Exploratory Data Analysis (EDA) Report

Project Title: Coffee Sales
Author: Kayla Boyce
Date: 27 Dec, 2025

1. Introduction

This EDA will delve into my thinking process in how transactional coffee shop sales data was examined. I looked at the data to understand revenue patterns, product performance, time-based demand trends, and customer purchasing behaviour.

The prime objective of this analysis is to identify the key revenue drivers and operational insights that could support decision-making in terms of inventory planning, staffing, and payment method strategies.
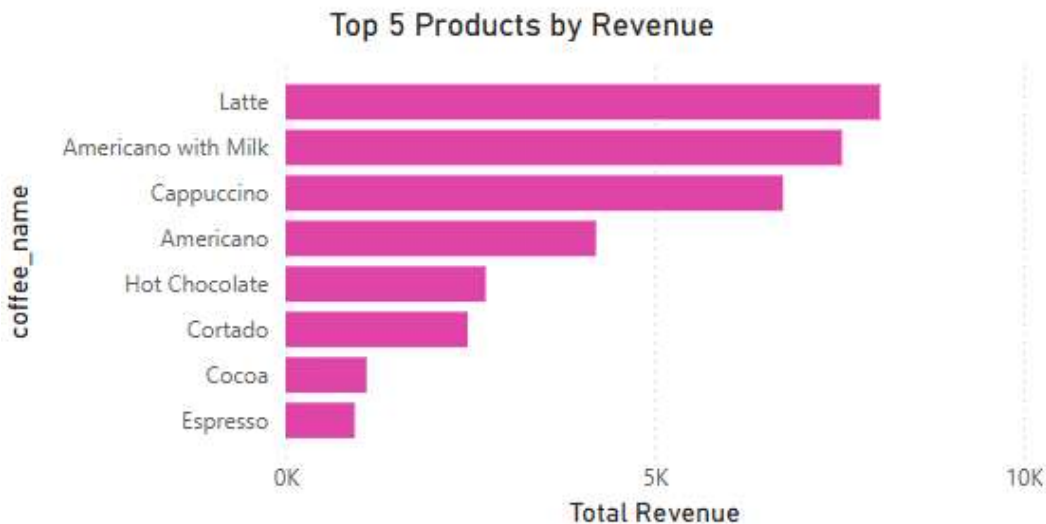
2. Data Overview

The dataset used was sourced from a publicly available Kaggle dataset containing transactional coffee shop sales records. There were approximately 3600 individual transactions in this dataset, with variables such as transaction date and time, product name, payment method, and transactional value. Key variables analysed include coffee product type, hour of day, weekday, payment method, and transaction amount.

3. Data Preparation Summary

Before the analysis stage of this project, the dataset underwent thorough data cleaning and preparation using SQL (MySQL). The monetary values were originally stored as text with currency symbols, however I converted them into numeric format (decimal) to enable accurate aggregation and analysis. Additionally, the date and time fields were validated and standardized, and variables such as product names, weekdays, and payment methods were reviewed to ensure there was consistency.
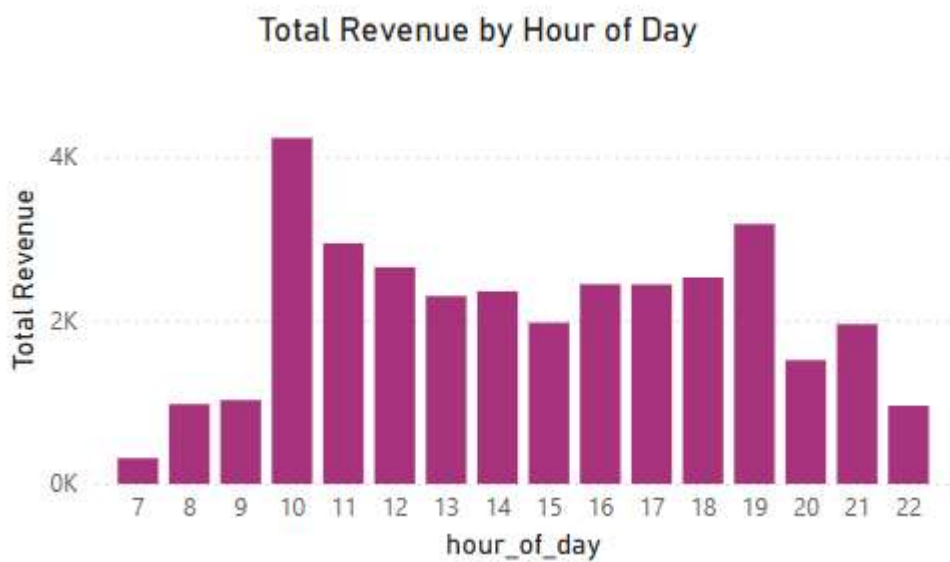
4. Exploratory Data Analysis Findings

4.1 Product Performance Analysis

Top 5 Products by Revenue

Upon reviewing the revenue data of the coffee sales, analysis revealed that a small number of products provide a disproportionate share of total sales. Particularly, Latte, Americano and Cappuccino drinks emerged as the highest revenue-generating items, which indicates a strong customer preference for these drinks.
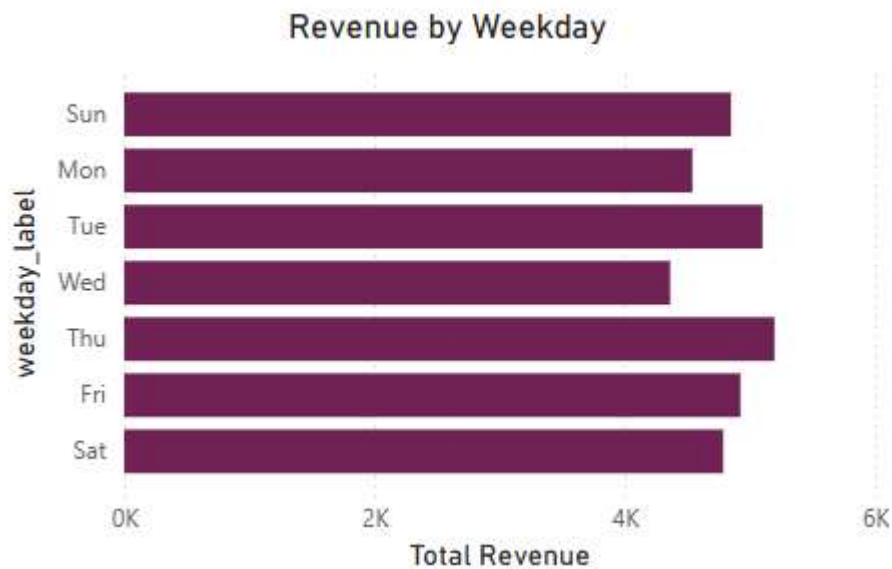
4.2 Time-Based Sales Patterns

Total Revenue by Hour of Day

Following the revenue patterns of the data by the hour of day, it shows that there was a very pronounced peak during the morning hours, particularly at 10 AM, consistent with

commuter-driven demand. Another, smaller increase is noted during the early evening, while later hours show relatively reduced activity.

4.3 Weekly Revenue Distribution

Evaluation of the revenue patterns by the weekday, revenue remained relatively stable across the days of the week with marginally higher performance observed specifically on Tuesdays and Thursdays. This indicates steady customer demand and suggests that sales are not entirely reliant on weekend traffic alone. Such consistency supports reliable staffing, inventory planning, and revenue forecasting.

4.4 Customer Payment Behaviour

My comparison of average transaction value by payment type shows modest differences in customer spending behaviour. Card and cash-based transactions vary in average spend levels, which suggests the payment method may influence purchasing patterns.

4.5 Cross-Indicator Observations

When I examined the findings as a whole, product performance and time-based demand patterns suggest that revenue concentration is driven by both product popularity as well as peak purchasing periods. The morning demand combined with the top-performing products contributes significantly to the overall revenue generation of the coffee business.

5. Key Insights Summary

- High revenue concentration among a small subset of coffee products.
- Morning hours represent a large, significant portion of the revenue.
- Weekly revenue patterns are stable and predictable.
- Customer spending behaviour varies by payment method.

6. Conclusion & Business Implications

The findings derived from this EDA highlight clear opportunities for optimization of operations. The coffee shop operations achieve such optimization by prioritizing top-performing products, aligning staffing with peak demand periods, and encouraging higher-value payment methods.

7. Limitations

The analysis is based off of a static, publicly available dataset and does not cater for seasonal effects, pricing changes or any marketing / promotional activity. Additionally, the findings are limited to observed transactional data and does not include any qualitative data or customer demographic information.

8. References

Kaggle. Coffee Shop Sales Transactions Dataset.