# Project 1

# Portfolio Investigation and Top-Up Loan Strategy

# Project 2

# One-Touch Cross Product Financial Impact Calculator

A Project Report Submitted in partial fulfillment of the degree of

Master of Computer Applications

By
**Shravan Kumar**


School of Computer and Information Sciences
University of Hyderabad, Gachibowli
Hyderabad - 500046, India

June 2020


# CERTIFICATE

This is to certify that the Project Report entitled "**Project 1: Portfolio Investigation and Top-Up Loan Strategy & Project 2: One-Touch – Cross Product Financial Calculator**" submitted by **Shravan Kumar** bearing Reg. No. 17MCMC56 in partial fulfillment of the requirements for the award of Master of Computer Applications is a bonafide work carried out by them under my supervision and guidance.

The Project Report has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Dr. Y.V Subba Rao
Internal Supervisor
SCIS, University of Hyderabad

Mr. Kaustav Mukherjee
External supervisor
Lead AVP, HSBC - HDPI

Dean, School of Computer and Information Sciences
University of Hyderabad

# DECLARATION

I, **Shravan Kumar** hereby declare that this dissertation entitled "**Project 1: Portfolio Investigation and Top-Up Loan Strategy & Project 2: One-Touch – Cross Product Financial Calculator**" submitted by us under the guidance and supervision of Dr. Y.V Subba Rao is a bonafide work. We also declare that it has not been submitted previously in part or in full to this or any other University or Institution for the award of any degree or diploma.

Date: 10-June-2020

Shravan Kumar
Reg. No.: 17MCMC56

Signature of the Student

# Acknowledgment

I would like to thank our Dean for supporting and encouraging me to take up an internship. I would also like to thank my internal supervisor, Dr. Y.V. Subba Rao for guiding me and providing valuable support during the internship period. I would also like to thank the University of Hyderabad for everything that they have provided us for the tenure of our degree.

I would like to thank HSBC Global Analytics Centre, Kolkata for giving me this opportunity to intern with them and get hands-on experience on how corporate companies function and the work they do. I would like to thanks Ipsita Ghosh (VP), Kaustav Mukherjee (Lead AVP), Debolina Ghosh (AVP), Abhishek Mishra (Manager) and Ritobrata Sarkar (Asst. Manager) for providing me valuable insights, guiding me, supporting me, and helping me with my work whenever needed.

I take this opportunity to express our gratitude to my family members for their constant encouragement and moral support.

**Shravan Kumar**

**17MCMC56**

# Abstract

# Portfolio Investigation and Top-Up Loan Strategy

In this project, an effort had been made to categorize the customers and identify the risk based on a given dataset that contains all the accounts booked on a bank's portfolio until Feb'19. Various customer demographics, bureau, as well as intrinsic loan information, are present in the dataset which have been studied. To create a thorough understanding of the portfolio, a report has been prepared called Management information (MI) which consists of dashboards for:

- Portfolio view – To understand the current health and characteristics of the bank standing at Feb'19
- Vintage view – To study the performance of the booked loans in previous years, starting from 2007

Thereafter, the profiling technique was applied to classify the customers into bands based on different characteristics. All the combinations of bands were also created to study the segments where performance was bad, and credit risk was high.

This exercise helps to identify the most suitable customer to whom banks can offer top-up loans for the open loan's portfolio. Top-up loans are provided by banks, financial institutions, and housing finance companies that offer the customer to borrow an additional amount of money near the end of his current term, depending on his performance throughout the years as well as other risk characteristics.

# Abstract

# One-Touch - Cross Product Financial Impact Calculator

Detailed Financial Impacts across various metrics is essential for a strategy change submission to the business management committee. Idea is to create a simplified GUI for calculating the financial impacts of a strategy change based on defined inputs from users ensuring we capture the commercial benefits during every strategy development.

It helps the business management committee to boost profitability and performance. It focuses on increasing visibility, streamlining processes, minimizing cost, and the improvement on cash flow. Using the calculator, the user can calculate the financial benefits of strategy change in quick time with minimal effort

# Contents

## One-Touch – Cross Product Financial Calculator

# List of Figures

**Portfolio Investigation and Top-Up Loan Strategy**



**One-Touch – Cross Product Financial Calculator**

# List of Tables

## Portfolio Investigation and Top-Up Loan Strategy



## One-Touch – Cross Product Financial Calculator

# Chapter 1

# Introduction

People require additional financial support for various purposes like higher education, buying a car/home, etc. Any bank or financial institution has a large loan portfolio, which is the case study here. A thorough study on the nature of the portfolio as well as identifying the risk through characteristics and performance is the objective here. Once the risk is identified, we are looking to devise top-up strategies, assign risk grades, etc.

## 1.1 Context and background

Management information (MI) is a summarization of loan portfolio data which helps in identifying the potential risk in the portfolio and quick strategic intervention to mitigate it. Regular strategic intervention is required to remain competitive in the market by making any changes in strategy. Strategy change may require to increase or decrease the limits of cards, score cut-off modification, delinquency pattern, top-up offering, planning of collection, etc. We have to monitor the strategy regularly, whether the risk portfolio has been increased or it is becoming better accordingly, we can modify or redevelop the strategy if at some point it is not optimal.

If we identify any high-risk segment in the MI report for that we can put more robust checking process scrutiny therefore we can quickly mitigate and the bank can generate a good profit, many times a big loss happens because of a higher risk portfolio. MI report gives an idea about the volume of delinquent and charged-off customers.

A top-up loan is issued by banks on existing loans if the customer has good repaying history and credit scores. Before availing top-up to the customer, banks check the customer repayment capabilities from many dimensions. If the customer fulfills the all eligibility criteria, then only the bank grants the loan by re-adjusting his EMI and term. Borrowers can utilize a top-up loan for purposes of modification or construction of your house or to meet your expenses like funding for children's education, etc. However, they cannot utilize the top-up loan amount for a speculative motive.

## 1.2 Project plan to carry out

To offer top-up, banks have to develop a strategy so that banks can cater to the requirement of customer satisfaction. In this project work, we will first try to understand the given loan portfolio dataset by creating a Management Information (MI) report which will give the insights of the portfolio, and accordingly, we can take the rational decision while developing the top-up strategy.

After creating the MI, we will categorize the customers in different bands according to their Credit score, debt to income ration, and other demographics. By combining the created band's, risk segmentation will be done which is more or less prerequisite of strategy development. While diving into the loan portfolio data we will also collect the pearls in the form of GUI based dashboards.

## 1.3 The rationale of the project

As the current situation of pandemic COVID – 19, the volatile nature of the market leads to uncertainty and a high risk of loss. In the field of finance, the organization has to cop up many types of unexpected risks so that they can minimize the risk and can create a favorable environment for the organization.

Ignoring the different risk aspects can lead to huge financial loss to any bank. To sustain in such type of market where dynamism is an unchanging principle, the organization has to investigate the portfolio regularly and evolve the strategy continuously to minimize the risk to its lowest. MI report gives the idea of a strategy change requirement.

For data analysis, many banks use SAS (Statistical Analysis Software) which is an advanced analytical tool. SAS is paid software but also provides full customer care support. Nowadays Python is becoming popular for data analytics, because of its easiness and it is providing better data visualization, libraries like NumPy, Pandas, and data analysis algorithms.

A python is an open-source software it has strong community support but doesn't have any reliable and dedicated support system. In this project, we are using the Python Programming language to explore its capabilities in the context of analysis, and its ability to replace SAS in the future to save the banks the cost of purchasing SAS.

# Chapter 2

# Proposed Solution

## 2.1 Overview

Dataset Initially had a huge amount of information that is not required for visualization, so the dataset needs to be in the summarized form. I have used Python for summarizing the dataset and bring it in a form that is effective for visualization. According to the summarized data dashboard is created in Microsoft Excel using different charts, graphs, and tables. This also allows slicing and dicing the data to get hidden patterns which can provide key insights for future business changes to improve portfolio performance

## 2.2 Introduction to Visualization

Data visualization is the presentation of data in a pictorial or graphical format. It enables decision-makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data we see and how it's processed.

## 2.3 Data Manipulation and Understanding

The collection and organization of data are an integral and critical part of the visualization process. In business intelligence applications the idea of bringing in the right data helps in visualize the trend, find insights, and further helps in decision making, thus data understanding plays an important role. Data Manipulation is helpful to clean up an existing dataset.

## 2.4 Data Preparation

The collection and organization of data are an integral and critical part of the visualization process. In business intelligence applications the idea of bringing in the right data helps in visualize the trend, find insights, and further helps in decision making, thus data understanding plays an important role. Or we'll simply need to "dig deeper" into the data we already have. It will allow us to subset our data, aggregate it, and transform our data to help us find more insights.

## 2.5 Managing Information (MI) Report

In business to identify the potential risk or high-risk portfolio segment, the Management Information report is a helpful way of understanding the live book. The Managing information report gives the idea of which segment is not performing well so that proper risk mitigation techniques can be applied. Without identification of risk, we can't judge so MI report becomes important.

### 2.5.1 Portfolio View

In the Managing information report, we calculate the volume of account, balance, bad rate, bad loan amount, and delinquency rate according to different customer demographics. By calculating these things, it gives insights into the distribution of customers in a different category as per the deferent demographics.

- **Customer Demographics**: The statistical information related to the unique customer is called customer demographics. Example: Income
- **The volume of Account:** Total no of accounts in different bands.
- **Balance:** The left amount of loan which has to be paid by the customer.
- **Bad rate:** The percentage of the loans which are charged-off (not repaid)
- **Bad amount**: The original loan amount which was charged-off
- **Delinquency rate:** The percentage of customers who have failed to pay the minimum monthly due amount.

### 2.5.2 Vintage View

Vintage view analysis is a method of evaluating loan portfolio credit quality in a given loan pool for a particular period. In our case, the period is yearly but we have summed up the values for some years where the number of customers is very low as compare to the other years.

The vintage view gives the trend of yearly acquisitions or exposure. It gives insight into the distribution of demographics across vintage (Profile of bookings). We are calculating closure rate, bad rate, the delinquency rate for every year. Bad rate, Delinquency rate across vintage by demographics and bureau attributes. Example of bureau attributes: DTI, no of total enquires, Bureau utilization.

- Closure rate: the percentage of accounts that have been closed or fully paid in a given pool of loan data is called closure rate.
- Bureau attributes: This is the details of customer which describe his credibility which is analyzed by different bureau agency according to the customer past behavior and income index.
- DTI (Debt to income): The percentage of the monthly gross income of customers that goes toward going existing debts.

- Total no of inquiry: Number of credits enquires done by the customer.
- Bureau Utilization: It is the balance to the credit limit on all trades.

## 2.6 Segment identification for Top-Up Strategy

Segment Identification is a crucial step in the top-up offering strategy. In this step, the analyst decides which segment will be the least risky and in turn, most suitable for a top-up loan. According to the MI, one can understand the customer base is performing well and which is not. Here the meaning of the customer performance is that the repayment pattern of loan amount by the customer.

In the segment identification step, we considered the bureau information like debts burden, credit inquiries, the term left, loan amount paid, customer income, etc. A high debt burden shows the customer is financially weak in the current time if he is already burdened by other debts than there is less a chance that the customer will be able to pay the additional amount of top-up loan. If any customer has a high credit inquiry then maybe the reason is that he is seeking credit on an urgent basis but is being declined, which indicated a risky profile. The term left helps in identifying loans that are near their end of term, and could potentially be offered a top-up. The past repayment pattern of a customer gives an understanding of customer behavior.

For the segmentation process, we are calculating the percentage of delinquency rate including the grace period, delinquency excluding the grace period, bad rate, bureau bad rate, overall bad rate including the grace period, and overall excluding grace period. Waterfall model is also used to understand the decreasing trend in the number of accounts(customer) by excluding the accounts which have been closed or fully paid, charged-off or bad, bureau bad accounts and newly added account which is less than 6 months. After this combined the different demographics DTI band, Income band, Inquiry band, and risk grade and find the total accounts, overall bad rate including grace period and excluding grace period.

# Chapter 3

# Data Preparation

## 3.1 Data Acquisition

In the first step, we identify the data or the source of information and from that, we should be able to understand what information we want to consider for our study to retrieve data from. For this operation, we are using Python in Google Colab.

- Google Colab: We are using Google Colab which is a python development environment that runs in the browser using Google Cloud.

### 3.1.1 Importing Data from Kaggle

- Kaggle: Kaggle is a platform where the organization posts the data and analysts & data miners compete for the development of the best analytical and predictive model.

  We are importing a dataset from Kaggle which is provided by "Lending Club loan data". For importing the dataset, we have created JSON API which has to be uploaded in Google Colab by following line of python code:

Figure 3.1: Data import from Kaggle

```
#create upload link
uploaded = files.upload()
```

- **kaggle.json**(application/json) - 71 bytes, last modified: n/a - 100% done

```
Saving kaggle.json to kaggle.json
```

```
#copy json file in a directory
!mkdir -p ~/.kaggle
!cp kaggle.json ~/.kaggle/

#permission to access kaggle data
!chmod 600 ~/.kaggle/kaggle.json

#Download kaggle data
!kaggle datasets download -d wendykan/lending-club-loan-data

#Unzip kaggle data
from zipfile import ZipFile
file_name = "lending-club-loan-data.zip"

with ZipFile(file_name, 'r') as zip:
  zip.extractall()
  print('Unzip Done')

#Create DataFrame with all data
data = pd.read_csv('loan.csv')
total_rows_Data=len(data.axes[0])
print('Total Rows in Data:', total_rows_Data)
```

In python using Panda's library, we are creating a new dataset that will contain the only subset of the original dataset.

Figure 3.2: Creating Dataset

```
#Create Working Dataset
df = data.loc[:, ['annual_inc', 'loan_amnt', 'annual_inc_joint', 'is
sue_d', 'grade', 'loan_status', 'dti', 'dti_joint', 'acc_now_delinq'
, 'inq_last_12m', 'inq_fi', 'hardship_flag', 'collections_12_mths_ex
_med']]
```

## 3.2 Data Parsing

In this step, we are providing a structure for the meaning of the data and dividing it into different categories. The amount of data one can collects and analyze in immense. It is necessary to structure the collected data so that others can understand the format, tags, and indices of data you used.

Figure 3.2: Data Parsing

```
#Generate Bands according to Total income
inc_bands = [0, 34000, 42000, 50000, 56000, 65000, 72096, 83806, 988
00, 125000, 10000000]
inc_bands_name =[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]
df['inc_band'] = pd.cut(df['total_inc'], inc_bands, labels = inc_ban
ds_name)

#Banding of Employee Length
df['emp_len_band'] = df['emp_length']
df['emp_len_band'].replace(to_replace = ['< 1 year', '1 year', '2 ye
ars', '3 years', '4 years', '5 years', '6 years', '7 years'

value = ['<=3 year', '<=3 year', '<=3 year', '<=3 year', '<=6 years'
, '<=6 years', '<=6 years', '>7 years'],inplace = True)
```

## 2.1 Data Filtration

Now we remove all but the data of interest. After structuring the dataset, we are filtering out the data which is not necessary for our data analysis. If we are analyzing a particular gender, we have to remove the genders which are irrelevant for our analysis purpose.

Diagram 2.4: Data filtration

```
#Drop All Rows with any Null/NaN/NaT Values
df = df.dropna()
```

## 2.2 Data Mining

Now our prime focus will be understanding the data and producing the desired information. We will use the analytical methodologies to process the data as a way to discern patterns or place the data in a mathematical context. This step will help us to increase the basic understanding of data before doing the representational steps.

# Chapter 4

# Result and Analysis

## 4.1 Management Information Report

### 4.1.1 Portfolio View

In the portfolio view, we dashboard according to the summary table for different demographics which represents the distribution of accounts and their balances. In the following table as an example, demographics and their different categories are mentioned:

Table 4.1: Demographics and their categories

| Category | Term | Risk Grade | Purpose | Income band | Interest rate |
|---|---|---|---|---|---|
| 1 | 36 Months | A | Debt Consolidation | < 50000 | 5%-<11% |
| 2 | 60 Months | B | Credit card | >=50000 - < 100000 | <=11%-<16% |
| 3 | | C | Home improvement | >=100000 - < 150000 | <=16%-21% |
| 4 | | D | Other | >=150000 - < 10000000 | <=21%-<26% |
| 5 | | E | Major purchase | | <=26%-<31% |
| 6 | | F | Medical | | |
| 7 | | G | Small Business | | |

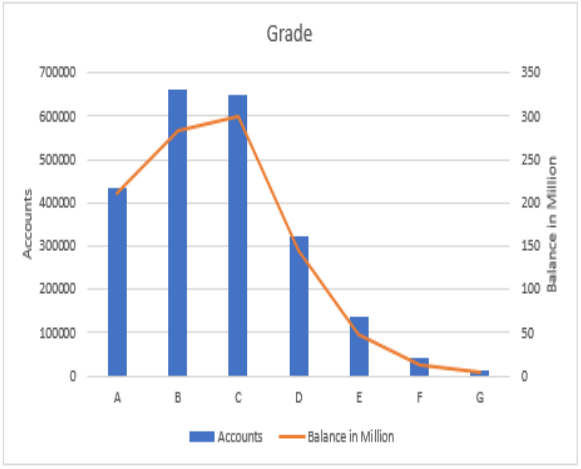
In our analysis, we have considered five demographics but here we will be presenting only one demographics for demonstration purpose. Using Python programming language, we processed the data based on Risk grade therefore following summary is generated:

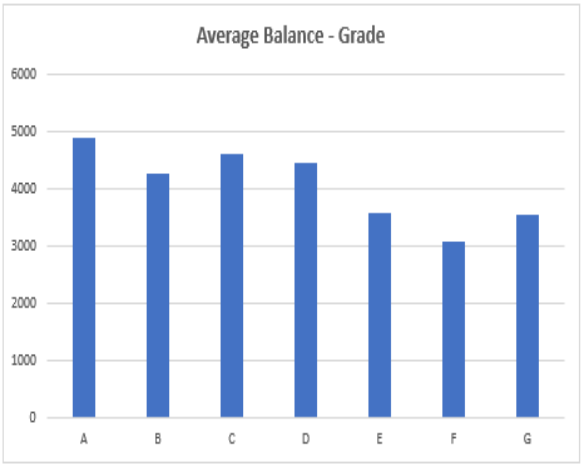
#### 4.1.1.1 Distribution of Accounts and balances by demographics

Table 4.2: Distribution of Accounts and balances of Risk grade

| *Grade* | *Accounts* | *Balance* | *Balance in Million* | *Average Balance* | *% Accounts* | *% Balance* |
|---|---|---|---|---|---|---|
| **A** | 433027 | 2,11,88,25,163 | 211.88 | 4893 | 19.2 | 21.1 |
| **B** | 663557 | 2,82,46,56,386 | 282.47 | 4257 | 29.4 | 28.1 |
| **C** | 650053 | 3,00,42,38,701 | 300.42 | 4622 | 28.8 | 29.9 |
| **D** | 324424 | 1,44,72,13,458 | 144.72 | 4461 | 14.4 | 14.4 |
| **E** | 135639 | 48,45,64,314 | 48.46 | 3572 | 6.0 | 4.8 |
| **F** | 41800 | 12,88,84,967 | 12.89 | 3083 | 1.8 | 1.3 |
| **G** | 12168 | 4,32,09,051 | 4.32 | 3551 | 0.5 | 0.4 |

Figure 4.1: Graphs of Risk Grade



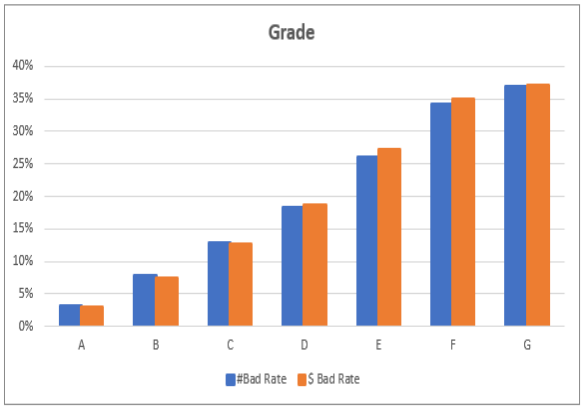

Here "A" means lower risk and "G" means higher risk. As we can see the higher number of customers and balance is in "B" and "C" risk grade and lowest in "F" and "G" risk grade. The average balance is highest in grade "A".

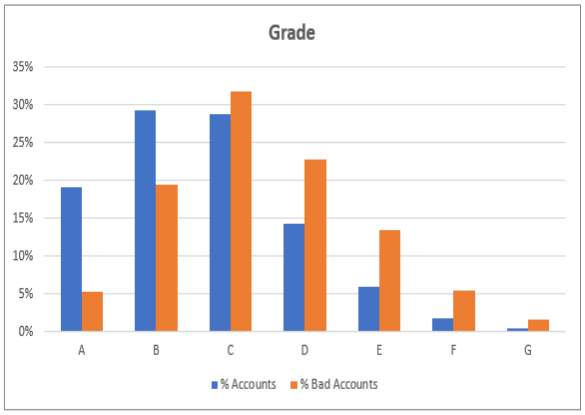
### 4.1.1.2 Distribution of Bad Accounts and Bad balances by demographics

Table 4.2: Bad rate and volume of Bad accounts of Risk grade

| Grade | Total Account | Bad Account | Total Loan Amount | Bad Loan Amount | #Bad Rate | $ Bad Rate | % Accounts | % Bad Accounts | % Loan Amount | % Bad loan Amount |
|---|---|---|---|---|---|---|---|---|---|---|
| A | 433,023 | 13,766 | 6,323,616,700 | 188,469,225 | 3 | 3 | 19 | 5 | 19 | 5 |
| B | 663,557 | 51,077 | 9,404,817,775 | 691,605,075 | 8 | 7 | 29 | 20 | 28 | 17 |
| C | 650,053 | 83,262 | 9,775,551,175 | 1,225,099,850 | 13 | 13 | 29 | 32 | 29 | 30 |
| D | 324,424 | 59,441 | 5,097,344,375 | 950,433,950 | 18 | 19 | 14 | 23 | 15 | 23 |
| SE | 135,639 | 35,364 | 2,367,318,100 | 641,854,300 | 26 | 27 | 6 | 14 | 7 | 16 |
| F | 41,800 | 14,263 | 799,410,225 | 279,259,225 | 34 | 35 | 2 | 5 | 2 | 7 |
| G | 12,168 | 4,482 | 248,032,375 | 91,746,575 | 37 | 37 | 1 | 2 | 1 | 2 |

Figure 4.2: Graphs for bad rate and percentage of bad accounts of Risk grade




In the above charts, we can see that Risk grade "G" has higher Bad accounts and a lower percentage of loan amount because "G" it is higher risk grade so it has higher Bad accounts and because of risk grade bank grants lower amount that is why it has a lower percentage of the loan amount.

### 4.1.2 Vintage View

Designed a dashboard for the Vintage view. In Vintage view we have visualized the trend of yearly acquisition, Distribution of demographics across the vintage, Closure Rate, Bad Rate, Delinquency rate across vintages and Bad Rate, Delinquency Rate across vintage by Demographics and Bureau attributes but for the demonstration purpose we will consider few of them.

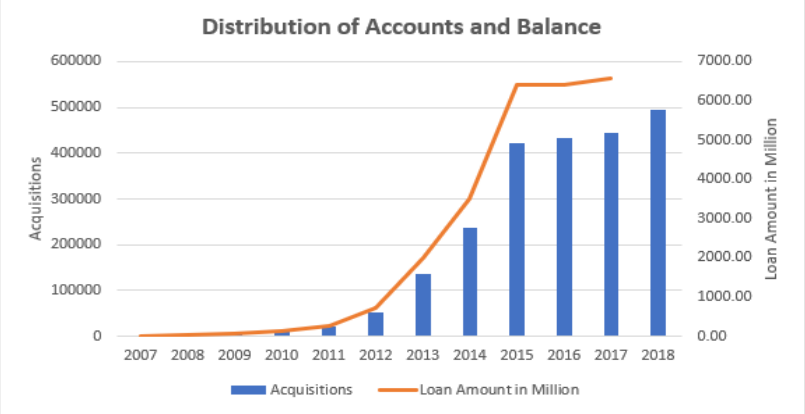
#### 4.1.2.1 Trend of yearly acquisitions

Designed a dashboard for different demographics which represents the yearly trend of acquisitions or Exposure and their loan amount for vintage View

Table 4.: Trend of Yearly acquisitions

| Year | Acquisitions | Loan Amount | Loan Amount in Million |
|---|---|---|---|
| 2007 | 603 | 4,977,475 | 4.98 |
| 2008 | 2,393 | 21,119,250 | 21.12 |
| 2009 | 5,281 | 51,928,250 | 51.93 |
| 2010 | 12,537 | 131,992,550 | 131.99 |
| 2011 | 21,721 | 261,683,825 | 261.68 |
| 2012 | 53,367 | 718,411,025 | 718.41 |
| 2013 | 134,814 | 1,982,765,275 | 1982.77 |
| 2014 | 235,629 | 3,503,840,175 | 3503.84 |
| 2015 | 421,095 | 6,417,608,175 | 6417.61 |
| 2016 | 434,407 | 6,400,569,700 | 6400.57 |
| 2017 | 443,579 | 6,584,957,075 | 6584.96 |
| 2018 | 495,242 | 7,936,263,150 | 7936.26 |

Figure 4.3: Trend of yearly acquisitions




The lending club was founded in 2007 and it's interesting to see that during the height of the Great Recession in the US, the dip in the growth of the loan issuance in 2008. Thereafter, it began growing rapidly as shown in the above chart for the volume of loans issued each year. 421095 loans were issued in 2015.

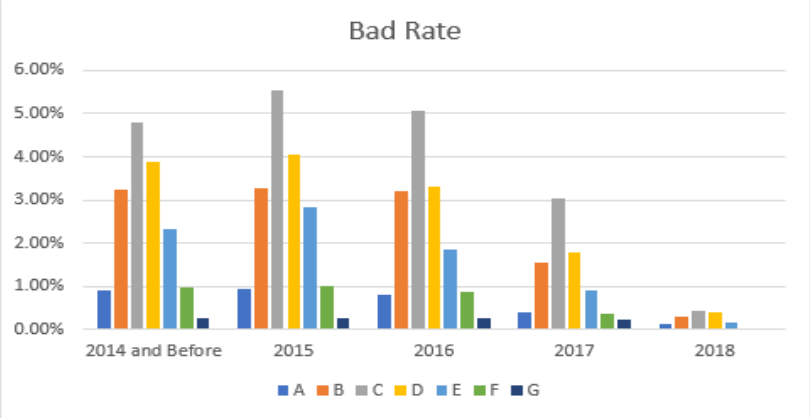
#### 4.1.2.1 Bad rate across vintage by demographics

Designed a dashboard for different demographics which represents the yearly trend of bad rate and delinquency rate

Table: Bad rate of Grade across the vintage

| Bad Rate | A | B | C | D | E | F | G |
|---|---|---|---|---|---|---|---|
| 2014 and Before | 0.89 | 3.24 | 4.80 | 3.89 | 2.31 | 0.98 | 0.27 |
| 2015 | 0.94 | 3.29 | 5.53 | 4.06 | 2.82 | 0.99 | 0.24 |
| 2016 | 0.79 | 3.21 | 5.08 | 3.30 | 1.85 | 0.86 | 0.26 |
| 2017 | 0.38 | 1.53 | 3.02 | 1.78 | 0.90 | 0.37 | 0.23 |
| 2018 | 0.10 | 0.28 | 0.43 | 0.40 | 0.14 | 0.04 | 0.01 |

Figure 4.4: Bad rate for Vintage View




## 4.2 Segment Identification for Top-up strategy

### 4.2.1 Summarization of data

For segment Identification, we have created some key performance indicators (KPI) and summarized the data according to bad rate, delinquency rate, bureau bad, and overall bad for different demographics like Debt to income (DTI) ration, Credit inquiry, Income band, Risk grade. In this analysis, we have considered the grace period so that we can understand the data deeply by including and excluding the grace period. In the following table the summarized data is provided for DTI:

Table: Representation based on different KPI of DTI

| Demographic – DTI | Delinquent (Excluding Grace Period) | Delinquent (Including Grace Period) | Bad Rate | Bureau Bad Rate | Over All Bad Rate (inc grace) | Over All Bad Rate (ex grace) |
|---|---|---|---|---|---|---|
| < 10 | 1.05 | 1.42 | 8.84 | 2.14 | 12.09 | 11.74 |
| >=10 -< 15 | 1.02 | 1.37 | 10.07 | 2.15 | 13.25 | 12.91 |
| >=15 – <20 | 1.09 | 1.47 | 11.48 | 2.06 | 14.63 | 14.26 |
| >=20 - <26 | 1.20 | 1.63 | 12.99 | 1.99 | 16.23 | 15.82 |
| >= 26 | 1.35 | 1.81 | 15.21 | 1.96 | 18.54 | 18.09 |

- **Overall Bad Rate -** This is a binary KPI that is generated by considering bad flag, delinquency flag, bureau flag. If any of these is flag hit, then the account is considered overall bad.

### 4.2.2 Waterfall Model

This model describes the exclusion and inclusion of a subset of data in a particular dataset. In our case, it explains how data is getting trimmed based on the different criteria and what is the final number of accounts on which we can apply our analysis methodology for strategy development. In the following table, it is illustrated how we excluding the subset of data from the actual dataset:

Table: Waterfall model

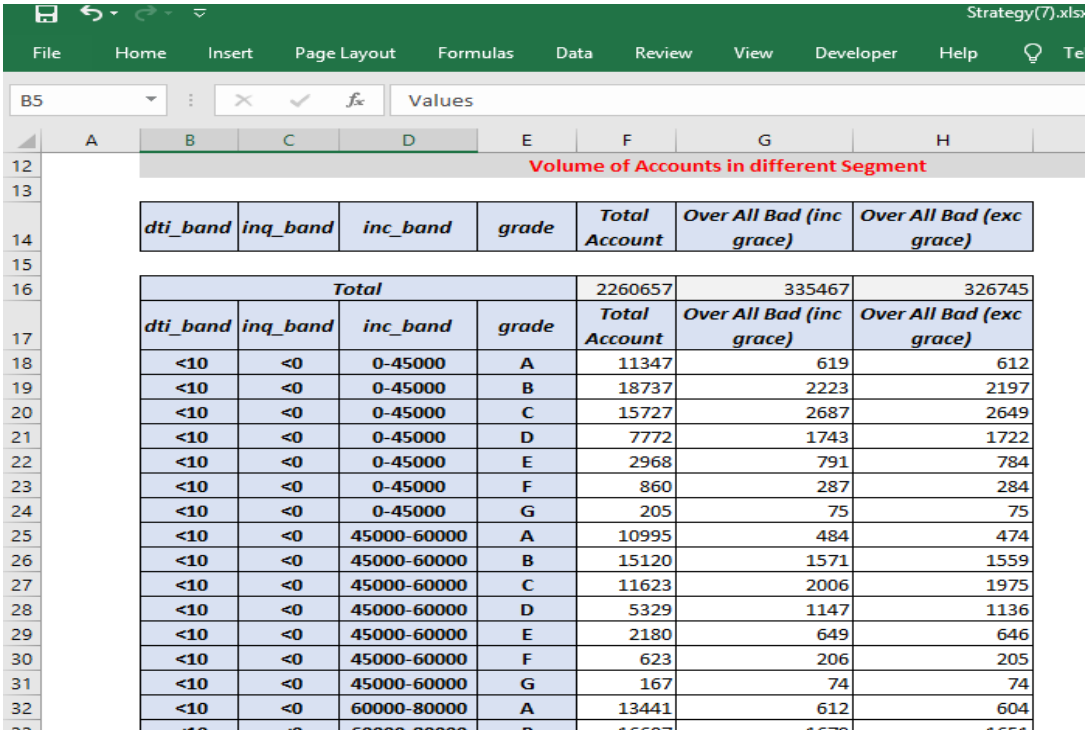
| **Values** | **Exclude** | **Include** |
|---|---|---|
| Total | - | 2,260,668 |
| Fully paid Account | 1,041,952 | 1,218,716 |
| Charged Off Accounts | 261,655 | 957,061 |
| Bureau Bad Accounts | 51,066 | 905,995 |
| last 6 Month Data | 237,318 | 668,677 |

Here we can see that the total number of accounts was 2260668 but after excluding the accounts on various parameters, we have 668677 accounts.

### 4.2.3 Risk Segmentation

In this process, we are creating a combination of different demographics based on the key point indicator so that we can recognize the segment of the customer which has good customers so that we can proceed with our strategy development process keeping that customers in mind. Below we have provided a snapshot of customer segmentation for better understanding:

Figure: Snapshot of risk segmentation



Strategy(7).xlsx

File Home Insert Page Layout Formulas Data Review View Developer Help

B5 Values

Volume of Accounts in different Segment

| dti_band | inq_band | inc_band | grade | Total Account | Over All Bad (inc grace) | Over All Bad (exc grace) |
|---|---|---|---|---|---|---|
| Total | | | | 2260657 | 335467 | 326745 |
| dti_band | inq_band | inc_band | grade | Total Account | Over All Bad (inc grace) | Over All Bad (exc grace) |
| <10 | <0 | 0-45000 | A | 11347 | 619 | 612 |
| <10 | <0 | 0-45000 | B | 18737 | 2223 | 2197 |
| <10 | <0 | 0-45000 | C | 15727 | 2687 | 2649 |
| <10 | <0 | 0-45000 | D | 7772 | 1743 | 1722 |
| <10 | <0 | 0-45000 | E | 2968 | 791 | 784 |
| <10 | <0 | 0-45000 | F | 860 | 287 | 284 |
| <10 | <0 | 0-45000 | G | 205 | 75 | 75 |
| <10 | <0 | 45000-60000 | A | 10995 | 484 | 474 |
| <10 | <0 | 45000-60000 | B | 15120 | 1571 | 1559 |
| <10 | <0 | 45000-60000 | C | 11623 | 2006 | 1975 |
| <10 | <0 | 45000-60000 | D | 5329 | 1147 | 1136 |
| <10 | <0 | 45000-60000 | E | 2180 | 649 | 646 |
| <10 | <0 | 45000-60000 | F | 623 | 206 | 205 |
| <10 | <0 | 45000-60000 | G | 167 | 74 | 74 |
| <10 | <0 | 60000-80000 | A | 13441 | 612 | 604 |

# Chapter 5

# Conclusion and future work

Creating a MI report helps in identifying for high-risk segments or potential risks in the portfolio which is very critical to understand. If any segment is a high-risk segment, we would not consider that segment for top-up loan issuance and if the segment is low risk or has a good score so we can proceed with them. Therefore, we can minimize the risk of the top-up loan and maximize the profit of the company. In the future, we will also calculate the impact of the estimation of profitability.

# Chapter 6

# Introduction

## 6.1 Overview

A financial calculator or business calculator is an electronic calculator that performs financial functions commonly needed in business and commerce communities (simple interest, interest, cash flow, amortization, conversion, cost/sell/margin, etc.). It uses standalone keys for several financial functions and calculations, making such calculations more direct than on standard calculators. It may be user-programmable, allowing the user to add functions that the manufacturer has not provided by default.

For economics or cost - sell margins some financial calculators also include the ability of graph financial calculations. One of the best uses of a financial calculator is in financial problems. Most financial calculators will be able to calculate problems for discounted cash flows, internal rates of returns, loan formulas, NPV (net present value), and markup calculations.

Whenever we propose a new strategy or modification on the existing strategy, we have to calculate the impact of a new strategy in business and submit to the business management committee (Common risk mitigation mechanism) so that the forum can understand the need for a new strategy.

## 6.2 Motivation

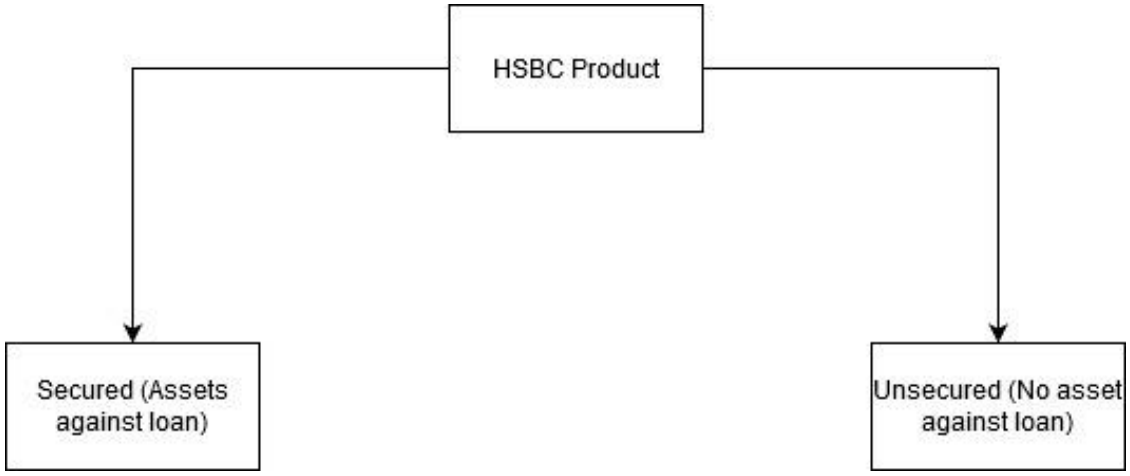
Primarily the idea has been initiated by GAC (Global Analytical Centre) credit risk team. Currently, the financial impact is calculated as an analytical extension of the project. Different projects/teams rely on internal financial KPIs and output to generate the financial benefit from the strategy change, however, there is no common methodology defined for doing the calculation of financial benefit. A detailed assessment of financial impacts is essential after the development of risk strategy and seek approval at the business management committee.

## 6.3 Bank and Its Portfolio

Bank caters to millions of customers worldwide. It offers various products and maintains a portfolio for each product offered. Some of the important products are as follows:

- Loans
- Credit Cards
- Current and Saving Account
- Mortgage

Figure 1.1: Bank Product Type



## 6.4 Personal loan

A personal loan is a type of unsecured loan and helps customers to meet their current financial needs. Customers don't usually need to pledge any security or collateral while availing a personal loan and the customer's lender provides customers with the flexibility to use the funds as per their need. It can serve the customers as a solution for managing their travel costs and wedding expenses as well as the expenses of a medical emergency, home renovation, debt consolidation, and others.

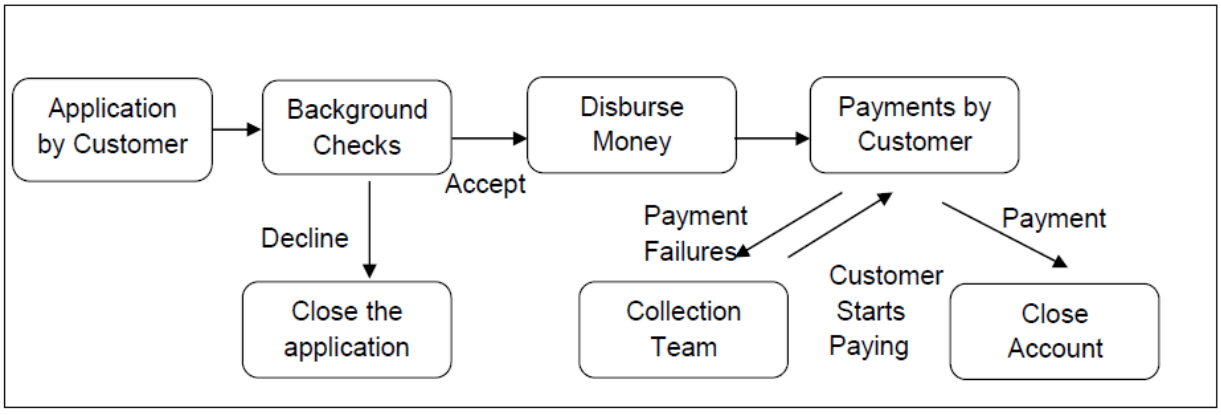
My team works for the UK region, handling the Loan portfolio. It deals with the credit risk management of loans portfolio and the development of the acquisition strategy.

## 6.5 Loan life cycle

The loan life cycle consists of the following stages:

- Application for a loan by a customer.
- Verification, background checks, and eligibility checks done by Bank.
- Disbursement of the requested amount.
- Loan repayment on fixed terms.
- Collection strategies in case of defaulters
- Closing of account on completion.

Figure 1.2: Loan life cycle



## 6.6 Data Acquisition

Acquisition data refers to the pool of information related to applications, covering the number and quality of the same. This allows the development of metrics to monitor factors including the decisions for these applications, the offered amount, the date of request, and the purpose of the loan. Monitoring of this data allows the bank to evaluate the type of applications being received by the bank and verify, realign, or develop strategies to improve the customer base to make it more profitable. This data is used to explore and look into the loan scenario from various aspects and allow us to make decisions that provide for better prospective returns.

## 6.7 Data sources

The applications for loans in the UK region come from Direct/Indirect Sources.

## 6.8 Loan Data types

The loan applications can be classified to be in any of the following segments:
The decision not given: All those applications that come and for which the bank has not given any decision, falls into this category.
Decision completed: The applications for which a decision (accept/reject/customer decline) has been given fall into this category.
Live Performing: The applications which are accepted by the bank and for which money is disbursed fall in this segment.
Closed: The applications whose tenure is over and the lent-out amount is paid back fall in this category.
Delinquent: The applications defaulting in payment fall into this category.

## 6.9 Problem statement

All users involved with strategy development required to create a business management committee submission every time, whenever a change in strategy takes place. Calculating the financial impact manually is a big hassle. The strategy team required a Financial Tool that can be used to calculate the financial impact of strategy with all useful features. To reduce long complex complications for finance-related questions the finance calculator is needed in every business. Performing all his job manually decreases the user's efficiency. An automated financial calculator was needed which can do all the intermediary jobs automatically.

# Chapter 7

# Related Work

## 9.1 General Idea

Now a day's financial calculator became very common in industries because of its usefulness, many organizations develop the generic financial calculator for a particular type of industry so the companies can customize the calculator according to their need. Some companies develop the financial calculator on their own to meet their dynamic requirement.

## 9.2 Related work in the bank

In the bank, financial impacts were calculated basis a generic approach by creating numerous tables and intermediary excel spreadsheet-based formulae. These calculations were done in silos across different teams post every strategy development. The results generated required a lot of manual inputs from the user consequently it was prone to human error. Therefore, impacting the accuracy of the financials.

The excel spreadsheet didn't have any controls on the nature of inputs given by the user. Moreover, a new user needs to give extra attention to learn the nuances of generating financials which leads to the additional requirement of time and effort. The process was taking much more time for processing the data and producing the result.

## 9.3 Our Contribution

The new approach of the financial calculator is more focused on automation, accuracy, and user experience. Previously financial impacts used to calculate with all the available data irrespective of the record's correctness but this application automatically performs the data cleansing operation. It auto-detects the incorrect or incomplete records and removes those records from the data set. A separate data set is created to keep all the dirty data so that the user can review them later. The application removes the erroneous data before analysis therefore the accuracy of impact calculation is achieved. It requires very little manual work from the user to make the user experience good.

To accomplish this solution, we used the VBA (Visual Basic for Application) programming language - Description of the technology is below. We changed the techniques of data processing so it becomes much more efficient. We enabled the user to control the application with a graphical user interface so it became a more user-centric application.

## 9.4 Technology aspects

To cater to such type of problem many types of programming languages or tools can be used which is developed by different organizations. Following are the example of programming languages and tools:

### 9.4.1 Microsoft Excel

Microsoft Excel is a spreadsheet developed by Microsoft for Android, Windows, macOS, and iOS. It features calculation, pivot tables, graphing tools, pivot tables, and a macro programming language called Visual Basic for Applications.

### 9.4.2 VBA

VBA (Visual Basic for Application) is an event-driven programming language from Microsoft which provides the facility of the macro. It is now predominantly used with Microsoft Office applications such as MS-Word, MS Excel, and MS-Access.

### 9.4.3 SAS

### 9.4.4 Python

# Chapter 8

# Proposed Solution

## 8.1 Idea

Idea is to create a simple excel based GUI for calculating the financial impacts of a strategy change, given application will be independent of the existing system, no implementation challenge is expected. It automates the process of financial calculation. It minimizes the processing time so that the user can utilize his important time on other important aspects of the business.

## 8.2 Workflow

Once the user places data in the application and clicks the button for calculating the impacts, it automatically cleans data and separates the garbage data in a particular sheet, given the option to the user to review the garbage data. It does not require to validate the data and define the formula every time manually. The application would enable the user to calculate all the relevant financial insights or summary in one – go with minimum attention of the user. It takes minimum time to do all processing.

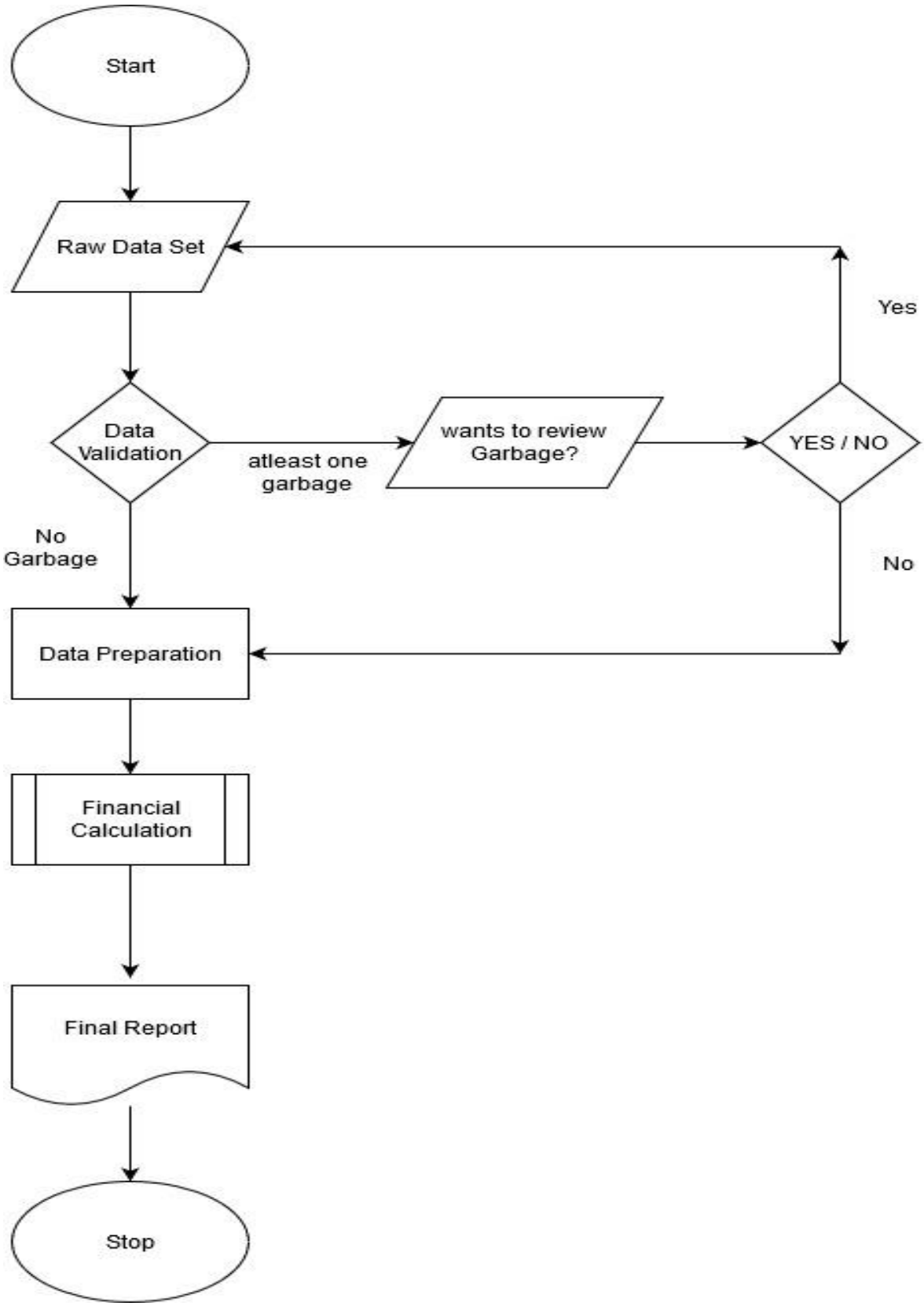



Figure 2.1: Application Flowchart

## 8.3 Key areas

Tool drives consistency of calculations across all products in strategy change and helps avoid errors or localized assumptions. At the different stages of data processing we applied the following techniques to make it efficient and accurate:

### 8.3.1 Data Validation

Data validation helps to ensure data quality. It confirms that data has been undergone a thorough cleaning process. Data quality plays a major role in the accuracy of the analysis.

### 8.3.2 Optimized formulas

Financial analysis requires a lot of calculations based on formulas. Hence, simplification of formula can play a vital role in memory and processing time conservation. After applying simplified formulas, we were able to save a margin of system memory and processing time.

### 8.3.3 Macro

Developed a Macro which reduced the possibility of human error that increases with many, repetitive keystrokes and tasks. It also reduced the amount of time that must be spent performing basic computing tasks, freeing users up for idea-generating, and more complex problem-solving activities.

### 8.3.4 Code Structure

The creation of the functions for repetitive work contributed to reducing the processing time. We focused on decreasing the interaction between VBA macro code and the spreadsheet as a result we got our results in very little time.

### 8.3.5 GUI

The GUI (Graphical User Interface) is a form of user interface which allows end-users to interact with the application through icons in the graphical format and an audio indicator such as primary notation, instead of text-based user interfaces, text navigation or typed command labels.

### 8.3.6 Accuracy

By reducing the human job, we achieved accuracy or reduced human error.

# Chapter 9

# Results and Analysis

## 9.1 Input Dataset

Raw data consists of the details of the applicant (customer) who have applied for a loan or any other product. Raw data should have defined attributes. We have applied a methodology to validate the raw data and to transform the raw data into the input dataset so that it can readily and accurately be analyzed in the context of our desired business. If a customer application doesn't pass the validation criteria it will be separated from the input dataset. Input dataset attributes and their formats:

Table 9.1: Input data variables and their criteria

| Variable | Variable Types | Variable Size |
|---|---|---|
| Application Number | Numeric \| | 10<n<16 |
| Application Score | Numeric | N/A |
| Offer APR | Numeric | N/A |
| Term | Numeric | N/A |
| Request Amount | Numeric | N/A |
| Initial Decision | Character | 1 |
| Final Decision | Character | 1 |
| Gross Available offer | Numeric | N/A |
| Date of Application | Date | N/A |

## 9.2 Result

After data preparation, it comes to complex financial calculations. To calculate the impact of strategy change, basis bank's internal financial KPIs financial impacts are generated. The calculation of the internal financial KPIs are programmed within the VBA code, which allows the user to have a one-click experience to create a financial summary from a strategy change:

Table: 9.2: Impact matrix

| **Per Account** | **Current** | **Proposed** | **Difference** |
|---|---|---|---|
| ROE | | | |
| RORWA | | | |
| ALR | | | |

| **Per Account (Month)** | **Current%** | **Proposed%** | **Difference%** |
|---|---|---|---|
| Total Income | | | |
| Total Losses (Excluding Fraud) | | | |
| EP (Economic Profit) | | | |
| FTE (Full-time employee) Saving | | | |
| **Total** | | | |
| **Per Account (Annual)** | **Current $** | **Proposed $** | **Difference $** |
| Total Income | | | |
| Total Losses (Excluding Fraud) | | | |
| EP (Economic Profit) | | | |
| FTE (Full-time employee) Saving | | | |
| **Total** | | | |

According to the above summary table, one can understand the impact of strategy change, if any cell in the difference column shows negative value then the strategy is not viable else some good impacts can be seen.

- **ROE (Return on Equity)**

  In business it is a parameter of the profitability concerning the equity because equity of shareholder can be calculated as a return on assets minus liabilities, ROE can also be thought of as taking all assets and subtracting all liabilities.

- **RORWA (Return on Risk-weighted assets)**

  It is a measure of profit per unit of risk. Return on risk-weighted assets generally measured as profit before tax as a percentage of risk-weighted assets.

- **ALR (Annual loss rate)**

  The loss can be beard by the organization it is also known as acceptable damage. This is the loss rate that can be accepted by the company.

- **EP (Economic Profit)**

  An economic profit is a difference between the revenue received from the sales of output and the costs of all inputs used as well as any opportunity costs. While calculating economic profit, opportunity costs and explicit costs are deducted from revenues earned by the company.

- **FTE (Full-Time equivalent) savings**

  The calculation of full-time equivalent (FTE) is an employee's scheduled hours divided by the employer's hours for a full-time workweek. When an employer has a 40-hour workweek, employees who are scheduled to work 40 hours per week are 1.0 FTEs. Employees scheduled to work 20 hours per week are 0.5 FTEs.

In the last two table analysis is same but the presentation of data is different. In the second table of the impact matrix, the impacts are presented monthly, and in the third table, impacts are shown in the form of an annual basis.

# Chapter 10

# Conclusion and future work

The bank is taking a step towards the area of advanced cross-product financial calculator and this project helps to realize the potential of different languages or tools such that the data analyst, stakeholders, as well as end-users, gets the data as a story to find out the impacts and insights.

Microsoft Excel and VBA can perform the analysis task and automation. Bank has the license of Microsoft excel hence we can expect the commercial deployment of this tool in the bank to understand the impact of strategy modification or change. This application can also be applied across the product by changing the KPI's parameter according to the different products.

In the future, more tools and languages for advanced financial analysis will be explored such that to learn and understand the new analysis technique and thus trying to analyze the effect of strategy change. We can incorporate this application in the cloud so that the availability and computing power of application can be increased.

# Bibliography


**[1]** **https://support.office.com/en-us/excel**

**[2]** **https://www.wikipedia.org/**

**[3]** **https://www.tutorialspoint.com/index.htm**

**[4]** **https://www.kaggle.com/wendykan/lending-club-loan-data**

**[5]** **https://rpubs.com/culight/310410**

**[6]** **https://docs.python.org/3/tutorial/**

**[7]** **https://colab.research.google.com/notebooks/intro.ipynb**