ANSWERS TO OBJECTIVE QUESTIONS:

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

By the use of Power Query Editor, we clean the data for better analysis. First, we transform the data into power query editor and start cleaning our data.

In data cleaning, I make sure that all the columns do not have any empty rows. After removing the empty rows, I changed the data type of order date and delivery date column to “Date” type.

In the reason column, there were some empty rows without any reason for return of the product and for delivered products. So, I replaced those rows with “Unknown reason” for empty rows which have returned products and “Not Applicable” reason for empty rows whose products were delivered.

By looking at the data set, there were 5 extra columns which are not necessary, so I removed them from the data set.

* **Total Revenue Generated by all the Sales.**

To calculate the Total Revenue, I have created a New Measure to sum the Sale Price Column. The DAX formula is **Total Revenue = SUM(Data[Sale Price])** which gives the answer Total Revenue of **107.23M**.



* **Total number of Unique Customers who made purchases in each year and increase in the numbers over the years.**

Total unique customers are **0.11M** who made purchases. The approach is by creating a visual card and counting the Customer ID to get the values.



In the below Bar chart, we can see how many customers made purchases each year. In x-axis Count of Customer ID and in y-axis Year. As we can see in the chart, there is an increase trend over the years.



* **Total number of Unique Products available in the company.**

There are a total of **44** unique products available in the company. To show this value, I used card visual and by the Count of product column in the field we can determine the total number of unique products.



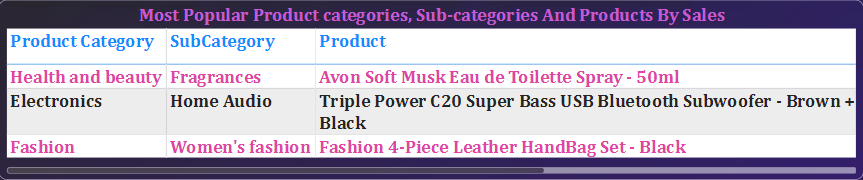
* **Average Delivery Time for products to be delivered**

The average delivery time for a product to be delivered for “Delivered Products” is 9.41 days. The DAX Functions which I used are **Delivered Orders = CALCULATE(SUM(Data[Delivery Time]), Data[Status] = "Delivered") and Average Delivery Time for Delivered Orders = AVERAGEX(Data, [Delivered Orders]).**

By using the card visual I presented the Average Delivery Time for Delivered orders inserting calculated column Average Delivery Time for Delivered Orders.

****

* **Most popular Product categories, Sub Categories and Products name.**

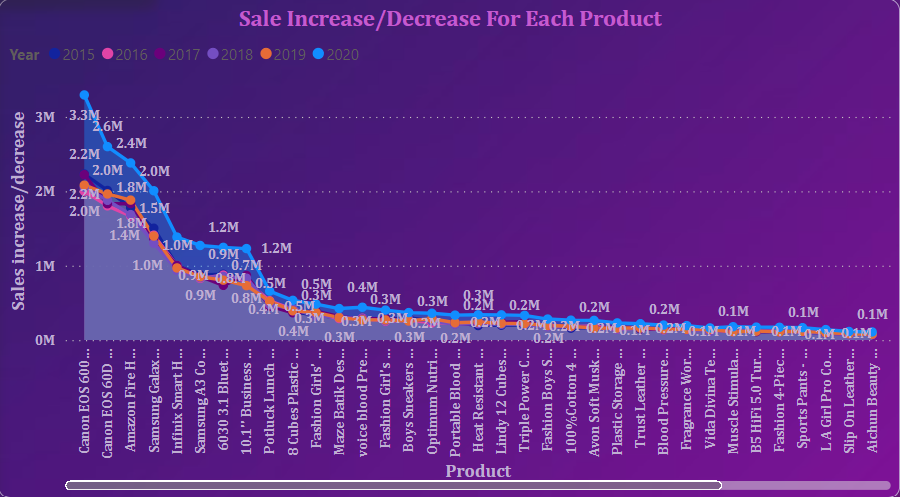
****

In the above visual, we can see the most popular Products, Product Categories and Sub Categories by Highest Sales. By using the Table chart and Filter Pane obtained the results. In the Filter pane, I filtered the product with Top 3 in Sum of Order Quantity.

* **Products that have seen an increase or decrease in sales over the year.**

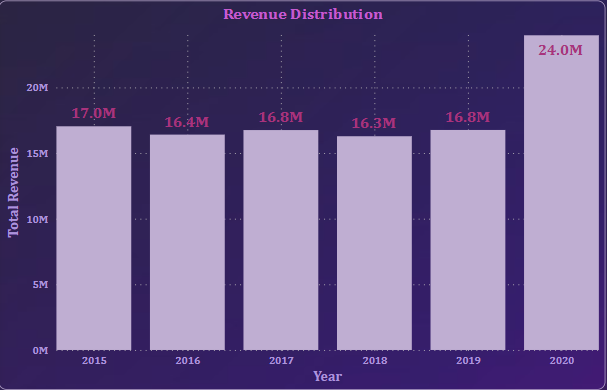
To calculate the Sales increase or decrease by year, first I calculated Previous year sales with the DAX formula **Previous Year Sales = VAR Previous = CALCULATE(SUM(Data[Sales]),PREVIOUSYEAR(Data[OrderDate].[Date]))RETURN IF(ISBLANK(Previous),0,Previous)** and Calculated the Sales Increase/Decrease with **Sales increase/decrease = [Total Sales] - [Previous Year Sales].**

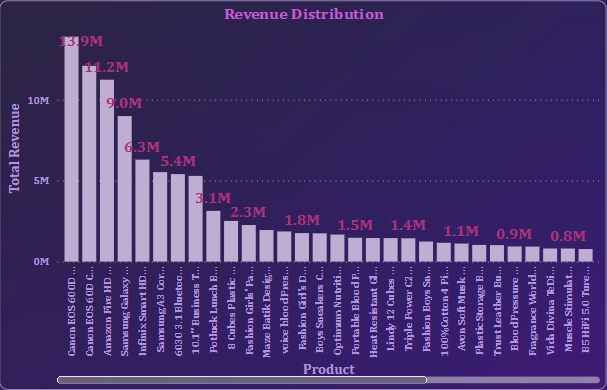
I have used Area chart to show Sales Increase/Decrease for each product over the period of 6 years.

****

ANSWERS TO SUBJECTIVE QUESTIONS:

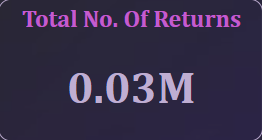
* **How does revenue break down by year and by-product? Evaluate how different products contribute to annual revenue and come up with suggestions to increase the sales of the low-selling items.**





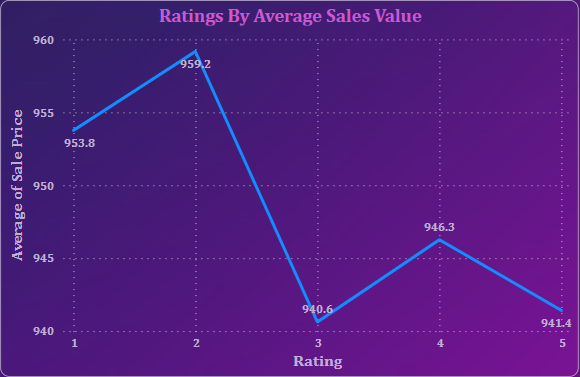
* In the first column chart, we can see how revenue break down by each year. In the second column chart, it is noticeable how different products contribute to annual.
* To increase the sales of low-selling items we can use some strategies like Selling in a bundle of low-priced items to increase their revenue, marketing promotions for low selling items, giving more points on the purchase of that product etc.
* **How many products were returned? Examine the possible reasons for returns and consider how this metric could inform improvements in product descriptions or quality control.**

The total number of returns is 0.03M with 44 unique products.

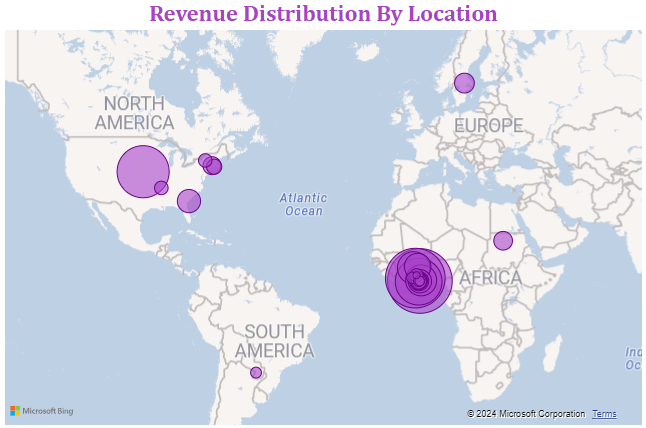


The possible reasons for the returns are:

* Defective items - Delivering the defective items makes customers lose interest in the quality of the brand. So, make sure before going for a package the quality of the product needs to be checked.
* Delivery Missing item/part - Product description should mention clearly all the components and should provide a checklist of what's included.
* Onsite Description Mismatch - Product description should match with the products that are going to deliver otherwise it leads to customer dissatisfaction. Make sure to regularly update the description in the online platform.
* Product not fitting expectations - Make sure the size chart provided in the product description shows multiple options for customers to choose perfect size.
* Delivery wrong item - Have multiple checkpoints to see if the right product is going for a package before delivering the product.
* Unknown Reasons - There are some unknown reasons for the product to be returned. Make sure while collecting the return packages take note of the reasons for return for better understanding.
* **Whenever a customer goes to Amazon, they’ll filter the most rated products in order to buy the better category. Can you verify this using any visualization or table that the ratings of products impact their sales value?**



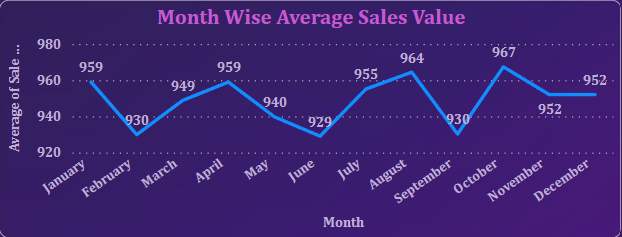
* From the given data, it is evident that products with lower ratings have higher sales. This could be for several reasons like immediate need of the product, low price of the product, availability of the specific product in the platform, specific brand etc.
* **Investigate how revenue distribution varies 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 approach to market segmentation and resource allocation?**

****

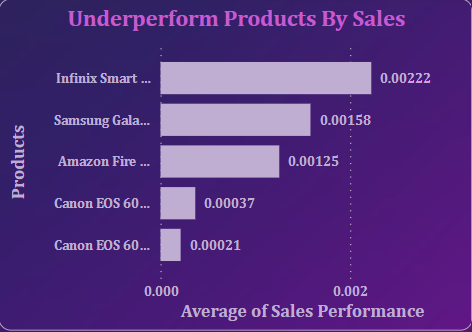
From the above Map chart, we can see the revenue distribution across different locations. As we can see North America and African regions contribute most to sales, so tailor the marketing strategies accordingly. For example, North American regions, based on the regional demands adjust the product categories and price to meet the demands. As for distribution, collaborate with regional delivery partners to accommodate smooth delivery. Regional marketing should involve the traditions of that particular region so that it affects the sentiment of the customers and it attracts more customers.

Location based trends help the company to identify key locations where they can focus more on marketing promotions as they can generate more revenue compared to other locations. Company can also increase distribution channels and establish more warehouses for fast delivery of the products.

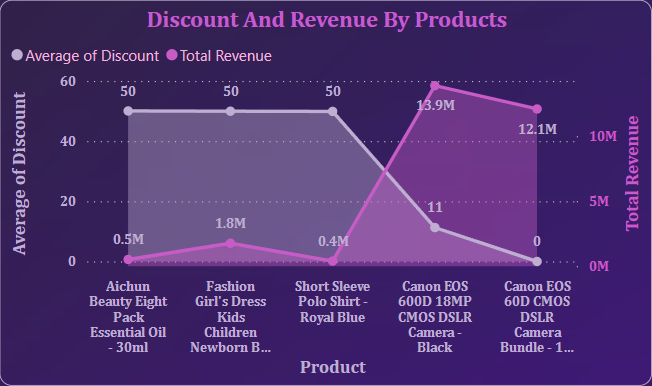
* **Determine which month could benefit from enhanced promotional offers to boost sales. Can you suggest some targeted marketing strategies here?**



* From the above Line chart, we can see the Month wise sales in the company. As we observe the chart February, June and September sales are slightly less compared to other months. From enhanced promotional offers these months could benefit to boost sales.
* Targeted marketing strategies like Time limited sales, special discounts and incentives on that particular month could increase the sales. Go for more advertising on that particular month with the special deals.
* **Identify which products may require increased marketing efforts. Which items have high prices yet underperform in sales?**

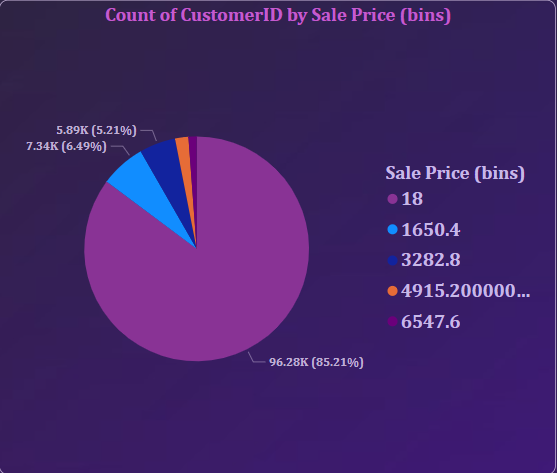
****

* From the above Bar chart, these are the 5 underperformed products which may require increased marketing efforts.
* First I calculated sales performance by using DAX formula “DIVIDE” i.e. **Sales Performance = DIVIDE(Data[Order Quantity], Data[Sale Price])**.
* With the average of sales performance against products by using Bar chart and using filters for lowest 5 products I got the results.
* Despite having high price, the sales of these products are very low.
* **Assess which products should have discounts. How can targeted incentives drive sales and customer loyalty for specific products?**

****

As we see in the above Area chart, these are the products with lowest discounts. If we have more discounts for these products, we can drive more sales which will result in more revenue. I calculated discount for the products using **Discount = Data[Cost Price] - Data[Sale Price].** With average of discount in y- axis, total revenue in secondary axis and products in x-axis, we can see the

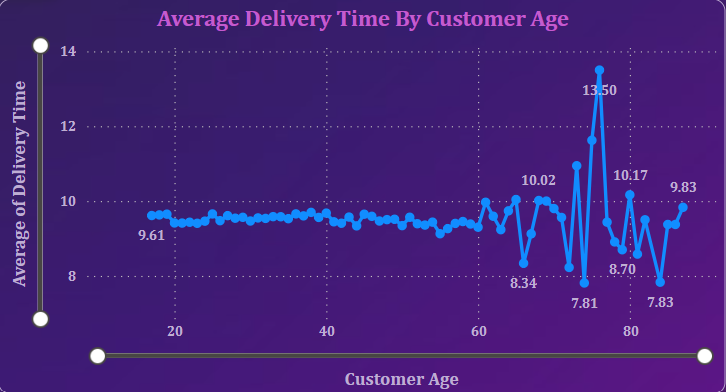
* **Come up with a loyalty program to benefit the company's customers. From the available lot of customers come up with the strategies to bucket them and provide benefits under different loyalty programs.**

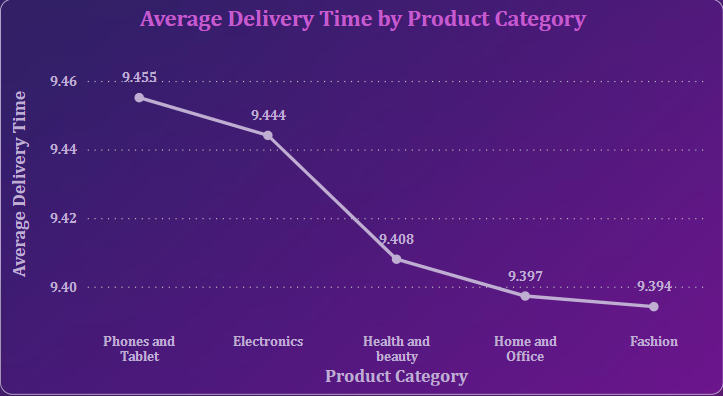


I have grouped the customers into 5 categories based on the spending capacity.

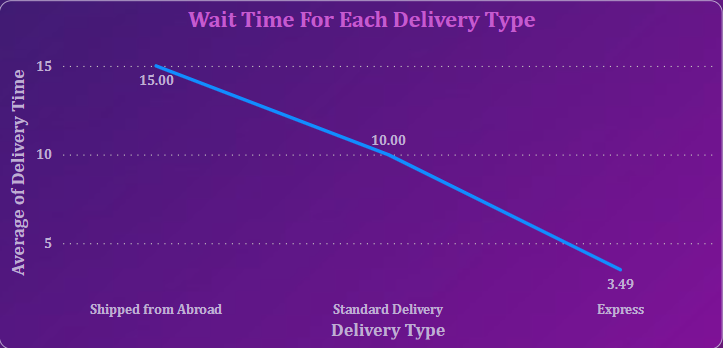
1. VIP Program for High End Customers - These customers are the highest spenders with minimum sale price of 6547. We can provide premium services tor these customers like early access to big sales.
2. Reward Points Premium Customers - Premium customers minimum spending amount is 4915. Giving points on every spend will be a reward for their loyalty.
3. Special Discounts for Standard Customers – Standard customers usual spend minimum of 3282. Giving rewards on special occasions like birthdays will make the customers happy.
4. Card discounts for Economy Customers – Customers with minimum spending amount of 1650.4 are the economy customers. We can provide Card discounts for them if they pay money through their debit or credit cards.
5. Cashback and High Discounts for Budget Friendly Customers – These customers mainly focus on their budget. If the sale price is less then only, they will purchase the products. So, giving cashback and high discounts will attract these customers.

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

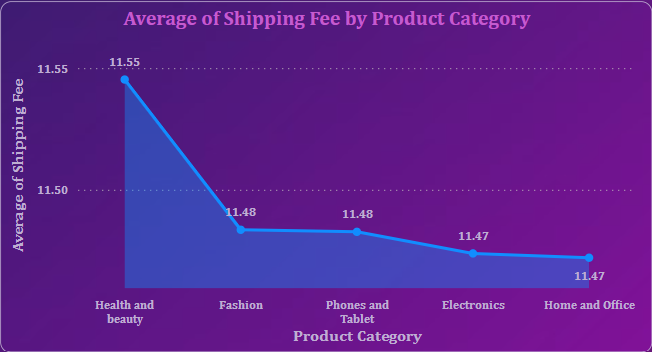




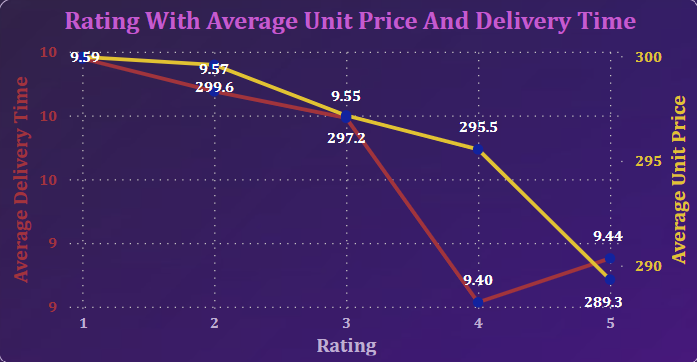
* Here Wait Time means Delivery Time from my dataset. In the first Line chart, we can see how wait time correlate with demographic i.e. Customer age. In the graph, above 60 years age group the average delivery time varies from 8.34 days to 13.50 days. Below 60 years there is no much variation in the average delivery time.
* In the second Line chart, we can see how average wait time vary across different product categories. For “Phones and Tablet” the average delivery time is 9.455 which is more compared to other product categories. We can do the scheduling and staffing accordingly.
* **Explore if there is any relationship between the Delivery type and waiting time between ordering and receiving an item.**



* As we can see in the above Line chart, there are mainly 3 different delivery types. Each delivery type has a different delivery time. Express delivery has an average of 3.49 days. Standard delivery type has an average of 10 days to deliver the product. Products from abroad will take more days to deliver.
* **Is there any relationship between shipping charges and product type?**



* By using Area chart with Product Category in X-axis and Average of Shipping fee in Y-axis, we can see the relationship between Shipping fee and product category. As we can see in the chart, Health and Beauty have high average shipping fee.
* **Come up with strategies to decrease the low rating orders after analyzing different factors like waiting time, shipping type, unit price, etc.**



* To come up with different strategies to decrease the low rating orders, I have analyzed the waiting time, unit price for each rating.
* Here I used Line chart to display the waiting time and unit price for each rating.
* In the above chart, we can see for 5 rating the average delivery time and average unit price is low compared to other ratings.
* Strategies to decrease low rating orders:
* We can decrease the low rating orders by decreasing the delivery time. This can be done by offering them different delivery options like Express delivery, Standard delivery etc.
* By lowering the unit price, low rating orders can also be reduced.