To,

Client

Sprocket Central Pty Ltd

Subject- Data quality issues and strategies to mitigate these issues.

Respected Sir/Ma'am,

Thank you for sending us the datasets from Sprocket Central Pty Ltd. While working on the datasets, we found some issues that are listed below-

1. Blank spaces in columns like
   Online order, brand, product_line, product_class, product_size, standard cost, product first sold date, DOB, job title, job industry, DOB, job title, etc.
   (Try to tally the data and fill in the empty spaces)
2. Unknown gender U in the "NewCustomerList" dataset.
   (Tally from the old data and try to figure out "U" gender)
3. Count of Transaction I'd, customers I'd, and unique customer I'd doesn't match in the "Transactions dataset".

Column Name	Count
transaction_id	20002
customer_id	20001
Unique customer_id	3495

4. Inappropriate Date of birth "1843-12-21" in CustomerDemographic dataset.
5. Unreadable Default column in CustomerDemographic dataset.
6. Inappropriate Date format in the "product_first_sold_date" column in the Transactions dataset.
   (Look for the data type of column)

   Regards,
   Unnati Singh