

Assignment

Personal loan data analysis – Exploratory data analysis

Problem Statement (Situation):

“Finding out potential target customers for personal loans”

In this case study, the goal of HBFC bank is to sell more personal loans to their savings account holders. The bank wants to start a campaign to sell the personal loans, but before that they want to analyze last marketing campaign data to understand the profile of potential loan customers. This will help them in doing a targeted approach to the prospective customers in future. The bank has approached you, to help them with the analysis of the previous campaign data. The bank basically has two customers,

- **Liability customers** – They deposit the money in the bank and pays interest against the deposited money.
- **Asset customers** - They borrow money (take different types of loan) from the bank and the bank charges interest against the borrowed money.

At present the bank has small number of asset/loan customers. The bank wants to increase their income by increasing the customer base of “asset customers”. Last year the bank ran a campaign where they successfully converted 9% of the existing “liability customers” to “asset customers”. This has encouraged the bank to have a better targeted marketing campaign to Increase the success ratio with minimum budget.

You would receive a file Bank.xls which contains data of 5000 customers. It includes,

- Customer demographic information (age, income, etc.),
- Relationship with the bank (mortgage, securities account, etc.)
- Customer response to the last personal loan campaign (Personal Loan).
- Among the 5000 customers, only 480 (9.6%) customers have accepted the personal loan in the last year campaign.

Following is the description of the columns in excel file.

ID	Customer ID
Age	Customer's age in years
Experience	Years of professional experience
Income	Annual income of the customer (\$000)
ZIPCode	Home Address ZIP code.
Family	Family size of the customer
CCAvg	Avg. spending on credit cards per month (\$000)
Education	Education Level. 1: Undergrad; 2: Graduate; 3: Advanced/Professional
Mortgage	Value of house mortgage if any. (\$000)
Personal Loan	Did this customer accept the personal loan offered in the last campaign?
Securities Account	Does the customer have a securities account with the bank?

TD Account	Does the customer have a Term deposit (Including Fixed and Recurring Deposits) account with the bank?
Online	Does the customer use internet banking facilities?
CreditCard	Does the customer use a credit card issued by the bank?
Income Categorical	Column Created as an example

Additionally, a column Income Categorical has already been created for your reference.

Objective (Task):

- As a consultant, you must perform preliminary data analysis (EDA) and visualization to understand the profile of customers having savings account, who took personal loan in the last marketing campaign VS customers who didn't take it up.
- Using EDA identify profile of customers whom bank can target for selling personal

Note: - Read the column descriptions and ensure you understand each attribute well before starting the analysis.

In order to do this analysis, you are expected to solve these questions:

- What percentage of the bank's customers (according to the data) have availed Personal Loans vs the ones who have not availed it?
- Generate a table with min, max, median & average for all numeric variables (age, experience, income, family members, CCAvg, Mortgage)
- Create a new categorical variable for Experience using 4 categories –
 - 0 to 10 years
 - 11 to 20 years
 - 21 to 30 years
 - 30+ years.

Plot a bar graph for this new categorical variable

[Hint – You may make use of if else/nested if statements to accomplish this task. You can refer how Income_Category has been created in the dataset]

- Create a scatter plot of the Age and the Experience variable. What do you observe?
- What are the top 3 areas (ZIP Codes) where the bank's customers are located?
- How many customers have a combination of Fixed Deposits and Credit Cards but not Personal Loan?
- What is the median income of the customers who have availed personal loans and compare it with the median income of those customers who have not availed personal loans? What do you infer?
- Create 4 separate Pivot Tables. Summarize your data by percentage values.
 - Education vs Personal Loan
 - TD Account Vs Personal Loan
 - Online vs Personal Loan
 - Income_Category vs Personal Loan

[Hint: Please drag Personal Loan to the Columns area while creating the Pivot Table to get the required values]

- 9) Analyze the Pivot tables created in the previous question and state any anomaly that you observe. Which categorical variables appear most important for your further study if you want to analyze which customers are most likely to take personal loans and why?
- 10) In the last campaign, bank reached out to 5000 customers out of which 480 customers accepted the personal loan offer. The bank incurred a huge cost in running a marketing campaign to reach out to so many customers. This is where you as a strategic business consultant step in. You are tasked to optimize the cost of this campaign by identifying the correct target base (without significant reduction in number of acceptances of offers). The bank can then send Personal Loan offers to these target customers who have a higher chance of accepting the offer. Based on your analysis, what strategy would you suggest to the management of HBFC bank?

Learning Outcome (Result):

- Understand implementation of Exploratory Data Analysis & Data Visualization to understand the nature of different data-attributes
- The learners will understand how to use various tools like pivot tables, charts, basic statistical/mathematical/analytical tools in MS Excel.