Shopeasy: Retail Sales & Inventory Optimization Report (2024)

Executive Summary

This report details a comprehensive analysis of Shopeasy's 2024 sales and inventory data. The primary objective was to analyze historical performance, identify key business drivers and visualize performance indicators to address challenges in sales fluctuations and inventory management.

The analysis involved three stages:

1. **Data Processing:** Cleaning and enriching raw sales data using Python (Jupyter Notebook).
2. **Database Analysis:** Loading the prepared data into a MySQL database for complex, SQL-based querying.
3. **Visualization:** Building an interactive Power BI dashboard to monitor

1. Introduction: The Business Problem

Shopeasy, a retail chain with stores across India, is facing significant operational challenges, including fluctuating sales, inventory overstocking and lost revenue due to seasonal demand changes.

This project was commissioned to perform an in-depth analysis of their 2024 performance. The goals were to:

- Analyze historical sales, inventory, and promotional data.
- Answer key business questions regarding performance by store, product, and region.
- Visualize these findings in an interactive dashboard to empower data-driven decision-making.

2. Data Analysis & Visualization Workflow

This section details the end-to-end process of transforming raw data into actionable insights.

2.1 Data Processing (Jupyter Notebook)

The foundation of the analysis was built in a Jupyter Notebook using Python, pandas, and SQLAlchemy.

- **Data Loading:** The ShopEase_Sales_Inventory_2024.csv dataset (36,600 rows, 15 columns) was loaded.
- **Data Cleaning:** An initial check confirmed **zero null values**, ensuring data integrity. Column names were standardized (e.g., units_sold) for database compatibility.
- **Feature Engineering:** To add analytical depth, two new columns were created:

    - **inventory_used:** A categorical feature ('low', 'medium', 'high') created by binning the inventory_available column. This allows for high-level analysis of sales performance based on stock levels.
    - **seasons:** A column (Winter, Spring, Summer, Autumn) was mapped from the date column to facilitate seasonal analysis.
- **Database Export:** The final, cleaned DataFrame was exported to a MySQL database named shopeasy_updated into a table called mytable using SQLAlchemy.

2.2 Database Analysis (MySQL)

The centralized mytable in MySQL was queried to answer key business questions from management:

- **Top Performers:** Identified top-performing store_id and product_category by aggregating SUM(units_sold).
- **Geospatial & Seasonal Trends:** Analyzed sales patterns by city, month, and seasons to reveal hotspots and seasonal peaks.
- **Inventory vs. Sales:** Calculated an **Inventory_to_Sales_Ratio** (AVG(inventory_available) / AVG(units_sold)) for each product. This metric pinpoints:
    - **Overstocked items:** High ratio (high inventory, low sales).
    - **Understocked items:** Low ratio (high sales, low inventory), posing a risk of stock-outs.
- **Growth & Decline:** Used a Common Table Expression (CTE) to compare sales in the first half of the year vs. the second half, calculating a growth_percentage for each product.
- **External Factors:** Correlated sales data against holiday status, promotion status, and temperature buckets to quantify their impact on total_sales

2.3 Business Intelligence Dashboard (Power BI)

A "Shopeasy Dashboard" was built in Power BI, connected directly to the MySQL database, to provide an interactive and at-a-glance view of performance.

- **Key Performance Indicators (KPIs) for 2024:**
    - **Total Sales:** 767.95M
    - **Units Sold:** 732K
    - **Units Established:** 38.40M
- **Key Visual Insights:**
    - **Total Sales by Inventory Used:** This chart provides the most critical insight of the analysis. Sales are highest for products in the **'medium'** inventory category and drop significantly for items in the **'low'** category. This is strong evidence that **stock-outs are causing a direct and significant loss of revenue.**

- **Unit Price Trends:** A line chart tracks unit price fluctuations throughout 2024, showing peaks in March and November, likely corresponding to seasonal product launches or pricing strategies.
- **Store Performance:** Donut charts provide a clear breakdown of sales contribution by store, allowing for easy comparison of store-level performance.

3. Key Findings & Strategic Recommendations

This comprehensive analysis has produced several actionable insights for Shopeasy.

- **Finding 1:** There is a direct, negative impact of "low" inventory on sales. The Power BI dashboard confirms that stock-outs are a primary source of lost revenue, as the "medium" inventory bracket generates the most sales.
- **Recommendation 1: Prioritize maintaining a "medium" stock level** for high-performing products. Use the "low" inventory category as a critical alert to trigger re-ordering before sales are lost.
- **Finding 2:** Sales are clearly influenced by external factors such as seasons, holidays, and promotions, as identified in the SQL analysis.
- **Recommendation 2: Align marketing and inventory strategies.** The seasons and holiday data should be used to plan promotions and stock levels *in advance* of these high-demand periods to maximize sales.
- **Finding 3:** The Inventory_to_Sales_Ratio calculated in MySQL is highly effective at identifying inventory inefficiencies.
- **Recommendation 3: Create two automated inventory reports** based on the SQL query:
  - **"Top 10 Overstocked Products":** Flag these items (highest ratio) for a targeted promotion or discount to clear inventory and free up capital.
  - **"Top 10 At-Risk Products":** Flag high-selling items (lowest ratio) to prioritize their re-stocking and prevent stock-outs.
- **Finding 4:** Sales performance varies significantly by city and store, as shown in the Power BI dashboard.
- **Recommendation 4: Adopt a localized inventory strategy.** Do not use a single, chain-wide stocking policy. Managers should use the city and store_id filters in the dashboard to make stocking decisions based on their specific store's performance and regional demand.