Business Report:

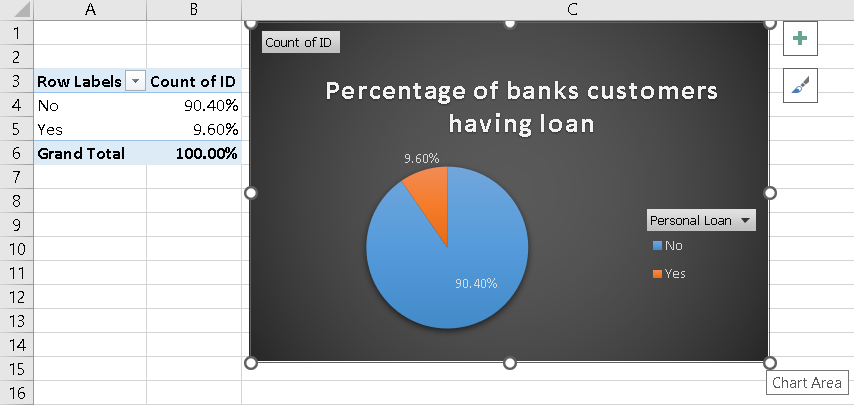
## Questions:

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

* ***Approach:***

Here I used Pivot Table for finding the aggregate results and used the “Show value as” option to get the result in Percentage format

* ***Output:***

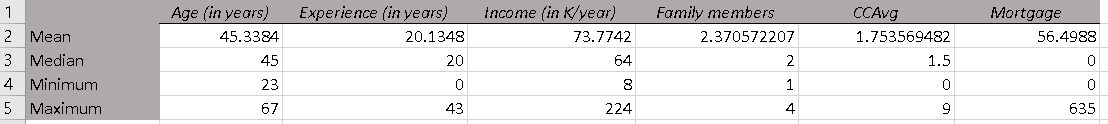


* ***Insights:***
  + Around 10% of customers having bank account has taken a personal loan
* ***Inference:***
  + There are around 90% of customers had not taken loan

## 

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

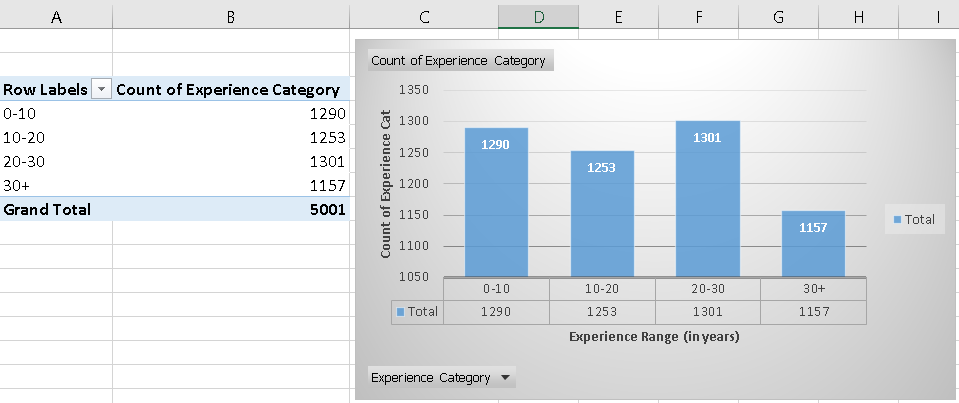
* ***Approach:*** 
  + Applied Descriptive Data Analysis on above columns
* ***Output:***



* ***Insights:***
  + We could see the maximum income and experience is 224k/yr and 43 yrs respectively

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]

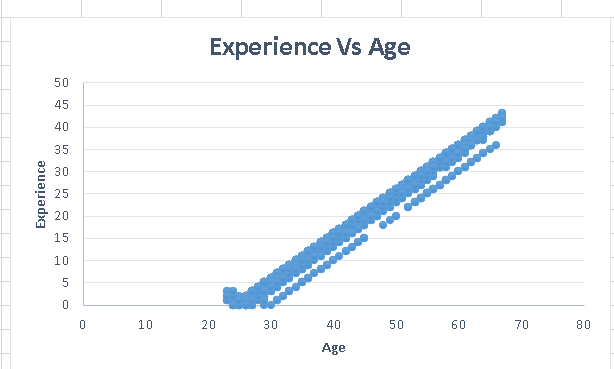
* ***Approach:*** 
  + Step 1: Create a New “Experience Category” using IF Else/And commands, with 4 Categories (0-10,10-20, 20-30, 30+).
  + Step 2: Used the Pivot table to see the aggregate result (Count of each category).
  + Step 3: Plotted the Bar graph
* ***Output:***



* ***Insights:***
  + The Category “0-10” and “20-30” is almost equal.
  + The Category 30+ experience is the minimum.
* ***Inference:***
  + The experience range of 20-30 is the maximum, with which we can say in the next 10 years Category 30+ will be maximum

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

* ***Approach:*** 
  + Selected Age and Experience column and selected Scatter Plot
* ***Output:***



* ***Insights:***
  + The value of Age Increases as the value of Experience also increasing
* ***Inference:***
  + Experience and Age are Linearly related if there is an increase in Experience there is an increase in Age

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

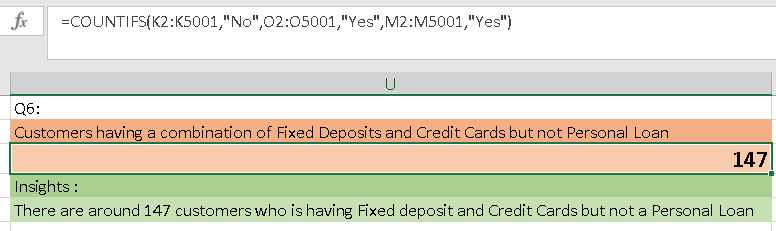
* ***Approach:*** 
  + Step 1: Inserted a pivot table, where I dragged “Zipcode” to row and “Id” to values
  + Step 2: Then used “Filter by values” and used “top 10” where I chose the top 3 values.
* ***Output:***



* ***Insights:***
  + The max Bank customers are located in ZipCode 94720
* ***Inference:***
  + The Top 3 zipcode areas contribute about 8% of total Bank Customer

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

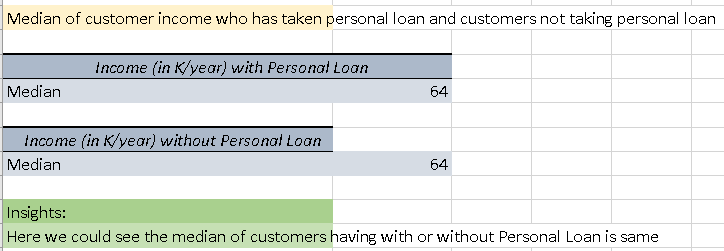
* ***Approach:*** 
  + Used “Countifs” approach to find the count of Customers who is having Fixed deposits and Credit Cards but not a Personal loan
  + *=COUNTIFS(K2:K5001,"No",O2:O5001,"Yes",M2:M5001,"Yes")*
* ***Output:***



* ***Insights:***
  + There are around 147 customers who is having Fixed deposit and Credit Cards but not a Personal Loan
* ***Inference:***
  + Since the customer is has both a Fixed deposit and Credit Cards but not a personal loan, he can be eligible for another Credit card since he has an FD.
  + The above customer is also eligible for a Personal loan since he has a FD

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?

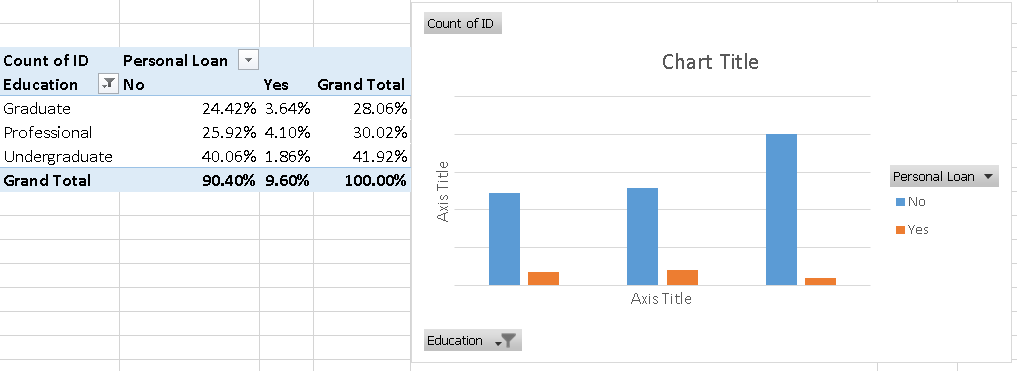
* ***Approach:*** 
  + Step 1: Applied filter to the customers having a personal loan or not
  + Step 2: Applied descriptive data analysis technique to find the median for customers having personal loans or not
  + Step 3: Kept only the median row and deleted rest of rows
* ***Output:***

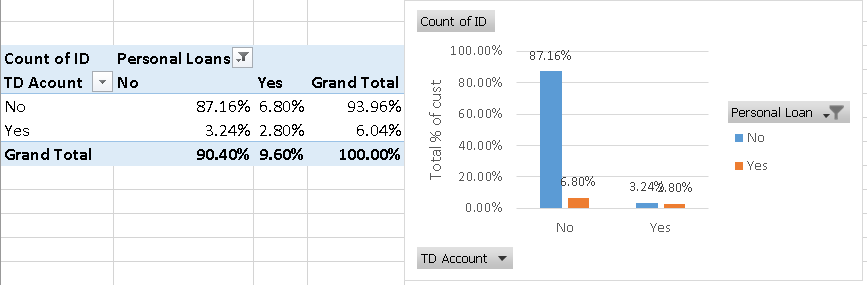


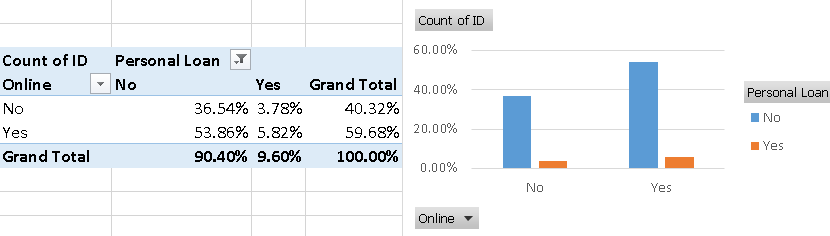
* ***Insights:***
  + Here we could see the median for customers with or without Personal Loans is the same
* ***Inference:***
  + We could say that the customer having an income of 64K/year can either take or not take the personal loan, since the median of customers with or without personal loan is same

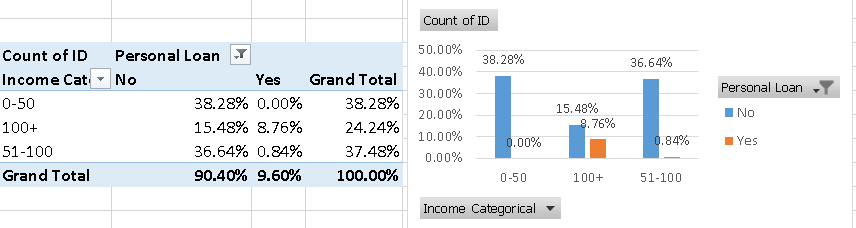
8. Create 4 separate Pivot Tables. Summarise your data by percentages. Education vs Personal Loan TD Account vs Personal Loan Online vs Personal Loan Income\_Category vs Personal Loan [Hint: Please drag Personal Loan to the Columns area while creating the Pivot Table to get the required values]

* ***Approach:*** 
  + Used pivot table in getting the insights
* ***Output:***









* ***Insights:***
  + Almost 40% of Undergraduates has not taken a Personal Loan
  + We could see around 90% of Customers has not taken a personal Loan
  + The Education has Professional has take more loans when compare to others.
  + We see that around 87% of customers don’t haveTD Account and has not taken the personal loan
  + Around 3% of customer has both TD account and Personal Loan
  + Around 7% of Customers is have Personal Loan but not TD Account
  + Around 6% of customers who are having the Online has taken Personal Loan
  + Around 54% of customers has Online but not taken the personal loan
  + We could see none of customers from 0-50 k income category has taken a personal loan
  + Around 1% of customers in the category of 50-100k has taken a personal loan
  + Income category of 100+ has the majority contribution in taking a personal loan.
* ***Inference:***
  + Target Professionals and Graduates, they are more likely to take personal loan
  + Don’t focus more on Undergraduates, they have less chance of taking loan
  + Target more of customers who don’t have TD Account, they are more keen to take loan
  + Stop Approaching customers who don’t have TD Account, the success rate here is less than 10%
  + Around 50% of customers having TD account has a chance of taking personal loan, this would be more easy and time effective since the count of customers having TD account is less. We have almost 50% of succes rate targeting them
  + Around 10% of customers having Online account has the chance of taking personal loan
  + Around 10% of customers not having Online account has the chance of taking personal loan.
  + Overall there is around 10% success rate in targeting the customers who have or not have an Online account
  + Concentrate on giving loan to the category of 100+ where around 40% of this category has not taken loan
  + Target customers having a income category of 100+
  + We could see the success rate for customers of income category 100+ is more than 50% for taking personal loan
  + Stop Targeting income category of 0-50 since they have 0% chance of taking personal loan

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

* Anomaly Observed:
  + Don’t focus more on Undergraduates, they have less chance of taking loan
  + Stop Approaching customers who don’t have TD Account, the success rate here is less than 10%
  + Stop Targeting income category of 0-50 since they have 0% chance of taking personal loan
* Customers who are most likely to take personal loan
  + Target Professionals and Graduates, they are more likely to take personal loan
  + Around 50% of customers having TD account has a chance of taking personal loan, this would be more easy and time effective since the count of customers having TD account is less. We have almost 50% of succes rate targeting them
  + Overall there is around 10% success rate in targeting the customers who have or not have an Online account
  + Concentrate on giving loan to the category of 100+ where around 40% of this category has not taken loan
  + We could see the success rate for customers of income category 100+ is more than 50% for taking 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?

* Target:
  + Professional and Graduates
    - Reason: Around 8 percent of 58% entire population are professional and Graduates who have taken personal loan
  + TD Account Holder
    - Reason: Around 50% of customers having TD account has a chance of taking personal loan, this would be more easy and time effective since the count of customers having TD account is less. We have almost 50% of succes rate targeting them
  + Online Account holder
    - Reason: There is a 10% success rate for targetting both customers having or not having a Online account
  + Income category of 100+
    - Reason: We could see the success rate for customers of income category 100+ is more than 50% for taking personal loan
  + 147 customers having both FD and Credit Card but not personal loan
    - Reason: Target These 147 are more likely and eligible to take personal loan, since they have enough credit score and a deposit
  + Concentrate on top 3 zip code areas
    - Reason: The Top 3 zipcode areas contribute about 8% of total Bank Customer
  + 50% of population whose whose income is greater than 64k which is median
    - Reason: we have seen that income cat of 100+ has more probability of getting Personal Loan, and 50% of customers is having income less than 64K