**COFFEE SALES DATA ANALYSIS**

**Introduction**  
This project provides an in-depth analysis of coffee sales data spanning multiple months, focusing on sales volume, revenue, customer distribution, and payment preferences. Leveraging data cleaning, exploratory data analysis, and interactive dashboard development, it aims to support business decision-making and operational improvements in a coffee retail environment.

The dataset, sourced from publicly available records, each record includes rich attributes such as:

* **Date, time, and hour of purchase**
* **Coffee Product Name**
* **Payment Method** (Card or Cash)
* **Customer IDs** (anonymized)
* **Sales Amount**

To transform this raw dataset into actionable intelligence, the project employed a multi- tool, multi-phase methodology involving:

* **Python** (Pandas, NumPy, Matplotlib, Seaborn) for data cleaning, wrangling, and exploratory data analysis.
* **Power BI** for developing interactive dashboards and visual storytelling.

The project’s primary goals include:

1. Identifying sales and revenue performance trends across different coffee categories, time periods, and customer segments.
2. Detecting dominance patterns where specific coffee categories or products lead in sales or revenue within targeted time frames.
3. Creating an interactive and dynamic dashboard to enable non-technical stakeholders to explore sales data easily and make informed business decisions.

By integrating meticulous statistical analysis with modern visualization techniques using Python and Power BI, this project delivers a comprehensive view of coffee sales performance, empowering retail managers, marketing teams, and decision-makers with actionable insights to optimize product offerings, sales strategies, and customer engagement.

**Project** **Objectives**

The primary objectives of this project are to deliver a comprehensive, data-driven understanding of Coffee Sales. Specifically, the project aims to:

1. Analyse coffee sales patterns by category and time segments
2. Assess customer behaviour and order distribution across weekdays and weekends
3. Examine revenue trends and payment method preferences
4. Develop a dynamic, user-friendly dashboard for ongoing sales monitoring and stakeholder insights
5. Provide actionable recommendations to optimize sales strategies and product focus

**Dataset Overview**

The dataset used in this project provides a comprehensive record of coffee sales transactions over several months, capturing detailed customer, product, and temporal information.

* **Time Period Covered:** Multiple months covering various weekdays and weekend dates.
* **Total Records:** Thousands of transaction entries, each representing a single coffee purchase.
* **Number of Features:** 11 well-defined attributes, including:
  1. **Date** – The date of the transaction.
  2. **Time** – Specific time of the purchase.
  3. **Hour** – Hour of the day of the transaction.
  4. **Time of Day** – Categorized into Morning, Afternoon, Evening, and Night.
  5. **Payment Category** – Mode of payment (Card or Cash).
  6. **Customer ID** – Anonymized identifier of the purchaser.
  7. **Money** – The sale amount for each transaction.
  8. **Coffee Name** – The specific coffee product purchased.
  9. **Coffee Category** – Broad category such as Milk-based, Espresso-based, or Chocolate-based.
  10. **Day of Week** – The weekday of the transaction.
  11. **Is Weekend** – Boolean indicating whether the transaction occurred on a weekend.

**Initial Observations:**

* Presence of null or missing values was checked and appropriately handled during cleaning.
* Duplicate transactions and anomalies were scrutinized to maintain data accuracy.
* Categorical attributes were standardized for better analytical consistency.
* Temporal information required conversion and standardization to support time-based trend analysis.

This dataset offers valuable opportunities to analyse sales performance trends across products, customer segments, payment modes, and time periods, supporting actionable business insights in the coffee retail context.

**Data Cleaning & Preprocessing**

A critical step in ensuring the accuracy and reliability of insights was the systematic cleaning and preprocessing of the coffee sales dataset. The following actions were undertaken to prepare the data for exploratory analysis and modelling:

1. **Handling Missing Values**

* Conducted thorough assessment of null and missing entries across all relevant fields.
* Applied appropriate treatment measures including removal of records with irretrievable missing data and imputing values where feasible to maintain dataset completeness.

1. **Removal of Redundant Features**

* Eliminated unnecessary or duplicate columns, streamlining the dataset by removing redundant attributes such as repeated categorical labels.

1. **Data Type Standardization**

* Corrected data type inconsistencies, such as converting date and time fields into standardized datetime formats to support accurate temporal analyses.

1. **Duplicate Record Management**

* Identified and removed duplicate transactions or entries to preserve data integrity and avoid analytical bias.

1. **Null Row Elimination**

* Removed rows containing entirely null values to prevent distortions in statistical summaries and visualizations.

1. **Final Data Verification**

* Performed validation checks confirming that the cleaned dataset was free from critical missing values, structurally consistent, and ready for subsequent exploratory data analysis (EDA) and predictive modelling.

By implementing these rigorous data cleaning and preprocessing steps, the dataset was transformed into a reliable, standardized, and analysis-ready format, ensuring robustness and validity of the descriptive and predictive outcomes.

**Exploratory Data Analysis (EDA)**

The Exploratory Data Analysis phase aimed to uncover patterns, trends, and relationships within the coffee sales dataset through statistical summaries and visualizations. The key areas of analysis included:

1. **Sales Trends Over Time**

* Generated line charts illustrating daily sales and revenue progression across different months and time periods.
* Highlighted fluctuations in sales volumes and revenue at different hours of the day, segmented by coffee category.

1. **Top Coffee Categories and Products**

* Created bar charts ranking coffee categories and individual products by total sales and revenue.
* Analysed dominance of milk-based coffees compared to espresso and chocolate-based ones, identifying consistently high-performing products.

1. **Customer Behaviour Analysis**

* Identified customer order distributions across weekdays versus weekends to understand buying patterns.
* Highlighted repeat customers and their contribution to total sales, emphasizing customer retention metrics.

1. **Payment Method Distribution**

* Produced pie and bar charts to visualize the proportion of transactions by payment type (Card vs Cash) for various time periods.
* Investigated revenue and order count variations influenced by payment mode preferences.

1. **Temporal and Seasonal Variations**

* Assessed sales patterns by time of day (Morning, Afternoon, Evening, Night) and hour to optimize staffing and marketing efforts.
* Evaluated sales differences among weekdays and weekends to guide promotional timing.

This phase provided a solid foundation for understanding the sales dynamics, identifying high-demand products and peak sales periods, and informed the design of an interactive Power BI dashboard to enable ongoing monitoring and strategic decision-making

.

**Power BI Dashboard**

The project included the design and development of an interactive Power BI dashboard that transformed analytical insights into an intuitive, visually appealing interface for stakeholders. Key features of the dashboard design included:

1. **Interactive Filtering Capabilities**

* Integrated slicers and filters for coffee categories, day types (weekday/weekend), payment methods (card/cash), and time periods (Morning, Afternoon, Evening, Night).
* Enabled stakeholders to customize views and focus on specific segments or time frames, enhancing exploratory analysis.

1. **Comprehensive Visualizations**

* Developed clear and engaging charts illustrating coffee sales trends by category, temporal segments, and payment method distributions.
* Showcased customer behaviour insights, popular products, and revenue patterns with easy-to-understand graphics.

1. **Drill-Down and Hierarchical Navigation**

* Implemented drill-down features allowing users to move from aggregate sales metrics to granular details such as individual product sales or specific time intervals.
* Facilitated deep analysis without losing the overall context of sales performance.

1. **Key Performance Indicators (KPIs)**

* Integrated KPI cards highlighting total revenue, total orders, average order value, and unique customer counts.
* Provided instant access to concise, impactful summary metrics that support rapid business decisions.

By combining interactive elements, compelling visuals, and real-time filtering, the Power BI dashboard served as a powerful tool for both technical and non-technical audiences to effortlessly explore coffee sales dynamics. This effectively supports strategic planning, operational monitoring, and informed decision-making in the coffee retail business.

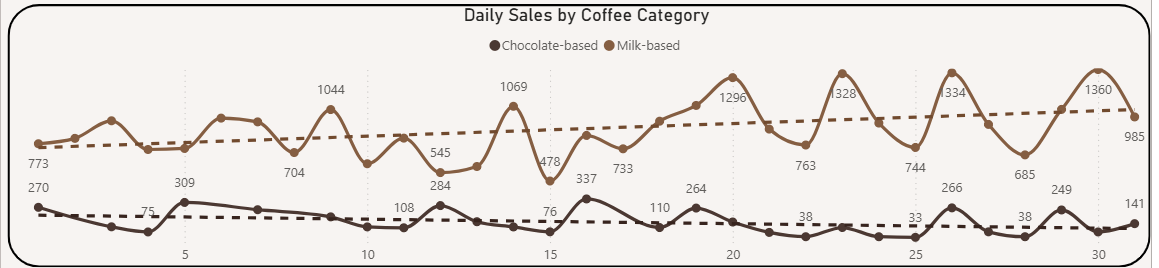
**Key Findings & Patterns**

The dashboard analysis provides a holistic overview of coffee sales performance by visually mapping patterns across product categories, customer behaviors, revenue trends, and payment preferences. By combining interactive data views and statistical summaries, the report uncovers actionable insights that drive strategic business decisions and operational improvements.

The following key findings and patterns were identified from the integrated analysis of sales, customer segments, and temporal order trends

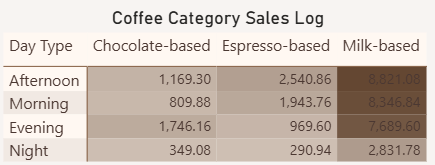
1. **Category Dominance**

* Milk-based coffees consistently achieved the highest sales and revenue throughout the reporting period, significantly outperforming both chocolate-based and espresso-based options.
* Chocolate-based products showed steady but lower sales, while espresso-based coffees demonstrated periodic peaks, especially during specific times of day.



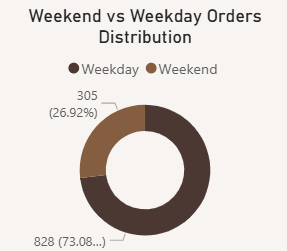
2. **Temporal Sales Trends**

* Sales peaked during the **afternoon**, with revenue figures in this time slot reaching 8.45K for weekdays and 4.08K for weekends, confirming afternoon as the preferred coffee consumption period.
* Morning sales were also strong and competitive, whereas night-time orders were consistently lower across categories.



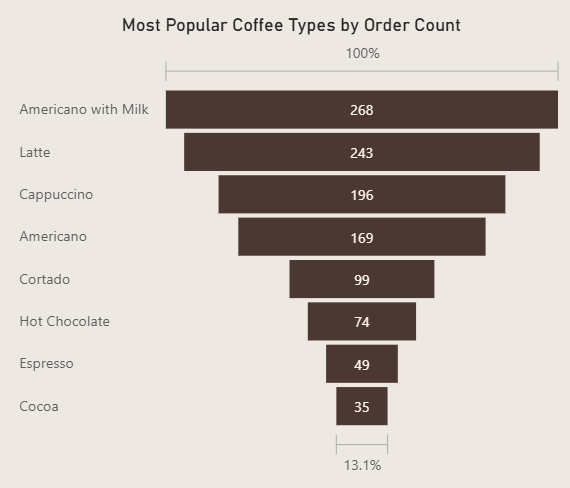
3. **Weekday vs Weekend Patterns**

* **Weekday orders** formed the majority, accounting for 73% of all sales, while weekend orders represented 27%, indicating clear demand differences and regular working-day customer engagement.
* Revenue and order heatmaps revealed that Mondays and Sundays held the highest daily revenue, while Saturday sales slightly declined compared to other days.



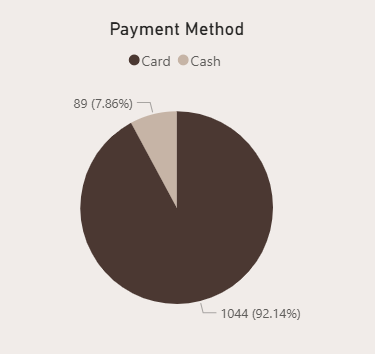
4. **Product Popularity**

* **Americano with Milk, Latte, and Cappuccino** emerged as the most popular coffee options by order quantity, driven by strong customer preference for milk-based drinks.
* Americano and Cortado showed moderate popularity, while Cocoa and Espresso were niche choices.



5. **Payment Method Preference**

* **Card payments** dominated overall sales, particularly during weekdays and busy hours, whereas **cash transactions** increased during weekend evenings and nights, possibly signalling leisure spending or alternate customer segments.

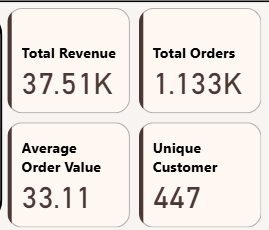


6. **Customer Loyalty & Contribution**

* Repeat customers (447 unique customers) made considerable contributions to sales volume and revenue, with top customers placing as many as 88 orders and generating up to 2,593.18 in revenue.
* The spread in unique customer order frequency highlights opportunities for targeted loyalty initiatives and customer retention strategies.

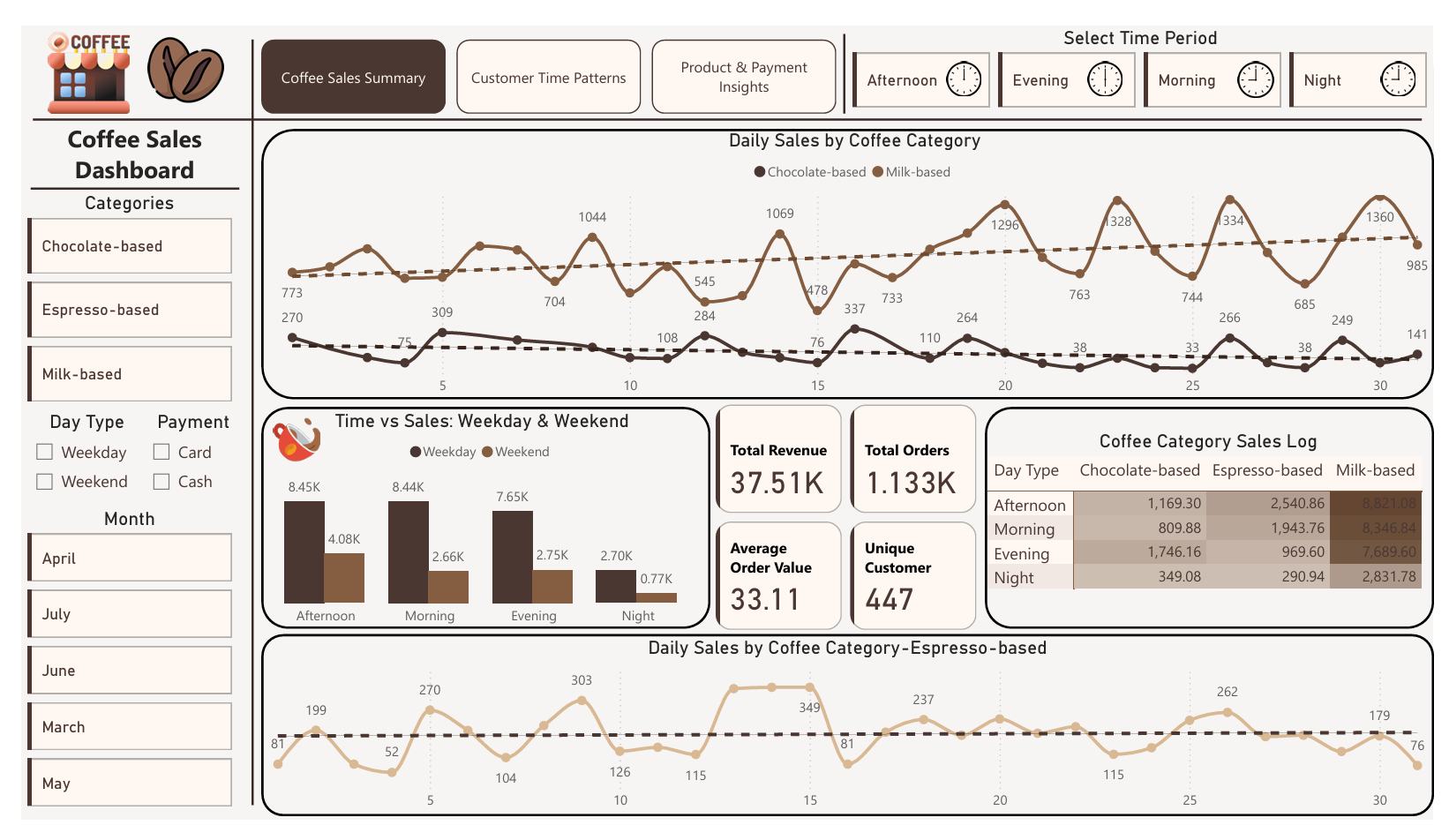
7. **Revenue Insights**

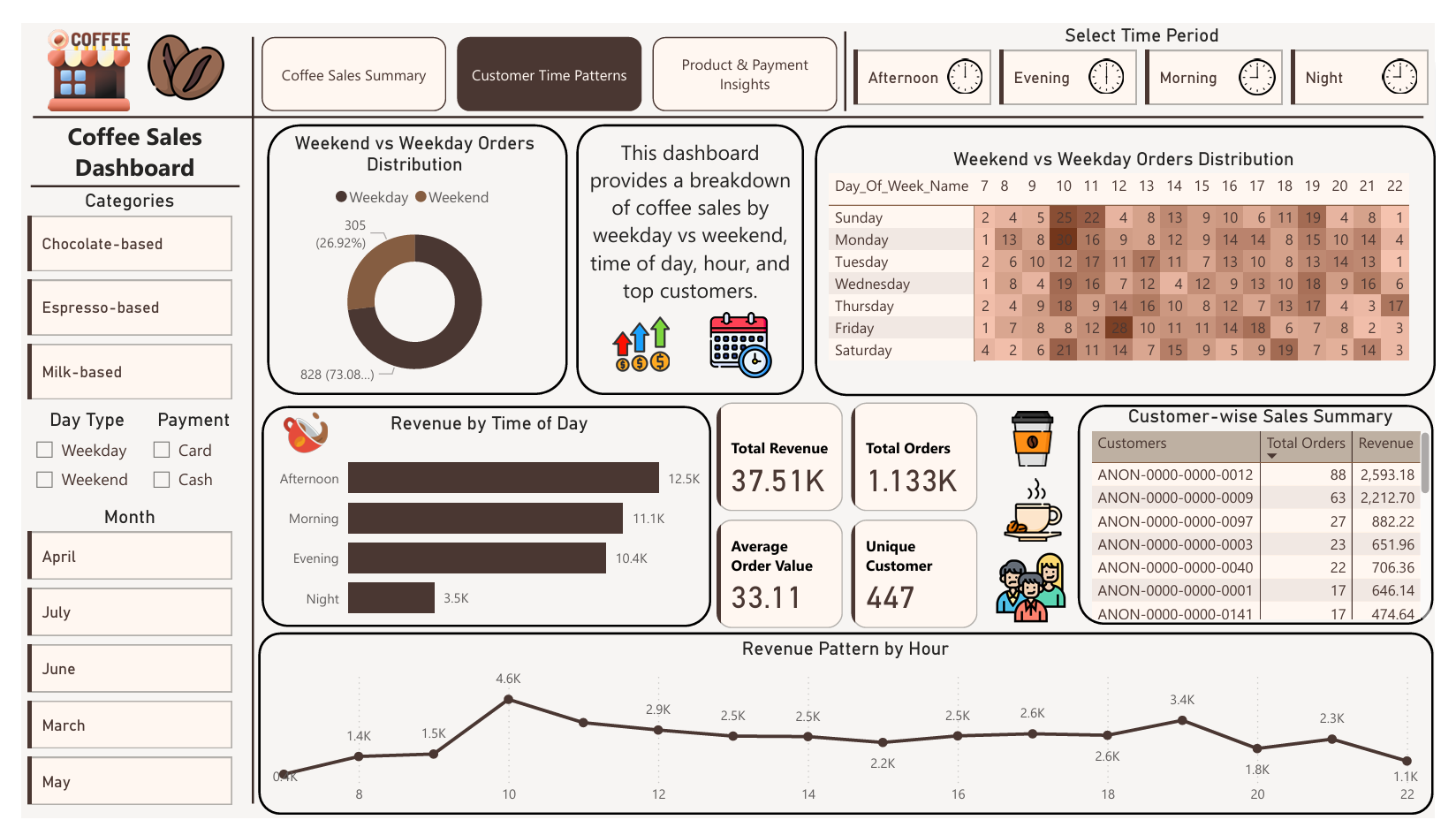
* Total revenue during the period reached **37.51K** over 1,133 total orders, with an average order value of 33.11 units, confirming steady sales turnover and transactional consistency.

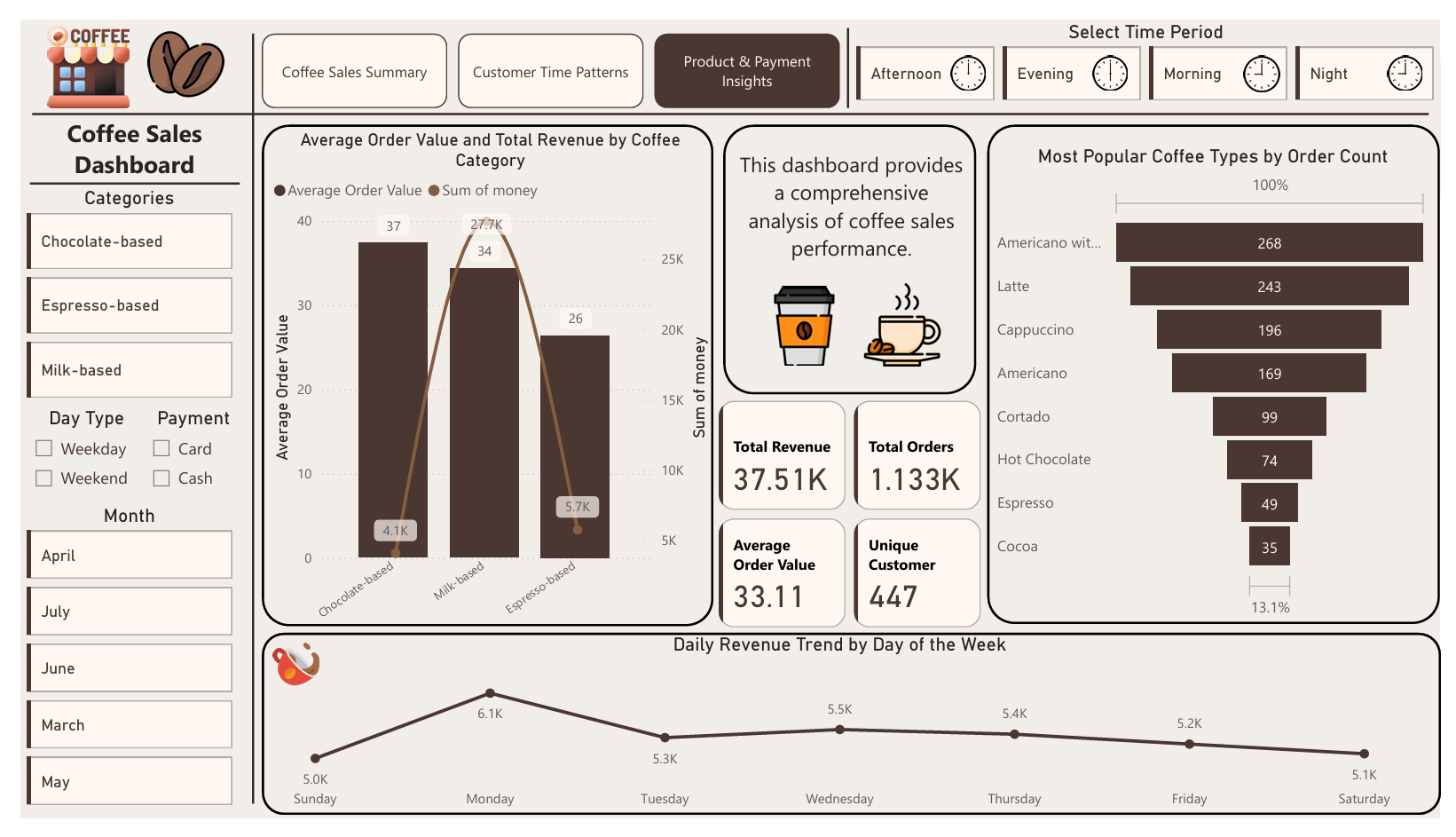


These findings not only validate well-known customer preferences and sales patterns but also provide quantifiable evidence of competitive dynamics within the coffee retail market. They inform strategic planning and operational decisions for marketing, inventory, and customer engagement, supporting efforts to optimize product offerings and enhance revenue growth in an increasingly competitive landscape.

**Dashboard Visualizations**

*Figure 1.* Daily Sales by Coffee Category and Sales Overview Dashboard.

**

*Figure 2.* Customer Time Patterns and Sales Distribution Dashboard.  


*Figure 3.* Product & Payment Insights and Revenue Trends Dashboard.

**Conclusion**

This coffee sales analysis project successfully transformed raw transactional data into actionable business insights through detailed exploratory data analysis and interactive dashboard visualization. The findings highlighted strong sales performance of milk-based coffee products, clear temporal patterns with peak sales occurring in afternoons and mornings, and the dominance of card payments in transactions. Repeat customer behavior was identified as a key driver of sales volume, underscoring opportunities for targeted loyalty programs. The project demonstrates the power of data-driven decision-making in optimizing product offerings, marketing strategies, and operational efficiency, contributing to sustained growth and competitive advantage in the coffee retail sector.

**References**

* The Microsoft Power BI Documentation. “Designing Effective Dashboards.” Microsoft, 2023.
* McKinney, Wes. *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython.* O'Reilly Media, Inc., 2018.
* Kaggle. “Coffee Shop Sales Analysis Dataset.” Kaggle, 2024.
* Grand View Research. “Coffee Market Size, Share & Trends Analysis Report, 2030.” Grand View Research, Dec. 2024.
* Industry Reports on Coffee Retail Trends and Consumer Behaviour, various sources, 2023–2025.
* Microsoft Excel and Python Integration Documentation. Microsoft, 2023.