

Coffee Shop Sales Analysis: Trends, Insights, and Predictive Modeling

📍 Executive Summary

This analysis dives into the sales data from a coffee shop chain between January and June 2023. It focuses on identifying trends, customer behaviors, and opportunities for improving sales. Using data cleaning, exploratory analysis, and predictive modeling, we uncover key insights to help make informed business decisions.

📊 Dataset Overview

- **Total Transactions:** 149,116
- **Key Features:** Product, price, quantity, category, store location, and timestamps
- **Time Period:** January–June 2023
- **Analysis Tools:** Python (Pandas, Matplotlib, Seaborn), Jupyter Notebook, and scikit-learn for predictive modeling

📊 Key Insights

📋 Sales Composition

- **Coffee Products:** Coffee leads in both revenue and quantity sold, making it the top performer.
- **Secondary Products:** Tea and Drinking Chocolate contribute significantly to sales, with premium items like **Gourmet Coffee** being a key favorite.
- **Revenue Drivers:** Higher-priced items (like **Drinking Chocolate**) and frequent small purchases drive total revenue.

Sales Trends (Time-Based Insights)

- **Morning Rush:** Peak sales occur during **7–9 AM**, driven by coffee sales.
- **Weekdays vs. Weekends:** **Weekdays** (especially Tuesday–Thursday) have more sales than weekends, with a drop on Mondays.
- **Seasonal Variations:** Higher sales in the winter months, suggesting a preference for hot beverages during colder weather.

Product Insights

- **Top Sellers:** **Gourmet Coffee** and specialty drinks like **Spicy Eye Opener Chai** consistently lead in revenue.
- **Price Sensitivity:** High-price items like **Drinking Chocolate** bring in more revenue despite lower transaction volumes.

Store Performance

- **Top Locations:** **Lower Manhattan** and **Hell's Kitchen** show the highest sales.
- **Location-Based Preferences:** Different stores prefer different product mixes, based on location (e.g., coffee-heavy in some areas, tea in others).

Customer Behavior

- **Peak Hours:** Customers buy most in the morning, and coffee is the most popular choice.
- **Combo Purchases:** Many customers buy coffee with other items, like bakery goods or additional flavors, indicating a chance to offer combo deals.

Methodology

Data Cleaning & Preparation

- Merged date and time columns into a single `transaction_datetime` field for accurate time-based analysis.
- Extracted features like `hour`, `weekday`, and `month` to analyze sales by time.
- Standardized product names and store locations to ensure consistency.



Exploratory Data Analysis (EDA)

- **Sales Trends:** Analyzed daily and monthly sales to identify peak times and seasonal patterns.
- **Product Breakdown:** Coffee dominates sales, with premium items showing a distinct contribution.
- **Hourly Sales Patterns:** Identified the early morning rush hours as the busiest time.



Predictive Modeling

- Built a **regression model** using scikit-learn to predict future sales trends.
 - Measured model accuracy using **R²**, **MAE**, and **RMSE** metrics, and compared predicted vs. actual sales in a scatter plot.
-



Visualizations Overview

Key Visual Insights:

- **Sales Trends Over Time:** A line chart showing daily or monthly revenue trends.
 - **Product Breakdown:** A bar or pie chart illustrating which products bring in the most sales.
 - **Peak Hours:** A histogram or line chart displaying transactions by hour, highlighting busy times (7–9 AM).
 - **Store Performance:** A heatmap or bar chart comparing sales across different store locations.
 - **Profit vs. Cost:** A scatter plot showing the relationship between product profitability and costs.
-



Business Applications



Inventory Optimization

- Focus on stocking more **gourmet coffee** and premium items during peak morning hours.
- Adjust product availability by store, ensuring popular items are always in stock, especially in high-performing stores like **Lower Manhattan** and **Hell's Kitchen**.

Staff Scheduling

- Increase staff during peak hours (7–9 AM) and ensure stores with high foot traffic have more employees.
- Align staff expertise to customer preferences for coffee and related items during busy hours.

Marketing Strategies

- Promote **tea** and **drinking chocolate** in the afternoons to boost sales during slower periods.
- Offer combo deals like coffee with bakery items to increase the value of each transaction.

Revenue Growth Opportunities

- Expand product lines, particularly around **Ethiopian coffee** and other popular premium coffee choices.
- Implement **loyalty programs** to reward regular customers, encouraging repeat purchases.

Operational Improvements

- Streamline morning operations to handle the **peak demand** efficiently.
 - Consider **store layout changes** to feature high-demand products more prominently and ensure a smoother customer experience.
-

Interactive Dashboard (Plotly)

An interactive dashboard lets stakeholders explore the insights visually:

- **Sales by Product Category:** Easily view the share of each product category in total sales.
- **Sales by Location:** See which stores are performing the best.
- **Top Selling Products:** Quickly identify the best-performing items.
- **Monthly Sales Trend:** Track revenue growth or dips month by month.

KPI Dashboard

