

Hello [Sprocket Central Pty Ltd](#) ,

*Thank you for providing us with the dataset from Sprocket. After reviewing the dataset, we have identified several data quality issues and have taken relevant steps to address them.*

*The table outlines the data quality issues with the [Sprocket Central Pty Ltd](#) dataset.*

<b>Sheet Name</b>	<b>Data quality issue</b>
Transactions	Completeness and Relevancy.
NewCustomerList	Completeness and Relevancy.
CustomerDemographics	Completeness, Consistency and Relevancy

*Below is a summary of the issues.*

- 1. In the Transaction worksheet, we identified blank values for the "online\_order" and "brand" columns, and converted the "product first sold date" column into a date/time format. Removing blank values from the dataset is important for completeness and accuracy during modeling.*
- 2. In the New Customer List worksheet, we found blank values in the "second\_name" column and inconsistent values for gender. While blank values in "second\_name" are not a significant issue, inconsistent gender values were changed to "M" for male and "F" for female. The irrelevant "U" value was also removed. These were followed by more blank and null values in columns "job\_ttle" and "industry"*
- 3. In the Customer Demographic worksheet, we found inconsistency in gender values, missing values, and an irrelevant field called "default". We removed null values from "job\_title" and "job\_industry", and removed the "default" field.*