**Issues of Consistency and Validity in Transactions**

1) One is the date datetime64 in format MM/DD/YYYY. Whereas, YYYY-MM-DD date format is used to capture DOB of customers. It would be better if it is kept consistent.

2) product\_first\_sold\_date is a date and should be datetime64 and must be in the standard format.

3) Another one is Online Order which is captured in a column of float64 datatype, however the values are boolean, that is true and false.

**Issues of Completeness in Transactions**

1) Order Online columns has about 1.83% of null values. There are 360 records in which order\_online was not captured.

2) Columns brand, product\_line, product\_class, product\_size, standard\_cost, product\_first\_sold\_date also has a percentage of 0.995% missing values that is 197 null values, which should not be missing if product\_id is inherited and the details of the product cannot be missing.

**Highlights of Accuracy in Transactions**

A single product ID should be referencing a single product with unique values.

**Issues of Consistency and Validity in NewCustomerList, Customer Demographic and Customer Address**

1) Structure format of NewCustomerList must be consistent with Customer Demographic and Customer Address.

2) There is no customer\_id in NewCustomerList.

3) Number of columns are inconsistent because in NewCustomerList there are 4 columns which are Unnamed and they contain some values as well, however are not labeled so cannot be identified.

4) There is one column in NewCustomerList which is 'Value', it is captured in a column of float64 datatype but this was not captured before and is not present in CustomerDemographic or CustomerAddress.

5) There is one column named default in CustomerDemographic, it is captured in a column of object datatype, some values are observed to be date values but this was not captured after and is not present in NewCustomerList.

6) DOB is the date column datetime64 in format YYYY-MM-DD. The date format used to capture transaction date in Transactions is MM/DD/YYYY. It would be better if it is kept consistent.

7) Rest of the columns are in object data type values but, deceased\_indicator must have contain boolean like True and False.

Data Captured in Gender column in the dataset CustomerDemographic is not consistent. It should be "Male", "Female" and "U" as per the NewCustomerList.

**Issues of Completeness in NewCustomerList, Customer Demographic and Customer Address**

1) In NewCustomerList 19.76% of job\_industry\_category values are missing almost similar to CustomerDemographic which is 19.61%.

2) 11.85% of job\_title values are missing in NewCustomerList a little less as compared to CustomerDemographic that has 14.48% of missing values.

3) 3.22% of last\_name values were missing in CustomerDemographic yet 2.98% of last\_name values are missing in NewCustomerList.

4) CustomerDemographic has 2.22% of missing DOB values which is slighlty decreased to 1.72% NewCustomerList.

5) There is a 2.22% of missing tenure values in CustomerDemographic but there is no missing values of tenure in NewCustomerList.

6) There is 1 missing record of address of customer\_id = 3 in CustomerAddress, as per identified by the shape of the datasets.

**Issue of Accuracy in NewCustomerList, Customer Demographic and Customer Address**

1) One date value in the column DOB is wrong. 1843 year is not possible.

**Highlights of Uniqueness in All Datasets**

All records are unique.