

Assignment-1

Q1) What percentage of the bank's customers (according to the data) have availed Personal Loans vs the ones who have not availed it?

From the given data **9.60%** of the Bank's Customers have availed for Personal Loans. and the remaining **90.40%** of Customers have not availed for Personal Loans.

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

This Percentage is figured out by using pivot table taking Personal Loan as row and count of personal loan and represented in percentage.

Q2) Generate a table with min, max, median & average for all numeric variables (age, experience, income, family members, CCAvg, Mortgage)

	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	2.397230028	1.937938	56.4988

The table for Min, Max, Median & average for all numeric variables as such age, experience, income, family members, CCAvg, Mortgage were tabulated using the required formuale

Q3) 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

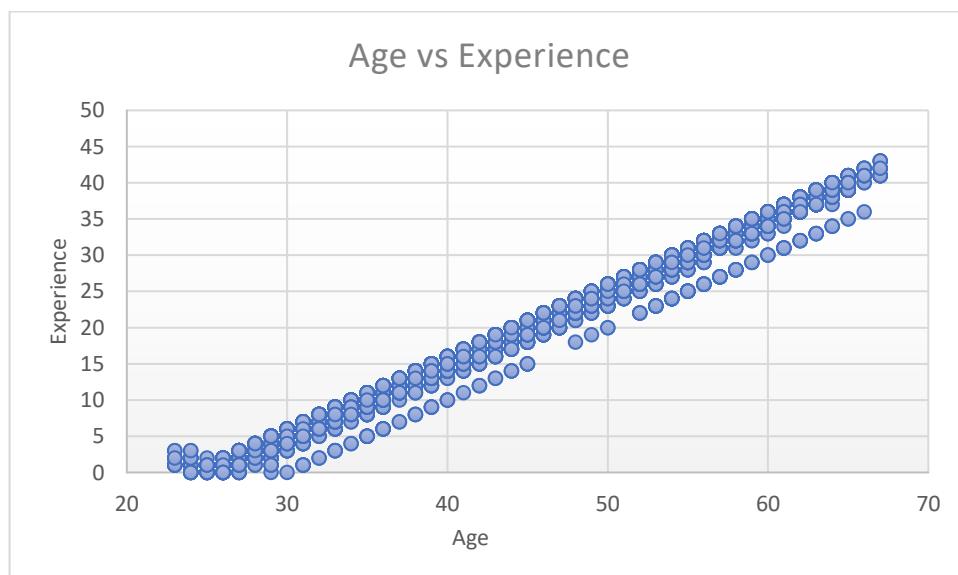
A new categorical variable for Experience using 4 categories was created using the nested IF formula.

Formula used: [=IF(C2<=10,"0 to 10 Years",IF(C2<=20,"11 to 20 Years",IF(C2<=30,"21 to 30 Years","30+ Years")))]

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



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



From the scatter plot we can deduce that as Age increases Experience also increases. This plot is taken by selecting the two required columns separately and insert a Scatter plot chart.

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

The top 3 areas where the bank's Customers located are as follows,

ZIP CODE	Count
94720	169
94305	127
95616	116

This problem is solved by using countif in the Zip code column and then by removing the duplicate names in the data and finally sorting data in largest to smallest order.

Formula used: =COUNTIF(\$I\$6:\$I\$5005,I6)

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

Count of Customer with FD and CC but No Personal Loan is **147** and percentage is **2.94** of the total Customers. This Is done by creating a new column and using the AND function the identifying customers have a combination of Fixed Deposits and Credit Cards but not Personal Loan.

Formula used: =COUNTIF(D6:D5005,"Yes")

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?

The Median Income of Customers with Personal Loan is **142.5K** and of Customers without Personal Loan is **59K**.

From this Data we have infer that the Income of Customers without Personal Loan is Significantly lesser than ones with Personal Loan.

In this we take the median of the Customers with Personal Loan and of the Customers without Personal Loan and find the median separately.

Formula used: =MEDIAN(B7:B486)

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

1. Education vs Personal Loan			
Education	Personal Loan		Grand Total
	No	Yes	
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%

2. TD Account Vs Personal Loan			
TD Account	Personal Loan		Grand Total
	No	Yes	
No	92.76%	7.24%	100.00%
Yes	53.64%	46.36%	100.00%
Grand Total	90.40%	9.60%	100.00%

3. Online vs Personal Loan			
Online	Personal Loan		Grand Total
	No	Yes	
No	90.63%	9.38%	100.00%
Yes	90.25%	9.75%	100.00%
Grand Total	90.40%	9.60%	100.00%

4. Income_Category vs Personal Loan			
Income_Category	Personal Loan		Grand Total
	No	Yes	
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) 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?

From the pivot table of we can very evidently understand that not a single customer with the income category of 0-50k/year has availed a Personal Loan.

Income_Category(k/year)	Personal Loan		Grand Total
	No	Yes	
	100.00%	0.00%	
0-50			100.00%

The Customers With TD Account have the highest percentage of Personal Loan acceptance of 46.36% followed by customers of income category of 100+K/year. Hence we can say that focusing on Customers with TD Account will result in raise in the Personal Loan count in the Bank.

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

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?

According to many analysis I find that targeting customers with TD Account and Customers having Income of 100K/per year will get the bank to see more Personal Loan Acceptance.

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_Category	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%