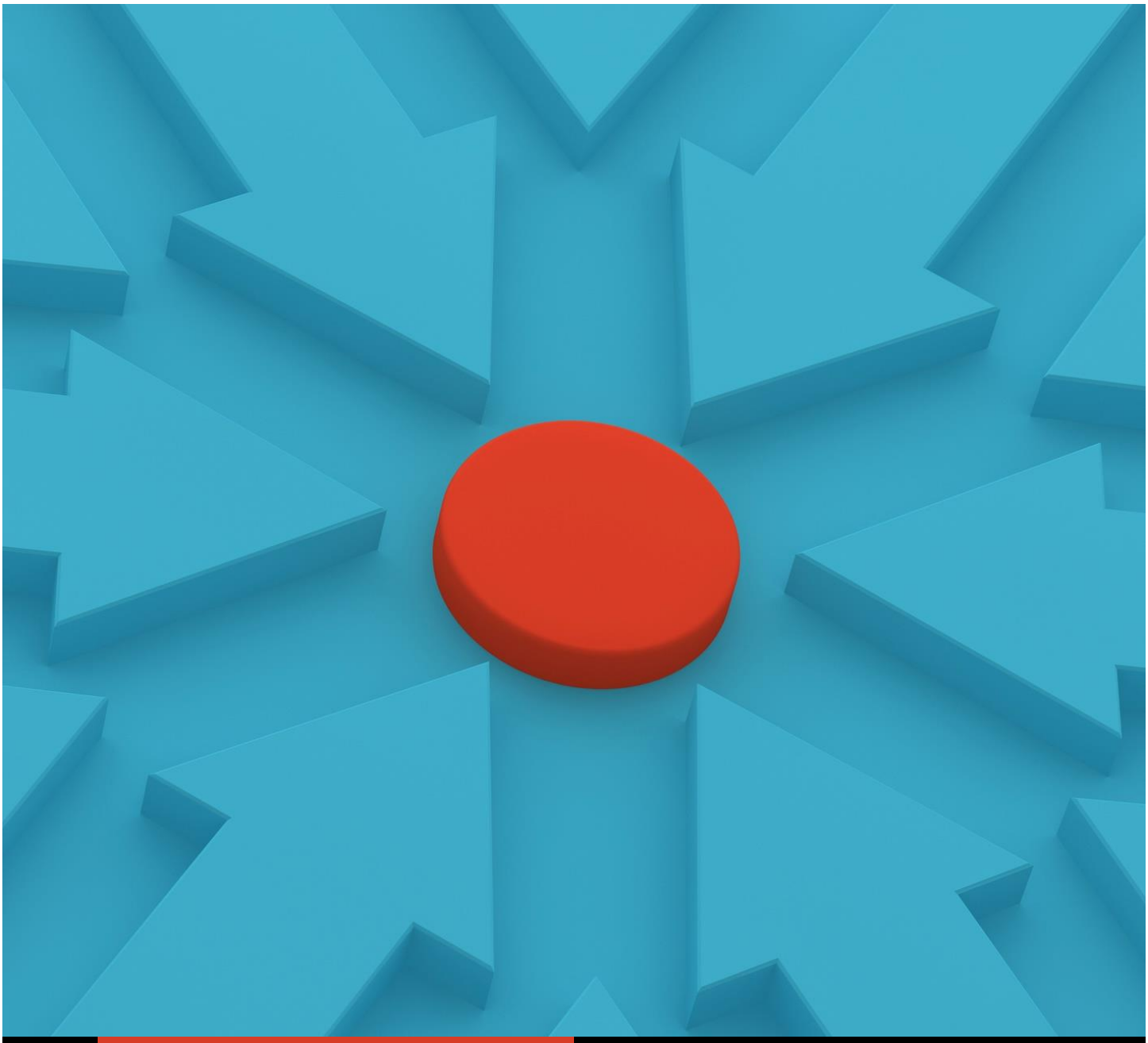


HBFC BANK PERSONAL LOANS

2023

Great Learning

Authored by: Nutan Parab

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

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

DATA DICTIONARY:

ID	Customer ID
Age	Customer's age in years
Experience	Years of professional experience
Income	Annual income of the customer ($000)
Zip Code	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?
Credit Card	Does the customer use a credit card issued by the bank?
Income Categorical	Column Created as an example

1. What percentage of the bank's customers (according to the data) have availed Personal Loans vs the ones who have not availed it?

Row Labels	Count of Personal Loan
No	90.40%
Yes	9.60%
Grand Total	**100.00%**

- We can see that approximately 9.60% of the bank's customers have availed personal loans, while about 90.40% have not availed them. This gives us an idea of the percentage distribution between customers who have and have not availed personal loans.

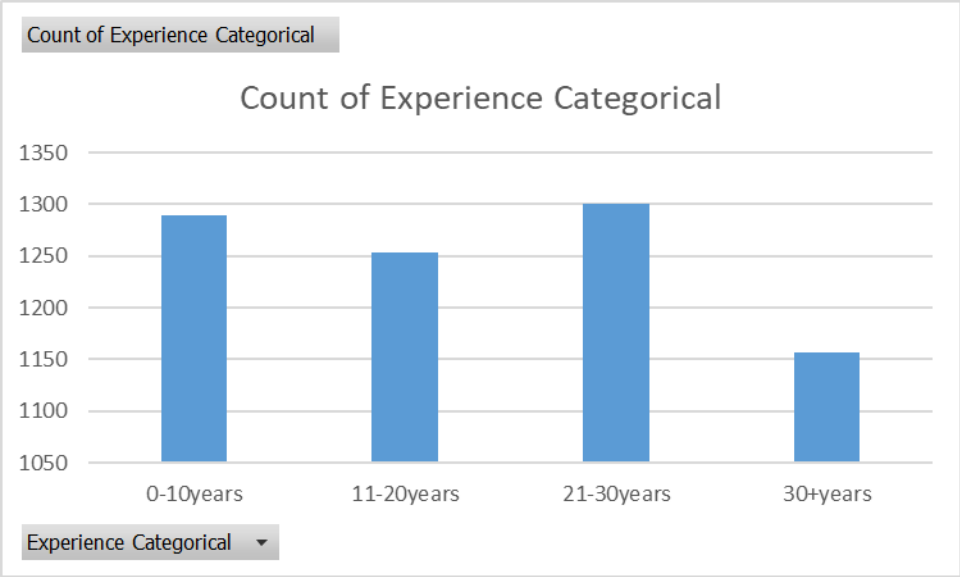
2. Generate a table with min, max, median & average for all numeric variables (age, experience, income, family members, CCAvg, Mortgage)

Column1	Age (in years)	Experience (in years)	Income (in K/year)	Family members	CCAvg	Mortgage
Min	23	0	8	1	0	0
Max	67	43	224	4	10	635
Median	45	20	64	2	1.5	0
Average	45.3384	20.1348	73.7742	2.370572207	1.93794	56.4988

- Age ranges from 23 to 67, with an average of around 45.
- Experience ranges from 0 to 43 years, with an average of around 20.
- Income ranges from 8,000 K/year to 224,000 K/year, with an average of around 73,000 K/year.
- Family members range from 1 to 4, with an average of around 2.
- CCAvg ranges from 0 to 10, with an average of around 1.9.
- Mortgage ranges from 0 to 635, with an average of around 56,000.

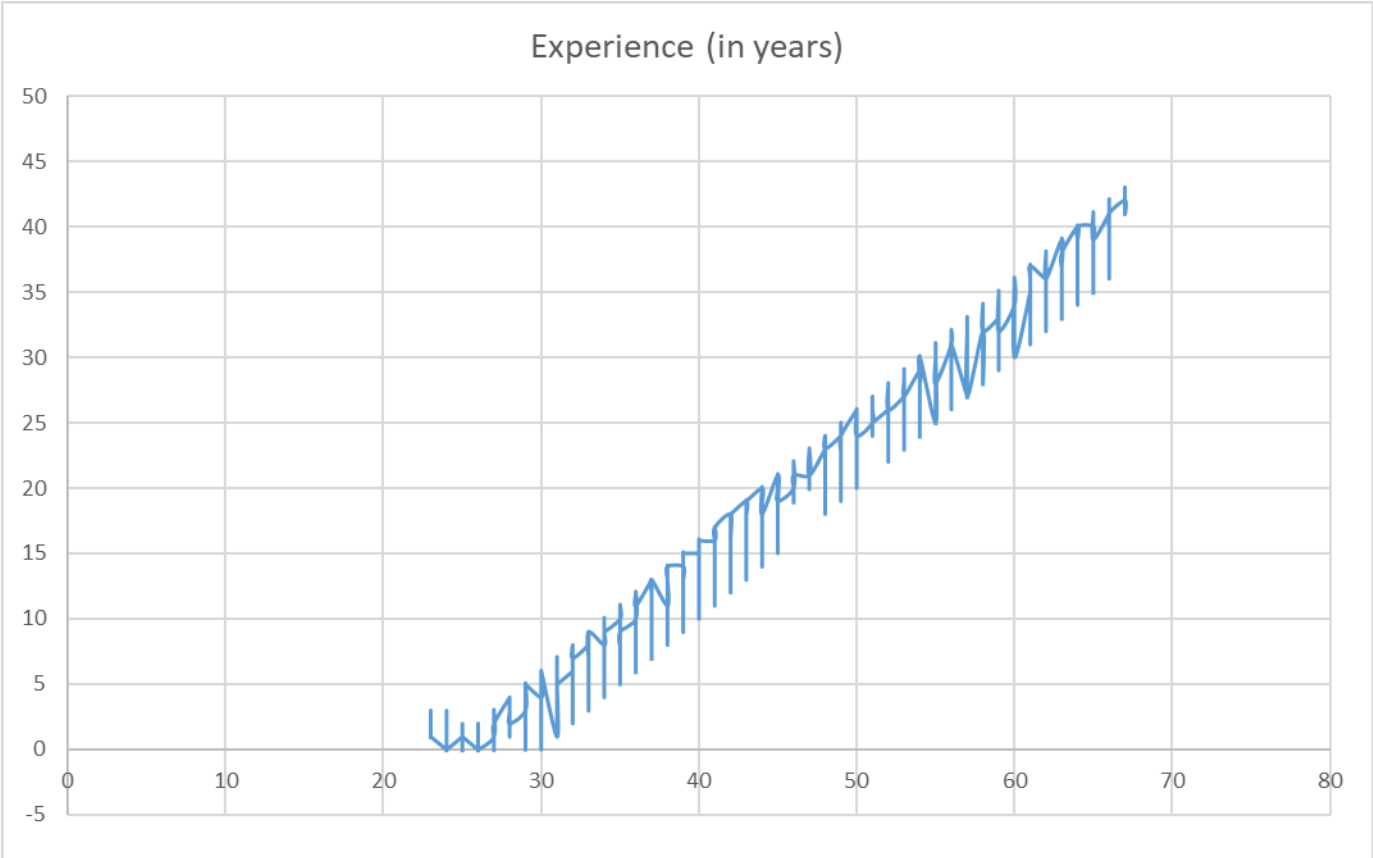
3. Create a new categorical variable for Experience using 4 categories –
• 0 to 10 years
• 11 to 20 years
• 21 to 30 years
• 30+ years.
Plot a bar graph for this new categorical variable
[Hint – You may make use of if else/nested if statements to accomplish this task. You can refer how Income_Category has been created in the dataset]

Experience Categorical	Count of Experience Categorical
0-10years	1289
11-20years	1253
21-30years	1301
30+years	1157



- Majority of the individuals in the dataset have experience in the range of 21-30 years, with a count of 1301.
- The other categories also have significant counts: 0-10 years (1289), 11-20 years (1253), and 30+ years (1157).
- This distribution gives us an idea of the experience levels among the individuals in the dataset

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



From the above graph we can infer,

- Age is directly proportional to experience
- As age increases, experience also increases

5. What are the top 3 areas (ZIP Codes) where the bank's customers are located?

Zip Code	Count of ZIP Code
94720	169
94305	127
95616	116

From the given dataset, majority number of customers of bank are located in following Zip Codes :

- 94720
- 94305
- 95616

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

Row Labels	Count of TD Account	Count of Credit Card	Count of Personal Loan
Yes	**147**	**147**	**147**
Yes	**147**	**147**	**147**
No	147	147	147
Grand Total	**147**	**147**	**147**

There are a total of 147 customers that have combination of both Fixed Deposits and Credit Cards but no personal loans.

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?

Median for people who took loan	Median for people who did not took loan
142.5	59

- The median for people who took loan is 142.5 and median for people who did not took loan is 59.
- According to the records, people who have less income did not took the loan whereas people who have high income have taken loan maybe as an investment.

8. 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]

Education vs Personal Loan			
Count of Personal Loan	**Column Labels**		
Row Labels	**No**	**Yes**	**Grand Total**
Graduate	24.42%	3.64%	28.06%
Professional	25.92%	4.10%	30.02%
Undergraduate	40.06%	1.86%	41.92%
Grand Total	**90.40%**	**9.60%**	**100.00%**

We can see distribution of personal loans based on education level. Here are the percentages:

- Among graduates, 24.42% have not taken a personal loan, while 3.64% have.
- For professionals, 25.92% have not taken a personal loan, while 4.10% have.
- Among undergraduates, the majority (40.06%) have not taken a personal loan, while only 1.86% have.

Online vs Personal Loan			
Count of Personal Loan	**Column Labels**		
Row Labels	**No**	**Yes**	**Grand Total**
No	36.54%	3.78%	40.32%
Yes	53.86%	5.82%	59.68%
Grand Total	**90.40%**	**9.60%**	**100.00%**

- Majority of people, around 53.86%, have taken personal loans, while 5.82% have taken online loans.
- On the other hand, about 36.54% have not taken any loans, and 3.78% have taken personal loans but not online loans.
- These numbers give us an idea of the distribution of loan types.

TD Account Vs Personal Loan			
Count of Personal Loan	**Column Labels**		
Row Labels	**No**	**Yes**	**Grand Total**
No	87.16%	6.80%	93.96%
Yes	3.24%	2.80%	6.04%
Grand Total	**90.40%**	**9.60%**	**100.00%**

- The majority, around 87.16%, do not have a personal loan or a TD account. About 6.80% have a TD account but no personal loan, while 2.80% have both a TD account and a personal loan.
- These numbers give us an idea of the distribution between TD accounts and personal loans

Income_Category vs Personal Loan			
Count of Personal Loan	**Column Labels**		
Row Labels	**No**	**Yes**	**Grand Total**
0-50	38.28%	0.00%	38.28%
100+	15.48%	8.76%	24.24%
51-100	36.64%	0.84%	37.48%
Grand Total	**90.40%**	**9.60%**	**100.00%**

- An income of 0-50, 38.28% do not have a personal loan, while none of them have a personal loan.
- For the income category of 100+, 15.48% do not have a personal loan, while 8.76% do have a personal loan.
- In the 51-100 income category, 36.64% do not have a personal loan, while 0.84% do have a personal loan.
- These numbers give us an idea of the distribution of personal loans across different income categories.

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?

- Professionals have a higher count of "Yes" for personal loans compared to other categories. This could indicate that professionals have specific financial needs or circumstances that make them more inclined to take personal loans. Further analysis could explore factors such as income levels, career stability, or specific loan requirements for professionals.
- Moving on to the second pivot table, we noticed that the count of "Yes" for personal loans among TD account holders is unexpectedly lower. This suggests that TD account holders may have alternative financial options or may be less inclined to take personal loans due to their existing banking relationship. Investigating factors like interest rates, loan offerings, or customer loyalty programs for TD account holders could provide insights into this pattern.
- To determine the most important categorical variables for further study, we should consider the variables that show significant differences in the

likelihood of taking personal loans. In this case, education level and TD account status stood out as potential influential factors. Analyzing these variables can help us understand why professionals and TD account holders have different tendencies when it comes to personal loans.

- Income also plays crucial role in loans as various factors such as amount for loan, interest depends on it.

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?

- Conduct a comprehensive analysis: Evaluate the data of the 480 customers who accepted the personal loan offer. Look for common characteristics, such as demographics, income levels, employment types, or education levels. This analysis will help identify the key traits of customers who are more likely to accept the offer.
- Build a targeted customer profile: Based on the analysis, create a detailed profile of the ideal customer who is more likely to accept the personal loan offer. This profile will act as a guide in identifying the target base for future campaigns.
- Refine the customer selection process: Use the customer profile to refine the selection process for future campaigns. Leverage data-driven techniques like predictive modeling or machine learning algorithms to identify potential customers with a higher chance of accepting the offer. This will help optimize the cost by focusing on a more targeted customer base.
- Personalize marketing communication: Tailor the marketing communication, including the messaging and channels used, to resonate with the identified target customers. Personalization can increase engagement and the likelihood of acceptance.
- Monitor and iterate: Continuously monitor the results of the campaigns and analyze the acceptance rates. Making adjustments is as necessary based on the

- feedback and insights gained. This iterative approach will help refine the targeting strategy over time.
- By implementing these strategies, HBFC bank can optimize the cost of the campaign while improving the acceptance rate of personal loan offers. It's all about finding the right balance between cost optimization and maximizing customer acceptance.

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.