Hello Team

Sprocket Central Pty Ltd ,

Below are details of data quality issues :

**Common Data Quality Issues:**

Check for duplicacy

|  |  |  |
| --- | --- | --- |
| **CustomerDemographic** | | |
| **Column name** | **Issues (Yes/No)** | **Issue Description** |
| customer\_id | No |  |
| first\_name | No |  |
| last\_name | Yes | Missing values can be **ignored** as there are cases where customers use first names only |
| gender | Yes | Data is inconsistent with Male, Female, M, F, U. It must be either F,M,U or Male, Female, Unmarried. |
| past\_3\_years\_bike\_related\_purchases | No |  |
| DOB | Yes | DOB is an important client detail, with its missing values, we can’t focus on the type of population we are interested in. |
| job\_title | Yes | Missing values can be **ignored as this field is not as important as other fields** |
| job\_industry\_category | Yes |
| wealth\_segment | No |  |
| deceased\_indicator | No |  |
| default | Yes | The column is unimportant as the values can’t be understood. |
| owns\_car | No |  |
| tenure | Yes | Missing values (NaN) are replace by 0 as no tenure data will be considered as no tenure. |