

E-commerce Sales Analysis Report

Insights from Transaction Data 2010–2011

Prepared By

Tahmid Rahman Siddiki

Data Analyst

Prepare on September 2025

2025

E-Commerce Sales Analysis Business Report

Executive Summary

This report presents a comprehensive analysis of e-commerce sales data, identifying key trends, customer behaviors, and product performance to inform strategic business decisions. The analysis, conducted on approximately 392,000 transactions from 2010–2011, reveals significant insights into sales seasonality, top-performing products and markets, and distinct customer segments. Key findings include a dominant sales contribution from the United Kingdom, strong international growth opportunities in European markets and Australia, and a pronounced seasonal sales spike from September to November. Customer segmentation using RFM (Recency, Frequency, Monetary) analysis categorizes customers into VIP, Loyal, Potential, At Risk, and Lost groups, enabling targeted marketing and retention strategies. The report concludes with actionable recommendations aimed at optimizing inventory, enhancing marketing efforts, and improving customer relationship management to drive revenue growth and mitigate business risks.

1. Introduction

The objective of this project is to leverage data analytics to extract meaningful insights from an e-commerce sales dataset. The primary goal is to understand sales patterns, identify high-value products and customer segments, and provide data-driven recommendations for business optimization. The dataset comprises transactional records from an online retail store, detailing sales activities over a period spanning from December 2010 to December 2011.

2. Data Cleaning and Preprocessing

Before analysis, the raw dataset underwent a rigorous cleaning and preprocessing phase to ensure data quality and reliability. The initial dataset contained 541,909 entries and 8 columns. The following steps were executed:

- **Duplicate Removal:** A total of 5,268 exact duplicate rows were identified and removed.
- **Handling Missing Values:** Rows with missing Description and CustomerID were dropped, as these are crucial for product identification and customer segmentation, respectively. Initially, 1,454 descriptions and 135,080 customer IDs were missing.
- **Invalid Transaction Filtering:** Transactions with non-positive Quantity (indicating returns or errors) and non-positive UnitPrice were excluded. This step addressed 10,624 rows where Quantity was less than or equal to zero.
- **Date Format Standardization:** The InvoiceDate column was converted to a proper datetime format, ensuring consistency and enabling time-series analysis. The initial

parsing attempt resulted in 308,950 missing parsed dates due to incorrect format assumption, which was corrected to %m/%d/%Y %H:%M for successful parsing.

- **Index Reset:** The DataFrame index was reset after cleaning operations.

Following these cleaning steps, the refined dataset consisted of 392,692 transactions and 9 columns, ready for in-depth analysis.

3. Key Findings and Analysis

3.1 Product Performance

Analysis of product sales revealed a concentration of revenue generated by a few key items, often referred to as "hero products." These products are critical drivers of sales volume and monetary value.

Top 10 Products by Total Sales:

Description	Total Sales (£)
PAPER CRAFT , LITTLE BIRDIE	168,469.60
REGENCY CAKESTAND 3 TIER	142,264.75
WHITE HANGING HEART T-LIGHT HOLDER	100,392.10
JUMBO BAG RED RETROSPOT	85,040.54
MEDIUM CERAMIC TOP STORAGE JAR	81,416.73
POSTAGE	77,803.96
PARTY BUNTING	68,785.23
ASSORTED COLOUR BIRD ORNAMENT	56,413.03
Manual	53,419.93
RABBIT NIGHT LIGHT	51,251.24

Notably, "POSTAGE" also appears as a significant revenue generator, indicating the importance of shipping services within the business model.

3.2 Geographical Sales Analysis

The geographical distribution of sales highlights the United Kingdom as the primary market, accounting for over 90% of total sales. However, several international markets demonstrate strong potential for growth.

Top 10 Countries by Total Sales (Excluding UK):

Country	Total Sales (£)
Netherlands	285,446.34
EIRE	265,262.46
Germany	228,678.40
France	208,934.31
Australia	138,453.81
Spain	61,558.56
Switzerland	56,443.95
Belgium	41,196.34
Sweden	38,367.83
Japan	37,416.37

The Netherlands, EIRE (Ireland), Germany, and France represent the strongest European markets, while Australia stands out as a significant overseas contributor, despite potentially higher logistical costs.

3.3 Monthly Sales Trend

The monthly sales trend reveals a distinct seasonal pattern, crucial for inventory management and marketing campaign planning. The provided chart illustrates this trend:

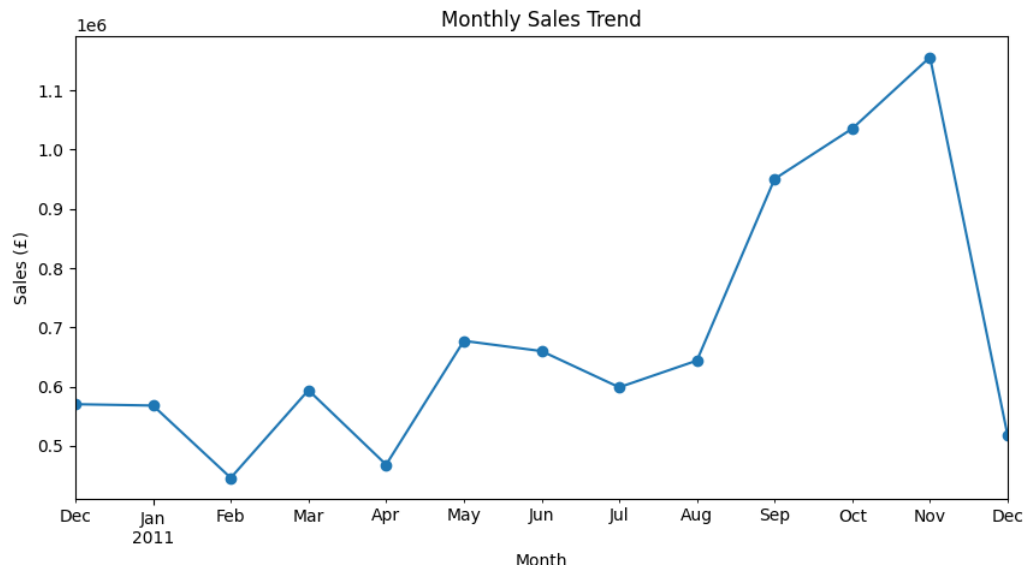


Figure 1: Monthly Sales Trend (December 2010 - December 2011)

The trend shows relatively stable sales in early 2011, followed by a significant increase from September to November, peaking in November. This surge is characteristic of holiday shopping seasons. A sharp decline is observed in December, which can be attributed to the natural end of the holiday peak and potentially incomplete data for the month.

3.4 Customer Behavior

The customer base consists of 4,338 unique customers, with an average order value of £68.38. A critical insight is the substantial reliance on a small number of high-value customers. The top three customers alone contributed over £700,000 in sales, underscoring the importance of customer loyalty and retention strategies for these key accounts.

Top 10 Customers by Total Spend:

CustomerID	Total Sales (£)
14646.0	280,206.02
18102.0	259,657.30
17450.0	194,390.79
16446.0	168,472.50
14911.0	143,711.17
12415.0	124,914.53
14156.0	117,210.08
17511.0	91,062.38
16029.0	80,850.84
12346.0	77,183.60

3.5 Customer Segmentation (RFM Analysis)

To better understand and target customer groups, an RFM (Recency, Frequency, Monetary) analysis was performed. This method segments customers based on how recently they made a purchase, how often they purchase, and how much they spend. The customers were categorized into five distinct segments:

- **VIPs:** Customers with high recency, frequency, and monetary values.
- **Loyal:** Regular customers with good frequency and monetary values.
- **Potential:** Customers who have made recent purchases but may not have high frequency or monetary value yet.
- **At Risk:** Customers whose purchasing activity has declined.
- **Lost:** Customers who have not made purchases for a significant period.

Customer Segment Distribution and Average Monetary Value:

Segment	Number of Customers	Average Monetary Value (£)
VIP	933	6,689.76
Loyal	1,010	1,389.11
Potential	1,088	807.02
At Risk	1,124	300.90
Lost	183	144.31

This segmentation provides a clear framework for developing tailored engagement strategies for each group.

4. Business Recommendations

Based on the comprehensive data analysis, the following strategic recommendations are proposed:

- 1 **Product Strategy:** Prioritize securing consistent stock levels for "hero products" (e.g., Paper Craft, Regency Cakestand). Explore bundling strategies to pair best-selling items with mid-tier products to boost overall sales and introduce customers to a wider range of offerings.
- 2 **Market Expansion:** Intensify focus on expanding market share in high-potential European countries such as the Netherlands, Ireland, Germany, and France. Develop targeted marketing campaigns and localized strategies for these regions. Furthermore, capitalize on the strong performance in Australia by optimizing logistics and potentially establishing regional distribution hubs to reduce shipping costs and improve delivery times.
- 3 **Marketing and Seasonality:** Design and execute pre-holiday marketing campaigns from September to November to maximize sales during the peak season. Implement post-December clearance sales or promotional events to manage inventory and mitigate the sharp decline observed after the holiday rush.
- 4 **Customer Relationship Management (CRM):** Implement a tiered CRM program based on the RFM segments:
 - **VIPs:** Offer exclusive benefits, personalized communications, and early access to new products to reinforce loyalty.
 - **Loyal Customers:** Encourage repeat purchases through loyalty programs, special discounts, and personalized recommendations.
 - **Potential Customers:** Develop campaigns to increase purchase frequency and average order value, such as targeted promotions or product suggestions based on their initial purchases.
 - **At Risk Customers:** Initiate re-engagement campaigns, including win-back offers, feedback surveys, and personalized outreach to understand and address their declining activity.
 - **Lost Customers:** Conduct deep-dive analysis to understand reasons for churn and explore last-ditch re-activation efforts with highly attractive offers.

5. Conclusion

This e-commerce sales analysis underscores the profound value of data-driven insights in shaping effective business strategies. The findings highlight the critical role of specific products and markets, the predictable seasonality of sales, and the diverse nature of the customer base. By implementing the recommended strategies in inventory management, market development, seasonal marketing, and customer relationship management, the business can enhance operational efficiency, foster customer loyalty, and achieve sustainable revenue growth. Continuous monitoring and iterative analysis will be essential to adapt to evolving market dynamics and customer preferences.