Hi,

Based on the dataset you have sent us,here is a report that I have found out about the data quality.

Regards

Araf Mustavi

**Report on the Customer Dataset**

|  |  |
| --- | --- |
| Customer Demography | |
| Customer ID | Total 4000 Entry and all are good |
| Full Name | 2 Duplicate Values, Last name is missing in some case |
| Gender | 7 of the Entries are Blank |
| Date of Birth | Customer ID 34 has bad date of Birth, 87 Entries are empty |
| Job Title | 507 Entries have no job title |
| Default | Irrelevant data in every cell,240 empty |
| Tenure | 88 empty |

|  |  |
| --- | --- |
| Transactions | |
| Customer ID | 3 of the entry has 5034 ID whereas exists 4000 |
| Online Order | 360 Empty |
| Brand + line class + size + standard cost | 197 Entries are Empty |
| First Sold Date | Irrelevant Data, shows number |

|  |  |
| --- | --- |
| New Customer List and Addess | |
| Last Name | 29 Empty row |
| Past 3 year bike related purchase | Data Stored as Text, Irrelevant data |
| Postcode | Non-homogenous data |
| Date of Birth | 17 Entries are empty |
| Customer ID | 4001-4003 is extra 3 ID |