APPLICATION SCORECARD DEVELOPMENT

Overview

- Introduction
- Data Description
- Data Exploration
- Univariate Analysis
- Bivariate Analysis
- Pre-processing data
- Modelling Process
- Conclusion & Discussion
- References
- Appendix

Introduction

This report presents a comprehensive analysis of the Scorecard development process. The objective of this project is to create a scorecard model that can be utilized to assess the initial risk associated with granting the loan to a new customer based on their personal details, past credit history, and behavioral data. The score generated from the model will be used to evaluate the probability of default and assess the risk for new credit contracts.

Data Description

Exploring the dataset, we have 7 columns and 1000 rows in which there are 6 independent variables and 1 dependent/ target variable which is Default variable. The output variable is in binary format.

IDcli – ID number

Age: Age of borrower [years]

Years_current_job : Duration of working in the current job [years]

Residential_status : Residential status of client [Renter, Owner, Other]

Monthly_Net_Income: Monthly net income of each client.

DebtRatio: ratio of the loan payment with income.

Default: Classification of defaulted client using a 12-month window of observation period in a binary variable: 1 - Default (Bad client), 0 - Not Default (Good client).

Data Exploration

In this dataset, we have 5 numerical data and 2 categorical data. And we have a total of 1000 rows and 7 columns. (Here, I am considering any column contains more than 3 subcategories are numerical data)

IDCli	Default	Age	Years_current_job	Residential_status	Monthly_	DebtRatio
8191	1	35	0	Owner	1171.98	0.75834
8217	1	26	0	Renter	2107.32	0.25641
8372	1	36	5	Renter	1654.44	0.29601
8466	1	35	11	Owner	2303.16	0.31977

Figure 1

- **Data sampling**

Data sampling is an essential component of scorecard development, in this process, we apply stratified sampling technique which divide the population into relevant strata based on key variables such as Default. Basically, this technique improves the accuracy of the scorecard model (Siddiqi, 2016).

- **Missing Values**

Now, we are going to check for missing values in the dataset.

1. Numerical variables:

The MEANS Procedure

Variable	Label	N Miss
Age	Age	0
Years_current_job	Years_current_job	195
Monthly_Net_Income	Monthly_Net_Income	0
DebtRatio	DebtRatio	0

Figure 2

From fig (2), we found that only Years_current_job contains the missing values. Here, we are not going to apply imputation technique but, instead we are going to use this as a separate attribute in the scorecard development process and all missing values will be replaced with value 9999 for the analysis as suggested by Siddiqi (2016).

2. Categorical variables:

From the analysis, it is found that there are no missing values present in either Residential status or Default variables.

- **Outlier Detection:**

Now, we will observe the outliers in each variable.

From the analysis, we found that only Monthly_Net_Income contains extreme observations.

From fig (3), we can observe that observation 363 contains extreme values which could be an outlier. According to Siddiqi (2016), for scorecard development we are going to carry out a binning process which will reduce the impact of outliers, but we still need to do further analysis.

Monthly_Net_Income

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
878.22	883	5593.68	707
878.22	802	5913.96	556
878.22	757	7485.78	706
1099.56	496	8263.02	280
1099.56	311	18980.16	363

Figure 3

- **Unique values**

Now, we will observe unique values in the categorical variables.

So, Residentail status contains 3 subcategories – Renter, owner and Other

Default – 0 & 1

Univariate Analysis

Numerical Data:

1. Age

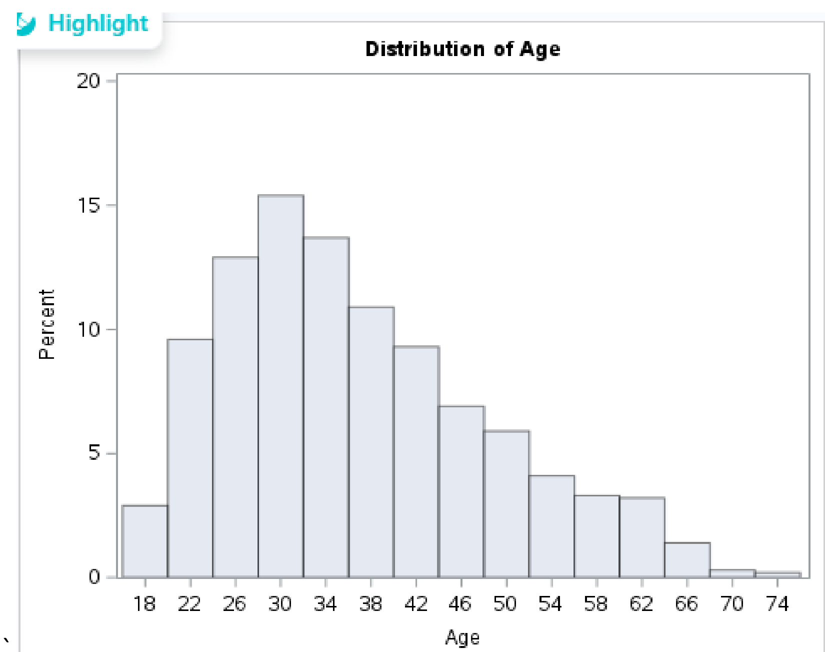


Figure 4

The distribution of age variable is normally distributed with the mean age of 36, and the range lies between 18 to 74 years.

2. Years_current_job

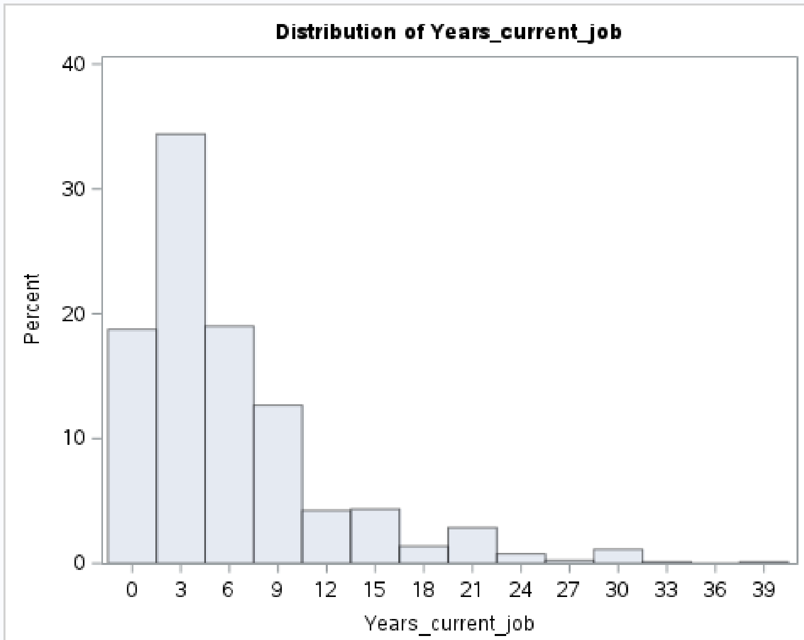


Figure 5

The distribution of years of current job is right skewed with the mean of 5.9 years, majority of the data is concentrated between 0 to 9 years.

3. Monthly_Net_Income

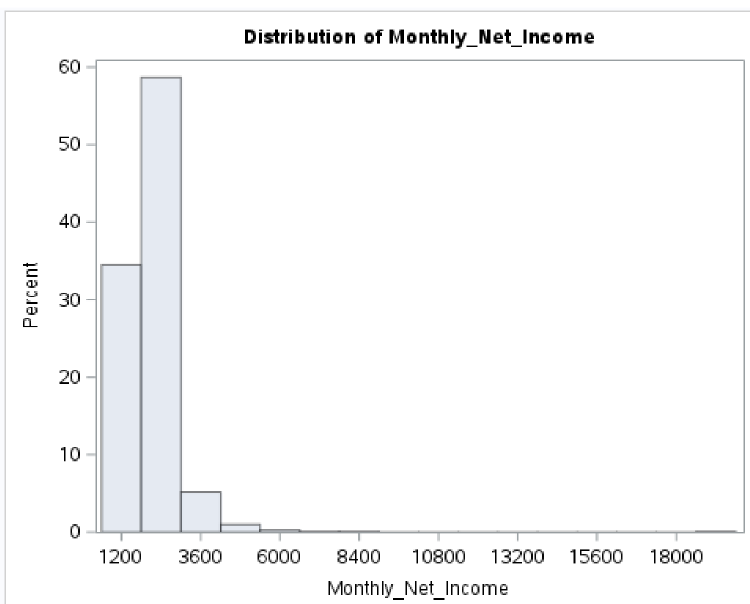


Figure 6

The distribution of net monthly income is right skewed with the mean of 2145, majority of the data is concentrated between 1718 to 2395. And outlier is clearly detected from figure 6.

4. Debt Ratio

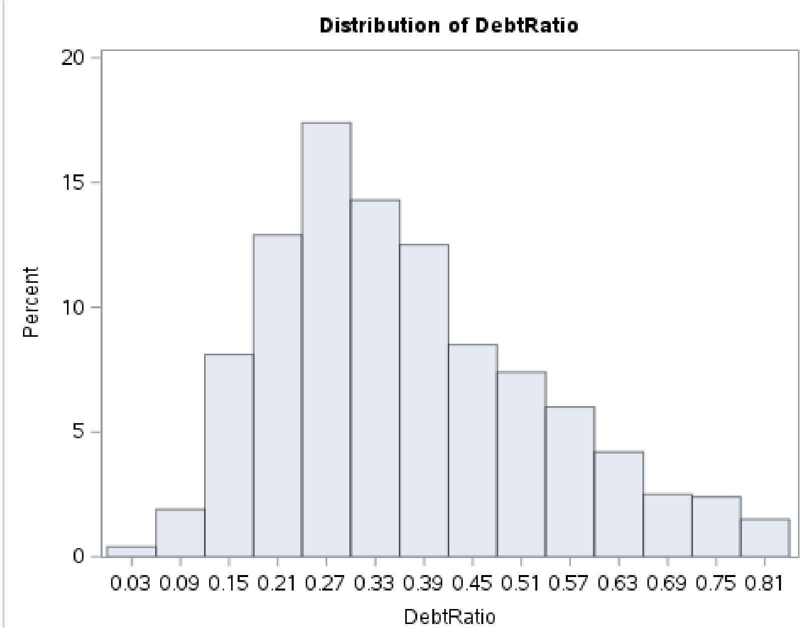


Figure 7

The distribution of Debt to ratio is normally distributed.

Categorical Data:

1. Residentail Status

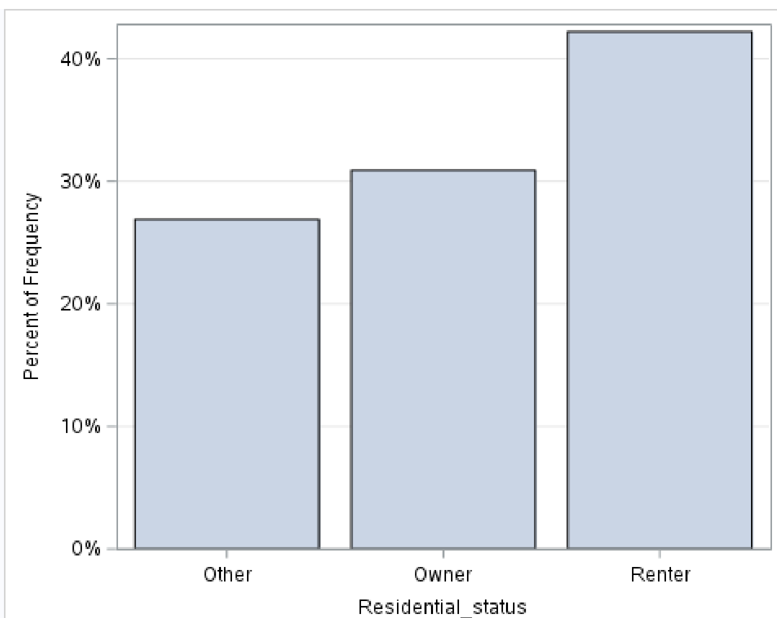


Figure 8

From fig 8, we can observe that residential status contains three categories which is Renter which is dominating with the 40%. Followed by Owner and Other.

2. Default

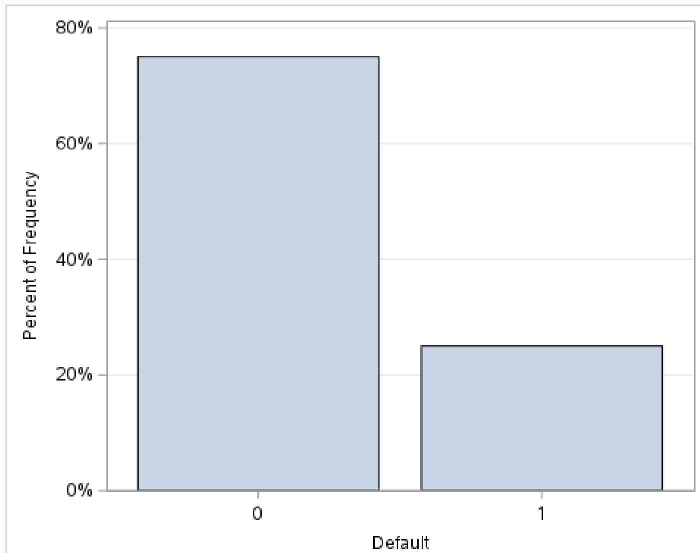


Figure 9

The ratio between the default and non-default is well distributed, whereas client with default is around 23% compared to non-default at 77%.

Bivariate Analysis:

- **Scatter plot**

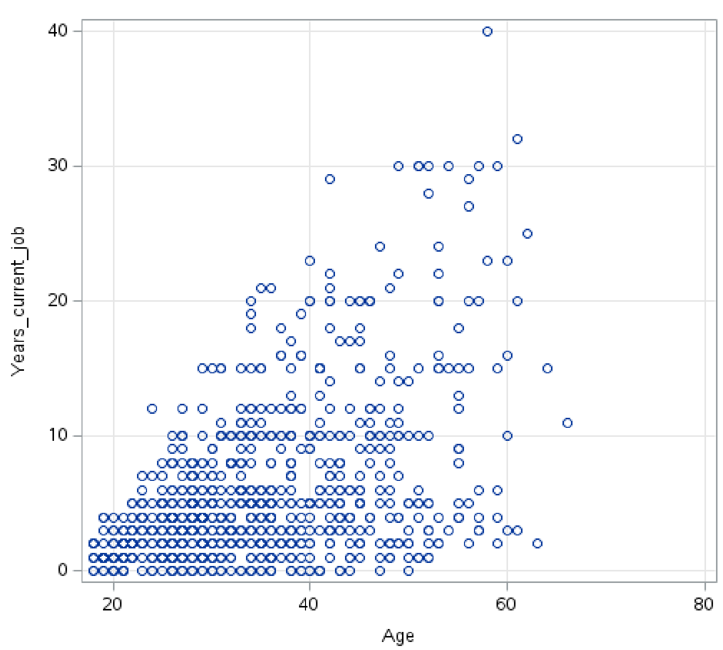


Figure 10

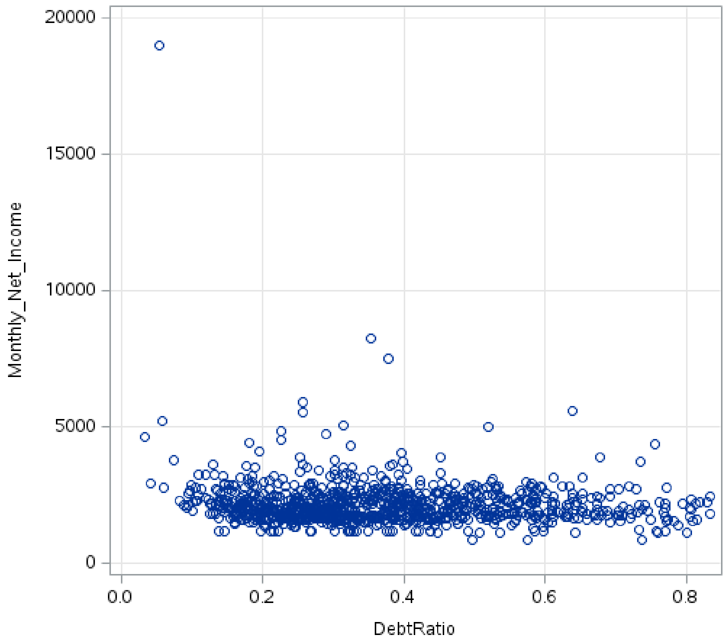


Figure 11

- **Correlation analysis**

Correlation Analysis

The CORR Procedure

4 Variables:	WoE_Residential_status WoE_DebtRatio WoE_Years_current_job WoE_Monthly_Net_Income

Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
WoE_Residential_status	1000	0.13211	0.72450	132.10705	-0.43025	1.20754
WoE_DebtRatio	1000	0.00882	0.12370	8.82141	-0.10958	0.31876
WoE_Years_current_job	1000	-0.06185	0.14060	-61.85180	-0.18174	0.21044
WoE_Monthly_Net_Income	1000	0.01113	0.13852	11.13471	-0.09988	0.30103

Pearson Correlation Coefficients, N = 1000 Prob > |r| under H0: Rho=0				
	WoE_Residential_status	WoE_DebtRatio	WoE_Years_current_job	WoE_Monthly_Net_Income
WoE_Residential_status	1.00000	0.27113 <.0001	0.19204 <.0001	0.40952 <.0001
WoE_DebtRatio	0.27113 <.0001	1.00000	0.04967 0.1165	-0.01822 0.5650
WoE_Years_current_job	0.19204 <.0001	0.04967 0.1165	1.00000	0.14646 <.0001
WoE_Monthly_Net_Income	0.40952 <.0001	-0.01822 0.5650	0.14646 <.0001	1.00000

Figure 12

I. There is a weak positive relationship between WoE_Residential_status and WoE_DebtRatio and p-value suggest that correlation is statistically significant.
II. There is a a weak positive relationship between WoE_Residential_status and WoE_Years_current_job and p-value suggest that correlation is statistically significant.
III. There is a negligible relationship between WoE_DebtRatio and WoE_Years_current_job.
IV. There is a negligible relationship between WoE_DebtRatio and WoE_Monthly_Net_Income.

The strongest relation between the two independent variables is 0.4905 which is still moderate positive relationship which suggests that there is no problem for multicollinearity between these variables (Baesens, Roesch and Scheule, 2016).

Next, we are going to group similar indicators together to reduce the complexity of the data and make it easier to understand and interpret the result.

Pre-processing data:

- **Grouping of variables**

In this process we will basically group the large amount of data into meaningful clusters. So that grouping helps us to focus on smaller numbers of groups with better understanding (Siddiqi, 2016).

1. Years_current_job

Now, using the bivariate analysis code we are going to create the table which consist of different clusters, Total number and percentage of observation, Total number and percentage of goods, bads, weight of evidence and Information value.

The classes are assigned based on the weight of evidence and the percentage of observations.

CLUSTER	Total	Total_Ba	Total_Go	N_Class	%class	Min	Max	PCT_B	PCT_G	Bads	Goods	WOE	IVi	class
1	1000	250	750	151	15.10%	0	1	21.2	13.07	53	98	-0.48394	3.936019	1
3	1000	250	750	194	19.40%	2	3	25.2	17.47	63	131	-0.36655	2.834651	1
5	1000	250	750	168	16.80%	4	5	15.2	17.33	38	130	0.131336	0.280183	2
7	1000	250	750	68	6.80%	6	7	8.8	6.13	22	46	-0.36101	0.962702	2
9	1000	250	750	34	3.40%	8	8	3.2	3.47	8	26	0.080043	0.021345	2
10	1000	250	750	68	6.80%	9	10	4.8	7.47	12	56	0.441833	1.178221	2
11	1000	250	750	69	6.90%	11	16	2	8.53	5	64	1.450833	9.478775	2
12	1000	250	750	40	4.00%	17	25	2.4	4.53	6	34	0.635989	1.356776	4
2	1000	250	750	11	1.10%	27	30	0.8	1.2	2	9	0.405465	0.162186	4
4	1000	250	750	1	0.10%	32	32	0	0.13	0	1			4
6	1000	250	750	1	0.10%	40	40	0	0.13	0	1			4
8	1000	250	750	195	19.50%	9999	9999	16.4	20.53	41	154	0.224768	0.929042	5

Figure 13

Next, we will group the clusters based on the weight of evidence and also ensure that each cluster must contains atleast 5% of observations (Siddiqi, 2016).

Row Labels	Sum of Goods	Sum of Bads	%G	%B	WoE	IVi
1	229	116	0.305333	0.464	-0.18174	0.028837
2	322	85	0.429333	0.34	0.101316	0.009051
4	45	8	0.06	0.032	0.273001	0.007644
5	154	41	0.205333	0.164	0.097616	0.004035
Grand Total	750	250				0.049566

Figure 14

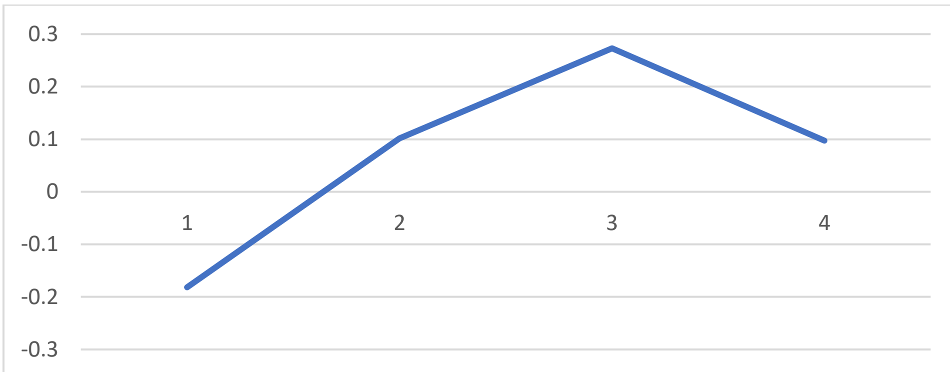


Figure 15

Similar Technique is applied to all variables and the calculated predictive power of each feature are shown in the table 2.

Note: for each step of all variables please refer appendix

For any feature to be considered for modelling it needs to between weak to strong predictor which is basically from the range of 0.02 to 0.5 (Siddiqi, 2016).

Model 1:

Table 2

Feature	IV
Age	0.049563898
DebtRatio	0.028596976
Years_current_job	0.049566359
Monthly_Net_Income	0.040055363
Residential_status	0.086772577

From table 2, It is clear that all the variables are eligible for modelling process.

Model 2:

Table 3

Feature	IV	
Age	0.008734901	(Not satisfied)

DebtRatio	0.028596976
Years_current_job	0.049566359
Monthly_Net_Income	0.040055363
Residential_status	0.086772577

From table 3, Age variables is not eligible for modelling process because the IV value is less than 0.02.

Modelling Process:

- **Multiple Logistic Regression model**

Model 1:

As we know, the variable Age, Years_current_job , Residential_status , Monthly_Net_Income , DebtRatio all satisfy the condition of having the p value greater than 0.02, which suggest that there is strong relationship. Hence, we are going to use all these variables for the modelling process (Siddiqi, 2016).

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1827	0.0912	168.1593	<.0001
WoE_Residential_stat	1	-2.2898	0.3281	48.7155	<.0001
WoE_Years_current_jo	1	-1.3228	0.5767	5.2614	0.0218
WoE_Age	1	-2.4362	0.5251	21.5250	<.0001

Figure 16

From fig (16), we can observe that the weight of evidence of Debt Ratio, Age, and years of current job is considered significant for scorecard development.

Model 2:

the variable Age does not satisfy the condition of having the p value greater than 0.02. Hence, we are not going to input age variables in the algorithm.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.1582	0.0895	167.5284	<.0001
WoE_Residential_stat	1	-2.2730	0.3273	48.2150	<.0001
WoE_Years_current_jo	1	-1.0158	0.5651	3.2318	0.0722

Figure 17

From fig (18), we can observe that the weight of evidence of Residential status, and years of current job is considered significant for scorecard development. Here, we are taking the significance level of 10%.

- **Performance of the model**

Evaluating using Gini index: (AUC*2-1)

Model 1:

Gini index - 0.6990*2-1 = 0.398 (which is considered decent because it is very close to 0.4) (Siddiqi, 2016).

- Percent Concordant: 66.4%
- Somers' D: 0.398
- Percent Discordant: 26.6%
- Gamma: 0.428
- Percent Tied: 7.1%
- Tau-a: 0.152
- Pairs: 170610
- c: 0.699

Model 2:

Gini index - 0.6542*2-1 = 0.3084 (very weak model because it is way lower than 0.4).

- Percent Concordant: 56.2%
- Somers' D: 0.308
- Percent Discordant: 25.4%
- Gamma: 0.378
- Percent Tied: 18.4%
- Tau-a: 0.117
- Pairs: 170610
- c: 0.654

Finally, Model 1 outperforms Model 2 across a wide range of association metrics. Model 1 indicates that, higher percentages of concordant pairs, higher Somers' D and Gamma coefficients, lower percentages of discordant and tied pairs, higher Tau-a coefficient, and a higher c statistic it shows greater connections and better discrimination between expected probability and observed responses.

- **Building a Scorecard Model**

According to (siddiqi 2016), The score is calculated using the intercept and Estimate value.

Model 1:

Table 4

Characteristic	Attribute	Score points
Age	Age <25	52
Age	Age >=25 AND Age <26	43
Age	Age >=26 AND Age <32	24
Age	Age >=32	45
Age	Neutral	41
Residential_status	Renter	28
Residential_status	Other	34
Residential_status	Owner	75
Residential_status	Neutral	44
Years_current_job	Years_current_job <4	33
Years_current_job	Years_current_job >=4 AND Years_current_job <17	44
Years_current_job	Years_current_job >=17	51
Years_current_job	missing(Years_current_job) OR Years_current_job =.	44
Years_current_job	Neutral	41

Table 5

Client A			Client B		
Age	40	45	Age	26	24
Residential status	Owner	75	Residential status	Renter	28
Years of current job	18	51	Years of current job	2	33
Total		171	Total		85
Decision		Granted	Decision		Not Granted

In table 5, client A has higer score than the weighted average /Neutral score (Table 4) for each characteristics which makes them eligible for loan.

For Client B, we can observe the score achieved is lower than weighted average /Neutral score for each characteristics which makes them not eligible for loan.

Business Rule: All clients needs to have Atleast 2 characteristics which has achieved score more than neutral / weighted average score.

Model 2:

Table 6

Characteristic	Attribute	Score points

Residential_status	Renter	48
Residential_status	Other	54
Residential_status	Owner	95
Residential_status	Neutral	64
Years_current_job	Years_current_job <4	55
Years_current_job	Years_current_job >=4 AND Years_current_job <17	63
Years_current_job	Years_current_job >=17	68
Years_current_job	missing(Years_current_job) OR Years_current_job =.	63
Years_current_job	Neutral	61

Table 7

Client A			Client B		
Residential status	Owner	95	Residential status	Renter	48
Years of current job	18	63	Years of current job	2	55
Total		158	Total		103
Decision		Granted	Decision		Not Granted

In table 7, client A has higer score than the weighted average /Neutral score (table 6) for each characteristics which makes them eligible for loan (Siddiqi, 2016).

For Client B, we can observe the score achieved is lower than weighted average /Neutral score for each characteristics which makes them not eligible for loan (Siddiqi, 2016).

Conclusion & Discussion:

The Scorecard is created from a predictive model that assigns scores to individuals based on numerous different characteristics. Typically, this method begins with data preparation, such as cleaning and dealing with outliers, followed by the selection of strong variables using information value and other metrics for the modelling process, and ultimately using a logistic regression model to predict the likelihood of the target variable. In the end, the scorecard's performance is evaluated using various key metrics and scorecard will be created. Finally, The scorecard is used to decide whether or not a loan should be approved of the individual based on their characteristics.

References:

Siddiqi, N. (2016) *Intelligent Credit Scoring: Building and Implementing Better Credit Risk Scorecards*. Available at: https://lib.ugent.be/en/catalog/rug01:002333325

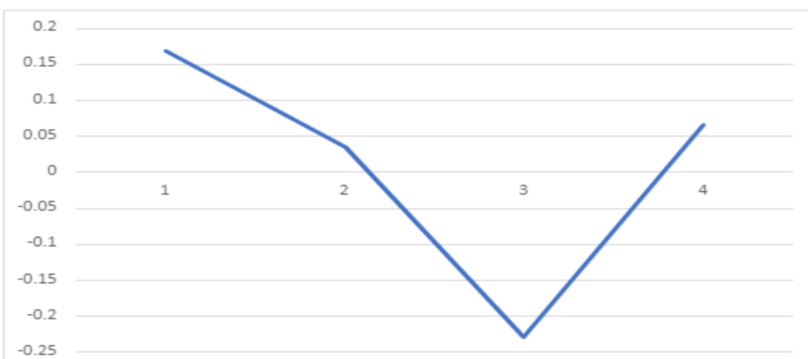
Baesens, B., Roesch, D. and Scheule, H. (2016) *Credit risk analytics : measurement techniques, applications, and examples in SAS*. Available at: https://openlibrary.org/books/OL29315061M/Credit_Risk_Analytics

Appendix

Age (model 1)

CLUSTER	Total	Total_Bad:	Total_Goo	N_Class	%obs	Min	Max	PCT_B	PCT_G	Bads	Goods	WOE	IVi	class
1	1000	250	750	29	2.90%	18	19	2.8	2.93	7	22	0.04652	0.006203	1
3	1000	250	750	47	4.70%	20	21	2.8	5.33	7	40	0.644357	1.632371	1
5	1000	250	750	23	2.30%	22	22	2	2.4	5	18	0.182322	0.072929	1
6	1000	250	750	26	2.60%	23	23	1.6	2.93	4	22	0.606136	0.808181	1
7	1000	250	750	16	1.60%	24	24	1.2	1.73	3	13	0.367725	0.19612	1
8	1000	250	750	34	3.40%	25	25	3.2	3.47	8	26	0.080043	0.021345	3
9	1000	250	750	44	4.40%	26	26	6	3.87	15	29	-0.43937	0.937316	3
10	1000	250	750	189	18.90%	27	31	27.6	16	69	120	-0.54523	6.324634	4
11	1000	250	750	379	37.90%	32	45	34.4	39.07	86	293	0.127213	0.593661	5
12	1000	250	750	186	18.60%	46	62	14.8	19.87	37	149	0.294416	1.491708	5
2	1000	250	750	26	2.60%	63	72	3.6	2.27	9	17	-0.46262	0.616831	5
4	1000	250	750	1	0.10%	75	75	0	0.13	0	1			5

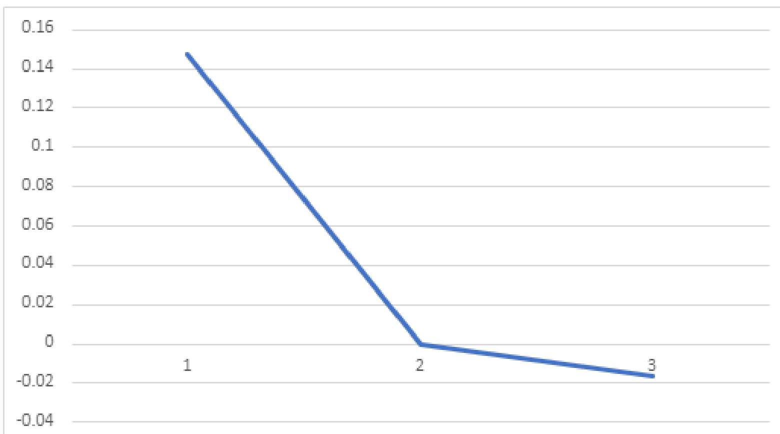
Row Labels	Sum of Bads	Sum of Goods	%G	%B	WOE	IV	
1	26	115	0.153333	0.104	0.168603238	0.008318	IF Age <25 then WoE_Age =0.168603237663131;Else
3	8	26	0.034667	0.032	0.034762106	9.27E-05	IF Age >=25 AND Age <26 then WoE_Age =0.0347621062592119;Else
4	84	149	0.198667	0.336	-0.228214272	0.031341	IF Age >=26 AND Age <32 then WoE_Age =-0.22821427236927;Else
5	132	460	0.613333	0.528	0.065062646	0.005552	IF Age >=32 then WoE_Age =0.0650626457560617
Grand Total	250	750			Sum of IV	0.045304	



Age (model 2)

CLUSTER	Total	Total_Bad:	Total_Goo	N_Class	%obs	Min	Max	PCT_B	PCT_G	Bads	Goods	WOE	IVi	class
1	1000	250	750	29	2.90%	18	19	2.8	2.93	7	22	0.04652	0.006203	1
3	1000	250	750	47	4.70%	20	21	2.8	5.33	7	40	0.644357	1.632371	1
5	1000	250	750	23	2.30%	22	22	2	2.4	5	18	0.182322	0.072929	1
6	1000	250	750	26	2.60%	23	23	1.6	2.93	4	22	0.606136	0.808181	2
7	1000	250	750	16	1.60%	24	24	1.2	1.73	3	13	0.367725	0.19612	2
8	1000	250	750	34	3.40%	25	25	3.2	3.47	8	26	0.080043	0.021345	2
9	1000	250	750	44	4.40%	26	26	6	3.87	15	29	-0.43937	0.937316	2
10	1000	250	750	189	18.90%	27	31	27.6	16	69	120	-0.54523	6.324634	3
11	1000	250	750	379	37.90%	32	45	34.4	39.07	86	293	0.127213	0.593661	3
12	1000	250	750	186	18.60%	46	62	14.8	19.87	37	149	0.294416	1.491708	3
2	1000	250	750	26	2.60%	63	72	3.6	2.27	9	17	-0.46262	0.616831	3
4	1000	250	750	1	0.10%	75	75	0	0.13	0	1			3

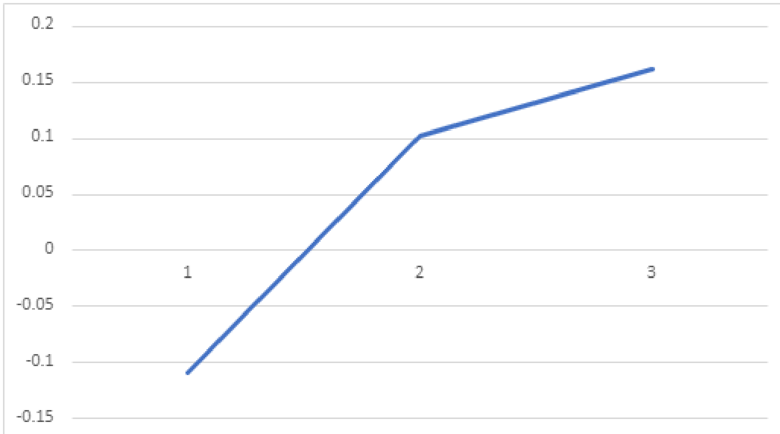
Row Labels	Sum of Bads	Sum of Goods	%G	%B	WOE	IV
1	19	80	0.106667	0.076	0.147215131	0.004515
2	30	90	0.12	0.12	0	0
3	201	580	0.773333	0.804	-0.016889319	0.000518
Grand Total	250	750				
					Sum of IV	0.005033



Debt ratio (model 1,2)

CLUSTER	Total	Total_Bad:	Total_Goo	N_Class	Min	Max	PCT_B	PCT_G	Bads	Goods	WOE	IVi	factor	%obs	class
1	1000	250	750	2	0.03366	0.04158	0	0.27	0	2			1	0.20%	1
5	1000	250	750	3	0.05346	0.06039	0	0.4	0	3			1	0.30%	1
6	1000	250	750	2	0.07524	0.08316	0.4	0.13	1	1	-1.09861	0.292963	1	0.20%	1
9	1000	250	750	68	0.08811	0.16335	5.2	7.33	13	55	0.343772	0.733379	2	6.80%	1
4	1000	250	750	398	0.16434	0.32274	51.2	36	128	270	-0.35222	5.353753	3	39.80%	1
12	1000	250	750	313	0.32373	0.50094	28.8	32.13	72	241	0.109519	0.365062	3	31.30%	2
10	1000	250	750	123	0.50292	0.61776	7.6	13.87	19	104	0.60134	3.768395	3	12.30%	2
8	1000	250	750	33	0.6237	0.6732	3.6	3.2	9	24	-0.11778	0.047113	3	3.30%	5
3	1000	250	750	19	0.67419	0.71775	0.8	2.27	2	17	1.041454	1.527466	3	1.90%	5
2	1000	250	750	13	0.72072	0.75636	0.8	1.47	2	11	0.606136	0.404091	3	1.30%	5
7	1000	250	750	14	0.75834	0.79497	0.8	1.6	2	12	0.693147	0.554518	3	1.40%	5
11	1000	250	750	12	0.80091	0.83358	0.8	1.33	2	10	0.510826	0.27244	3	1.20%	5

Row Labels	Sum of Goods	Sum of Bads	%G	%B	WoE	IVi	
1	331	142	0.441333	0.568	-0.109581605	0.013880337	IF DebtRatio <0.32373 then WoE_DebtRatio =-0.109581605327;Else
2	345	91	0.46	0.364	0.101656448	0.009759019	IF DebtRatio >=0.32373 AND DebtRatio <0.6237 then WoE_DebtRatio =0.101656448032518;Else
5	74	17	0.098667	0.068	0.161661544	0.004957621	IF DebtRatio >=0.6237 then WoE_DebtRatio =0.16166154363304
Grand Total	750	250					
					Sum of IV	0.028596976	> 0.02



Monthly Net Income (model 1,2)

CLUSTER	Total	Total_Bad	Total_Goo	N_Class	%class	Min	Max	PCT_B	PCT_G	Bads	Goods	WOE	IVi	class
1	1000	250	750	44	4.40%	878.22	1367.82	7.6	3.33	19	25	-0.82418	3.516482	1
5	1000	250	750	494	49.40%	1369.86	2042.04	56	47.2	140	354	-0.17096	1.504429	1
12	1000	250	750	329	32.90%	2044.08	2707.08	28.8	34.27	72	257	0.173798	0.950094	3
10	1000	250	750	94	9.40%	2711.16	3297.66	4.8	10.93	12	82	0.8232	5.048962	4
8	1000	250	750	16	1.60%	3370.08	3730.14	1.6	1.6	4	12	0	0	4
6	1000	250	750	7	0.70%	3791.34	4121.82	0.8	0.67	2	5	-0.18232	0.02431	4
4	1000	250	750	6	0.60%	4290.12	4741.98	0	0.8	0	6			4
9	1000	250	750	4	0.40%	4834.8	5191.8	0	0.53	0	4			4
3	1000	250	750	3	0.30%	5553.9	5913.96	0.4	0.27	1	2	-0.40547	0.054062	4
2	1000	250	750	1	0.10%	7485.78	7485.78	0	0.13	0	1			4
11	1000	250	750	1	0.10%	8263.02	8263.02	0	0.13	0	1			4
7	1000	250	750	1	0.10%	18980.16	18980.16	0	0.13	0	1			4

Row Labels	Sum of Goods	Sum of Bads	%G	%B	WoE	IVi	
1	379	159	0.505333	0.636	-0.099879169	0.013051	IF Monthly_Net_Income <2044.08 then WoE_Monthly_Net_Income =-0.0998791690720416;Else
3	257	72	0.342667	0.288	0.075479372	0.004126	IF Monthly_Net_Income >=2044.08 AND Monthly_Net_Income <2711.16 then WoE_Monthly_Net_Income =0.0754793721803637;Else
4	114	19	0.152	0.076	0.301029996	0.022878	IF Monthly_Net_Income >=2711.16 then WoE_Monthly_Net_Income =0.301029995663981
Grand Total	750	250					
					Sum of IV	0.040055	

Residential_status: (model 1,2)

Var	Total	Total_Bads	Total_Goods	N_Class	perct_obs	PCT_B	PCT_G	Bads	Goods	WOE	IVi	class
Renter	1000	250	750	422	42.20%	57.2	37.2	143	279	-0.43025	8.604903	1
Other	1000	250	750	269	26.90%	31.6	25.33	79	190	-0.22104	1.385159	2
Owner	1000	250	750	309	30.90%	11.2	37.47	28	281	1.207538	31.71799	3

Row Labels	Sum of Goods	Sum of Bads	%G	%B	WOE	IV	
1	279	143	0.372	0.572	-0.186853089	0.037371	IF Residential_status IN ('Renter') then WoE_Residential_status=-0.186853088911127;Else
2	190	79	0.253333	0.316	-0.095994745	0.006016	IF Residential_status IN ('Other') then WoE_Residential_status=-0.0959947450572749;
3	281	28	0.374667	0.112	0.524427034	0.043386	IF Residential_status IN ('Owner') then WoE_Residential_status=0.524427033843198;
Grand Total	750	250					
					Sum of IV	0.086773