HBFC Bank Analysis

Name: Nagesh S R
Data Analyst

INRODUCTION:

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 analyse 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.

Data description:

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

OBJECTIVE:

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

ANS:

Personal Loan	Count of Personal Loan
No	4520
Yes	480
Grand Total	**5000**

Total Customer's have not availed Personal Laon **:4520**
Total Customer's have availed Personal Loan **:480**
Total Customers' **: 5000**

When we convert to Percentage it will look like bellows table

Personal Loan	Count of Personal Loan
No	90.4%
Yes	9.6%
Grand Total	**100.00%**

Percentage of customers have not availed for Personal Loan **: 90.4%**
Percentage of customers have availed for Personal Loan **: 9.6%**
Total Percentage **: 100%**

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

ANS:

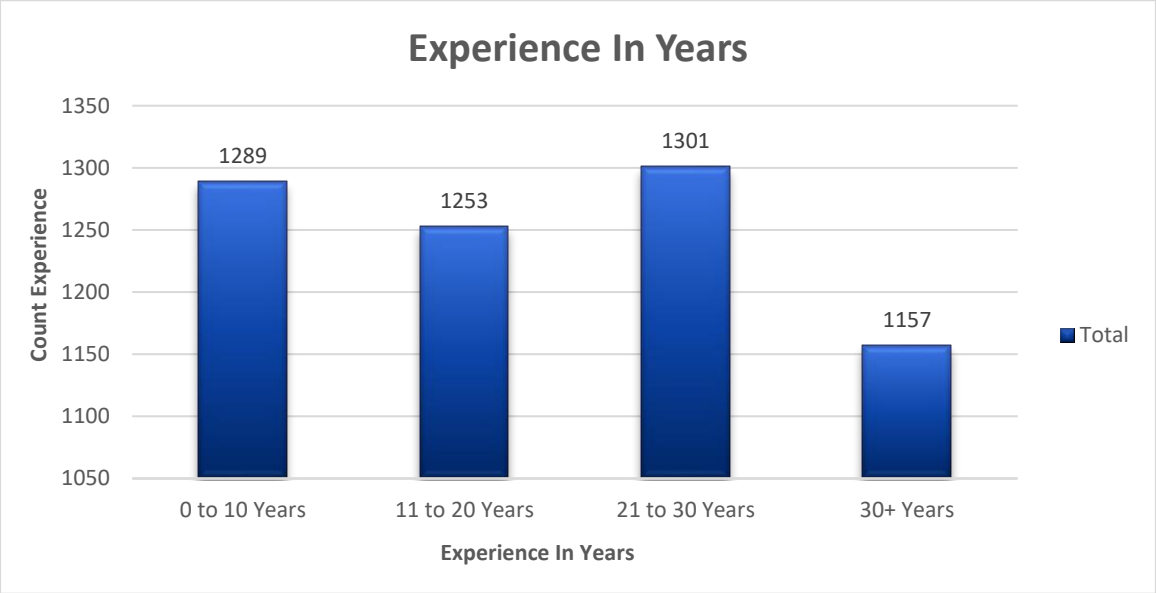
Criteria	Age (in years)	Experience (in years)	Income (in K/year)	Family_Members	CCAvg	Mortgage
Min	23	0	8	1	0	0
Max	67	43	224	4	10	635
Median	45	20	64	2	1.5	0
Average	45.3	20.1	73.8	2.4	1.9	56.5

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.

ANS:

Experience	Count of Experience Categorical
0 to 10 Years	1289
11 to 20 Years	1253
21 to 30 Years	1301
30+ Years	1157
Grand Total	**5000**

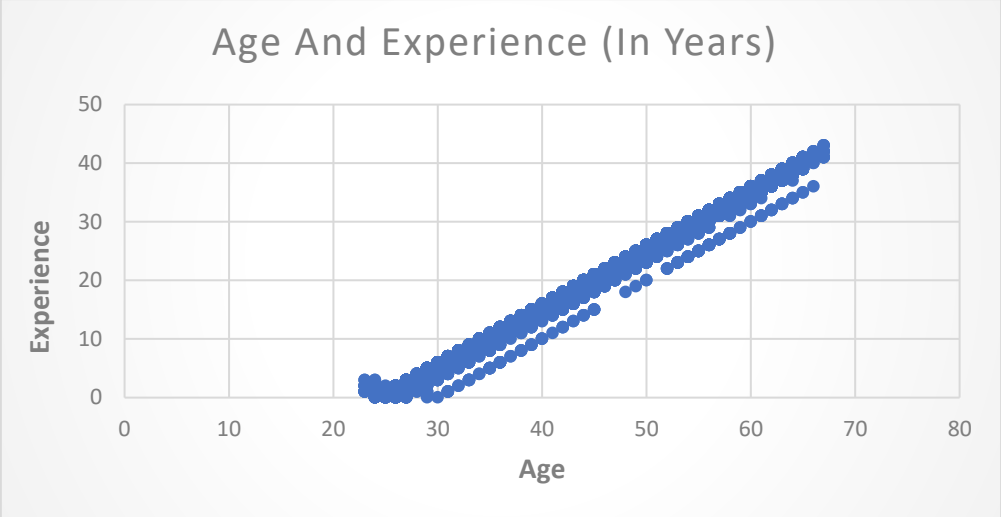


=IF([@[Experience (in years)]]<=10,"0 to 10 Years",IF([@[Experience (in years)]]<=20,"11 to 20 Years",IF([@[Experience (in years)]]<=30,"21 to 30 Years","30+ Years")))

Using this formula I divided experience in 4 Categories
0 to 10 Years
11 to 20 Years
21 to 30 Years
30+ Years.

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

ANS:



Observation:
The plot Showing a upward curve which indicates the Positive correlation of Age And Experience.
Correlation of Age And Experience is : **0.99399**

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

ANS:

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

These are the 3 areas where bank's customers are located.

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

ANS:

Personal Loan	No

Count of Credit Card Credit Card	TD Account Yes
Yes	147

147 Customers have a combination Fixed Deposits And Credit Cards but not Personal Loan
Find the number of counts using formula is:

=COUNTIFS (HBFC[TD Account],"Yes", HBFC[CreditCard],"Yes", HBFC[Personal Loan],"No")

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?

ANS:

Median income of the customers who have availed personal loans
142.5

Median income of those customers who have not availed personal loans
59

Inference:
Median income of the customers who have availed personal loans is **142.5** compare to Median income of those customers who have not availed personal loans **59**

=MEDIAN(IF(HBFC[Personal Loan]="Yes", HBFC[Income (in K/year)])) Using this formula I Find the ***Median income of the customers who have availed personal loans***

=MEDIAN(IF(HBFC[Personal Loan]="NO",HBFC[Income (in K/year)])) Using this formula I Find the ***Median income of those customers who have not availed personal loans***

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.**

ANS:

Education vs Personal Loan

Count of Personal Loan Education	Personal Loan No	Yes	Grand Total
Graduate	87.03%	12.97%	100.00%
Professional	86.34%	13.66%	100.00%
Undergraduate	95.56%	4.44%	100.00%
Grand Total	**90.40%**	**9.60%**	**100.00%**

Online vs Personal Loan

Count of Personal Loan Online	Personal Loan No	Yes	Grand Total
No	90.63%	9.38%	100.00%
Yes	90.25%	9.75%	100.00%
Grand Total	**90.40%**	**9.60%**	**100.00%**

TD Account Vs Personal Loan

Count of Personal Loan	Personal Loan		
TD Account	No	Yes	Grand Total
No	92.76%	7.24%	100.00%
Yes	53.64%	46.36%	100.00%
Grand Total	**90.40%**	**9.60%**	**100.00%**

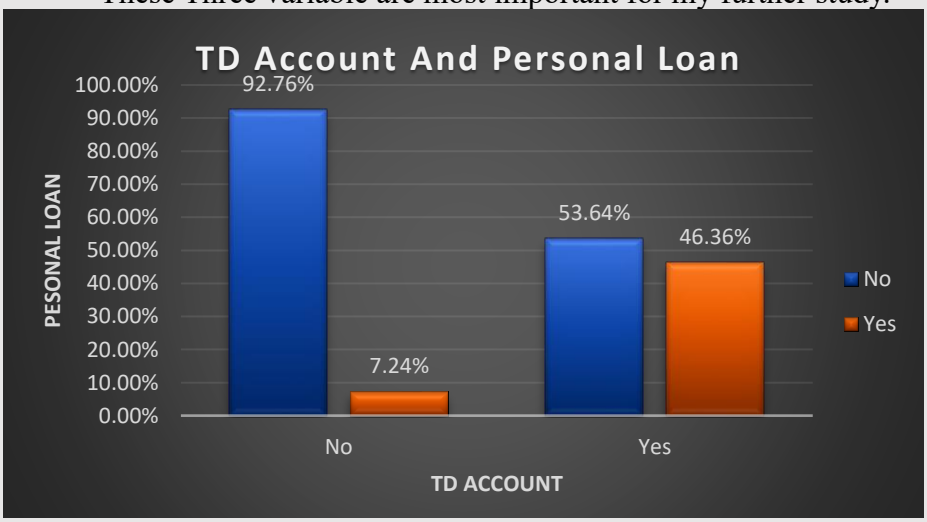
Income Category vs Personal Loan

Count of Personal Loan	Personal Loan		
Income Categorical	No	Yes	Grand Total
0-50	100.00%	0.00%	100.00%
100+	63.86%	36.14%	100.00%
51-100	97.76%	2.24%	100.00%
Grand Total	**90.40%**	**9.60%**	**100.00%**

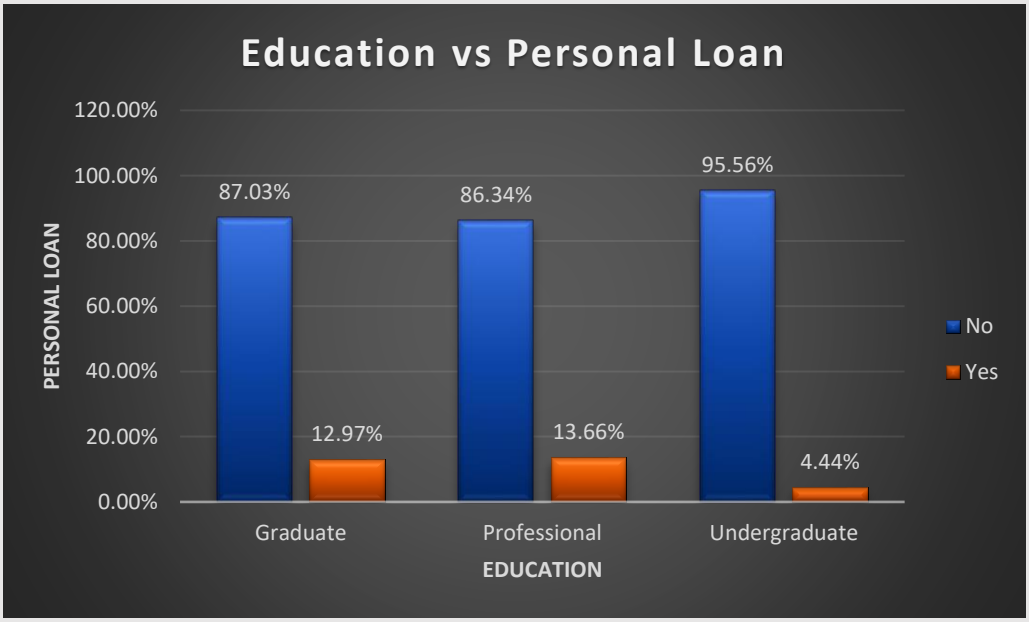
9) Analyse the Pivot tables created in the previous question and state any anomaly that you observe. Which categorical variables appear most important for your further study if you want to analyse which customers are most likely to take personal loans and why?

ANS:

- According to the previous questions it appears that the variables ***Education*** ***Termed Deposit Account*** *and* ***Income Categorical*** are more important than another variable.
- The Customers having a ***TD Account*** are taken a Personal Loan, Customers Having a ***TD account*** are higher chance to get a Personal Loan than others.
- In the other hand the customers having more than ***100+*** income are higher chances of getting a Personal Loan compared to ***0-50, 51-100*** of income.
- Thirdly is **education** according to this variable ***Professional*** and ***Graduates*** are taking more loan, we Need not to Worry about Undergraduates.
- These Three variable are most important for my further study.




Income_Category And Personal Loan
120.00%
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
PERSONAL LOAN
0-50
100+
51-100
INCOME_CATEGORY VS PERSONAL LOAN
No
Yes


Education vs Personal Loan
120.00%
100.00%
80.00%
60.00%
40.00%
20.00%
0.00%
PERSONAL LOAN
87.03%
12.97%
86.34%
13.66%
95.56%
4.44%
Graduate
Professional
Undergraduate
EDUCATION
No
Yes

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?

ANS:

TD Account	(All)
Income Categorical	(All)

Count of ZIP Code	Education		
ZIP code and Personal Loan	Graduate	Professional	Undergraduate
94720	**28.99%**	**30.77%**	**40.24%**
No	25.33%	30.67%	44.00%
Yes	57.89%	31.58%	10.53%
94305	**31.50%**	**28.35%**	**40.16%**
No	30.70%	27.19%	42.11%
Yes	38.46%	38.46%	23.08%
95616	**31.03%**	**28.45%**	**40.52%**
No	31.82%	27.27%	40.91%
Yes	16.67%	50.00%	33.33%

These ZIP codes are for example there are some other ZIP code is Available you can my work in excel

- In Next campaign we only target Particular cities rather than entire population.
- Reach to **Graduates** and **Professional** in above-mentioned Cities, these Customers are more likely to be take a Personal Loan.
- Customers having **more than 100K** will have more possibilities of taking Personal Loan.
- And Customers are having **Termed Deposit Account** will have More Possibilities of taking Personal Loan.
- In the above Table you choose Who are having **TD Account (YES)** And also choose **Income Category 51-100 And 100+**