A Data-Driven Analysis of Sales and Customer Behavior

A Proposal report for the BDM capstone Project

Submitted by

Name: Kapishankar Ashtankar

Roll number: 21F3000371

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 "**A Data-Driven Analysis of Sales and Customer Behavior**". I extend my appreciation to Balaji Restaurant, 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 from primary 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 thorough 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: Kapishankar Ashtankar

Date: 12-06-2025

1 Executive Summary and Title

This project aims to analyze the sales performance of a local food retail business using descriptive and predictive analytics. The primary objective is to identify the most frequently sold items, sales patterns and uncover time-based patterns such as seasonal trends and peak sales periods. Historical sales data was collected and processed using tools such as Microsoft Excel and Google Colab, enabling data cleaning, preparation, visualization, and statistical analysis.

Descriptive analytics was used to evaluate overall sales trends, segment sales by item, and highlight customer preferences through visualizations such as bar and pie charts. Time series analysis was applied to detect recurring patterns and shifts in demand. Predictive modeling techniques were used to forecast future sales, allowing for more effective planning.

The expected outcomes of this project include data-driven insights into top-performing products, customer purchasing behaviors, and demand fluctuations over time. These findings support inventory planning and menu optimization. The project also provides actionable recommendations to help the business improve operational efficiency and increase profitability. This dataset was taken from the kaggle website.

By leveraging data, this project empowers the business to make informed decisions that align with customer demand and market behavior, ensuring improved service delivery and sustainable growth.

2 Organization Background

Balaji Fast Food is a quick-service restaurant located in a bustling commercial neighborhood in Nagpur, Maharashtra. Operating in the fast-food and casual dining segment, the restaurant serves a variety of popular Indian snacks and beverages, including items like cold coffee, panipuri, vada pav, aloo puri and so on. The restaurant is known for its affordable pricing, fast service, and convenient takeaway options, which have made it a popular choice for nearby office-goers, students, and families.

Despite a loyal customer base, Balaji Fast Food faces operational challenges common to many small food businesses — such as managing fluctuating demand, inventory planning, and limited insights into customer preferences. With increasing competition from nearby vendors and online delivery platforms, the business recognizes the need to become more data-driven.

This project is based on real sales data (sourced from Kaggle) that reflects daily transactional details including items sold, quantity, time of sale and payment method. The goal is to analyze this data to identify fast-moving items, uncover sales trends, and forecast demand — ultimately leading to smarter business decisions and improved operational efficiency.

3 Problem Statement

3.1 Identify which food items are sold the most and predict future demand.

3.2 To analyze historical sales data in order to identify time-based trends and seasonal patterns that impact overall business performance.

4 Background of the Problem

In the dynamic and highly competitive fast-food industry, staying responsive to changing customer preferences and consumption trends is essential for operational success. Balaji Fast Food, a local quick-service restaurant, is currently facing challenges in managing demand fluctuations and optimizing resources. Over time, the establishment has noticed irregular sales volumes across different days, times, and food categories. These inconsistencies have made inventory planning inefficient and overall revenue performance difficult to sustain.

The root of the problem lies in the absence of a data-driven approach. Without proper analysis of historical sales patterns, the restaurant is unable to accurately identify its most popular items or forecast future demand. This lack of insight hampers timely decision-making and prevents the business from adapting to customer behavior or market shifts.

4.1 Internal Problems:

1. The dataset contains missing values, indicating inconsistencies in data recording.
2. Sales are highly concentrated on a few specific items, which may reflect over-reliance on limited menu options.

4.2 External Problems:

1. Menu item prices appear static, with no adjustments based on seasonality or demand, which may hinder dynamic pricing strategy.
2. Nearby food stalls offering similar items at competitive prices may affect sales.
3. Customer volume may vary significantly day-to-day, making it hard to plan staff or stock levels.
4. Fast food vendors now face increasing competition from online food delivery platforms like Zomato, Swiggy, and food chains that provide home delivery and discounts.

5 Problem Solving Approach

To address the challenges faced by Balaji Fast Food Restaurant in understanding item demand and sales trends, a structured data-driven approach will be followed. The primary objective is to identify fast-moving food items and uncover patterns in sales over time. This will enable better decision-making related to inventory management, menu planning, and resource allocation.

5.1 Methods: Analytical Approach and Justification

The analysis begins with data cleaning and preparation, which is essential to ensure consistency, remove discrepancies, and prepare the dataset for downstream analysis. Missing values, inconsistent date formats, and duplicated entries are handled during this stage.

Next, Exploratory Data Analysis (EDA) is conducted to extract descriptive insights. Grouping by item name helps determine which products contribute the most to overall sales. Visualization techniques such as bar charts and pie charts help to effectively communicate customer preferences and item popularity.

For time-based analysis, the data is aggregated daily and monthly to detect seasonal trends. Time series decomposition is applied to identify trend, seasonality, and noise components. These insights form the basis for demand forecasting.

Finally, forecasting models like ARIMA and Facebook Prophet are implemented to predict future sales. These models are chosen due to their reliability in capturing seasonality, trends, and short-term fluctuations. Forecasting enables better planning for raw materials and labor, and reduces operational inefficiencies.

5.2 Data Source and Justification

The dataset was sourced from Kaggle, a well-known open data platform that hosts numerous real-world datasets for analytical projects. The dataset captures daily sales transactions from a fast-food outlet and includes attributes such as:

1. **order_id**
2. **date**
3. **item_name**
4. **item_type**
5. **item_price**
6. **Quantity**
7. **transaction_amount**
8. **transaction_type**
9. **received_by**
10. **time_of_sale**.

This dataset was chosen because it closely reflects the operations of Balaji Fast Food, making it highly relevant for the defined problem statement. It offers a comprehensive view of both internal business operations and customer behavior, thus serving as a suitable foundation for sales trend analysis and forecasting.

5.3 Tools and Technologies Used

Google Colab: Cloud-based platform for coding and analysis using Python.

Python: Main programming language used due to its rich data analytics libraries.

Microsoft Excel: Utilized for initial data review and verification.

Pandas & NumPy: Used for efficient data wrangling and numerical computations.

Matplotlib & Seaborn: Employed for static visualizations like bar and line graphs.

Plotly: Used for interactive Gantt charts and trend analysis.

Facebook Prophet & ARIMA: Applied for robust and interpretable time series forecasting.

6 Expected Timeline

6.1 Work Breakdown Structure:

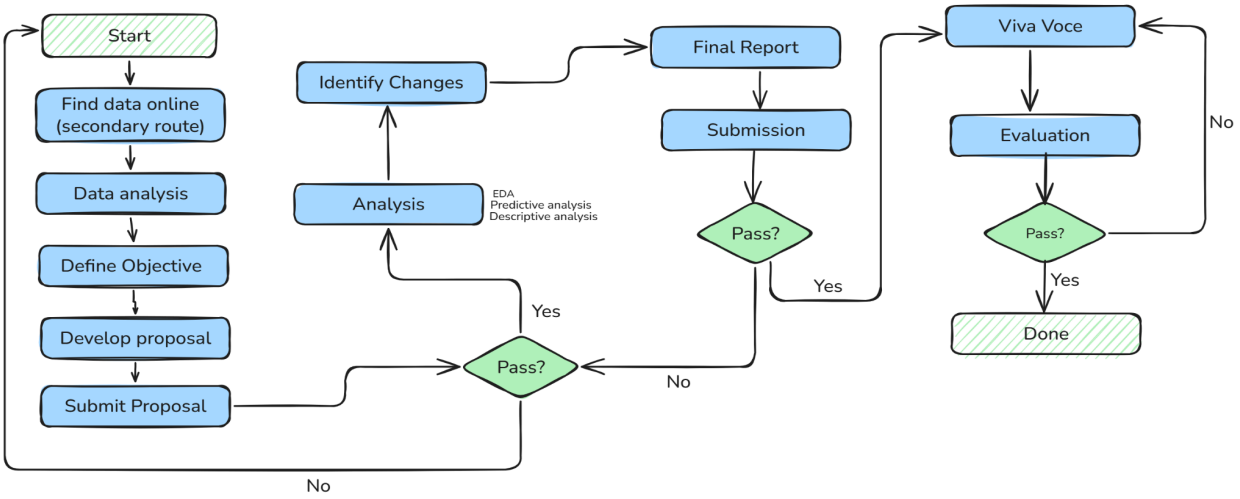

fig.1: Work Breakdown Structure

6.2 Gantt chart:

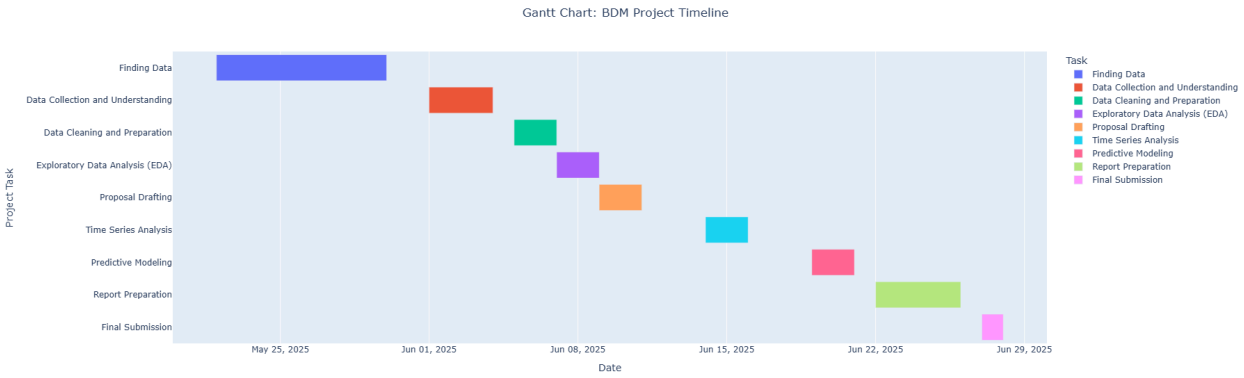

fig.1: Gantt Chart

For high resolution images refer:

1) wbs: Click Here
2) Gantt Chart: Click Here

7 Expected Outcome

1. Comprehensive analysis of sales trends through time series techniques, enabling the identification of consistent patterns, peak sales periods, and fluctuations in customer demand over time.
2. Enhanced decision-making through integration of descriptive and predictive analytics.
3. Insight-driven strategies to enhance inventory control and menu planning based on sales performance.