Objective Questions

1. **What is the total number of attributes in the customer table?**
   * The Customer Table includes 3 following attributes:
     + **CustomerID:** A unique identifier for each customer.
     + **Customer Age:** The age of the customer.
     + **Customer Gender:** The gender of the customer (M for male, F for female).

# How will you get the “Customer’s” ages in the “Order” tables according

**to customer IDs?**

* + **Steps:**
    - In the Data View Clicked on the Order Table in the right-side of Fields panel.
    - Click on Modelling in the top menu and then select New column.
    - Enter the following formula:

***CustomerAge = RELATED(Customers[Customer Age])***

* + - After pressing Enter, a new column CustomerAge appeared in the Order Table.

# In analysing the dataset with Power BI, ensure data cleaning to address inconsistencies and missing values before further analysis.

* + In Power BI, we use Power Query Editor to:
    - Removing duplicates.
    - Handling blank rows by removing it.
    - Ensuring data types are correct.
    - Removing extra blank columns.

# How can we calculate the total revenue generated by all the sales?

* + I have created a Measure for Total Revenue by following steps:
    - Data View or Model View.
    - Clicked on the Order Table.
    - In the ribbon, clicked on New Measure and entered the following DAX formula:



# What is the total number of unique customers who made purchases each year? Is there any increase in the number over the years?

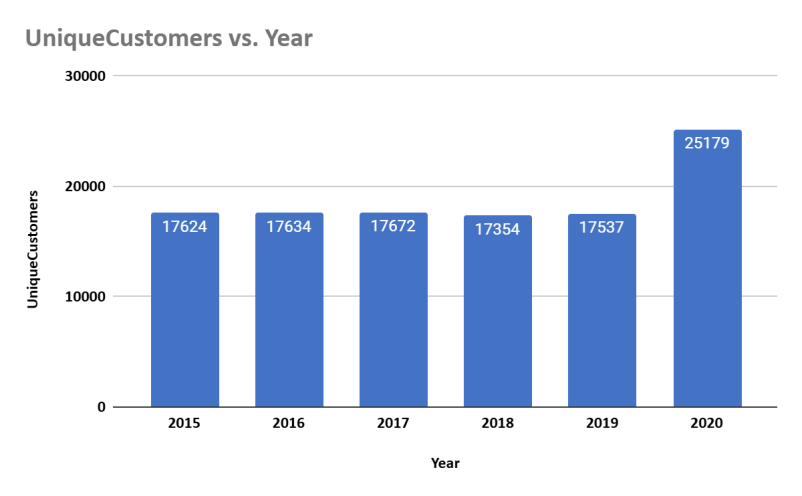
* + **Extract the Year from Order Date:**
    - Transform Data > Order Date > Add Column > Date > Year
    - This step has created a new column name Year representing the year of the order.

**Create a Measure for Unique Customers Per Year:**

* + - Data View > Order Table > New Measure
    - Entered the DAX formula:

***UniqueCustomersPerYear = DISTINCTCOUNT(Orders[CustomerID])***

**Visualization:**



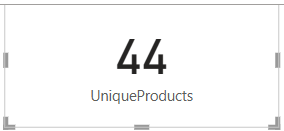
**Observations:**

* + - The unique customer count shows a steady increase from 2019 to 2020, suggesting effective marketing strategies or an expanding product range that attracts new customers.
    - A slight decline in customer count from 2017 to 2018 may indicate potential issues such as product availability, pricing changes, or increased competition.

# How can we determine the total number of unique products available in the company?

* + Followed these steps:
    - Data View > Order Table > New Measure
    - Entered the DAX formula:

***UniqueProducts = DISTINCTCOUNT(Orders[Product])***

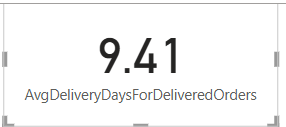


# What is the average number of days it takes for products to be delivered, get the metric for only the delivered orders.

* + Followed these steps:
    - **Created a New Column for Delivery Duration**
      * Data View > Select Orders Table > New Column
      * Enter the DAX formula:

***DeliveryDuration = DATEDIFF(Orders[OrderDate], Orders[Delivery Date], DAY)***

* + - * This formula calculates the number of days between the OrderDate and Delivery Date.
    - **Filter for Delivered Orders**
      * New Measure
      * Enter the DAX formula: ***AvgDeliveryDaysForDeliveredOrders=CALCULATE(AVERAGE(Or ders[DeliveryDuration]),Orders[Status] = "Delivered")***
      * This measure calculates the average delivery duration for orders where the status is "Delivered."

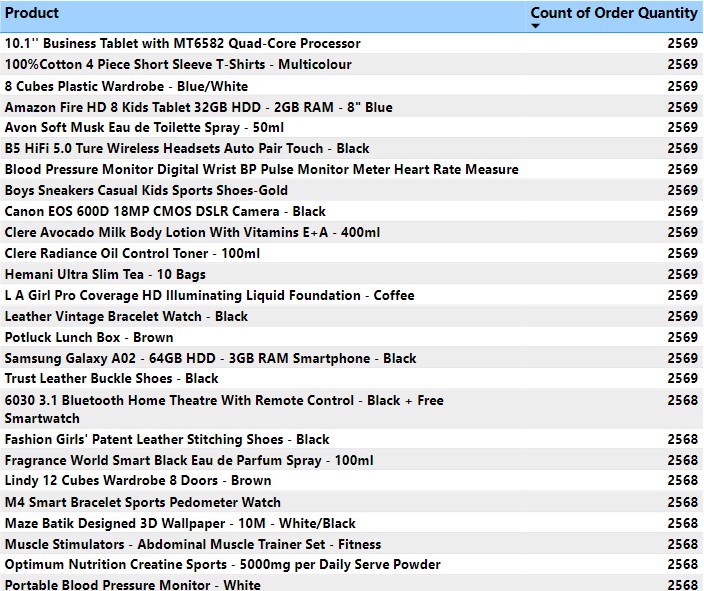


# Which products, categories, and subcategories are the most popular?

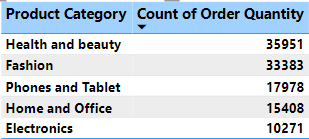
* + **Popularity by Order Quantity:**
    - Modeling > New Measure

***TotalOrderQuantity = SUM(Orders[Order Quantity])***

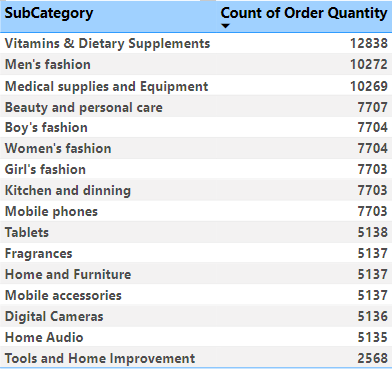
1. **For Most Popular Products:**
   * Add a Table Visual to the canvas
   * Drag the Product column into the table
   * Drag the TotalOrderQuantity
   * Sort the table by TotalOrderQuantity in descending order.



1. **For Most Popular Categories :**
   * Add a Table Visual to the canvas
   * Drag the Product Category column into the table
   * Drag Total Order Quantity
   * Sort the table by TotalOrderQuantity in descending order.

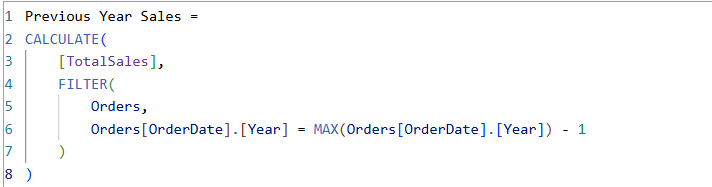


1. **For Most Popular Subcategories**
   * Add a Table Visual to the canvas
   * Drag the Subcategories column into the table
   * Drag Total Order Quantity
   * Sort the table by TotalOrderQuantity in descending order.



# Which products have seen an increase or decrease in sales over the year?

* + To analyse which products have seen an increase or decrease in sales over the years, I followed these steps:
    - Create a measure for Previous Year Sales

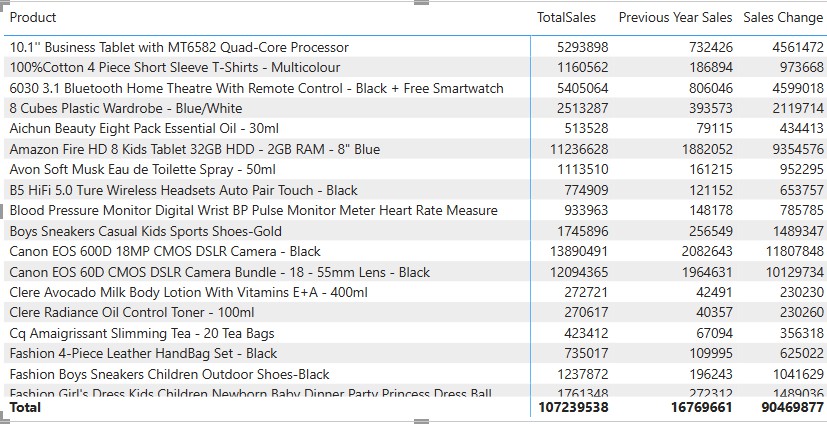


* + - Create a measure for Sale Change



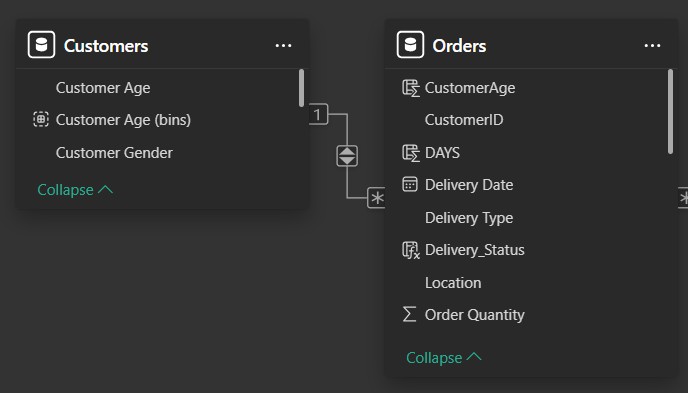
**Matrix Visualization**

* + - Rows – Product
    - Values – Total Sales, Previous Year Sales, Sales Change



# While modelling the data relationships, what will be the type of relationship between the customer ID of Orders and customer tables?

* + Model view > Drag the CustomerID column from the Orders Table to the CustomerID column in the Customers Table



While modelling the data relationships, the type of relationship between the customer ID of Orders and Customer table is “**one to many”** relationship.

# How have you handled the null values in the data?

* + **Using Power Query Editor to Filter or Remove Null Values**
    - Go to Power BI Desktop
    - Click on **Transform Data**
    - Select the column where I’ve to handle null values
    - Click on the **Filter Dropdown**
    - If null values exist, then uncheck "(null)" in the filter
    - This removes rows containing null values

**Handling Blank Values Using DAX**

* + - Open Power BI Report View
    - Click on the "**Modeling**" tab
    - Click **New Measure**
    - Create a DAX Measure for Null Handling. Use this formula to replace blank values in the "Sale Price" column"

***Sale Price Cleaned = IF(ISBLANK(orders[Sale Price]), 0, orders[Sale Price])***

**Removing Blank Rows**

* + - In Power Query, select the dataset
    - Click on Transform > Go to Row
    - Check if there are any fully blank rows
    - Click Remove Rows > Remove Blank Rows
    - This removes any row where all columns are blank

# Were there any data format issues in the data, and if there were/are how you would handle them?

* + In our data, there was no need to change the data types.

If there was such an issue to format the datatypes then I can check and transform data types in the Power Query editor.

# When we add a column in Power Query what’s the code that comes in M

**language in the formula bar? What do you know about M-query?**

* + **M Query** is the functional language used in Power Query to transform and prepare data before loading it into the Power BI data model.

When we add a column in Power Query, Power BI generates M query that reflects the action.

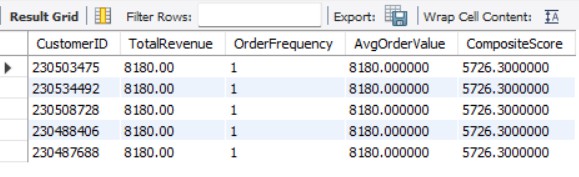
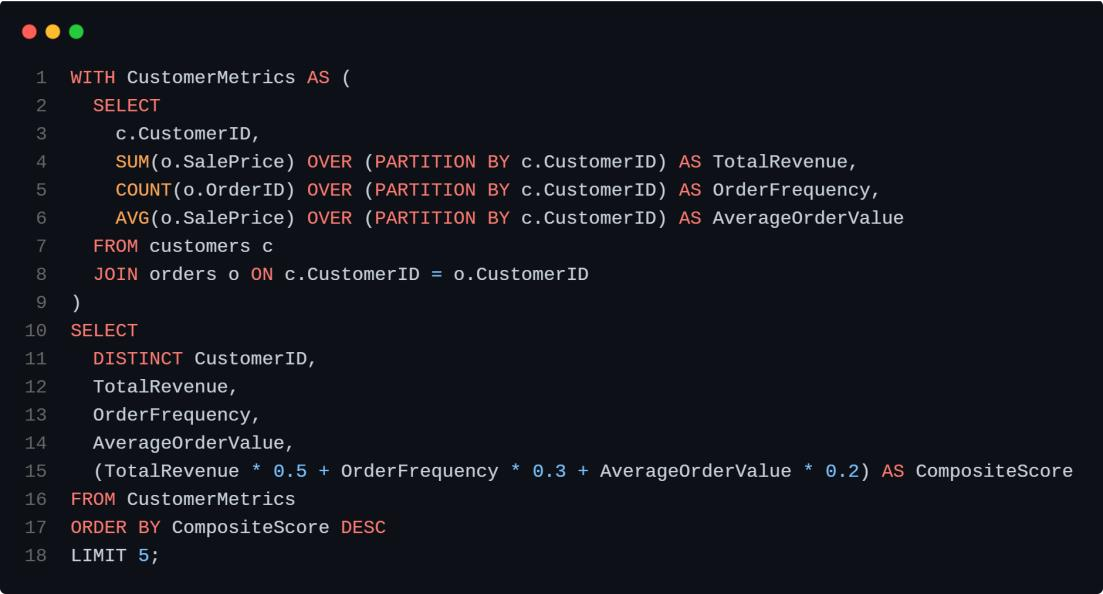
**Example of M Code for Adding a Column**

***= Table.AddColumn(Source, "NewColumnName", each [ExistingColumn] \* 2, type number)***

**Explanation:**

* + - Table.AddColumn: A function that adds a new column to a table.
    - Source: Refers to the table or step we’re working with.
    - "NewColumnName": The name of the new column.
    - each [ExistingColumn] \* 2: The calculation or logic for the new column (in this case, doubling the values of ExistingColumn).
    - type number: Specifies the data type of the new column (e.g., number, text, date).

# Identify the top 5 most valuable customers using a composite score that combines three key metrics: (SQL)



1. Total Revenue (50% weight): The total amount of money spent by the customer.
2. Order Frequency (30% weight): The number of orders placed by the customer, indicating their loyalty and engagement.
3. Average Order Value (20% weight): The average value of each order placed by the customer, reflecting the typical transaction size.

* The Composite Score will be calculated as:

***(0.5×Total Revenue)+(0.3×Order Frequency)+(0.2×Average Order Value)***

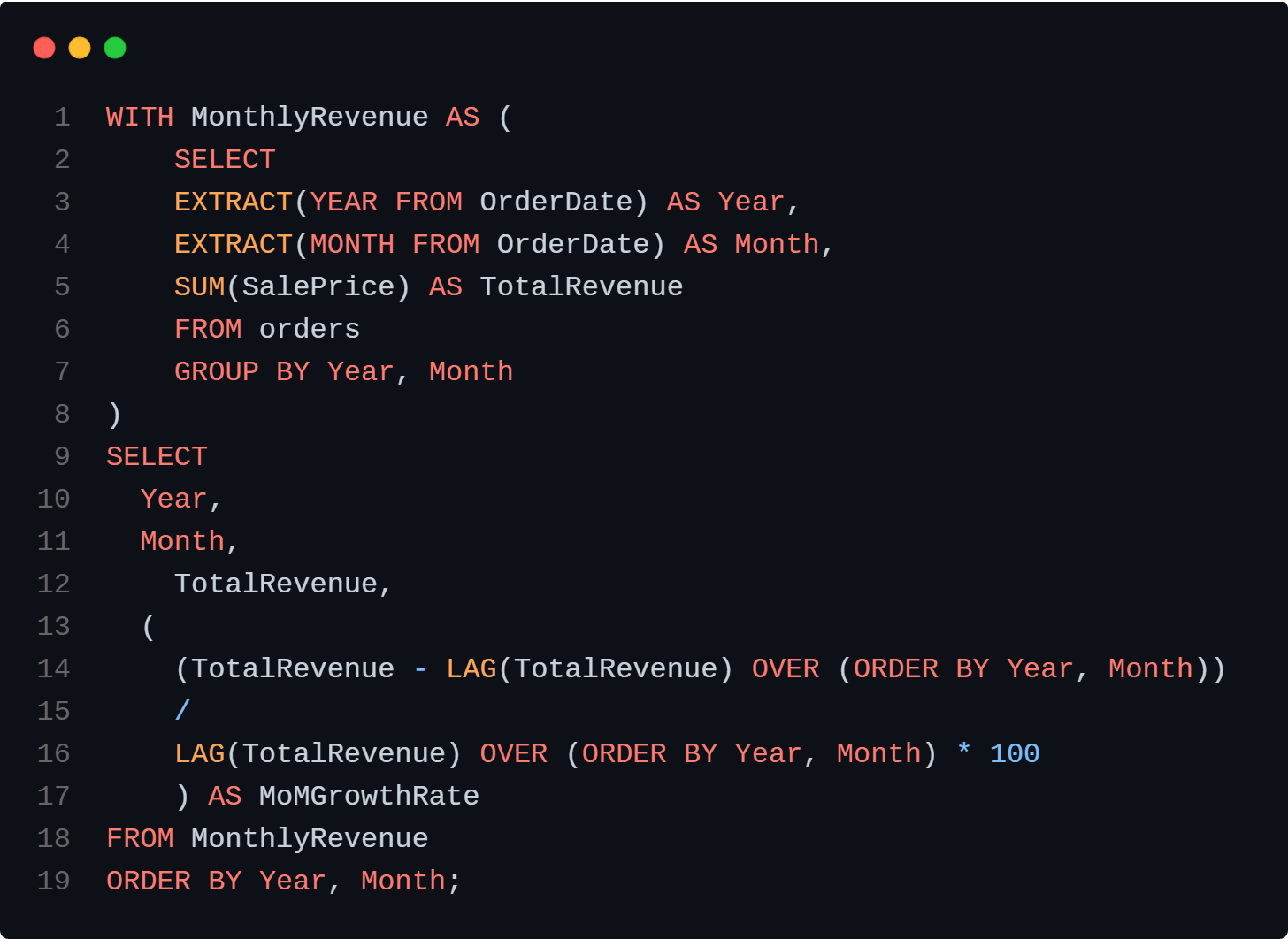
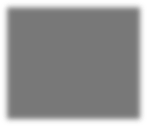
**Query:**

**Output:**

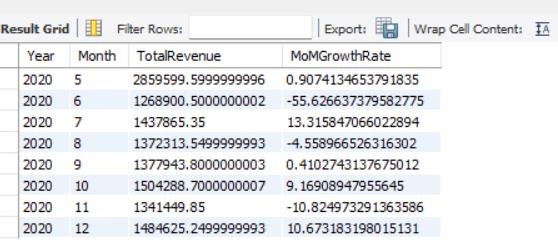
The top 5 valuable customers as follows:

# Calculate the month-over-month growth rate in total revenue across the entire dataset. (SQL)

* + **Query:**

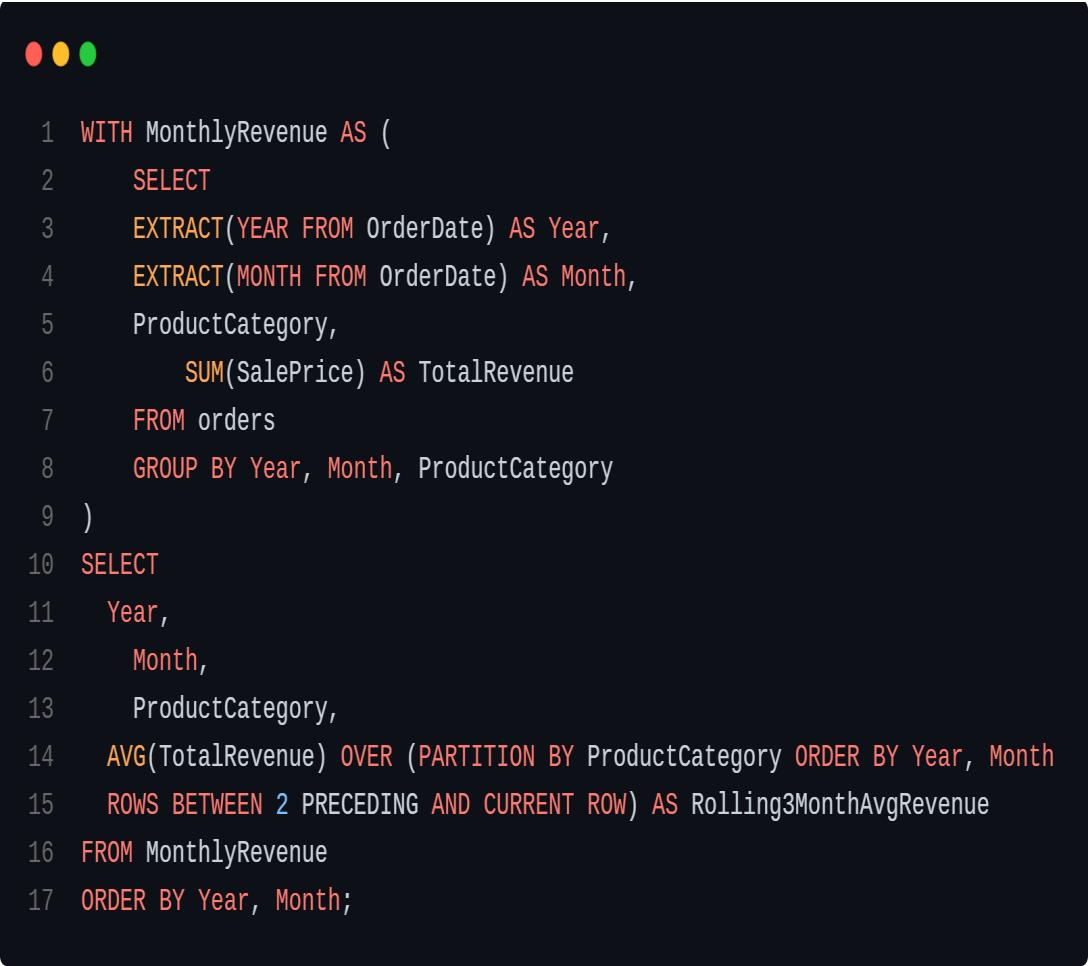
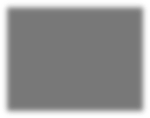


**Output:**

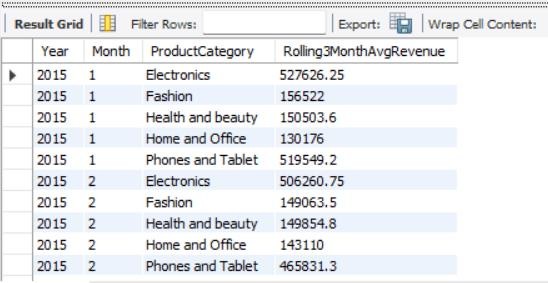


# Calculate the rolling 3-month average revenue for each product category. (SQL)

* + **Query:**



**Output:**



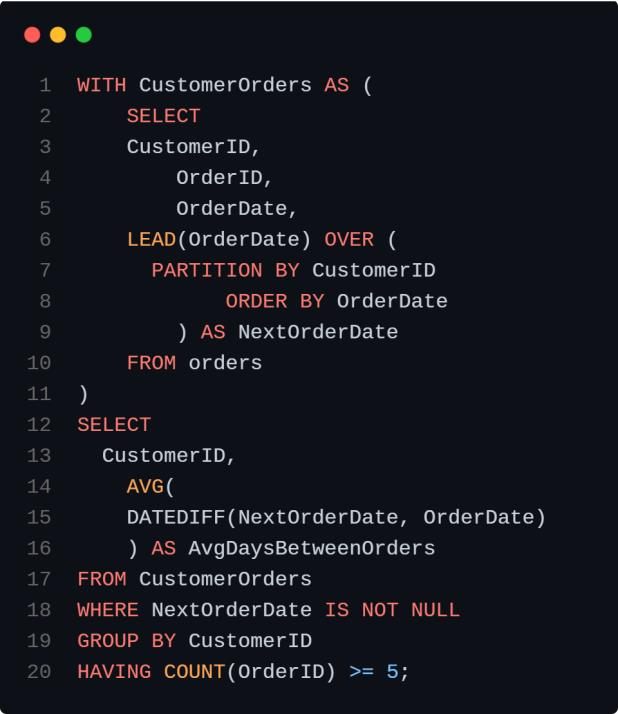
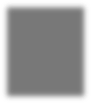
# Update the orders table to apply a 15% discount on the `Sale Price` for orders placed by customers who have made at least 10 orders. (SQL)

* + **Query:**

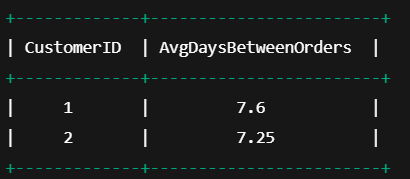


# Calculate the average number of days between consecutive orders for customers who have placed at least five orders. (SQL)

* + **Query:**



**Output:**

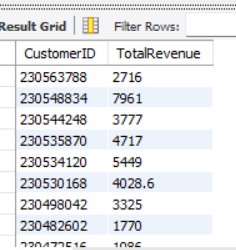


# Identify customers who have generated revenue that is more than 30% higher than the average revenue per customer. (SQL)

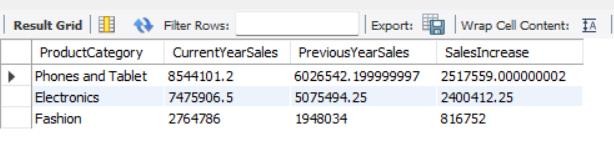
* + **Query:**



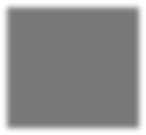
**Output:**



# Determine the top 3 product categories that have shown the highest increase in sales over the past year compared to the previous year. (SQL)



* + **Query:**



**Output:**