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).

2. 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.

3. 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.

4. 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:

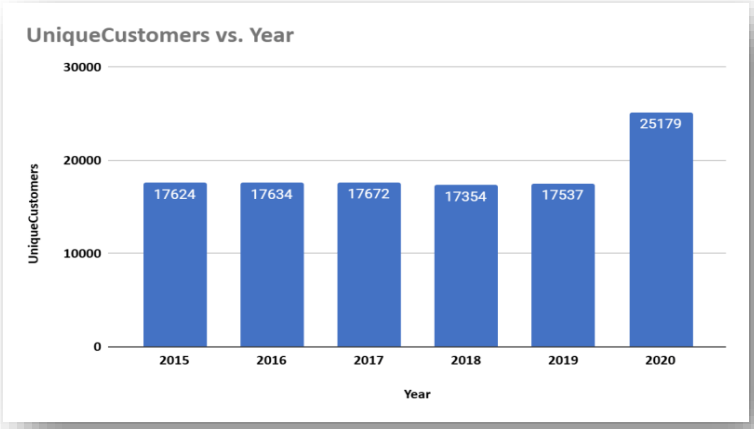
```
TotalRevenue = SUM(Orders[Sale Price])
```

5. 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])

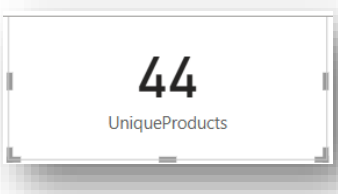
Visualization:



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.

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

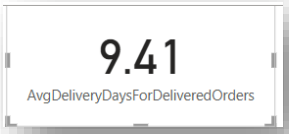
➢ Followed these steps:
- Data View > Order Table > New Measure
- Entered the DAX formula:
  UniqueProducts = DISTINCTCOUNT(Orders[Product])



7. 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(Orders[DeliveryDuration]),Orders[Status] = "Delivered")
    - This measure calculates the average delivery duration for orders where the status is "Delivered."



8. Which products, categories, and subcategories are the most popular?

- **Popularity by Order Quantity:**
  - Modeling > New Measure
    TotalOrderQuantity = SUM(Orders[Order Quantity])

a) 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.

Product	Count of Order Quantity
10.1'' Business Tablet with MT6582 Quad-Core Processor	2569
100%Cotton 4 Piece Short Sleeve T-Shirts - Multicolour	2569
8 Cubes Plastic Wardrobe - Blue/White	2569
Amazon Fire HD 8 Kids Tablet 32GB HDD - 2GB RAM - 8" Blue	2569
Avon Soft Musk Eau de Toilette Spray - 50ml	2569
B5 HiFi 5.0 Ture Wireless Headsets Auto Pair Touch - Black	2569
Blood Pressure Monitor Digital Wrist BP Pulse Monitor Meter Heart Rate Measure	2569
Boys Sneakers Casual Kids Sports Shoes-Gold	2569
Canon EOS 600D 18MP CMOS DSLR Camera - Black	2569
Clere Avocado Milk Body Lotion With Vitamins E+A - 400ml	2569
Clere Radiance Oil Control Toner - 100ml	2569
Hemani Ultra Slim Tea - 10 Bags	2569
L A Girl Pro Coverage HD Illuminating Liquid Foundation - Coffee	2569
Leather Vintage Bracelet Watch - Black	2569
Potluck Lunch Box - Brown	2569
Samsung Galaxy A02 - 64GB HDD - 3GB RAM Smartphone - Black	2569
Trust Leather Buckle Shoes - Black	2569
6030 3.1 Bluetooth Home Theatre With Remote Control - Black + Free Smartwatch	2568
Fashion Girls' Patent Leather Stitching Shoes - Black	2568
Fragrance World Smart Black Eau de Parfum Spray - 100ml	2568
Lindy 12 Cubes Wardrobe 8 Doors - Brown	2568
M4 Smart Bracelet Sports Pedometer Watch	2568
Maze Batik Designed 3D Wallpaper - 10M - White/Black	2568
Muscle Stimulators - Abdominal Muscle Trainer Set - Fitness	2568
Optimum Nutrition Creatine Sports - 5000mg per Daily Serve Powder	2568
Portable Blood Pressure Monitor - White	2568

b) 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.

Product Category	Count of Order Quantity
Health and beauty	35951
Fashion	33383
Phones and Tablet	17978
Home and Office	15408
Electronics	10271

c) 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.

SubCategory	Count of Order Quantity
Vitamins & Dietary Supplements	12838
Men's fashion	10272
Medical supplies and Equipment	10269
Beauty and personal care	7707
Boy's fashion	7704
Women's fashion	7704
Girl's fashion	7703
Kitchen and dinning	7703
Mobile phones	7703
Tablets	5138
Fragrances	5137
Home and Furniture	5137
Mobile accessories	5137
Digital Cameras	5136
Home Audio	5135
Tools and Home Improvement	2568

9. 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

```
Previous Year Sales =
CALCULATE(
    [TotalSales],
    FILTER(
        Orders,
        Orders[OrderDate].[Year] = MAX(Orders[OrderDate].[Year]) - 1
    )
)
```

- Create a measure for Sale Change

```
Sales Change = [TotalSales] - [Previous Year Sales]
```

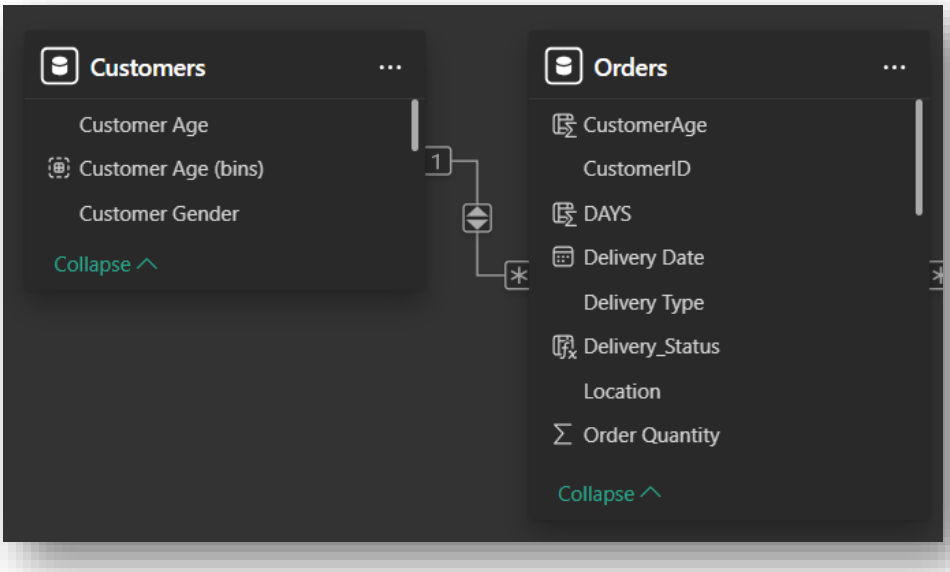
Matrix Visualization

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

Product	TotalSales	Previous Year Sales	Sales Change
10.1'' Business Tablet with MT6582 Quad-Core Processor	5293898	732426	4561472
100%Cotton 4 Piece Short Sleeve T-Shirts - Multicolour	1160562	186894	973668
6030 3.1 Bluetooth Home Theatre With Remote Control - Black + Free Smartwatch	5405064	806046	4599018
8 Cubes Plastic Wardrobe - Blue/White	2513287	393573	2119714
Aichun Beauty Eight Pack Essential Oil - 30ml	513528	79115	434413
Amazon Fire HD 8 Kids Tablet 32GB HDD - 2GB RAM - 8" Blue	11236628	1882052	9354576
Avon Soft Musk Eau de Toilette Spray - 50ml	1113510	161215	952295
B5 HiFi 5.0 Ture Wireless Headsets Auto Pair Touch - Black	774909	121152	653757
Blood Pressure Monitor Digital Wrist BP Pulse Monitor Meter Heart Rate Measure	933963	148178	785785
Boys Sneakers Casual Kids Sports Shoes-Gold	1745896	256549	1489347
Canon EOS 600D 18MP CMOS DSLR Camera - Black	13890491	2082643	11807848
Canon EOS 60D CMOS DSLR Camera Bundle - 18 - 55mm Lens - Black	12094365	1964631	10129734
Clere Avocado Milk Body Lotion With Vitamins E+A - 400ml	272721	42491	230230
Clere Radiance Oil Control Toner - 100ml	270617	40357	230260
Cq Amaigrissant Slimming Tea - 20 Tea Bags	423412	67094	356318
Fashion 4-Piece Leather HandBag Set - Black	735017	109995	625022
Fashion Boys Sneakers Children Outdoor Shoes-Black	1237872	196243	1041629
Fashion Girl's Dress Kids Children Newborn Baby Dinner Party Princess Dress Ball	1761348	272312	1489036
Total	**107239538**	**16769661**	**90469877**

10. 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.

11. 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

12. 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.

13. 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).

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

a. Total Revenue (50% weight): The total amount of money spent by the customer.
b. Order Frequency (30% weight): The number of orders placed by the customer, indicating their loyalty and engagement.
c. 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:

```sql
WITH CustomerMetrics AS (
  SELECT
    c.CustomerID,
    SUM(o.SalePrice) OVER (PARTITION BY c.CustomerID) AS TotalRevenue,
    COUNT(o.OrderID) OVER (PARTITION BY c.CustomerID) AS OrderFrequency,
    AVG(o.SalePrice) OVER (PARTITION BY c.CustomerID) AS AverageOrderValue
  FROM customers c
  JOIN orders o ON c.CustomerID = o.CustomerID
)
SELECT
  DISTINCT CustomerID,
  TotalRevenue,
  OrderFrequency,
  AverageOrderValue,
  (TotalRevenue * 0.5 + OrderFrequency * 0.3 + AverageOrderValue * 0.2) AS CompositeScore
FROM CustomerMetrics
ORDER BY CompositeScore DESC
LIMIT 5;
```

Output:
The top 5 valuable customers as follows:

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

CustomerID	TotalRevenue	OrderFrequency	AvgOrderValue	CompositeScore
230503475	8180.00	1	8180.000000	5726.3000000
230534492	8180.00	1	8180.000000	5726.3000000
230508728	8180.00	1	8180.000000	5726.3000000
230488406	8180.00	1	8180.000000	5726.3000000
230487688	8180.00	1	8180.000000	5726.3000000

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

➢ **Query:**

```sql
WITH MonthlyRevenue AS (
    SELECT
    EXTRACT(YEAR FROM OrderDate) AS Year,
    EXTRACT(MONTH FROM OrderDate) AS Month,
    SUM(SalePrice) AS TotalRevenue
    FROM orders
    GROUP BY Year, Month
)
SELECT
  Year,
  Month,
    TotalRevenue,
  (
    (TotalRevenue - LAG(TotalRevenue) OVER (ORDER BY Year, Month))
    /
    LAG(TotalRevenue) OVER (ORDER BY Year, Month) * 100
    ) AS MoMGrowthRate
FROM MonthlyRevenue
ORDER BY Year, Month;
```

Output:

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Year	Month	TotalRevenue	MoMGrowthRate
2020	5	2859599.5999999996	0.9074134653791835
2020	6	1268900.5000000002	-55.626637379582775
2020	7	1437865.35	13.315847066022894
2020	8	1372313.5499999993	-4.558966526316302
2020	9	1377943.8000000003	0.4102743137675012
2020	10	1504288.7000000007	9.16908947955645
2020	11	1341449.85	-10.824973291363586
2020	12	1484625.2499999993	10.673183198015131

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

➢ **Query:**

```sql
WITH MonthlyRevenue AS (
    SELECT
    EXTRACT(YEAR FROM OrderDate) AS Year,
    EXTRACT(MONTH FROM OrderDate) AS Month,
    ProductCategory,
        SUM(SalePrice) AS TotalRevenue
    FROM orders
    GROUP BY Year, Month, ProductCategory
)
SELECT
  Year,
    Month,
    ProductCategory,
  AVG(TotalRevenue) OVER (PARTITION BY ProductCategory ORDER BY Year, Month
  ROWS BETWEEN 2 PRECEDING AND CURRENT ROW) AS Rolling3MonthAvgRevenue
FROM MonthlyRevenue
ORDER BY Year, Month;
```

Output:

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Year	Month	ProductCategory	Rolling3MonthAvgRevenue
2015	1	Electronics	527626.25
2015	1	Fashion	156522
2015	1	Health and beauty	150503.6
2015	1	Home and Office	130176
2015	1	Phones and Tablet	519549.2
2015	2	Electronics	506260.75
2015	2	Fashion	149063.5
2015	2	Health and beauty	149854.8
2015	2	Home and Office	143110
2015	2	Phones and Tablet	465831.3

17. 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:**

```sql
UPDATE orders
SET SalePrice = SalePrice * 0.85
WHERE CustomerID IN (
    SELECT CustomerID
    FROM orders
    GROUP BY CustomerID
    HAVING COUNT(OrderID) >= 10
);
```

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

➢ **Query:**

```sql
WITH CustomerOrders AS (
    SELECT
    CustomerID,
        OrderID,
        OrderDate,
    LEAD(OrderDate) OVER (
      PARTITION BY CustomerID
            ORDER BY OrderDate
        ) AS NextOrderDate
    FROM orders
)
SELECT
  CustomerID,
    AVG(
    DATEDIFF(NextOrderDate, OrderDate)
    ) AS AvgDaysBetweenOrders
FROM CustomerOrders
WHERE NextOrderDate IS NOT NULL
GROUP BY CustomerID
HAVING COUNT(OrderID) >= 5;
```

Output:

CustomerID	AvgDaysBetweenOrders
1	7.6
2	7.25

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

➢ **Query:**

```sql
WITH TotalCustomerRevenue AS (
    SELECT
    CustomerID,
        SUM(SalePrice) AS TotalRevenue
    FROM orders
    GROUP BY CustomerID
),
AverageRevenue AS (
    SELECT AVG(SalePrice) AS AvgRevenue
    FROM orders
)
SELECT
  t.CustomerID,
  t.TotalRevenue
FROM TotalCustomerRevenue t, AverageRevenue a
WHERE t.TotalRevenue > a.AvgRevenue * 1.30;
```

Output:

Result Grid | Filter Rows:

CustomerID	TotalRevenue
230563788	2716
230548834	7961
230544248	3777
230535870	4717
230534120	5449
230530168	4028.6
230498042	3325
230482602	1770

20. 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:**

```sql
SELECT
  a.ProductCategory,
  a.TotalSales AS CurrentYearSales,
  b.TotalSales AS PreviousYearSales,
  a.TotalSales - b.TotalSales AS SalesIncrease
FROM (
  SELECT ProductCategory, SUM(SalePrice) AS TotalSales
  FROM orders
  WHERE YEAR(OrderDate) = 2020
  GROUP BY ProductCategory
) as a
JOIN (
  SELECT ProductCategory, SUM(SalePrice) AS TotalSales
  FROM orders
  WHERE YEAR(OrderDate) = 2019
  GROUP BY ProductCategory
) as b
ON a.ProductCategory = b.ProductCategory
ORDER BY SalesIncrease DESC
LIMIT 3;
```

Output:

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

ProductCategory	CurrentYearSales	PreviousYearSales	SalesIncrease
Phones and Tablet	8544101.2	6026542.199999997	2517559.000000002
Electronics	7475906.5	5075494.25	2400412.25
Fashion	2764786	1948034	816752