

Sprocket Central Pty Ltd Data Quality Assessment

Dear Sir/Madam,

Following our last conversation about the quality of the Sprocket Central Pty Ltd Data Set, I was able to identify a few issues with the data and made a few recommendations for change. Here are the quality issues and how they can impact our analysis:

1. Consistency issues

Some of the values were contradicting. For instance, the **Gender** column on the CustomerDemographic sheet had to contradict values for Male and Female values. Also, the **State** column had a few contradicting values representing the same states, with some bearing abbreviated values.

2. Accuracy

Some of the columns had non-accurate values, as defined by the data type. For instance, naming **list_price** and **standard_cost** with a text datatype is wrong, and could impact our analysis.

3. Completeness

Some of the column bearings – what in a DMBS, we would say is a primary key – were having empty values. For instance, the column **customer_id** under the CustomerDemographic sheet had blanks.

4. Validity

The **default** column under the CustomerDemographic's sheet had values that should not be allowed during an analysis.

5. Relevancy

The data had information about deceased individuals, which may not be valuable when making insights about the future of the company.

Conclusion

Poor data quality could flaw our analysis and lead to lost revenue. For instance, if Sprocket Central Pty Ltd comes up with a marketing campaign based on the current data (which has flaws), they may end up missing potential customers, hence missing conversions.