Dear Sir,
I have gone through the data provided by your team. As discussed, I have thoroughly searched and scanned the data and have found several data quality issues. These issues need to be solved before we perform any further analysis on the data sets. I have listed the issues below kindly look into it.

The data quality issues are:

1) Transaction Sheet:
    - The **product_first_sold_date** column's data type was not in the correct format. It was in integer format whereas it should have been date
    - The **list_price** column's data type was not in currency mode.
    - The standard_cost column's data type was not standard to 2 place decimals
    - There are blanks in the column **online_order** which must be removed
2) Customer Demographic sheet:
    - Blanks in **last_name** column
    - Multiple names for Female and Male. No standard convention followed.
    - Several blanks in **job_title** column
    - Several **job industry** categories are declared as n/a. Rather it should be declared with some category
    - **Blanks** in tenure column
3) Customer Address sheet:
    - The **state** column still had full form names which are yet to be changed to their short form abbreviations

Further there is a discrepancy in duplication of data in the three sheets:

Sheet	**Total Records**	**Distinct Customer id's**
Transaction	20001	3495
Customer Demographic	4002	4002
Customer Address	4001	4001

As you can see from the above table, there are additional customers in the Transaction table than the Customer Demographic and Customer Address Table. It would be best of you, if you provide us with a final list of customers which is common to all the three sheets where we would finally run our models.

We would perform several data cleaning tasks to clean the data as per the issues and further would be performing several analyzing tasks on it.

Regards
Ved Thakur