Dear Sir/ Madam,

I hope this email finds you well. Thank you for providing us with the three datasets from Spocket Central Pty Ltd. I am pleased to present you with the summary of our recent data analysis, which aimed to uncover valuable insights and identify actionable recommendations based on the findings. Please let us know if you have any queries surrounding the issues presented.

**Summary Table**

|  | **Filtered Columns** | **Added Columns** | **Deleted Columns** |
| --- | --- | --- | --- |
| **Customer Demographic** | * DOB - blanks * Job title - blanks * Gender - data format * Deceased - filtered ‘Y’ | * Age | * Default |
| **Customer Address** | * State - data format |  |  |
| **Transactions** | * Online order - blanks * Brands - Blanks * Order status - filtered * List Price - data format * Product sold date - data format | * Profit |  |

**Descriptions:**

Based on the insights gained from the data analysis, we have formulated the following Explanations and recommendations:

**Customer Demographic**

* **DOB** - There were missing dates and these outliers were filtered out.
* **Job Title** - blanks are treated as incomplete data and can skew further analysis, so these were filtered out.
* **Customer id** - Filtered Customer IDs remaining with 3,413 of the 4000.
* **Gender** - Filtered all ‘M’ under category ‘Male’ and all F’ and ‘Femal’ under category ‘Female’.
* **Deceased** - People who are deceased ‘Y’ are not current customers hence they were filtered out.
* **Default** - Column delighted because it did not have useful information.

**Recommendations**

* **Age column** was created to easily see outliers of abnormal ages and these were filtered out.
* Ensure tables are up to date from the same time period.
* Create a drop down menu for gender. This minimises manual entry and human error.
* Check for incomprehensive data and delete or format like the Default column that was deleted.

**Customer Address**

* **States** - Filtered all ‘New South Wales’ under category ‘NSW’ and all Victoria’ under category ‘VIC’.

**Recommendations**

* Create a drop down menu for States. This minimises manual entry and human error.

**Transactions**

* **Customer id** - Filtered Customer IDs
* **Online Orders** - Filtered out ‘cancelled’ orders because we need information about orders that were made
* **Dates** - Format all dates to match in terms of data type
* **Brand** - Filtered out the blanks
* **List Price** - Format column to currency data type.

**Recommendations**

* **Profit column** was created to easily see outliers of abnormal ages and these were filtered out. This will assist in future monetary analysis.
* ‘Cancelled’ Order status is irrelevant because it can hinder the total number of customers over time since the analysis will be focused on Frequency, Recency and Monetary terms.
* Set the columns to the right data type format. This will help even when entering new data.

Please let us know your thoughts on the findings and recommendations presented in this summary. The mitigation strategies suggested are simple and effective ways of improving data quality for future analysis. We value your input and would appreciate any feedback or additional information that might assist us in refining our analysis.

Thank you for your time and consideration. We are confident that these insights and recommendations will contribute to improving Spocket Central Pty Ltd’s performance. Should you have any queries or require further assistance, please do not hesitate to reach out to us.

Best regards,

Shirley Nakalema