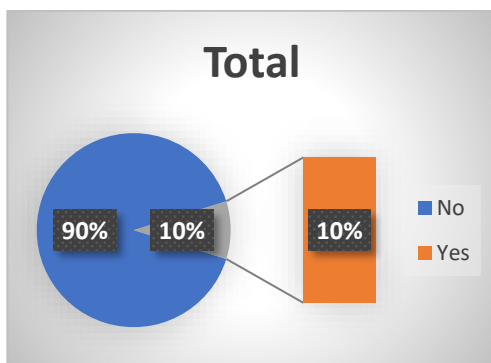


# HBFC Bank Project

1. What percentage of the bank's customers (according to the data) have availed Personal Loans?

Ans: To determine the percentage of bank customers who have availed personal loan using pivot table:

- Select the data range that contains the customer information.
- Create a pivot table.
- Add the "Personal Loan" field to the "Rows" area of the pivot table.
- Add the "Personal Loan" field to the "Values" area of the pivot table and select the show values as percentage of the grand total.
- The percentage of bank customers who have availed for personal loans will be displayed in the "Count of Personal Loan" cell.



Row Labels	Count of Personal Loan
No	90.40%
Yes	9.60%
<b>Grand Total</b>	<b>100.00%</b>

The Pie chart shows that 90.4% of the bank's customers have not availed for personal loans. This is equivalent to 90%. The remaining 9.6% of the customers have availed for personal loans.

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

Ans:

Category	Min	Max	Median	Average
Age (in years)	23	67	45	45.3384
Experience (in years)	0	43	20	20.1348
Income (in K/year)	8	224	64	73.7742
Family members	1	4	2	2.397230028
CCAvg	0	10	1.5	1.937938
Mortgage	0	635	0	56.4988

Here are some of the most important observations I made about the table in the image.

- The median age of the customers is 45 years old. This suggests that most of the customers are in the prime working age range.
- The median experience of the customers is 10 years. This suggests that most of the customers have some experience in the workforce.
- The median income of the customers is 200K/year. This is a very high income, which suggests that the customers are likely to be able to afford a mortgage.
- The median family size of the customers is 2. This is a small family size, which suggests that the customers may have fewer expenses than those with larger families.
- The median CCAvg (spending on credit cards) of the customers is 0. This suggests that the customers are not spending a lot of money on credit cards, which could be a good sign for their creditworthiness.
- The median mortgage of the customers is 0. This suggests that most of the customers do not have a mortgage, which could be a good sign for their financial situation.
- Overall, the customers in this table seem to be in a good financial position and are likely to be able to afford a mortgage.

3. Create a new categorical variable for Experience using 4 categories.

- 0 to 10 years
- 11 to 20 years
- 21 to 30 years and
- 30+ years.

Plot a bar graph for this new categorical variable

Ans:

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

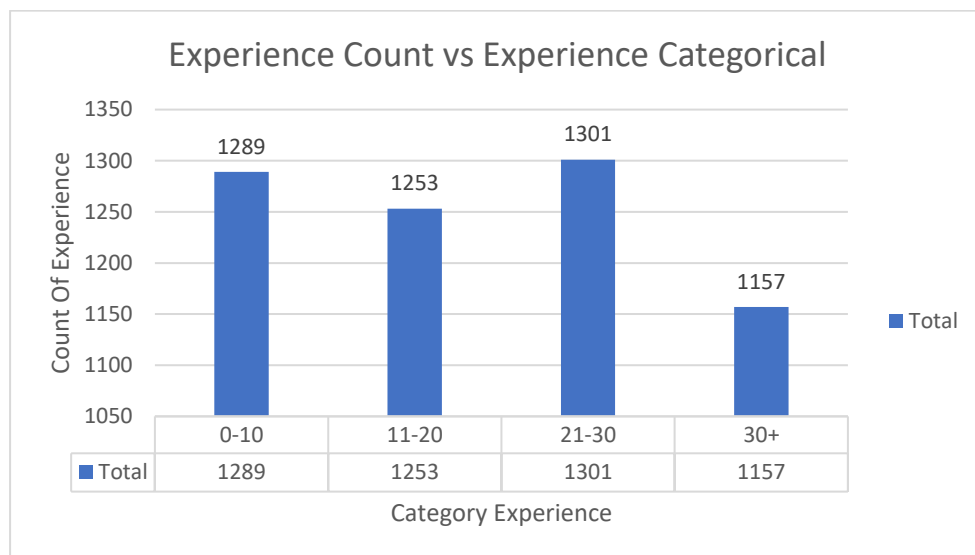
First, create a new column in your Excel spreadsheet and name it "Experience\_cat". In the first cell of the "Experience\_cat" column, enter the following formula: =IF (Experience<10, "0-10 years", IF(Experience<20, "11-20 years", IF(Experience<30, "21-30 years", "30+ years"))).

This formula will use an IF-ELSE statement to assign each row in the "Experience" column to one of the four categories. For example, if the value in the "Experience" column is less than 10, the formula will return the value "0-10 years". If the value in the "Experience" column is greater than 9 and less than 20, the formula will return the value "11-20 years". And so on.

Copy the formula down the "Experience\_cat" column. This will create a new categorical variable for Experience with 4 categories: 0-10 years, 11-20 years, 21-30 years, and 30+ years. Once you have created the new categorical variable for Experience, you can plot it using a bar chart. To do this, follow these steps:

- Select the "Experience\_cat" column.
- Click on the "Insert" tab and then click on the "Column" chart type.
- In the "Chart Tools" ribbon, click on the "Design" tab and then click on the "Change Chart Type" button.
- In the "Change Chart Type" dialog box, select the "Bar" chart type and then click on the "OK" button.

This will create a bar chart with the different categories of Experience on the x-axis and the number of customers in each category on the y-axis.



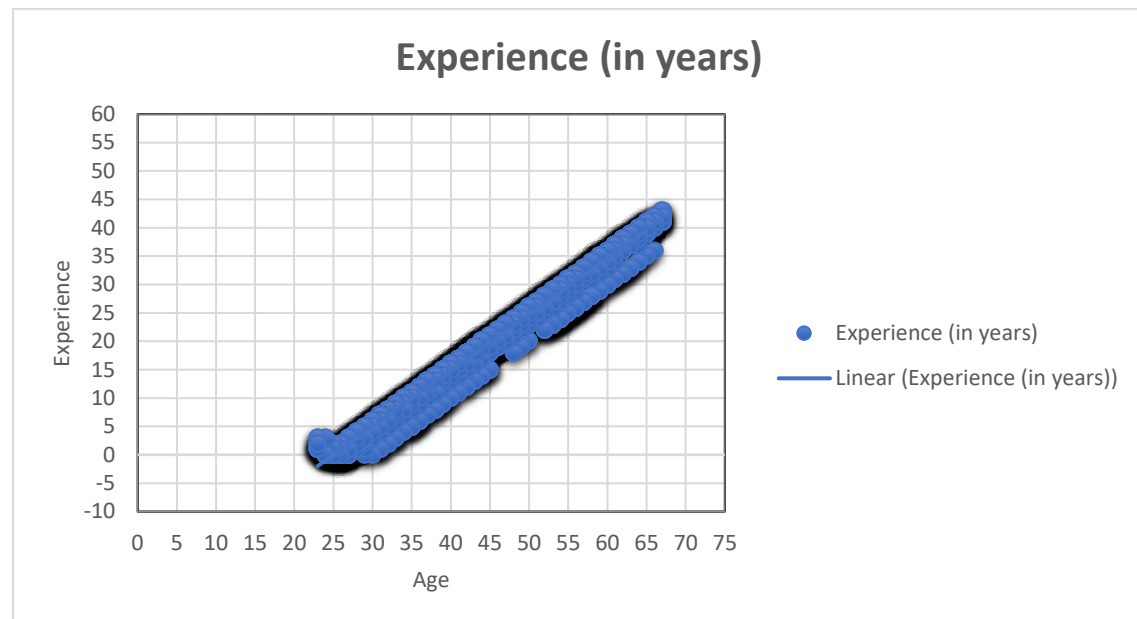
4. Create a scatter plot of the Age and the Experience variable. What do you observe?

Ans:

- Select the Age and Experience data.
- Click on the "Insert" tab.
- In the "Charts" group, click on the "Scatter" button.
- Click on the "Scatter with only markers" chart type.

I observed that the line graph shows that experience in years increases with age. The person with the most experience has 55 years of experience, followed by the person with 40 years of experience. The person with the least experience has only 2 years of experience. I also observed that the graph is not smooth but has

some fluctuations. This could be due to several factors, such as the person taking time off from work, changing jobs, or taking on new responsibilities.



Overall, the graph shows that experience in years generally increases with age.

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

Ans:

Zip Codes	Count of ID(Customers)
94305	127
94720	169
95616	116
<b>Grand Total</b>	<b>5000</b>

- Select the ZIP code data.
- Click on the "PivotTable" button.
- Click on the "New sheet" option.
- Drag the "ZIP code" field to the "Rows" area and the "Count" field to the "Values" area.
- Click on the "Top 3" filter button.

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

Ans: Here are the steps on how to create a pivot table for the combination of fixed deposits and credit cards with personal loan filter.

- Select the TD account and Credit card data.
- Use the AND function to create a new column called "Combination of fixed deposits and credit cards".
- Create a pivot table with the "Combination of fixed deposits and credit cards" column in the rows and the "personal loan" column in the filters.
- Filter the pivot table to show only rows where Personal loan is "No".

This will create a pivot table that shows the count of people who have a combination of fixed deposits and credit cards, but do not have a personal loan.

Personal Loan		No
		<b>Count of Combined of TD and Credit Card</b>
<b>Row Labels</b>		
FALSE		4373
TRUE		147
<b>Grand Total</b>		<b>4520</b>

The count of persons who have both TD account and credit card but do not have personal loans is 147.

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: Here are the steps on how to filter the personal loan column and calculate the median for customers who have and have not availed of a personal loan, with the steps reduced:

- Open the Excel spreadsheet that contains the data.
- Filter the personal loan column to show only rows where personal loan is equal to "Yes".
- Copy the income (k/year) and personal loan data for the rows where personal loan is equal to "Yes".
- Paste the income (k/year) and personal loan data in a new sheet.
- Repeat steps 2-4 for the rows where personal loan is equal to "No".
- Calculate the median for the income (k/year) data for the rows where personal loan is equal to "Yes".
- Calculate the median for the income (k/year) data for the rows where personal loan is equal to "No".

This will calculate the median income for customers who have and have not availed of a personal loan.

Here are the observations:

- Only a small number of customers (480) have availed of a personal loan.
- Customers who have availed of a personal loan have a significantly higher median income (₹1,42,500) than customers who have not availed of a personal loan (₹59,000).
- This suggests that customers with higher incomes are more likely to be approved for a personal loan.
- Overall, the data suggests that there is a positive correlation between income and the likelihood of being approved for a personal loan.

8. Create 4 separate Pivot Tables. Summarize your data by percentages.

- Education vs Personal Loan
- TD Account Vs Personal Loan
- Online vs Personal Loan
- Income Category vs Personal Loan

Ans: Education vs Personal Loan

Here are the steps on how to create a separate pivot table to summarize data by percentages for education vs personal loan, with some points reduced:

- Select the Education and Personal Loan data.
- Create a pivot table with the Education field in the rows and the Personal Loan field in column and values.
- Click on the "Percentage of total" button in the values area.

This will create a pivot table that shows the percentage of people who have a personal loan by education level.

Education vs Personal Loan

Count of Personal Loan	% Count of Personal Loan		
	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%
<b>Grand Total</b>	<b>90.40%</b>	<b>9.60%</b>	<b>100.00%</b>

### TD Account Vs Personal Loan

Count of Personal Loan	Column Labels			
	TD Account	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%

- Select the TD account and Personal Loan data.
- Create a pivot table with the TD account field in the rows and the Personal Loan field in columns and values.
- Click on the "Percentage of total" button in the values area.

This will create a pivot table that shows the percentage of people who have a personal loan by education level.

### Online vs Personal Loan

Count of Personal Loan	Column Labels			
	Online	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%

- Select the Online and Personal Loan data.
- Create a pivot table with the Online field in the rows and the personal loan field in columns and values.
- Click on the "Percentage of total" button in the values area.

This will create a pivot table that shows the percentage of people who have a personal loan by education level.

### Income Category vs Personal Loan

- Select the Income Category and Personal Loan data.

Count of Personal Loan	% Count Personal Loan		
	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%
<b>Grand Total</b>	<b>90.40%</b>	<b>9.60%</b>	<b>100.00%</b>

- Create a pivot table with the Income Category field in the rows and the Personal Loan field in columns and values.
- Click on the "Percentage of total" button in the values area.

This will create a pivot table that shows the percentage of people who have a personal loan by education level.

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?

Ans: After analyzing the pivot table created in the previous question, I observed that.

- Graduate: 24.42% with No personal loan, 3.64% with Yes personal loan.
- Professional: 25.92% with No personal loan, 4.10% with Yes personal loan.
- Undergraduate: 40.06% with No personal loan, 1.86% with Yes personal loan.

It appears that undergraduates have the lowest percentage of personal loan uptake among the education levels mentioned. However, professionals have the highest percentage of personal loan uptake.

- Online (No): 36.54% with No personal loan, 3.78% with Yes personal loan.
- Online (Yes): 53.86% with No personal loan, 5.82% with Yes personal loan.

A potential anomaly or interesting observation could be the notably higher percentage of individuals with "Yes" personal loans when internet banking is available ("Yes" category) compared to when it is not available ("No" category). This pattern suggests that there might be a positive correlation between having internet banking and being approved for personal loans.



- Customers without a TD Account (No): 93.96% of them have not taken a personal loan, while 6.80% have taken a personal loan.
- Customers with a TD Account (Yes): Only 6.04% of them have taken a personal loan, while 2.80% have taken a personal loan.

TD Account and Personal Loan Uptake: The data shows that customers with a TD Account are less likely to take a personal loan.

- Customers in the "0-50" income category: 38.28% of them have not taken a personal loan, and no customers in this category have taken a personal loan.
- Customers in the "100+" income category: 24.24% of them have taken a personal loan, while 8.76% have not.
- Customers in the "51-100" income category: 37.48% of them have not taken a personal loan, and 0.84% have taken a personal loan.

People with an income of less than 50 are less likely to take out a personal loan, while people with an income of 100 or more are more likely to take out a personal loan. People with a lower income may not be able to afford the monthly payments on a personal loan.

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 acceptance 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: Utilize data analytics and customer profiling to identify segments that are more likely to respond positively to personal loan offers. This involves analyzing historical data to identify patterns and characteristics of customers who have accepted offers in the past. Factors such as income, credit score, age, employment status, and previous transaction history can be important indicators.

Predictive Modeling: Build predictive models using machine learning algorithms to estimate the likelihood of each customer accepting the personal loan offer. This will help you rank customers based on their probability of acceptance. Models like logistic regression or decision trees can be used for this purpose.

Scoring and Ranking: Assign a score to each customer based on the predictive model's outcome. Rank customers in descending order of their likelihood of acceptance.

By implementing this strategy, the bank can significantly reduce the cost of the marketing campaign by focusing its efforts on a more receptive and targeted audience, while still maintaining an acceptable level of personal loan offer acceptance. It's important to note that data privacy and ethical considerations should be followed throughout the process of customer segmentation and targeting.