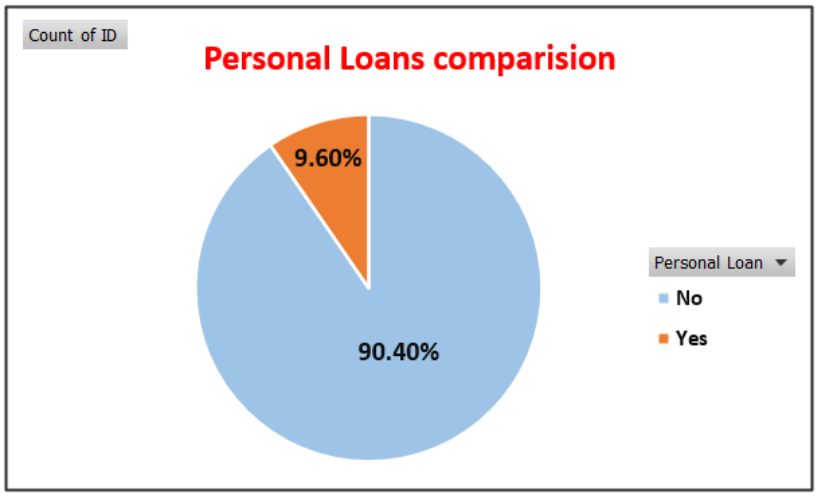
HDFC Bank Personal Loans

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

Personal Lo	Count of ID
No	90.40%
Yes	9.60%
Grand Total	**100.00%**



Here we take customer ID and personal loan data to design pie chat by using pivot table. In above graph represents that 90.4% of population does not have personal loans while only 9.6% of the population has taken personal loans. Therefore, non-personal loan population is very high compared with the population who take 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?

Age (in years)	Experience (in years)	Income (in K/year)	Family members	CCAvg	Mortgage
67	41	112	1	2	0
67	42	32	1	1.1	0
67	41	20	2	0.4	80
67	43	105	4	1.7	0
67	41	114	4	2.4	0
67	42	21	3	0.1	0
67	41	78	4	2.4	0
67	43	79	4	1.7	215
67	42	75	4	0.1	182
67	43	41	2	1.1	0
67	41	18	2	0.4	0
67	42	51	3	2.2	0
66	41	15	3	0.1	0
66	40	42	2	0.7	138
66	41	18	3	0.5	0

Age (in years)		*Experience (in years)*		*Income (in K/year)*		*Family members*		*CCAvg*		*Mortgage*	
Mean	45.3384	Mean	20.1348	Mean	73.7742	Mean	2.39723003	Mean	1.937938	Mean	56.4988
Median	45	Median	20	Median	64	Median	2	Median	1.5	Median	0
Minimum	23	Minimum	0	Minimum	8	Minimum	1	Minimum	0	Minimum	0
Maximum	67	Maximum	43	Maximum	224	Maximum	4	Maximum	10	Maximum	635
Count	5000	Count	5000	Count	5000	Count	4982	Count	5000	Count	5000

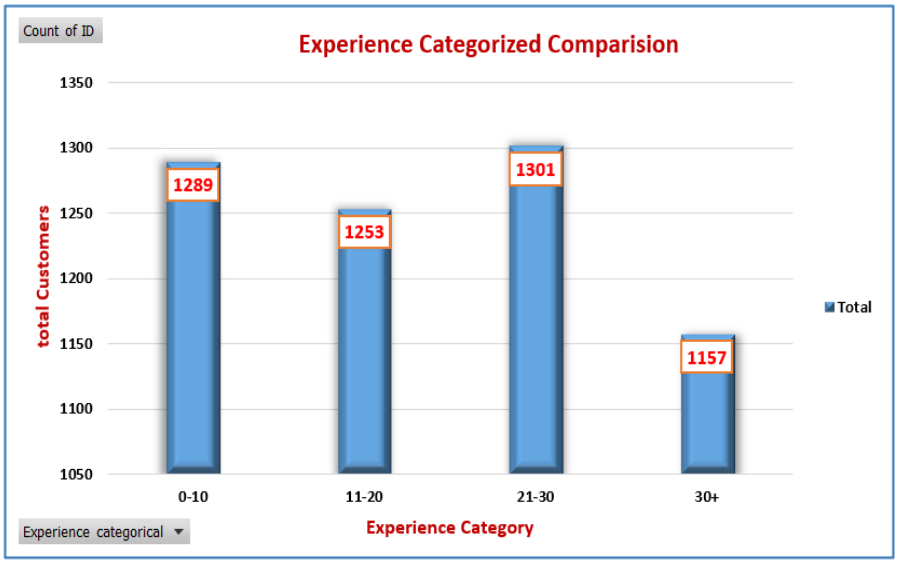
In the above table we find the Avg., Min, Max and Median values for Age, Experience, Income, Family member, CCAvg and Mortgage by using descriptive statistics. The descriptive statistics are found in Data ➡ Data Analysis ➡ Descriptive Statistics. The average age is 45 years old with 20 years of experience and an annual salary of around 74k with an average of 2 members in a family. Here the maximum age is 67 years with 43 years of experience and 224k/year of max income with maximum 4 members in a family.

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 task. You can refer how Income_Category has been created in the dataset]

ID	Experience (in years)	Experience categoric
15	41	30+
1481	42	30+
1860	41	30+
2847	43	30+
3265	41	30+
3332	42	30+
3704	41	30+
3887	43	30+
4173	42	30+
4361	43	30+
4452	41	30+
4469	42	30+
100	41	30+
256	40	30+
258	41	30+
466	42	30+
669	41	30+

In the above table we use experience data and that data can be categorized into four different groups by using “Nested if function” (IF(B2<=10,"0-10",IF(B2<=20,"11-20",IF(B2<=30,"21-30","30+"))). This is the formula that is used to create experience category.

Experienc	Count of ID
0-10	1289
11-20	1253
21-30	1301
30+	1157
Grand Total	**5000**



In above graph shows that which experience category have high number of customers. Here, the customers who have between 21-30 years of experience category are in high.

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

We use Age vs Experience in above graph. The graph can be representing like, if Age increases the Experience also increases dramatically.

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

ZIP Code	result
91741	6
93943	54
91741	6
93711	9
95616	116
94607	7
94301	27
95616	116
90041	10
90024	50
92130	18
94117	20
91711	52

Top 3 Zip codes

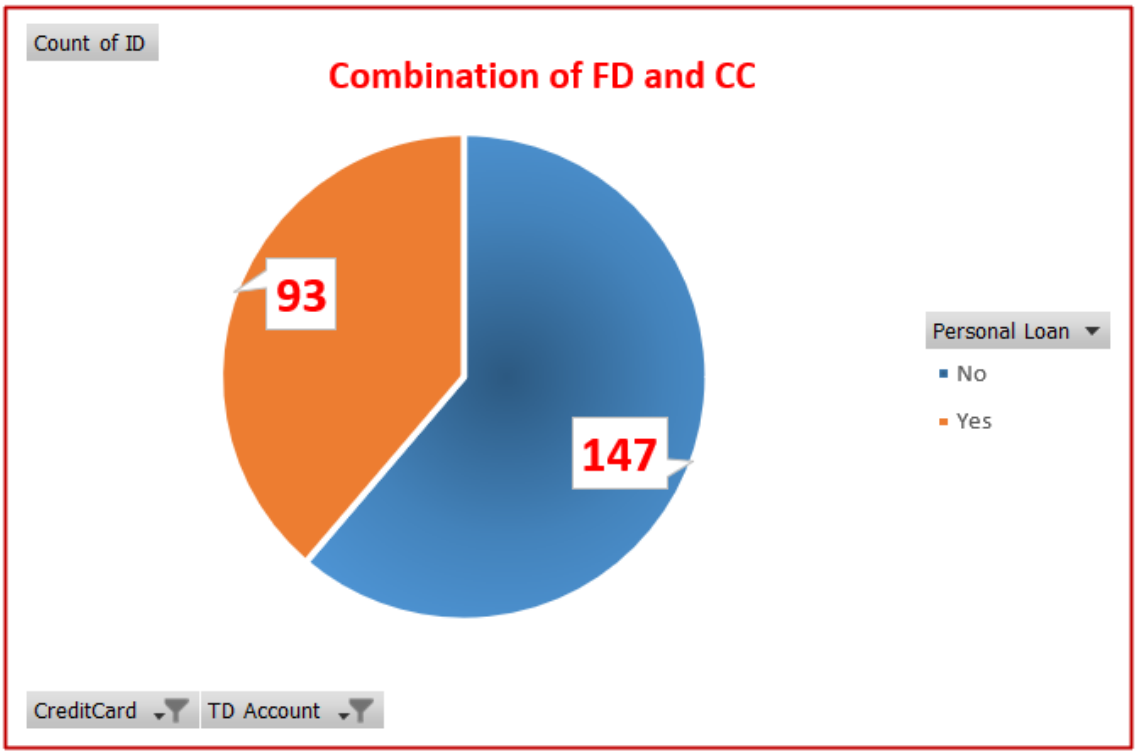
ZIP Code	total count
94720	169
94305	127
95616	116

Here we find top 3 zip codes by using countif function to overall zip code data. Firstly, we using countif function to find the count of each zip code. Secondly, we use filter to sort the data

by descending order and again we use countif function to find the top 3 areas zip codes where bank customers are located. The top 3 zip codes are 94720, 94305 and 95616 which are the total count of 169, 127 and 116 respectively.

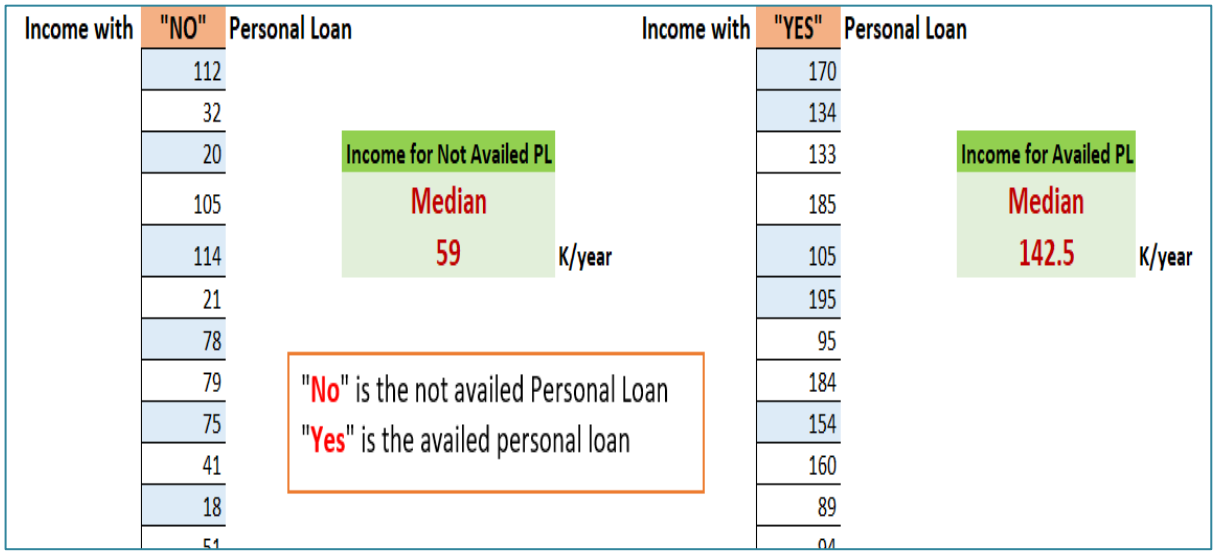
6. How many customers have a combination of Fixed Deposits and Credit Cards but not Personal Loan?

Count of I Column Lab			
	Yes	Yes Total	Grand Total
Row La	Yes		
No	147	147	147
Yes	93	93	93
Grand Tot	**240**	**240**	**240**



Here we use Customer ID, Personal loan, credit card and TD account data. In the above graph we use combination of Credit card and TD account as Yes i.e.; the customer who have Credit card and TD account. The Pie Graph shows that 93 customers who take personal loans and also have Credit Card and TD accounts, while compared to 147 customers who do not take personal loans and have both Credit Card and TD accounts.

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?



Here, we use income and personal loan data to find the median of the income. Firstly, we use customers without personal loan to determine the median salary. Here “No” is the not availed personal loan and “Yes” is the availed personal loan. The customers without personal loan make 59k/year as median salary. Secondly, we use customer who have personal loans and find the median salary of the income. The median income of the customers who have availed personal loan is 142.5k/year.

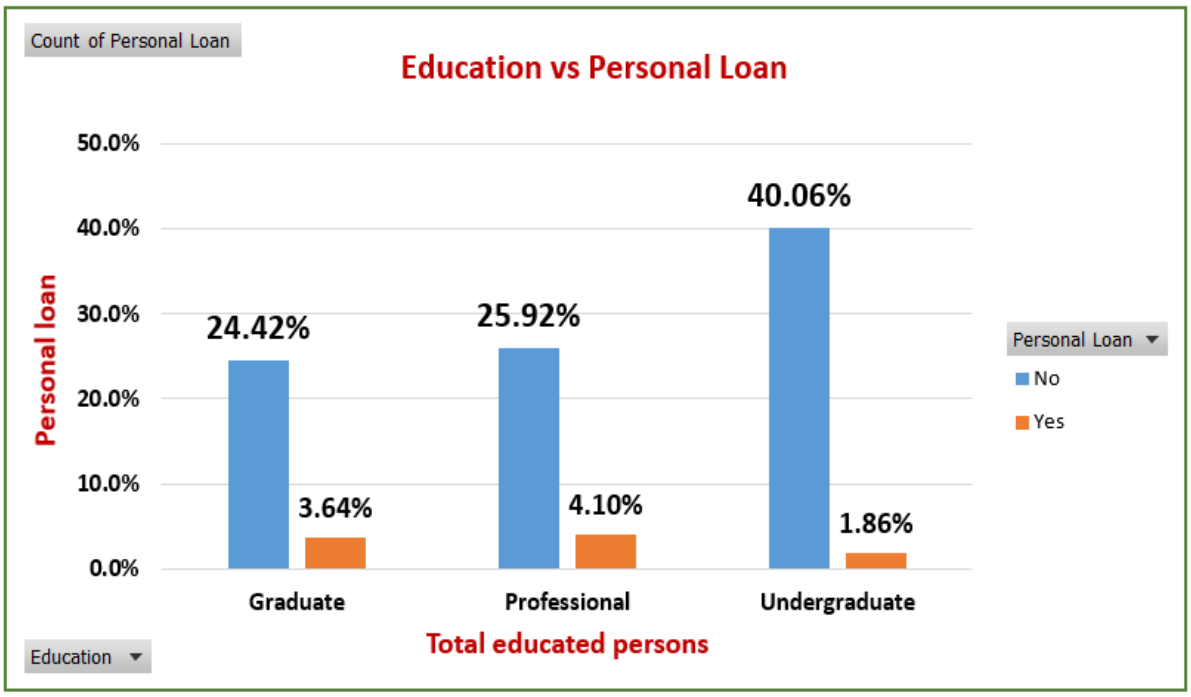
8. Create 4 separate Pivot Tables. Summarise your data by percentages. Education vs Personal Loan TD Account vs Personal Loan Online vs Personal Loan Income_Category vs Personal Loan [Hint: Please drag Personal Loan to the Columns area while creating the Pivot Table to get the required values]

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%**

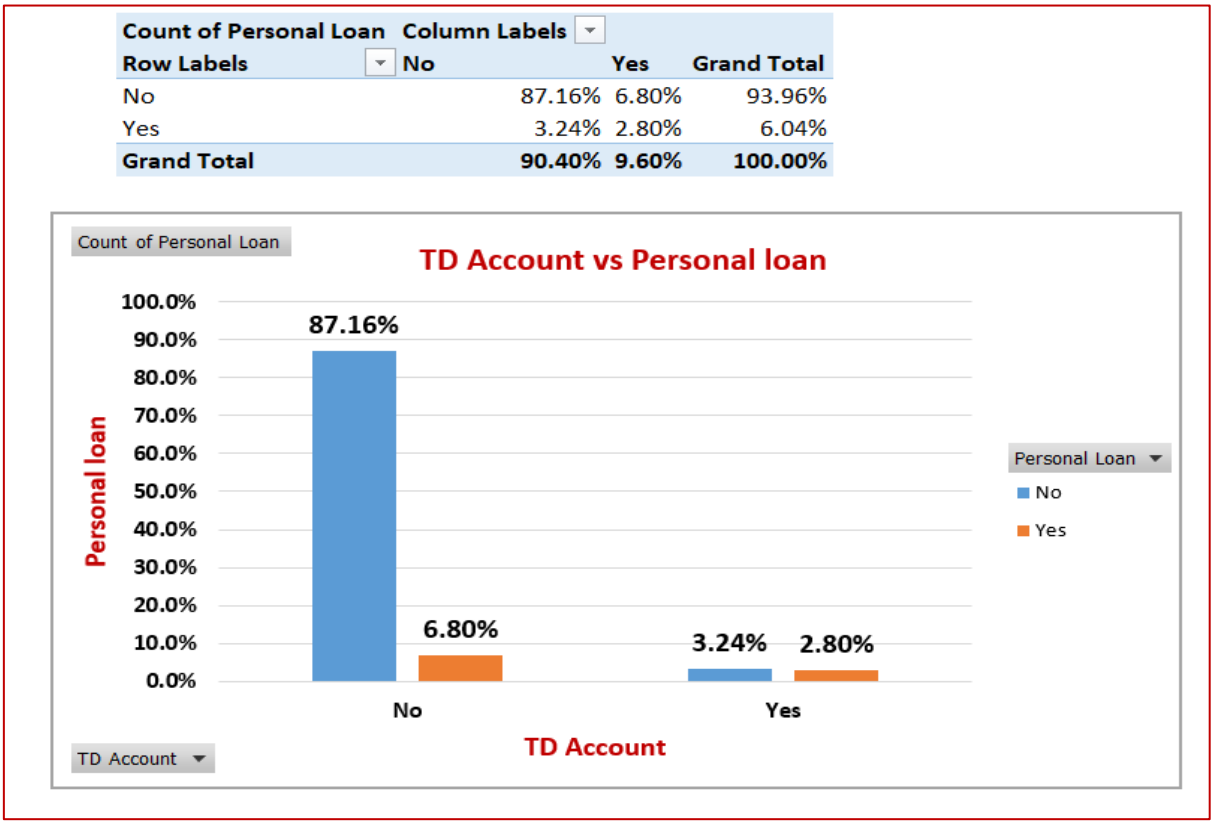
In the above table we use Education vs Personal loan. The graph represents that 40% of undergraduates does not have personal loans and only 3.6% of graduates, 4.1% of professionals

and 1.8% of undergraduates have personal loans. Finally, we conclude that there will be little bit changes of professional to take personal loans.



TD Account vs Personal Loan

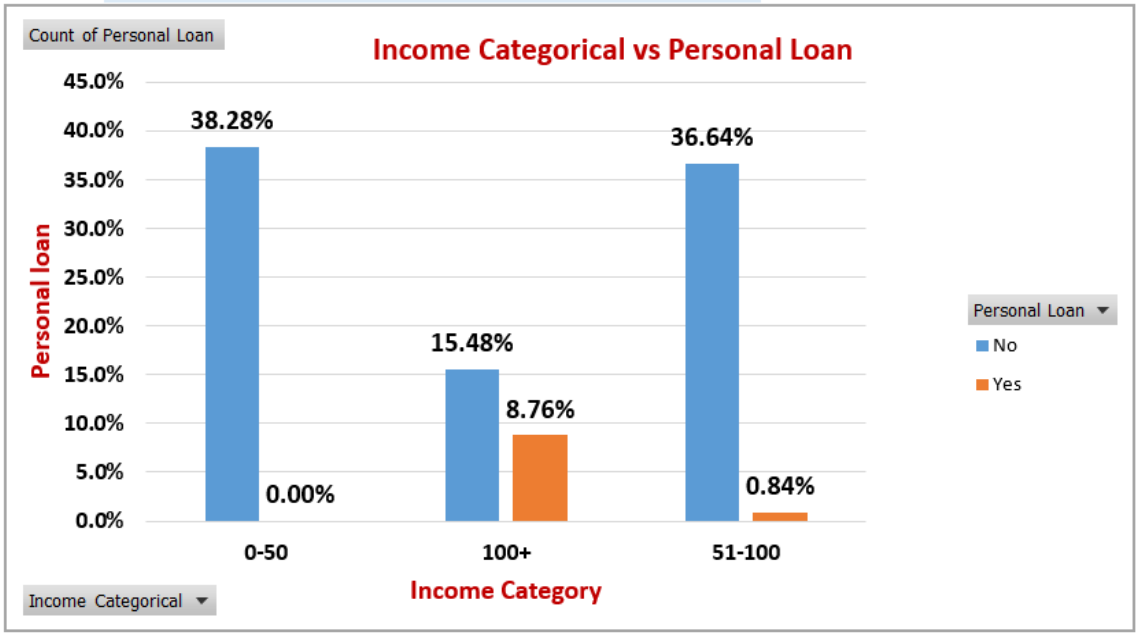
Count of Personal Loan	Column Labels		
Row Labels	No	Yes	Grand Total
No	87.16%	6.80%	93.96%
Yes	3.24%	2.80%	6.04%
Grand Total	**90.40%**	**9.60%**	**100.00%**



In this graph, we use TD account vs personal loan. Here 87.16% have No Personal Loan and No TD account. 6.8% have Personal Loan and No TD account. 3.24% have No Personal Loan and have TD account. 2.8% have both Personal Loan and TD account.
Finally, we found that there are high percentage of customers who does not have TD account and personal loans.

Income Category vs Personal Loan

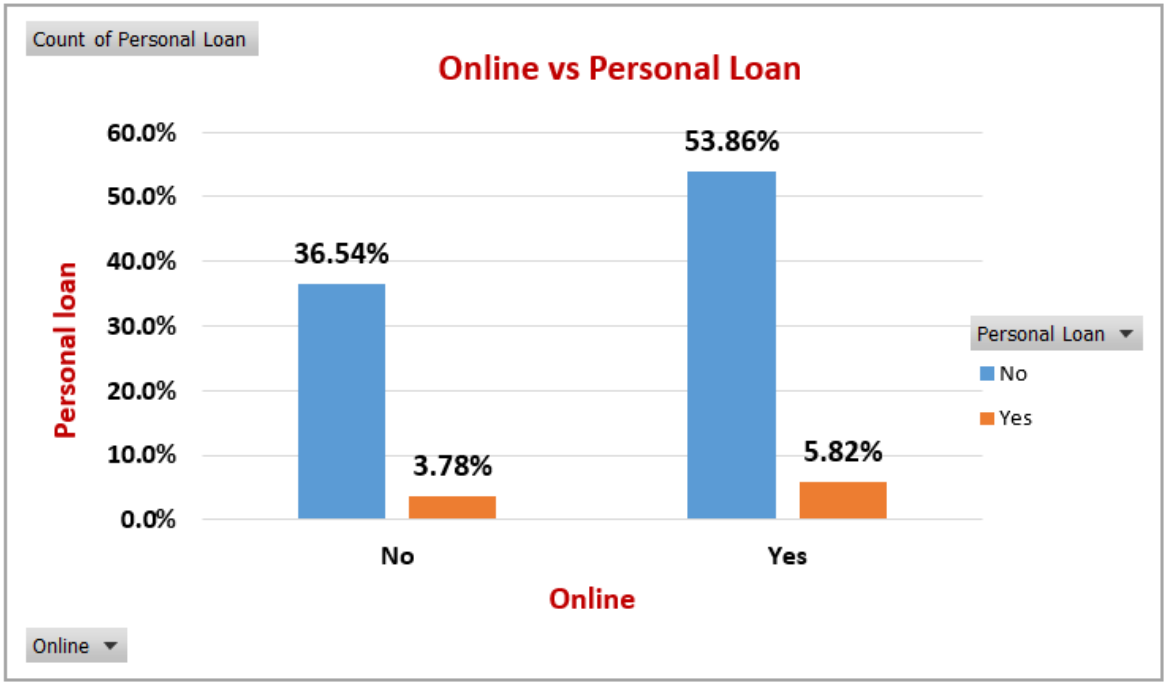
Count of Personal Loan	Column Labels		
Row Labels	No	Yes	Grand Total
0-50	38.28%	0.00%	38.28%
100+	15.48%	8.76%	24.24%
51-100	36.64%	0.84%	37.48%
Grand Total	**90.40%**	**9.60%**	**100.00%**



In the above graph, we use Income Category vs Personal loan. Here 0-50K/year income category has high percentage i.e. 38.2% of customers with no personal loan and 100+K/year of income category having a personal loan of 8.76%. Therefore, the annual salary of 50-100 K income category have a change to take personal loan.

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%**



In the above graph, we use online vs personal loans. According to the graph around 54% of the online customers do not have personal loans and just only 6% of the online customers have the personal loans. Finally, we determined that majority of the customers do not use personal loans from the online.

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?

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%**

→ Education 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%**

→ TD Account 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%**

→ Income Categorical 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%**

→ Online vs Personal Loan

Here we created 4 different pivot tables i.e. Education vs personal loan, TD account vs personal loans, Income category vs Personal loans, Online vs personal loans.
According to the above pivot tables, there will be a significant change in the professional and 50-100K yearly income category customers need to take personal loans.

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?

Education	TD Account	Income Categorical	PersonalLoan(Yes)
Graduate	Yes	100+	96.15%
Professional	Yes	100+	94.44%
Graduate	No	100+	73.75%
Professional	No	100+	72.92%
Undergraduate	Yes	100+	49.06%
Graduate	Yes	51-100	29.41%
Undergraduate	Yes	51-100	18.18%
Professional	Yes	51-100	9.09%
Undergraduate	No	100+	7.56%
Professional	No	51-100	2.02%
Graduate	No	51-100	1.73%
Undergraduate	No	51-100	1.16%
Undergraduate	No	0-50	0.00%
Undergraduate	Yes	0-50	0.00%
Professional	No	0-50	0.00%
Professional	Yes	0-50	0.00%
Graduate	No	0-50	0.00%
Graduate	Yes	0-50	0.00%
Grand Total			**9.60%**

According to the table the graduates and professionals with high income of 100+ K/year categorized customers received 96% and 94% of personal loans. The customers with income of 0-50 K/year category have very less changes to take personal loans. The customers who are 50-100 K/year of income category with TD account have high changes to take personal loan. We targeted those kind of customers who got 50-100K/year income with having TD account. Finally, we determined that just 9.6% of the overall customers took out personal loans.

HDFC Bank Personal Loan Project

Done by

Siliveri Mohan