Documentation

By: Riya Mhatre

1. Dataset

The dataset being used is the UCI Repository Online Retail dataset, which contains transactional data for a UK-based, non-store online retailer. The dataset spans transactions that occurred between December 2010 and December 2011. The company primarily sells unique all-occasion gifts, with a significant portion of its customers being wholesalers. This dataset includes information on customer purchases, product details, transaction dates, and sales values, offering valuable insights into sales trends, customer behavior, and purchasing patterns. It serves as a rich resource for analyzing various aspects of retail sales and customer segmentation.

2. Problem Definition and Objective

Using this dataset, I will analyze sales trends and perform customer segmentation to identify distinct purchasing patterns and target key customer groups effectively. By leveraging advanced analytics techniques, I will uncover drivers of high sales and pinpoint low-performing categories, providing actionable insights for improvement. Incorporating external factors such as holidays, weather, and location demographics into my analysis will offer a comprehensive understanding of their impact on sales, enabling me to make data-informed decisions. Additionally, I will provide tailored suggestions to enhance sales performance and optimize product placements, ensuring businesses can adapt strategies to maximize revenue and meet customer demands efficiently.

The main questions I aim to answer in this analysis revolve around understanding sales performance and customer behavior in-depth. I will explore the following key points:

- What are the primary factors influencing sales?
- How do customer demographics and behaviors vary across different locations?
- How do external factors, such as weather and holidays, impact sales?

To guide this analysis, I will track and analyze the following key performance indicators (KPIs):

- Revenue Growth
- Sales by Location
- Customer Retention Rates
- Impact of Holidays

These KPIs will help derive actionable insights and ensure our strategy aligns with business objectives.

3. Summary of Key Findings

1. Time Periods as Predictors of Revenue

Across all locations, it seems that time periods such as the month, day of week, and hour of the day can be used as a predictor for how much revenue the store gets. There is some variation across certain countries, which can be attributed to country-specific factors such as general work schedules or even country-specific holidays. Countries with strong weekday sales suggest alignment with workday shopping routines, while weekend peaks indicate leisure-driven purchases.

Recommendation:
Leverage time-based insights to optimize store hours, inventory levels, and marketing campaigns. For example, launch targeted promotions during holidays or high-traffic hours to maximize revenue.

2. United Kingdom Dominance

The United Kingdom stands out as the top-performing country by far in both revenue and product sales, which is expected given that the store is UK-based. This dominance can be attributed to the strong local presence, brand recognition, and customer loyalty of the store within its home market.

Recommendation: By analyzing the UK's performance, we can identify strategies and practices that resonate with customers in the home market and consider how these might be adapted for success in other regions. This insight is vital for understanding how local market dynamics influence broader trends and for replicating this success globally.

3. Product Performance Insights

Analyzing the top 10 most sold products and the top 10 products by revenue for each country is essential for uncovering valuable customer insights and refining business strategies. The most sold products reveal what customers frequently purchase, reflecting consumer preferences, demand patterns, and price sensitivity in each market. On the other hand, examining the top products by revenue highlights high-value items that contribute significantly to profitability, often revealing opportunities for upselling or premium product positioning. These two perspectives, when combined, offer a nuanced understanding of what drives both volume and

revenue in different regions. By segmenting this analysis by country, the store can tailor product offerings, pricing strategies, and marketing efforts to align with local needs, maximize customer satisfaction, and optimize financial outcomes. This country-specific approach ensures that businesses can effectively address regional variations and maintain a competitive edge globally.

Examining the bottom 10 most sold products and the bottom 10 products by revenue for each country is equally critical, as it helps identify underperforming products and areas for improvement. Products with low sales volumes may indicate a lack of customer interest, ineffective marketing, or misaligned pricing strategies, while low-revenue items could point to products with insufficient profit margins or poor perceived value. By analyzing these bottom performers, businesses can uncover trends such as regional preferences or market-specific challenges that might not be apparent from top-performing products alone. This insight enables informed decisions about whether to discontinue, reposition, or enhance these products to better meet customer needs. Additionally, addressing underperforming items can free up resources for more promising opportunities, ensuring that the product portfolio remains competitive and optimized for each market. This approach fosters a proactive strategy for improving overall profitability and customer satisfaction.

Understanding which products were returned is essential for identifying potential issues that may be impacting customer satisfaction and operational efficiency. High return rates can signal problems such as poor product quality, misleading product descriptions, sizing or functionality issues, or unmet customer expectations. Analyzing returned products provides valuable insights into why customers are dissatisfied, enabling businesses to address these concerns proactively by improving product design, refining marketing materials, or enhancing quality control processes. Furthermore, tracking return patterns can highlight trends, such as specific categories or regions with high return rates, which may require targeted interventions. Minimizing returns not only reduces costs associated with reverse logistics and restocking but also builds customer trust and loyalty by demonstrating a commitment to delivering value and meeting expectations.

Recommendation:

- Promote high-revenue products through targeted campaigns, emphasizing their premium value.
- Use high-volume items to attract customers and bundle them with higher-margin products to increase overall profitability.
- **For Bottom Products by Revenue or Sales**: Low-performing items may highlight pricing issues, lack of appeal, or ineffective marketing.
    - Reassess the viability of these products: consider discontinuing, rebranding, or repositioning them.
    - Use customer feedback to determine whether product improvements or targeted promotions can boost their performance.

4. Customer Insights

Customer segmentation has identified three primary groups: Budget-Conscious Regular Shoppers, who buy frequently but prefer lower-cost items and drive sales volume; High Value Buyers, a smaller yet crucial segment that makes high-value purchases and significantly impacts revenue; and Occasional Buyers, who shop less often but still contribute positively to sales. Key insights reveal that increasing purchase frequency has a greater compounding effect on total revenue compared to simply increasing the quantity of items purchased per transaction, though both strategies contribute positively.

Recommendation:

- Focus on **encouraging repeat purchases** by offering loyalty programs, personalized promotions, or discounts for returning customers.
- Use upselling strategies to increase the number of items purchased per transaction, targeting high-spending and occasional buyers with premium offerings.
- Regularly segment customers to ensure marketing efforts align with evolving purchase behaviors and preferences.

5. External Factors: Holidays

Statistical analysis reveals that holidays do not significantly impact revenue, as indicated by a high p-value (0.698) and a negligible t-statistic (-0.39), suggesting any difference in mean revenue is likely due to random variation. However, holidays do significantly affect the number of products purchased, with a t-statistic of 2.47 and a p-value of 0.013. This indicates a meaningful increase in purchase frequency on holidays, highlighting a potential opportunity for businesses to capitalize on higher product demand during these periods.

Recommendation:

- **Holiday Strategies**: Launch holiday-specific campaigns with curated product bundles or discounts to capitalize on increased demand.
- **Non-Holiday Strategies**: Identify and target key non-holiday periods, such as the start of a month or post-weekend spikes, to maintain consistent revenue.