Dear Client,

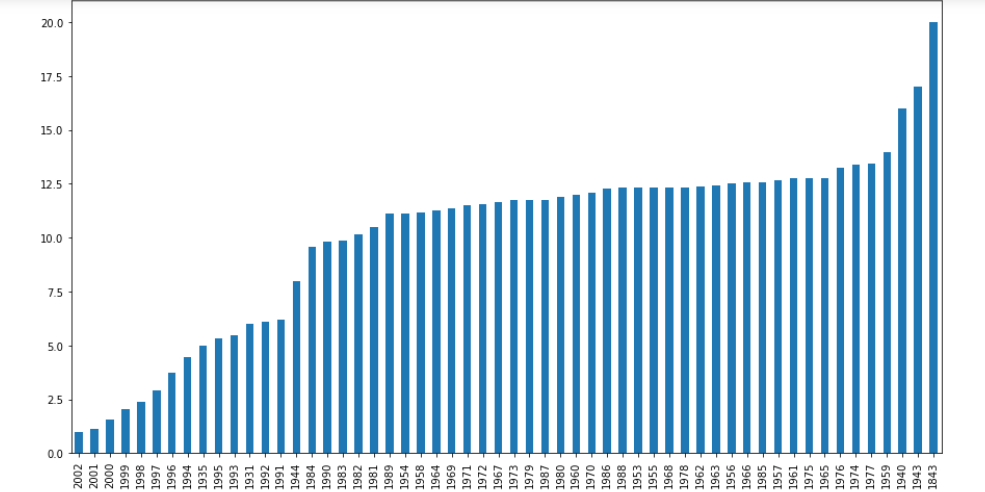
We as an organisation thank you for your recognised trust on us for betterment of your business model. Going through your data, I found following disparities which requires your attention for better prediction purpose.

This is the primary observation for the tables given.

|  |  |  |  |
| --- | --- | --- | --- |
| Name of the Table | Number of records | Number of unique records | Number of unique customer id |
| Customer Demographic (CD) | 4000 | 4000 | 4000 |
| Customer Address (CA) | 4000 | 3999 | 3999 |
| Transaction (TR) | 20000 | 20000 | 3494 |

Disparities:

1. There are 3 such customer addresses present in CA table which customers are not present in CD table.
2. There are missing values in different columns which can be indication of data losses. If not so, mention it so that we can derive a method to treat them. (Refer report for finding these columns and missing values.)



1. It seems like there are some mistyping regarding gender of the customer in ‘gender’ column. (Refer to report to find such mistyping)
2. There is one customer’s date of birth present in ‘DOB’ column of CD table which seems impossible. (Refer report for detailed analysis)
3. The column ‘default’ in CD table which seems arbitrary. If not provide some description for that.
4. More explanations required for certain columns. (Refer report for better understanding)
5. Other problems are given in detailed report which is attached below, if my understanding regarding them is not consistent with your data collection mention that.

Thank you

Sai Sankar Chakradhar Maddali

KPMG