**Task 2: EDA Business Analysis and notes**

1. **Analysis**
2. The dataset has 100000 row and 28 column which does not include any duplicated rows, however, it has 60071 missing cells which forms about 2.1% of the dataset.
3. The data includes 8 numerical features, and 20 categorical features which will need to be handled in the pre-processing task.
4. Some columns could be deleted such as the ID, Name ,and SSN columns as they will not benefit our model by any means.
5. The Customer\_ID column divides the dataset into 12500 customers, each Customer\_ID maps and strongly relates to some other features, this is why this column will be vital while filling the missing values.
6. Some columns such as Num\_Bank\_Accounts and Num\_Credit\_Cards have un-logical values which could be considered as outliers such as one person has over 1700 account in the same bank.
7. Some columns also have negative values such as the age column which includes a row with a person aging -500 years.
8. Some columns such as the Num\_Bank\_Accounts, Num\_Credit\_Cards, Intrest\_Rate, Num\_Credit\_Iniquires and Total\_EMI\_Per\_Month are skewed.
9. **Business Problems:**
10. The data entry team has some serious problems, as they created some row values that are not meaningful, such as the value named “NM” in the Payment\_of\_Min\_amount column and the “-“ value in the Credit\_Mix, which will mislead other teams and affect the quality of the dataset which automatically negatively affects the model output.
11. Another more problem done by the data entry team is in the amount\_invested\_monthly column that includes 4300 entry with the value of “\_\_10000\_\_” which is a string instead of float.
12. The “Good” classified customers of the bank are around 17800 out of 100000 customers, which clearly indicates that the bank has a problem in its target audience, and in the marketing team.
13. We can also state that the marketing team needs to change its marketing campaigns and strategies to attract more “Good” and appropriate customers.
14. The Loan Office team also needs to reconsider its strategy of giving loans as most of the customers doesn’t pay back their loans in the specified due dates.