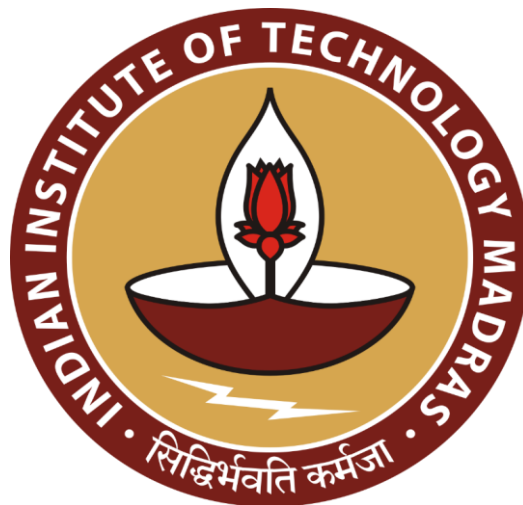# Sundarayatan Store: A Strategic Analysis for Business Enhancement

A Mid Term report for the BDM capstone Project

Submitted by

Name: Avinash Kumar

Roll Number: 21f1006385

IITM Online BS Degree Program,

Indian Institute of Technology, Madras, Chennai

Tamil Nadu, India, 600036

Contents

Sundarayatan Store: A Strategic Analysis for Business Enhancement

1. Executive Summary:

Sundarayatan Store is a medium-scale neighbourhood grocery outlet located near NIT More, Patna, and has served the local community since 2006. Despite a loyal customer base and an average daily footfall of 130–160 customers, the store has experienced declining profitability and operational challenges. This decline is primarily attributed to the rise of online retail platforms, overstocking issues, and insufficient data-driven decision-making. The project aims to investigate these problems and suggest data-backed strategies for inventory optimization and profit margin improvement.

To support this objective, 61 days of data (from March 1 to April 30, 2025) were manually collected from handwritten sales registers, inventory logs, and past purchase invoices. The dataset covers 12 commonly sold grocery items, including rice, flour, pulses, cooking oil, dairy products, ghee, dry fruits, and beverages. Key variables captured include cost price, selling price, quantity sold, and purchase quantity. Descriptive statistics revealed that Dry Fruits and Rice were the most profitable items, generating ₹34,020 and ₹32,591 respectively, whereas Sugar and Moong Daal had the lowest profits at ₹5,654 and ₹4,726.

A range of analytical methods is being applied to address the identified problems. For inventory optimization, ABC analysis and demand forecasting (using moving averages) are used. Rice, Atta, and Cooking Oil were classified as 'A' category products. For margin analysis, gross profit evaluation showed an average daily profit of ₹2,819 on sales worth ₹24,835. Customer loyalty and behavior are also being examined. RFM segmentation and survey data revealed that 37.9% of customers preferred online purchases, while 14 customers scored 8 on the loyalty index, indicating moderate loyalty levels. This analysis is expected to drive better inventory control and margin growth in a competitive market.

2. Proof of Originality of Data:

You can find the supporting evidence verifying the originality and authenticity of the datasets used in the Sundarayatan Store – Business Data Management Project through the following link:

LINK: Google Drive Folder

As proof of originality, the following materials have been included:

- A formal letter of permission from the store owner.
- A link to the dataset used for this project.
- Photographic evidence of the store premises.
- A short video interaction with the shop owner validating the data collection process.

3. Metadata and Descriptive Statistics:

3.1. Metadata: The dataset encompasses information on product-wise sales, revenue, expenses, profits, and inventory for the period between March 1, 2025, and April 30, 2025. It includes records for 12 key items, namely: Rice, Sugar, Cooking Oil, Atta (Wheat Flour), Milk and Dairy Products, Dry Fruits, Ghee, Toor Daal, Urad Daal, Moong Daal, Ice Cream, and Cold Drinks.

- **Data Collection Period:** The data collection period for this project spans two consecutive months, specifically from March 1, 2025, to April 30, 2025.

- **Data Sources:** The data used in this project was sourced directly from Sundarayatan Stores' internal records. It includes two months of transactional and operational data, covering key metrics such as sales, revenue, expenses, profit, inventory levels, cost of goods sold, cost price, and purchase quantities. The dataset is provided in Excel format and reflects daily product-wise performance, allowing for granular analysis of store operations and financial health.

3.1.1. Store's Operational Data: The operational dataset of the store captures essential product-level information, including sales, daily pricing, quantities purchased, inventory status, cost of goods sold, and profit margins. This data provides a comprehensive view of

the store's day-to-day functioning and supports analysis. The dataset consists of 14 columns and 62 rows for sales, inventory and purchase data.

Column Name	Description	Data Type	Purpose
Date	The date of record from (March- May 25)	Date	Enables time-series tracking and daily comparisons between different products.
Rice	Value related to Rice (e.g., Sales, Profit, Inventory, etc.).	Numeric	Analyze trends specific to rice in the chosen metric (sales, inventory, etc.).
Sugar	Value related to Sugar.	Numeric	Helps assess sugar-specific performance or stock status.
Cooking Oil	Value related to Cooking Oil.	Numeric	Useful for tracking a frequently purchased kitchen essential.
Atta	Value related to Wheat Flour (Atta).	Numeric	Evaluate supply and performance of staple products.
Milk and Dairy	Value related to Milk and Dairy items.	Numeric	Important for perishable inventory and freshness-sensitive products.
Ice Cream	Value related to Ice Cream.	Numeric	Seasonal or high-margin product performance tracking.
Ghee	Value related to Ghee.	Numeric	Traditional product with significant cultural and festive value.
Dry Fruits	Value related to Dry Fruits.	Numeric	Premium items, useful for analyzing festive/seasonal demand.
Cold Drinks	Value related to Cold Drinks.	Numeric	Evaluate beverage sales/stock patterns, especially in warmer months.
Toor Daal	Value related to Toor Daal.	Numeric	Helps assess staple pulse item turnover and profitability.
Moong Daal	Value related to Moong Daal.	Numeric	Protein-rich pulse often monitored for daily consumption trends.
Urad Daal	Value related to Urad Daal.	Numeric	Key ingredient in various recipes—useful for daily stock and pricing decisions.
Total Value	Aggregated total for the day (e.g., Sales, Profit, Revenue etc.).	Numeric	Overall summary metric per day—used for forecasting, profitability, or inventory planning based on the sheet type.

Table 1: *Metadata of Store's Collected Data for 2 Months*

3.1.2. Customer Data: The customer survey dataset comprises responses from 95 individual customers, consisting of 14 columns and 95 rows capturing key information to evaluate shopping behavior and satisfaction levels. It includes variables such as customer ID, name, age group, gender, income bracket, visit frequency, satisfaction score, and open-ended feedback.

This dataset supports demographic segmentation, helps identify customer preferences, and provides insights to enhance service quality and tailor marketing strategies.

Column Name	Description	Data Type	Purpose
Customer_ID	Unique identifier for each customer.	Categorical	Used to distinguish individual respondents.
Customer_Name	Name of the customer.	Categorical	For personalization and customer-specific insights.
Age_Group	Customer's age group (e.g., 18–25, 26–35).	Categorical	Useful for demographic segmentation and targeting.
Gender	Customer's gender (e.g., Male, Female, Other).	Categorical	Enables gender-based behavior and preference analysis.
Income_Level	Income category of the customer (e.g., Low, Medium, High).	Categorical	Useful for socioeconomic segmentation and offer customization.
Shopping_Frequency	Frequency of shopping (e.g., Weekly, Monthly).	Categorical	Helps identify high-engagement customers and optimize campaign timing.
Preferred_Purchase_Channel	Mode of preferred shopping (e.g., Online, In-store).	Categorical	Assists in optimizing channel strategy and resource allocation.
Price_Sensitivity	Customer's sensitivity to price (e.g., High, Medium, Low).	Categorical	Important for pricing strategy and discount targeting.
Loyalty_Score	Numerical score indicating customer loyalty.	Numerical	Used for loyalty segmentation and retention strategy planning.
Switching_Reason	Stated reason why a customer might switch to a competitor.	Categorical	Identifies pain points and areas of improvement.
Product_Preferences	Categories or types of products preferred by the customer.	Categorical	Useful for recommendation systems and inventory planning.
Offers_Response	How the customer reacts to promotional offers (e.g., Accepts, Ignores).	Categorical	Helps measure campaign effectiveness and refine targeting.
Competition_Visited	Indicates if the customer visited a competitor (Yes/No).	Categorical	Used to assess competitive threats and brand stickiness.
Retention_Suggestion	Customer's suggestion to improve retention.	Categorical	Valuable for customer feedback analysis and service improvement initiatives.

Table 2: *Metadata of Customer's Survey Sheet*

3.2. Descriptive Statistics: A comprehensive examination of the dataset and graphical trends from the past two months at Sundarayatan Stores has led to the key insights and conclusions summarized in Table 3: -

Metric	Sales	Revenue	Expense	Profit	Inventory
Mean	310.78	24,835.52	22,016.52	2,819.00	2,783.59
Standard Error	6.02	469.52	421.01	61.10	85.08
Median	305	24,270	21,608.50	2,790.00	2,654.00
Mode	357	32,615	29,034.50	3,781.50	2,706.00
Standard Deviation	47.00	3,667.06	3,288.23	477.22	664.52
Sample Variance	2,208.95	13,447,313.45	10,812,452.30	227,735.95	441,585.90
Kurtosis	-0.28	-0.71	-0.72	-0.72	0.89
Skewness	0.63	0.35	0.39	-0.02	0.75
Range	186	13,670	11,991	1,880	3,195.50
Minimum	238	18,945	17,043.50	1,901.50	1,598.50
Maximum	424	32,615	29,034.50	3,781.50	4,794.00
Sum	18,957.50	15,14,967	13,43,008	1,71,959	1,69,799
Count	61	61	61	61	61
Largest (1)	424	32,615	29,034.50	3,781.50	4,794.00
Smallest (1)	238	18,945	17,043.50	1,901.50	1,598.50
Confidence Level (95%)	12.04	939.18	842.15	122.22	170.19

Table 3: *Descriptive Statistics*

The descriptive statistics reveal that the business maintained consistent daily operations over 61 days, with average sales of 311 units and revenue of ₹24,836. Expenses averaged ₹22,017, leaving a steady profit of around ₹2,819 per day. While revenue and inventory showed noticeable variability, profit remained relatively stable, indicating efficient cost control. Sales and inventory distributions were slightly right skewed, suggesting occasional high spikes. Overall, the business performed reliably with some fluctuations in demand and stock levels, offering opportunities to optimize inventory and boost high-sales days for increased profitability.

4. Detailed Explanation and Analysis Process/Method:

The analysis for Sundarayatan Stores was carried out in a structured and data-driven manner to address the key business challenges: inventory optimization, profit margin

evaluation, and customer retention. The following steps outline the analytical methods used, along with their rationale and justification:

4.1. Data Cleaning and Preprocessing: The data preprocessing phase involved a series of cleaning operations to ensure consistency and reliability across all datasets, including Sales, Revenue, Expense, Profit, Inventory, and Purchase Quantity. Missing, duplicate, and inconsistent values were identified and removed to maintain data integrity. Date formats were standardized for temporal alignment, and numeric columns were properly formatted by stripping currency symbols and converting textual numbers into float types. Additionally, daily records were aggregated at the product level to ensure uniformity and enable seamless integration across various sheets during analysis.

4.2. ABC Analysis for Inventory Optimization: To identify which products, contribute most to the store's revenue and require close inventory monitoring, ABC analysis was applied. Items were categorized into A, B, and C classes based on their annual consumption value (sales quantity × cost price). This helped prioritize high-impact products and guide stocking decisions to reduce wastage and overstocking.

- Products were categorized into:
    1. **A-category:** Top ~70–80% of value (critical)
    2. **B-category:** Next ~15–20%
    3. **C-category:** Bottom ~5–10% (low impact)
- **Focuses Problem 1:** This method aligns directly with the goal of inventory optimization by focusing attention on high-value items.

4.3. Time Series Forecasting for Demand Prediction: ARIMA (AutoRegressive Integrated Moving Average) was employed to forecast daily product sales. This statistical model, widely used for time series forecasting, helps in capturing trends and seasonality in sales data. By applying ARIMA, future demand for each product was projected, enabling better alignment of procurement schedules with actual demand. This approach minimizes the risk of stockouts and excess inventory, ensuring optimal stock levels and improved inventory turnover. **Abstract Model:** $Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + ... + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \varepsilon_t$

- **Focuses Problem 1:** Prevents stockouts and excess inventory by aligning ordering with demand projections.

4.4. Profit Margin Evaluation: To evaluate profitability, a detailed profit margin analysis was conducted at the individual product level. Each item's profit margin was calculated and ranked to distinguish high-margin from low-margin products. Additionally, trends in Contribution Margin and Cost of Goods Sold (COGS) were examined to gain insights into underlying cost structures.

- **Focuses Problem 2:** This method helps in pinpointing products that contribute most significantly to overall profitability. The analysis serves as a foundation for strategic decisions related to pricing aimed at maximizing profit margins.
- **Formula** - Profit Margin (%) = (Profit ÷ Revenue) × 100

Product	Profit (₹)	Revenue (₹)	Profit %	Profit Margin
Rice	₹32,591.00	₹170,040.00	18.95%	19.17%
Sugar	₹5,654.00	₹125,521.00	3.29%	4.50%
Cooking Oil	₹18,932.00	₹336,860.00	11.01%	5.62%
Atta	₹12,894.50	₹124,441.00	7.50%	10.36%
Milk and Dairy	₹7,794.00	₹155,880.00	4.53%	5.00%
Ghee	₹12,297.50	₹134,010.00	7.15%	9.18%
Dry Fruits	₹34,020.00	₹151,200.00	19.78%	22.50%
Cold Drinks	₹11,870.00	₹59,350.00	6.90%	20.00%
Ice Cream	₹14,350.00	₹57,400.00	8.35%	25.00%
Toor Daal	₹9,870.00	₹67,060.00	5.74%	14.72%
Moong Daal	₹4,726.00	₹72,885.00	2.75%	6.48%
Urad Daal	₹6,960.00	₹60,320.00	4.05%	11.54%
Sum	**₹1,71,959.00**	**₹15,14,967.00**	**Average**	**12.84%**

Table 4: Profit Margin Evaluation

4.5. Customer Survey Analysis: Data from 95 customer responses was analyzed to understand preferences, satisfaction levels, and expectations. Descriptive statistics were used to quantify ratings, while feedback comments were categorized into key themes. Additionally, RFM (Recency, Frequency, Monetary) analysis was used to segment customers by their shopping behavior, which helped design targeted retention strategies.

- Applied RFM (Recency, Frequency, Monetary) scoring model:
- **Score** = R + F + M (where each is normalized on a 1–5 scale)
- **Focuses Problem 3:** Supports retention strategies by identifying loyal, at-risk, and high-value customer groups.

5. Results and Findings:

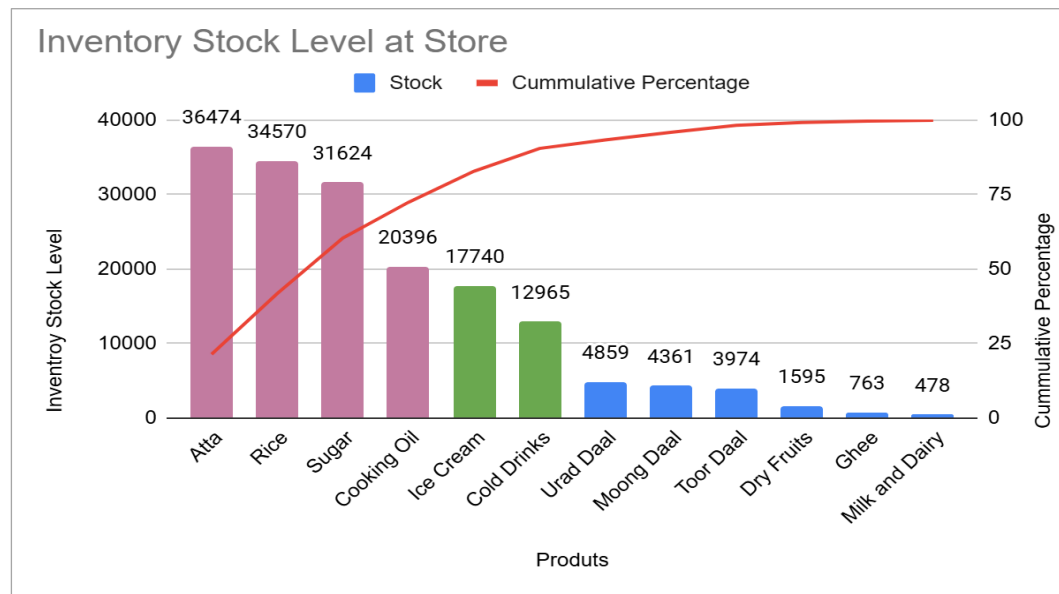### 5.1. ABC Analysis for Inventory Optimization:

Figure 1: *Pareto Chart for ABC Analysis*

The chart in figure 1 illustrates an ABC analysis of inventory stock levels at Sundarayatan Store, highlighting how different products contribute to total inventory. Items like Atta, Rice, Sugar, and Cooking Oil fall under Category A, holding the largest stock quantities and accounting for nearly 70% of total inventory, requiring strict control and frequent monitoring.

Category B, which includes Ice Cream and Cold Drinks, holds moderate stock and contributes up to 90% cumulatively, needing regular but less intensive oversight. The remaining items, like Daal varieties, Dry Fruits, Ghee, and Milk, make up Category C with minimal stock levels and account for the last 10%, needing minimal control.

This classification helps the store prioritize inventory management, focusing efforts on high-value items to minimize waste and optimize procurement.

5.2. Sales Demand Forecasting:

Statistic	Value
Alpha	0.13
MASE	0.74
SMAPE	0.13
MAE	39.84
RMSE	49.77

The Sales Demand Forecast chart illustrates historical and predicted sales from March to May 2025. The actual sales (in dark blue) show daily fluctuations, while the forecasted sales (in orange) for May are accompanied by upper and lower confidence bounds (light blue and green), indicating the expected range of variability. The forecast performance metrics show strong accuracy: a MASE of 0.74 and SMAPE of 13% reflect

reliable predictions, while the average error (MAE) is around 40 units. The model's narrow confidence intervals and low error rates suggest it is effective for guiding inventory planning and demand-driven procurement.
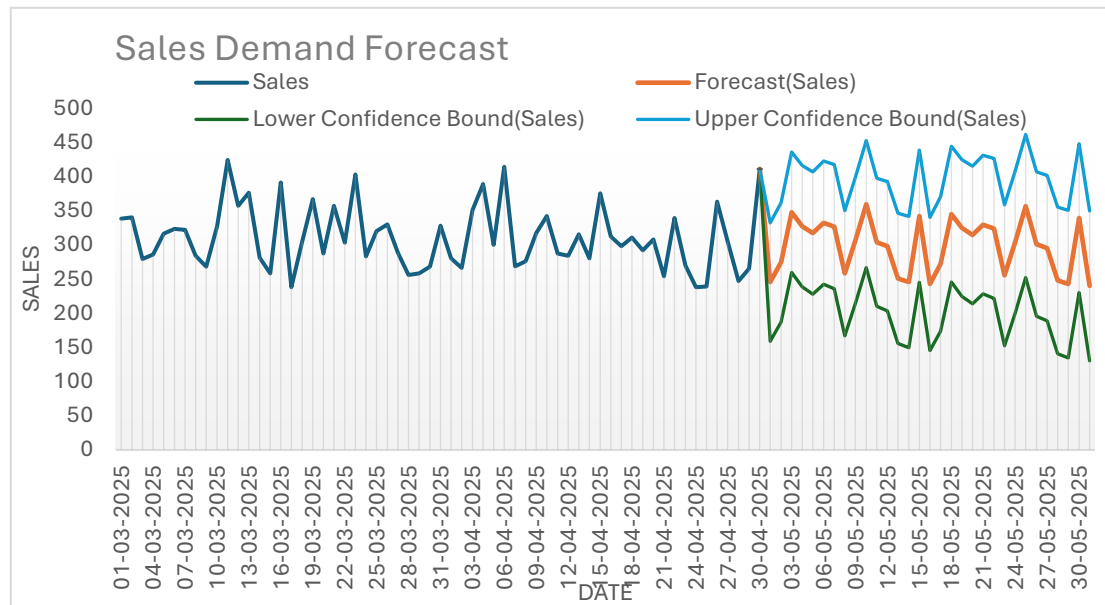

Figure 2: *Sales Demand Forecasting*
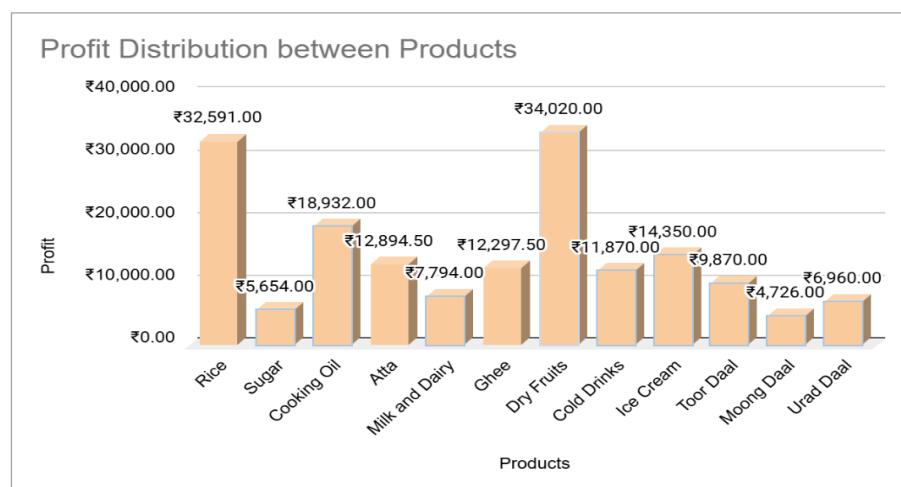
5.3. Profit and Profit Margin Analysis:

Figure 3: *Profit Distribution of Products*

The bar chart in figure 3 shows how different products contribute to overall profit. It clearly highlights that Dry Fruits (₹34,020) and Rice (₹32,591) are the top profit-generating items. Mid-range contributors include Cooking Oil (₹18,932), Ice Cream (₹14,350), and Atta (₹12,894.50). On the lower end, Sugar (₹5,654), Toor Daal (₹4,726), and Moong Daal (₹9,870) yield relatively less profit. This distribution helps identify high-margin products that can be

prioritized in pricing and promotional strategies, while low-profit items may need cost review or bundling tactics.

5.4. Customer Segmentation based on Preffered Purchase Channel:

Figure 4: Pie chart on preferred purchase channel
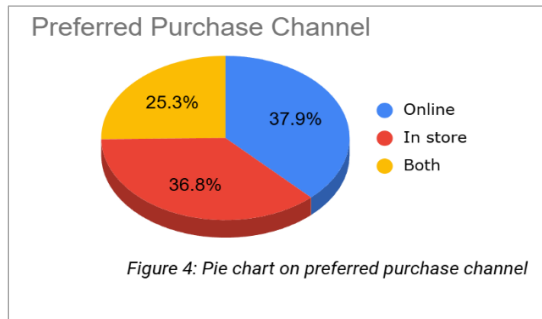
The pie chart in figure 4 illustrates how customers of the store prefer to shop—online, in-store, or both. The largest segment, 37.9% (35 customers), prefer online shopping, indicating a strong digital shopping trend. Close behind, 36.8% (36 customers) still prefer the in-store experience, highlighting the importance of maintaining physical retail operations. Meanwhile, 25.3% (24 customers) use both channels, showing a growing interest in an omnichannel approach. This distribution suggests that while online shopping is slightly more popular, a balanced strategy catering to both physical and digital preferences will be most effective for customer retention and satisfaction.

5.5. Customer Segmentation Analysis based on Loyalty Score:

The bar chart in figure 5 shows how customers are distributed across different loyalty score levels, ranging from 1 to 10. The highest number of customers (14) have a loyalty score of 8,
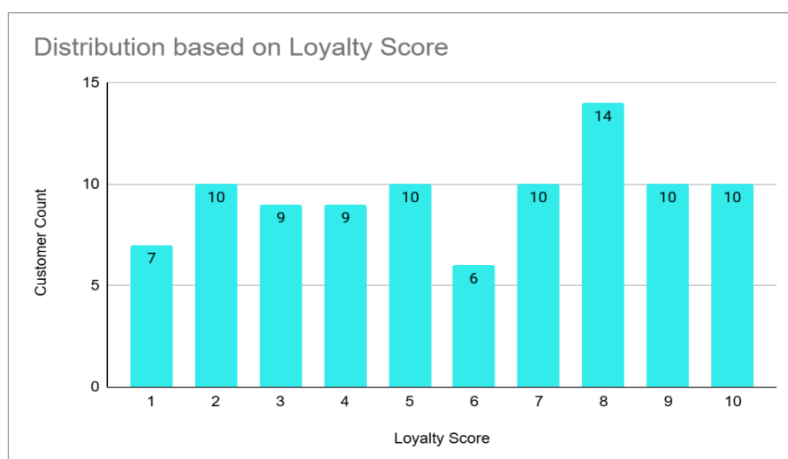
indicating strong loyalty in a significant portion of the customer base. Scores 2, 5, 7, 9, and 10 each have 10 customers, reflecting consistent engagement at various loyalty levels. On the lower end, fewer customers

Figure 5: *Bar Chart based on Loyalty Score*

scored 1 (7 customers) and 6 (6 customers), suggesting weaker engagement or newer customers in those categories. This distribution can help tailor loyalty programs to boost engagement in the lower-score segments while retaining high-score customers.

Reference Link: Dataset