Questions to be answered

Agenda: Given the sales transaction data we want to assess the performance of the company.

1. Is the company's performance improving or degrading over time?
2. Examine and highlight important trends visible in the sales data and insights.
3. How can we measure our performance in terms of customer acquisition and building customer loyalty?
    a. What kind of customer do typically buy from us?
    b. Identify relationships and drivers of sales that might be hidden in the dataset.
4. Can we take some initiatives based on the data to increase the sales? Also mention, based on data can we avoid out of stock situations?

A good way to present the analysis:

1. First give an overview of the data, then go on to formalise the question(s) you want to answer.
2. Mention your approach (metrics/visualizations/analysis) to answer the above questions.
3. Explain why you chose the metrics/visualizations/analysis you chose and how will it help to answer the questions raised.
    a. In addition, explain why it is important to track this metric for the business.
4. Finally, present the analysis with metrics, visualizations and insights answering the questions. Also mention any sort of data cleansing you had to do (if any).

The expected output is a google slide/ppt containing the overall structure including questions, reasonings, visualizations & insights around the analysis.

Tables

sales_data.xlsx

transaction id: a 6-digit integer to uniquely identify each transaction. A prefix of 'c' in the transaction id indicates a cancelled transaction.

product id: Unique identifier of the item sold.

product description: Name of the product being sold.

quantity sold: quantity of the item sold.

transaction timestamp: the time at which the transaction happened. This column can be used to join this table with the date.xlsx to obtain attributes of timestamp.

unit price: per piece price of the item sold.

customer id: Unique identifier of the customer who made the purchase.

transaction country: Country where the product was sold.

date.xlsx

timestamp: self explanatory

date: self explanatory

day_name: self explanatory

day_of_month: self explanatory

month_of_year: self explanatory

time_of_day (hh:mm:ss): self explanatory