

Dear Client,

Based on the dataset that you have provided, I have found a couple of issues with the quality and have some suggestions.

1) Completeness

- a) In the Transactions Sheet: online_order, brand, product_line, product_class, product_size, standard_cost, and product_first_sold_date contain null values.
- b) In the NewCustomerList sheet: DOB, job title, and job industry contain null values.
- c) In the CustomerDemographic sheet: DOB, job title, job industry contain null values.

To fix this, we can try to look at past records to fill in the null values, if possible.

2) Relevancy (We can get rid of some irrelevant values)

- a) In NewCustomerList and CustomerDemographics, we can delete "U" under gender because it doesn't identify whether they are a female or a male.
- b) We can also delete entries that are not alive under deceased_indicator because we are examining alive customers.

We can try to determine the meaning of "U" for gender to find the correct values.

3) Accuracy (All numbers are in the correct format)

- a) In the Transactions sheet, the product_first_sold_date can be changed from number to date to accurately represent the data type.

Checking the data types of each column can fix this issue.

4) Consistency

- a) In the CustomerDemographics sheet, we can change the values of gender to be consistent - "F" is female, and "M" is male.
- b) In CustomerAddress, we can change "New South Wales" to "NSW" and "Victoria" to "VIC"

It would be best practice to check the entries that are in the field to remain consistent.

I hope this information helps.

Thanks,
Willa