**Final Project**

Contents

[Cleaning Data: 2](#_Toc171090456)

[Question 1) 4](#_Toc171090457)

[Final Visualisation 6](#_Toc171090458)

[Question 2) 7](#_Toc171090459)

[Final Visualisation 10](#_Toc171090460)

[Question 3) 10](#_Toc171090461)

[Final Visualisation 12](#_Toc171090462)

[Question 4) 12](#_Toc171090463)

[Final Visualisation 14](#_Toc171090464)

# Cleaning Data:

* In the Quantity column, unchecked values below 1 to remove invalid quantities.

A screenshot of a computer

Description automatically generated

* In the UnitPrice column, filtered to select rows where UnitPrice is greater than or equal to 0.

A screenshot of a computer

Description automatically generated

* Selected the StockCode column.
* Clicked on Split Column > By Digit to Non-Digit to separate text values from numeric values.

A screenshot of a computer

Description automatically generated

* Removed errors and null values from the whole dataset.

A screenshot of a computer

Description automatically generated

# Question 1)

* Went to the Add Column tab.
* Clicked on Custom Column.
* Created a new column named Revenue with the formula: [Quantity] \* [UnitPrice].

A screenshot of a computer

Description automatically generated

* Selected the InvoiceDate column.
* Went to the Add Column tab and wrote Year (2010) to create a separate column for the year.

A screenshot of a computer

Description automatically generated

* To meet the CEO's requirement to view granular data by looking into revenue for each month, I created a line chart visualization by dragging InvoiceDate to the X-axis and Sum of Revenue to the Y-axis. From the InvoiceDate hierarchy, Month is selected to display the data as required by the CEO.

A screenshot of a computer

Description automatically generated

* Additionally, the separately created Year column is dragged to the filters pane and checked 2011, as the CEO was interested in viewing data for the year 2011 only. This setup allowed the CEO to view the monthly revenue trends and analyze seasonal patterns effectively.

A screenshot of a graph

Description automatically generated

## Final Visualisation

A graph on a computer screen

Description automatically generated

# Question 2)

* The Line and Column chart visual was selected to represent the data.
* Country was dragged to the X-axis, and Sum of Revenue was dragged to the Y-axis to display revenue per country.

A screenshot of a computer

Description automatically generated

* To filter the top 10 countries by revenue, a Top N filter type was applied based on Sum of Revenue. This ensures that only the top 10 revenue-generating countries are shown in the visual.

A screenshot of a computer

Description automatically generated

* As per the CMO's requirement to exclude the United Kingdom from the data, the United Kingdom was unchecked in the filter options.

A screenshot of a graph

Description automatically generated

* Sum of Quantity was dragged to the Line Y-axis. This step was essential to display the quantity sold alongside the revenue generated, providing a comprehensive view of both metrics together.

A screenshot of a graph

Description automatically generated

* This setup helps the CMO analyze revenue and sales performance across different countries effectively.

A graph showing the growth of a country

Description automatically generated with medium confidence

## Final Visualisation

A graph of a graph

Description automatically generated with medium confidence

# Question 3)

* Chose a Clustered Column chart to represent each customer's revenue as individual bars. Organized customers horizontally by placing CustomerID on the X-axis, and Revenue on the Y-axis to display each customer's total revenue vertically, as required.

A screenshot of a computer

Description automatically generated

* Selected the categorical type in x-axis to view the bras effectively.

A screenshot of a computer

Description automatically generated

* Applied a Top N filter to show the top 10 customers by revenue, sorted in descending order, as requested by CMO. Ensured data integrity by excluding any blank CustomerID values.

A screenshot of a computer

Description automatically generated

## Final Visualisation

A graph of a bar

Description automatically generated with medium confidence

# Question 4)

* The CEO aims to analyze countries with the highest demand for their products. For this purpose, a map visual was chosen where countries are mapped based on their location, and the size of each bubble represents the quantity of products sold.

A screenshot of a computer

Description automatically generated

* United Kingdom is unchecked as the CEO is more interested in viewing the countries that have expansion opportunities.

A screenshot of a computer

Description automatically generated

A screenshot of a map

Description automatically generated

## Final Visualisation

A map of the world with blue circles

Description automatically generated