Subject: Data Quality Assessment and Recommendations for Sprocket Central Pty Ltd

Dear Sprocket Central Rep,

Thank you for providing us with the three datasets from Sprocket Central Pty Ltd. Below is a summary of the datasets received:

Table Name	No. of Records	Distinct Customer IDs	Date Data Received
Customer Demographic	-insert value-	-insert value-	-insert value-
Customer Address	-insert value-	-insert value-	-insert value-
Transaction Data	-insert value-	-insert value-	-insert value-

Notable Data Quality Issues and Mitigations

1. Additional Customer IDs

Issue: Extra customer IDs in Transactions and Customer Address tables not present in Customer Demographic. **Mitigation**: Ensure all tables are synchronized. Use only customers in the Customer Demographic for analysis to avoid skewed results.

2. Missing Data

NewCustomerList Sheet: Last names, DOBs, job titles, job industry ('n/a'). **Transactions Sheet**: Online order, brand, product line, product class, product size, standard cost, product first sold date. **CustomerDemographic Sheet**: Last names, job titles, tenure, job industry ('n/a'), confusing default column data. **Mitigation**: Make critical fields mandatory, gather missing information, and impute missing values where necessary.

3. Inconsistent Gender Entries

Issue: Gender entries include 'Male', 'Female', and 'U'. **Mitigation**: Standardize entries to 'Male' and 'Female', correcting existing inconsistencies.

4. Invalid DOB Entry

Issue: DOB of 1843 in CustomerDemographic. **Mitigation**: Implement validation rules for DOBs.

5. Inconsistent Attribute Values

Issue: Different representations of the same value (e.g., address data like 'Victoria', 'V', 'Vic'). **Mitigation**: Use regular expressions for consistency and enforce dropdown lists for data entry.

6. Inconsistent Data Types

Issue: Numeric values mixed with strings. **Mitigation**: Convert strings to numeric values where appropriate and ensure data type constraints in the database.

Next Steps

We will continue with data cleaning, standardization, and transformation. We plan to consult with your data SME to ensure all assumptions align with Sprocket Central's understanding.

Please let us know if you have any questions or require further assistance.

Kind Regards,

Onyero Walter Ofuzim
Data Quality Assessment Officer
KPMG's Lighthouse & Innovation Team