

## FRA Project (Milestone-1)

By Akansha Pruthi

PGP – DSBA Jan 2021

## **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 Net worth of the company in the following year (2016) is provided which can be used to drive the labelled field.

## **Data Dictionary:**

Co_Code	Company Code
Co_Name	Company Name
Networth Next Year	Value of a company as on 2016 - Next Year (difference between the value of total assets and total liabilities)
Equity Paid Up	Amount that has been received by the company through the issue of shares to the shareholders
Networth	Value of a company as on 2015 - Current Year
Capital Employed	Total amount of capital used for the acquisition of profits by a company
Total Debt	The sum of money borrowed by the company and is due to be paid
Gross Block	Total value of all of the assets that a company owns
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).
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.
Current Liabilities and Provisions	Short-term financial obligations that are due within one year (includes amount that is set aside cover a future liability)
Total Assets/Liabilities	Ratio of total assets to liabilities of the company
Gross Sales	The grand total of sale transactions within the accounting period
Net Sales	Gross sales minus returns, allowances, and discounts

Other Income	Income realized from non-business activities (e.g., sale of long-term asset)
Value Of Output	Product of physical output of goods and services produced by company and its market price
Cost of Production	Costs incurred by a business from manufacturing a product or providing a service
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)
PBIDT	Profit Before Interest, Depreciation & Taxes
PBDT	Profit Before Depreciation and Tax
PBIT	Profit before interest and taxes
PBT	Profit before tax
PAT	Profit After Tax
Adjusted PAT	Adjusted profit is the best estimate of the true profit
CP	Commercial paper, a short-term debt instrument to meet short-term liabilities.
Revenue earnings in forex	Revenue earned in foreign currency
Revenue expenses in forex	Expenses due to foreign currency transactions
Capital expenses in forex	Long term investment in forex
Book Value (Unit Curr)	Net asset value
Book Value (Adj.) (Unit Curr)	Book value adjusted to reflect asset's true fair market value
Market Capitalisation	Product of the total number of a company's outstanding shares and the current market price of one share
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
Cash Flow from Operating Activities	Use of cash from ongoing regular business activities
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
Cash Flow from Financing Activities	Net flows of cash that are used to fund the company (transactions involving debt, equity, and dividends)
ROG-Net Worth (%)	Rate of Growth - Networth
ROG-Capital Employed (%)	Rate of Growth - Capital Employed
ROG-Gross Block (%)	Rate of Growth - Gross Block
ROG-Gross Sales (%)	Rate of Growth - Gross Sales
ROG-Net Sales (%)	Rate of Growth - Net Sales
ROG-Cost of Production (%)	Rate of Growth - Cost of Production
ROG-Total Assets (%)	Rate of Growth - Total Assets
ROG-PBIDT (%)	Rate of Growth- PBIDT
ROG-PBDT (%)	Rate of Growth- PBDT
ROG-PBIT (%)	Rate of Growth- PBIT
ROG-PBT (%)	Rate of Growth- PBT

ROG-PAT (%)	Rate of Growth- PAT
ROG-CP (%)	Rate of Growth- CP
ROG-Revenue earnings in forex (%)	Rate of Growth - Revenue earnings in forex
ROG-Revenue expenses in forex (%)	Rate of Growth - Revenue expenses in forex
ROG-Market Capitalisation (%)	Rate of Growth - Market Capitalisation
Current Ratio[Latest]	Liquidity ratio, company's ability to pay short-term obligations or those due within one year
Fixed Assets Ratio[Latest]	Solvency ratio, the capacity of a company to discharge its obligations towards long-term lenders indicating
Inventory Ratio[Latest]	Activity ratio, specifies the number of times the stock or inventory has been replaced and sold by the company
Debtors Ratio[Latest]	Measures how quickly cash debtors are paying back to the company
Total Asset Turnover Ratio[Latest]	The value of a company's revenues relative to the value of its assets
Interest Cover Ratio[Latest]	Determines how easily a company can pay interest on its outstanding debt
PBIDTM (%) [Latest]	Profit before Interest Depreciation and Tax Margin
PBITM (%) [Latest]	Profit Before Interest Tax Margin
PBDTM (%) [Latest]	Profit Before Depreciation Tax Margin
CPM (%) [Latest]	Cost per thousand (advertising cost)
APATM (%) [Latest]	After tax profit margin
Debtors Velocity (Days)	Average days required for receiving the payments
Creditors Velocity (Days)	Average number of days company takes to pay suppliers
Inventory Velocity (Days)	Average number of days the company needs to turn its inventory into sales
Value of Output/Total Assets	Ratio of Value of Output (market value) to Total Assets
Value of Output/Gross Block	Ratio of Value of Output (market value) to Gross Block

## **Exploratory Data Analysis of the Dataset**

Top 5 rows of the data:

	Co_Code	Co_Name	Networth Next Year	Equity Paid Up	Networth	Capital Employed	Total Debt	Gross Block	Net Working Capital	Current Assets	...	PBIDTM (%) [Latest]	PBITM (%) [Latest]
0	16974	Hind.Cables	-8021.60	419.36	-7027.48	-1007.24	5936.03	474.30	-1076.34	40.50	...	0.00	0.00
1	21214	Tata Tele. Mah.	-3986.19	1954.93	-2968.08	4458.20	7410.18	9070.86	-1098.88	486.86	...	-10.30	-39.74
2	14852	ABG Shipyards	-3192.58	53.84	506.86	7714.68	6944.54	1281.54	4496.25	9097.64	...	-5279.14	-5516.98
3	2439	GTL	-3054.51	157.30	-623.49	2353.88	2326.05	1033.69	-2612.42	1034.12	...	-3.33	-7.21
4	23505	Bharati Defence	-2967.36	50.30	-1070.83	4675.33	5740.90	1084.20	1836.23	4685.81	...	-295.55	-400.55

Data has 3586 rows and 67 columns.

## Data Information:

```
<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/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 (%)                  3586 non-null   float64
36  ROG-Capital Employed (%)           3586 non-null   float64
37  ROG-Gross Block (%)                3586 non-null   float64
38  ROG-Gross Sales (%)                3586 non-null   float64
39  ROG-Net Sales (%)                  3586 non-null   float64
40  ROG-Cost of Production (%)          3586 non-null   float64
41  ROG-Total Assets (%)               3586 non-null   float64
42  ROG-PBIDT (%)                      3586 non-null   float64
43  ROG-PBDT (%)                      3586 non-null   float64
44  ROG-PBIT (%)                      3586 non-null   float64
45  ROG-PBT (%)                       3586 non-null   float64
46  ROG-PAT (%)                       3586 non-null   float64
47  ROG-CP (%)                        3586 non-null   float64
48  ROG-Revenue earnings in forex (%)  3586 non-null   float64
49  ROG-Revenue expenses in forex (%)  3586 non-null   float64
50  ROG-Market Capitalisation (%)       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 (%) [Latest]                3585 non-null   float64
58  PBITM (%) [Latest]                 3585 non-null   float64
59  PBDTM (%) [Latest]                 3585 non-null   float64
60  CPM (%) [Latest]                   3585 non-null   float64
61  APATM (%) [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/Total Assets        3586 non-null   float64
66  Value of Output/Gross Block         3586 non-null   float64
dtypes: float64(63), int64(3), object(1)
memory usage: 1.8+ MB
```

Data has 103 missing values in Inventory Velocity (Days).

It has 0 duplicated values.

### Descriptive Summary of the data:

	count	mean	std	min	25%	50%	75%	max
Co_Code	3586.0	16065.388734	19776.817379	4.00	3029.2500	6077.500	24269.5000	72493.00
Networth Next Year	3586.0	725.045251	4769.681004	-8021.60	3.9850	19.015	123.8025	111729.10
Equity Paid Up	3586.0	62.966584	778.761744	0.00	3.7500	8.290	19.5175	42263.46
Networth	3586.0	649.746299	4091.988792	-7027.48	3.8925	18.580	117.2975	81657.35
Capital Employed	3586.0	2799.611054	26975.135385	-1824.75	7.6025	39.090	226.6050	714001.25
...	...	...	...	...	...	...	...	...
Debtors Velocity (Days)	3586.0	603.894032	10636.759580	0.00	8.0000	49.000	106.0000	514721.00
Creditors Velocity (Days)	3586.0	2057.854992	54169.479197	0.00	8.0000	39.000	89.0000	2034145.00
Inventory Velocity (Days)	3483.0	79.644559	137.847792	-199.00	0.0000	35.000	96.0000	996.00
Value of Output/Total Assets	3586.0	0.819757	1.201400	-0.33	0.0700	0.480	1.1600	17.63
Value of Output/Gross Block	3586.0	61.884548	976.824352	-61.00	0.2700	1.530	4.9100	43404.00

Cleaning the data and correcting the formatting of the columns to get good results from data.

### Data info after formatting:

#	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_by_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

### Treating Missing Values:

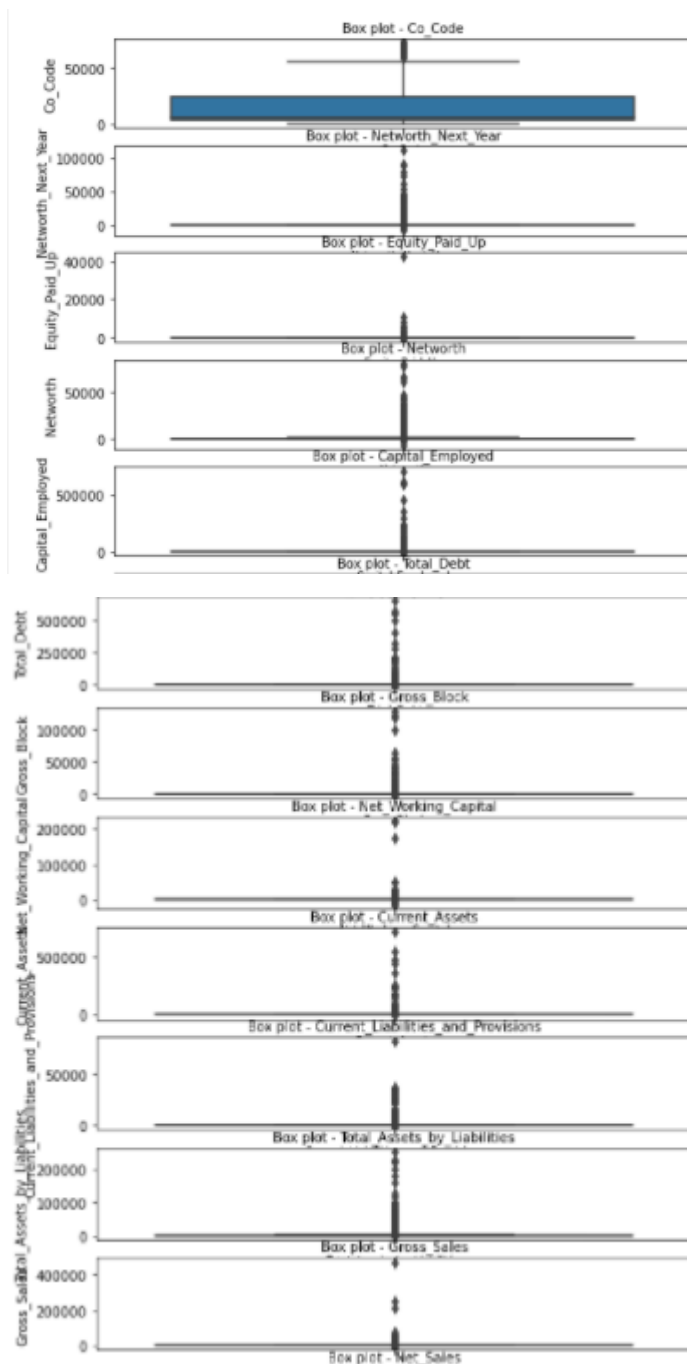
Treating the missing values by imputing using sklearn library using median strategy.

```

Co_Code      0
Networth_Next_Year  0
Equity_Paid_Up    0
Networth        0
Capital_Employed  0
..
Debtors_Velocity_Days  0
Creditors_Velocity_Days  0
Inventory_Velocity_Days  0
Value_of_Output_by_Total_Assets  0
Value_of_Output_by_Gross_Block  0
Length: 66, dtype: int64

```

## Checking for Outliers:



## Treating the outliers:

463.75 3029.25 24269.5 67803.75 <built-in function min> <built-in function max>  
-17.445 3.985 123.80250000000001 1978.82250000000002 <built-in function min> <built-in function max>  
0.5 3.75 19.517500000000002 131.24 <built-in function min> <built-in function max>  
-11.697499999999998 3.8925 117.2975 1829.0825 <built-in function min> <built-in function max>  
0.4225 7.602499999999999 226.60500000000002 3634.915 <built-in function min> <built-in function max>  
0.0 0.03 72.35000000000001 1572.61 <built-in function min> <built-in function max>

0.0 0.57 131.895 1409.325 <built-in function min> <built-in function max>  
-11.944999999999999 0.9424999999999999 61.175 827.735 <built-in function min>  
<built-in function max>  
0.14 4.0 135.2775 2014.74 <built-in function min> <built-in function max>  
0.02 0.7324999999999999 65.65 1021.03 <built-in function min> <built-in function m  
ax>  
1.22 10.555 310.54 4568.73000000000005 <built-in function min> <built-in function m  
ax>  
0.0 1.4425000000000001 242.25 2845.3725 <built-in function min> <built-in function  
max>  
0.0 1.44 234.44000000000003 2780.1400000000003 <built-in function min> <built-in  
function max>  
0.0 0.02 3.635 78.80250000000001 <built-in function min> <built-in function max>  
0.0 1.4124999999999999 235.8375 2803.74 <built-in function min> <built-in function  
max>  
0.01 0.94 189.54999999999998 1981.68 <built-in function min> <built-in function ma  
x>  
0.0 0.0 3.8825 74.97999999999999 <built-in function min> <built-in function max>  
-3.995 0.04 23.525 455.5025 <built-in function min> <built-in function max>  
-17.872500000000002 0.0 12.945 290.0175 <built-in function min> <built-in function  
max>  
-10.834999999999999 0.0 16.6675 365.68 <built-in function min> <built-in function m  
ax>  
-32.7575 -0.06 7.422499999999999 219.055 <built-in function min> <built-in function  
max>  
-30.575 -0.06 5.54 171.6925 <built-in function min> <built-in function max>  
-29.557499999999997 -0.09 5.342499999999999 153.3 <built-in function min> <built  
-in function max>  
-15.584999999999999 0.0 10.91 231.39 <built-in function min> <built-in function max  
>  
0.0 0.0 7.2 360.725 <built-in function min> <built-in function max>  
0.0 0.0 6.9875 304.665 <built-in function min> <built-in function max>  
0.0 0.0 0.0 11.322500000000002 <built-in function min> <built-in function max>  
-13.567499999999999 7.962499999999995 71.66749999999999 349.655 <built-in f  
unction min> <built-in function max>  
-13.567499999999999 7.064999999999995 59.96 298.08750000000003 <built-in fu  
nction min> <built-in function max>  
0.0 0.0 111.45750000000001 3984.7225000000003 <built-in function min> <built-in f  
unction max>  
-7.475 0.0 8.772499999999999 59.3 <built-in function min> <built-in function max>  
-21.95 -0.3075 12.6475 229.84 <built-in function min> <built-in function max>  
-143.6425 -5.1175 0.12 13.997500000000002 <built-in function min> <built-in functio  
n max>  
-130.235 -5.847499999999999 0.4575 52.332499999999996 <built-in function min>  
<built-in function max>  
-55.6275 -1.487499999999998 11.3625 90.20750000000001 <built-in function min>  
<built-in function max>  
-33.382499999999999 -3.835 12.587499999999999 73.77 <built-in function min> <bui  
lt-in function max>  
-32.5875 0.0 6.72 51.365 <built-in function min> <built-in function max>



-73.18499999999999 -8.0775 21.525000000000002 227.41 <built-in function min> <built-in function max>  
-73.17249999999999 -8.1175 21.567500000000003 227.41 <built-in function min> <built-in function max>  
-66.4 -7.2425 23.122500000000002 225.93 <built-in function min> <built-in function max>  
-26.4375 -3.9725 12.5 64.1575 <built-in function min> <built-in function max>  
-200.0 -23.3625 47.875 347.0375 <built-in function min> <built-in function max>  
-313.59749999999997 -30.5975 52.915 350.0 <built-in function min> <built-in function max>  
-250.0 -31.3525 50.1425 371.665 <built-in function min> <built-in function max>  
-513.9475 -41.235 61.957499999999996 372.3775 <built-in function min> <built-in function max>  
-501.2525 -43.7325 65.3475 422.02500000000003 <built-in function min> <built-in function max>  
-321.7375 -29.505 52.9075 347.12749999999994 <built-in function min> <built-in function max>  
-60.907500000000006 0.0 0.0 56.96 <built-in function min> <built-in function max>  
-76.4 0.0 0.0 122.17999999999999 <built-in function min> <built-in function max>  
-41.0825 0.0 47.515 237.875 <built-in function min> <built-in function max>  
0.1 0.88 2.77 39.3325 <built-in function min> <built-in function max>  
0.0 0.27 4.74 54.667500000000004 <built-in function min> <built-in function max>  
0.0 0.0 8.9375 58.0325 <built-in function min> <built-in function max>  
0.0 0.42 8.5175 45.832499999999996 <built-in function min> <built-in function max>  
0.0 0.07 1.55 3.98 <built-in function min> <built-in function max>  
-6.7075 0.0 3.71 59.75 <built-in function min> <built-in function max>  
-58.682500000000005 0.0 18.9875 72.23249999999999 <built-in function min> <built-in function max>  
-82.2225 0.0 14.285 66.4725 <built-in function min> <built-in function max>  
-90.795000000000002 0.0 14.1 57.9775 <built-in function min> <built-in function max>  
-87.21249999999999 0.0 11.387500000000001 48.2475 <built-in function min> <built-in function max>  
-117.12 0.0 7.407500000000001 40.692499999999995 <built-in function min> <built-in function max>  
0.0 8.0 106.0 715.25 <built-in function min> <built-in function max>  
0.0 8.0 89.0 615.25 <built-in function min> <built-in function max>  
0.0 0.0 93.0 324.75 <built-in function min> <built-in function max>  
0.0 0.07 1.16 2.5175 <built-in function min> <built-in function max>  
0.0 0.27 4.91 57.9625 <built-in function min> <built-in function max>

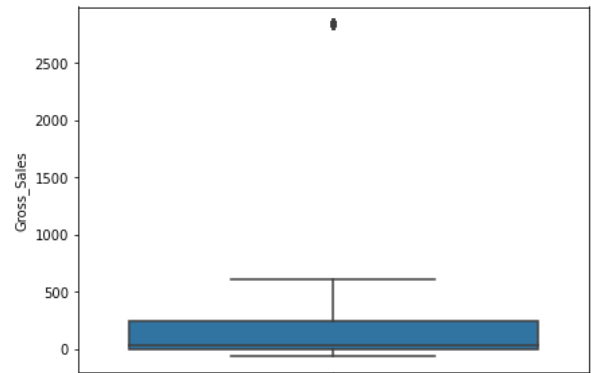
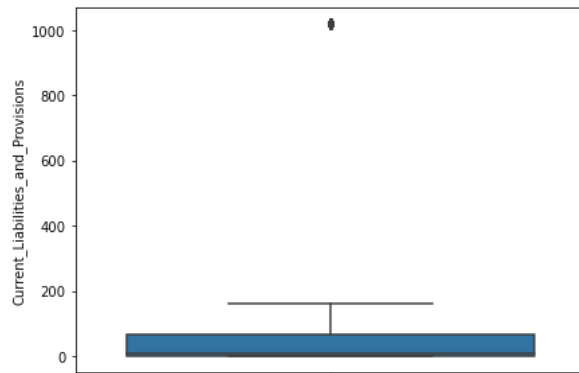
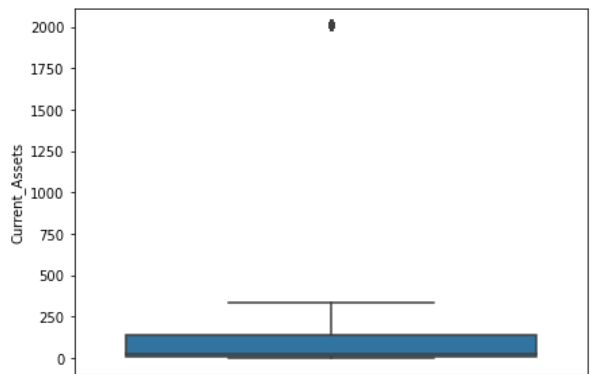
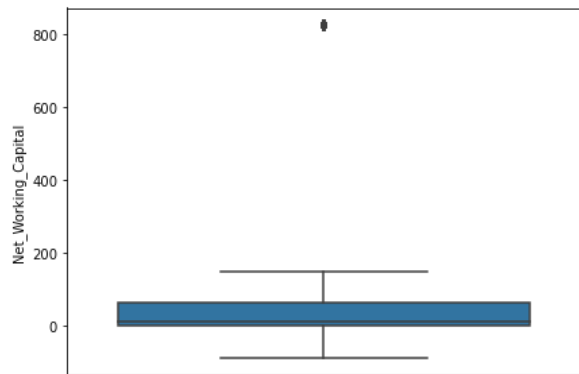
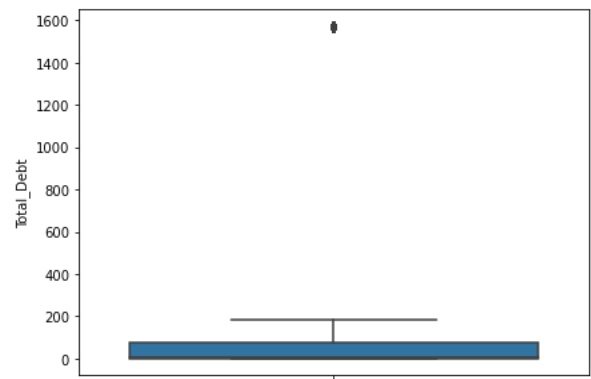
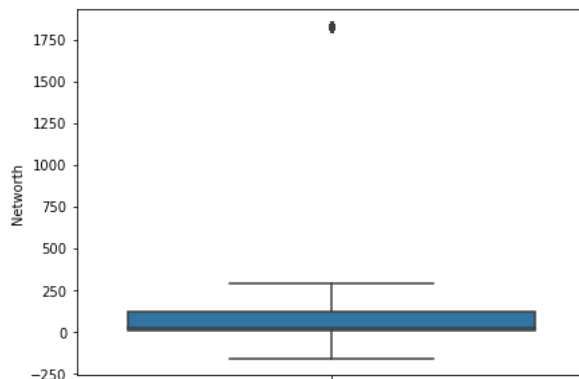
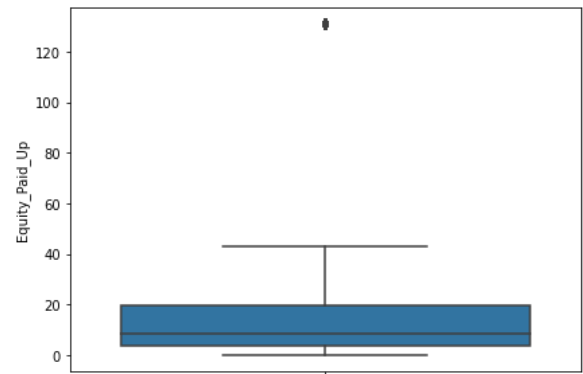
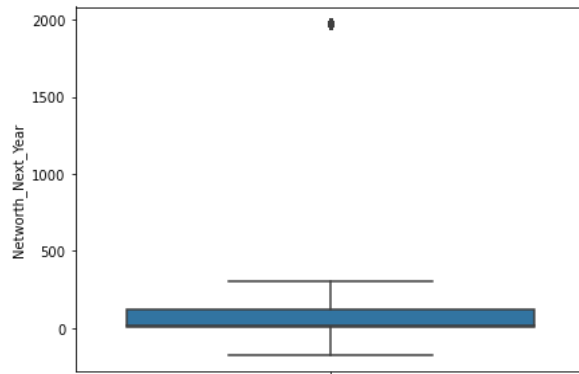
	default	Networth_Next_Year
0	1	-17.445
1	1	-17.445
2	1	-17.445
3	1	-17.445
4	1	-17.445

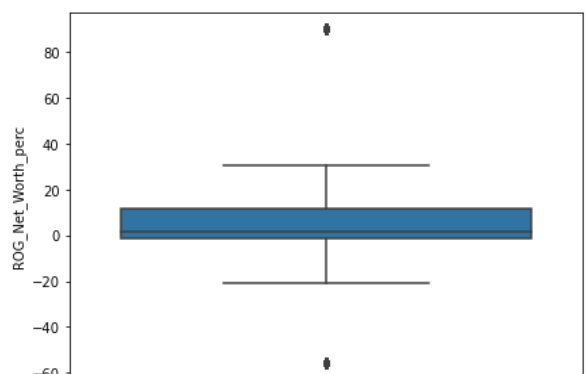
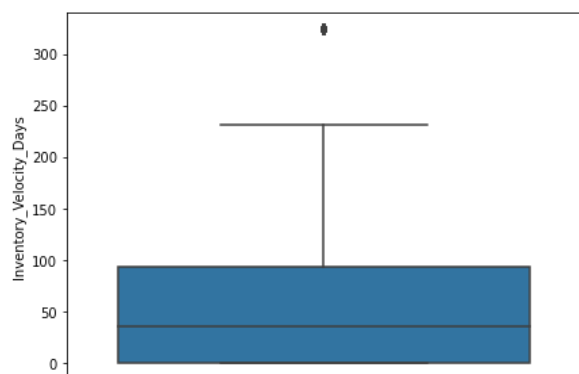
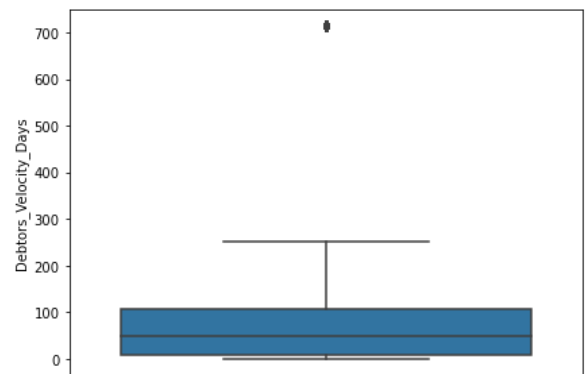
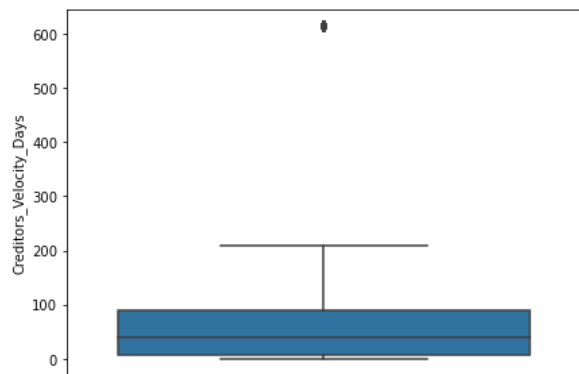
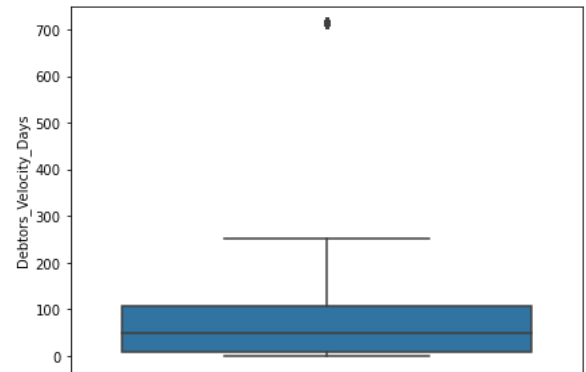
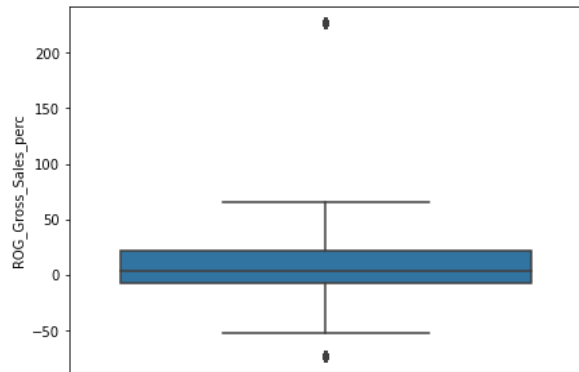
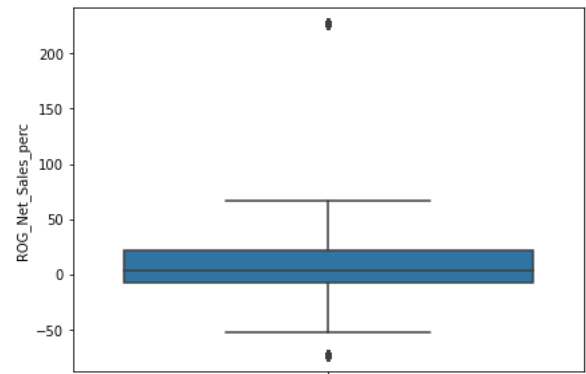
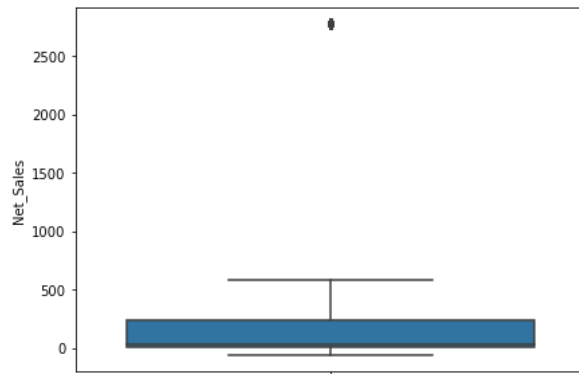
proportion of default: 0.10819854991634133

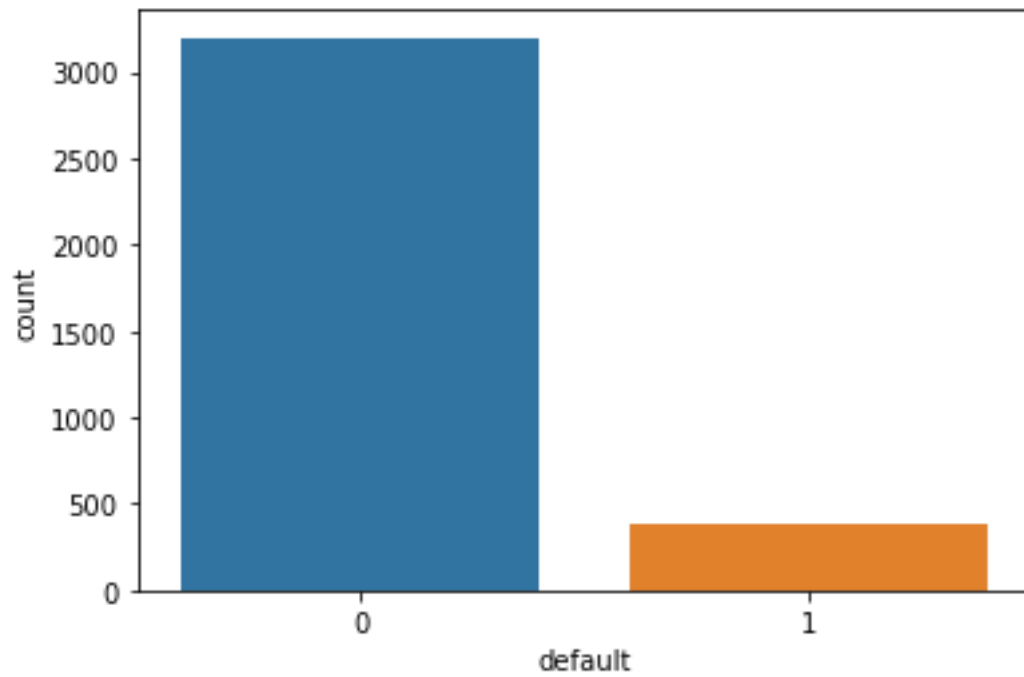
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523</
--	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------

## Univariate and Bi-Variate Analysis [1](#)

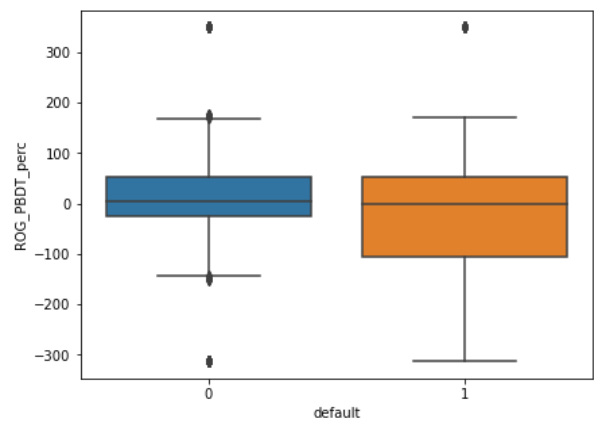
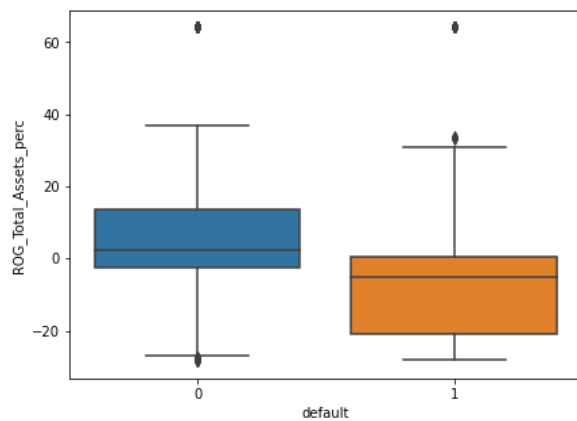
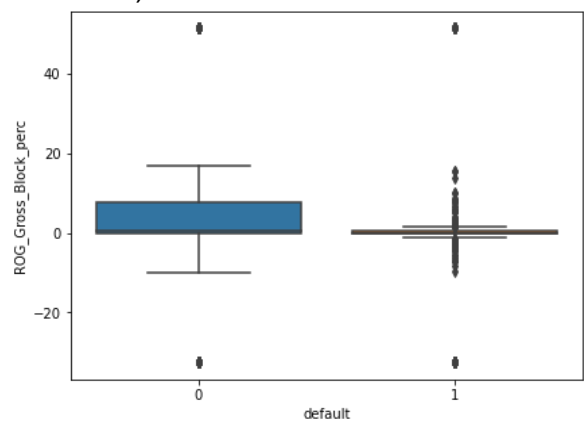
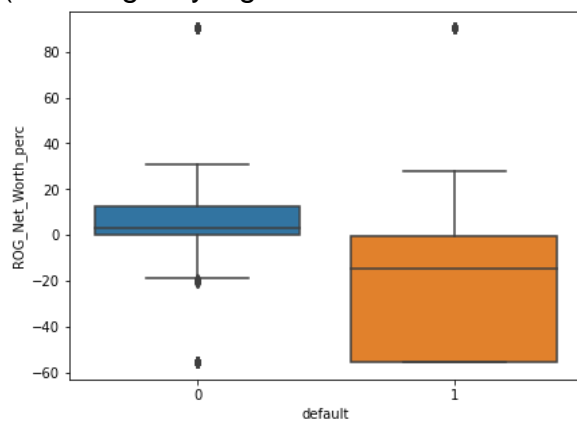
(only those variables which were significant in the model building)

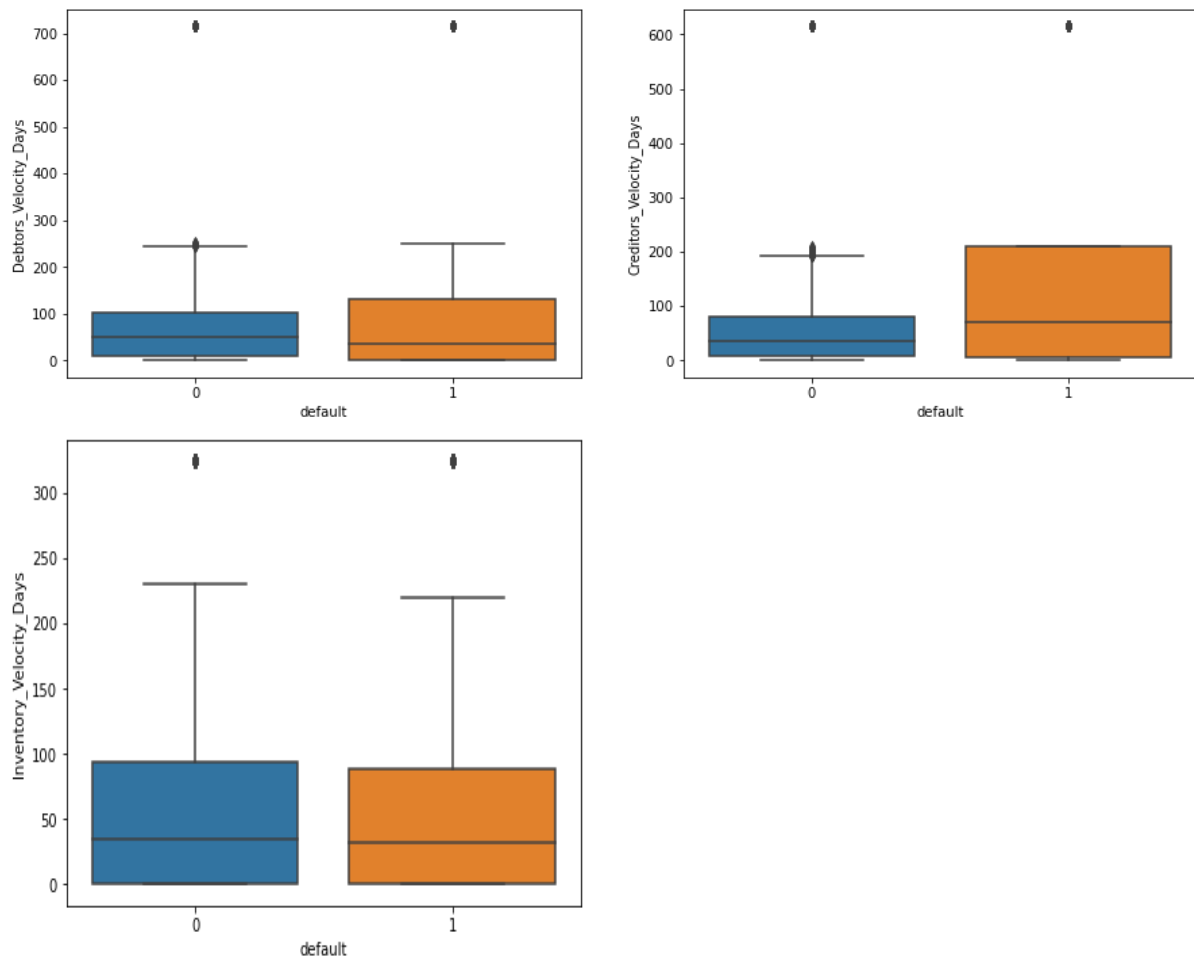






### Bi- Variate Analysis (including only significant variables relevant to model)





### Splitting the data into Train and Test Data:

Splitting the data into 67:33 ratio and keeping the random state 42.

Shape of the data:

(2402, 66)

(1184, 66)

### Building the logistic regression model:

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

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

## Summary: Model 1

Logit Regression Results

Dep. Variable:	default	No. Observations:	2402
Model:	Logit	Df Residuals:	2337
Method:	MLE	Df Model:	64
Date:	Wed, 08 Dec 2021	Pseudo R-squ.:	0.6385
Time:	01:12:57	Log-Likelihood:	-288.11
converged:	True	LL-Null:	-791.34
Covariance Type:	nonrobust	LLR p-value:	3.166e-170

	coef	std err	z	P> z	[0.025	0.975]
Intercept	-0.5073	0.226	-2.247	0.025	-0.950	-0.065
Equity_Paid_Up	-0.0091	0.005	-1.969	0.049	-0.018	-4.19e-05
Networth	-0.0007	0.001	-0.974	0.330	-0.002	0.001
Capital_Employed	-7.708e-05	0.001	-0.143	0.887	-0.001	0.001
Total_Debt	0.0016	0.001	2.721	0.007	0.000	0.003
Gross_Block	-0.0007	0.001	-0.950	0.342	-0.002	0.001
Net_Working_Capital	-0.0004	0.001	-0.284	0.777	-0.003	0.002
Current_Assets	0.0005	0.001	0.685	0.493	-0.001	0.002
Current_Liabilities_and_Provisions	0.0009	0.001	0.869	0.385	-0.001	0.003
Total_Assets_by_Liabilities	0.0002	0.000	0.478	0.632	-0.001	0.001
Gross_Sales	0.0032	0.003	1.096	0.273	-0.003	0.009
Net_Sales	0.0180	0.005	3.920	0.000	0.009	0.027
Other_Income	0.0059	0.008	0.734	0.463	-0.010	0.022
Value_Of_Output	-0.0211	0.004	-5.704	0.000	-0.028	-0.014
Cost_of_Production	-0.0008	0.001	-0.946	0.344	-0.002	0.001

## Validating the model on train set

### Confusion matrix:

```
[[2137  20]
 [  74 171]]
```

### Classification report:

```
              precision    recall  f1-score   support

     0               0.97       0.99       0.98       2157
     1               0.98       0.78       0.78       245

 accuracy              0.96       0.96       0.96       2402
 macro avg              0.93       0.84       0.88       2402
 weighted avg           0.96       0.96       0.96       2402
```

## Validating the model on test set

### Confusion matrix:

```
[[1021  20]
 [  41 102]]
```

## Classification report:

	precision	recall	f1-score	support
0	0.96	0.98	0.97	1041
1	0.84	0.71	0.77	143
accuracy			0.95	1184
macro avg	0.90	0.85	0.87	1184
weighted avg	0.95	0.95	0.95	1184

## Summary : Model 2

### Logit Regression Results

Dep. Variable:	default	No. Observations:	3774
Model:	Logit	Df Residuals:	3709
Method:	MLE	Df Model:	64
Date:	Wed, 08 Dec 2021	Pseudo R-squ.:	0.7440
Time:	01:16:56	Log-Likelihood:	-659.82
converged:	True	LL-Null:	-2577.2
Covariance Type:	nonrobust	LLR p-value:	0.000

	coef	std err	z	P> z	[0.025	0.975]
Intercept	1.3600	0.159	8.569	0.000	1.049	1.671
Equity_Paid_Up	-0.0141	0.004	-3.824	0.000	-0.021	-0.007
Networth	-0.0024	0.001	-4.226	0.000	-0.004	-0.001
Capital_Employed	0.0003	0.000	0.597	0.550	-0.001	0.001
Total_Debt	0.0022	0.001	4.129	0.000	0.001	0.003
Gross_Block	-0.0017	0.001	-2.649	0.008	-0.003	-0.000
Net_Working_Capital	0.0035	0.001	3.368	0.001	0.001	0.006
Current_Assets	-0.0010	0.001	-1.693	0.090	-0.002	0.000
Current_Liabilities_and_Provisions	0.0048	0.001	5.291	0.000	0.003	0.007
Total_Assets_by_Liabilities	0.0004	0.000	0.987	0.323	-0.000	0.001
Gross_Sales	0.0056	0.002	2.681	0.007	0.001	0.010
Net_Sales	0.0183	0.003	5.978	0.000	0.012	0.024
Other_Income	0.0095	0.006	1.530	0.126	-0.003	0.022
Value_Of_Output	-0.0251	0.002	-10.524	0.000	-0.030	-0.020
Cost_of_Production	-9.527e-06	0.001	-0.015	0.988	-0.001	0.001

## Validating on resampled train set

### Confusion matrix:

```
[[2012 145]
 [ 86 1531]]
```



### Classification report:

	precision	recall	f1-score	support
0	0.96	0.93	0.95	2157
1	0.91	0.95	0.93	1617
accuracy			0.94	3774
macro avg	0.94	0.94	0.94	3774
weighted avg	0.94	0.94	0.94	3774

### Validating on test set

#### Confusion matrix:

```
[[ 957   84]
 [  14 129]]
```

### Classification report:

	precision	recall	f1-score	support
0	0.99	0.92	0.95	1041
1	0.61	0.90	0.72	143
accuracy			0.92	1184
macro avg	0.80	0.91	0.84	1184
weighted avg	0.94	0.92	0.92	1184

This shows that the model is highly accurate on predicting whether the net-worth is positive or negative for the company and we can confidently deploy the model for production.

The End.

