Dear Client,

I would like to thank you for giving these datasets, during the assessment of these datasets we got the following overview.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Datasets | No of records | Unique id’s | Accuracy | Completeness | Consistency | format |
| Transaction Dataset | 20002 | 3495 | Profit missing | Online order: Blanks  Brand: Blanks |  | Product Sold date: Format.  List Price:  Format |
| Customer  Demographic | 4002 | 4002 | DOB inaccurate | Job titles: Blanks | Gender:  inconsistency |  |
| Customer  Address | 4001 | 4001 |  |  | States:  Inconsistency |  |

Please find the below in-depth analysis of the datasets

Transaction dataset:

1. Customer\_id column might be facing some data loss.
2. There are blank cells in Online order and Brand column.
3. Product sale date and list price doesn’t have the right format.

Customer Demographic:

1. Inaccuracy in the date column. Creating additional column of Age can help the find the errors.
2. There are Blank cells in Job titles.
3. Inconsistency observed in Gender column.

Customer Address:

1. Inconsistency observed in States column.

That summarizes the above mentioned all the data quality issues that we found during the first stage of data quality analysis.

Please let me know if there are any issues related to the data analysis that is done.

Kind regards,

Akash Nair.