Dear Client

Thank you for providing us with three datasets from Sprocket Central Pt Ltd.

Notable data quality issues that were encountered and the methods used to mitigate the identified data inconsistencies are as follows.

* Transactions Table:
  + There are lots of null values in 7 columns including online\_order, brand ,product\_line, product\_class etc
    - They can be imputed using statistical approaches .
  + Inconsistent data type as some are numerical like online\_order whereas categorical features like order\_status, brand, product\_class.
    - Categorical features must be converted to dummy variables.
* CustomerDemographic Table:
  + There are 6 columns with null values including DOB, job\_title, job\_industry\_category.
    - If there are small number of empty rows, they can be deleted.
  + Invalid type of data present in the column ‘default’.
    - Should be removed.
  + Gender column have inconsistency F, M, Male, Female are used.
    - Should be made consistent using only F and M.
* CustomerAddress Table:
  + Address column is of no use if data type remained object only.
  + It should be splitted into two parts numeric and string so that can be further used to analyze dataset.