

Coffee Sales Analysis Project

Prepared by : Mikateko Mpapele
Date: 12 October 2025

Abstract

Project Overview

1. Sales Trends Over Time - Monthly Revenue Trends

1. Revenue by Store Location and Units Sold

2. Customer Behavior and Patterns

3. Product Category Insights

Abstract

Sales dataset found on Kaggle, assumed to be of a multistore coffee shop. The dataset consists of 149K distinct orders that have generated about \$712K in total revenue, with nine (9) unique products being offered across three (3) coffee shops. The primary objective of this project is to analyse and track regional store performance in terms of revenue and quantity of goods sold over time.

Project Overview

The key objective of this project is to draw insight from the data to understand which products generate the most revenue, the time intervals at which stores perform best. To provide an in-depth analysis of sales trends across different product offerings. Next, based on these findings, propose recommendations to improve sales.

Key Steps Involved:

1. Data Storage : organize raw sales data into structured CSV format for easier access and analysis, on-hand.
2. ETL Process: Implement an automated ETL pipeline using Hadoop to load data into DataBricks.
3. Data Cleaning and Transformation: Ensure that data is consistent, reliable to draw meaningful insight.
4. Visualization and Reporting: create an insightful visualization to track sales trends and performance.
5. Explanatory Data Analysis (EDA) : to gain deeper insight into sales patterns and business performance.
6. Tech Stack: SQL, Excel and Snowflake
7. Visualization Stack: PowerBI
8. Data Processing: Data cleaning and transformation.
9. Data Storage: Data Insertion and Updation
10. Data Analysis: SQL query for data preparation, data filtering and PowerBI for Dashboards.

Sales Trends Over Time - Monthly Revenue Trends

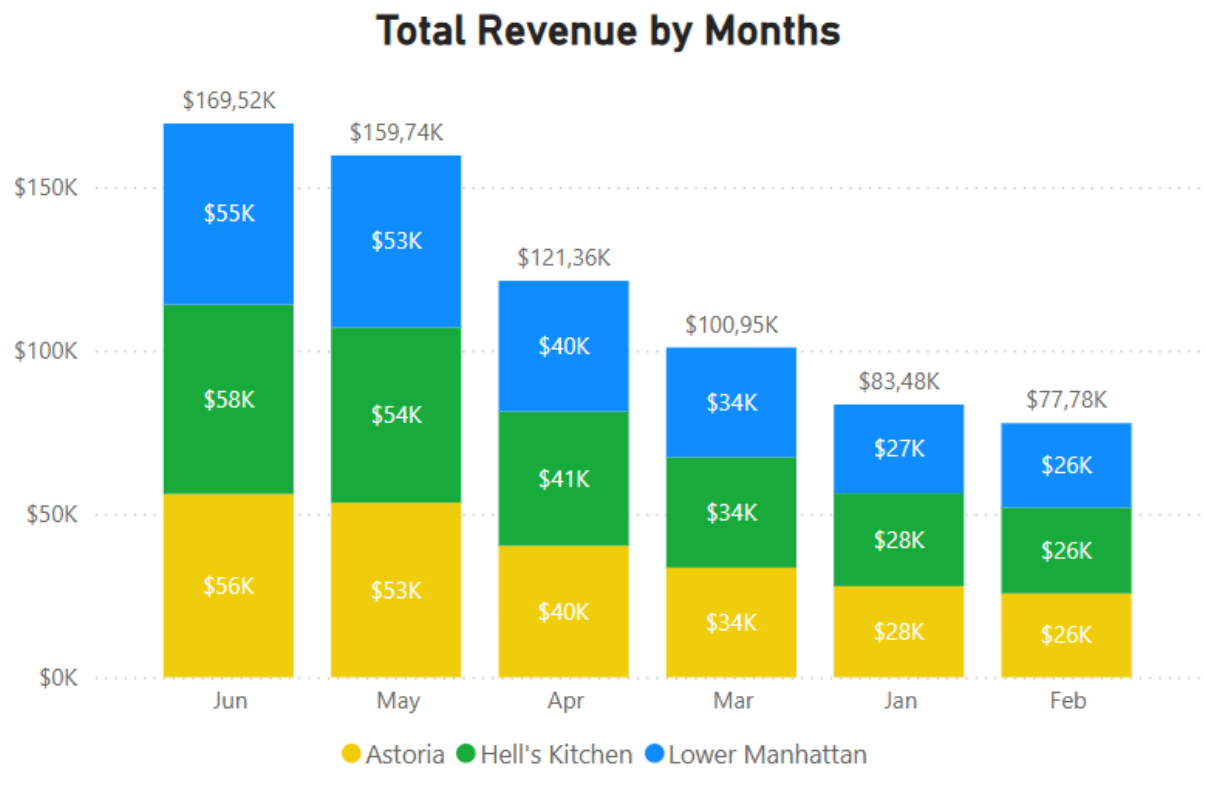


Figure 1: Revenue by Store Location Over Time

- Figure 1 offers a glimpse of how all three coffee shops performed over time, from the beginning of January to the end of June. Sales have grown consistently over the 6 months starting in February. Revenue doubled from January (\$83.48K) to June, sales stand at a staggering \$169.52K. This trend is prevalent among all stores. In Astoria, total revenue across all products was a mere \$26K but jumped to \$56K in June. Credit may be given to significant seasonal demand shifts or targeted and effective promotional efforts at the beginning of the year.

Revenue by Store Location and Units Sold

Revenue Distribution by Store Location

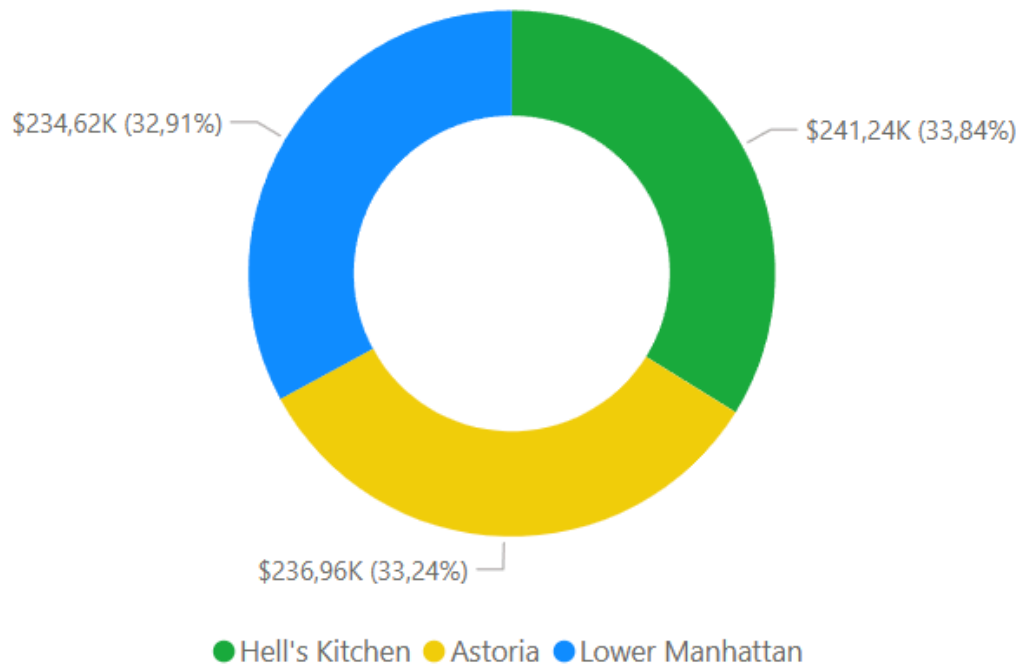


Figure 2: Revenue by Location

Revenue by Quantity Sold

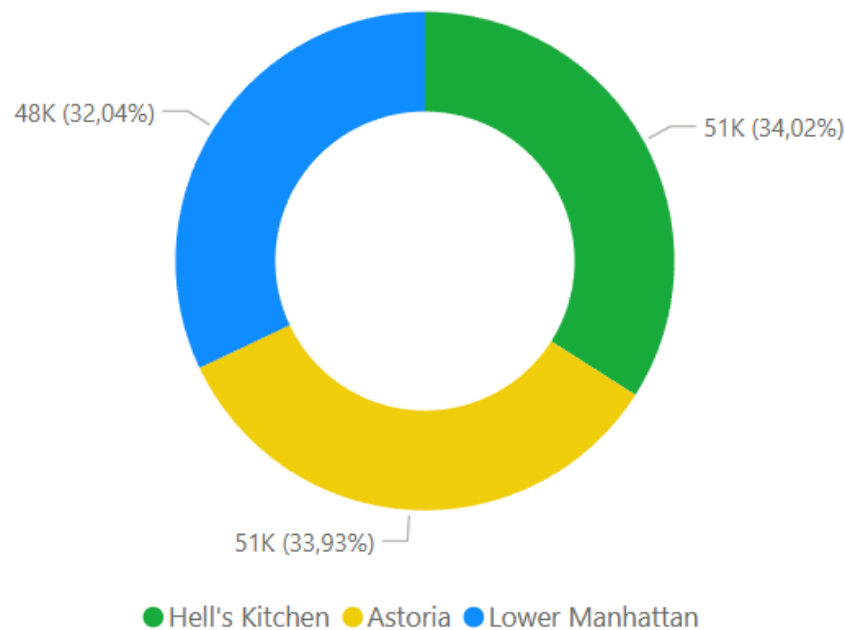


Figure 3: Revenue by Quantity Sold

- Looking at the two figures (Figure 2 and Figure 3) above, we see a detailed distribution of revenue across all coffee shops and the contribution of quantities sold to total revenue. Hell's Kitchen sales represent 34.02% of total revenue (\$712K), while Astoria and Lower Manhattan reflect about 33.93% and 32.04% respectively.
- **Balanced Sales Across Locations:** All stores indicate a similar pattern of sales performance, demand is not heavily concentrated on one location, but remains evenly distributed across all store locations.
- A study of customer behaviour (Figure 4: Revenue by Hour of Day), in terms of their spending patterns, sales begin to spike around 7 am until 10 am, considering this to be peak-time for individual stores, eventually declining and remaining constant for the remainder of the day.

Customer Behavior and Patterns

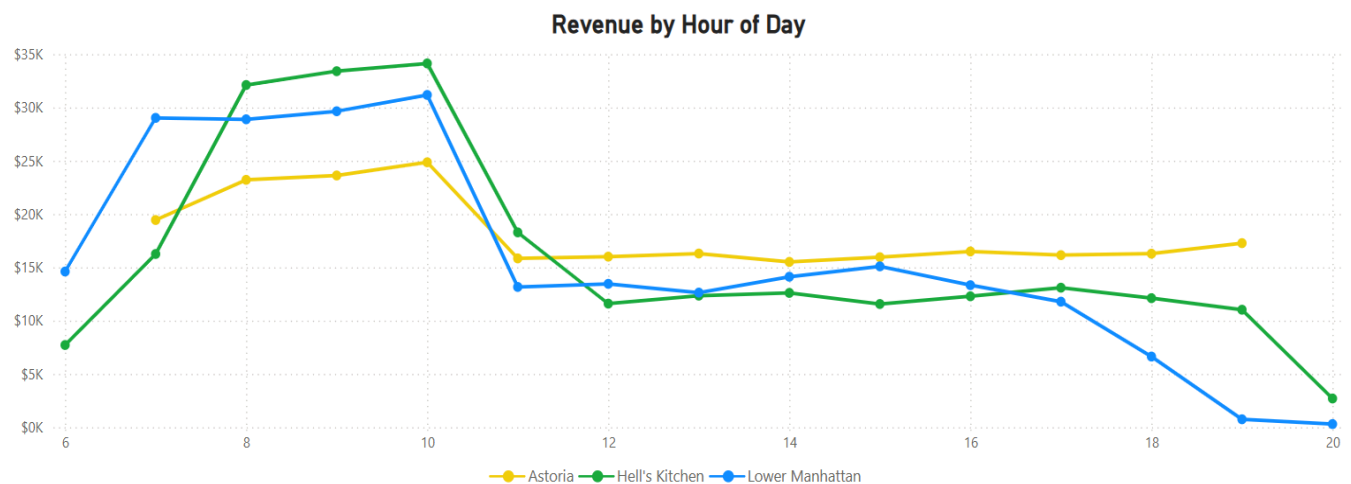


Figure 4: Revenue by Hour of Day

- **Single-item purchasing** rather than bulk purchases are observed across all store locations. A majority of transactions consists of one-to-two items.
- **High Purchase Volumes are Rare:** the purchase of three-to-four items is a rare occurrence, this suggests that customers usually visit the coffeeshop for individual consumption rather than bulk purchases.
- **Potential to Increase Sales through A Customer Loyalty Program, Upsell and bundle:** offer combo deals like a coffee and pastry for a fixed price and/or reward repeat customers with points/discounts after a number of certain purchases.
- **Sales volume surges between 8 AM and 10 AM**, indicating a high morning demand, while **activity slows but remains steady from 11 AM to 7 PM.**

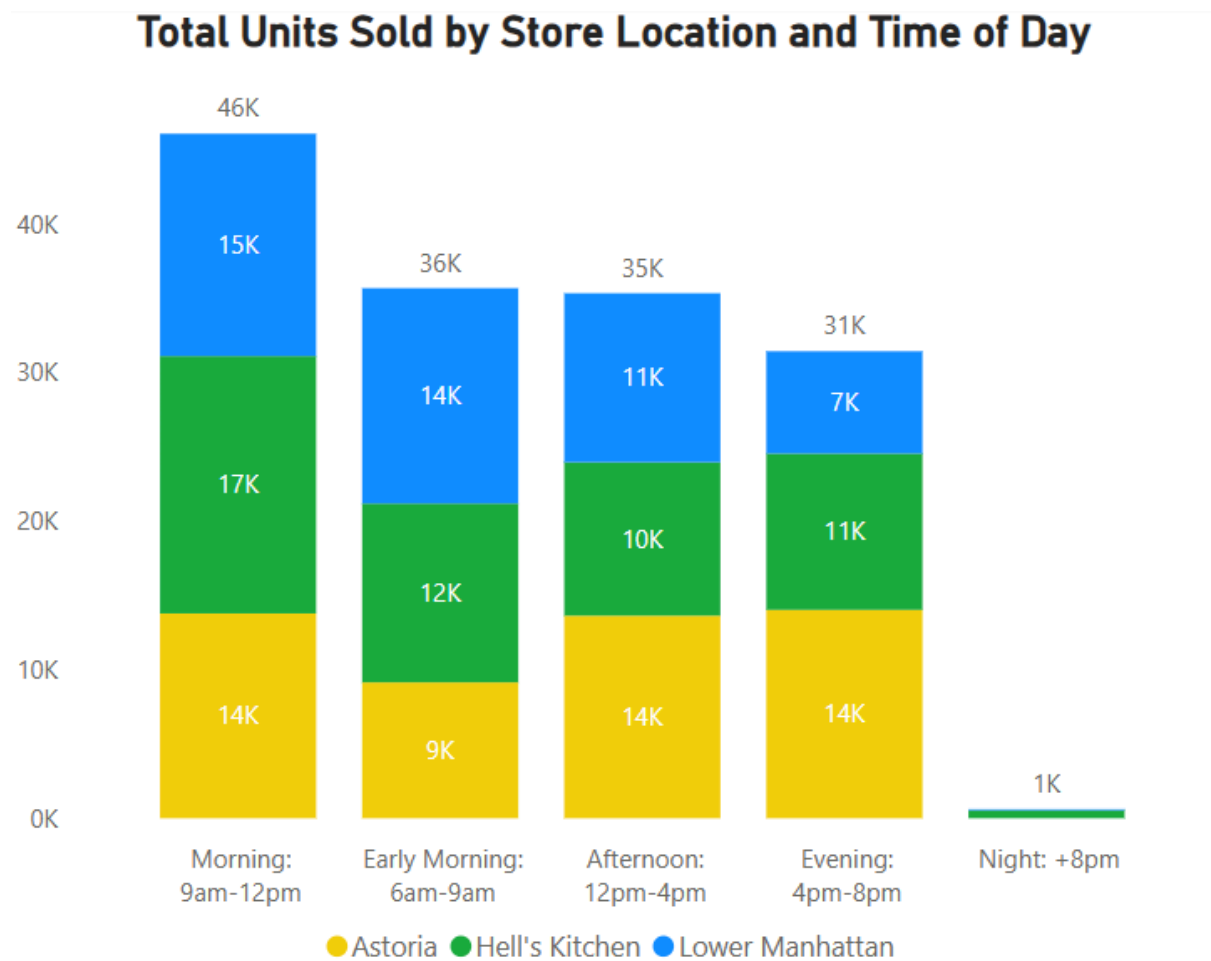


Figure 5: Total Units Sold by Store Location and Time of Day

- Astoria: Sales remain constant without significant drop in sales during the day. Sales rose by 35,71% between 6 am and 12 pm and remained at \$14K by the end of 8 pm.
- Hell's Kitchen: the first significant spike in sales is 41.7%, from \$12K to \$17K, dropping to \$10K (41.18% decrease) and settling at \$11K by evening time.
- Lower Manhattan: Lower Manhattan starts the day at a high, \$14K, compared to other stores, sales decline by 26.7% by 4 pm and closing in at just \$7K, the lowest compared to Astoria and Hell's Kitchen.

Product Category Insights

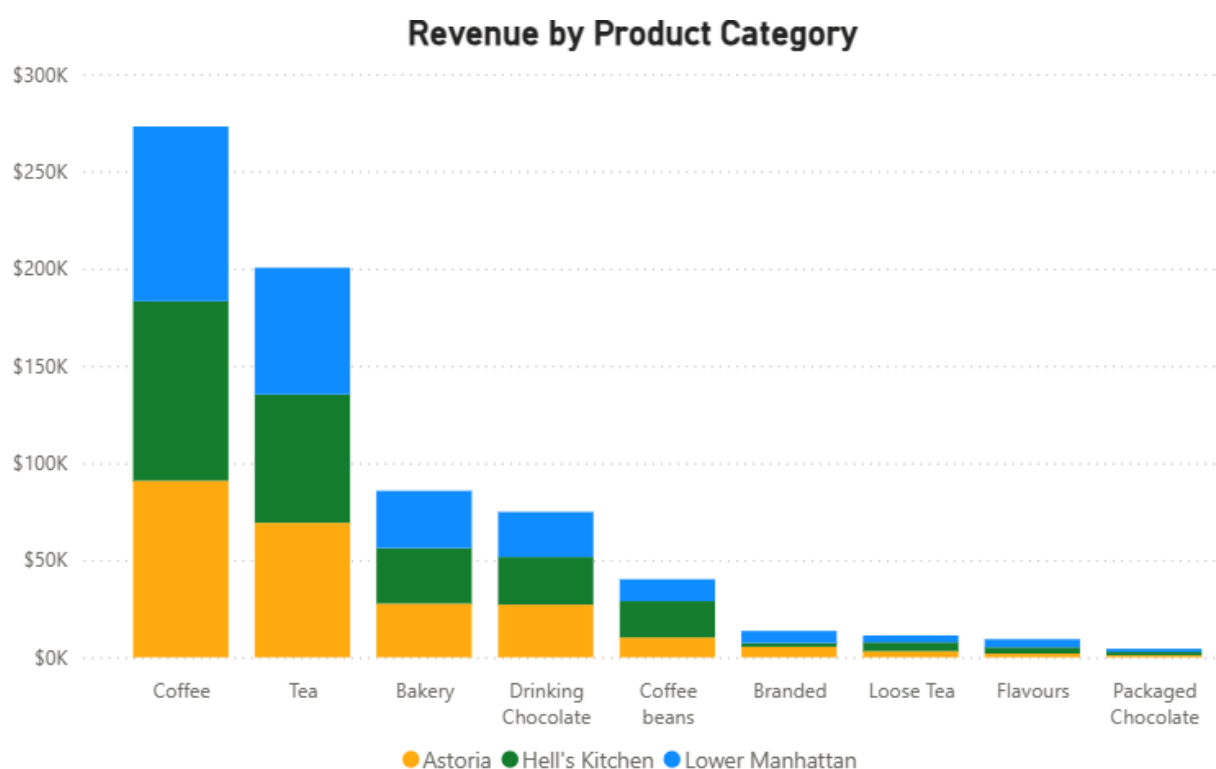


Figure 6: Revenue by Product Category

- **Coffee and Tea** generate the most revenue, accounting for about 70% of sales, making them the **primary revenue drivers**.
- **High-grade products** like loose tea, flavours, packaged chocolate struggle to sell, they are either pocket breakers or rather not appealing to the current customer base.
- **Recommendation:** remove the products off our shelves (loose tea) and sell packaged chocolate as part of a combo deal (coffee, pastry and snack) to drive sales, at a fixed price.