Subject: Data Quality Assessment and Mitigation Strategies

Dear KPMG AU Team,

I hope this email finds you well. As part of our ongoing efforts to ensure data accuracy and reliability, I would like to bring to your attention some data quality issues identified during our recent review. Alongside these issues, I am also proposing strategies to mitigate them effectively.

|  |  |  |  |
| --- | --- | --- | --- |
|  | No. of records | Distinct Customer IDs | Date Data Received |
| Customer Demographic | 4002 | 4002 | 26/02/2024 |
| Customer Address | 4001 | 4001 | 26/02/2024 |
| Transaction Data | 20002 | 3495 | 26/02/2024 |

Below are the identified data quality issues and corresponding mitigation strategies for each:

1. Transaction Data:

Issue: Presence of blank values in the transaction table.

Mitigation Strategy: We have addressed this issue by removing the blank entries and introducing a new column for transaction date and profit. This enhances the completeness and accuracy of our transaction records.

2. Customer Demographic Data:

Issue: Inconsistent gender representation (e.g., 'F' for female and 'M' for male, 'femaleale' and 'maleeale').

Mitigation Strategy: We have standardized the gender representation by replacing 'F' with 'Female', 'M' with 'Male', 'femaleale' with 'Female', and 'maleeale' with 'Male'. This ensures consistency and clarity in gender classification across the dataset.

3. Customer Address Data:

Issue: Presence of blank values and inconsistent state abbreviations (e.g., 'VIC' and 'NSW').

Mitigation Strategy: We have addressed the blank entries and standardized state abbreviations by replacing 'VIC' with 'Victoria' and 'NSW' with 'New South Wales'. This improves the completeness and consistency of our customer address records.

Data Quality Framework:

To ensure comprehensive evaluation and improvement of data quality, we have utilized the Data Quality Framework Table and relevant criteria and dimensions provided in our resources.

By implementing these mitigation strategies, we aim to enhance the overall quality, consistency, and reliability of our data, thereby facilitating more accurate analysis and informed decision-making processes.

Should you have any questions or require further clarification on these matters, please feel free to reach out. Your collaboration and support in ensuring data integrity are greatly appreciated.

Best regards,

Umair Shaikh