Dear client,

Upon review of the customer spreadsheet file, several data quality issues were raised and recommendations have been issued to ensure that the accuracy and understanding of the data is maintained to a suitable level.

* In the customer demographic sheet, the DOB entry of one customer was 1843-12-21, this customer entry needs to be removed as well as the customers who have been deceased as indicated by the Y on the deceased indicator column.
* Out of 20,000 transactions, there are 198 transactions which do not have an entry for the product size. The interesting thing is that out of these 198 transactions, all of them do not contain any information regarding the bike and also refer to the same product (product ID=0). These fields will be removed from the dataset as they only constitute to approximately 1% of the total number of transaction records and the revenue.
* There are inconsistent entries for the gender (F, female, M, male). Ensure that the data entry options are consistent whether you use long or short text. A drop-down list will help this cause instead of entering the data manually. This is also the case for the state column in the customer address sheet (NSW, New South Wales, VIC, Victoria).
* Some data entries for the job title are blank whilst the job category for some of the blank entries are filled. For blank job title entries, ensure unemployed, student and retired are options as this could be the cause for the blank entries.
* On the customer demographic sheet, the default column contains cryptic text. This could be the cause of the wrong data entry format but instead the decision was made to remove this column as the column header did not specify what information was to be contained.
* For the tenure column, some entries are blank which could represent clients which have owned a property for less than a year.

Some data entries for the last name are blank, if the last name is not applicable for a client, enter N/A.

I hope you find these problems and the subsequent solutions useful and please do not hesitate to email me if there is an issue.

Many thanks

Ansar lala