

EDA for Retail Sales Analysis

This report aims to answer the business question:

“How can the business improve its decision-making regarding stock management, promotions, and regional targeting when it lacks visibility into which products, customer segments, and regions are driving revenue or losses?”

Using transactional data from `Superstore Retail Dataset`, this Analysis applies cleaning and pre-processing of raw data followed by EDA in Python to explore product performance, profit trends, shipping behaviour, and customer segments. Built a 3-page interactive Power BI dashboard for business intelligence.

This Report focuses on four core areas:

1. **Descriptive Statistics for both Numerical and Categorical features**
2. **Time Series Analysis**
3. **Customer Segmentation Analysis**
4. **Product Performance Analysis**

These insights inform data-driven recommendations to support customer retention, increase marketing ROI, and align promotional strategies with actual buyer behaviour.

Executive Summary

This analysis explores retail sales data from 2011 to 2014 to uncover customer behaviour, sales patterns, and product performance. The goal is to provide business-ready insights that inform marketing, inventory, and customer segmentation strategies.

Key Insights:

- **Seasonal Trends:** Sales peak consistently in **Q4**, suggesting a need to **intensify promotions in Q3**.
- **Customer Distribution:** The **top 15% of customers** contribute to over **65% of revenue**, indicating a strong opportunity for **VIP loyalty programs**.
- **Product Performance:** Sub-categories like **“Phones”** and **“Chairs”** are top performers in both sales and profit.

- **Inventory Risk:** Some products show **high quantity but low sales**, signalling potential overstock or low demand.
- **Category Performance: Office Supplies** has high sales volume but low profit margins — consider reviewing pricing or discounting strategy.

Recommendations:

- **Increase marketing spend in Q3** to maximize Q4 peak season.
- Reassess **inventory strategy** for slow-moving products.
- Consider **re-pricing or bundling** low-margin sub-categories.
- Implementing a tiered **VIP loyalty program** is a strong recommendation with the schemes like **Exclusive Access, Personalized Offers, Personalized Communication, Track of Key Metrics**.

1. Descriptive Statistics for both Numerical and Categorical features

Key Insights:

- Consumer segment dominates sales volume.
- Most orders are shipped via Standard Class.
- Office Supplies is the most frequent product category.
- West region has the highest transaction volume.

Recommendations:

- While **Consumers dominate**, explore Corporate/Business segment opportunities through targeted offers.
- Audit **Standard Class** performance and costs. Explore ways to shift customers to more cost-efficient modes.
- Promote high-margin products in frequently ordered categories like **Office Supplies**.
- Use regional data to expand into less saturated markets like **Central** and **South** with focused marketing or regional pricing strategies.

2. Time Series Analysis

Key Insights:

- Sales consistently **peak in Q4** (October to December), likely due to seasonal promotions or holidays.
- The trend shows **year-over-year** growth, especially between **2012 and 2014**.
- Minor dips in mid-year months (May–July) suggest opportunities to boost promotions in those periods.
- Slower average performance in months like **July and September**.
- **Technology** shows the most **consistent and high sales**, especially in Q4

Recommendations:

- **Increase marketing** spend in **Q3** to capitalize on predictable Q4 spikes.
- Consider adding **campaigns** or **discounts** in slower months to lift the baseline.
- Prioritize **inventory and marketing focus on Technology** during Q4.
- **Monitor Office Supplies** trends for school/business seasonality to time promotions better.

3. Customer Segmentation Analysis

Key Insights:

- The **Consumer segment** contributes the **highest number of orders and sales**, indicating it's the primary revenue driver.
- **Corporate** and **Home Office** segments have **lower volume** but may represent higher-value transactions per order.
- Revenue per order may differ across segments — worth deeper profitability analysis.
- The segment mix may vary across regions or product categories, offering room for personalization.

Recommendations:

- **Consumer Loyalty Programs:** Introduce a reward or referral system to maintain dominance in this segment.
- **Corporate Upselling:** Bundle offers and enterprise packages can unlock greater value from corporate clients.
- **Nurture Home Office Segment:** Target this underutilized group with flexible pricing or starter kits for small businesses.
- **Segmented Marketing Campaigns:** Run personalized marketing (e.g., email campaigns) by segment to improve engagement and retention.

4. Product Performance Analysis

Key Insights:

- **Office Supplies** category has the **highest number of orders** but relatively **lower profit margins**.
- **Technology** products (e.g., phones, copiers) bring **higher revenue and profitability** despite lower volume.
- Certain **sub-categories like Binders and Paper** are frequently sold but may yield **low or negative margins**.

Recommendations:

- **Focus on High-Margin Tech Products:** Prioritize and promote top-performing tech items (e.g., phones, accessories) through bundling or featured placement.

- **Bundle Low-Profit Items:** Combine items like paper and binders with higher-value tech products to increase average order profitability.
- **Seasonal Product Strategy:** If certain products peak at specific times (e.g., office chairs), plan campaigns or procurement accordingly.

Tools Used:

- Python
- Pandas
- Matplotlib
- Seaborn
- Jupyter Notebook
- Power BI

Interested in a similar dashboard and Analysis for your business?

Let's talk: shubh.pal753@gmail.com