Final Project Presentation

Optimizing Footfall, Sales & Resource Allocation at North Cafeteria

Machine Learning & Pattern Recognition

The Team

Atharva Sawant
BSE

Tanisha Saraf
RCPS

Udhav Shankar
DSEB

Our Problem?

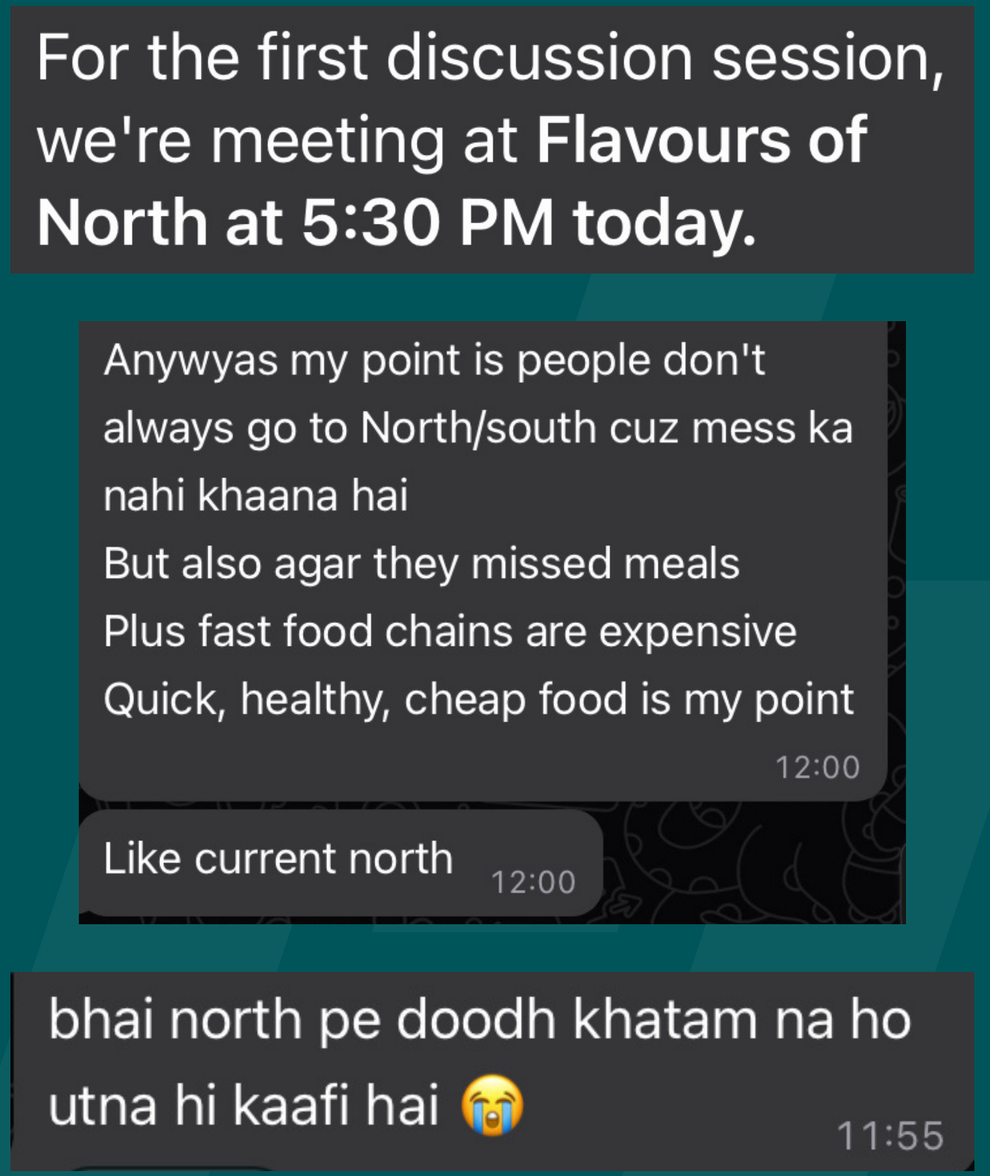
The “North Cafeteria” kiosk owner faces significant challenges in managing **inventory and staffing.** He struggles to determine the **precise quantities of raw materials** required for individual dishes, leading to **overstocking** and **understocking**, incurring potential **financial losses**, and risking **ingredient shortages** during peak hours. Staffing decisions lack data-driven insights, resulting in occasional **understaffing** during high-demand periods, compromising customer service, and leading to unnecessary labor costs. In essence, the core issue is **the absence of data-driven decision-making processes, impacting inventory, staffing, and customer service and ultimately sales.**

Motivation

- **Student Experience:** Students at the kiosk encounter issues such as delays, ingredient shortages, and long wait times. This leads to frustration and possible class lateness.

- **Reputation and Customer Retention:** The kiosk's failure to meet expectations harms its reputation and deters student loyalty. Negative feedback circulates, impacting the kiosk's business and the campus experience.

- **Community Benefit:** Efficient kiosk management enhances the entire campus community's access to reliable and convenient food services.



Potential Applications

- **Small Businesses:** Our solution will be applicable to other small businesses and local eateries facing similar operational challenges in inventory management, staffing, and customer service.

- **Food Industry:** The application of machine learning and pattern recognition can be extended to various segments of the food industry, aiding in better resource allocation and enhanced customer experiences.

- **Service Sector:** Businesses in the service industry, such as salons or repair shops, can also employ data-driven insights to enhance customer service and reduce operational costs.

Potential Impact

- **Cost Reduction:** Our solution has the potential to significantly reduce operational costs for the kiosk by preventing overstocking and optimizing staffing, leading to savings in labor and storage expenses and consequently potentially increasing profit.

- **Sustainability:** By reducing food wastage and labor inefficiencies, our solution promotes sustainability, aligning with broader environmental and economic sustainability goals.

- **Enhanced Customer Experience:** Efficient staffing and inventory management can lead to improved customer service and satisfaction, contributing to repeat business and positive word-of-mouth marketing.

- **Modernization:** The adoption of machine learning and pattern recognition techniques modernizes the kiosk's operations.

Literature Survey

Inventory Forecasting

Data-Intensive Inventory Forecasting with Artificial Intelligence Models for Cross-Border E-Commerce Service Automation

- **Problem:** This paper addresses the the issue of accurately predicting how much inventory is needed for cross-border e-commerce, which is challenging due to the complexity of international sales and supply chain management.

- **Solution:** Proposes the use of AI models, particularly the XGBoost algorithm, to handle data-intensive inventory forecasting. It demonstrates that XGBoost outperforms other methods in terms of accuracy and efficiency.

- **Conclusion:** Research concludes that AI models, such as XGBoost, significantly enhance inventory forecasting accuracy and efficiency in cross-border e-commerce, offering a promising solution to supply chain management challenges in this context.

Source: https://www.mdpi.com/2076-3417/13/5/3051

Inventory Management

Inventory Management using Machine Learning

- **Problem:** Focuses on addressing the challenge of inventory management, specifically demand forecasting, for small and medium-sized businesses, aiming to predict product demand and reduce overstock and out-of-stock situations.

- **Solution:** Employs the XGBoost regression model, a decision tree-based machine learning algorithm, to predict product demand and facilitate more efficient inventory management.

- **Conclusion:** The paper concludes that demand forecasting using algorithms like the XGBoost regression model can help reduce manual labor, minimize overstock and stock-out situations, and improve profitability for businesses, highlighting the potential for further enhancements through the incorporation of categorical embeddings in neural networks.

Source: https://www.researchgate.net/profile/Madhuri-Javaregowda/publication/342529859_Inventory_management_using_Machine_Learning/links/5f00bc48299bf18816008a64/Inventory-management-using-Machine-Learning.pdf

AI in Food Industry

Opportunities of Artificial Intelligence and Machine Learning in the Food Industry

- **Problem:** The paper addresses the challenge of enhancing various aspects of the food industry, such as customer satisfaction, recipe development, food delivery, and food safety, by leveraging artificial intelligence (AI) and machine learning (ML) technologies.
- **Solution:** The authors recommend AI-based customer feedback systems, predictive analytics for food delivery.
- **Conclusion:** The paper emphasizes the potential of AI and ML in transforming the food industry, making it more efficient, cost-effective, and customer-centric. These technologies are presented as key drivers of innovation and improvement across various food-related processes.

Source: https://downloads.hindawi.com/journals/jfq/2021/4535567.pdf

Current Implementations

Walmart

- Operational Flow
- Customer-Centric Operations
- Predictive Analytics
- Deep Learning for Recommendations
- Neural Networks

Source: https://corporate.walmart.com/news/2019/04/25/walmarts-new-intelligent-retail-lab-shows-a-glimpse-into-the-future-of-retail-irl

amazon

- Product Recommendations
- Inventory Forecasting
- Amazon Forecast-Deep Learning Algorithm
- Neural Networks
- Clustering

Source: https://aws.amazon.com/blogs/aws/amazon-forecast-now-generally-available/

- Optimized Operations
- Predictive Analysis
- Neural Networks
- Regression Models
- Decision Trees

Source: https://emerj.com/ai-sector-overviews/artificial-intelligence-at-mcdonalds/

- Sales Prediction & Demand Forecasting
- Optimizing menu & Pricing Strategies
- Inventory Management
- Predictive Analytics
- Dynamic Pricing

Source: https://www.forbes.com/sites/bernardmarr/2018/05/28/starbucks-using-big-data-analytics-and-artificial-intelligence-to-boost-performance/?sh=7acdec7765cd

Flipkart

- Demand Forecasting
- Supply Chain Optimization
- Neural Networks
- Regression Models
- Decision Trees

Source: https://www.itnews.asia/news/flipkart-uses-machine-learning-to-boost-performance-592958

zomato

- Menu Digitization & Optimization
- Active DP Dispatch
- Predicting Food Preparation Time (FPT)
- CNNs
- LSTMs

Source: https://blog.zomato.com/elements-of-scalable-machine-learning

Inspiration

- **Predictive Analytics:** Implement predictive analytics to forecast demand and optimize inventory levels based on historical data and external factors like weather, day of the week, and special events.
- **Dynamic Pricing:** Employ dynamic pricing strategies based on demand fluctuations, similar to Starbucks, to maximize revenue and reduce excess inventory.
- **Machine Learning Models:** Develop machine learning models, including neural networks, regression, and decision trees, as used by McDonald's, for optimized staff scheduling based on historical sales patterns and foot traffic.
- **Recommendation Systems:** Utilize recommendation systems, as done by Amazon, to suggest complementary items, increasing sales and helping in inventory management.
- **Inventory Forecasting Algorithms:** Explore advanced inventory forecasting algorithms like Amazon's Amazon Forecast, a deep learning solution to predict stock levels more accurately, minimizing overstocking and understocking issues.

Dataset &
Features Pre-Processing

Data Collection

- We obtained the dataset from the kiosk owner, capturing 1 month of transactional information.
- This dataset because it aligns with the project's objective of analyzing and optimizing the kiosk's performance, offering insights into product sales, popularity, and financial aspects.
- To collect this data, we interviewed the kiosk owner to understand day-to-day operations, challenges, and potential areas of improvement. We faced some setbacks during this process, as a technical issue resulted in the loss of data. Additionally, a change in the North Cafeteria contract led to a new kiosk owner taking over, necessitating an adaptation of our data collection strategy.
- To safeguard sensitive financial information, we implemented measures, including the signing of an NDA with the owners that limits the use of data to research purposes within our team.

Pre-Processing

Columns	Importance	Justification
Category of the Item	Grouping items into categories enhances understanding, and reveals food type popularity.	Grouping items into categories simplifies the analysis process and provides a clear picture of which types of items are performing well or may need improvement.
Price Range of the Item	Segmenting by price (e.g., low, mid, high) optimizes pricing strategy and caters to diverse customer segments.	Tailoring pricing based on different customer segments and price sensitivity can enhance overall profitability and sales.
Sales Rank	Calculating the "Sales Rank" helps identify the best and worst-selling items, aiding inventory decisions to optimize the kiosk offerings.	Prioritizing top-sellers ensures popular items are consistently available, while addressing low-sellers allows for adjustments or potential removal from the menu.
Quantity Rank	Calculating the "Quantity Rank" is crucial for understanding the popularity and demand for different items based on the quantity sold.	Used for optimizing inventory management, ensuring customer satisfaction, and helping the kiosk identify which items are frequently ordered in larger quantities and which ones are less popular.

Raw Data

SERIAL NO	PRODUCT NAME	QTY	PRICE	TOTAL AMOUNT
2	ADRAK CHAI - SML	516	20	10320
3	PANEER PATTIES	22	70	1540
4	G GRILLED SANDWICH	80	70	5600
5	LEMONADE	17	50	850
6	CHEESE GARLIC BREAD	14	95	1330
7	PLAIN GARLIC BREAD	8	60	480
8	LOADED FRIES	5	99	495
9	CRISPY PIZZA TOAST	14	95	1330
10	LEMON SODA	15	60	900
11	VEGGIE LOVER PIZZA	5	159	795
12	LEMON RICE	1	95	95
13	EER TIKKA SANDWICH	60	95	5700
14	KESAR CHAI - SML	26	25	650
15	DOUBLE CHEESE PIZZA	64	129	8256
16	)N GINGER CHAI - SML	4	25	100
17	)COLATE COFFEE - SML	48	25	1200
18	WHITE SAUCE PASTA	27	109	2943
19	TANDOORI SANDWICH	21	75	1575
20	GREEN APPLE MOJITO	4	80	320
21	FRIED RICE	19	85	1615
22	KIWI ICE CRUSHER	2	60	120
23	MIX VEG	2	100	200

Analysed Data

SERIAL NO	PRODUCT NAME	QTY	PRICE	TOTAL AMOUNT	CATEGORY	Price Range	SALES RANK	QUANTITY RANK
2	ADRAK CHAI SMALL	516	20	10320	BEVERAGE	Low-Priced	1	1
97	NORMAL COFFEE SMALL	121	20	2420	BEVERAGE	Low-Priced	11	2
110	VEG CHEESE BURGER	105	70	7350	MAINCOURSE	Mid-Priced	4	3
83	PERI PERI FRIES	94	85	7990	APPETIZER	Mid-Priced	3	4
53	ALOO PATTIES	85	35	2975	APPETIZER	Low-Priced	8	5
4	VEG GRILLED SANDWICH	80	70	5600	MAINCOURSE	Mid-Priced	6	6
115	ELAICHI CHAI SMALL	77	25	1925	BEVERAGE	Low-Priced	16	7
15	DOUBLE CHEESE PIZZA	64	129	8256	MAINCOURSE	High-Priced	2	8
13	PANEER TIKKA SANDWICH	60	95	5700	MAINCOURSE	Mid-Priced	5	9
72	PLAIN MAGGI	58	40	2320	APPETIZER	Low-Priced	12	10.5
62	MASKA BUN	58	30	1740	APPETIZER	Low-Priced	21	10.5
94	HAZELNUT COFFEE SMALL	55	30	1650	BEVERAGE	Low-Priced	23	12.5
57	MASALA CHAI SMALL	55	25	1375	BEVERAGE	Low-Priced	31	12.5
95	VEG BURGER	49	60	2940	MAINCOURSE	Mid-Priced	10	14
17	CHOCOLATE COFFEE SMALL	48	25	1200	BEVERAGE	Low-Priced	37	15
52	ROSE CHAI SMALL	41	20	820	BEVERAGE	Low-Priced	50	16
99	PAAN CHAI SMALL	40	25	1000	BEVERAGE	Low-Priced	39	17
36	STRONG COFFEE SMALL	34	25	850	BEVERAGE	Low-Priced	48.5	18
104	ALOO PRANTHA	31	50	1550	MAINCOURSE	Mid-Priced	28	19
82	GARDEN PIZZA	27	139	3753	MAINCOURSE	High-Priced	7	21
18	WHITE SAUCE PASTA	27	109	2943	MAINCOURSE	High-Priced	9	21
63	CARMAL COFFEE SMALL	27	30	810	BEVERAGE	Low-Priced	51	21
60	CHILLY CHEESE TOAST	26	85	2210	APPETIZER	Mid-Priced	13	24
54	CHEESE GARLIC BREAD	26	80	2080	APPETIZER	Mid-Priced	14	24
14	KESAR CHAI SMALL	26	25	650	BEVERAGE	Low-Priced	55	24
46	VEG PANEER BURGER	25	80	2000	MAINCOURSE	Mid-Priced	15	26
71	PLAIN COLD COFFEE	24	80	1920	BEVERAGE	Mid-Priced	17	27
67	FRENCH FRIES	23	80	1840	APPETIZER	Mid-Priced	18	28.5
32	GARLIC MASKA BUN	23	40	920	APPETIZER	Low-Priced	42	28.5
3	PANEER PATTIES	22	70	1540	APPETIZER	Mid-Priced	29	30.5
74	CHOCOLATE CHAI SMALL	22	20	440	BEVERAGE	Low-Priced	73	30.5
19	TANDOORI SANDWICH	21	75	1575	MAINCOURSE	Mid-Priced	27	32
80	ONION RINGS	20	90	1800	APPETIZER	Mid-Priced	19	33.5
50	LEMON ICE TEA	20	80	1600	BEVERAGE	Mid-Priced	26	33.5
47	OREO SHAKE	19	90	1710	BEVERAGE	Mid-Priced	22	35.5
21	FRIED RICE	19	85	1615	MAINCOURSE	Mid-Priced	25	35.5
91	ALOO CHILLI MASKA BUN	17	55	935	APPETIZER	Mid-Priced	41	37.5
5	LEMONADE	17	50	850	BEVERAGE	Mid-Priced	48.5	37.5
51	RED SAUCE PASTA	15	109	1635	MAINCOURSE	High-Priced	24	41
35	BLUEBERRY SHAKE	15	99	1485	BEVERAGE	Mid-Priced	30	41
41	CHEESE FRIES	15	90	1350	APPETIZER	Mid-Priced	32	41
10	LEMON SODA	15	60	900	BEVERAGE	Mid-Priced	44	41

Summary Statistics

- **Quantity Sales:** Averages around 22.5 units with significant variation. Ranges from very low (min = 1) to exceptionally high sales (max = 516).

- **Total Amount:** Averages approximately Rs. 1180, showcasing considerable variability. Spans from low (min = 10) to high values (max = 10320).

	SERIAL NO	QTY	PRICE	TOTAL AMOUNT	SALES RANK	\
count	116.000000	116.000000	116.000000	116.000000	116.000000	
mean	59.500000	22.508621	68.491379	1180.094828	58.500000	
std	33.630343	51.657483	32.596478	1708.108957	33.627111	
min	2.000000	1.000000	10.000000	10.000000	1.000000	
25%	30.750000	4.750000	40.000000	300.000000	29.750000	
50%	59.500000	9.500000	70.000000	550.000000	59.000000	
75%	88.250000	22.250000	86.250000	1498.750000	87.000000	
max	117.000000	516.000000	179.000000	10320.000000	116.000000	

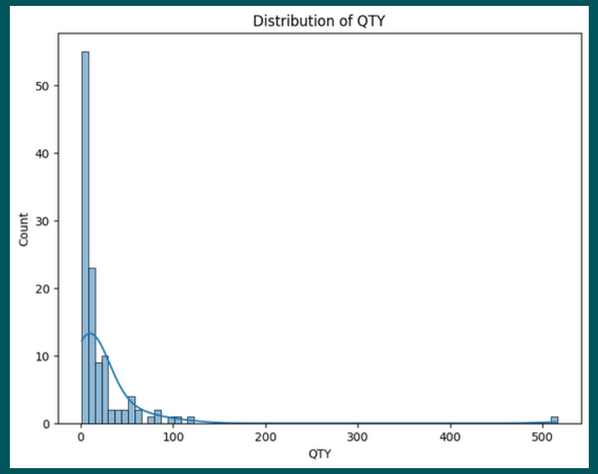
	QUANTITY RANK
count	116.000000
mean	58.500000
std	33.586229
min	1.000000
25%	30.000000
50%	58.500000
75%	85.000000
max	111.500000

- **Price Per Unit:** Averages about Rs. 68.49 with moderate pricing variation. Since the standard deviation is moderate and ranges from Rs. 10 to Rs. 179, illustrating diverse pricing.

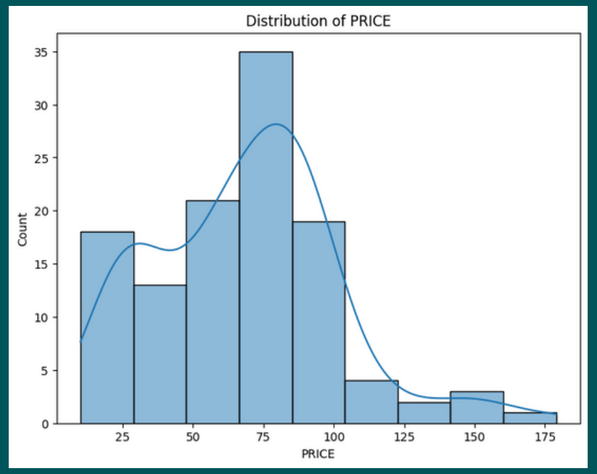
The presence of outliers in 'Quantity Sales' and 'Total Amount' suggests there might be a few items with exceptionally high demand and sales value.

Exploratory Data Analysis

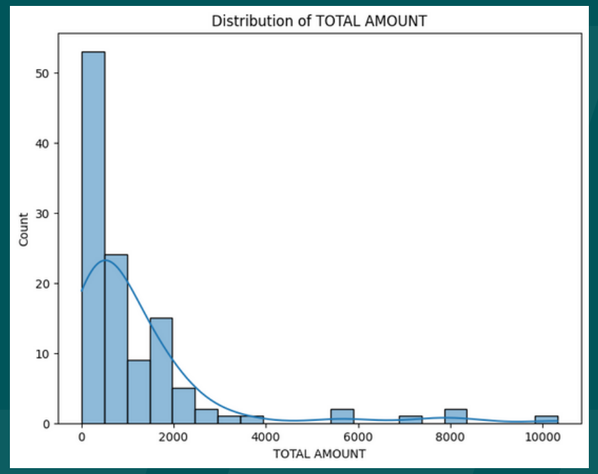
Distribution of Quantity



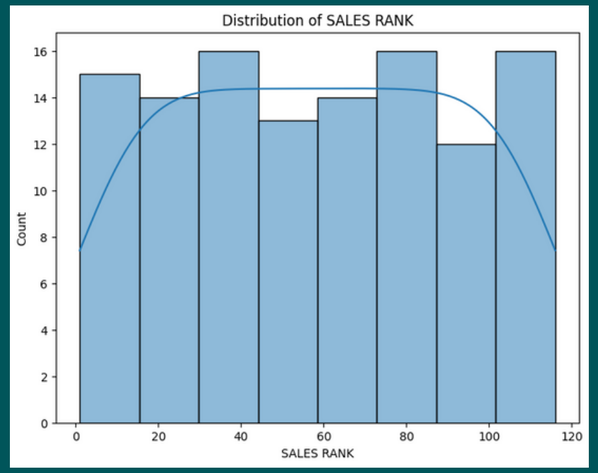
Distribution of Price



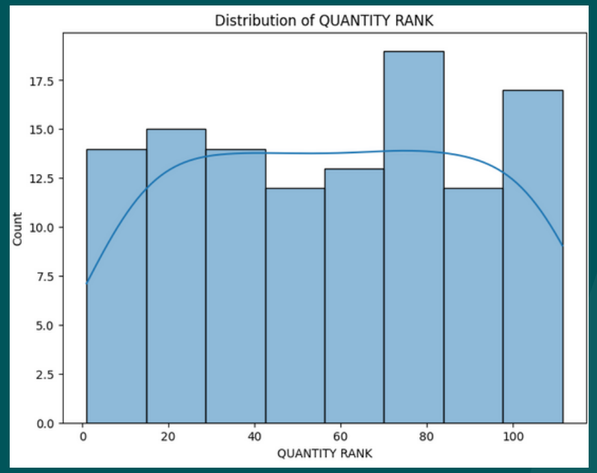
Distribution of Total Amount



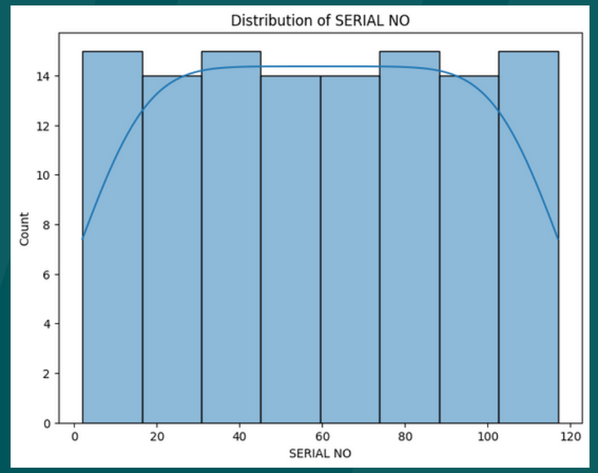
Distribution of Sales Rank



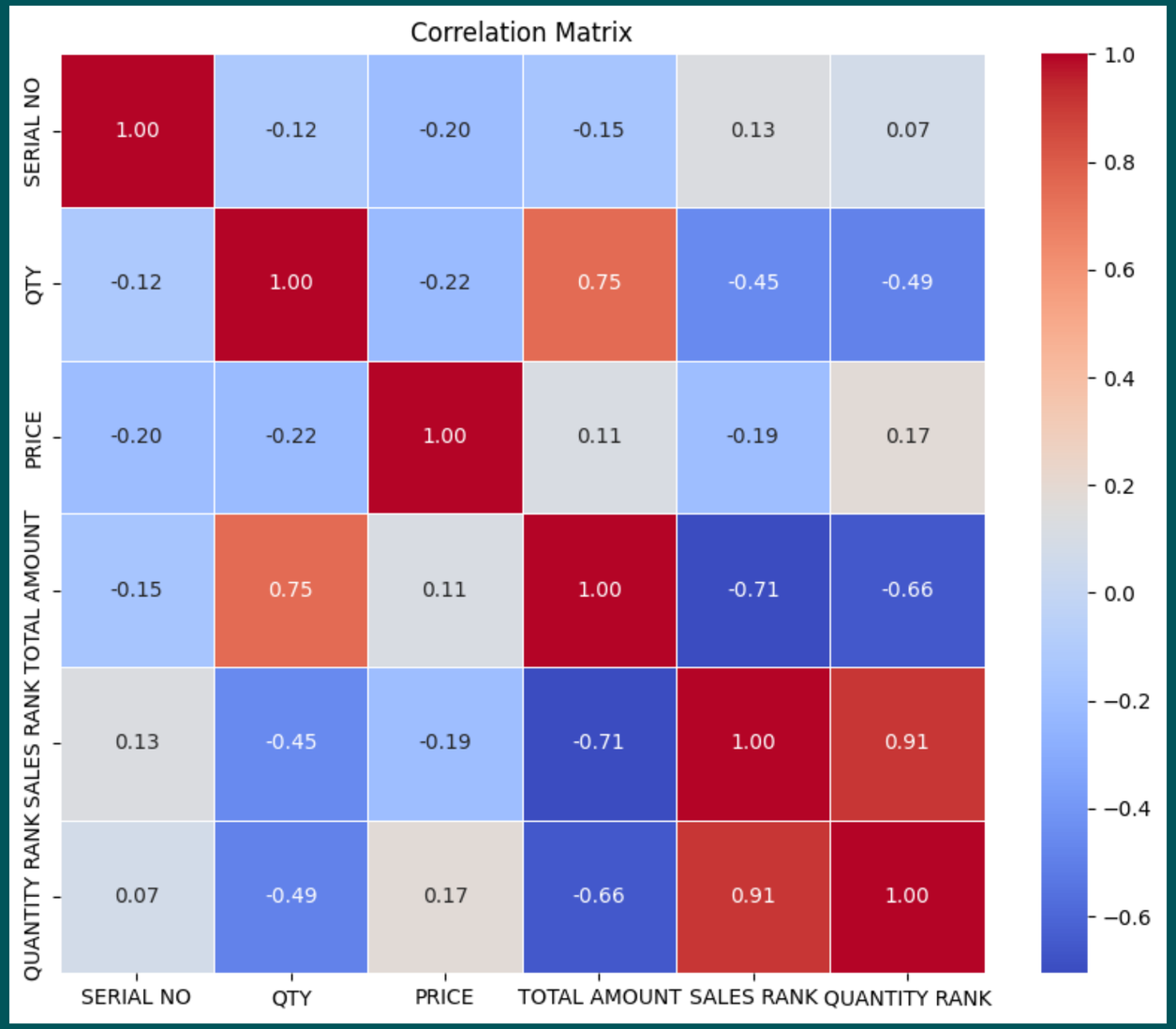
Distribution of Quantity Rank



Distribution of Serial No



Correlation Analysis



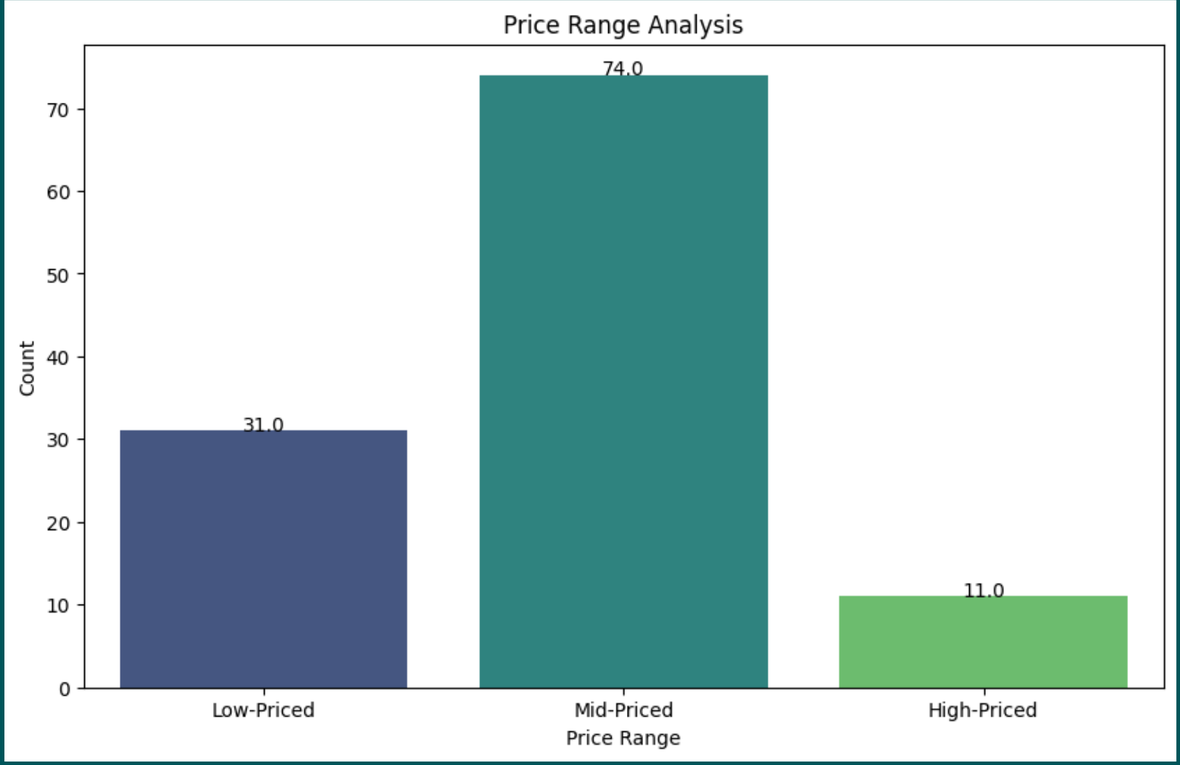
The negative correlation of -0.66 between 'Total Amount' and 'Quantity Rank' is significant. This suggests that changes in the quantity rank are associated with a considerable impact on the total amount, offering valuable insights for decision-making.

- **Decision-Making Strategy:** Utilize the negative correlation between 'Total Amount' and 'Quantity Rank' to inform decisions on quantity-related changes. Consider adjustments such as promotions or discounts for items with lower quantity ranks to potentially boost overall revenue.

- **Strategic Inventory Management:** Leverage this correlation analysis to optimize inventory, ensuring that items with higher quantity ranks are adequately stocked to meet customer demand and maximize sales.

- **Business Performance Optimization:** Implement data-driven decision-making strategies in inventory, pricing, and marketing based on the analysis of correlations. This approach enhances overall business performance and contributes to increased customer satisfaction.

Price Range Analysis



- **Price Range Distribution:** "Mid-Priced" items have the highest count, followed by "low-priced" and "high-priced" categories.

Mid-priced items dominate the kiosk's menu, ensuring affordability with perceived value. This strategic approach aims for broad customer appeal, accommodating diverse budgets, and enhancing overall satisfaction.

- **Market Positioning:** The kiosk focuses on striking balance between, between affordability and perceived value appealing to a diverse customer range, accommodating different budget preferences and enhancing overall satisfaction.

A high count of mid-priced items indicates the kiosk's strategic focus on catering to a broad customer base with varying budget considerations.

- **Menu Diversification:** Addressing gaps in price range representation can guide menu expansion to attract diverse customer segments.

A lower representation of high-priced items suggests an opportunity to introduce new menu items catering to this range, potentially attracting a different customer segment.

- **Profit Optimization:** By analyzing the popularity of items in different price ranges, the owner can optimize the menu for both customer satisfaction and profitability.

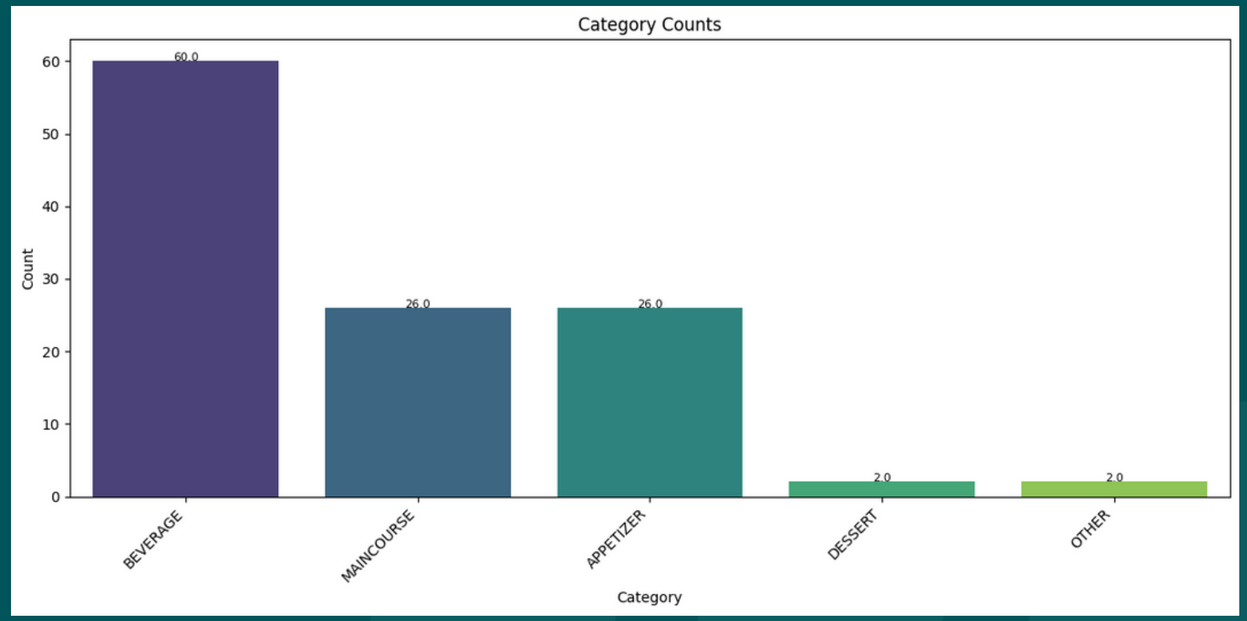
Category Counts

Menu Diversity: Diverse menu, especially in "Main Course" and "Beverages," attracts a wide customer base, potentially boosting customer retention.

Opportunities for Expansion: Expanding the limited "Desserts" category can attract sweet-toothed customers, enhancing overall sales.

Resource Allocation: Balancing resources based on category popularity optimizes inventory, marketing, and staffing decisions. Focusing on promoting the abundant "Beverages" items can be strategic.

Customer Preferences: Category counts reveal customer preferences, e.g., a high "Beverages" count indicates a preference for drinks with meals, guiding marketing and upselling strategies.



This output is significant because it provides insights into the diversity of the kiosk's menu. It helps the kiosk owner make decisions about the allocation of resources, marketing strategies, and potential expansions. For instance, if the kiosk owner sees that 'Desserts' has a lower count, they might consider introducing new dessert items to attract customers and boost sales in that category.

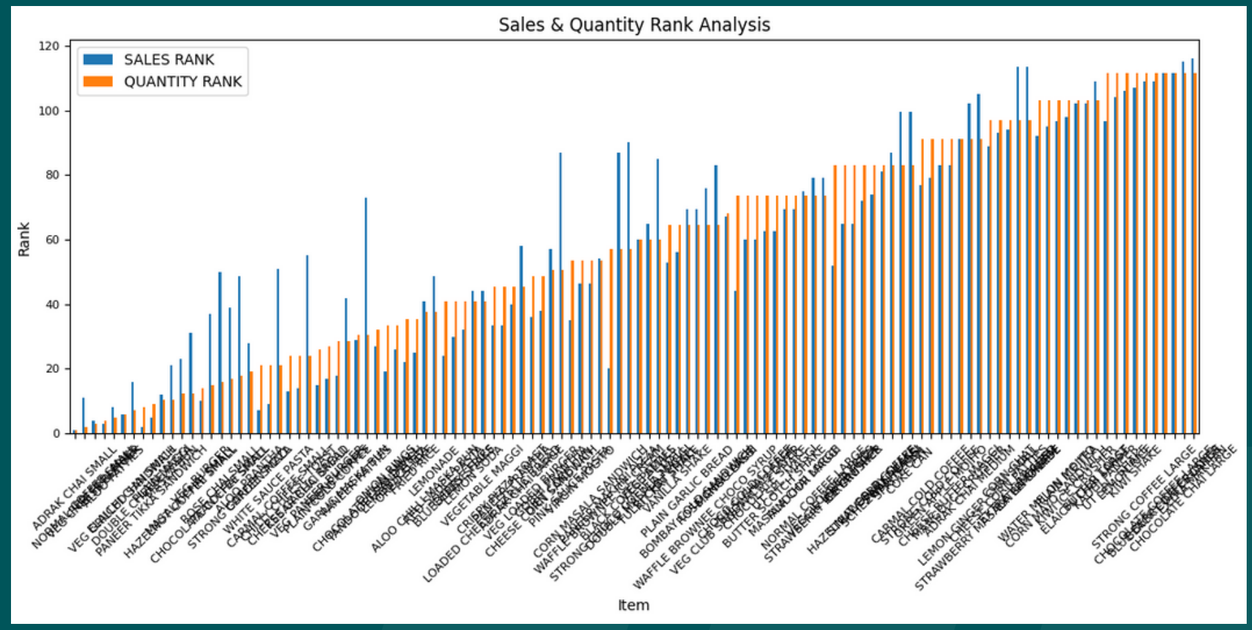
Sales & Quantity Rank Analysis

- **Resource Allocation:** The distribution of total sales rank guides resource allocation. The kiosk owner can focus marketing efforts, inventory management, and promotions on items in the most profitable sales and quantity rank ranges.

- **Customer-Centric Approach:** Understanding which items are top-sellers provides insights into customer preferences. This knowledge can lead to a more customer-centric approach in product offerings.

- **Stock Management:** The owner can optimize inventory management by reducing the stock of less popular items, preventing overstocking, and maintaining optimal levels of top-performing items.

- **Revenue Growth:** By strategically addressing the sales performance of items across different ranges of sales and quantity ranks, the kiosk owner can enhance overall revenue growth and profitability.



- The histogram comprises 232 bars, with each item having a blue and orange bar. Blue bars represent the Sales Rank, whereas orange bars represent the Quantity Rank.

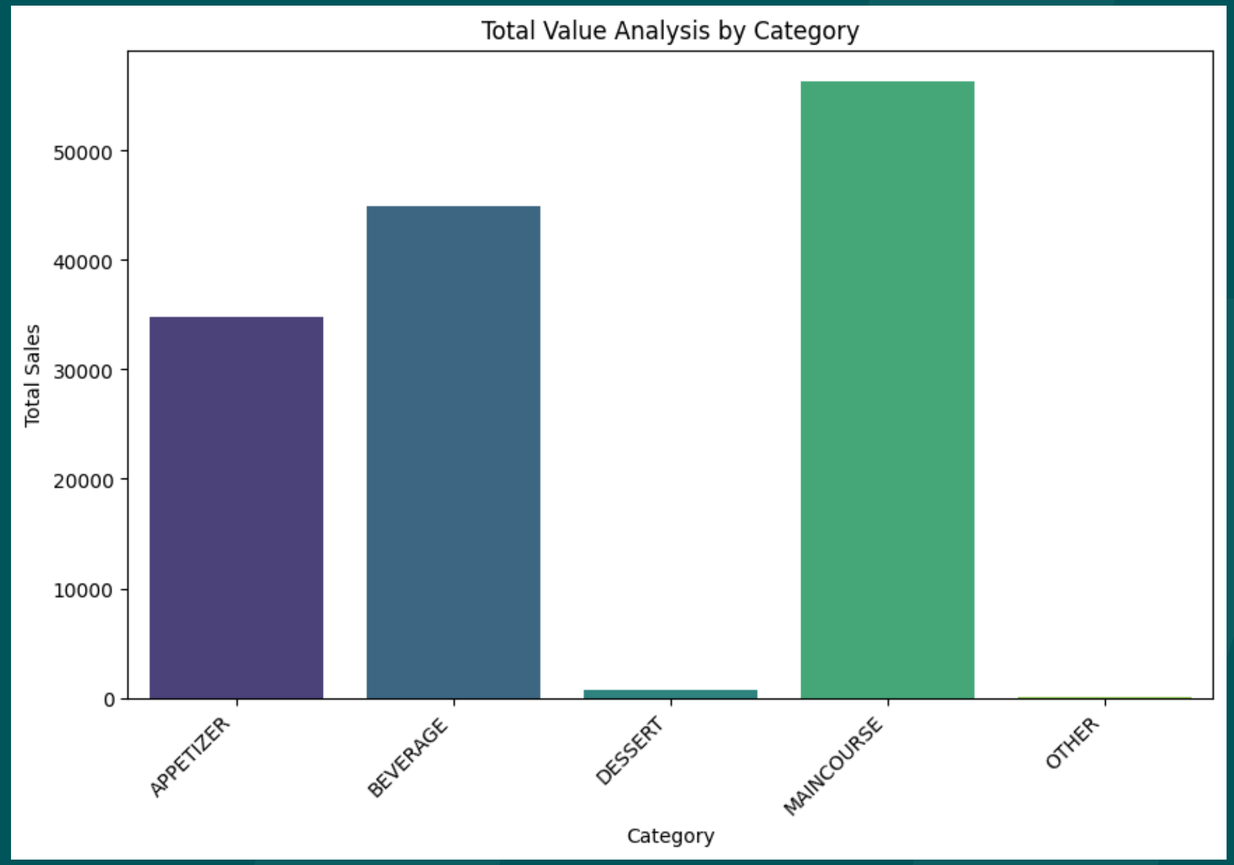
- The distribution indicates a substantial number of items in the low range of total sales ranks, suggesting they are moderately popular but not top-sellers.

Total Value Analysis by Category

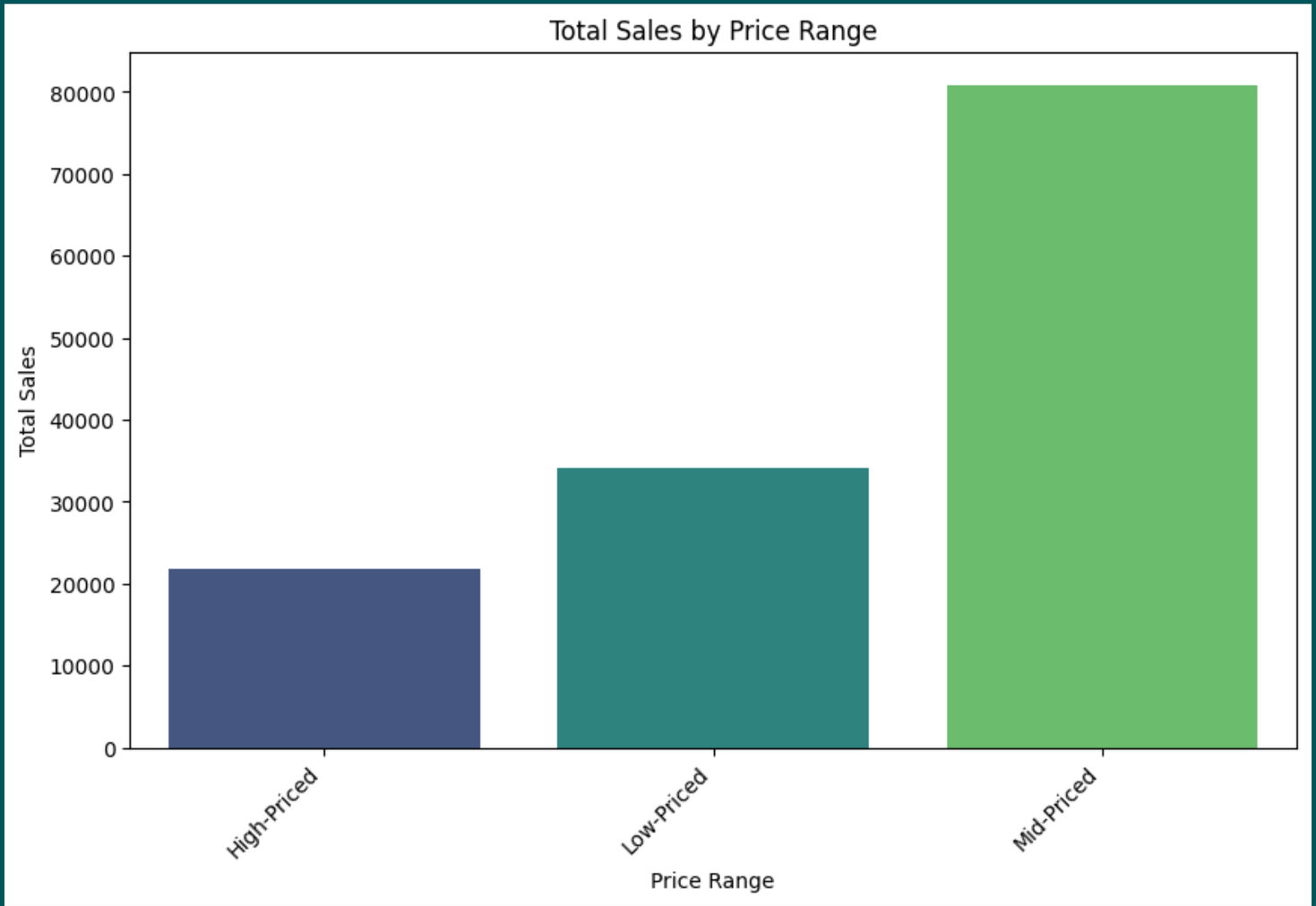
- **Strategic Decision-Making:** These inferences and conclusions guide the kiosk owner in making strategic decisions about product offerings. They provide clear insights into which categories to emphasize and develop.

- **Revenue Maximization:** By concentrating on high-value categories, the owner can maximize revenue and profitability. This data is invaluable for resource allocation.

- **Inventory Management:** The owner can optimize inventory management by ensuring that the most popular categories have sufficient stock while reducing unnecessary inventory costs in less-performing categories.



Total Value Analysis by Price Range



- **Pricing Strategy Optimization:** Guides the kiosk owner in optimizing pricing strategies. For example: promoting mid-priced items or expanding more options in this range.

- **Resource Allocation:** The kiosk can ensure adequate stock for top-performing price ranges and reduces excess inventory.

- **Maximizing Revenue:** Focuses on high-performing price ranges, the kiosk can maximize revenue and profitability.

- **Customer Preferences:** Provides insights into customer preferences based on price range, helping the owner understand what customers are willing to pay for.

ML Methodology

Pre-Processing

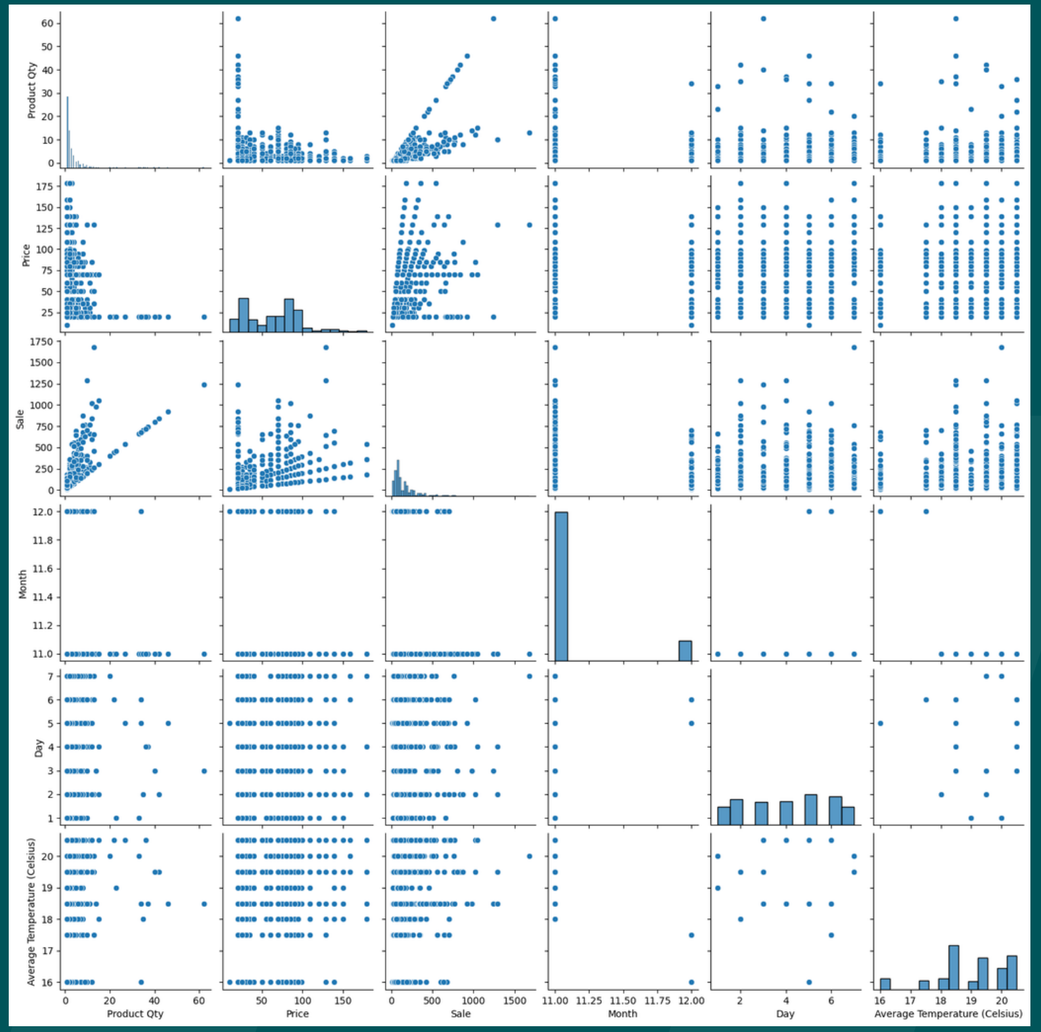
Columns	Work Done	Significance
Price	Added price for each product for each date.	Incorporating the 'Price' column allows for a detailed analysis of pricing trends, aiding in strategic pricing decisions and identifying optimal price points.
Sale	Computed sale for that specific product based product quantity and price.	The 'Sale' column provides insights into the sales performance of each product, guiding decisions on product promotions, discounts, and inventory management.
Month	Label Encoded Month	Labeling and encoding the 'Month' column facilitates seasonal analysis, enabling the identification of monthly patterns and trends.
Date (Reformatting)	Reformatted date from text to date format for better analysis	Reformatting the 'Date' column simplifies temporal analysis, aiding in the identification of trends, seasonality, and the impact of time on product performance.
Day	Label encoded day based on 7 days of the week	Label encoding the 'Day' column supports weekly analysis, providing insights into the popularity of products on different days of the week.
Average Temperature	Computed average temperature in Celsius	Incorporating the 'Average Temperature' column allows for analyzing the impact of weather on product sales, helping in seasonal stocking and marketing decisions.

ML Methods

Multi-Layer Perceptron: MLP, a type of neural network, is ideal for non-linear patterns and complex relationships. With 794 rows and 8 columns, MLP excels in capturing intricate dependencies, providing flexibility for diverse dynamics in product quantities. It's well-suited for scenarios where linear techniques fall short. MLP's adaptability is particularly beneficial for time-series data, incorporating temporal dependencies effectively.

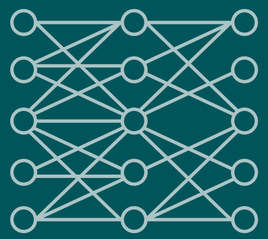
Gradient Boosting: GB, a robust ensemble method, excels in handling large datasets (794 rows) and capturing complex feature interactions. It sequentially improves predictions, making it effective for nuanced forecasting. With date-related features, GBR recognizes temporal patterns, providing valuable insights into product quantity dynamics over time.

Random Forest: Random Forest, an ensemble algorithm like GBR, is well-suited for handling non-linear relationships and large datasets (794 rows). It reduces overfitting, enhancing generalization through aggregated predictions from multiple trees. With diverse patterns in the data, RandomForest ensures robust predictions for product quantities. Date-related features enhance its ability to capture temporal variations, contributing to a comprehensive understanding of influencing factors over time.

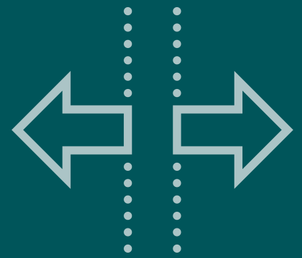


Multi-Layer Perceptron

Data Preparation



MLP Model Training

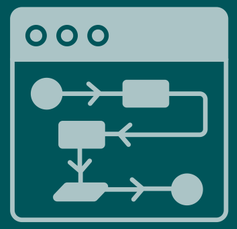


Prediction on Testing Data

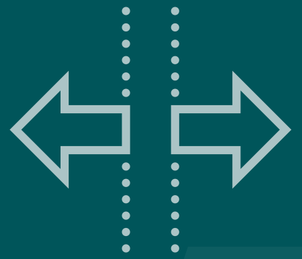
Future Date Predictions

Gradient Boosting

Data Preparation



Training the GB Model

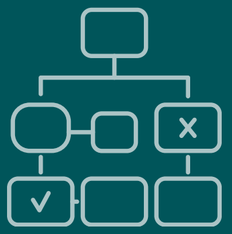


Prediction on Testing Data

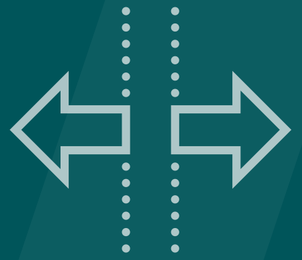
Future Date Predictions & Results

Random Forest

Data Preparation



Training the Random Forest Model



Prediction on Testing Data

Future Date Predictions & Results

Challenges

Inadequate Quantity of Data: Limited availability of data posed a challenge, restricting the model's ability to comprehensively learn patterns.

Lack of Data Quality: The data provided by Flavours of North Cafe was of poor quality, introducing noise and uncertainties into the training process.

Project Initiation Setback: The project initially started with Flavours of North Cafe, but due to its sudden closure, the team had to pivot to Uncle Tony's, causing a disruption in data collection and project continuity.

Data Collection Delay: Waiting for Uncle Tony's to collect sufficient data added to the project timeline, leading to delays in initiating and restarting the modeling process.

Limited Temporal Coverage: Training and testing the models on only 15 days of available data imposed temporal limitations, potentially hindering the model's understanding of long-term patterns.

Inefficiency of 80-20 Split: Despite using an 80-20 train-test split, the limited dataset might have affected the model's robustness, and a larger dataset would likely have provided more reliable insights.

Performance Metrics & Deployability

Performance Metrics

Multi-Layer Perceptron

```
Model Evaluation on Testing Data:
Mean Squared Error: 0.9887358490566036
Mean Absolute Error: 0.24528301886792453
R-squared: 0.9661449984559791
```

```
Predictions for the future date (2023-12-03):
          Product Name  Prediction
0      ADRAK CHAI LARGE        1.02
1     ADRAK CHAI MEDIUM        2.00
2      ADRAK CHAI SMALL       13.12
3       CHOCOLATE SHAKE        1.01
4   NORMAL COFFEE SMALL        7.12
..                  ...         ...
88           COKE CAN         1.14
89        ALOO PRANTHA       12.31
90        GOBHI PRANTHA        4.21
91              MIX VEG        2.00
92       PANEER PRANTHA        6.75

[93 rows x 2 columns]
```

Gradient Boosting

```
Model Evaluation on Testing Data:
Mean Squared Error: 0.40931066298589974
Mean Absolute Error: 0.1877682606889118
R-squared: 0.9859849188834474
```

```
Predictions for the future date (2023-12-03):
          Product Name  Prediction
0      ADRAK CHAI LARGE    1.000020
1     ADRAK CHAI MEDIUM    2.000000
2      ADRAK CHAI SMALL   13.000473
3       CHOCOLATE SHAKE    1.000005
4   NORMAL COFFEE SMALL    7.000022
..                  ...         ...
89           COKE CAN     1.000007
90        ALOO PRANTHA   12.999929
91        GOBHI PRANTHA    4.000000
92              MIX VEG    2.000000
93       PANEER PRANTHA    7.999911

[94 rows x 2 columns]
```

Random Forest

```
MLP Model Evaluation on Testing Data:
Mean Squared Error: 0.09521491476472052
Mean Absolute Error: 0.14063674725888692
R-squared: 0.9967397752499324
```

```
Predictions for the future date (2023-12-03):
                Product Name  Prediction_MLP
0          ADRAK CHAI MEDIUM        1.809260
1           ADRAK CHAI SMALL       13.217951
2         NORMAL COFFEE SMALL        7.045302
3   WAFFLE BROWNEE CHOCO SYRUP        1.097750
4       ALOO CHILLI MASKA BUN        2.022592
5                ALOO PATTIES        2.061846
6          BLACK COFFEE SMALL        3.608438
7               CHEESE FRIES        1.859426
8          CHILLY CHEESE TOAST        1.943943
9       CHOCOLATE COFFEE SMALL        1.763233
10       DOUBLE CHEESE PIZZA        4.906466
11        ELAICHI CHAI SMALL        4.703451
12              EXTRA CHEESE        1.199423
13              FRENCH FRIES        1.924768
14              GARDEN PIZZA        1.346855
15      HAZELNUT COFFEE SMALL        4.851659
16         KESAR CHAI SMALL        1.822317
17               OREO SHAKE        1.859426
18           PANEER PATTIES        4.946000
19      PANEER TIKKA SANDWICH        1.855480
20          PERI PERI FRIES        4.015175
21              PLAIN ICE TEA        1.005542
...
49              GOBHI PRANTHA        3.993146
50                    MIX VEG        1.863388
51             PANEER PRANTHA        7.587159
52              EXTRA MAKHAN        1.678481
```

Deployment Analysis

Solution Effectiveness: The solution is functional and yields results, but optimal deployment is recommended after the current semester concludes to leverage a larger dataset for enhanced accuracy.

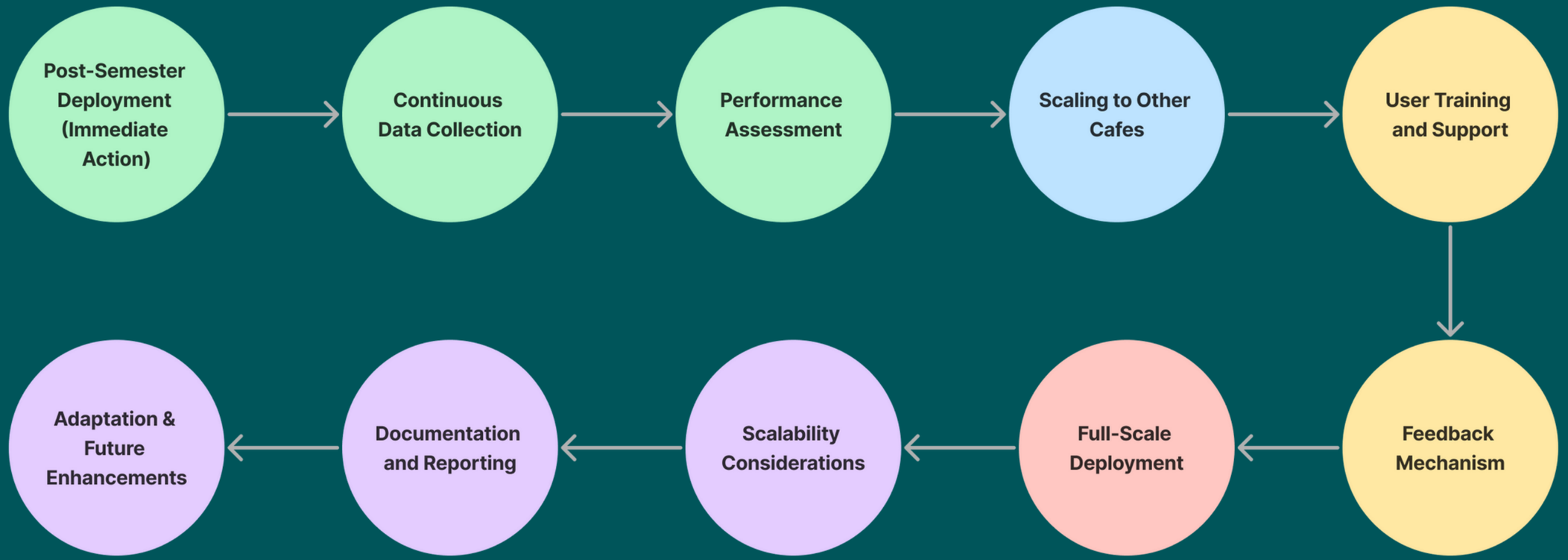
Data Collection Impact: Waiting for more data will provide a clearer understanding of patterns and behaviors, allowing for a more informed deployment decision.

Scalability: The solution exhibits scalability, allowing for swift deployment to newly inaugurated cafes like Arihant and Friends once sufficient data is available.

Project Efficiency: Compared to many other projects, our solution can be scaled rapidly, showcasing its efficiency in meeting high demand.

Deployment Consideration: The decision to deploy should be contingent on the volume of data available, ensuring a robust and reliable predictive model for cafe management.

Deployment Solution



Future Challenges

- **Increased Data Volume:** Handling a larger volume of daily data entries may strain the existing infrastructure.
- **Model Generalization:** Ensuring the predictive model generalizes well to diverse cafes with varying menu items and customer preferences.
- **User Adoption:** Cafes may face resistance or difficulties in adopting the solution, especially if there are changes in operational workflows.
- **Integration Complexity:** Integrating the solution into existing cafe management systems may be complex, leading to potential disruptions.
- **Cafe-Specific Adaptation:** Adapting the model to the unique characteristics of each cafe, such as seasonal trends or special events.
- **Data Quality Assurance:** Ensuring the ongoing quality and accuracy of the collected data, especially as the number of cafes and data points grows.
- **Predictive Model Performance:** Sustaining accurate predictions over time, especially if cafe dynamics change or new products are introduced.


Thank You!