Subject: Data Quality Assessment and Cleanup Recommendations.

Dear Sir/Ma'am,

I hope this message finds you well. I am Jyoti and I wanted to share the results of my recent data quality assessment for the dataset provided by Sprocket Central PTY LTD. My objective was to identify and address data quality issues to ensure the data accuracy and usability for analysis and reporting requirements.
Here's a summary of my findings and the corresponding recommendations for each worksheet:

Transactions Worksheet:
1) The Brand, Product Line, Product Class, Standard Cost, and Product First Sold Date columns had missing   values.
Recommendation: Delete rows with missing values in these columns.
2) Some values in the **Standard Cost** column were not in the right currency format.
Recommendation: Standardize the currency format in the Standard Cost column.
3) The **Product First Sold Date** column was in the wrong date format.
Recommendation: Convert the date format in the Product First Sold Date column to the correct format.

Customer Demographics Worksheet:
 1)   The **Last Name** and **Tenure** columns had missing values.
 Recommendation: Delete rows with missing values in these columns.
 2)   The **Gender** column had inconsistent letters like "m," "f," "u," and "femal."
Recommendation: Standardize the gender labels to "M" for Male, "F" for Female, and handle any inconsistencies.
 3)   The **DOB** column had missing values and an invalid date value (e.g., 1843-12-21).
Recommendation: Delete rows with missing or invalid date values.
 4)   The **Deceased Indicator column** had values "n" and "y" instead of "No" and "Yes."
Recommendation: Standardize values in the Deceased Indicator column to "No" and "Yes."
5) The **Job Title** column had missing values and unnecessary letters like "I," "II," "III" after them.
Recommendation: Remove unnecessary letters and delete rows with missing job titles we can use nested Substitute to remove these unnecessary letters.
 6)   The **Job Industry** column had a value spelled as "Argiculture" instead of "Agriculture" and "n/a."
Recommendation: Correct the spelling and replace "n/a" with null values.
 7)   The **Default** column with different characters is meaningless for the analysis.
Recommendation: Consider removing the Default column if it does not provide valuable information.

Customer Address Worksheet:
1) The **State column** had consistency issues.
 Recommendation: Standardize the values in the State column for consistency.

Implementing these recommendations will significantly enhance the data quality and reliability for analysis.

Thank you & Best Regards,

Jyoti