Analysis Report By Muhammad Shabbar

shabbaraza26@gmail.com | Github | Linkedin

Online Retail Data Analysis Report

Executive Summary

This analysis explores transactional data from an online retail store from 2010 to 2011, focusing on identifying sales trends, customer behavior patterns, and popular products. The dataset contains information about customer purchases, including product details, quantities, prices, and timestamps. Through exploratory data analysis, we've uncovered key insights that can guide strategic business decisions and enhance the store's performance in the competitive online retail market.

Data Overview

The analysis was performed on a dataset containing the following fields:

- **InvoiceNo**: Invoice number of the transaction
- **StockCode**: Unique code of the product
- **Description**: Description of the product
- **Quantity**: Quantity of the product in the transaction
- **InvoiceDate**: Date and time of the transaction
- **UnitPrice**: Unit price of the product
- **CustomerID**: Unique identifier of the customer
- **Country**: Country where the transaction occurred

Data Preparation and Cleaning

Several data preprocessing steps were taken to ensure data quality:

- Removed rows with missing CustomerID, StockCode, or InvoiceNo
- Filtered out transactions with negative or zero quantities and prices
- Removed duplicate records (5,192 duplicated rows identified)
- Created additional time-based features including Year, Month, and DayOfWeek

Key Findings

1. Sales Trends

Monthly Patterns:

- The highest sales volume occurred in January 2011, with 349,098 units sold
- December 2010 and March 2011 also showed strong sales performance
- A noticeable sales pattern emerged across the year, indicating potential seasonal trends

```
monthly_sales columns: ['Year', 'Month', 'Quantity']
    Year  Month  Quantity
0   2010     12    312265
1   2011      1    349098
2   2011      2    265622
3   2011      3    348503
4   2011      4    292222
```

Monthly Sales Trends by Year
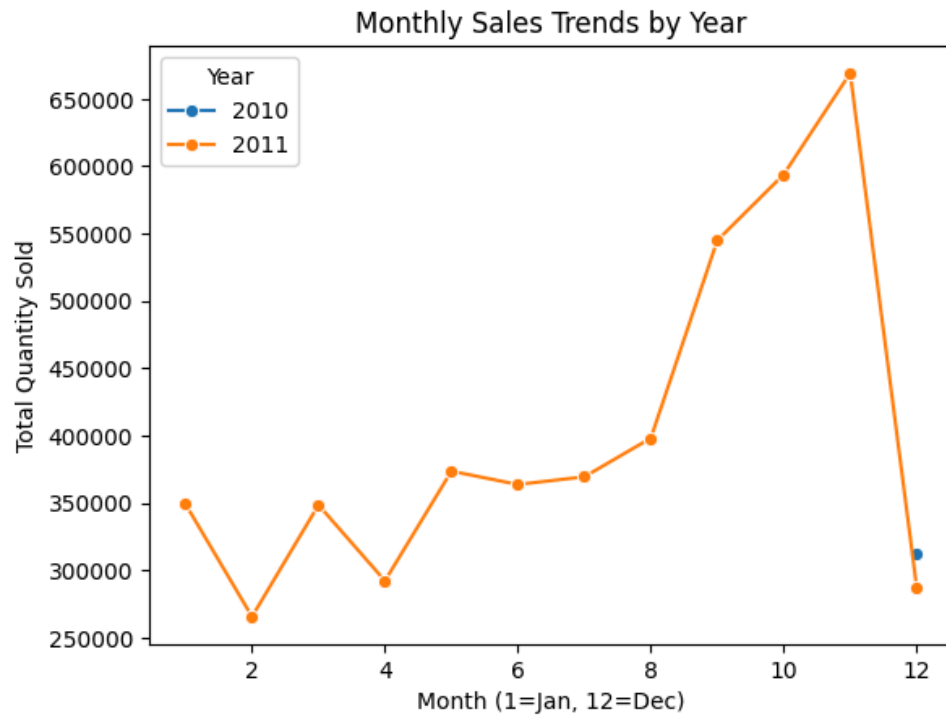
Weekly Patterns:

- Thursdays showed the highest sales volume among all days of the week
- 2,007 unique customers place orders specifically on Thursdays
- This suggests an opportunity for mid-week targeted promotions

Sales by Day of Week
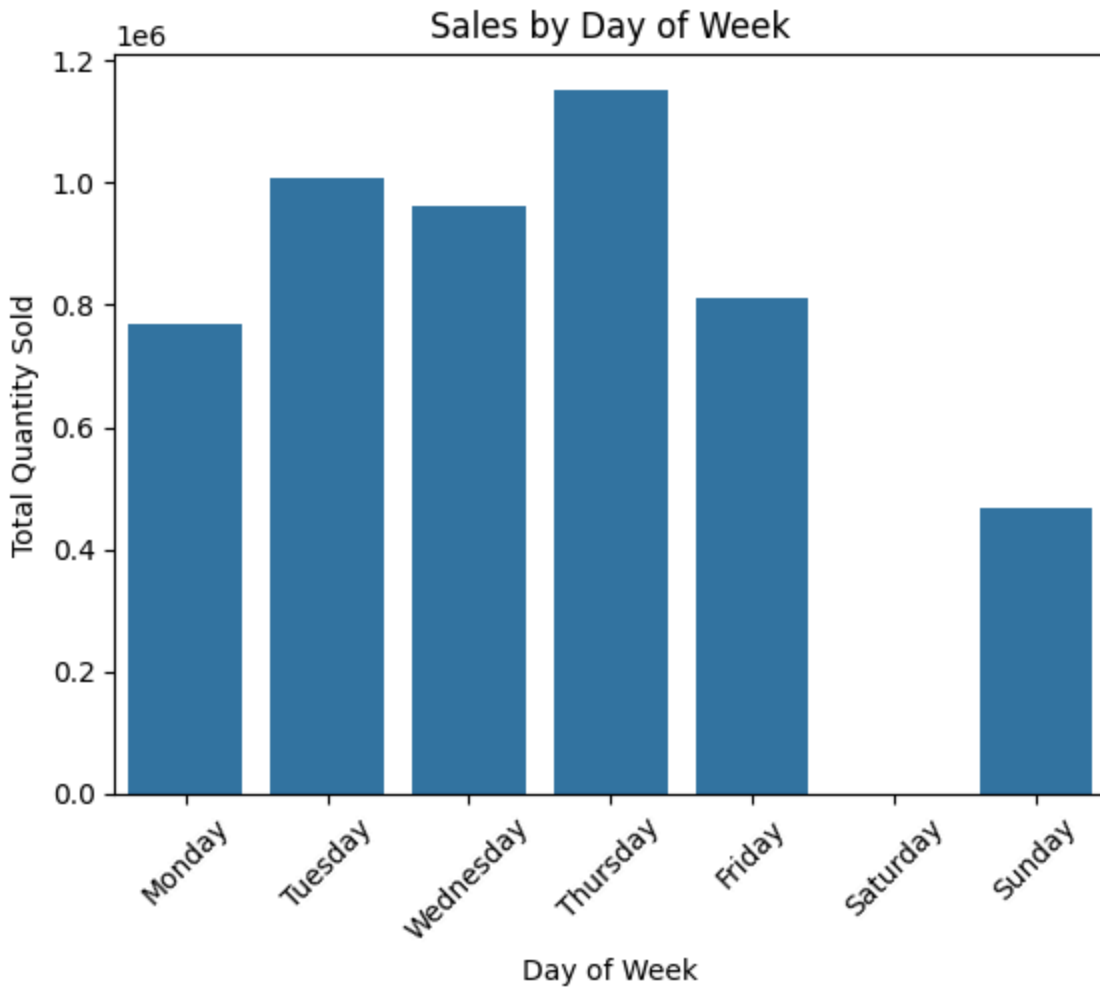
2. Product Analysis

Top-Selling Products:

- "PAPER CRAFT, LITTLE BIRDIE" was the highest-volume product with 80,995 units sold
- "WORLD WAR 2 GLIDERS ASSTD DESIGNS" ranked second with impressive sales
- Gift items and decorative products dominate the top-selling list

December (Holiday Season) Favorites:

- Top December products include "PAPER CRAFT, LITTLE BIRDIE" and "WORLD WAR 2 GLIDERS"
- Seasonal items like "WHITE HANGING HEART T-LIGHT HOLDER" and "DISCO BALL CHRISTMAS DECORATION" perform well in December
- This information is crucial for Q4 inventory planning

Top 10 Products by Quantity Sold
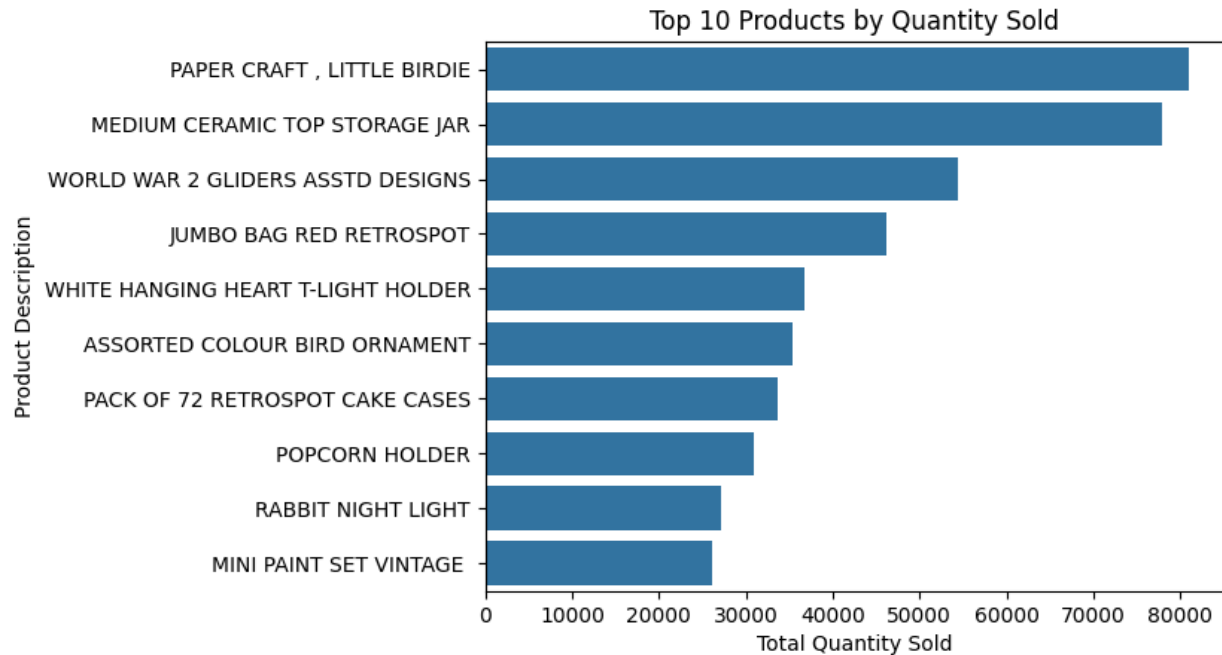
3. Geographic Analysis

Market Distribution:

- United Kingdom is the primary market with significantly higher sales than other countries
- Outside the UK, the Netherlands leads international sales with 200,361 units
- Other promising markets include Ireland (140,275 units), Germany (119,261 units), and France (111,471 units)
- Australia shows strong potential with 83,901 units despite geographic distance

Top 10 Countries by Quantity Sold
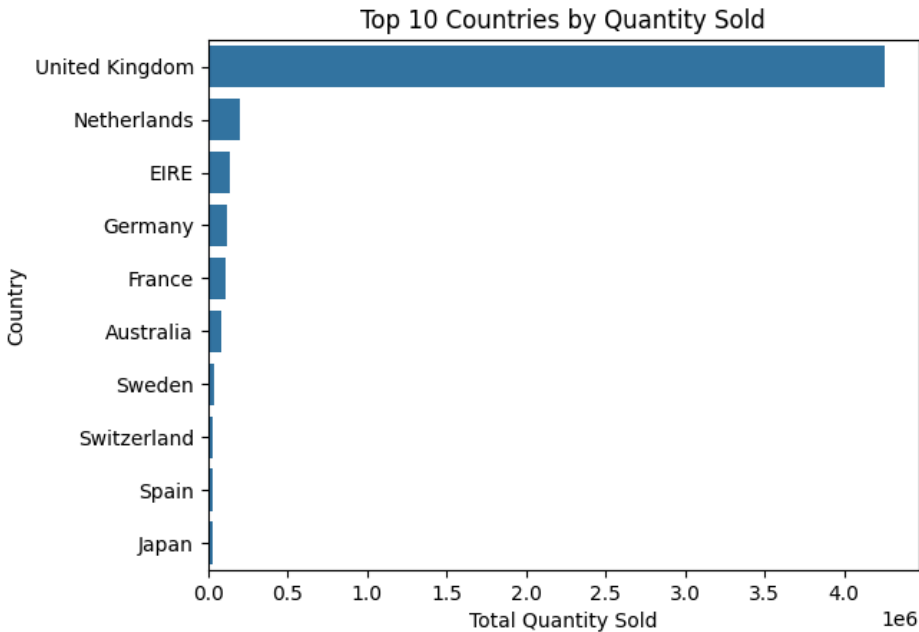
4. Outlier Analysis

- 25,647 transactions were identified as quantity outliers (above the upper bound threshold)
- The largest single transaction was for 80,995 units of "PAPER CRAFT, LITTLE BIRDIE"
- Several very large orders (over 2,000 units) were placed by specific customers
- These outliers may represent wholesale purchases or special orders that should be analyzed separately

```
Top 20 high-quantity transactions for manual review:
        InvoiceNo                      Description  Quantity  CustomerID
540421     581483       PAPER CRAFT , LITTLE BIRDIE     80995     16446.0
61619      541431     MEDIUM CERAMIC TOP STORAGE JAR    74215     12346.0
421632     573008   WORLD WAR 2 GLIDERS ASSTD DESIGNS    4800     12901.0
206121     554868            SMALL POPCORN HOLDER        4300     13135.0
97432      544612           EMPIRE DESIGN ROSETTE        3906     18087.0
270885     560599  ESSENTIAL BALM 3.5g TIN IN ENVELOPE   3186     14609.0
160546     550461   FAIRY CAKE FLANNEL ASSORTED COLOUR   3114     15749.0
52711      540815   FAIRY CAKE FLANNEL ASSORTED COLOUR   3114     15749.0
433788     573995       SMALL CHINESE STYLE SCISSOR      3000     16308.0
291249     562439       ASSORTED COLOUR BIRD ORNAMENT    2880     12931.0
4945       536830   WORLD WAR 2 GLIDERS ASSTD DESIGNS    2880     16754.0
201149     554272   PACK OF 60 PINK PAISLEY CAKE CASES   2700     12901.0
80742      543057   WORLD WAR 2 GLIDERS ASSTD DESIGNS    2592     16333.0
87631      543669   GROW A FLYTRAP OR SUNFLOWER IN TIN   2400     16029.0
421601     573003               RABBIT NIGHT LIGHT       2400     14646.0
91772      544152  ESSENTIAL BALM 3.5g TIN IN ENVELOPE   2400     14609.0
32671      539101   GROW A FLYTRAP OR SUNFLOWER IN TIN   2400     16029.0
125282     547037          PACK OF 12 SKULL TISSUES      2160     14101.0
437480     574294           RED  HARMONICA IN BOX        2100     16333.0
511421     579498               RABBIT NIGHT LIGHT       2040     12798.0
```

Business Insights and Recommendations

1. Sales Strategy

Apply the Pareto Principle:

- Focus resources on top-performing customers and products that generate the majority of revenue
- Develop targeted retention strategies for high-value customers

Leverage Seasonal Trends:

- Implement strategic inventory planning for Q4/holiday season
- Prepare marketing campaigns in advance of peak months (January, March, December)
- Consider running special promotions during slower months to boost sales

Optimize for Day-of-Week Patterns:

- Launch "Thursday Special" flash sales to capitalize on the highest-volume day of the week
- Consider different marketing messages and offers for different days of the week

2. Inventory Management

Prioritize Core Products:

- Ensure consistent availability of top-selling products
- Monitor stock levels closely for items that appear frequently in the top-seller lists
- Consider bundling popular items for promotional offers

Seasonal Planning:

- Increase inventory of seasonal products before anticipated demand peaks
- Develop a seasonal forecasting model based on historical monthly patterns

3. International Expansion

Target High-Potential Markets:

- Develop tailored marketing strategies for the Netherlands, Ireland, Germany, and France
- Consider localization efforts (language, payment methods, shipping options) for these markets
- Investigate the success factors in the Australian market despite distance challenges

4. Data Quality and Business Processes

Address Anomalies:

- Investigate the causes of large-quantity outlier transactions
- Review and improve return/cancellation processes to reduce negative impacts on margins
- Implement better data validation procedures at the point of entry

5. Customer Segmentation

Develop Targeted Campaigns:

- Use RFM (Recency, Frequency, Monetary) analysis to segment customers
- Create personalized marketing strategies for different customer segments
- Focus retention efforts on high-value, repeat customers

Future Analysis Recommendations

1. **Pricing Analysis:** Study the relationship between pricing and sales volume
2. **Basket Analysis:** Identify frequently co-purchased items for cross-selling opportunities
3. **Customer Lifetime Value:** Calculate CLV to better prioritize customer acquisition and retention efforts
4. **Return Rate Analysis:** Investigate patterns in returns to address potential product or service issues
5. **Seasonality Modeling:** Develop predictive models for seasonal demand forecasting

Conclusion

The online retail business demonstrates clear patterns in sales distribution, product popularity, and geographic performance. By focusing on high-performing products, optimizing for day-of-week patterns, and strategically approaching international markets, the company can enhance overall performance. Additionally, addressing data anomalies and implementing customer segmentation strategies will contribute to improved business processes and more targeted marketing efforts.

The data reveals both challenges and opportunities. While sales are concentrated in specific products and geographies, there is significant potential for growth through strategic inventory management, targeted promotions, and international expansion. By leveraging these insights, the online retail store can make data-driven decisions that enhance customer satisfaction and drive business growth.