Great Learning

HBFC Bank Personal Loans Project

BUSINESS REPORT

Lohitha Mada

GLCA DA Online August 2023

Abstract

This case study focuses on HBFC Bank's strategy to enhance its loan business by converting its predominantly deposit-focused customer base into personal loan customers, while maintaining their deposit relationships. With a limited number of existing borrowers, the bank is eager to rapidly expand this segment and leverage the interest income generated by loans. To achieve this goal, the bank plans to construct a predictive model capable of identifying potential customers with a higher likelihood of availing personal loans. The initial step in this initiative involves conducting a thorough analysis of historical customer data to extract valuable insights.

The provided dataset contains information on 5000 customers, including their demographic characteristics (such as age and income), their interactions with the bank (including mortgages and securities accounts), and their responses to the most recent personal loan campaign. Utilizing a data-driven approach, HBFC Bank aims to formulate a targeted strategy for acquiring loans and retaining customers, ultimately promoting sustainable growth in its loan portfolio while preserving its deposit base.

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

Answer: According to the given data, 9.60% of the bank's customers have availed Personal loans.

Yes or no	Count of Personal Loan
Yes	9.60%
No	90.40%

Inference:

According to the given data, 9.60% of the bank's customers have availed Personal Loans.

Approach Used:

1. Selected the necessary data and generated a pivot table using the "Insert" tab.
2. Then, I took the "Personal Loan" column and placed it in both the rows and values fields. I adjusted the values to be displayed as a count and utilized the "Show Values" feature (By right-click) to present the numerical outcomes as column percentages.

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?

Answer: Please find the below table with a structured description of the requested observations for each numeric variable:

Requested Inputs	Min	Max	Median	Average
Age	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	2.370572207
CCAvg	0	10	1.5	1.937938
Mortgage	0	635	0	56.4988

Approach Used:

1. Selected the required data.
2. Selected Data analysis ribbon in Data tab.
3. From the displayed analysis tool, selected Descriptive analysis and clicked ok.
4. Inside the Descriptive analysis, selected the input range, summary statistics, Labels in first row (As I selected labels) and selected output range as new worksheet and clicked ok.
5. In the new worksheet, I got all the summary statistics of all the columns. As per the question, I created the new table with the min, max, median, average for all numeric variables.

My Observations/ Inference:

Age:
The age of individuals in the dataset ranging from 23 to 67 years.
The median age is 45 years, with an average age of 45.34 years.

Experience (in years):
Work experience among the customers according to the given dataset ranges from 0 to 43 years.
The median work experience for these customers is 20 years, with an average experience of 20.13 years.

Income (in K/year):
The annual income of customers based on the dataset ranges from 8,000 to 224,000.
The median income level for these customers is 64,000 with an average income of approximately 73,774

Family members:
The number of family members in this dataset ranges from 1 to 4.
The median family size is 2 members, with an average family size of 2.37 members.

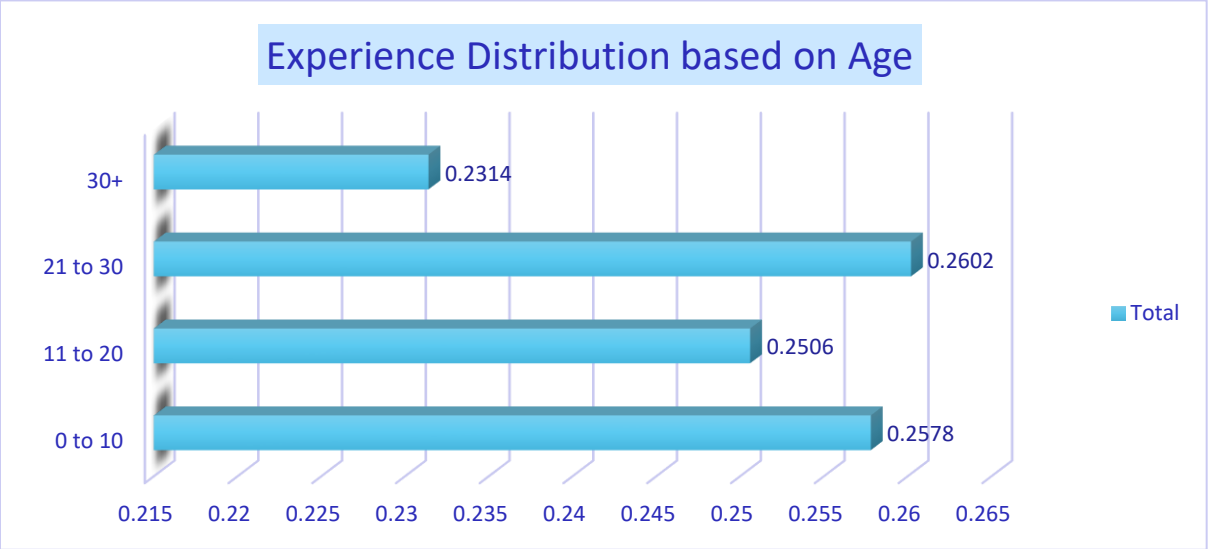
CC Avg:
Average monthly credit card spending among the given dataset varies from 0 to 10.
The median monthly spending is 1.5, with an average monthly spending of 1.93

Mortage:
Mortgage amounts range from 0 to 635.

The median mortgage amount is 0, while the average mortgage amount is approximately 56.5.

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

Answer: Below is the bar graph of the customers whose experience is categorized based on their ages.



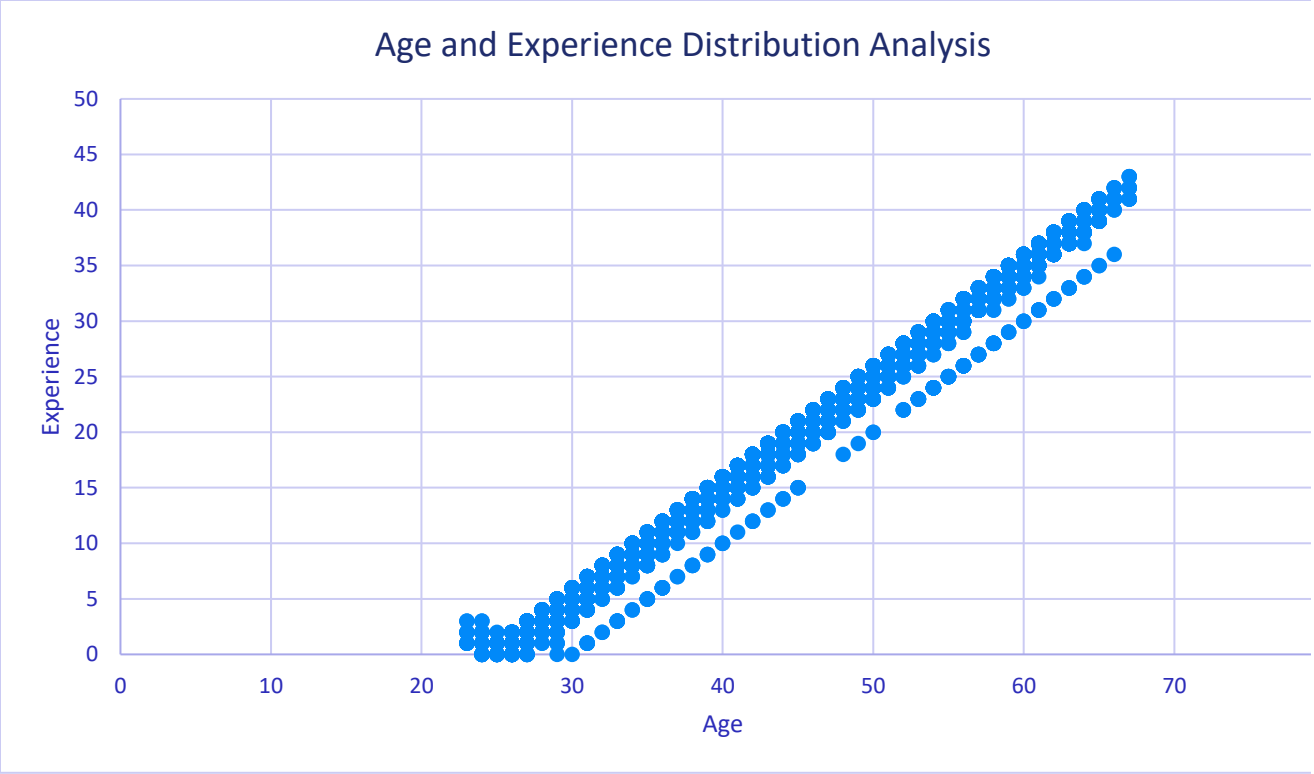
Approach Used:

1. To create a new categorical variable for Experience using 4 categories, I used IF and AND functions. This is the formula used:
   =IF (C2<=10,"0 to 10", IF(AND(C2>10, C2<=20),"11 to 20", IF(AND(C2>20, C2<=30),"21 to 30","30+"))).
2. With the result of that column, created a pivot table and added bar graph using pivot chart ribbon from pivot table analyze tab.
3. Renamed the title on the chart and included data labels to display percentage.

Inference:
Based on the bar graph, People of age 21-30 years were more experienced.

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

Answer: Please find the below scatter plot of the Age and the experience variable.



Approach Used:

1. Selected Age, Experience columns and added(pasted) those two columns in a new sheet.
2. Using Insert tab, selected scatter chart ribbon.
3. Added axis title, chart titles, gridlines.

My Observation/ Inference:

Based on the scatter plot, experience is directly correlated with age. Which results that there is a strong, positive relationship between a person's age and their level of experience.

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

Answer: The top 3 areas (ZIP Codes) where the bank's customers are located were:

1. ZIP Code 94720 with 169 customers.
2. ZIP Code 94305 with 127 customers.
3. ZIP Code 95616 with 116 customers

Zip Codes	Count of ZIP Code
94720	169
94305	127
95616	116
90095	71
93106	57
93943	54
92037	54
91320	53
91711	52
94025	52
92093	51
90024	50
90245	50
91330	46
90089	46
92121	45
94304	45
94143	27

Approach Used:

1. Selected zip codes and created pivot table.
2. Calculated the count of zip code and sorted with largest to smallest.

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

Answer: There were 147 customers who have both Fixed Deposits and Credit Cards but do not have a Personal Loan.

Approach Used:

Selected the required columns (Personal Loan, TD Account, Credit Card) and used the formula below to get the required outcome.

=IF (AND (B2="Yes", C2="Yes", A2="No"),"Yes", "No")

=COUNTIFS(D2:D5001,"Yes")

	A	B	C	D
1	Personal Loan	TD Account	CreditCard	Fixed+Credit-Personal Loan
2	No	Yes	Yes	Yes
3	No	Yes	Yes	Yes
4	No	Yes	Yes	Yes
5	No	Yes	Yes	Yes
6	No	Yes	Yes	Yes
7	No	Yes	Yes	Yes
8	No	Yes	Yes	Yes
9	No	Yes	Yes	Yes
10	No	Yes	Yes	Yes
11	No	Yes	Yes	Yes
12	No	Yes	Yes	Yes
13	No	Yes	Yes	Yes
14	No	Yes	Yes	Yes
15	No	Yes	Yes	Yes

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?

Answer: Median income of customers who have availed personal Loans is **142.5**K/Year whereas median income of customers who have not availed personal loans is **59**K/Year

Income (in K/year)	Personal Loan	Income (in K/year)	Personal Loan	Income (in K/year)	Personal Loan		
170	Yes	170	Yes	112	No		
134	Yes	134	Yes	32	No		
133	Yes	133	Yes	20	No		
185	Yes	185	Yes	105	No		
105	Yes	105	Yes	114	No		
195	Yes	195	Yes	21	No		
95	Yes	95	Yes	78	No	142.5	
184	Yes	184	Yes	79	No		Median income of customers who have availed personal Loans: 142.5
154	Yes	154	Yes	75	No	59	Median income of customers who have not availed personal loans: 59
160	Yes	160	Yes	41	No		
89	Yes	89	Yes	18	No		

Approach Used:

1. Selected the required columns (Income and Personal loan) and added it in a new worksheet.
2. Sorted the Personal Loan column in ascending order in order to display "Yes" first and "No" second.
3. By using a median formula, calculated the median of the incomes of the customers who have taken personal loans and customers who have not taken personal loans. Formula used:
   =MEDIAN(D2:D481), =MEDIAN(G2:G4521)

Inference:

People who borrowed personal loans are having higher incomes, indicating that they are in a good financial position and eligible for loans.
People who didn't borrowed personal loans has lower incomes, indicating they might be facing financial challenges or might not meet the requirements to get a loan because of their financial history.

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

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

Answer: Below are the pivot tables you have asked for:

1. **Education Vs Personal Loan**

Education Vs Personal Loan		
Count of Personal Loan	Personal Loan	
Education	Yes	No
Graduate	37.92%	27.01%
Professional	42.71%	28.67%
Undergraduate	19.38%	44.31%

2. **TD Account Vs Personal Loan**

TD Account Vs Personal Loan		
Count of Personal Loan	Personal Loan	
TD Account	Yes	No
Yes	29.17%	3.58%
No	70.83%	96.42%

3. **Online Vs Personal Loan**

Online Vs Personal Loan		
Count of Personal Loan	Personal Loan	
Online	Yes	No
Yes	60.63%	59.58%
No	39.38%	40.42%

4. **Income_Category Vs Personal Loan**

Income_Category Vs Personal Loan		
Count of Personal Loan	Personal Loan	
Income Categorical	Yes	No
0-50	0.00%	42.35%
100+	91.25%	17.12%
51-100	8.75%	40.53%

Approach Used:

1. Selected the required data and created pivot table using the insert tab. Next, I selected the Education column to display in rows and Personal Loan values to display in columns and added the same to values field. Converted values to be

visible as a count and used the show values feature (By right click) to change the numerical result to be displayed in the form of column percentage.
2. Repeated the same steps (1,2) to create pivot tables for TD Account Vs Personal Loan, Online vs Personal Loan and Income Category Vs Personal Loan.
3. Highlighted the Values which has more significance.

Question 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?

Answer:

1. Based on the Education Vs Personal Loan pivot table Analysis, it's evident that Professional degree holders has significantly higher percentage of availed personal loans (42.71%) compared to graduates (37.92%) and undergraduates (19.38%).
2. The second pivot table, i.e., TD Account Vs Personal Loan, says that Customers without a TD account (No) have a higher percentage of personal loans (70.83%) compared to those with a TD account (29.17%).
3. The third pivot table, i.e., based on Online Vs Personal Loan, says that customers using internet facility has higher percentage of personal loans (60.63%) compared to those who are not using internet facilities (39.38%).
4. The fourth pivot table, which is based on Income Category Vs Personal Loan, shows that customers with the highest income category (100+) have the highest percentage of personal loans (91.25%) compared to other income categories. The lowest-income category (0-50) has no personal loans (0.00%).

Based on the above inferences, customers of Income Category seem to be the most important for further study when analyzing which customers are most likely to take personal loans. People with higher incomes are more inclined to get personal loans. To better serve these customers, we should understand why they choose personal loans and what they need.

Additionally, it's interesting to investigate why people with professional degrees, Online users, TD account holders tend to get personal loans. This can give us insights into what this group of customers prefers choosing Personal loans.

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 optimize 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?

Answer: Based on the insights obtained from the analysis of the provided data, I would recommend the following strategy to reduce the cost of the personal loan campaign for HBFC Bank while maintaining a high rate of acceptance:

1. **Target on High-Income customers:**

   As high-income customers are more likely to accept personal loan offers, first target on customers whose income is above 100+. Make special ads that speak directly to these rich customers. Tell them why our loans are great for them.
2. **Focus on Professional Graduates:**

   Customers with professional degrees have shown a higher significance to get personal loans. Try reaching to these people as they are likely in the age to get luxury life and may need money for higher studies, highlight the benefits of personal loans that are particularly attractive to this group, such as flexibility, interest rates and investment opportunities.
3. **Cost-Efficient Campaigns:**

   To save money on our marketing campaign, we should be pickier about who we contact. Instead of reaching out to all 5000 customers, we should concentrate our efforts on the high-income and professional degree-holding customers we identified earlier. By doing this, we can spend less on outreach while still having a good chance of getting a positive response.

4. **Use Online:**

   Use internet to tell people about our loans, maintain a strong online presence. Ensure that the bank's online marketing efforts are well-targeted, clear, and user-friendly, providing easy access to personal loan information and application processes. This also help us to target the customers who use our internet banking facilities.

5. **Request Feedback:**

   Set up a way to get feedback from customers who say yes or no to our loan offers. We can use this feedback to make our future marketing plans better and to make sure our customers have a good experience with us.

6. **Monitor Cost and Budget Allocation:**

   Keep a close eye on how much we spend and make smart decisions about our budget. Figure out exactly how much we can afford to spend, and regularly check if our marketing is giving us good value for that money. If some groups of customers are doing better than others, change how we spend our money to match them.

By using these strategies, HBFC Bank can spend less money on personal loan ads while still finding customers who are likely to say yes. This helps the bank save money and also makes the bank's chances of increasing its loan portfolio.

…….. **The End** ……..……