Project 2

25982672

How do customer spending trends, product performance, and seasonal trends influence profitability and customer retention in retail stores?

Introduction

In todays fast changing retail environment, insights into retail data sales analytics are a cornerstone to the success of businesss. By executing the evaluation of online retail purchase data this project will reveal insights into customer spending behaviors, product performance, seasonal trends, and customer loyalty. Each analysis focuses on a key aspect that, when combined, provides a holistic view of how these factors influence business performance. Using data-driven analysis, we would be able to identify the profitability drivers and customer retention that would be capable of usefully informing the optimized business operations and revenue generation strategies. The analyses cover everything from identifying top customers and best-selling products to understanding customer retention and the effects of seasonality on sales.

The following key questions will be explored:

1. With different analyses, who are the biggest spending customers, and how do we support them better?

2. What products best sell, and how do they drive overall sales?

3. At what time of the year do sales peak, being related to seasonal factors?

4. To what extent would customer loyalty be reflected in overall spending, and what is the retention rate?

This project integrates cleaning and transforming data, Exploratory data analysis , and visualization to provide actionable insights that will drive business decisions while strengthening relationships with customers

1

Data Preparation

For the purpose of this project, the dataset used is "Online Retail.xlsx," containing transactional data from an online retail store. This dataset includes information such as product details, sales quantities, unit prices, and customer identifiers.

Before conducting any analysis, it was essential to perform key data cleaning and preparation steps to ensure that the analysis would be accurate and reliable. The data cleaning process followed these key steps:

Removing Duplicates

Duplicate records were identified and removed from the dataset to ensure that each row represented a unique transaction. By removing duplicates, we prevented any inflation in sales statistics due to repeated entries.

- **Duplicates removed:** A total of 5268 of duplicates were identified and removed while a total of 135037 with missing CustomerIDs were identified and removed. These as customer ID are essential for the accuracy of the analysis. While this is a significant portion of the data, it was necessary because customer-specific analysis required valid IDs. Other missing values were not addressed in this round of cleaning as they did not pertain to key fields for the current stage.

Data Type Conversion and Feature Engineering

To facilitate the analysis of transaction trends over time, several transformations were applied:

1. Date Conversion: The InvoiceDate field, initially in string format, was converted to datetime format to enable time-based analyses.

2. **Feature Creation:** New time-related features such as `Year`, `Month`, `Day`, and `Hour` were extracted from the `InvoiceDate` field, allowing for deeper analysis of seasonal or time-based patterns.

Filtering Invalid Quantities

Transactions with negative quantities were filtered out, as they are typically associated with returns or errors that are not relevant to the current analysis. Only transactions with positive quantities were retained for further exploration.
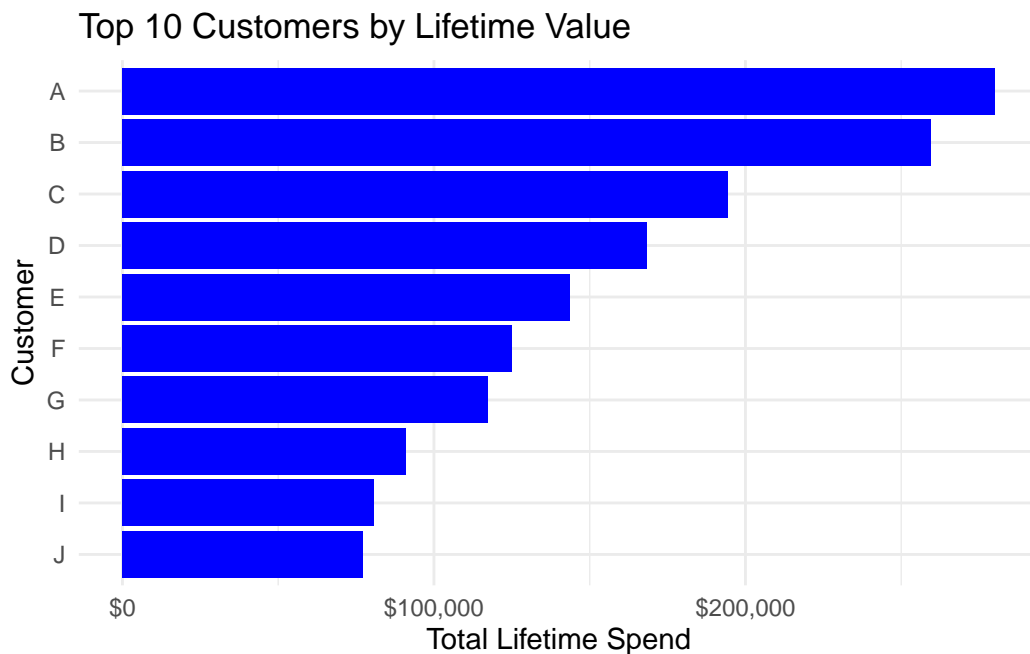
By conducting these data preparation steps, we ensured the dataset was clean, consistent, and ready for the analysis phase.

```
Removed duplicates: 5268

Removed missing CustomerIDs: 135037
```
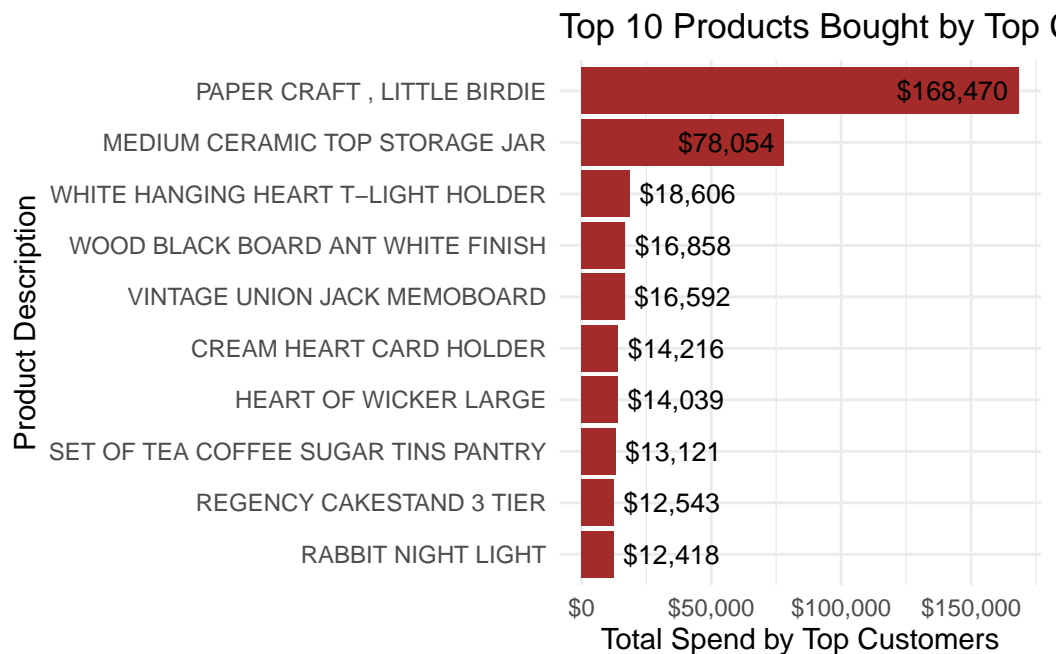
Exploratory Data Analysis and Reporting

What do the customer lifetime values (CLV) indicate about customer profitability?

Top 10 Customers by Lifetime Value

Interpretation: The first analysis delves into Customer Lifetime Value (CLV), a crucial metric for assessing the long-term revenue potential of customers. Our focus is on the top 10 highest-value customers, who significantly contribute to overall profitability. This highlights the critical need for effective retention strategies. By pinpointing these pivotal customers, businesses can strategically customize their services and engagement efforts, ultimately enhancing customer satisfaction and sustaining long-term profitability. Thus, understanding CLV not only informs business strategies but also reveals valuable insights into consumer behavior, which is essential for fostering lasting relationships and maximizing revenue potential.

What products do these customers spend the most of their money on ?

Top 10 Products Bought by Top (

Product Description	Total Spend
PAPER CRAFT , LITTLE BIRDIE	$168,470
MEDIUM CERAMIC TOP STORAGE JAR	$78,054
WHITE HANGING HEART T–LIGHT HOLDER	$18,606
WOOD BLACK BOARD ANT WHITE FINISH	$16,858
VINTAGE UNION JACK MEMOBOARD	$16,592
CREAM HEART CARD HOLDER	$14,216
HEART OF WICKER LARGE	$14,039
SET OF TEA COFFEE SUGAR TINS PANTRY	$13,121
REGENCY CAKESTAND 3 TIER	$12,543
RABBIT NIGHT LIGHT	$12,418

Total Spend by Top Customers

$0 $50,000 $100,000 $150,000

Interpretation: The analysis above shows the top 10 products bought by the top 10 customers.If certain products consistently drive high revenue, it's smart to focus on those.By identifying the top customers based on CLV and identifying which they buy the most, businesses can tailor marketing and retention strategies to ensure these high-value customers remain engaged. This helps focus efforts where the most value is generated.

What is the customer retention rate for the top 10 customers ?

## Customer Retention Rate for Top 10 Customers
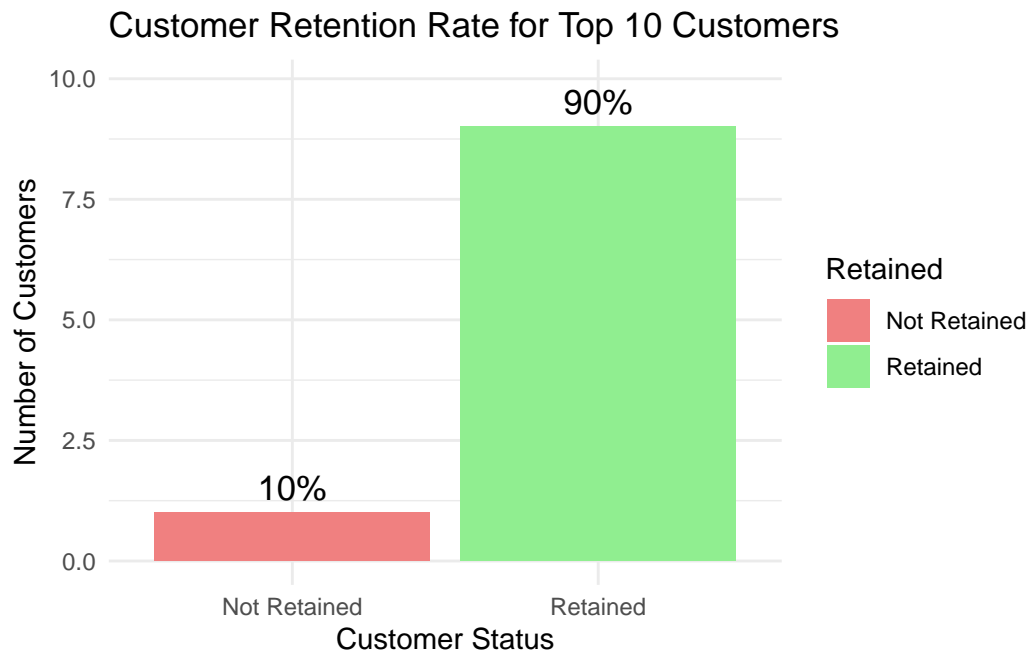
Interpretation: Now that we know who these top 10 customer we can determine the businesses engagement with them by determining the retention rate. According to the analysis above retention rate for high values customers is 90%,while the churn rate is 10 %. which signals high engagement with the customers that generate the most profit for the retail business.

How does the overall customer retention rate reflect the retail business performance?

Customer Retention Rate
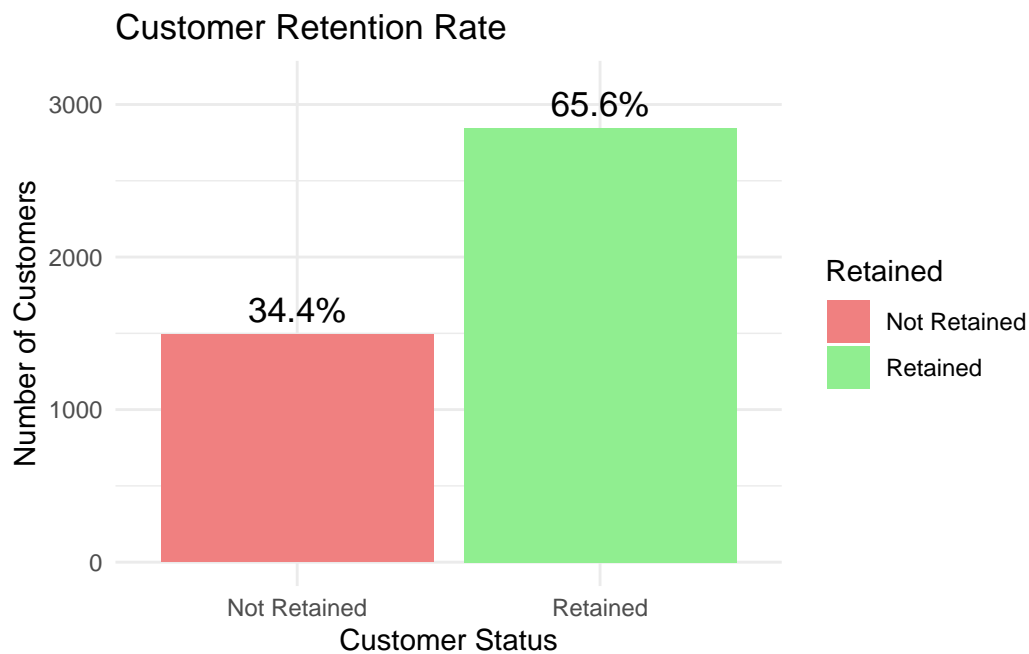
Interpretation: While the retention rate for the top 10 customers is 90 % what the analysis above reveals is that the overall customer retention rate is 65.6% which reveals that there is room for improvement. Understanding why customers remain loyal or disengage can help businesses enhance satisfaction and develop loyalty programs that target a broader audience.

Do loyal customers (frequent buyers) spend more on average?

Relationship between Purchase Frequency and Total Sper
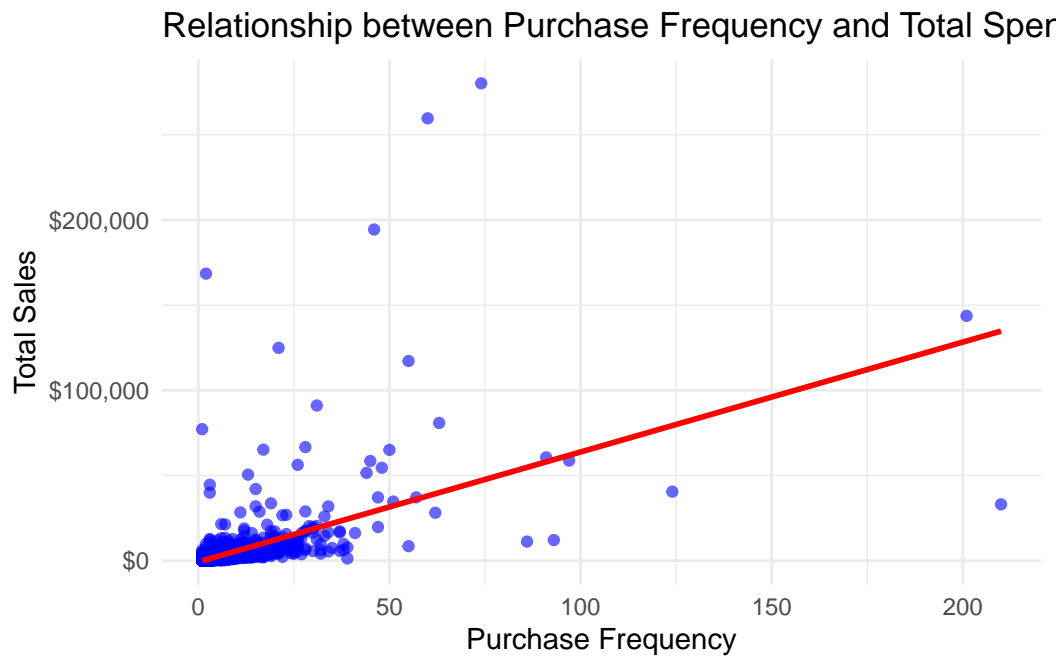
Interpretation: According to the analysis of purchase frequency and total spending there's a positive correlation between how often customers purchase and their overall spending. Frequent buyers contribute more to total revenue, making them a key segment for businesses to focus on. Promoting loyalty through tailored product recommendations, discounts, and loyalty programs can encourage repeat purchases and maximize customer lifetime value.

Are there differences in spending behavior based on geographic location?

Top 5 Sales by Country
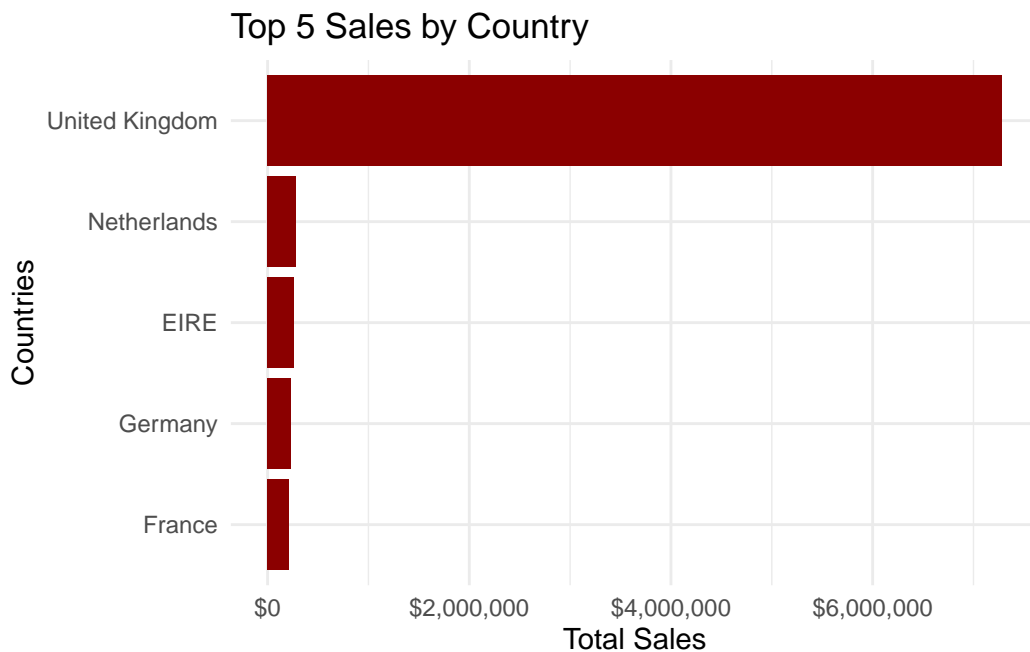
Interpretation: Besides detemining the top customers, determining the top countries based on sales is also important. The geographic analysis above reveals that the United Kingdom is the dominant leader in total sales, highlighting a significant market presence. However, it's not solely about the location of customers, but the variance in their preferences across regions. These geographic differences in product choices imply that businesses must localize their strategies rather than adopting a blanket approach. A "one-size-fits-all" tactic is ineffective; targeted regional marketing can drive significantly improved outcomes. This localized tailoring allows companies to better align with regional consumer behavior, ultimately enhancing performance and customer satisfaction.

What are the top 3 best selling products in countries that generate the most sales?

Top 3 Products Sold by Country
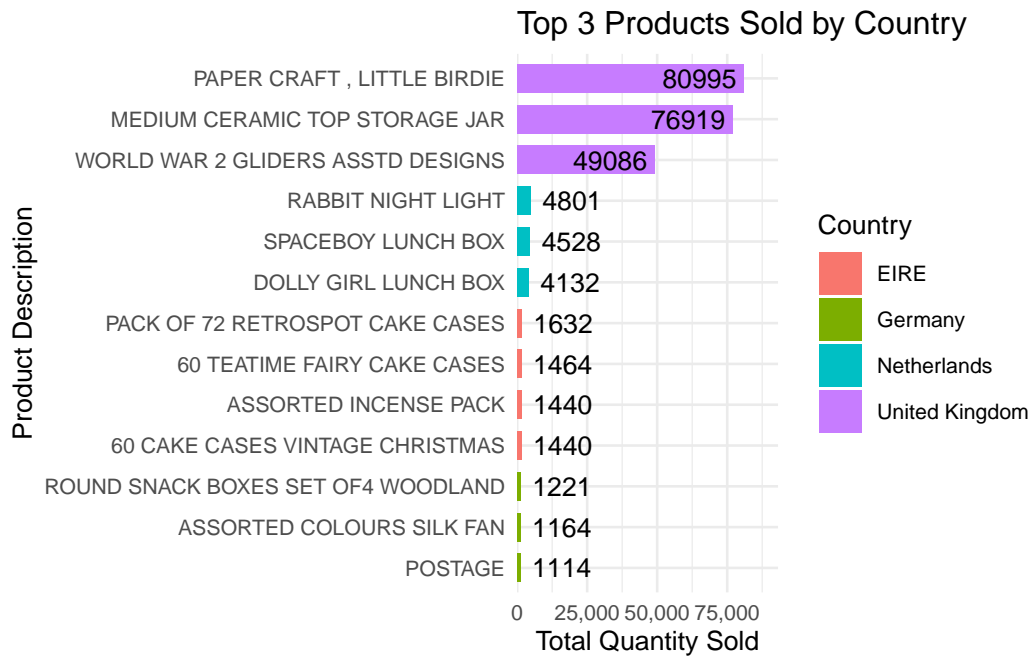
Interpretation: The graph vividly illustrates the top-selling products across various countries, with "PAPER CRAFT, LITTLE BIRDIE" consistently emerging as a bestseller across all regions, indicating its broad, cross-cultural appeal. Yet, country-specific preferences diverge. For example, the "MEDIUM CERAMIC TOP STORAGE JAR" performs exceptionally well in the UK, while it garners less popularity in the Netherlands and the EIRE. This diversity in product sales, ranging from paper crafts to ceramics, toys, and novelty items, reflects retail store's ability to attract a wide customer base. Leveraging these insights could involve tailoring product offerings to align more closely with the preferences of individual countries whether by introducing new items or promoting those already favored in specific regions. Moreover, refining inventory management based on product demand will help maintain optimal stock levels. Lastly, executing targeted marketing campaigns that address regional preferences can significantly drive sales for the top products in each market.

In which month does the retail business generate the highest total sales, and what factors could contribute to seasonal variations in customer spending?

Total Sales by Month
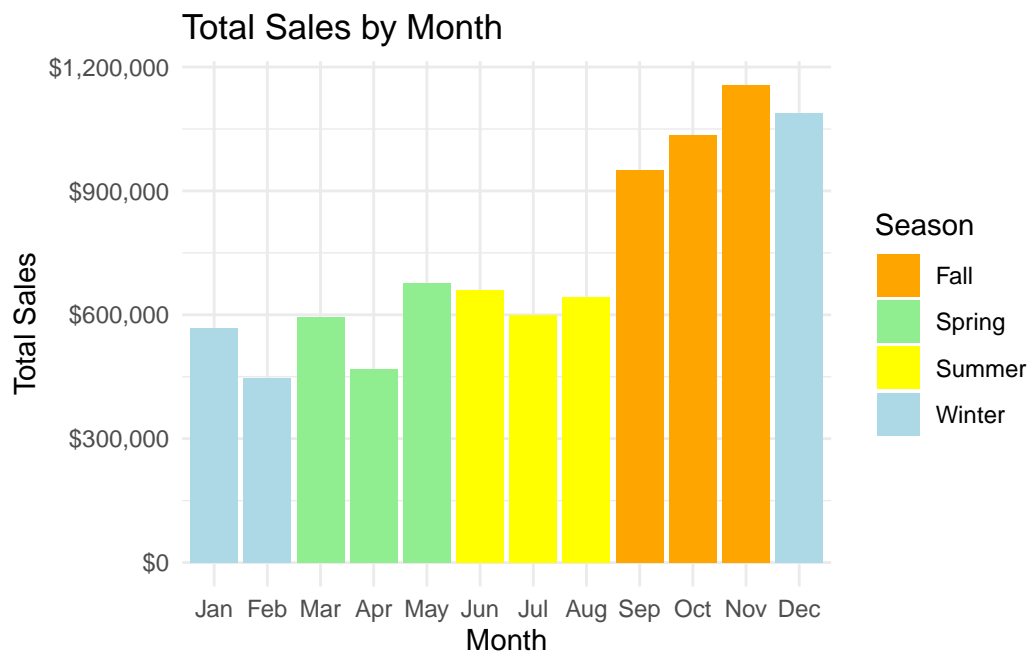
Interpretation: Seasonality has a noticeable influence on retail dynamics, shaping sales trends in discernible ways. The data reveals a pronounced sales surge during the fall, specifically in October and November, likely a consequence of holiday shopping. This seasonal spike underscores an opportunity for businesses to fine-tune marketing efforts and inventory planning, capitalizing on peak demand periods for optimal outcomes. Conversely, the slower seasons present a challenge, yet also a potential area for innovation. Strategic interventions, such as targeted promotions or exclusive offers, could mitigate these lulls, fostering a steadier and more predictable sales trajectory throughout the year.

Conclusion

Through data analysis, this project has revealed key insights into various dynamics, including Customer Lifetime Value (CLV), retention rates, purchase frequency, spending patterns, geographic variability, and seasonal trends. These findings empower businesses to refine strategies that boost loyalty, streamline inventory, and elevate customer satisfaction, driving profitability and securing their long-term position in a competitive market. Future research could explore more granular areas like customer churn, marketing effectiveness, and the influence of emerging trends. Continuous data-driven strategy adaptation is vital for retail success in an ever-changing landscape.