700.0			
	ROG_PBIT_perc	3586.0	NaN	NaN	NaN	374.699958	22462.789381	-58500.0	-31.3525	2.13	50.1425	1338000.0			
	ROG_PBT_perc	3586.0	NaN	NaN	NaN	224.070248	19659.232661	-78900.0	-41.235	0.025	61.9575	1160500.0			
	ROG_PAT_perc	3586.0	NaN	NaN	NaN	112.231654	13480.515287	-114500.0	-43.7325	0.0	65.3475	774200.0			
	ROG_CP_perc	3586.0	NaN	NaN	NaN	221.091523	13980.202791	-52200.0	-29.505	4.615	52.9075	822400.0			

ROG_Revenue_earnings_in_forex_perc	3586.0	NaN	NaN	NaN	37.227844	658.666041	-100.0	0.0	0.0	0.0	29084.77
ROG_Revenue_expenses_in_forex_perc	3586.0	NaN	NaN	NaN	364.863221	15233.643027	-100.0	0.0	0.0	0.0	894591.69
ROG_Market_Capitalisation_perc	3586.0	NaN	NaN	NaN	63.68222	1047.928144	-98.05	0.0	0.0	47.515	61865.26
Current_Ratio_Latest_	3585.0	NaN	NaN	NaN	12.056603	108.410131	0.0	0.88	1.36	2.77	4813.0
Fixed_Assets_Ratio_Latest_	3585.0	NaN	NaN	NaN	51.53884	681.15091	0.0	0.27	1.56	4.74	22172.0
Inventory_Ratio_Latest_	3585.0	NaN	NaN	NaN	37.798946	458.189394	0.0	0.0	3.56	8.94	15472.0
Debtors_Ratio_Latest_	3585.0	NaN	NaN	NaN	33.026996	489.563498	0.0	0.42	3.82	8.52	22992.67
Total_Asset_Turnover_Ratio_Latest_	3585.0	NaN	NaN	NaN	1.237236	2.673228	0.0	0.07	0.6	1.55	57.75
Interest_Cover_Ratio_Latest_	3585.0	NaN	NaN	NaN	16.387894	351.73784	-5450.0	0.0	1.08	3.71	18639.4
PBIDTM_perc_Latest_	3585.0	NaN	NaN	NaN	-51.16289	1795.131025	-78870.45	0.0	8.07	18.99	19233.33
PBITM_perc_Latest_	3585.0	NaN	NaN	NaN	-109.213414	3057.63587	-141600.0	0.0	5.23	14.29	19195.7
PBDTM_perc_Latest_	3585.0	NaN	NaN	NaN	-311.570357	10921.592639	-590500.0	0.0	4.69	14.11	15640.0
CPM_perc_Latest_	3585.0	NaN	NaN	NaN	-307.005632	10676.149629	-572000.0	0.0	3.89	11.39	15640.0
APATM_perc_Latest_	3585.0	NaN	NaN	NaN	-365.056187	12500.051387	-688600.0	0.0	1.59	7.41	15266.67
Debtors_Velocity_Days	3586.0	NaN	NaN	NaN	603.894032	10636.75958	0.0	8.0	49.0	106.0	514721.0
Creditors_Velocity_Days	3586.0	NaN	NaN	NaN	2057.854992	54169.479197	0.0	8.0	39.0	89.0	2034145.0
Debtors_Velocity_Days	3586.0	NaN	NaN	NaN	603.894032	10636.75958	0.0	8.0	49.0	106.0	514721.0
Creditors_Velocity_Days	3586.0	NaN	NaN	NaN	2057.854992	54169.479197	0.0	8.0	39.0	89.0	2034145.0
Inventory_Velocity_Days	3483.0	NaN	NaN	NaN	79.644559	137.847792	-199.0	0.0	35.0	96.0	996.0
Value_of_Output_to_Total_Assets	3586.0	NaN	NaN	NaN	0.819757	1.2014	-0.33	0.07	0.48	1.16	17.63
Value_of_Output_to_Gross_Block	3586.0	NaN	NaN	NaN	61.884548	976.824352	-61.0	0.27	1.53	4.91	43404.0

From the description we can see that:

- **Co_Code:** ranges from 4.0 to 72493.0
- **Co_Name:** We have 3586 unique company names
- **Networth_Next_Year:** Value of a company as on 2016 - Next Year(difference between the value of total assets and total liabilities) .this variable has minimum value of -8021.6 and maximum value of 111729.1 with average of 725.045251
- **Equity_Paid_Up:** Amount that has been received by the company through the issue of shares to the shareholders has minimum value of 0.0 and maximum of 42263.46 and mean of 62.966584
- **Networth:** Value of a company as on 2015 - Current Year ranges from -7027.48 to 81657.35 with average of 649.746299
- **Capital_Employed:** Total amount of capital used for the acquisition of profits by a company which ranges from -1824.75 to 714001.25 with average being 2799.611054
- **Total_Debt:** The sum of money borrowed by the company and is due to be paid ranges from -0.72 to 652823.81 with average of 1994.823779
- **Gross_Block_:** Total value of all of the assets that a company owns ranges from -41.19 to 128477.59 with average of 594.178829
- **Net_Working_Capital_:** The difference between a company's current assets (cash, accounts receivable, inventories of raw materials and finished goods) and its current liabilities (accounts payable). Ranges from -13162.42 to 223257.56 with average being 410.809665
- **Current_Assets_:** All the assets of a company that are expected to be sold or used as a result of standard business operations over the next year. This value ranges from -0.91 to 721166.0 and average is 1960.349172
- **Current_Liabilities_and_Provisions_:** Short-term financial obligations that are due within one year (includes amount that is set aside cover a future liability) starts from -0.23 to 83232.98 with average of 391.992078
- **Total_Assets_to_Liabilities_:** Ratio of total assets to liabilities of the company ranges from -4.51 to 254737.22 with average being 1778.453751

- **Gross_Sales:** Total value of all of the assets that a company owns starts from -62.59 to 474182.94 and average is 1123.738985
- **Net_Sales:** Gross sales minus returns, allowances, and discounts ranges from -62.59 to 443775.16 and average is 1079.702579
- **Other_Income:** Income realized from non-business activities (e.g. sale of long term asset) starts from -448.72 to 14143.4 and average being 48.729824
- **Value_Of_Output:** Product of physical output of goods and services produced by company and its market price ranges from -119.1 to 435559.09 and 50% values lies below 30.895
- **Cost_of_Production:** Costs incurred by a business from manufacturing a product or providing a service starts from -22.65 to 419913.5 and average is 798.544621
- **Selling_Cost:** Costs which are made to create the demand for the product (advertising expenditures, packaging and styling, salaries, commissions and travelling expenses of sales personnel, and the cost of shops and showrooms) ranges from 194.244466 to 5283.91 with average being 25.554997
- **PBIDT:** Profit Before Interest, Depreciation & Taxes starts from -4655.14 to 42059.26 and 50% companies lie below 2.045
- **PBDT:** Profit Before Depreciation and Tax ranges from -5874.53 to 23215.0 and 50% companies lie below 0.795
- **PBIT:** Profit before interest and taxes ranges from -4812.95 to 41402.96 and 50% companies lie below 1.15
- **PBT:** Profit before tax ranges from -6032.34 to 16798.0 and 50% companies lie below 0.31
- **PAT:** Profit After Tax ranges from -6032.34 to 13383.39 and 50% companies lie below 0.255 and average is 61.218313
- **Adjusted_PAT:** Adjusted profit is the best estimate of the true profit ranges from -4418.72 to 13384.11 and average is 60.058963 where 50% values lie below 0.21
- **CP:** Commercial paper , a short-term debt instrument to meet short-term liabilities. Ranges from -5874.53 to 20760.2 and average is 91.7342
- **Revenue_earnings_in_forex:** Revenue earned in foreign currency ranges from 0.0 to 46158.0 and 50% values are below 0.0
- **Revenue_expenses_in_forex:** Expenses due to foreign currency transactions . It ranges from 0.0 to 193979.73 where average is 256.327002 and 50% companies lie below 0.0
- **Capital_expenses_in_forex:** Long term investment in forex. It ranges from 0.0 to 3722.1 where average is 7.655689 and 50% companies lie below 0.0
- **Book_Value_Unit_Curr:** Net asset value . It ranges from -3371.57 to 75790.0 where average is 157.237836 and 50% companies lie below 21.665
- **Book_Value_Adj__Unit_Curr:** Book value adjusted to reflect asset's true fair market value. It ranges from -33715.7 to 7677600.29 where average is 2243.152917 and 50% companies lie below 18.925
- **Market_Capitalisation:** Product of the total number of a company's outstanding shares and the current market price of one share. It ranges from 0.0 to 260865.08 where average is 1664.092387 and 50% companies lie below 8.37
- **CEPS_annualised_Unit_Curr:** Cash Earnings per Share, profitability ratio that measures the financial performance of a company by calculating cash flows on a per share basis. It ranges from -1808.0 to 45438.44 where average is 36.018709 and 50% companies lie below 1.145
- **Cash_Flow_From_Operating_Activities:** Use of cash from ongoing regular business activities. It ranges from -25469.23 to 44529.4 where average is 65.77075 and 50% companies lie below 0.45
- **Cash_Flow_From_Investing_Activities:** Cash used in the purchase of non-current assets—or long-term assets—that will deliver value in the future. It ranges from -23843.45 to 3732.98 where average is -60.870365 and 50% companies lie below -0.12
- **Cash_Flow_From_Financing_Activities:** Net flows of cash that are used to fund the company (transactions involving debt, equity, and dividends) It ranges from -38374.04 to 28846.0 where average is 11.436453 and 50% companies lie below 0.00

- **ROG_Net_Worth_perc:** Rate of Growth – Networth It ranges from -14485.71 to 2144020.0 where average is 1237.624576 and 50% companies lie below 1.84
- **ROG_Capital_Employed_perc:** Rate of Growth - Capital Employed It ranges from -8614.63 to 7412700.0 where average is 2988.884612 and 50% companies lie below 1.375
- **ROG_Gross_Block_perc:** Rate of Growth - Gross Block It ranges from -116.12 to 47400.0 where average is 37.554306 and 50% companies lie below 0.25
- **ROG_Gross_Sales_perc:** Rate of Growth - Gross Sales It ranges from -5503.7 to 320200.0 where average is 242.672962 and 50% companies lie below 3.31
- **ROG_Net_Sales_perc:** Rate of Growth - Net Sales It ranges from -5503.7 to 320200.0 where average is 242.58853 and 50% companies lie below 3.205
- **ROG_Cost_of_Production_perc:** Rate of Growth - Cost of Production. It ranges from -2130.23 to 267150.0 where average is 310.488405 and 50% companies lie below 4.415
- **ROG_Total_Assets_perc:** Rate of Growth - Total Assets. It ranges from -136.13 to 7422120.0 where average is 2793.282621 and 50% companies lie below 1.475
- **ROG_PBIDT_perc:** Rate of Growth- PBIDT . It ranges from -52200.0 to 1386200.0 where average is 375.852181 and 50% companies lie below 4.57
- **ROG_PBDT_perc:** Rate of Growth- PBDT. It ranges from -52200.0 to 1208700.0 where average is 336.379947 and 50% companies lie below 3.365
- **ROG_PBIT_perc:** Rate of Growth- PBIT. It ranges from -58500.0 to 1338000.0 where average is 374.699958 and 50% companies lie below 2.13
- **ROG_PBT_perc:** Rate of Growth- PBT. It ranges from -78900.0 to 1160500.0 where average is 224.070248 and 50% companies lie below 0.025
- **ROG_PAT_perc:** Rate of Growth- PAT. It ranges from -114500.0 to 774200.0 where average is 112.231654 and 50% companies lie below 0.0
- **ROG_CP_perc:** Rate of Growth- CP. It ranges from -52200.0 to 822400.0 where average is 221.091523 and 50% companies lie below 4.615
- **ROG_Revenue_earnings_in_forex_perc:** Rate of Growth - Revenue earnings in forex. It ranges from -100.0 to 29084.77 where average is 37.227844 and 50% companies lie below 0.0
- **ROG_Revenue_expenses_in_forex_perc:** Rate of Growth - Revenue expenses in forex. It ranges from -100.0 to 894591.69 where average is 364.863221 and 50% companies lie below 0.0
- **ROG_Market_Capitalisation_perc:** Rate of Growth - Market Capitalisation. It ranges from -98.05 to 61865.26 where average is 63.68222 and 50% companies lie below 0.0
- **Current_Ratio_Latest_:** Liquidity ratio, company's ability to pay short-term obligations or those due within one year. It ranges from 0.0 to 4813.0 where average is 12.056603 and 50% companies lie below 1.36
- **Fixed_Assets_Ratio_Latest_:** Solvency ratio, the capacity of a company to discharge its obligations towards long-term lenders indicating. It ranges from 0.0 to 22172.0 where average is 51.53884 and 50% companies lie below 1.56
- **Inventory_Ratio_Latest_:** Activity ratio, specifies the number of times the stock or inventory has been replaced and sold by the company. It ranges from 0.0 to 15472.0 where average is 37.798946 and 50% companies lie below 3.56
- **Debtors_Ratio_Latest_:** Measures how quickly cash debtors are paying back to the company. It ranges from 0.0 to 22992.67 where average is 33.026996 and 50% companies lie below 3.82
- **Total_Asset_Turnover_Ratio_Latest_:** The value of a company's revenues relative to the value of its assets. It ranges from 0.0 to 57.75 where average is 1.237236 and 50% companies lie below 0.6
- **Interest_Cover_Ratio_Latest_:** Determines how easily a company can pay interest on its outstanding debt. It ranges from -5450.0 to 18639.4 where average is 16.387894 and 50% companies lie below 1.08
- **PBIDTM_perc_Latest_:** Profit before Interest Depreciation and Tax Margin. It ranges from -78870.45 to 19233.33 where average is -51.16289 and 50% companies lie below 8.07
- **PBITM_perc_Latest_:** Profit Before Interest Tax Margin. It ranges from -109.213414 to 19195.7 where average is -109.213414 and 50% companies lie below 5.23

- **PBDM_perc_Latest**: Profit Before Depreciation Tax Margin. It ranges from -590500.0 to 15640.0 where average is -311.570357 and 50% companies lie below 4.69
- **CPM_perc_Latest**: Cost per thousand (advertising cost). It ranges from -572000.0 to 15640.0 where average is -307.005632 and 50% companies lie below 3.89
- **APATM_perc_Latest**: After tax profit margin. It ranges from -688600.0 to 15266.67 where average is -365.056187 and 50% companies lie below 1.59
- **Debtors_Velocity_Days**: Average days required for receiving the payments. It ranges from 0.0 to 514721.0 where average is 603.894032 and 50% companies lie below 49.0
- **Creditors_Velocity_Days**: Average number of days company takes to pay suppliers. It ranges from 0.0 to 2034145.0 where average is 2057.854992 and 50% companies lie below 39.0
- **Inventory_Velocity_Days**: Average number of days the company needs to turn its inventory into sales. It ranges from -199.0 to 996.0 where average is 79.644559 and 50% companies lie below 35.0
- **Value_of_Output_to_Total_Assets**: Ratio of Value of Output (market value) to Total Assets. It ranges from -0.33 to 17.63 where average is 0.819757 and 50% companies lie below 0.48
- **Value_of_Output_to_Gross_Block**: Ratio of Value of Output (market value) to Gross Block. It ranges from -61.0 to 43404.0 where average is 61.884548 and 50% companies lie below 1.53

Dependent variable - We need to create a default variable that should take the value of 1 when net worth next year is negative & 0 when net worth next year is positive.

We have got 3198 non defaulters and 388 defaulters. In other words, 11% are defaulters.

Missing Value Treatment:

Now let us look at the missing values counts:

Co_Code	0
Co_Name	0
Nerworth_Next_Year	0
Equity_Paid_Up	0
Nerworth	0
Capital_Employed	0
Total_Debt	0
Gross_Block_	0
Net_Working_Capital_	0
Current_Assets_	0
Current_Liabilities_and_Provisions_	0
Total_Assets_to_Liabilities_	0
Gross_Sales	0
Net_Sales	0
Other_Income	0
Value_Of_Output	0
Cost_of_Production	0
Selling_Cost	0
PBIDT	0
PBDT	0
PBIT	0
PBT	0
PAT	0
Adjusted_PAT	0
CP	0
Revenue_earnings_in_forex	0

Revenue_earnings_in_forex	0
Revenue_expenses_in_forex	0
Capital_expenses_in_forex	0
Book_Value_Unit_Curr	0
Book_Value_Adj__Unit_Curr	4
Market_Capitalisation	0
CEPS_annualised_Unit_Curr	0
Cash_Flow_From_Operating_Activities	0
Cash_Flow_From_Investing_Activities	0
Cash_Flow_From_Financing_Activities	0
ROG_Net_Worth_perc	0
ROG_Capital_Employed_perc	0
ROG_Gross_Block_perc	0
ROG_Gross_Sales_perc	0
ROG_Net_Sales_perc	0
ROG_Cost_of_Production_perc	0
ROG_Total_Assets_perc	0
ROG_PBIDT_perc	0
ROG_PBDT_perc	0
ROG_PBIT_perc	0
ROG_PBT_perc	0
ROG_PAT_perc	0
ROG_CP_perc	0
ROG_Revenue_earnings_in_forex_perc	0
ROG_Revenue_expenses_in_forex_perc	0
ROG_Market_Capitalisation_perc	0
ROG_Revenue_earnings_in_forex_perc	0
ROG_Revenue_expenses_in_forex_perc	0
ROG_Market_Capitalisation_perc	0
Current_Ratio_Latest_	1
Fixed_Assets_Ratio_Latest_	1
Inventory_Ratio_Latest_	1
Debtors_Ratio_Latest_	1
Total_Asset_Turnover_Ratio_Latest_	1
Interest_Cover_Ratio_Latest_	1
PBIDTM_perc_Latest_	1
PBITM_perc_Latest_	1
PBDTM_perc_Latest_	1
CPM_perc_Latest_	1
APATM_perc_Latest_	1
Debtors_Velocity_Days	0
Creditors_Velocity_Days	0
Inventory_Velocity_Days	103
Value_of_Output_to_Total_Assets	0
Value_of_Output_to_Gross_Block	0
default	0

In total there are 118 missing values. We have used KNN imputer to treat missing values which uses the 10 neighboring values to fill up the missing values.

Univariate analysis:

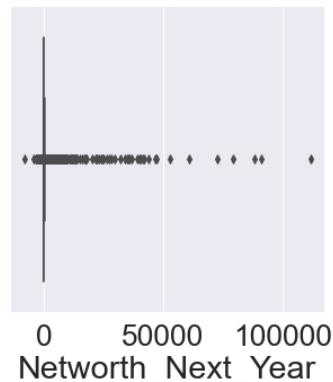
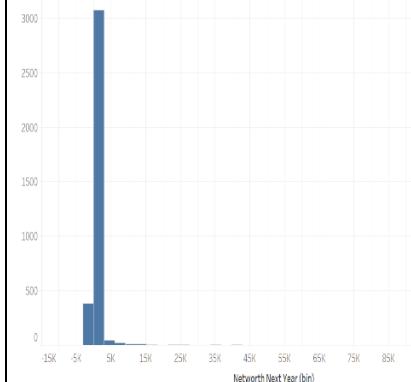
Let us now start the uni-variate analysis by looking at description, histogram and boxplot of each variable:

Description of Networth_Next_Year

```

count      3586.00
mean       725.05
std        4769.68
min      -8021.60
25%        3.98
50%       19.02
75%      123.80
max     111729.10
Name: Networth_Next_Year, dtype: float64

```

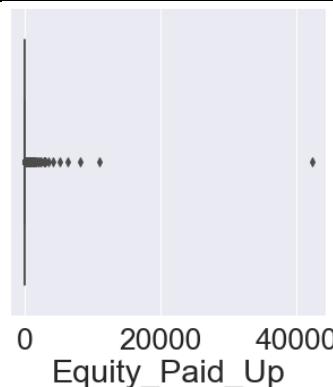
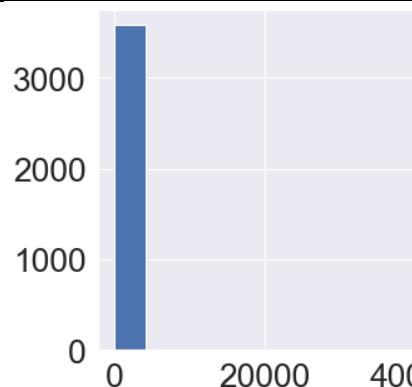


Description of Equity_Paid_Up

```

count      3586.00
mean       62.97
std        778.76
min       0.00
25%        3.75
50%       8.29
75%      19.52
max     42263.46
Name: Equity_Paid_Up, dtype: float64

```

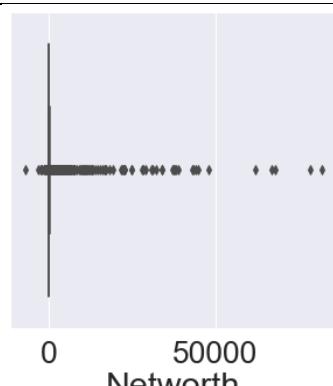
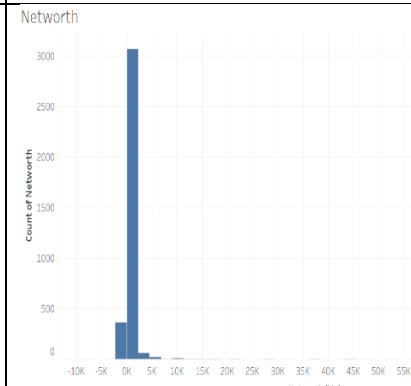


Description of Networth

```

count      3586.00
mean       649.75
std        4091.99
min      -7027.48
25%        3.89
50%       18.58
75%      117.30
max     81657.35
Name: Networth, dtype: float64

```

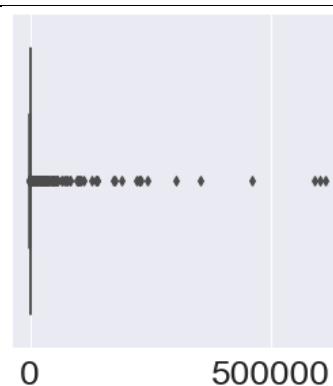
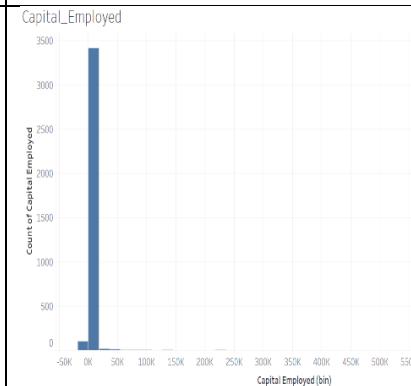


Description of Capital_Employed

```

count      3586.00
mean      2799.61
std        26975.14
min      -1824.75
25%        7.60
50%       39.09
75%      226.60
max    714001.25
Name: Capital_Employed, dtype: float64

```

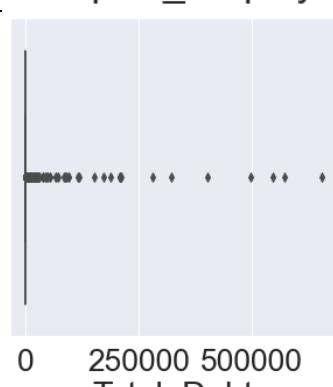
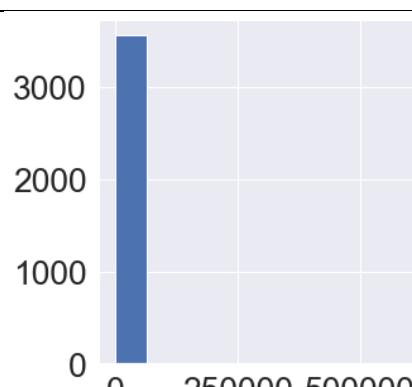


Description of Total_Debt

```

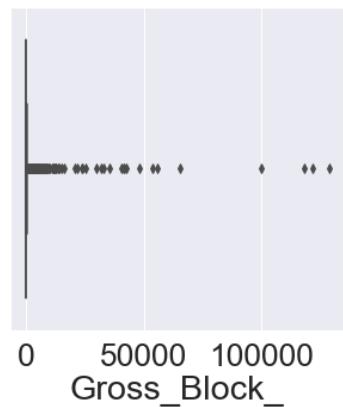
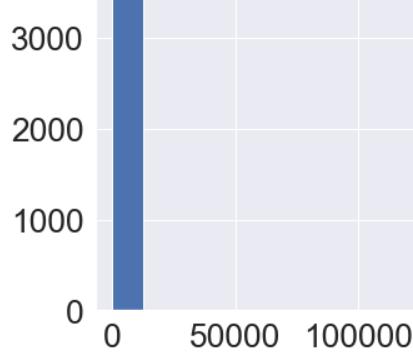
count      3586.00
mean      1994.82
std        23652.84
min       -0.72
25%        0.03
50%       7.49
75%      72.35
max    652823.81
Name: Total_Debt, dtype: float64

```



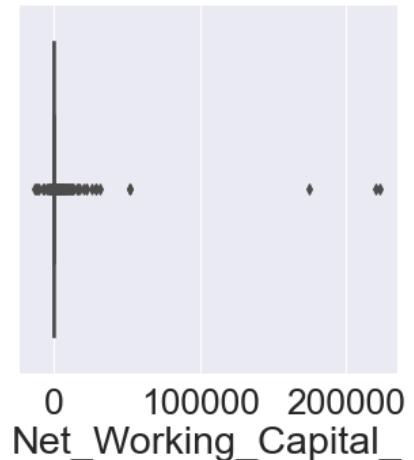
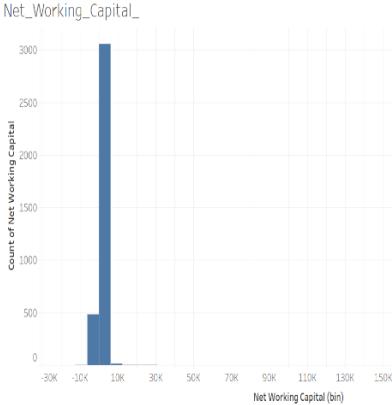
Description of Gross_Block_

```
count      3586.00
mean       594.18
std        4871.55
min       -41.19
25%        0.57
50%       15.87
75%      131.90
max     128477.59
Name: Gross_Block_, dtype: float64
```



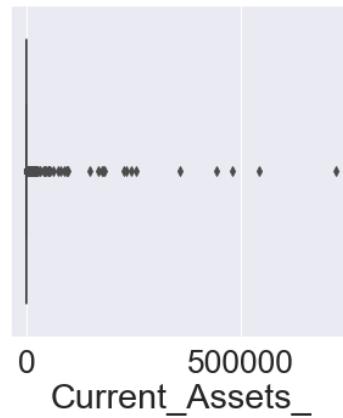
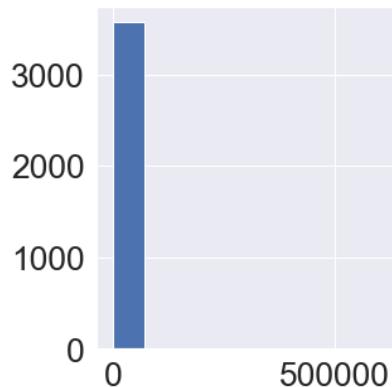
Description of Net_Working_Capital_

```
count      3586.00
mean       410.81
std        6301.22
min      -13162.42
25%        0.94
50%       10.14
75%       61.17
max     223257.56
Name: Net_Working_Capital_, dtype: float64
```



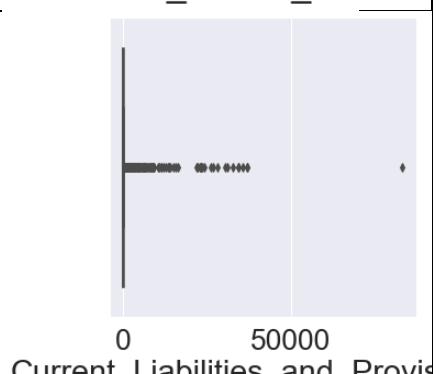
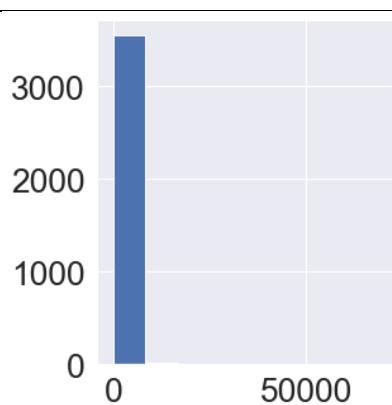
Description of Current_Assets_

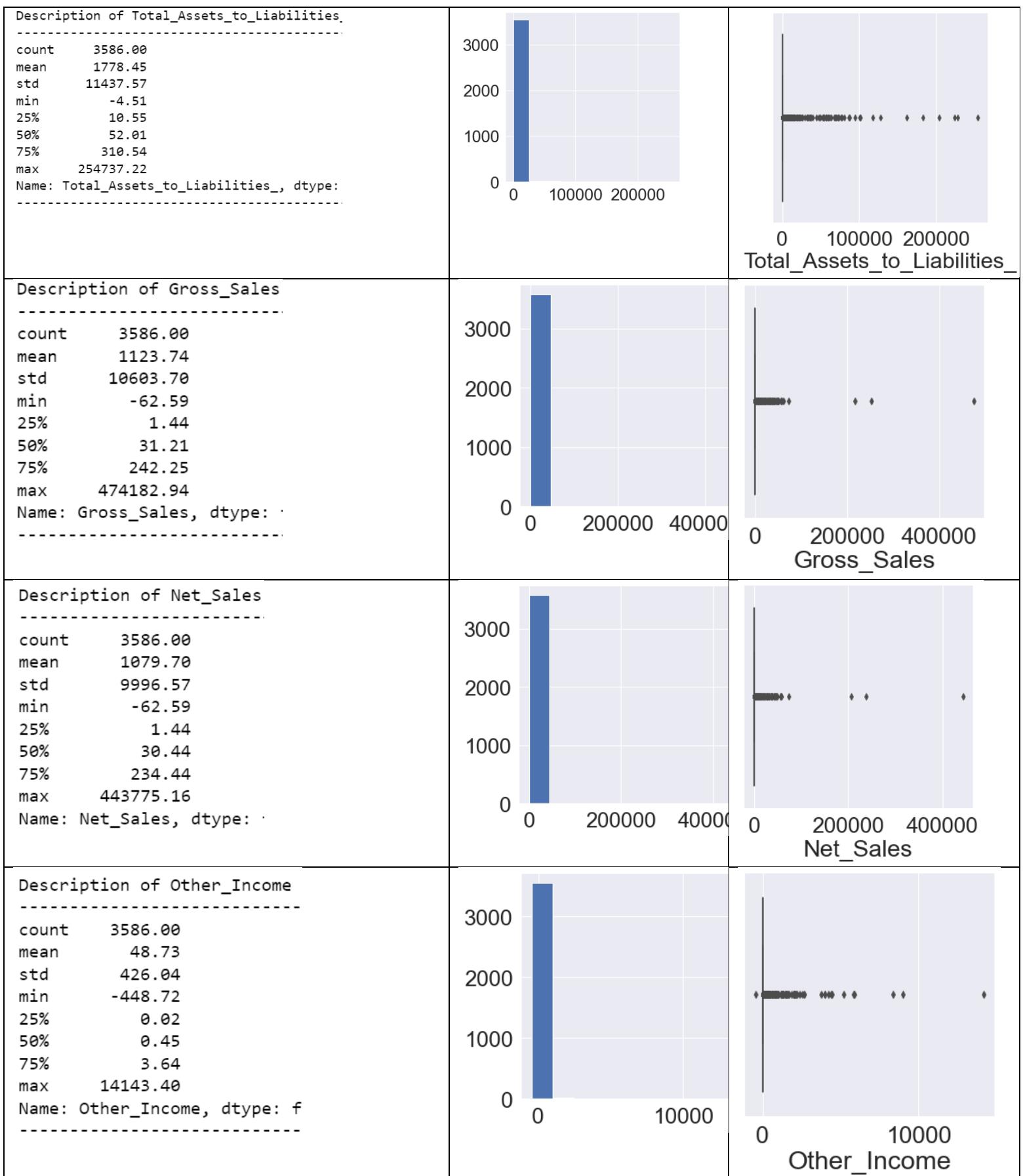
```
count      3586.00
mean      1960.35
std       22577.57
min       -0.91
25%        4.00
50%       24.54
75%      135.28
max     721166.00
Name: Current_Assets_, dtype: float64
```



Description of Current_Liabilities_and_Provisions_

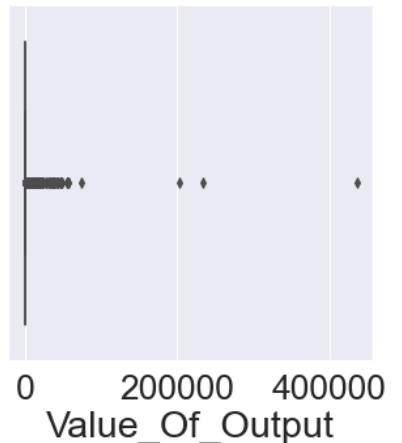
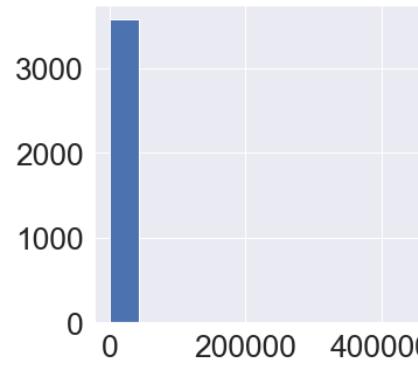
```
count      3586.00
mean      391.99
std       2675.00
min       -0.23
25%        0.73
50%       9.23
75%      65.65
max     83232.98
Name: Current_Liabilities_and_Provisions_, dtype: float64
```





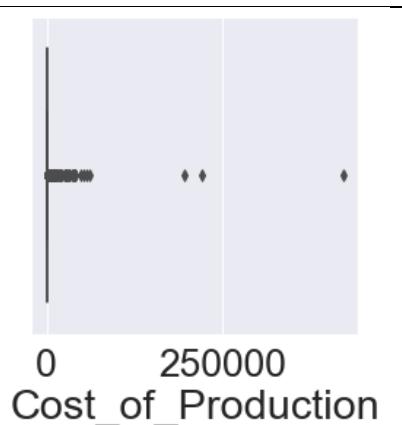
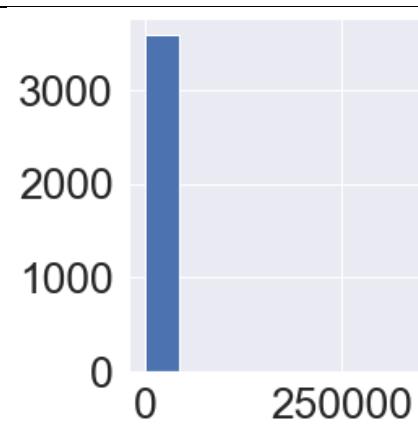
Description of Value_Of_Output

```
count      3586.00
mean       1077.19
std        9843.88
min       -119.10
25%        1.41
50%       30.89
75%      235.84
max     435559.09
Name: Value_Of_Output, dtype: float64
```



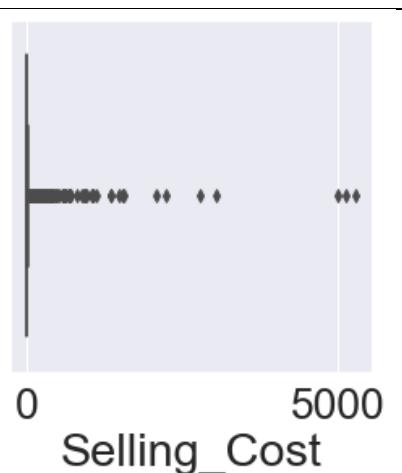
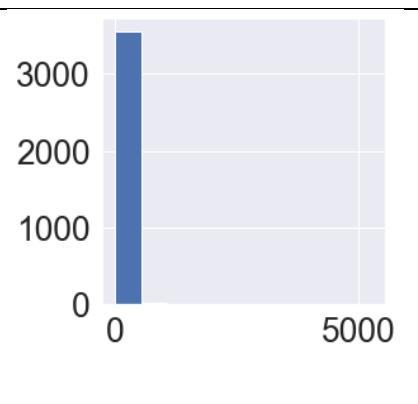
Description of Cost_of_Production

```
count      3586.00
mean       798.54
std        9076.70
min       -22.65
25%        0.94
50%       25.99
75%      189.55
max     419913.50
Name: Cost_of_Production, dtype: float64
```



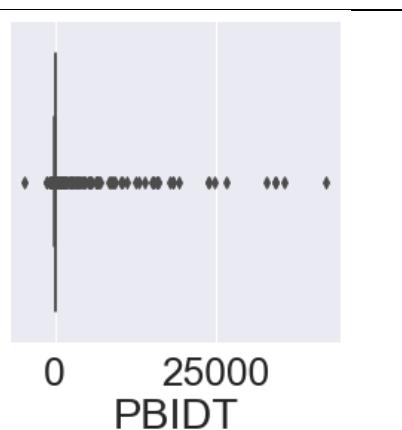
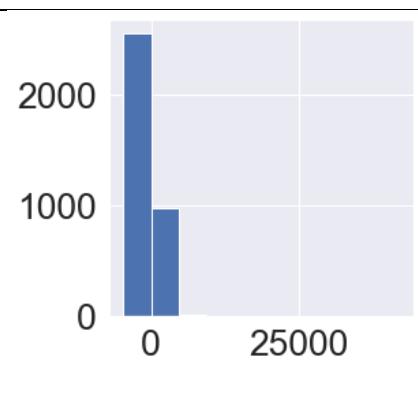
Description of Selling_Cost

```
count      3586.00
mean       25.55
std        194.24
min       0.00
25%        0.00
50%        0.16
75%       3.88
max     5283.91
Name: Selling_Cost, dtype: float64
```



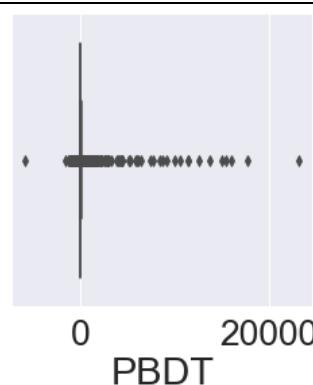
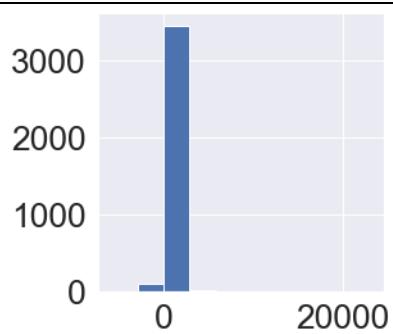
Description of PBIDT

```
count      3586.00
mean       248.18
std        1949.59
min      -4655.14
25%        0.04
50%        2.04
75%       23.52
max     42059.26
Name: PBIDT, dtype: float64
```



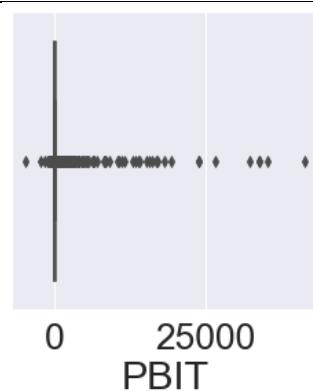
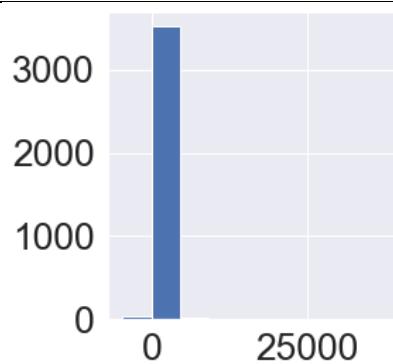
Description of PBDT

```
count    3586.00
mean     116.27
std      956.20
min     -5874.53
25%      0.00
50%      0.80
75%     12.95
max     23215.00
Name: PBDT, dtype: float64
```



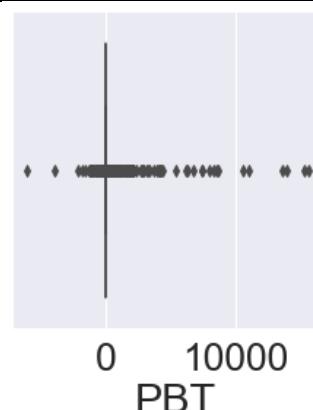
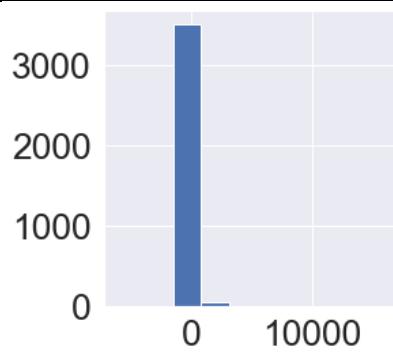
Description of PBIT

```
count    3586.00
mean     217.66
std      1850.97
min     -4812.95
25%      0.00
50%      1.15
75%     16.67
max     41402.96
Name: PBIT, dtype: float64
```



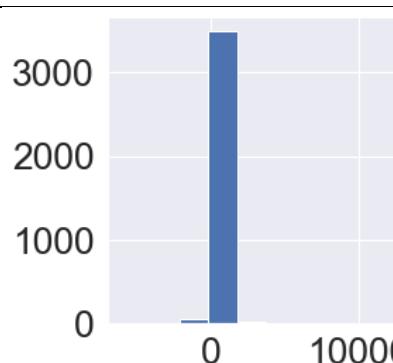
Description of PBT

```
count    3586.00
mean     85.75
std      799.93
min     -6032.34
25%     -0.06
50%      0.31
75%      7.42
max     16798.00
Name: PBT, dtype: float64
```



Description of PAT

```
count    3586.00
mean     61.22
std      620.30
min     -6032.34
25%     -0.06
50%      0.26
75%      5.54
max     13383.39
Name: PAT, dtype: float64
```

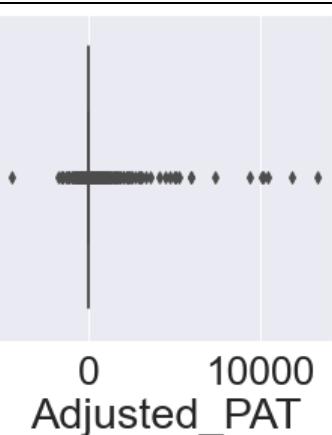
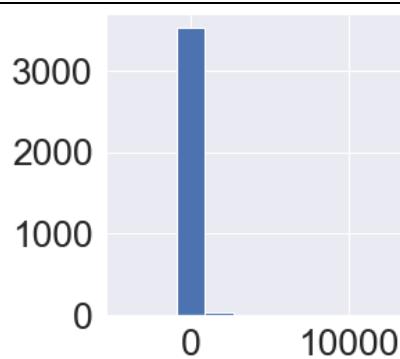


Description of Adjusted_PAT

```

count    3586.00
mean     60.06
std      580.43
min     -4418.72
25%     -0.09
50%      0.21
75%      5.34
max    13384.11
Name: Adjusted_PAT, dtype: float64

```

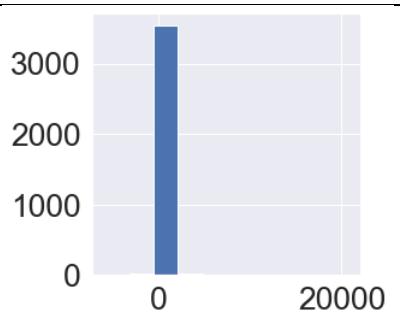


Description of CP

```

count    3586.00
mean     91.73
std      780.79
min     -5874.53
25%      0.00
50%      0.74
75%     10.91
max    20760.20
Name: CP, dtype: float64

```

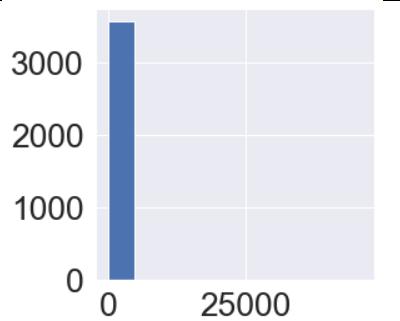


Description of Revenue_earnings_in_forex

```

count    3586.00
mean    131.17
std     1150.73
min      0.00
25%      0.00
50%      0.00
75%     7.20
max   46158.00
Name: Revenue_earnings_in_forex, dtype: float64

```

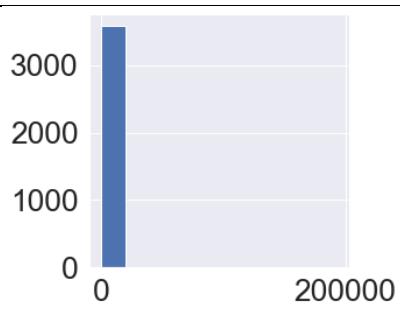


Description of Revenue_expenses_in_forex

```

count    3586.00
mean    256.33
std     4132.34
min      0.00
25%      0.00
50%      0.00
75%     6.99
max  193979.73
Name: Revenue_expenses_in_forex, dtype: float64

```

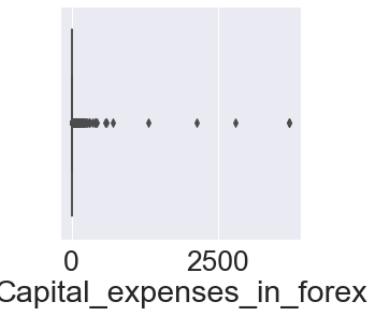
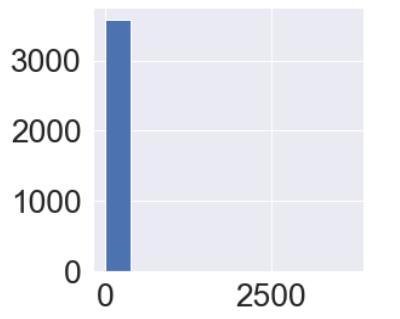


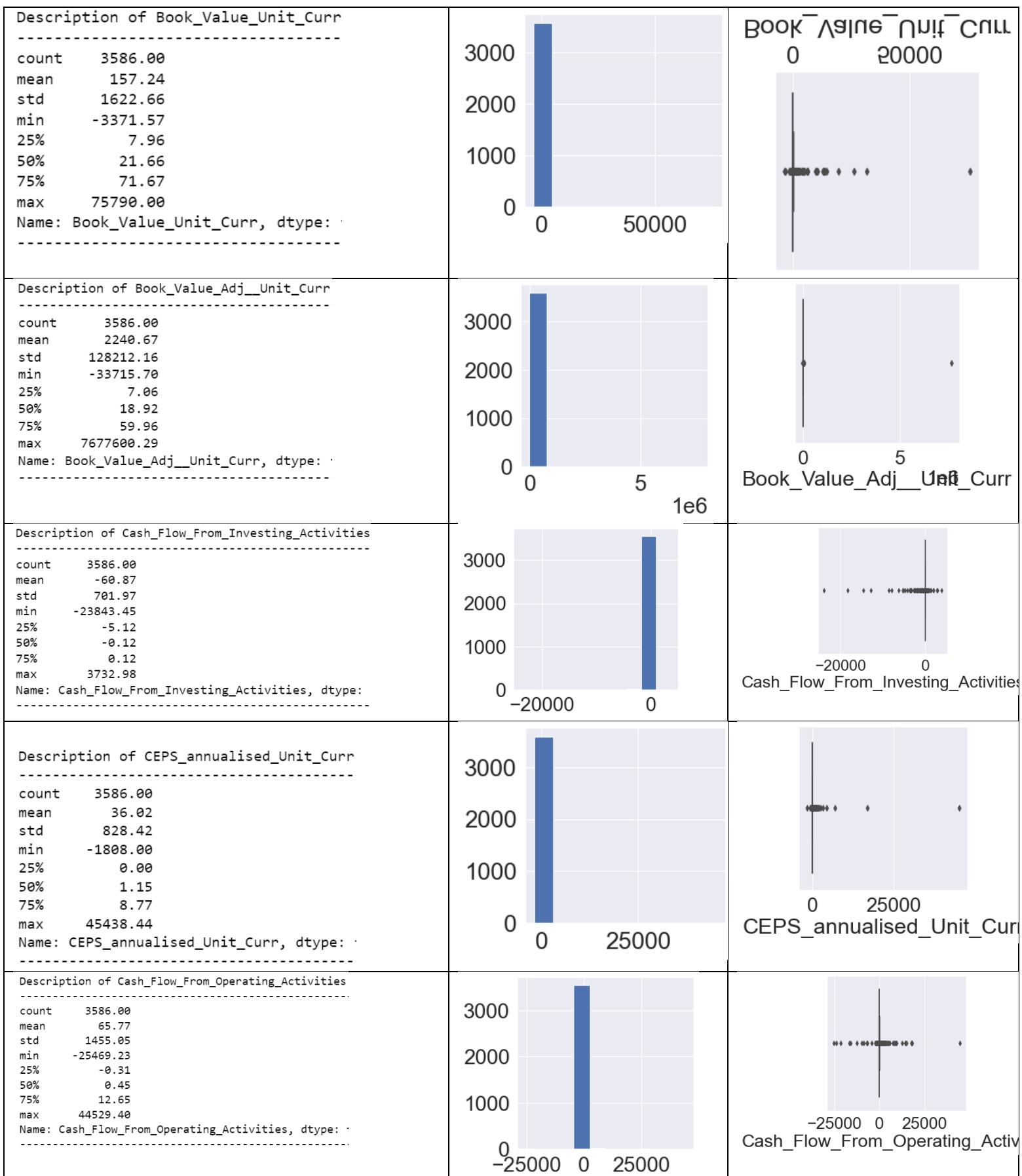
Description of Capital_expenses_in_forex

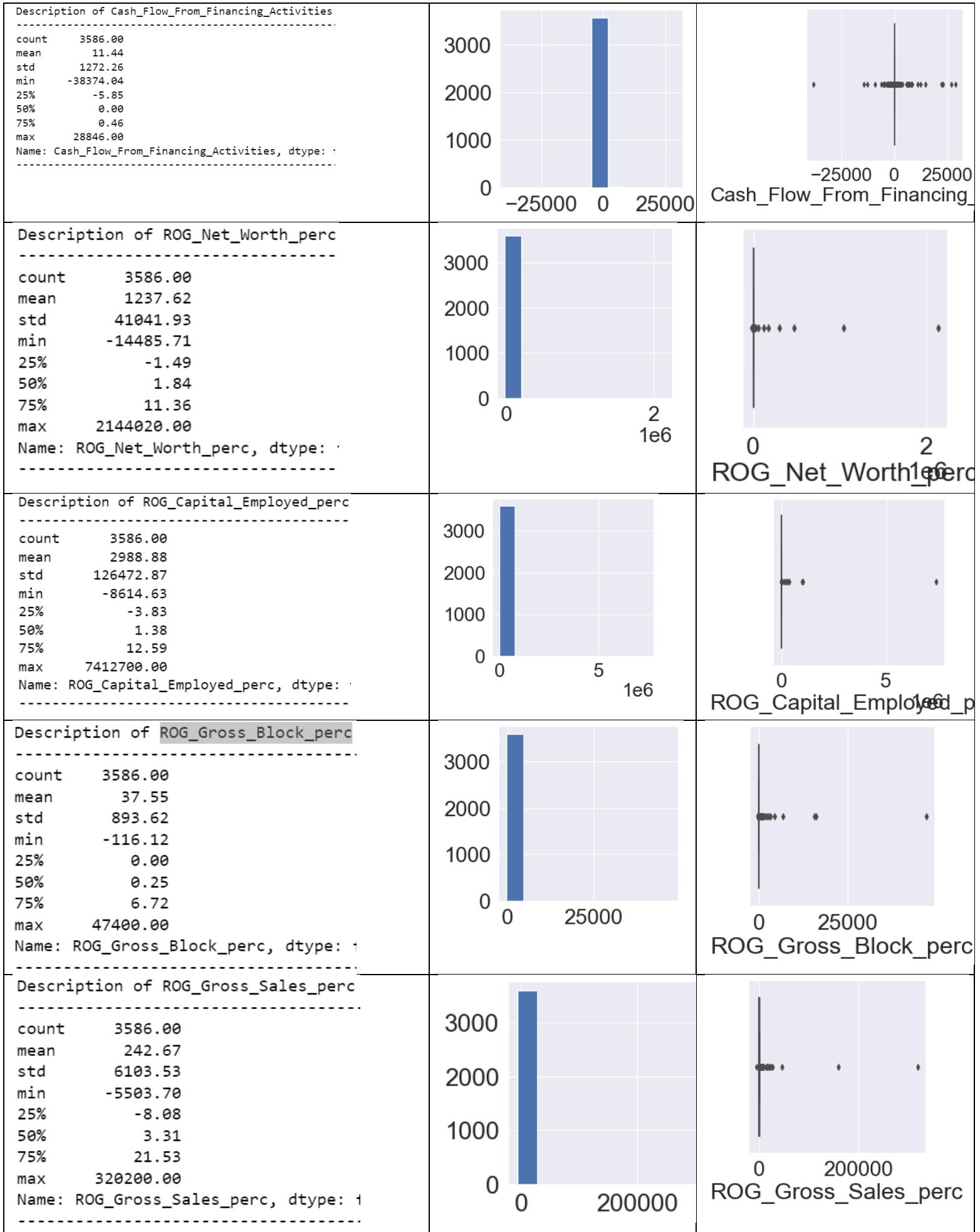
```

count    3586.00
mean     7.66
std     111.43
min      0.00
25%      0.00
50%      0.00
75%      0.00
max    3722.10
Name: Capital_expenses_in_forex, dtype: float64

```

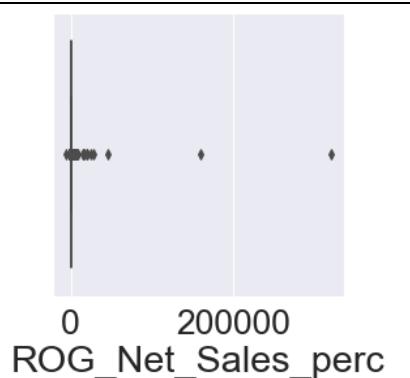
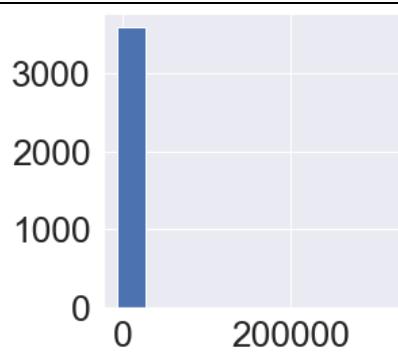






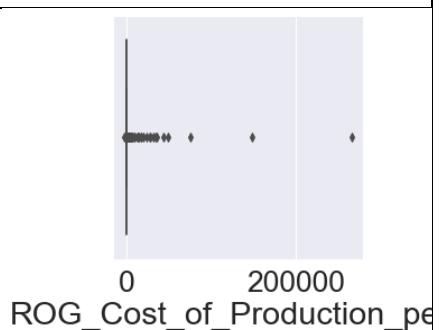
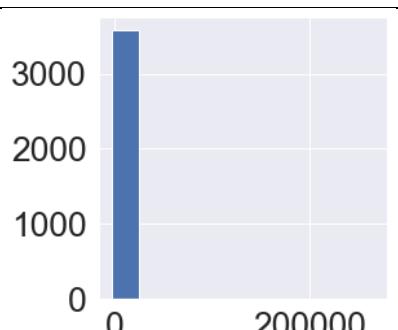
Description of ROG_Net_Sales_perc

```
count      3586.00
mean       242.59
std        6103.49
min      -5503.70
25%       -8.12
50%        3.21
75%       21.57
max     320200.00
Name: ROG_Net_Sales_perc, dtype: float64
```



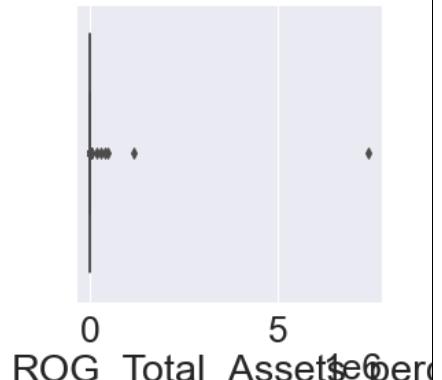
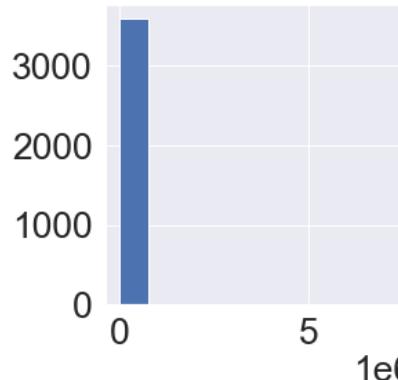
Description of ROG_Cost_of_Production_perc

```
count      3586.00
mean       310.49
std        5573.22
min      -2130.23
25%       -7.24
50%        4.42
75%       23.12
max     267150.00
Name: ROG_Cost_of_Production_perc, dtype: float64
```



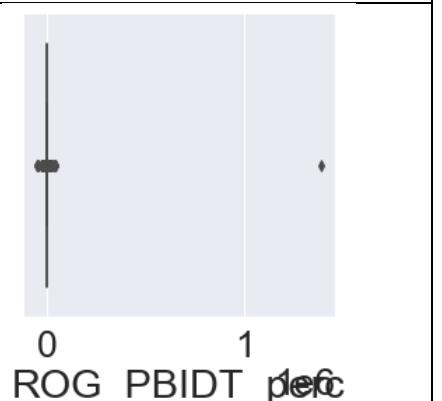
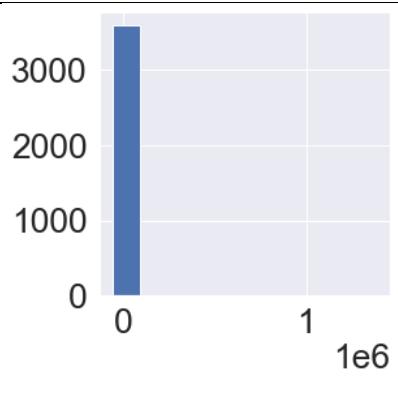
Description of ROG_Total_Assets_perc

```
count      3586.00
mean       2793.28
std       125941.65
min      -136.13
25%       -3.97
50%        1.48
75%       12.50
max     7422120.00
Name: ROG_Total_Assets_perc, dtype: float64
```



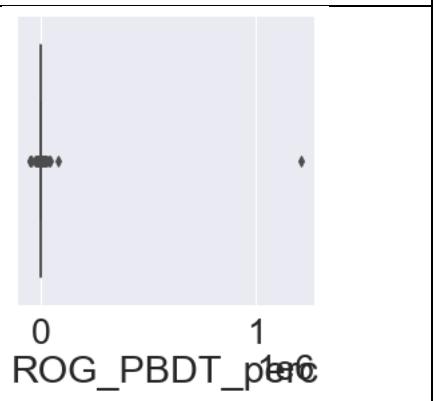
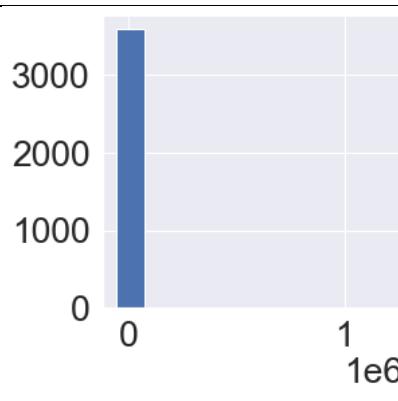
Description of ROG_PBIDT_perc

```
count      3586.00
mean       375.85
std        23278.40
min      -52200.00
25%       -23.36
50%        4.57
75%       47.88
max     1386200.00
Name: ROG_PBIDT_perc, dtype: float64
```

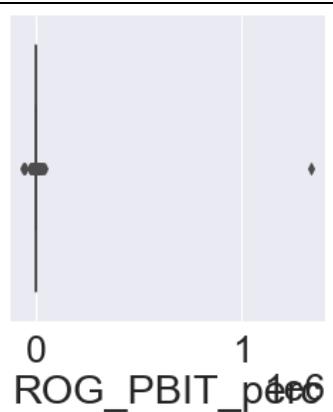
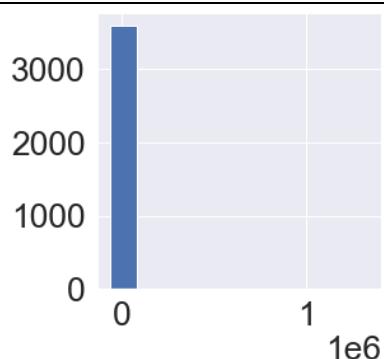


Description of ROG_PBDT_perc

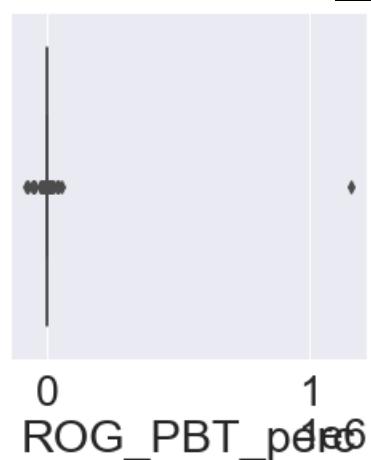
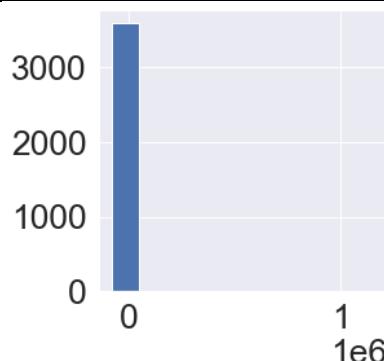
```
count      3586.00
mean       336.38
std        20353.40
min      -52200.00
25%       -30.60
50%        3.37
75%       52.91
max     1208700.00
Name: ROG_PBDT_perc, dtype: float64
```



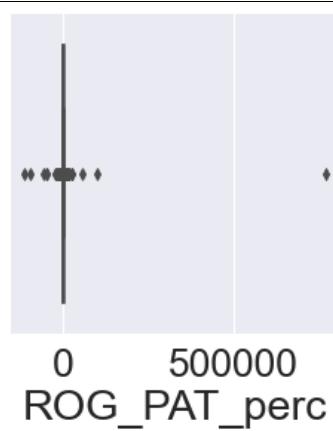
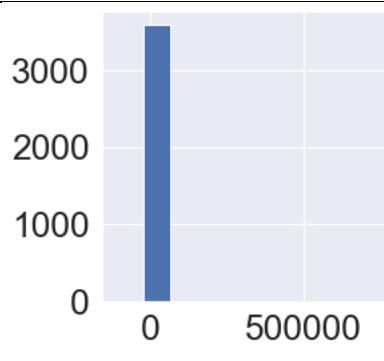
```
Description of ROG_PBIT_perc
-----
count      3586.00
mean       374.70
std        22462.79
min     -58500.00
25%      -31.35
50%       2.13
75%      50.14
max    1338000.00
Name: ROG_PBIT_perc, dtype: float64
```



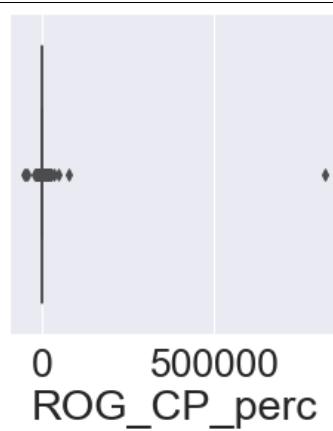
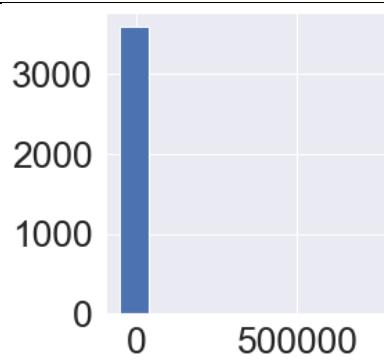
```
Description of ROG_PBT_perc
-----
count      3586.00
mean       224.07
std        19659.23
min     -78900.00
25%      -41.23
50%       0.03
75%      61.96
max    1160500.00
Name: ROG_PBT_perc, dtype: float64
```

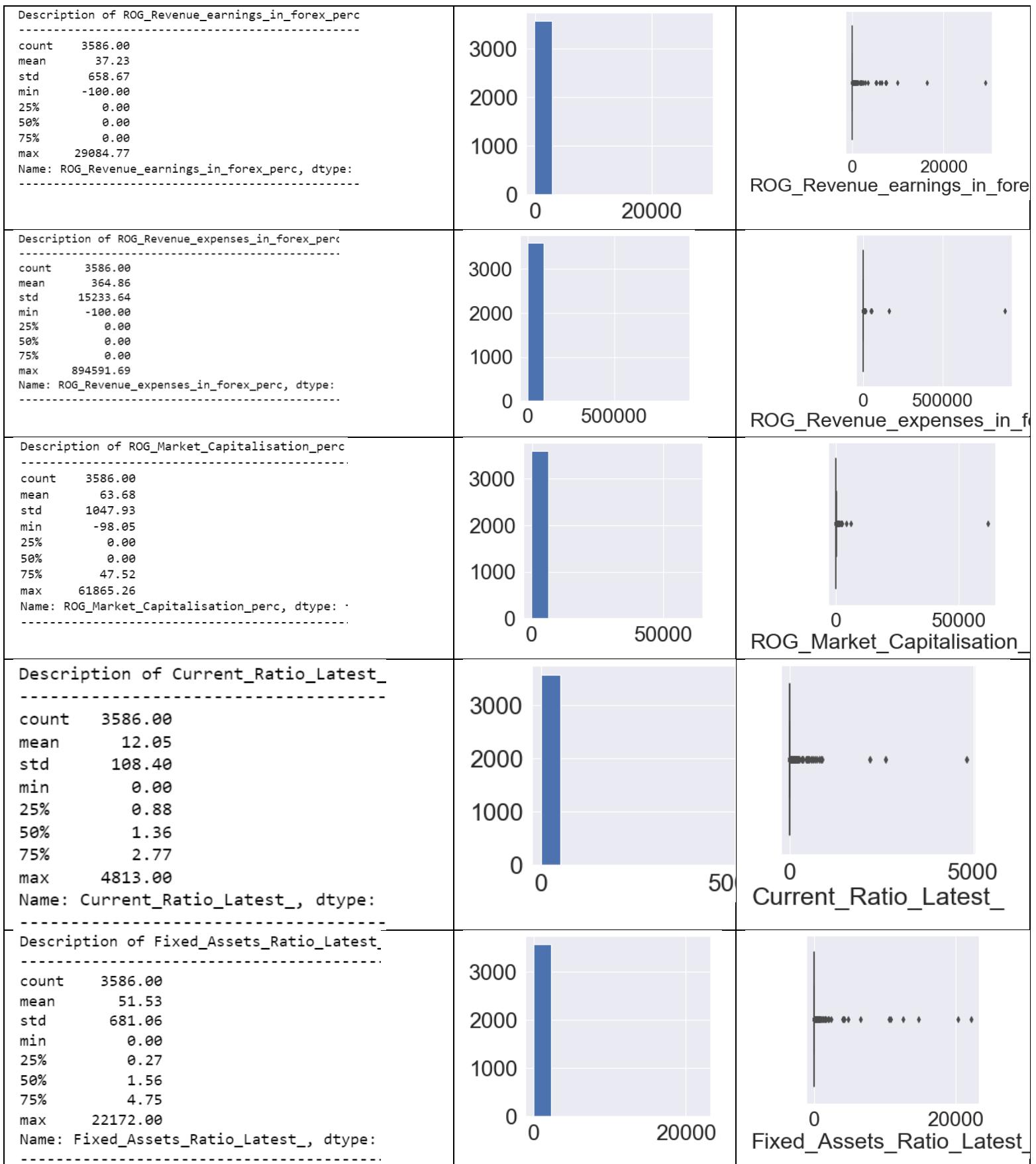


```
Description of ROG_PAT_perc
-----
count      3586.00
mean       112.23
std        13480.52
min     -114500.00
25%      -43.73
50%       0.00
75%      65.35
max    774200.00
Name: ROG_PAT_perc, dtype: float64
```



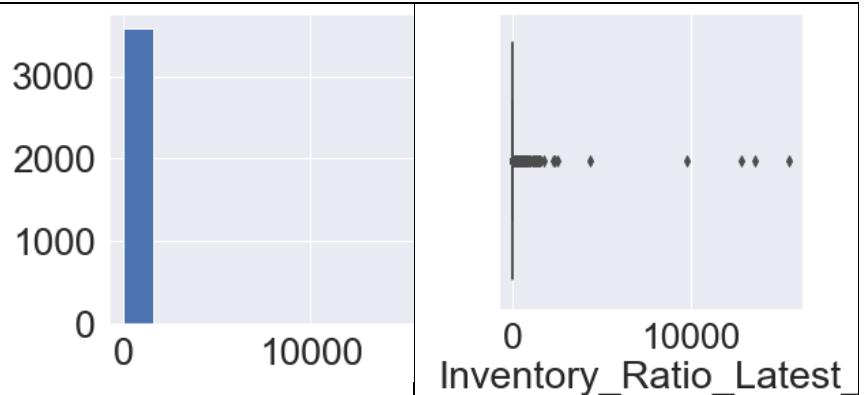
```
Description of ROG_CP_perc
-----
count      3586.00
mean       221.09
std        13980.20
min     -52200.00
25%      -29.51
50%       4.62
75%      52.91
max    822400.00
Name: ROG_CP_perc, dtype: float64
```





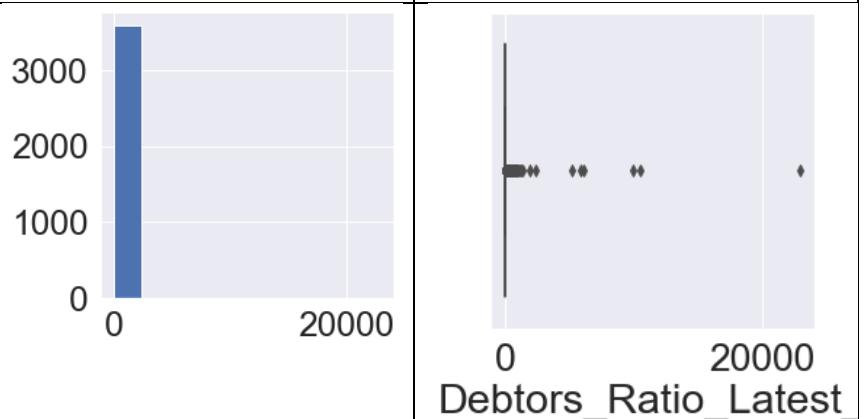
Description of Inventory_Ratio_Latest_

```
count    3586.00
mean     38.15
std      458.61
min      0.00
25%     0.00
50%     3.56
75%     8.94
max    15472.00
Name: Inventory_Ratio_Latest_, dtype: float64
```



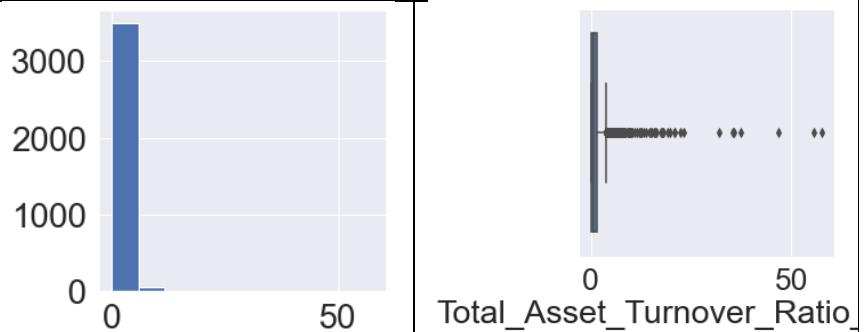
Description of Debtors_Ratio_Latest_

```
count    3586.00
mean     33.02
std      489.50
min      0.00
25%     0.42
50%     3.82
75%     8.54
max    22992.67
Name: Debtors_Ratio_Latest_, dtype: float64
```



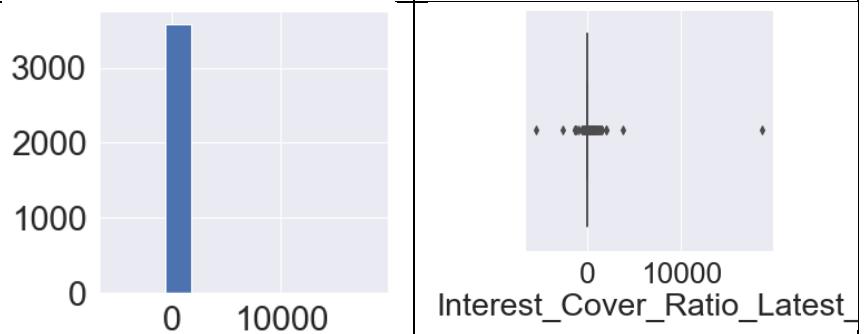
Description of Total_Asset_Turnover_Ratio_Latest_

```
count    3586.00
mean     1.24
std      2.67
min      0.00
25%     0.07
50%     0.60
75%     1.55
max    57.75
Name: Total_Asset_Turnover_Ratio_Latest_, dtype: float64
```



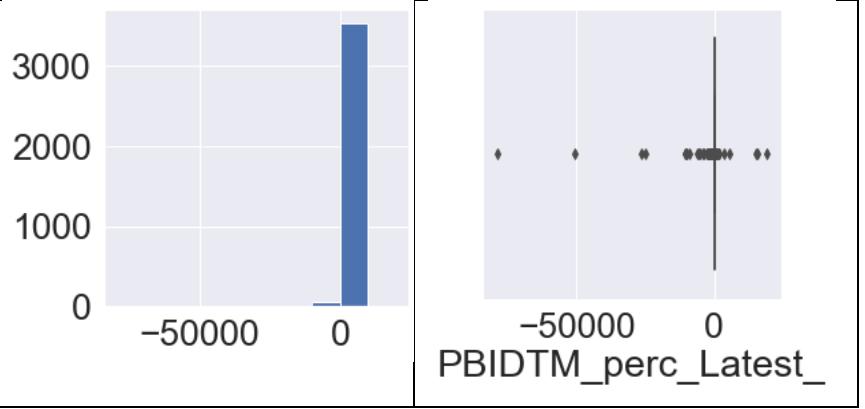
Description of Interest_Cover_Ratio_Latest_

```
count    3586.00
mean     16.38
std      351.69
min    -5450.00
25%     0.00
50%     1.08
75%     3.73
max    18639.40
Name: Interest_Cover_Ratio_Latest_, dtype: float64
```



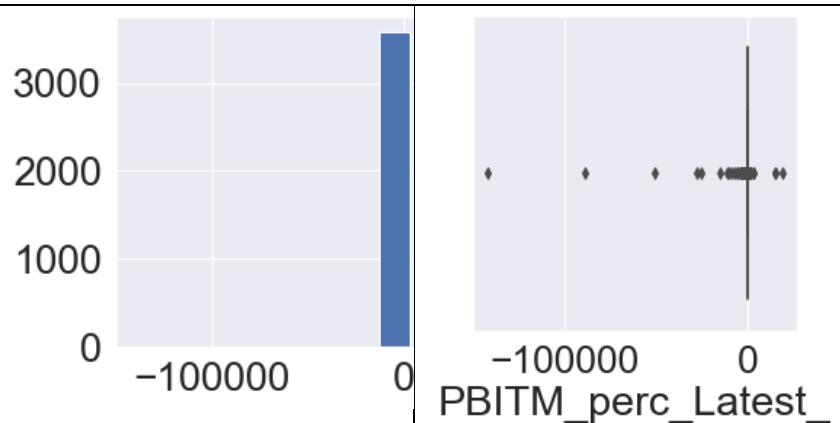
Description of PBIDTM_perc_Latest_

```
count    3586.00
mean    -51.15
std     1794.88
min   -78870.45
25%     0.00
50%     8.07
75%    18.99
max   19233.33
Name: PBIDTM_perc_Latest_, dtype: float64
```



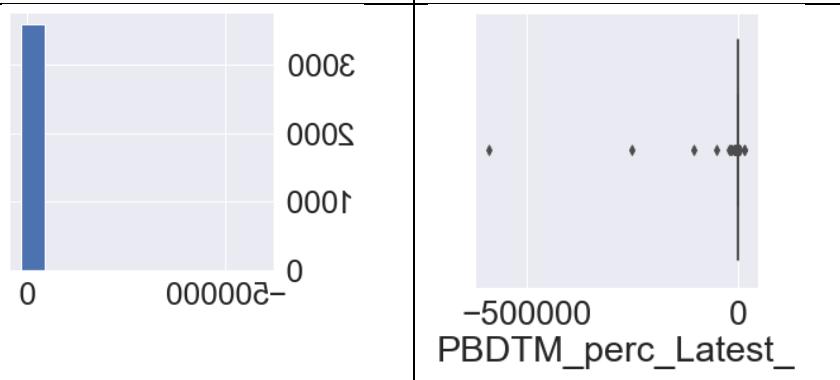
Description of PBITM_perc_Latest_

```
count      3586.00
mean       -109.18
std        3057.21
min     -141600.00
25%        0.00
50%        5.23
75%       14.29
max      19195.70
Name: PBITM_perc_Latest_, dtype: int64
```



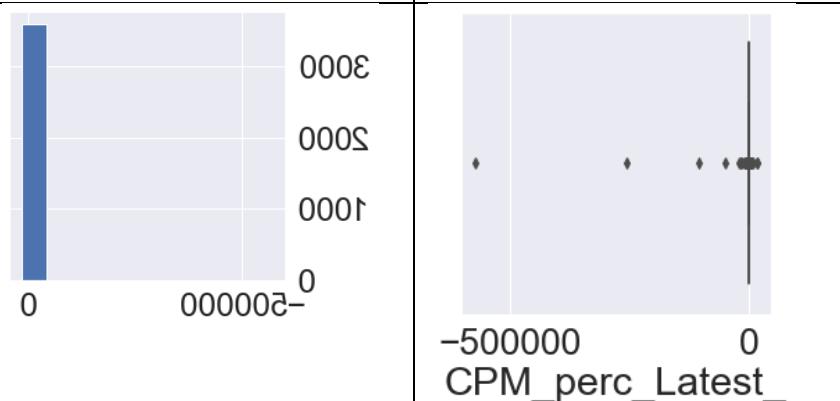
Description of PBDTM_perc_Latest_

```
count      3586.00
mean      -311.48
std       10920.07
min    -590500.00
25%        0.00
50%        4.69
75%       14.10
max      15640.00
Name: PBDTM_perc_Latest_, dtype: float64
```



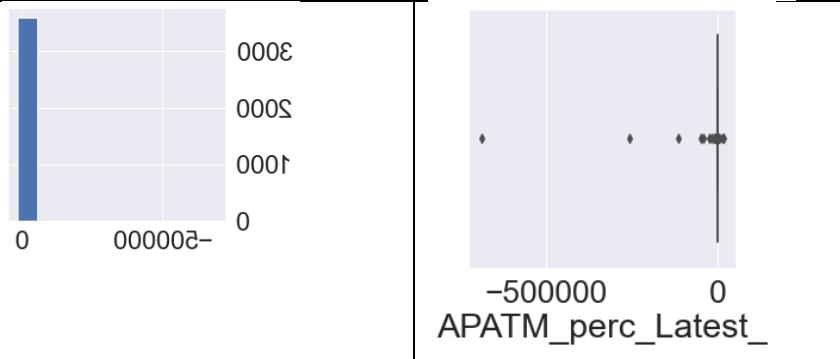
Description of CPM_perc_Latest_

```
count      3586.00
mean      -306.92
std       10674.66
min    -572000.00
25%        0.00
50%        3.89
75%       11.39
max      15640.00
Name: CPM_perc_Latest_, dtype: float64
```



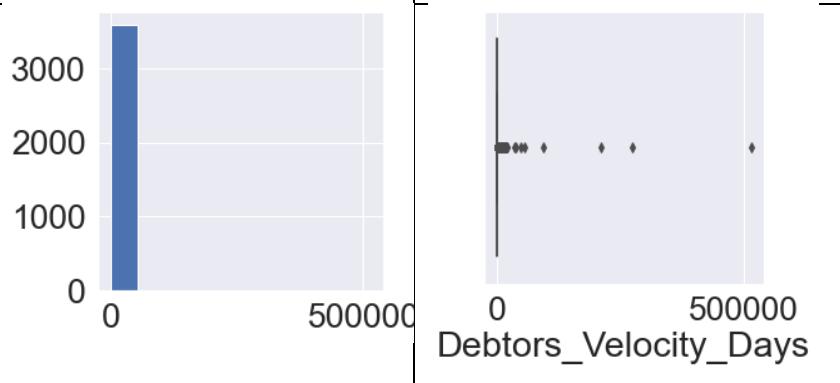
Description of APATM_perc_Latest_

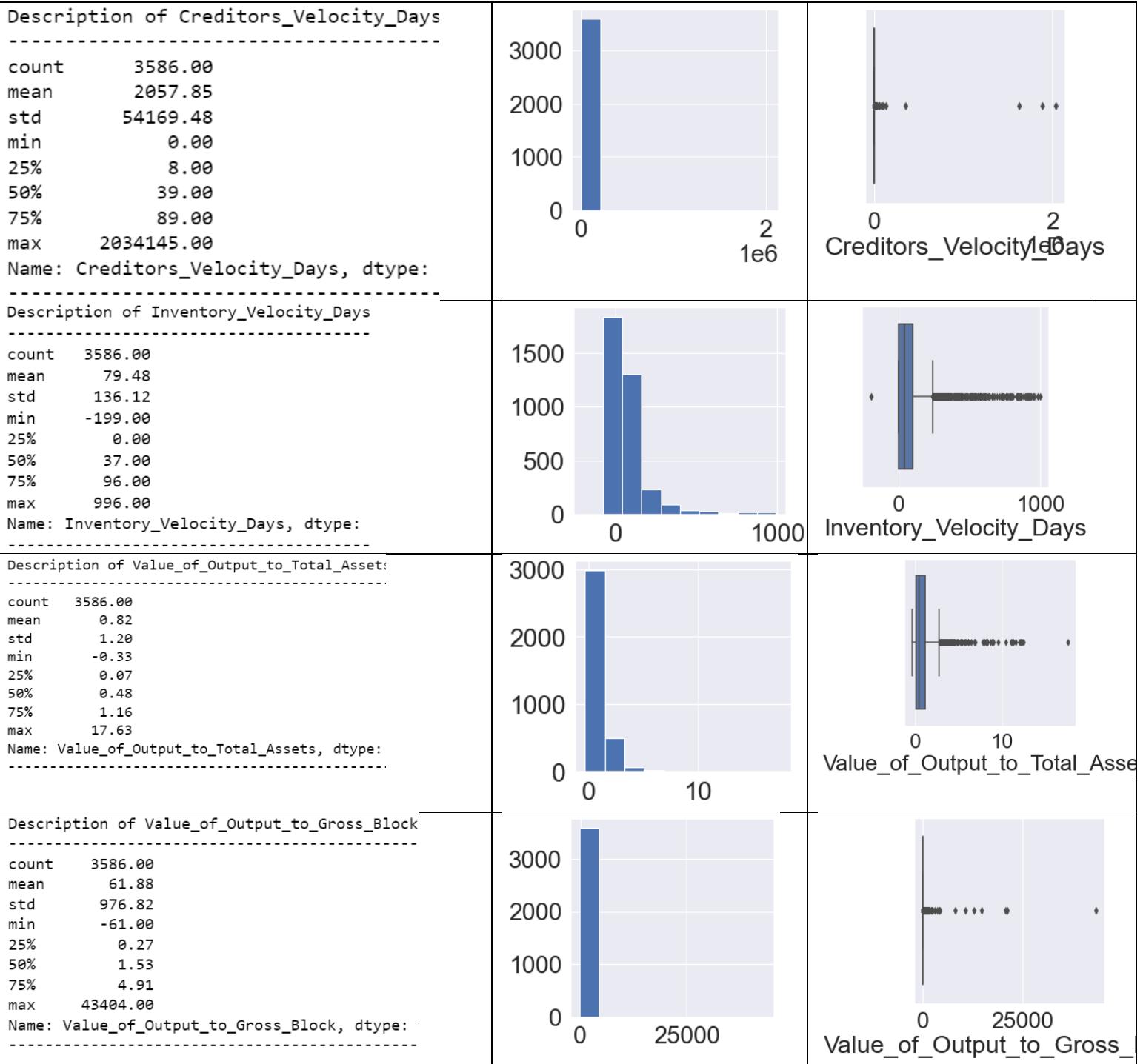
```
count      3586.00
mean      -364.95
std       12498.31
min    -688600.00
25%        0.00
50%        1.59
75%       7.41
max      15266.67
Name: APATM_perc_Latest_, dtype: float64
```



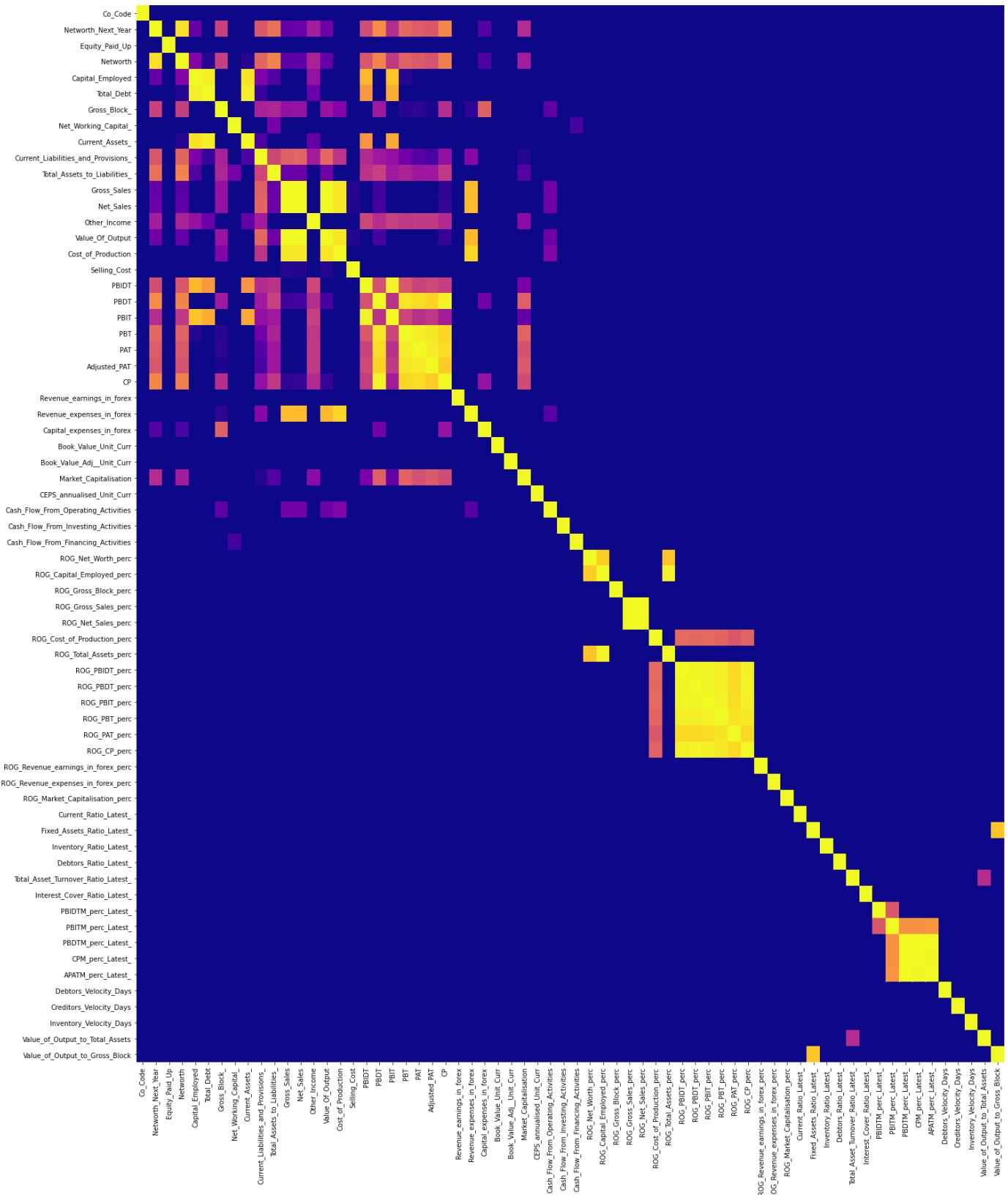
Description of Debtors_Velocity_Days

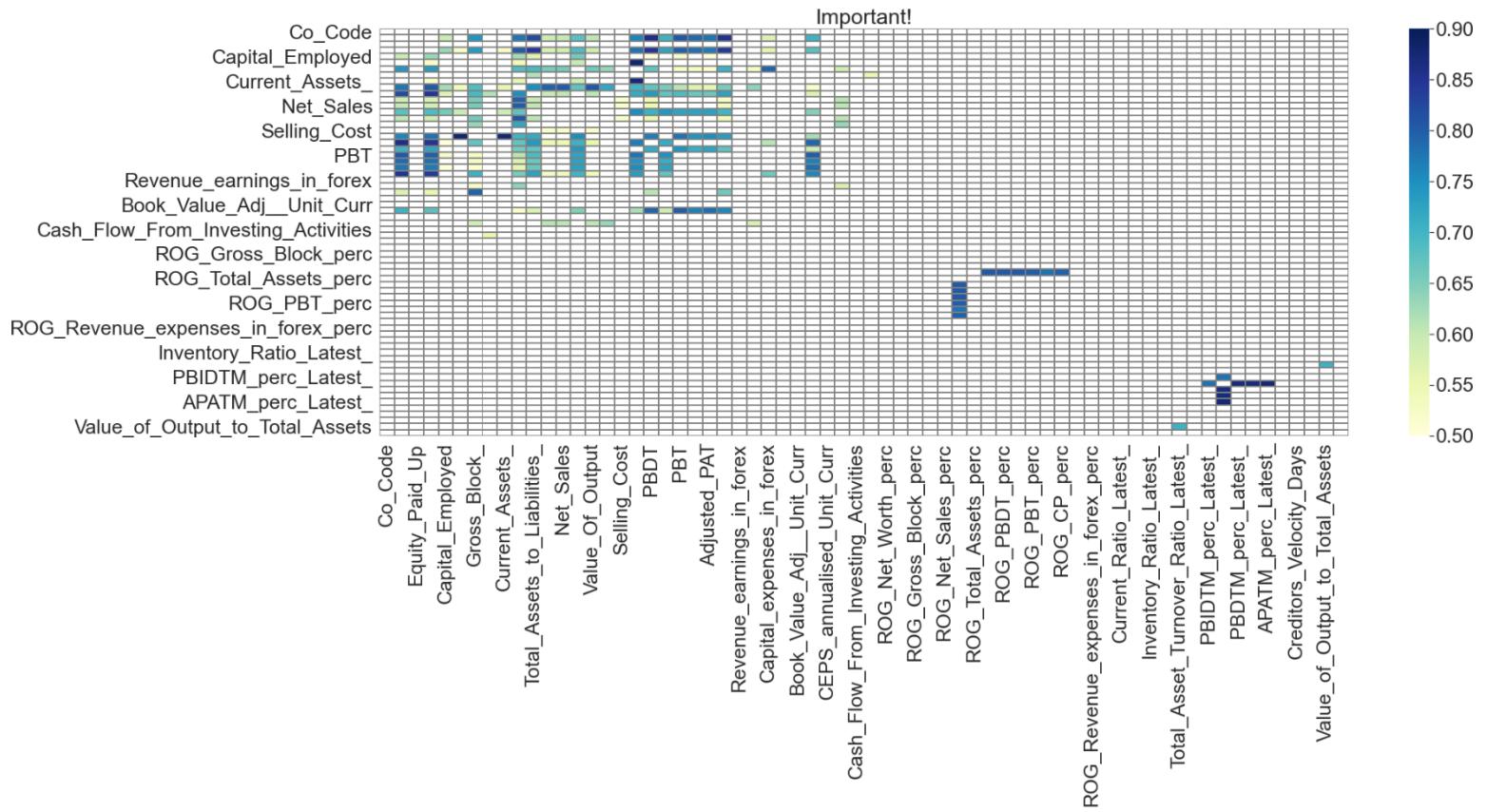
```
count      3586.00
mean       603.89
std        10636.76
min        0.00
25%        8.00
50%       49.00
75%      106.00
max     514721.00
Name: Debtors_Velocity_Days, dtype: float64
```





Bivariate analysis:





We have plotted heatmap with more than 0.5 correlation since we have large number of variables. The following variables have high co-relation:

Networth_Next_Year	Networth_Next_Year	1.00
ROG_Net_Worth_perc	ROG_Net_Worth_perc	1.00
Cash_Flow_From_Investing_Activities	Cash_Flow_From_Investing_Activities	1.00
Cash_Flow_From_Operating_Activities	Cash_Flow_From_Operating_Activities	1.00
CEPS_annualised_Unit_Curr	CEPS_annualised_Unit_Curr	1.00
Market_Capitalisation	Market_Capitalisation	1.00
Book_Value_Adj_Unit_Curr	Book_Value_Adj_Unit_Curr	1.00
Book_Value_Unit_Curr	Book_Value_Unit_Curr	1.00
Capital_expenses_in_forex	Capital_expenses_in_forex	1.00
Revenue_expenses_in_forex	Revenue_expenses_in_forex	1.00
Revenue_earnings_in_forex	Revenue_earnings_in_forex	1.00
CP	CP	1.00
Adjusted_PAT	Adjusted_PAT	1.00
PAT	PAT	1.00
PBT	PBT	1.00
PBIT	PBIT	1.00
PBDT	PBDT	1.00
PBIDT	PBIDT	1.00
Selling_Cost	Selling_Cost	1.00
Cost_of_Production	Cost_of_Production	1.00
Value_Of_Output	Value_Of_Output	1.00
Other_Income	Other_Income	1.00
Net_Sales	Net_Sales	1.00
Gross_Sales	Gross_Sales	1.00
Total_Assets_to_Liabilities_	Total_Assets_to_Liabilities_	1.00
Current_Liabilities_and_Provisions_	Current_Liabilities_and_Provisions_	1.00
Current_Assets_	Current_Assets_	1.00
Net_Working_Capital_	Net_Working_Capital_	1.00

Gross_Block_	Gross_Block_	1.00
Total_Debt	Total_Debt	1.00
Capital_Employed	Capital_Employed	1.00
Cash_Flow_From_Financing_Activities	Cash_Flow_From_Financing_Activities	1.00
ROG_Capital_Employed_perc	ROG_Capital_Employed_perc	1.00
Equity_Paid_Up	Equity_Paid_Up	1.00
ROG_Gross_Block_perc	ROG_Gross_Block_perc	1.00
Value_of_Output_to_Gross_Block	Value_of_Output_to_Gross_Block	1.00
Value_of_Output_to_Total_Assets	Value_of_Output_to_Total_Assets	1.00
Inventory_Velocity_Days	Inventory_Velocity_Days	1.00
Creditors_Velocity_Days	Creditors_Velocity_Days	1.00
Debtors_Velocity_Days	Debtors_Velocity_Days	1.00
APATM_perc_Latest_	APATM_perc_Latest_	1.00
CPM_perc_Latest_	CPM_perc_Latest_	1.00
PBDM_perc_Latest_	PBDM_perc_Latest_	1.00
PBITM_perc_Latest_	PBITM_perc_Latest_	1.00
PBIDTM_perc_Latest_	PBIDTM_perc_Latest_	1.00
Interest_Cover_Ratio_Latest_	Interest_Cover_Ratio_Latest_	1.00
Total_Asset_Turnover_Ratio_Latest_	Total_Asset_Turnover_Ratio_Latest_	1.00
Debtors_Ratio_Latest_	Debtors_Ratio_Latest_	1.00
Inventory_Ratio_Latest_	Inventory_Ratio_Latest_	1.00
Fixed_Assets_Ratio_Latest_	Fixed_Assets_Ratio_Latest_	1.00
Current_Ratio_Latest_	Current_Ratio_Latest_	1.00
ROG_Market_Capitalisation_perc	ROG_Market_Capitalisation_perc	1.00
ROG_Revenue_expenses_in_forex_perc	ROG_Revenue_expenses_in_forex_perc	1.00
ROG_Revenue_earnings_in_forex_perc	ROG_Revenue_earnings_in_forex_perc	1.00
ROG_CP_perc	ROG_CP_perc	1.00
ROG_PAT_perc	ROG_PAT_perc	1.00
ROG_PBT_perc	ROG_PBT_perc	1.00
ROG_PBIT_perc	ROG_PBIT_perc	1.00
ROG_PBDT_perc	ROG_PBDT_perc	1.00
ROG_PBIDT_perc	ROG_PBIDT_perc	1.00
ROG_Total_Assets_perc	ROG_Total_Assets_perc	1.00
ROG_Cost_of_Production_perc	ROG_Cost_of_Production_perc	1.00
ROG_Net_Sales_perc	ROG_Net_Sales_perc	1.00
ROG_Gross_Sales_perc	ROG_Gross_Sales_perc	1.00
Networth	Networth	1.00
default	default	1.00
ROG_Net_Sales_perc	ROG_Gross_Sales_perc	1.00
ROG_Gross_Sales_perc	ROG_Net_Sales_perc	1.00
Net_Sales	Value_Of_Output	1.00
Value_Of_Output	Net_Sales	1.00
PBDM_perc_Latest_	CPM_perc_Latest_	1.00
CPM_perc_Latest_	PBDM_perc_Latest_	1.00
Gross_Sales	Net_Sales	1.00
Net_Sales	Gross_Sales	1.00
Value_Of_Output	Gross_Sales	1.00
Gross_Sales	Value_Of_Output	1.00
APATM_perc_Latest_	PBDM_perc_Latest_	1.00
PBDM_perc_Latest_	APATM_perc_Latest_	1.00
ROG_PBDT_perc	ROG_PBIDT_perc	1.00
ROG_PBIDT_perc	ROG_PBDT_perc	1.00
APATM_perc_Latest_	CPM_perc_Latest_	1.00
CPM_perc_Latest_	APATM_perc_Latest_	1.00
ROG_PBIT_perc	ROG_PBIDT_perc	1.00
ROG_PBIDT_perc	ROG_PBIT_perc	1.00

ROG_PBIT_perc	ROG_PBDT_perc	1.00
ROG_PBDT_perc	ROG_PBIT_perc	1.00
ROG_CP_perc	ROG_PBDT_perc	0.99
ROG_PBDT_perc	ROG_CP_perc	0.99
PBIDT	PBIT	0.99
PBIT	PBIDT	0.99
PBDT	CP	0.99
CP	PBDT	0.99
ROG_PBT_perc	ROG_PBIT_perc	0.99
ROG_PBIT_perc	ROG_PBT_perc	0.99
ROG_Total_Assets_perc	ROG_Capital_Employed_perc	0.99
ROG_Capital_Employed_perc	ROG_Total_Assets_perc	0.99
ROG_CP_perc	ROG_PBIDT_perc	0.99
ROG_PBIDT_perc	ROG_CP_perc	0.99
ROG_PBT_perc	ROG_PBIDT_perc	0.99
ROG_PBIDT_perc	ROG_PBT_perc	0.99
ROG_PBT_perc	ROG_PBDT_perc	0.99
ROG_PBDT_perc	ROG_PBT_perc	0.99
ROG_PBIT_perc	ROG_CP_perc	0.99
ROG_CP_perc	ROG_PBIT_perc	0.99
PAT	PBT	0.99
PBT	PAT	0.99
Total_Debt	Capital_Employed	0.99
Capital_Employed	Total_Debt	0.99
Total_Debt	Current_Assets_	0.99
Current_Assets_	Total_Debt	0.99
Adjusted_PAT	PAT	0.98
PAT	Adjusted_PAT	0.98
ROG_CP_perc	ROG_PBT_perc	0.98
ROG_PBT_perc	ROG_CP_perc	0.98
Gross_Sales	Cost_of_Production	0.98
Cost_of_Production	Gross_Sales	0.98
Net_Sales	Cost_of_Production	0.98
Cost_of_Production	Net_Sales	0.98
Adjusted_PAT	PBT	0.98
PBT	Adjusted_PAT	0.98
PBDT	PBT	0.98
PBT	PBDT	0.98
Cost_of_Production	Value_Of_Output	0.98
Value_Of_Output	Cost_of_Production	0.98
ROG_PBT_perc	ROG_PAT_perc	0.98
ROG_PAT_perc	ROG_PBT_perc	0.98
Current_Assets_	Capital_Employed	0.98
Capital_Employed	Current_Assets_	0.98
Networth_Next_Year	Networth	0.97
Networth	Networth_Next_Year	0.97
PBDT	PAT	0.97
PAT	PBDT	0.97
CP	CP	0.97
CP	PAT	0.97
ROG_PAT_perc	ROG_PBIT_perc	0.97
ROG_PBIT_perc	ROG_PAT_perc	0.97
ROG_PAT_perc	ROG_PBDT_perc	0.96
ROG_PBDT_perc	ROG_PAT_perc	0.96
ROG_PBIDT_perc	ROG_PAT_perc	0.96
ROG_PAT_perc	ROG_PBIDT_perc	0.96

	ROG_CP_perc	0.96
ROG_CP_perc	ROG_PAT_perc	0.96
Revenue_expenses_in_forex	Cost_of_Production	0.96
Cost_of_Production	Revenue_expenses_in_forex	0.96
CP	PBT	0.96
PBT	CP	0.96
PBDT	Adjusted_PAT	0.95
Adjusted_PAT	PBDT	0.95
ROG_Capital_Employed_perc	ROG_Net_Worth_perc	0.95
ROG_Net_Worth_perc	ROG_Capital_Employed_perc	0.95
Adjusted_PAT	CP	0.95
CP	Adjusted_PAT	0.95
Fixed_Assets_Ratio_Latest_	Value_of_Output_to_Gross_Block	0.94
Value_of_Output_to_Gross_Block	Fixed_Assets_Ratio_Latest_	0.94
Capital_Employed	PBIT	0.94
PBIT	Capital_Employed	0.94
ROG_Net_Worth_perc	ROG_Total_Assets_perc	0.94
ROG_Total_Assets_perc	ROG_Net_Worth_perc	0.94
Gross_Sales	Revenue_expenses_in_forex	0.93
Revenue_expenses_in_forex	Gross_Sales	0.93
Net_Sales	Revenue_expenses_in_forex	0.93
Revenue_expenses_in_forex	Net_Sales	0.93
Value_Of_Output	Revenue_expenses_in_forex	0.93
Revenue_expenses_in_forex	Value_Of_Output	0.93
Capital_Employed	PBIDT	0.92
PBIDT	Capital_Employed	0.92
PBIT	Total_Debt	0.91
Total_Debt	PBIT	0.91
Current_Assets_	PBIT	0.90
PBIT	Current_Assets_	0.90
PBIDT	Total_Debt	0.88
Total_Debt	PBIDT	0.88
PBITM_perc_Latest_	APATM_perc_Latest_	0.88
APATM_perc_Latest_	PBITM_perc_Latest_	0.88
Current_Assets_	PBIDT	0.87
PBIDT	Current_Assets_	0.87
PBDTM_perc_Latest_	PBITM_perc_Latest_	0.87
PBITM_perc_Latest_	PBDTM_perc_Latest_	0.87
CPM_perc_Latest_	PBITM_perc_Latest_	0.87
PBITM_perc_Latest_	CPM_perc_Latest_	0.87
PBDT	Networth_Next_Year	0.86
Networth_Next_Year	PBDT	0.86
CP	Networth_Next_Year	0.85
Networth_Next_Year	CP	0.85
PBDT	Networth	0.85
Networth	PBDT	0.85
	Total_Assets_to_Liabilities_	0.85
Total_Assets_to_Liabilities_	Networth	0.85
Networth	CP	0.84
CP	Networth	0.84
Total_Assets_to_Liabilities_	Networth_Next_Year	0.82
Networth_Next_Year	Total_Assets_to_Liabilities_	0.82
ROG_Cost_of_Production_perc	ROG_PBIDT_perc	0.81
ROG_PBIDT_perc	ROG_Cost_of_Production_perc	0.81
ROG_Cost_of_Production_perc	ROG_PBIT_perc	0.81
ROG_PBIT_perc	ROG_Cost_of_Production_perc	0.81

ROG_PBDT_perc	ROG_Cost_of_Production_perc	0.81
ROG_Cost_of_Production_perc	ROG_PBDT_perc	0.81
Value_Of_Output	Current_Liabilities_and_Provisions_	0.81
Current_Liabilities_and_Provisions_	Value_Of_Output	0.81
Networth_Next_Year	PBT	0.81
PBT	Networth_Next_Year	0.81
Current_Liabilities_and_Provisions_	Networth	0.80
Networth	Current_Liabilities_and_Provisions_	0.80
Current_Liabilities_and_Provisions_	Net_Sales	0.80
Net_Sales	Current_Liabilities_and_Provisions_	0.80
ROG_Cost_of_Production_perc	ROG_PBT_perc	0.80
ROG_PBT_perc	ROG_Cost_of_Production_perc	0.80
PBT	Market_Capitalisation	0.80
Market_Capitalisation	PBT	0.80
ROG_CP_perc	ROG_Cost_of_Production_perc	0.80
ROG_Cost_of_Production_perc	ROG_CP_perc	0.80
PBT	Networth	0.80
Networth	PBT	0.80
Current_Liabilities_and_Provisions_	Gross_Sales	0.80
Gross_Sales	Current_Liabilities_and_Provisions_	0.80
Capital_expenses_in_forex	Gross_Block_	0.79
Gross_Block_	Capital_expenses_in_forex	0.79
PBDT	Market_Capitalisation	0.79
Market_Capitalisation	PBDT	0.79
Networth_Next_Year	PAT	0.79
PAT	Networth_Next_Year	0.79
Adjusted_PAT	Market_Capitalisation	0.78
Market_Capitalisation	Adjusted_PAT	0.78
Networth	PAT	0.78
PAT	Networth	0.78
PBITM_perc_Latest_	PBIDTM_perc_Latest_	0.78
PBIDTM_perc_Latest_	PBITM_perc_Latest_	0.78
ROG_Cost_of_Production_perc	ROG_PAT_perc	0.78
ROG_PAT_perc	ROG_Cost_of_Production_perc	0.78
PBT	PBIDT	0.78
PBIDT	PBT	0.78
	Networth	0.78
Networth	PBIDT	0.78
Networth_Next_Year	Current_Liabilities_and_Provisions_	0.78
Current_Liabilities_and_Provisions_	Networth_Next_Year	0.78
Networth_Next_Year	Adjusted_PAT	0.78
Adjusted_PAT	Networth_Next_Year	0.78
Networth	Adjusted_PAT	0.77
Adjusted_PAT	Networth	0.77
PBIDT	PBDT	0.77
PBDT	PBIDT	0.77
PAT	Market_Capitalisation	0.77
Market_Capitalisation	PAT	0.77
Networth_Next_Year	PBIDT	0.76
PBIDT	Networth_Next_Year	0.76
Market_Capitalisation	CP	0.76
CP	Market_Capitalisation	0.76
Adjusted_PAT	PBIDT	0.75
PBIDT	Adjusted_PAT	0.75
	Other_Income	0.75
Other_Income	PBIDT	0.75

Current_Liabilities_and_Provisions_	Total_Assets_to_Liabilities_	0.75
Total_Assets_to_Liabilities_	Current_Liabilities_and_Provisions_	0.75
PBT	PBIT	0.75
PBIT	PBT	0.75
PAT	PBIDT	0.74
PBIDT	PAT	0.74
Gross_Block_	Networth_Next_Year	0.74
Networth_Next_Year	Gross_Block_	0.74
PBIT	Other_Income	0.74
Other_Income	PBIT	0.74
PBDT	Total_Assets_to_Liabilities_	0.74
Total_Assets_to_Liabilities_	PBDT	0.74
Gross_Block_	Networth	0.74
Networth	Gross_Block_	0.74
PBIDT	CP	0.73
CP	PBIDT	0.73
	Total_Assets_to_Liabilities_	0.73
Total_Assets_to_Liabilities_	CP	0.73
Adjusted_PAT	Other_Income	0.73
Other_Income	Adjusted_PAT	0.73
	PAT	0.73
PAT	Other_Income	0.73
PBIT	Networth	0.73
Networth	PBIT	0.73
PBIT	Adjusted_PAT	0.73
Adjusted_PAT	PBIT	0.73
PBT	Other_Income	0.72
Other_Income	PBT	0.72
Cost_of_Production	Current_Liabilities_and_Provisions_	0.72
Current_Liabilities_and_Provisions_	Cost_of_Production	0.72
PBIDT	Total_Assets_to_Liabilities_	0.72
Total_Assets_to_Liabilities_	PBIDT	0.72
PBIT	PBDT	0.71
PBDT	PBIT	0.71
PAT	PBIT	0.71
PBIT	PAT	0.71
PBDT	Other_Income	0.71
Other_Income	PBDT	0.71
CP	Gross_Block_	0.71
Gross_Block_	CP	0.71
Value_of_Output_to_Total_Assets	Total_Asset_Turnover_Ratio_Latest_	0.71
Total_Asset_Turnover_Ratio_Latest_	Value_of_Output_to_Total_Assets	0.71
PBIT	Networth_Next_Year	0.70
Networth_Next_Year	PBIT	0.70
Market_Capitalisation	Networth_Next_Year	0.70
Networth_Next_Year	Market_Capitalisation	0.70
Current_Liabilities_and_Provisions_	PBIDT	0.70
PBIDT	Current_Liabilities_and_Provisions_	0.70
CP	Other_Income	0.70
Other_Income	CP	0.70
Gross_Block_	Total_Assets_to_Liabilities_	0.69
Total_Assets_to_Liabilities_	Gross_Block_	0.69
PBT	Total_Assets_to_Liabilities_	0.69
Total_Assets_to_Liabilities_	PBT	0.69
Networth	Other_Income	0.69
Other_Income	Networth	0.69

Gross_Block_	Current_Liabilities_and_Provisions_	0.69
Current_Liabilities_and_Provisions_	Gross_Block_	0.69
Networth_Next_Year	Other_Income	0.68
Other_Income	Networth_Next_Year	0.68
Networth	Market_Capitalisation	0.68
Market_Capitalisation	Networth	0.68
Other_Income	Current_Liabilities_and_Provisions_	0.67
Current_Liabilities_and_Provisions_	Other_Income	0.67
CP	PBIT	0.67
PBIT	CP	0.67
Gross_Block_	PBDT	0.67
PBDT	Gross_Block_	0.67
Total_Assets_to_Liabilities_	PBIT	0.67
PBIT	Total_Assets_to_Liabilities_	0.67
PBDT	Current_Liabilities_and_Provisions_	0.67
Current_Liabilities_and_Provisions_	PBDT	0.67
Adjusted_PAT	Total_Assets_to_Liabilities_	0.67
Total_Assets_to_Liabilities_	Adjusted_PAT	0.67
PAT	Total_Assets_to_Liabilities_	0.67
Total_Assets_to_Liabilities_	PAT	0.67
Gross_Block_	Value_Of_Output	0.66
Value_Of_Output	Gross_Block_	0.66
CP	Capital_expenses_in_forex	0.66
Capital_expenses_in_forex	CP	0.66
Net_Sales	Gross_Block_	0.66
Gross_Block_	Net_Sales	0.66
Other_Income	Capital_Employed	0.66
Capital_Employed	Other_Income	0.66
Gross_Sales	Gross_Block_	0.66
Gross_Block_	Gross_Sales	0.66
PBIT	Current_Liabilities_and_Provisions_	0.66
Current_Liabilities_and_Provisions_	PBIT	0.66
CP	Current_Liabilities_and_Provisions_	0.65
Current_Liabilities_and_Provisions_	CP	0.65
Other_Income	Market_Capitalisation	0.65
Market_Capitalisation	Other_Income	0.65
Revenue_expenses_in_forex	Current_Liabilities_and_Provisions_	0.64
Current_Liabilities_and_Provisions_	Revenue_expenses_in_forex	0.64
Gross_Block_	Cost_of_Production	0.63
Cost_of_Production	Gross_Block_	0.63
	Cash_Flow_From_Operating_Activities	0.63
Cash_Flow_From_Operating_Activities	Cost_of_Production	0.63
Networth	Capital_Employed	0.63
Capital_Employed	Networth	0.63
Current_Liabilities_and_Provisions_	Capital_Employed	0.62
Capital_Employed	Current_Liabilities_and_Provisions_	0.62
Market_Capitalisation	PBIDT	0.62
PBIDT	Market_Capitalisation	0.62
Total_Assets_to_Liabilities_	Net_Working_Capital_	0.61
Net_Working_Capital_	Total_Assets_to_Liabilities_	0.61
Gross_Sales	Cash_Flow_From_Operating_Activities	0.61
Cash_Flow_From_Operating_Activities	Gross_Sales	0.61
PBDT	Capital_expenses_in_forex	0.61
Capital_expenses_in_forex	PBDT	0.61
PBT	Current_Liabilities_and_Provisions_	0.61
Current_Liabilities_and_Provisions_	PBT	0.61

Net_Sales	Cash_Flow_From_Operating_Activities	0.61
Cash_Flow_From_Operating_Activities	Net_Sales	0.61
Total_Assets_to_Liabilities_	Value_Of_Output	0.61
Value_Of_Output	Total_Assets_to_Liabilities_	0.61
Cash_Flow_From_Operating_Activities	Value_Of_Output	0.61
Value_Of_Output	Cash_Flow_From_Operating_Activities	0.61
Networth_Next_Year	Value_Of_Output	0.61
Value_Of_Output	Networth_Next_Year	0.61
Net_Sales	Total_Assets_to_Liabilities_	0.60
Total_Assets_to_Liabilities_	Net_Sales	0.60
Networth_Next_Year	Net_Sales	0.60
Net_Sales	Networth_Next_Year	0.60
Total_Debt	Other_Income	0.60
Other_Income	Total_Debt	0.60
Total_Assets_to_Liabilities_	Gross_Sales	0.60
Gross_Sales	Total_Assets_to_Liabilities_	0.60
Capital_Employed	Networth_Next_Year	0.59
Networth_Next_Year	Capital_Employed	0.59
Other_Income	Current_Assets_	0.59
Current_Assets_	Other_Income	0.59
Networth_Next_Year	Gross_Sales	0.59
Gross_Sales	Networth_Next_Year	0.59
Value_Of_Output	Networth	0.59
Networth	Value_Of_Output	0.59
Market_Capitalisation	PBIT	0.59
PBIT	Market_Capitalisation	0.59
Cash_Flow_From_Operating_Activities	Gross_Block_	0.59
Gross_Block_	Cash_Flow_From_Operating_Activities	0.59
Networth	Net_Sales	0.59
Net_Sales	Networth	0.59
Gross_Sales	Networth	0.58
Networth	Gross_Sales	0.58
Revenue_expenses_in_forex	Cash_Flow_From_Operating_Activities	0.58
Cash_Flow_From_Operating_Activities	Revenue_expenses_in_forex	0.58
Current_Liabilities_and_Provisions_	PAT	0.57
PAT	Current_Liabilities_and_Provisions_	0.57
Market_Capitalisation	Total_Assets_to_Liabilities_	0.57
Total_Assets_to_Liabilities_	Market_Capitalisation	0.57
Networth_Next_Year	Capital_expenses_in_forex	0.57
Capital_expenses_in_forex	Networth_Next_Year	0.57
Capital_Employed	Total_Assets_to_Liabilities_	0.57
Total_Assets_to_Liabilities_	Capital_Employed	0.57
Capital_expenses_in_forex	Networth	0.56
Networth	Capital_expenses_in_forex	0.56
Current_Assets_	Current_Liabilities_and_Provisions_	0.56
Current_Liabilities_and_Provisions_	Current_Assets_	0.56
Adjusted_PAT	Current_Liabilities_and_Provisions_	0.56
Current_Liabilities_and_Provisions_	Adjusted_PAT	0.56
Net_Working_Capital_	Cash_Flow_From_Financing_Activities	0.56
Cash_Flow_From_Financing_Activities	Net_Working_Capital_	0.56
Value_Of_Output	PBDT	0.56
PBDT	Value_Of_Output	0.56
Net_Sales	PBDT	0.55
PBDT	Net_Sales	0.55
Gross_Sales	PBDT	0.55
PBDT	Gross_Sales	0.55

CP	Value_Of_Output	0.54
Value_Of_Output	CP	0.54
Total_Debt	Current_Liabilities_and_Provisions_	0.54
Current_Liabilities_and_Provisions_	Total_Debt	0.54
CP	Net_Sales	0.53
Net_Sales	CP	0.53
Gross_Sales	CP	0.53
CP	Gross_Sales	0.53
Gross_Block_	PAT	0.53
PAT	Gross_Block_	0.53
Revenue_expenses_in_forex	Gross_Block_	0.53
Gross_Block_	Revenue_expenses_in_forex	0.53
	PBT	0.52
PBT	Gross_Block_	0.52
Current_Assets_	Networth	0.52
Networth	Current_Assets_	0.52
PBT	Capital_Employed	0.52
Capital_Employed	PBT	0.52
Current_Liabilities_and_Provisions_	Market_Capitalisation	0.52
Market_Capitalisation	Current_Liabilities_and_Provisions_	0.52
Selling_Cost	Value_Of_Output	0.52
Value_Of_Output	Selling_Cost	0.52
Gross_Sales	Selling_Cost	0.52
Selling_Cost	Gross_Sales	0.52
Total_Debt	Networth	0.52
Networth	Total_Debt	0.52
Selling_Cost	Net_Sales	0.52
Net_Sales	Selling_Cost	0.52
Gross_Block_	Adjusted_PAT	0.51
Adjusted_PAT	Gross_Block_	0.51

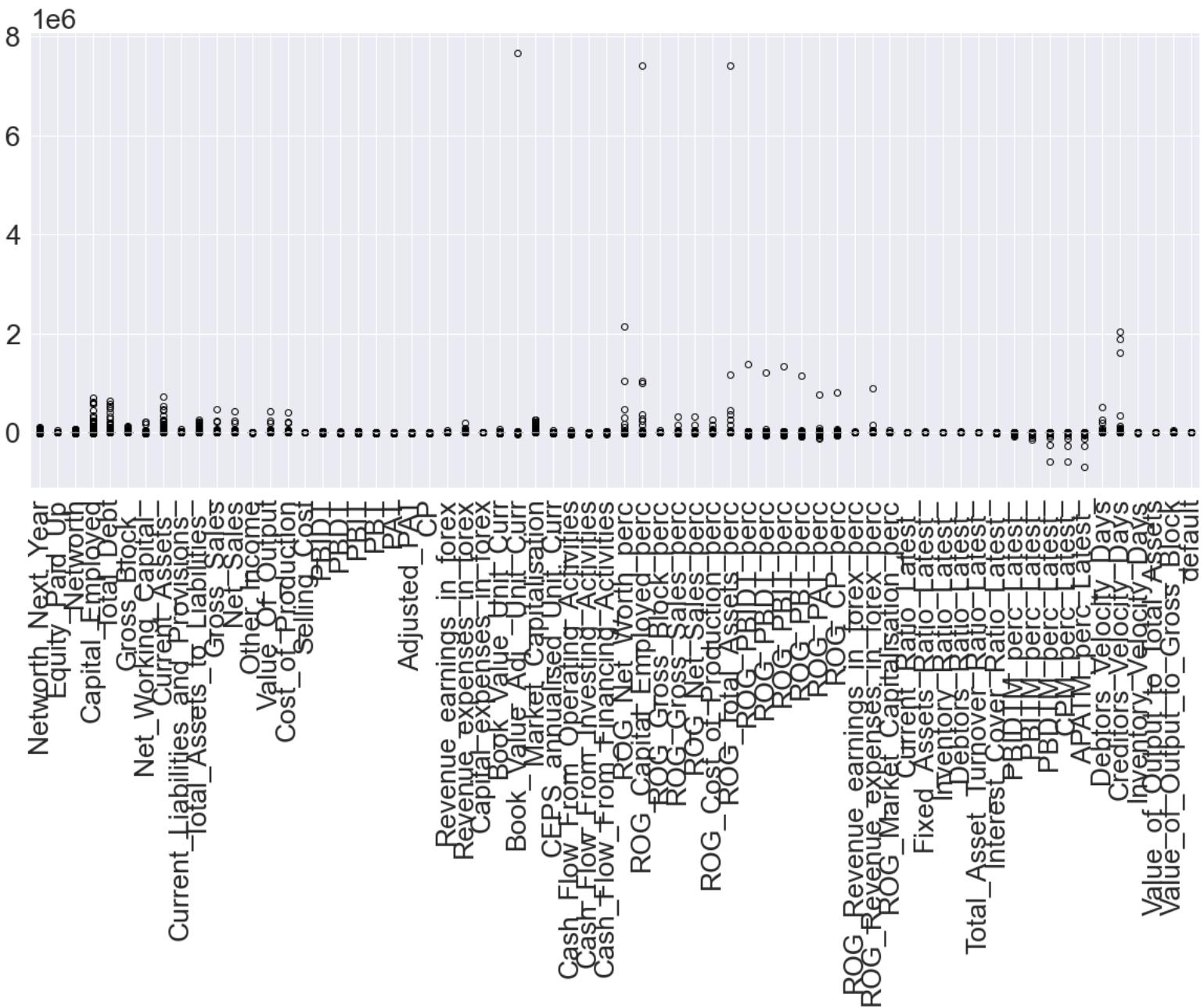
Outliers:

Let us look the outliers in each column:

Networth_Next_Year	676
Equity_Paid_Up	448
Networth	650
Capital_Employed	596
Total_Debt	583
Gross_Block_	540
Net_Working_Capital_	625
Current_Assets_	577
Current_Liabilities_and_Provisions_	581
Total_Assets_to_Liabilities_	574
Gross_Sales	554
Net_Sales	556
Other_Income	603
Value_Of_Output	559
Cost_of_Production	560
Selling_Cost	605
PBIDT	671
PBDT	815
PBIT	720
PBT	941
PAT	959
Adjusted_PAT	954
CP	816

Revenue_earnings_in_forex	738
Revenue_expenses_in_forex	693
Capital_expenses_in_forex	694
Book_Value_Unit_Curr	485
Book_Value_Adj__Unit_Curr	486
Market_Capitalisation	639
CEPS_annualised_Unit_Curr	602
Cash_Flow_From_Operating_Activities	801
Cash_Flow_From_Investing_Activities	876
Cash_Flow_From_Financing_Activities	1005
ROG_Net_Worth_perc	747
ROG_Capital_Employed_perc	572
ROG_Gross_Block_perc	830
ROG_Gross_Sales_perc	671
ROG_Net_Sales_perc	667
ROG_Cost_of_Production_perc	675
ROG_Total_Assets_perc	483
ROG_PBIDT_perc	611
ROG_PBDT_perc	628
ROG_PBIT_perc	616
ROG_PBT_perc	611
ROG_PAT_perc	598
ROG_CP_perc	637
ROG_Revenue_earnings_in_forex_perc	1317
ROG_Revenue_expenses_in_forex_perc	1615
ROG_Market_Capitalisation_perc	497
Current_Ratio_Latest_	565
Fixed_Assets_Ratio_Latest_	495
Inventory_Ratio_Latest_	376
Debtors_Ratio_Latest_	370
Total_Asset_Turnover_Ratio_Latest_	201
Interest_Cover_Ratio_Latest_	723
PBIDTM_perc_Latest_	595
PBITM_perc_Latest_	718
PBDTM_perc_Latest_	696
CPM_perc_Latest_	720
APATM_perc_Latest_	934
Debtors_Velocity_Days	398
Creditors_Velocity_Days	391
Inventory_Velocity_Days	263
Value_of_Output_to_Total_Assets	150
Value_of_Output_to_Gross_Block	481

Since all the columns have high number of outliers, we have not treated the outliers because this will affect our analysis significantly.



Removing outliers is legitimate only for specific reasons. Outliers can be very informative about the subject-area and data collection process. ... Outliers increase the variability in your data, which decreases statistical power. Consequently, excluding outliers can cause your results to become statistically significant.

Now, let us Split the data into Train and Test dataset in a ratio of 67:33 and using random_state =42. Before we build the model using statsmodel library, first let us calculate variance_inflation_factor. A rule of thumb for interpreting the variance inflation factor:

- 1 = not correlated.
 - Between 1 and 5 = moderately correlated.
 - Greater than 5 = highly correlated.

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

We have created logistic regression equation & stored it in f_1

```
model = SM.logit(formula='Dependent Variable ~ ΣIndependent Variables (k)'
```

```
    data = 'Data Frame containing the required values').fit()
```

One way to measure multicollinearity is the variance inflation factor (VIF), which assesses how much the variance of an estimated regression coefficient increases if your predictors are correlated. A VIF between 5 and 10 indicates high correlation that may be problematic. That is why we are considering VIF less than 5.

	variables	VIF
11	CEPS_annualised_Unit_Curr	4.59
10	Market_Capitalisation	4.11
4	Selling_Cost	4.08
1	Total_Debt	4.06
16	ROG_Capital_Employed_perc	3.86
0	Equity_Paid_Up	3.82
3	Other_Income	3.67
7	Revenue_expenses_in_forex	3.66
5	Adjusted_PAT	3.64
12	Cash_Flow_From_Operating_Activities	3.60
9	Book_Value_Adj__Unit_Curr	3.31
29	Total_Asset_Turnover_Ratio_Latest_	3.30
22	ROG_CP_perc	3.26
21	ROG_PBIT_perc	3.23
20	ROG_Total_Assets_perc	3.23

6	Revenue_earnings_in_forex	3.04
2	Net_Working_Capital_	2.86
14	Cash_Flow_From_Financing_Activities	2.84
15	ROG_Net_Worth_perc	2.83
13	Cash_Flow_From_Investing_Activities	2.42
28	Debtors_Ratio_Latest_	2.40
32	Debtors_Velocity_Days	2.40
35	Value_of_Output_to_Gross_Block	2.36
27	Inventory_Ratio_Latest_	2.29
30	Interest_Cover_Ratio_Latest_	2.28
26	Current_Ratio_Latest_	2.25
33	Creditors_Velocity_Days	2.14
18	ROG_Net_Sales_perc	2.12
19	ROG_Cost_of_Production_perc	2.00
34	Inventory_Velocity_Days	1.87
31	PBITM_perc_Latest_	1.83
25	ROG_Market_Capitalisation_perc	1.63
17	ROG_Gross_Block_perc	1.53

So we have considered variables with VIF less than 5 for model building. Now let us look at the model 1 results:

Logit Regression Results

Dep. Variable:	default	No. Observations:	3586				
Model:	Logit	Df Residuals:	3552				
Method:	MLE	Df Model:	33				
Date:	Sat, 05 Feb 2022	Pseudo R-squ.:	0.6576				
Time:	04:43:14	Log-Likelihood:	-420.80				
converged:	True	LL-Null:	-1229.0				
Covariance Type:	nonrobust	LLR p-value:	0.000				
		coef	std err	z	P> z	[0.025	0.975]
	Intercept	-0.0474	0.203	-0.234	0.815	-0.444	0.350
	ROG_Gross_Block_perc	-0.0247	0.017	-1.462	0.144	-0.058	0.008
	ROG_Market_Capitalisation_perc	-0.0007	0.002	-0.310	0.756	-0.005	0.004
	PBITM_perc_Latest_	-0.0279	0.008	-3.602	0.000	-0.043	-0.013
	Inventory_Velocity_Days	-0.0002	0.001	-0.115	0.908	-0.003	0.002
	ROG_Cost_of_Production_perc	-0.0037	0.003	-1.122	0.262	-0.010	0.003
	ROG_Net_Sales_perc	-0.0015	0.004	-0.418	0.676	-0.008	0.005
	Creditors_Velocity_Days	0.0006	0.001	0.534	0.593	-0.002	0.003

Current_Ratio_Latest_	-0.5146	0.075	-6.823	0.000	-0.662	-0.367
Interest_Cover_Ratio_Latest_	-0.0746	0.039	-1.911	0.056	-0.151	0.002
Inventory_Ratio_Latest_	-0.0308	0.017	-1.786	0.074	-0.065	0.003
Value_of_Output_to_Gross_Block	-0.0478	0.035	-1.367	0.172	-0.116	0.021
Debtors_Velocity_Days	-0.0033	0.001	-2.863	0.004	-0.006	-0.001
Debtors_Ratio_Latest_	-0.0348	0.019	-1.856	0.063	-0.072	0.002
Cash_Flow_From_Investing_Activities	-0.0089	0.035	-0.258	0.796	-0.077	0.059
ROG_Net_Worth_perc	-0.0205	0.009	-2.283	0.022	-0.038	-0.003
Cash_Flow_From_Financing_Activities	0.0072	0.031	0.232	0.816	-0.053	0.068
Net_Working_Capital_	-0.0038	0.003	-1.242	0.214	-0.010	0.002
Revenue_earnings_in_forex	-0.0171	0.026	-0.662	0.508	-0.068	0.033
ROG_Total_Assets_perc	-0.0141	0.008	-1.794	0.073	-0.030	0.001
ROG_PBIT_perc	0.0018	0.002	1.037	0.300	-0.002	0.005
ROG_CP_perc	-0.0010	0.002	-0.544	0.587	-0.004	0.002
Total_Asset_Turnover_Ratio_Latest_	0.0994	0.109	0.911	0.362	-0.114	0.313
Book_Value_Adj_Unit_Curr	-0.1233	0.010	-12.680	0.000	-0.142	-0.104
Cash_Flow_From_Operating_Activities	-0.0052	0.017	-0.300	0.764	-0.039	0.029
Adjusted_PAT	0.0056	0.037	0.150	0.880	-0.068	0.079
Revenue_expenses_in_forex	0.0500	0.028	1.801	0.072	-0.004	0.104
Other_Income	0.0305	0.051	0.597	0.550	-0.070	0.131
Equity_Paid_Up	-0.0108	0.010	-1.125	0.261	-0.030	0.008
ROG_Capital_Employed_perc	0.0182	0.008	2.221	0.026	0.002	0.034
Total_Debt	0.0117	0.003	3.542	0.000	0.005	0.018
Selling_Cost	-0.0266	0.060	-0.441	0.659	-0.145	0.091
Market_Capitalisation	-0.0059	0.002	-2.795	0.005	-0.010	-0.002
CEPS_annualised_Unit_Curr	-0.0869	0.030	-2.883	0.004	-0.146	-0.028

Possibly complete quasi-separation: A fraction 0.31 of observations can be perfectly predicted. This might indicate that there is complete quasi-separation. In this case some parameters will not be identified.

Let us remove all the insignificant variables with more than 0.05 p-value and rebuild the model. We get the following result:

Logit Regression Results

Dep. Variable:	default	No. Observations:	2402			
Model:	Logit	Df Residuals:	2391			
Method:	MLE	Df Model:	10			
Date:	Sat, 05 Feb 2022	Pseudo R-squ.:	0.6304			
Time:	04:43:14	Log-Likelihood:	-292.52			
converged:	True	LL-Null:	-791.34			
Covariance Type:	nonrobust	LLR p-value:	6.007e-208			
	coef	std err	z	P> z	[0.025	0.975]
Intercept	-0.3387	0.184	-1.840	0.066	-0.699	0.022
PBITM_perc_Latest_	-0.0399	0.010	-4.169	0.000	-0.059	-0.021
Current_Ratio_Latest_	-0.5319	0.091	-5.830	0.000	-0.711	-0.353
Interest_Cover_Ratio_Latest_	-0.0500	0.041	-1.232	0.218	-0.130	0.030
Debtors_Velocity_Days	-0.0028	0.001	-2.257	0.024	-0.005	-0.000
ROG_Net_Worth_perc	-0.0199	0.009	-2.126	0.034	-0.038	-0.002
Book_Value_Adj_Unit_Curr	-0.1187	0.011	-10.949	0.000	-0.140	-0.097
ROG_Capital_Employed_perc	0.0105	0.007	1.463	0.143	-0.004	0.024
Total_Debt	0.0093	0.002	3.975	0.000	0.005	0.014
Market_Capitalisation	-0.0055	0.002	-2.678	0.007	-0.010	-0.001
CEPS_annualised_Unit_Curr	-0.0816	0.032	-2.579	0.010	-0.144	-0.020

Possibly complete quasi-separation: A fraction 0.30 of observations can be perfectly predicted. This might indicate that there is complete quasi-separation. In this case some parameters will not be identified.

We still have 2 insignificant values **Interest_Cover_Ratio_Latest_** and **ROG_Capital_Employed_perc**. After removing these variables let us build the model again:

Logit Regression Results

Dep. Variable:	default	No. Observations:	2402			
Model:	Logit	Df Residuals:	2393			
Method:	MLE	Df Model:	8			
Date:	Sat, 05 Feb 2022	Pseudo R-squ.:	0.6282			
Time:	04:43:14	Log-Likelihood:	-294.20			
converged:	True	LL-Null:	-791.34			
Covariance Type:	nonrobust	LLR p-value:	2.567e-209			
	coef	std err	z	P> z	[0.025	0.975]
Intercept	-0.3357	0.183	-1.830	0.067	-0.695	0.024
PBITM_perc_Latest_	-0.0444	0.009	-5.160	0.000	-0.061	-0.028
Current_Ratio_Latest_	-0.5391	0.091	-5.910	0.000	-0.718	-0.360
Debtors_Velocity_Days	-0.0030	0.001	-2.374	0.018	-0.005	-0.001
ROG_Net_Worth_perc	-0.0137	0.008	-1.666	0.096	-0.030	0.002
Book_Value_Adj_Unit_Curr	-0.1175	0.011	-11.069	0.000	-0.138	-0.097
Total_Debt	0.0096	0.002	4.122	0.000	0.005	0.014
Market_Capitalisation	-0.0055	0.002	-2.656	0.008	-0.009	-0.001
Market_Capitalisation	-0.0055	0.002	-2.656	0.008	-0.009	-0.001
CEPS_annualised_Unit_Curr	-0.0834	0.031	-2.671	0.008	-0.145	-0.022

Possibly complete quasi-separation: A fraction 0.29 of observations can be perfectly predicted. This might indicate that there is complete quasi-separation. In this case some parameters will not be identified.

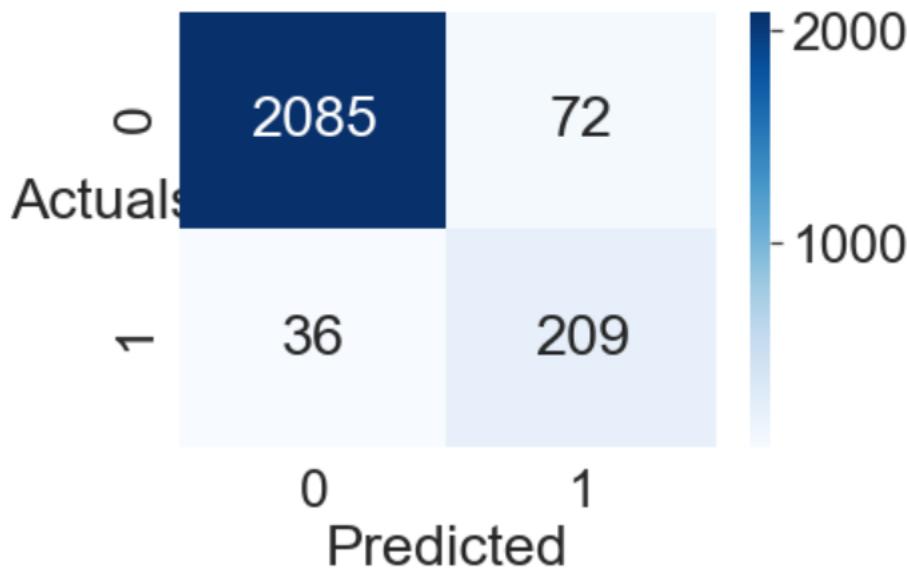
ROG_Net_Worth_perc is a insignificant variable. After removing it we have built the fourth model which can be summarized as follows:

Logit Regression Results

Dep. Variable:	default	No. Observations:	2402			
Model:	Logit	Df Residuals:	2394			
Method:	MLE	Df Model:	7			
Date:	Sat, 05 Feb 2022	Pseudo R-squ.:	0.6264			
Time:	04:43:14	Log-Likelihood:	-295.62			
converged:	True	LL-Null:	-791.34			
Covariance Type:	nonrobust	LLR p-value:	8.458e-210			
	coef	std err	z	P> z	[0.025	0.975]
Intercept	-0.2986	0.181	-1.648	0.099	-0.654	0.056
PBITM_perc_Latest_	-0.0462	0.009	-5.398	0.000	-0.063	-0.029
Current_Ratio_Latest_	-0.5499	0.092	-6.007	0.000	-0.729	-0.370
Debtors_Velocity_Days	-0.0029	0.001	-2.344	0.019	-0.005	-0.000
Book_Value_Adj_Unit_Curr	-0.1193	0.011	-11.167	0.000	-0.140	-0.098
Total_Debt	0.0098	0.002	4.176	0.000	0.005	0.014
Market_Capitalisation	-0.0054	0.002	-2.634	0.008	-0.009	-0.001
CEPS_annualised_Unit_Curr	-0.1085	0.029	-3.786	0.000	-0.165	-0.052

Now let us Validate the Model on Train and Test Dataset and state the performance matrices:

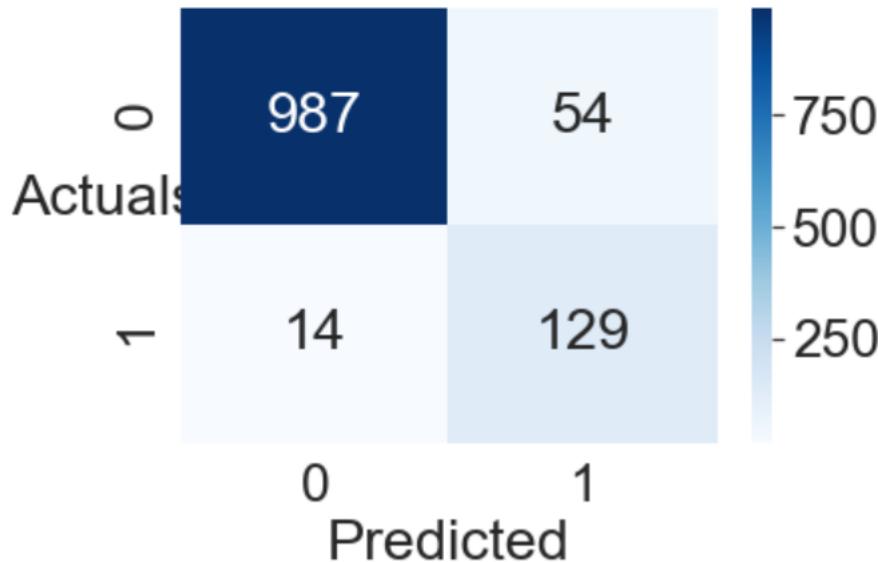
The confusion matrix and classification report for train data can be summarized as follows:



	precision	recall	f1-score	support
0	0.983	0.967	0.975	2157
1	0.744	0.853	0.795	245
accuracy			0.955	2402
macro avg	0.863	0.910	0.885	2402
weighted avg	0.959	0.955	0.956	2402

With recall of 85.3% and precision of 74.4% Precision measures how good our model is when the prediction is positive. Recall measures how good our model is at correctly predicting positive cases. We have got a good model

The confusion matrix and classification report for test data can be summarized as follows:



	precision	recall	f1-score	support
0	0.986	0.948	0.967	1041
1	0.705	0.902	0.791	143
accuracy			0.943	1184
macro avg	0.845	0.925	0.879	1184
weighted avg	0.952	0.943	0.946	1184

90.2% recall and 70.5% precision on test data is an excellent model . Finally, we are able to achieve a descent recall value without overfitting. Considering the opportunities such as outliers, missing values and correlated features this is a fairly good model. It can be improved if we get better quality data where the features explaining the default are not missing to this extent. Of course we can try other techniques which are not sensitive towards missing values and outliers.