



Decathlon Case Study

Problem Statement

- A leading retail company aims to enhance its understanding of customer purchasing behaviors and market dynamics to optimize inventory management, boost sales, and improve overall customer satisfaction. By analyzing transaction data from 2009 to 2011, the company seeks to identify key trends in product popularity, revenue generation, and purchasing patterns across different markets to inform targeted marketing strategies and operational decisions.

Dataset Information

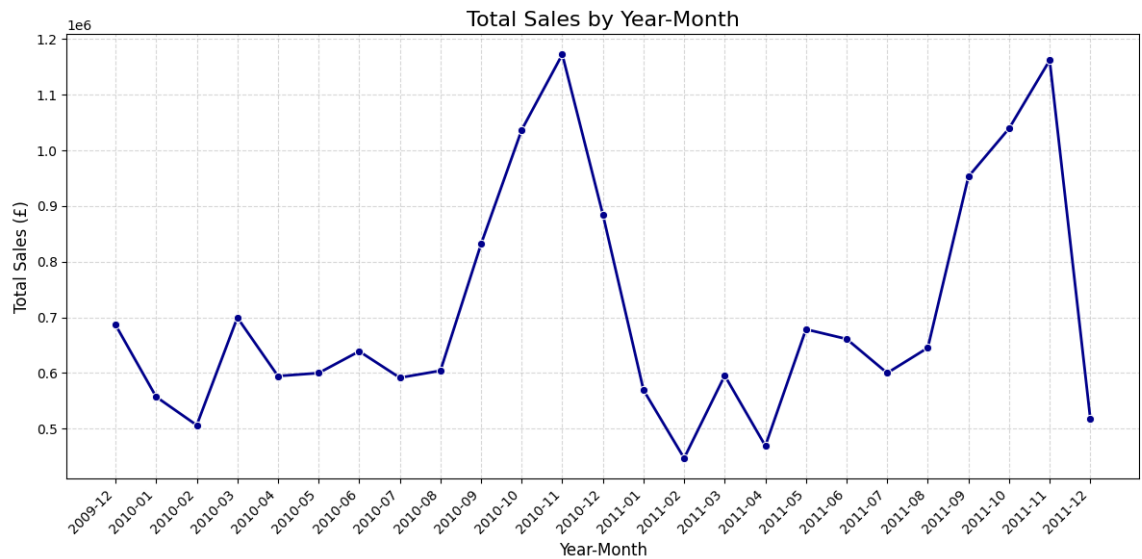
- The **dataset** includes over 1 million transaction records from 2009 to 2011, capturing details like product, quantity, price, customer ID, and country, enabling detailed analysis of purchasing trends.
- **Shape:** (1067371, 8)
- **Columns:**
 - Invoice No - Invoice number. Nominal. A 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancellation.
 - Stock Code - Product (item) code. Nominal. A 5-digit integral number uniquely assigned to each distinct product.
 - Description - Product (item) name. Nominal.
 - Quantity - The quantities of each product (item) per transaction. Numeric.
 - Invoice Date - Invoice date and time. Numeric. The day and time when a transaction was generated.
 - Unit Price - Unit price. Numeric. Product price per unit in sterling (Â£).
 - Customer ID- Customer number. Nominal. A 5-digit integral number uniquely assigned to each customer.
 - Country - Country name. Nominal. The name of the country where a customer resides.

Data Processing & Cleaning Steps

- **Merged Data:** Combined transaction records from 2009-2010 and 2010-2011 into a single dataset.
- **Removed Cancellations:** Excluded transactions where the InvoiceNo starts with 'C', indicating cancellations.
- **Filtered Customer IDs:** Removed records with missing CustomerID to ensure data completeness.
- **Formatted Dates:** Converted InvoiceDate to DateTime format for more accurate temporal analysis.
- **Extended Date Attributes:** Added columns for 'date', 'month', 'week', and 'time' to facilitate detailed temporal analysis.

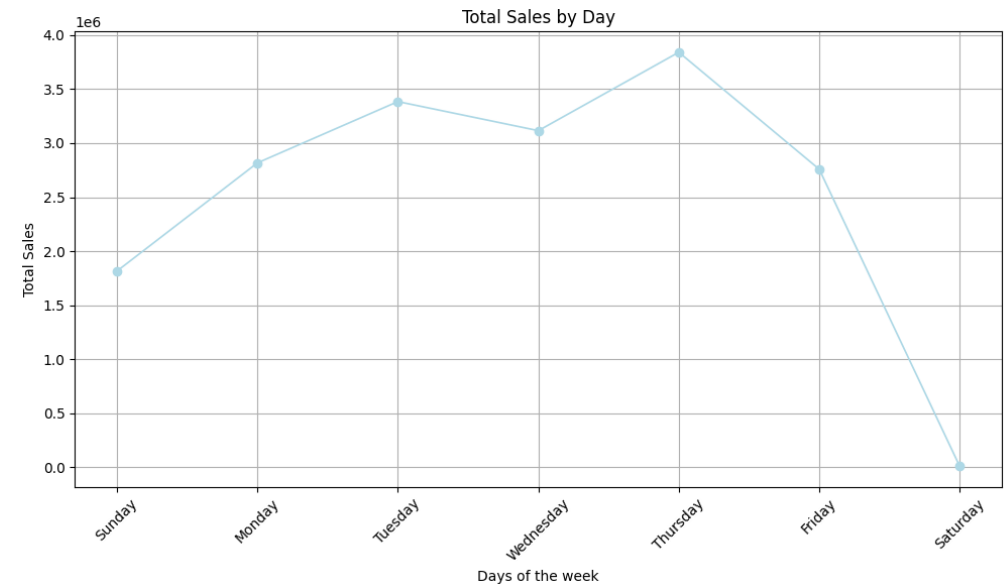
Analysis – Total Sales by Year-Month

- **Seasonal Peaks:** There are noticeable peaks in sales volume around certain times of the year. This could align with seasonal shopping periods such as holidays or specific sales events (like Black Friday or Christmas).
- **Growth Trends:** There seems to be an upward trend in sales volumes towards the end of each year (Q4), suggesting increased consumer buying activity during the holiday season. This trend repeats annually, indicating consistent consumer behavior.
- **Mid-Year Variations:** There are fluctuations in sales throughout the middle of the year. These could be influenced by various factors such as promotional events, changes in stock availability, or market conditions.
- **End of Year Spike:** The significant spikes at the end of each year could be effectively utilized for strategic planning. The company might consider increasing stock levels, launching marketing campaigns, or offering special promotions during these periods to maximize revenue.



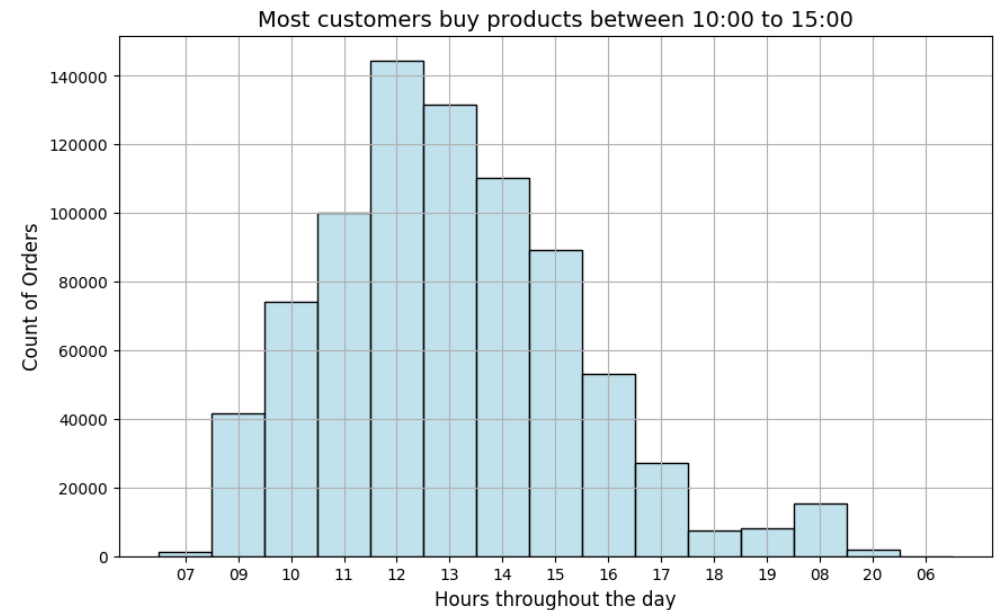
Analysis – Total Sales by Day

- **Weekday Dominance:** Sales are higher on weekdays compared to weekends. This suggests that customer purchasing activity is concentrated during the regular workweek.
- **Mid-Week Peak:** There appears to be a peak around the middle of the week, particularly on Tuesday and Thursday. This could indicate when customers are most likely to engage in purchasing activities, possibly related to mid-week promotions or shopping habits.
- **Weekend Drop:** Sales significantly drop during the weekend. This might be due to the nature of the business or possibly reduced hours of operation. This also might indicate that the target customer base prefers weekday shopping, perhaps due to work schedules.



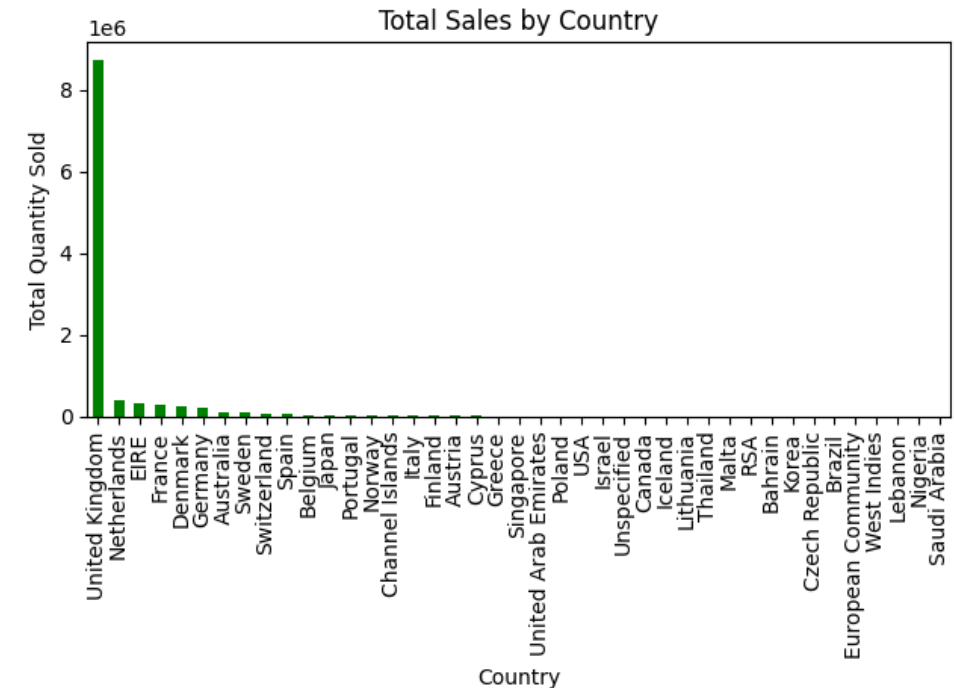
Analysis – Total Sales by Time of Day

- **Peak Purchase Hours:** Most orders occur between 11:00 and 15:00, indicating a significant surge in purchasing activity during these hours. This could be due to customers making purchases during lunch hours or taking breaks from work.
- **Off-Peak Opportunities:** The lower activity in the morning and late evening suggests potential for targeted promotions or marketing strategies to increase sales during these times, possibly by offering special discounts or incentives to shop during off-peak hours.



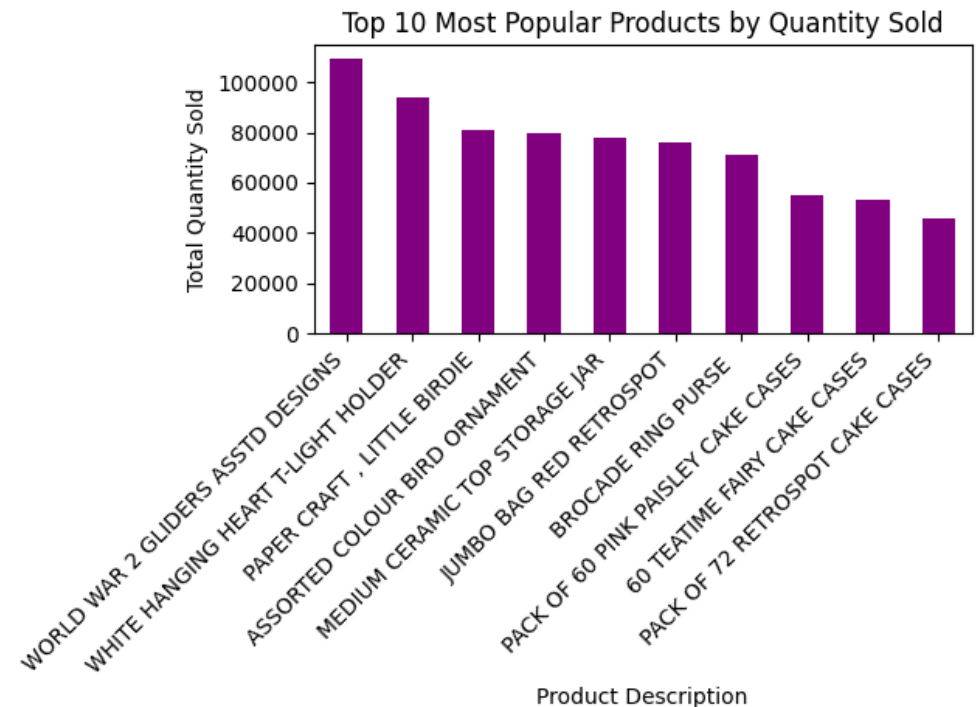
Analysis – Total Sales by Country

- **Dominant Market:** The United Kingdom overwhelmingly dominates the sales volume. This suggests that the primary market for the company is the UK, and most of the customer base is concentrated there.
- **Minor Markets:** Other countries contribute significantly less to the total sales. However, countries like Netherlands, Ireland, France & Germany show notable sales, indicating potential key secondary markets.
- **Expansion Opportunities:** The presence of other European countries in the list suggests that expanding marketing and sales efforts in these regions could be beneficial. Tailoring strategies to meet the specific needs of these markets could potentially increase sales.



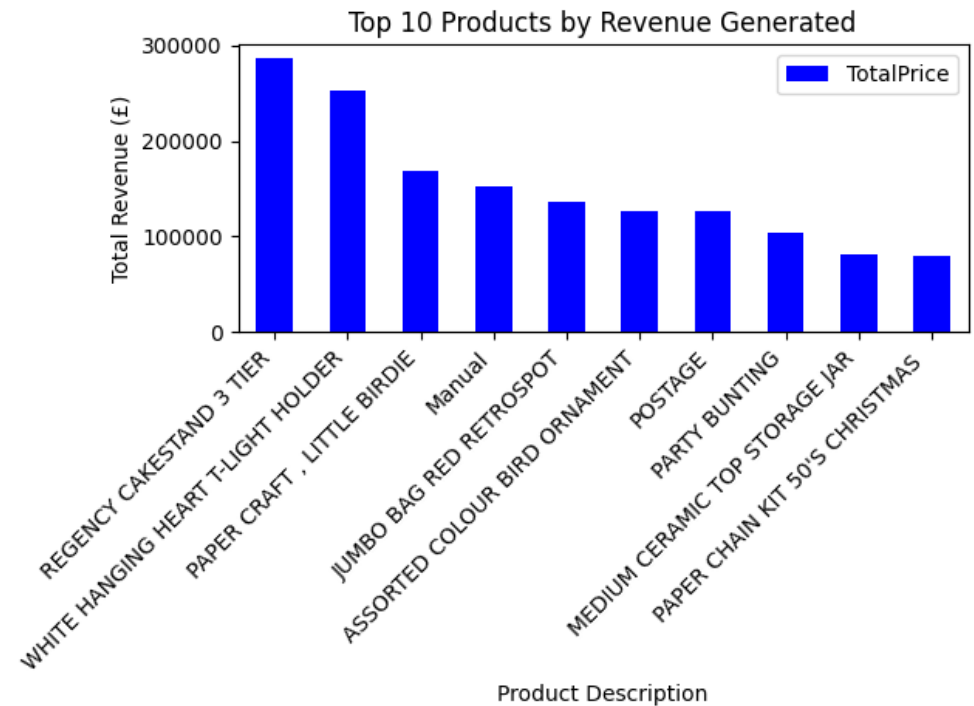
Analysis – Top 10 Most Popular Products by Quantity Sold

- **High Demand Products:** The leading products, as seen in the chart, likely have broad appeal to the customer base or may be essential items that customers purchase frequently.
- **Inventory Management:** Understanding which products are most popular can help optimize inventory management. Stock levels of these items should be carefully managed to avoid stockouts, especially during peak sales periods.
- **Promotional Opportunities:** Products that have high sales volumes are excellent candidates for promotions or upselling strategies. Bundling them with less popular items or offering discounts on bulk purchases could increase overall sales.



Analysis – Top 10 Products by Revenue Generated

- **Revenue Drivers:** The products leading in revenue may not necessarily be the same as those with the highest sales volumes. This distinction highlights the importance of considering both quantity sold and unit price in strategic decisions.
- **Pricing Strategy:** High-revenue products might benefit from pricing strategies that optimize profit margins without sacrificing sales. This could include tiered pricing, promotional discounts, or bundling with other products.
- **Supply Chain Management:** Ensuring the availability of these high-revenue products is critical. Supply chain efforts may need to be adjusted to cater to the demand patterns of these products, ensuring that they are always in stock.



RFM(Recency, Frequency, Monetary)



Recency: How recent was a customer's latest purchase? In this case we're considering the latest date of the dataset & the last purchase date of a specific Customer ID & calculating the recency wrt the latest date.



Frequency: How often does a customer make purchases? Counting the no of Invoice for each Customer ID



Monetary: How much does a customer spend? Total of all the transactions for each Customer ID i.e. Sum of Quantity * Price

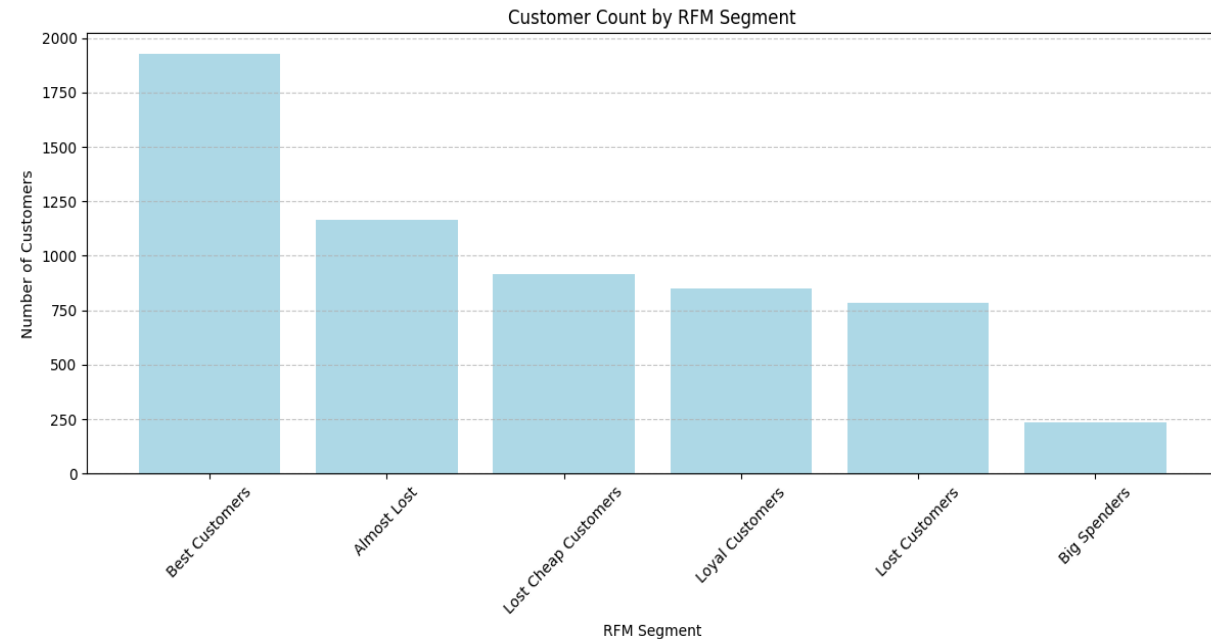
RFM Segmentations

- **Best Customers:** Bought most recently, more often & spends the most.
- **Loyal Customers:** Buys more frequently.
- **Big Spenders:** Spends the most.
- **Almost Lost:** Haven't purchased for some time but used to purchase frequently & spend most.
- **Lost Customers:** Haven't purchased for a long time but used to purchase frequently & spend most.
- **Lost Cheap Customers:** Purchased long time ago, purchased few & spend very little.



RFM Analysis & Recommendations

- **Best Customers:** Maintain engagement with exclusive rewards and recognition. Consider creating brand ambassador roles to capitalize on their potential influence.
- **Almost Lost:** Incentivize their return with time-sensitive offers. Conduct a survey to identify any service or product issues that led to their decreased engagement.
- **Lost Cheap Customers:** Offer small, targeted promotions or discounts to encourage another purchase. Assess whether retaining this segment is cost-effective.
- **Loyal Customers:** Encourage higher spending through bundle offers or benefits that add value to their frequent purchases, without eroding the profit margin.
- **Lost Customers:** Reconnect with these customers through personalized messages that acknowledge their absence and offer them an attractive reason to return.
- **Big Spenders:** Focus on high-value product recommendations and personalized service to continue to drive their high spending behavior



Top 3 Recommendations for CEO

- **Enhance Inventory Management:** Increase stock for top-selling and high-revenue products, especially during peak sales periods identified from monthly and daily trends.
- **Targeted Marketing Strategies:** Strengthen marketing efforts in the UK and expand into other promising markets like Germany and France, focusing on localized campaigns to increase market share.
- **Invest in Data Analytics:** Utilize data-driven insights to optimize product offerings and customer engagement, improving decision-making and operational efficiency across the company.

