Data Science Prodegree

Final Project

SAS

Evaluate the performance of Credit management system of bank.

Abstract

The goal of this project is to analyze the given credit management data of a bank and get insights to evaluate the performance of the credit management system. With given data we have generated reports in which bank is making profit w.r.t. Customer, City, Segment, Age group, spending habit of customer etc. Accordingly, we have forecasted and tried to get insights in which area bank can make more profit.

Table of Contents

Introduction

Credit Management

Credit management plays a vital role in the banking sector. As we all know bank is one of the major source of lending capital. Credit management is the process of monitoring and collecting payments from customers. A good credit management system minimizes the amount of capital tied up with debtors.

Credit Card Management

Credit Card Management helps you capitalize on revenue-generating opportunities and gain a competitive advantage in your market. The solution supports flexible rate and balance plans, a tiered late fee option, reward programs (including cash-back rewards), custom card images and other services your cardholders expect. You can deliver even more value to cardholders by enabling them to access current card balance, available credit, advance activity and payment activity information using Voice Response, or other interactive voice response (IVR) systems and online banking solutions.

In this project we are considering only Credit card management system as provided dataset is about credit card management system.

The Dataset

Number of sheets in the given dataset: 3

Total number of observations in the given dataset: 1500

Sheet 1: Customer Acquisition:

 Contains Customer demographics and credit card limit details involving Customer ID, Age, City, Credit card Product, Credit card limit, Company, Segment.

Sheet 2: Spend:

This sheet contains Customer's Spent data involving Customer ID, spend_Month, spend_Type, Spend_Amount.

Sheet 3: Repayment:

This sheet Contains Customer's Repayment data involving Customer ID, Repayment_Month, Repayment_Amount.

Bank's Profit from Credit Card:

Credit card issuing banks make money in the form of interest and fees as given below-

- Transaction Fees (Interchange Fees) - Earned when the issuing bank's card is used for merchant payments
- Interest - Charged on outstanding balance
- Late Payment Fees
- Cash Withdrawal Fees
- Annual Fees - Card account maintenance fees
- Currency Conversion Fees - For transactions done in a currency other than the card currency.

In This Project we are considering only Interest on outstanding balance as a bank's profit due to non-availability of data to consider other factors.

Purpose of Analytics

To identify trends and issues, uncover new insights, find and prevent future risks and efficient operations to make smarter decisions and meet business goals.

Pertaining to this project, the financial institution has data on customer acquisition as per the working segments of people and their spending and repayment habits.

The purpose is to basically analyse these customer habits and how they would influence the credit policy of the financial institution such that there would be a judicious balance in lending and repayment thus avoiding adverse situations like bankruptcy.

Data Exploration

Summary of the data/ Review of Literature

As We can see We have 3 Sheets of dataset "Customer Acquisition" ,"Spend" and "Repayment".

❖ ***Customer Acquisition:***

```
Proc contents data=Project.Customer_Acquisition;
title 'Customer_Acquisition';
run;
```

Customer_Acquisition
The CONTENTS Procedure

Data Set Name	PROJECT.CUSTOMER_ACQUISITION	Observations	100
Member Type	DATA	Variables	8
Engine	V9	Indexes	0
Created	03/24/2018 23:18:12	Observation Length	112
Last Modified	03/24/2018 23:18:12	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	65536
Number of Data Set Pages	1
First Data Page	1
Max Obs per Page	584
Obs in First Data Page	100
Number of Data Set Repairs	0
Filename	/folders/myshortcuts/sf_myfolders/SAS Project/customer_acquisition.sas7bdat
Release Created	9.0401M5
Host Created	Linux
Inode Number	52
Access Permission	rwxrwx---
Owner Name	root
File Size	128KB
File Size (bytes)	131072

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
3	Age	Num	8	4.1		Age
4	City	Char	15	$15.	$15.	City
7	Company	Char	15	$15.	$15.	Company
5	Credit_Card_Product	Char	22	$22.	$22.	Credit Card Product
2	Customer	Char	18	$18.	$18.	Customer
6	Limit	Num	8	BEST20.		Limit
8	Segment	Char	17	$17.	$17.	Segment
1	SI_No_	Num	8	BEST12.		SI No:

```
Proc means data=project.customer_acquisition;
title 'Customer_Acquisition';
run;
```

Customer_Acquisition
The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SI_No_	SI No:	100	50.5000000	29.0114920	1.0000000	100.0000000
Age	Age	100	38.2196006	23.0755226	0.2463658	79.0136998
Limit	Limit	100	284700.61	218494.97	10000.00	500000.00

> ➤ **Sl No:** which shows serial no of content. As it is not necessary in analysis, we will handle it later.

> ➤ We can see that there are total **100** unique **customer's ID** of bank.

> ➤ The **age** of customers ranges from 0.2 years old to a maximum of 79 years old. However, Credit card would not be offered to a person who is below 18 years old so, in this case we assume that data wasn't collected correctly. This will be fixed later.

> ➤ We have Customer's from BANGALORE, CALCUTTA, COCHIN, BOMBAY, DELHI, PATNA, CHENNAI and TRIVANDRUM **cities** So, we can treat city as a factor.

> ➤ We have **Credit card product**s of only 3 types i.e. Gold, Silver and Platinum- i.e. factor type.

> ➤ We have only 3 types of **limits** on credit card i.e. 10,000/-, 100,000/- and 500,000/- in Indian rupees.

> ➤ We have Customers from 41 different **companies**.

> ➤ We have categorized Customers into 5 **segments** i.e. Self Employed, Salaried_MNC, Salaried_Pvt, Govt and Normal Salary and this is of factor type.

❖ *Spend:*

The CONTENTS Procedure

Data Set Name	PROJECT.SPEND	Observations	1500
Member Type	DATA	Variables	5
Engine	V9	Indexes	0
Created	03/24/2018 23:18:12	Observation Length	56
Last Modified	03/24/2018 23:18:12	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	65536
Number of Data Set Pages	2
First Data Page	1
Max Obs per Page	1166
Obs in First Data Page	1122
Number of Data Set Repairs	0
Filename	/folders/myshortcuts/sf_myfolders/SAS Project/spend.sas7bdat
Release Created	9.0401M5
Host Created	Linux
Inode Number	54
Access Permission	rwxrwx---
Owner Name	root
File Size	192KB
File Size (bytes)	196608

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
4	Category	Char	18	$18.	$18.	Type
2	Customer	Char	14	$14.	$14.	Customer
3	Month	Num	8	MONYY7.		Month
1	SI_No_	Num	8	BEST12.		SI No:
5	Spend_Amount	Num	8	11.3		Amount

Spend

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SI_No_	SI No:	1500	750.5000000	433.1570154	1.0000000	1500.00
Month	Month	1500	16572.89	332.0377941	16073.00	17138.00
Spend_Amount	Amount	1500	253173.73	143824.58	53.7648869	499663.19

➢ Total no. of observations is **1500**.

➢ **SI No:** which shows serial no of content. As it is not necessary in analysis we will fix this later.

➢ **Customer** shows unique Customer ID.

➢ **Month** shows on which day/Month customer has spent.

➢ **Type** shows on which type of shopping done by customer.

➢ **Amount** shows data how much customer spent w.r.t. month and type.

❖ *Repayment*

The CONTENTS Procedure

Data Set Name	PROJECT.REPAYMENT	Observations	1523
Member Type	DATA	Variables	5
Engine	V9	Indexes	0
Created	03/24/2018 23:18:12	Observation Length	48
Last Modified	03/24/2018 23:18:12	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	65536
Number of Data Set Pages	2
First Data Page	1
Max Obs per Page	1360
Obs in First Data Page	1309
Number of Data Set Repairs	0
Filename	/folders/myshortcuts/sf_myfolders/SAS Project/repayment.sas7bdat
Release Created	9.0401M5
Host Created	Linux
Inode Number	56
Access Permission	rwxrwx---
Owner Name	root
File Size	192KB
File Size (bytes)	196608

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat	Label
2	Customer	Char	11	$11.	$11.	Customer
5	E	Char	9	$9.	$9.	E
3	Month	Num	8	MONYY7.		Month
4	Repayment_Amount	Num	8	11.3		Amount
1	SL_No_	Num	8	BEST12.		SL No.

Repayment

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
SL_No_	SL No.	1500	750.5000000	433.1570154	1.0000000	1500.00
Month	Month	1500	16580.05	328.3451965	16073.00	17138.00
Repayment_Amount	Amount	1500	254498.43	145872.98	107.3390498	499837.44

> **Sl No and E:** which shows serial no of content and the other is just a blank column. As it is not necessary in analysis we will fix this later.

> **Customer** shows unique respective Customer ID.

> **Month** shows on which day/Month customer has Repaid the amount.

> **Amount** shows data as to how much the customer Repaid w.r.t. month.

> We can see that total no. of observations are 1523 but in the 2nd screenshot it is showing only 1500. There might me some missing values, we will fix this later.

Transform data

Treating missing data

The missing data must be treated to ensure accurate analysis.

Repayment

The MEANS Procedure

Variable	Label	N Miss	N
SL_No_	SL No.	23	1500
Month	Month	23	1500
Repayment_Amount	Amount	23	1500

Here, N Miss shows total no. of missing values. As we have seen in data exploration chapter, Repayment dataset it was showing total **1523** observation but while counting w.r.t. variables it was showing only **1500** observations; those extra **23** records were added while importing data to SAS and all these variables are default empty slots, hence we decided to delete them as they have no impact on the data.

Dropping unnecessary column:

SI No. column is not necessary in our analysis as we have seen earlier. So, using the code below, we delete SI No. and E column from all 3 sheets:

```
data Project.Customer_Acquisition (drop=Sl_No_);
    set Project.Customer_Acquisition;
run;

data Project.Repayment (drop=Sl_No_ E);
    set Project.Repayment;
run;

data Project.Spend (drop=Sl_No_);
    set Project.Spend;
run;
```

Data Manipulation:

While considering the Customer_Aquisition data, we can see that there is no proper limit maintained so we deal with it as follows:-

```
data Project.Customer_Acquisition;
    set Project.Customer_Acquisition;

    if Limit in(10001, 10002) then
        Limit=10000;

    if Limit in(100001, 100002, 100003) then
        Limit=100000;
run;
```

Data Preparation

To get insights to evaluate Credit management system of the bank, we need customer's acquisition, spending and repayment data in one dataset. As we have seen that we have Customer's ID in every dataset, we can merge the data into one dataset after sorting with Customer ID.
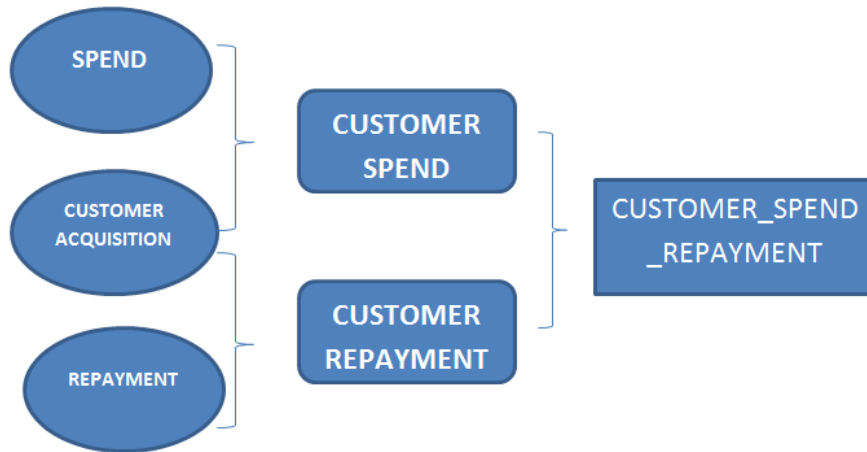
Steps Performed for merging data:

1. Sorted all 3 datasets with Customer in ascending order.
2. With sorted dataset we created 3 temporary datasets by grouping data on Spend month and Repayment Month and Category w.r.t. Customer named '*SpendMonthlyGroup*', '*RepaymentMonthlyGroup*', and '*SpendCategoryGroup*' respectively.
3. We merged '*Customer_Acquisition_Sorted*' and '*RepaymentMonthlyGroup*' by *Customer* to get "*Customer_Repayment*" dataset.
4. We merged '*Customer_Acquisition_Sorted*' and '*SpendMonthlyGroup*' by *Customer* to get "*Customer_Spend*" dataset.
5. Finally merged "*Customer_Repayment*" and "*Customer_Spend*" by *Customer* and *Month* and got "*MONTHLY_SPEND_REPAYMENT*" .

The table below shows the variables contained in the merged dataset.

| | Alphabetic List of Variables and Attributes | | | | | |
#	Variable	Type	Len	Format	Informat	Label
2	Age	Num	8	4.1		Age
3	City	Char	15	$15.	$15.	City
6	Company	Char	15	$15.	$15.	Company
4	Credit_Card_Product	Char	22	$22.	$22.	Credit Card Product
1	Customer	Char	18	$18.	$18.	Customer
5	Limit	Num	8	BEST20.		Limit
8	Month	Num	8	MONYY7.		Month
10	Repayment_Amount	Num	8	11.3		Amount
7	Segment	Char	17	$17.	$17.	Segment
9	Spend_Amount	Num	8	11.3		Amount

As we saw in Introduction Chapter, Bank's profit is depending on customer's extra spending than limit and interest on outstanding amount.

To calculate bank's profit and get insight on same we have added few more variable in 'MONTHLY_SPEND_REPAYMENT' dataset with below code:

```
data Project.MONTHLY_SPEND_REPAYMENT;
    set Project.MONTHLY_SPEND_REPAYMENT;

    if Age<18 then
        Age=38.2;   /**Mean of Age**/

    if Age >=18 and Age <=30 then
        Age_Group='18 - 30';
    else if Age> 30 and Age <=60 then
        Age_Group='31 - 60';
    else if Age >60 then
        Age_Group='60+';

    if Spend_Amount > Limit then
        Spend_Exceed_By=Spend_amount - Limit;
    else
        Spend_Exceed_By='';

    if Repayment_Amount > Limit then
        Repayment_Exceed_By=Repayment_amount - Limit;
    else
        Repayment_Exceed_By='';

    if Repayment_Amount < Spend_Amount then
        do;
            Pending_Pay=Spend_Amount - Repayment_Amount;
            Interest=Pending_Pay*(2.9/100);
        end;
    else if Repayment_Amount > Spend_Amount then
        Extra_Pay=Repayment_Amount - Spend_Amount;
    Bank_Profit=Spend_Exceed_By+Interest;
run;
```

Alphabetic List of Variables and Attributes						
#	Variable	Type	Len	Format	Informat	Label
2	Age	Num	8	4.1		Age
11	Age_Group	Char	7			
17	Bank_Profit	Num	8			
3	City	Char	15	$15.	$15.	City
6	Company	Char	15	$15.	$15.	Company
4	Credit_Card_Product	Char	22	$22.	$22.	Credit Card Product
1	Customer	Char	18	$18.	$18.	Customer
16	Extra_Pay	Num	8			
15	Interest	Num	8			
5	Limit	Num	8	BEST20.		Limit
8	Month	Num	8	MONYY7.		Month
14	Pending_Pay	Num	8			
10	Repayment_Amount	Num	8	11.3		Amount
13	Repayment_Exceed_By	Num	8			
7	Segment	Char	17	$17.	$17.	Segment
9	Spend_Amount	Num	8	11.3		Amount
12	Spend_Exceed_By	Num	8			

We have created 'Age_Group', 'Spend_Exceed_By', 'Repayment_Exceed_By', 'Pending_Pay', 'Interest', 'Extra_Pay' and 'Bank_Profit' variables in the merged dataset.

Reports

To Explore the fata more and find some trends and insights, we have churned the given data and generated the following reports.

Customer's Spending Habits

PDF

Customer Spending Habit.pdf

Obs	Customer	Category	Spend_Amount
1	A1	SANDALS	419767.586
2	A10	CLOTHES	452583.880
3	A100	BUS TICKET	276927.801
4	A11	CLOTHES	357222.728
5	A12	FOOD	380558.839
6	A13	CLOTHES	470083.384
7	A14	TRAIN TICKET	375784.598
8	A15	MOVIE TICKET	441247.830
9	A16	CAR	361828.497
10	A17	BIKE	423619.100
11	A18	RENTAL	471240.750
12	A19	MOVIE TICKET	499430.644
13	A2	FOOD	437729.318
14	A20	CAMERA	373695.356
15	A21	AUTO	385697.514
16	A22	BUS TICKET	470068.870
17	A23	JEWELLERY	407228.353
18	A24	AIR TICKET	316587.956
19	A25	AIR TICKET	467053.192
20	A26	CAR	379008.578
21	A27	RENTAL	382334.719
22	A28	JEWELLERY	470636.609
23	A29	BUS TICKET	321720.324
24	A3	FOOD	464047.032
25	A30	AIR TICKET	410069.022
26	A31	BUS TICKET	415080.364
27	A32	FOOD	472240.523

Analysis of customer Demographics in terms of:

i. Age group

ii. City

iii. Product Segment

Age_Group	Frequency
18 - 30	182
31 - 60	706
60+	214

City	
City	Frequency
BANGALORE	214
BOMBAY	161
CALCUTTA	182
CHENNAI	89
COCHIN	225
DELHI	78
PATNA	76
TRIVANDRUM	77

Segment	
Segment	Frequency
Govt	202
Normal Salary	327
Salaried_MNC	177
Salaried_Pvt	200
Self Employed	196

Monthly Spend of each Customer

Monthly Spend of each customer.pdf

Obs	Customer	Spend_Amount
1	A1	527780.251
2	A10	392845.437
3	A100	276927.801
4	A11	335117.033
5	A12	567470.643
6	A13	494109.341
7	A14	520577.103
8	A15	388481.355
9	A16	419412.421
10	A17	416193.438
11	A18	445013.202
12	A19	403925.744
13	A2	259057.819
14	A20	415853.346
15	A21	486093.486
16	A22	553711.389
17	A23	446576.864
18	A24	404869.325
19	A25	407574.120
20	A26	486685.645
21	A27	604604.600
22	A28	479732.816
23	A29	475386.444
24	A3	376153.171

Monthly Repayment of each Customer

Monthly Repayment of each customer.pdf

Obs	Customer	Repayment_Amount
1	A1	547362.849
2	A10	449147.761
3	A100	459684.083
4	A11	454092.535
5	A12	554495.460
6	A13	513657.230
7	A14	406582.175
8	A15	700431.608
9	A16	395250.554
10	A17	385213.837
11	A18	507159.757
12	A19	466446.804
13	A2	295147.669
14	A20	495864.302
15	A21	571188.242
16	A22	502762.595
17	A23	538697.125
18	A24	425129.884
19	A25	440364.472
20	A26	487283.118
21	A27	398924.457
22	A28	629553.561
23	A29	511872.064
24	A3	392087.627
25	A30	554254.905

Monthly Interest for Each Customer

Monthly Interest (rate
= 2.9) for each custom

| Obs	Customer	Interest
1	A1	8848.64
2	A10	7582.74
3	A100	.
4	A11	8953.36
5	A12	13860.32
6	A13	4469.21
7	A14	7666.13
8	A15	10955.39
9	A16	13031.65
10	A17	13827.95
11	A18	5102.52
12	A19	22997.17
13	A2	9253.44
14	A20	5384.96
15	A21	11022.00
16	A22	9587.83
17	A23	15132.61
18	A24	12842.51
19	A25	7662.59
20	A26	2169.64
21	A27	15030.83
22	A28	10261.32
23	A29	13060.84
24	A3	6307.40
25	A30	14585.41

Monthly Profit for the Bank

Monthly Profit for
bank.pdf

Obs	Repayment_Month	Bank_Profit
1	JAN2004	1044458.84
2	FEB2004	372539.23
3	MAR2004	362876.94
4	APR2004	566483.95
5	MAY2004	531133.92
6	SEP2004	469951.87
7	NOV2004	543592.61
8	JAN2005	689483.49
9	FEB2005	553830.64
10	APR2005	660526.48
11	MAY2005	566064.09
12	JUN2005	557635.37
13	JUL2005	565585.77
14	AUG2005	420926.92
15	SEP2005	367268.73
16	OCT2005	369203.72
17	NOV2005	260831.09
18	DEC2005	468250.51
19	JAN2006	248798.09
20	FEB2006	492288.67
21	MAR2006	492738.82
22	APR2006	633808.84
23	MAY2006	399948.51
24	JUN2006	408986.54
25	JUL2006	327359.71

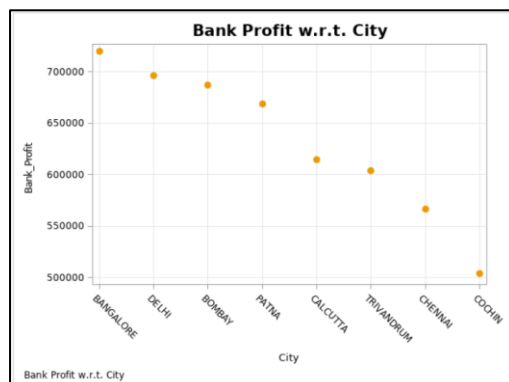## Category in which customers are Spending More Money

Obs	Category	Spend_Amount
1	BUS TICKET	295469.908

Graphical representation:
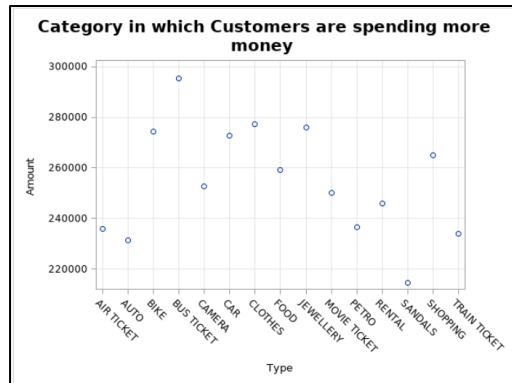
City in which Bank in making More Profit

| Obs	City	Bank_Profit
1	BANGALORE	720037.03

Customers Spending More Than Limit

customer is spending
 More than limit.pdf

| Obs	Customer	Spend_Exceed_By
1	A1	955671.22
2	A10	352182.23
3	A100	176927.80
4	A11	332466.65
5	A12	515581.47
6	A13	403516.17
7	A14	686989.02
8	A15	300607.20
9	A16	806678.71
10	A17	580935.36
11	A18	345013.20
12	A19	430641.38
13	A2	219509.76
14	A20	405853.35
15	A21	476093.49
16	A22	453711.39
17	A23	412949.18
18	A24	329096.15
19	A25	356007.32
20	A26	246476.26
21	A27	664906.35
22	A28	525317.66
23	A29	468181.91
24	A3	366153.17
25	A30	521203.16

Age Group Spending More Money

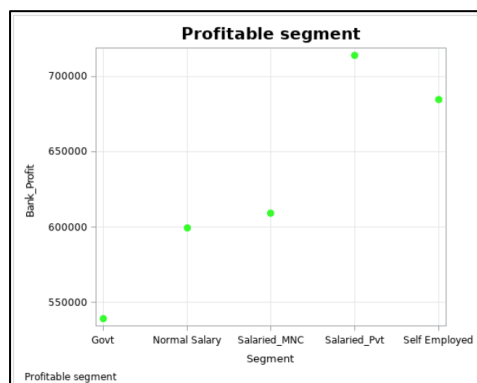| Obs	Age_Group	Spend_Amount
1	60+	494351.190

Spending money w.r.t. Age_Group
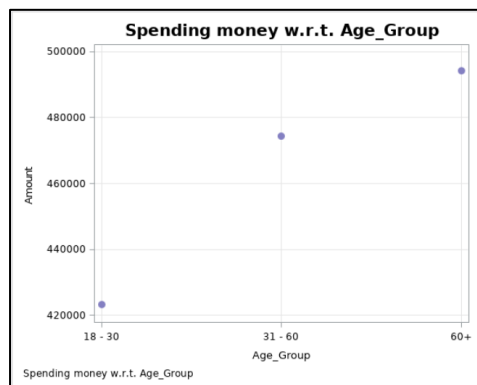
Which Segment people are Spending More Money

Obs	Segment	Spend_Amount
1	Self Employed	497059.252

Segment in which Customers are spending more money
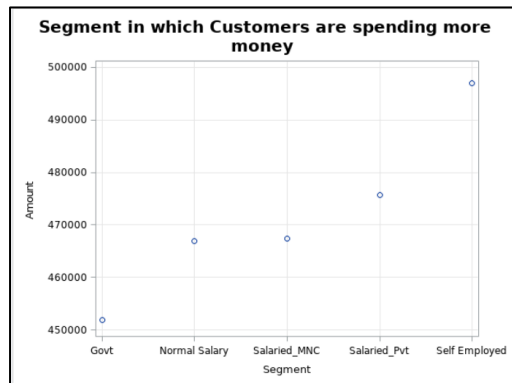
Profitable Segment

Obs	Segment	Bank_Profit
1	Salaried_Pvt	714027.26

Profitable segment

Highest Paying 10 Customers

Please find below list of customers who are top 10 in paying

Obs	Customer	Repayment_Amount
1	A72	469945.866
2	A79	462834.473
3	A100	459684.083
4	A87	454070.293
5	A76	452466.196
6	A81	418250.952
7	A97	411156.166
8	A78	406001.103
9	A64	403832.558
10	A98	400810.216

Forecast

After generating reports wrt Spending habits, segments, city, age group, and so on, we are now building a forecasting model for Bank profit for the next 12 months and the result is as follows: -

Forecast of Bank's Profit

Obs	MONTH	_TYPE_	_LEAD_	Bank_Profit
1	JAN2007	FORECAST	1	393592.80
2	FEB2007	FORECAST	2	392008.72
3	MAR2007	FORECAST	3	390424.64
4	APR2007	FORECAST	4	388840.56
5	MAY2007	FORECAST	5	387256.49
6	JUN2007	FORECAST	6	385672.41
7	JUL2007	FORECAST	7	384088.33
8	AUG2007	FORECAST	8	382504.25
9	SEP2007	FORECAST	9	380920.18
10	OCT2007	FORECAST	10	379336.10
11	NOV2007	FORECAST	11	377752.02
12	DEC2007	FORECAST	12	376167.94

Conclusion

Credit card management system is a very crucial part of the banking sector and it plays a vital role in the growth of a bank's profit. It goes hand -in-hand with the customer spending and repayment habits.

There are some elements which can be improved so that the bank can make more profit such as, giving promotional offer to customer w.r.t. their spending habit, and encourage customers to upgrade their card in terms of credit card limit.