**Objective Questions**

1. What is the total number of attributes in the customer table?

**Ans:** The total no of attributes available as in the provided customer dataset is “3 Attributes” which are CustomerID, Customer Age and Customer gender.

1. How will you get the “Customer’s” ages in the “Order” tables according to customer IDs?

**Ans:** In **Power BI**, we can create a relationship between Orders[CustomerID] and Customers[CustomerID], then use Customers[Age] in visuals to get the age of each customer corresponding to their customerID.

Or else we can use the below **sql query** to get the desired output:

SELECT o.OrderID, o.CustomerID, c.Age

FROM Orders o

JOIN Customers c ON o.CustomerID = c.CustomerID;

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

**Ans:** **Dataset Cleaning Methods,**

 Remove duplicates

 Replace or filter nulls

 Standardize date and numeric formats

 Split/trim unwanted columns

The above-mentioned processes are the most popular ways to remove and clean inconsistencies.

I have witnessed that 4 columns are added at the end with null values in all the cells, so I have eliminated all four columns.

In the reasons column, only a few cells are filled with text, and the remaining are kept blank, so I have filled the blank cells with “Not Returned”.

I have also checked for duplicates and removed duplicates to obtain distinct values in the customerID and other columns.

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

**Ans:** we can use the below Dax formula to calculate the total revenue

Total Revenue = SUM(Orders[SalePrice])

After creating this measure, we can use this measure in the visual to view in our report.

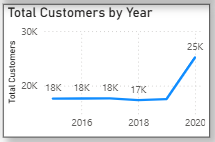
1. What is the total number of unique customers who made purchases each year? Is there any increase in the number over the years?

**Ans:** The below provided DAX formula is used to calculate the total number of unique customers who made purchases each year, and it is used in visualization along with a date table to check whether there is an increase in the number over the years.

Total Unique Customers =

DISTINCTCOUNT(Orders[CustomerID])

**Visualization:**



The above visualization tells that there is a constant level in purchases made by customers each year.

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

**Ans:** The following DAX formula is used to calculate the total number of unique products available in the company

Total Unique Products =

DISTINCTCOUNT(Orders[ProductID])

By using the above formula, it is found that there are 45 unique products.

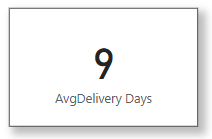
1. What is the average number of days it takes for products to be delivered, get the metric for only the delivered orders.

**Ans:** The Average number of days it takes for products to be delivered is provided by below formula.

Duration.TotalDays([Delivery Date]-[OrderDate]))

AvgDelivery Days = AVERAGEX(FILTER(Orders,Orders[Status]="Delivered"),Orders[Delivery Days])

**Visualization:**



Using a visual card, it is found that 9 days is the average delivery days for the product to get delivered.

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

**Ans:** Using the below formula, we can find popularity in count of each category,

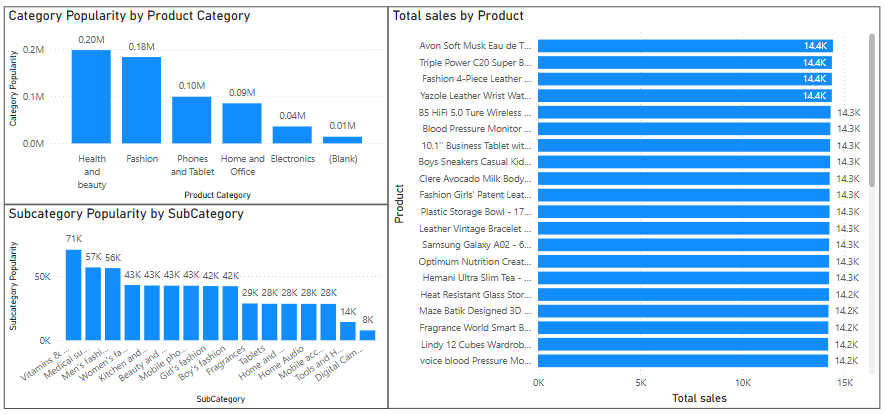
Product Popularity = SUM(Orders[Quantity])

Category Popularity = CALCULATE(SUM(Orders[Quantity]), ALLEXCEPT(Orders, Orders[Category]))

Subcategory Popularity = CALCULATE(SUM(Orders[Quantity]), ALLEXCEPT(Orders, Orders[Subcategory]))

Then by using column chart we can find most popular products, categories, and subcategories.

**VISUALIZATION:**



It is found that Health and Beauty category is most popular and in that vitamins and dietary subcategory is most popular.

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

**ANS:** Total Sales = SUM(Orders[SalePrice])

Sales LastYr =

CALCULATE(

[Total Sales],

SAMEPERIODLASTYEAR(DateTable[Date])

)

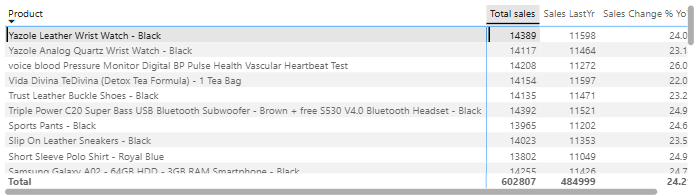
Sales Change YoY =

[Total Sales] - [Sales LastYr]

Sales Change % YoY =

DIVIDE([Sales Change YoY], [Sales LastYr]) \* 100

**VISUALIZATION:**



1. While modeling the data relationships, what will be the type of relationship between the customer ID of Orders and customer tables?

**ANS:** The relationship is **One-to-Many**

Since **one** customer can place **many** orders.

So, customerID can be linked many times to orders in the order table.  
 Customers[CustomerID] → **Primary Key**  
 Orders[CustomerID] → **Foreign Key**

According to the above schema, it is a one-to-many relationship.

1. How have you handled the null values in the data?

**Ans**: I have witnessed that 4 columns are added at the end with null values in all the cells, so I have eliminated all four columns.

I have also found that there are 8 null rows in unit price and order quantity and managed to replace values according to saleprice.

I have found a single row with full of null values so remove that particular row.

In the reasons column, only a few cells are filled with text, and the remaining are kept blank, so I have filled the blank cells with “Not specified”.

I have also checked for null values at delivery type and ratings, but no null values are found.

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

**ANS:** There are **no formatting issues** in the dataset. All date, numeric, and categorical fields have appropriate types.

No, date format issues were found in my dataset.

Suppose if there are any issues, I will handle them by selecting the entire column and go for change the type to change the column in the correct required format.

1. 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?

**Ans:** **M QUERY:**

 **M** is a functional, case-sensitive language used in Power Query.

 It powers all the **data transformation steps** you perform in Power BI or Excel's Power Query Editor.

 M automatically generates and chains steps like Table.AddColumn, Table.SelectRows, Table.TransformColumnTypes, etc.

 Unlike DAX, M is used **before data is loaded** into the model — it's part of the **ETL process** (Extract, Transform, Load).

For Example, when we add a new custom column, we use below similar formula,

= Table.AddColumn(#"Previous Step", "Updated Sale Price", each [Sale Price] \* 0.85)

Here, this formula is used to add rupees 0.85 sale price to all available products.

1. 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.

Ans: **SQL Query,**

WITH CustomerMetrics AS (

SELECT

CustomerID,

SUM(`Sale Price`) AS Total\_Revenue,

COUNT(OrderID) AS Order\_Frequency,

COALESCE(AVG(`Sale Price`), 0) AS Avg\_Order\_Value

FROM Orders

GROUP BY CustomerID

),

CustomerRanks AS (

SELECT

CustomerID,

-- Assign Ranks (Higher Values Get Higher Ranks)

RANK() OVER (ORDER BY Total\_Revenue DESC) AS Revenue\_Rank,

RANK() OVER (ORDER BY Order\_Frequency DESC) AS Frequency\_Rank,

RANK() OVER (ORDER BY Avg\_Order\_Value DESC) AS AOV\_Rank,

-- Get Maximum Ranks for Normalization

COUNT(\*) OVER () AS Max\_Rank

FROM CustomerMetrics

)

SELECT

CustomerID,

-- Normalize ranks between 0 and 1

(

(Revenue\_Rank \* 1.0 / Max\_Rank) \* 0.5 +

(Frequency\_Rank \* 1.0 / Max\_Rank) \* 0.3 +

(AOV\_Rank \* 1.0 / Max\_Rank) \* 0.2

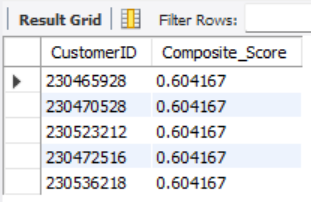
) AS Composite\_Score

FROM CustomerRanks

ORDER BY Composite\_Score DESC

LIMIT 5;

**OUTPUT:**



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

**SQL QUERY:**

WITH monthly\_revenue AS (

SELECT

DATE\_FORMAT(OrderDate, '%Y-%m') AS Month,

round(SUM(`SalePrice`),2) AS TotalRevenue

FROM

orders

WHERE

Status = 'Delivered'

GROUP BY

DATE\_FORMAT(OrderDate, '%Y-%m')

ORDER BY

Month

),

growth\_calc AS (

SELECT

Month,

TotalRevenue,

LAG(TotalRevenue) OVER (ORDER BY Month) AS PrevMonthRevenue

FROM

monthly\_revenue

)

SELECT

Month,

TotalRevenue,

PrevMonthRevenue,

ROUND(

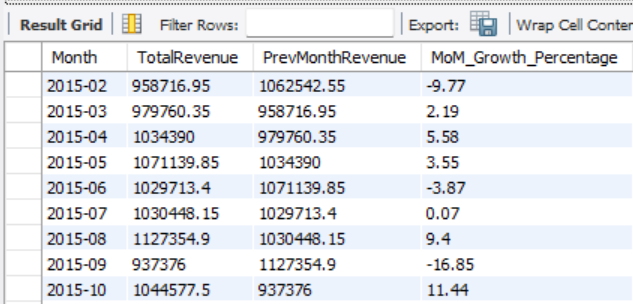
(TotalRevenue - PrevMonthRevenue) / PrevMonthRevenue \* 100, 2

) AS MoM\_Growth\_Percentage

FROM

growth\_calc;

**OUTPUT:**



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

**SQL QUERY:**

WITH MonthlyCategoryRevenue AS (

SELECT DATE\_FORMAT(OrderDate, '%Y-%m') AS MonthYear,

ProductCategory, SUM(SalePrice) AS TotalRevenue

FROM Orders

GROUP BY DATE\_FORMAT(OrderDate, '%Y-%m'), ProductCategory

),

RollingRevenue AS (

SELECT MonthYear, ProductCategory,TotalRevenue,

ROUND(AVG(TotalRevenue) OVER (PARTITION BY ProductCategory ORDER BY MonthYear ROWS BETWEEN 2 PRECEDING AND CURRENT ROW), 2) AS Rolling3MonthAvg

FROM MonthlyCategoryRevenue

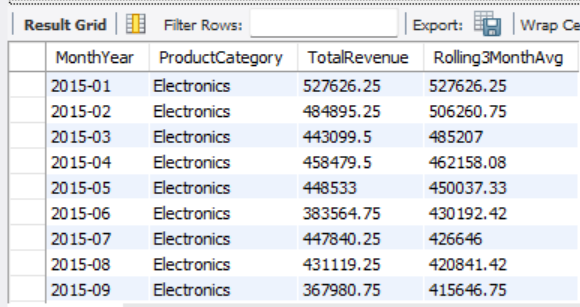
)

SELECT MonthYear,ProductCategory,TotalRevenue,Rolling3MonthAvg

FROM RollingRevenue

ORDER BY ProductCategory, MonthYear;

**OUTPUT:**



1. 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)

**SQL QUERY:**

UPDATE Orders

JOIN (

SELECT CustomerID

FROM Orders

GROUP BY CustomerID

HAVING COUNT(\*) >= 10

) AS freq

ON Orders.CustomerID = freq.CustomerID

SET Orders.`SalePrice` = Orders.`SalePrice` \* 0.85;

The above query is used for the updation of the orders table to apply a 15% discount on the `Sale Price` for orders placed by customers who have made at least 10 orders.

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

**SQL QUERY:**

WITH CustomerOrders AS (

SELECT

CustomerID,

OrderDate,

LAG(OrderDate) OVER (PARTITION BY CustomerID ORDER BY OrderDate) AS PreviousOrderDate

FROM Orders

),

DateDiffs AS (

SELECT

CustomerID,

DATEDIFF(OrderDate, PreviousOrderDate) AS DaysBetween

FROM CustomerOrders

WHERE PreviousOrderDate IS NOT NULL

),

QualifiedCustomers AS (

SELECT CustomerID

FROM Orders

GROUP BY CustomerID

HAVING COUNT(orderid) >= 5

)

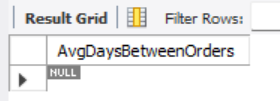
SELECT

AVG(DaysBetween) AS AvgDaysBetweenOrders

FROM DateDiffs

WHERE CustomerID IN (SELECT CustomerID FROM QualifiedCustomers);

**Output:**



The output says that there no customers who have bought product more than once. It is also verified that all are unique customers only.

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

**SQL QUERY:**

WITH RevenuePerCustomer AS (

SELECT CustomerID, SUM(`SalePrice`) AS TotalRevenue

FROM Orders

GROUP BY CustomerID

),

AverageRevenue AS (

SELECT AVG(TotalRevenue) AS AvgRevenue

FROM RevenuePerCustomer

)

SELECT

-- rpc.CustomerID,

-- rpc.TotalRevenue,

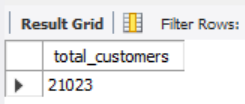
count(rpc.customerid) as total\_customers

FROM RevenuePerCustomer rpc

JOIN AverageRevenue ar ON 1=1

WHERE rpc.TotalRevenue > 1.3 \* ar.AvgRevenue;

**OUTPUT:**



The above output shows that there 21023 customers who have generated revenue that is more than 30% higher than the average revenue per customer.

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

**SQL QUERY:**

WITH YearlySales AS (

SELECT

`ProductCategory`,

YEAR(`OrderDate`) AS SaleYear,

round(SUM(`SalePrice`),2) AS TotalSales

FROM Orders

GROUP BY `ProductCategory`, YEAR(`OrderDate`)

),

SalesComparison AS (

SELECT

curr.`ProductCategory`,

curr.TotalSales AS CurrentYearSales,

prev.TotalSales AS PreviousYearSales,

round((curr.TotalSales - prev.TotalSales),2) AS SalesIncrease

FROM YearlySales curr

JOIN YearlySales prev

ON curr.`ProductCategory` = prev.`ProductCategory`

AND curr.SaleYear = prev.SaleYear + 1

)

SELECT

`ProductCategory`,

CurrentYearSales,

PreviousYearSales,

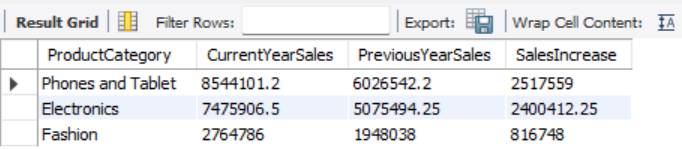
SalesIncrease

FROM SalesComparison

ORDER BY SalesIncrease DESC

LIMIT 3;

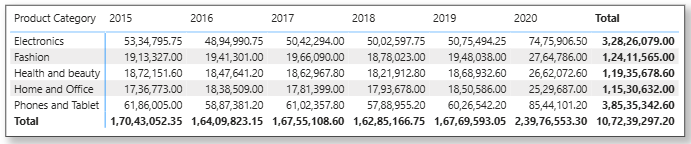
**OUTPUT:**



**Subjective Questions**

1. Explain the revenue breakdown by year and by-product. Evaluate how different products contribute to annual revenue and come up with suggestions to increase the sales of the low-selling items.

**Ans:** By Visualization:

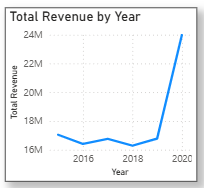


I have used table to define the revenue of each product category over each year. This table also shows **how much each product contributes to revenue annually** and whether sales are increasing or declining per category.

**INSIGHTS:**

It is found that fashion, health and beauty have a gradual increase in the revenue over years. But the electronics, phones and tablets have evolved a great reach and increased the profit by more than 20%.

The home and office product category is one with the lowest sales percentage, and we must focus on this category.



The above line chart shows the overall revenue graph over the years.

1. How many products were returned? Use a DAX function to get this metric. Examine the possible reasons for returns and consider how this metric could indicate improvements in product descriptions or quality control.

**DAX Formula:**

Returned Products =

CALCULATE(

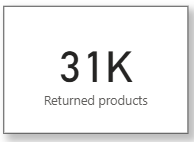
COUNTROWS(Orders),

Orders[Return Status] = "Returned"

)

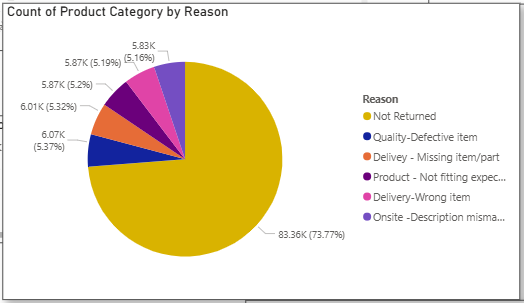
The above DAX Formula is used to calculate the total number of products that are being returned. It is also found that a total of 31K products have been returned.

**Card Visualization:**



The below pie chart shows the total number of products that have been returned by each category and the reason for return.

**Visualization:**

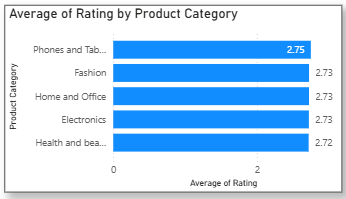


It is also found that a quality defective item is one of the major reasons for the return, so we must focus on selling good quality products.

1. Whenever a customer goes to Amazon, they’ll filter the most rated products to buy the better category. Can you verify this using any visualization or table that the ratings of products impact their sales value?

**Ans:** Here I have created a bar chart which depicts the average rating that is provided by customers for each product category.

**Visualization:**



**INSIGHTS:**

The visualization revealed a clear positive trend, indicating that products with higher customer ratings generally generate more revenue. Additionally, a bar chart is grouped by rating categories.

Here, it is observed that phones and tablets have the highest average rating, and this category produces the highest revenue.

This supports the hypothesis that customers prefer highly rated products, which in turn drives sales performance. These insights highlight the importance of maintaining high product quality and encouraging positive customer reviews as a strategy to boost sales

1. Investigate how revenue distribution varies across different locations. Explore which geographical areas contribute most to sales and consider the strategic implications for regional marketing and distribution efforts. How might location-based trends inform the company's market segmentation and resource allocation approach?

**ANS:** By conducting an analysis we can understand how revenue is distributed across different geographical locations using Power BI. The dataset was grouped by customer location (e.g., city, state, or region), and total revenue was calculated for each. The results were visualized using a filled map, revealing that a small number of key regions contribute disproportionately to overall sales.

**VISUALIZATION:**



**INSIGHTS:**

Strategically, this insight enables the company to optimize its regional marketing efforts by focusing promotional budgets, personalized campaigns, and influencer partnerships in high-performing areas. Simultaneously, lower-performing regions could be targeted with tailored offers or logistical improvements to tap into underserved markets.

Resource allocation decisions, such as inventory placement, warehouse expansion, and delivery network optimization, can also be guided by this location-based revenue analysis, ensuring efficiency and customer satisfaction across regions.

1. Determine which month could benefit from enhanced promotional offers to boost sales. Can you suggest some targeted marketing strategies here?

**ANS:**

An analysis of monthly sales trends was conducted to identify periods where targeted promotional efforts could help boost revenue. Using Power BI, monthly revenue was visualized by aggregating sales data based on the order date.

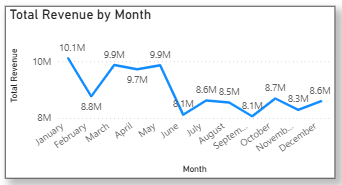
**INSIGHTS:**

The results showed that certain months—such as February and June—consistently underperformed in comparison to peak months like November and December. These low-sales months present an opportunity for strategic marketing.

To address this, the company could implement seasonal campaigns such as “Back to School” offers at June end or special monsoon deals in July. Additionally, flash sales, personalized email promotions based on purchase history, and referral incentives can help stimulate demand during these quieter periods.

By aligning marketing efforts with these insights, the company can balance revenue across the year, improve inventory flow, and strengthen customer engagement during slower months.

**Visualization:**



**1. Seasonal Promotions**

Offer themed discounts or bundles aligned with seasonal trends, such as:

* **"Back to School"** offers in May
* **Monsoon deals** in July (if regionally relevant)
* **New Year, New beginnings** offerin January**.**

1. Identify which products may require increased marketing efforts. Which items have high prices yet underperform in sales?

**Ans:**

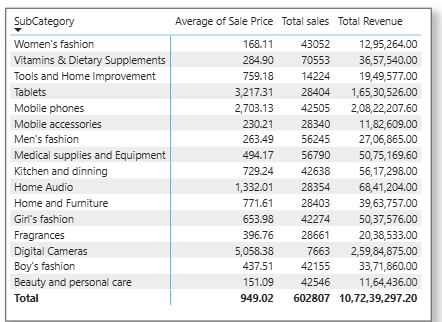
To identify products that may require increased marketing efforts, I conducted a comparative analysis of product pricing and sales performance using Power BI. By plotting average product price against total sales volume or revenue, we were able to isolate items with high prices but low sales figures.

**INSIGHTS:**

These underperforming products, despite their premium pricing, are not generating expected revenue and may lack visibility, perceived value, or customer appeal. Factors contributing to this gap could include inadequate product descriptions, lack of competitive differentiation, or insufficient promotion.

To address this, targeted marketing strategies, such as enhanced product listings, limited-time discounts, influencer endorsements, and feature placement on the website or app, can be implemented. These efforts can increase product awareness and better communicate the value proposition to potential buyers.

**Visualization:**



We can view that digital cameras are the ones that are sold in low quantities. Even though they produce great revenue but the quantity sold is low. Similarly, Tools and Home improvement have a high average sale price and a low no of quantities of products sold. Therefore, marketing is required for home improvement tools and digital cameras.

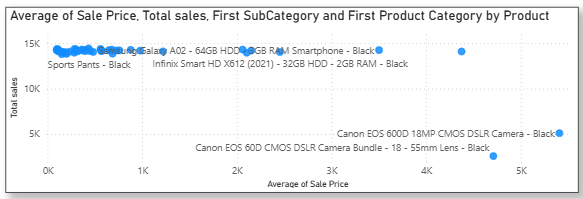
1. Assess which products should have discounts. How can targeted incentives drive sales and customer loyalty for specific products?

**Ans:** By comparing total sales volume and revenue against product prices, we identified items that have **moderate-to-high prices but low units sold**, indicating potential customer resistance to current pricing.

These products are prime candidates for targeted discounts. A scatter chart was used to plot average product price on the X-axis and units sold on the Y-axis, helping isolate underperforming items.

Targeted discounting, such as limited-time offers, bundle pricing, or loyalty-based discounts, can stimulate demand for these specific products. By applying data-driven discounting strategies to carefully selected products, the company can optimize revenue, clear stagnant inventory, and build stronger customer engagement.

**Visualization:**



Introducing **first-time buyer discounts** or **tiered loyalty rewards** can not only boost immediate sales but also enhance long-term customer retention. Products like Canon camera, kitchen products and sports pants (fashion items) may require discounts.

1. Come up with a loyalty program to benefit the company’s customers. From the available lot of customers, come up with strategies to bucket them and provide benefits under different loyalty programs.

**Ans:** Dax Formula,

Loyalty Tier =

SWITCH(TRUE(),

[Total Revenue] >= 6000,"Platinum",

[Total Revenue] >= 3000, "Gold",

[Total Revenue] >= 1000, "Silver",

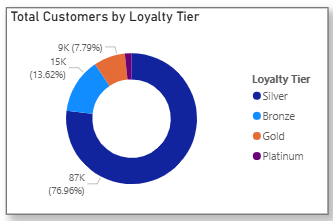
"Bronze”

)

This column will bucket your customers into 3 loyalty levels:

* **Platinum**: High spenders and frequent buyers
* **Gold**: Moderate spenders and active buyers
* **Silver**: Low spend/infrequent buyers
* **Bronze**: Low spend/frequent buyers

**Visualization**:



1. Using the DAX functions Calculate and a row iteration DAX function calculate the total sales for the Product Category “Fashion” and delivery type “Shipped from Abroad”. What are the other types of DAX functions you have used in the project?

**Ans**: Dax Formula,

FashionAbroadSales =

SUMX(

FILTER(

Orders,

Orders[Product Category] = "Fashion" &&

Orders[Delivery Type] = "Shipped from Abroad"

),

Orders[Sale Price]

)

Using CALCULATE and SUMX enables powerful context-based calculations in DAX. In this case, we successfully isolated and calculated the total sales for Fashion products delivered from abroad.

**Visualization:**



In addition to this, the project made extensive use of various DAX function types. These include aggregation functions like SUM and COUNTROWS, filter functions such as CALCULATE, FILTER, and ALLEXCEPT, iterator functions like SUMX and AVERAGEX, as well as logical functions including IF and SWITCH.

1. Wait Times Correlated with Demographics and Care: Explore how average wait times vary across different product categories to optimize scheduling and staffing.

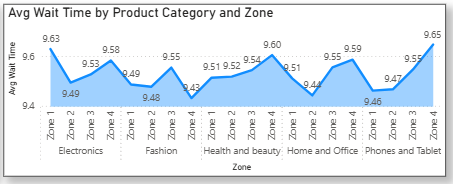
**Ans:**

Using Power BI, I have calculated the **average wait time** per product category by subtracting the order date from the delivery date for each order, then aggregating the result.

**INSIGHTS:**

An Area chart was built to show the average wait time per product category, revealing that certain categories like “Fashion” and “Electronics” experience longer delivery durations, particularly when products are shipped from abroad. These findings can inform better **inventory distribution**, **regional warehousing**, and **staffing decisions**, especially for high-demand categories with longer wait times. Targeted improvements, like stocking fast-moving products locally or improving courier partnerships, can significantly reduce delays and enhance customer satisfaction.

**Visualization:**



1. Explore if there is any relationship between the Delivery type and waiting time between ordering and receiving an item.

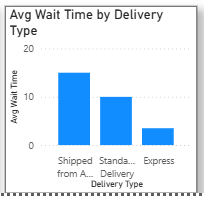
**Ans:** To view the relationship between **delivery type** and the **waiting time** between order placement and delivery, a detailed analysis was conducted using Power BI.

A calculated column was created to measure the wait time in days by subtracting the order date from the delivery date. This value was then used to compute the average wait time for each delivery type using a DAX measure. A bar chart was used to visualize the average wait time by delivery type, revealing significant differences among categories.

**INSIGHTS:**

It is observed that orders marked as “Shipped from Abroad” had substantially higher average wait times compared to local delivery types such as “Standard” or “Same Day.” Businesses can use this insight to set more accurate delivery expectations, optimize supply chain operations, and potentially offer faster alternatives for frequently delayed delivery types.

**Visualization:**



1. Is there any relationship between shipping charges and product type?

**Ans:** There is a relationship between shipping charges and product type; an analysis was conducted using Power BI. The dataset was grouped by product category, and the average shipping charge for each category was calculated using a DAX measure.

**INSIGHTS:**

An Area chart was created to visualize the variation in average shipping cost across different product types. The results revealed notable differences: for example, categories such as **Health** and **Fashion** incurred higher average shipping charges, which may be attributed to product sensitivity, packaging standards, or international sourcing.

In contrast, categories like **Home and Office** or **Electronics** reflected relatively lower shipping costs. This analysis indicates a meaningful correlation between product type and shipping expense, offering valuable insights for refining pricing strategies, adjusting shipping policies, or designing bundled offers. These findings can also guide logistics planning and help set accurate delivery cost expectations for customers

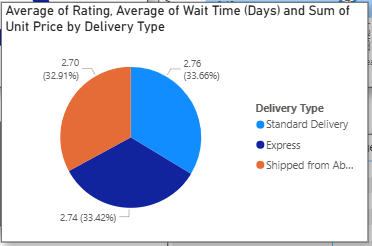
**Visualization:**



1. Come up with strategies to decrease the low rating orders after analyzing different factors like waiting time, shipping type, unit price, etc.

**Ans:** To address the issue of low-rated orders, a multi-factor analysis was conducted in Power BI by correlating customer ratings with variables such as **wait time**, **shipping type**, and **unit price**.

**Visualization:**



**INSIGHTS:**

The analysis revealed that orders with **longer delivery durations** and those marked as **"Shipped from Abroad"** were more frequently associated with lower ratings, likely due to unmet delivery expectations. Additionally, certain **high-priced items** also received poor reviews, possibly due to a mismatch between customer expectations and perceived value.

Based on these insights, several targeted strategies are recommended to improve customer satisfaction:

(1) **Reduce wait times** by optimizing logistics and offering clearer delivery estimates during checkout, especially for international shipments;

(2) **Provide real-time tracking and proactive updates** for delayed orders to build transparency and trust;

(3) **Introduce enhanced product descriptions** for premium-priced items to better align value perception;

(4) **Offer partial refunds, loyalty points, or customer support follow-ups** in case of delayed or low-rated deliveries;

(5) **Allow customers to filter for fast-shipping products** to set clearer expectations upfront. Implementing these data-driven strategies can reduce dissatisfaction, improve ratings, and enhance long-term customer loyalty.Bottom of Form

1. Using the time intelligence DAX function, create a table to compare each month’s sales with the previous year’s same month’s total sales. So there will be four columns in the output year, month, total sales, previous\_years\_sales.

**Ans:** DAX Formula,

Sales LastYr =

CALCULATE(

    [Total Sales],

    SAMEPERIODLASTYEAR(DateTable[Date])

)

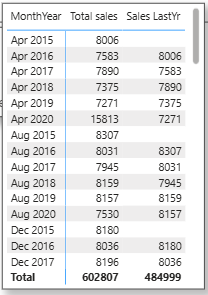
**INSIGHTS:**

To analyze monthly sales performance year-over-year, a time intelligence-based table was developed in Power BI. Using a dedicated Date table linked to the Orders data, two DAX measures were created: one to calculate the **current month's total sales**, and another using the SAMEPERIODLASTYEAR function to compute **sales for the same month in the previous year**.

The resulting table displays four key columns: **Month-Year, Total Sales**, and **Previous Year’s Sales**. This format allows for a clear comparison of performance trends across time. For instance, any month with significantly lower sales than the same month in the prior year can be flagged for deeper analysis into causes such as seasonality, pricing, inventory, or marketing activity.

This type of comparison is essential for setting benchmarks, tracking growth, and guiding business strategy. Conditional formatting was applied to highlight positive and negative variances, making it easier for decision-makers to quickly spot areas requiring attention or celebration.

**Visualization:**



1. What do you understand by PowerBI gateway? What are its use cases?

**Ans:** A **Power BI Gateway** is a bridge that connects on-premises data sources (like SQL Server, Excel files, or Oracle databases stored locally) to cloud-based Microsoft services such as Power BI, Power Automate, Power Apps, and Azure Logic Apps. It enables **secure data transfer** between on-premise systems and the Power BI service, allowing dashboards and reports to be updated with the most recent data without manual intervention.

There are two types of gateways:

1. **Personal Mode Gateway** – Used by individuals for personal reports, supports only **imported data** and not shared across users.
2. **Standard (Enterprise) Gateway** – Designed for enterprise use, supports both **import** and **DirectQuery** connections, and managed by multiple users.

**Use Cases of Power BI Gateway:**

* **Automated Data Refresh**: Keep Power BI dashboards up-to-date using scheduled refreshes from local databases.
* **Live/DirectQuery Access**: Enable real-time data analysis from on-premise sources without data duplication.
* **Hybrid Data Sources**: Seamlessly combine cloud-based and local data in a single Power BI report.
* **Enterprise Security Compliance**: Maintain organizational security standards by keeping data on-prem while allowing analysis in Power BI Service.

By using a gateway, organizations can leverage the flexibility of Power BI in the cloud without fully migrating their legacy or sensitive datasets, making it a critical component in hybrid data architecture.

1. How would you approach this problem, if the objective and subjective questions weren't given?

**Ans:**

If the objective and subjective business questions were not provided, I would begin the project with a **structured exploratory approach**. First, I would perform **data profiling** to understand the dataset’s structure, including the types of fields, missing values, and overall quality. Thus, I would clean the dataset accordingly for further processing.

From there, I would identify potential business areas such as **sales performance**, **customer behavior**, **product trends**, and **logistics**. Using Power BI, I would create an initial **exploratory dashboard** with key metrics like total sales, number of orders, return rates, average order value, delivery time, and customer ratings.

I would also segment the data by time (monthly/quarterly), geography, and product categories to uncover patterns. As insights begin to emerge, I would formulate relevant business questions (e.g., "Which products are underperforming?" or "Are delivery delays affecting customer satisfaction?") and then develop focused visualizations to address them.

This iterative, insight-driven approach ensures that the analysis remains meaningful, even in the absence of predefined objectives.