**Data Quality**

Dear Client,

Thanks, for providing datasets and during processing of datasets, we get the following Over view.

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Number of Records** | **Unique Id’s** | **Dataset**  **Receiving date** |
| Transactions | 20, 000 | 3494 | 11 / 05 / 2021 |
| CustomerDemographic | 4000 | 4000 | 11 / 05 / 2021 |
| CustomerAddress | 3999 | 3999 | 11 / 05 / 2021 |

**Issues related with datasets**

* In-consistant data values with same attribute
* Duplication of data
* Missing values
* Incomplatness of Data

**Ex.** Gender features male , female and f

Where, f is incompateness of female

* **Transactions dataset**
* Transactions dataset have 16506 duplicatr rows found
* Transactions dataset
* After remove Duplicates rows then
* Numbre of features have missing values avalible
* **CustomerDemographic datset**
* Number of Features have Missing values avalible
* **CustomerAddress dataset**
* Its is good accuracy of data

We are moving forward for further process like data cleaning, standardlization, etc and it will be great to spend some time with your data.

Regards,

Alpesh Patel