



# **DATASPARK: ILLUMINATING INSIGHTS FOR GLOBAL ELECTRONICS**

**EDA REPORT**

K Geethika | [kgeethikaprasad@gmail.com](mailto:kgeethikaprasad@gmail.com) | 12th Aug'24



# SUMMARY

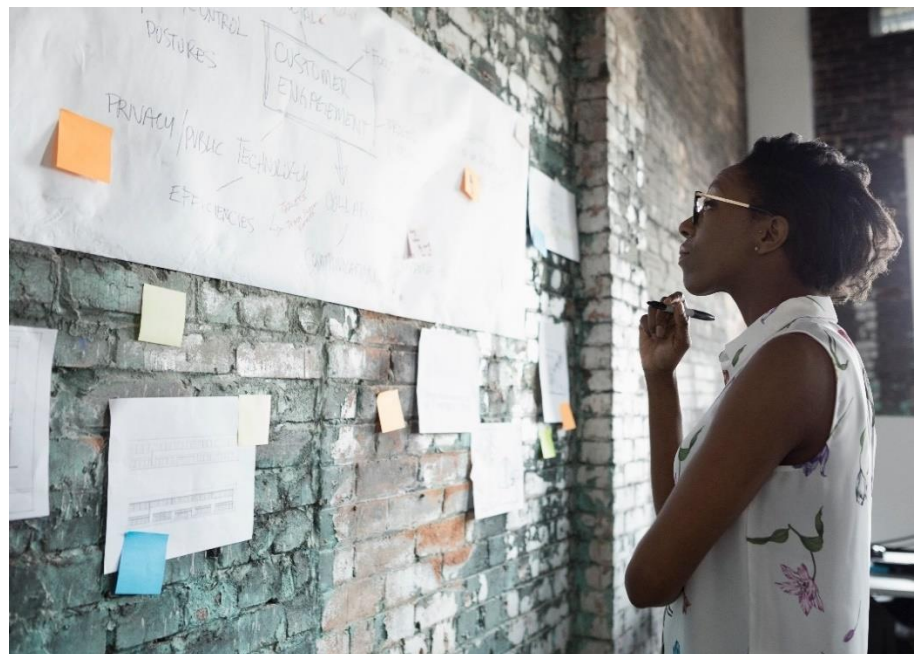
The EDA for Global Electronics provides a detailed exploration of the company's data, highlighting key areas such as customer demographics, purchasing behaviours, product performance, store operations, and the impact of currency fluctuations on sales. This analysis aims to deliver actionable insights to enhance marketing strategies, optimise inventory management, improve sales forecasting, guide product development, and inform store expansion and operational efficiency decisions.





# INTRODUCTION

This report provides an in-depth Exploratory Data Analysis (EDA) of Global Electronics' datasets, aiming to deliver valuable insights into customer demographics, purchasing behaviors, product performance, store operations, and the impact of currency fluctuations on sales. The insights generated from this analysis will inform actionable recommendations to enhance marketing strategies, optimize inventory management, improve sales forecasting, guide product development, and support store expansion and operational decisions. The ultimate goal is to increase customer satisfaction, maximize revenue, and drive business growth.





## 1. Data Overview and Defining Goals

The datasets given were observed to get an understanding of what is given and what need to be achieved. The business objectives were defined.

### 1.1 Data Sources

- Customer Details : Data includes information on customer key, gender, location.
- Sales Data: Covers transactions across all stores, including product details, quantities sold.
- Store Operations: Data on store sizes, locations.
- Exchange Rates: Historical exchange rate data for various currencies impacting sales.
- Product Details: Data Includes Product key, names, Categories, Sub-Categories.

### 1.2 Objectives:

- To enhance marketing Campaigns
- To develop better products
- To optimise inventory
- To plan effective promotions and store expansions
- To ultimately maximize revenues through increased customers' satisfaction and growth

## 2. Data Integration and Cleaning

The datasets were cleaned and integrated to ensure consistency and accuracy in the analysis.

Key variables include:

- **Customer Demographics:** Age, Gender, Location.
- **Purchasing Behaviors:** Frequency of purchases, Order values, Preferred categories.
- **Product Performance:** Sales volumes, Profit margins by product category.
- **Store Operations:** Sales by store location, Online vs. Offline sales.
- **Currency Impact:** Monthly sales trends, Currency fluctuation data.

### 2.1 Data Cleaning

- Missing Values: Addressed by imputing median values or removing incomplete records where necessary.
- Data Consistency: Ensured consistency across datasets by standardizing date formats, currency conversions, and product categorizations.
- Outlier Detection: Identified and reviewed outliers to determine their validity, especially in sales and customer demographics.

### 2.2 Data Intergration

- Joined the necessary datasets using join statements using common columns like Customer Key, StoreKey, Product Key, Exchange Rate

**Tools Used for this Process:** Python (Colab and VSCode) and SQL (TiDB Cloud)



### 3. Descriptive Analysis

To have a sense of the data given, Descriptive Analysis was performed on the integrated data set. The dataset contains 62,884 rows and 29 columns. Below is a summary of the data, including its structure, basic statistics, and the first few rows for context.

#### 3.1. Dataset Structure:

- Total Entries: 62,884
- Columns: 29
- Column Types:
- Integer (int64): 9 columns
- Float (float64): 4 columns
- Object (string): 16 columns

#### 3.2. Key Columns:

- OrderNo: Unique identifier for each order.
- LineItem: Represents the line item number within each order.
- Odate: Order date.
- DDate: Delivery date.
- CustKey: Customer identifier.
- StoreKey: Store identifier.
- ProdKey: Product identifier.
- Quant: Quantity ordered.
- CurCode: Currency code.
- Gender: Gender of the customer.
- CustName, City, State, Country, Continent: Customer details.
- ProdName, Brand, Color: Product details.
- UnitCostUSD, UnitPriceUSD: Cost and price of the product in USD.
- SubCatKey, Subcategory, CatKey, Category: Product categorization.
- SqMeters: Square meters (possibly related to store size).
- OpenDate: Store opening date.
- ExchangeRate: Exchange rate at the time of the order.

#### 3.3. Basic Statistics:

- OrderNo, CustKey, StoreKey, ProdKey: These columns show a wide range of values, indicating a diverse dataset.
- Quant: Quantity ordered varies from 1 to 8, with a mean of about 4.47.
- UnitCostUSD and UnitPriceUSD:  
Mean Unit Cost: \$1595.55  
Mean Unit Price: \$3706.52



### 3.4. Sample Data:

- The first few rows show a variety of orders, with details about the customer, product, and sales metrics. For example:
- Order 366000, line item 1 was for a "White" product in the "Cameras & Camcorders Accessories" subcategory, costing \$31.27 and priced at \$68.00.

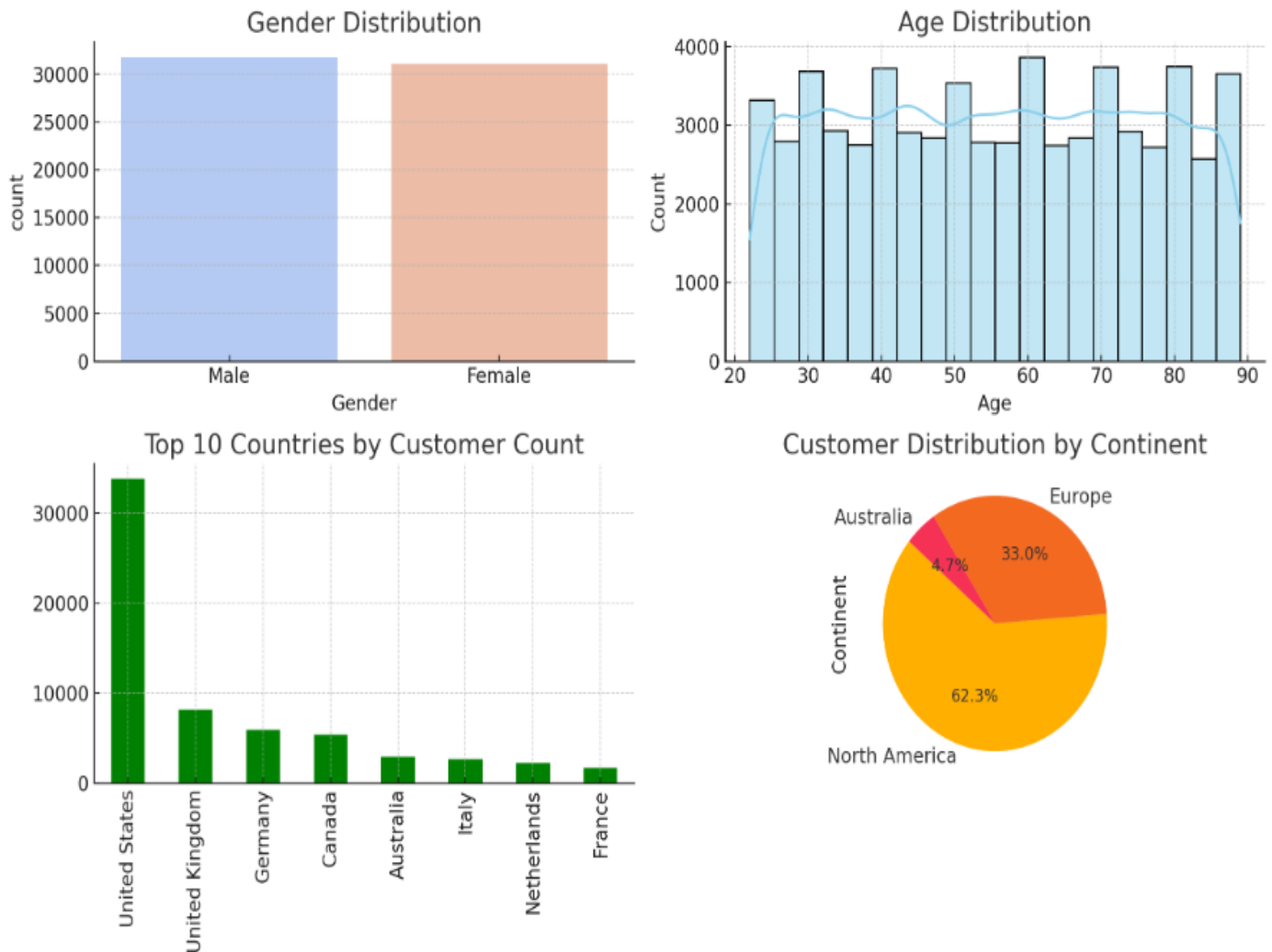


Figure 1: Descriptive Analysis

### 3.5. Tools Used:

- Visualization – PowerBI
- SQL – TiDB Cloud



## 4. Exploratory Data Analysis (EDA)

### 4.1 Customer Demographics Analysis

**Key Insights:**

- Age Distribution: Baby Boomers and Gen X customers are the largest contributors to revenue, with Baby Boomers leading, making them a critical focus for targeted marketing strategies. Gen Z customers, though currently a smaller segment, represent a growth opportunity.
- Gender Distribution: Sales are fairly evenly split between male and female customers, with a slight skew towards male customers. Gender-specific marketing campaigns could enhance engagement.
- Geographic Distribution: The United States and Canada are the primary markets, contributing the most significant share of revenue, while European markets show potential for growth with tailored marketing and operational strategies.

**Visualizations:**

- Bar charts illustrating revenue contributions by age group and gender.
- Heatmaps showing geographic revenue distribution.

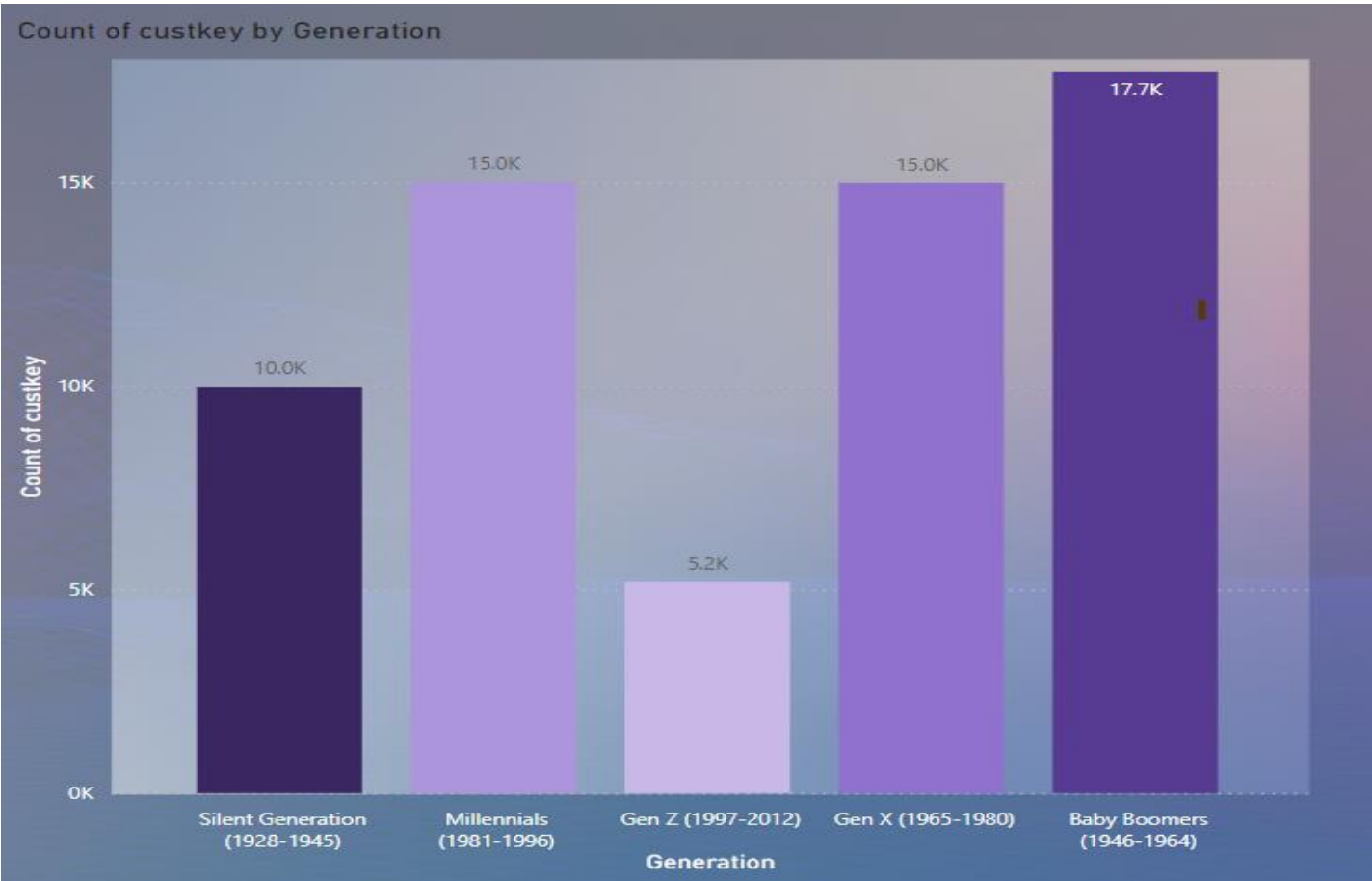


Figure 2: Age Distribution



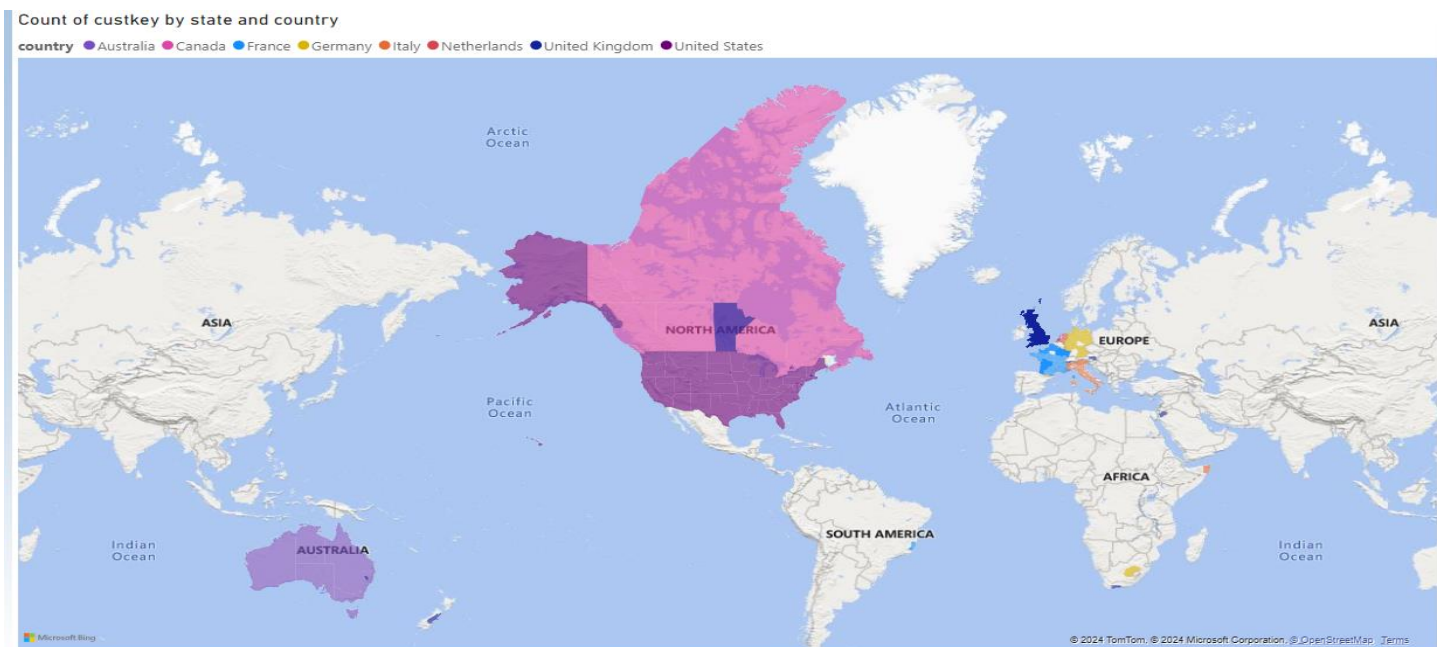


Figure 3: Customer Demographic distribution by Location

## 4.2 Purchasing Behaviors Analysis

### Key Insights:

- Customer Segmentation: The "Very Frequent" buyer segment, though small, generates a significant portion of revenue, highlighting the importance of loyalty programs. The majority are "Infrequent" buyers, presenting an opportunity to increase engagement through personalized promotions.
- Order Value Segmentation: Most customers fall into the "Medium" and "Low" order value segments, indicating potential for increasing average order value through upselling and bundling strategies.

### Visualizations:

- Pie charts showing customer segmentation by purchase frequency.
- Box plots representing order value distributions across different customer segments.

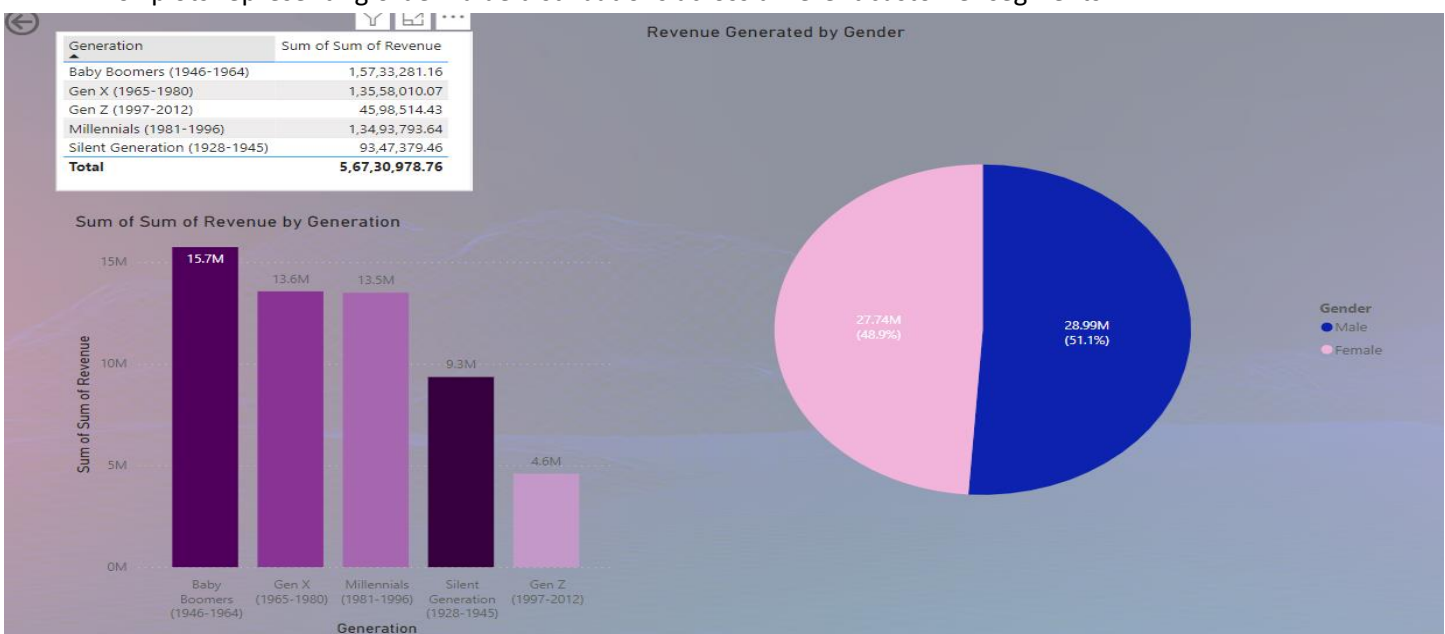


Figure 4: Revenue Distribution by Generation and Gender



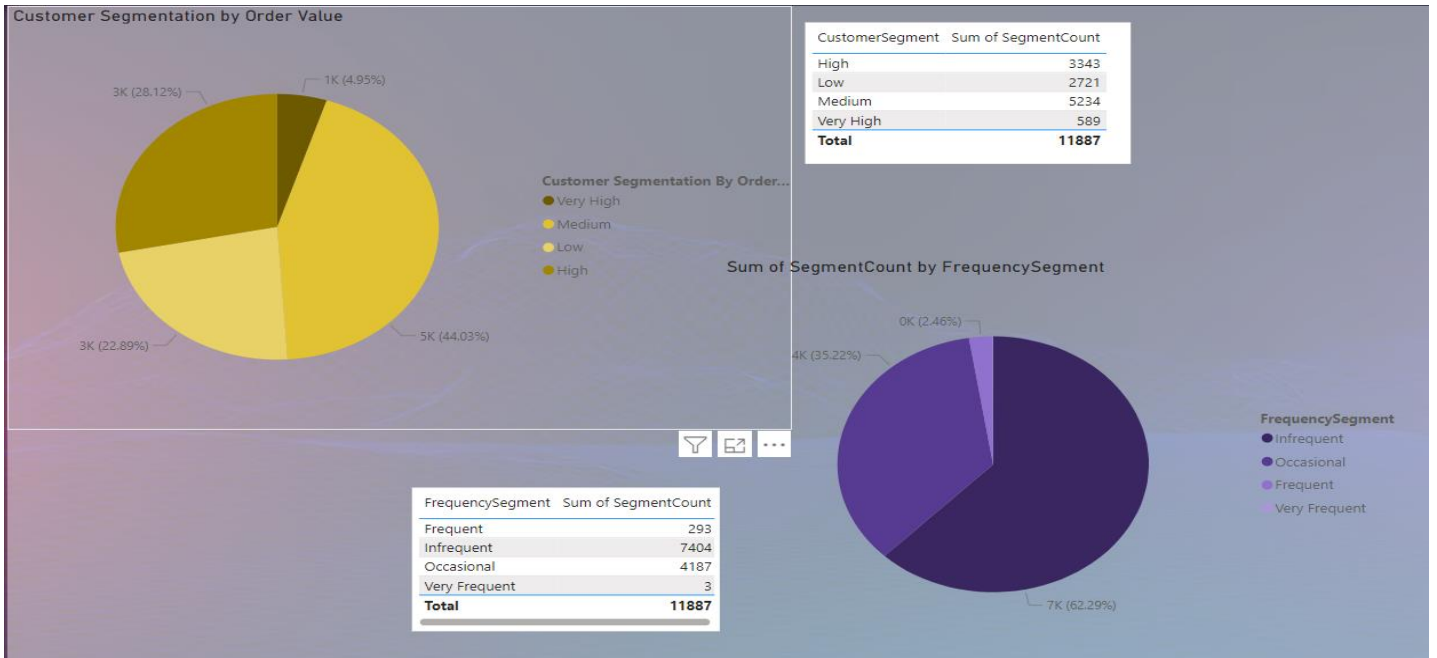


Figure 5: Customer Segmentation by Avg Order Value and Purchase Frequency

### 4.3 Product Performance Analysis

#### Key Insights:

- **Top-Performing Categories:** Categories such as Computers, Home Appliances, and TV and Video dominate sales, with the "Adventure Works Desktop PC" series being the top-selling product.
- **Profit Margins:** Categories like "Cameras and Camcorders" and "Computers" show strong profit margins, indicating a need to focus on these areas for profitability.

#### Visualizations:

- Tree maps depicting sales volumes and profit margins by product category.
- Trend lines showing sales performance of top products over time.

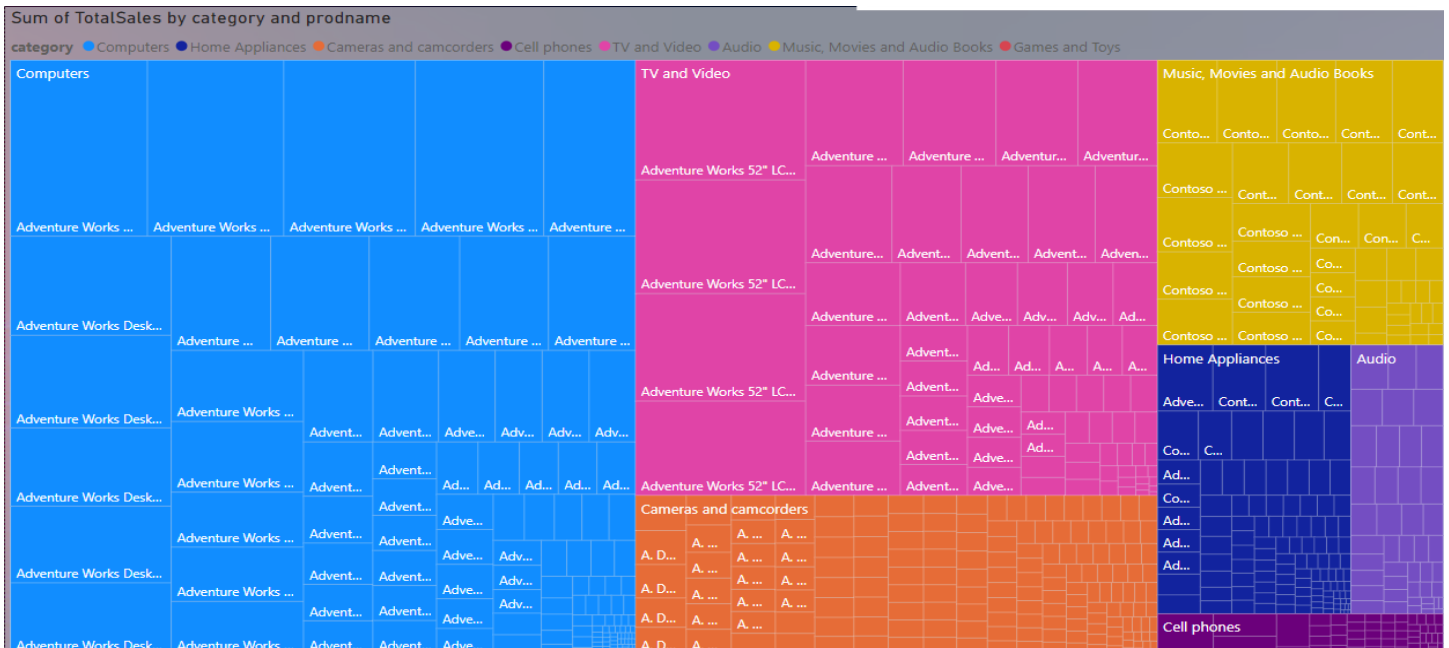


Figure 6: Total Sales by Product

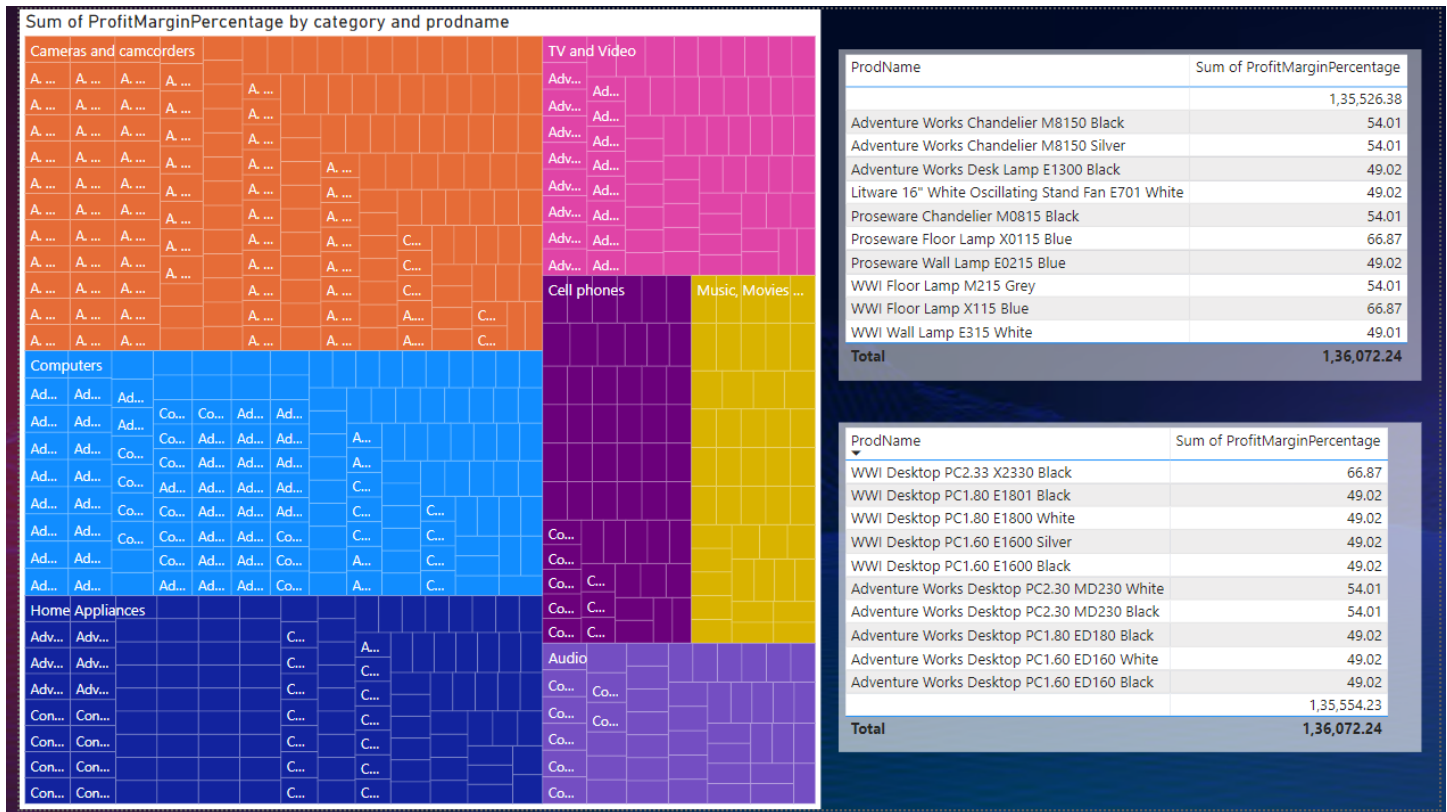


Figure 7: Profit Margin of each Product Category and Products

## 4.4 Store Operations Analysis

### Key Insights:

- Sales by Store: The online store significantly outperforms physical stores, but certain physical stores, especially in the U.S., also show strong performance, suggesting potential for targeted marketing and operations focus.
- Online vs. Offline Sales: The strong online presence indicates an opportunity to further invest in digital channels.

### Visualizations:

- Comparative bar charts for online vs. offline sales.
- Scatter plot of store performance.

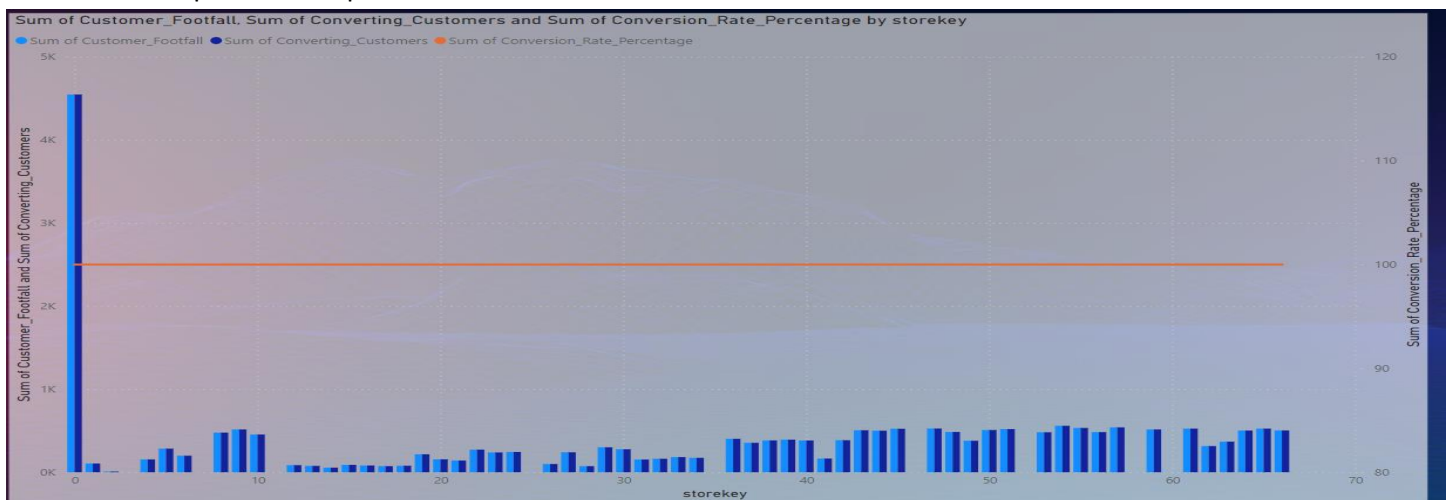


Figure 8: Comparison by Store Footfall Conversion Rate

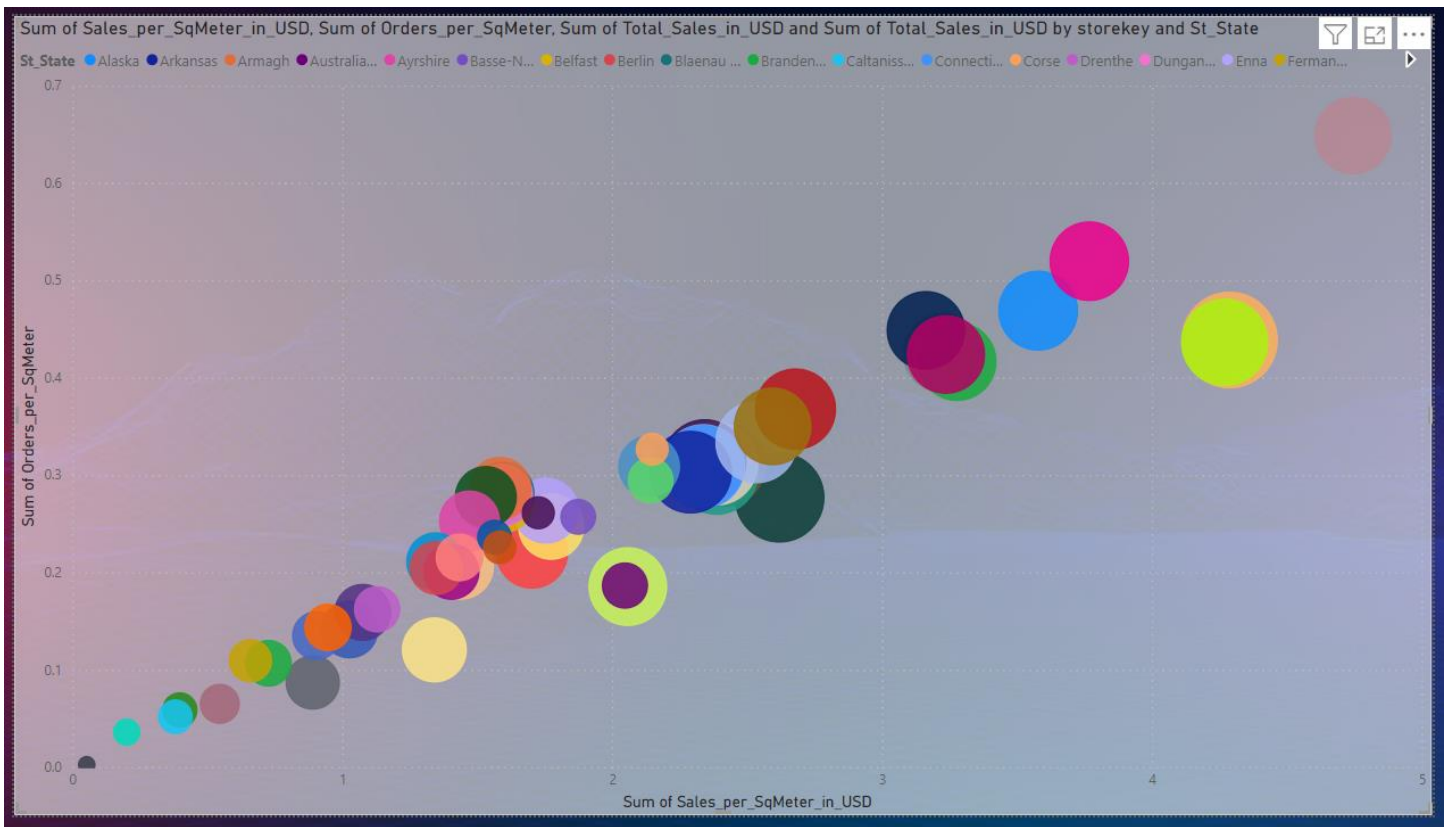


Figure 9: Operation Efficiency of Stores by Sum of Sales per m<sup>2</sup>

## 4.5 Currency Impact on Sales

### Key Insights:

- Seasonal and Currency Trends: Monthly sales data shows peaks in December, likely due to holiday shopping, and fluctuations that could be linked to currency movements. These factors should be incorporated into forecasting models.

### Visualizations:

- Line graphs showing monthly sales trends and overlaying currency exchange rates.
- Correlation plots between currency fluctuations and sales data.

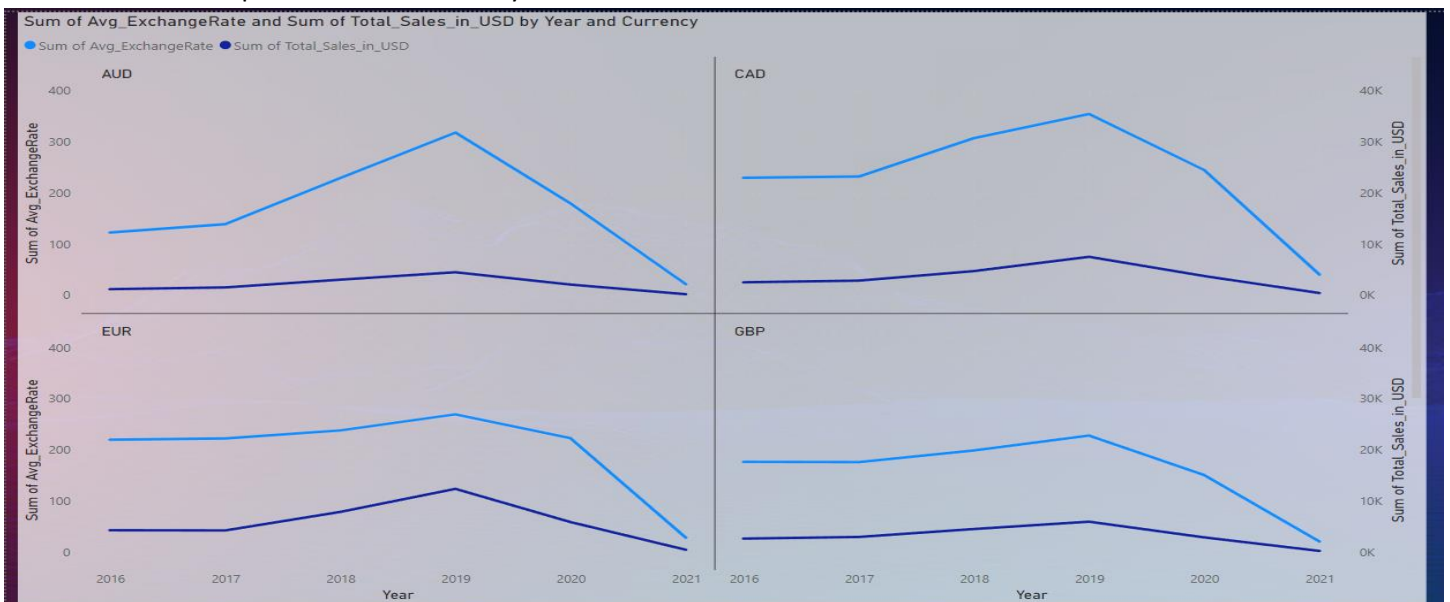


Figure 10: Exchange Rate Vs Sales

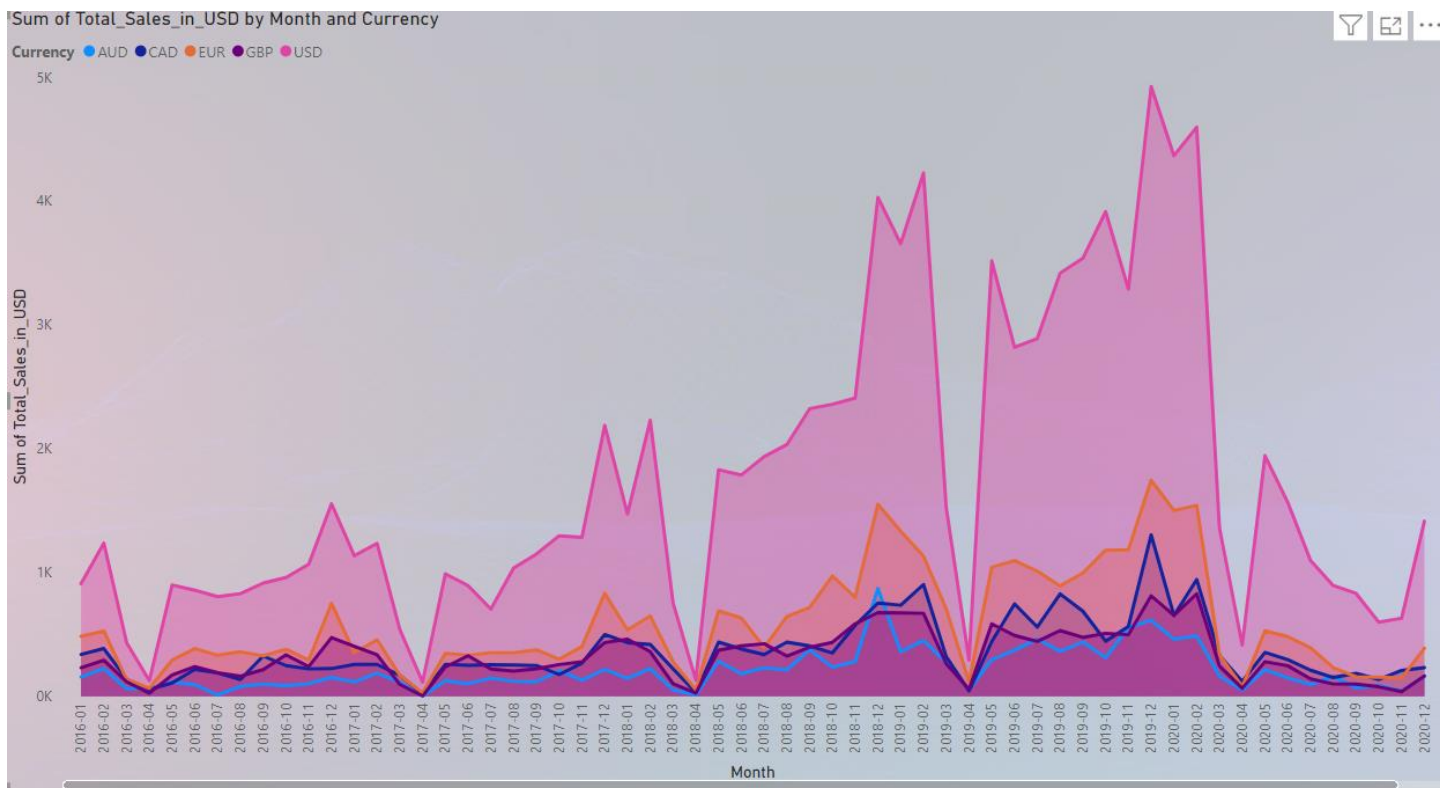


Figure 11: Montly Trend in Internasional Sales

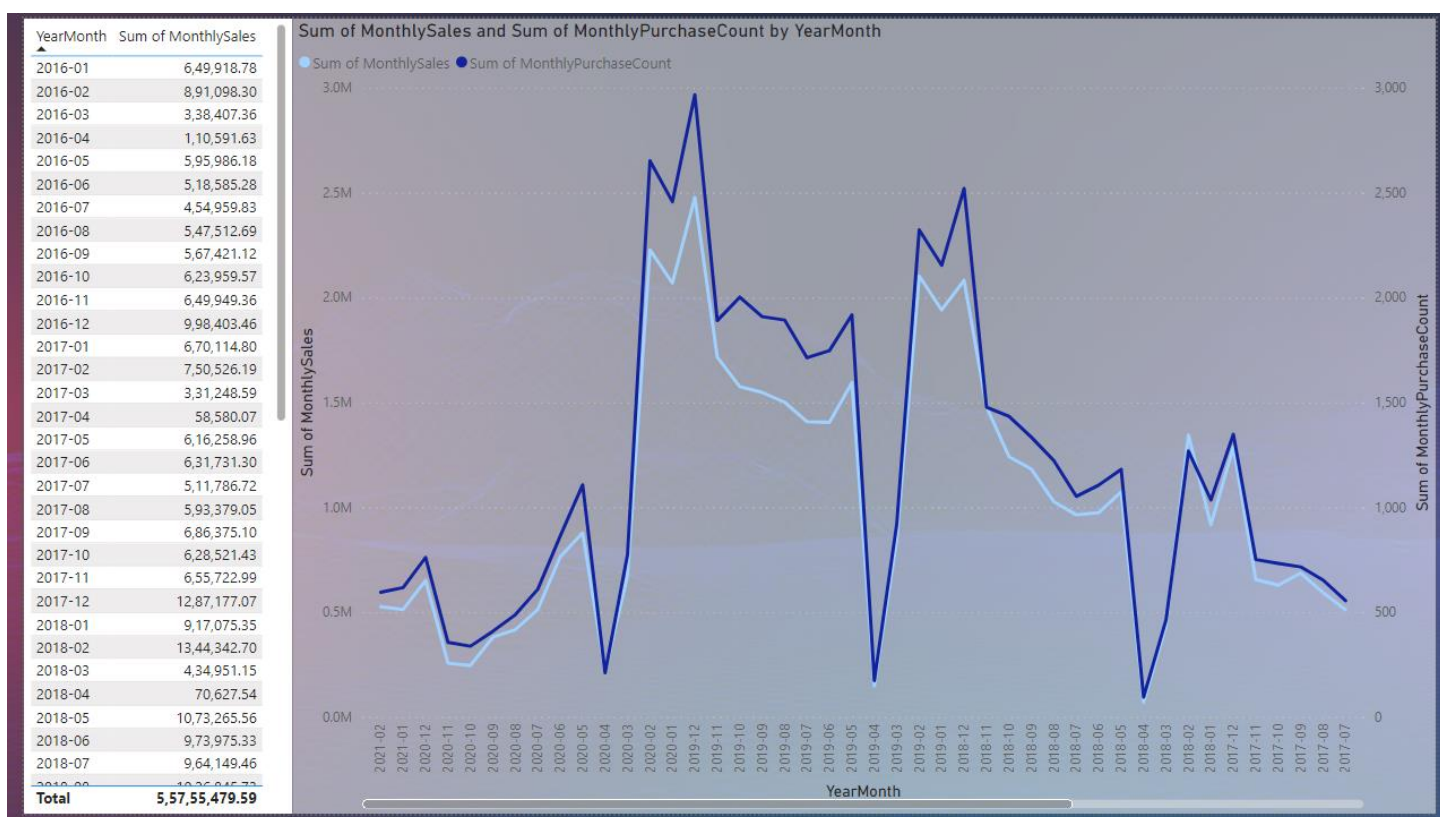


Figure 12 : Seasonal trend in Sales by Purchase Count





## 5. Recommendations

- **Marketing Strategies:** Develop targeted campaigns focusing on Baby Boomers and Gen X, and create growth strategies for Gen Z. Implement gender-specific marketing initiatives to better cater to the interests of male and female customers.
- **Geographic Expansion:** Evaluate potential in underperforming regions, particularly in Europe, to expand market share.
- **Loyalty Programs:** Implement or enhance loyalty programs aimed at converting infrequent buyers into frequent ones.
- **Upselling Strategies:** Develop strategies such as bundling and targeted upselling to increase the average order value among "Medium" and "Low" order value customers.
- **Product Development:** Expand successful product lines, especially in high-performing categories like Computers. Reevaluate or discontinue underperforming products to optimize product offerings.
- **Pricing Strategy:** Review pricing strategies for categories with lower margins to enhance profitability.
- **Digital Investment:** Continue to invest in and optimize the online store to maintain and grow the digital market share.
- **Store Optimization:** Focus on optimizing operations in high-performing physical stores and explore opportunities to enhance underperforming stores, particularly in regions with growth potential.
- **Sales Forecasting:** Integrate seasonality and currency fluctuation considerations into sales forecasts to better manage inventory and pricing strategies.
- **Risk Management:** Implement risk management strategies to hedge against adverse currency movements that could impact profitability



## 9. Conclusion

The EDA of Global Electronics has provided valuable insights across customer demographics, purchasing behaviors, product performance, store operations, and currency impacts. These insights will empower the company to refine its marketing strategies, optimize inventory management, enhance sales forecasting, guide product development, and make informed decisions about store operations and expansion. By acting on these recommendations, Global Electronics can increase customer satisfaction, maximize revenue, and drive overall business growth.

## 10. Next Steps

1. **Action Implementation:** Prioritize and implement the recommendations based on feasibility and impact.
2. **Continuous Monitoring:** Establish a framework for ongoing data analysis to monitor the effectiveness of implemented strategies and make adjustments as needed.
3. **Further Research:** Explore deeper analyses in specific areas like customer lifetime value, product lifecycle management, and market expansion opportunities.

This report serves as a foundation for driving data-informed decisions that will support Global Electronics in achieving its strategic goals and enhancing its competitive edge in the market.