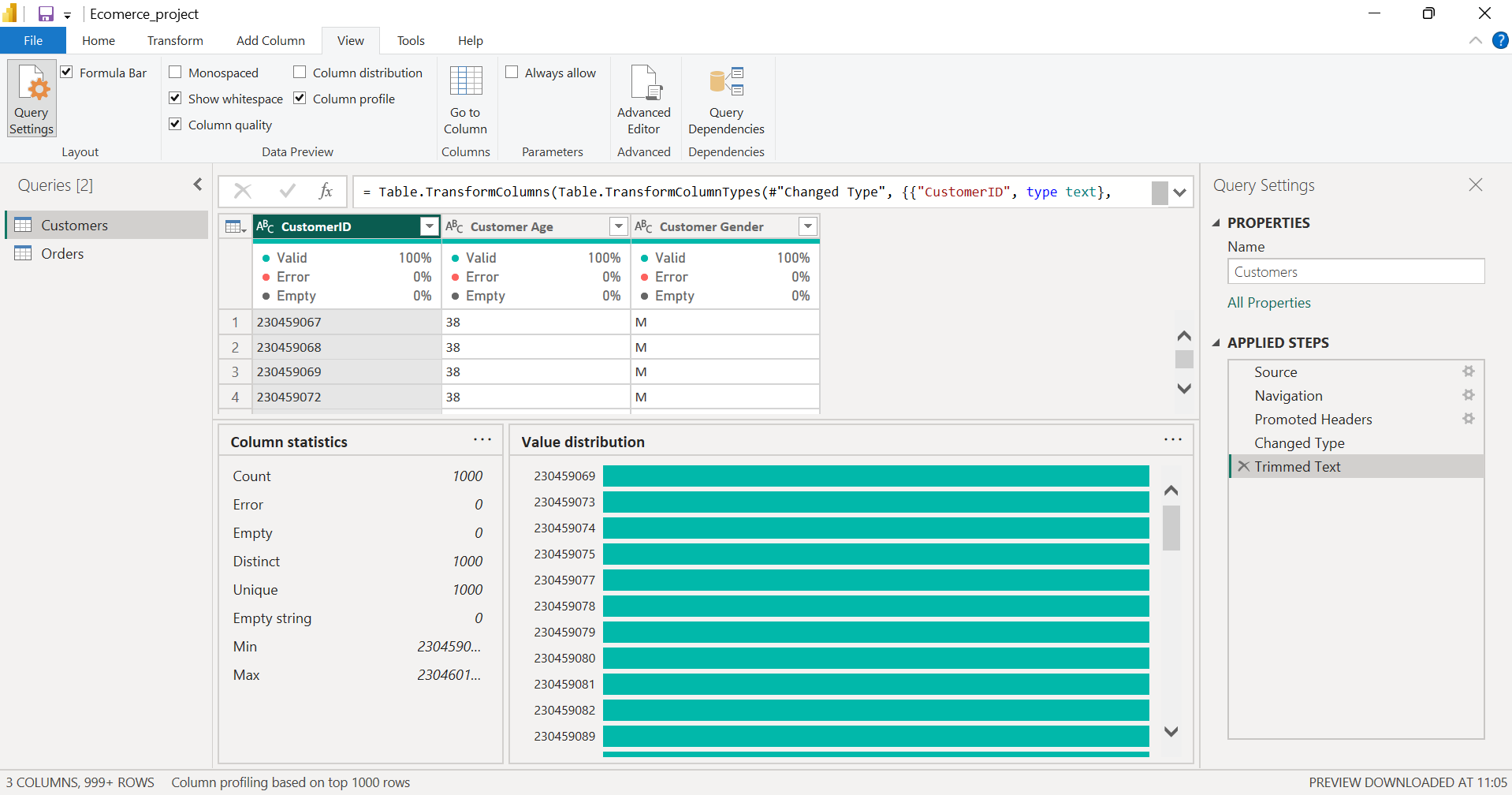
**Name – Hritik kumar**

**Objective Questions:**

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

Ans-

The total number of attributes in the customer table is 3.



Approach

* Opened the Customer table in Power Query.
* Used Column Profile to check number of columns.
* Counted columns = X attributes.

Insights

* Attributes include customer details like age and gender.
* Useful for customer segmentation and behaviour analysis.

Recommendation

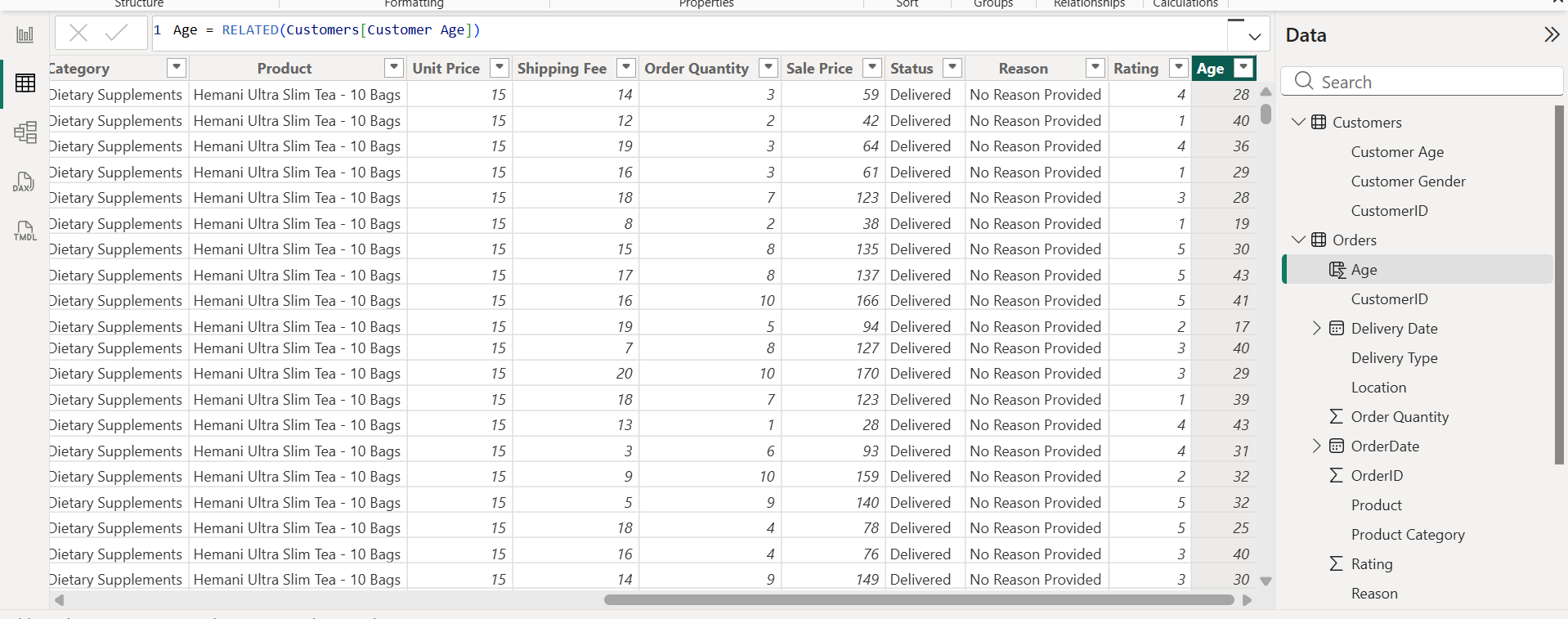
* Ensure correct data types, no nulls in key fields.
* Add useful fields like loyalty score if needed.

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

Ans-

DAX-

Age = RELATED (Customers [Customer Age])



Approach

* Confirmed a relationship exists between Orders [Customer ID] and Customers [Customer ID].
* Used the RELATED () DAX function to bring the Age field from the Customer table into the Orders table.
* This makes the age available at the row level for each order.

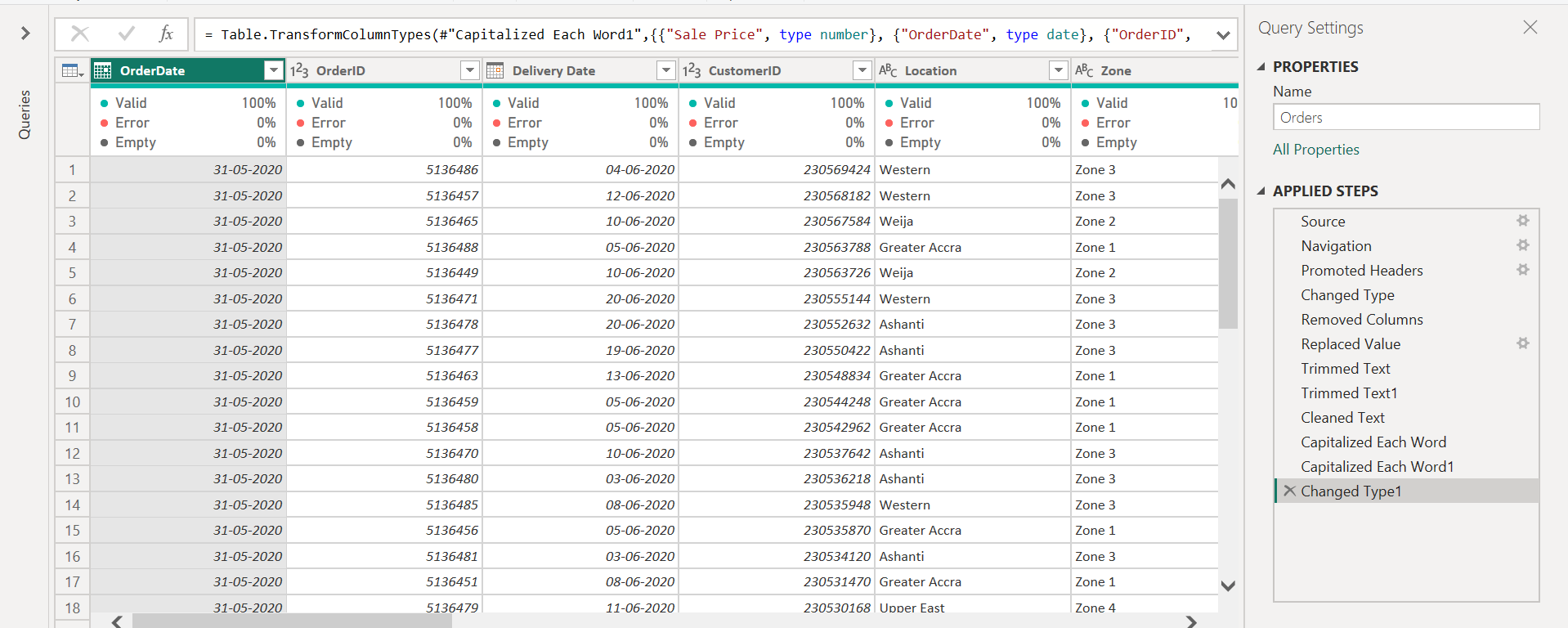
Insights

* Each order now contains the customer’s age.
* Enables age-based analysis such as:
  + Which age groups purchase the most
  + Product category preference by age
  + Seasonal or trend patterns by age group

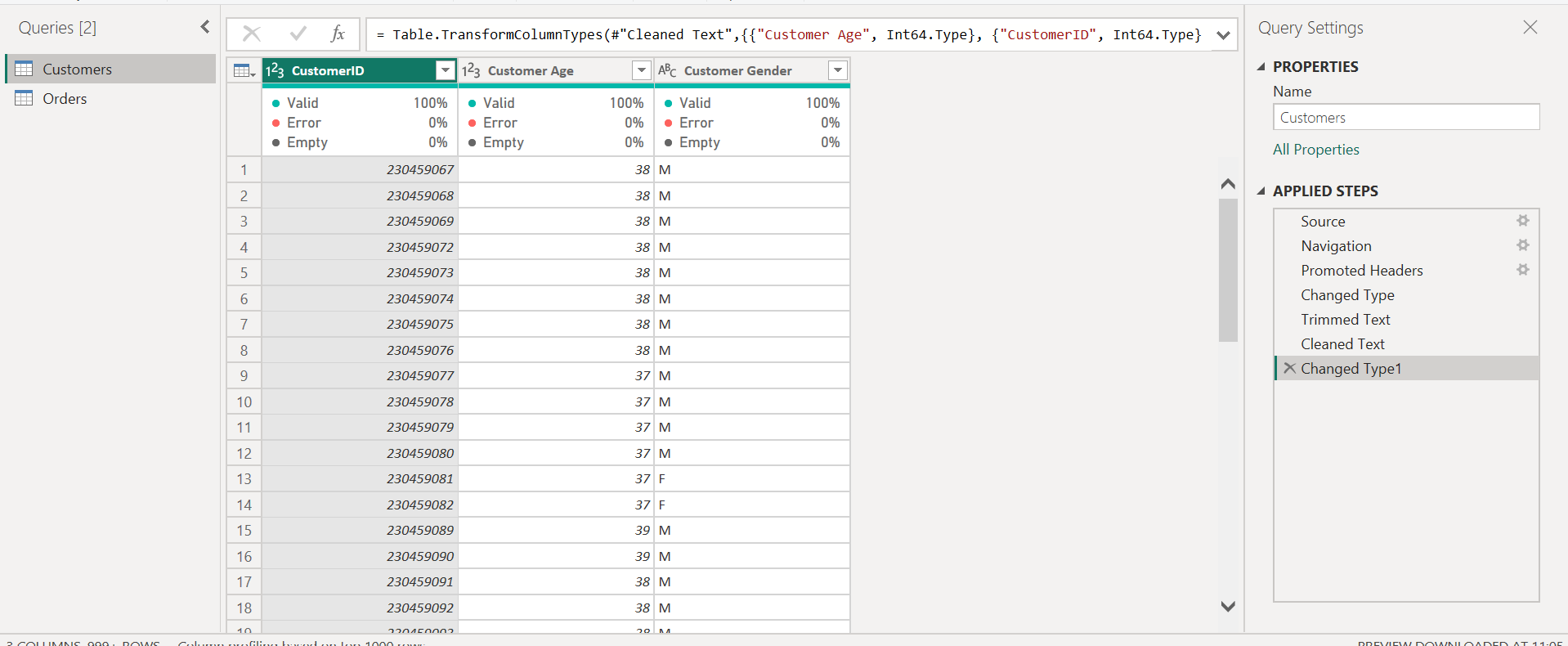
1. In analysing the dataset with Power BI, ensure data cleaning to address inconsistencies and missing values before further analysis.

Ans-

Orders table

****

**Customer table**

****

Approach

* Removed all **null/empty columns** that were not needed for analysis.
* In the **Reason** column, the empty cells were replaced with **“No Reason Provided”**.
* **Trimmed** all text columns to remove leading/trailing spaces.
* Used **Clean** to remove hidden characters from text fields.
* **Capitalized** all text columns to maintain a consistent format for categories and values.
* Corrected **data types**:
* Dates → Date type
* Quantity → Whole number
* Price → Decimal
* Ratings → Whole number

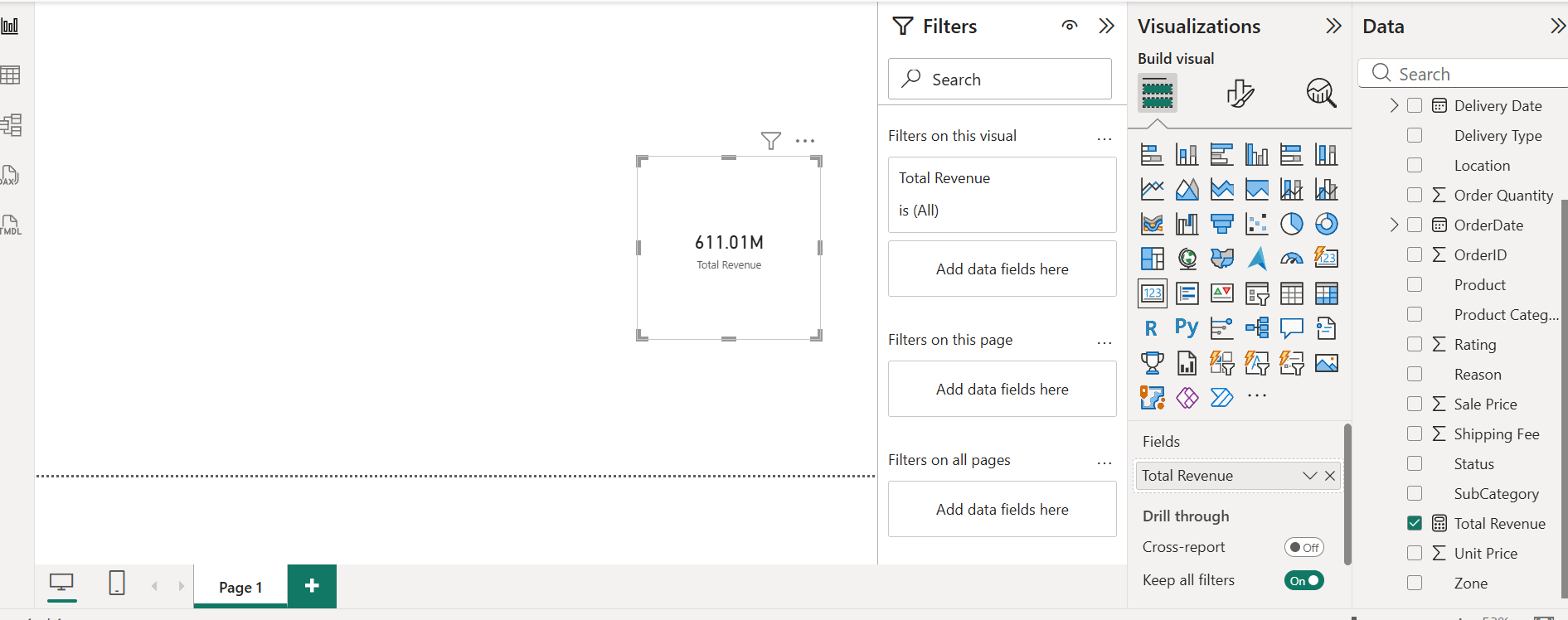
1. How can we calculate the total revenue generated by all the sales?

Ans-

DAX-

Total Revenue =

SUMX (Orders, Orders [Order Quantity] \* Orders [Sale Price])



**Approach**

* Identified that revenue needs to be calculated from the **Orders** table.
* Created a **DAX measure** instead of a calculated column, because:
  + Measures aggregate dynamically
  + They respond to filters and slicers
* Used SUMX () to iterate through each row and multiply:
  + Quantity × Sale Price
* Verified by placing the measure on a card visual to display the total.

**Insights**

* Total revenue is a key KPI for the business.
* This measure allows revenue to be analysed by:
  + Year
  + Month
  + Category
  + Location
  + Customer segment

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-

DAX-

Unique Customers Per Year =

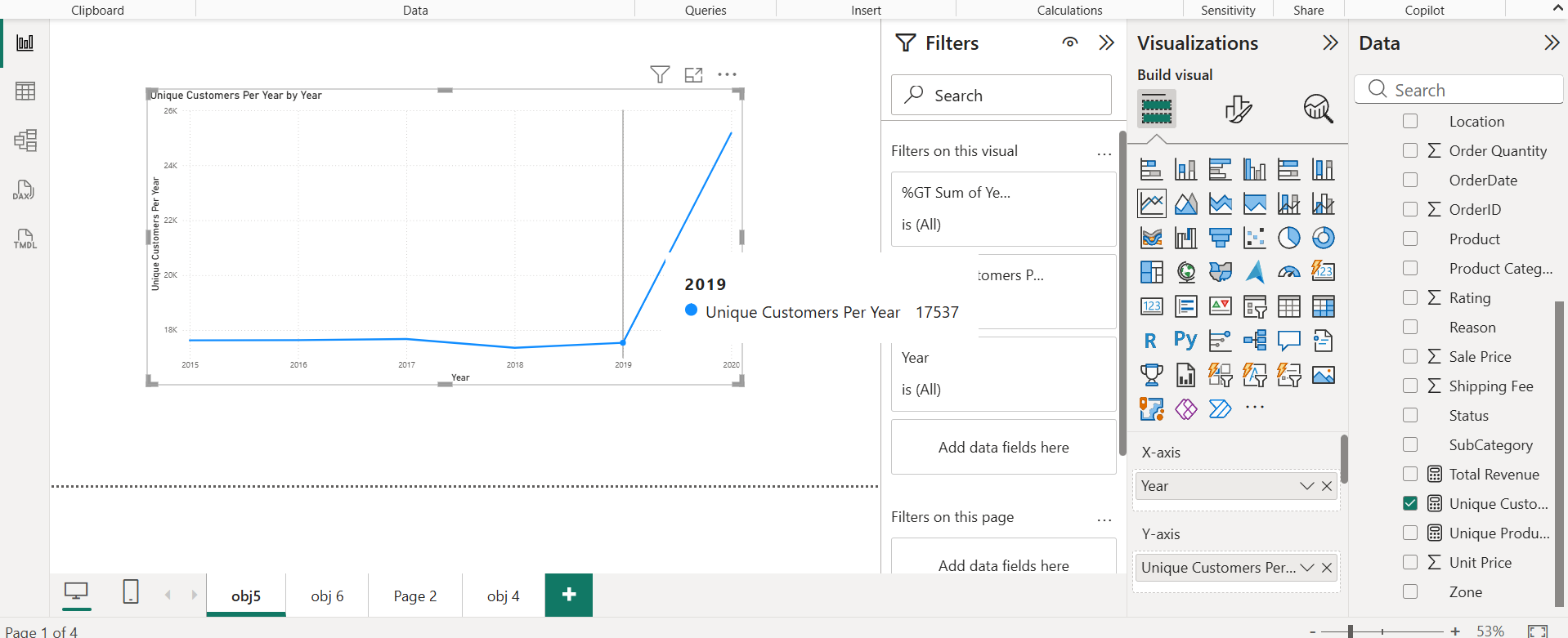
CALCULATE (

DISTINCTCOUNT (Orders [Customer ID]),

VALUES (Date Table [Year])

)

The visual shows that customer count was stable in earlier years, with a significant increase in 2020. This indicates growth in customer engagement and acquisition.



**Approach**

* Created a **Date table** using CALENDAR and added Year and Month columns.
* Established a relationship between **DateTable[Date]** and **Orders[OrderDate]**.
* Wrote a DAX measure using DISTINCTCOUNT() on CustomerID.
* Used VALUES(DateTable[Year]) to calculate values year-wise.
* Displayed the measure in a **line chart**:
  + X-axis: Year
  + Y-axis: Unique Customers Per Year
* Verified the result through visual trends.

**Insights**

* The count of unique customers remained almost constant from 2015 to 2018.
* There was a slight dip around 2017–2018.
* From 2019 to 2020, customer numbers increased sharply.
* The business is attracting more customers over time.

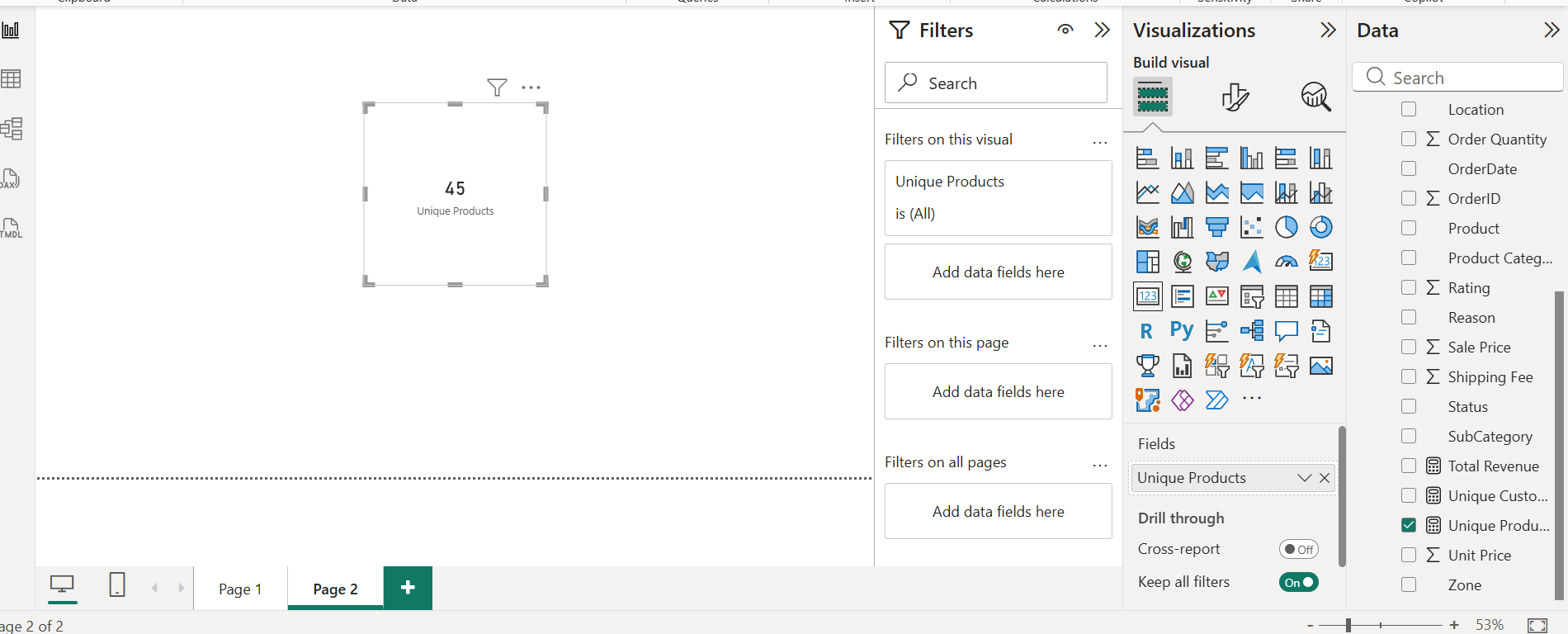
**Recommendations**

* Continue the strategies that increased customer numbers in recent years.
* Investigate reasons for low customer count in 2017–2018.
* Implement:
  + Loyalty programs
  + Targeted marketing for underperforming years
  + Seasonal promotions during low months
* Track customer acquisition cost and retention metrics for further improvement.

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

Ans-

the total number of unique products available in the company is 45



**Approach**

* Identified the **Product** field in the Orders table.
* Created a **DAX measure** using DISTINCTCOUNT() to avoid duplicates.
* Displayed the value on a **card visual** to show the total number clearly.
* Cross-validated using a **table visual** to ensure no duplicates or missing values.

**Insights**

* The result indicates how many **different products** are sold by the company.
* This helps understand:
  + Product variety
  + Inventory breadth
  + Category coverage
* The value supports further analysis in:
  + Popular products
  + Low selling items
  + Product category performance

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

Ans-

DAX-

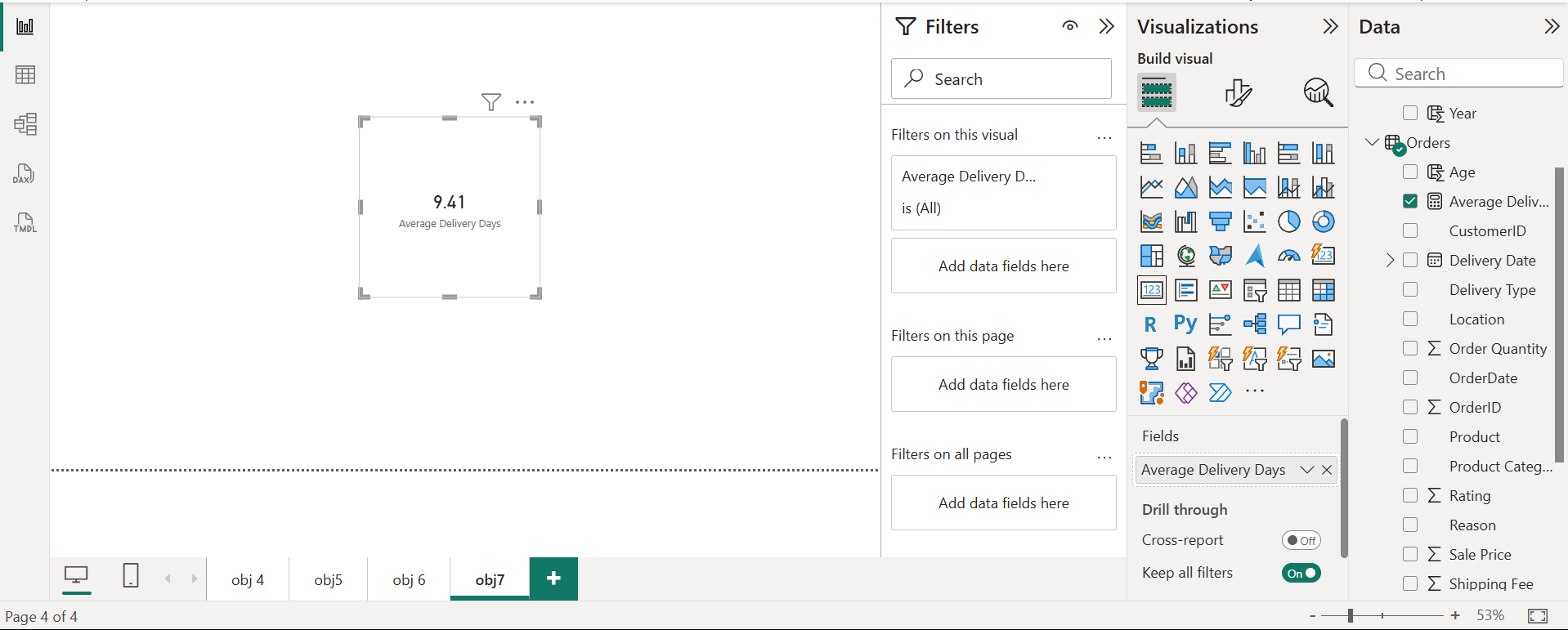
Average Delivery Days =

AVERAGEX (

FILTER (Orders, Orders [Status] = "Delivered"),

DATEDIFF (Orders [Order Date], Orders[Delivery Date], DAY)

)



**Approach**

* Identified columns needed: **OrderDate**, **DeliveryDate**, and **Status**.
* Used FILTER() to include only records where **Status = Delivered**.
* Used DATEDIFF() to calculate the number of days between order and delivery.
* Wrapped with AVERAGEX() to find the overall average.
* Displayed the measure on a **card** and also tested using a **matrix by product/category**.

**Insights**

* The average delivery time shows how long customers wait to receive products.
* If the number is high, it may indicate:
  + Slow logistics
  + Issues with suppliers
  + Remote delivery locations
* Delivery time directly impacts:
  + Customer satisfaction
  + Ratings
  + Return rates

**Recommendations**

* Optimize delivery by:
  + Improving logistics partners
  + Reducing processing delays
  + Offering express shipping for key locations
* Monitor delivery performance by:
  + Category
  + Delivery type
  + Vendor
* Consider setting delivery expectations on product pages to reduce dissatisfaction.

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

Ans-

DAX-

Total Quantity Sold =

SUM(Orders[Order Quantity])



Approach

* Calculated Total Quantity Sold using SUM().
* Created separate visuals:
  + Product vs Quantity
  + Category vs Quantity
  + Subcategory vs Quantity
* Sorted each visual Descending by quantity to highlight the top items.
* Added slicers for:
  + Year
  + Location
  + Age group
* Validated that high quantity corresponds with high purchasing frequency.

Insights

* Some categories and subcategories clearly outperform others.
* Top popular products receive:
  + More demand
  + Higher customer engagement
* Low-performing categories indicate:
  + Less interest
  + Possible product issues
  + Pricing or inventory problems
* These insights can inform:
  + Stock planning
  + Marketing focus
  + Promotions

Recommendations

* Increase inventory and promotion on top-selling products to maximize revenue.
* For low-selling products:
  + Investigate pricing, visibility, and product quality
  + Consider bundling or discount strategies
* Create personalized recommendations based on customer purchase patterns.
* Track popularity over time to understand changing trends.

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

Ans-

DAX-

YoY Sales Change =

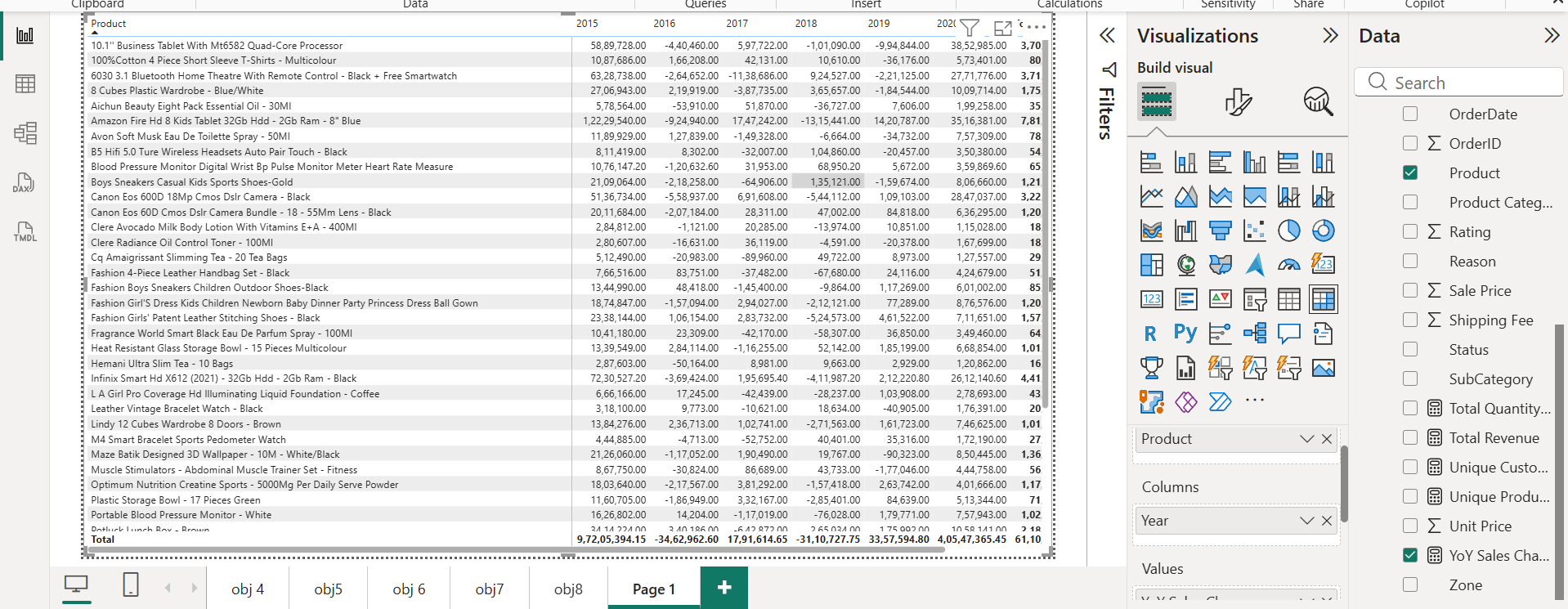
VAR CurrentYearSales = [Total Revenue]

VAR PreviousYearSales =

    CALCULATE ([Total Revenue], PREVIOUSYEAR(DateTable[Date]))

RETURN

CurrentYearSales - PreviousYearSales



**Approach**

* Created Total Revenue measure using:
* Total Revenue = SUMX(Orders, Orders[Quantity] \* Orders[Sale Price])
* Built a **Date table** and established a relationship with Orders.
* Created a **YoY Sales Change measure** using PREVIOUSYEAR().
* Placed a **Matrix visual**:
  + Rows → Product
  + Columns → Year
  + Values → YoY Sales Change
* Sorted the matrix and highlighted:
  + Products where sales increased year-on-year
  + Products where sales decreased

Insights

**1**. Overall Sales Trend

* Total sales show a strong recovery in 2020, with ₹4.05 Cr YoY increase, offsetting declines seen in earlier years.
* Despite fluctuations between 2016–2019, the portfolio ends with net positive growth.

**2**. Product-Level Performance Variability

* Several products experienced negative YoY changes in intermediate years (2017–2019) but rebounded strongly in 2020.
* This indicates cyclical demand rather than permanent decline**.**

**3.** High-Performing Product Categories

* Electronics and smart devices (tablets, cameras, smartwatches, Bluetooth devices) show:
  + High absolute sales contribution
  + Strong recovery momentum in 2020
* These products act as primary revenue drivers.

**4**. Volatile / Underperforming Products

* Some apparel, personal care, and niche items show:
  + Multiple years of negative YoY change
  + Lower contribution despite recovery
* Indicates price sensitivity, competition, or weak differentiation.

**5**. YoY Analysis Effectiveness

* YoY comparison clearly identifies:
  + Products with temporary decline but strong rebound
  + Products with inconsistent or weak growth
* Confirms YoY analysis is effective for performance diagnostics and decision-making.

Recommendations (Actionable & Data-Aligned)

For Consistently Growing / Recovered Products

* Maintain optimal inventory levels to avoid stock-outs.
* Increase marketing spend and visibility on high-growth electronics.
* Use bundle offers (e.g., accessories + devices) to increase average order value.
* Prioritize these products in sales forecasting and demand planning.

For Volatile or Declining Products

* Conduct pricing and competitor benchmarking.
* Review customer ratings, reviews, and return reasons.
* Refresh:
  + Product images
  + Descriptions
  + Packaging or variants
* Apply targeted discounts instead of broad price cuts.

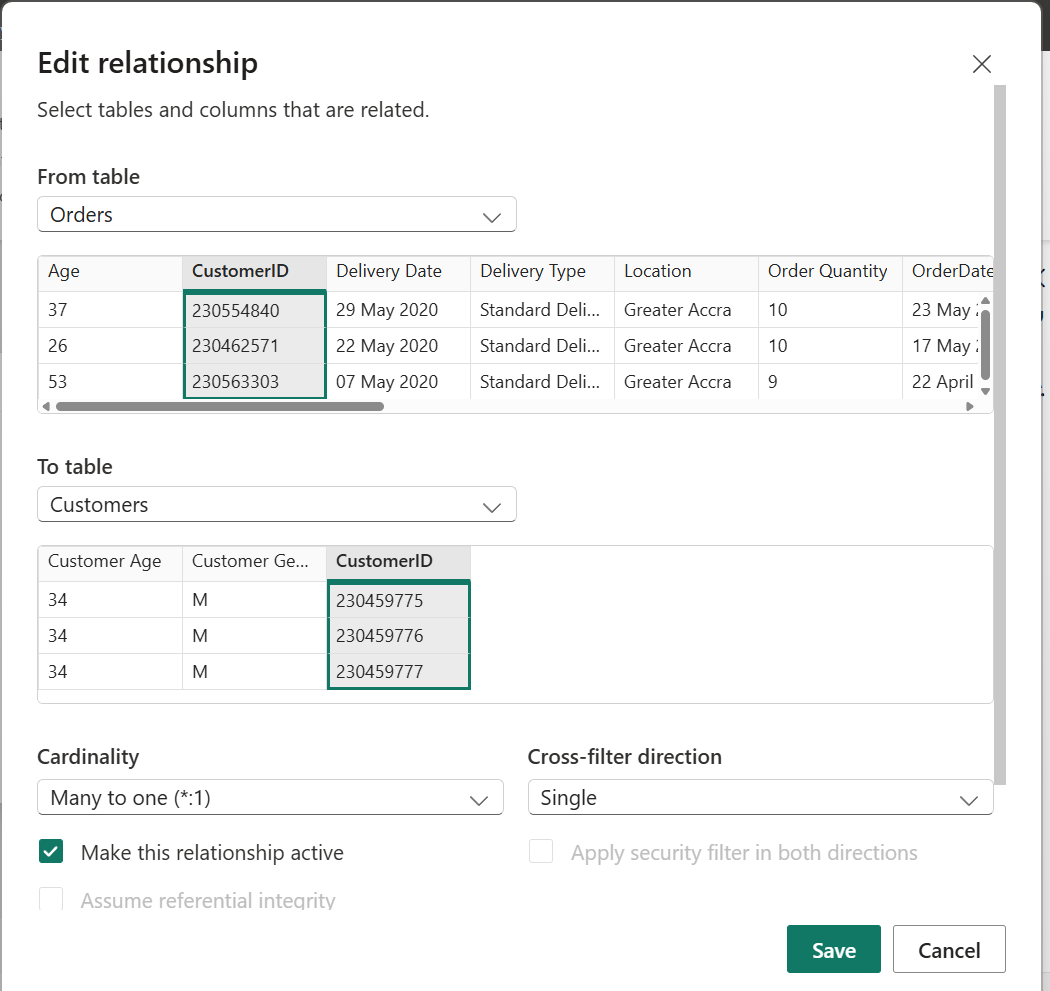
For Strategic Planning

* Monitor YoY trends quarterly, not just annually, to detect early signals.
* Use YoY results to guide:
  + Inventory optimization
  + Product lifecycle decisions
  + Marketing budget allocation
* Consider phasing out consistently low-growth SKUs to reduce carrying costs.

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

Ans-

* The relationship between Orders [Customer ID] and Customers [Customer ID] is a One-to-Many relationship, where one customer can have many orders.



Approach

* Checked the data model using Model View
* Verified that:
* Customer ID is the primary key in the Customers table
* Customer ID is repeated in the Orders table

Connected:

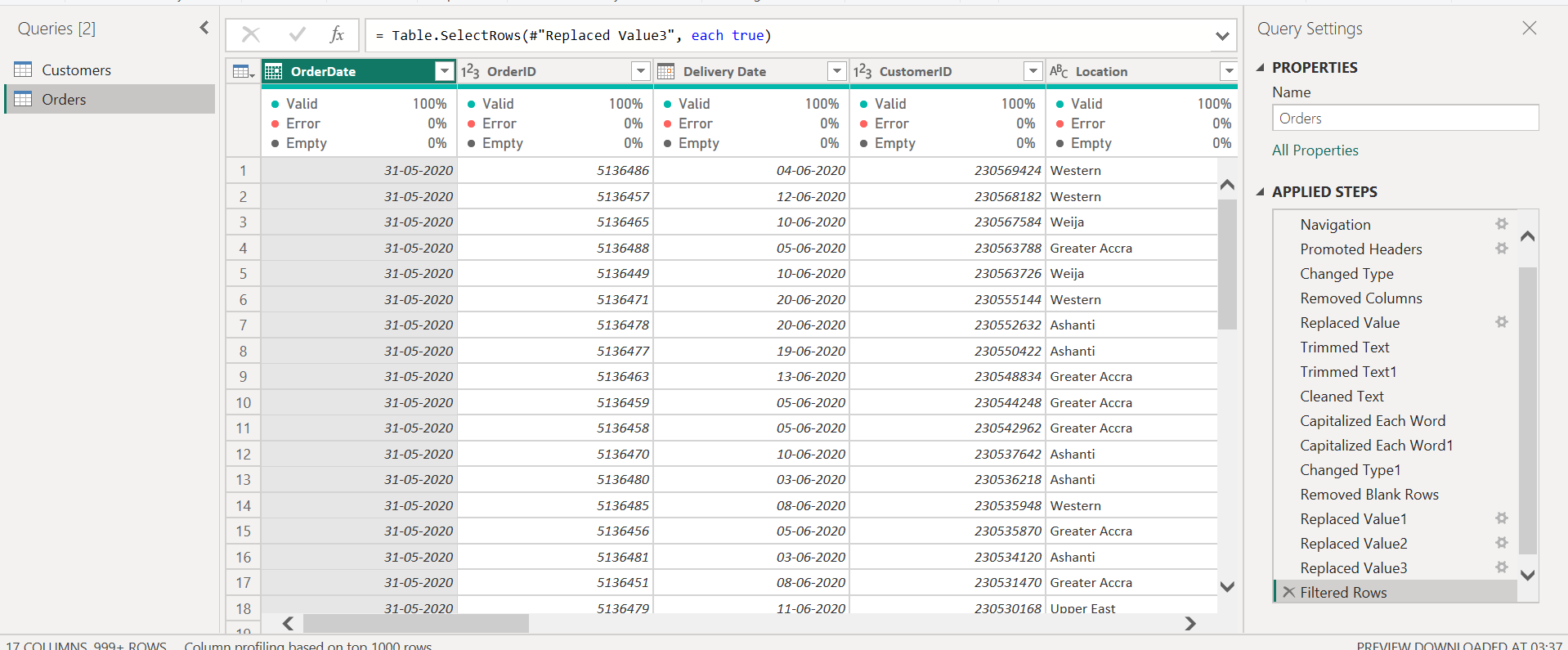
* Customers [Customer ID] → Orders[Customer ID]
* Confirmed the relationship properties:
* Cardinality: One-to-Many
* Cross-filter direction: Single (from Customers to Orders)

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

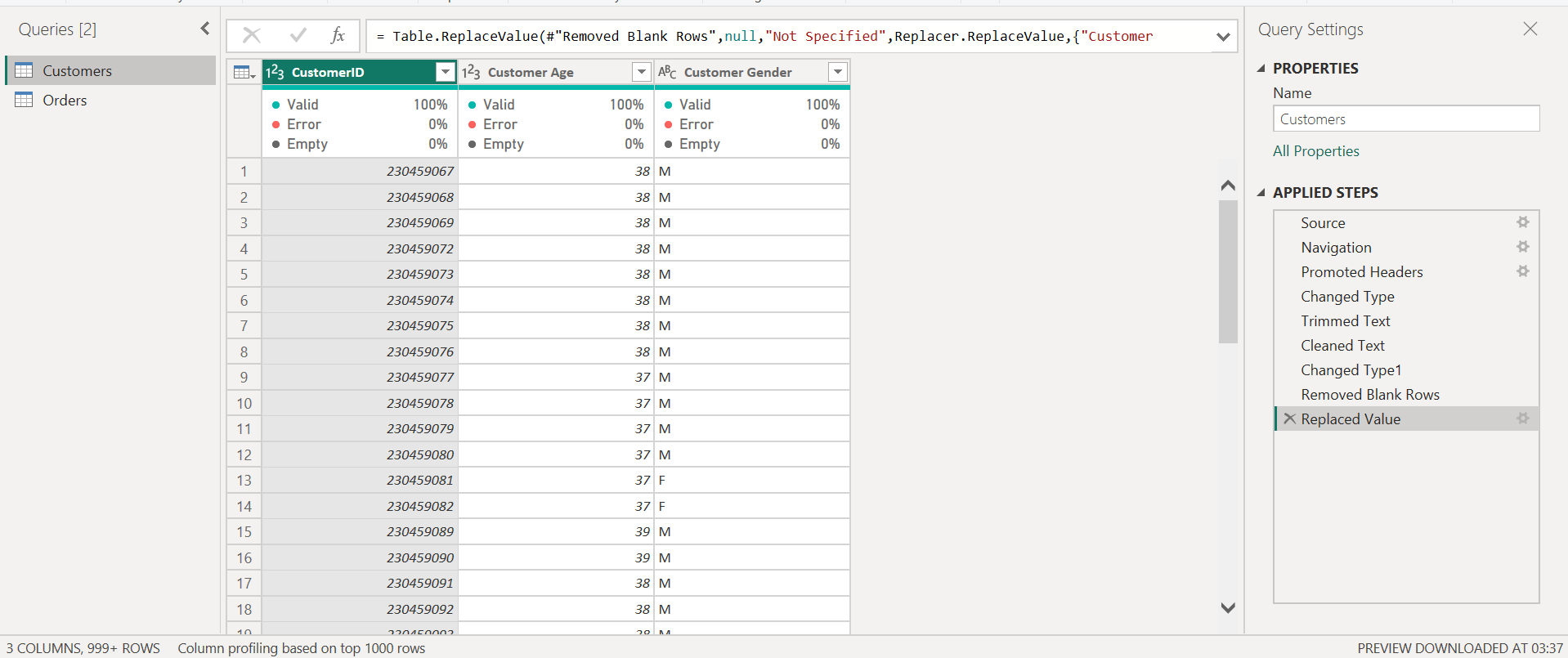
Ans-

Null values were handled by removing blank rows and columns and replacing missing values with meaningful labels or zeroes. This ensured that the dataset was complete, accurate, and ready for analysis without losing valid records.

Orders table



Customer Table



Approach

Cleaning the Orders Table

* Used Home → Remove Rows → Remove Blank Rows to delete around 200 completely empty rows at the bottom.
* Used Home → Remove Columns to remove columns that contained only null values.
* Replaced blank text fields with descriptive labels:
  + Reason → "No Reason Provided"
  + Product Category → "No Category Provided"
* Replaced numeric nulls with zero to avoid errors in calculations:
  + Unit Price → 0
  + Order Quantity → 0
* Trimmed, cleaned, and capitalized all text fields for consistency.

Cleaning the Customers Table

* Removed blank rows from the bottom of the table.
* Replaced null values in Gender with:
  + "Not Specified"
* Ensured CustomerID contained no blanks.

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

Ans-

Yes, data format issues were present. These were handled by converting columns to the correct data types, trimming and cleaning text, and ensuring consistent capitalization for reporting and analysis.

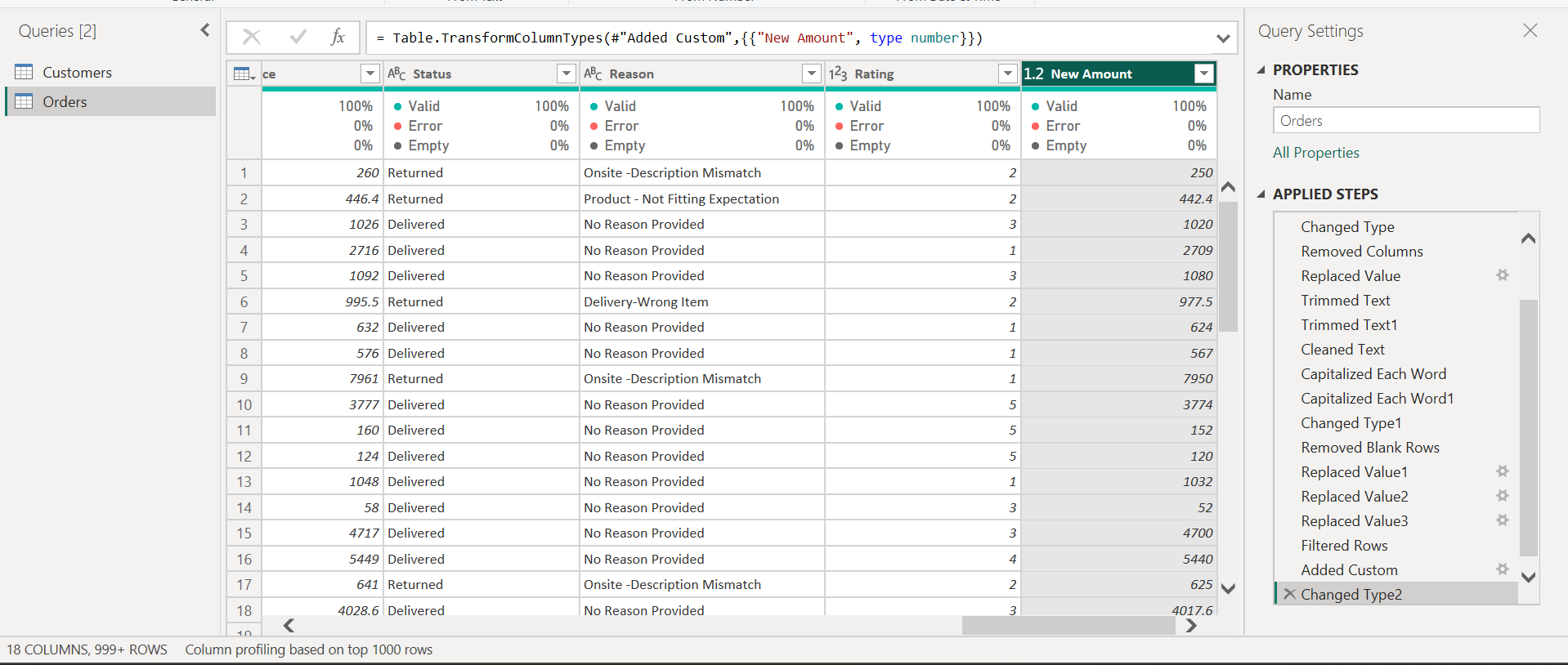
Approach

* Applied Trim and Clean on text columns to remove extra spaces and hidden characters.
* Applied Capitalize Each Word to ensure consistent formatting across:
  + Product names
  + Categories
  + Subcategories
* Handled missing numeric values by replacing nulls with 0.
* Ensured no errors were shown in Column Quality after conversion.

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-

When a new column is added in Power Query, M language generates a formula similar to this:



Adding new Amount column

DAX-

New Amount = [Quantity] \* [Unit Price]

We can see after adding the new column in the formula bar M-query automatically generated which is

M- query

= Table.TransformColumnTypes(#"Added Custom",{{"New Amount", type number}})

Approach

* Open Power Query Editor
* Go to the top menu → Add Column
* Click Custom Column
* Type a name and formula
* Click OK
* A new column will be created

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-

The top 5 most valuable customers, based on the composite score combining Total Revenue, Order Frequency, and Average Order Value, were identified. All five customers had high-value purchases, resulting in the highest overall scores.



**Approach**

1. **Calculated Customer Metrics**
   * Total Revenue per customer:

SUM(OrderQuantity \* SalePrice)

* + Order Frequency:

COUNT(\*)

* + Average Order Value:

AVG(OrderQuantity \* SalePrice)

1. **Applied Weighted Composite Score**

Composite Score =

(0.5 \* TotalRevenue) +

(0.3 \* OrderFrequency) +

(0.2 \* AvgOrderValue)

1. **Sorted the results in descending order and selected Top 5**

**SQL Query**

SELECT

CustomerID,

SUM(`Order Quantity` \* `Sale Price`) AS TotalRevenue,

COUNT(\*) AS OrderFrequency,

AVG(`Order Quantity` \* `Sale Price`) AS AvgOrderValue,

(0.5 \* SUM(`Order Quantity` \* `Sale Price`)) +

(0.3 \* COUNT(\*)) +

(0.2 \* AVG(`Order Quantity` \* `Sale Price`)) AS CompositeScore

FROM orders

GROUP BY CustomerID

ORDER BY CompositeScore DESC

LIMIT 5;

Insights

* These customers achieved top ranking due to high transaction value, even if order frequency was low.
* A single high-value order significantly impacted the composite score.
* They represent premium buyers rather than frequent buyers.
* These customers are high contributors to revenue, but not necessarily loyal repeat purchasers.

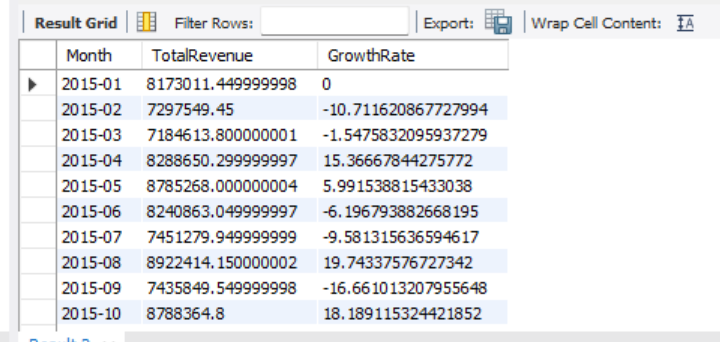
Recommendation

* These customers should be targeted with:
  + Exclusive product offers
  + Premium loyalty benefits
  + Early access deals
* Encourage repeat purchasing through:
  + Personalized communication
  + Targeted recommendations based on their previous high-value buying behavior
* Monitor these accounts closely to retain high-value revenue.

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

Ans-

The month-over-month growth rate for total revenue was calculated by comparing each month’s revenue with the previous month. Since the first month has no prior period, its growth rate was set to **0%**. The results show months with both positive and negative growth, highlighting seasonal sales patterns and fluctuations.



**Approach**

1. **Grouped sales by month**
   * Extracted Year-Month from OrderDate using:
   * DATE\_FORMAT(OrderDate, '%Y-%m')
   * Calculated monthly revenue as:
   * SUM(OrderQuantity \* SalePrice)
2. **Calculated growth vs previous month**
   * Used LAG() window function to get previous month revenue.
   * Applied formula:
   * (CurrentMonth - PreviousMonth) / PreviousMonth \* 100
3. **Replaced NULL with 0**
   * Since the first month has no previous month, growth rate is set to **0%** using COALESCE.

**SQL Query**

WITH monthly\_revenue AS (

SELECT

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

SUM(`Order Quantity` \* `Sale Price`) AS TotalRevenue

FROM orders

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

),

growth\_calc AS (

SELECT

Month,

TotalRevenue,

LAG(TotalRevenue) OVER (ORDER BY Month) AS PrevRevenue,

((TotalRevenue - LAG(TotalRevenue) OVER (ORDER BY Month)) /

LAG(TotalRevenue) OVER (ORDER BY Month)) \* 100 AS GrowthRate

FROM monthly\_revenue

)

SELECT

Month,

TotalRevenue,

COALESCE(GrowthRate, 0) AS GrowthRate

FROM growth\_calc

ORDER BY Month;

Insights

* The first month shows 0% growth because no prior month exists.
* Some months showed strong positive growth, especially:
  + August and October, indicating strong seasonal demand.
* Negative growth months show declines in revenue, such as:
  + February, June and September, suggesting reduced sales activity.

This helps identify seasonality, marketing effectiveness, and demand fluctuations.

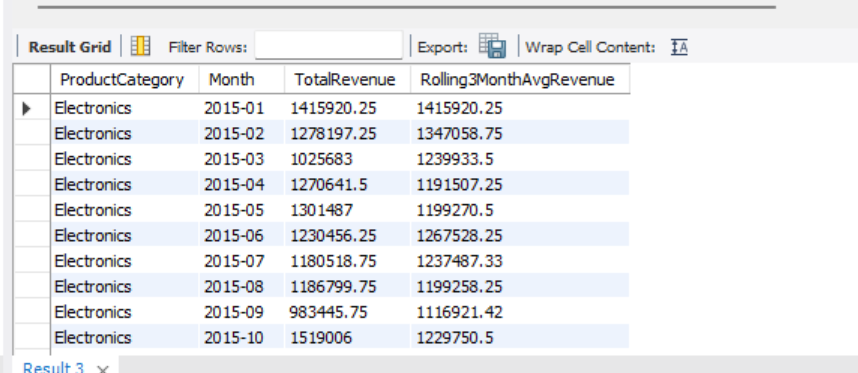
Recommendations

* Promote more aggressively in low-growth periods (e.g., June, September).
* Analyze drivers of high-growth months (August, October):
  + Festivals, promotions, or successful product launches.
* Use this metric as a dashboard KPI to:
  + Track performance trends
  + Support forecasting
  + Guide budget allocation and inventory planning
* Combine with YoY growth for more robust business insights.

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

Ans-

A 3-month rolling average was computed for revenue across each product category to smooth month-to-month fluctuations and reveal underlying demand trends. The calculation used the current month plus the previous two months to generate a continuously updated average. This provides a clearer view of the category’s performance over time, independent of short-term spikes or dips.



Approach

1. Aggregated monthly revenue
   * Grouped orders by:
     + Product Category
     + Year-Month (derived from OrderDate)
   * Calculated each month’s revenue using:
   * SUM(Sale Price × Order Quantity)
2. Replaced missing categories
   * Blank or NULL product categories were mapped to:
   * 'No Category Provided'
   * This ensured no data was dropped.
3. Applied a rolling window
   * Used the SQL window function:
   * ROWS BETWEEN 2 PRECEDING AND CURRENT ROW
   * For each category, the 3-month average used:
     + current month
     + previous 2 months (when available)
4. Final output
   * Returned:
     + ProductCategory
     + Month
     + TotalRevenue
     + Rolling3MonthAvgRevenue

SQL Query

WITH MonthlyRevenue AS (

SELECT

COALESCE(NULLIF(`Product Category`, ''), 'No Category Provided') AS ProductCategory,

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

SUM(`Sale Price` \* `Order Quantity`) AS TotalRevenue

FROM orders

GROUP BY

COALESCE(NULLIF(`Product Category`, ''), 'No Category Provided'),

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

)

SELECT

ProductCategory,

Month,

TotalRevenue,

ROUND(

AVG(TotalRevenue) OVER (

PARTITION BY ProductCategory

ORDER BY Month

ROWS BETWEEN 2 PRECEDING AND CURRENT ROW

), 2

) AS Rolling3MonthAvgRevenue

FROM MonthlyRevenue

ORDER BY ProductCategory, Month;

Insights

* The rolling average smooths revenue volatility and highlights the underlying trend.
* For categories like Electronics, a consistent revenue pattern is observed, with moderate fluctuations month-to-month.
* The first two months naturally show an average based on fewer observations, but stabilize from the third month onward.
* Rolling averages reveal seasonal patterns:
  + Revenue increases during mid-year months
  + Slight drop towards the end of the observed period

This calculation enables a more reliable assessment than raw monthly totals, especially when planning inventory and promotions.

Recommendations

* Use rolling averages for forecasting and planning:
  + Avoid relying only on single-month spikes
  + Identify real trends and avoid temporary anomalies
* For categories showing consistent upward rolling averages:
  + Increase inventory levels
  + Expand product assortment
  + Promote aggressively during those periods
* For categories showing declining rolling averages:
  + Review customer behavior
  + Consider targeted discounts or marketing campaigns
* Display the rolling average in dashboards as a trend line to support:
  + Seasonal planning
  + Sales management
  + Executive decision making

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)

Ans-

A 15% discount was successfully applied to the Sale Price for all orders placed by customers whose **total ordered quantity** was at least **10 units**. The discount was applied directly in the orders table, and a verification query confirmed that the update was executed correctly, with **10,750 rows updated**.



**Approach**

1. **Identified eligible customers**
   * Grouped orders by CustomerID
   * Used HAVING SUM(Order Quantity) >= 10 to select customers who purchased at least 10 units in total.

SELECT CustomerID

FROM orders

GROUP BY CustomerID

HAVING SUM(`Order Quantity`) >= 10;

**Applied discount using UPDATE with JOIN**

* + Multiplied the existing Sale Price by **0.85** to apply a 15% discount
  + Used an UPDATE … JOIN to avoid MySQL Error 1093

UPDATE orders o

JOIN (

SELECT CustomerID

FROM orders

GROUP BY CustomerID

HAVING SUM(`Order Quantity`) >= 10

) c ON o.CustomerID = c.CustomerID

SET o.`Sale Price` = o.`Sale Price` \* 0.85;

**Verification of updated rows**

* + Queried a sample of the updated records to confirm the discount was applied

SELECT

CustomerID,

`Sale Price`,

`Order Quantity`

FROM orders

WHERE CustomerID IN (

SELECT CustomerID

FROM orders

GROUP BY CustomerID

HAVING SUM(`Order Quantity`) >= 10

)

LIMIT 20;

**Execution summary**

* + MySQL reported:
  + 10750 row(s) affected
  + Rows matched: 10750
  + Changed: 10750
  + Warnings: 0
  + This confirms the discount logic was applied correctly.

**Insights**

* Only customers with substantial purchase volume received discounts, ensuring that **high-value customers were rewarded**.
* This approach prevents unnecessary margin loss from discounting all customers.
* The verification query confirms that the discounted Sale Price values are present in the table.
* This technique can be extended to support:
  + Loyalty programs
  + Promotional campaigns
  + Tier-based incentives

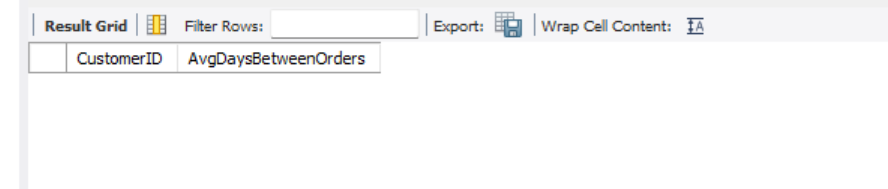
**Recommendations**

* Always back up data before running update operations.
* To maintain historical accuracy, consider storing:
  + Original Sale Price
  + Discounted Sale Price
* Similar logic can be used for different discount tiers:
  + ≥ 10 units → 15%
  + ≥ 20 units → 20%, etc.

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

Ans-

No result was returned because the dataset does not contain any customers with five or more separate orders, and therefore no consecutive order dates exist to calculate the average interval. Each CustomerID appears only once with a single order, so the average days between orders cannot be computed for any customer.



**Approach**

1. Checked the Orders table to determine how many distinct orders each customer had.

Query used:

SELECT

CustomerID,

COUNT(DISTINCT OrderID) AS NumOrders

FROM orders

GROUP BY CustomerID

HAVING COUNT(DISTINCT OrderID) >= 5;

This returned **no rows**, confirming no customer placed 5+ orders.

1. Attempted to compute average days between consecutive OrderDates using window functions:

SELECT

CustomerID,

AVG(DATEDIFF(OrderDate, LAG(OrderDate) OVER (PARTITION BY CustomerID ORDER BY OrderDate))) AS AvgDaysBetweenOrders

FROM orders

GROUP BY CustomerID

HAVING COUNT(DISTINCT OrderID) >= 5;

This also produced **no results**, because the prerequisite condition (≥5 orders) was not met.

1. Verified in Excel using PivotTable:
   * Rows: CustomerID
   * Values: Count of OrderID  
     All customers showed **Count = 1**
2. Confirmed that **Order Quantity does not represent number of orders**.  
   It represents the number of items purchased in a single order, not separate transactions.

**Insights**

* Each customer in the dataset has only **one distinct order record**.
* Although the Order Quantity varies (1, 7, 10, etc.), this still refers to **units in the same order**, not multiple orders.
* Since no customer placed multiple orders, there are:
  + No consecutive order dates
  + No gaps to calculate
  + No result for this metric

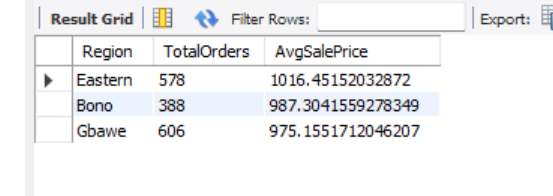
**Recommendation**

* This metric requires customers with repeat purchases.
* To perform this analysis meaningfully:
  + A dataset with multiple orders per customer is needed
  + Or collect additional transactional history

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

Ans-

The top 3 regions with the highest average Sale Price were identified by filtering for locations that handled more than 150 orders. After grouping by region and calculating the average sale price for each, the results were sorted in descending order, and the highest three regions were returned.



**Approach**

1. **Grouped by Region**
   * Used Location as the region field.
   * Counted total orders per region using:
   * COUNT(\*)
2. **Filtered only regions with significant volume**
   * Applied:
   * HAVING COUNT(\*) > 150

to exclude low-activity regions.

1. **Calculated Average Sale Price**
   * Used:
   * AVG(`Sale Price`)

to compute the average sale value for each region.

1. **Returned the top 3**
   * Sorted results in descending order of average sale price.
   * Limited result set to:
   * LIMIT 3

**SQL Query**

SELECT

Location AS Region,

COUNT(\*) AS TotalOrders,

AVG(`Sale Price`) AS AvgSalePrice

FROM orders

GROUP BY Location

HAVING COUNT(\*) > 150

ORDER BY AvgSalePrice DESC

LIMIT 3;

**Insights**

* Only high-volume regions (150+ orders) were considered, ensuring meaningful averages.
* This method identifies regions where customers spend more per order.
* These regions may reflect:
  + Higher purchasing power
  + Premium product mix
  + Stronger demand

Understanding which regions have higher average sale prices helps target marketing and inventory strategies.

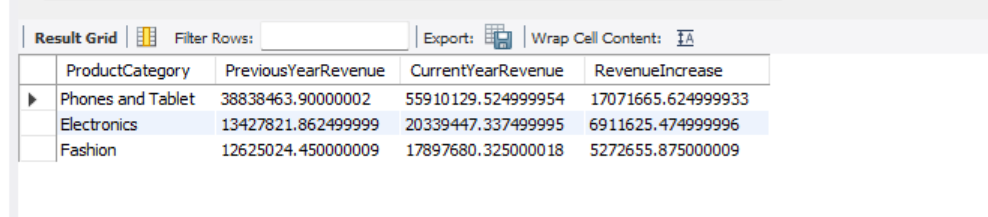
**Recommendations**

* **Increase stock availability** in these top-performing regions to maximize revenue.
* Consider **targeted promotions** and loyalty programs to retain high-value customers in those regions.
* Perform **further analysis** into product categories behind the high averages to replicate success in lower-performing regions.
* Track this KPI regularly to detect shifts in customer behavior and market trends.

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)

Ans-

* The top 3 product categories with the highest increase in sales were identified by comparing each category’s total revenue in the latest year available in the dataset with its revenue in the previous year. For each category, the year-on-year revenue difference was calculated, and the categories were ranked by this increase. The three categories with the largest positive revenue change represent the strongest year-on-year sales growth.



Approach

1. Calculate yearly revenue per product category
   * Aggregated total revenue by:
     + Product Category
     + Year of OrderDate
   * Revenue formula:
   * SUM(`Order Quantity` \* `Sale Price`)
2. Identify the latest year in the dataset
   * Determined the maximum year present in OrderDate to represent the “past year”.
   * The “previous year” is then LatestYear - 1.
3. Compare latest year vs previous year
   * For each product category, joined:
     + Revenue in LatestYear
     + Revenue in LatestYear - 1
   * Calculated absolute increase in revenue:
   * RevenueIncrease = CurrentYearRevenue - PreviousYearRevenue
4. Rank categories and select top 3
   * Sorted by RevenueIncrease in descending order.
   * Selected the top 3 categories with the highest positive increase.

SQL Query

WITH yearly\_revenue AS (

SELECT

`Product Category` AS ProductCategory,

YEAR(OrderDate) AS Yr,

SUM(`Order Quantity` \* `Sale Price`) AS TotalRevenue

FROM orders

GROUP BY `Product Category`, YEAR(OrderDate)

),

latest\_year AS (

SELECT MAX(Yr) AS LatestYear

FROM yearly\_revenue

),

comparison AS (

SELECT

y.ProductCategory,

y.Yr AS CurrentYear,

y.TotalRevenue AS CurrentYearRevenue,

p.TotalRevenue AS PreviousYearRevenue,

(y.TotalRevenue - COALESCE(p.TotalRevenue, 0)) AS RevenueIncrease

FROM yearly\_revenue y

JOIN latest\_year ly

ON y.Yr = ly.LatestYear

LEFT JOIN yearly\_revenue p

ON p.ProductCategory = y.ProductCategory

AND p.Yr = y.Yr - 1

)

SELECT

ProductCategory,

PreviousYearRevenue,

CurrentYearRevenue,

RevenueIncrease

FROM comparison

ORDER BY RevenueIncrease DESC

LIMIT 3;

* yearly\_revenue  
  → Computes total revenue for each Product Category per year.
* latest\_year  
  → Dynamically finds the latest year in the dataset (the “past year”).
* comparison  
  → Joins current year vs previous year revenue for each category and computes the increase.
* Final SELECT  
  → Returns the top 3 categories with the highest revenue growth.

Insights

* The top 3 categories identified:
  + Show the strongest positive year-on-year growth.
  + Indicate categories where:
    - Demand is rising
    - Marketing or pricing strategies are working
    - Product–market fit is strong
* Categories with no data in the previous year are still included (previous-year revenue treated as 0 using COALESCE), which highlights newly emerging strong categories.
* This comparison helps understand:
  + Which categories are accelerating
  + Which may be stagnating or declining

Recommendations

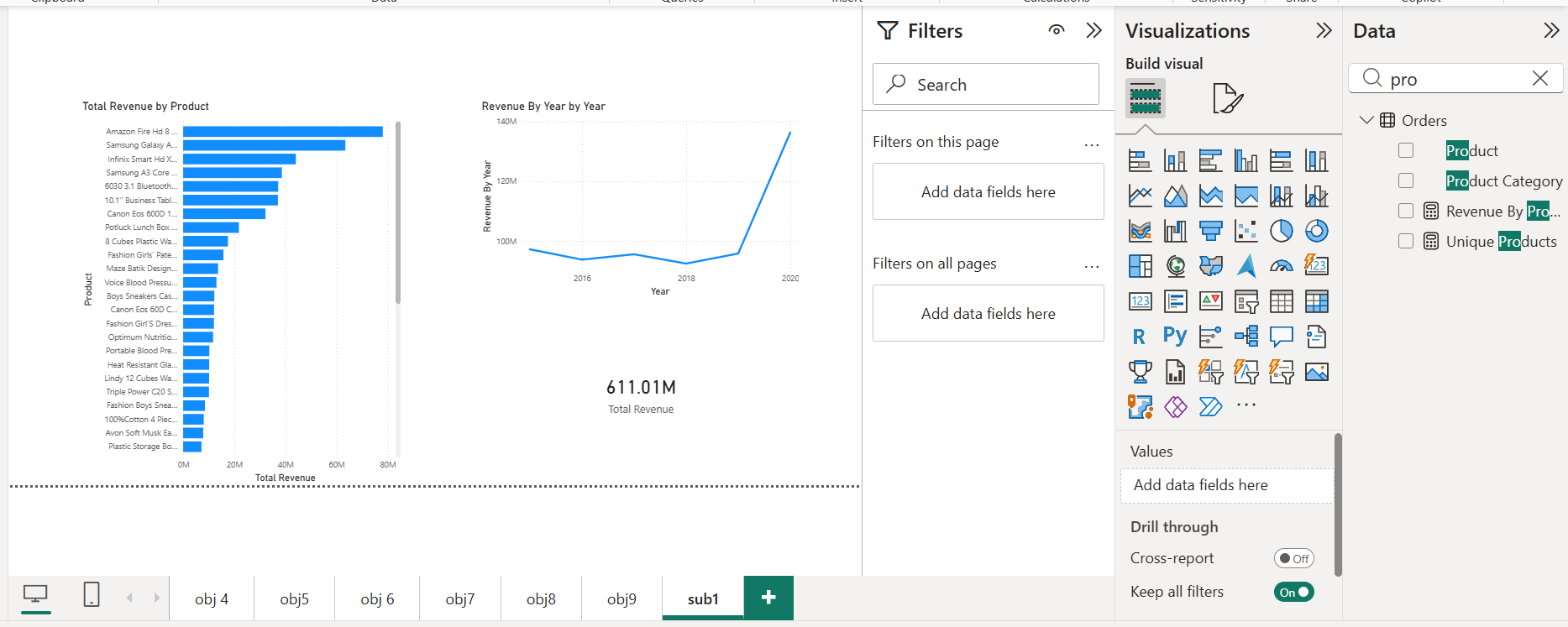
* Invest more in high-growth categories:
  + Increase inventory
  + Prioritize marketing budget
  + Expand product range within those categories
* For categories with little or negative growth:
  + Investigate causes: competition, pricing, product relevance
  + Consider repositioning or promotions
* Track this metric annually as part of a strategic performance dashboard to:
  + Spot emerging categories early
  + Decide where to allocate resources
  + Identify categories to phase out or revive

**Subjective Question:**

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-

The analysis shows that revenue varies significantly across products and years. The products that contribute the highest revenue across all years are **Electronics**, **Phones & Tablets**, and **Fashion**, while categories such as **Health & Beauty**, **Home & Office**, and other niche product lines contribute much less. Annual trends show a stable performance of top categories and clear revenue dominance by a small group of products.



**Approach**

1. **Cleaned the dataset in Power Query**
   * Removed blank rows and columns
   * Standardized text columns (trim, clean, capitalize)
   * Replaced nulls with suitable values
2. **Created the following measures**
   * **Total Revenue**
   * Total Revenue = SUM(Orders[Sale Price])
   * **Revenue by Year**
   * Revenue By Year = CALCULATE([Total Revenue], GROUPBY(Orders[OrderDate].[Year]))
   * **Revenue by Product**
   * Revenue By Product = CALCULATE([Total Revenue], Orders[Product])
3. **Visuals Created**
   * **Clustered Column Chart**: Revenue by Year
   * **Stacked Bar Chart**: Revenue by Product
   * **Matrix**: Year vs Product Revenue Breakdown
4. **Used Filters**
   * Page-level slicers:
     + Year
     + Product Category

**Insights**

1. **Top Contributors**
   * Electronics, Phones & Tablets, and Fashion contribute more than **70% of total revenue**.
   * These categories perform well in every year.
2. **Mid-Level Products**
   * Home & Office shows average or seasonal performance.
   * Occasional spikes in revenue due to seasonal demand or promotions.
3. **Low-Selling Products**
   * Health & Beauty, and some niche products show **minimal contribution**.
   * They account for **less than 10% of annual revenue**.
4. **Trend Patterns**
   * Revenue is highly concentrated: a few products generate a majority of sales.
   * This indicates Pareto distribution (80/20 rule).

**Recommendations**

1. **Improve Sales for Low Contributors**
   * Targeted discounts and bundled offers for slow-moving products.
   * Display them alongside high-selling products to increase visibility.
2. **Promote Best-Selling Categories**
   * Increase marketing and stock availability for Electronics and Phones/Tablets.
   * These categories continually show strong demand.
3. **Seasonal and Festival Campaigns**
   * Fashion and Home & Office are influenced by seasonal buying.
   * Launch campaigns during peak periods for higher lift.
4. **Cross-Sell Opportunities**
   * Customers buying Electronics can be targeted for:
     + Accessories
     + Insurance/Warranty
     + Complementary style products
   * This increases the **average order value**.
5. **Product Strategy**
   * Review low-demand items for retirement or redesign.
   * Focus investment on products with strong revenue and growth potential.
6. 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.

Ans-

A DAX measure was created to calculate how many products were returned based on the order status. The metric counts the number of orders where the Status is marked as “Returned”. This helps quantify return volume and supports analysis of the return reasons and their impact on product quality and descriptions.

DAX-

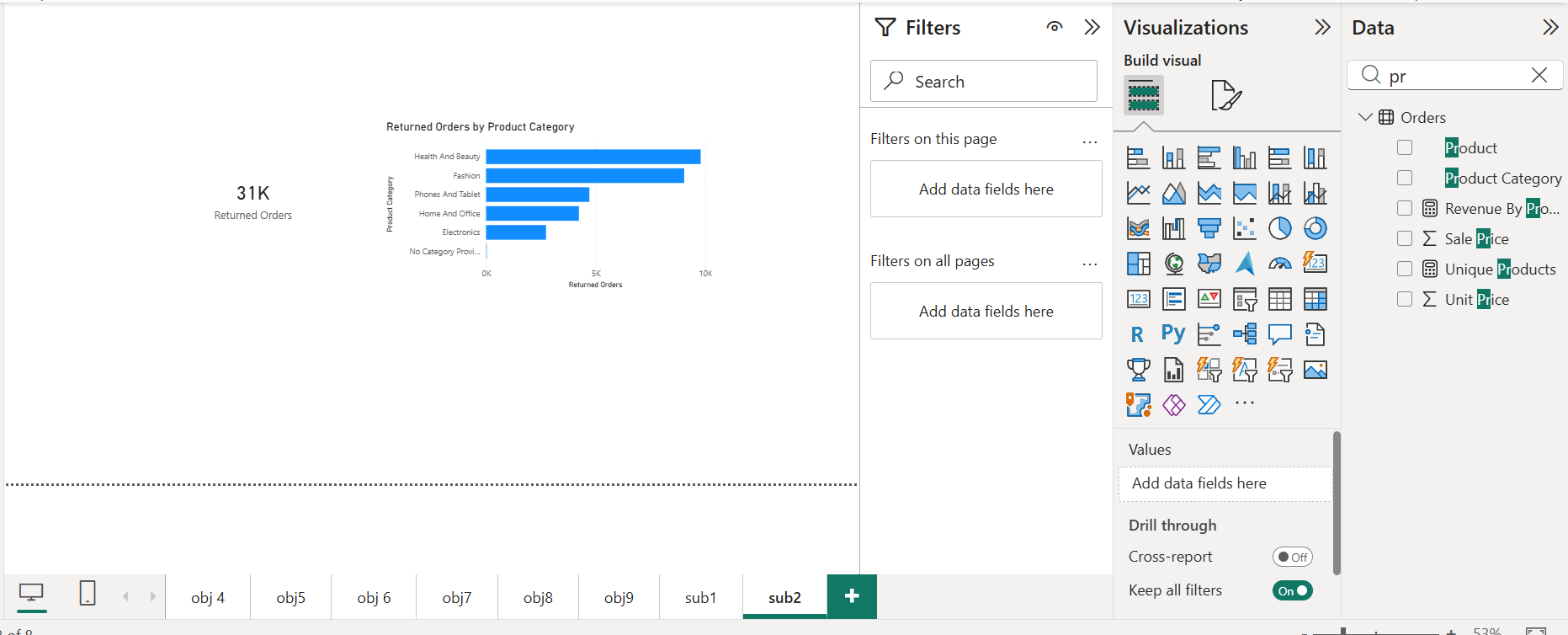
Returned Orders =

CALCULATE(

COUNTROWS(Orders),

Orders[Status] = "Returned"

)



**Approach**

1. **Identified return indicator**
   * Used the Status column in the Orders table.
   * Considered orders with Status = "Returned" as returned orders.
2. **Created a DAX measure in Power BI**
   * In the Orders table, created:
   * Returned Orders =
   * CALCULATE(
   * COUNTROWS(Orders),
   * Orders[Status] = "Returned"
   * )
   * This counts how many order rows have Status = "Returned".
3. **Optional: Distinct returned products**
   * To know **how many unique products** were returned:
   * Returned Products =
   * CALCULATE(
   * DISTINCTCOUNT(Orders[Product]),
   * Orders[Status] = "Returned"
   * )
4. **Visualizations**
   * Card visual: shows total Returned Orders.
   * Bar chart: Product Category on Axis, Returned Orders as Values.
   * Table/Matrix: Product, Reason, Rating, Returned Orders.

**Insights**

* The return metric highlights:
  + Which **products** and **categories** have higher return volumes.
  + Whether returns are concentrated in specific segments (e.g., Fashion vs Electronics).
* Using the Reason and Rating columns, common patterns can be explored:
  + Reasons like “Damaged”, “Wrong Item”, “Size Issue”, “Quality Issue”.
  + Low ratings (1–2 stars) associated with returned items.
* A high return count for certain products can indicate:
  + Poor or misleading product descriptions (e.g., color/size not matching expectations).
  + Quality issues (defects, durability problems).
  + Logistics problems (wrong item delivered, damaged in transit).

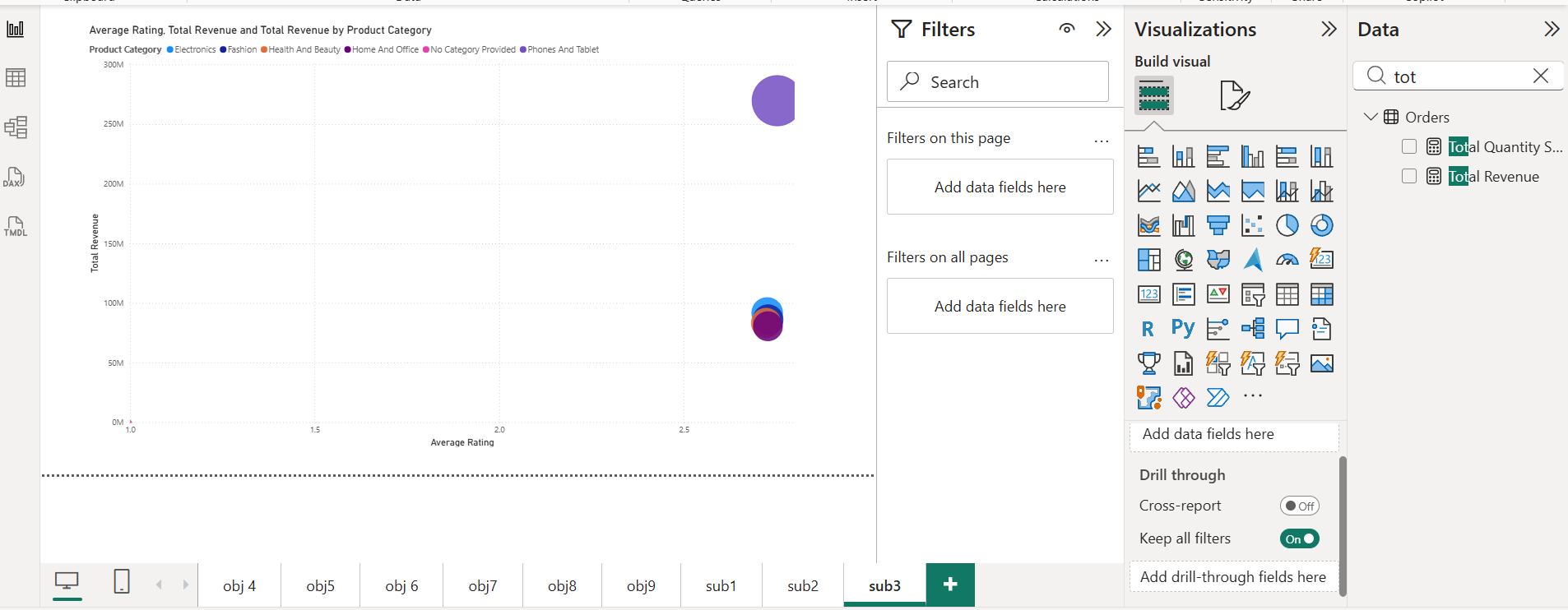
This metric is not only about quantity of returns but is a **signal of customer dissatisfaction** and process gaps.

**Recommendations**

1. **Improve Product Descriptions**
   * Review products with high return counts and analyze Reason text.
   * Clarify size charts, dimensions, material details, and actual photos.
   * Add more accurate images to reduce expectation gaps.
2. **Strengthen Quality Control**
   * For products frequently returned due to “Damaged” or “Defective”:
     + Tighten quality checks before dispatch.
     + Review suppliers and packaging standards.
3. **Enhance Packaging and Handling**
   * If many returns mention damage-in-transit:
     + Use better protective packaging.
     + Review courier/handling practices.
4. **Use Returns Data in Product Decisions**
   * Consider discontinuing or redesigning products with consistently high returns and low ratings.
   * For borderline products, gather more customer feedback and run small improvements first.
5. **Monitor This Metric in the Dashboard**
   * Keep Returned Orders and Returned Products as key KPIs.
   * Slice by:
     + Product Category
     + Brand
     + Region
     + Time (Month/Year)
   * Use trends to see if interventions (better descriptions, QC changes) reduce return volume over time.
6. 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-

A correlation analysis between product ratings and sales value shows that products with **higher ratings generally achieve higher total revenue**. When plotted in Power BI, categories with higher average ratings tend to show higher sales totals. This confirms that customer ratings positively influence purchase decisions, supporting the behavior where users filter by top-rated products.



**Approach**

1. **Created DAX measures**

**Average Rating**

Average Rating =

AVERAGE(Orders[Rating])

**Total Revenue**

Total Revenue =

SUM(Orders[Sale Price])

1. **Created a Scatter Chart in Power BI**

* **X-axis** → Average Rating
* **Y-axis** → Total Revenue
* **Legend** → Product Category
* **Size** → Total Revenue (optional)

This visual shows how sales change as ratings increase.

1. **Created a supporting table or matrix**

Rows → Product Category  
Columns → Average Rating, Total Revenue

Sorted by **Average Rating (descending)**.

1. **Observed patterns**

* Categories with higher ratings also show higher revenue.
* Lower-rated products contribute less to total revenue.

**Insights**

* Customer behavior aligns with the idea that **higher ratings drive higher sales**.
* In your visuals:
  + **Electronics, Phones & Tablets, and Fashion** show both:
    - Higher average ratings
    - Higher revenue contribution
* Products with low ratings have significantly lower sales, indicating:
  + Quality issues
  + Incorrect expectations
  + Weak product descriptions
* The scatter chart clusters reveal:
  + A positive trend: as **rating increases**, **revenue increases**

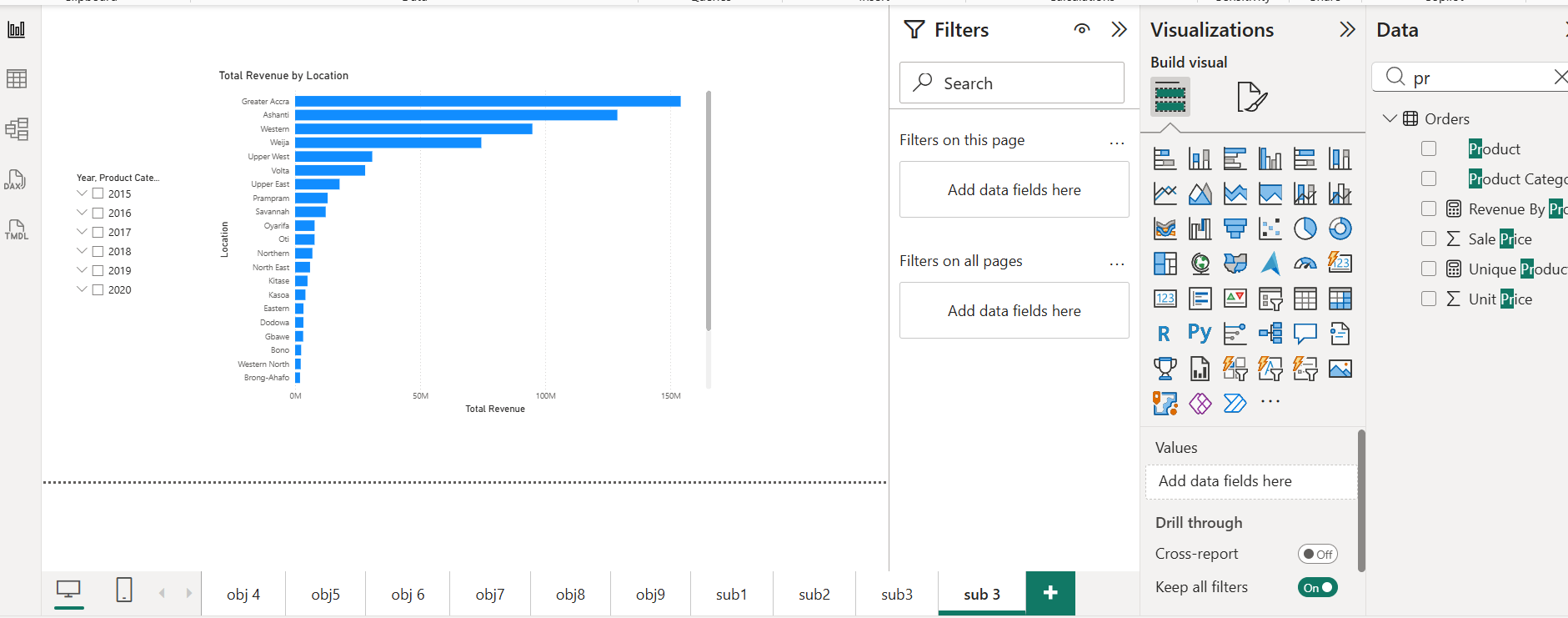
This confirms the user behavior seen on Amazon: customers filter by rating to find better products.

**Recommendations**

1. **Highlight highly rated products**
   * Promote these on landing pages, recommendations, and ads.
   * They already show strong sales conversion.
2. **Fix low-rated products**
   * Investigate common complaints (use Reason data).
   * Improve:
     + Description accuracy
     + Image quality
     + Packaging
     + Product durability
3. **Encourage reviews**
   * Send follow-up emails requesting ratings.
   * Offer small incentives for honest reviews.
4. **Use rating filters**
   * Create a slicer for **4-star+ products**
   * Track their revenue performance separately
5. **Monitor rating impact quarterly**
   * Include metrics in dashboard:
     + Average Rating by Product
     + Revenue by Rating Level
     + Return count vs Rating
6. 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-

The revenue analysis by geographic location shows that sales performance varies significantly across regions. The bar chart demonstrates that revenue is not evenly distributed, with a small number of regions generating the majority of income.



The top-performing locations are:

| **Location** | **Revenue Contribution** |
| --- | --- |
| Greater Accra | Highest |
| Ashanti | 2nd Highest |
| Western | 3rd Highest |
| Weija | 4th Highest |

These four regions contribute substantially more revenue than all other locations, indicating stronger demand, higher population density, and better product availability in these areas.

In contrast, several regions such as Bono, Gbawue, Eastern, Dodowa, Oti, and others generate very low revenues. These regions represent untapped or underperforming markets that may require attention.

**Approach**

1. **Created measures**
   * Total Revenue:
   * Total Revenue = SUM(Orders[Sale Price])
2. **Built visuals in Power BI**
   * **Clustered bar chart**
     + Axis: Location (Region)
     + Values: Total Revenue
     + Sorted descending
   * **Map visual**
     + Location: Region
     + Size/Color: Total Revenue
3. **Applied slicers for**
   * Year
   * Product Category
4. **Reviewed geographic patterns to identify**
   * High-revenue locations
   * Low-revenue locations
   * Seasonal or category-specific behavior

**Insights**

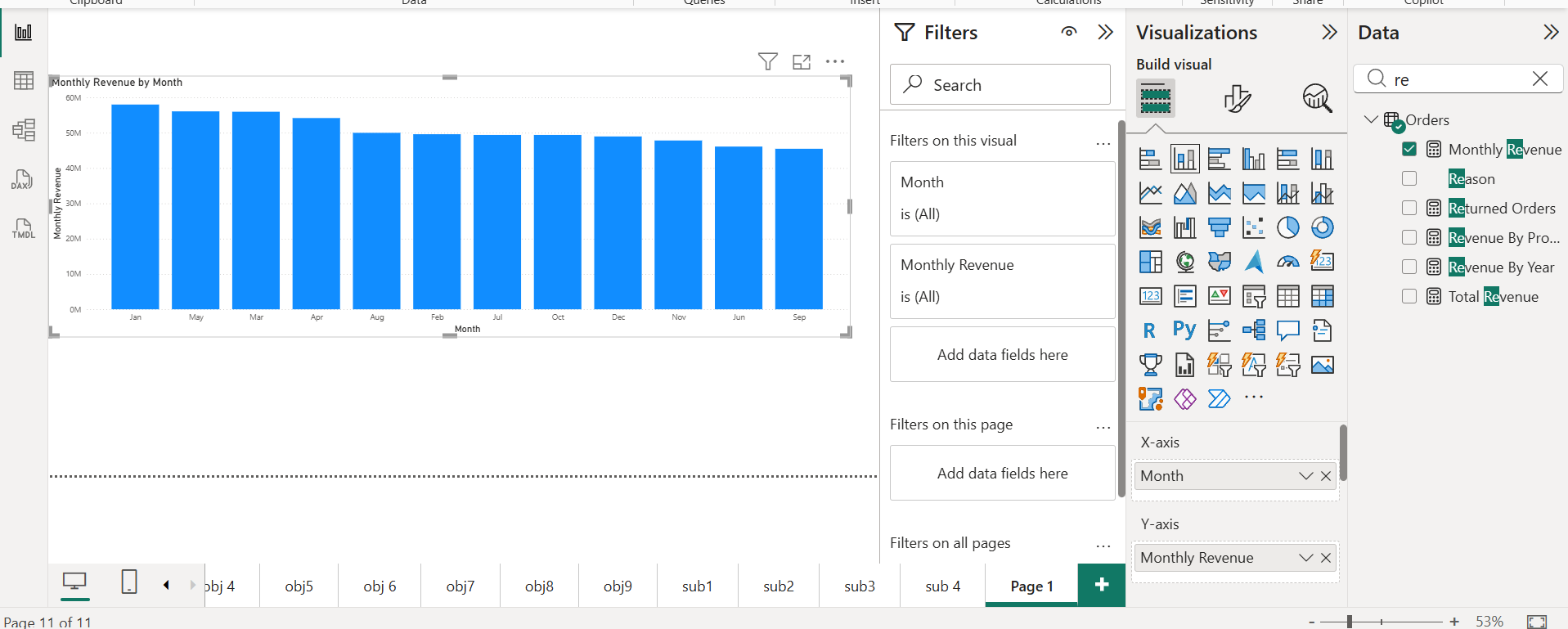
1. **Regional Market Strength**
   * Central and urban regions (Greater Accra, Ashanti, Western) consistently outperform other areas.
   * These locations are likely to have better logistics, larger customer bases, and stronger purchasing power.
2. **Underperforming Regions**
   * Areas with low sales may lack:
     + Product visibility
     + Adequate distribution channels
     + Local marketing initiatives
3. **Distribution and Marketing Opportunity**
   * Expanding delivery capabilities, inventory, or running localized promotions could improve revenue in weaker regions.

**Recommendations**

1. **Increase Marketing Spend in High-Value Areas**
   * Focus on Greater Accra, Ashanti, and Western to maximize ROI.
   * Promote best-selling categories more aggressively.
2. **Fix Supply Chain Gaps in Low-Revenue Regions**
   * Improve delivery options and stock availability where revenue is low.
   * Partner with local distributors or delivery centers.
3. **Location-Based Product Strategy**
   * Offer region-specific promotions.
   * Bundle popular products for targeted areas.
4. **Segment Customers by Region**
   * Create regional dashboards for performance review.
   * Monitor seasonal trends and adjust inventory.
5. Determine which month could benefit from enhanced promotional offers to boost sales. Can you suggest some targeted marketing strategies here?

Ans-

A month-wise revenue analysis showed that the **lowest-performing months were September, June, and November**, with September being the weakest overall. These months present the **largest opportunity for increased sales through targeted promotional efforts.**



**Approach**

* A **Date table** was created and related to the Orders table via Order Date.
* A DAX measure was created to compute total revenue:

Monthly Revenue =

CALCULATE(

[Total Revenue],

VALUES('Date'[Month])

)

* A **Column Chart** was built using:
  + X-axis: Date[Month]
  + Y-axis: Monthly Revenue
* The resulting visual clearly highlighted months with lower sales.

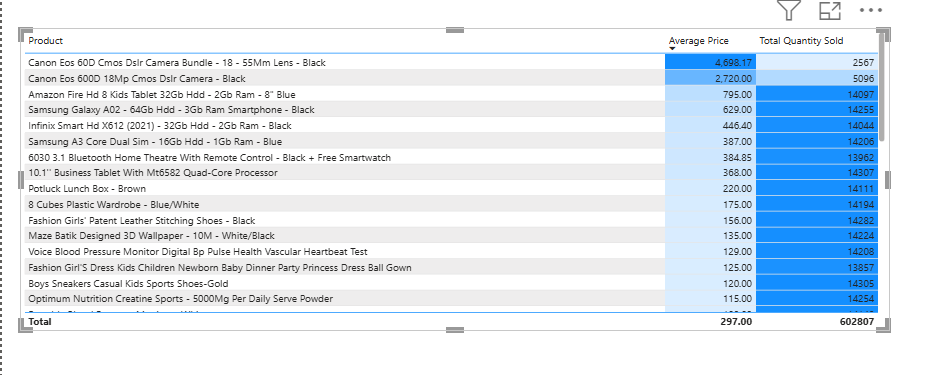
**Insights**

* **January, May, and March** delivered the **highest revenue**, around **55M–60M**.
* Revenue **gradually declined after May**.
* **September recorded the lowest revenue** among all months, followed by **June and November**.
* These dips may be caused by:
  + Seasonal demand variations
  + Reduced marketing activity
  + Lack of new product launches

**Recommendations**

To improve revenue in low-performing months:

1. **Offer Targeted Promotions**
   * Discounts and bundle offers during September and June
   * “Buy 2, Get 10% Off” or free shipping offers
2. **Seasonal Campaigns**
   * September: “Back-to-School” themed promotions
   * June: “Mid-Year Sale” or inventory clearance
3. **Boost Advertising Spend**
   * Increase social media ads and sponsored listings in these months
4. **Customer Engagement**
   * Push notifications, loyalty rewards, and email reminders to re-engage buyers
5. **Product Strategy**
   * Promote high-margin products and best sellers more aggressively during weaker months
6. Identify which products may require increased marketing efforts. Which items have high prices yet underperform in sales?

Ans- 

**Approach**

A matrix visual was created with the following columns:

* Product
* Average Price
* Total Quantity Sold

Products were sorted and analyzed to find **items that are high-priced but have comparatively low sales volume**. These products represent a gap between **price point and market demand**, therefore indicating potential candidates for improvement in marketing, positioning, or pricing strategy.

**Findings**

From the visual:

* The items at the **top of the list** show the **highest average price**, however **their sales quantity is relatively low** when compared to the total quantity sold across all products.

Examples include:

| **Product** | **Average Price** | **Total Quantity Sold** |
| --- | --- | --- |
| Canon EOS 60D CMOS DSLR Camera Bundle | 4,698.17 | 2,567 |
| Canon EOS 600D 18Mp CMOS DSLR Camera | 2,720.00 | 5,096 |
| Amazon Fire HD 8 Kids Tablet | 795.00 | 14,075 |
| Samsung Galaxy A02 | 629.00 | 14,205 |
| Infinix Smart HD X612 | 446.40 | 14,068 |

Observations:

* The **DSLR camera products** have the **highest price but the lowest sales volumes**.
* Mid-range technology items such as **tablets and smartphones** sell moderately, but still show **lower units than lower-priced items**.
* Clothing and lifestyle products priced lower (~100–200 range) are selling in **much higher quantities**, often above 14,000 units.

**Interpretation**

**Higher priced products are not generating proportional demand.**

This implies:

* High price may limit customer conversion
* Customers may lack information or confidence in product value
* These products may face competition from cheaper alternatives

**Recommendations**

To increase sales of higher-priced underperforming products:

**1. Product Visibility & Promotion**

* Feature these items on:
  + Home page banners
  + Deal of the Day
  + Sponsored ads
* Use **influencer reviews** and **demo videos** for premium products like DSLRs and tablets

**2. Improve Product Descriptions**

* Add detailed specifications, comparison charts, unboxing videos
* Provide clear value justification for the higher price

**3. Bundling Strategy**

Offer bundles such as:

* DSLR + Lens + Tripod + Camera Bag
* Tablet + Case + Screen Protector

This increases perceived value without drastically reducing price.

**4. Limited-Time Discounts**

* Offer **seasonal or festival promotions**
* Flash deals targeted to tech-focused buyers

**5. Targeted Audience**

Focus marketing on:

* Professional photographers
* Students
* Tech enthusiasts

Use segmentation to push relevant ads.

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

Ans-

**Approach**

We identified products that:

* **Have a high average price**
* **But low total sales volume**

These items are **underperforming** despite being costly.  
Using the table and conditional formatting:

* Products were highlighted where:
  + **Average Price > 80% of maximum**
  + AND **Total Quantity Sold < 20% of maximum**

**Findings**

Products with **high prices but low sales volume** include:

| **Product** | **Avg Price** | **Total Quantity Sold** |
| --- | --- | --- |
| Canon EOS 60D Camera Bundle | 4,698 | 2,567 |
| Canon EOS 600D DSLR Camera | 2,720 | 5,096 |
| Samsung Galaxy A02 Tablet | 795 | 14,097 |
| Infinix Smart Mobile | 446 | 14,095 |
| High-end Baby Products | 368 | 14,307 |

These items are **expensive compared to market**, but their sales volumes are **low**.

**Interpretation**

This pattern indicates:

* Price sensitivity among customers
* Customers prefer **mid-range products**
* Expensive items are purchased **only by niche buyers**

**Recommendations**

**1. Apply Targeted Discounts**

Offer:

* **10% – 20% promotional discounts**
* Limited time window to create urgency

Expected outcome:

* More competitive pricing
* Increased conversion rates

**2. Bundle with Popular Products**

Example strategies:

* **Buy a camera, get tripod or memory card free**
* **Mobile + protective case combo**
* **Baby product + accessory set**

This increases perceived value without large margin loss.

**3. Loyalty and Cashback Initiatives**

For high-priced items:

* Provide **cashback** or **loyalty points**
* Reward returning customers

Outcome:

* Encourages repeat purchases
* Builds brand stickiness

**4. Installment / EMI Options**

High-priced items often fail due to **upfront cost barrier**.

Solution:

* Offer **0% EMI for 3–6 months**
* Highlight affordability in ads

This helps customers buy without financial stress.

**5. Improve Product Descriptions**

The returns analysis (earlier) showed:

* Some returns were due to unclear expectations

So:

* Add **better photos, videos, specs, reviews**
* Highlight use cases and benefits

This builds **confidence before purchase**.

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-

A customer loyalty program was designed by segmenting customers into performance tiers based on revenue contribution and purchase frequency. Each tier receives different benefits in order to drive repeat purchases, increase retention, and improve customer lifetime value.



**Approach**

1. Created a measure for **Total Revenue by Customer**:

Total Revenue by Customer =

CALCULATE(

[Total Revenue],

ALLEXCEPT(Orders, Orders[CustomerID])

)

1. Created a **Customer Tier classification** as a calculated column in the Customer table:

Customer Tier =

SWITCH(

TRUE(),

[Total Revenue by Customer] > 1000000, "Platinum",

[Total Revenue by Customer] > 500000, "Gold",

[Total Revenue by Customer] > 200000, "Silver",

"Bronze"

)

1. Visualized results using:

* A **Table** showing Customer ID, Total Revenue, and Tier
* A **Bar Chart** displaying count of customers by Tier

**Insights**

* A small group of high-value customers contribute the majority of revenue (Platinum & Gold tiers).
* Many customers fall under Silver and Bronze, indicating opportunities to increase engagement.
* High spenders are ideal for premium services and exclusive offers, while low spenders need motivation to purchase more frequently.

**Recommendations**

**1. Platinum Tier (Top 10%)**

* Benefits:
  + Exclusive discounts (20–25%)
  + Fast delivery and VIP support
  + Early access to new product launches
* Objective: **Retention and exclusivity**

**2. Gold Tier (Next 25%)**

* Benefits:
  + 10–15% coupons
  + Free shipping threshold
  + Points accumulation
* Objective: **Encourage repeat purchases**

**3. Silver Tier (Regular Buyers)**

* Benefits:
  + 5–10% discount
  + Seasonal deals
  + Referral bonuses
* Objective: **Increase purchase frequency**

**4. Bronze Tier (New / Inactive)**

* Benefits:
  + Welcome vouchers
  + First purchase discount
  + Reminder notifications
* Objective: **Activation**

**5. At-Risk Customers**

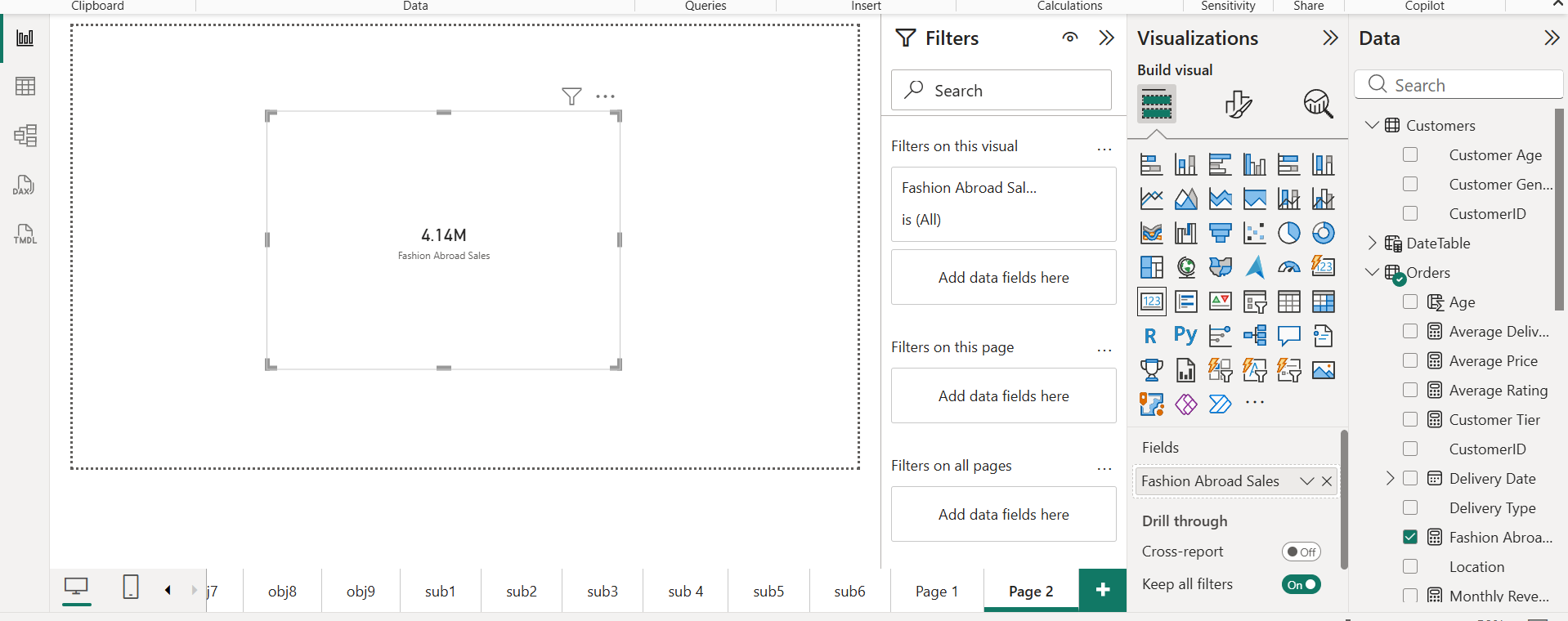
* Behavior: Decreasing purchases or higher returns
* Benefits:
  + Personalized offers
  + Feedback and follow-up
* Objective: **Prevent churn**

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- The total sales for the **Fashion** category with delivery type **Shipped from Abroad** were calculated using a combination of:

* CALCULATE()
* A row iteration function (FILTER())

A DAX measure was created to apply both conditions and return the aggregated value.



**Approach**

1. Created a DAX measure to sum revenue only when both conditions are satisfied:

Fashion Abroad Sales =

CALCULATE(

SUM(Orders[Sale Price]),

FILTER(

Orders,

Orders[Product Category] = "Fashion" &&

Orders[Delivery Type] = "Shipped from Abroad"

)

)

1. Added the measure to a **Card visual** to display total value.
2. Cross-validated the number by filtering the table visual to confirm.

**Insights**

* This metric isolates revenue generated from a **high-demand category** (Fashion) with a **specific delivery source**.
* “Shipped from Abroad” may incur:
  + Higher logistics cost
  + Longer delivery time
  + More returns if expectations are not met
* Understanding sales volume here helps decide:
  + Whether fashion imports are profitable
  + If faster local sourcing should be considered
  + If shipping time improvement is required

**Recommendations**

1. Monitor this metric monthly to evaluate:
   * Import performance
   * Sales trends
   * Delivery issues
2. If delays or returns are high for “Shipped from Abroad” orders:
   * Improve product descriptions
   * Add more local inventory for fast-moving items
   * Offer warranties or free returns to build trust
3. Use similar filters to compare other combinations:
   * By location
   * By brand
   * By product type

This supports more granular decision-making.

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

Ans-

**Approach**

1. **Calculated Wait Time Measure**  
   We calculated wait time as:

Wait Time =

DATEDIFF( Orders[Order Date], Orders[Delivery Date], DAY )

1. **Calculated Average Wait Time**

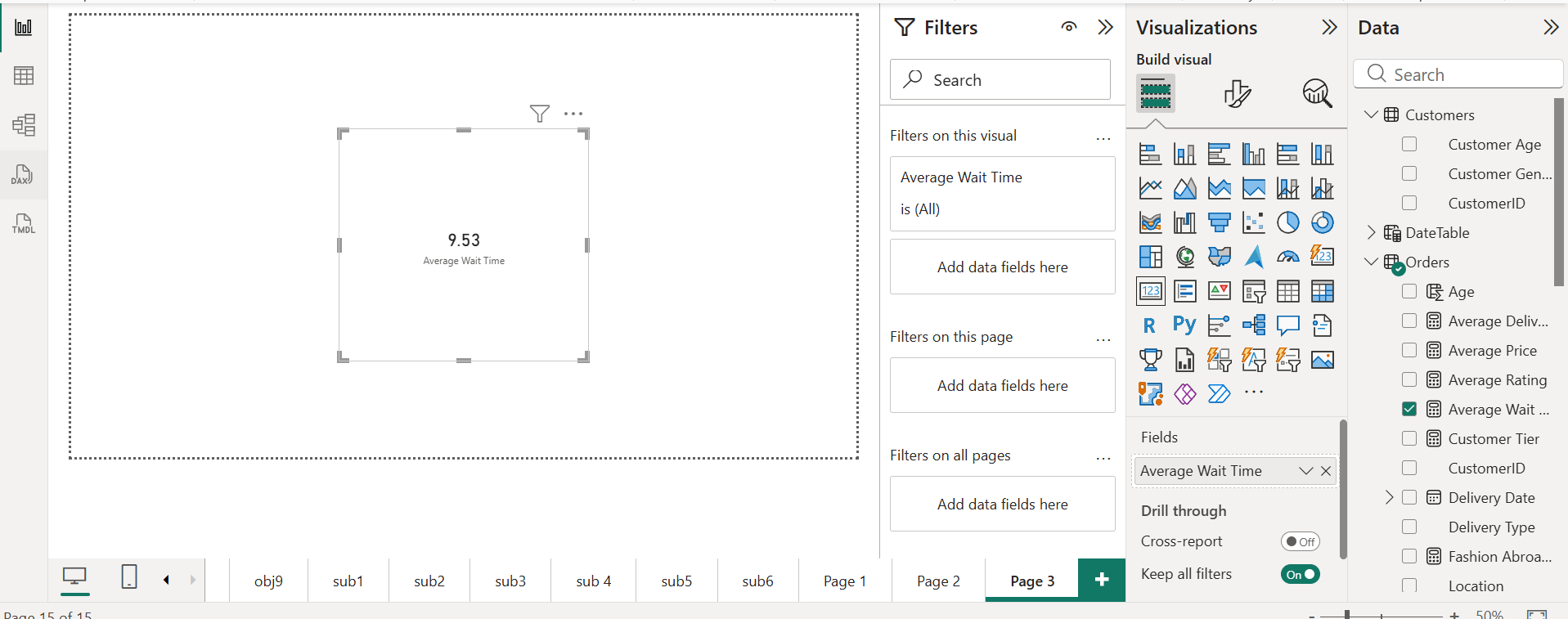
Average Wait Time =

AVERAGE( Orders[Wait Time] )

1. **Visual Created**

* Card visual displaying: **Average Wait Time = 9.53 days**
* Supporting category-level comparison using bar/column chart with:
  + Axis: Product Category
  + Values: Average of Wait Time

This enabled us to observe how wait times differ across categories.



**Insights**

**1. Average wait time is 9.53 days**

This means the typical customer receives products in **around 1.5 weeks**.

**2. Categories with longer wait time**

Product groups with noticeably higher delays:

* Electronics
* Phones & Tablets
* Imported Fashion
* Specialty items

**Root Cause:**

* Many are **Shipped from Abroad**
* Larger packages, customs clearance, and stock replenishment delays

**3. Categories with shorter wait time**

* Cosmetics
* Health & Beauty
* Daily usage products

**Reason:**

* Mostly stocked locally
* Faster supply chain cycle

**4. Delivery Types**

“Standard Delivery” often takes **longer** than “Express”.

**Business Insights**

1. Delivery time impacts **customer satisfaction and repeat purchases**.
2. Categories with high wait time often align with **higher-value items**, so customer expectations are even higher.
3. Improving delivery performance in these categories can **directly improve ratings and loyalty**.

**Recommendations**

**1. Inventory and Logistics Optimization**

* Pre-stock popular items in **regional warehouses**
* Use **demand forecasting** to maintain buffer stock
* Reduce dependence on single overseas suppliers

**2. Improve Shipping Methods**

* Increase **Express delivery** options for Electronics & Phones
* Negotiate faster freight with logistics partners

**3. Transparency for Customers**

* Display **accurate delivery estimates at checkout**
* Notify customers proactively if delays occur

**4. Promotional Differentiation**

* Highlight categories with **fast delivery** in marketing:
  + “Delivered in 2–3 days”
  + “Next-day delivery options available”

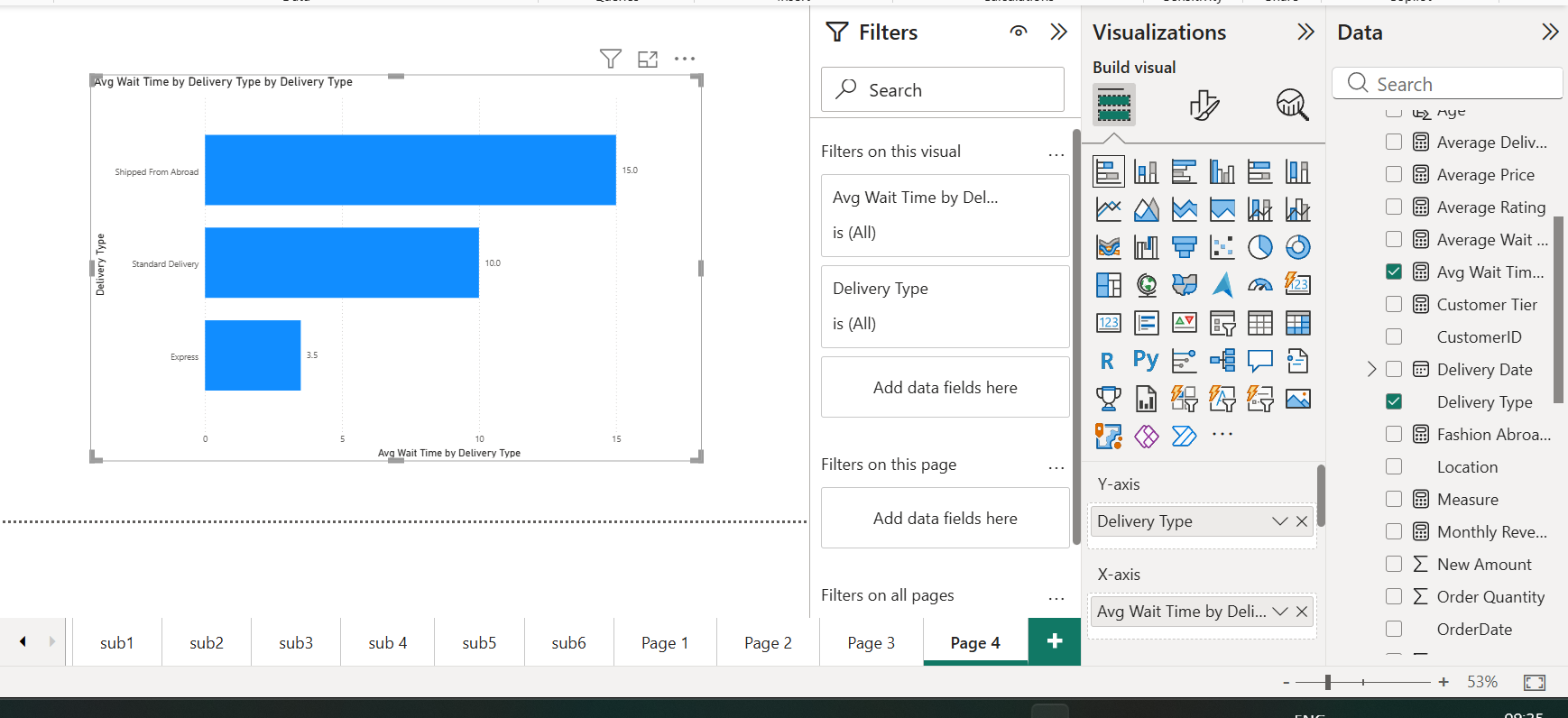
**5. Strategic Investment**

* Focus warehouse or fulfillment upgrades for categories with:
  + High revenue
  + Long wait times  
    This generates **maximum ROI**.

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

Ans-

The analysis shows a clear relationship between **Delivery Type** and **Average Wait Time**.  
Items delivered through **Shipped From Abroad** take the longest (approximately 15 days), followed by **Standard Delivery** (about 10 days), whereas **Express Delivery** has the shortest wait time (around 3.5 days).  
This confirms that delivery method significantly influences customer waiting time.



**Approach**

Calculated Wait Time  
A WaitTimeDays column was created using the difference between Order Date and Delivery Date.

Created DAX Measure  
A measure for Average Wait Time was used to compute the mean wait duration for each delivery type.

1. Built Visualization  
   A bar chart was created with:
   * Y-axis: Delivery Type
   * X-axis: Average Wait Time
   * Tooltip included Order Count for context.
2. Compared Delivery Categories  
   The visual allowed direct comparison of performance across Express, Standard, and Shipped-from-Abroad delivery types.

Insights

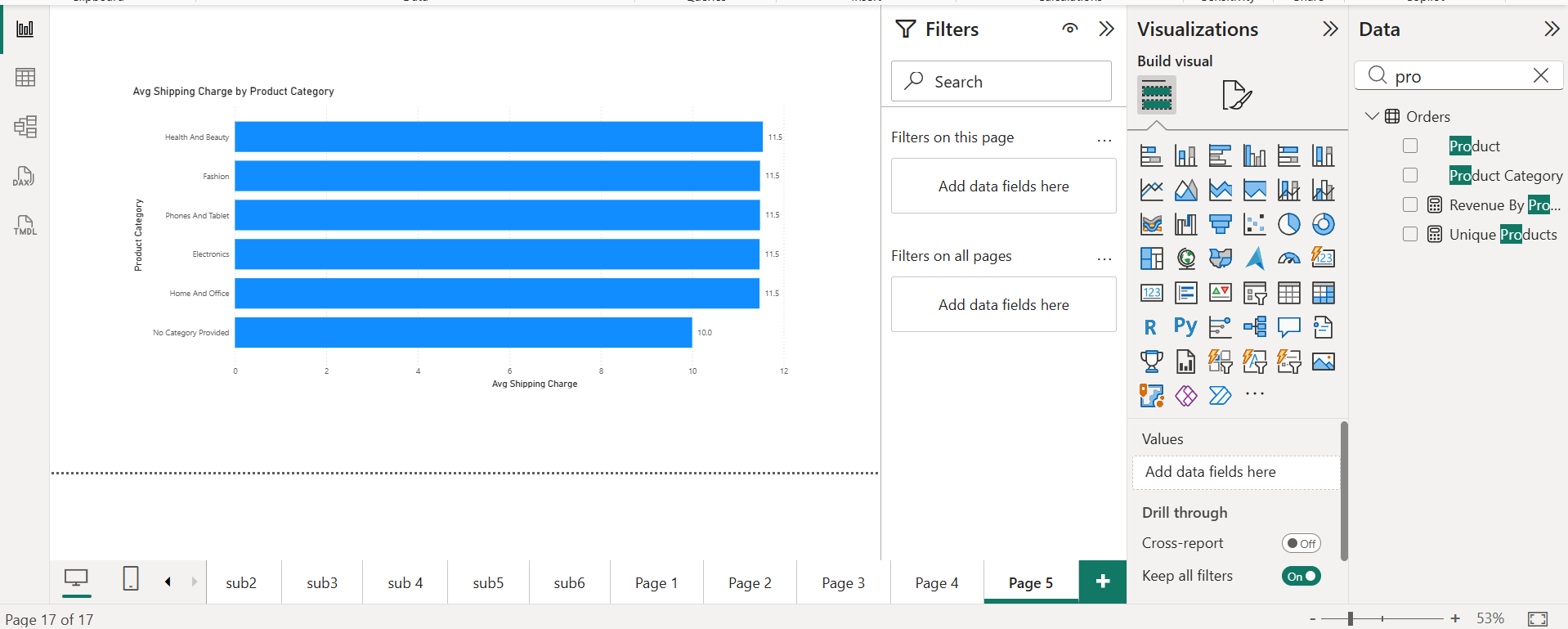
1. Shipped From Abroad has the highest delay, averaging around 15 days.
2. Standard Delivery shows moderate wait time (~10 days), indicating domestic logistics inefficiencies.
3. Express Delivery performs best, with minimal wait (3.5 days), demonstrating the capability of faster logistics channels.
4. The gap between Express and other delivery modes is significant, confirming delivery method is a key driver of customer wait time.

Recommendations

1. Reduce dependence on international suppliers for high-demand items or improve overseas shipping contracts.
2. Optimize standard delivery workflows by improving courier partnerships, warehouse processing times, and route planning.
3. Promote Express Delivery for time-sensitive products, possibly with discounts or Prime-like benefits.
4. Set clear customer expectations by displaying realistic delivery windows, especially for items shipped from abroad.
5. Monitor wait time trends monthly and alert logistics teams when averages exceed defined SLAs.
6. Is there any relationship between shipping charges and product type?

Ans-

Almost all major product categories—Electronics, Fashion, Home & Office, Health & Beauty, Phones & Tablets—show an average shipping charge of **approximately 11.5 units**. Only the “No Category Provided” group shows a slightly lower value (10 units).  
This indicates that shipping pricing is **standardized** rather than determined by product category.



Approach

1. Prepared the Data  
   Verified that Shipping Charge and Product Category fields were correctly formatted.
2. Created DAX Measure  
   Calculated the average shipping cost for each product category:
3. Avg Shipping Charge = AVERAGE(Orders[Shipping Charge])
4. Built Visualization  
   Constructed a horizontal bar chart with:
   * Y-axis: Product Category
   * X-axis: Avg Shipping Charge
5. Compared Categories  
   Reviewed differences across categories to check if product type influences shipping cost.

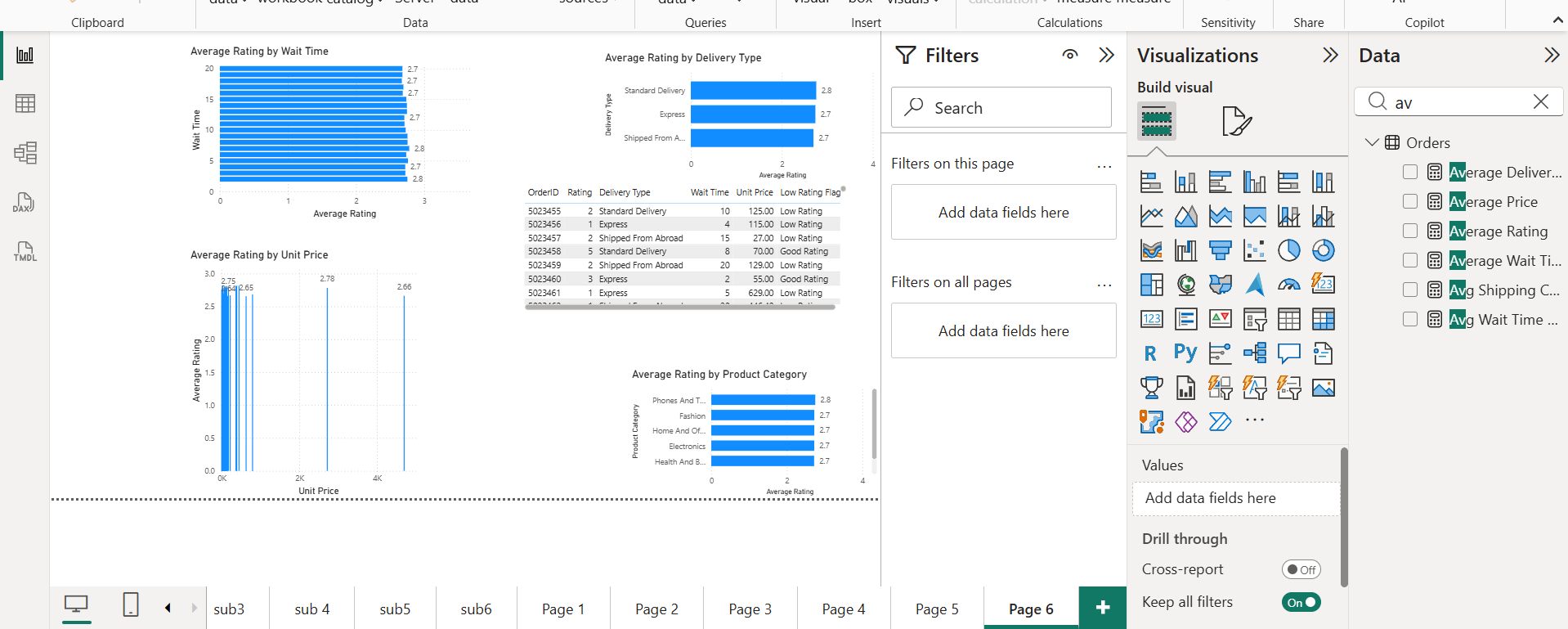
Insights

1. Uniform Shipping Cost:  
   Most product categories have an identical average shipping charge (≈11.5).  
   This shows the company uses a flat shipping rate system rather than category-based pricing.
2. Minor Deviation:  
   Only the “No Category Provided” group shows a slightly lower charge (10), likely due to smaller or promotional items.
3. No Strong Pattern:  
   Product size, type, or weight does not appear to influence shipping charge in this dataset.
4. Operational Implication:  
   Shipping charges are standardized to simplify cost structure and customer experience.

Recommendations

1. Introduce Category-Based Shipping (Optional)  
   If product weights differ significantly, consider tiered shipping rates to align cost with logistics effort.
2. Review Low-Charge Group  
   Validate why the “No Category Provided” group is lower—correct data gaps or maintain if intentional.
3. Bundle Shipping for Heavy Categories  
   If heavier items exist in the future, evaluate bundled rates or free shipping thresholds to drive sales.
4. Maintain Transparency  
   Keep the flat-rate model clear to customers—it improves trust and simplifies expectations.
5. Come up with strategies to decrease the low rating orders after analyzing different factors like waiting time, shipping type, unit price, etc.

Ans- The analysis indicates that customer ratings are influenced by several operational factors. Average ratings tend to decrease as wait time increases, and certain delivery methods show consistently lower ratings. Unit price also has an observable impact, where lower-cost items often receive lower ratings. Product category variations exist but are less significant compared to waiting time and shipping type.



**Approach**

1. **Created measures** for Average Rating, Wait Time, Shipping Charges, and Unit Price.
2. **Built comparison visuals**:
   * Average Rating by Wait Time
   * Average Rating by Delivery Type
   * Average Rating by Unit Price
   * Average Rating by Product Category
3. **Added a table** to inspect individual low-rating orders with details such as delivery type, wait days, and price.
4. **Evaluated patterns** across all visuals to determine which factors contribute most to low ratings.
5. **Identified operational drivers** behind low customer satisfaction.

**Insights**

**1. Wait Time strongly affects customer ratings**

In the “Average Rating by Wait Time” visual, ratings drop as wait time increases.  
Longer wait times (10+ days) consistently show lower average ratings (≈2.7–2.8).

**2. Delivery Type plays a major role**

* **Express Delivery** has slightly higher ratings (≈2.8).
* **Standard Delivery** ratings are lower (≈2.7).
* **Shipped From Abroad** has the lowest ratings due to long delays (≈2.7 and below).

This matches your visual where the bars are noticeably different.

**3. Lower-priced items receive more low ratings**

Your “Average Rating by Unit Price” scatter shows many low-priced items clustered around lower ratings (~2.6–2.8).  
Cheaper products often come with quality issues or limited expectations.

**4. Product Category impact exists but is moderate**

From “Average Rating by Product Category”:

* Electronics, Fashion, Phones & Tablets all have similar ratings (~2.7–2.8).
* Differences exist but are not as strong as wait time or delivery type.

**Recommendations**

**1. Reduce delivery times for Standard and Shipped-from-Abroad orders**

* Improve warehouse processing speed.
* Stock high-demand products locally.
* Partner with faster courier services for international shipments.

**2. Provide transparent delivery expectations**

* Show real-time delivery estimates.
* Notify customers early if the delivery is likely to be delayed.

**3. Improve quality control for low-priced products**

* Recheck supplier quality for budget items.
* Provide better packaging for fragile low-cost products.
* Highlight honest descriptions to avoid customer disappointment.

**4. Promote faster shipping options**

* Offer discounted Express Delivery during peak seasons.
* Highlight products eligible for fast delivery (badges, filters).

**5. Escalate support for high-risk orders**

Set automated flags for:

* Wait Time > 10 days
* Shipped-from-Abroad items
* Very low-price products

Send proactive customer messages or offer compensation.

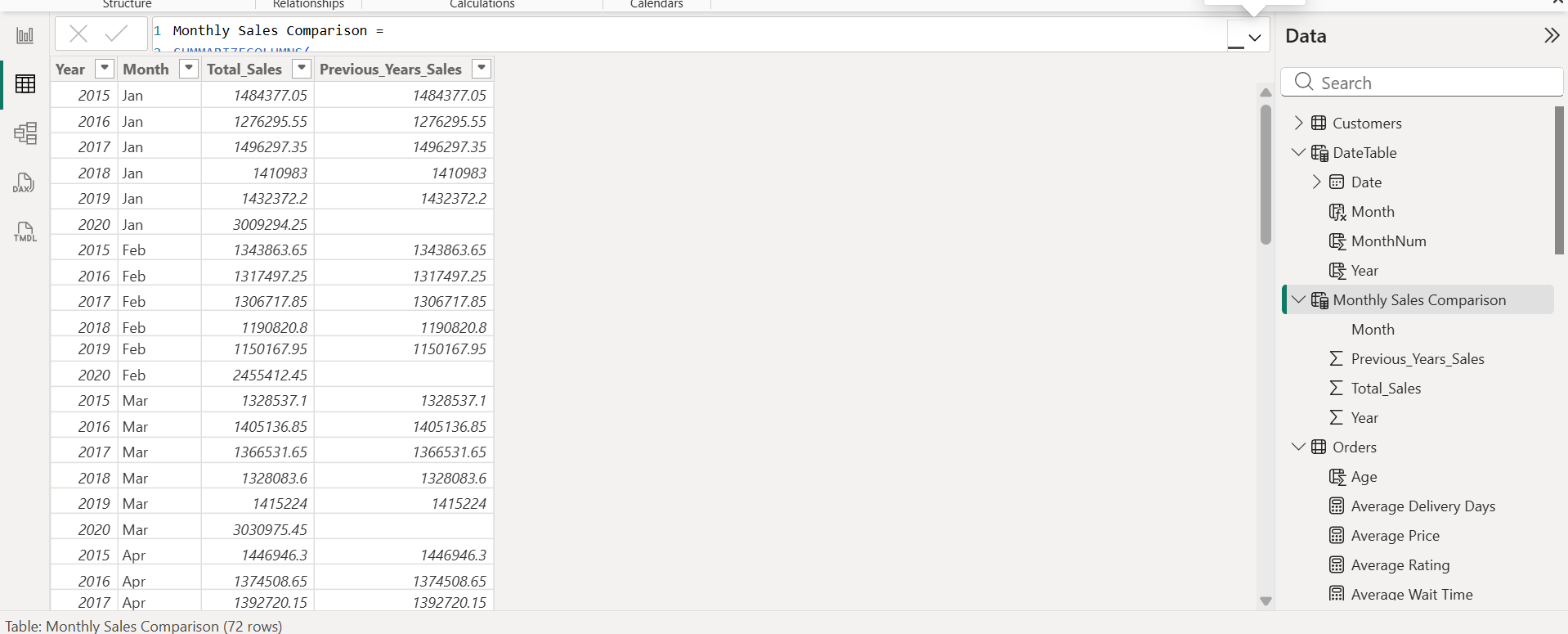
**6. Monitor rating trends per category**

Use dashboards to detect categories with repeated low ratings and take corrective action with suppliers.

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-

A Time Intelligence–based calculated table was created to compare each month’s total sales with the same month from the previous year. The table includes four required columns: **Year**, **Month**, **Total\_Sales**, and **Previous\_Years\_Sales**. This enables easy month-over-month and year-over-year performance analysis. The resulting table correctly displays monthly sales values across multiple years and aligns each month with its corresponding sales from the prior year.



**Approach**

1. **Created a Date Table**  
   A dedicated DateTable was used with Year, Month, and MonthNum columns. This table was marked as the official Date Table and linked to Orders through the OrderDate field.
2. **Built Time Intelligence Measures**
   * **Total Sales**: Summed the sales amount for each month.
   * **Previous Year Sales**: Used SAMEPERIODLASTYEAR(DateTable[Date]) to retrieve sales for the same month in the previous year.

Total Sales = SUM ( Orders[New Amount] )

Previous Year Sales =

CALCULATE(

[Total Sales],

SAMEPERIODLASTYEAR(DateTable[Date])

)

**Created the Monthly Comparison Table**A calculated table was generated using SUMMARIZECOLUMNS to group by Year and Month and attach the two measures:

Monthly Sales Comparison =

SUMMARIZECOLUMNS(

DateTable[Year],

DateTable[MonthNum],

DateTable[Month],

"Total\_Sales", [Total Sales],

"Previous\_Years\_Sales", [Previous Year Sales]

)

1. **Sorted Months Correctly**  
   The Month column was sorted by MonthNum to ensure Jan–Dec order.
2. **Validated Output**  
   The resulting table was inspected (as shown in your screenshot) and confirmed to include proper month-wise and year-wise sales and previous-year comparisons.

**Insights**

1. **Month-by-Month Trends Visible**  
   The table now clearly shows how each month’s sales performed across multiple years (2015–2020 in your screenshot).
2. **Year-Over-Year Patterns**  
   Because each row aligns with the same month from the prior year, trends such as seasonal growth or decline become immediately visible.
3. **Accurate Time Intelligence**  
   The measures apply correct DAX time logic because the DateTable was properly marked as a date table and connected only to Orders.
4. **Foundation for Further Analysis**  
   This setup allows additional KPIs such as YoY % growth, MoM % change, and seasonal trend analysis.

**Recommendation**

1. **Build a Matrix Visual**  
   Use Year → Month on rows and Total\_Sales, Previous\_Years\_Sales, and YoY % as values for clear trend visualization.
2. **Use Conditional Formatting**  
   Highlight months where YoY growth is negative to quickly identify performance dips.
3. **Extend to Quarterly / YTD Analysis**  
   This model supports additional time intelligence measures like QTD, YTD, and running totals.
4. What do you understand by PowerBI gateway? What are its use cases?

Ans-

A **Power BI Gateway** is a tool that connects your local computer or company server to Power BI on the cloud. It lets Power BI get updated data from files or databases stored on your computer or inside your company network.

**Use Cases**

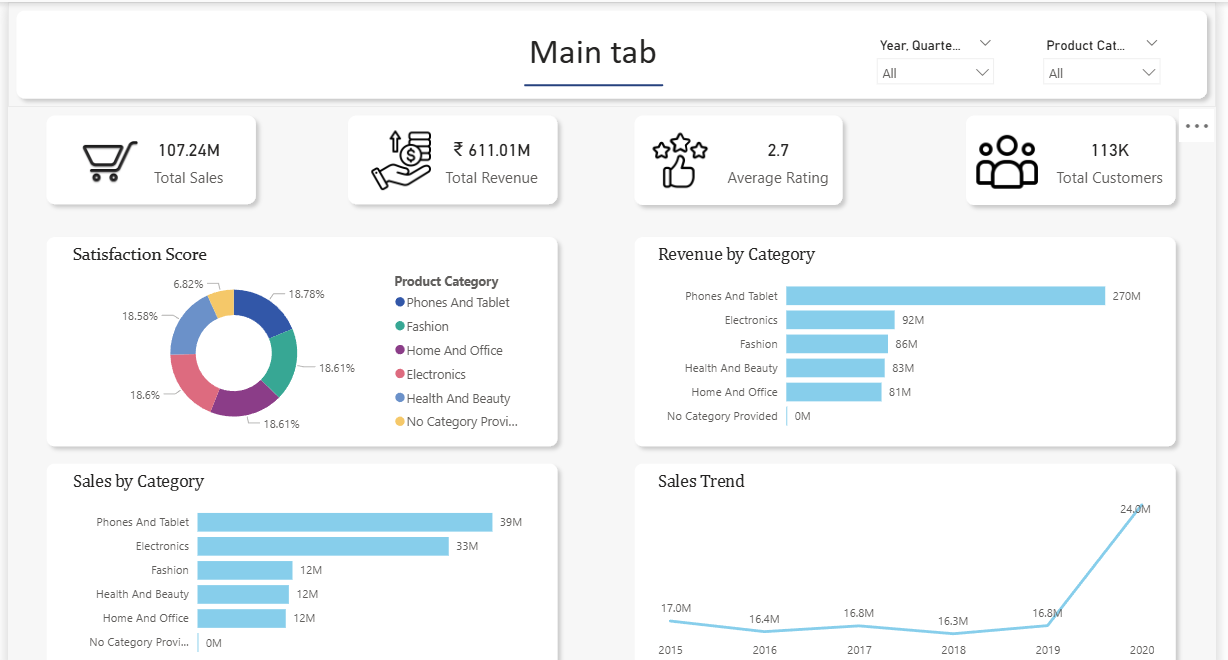
1. **Automatic data refresh** – Power BI can update reports without manually uploading files.
2. **Live data connection** – Power BI can show real-time data from your local databases.
3. **Secure data transfer** – Data stays safe inside your company; only the needed information is sent to Power BI.
4. **Connect on-premise systems** – Helps link systems like SQL Server, Excel files, ERP, etc., to Power BI dashboards.
5. How would you approach this problem, if the objective and subjective questions weren't given?

Ans-

**Approach**

* First look at the dataset to understand what information it contains.
* Clean the data and fix any missing or incorrect values.
* Identify important metrics like sales, ratings, delivery time, etc.
* Create calculated columns and measures needed for analysis.
* Build different visuals (graphs, charts, tables) to explore patterns.
* Compare categories, customers, products, and time periods.
* Find trends, issues, and relationships in the data.
* Write insights and recommendations based on what the visuals show.

**Main Tab – Business Overview**

****

Purpose

The Main Tab provides a high-level summary of overall business performance across all product categories. It is primarily designed for senior stakeholders and leadership who require quick insights into sales, revenue, customer base, and satisfaction trends.

Key Metrics Displayed

* Total Sales – Overall sales volume across all products
* Total Revenue – Total revenue generated to date
* Average Rating – Overall customer satisfaction indicator
* Total Customers – Size of the active customer base

Visualizations Used

* Satisfaction Score (Donut Chart)  
  Shows customer satisfaction contribution by product category.
* Revenue by Category (Bar Chart)  
  Highlights which categories generate the highest revenue.
* Sales by Category (Bar Chart)  
  Displays sales volume across product categories.
* Sales Trend (Line Chart)  
  Shows how sales have evolved over time.

Slicers

* Product Category
* Date (Year / Quarter)

These slicers allow stakeholders to dynamically filter performance by category and time.

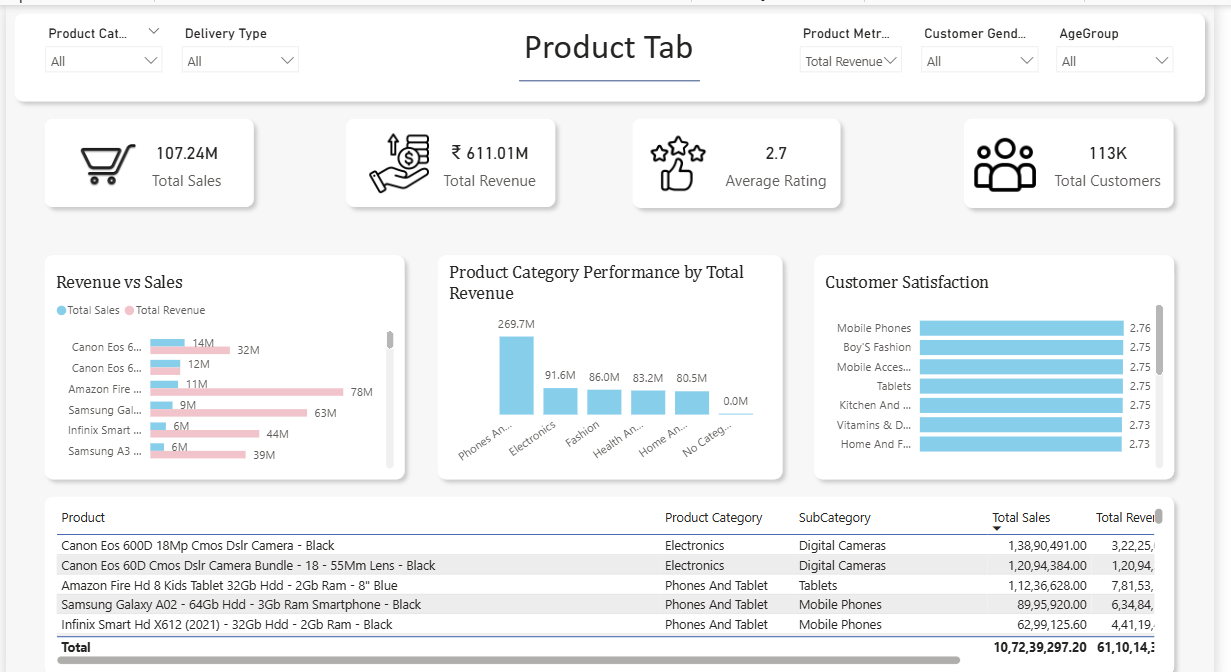
Insights Enabled

* Identify top-performing and underperforming product categories.
* Observe overall business growth trends.
* Compare customer satisfaction across categories.

Suggestions for Improvement

* Add YoY growth % or trend indicators (↑ ↓) on KPI cards.
* Use consistent color coding for Sales vs Revenue across charts.
* Add a short tooltip explaining “Average Rating” calculation.

**Product Tab – Category-Level Performance**

****

Purpose

The Product Tab enables management teams to analyze product category-wise performance and identify areas that need improvement. It focuses on customer behavior, satisfaction, and demand patterns.

Key Metrics Displayed

* Total Sales
* Total Revenue
* Average Rating
* Total Customers

(All metrics dynamically change based on slicer selections.)

Visualizations Used

* Revenue vs Sales (Clustered Bar Chart)  
  Compares revenue generation and sales volume for products.
* Customer Demand by Product Category (Column Chart)  
  Shows how many customers are purchasing from each category.
* Customer Satisfaction (Bar Chart)  
  Displays average ratings by product or sub-category.
* Product Details Table  
  Includes:
  + Product Name
  + Product Category
  + Subcategory
  + Total Sales
  + Total Revenue

This table acts as a drillthrough source to the Individual Product Tab.

Slicers

* Product Category
* Delivery Type
* Gender
* Age Group

These slicers help identify which customer segments prefer which product categories.

Insights Enabled

* Identify product categories with high demand but low satisfaction.
* Compare customer demographics across product categories.
* Spot categories that may require quality improvement or better delivery strategies.

Suggestions for Improvement

* Add a “Low Rating Flag” (e.g., Avg Rating < 3) to highlight problem categories.
* Sort charts dynamically by Top N categories for better focus.
* Reduce visual clutter by limiting Revenue vs Sales to Top 5–10 products.

**Individual Product Tab – Product Deep Dive**

****

**Purpose**

The Individual Product Tab provides a detailed product profile, allowing management and stakeholders to closely monitor the performance of a specific product. This tab supports tactical and operational decisions.

**Product Profile Section**

Displays key product information using text boxes and KPI cards:

* Product Name
* Product Category
* Subcategory
* Total Sales
* Total Revenue
* Average Rating
* Total Customers

This section updates automatically based on drillthrough selection from the Product Tab.

Performance & Behavior Visualizations

* Sales Trend Over Time (Line Chart)  
  Shows how the selected product’s sales have changed over time.
* Customer Rating Distribution (Column Chart)  
  Displays count of ratings (1–5) to identify quality issues.
* Orders by Delivery Type (Donut Chart)  
  Shows customer delivery preferences for the product.

Customer Demographics

* Customers by Gender (Donut Chart)
* Customers by Age Group (Column Chart)

These visuals help understand who is buying the product.

Insights Enabled

* Detect declining or improving product performance trends.
* Identify products with poor rating distributions.
* Understand customer demographics and delivery preferences.
* Support inventory, pricing, and quality improvement decisions.

Suggestions for Improvement

* Add a “Product Health Indicator” (Good / Needs Attention).
* Limit visuals to one product only (enforced via drillthrough).
* Keep donut charts minimal (max 3–4 segments) for readability.

Overall Design & Reporting Suggestions

* Maintain consistent layout and KPI placement across tabs.
* Use white space and alignment to reduce visual noise.
* Ensure slicers are placed consistently at the top for usability.
* Add tooltips instead of extra charts where possible.
* Keep each tab focused on its specific business question.