Dear client,

Subject: Regarding data quality issues

Hey I have been through the dataset that you provided. I explored the three dataset and I found some data quality issues. Talking about the 1st dataset Customer Demographic their are so many missing values like:

Last\_name - 125

DOB - 87

job\_title - 506

job\_industry\_category - 656

Default - 302

Tenure - 87

Even the gender column is not in standardised form like customer id 1 and 54 have female value as ‘F’ and ‘Femal’ respectively, unlike others then and their is a value ‘U’ in this column.

Issue with jephthah bachmann his DOB is in year 1843 which is practically not possible.

In the Transactions dataset the missing value column wise are:

Online\_order - 360

Brand - 197

Product\_line - 197

Product\_class -197

Product\_size - 197

Standarad\_cost - 197

product\_first\_sold\_date - 197

The product first sold date is not in the standardised date form.

Please go through the issues so the the data can be cleaned and data quality can be maintained.

Kind regards,

Namrata