TataO Online Retail Data Mining:

The senior management wants to understand how their business is performing and what areas are the key strengths of the company. They are also focused on identifying opportunities that would lead to growth and generate more revenue in the future.

1. Data Source: Available from: Daqing Chen, Sai Liang Sain, and Kun Guo, Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining, Journal of Database Marketing and Customer Strategy Management, Vol. 19, No. 3, pp. 197â€“208, 2012 (Published online before print: 27 August 2012. doi: 10.1057/dbm.2012.17).

<https://archive.ics.uci.edu/ml/citation_policy.html>

1. Software used:
   1. Excel
   2. PowerBI
2. Data Process
   1. Problem statement and understanding of need-
      1. Gathered requirements, and understood metrics/measures required.
   2. Data Inspection & Cleaning
      1. Checked for validity (I.e price should not be less than or equal to zero and quantity less than zero)

*Removed data that fell into this set*

* + 1. Checked for completeness (missing data)- removed null rows

-Customer ID and Description had empty rolls

-In the country column, some revenue where unspecified (not allocated to a country), there is not enough information to delete or say it is irrelevant. Would need to elicit more information in this regard.

* + 1. Data type conversation for validity-

Quantity (should be whole numbers not text), country (categorized as country- Geo), customer ID (from number to text)

* + 1. Checked for uniformity in format and measurements- Good
    2. Enriching, create new column- Revenue
  1. Data Integration – using Power BI- Power query
  2. extraction – ETL
  3. Descriptive Statistics
  4. Visualisation
  5. Dashboard reporting to generate insights
  6. Insights the CEO would be interested in:
     1. Which region is generating the highest revenue, and which region is generating the lowest?
     2. What is the monthly trend of revenue, which months have faced the biggest increase/decrease?
     3. Which months generated the most revenue? Is there a seasonality in sales?
     4. Who are the top customers and how much do they contribute to the total revenue? Is the business dependent on these customers or is the customer base diversified?
  7. Insights the CMO would be interested in:
     1. What is the percentage of customers who are repeating their orders? Are they ordering the same products or different?
     2. For the repeat customers, how long does it take for them to place the next order after being delivered the previous one?
     3. What revenue is being generated from the customers who have ordered more than once?
     4. Who are the customers that have repeated the most? How much are they contributing to revenue?

1. Power Point slides prepared
2. Presentation made to CEO and CMO with feedback received