Dear Sprocket Central Pty ltd

I hope this email finds you well. As a data analyst, I have assessed the data in your spreadsheet and identify different data quality issues. I would like to bring these issues to your attention and suggest strategies to mitigate them using Microsoft excel.

Transaction sheet:

- The "Product_id" and "customer_id" columns have a general data type, which should ideally be text. To address this,you can use excel's tesxt to column feature to convert data to appropriate format. Also, change the "transaction_date" and "Product_first_sold_date" is not in a proper date format and should be change to date format. The "list_price" and "standard_cost" column should be in currency data type.
- Approximately 360 rows in the "online_order" column and 197 in the "brand", "product_line", "product_size",columns are blank. Consider removing the rows or replace it with mean/median/mode.
- The entries in the "product_class" and "product_size" column should be in proper case . you can do so by using Excels PROPER function.
- It is ensured that there is no duplicates found in the data.

Customer Demographic Sheet:

- The gender column contains inconsistencies with entries such as "F", "M", "female" and "U". You can standardize the gender by replacing them with the consistent "Female", "Male" and "Unknown.
- There are 125 blank rows in the "last_name" column, you can either remove this row or impute missing values using suitable techniques.
- There are 87 blanks roe in the "DOB" column, with an outlier of 1843 as a date of birth. You may consider removing these rows or impute missing values using suitable techniques.
- There are 506 blank rows in "job_industry" column. You may consider removing these rows or impute missing values using suitable techniques.
- The default columns Is irrelevant and can be removed from the dataset to improve data cleanliness.

Customer Address Sheet :

- In the "state" column, there are inconsistences in the entries, such as "VIC" or "Victoria, "NSW", or "New Soth wales". To address this issue, you can use excel FIND AND REPLACE feature to standardize the state name uniformly

Best Regards,

Suliyat Oladimeji

Data Analyst