Dear Client,

First of all we would like to Thank you for providing the datasets, and during processing of these data sets, we get the following views.

No. of Records No. of Unique Receiving Data

Records

|  |  |  |  |
| --- | --- | --- | --- |
| **Transaction dataset** | **20002** | **3495** | **15/03/22** |
| **Customer Demogaphic** | **4002** | **4002** | **15/03/22** |
| **Customer Address** | **4001** | **4001** | **15/03/22** |

**Issues found in the dataset:**

* DOB column in “Customer Demographics” had few inappropriate values.
* Grammatical error in the ‘job\_industry\_category.’
* Null Values in many columns.
* There were many customer\_ids in Transactions which were not present in Customer Demographics and Customer Address sheets.

# Consistency should be maintained in all the values of each attribute.

# Inconsistent data type for the attributes

# Records with ‘Y’ in deceased\_indicator in “Customer Demographics” are irrelavent.

# Records with ‘Cancelled’ order\_status in “Transactions” data are irrelevant.

# Ensure all the columns are assigned proper data types.

# “Now will move forward to data cleaning etc”

# Regards,

# Potlacheruvu Satya Sai Varun