Dear Client,

Thanks for sharing the datasets, during preprocessing on datasets we have identified the following as of 27th October 2020

Below table depicts the Data Quality of datasets:

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Customer Demographic** | **Customer Address** | **Transaction Data** |
| **Accuracy** | DOB: Inaccuracy  Job Industry Category: Misspelling |  | Standard Cost: $ missing |
| **Completeness** | Last Name: Blanks  DOB: Blanks  Job Title: Blanks  Job Industry Category: Blanks  Tenure: Blanks  Customer Id: Not sync | Customer Id: Not sync | Online Order: Blanks  Brand: Blanks  Product Line: Blanks  Product Class: Blanks  Product Size: Blanks  Standard Cost: Blanks  Product First Sold Date: Blanks  Customer Id: Not sync |
| **Consistency** | Gender: Inconsistency | State: Inconsistency |  |
| **Currency** | Deceased Customers: Filter Out |  |  |
| **Relevancy** | Default: Exclude field |  | Order Status: Exclude Cancelled |
| **Validity** |  |  | Product First Sold Date: Format |

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

Regards,

Jaya