Dear Project Manager,

After reviewing the datasets provided from Sprocket Central Pty Ltd, there are some data quality issues that were encountered. The methods used to mitigate the inconsistent data are as follows;

**TRANSACTIONS DATA**

1. 360 blanks in the online order column.
2. 197 blanks in the brand column.
3. 197 blanks in the product line column.
4. 197 blanks in the product class column.
5. 197 blanks in the product size column.
6. List price isn’t in currency format
7. 197 blanks in the standard cost column.
8. Product first sold date column in a text format instead of date format.

**CUSTOMER DEMOGRAPHIC DATA**

1. 125 users had no surname data.
2. Inconsistent attribute representation in Gender (1 record of “Femal” instead of “Female”, 1 record of “M” instead of “Male”, and 1 record of “F” instead of “Female”). This was corrected for consistent attributes.
3. 87 blank entries in the DOB column.
4. A customer had ‘1843-12-21’ as their DOB which is illogical. This record was removed.
5. 506 blank entries in the job title column.
6. 240 blank entries in the default column.
7. Only 87 entries in the Tenure column.

**CUSTOMER ADDRESS DATA**

1. New South Wales and NSW in the state column refer to the same state. New South Wales was replaced with NSW.
2. VIC and Victoria in the state column refer to the same state. Victoria was replaced with VIC.

The team will continue in data cleaning and transformation processes in order to prepare the data for model analysis. This entire process is going to be documented.

Kind Regards,

Mevhare Afe.