Dear Client,

Thanks for providing the datasets from Sprocket Central Pty Ltd. It is indeed a good opportunity to learn something new. While working on the datasests the following overview is drawn and necessary corrections were made. The table below highlights the key quality issues with data set. Please let us know if you have any issues.

|  |  |  |  |
| --- | --- | --- | --- |
| Transactions | There are blanks in the brand column | List price and standard price values are not accurate | Out of 20000 records 19445 Unique |
| New Customer list | DOB column had blanks in it | There is no accuracy in the “past 3 years bike purchases related column” | In the gender column data is inconsistent |
| Customer Demographic | Job titles column had blanks in it  And DOB is inaccurate | Gender column had inconsistent data | 3494 unique records are identified out of 4000 records |
| Customer Address | Property values are  inaccurate | State column data is inconsistent | --- |

Below are more in depth descriptions of the data quality issues discovered.

1. Accuracy: DOB in Customer Demographic is inaccurate and Profit is missing in the Transactions set.
2. Completeness: Job title has blanks and Customer Id is Incomplete in Customer Demographic and even in the Customer Address and Transactions sets, Customer Id’s are incomplete and Online Order , Brand columns has blanks.
3. Consistency: Gender and States of Customer Demographic and Customer Address are Inconsistent.
4. Currency: Deceased customers were filtered out.
5. Relevancy: Cancelled order status in the Transactions set was filtered out
6. Validity: Transactions set has invalid columns such as List Price and Product sold date.

We are now moving ahead with further cleaning and standardization of data. Excited to be a part of this and looking forward to learn more.

Best Regards,

Nalliboyina Yuva Raja Phani Kumar