HBFC Bank Personal Loans

Arun Poochaiyan R

Great Learning

GLCA DA Online August 2023

Note

This contains the Excel Exploratory Data Analysis of the HBFC Bank Personal Loans

Abstract

In this Data Analysis Project, a comprehensive dataset of about 5000 existing customers’ details from the HBFC Bank were given. To refine its Personal Loan outreach, this dataset was taken into consideration and the analysis were done. This dataset consists of multiple factors, including not only demographic details such as age and income but also insights into each customer's existing relationship with the bank, denoted by factors like existing mortgages and accounts. Crucially, the dataset highlighted customer responses to the bank's previous personal loan campaign. As an external consultant, the primary objective was to analyze the data and finding the underlying patterns and correlations. Preliminary analysis indicated that certain factors and pre-existing bank relationships had a notable influence on a customer's inclination towards personal loans. The insights extracted from this exploration are suggested to empower the bank. By targeting individuals who consists of a higher chance for loan acquisition, based on the identified parameters, the bank can optimize its marketing strategies, ensuring both efficiency and efficacy in future loan campaigns. This targeted approach not only promises reduction of the total spent on the campaign, if might be held by next time but also provides more personalized customer banking experience.

HBFC BANK PERSONAL LOAN

Introduction:

This report consists of the Exploratory Data Analysis for the given dataset from HBFC Bank. The main aspect of the bank is to drill possible insights from given data so that the bank could leverage the insights to provide the personal loans to the appropriate customers and reducing the costs spent on the campaigns. Below mentioned, consisting inferences, methodology used, graphs, tables are the possible insights derived by Exploratory Data Analysis using Excel.

**Question 1**: What percentage of the bank’s customers (according to the data) have availed Personal Loans?

**Pivot Table**:





**Inference**: The overall percentage of the customers who availed the personal loan from the bank is 9.60%.

**Question** **2**: Generate a table with min, max, median & average for all numeric variables (age, experience, income, family members, CCAvg, Mortgage). What are your observations?

**Table**:



The above table consists of minimum, maximum, median, and average for all numeric variables.

**Observations**:

1.**Age**: The bank may focus on the customers who is between the age of 40 to 50 or the average age, as the needs for this group of customers might be high. And the chances of taking loan from the bank will also be high.

2.**Income**: People with average income can be targeted by the bank as they might be having more needs than the people who get high income. But the bank should also provide loans based on the income of the customer.

3.**Family** **Members**: The bank can also consider the total count of the customer’s family, as it one of the crucial factors of availing loan.

4.**Mortgage**: Bank should be avoiding the customer who has high mortgage as it may affect the loan repayment by those customers.

**Question 3:** Create a new categorical variable for Experience using 4 categories –

a. 0 to 10 years

b. 11 to 20 years

c. 21 to 30 years and

d. 30+ years.

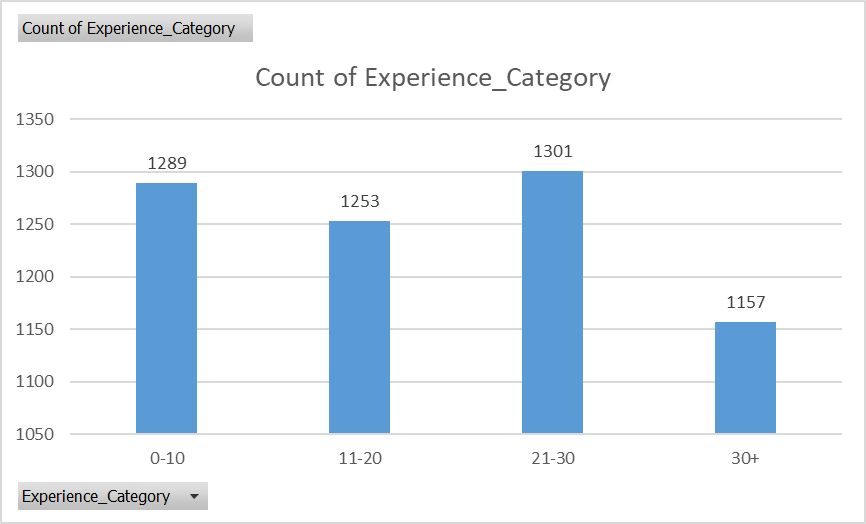
Plot a bar graph for this new categorical variable

**Pivot Table:**





**Graph**:

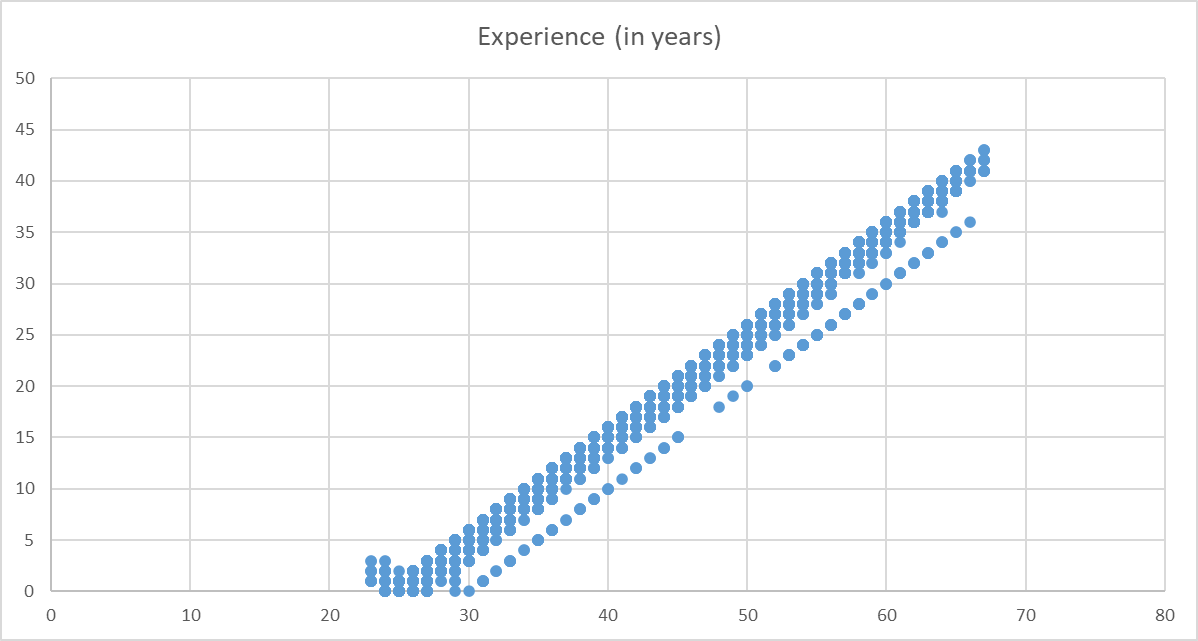


**Inference**: According the experience category data and graph, there were a greater number of people from the section 21-30 experience.

**Question** **4**: Create a scatter plot of the Age and the Experience variable. What do you observe?

**Table/Pivot Table:** No

**Graph**:



**Observation**: We could see that the Age is in accordance with the experience. That is, the customers have significant years of experience to that of their age.

**Question 5:** What are the top 3 areas (ZIP Codes) where the bank’s customers are located?

**Pivot Table:**



**Inference:** The top 3 areas (ZIP Codes) where the bank’s customers are located were

1. 94720 - 169 customers
2. 94305 – 127 customers
3. 95616 – 116 customers

**Question 6:** How many customers have a combination of Fixed Deposits and Credit Cards but not Personal Loan?

**Pivot Table:**





Count of (FD&CC)! PL = Count of the customers who has both Fixed deposits and Credit Card but not Personal Loan.

**Inference:** There were totally 147 customers who had both Fixed Deposits and Credit Cards but not availed Personal Loans.

**Question 7:** 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?

**Table**:





**Inference**: We could clearly see that customers with higher median of income availed personal loan than that of the customer who has lower median of income.

**Question 8:** Create 4 separate Pivot Tables. Summarise your data by percentages.

* Education vs Personal Loan
* TD Account Vs Personal Loan
* Online vs Personal Loan
* Income\_Category vs Personal Loan

**Pivot Table:**

1.Education vs Personal Loan





2.TD Account Vs Personal Loan





3.Online vs Personal Loan





4.Income\_Category vs Personal Loan





**Question 9:** Analyse 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 analyse which customers are most likely to take personal loans and why?

**Inference**: Among the four comparisons, the **Income\_Category vs Personal Loan** can be considered. We could clearly see that 91.25% of the customers who availed personal loan come under the income\_category of above $100000 per year, which is crucial instance that can be made use of, by the bank, to extend the personal loan to its customers in the upcoming loan campaign and it also improves the customer banking experience by targeting the specific audience who has higher chances of getting loan.

**Question 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 optimise 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?



The above picture is derived by considering income categorical, personal loan and experience. It tells clearly that the greatest number of people who have availed personal loan come under the category of $100K+ per year income category of number 438(91.25%) and followed by less number in 51K– 100K per year of income category of number 42 (8.75%). And further we could also see that, regardless of Experience, people who have higher income i.e., $100K+ per year has higher chance of getting personal loans from the bank.

**Conclusion:** So, the HBFC bank can first focus more on customer whose income is greater than $100K+ per year and then the bank can focus on people whose income is between $51K – $100K per year, thereby reducing the cost of campaign by focusing on the appropriate audience. This would not only focus on the right audience but also increases customer banking experience and reduces the work of the employees.

Thank You!