# Analysis and Forecasting of Coffee Sales Trends, Customer Loyalty, and Discount Impact

**A Proposal report for the BDM Capstone Project**


Submitted By

Name: Sudhanshu Prabhat

Roll Number: 23f1001796




IITM Online BS Degree Program,
Indian Institute of Technology, Madras, Chennai
Tamil Nadu, India, 600036

# Contents

# Declaration Statement

I am working on a Project titled “Analysis and Forecasting of Coffee Sales Trends, Customer Loyalty, and Discount Impact”. I extend my appreciation to Kaggle, for providing the necessary resources that enabled me to conduct my project.

I hereby assert that the data presented and assessed in this project report is genuine and precise to the utmost extent of my knowledge and capabilities. The data has been gathered through secondary sources and carefully analyzed to assure its reliability.

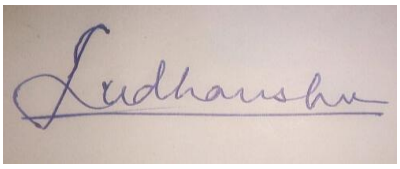
Additionally, I affirm that all procedures employed for the purpose of data collection and analysis have been duly explained in this report. The outcomes and inferences derived from the data are an accurate depiction of the findings acquired through analytical procedures.

I am dedicated to adhering to the principles of academic honesty and integrity, and I am receptive to any additional examination or validation of the data contained in this project report.

I understand that the execution of this project is intended for individual completion and is not to be undertaken collectively. I thus affirm that I am not engaged in any form of collaboration with other individuals, and that all the work undertaken has been solely conducted by me. In the event that plagiarism is detected in the report at any stage of the project's completion, I am fully aware and prepared to accept disciplinary measures imposed by the relevant authority.

I understand that all recommendations made in this project report are within the context of the academic project taken up towards course fulfillment in the BS Degree Program offered by IIT Madras. The institution does not endorse any of the claims or comments.

Signature of Candidate:



Name: Sudhanshu Prabhat

Date: Sep 30th, 2025

# 1 Executive Summary

QahwaCo ( **HYPOTHETICAL** ) is a coffee retailing company operating in Saudi Arabia, where they aim to provide premium specialty coffee beans to their customers. Since the start of the decade, there has been a rise in coffee enthusiasts shifting to speciality coffee beans over regular coffee powder for their unique tastes. Observing this a group of friends decided to turn their passion for coffee into a business venture, and that's how QahwaCo started.

The rising popularity of speciality coffee has led to coffee companies and retailers trying to figure out how different coffee bean varieties align with their customer base to tailor their sales to build recurring consumers. Coffee companies need insights into when their busy season is likely to occur and what factors cause repeat purchase rates; they also need to figure out to what degree the promotion or discounts are significant in driving customer behaviour and sales. Understanding these patterns are useful for important decisions including pricing, marketing, and inventory.

To solve these, the project utilizes secondary data from Kaggle which has coffee sales transactions during 2-year period. The data will be used to track trends in bean sales, customer loyalty and discounts and forecast sales. This can be utilised by the company on aspects like inventory management, supplier negotiation and ultimately aims to provide QahwaCo with, feasible and data-driven strategies in their sales, promotions and customer retention.

# 2 Organization Background

QahwaCo is a coffee company that has, since its small start in 2020, grown into a name recognizable by the people of Saudi Arabia. It began from a group of friends who decided to turn their passion for coffee beans into a venture. They started from a store in Riyadh, and over the years added several stores in more cities.

The company sources beans from many origins - Brazil, Ethiopia, Colombia, Costa Rica and Guatemala. These are presented as high quality, yet the focus is not only on the beans but also on what households and cafes actually prefer. Every season they offer seasonal promotions to grow their brand and build trust. These promotions include limited roasts, gift bundles, and occasional surprises to reach a wider market. This has led to QahwaCo becoming one of the leading players in the specialty coffee market in Saudi Arabia.

Beyond coffee they have opened a few cafes in major cities under a subsidiary company, this also helps in introducing customers to their coffee and test their new samples there along with customers provide feedback.

# 3 Problem Statement

3.1 **Analyzing Seasonal Coffee Sales Trends:** This study will examine how the sales of each variety of coffee bean fluctuate over time, to identify peak selling season and figure out demand. The goal is to help QahwaCo better anticipate customer demand.

3.2 **Exploring Customer Loyalty and Discount Influence:** This study is to investigate the purchasing habits of customers and understand what led to repeat orders, while analyzing the impact of discounts. This will support QahwaCo in customer retention strategies and scheduling promotional offers.

3.3 **Forecasting Coffee Sales for Inventory Planning:** This study will utilise sales data, to develop a predictive model to forecast future demand for each coffee variety. These forecasts will aid QahwaCo in making data driven inventory decisions.

# 4 Background of the Problem

The specialty coffee market in Saudi Arabia has been growing fast. Tastes keep shifting, and the demand for high grade beans has been rising. Yet for retailers such as QahwaCo, this growth is not a simple straight path. They are constantly trying to meet the needs of different groups of customers, trying to cater their demands that defines the twists and changes in the trends.

Companies intend on gaining a deeper understanding of these cycles as it has become a necessity. Numbers don't tell the whole story, they require context; just tracking overall sales numbers, in reality, often hides the why behind changes. Why does a certain roast suddenly sell out? Why does another bean linger on shelves? Answering these questions with numbers alone might not be possible but figuring out the trends might be.

To keep up, businesses have leaned on forecasting and behavioural analysis. These tools help refine inventory, pricing, and loyalty approaches, though they are not omnipotent. Still, studies suggest that sales analysis with models of customer loyalty and promotion effects can turn their nature from reactive to proactive.

This project aims to provide QahwaCo with insights that are more actionable than some vague instinct, giving them a competitive edge over their competitors.

# 5 Problem Solving Approach

- **Data Collection:** This study utilizes a secondary coffee sales dataset obtained from Kaggle, containing detailed transaction-level coffee sales data across different bean varieties, customers, and discount applications over two years. The comprehensive nature of this dataset aligns well with the project's objectives. (https://www.kaggle.com/datasets/halaturkialotaibi/coffee-bean-sales-dataset)

- **Data Cleaning and Preparation:** The collected raw data will undergo cleaning, normalization, and transformation to remove inconsistencies, missing values, and duplicates. Feature engineering techniques will be applied to enhance data quality, making it suitable for analysis.

- **Analytical Approach:**

  - Descriptive Analytics: Exploratory data analysis (EDA) will summarize sales patterns by coffee type, seasonality, and customer segments, using statistical summaries and visualizations.

  - Customer Behaviour and Loyalty Analysis: Cohort analysis and segmentation will help reveal repeat purchase trends and the relationship between discount offers and customer loyalty and revenue impact.

  - Demand Forecasting: Time series forecasting and regression modeling techniques will be applied to predict future sales volumes by coffee type.

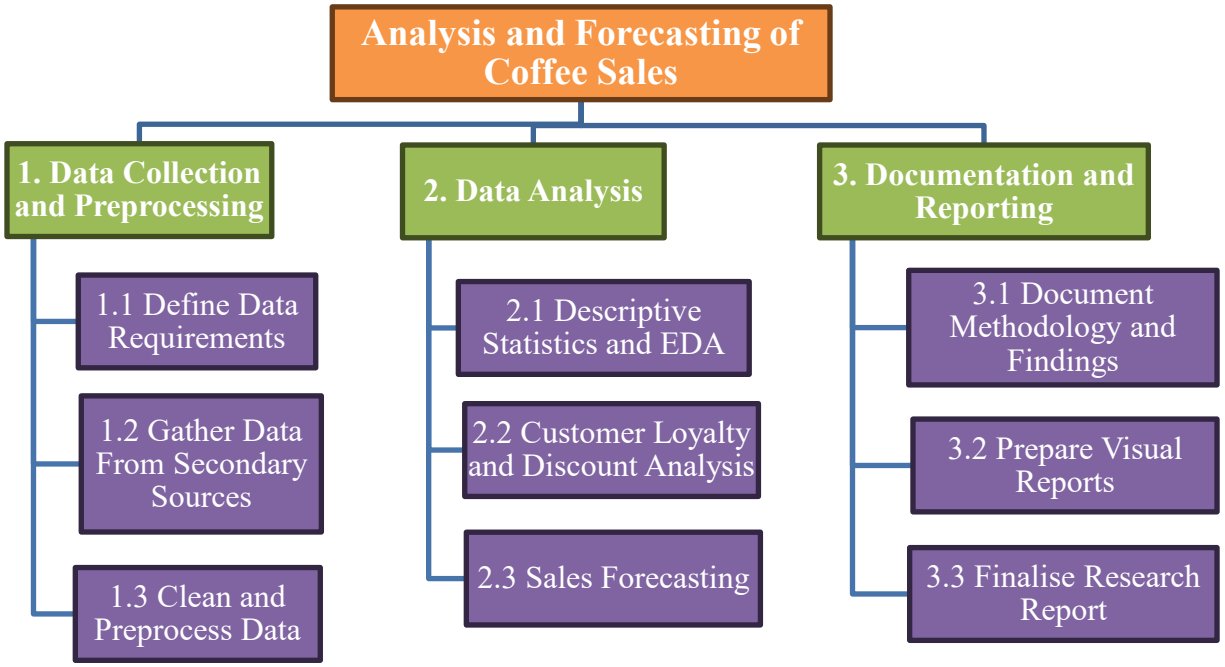
- **Tools and Software:** Python is the primary tool for analysis, utilizing libraries such as Pandas, NumPy, Scikit-learn, Statsmodels, Matplotlib, and Seaborn for data processing, modeling, and visualization.

- **Documentation:** All the steps will be documented from data collection to analysis and modeling to ensure consistency, transparency, reproducibility, and accuracy.

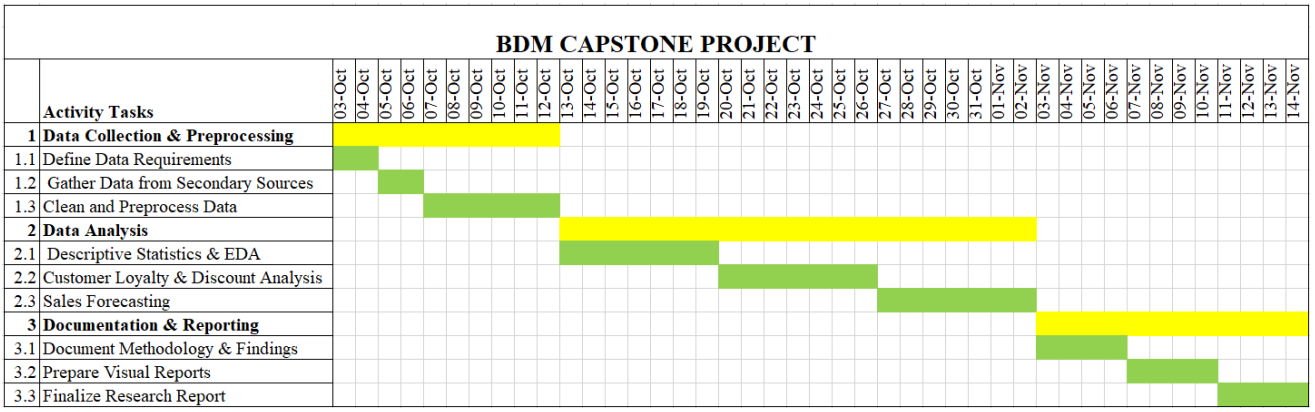
- **Validation:** Model validation techniques like R square, RMSE, etc will be used to grade the performance of the model.

# 6 Expected Timeline

## 6.1 Work Breakdown Structure:




## 6.2 Gantt Chart



## 7 Expected Outcome:

- The analysis is expected to deliver a comprehensive report featuring clear identification of sales peaks and seasonality for each coffee bean, helping retailers understand which products succeed at different times of year. Data-driven insights into customer loyalty will clarify the effectiveness of discounts in driving repeat business and increasing overall revenue.

- Actionable forecasting models will empower retailers to make informed inventory decisions, reducing the risk of stockouts or excess. These insights can also be utilized for optimal pricing, discount strategies, and retention programs.

- All findings will be firmly grounded in rigorous data analysis, allowing QahwaCo to enhance its business strategies with real evidence rather than intuition. This approach ensures recommendations are actionable, practical, and relevant to current market dynamics.