Dear Client,

After receiving the dataset provided by Sprocket Central Pty Ltd, I noticed some issues in the data that affect it's quality. So, I studied the data quality across 7 major criteria, which are shown in the below table:

	Accuracy	Validity
Transaction Table	product_first_sold_date contains values that are incorrect, contains number rather than date.	standard_cost there are different data types are storing the data, one by the dollar sign and the other by yuan sign.

	Consistence
CustomerAddress Table	State column have the same category but with another abridgment that affecting the understanding of the data.

	Consistence	Accuracy	Validity	Currency
CustomerDemographic Table	- Gender column has a lot of categories that seems like not correct and in a missy form. - Some data fields have Null values.	DOB there a column have unreal age, where it cannot make sense.	There is a default column, which have unrelated symbols where is no needs of it at all, there is no statistics, relationships or planning for it with any.	DOB one of the date haven't stored as the whole dates in the column, the format of data stored such : YYYY/MM/DD Whereas the least format stored is : YYYY-MM-DD .

Here are some suggestions and strategies to solve these issues and get ready to deal with a clean and accurate data : -
In Transaction table, the standard_cost column must changing the currency calculated with to be only one.
- In Costumer address table, In State column, preferable to select only one clear landmark/keyword for each state are exist.
- In CostumerDemographics table:
- Unreal age in DOB will be deleted.
- Use one data type for DOB that is YYYY/DD/MM.
- Gender can be replaced by 3 categories only (Female, Male, U).
- There is no need to have default column so, deleting it more appropriate.
- Add new column, that calculate the age of the customer, that will be more clearly.

Kind regards,
Maram M. Alelyani.