JOHNNY & JUGNU

DATA SCIENCE PROJECT

Waleed, Mustafa, Shanza, Amna, Hafsa.

Unveiling JJ's Success: *Local Flavors, Global Impact*

- **Overview**: JJ's Menu Diversity & Consumer Appeal
- **Dataset**: October & November 2022 Insights

- Goal: Improving Operational Efficiency
- Opportunities Identified:
Sales Boost & Brand Development
- Recommendations: Service Enhancement, Product Alignment with Trends
- Economic Contribution: Fostering Local Brands for Growth
- Vision: Strengthening the Economy via Local Brand Support

THE GOAL

RESEARCH QUESTIONS

1) What is the most likely delivery/queuing time for a given order, and how can predictive modeling provide more accurate delivery and queuing time?

2) To what extent does the mode of payment influence the value of customer orders, and is there a correlation between specific payment methods and order values? Additionally, should JJ consider offering discounts for particular payment modes to enhance revenue generation strategically?

~~a) How do promotional campaigns on Instagram affect customer behaviour, and are their specific promotions that lead to a significant increase in orders or revenue?~~

~~3) What are the most common reasons for customer complaints or service issues and what measures can be implemented to address these concerns and enhance customer retention?~~

DATA CLEANING:

Ensuring Data Accuracy:
Chronologically Ordered the data then combined the data frames into one combined data frame and separate ones for each branch

Adding Our Own Columns:
Total Order Time, Processed_Items, Hour were some of the columns we added

Making Data Consistent:
We had separate entires titled “generic_creditcard” and “Credit Card” in “Payment method” and we made that consistent.

Daily Sales Totals Extraction:
Creation of 'Date' Column for Sales Analysis

```
Data Types of Columns:
Restaurant name                 string
Order ID                        string
Platform Key                    string
Restaurant ID                   string
Order status                    string
Created at                datetime64[ns]
Accepted at                     string
Cancelled at                    string
Cancellation reason             string
Cancellation owner              string
Rejected at                    float64
Rejection reason               float64
Dispatched at                   string
Delivered at                    string
Address                        float64
Transport type                  string
Payment type                    string
Payment method                  string
Food value                       int64
Packaging charges                int64
Discount                         int64
Voucher                          int64
Minimum order value fee          int64
Delivery Fee                     int64
Items                           string
Preorder                        string
Billable Status                 string
```

- Four JJ Branches in Lahore
- **Elements Covered:** Restaurant Info, Order Details, Location, Payments, Financials, Order Contents

DATA SET OVERVIEW

Cleaning Highlights:

- **Data Unification:** Combining Diverse Branch Data for Comprehensive Analysis
- **Temporal Analysis:** Enabling Sales Pattern Examination via Timestamp Conversion
- **Data Transformation:** Conversion of 'Created at' Timestamps to Datetime Format

Importance:

- Cross-Location Analysis Preparation
- Understanding Sales Patterns over Time

	Restaurant name	Order ID	Platform Key	Restaurant ID
0	Johnny & Jugnu - Johar Town	p110-00ff	FP_PK	p110
1	Johnny & Jugnu - Johar Town	p110-01s3	FP_PK	p110
2	Johnny & Jugnu - Johar Town	p110-011h	FP_PK	p110
3	Johnny & Jugnu - Johar Town	p110-01sg	FP_PK	p110
4	Johnny & Jugnu - Johar Town	p110-01hu	FP_PK	p110
...	...	...	...	...
1746	Johnny & Jugnu - Johar Town	p110-xx80	FP_PK	p110
1747	Johnny & Jugnu - Johar Town	p110-xxux	FP_PK	p110
1748	Johnny & Jugnu - Johar Town	p110-xxv5	FP_PK	p110
1749	Johnny & Jugnu - Johar Town	p110-xxv4	FP_PK	p110
1750	Johnny & Jugnu - Johar Town	p110-xxr8	FP_PK	p110

	Order status	Created at	Accepted at	Cancelled at
0	Delivered	2022-10-08 19:09	2022-10-08 19:09	NaN
1	Delivered	2022-10-19 19:35	2022-10-19 19:36	NaN
2	Delivered	2022-10-13 16:31	2022-10-13 16:31	NaN
3	Delivered	2022-10-19 00:59	2022-10-19 01:00	NaN
4	Delivered	2022-10-16 19:13	2022-10-16 19:13	NaN
...	...	...	...	...
1746	Delivered	2022-10-11 21:11	2022-10-11 21:22	NaN
1747	Delivered	2022-10-20 14:59	2022-10-20 15:00	NaN
1748	Delivered	2022-10-12 21:02	2022-10-12 21:02	NaN
1749	Delivered	2022-10-11 15:51	2022-10-11 15:51	NaN
1750	Delivered	2022-10-12 21:14	2022-10-12 21:25	NaN

	Cancellation reason	Cancellation owner	...	Payment method
0	NaN	NaN	...	COD
1	NaN	NaN	...	COD
2	NaN	NaN	...	generic_creditcard
3	NaN	NaN	...	Credit Card
4	NaN	NaN	...	generic_creditcard
...	...	...	...	...
1746	NaN	NaN	...	COD
1747	NaN	NaN	...	COD
1748	NaN	NaN	...	COD
1749	NaN	NaN	...	COD
1750	NaN	NaN	...	COD

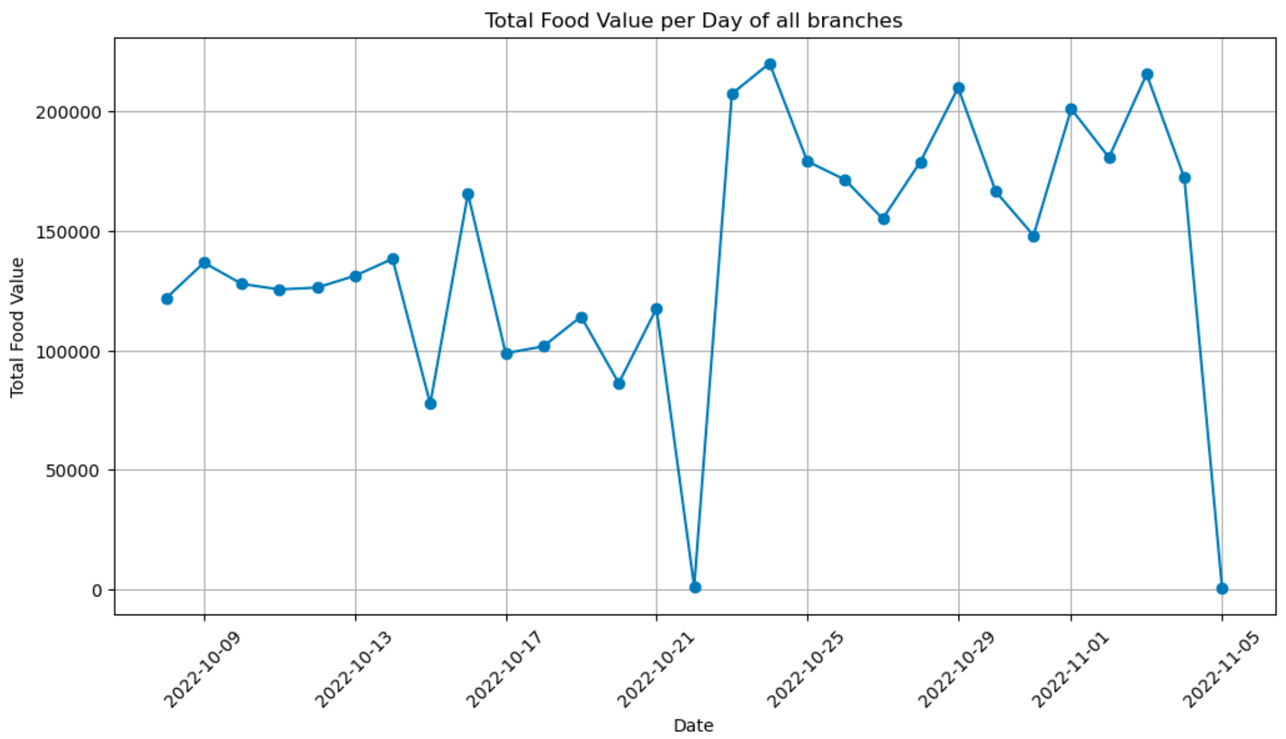
	Food value	Packaging charges	Discount	Voucher	Minimum order value	Fee
0	470	0	0	0		0
1	700	0	0	0		0
2	770	0	0	0		0
3	490	0	0	0		0
4	710	0	0	0		0
...	...	...	...	...	...	...
1746	3760	0	0	0		0
1747	1050	0	0	0		0
1748	670	0	0	0		0
1749	890	0	0	0		0
1750	900	0	0	0		0

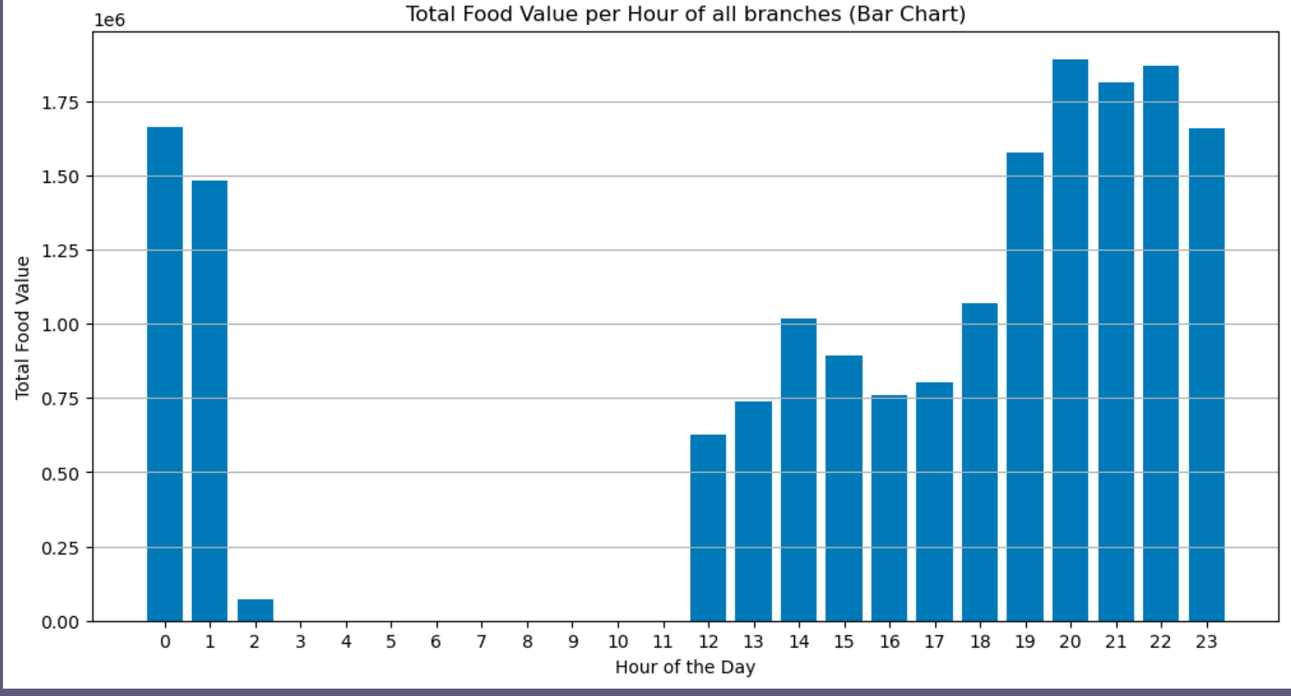
	Delivery Fee	Items
0	110	1 SMOL WRAP [1 CHIPOTLE, 1 ATOMIC, 1 NOT REQUI...
1	140	1 MEHSHI [1 CHIPOTLE, 1 CHIPOTLE, 1 NOT REQUIR...
2	110	1 CHICKEN FILLET [1 GARLIC, 1 GREEK, 1 NOT REQ...
3	110	1 NUGG WRAP [1 ATOMIC, 1 GREEK, 1 NOT REQUIRED...
4	110	1 TORTILLA WRAP [1 GARLIC, 1 GREEK, 1 NOT REQU...

Key Findings:

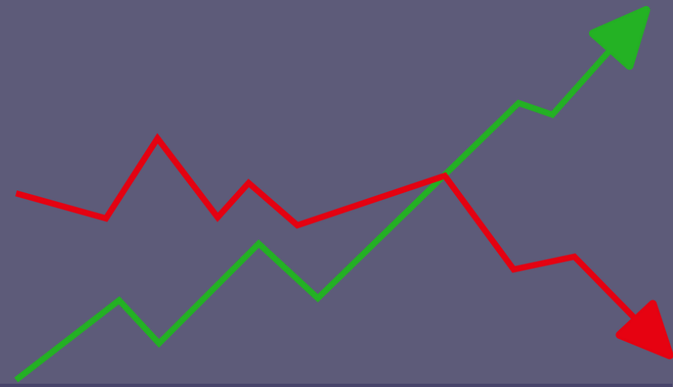
A. Exploratory Data Analysis

1. Order Volume Trends
2. Branch Performance
3. Payment methods





TOTAL FOOD VALUE ANALYSIS BY DAY OF THE WEEK



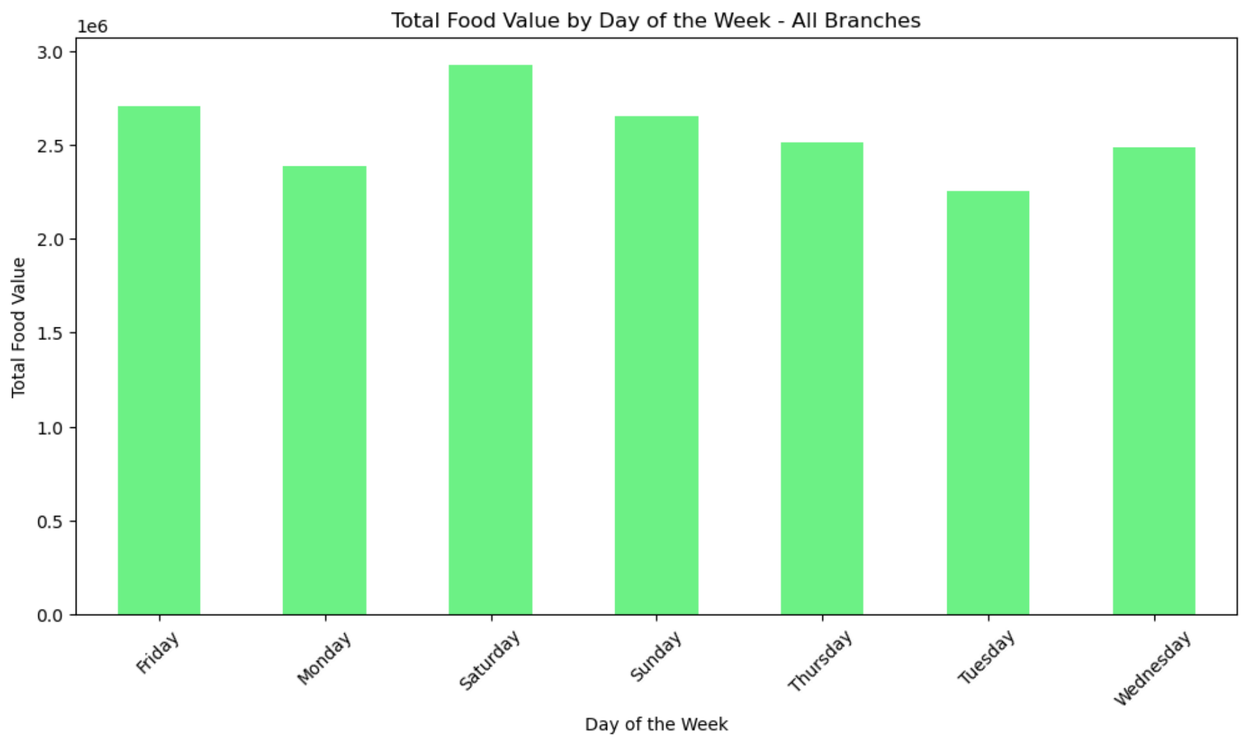
All Branches:
Mean: PKR 2,562,310
Std Dev: PKR 221,703
Min: PKR 2,257,570 | Max: PKR 2,926,830
25th: PKR 2,437,510 | 50th: PKR 2,511,550 | 75th: PKR 2,682,600

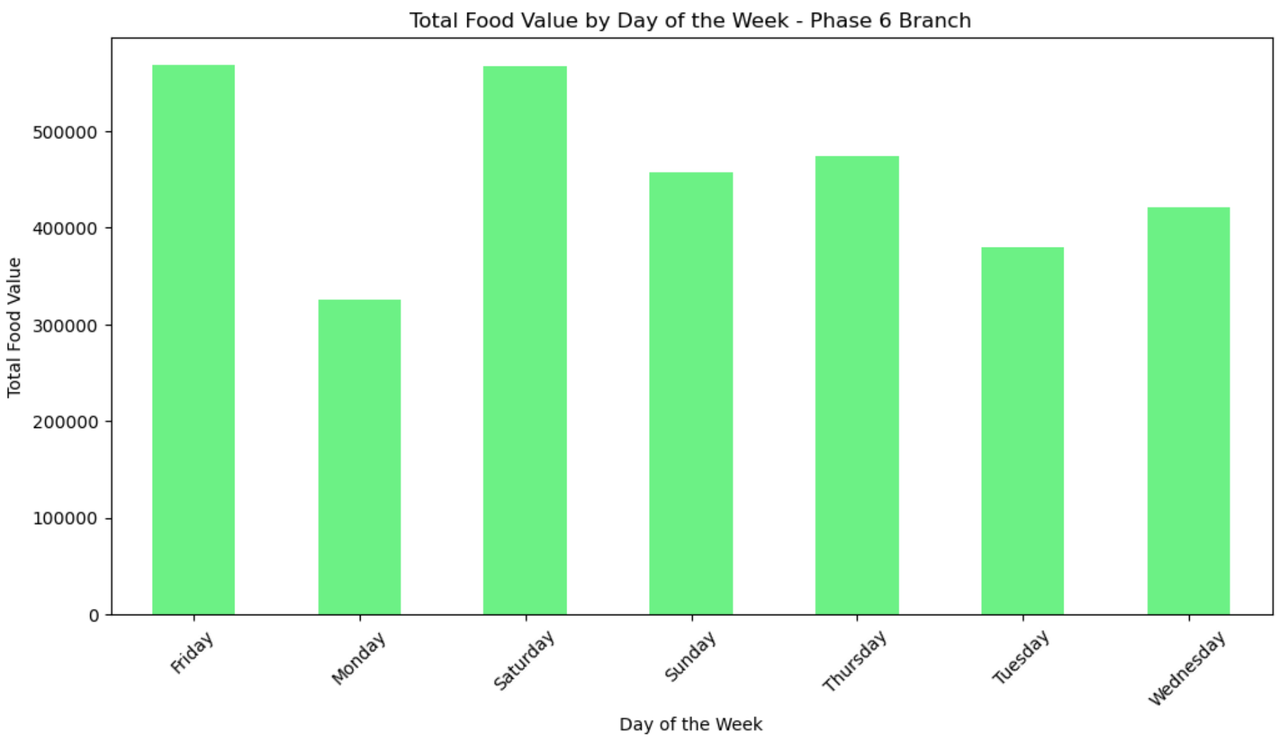
Bahria Town Branch:
Mean: PKR 474,566 | Std Dev: PKR 41,139
Min: PKR 432,600 | Max: PKR 532,500

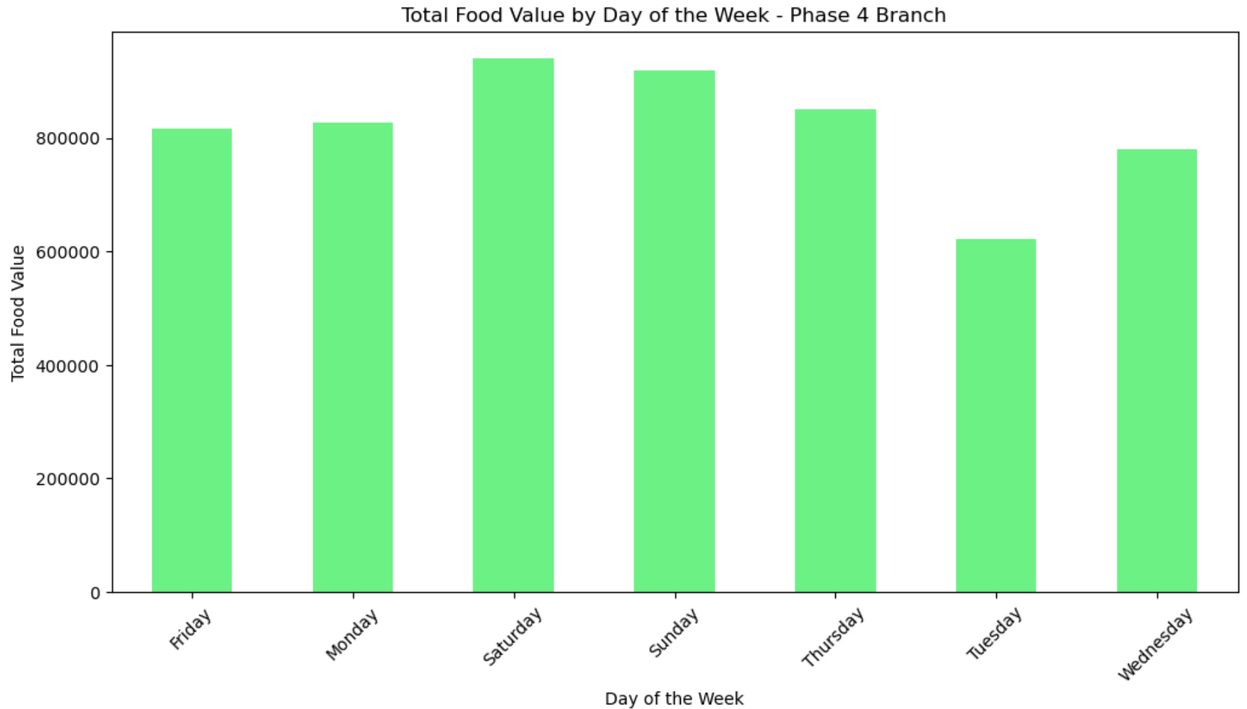
Phase 4 Branch:
Mean: PKR 822,015 | Std Dev: PKR 105,010
Min: PKR 621,680 | Max: PKR 940,330

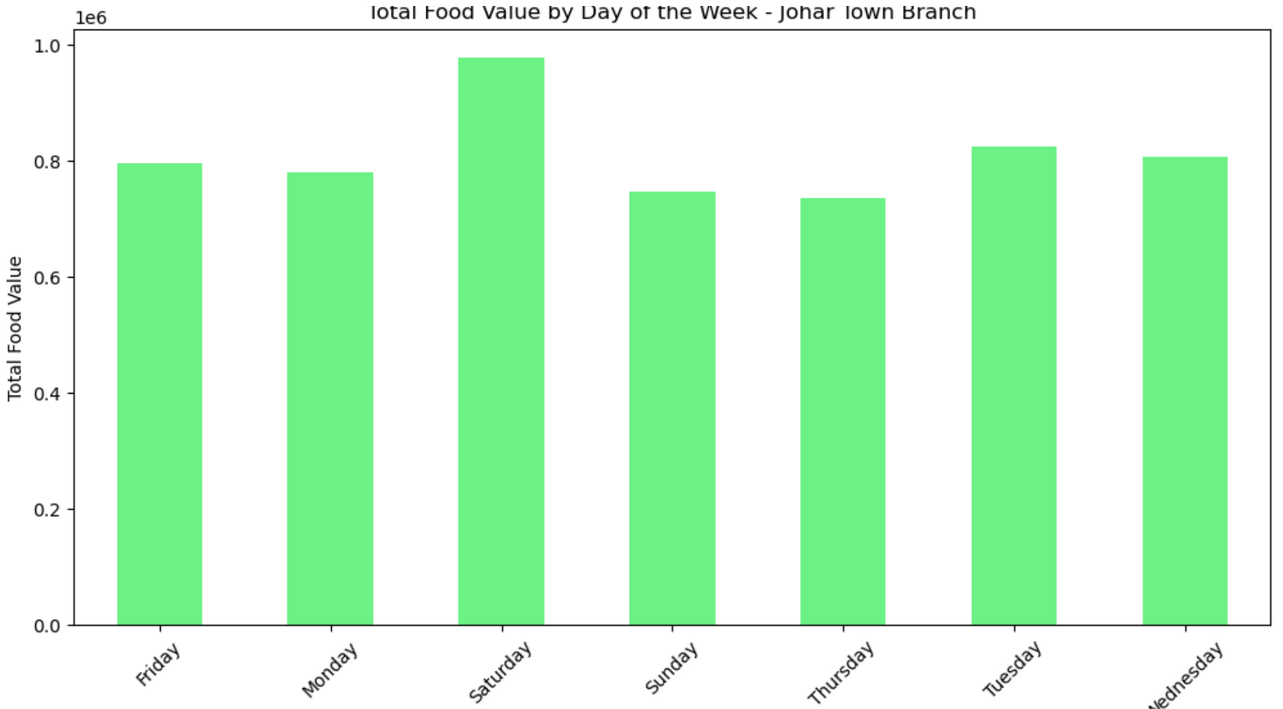
Phase 6 Branch:
Mean: PKR 456,313 | Std Dev: PKR 90,932
Min: PKR 324,990 | Max: PKR 568,430

Johar Town Branch:
Mean: PKR 809,416 | Std Dev: PKR 80,687
Min: PKR 735,010 | Max: PKR 977,880


Total Food Value by Day of the Week - All Branches
1e6
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Total Food Value
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
Day of the Week

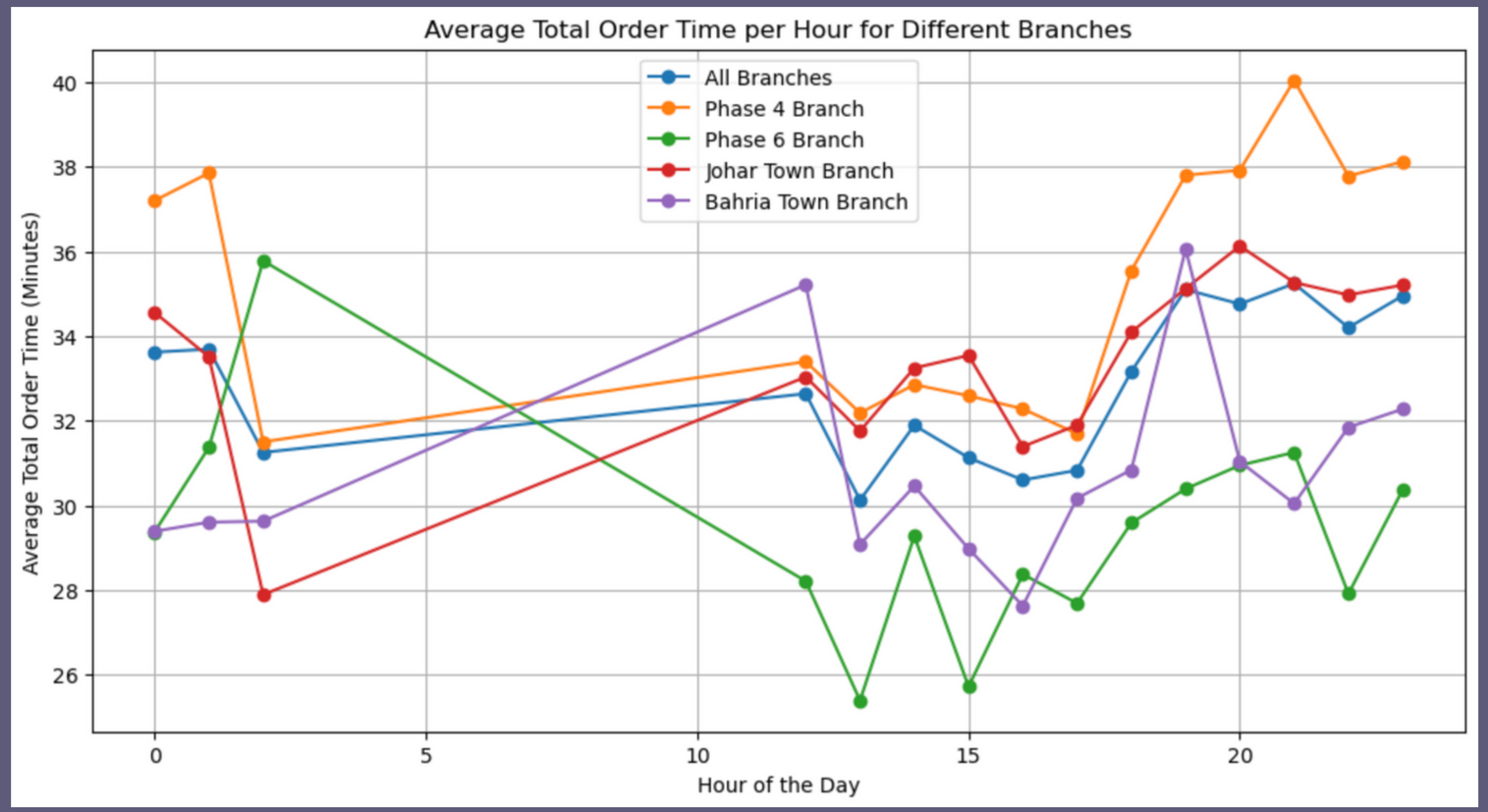

Total Food Value by Day of the Week - Phase 6 Branch
500000
400000
300000
200000
100000
0
Total Food Value
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
Day of the Week


Total Food Value by Day of the Week - Phase 4 Branch
800000
600000
400000
200000
0
Total Food Value
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday
Day of the Week


Total Food Value by Day of the Week - Johar Town Branch
1e6
1.0
0.8
0.6
0.4
0.2
0.0
Total Food Value
Friday
Monday
Saturday
Sunday
Thursday
Tuesday
Wednesday

ORDER TIME ANALYSIS

ORDERS ARE SLOWER TILL 2AM, QUICKER THROUGH THE DAY, SLOWER AGAIN AT NIGHT TIME. (NO ORDERS AFTER 2AM TILL 12PM)



Slowest: Phase 4 branch
Overall Average Time: 36.4 minutes
Peak order times extend into later evening, around 9 PM

Fastest: Phase 6 branch
Overall Average Time: 29.4 minutes
Consistent trend with a slight increase during the evening

PAYMENT METHODS

01 CASH ON DELIVERY

02 CREDIT CARD

03 BALANCE

03 INVOICE

VALUES

&

PERCENTAGES

COD
9768
Orders

60.5% of Orders

BALANCE
158
Orders

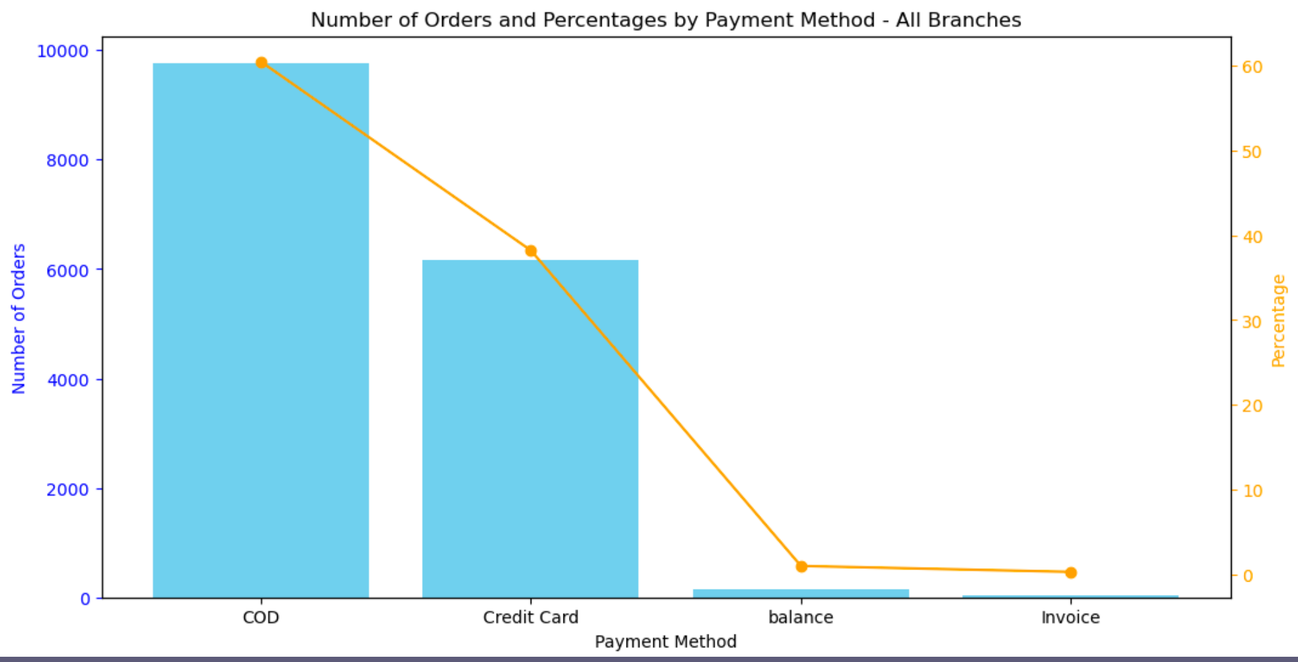
0.97% of Orders

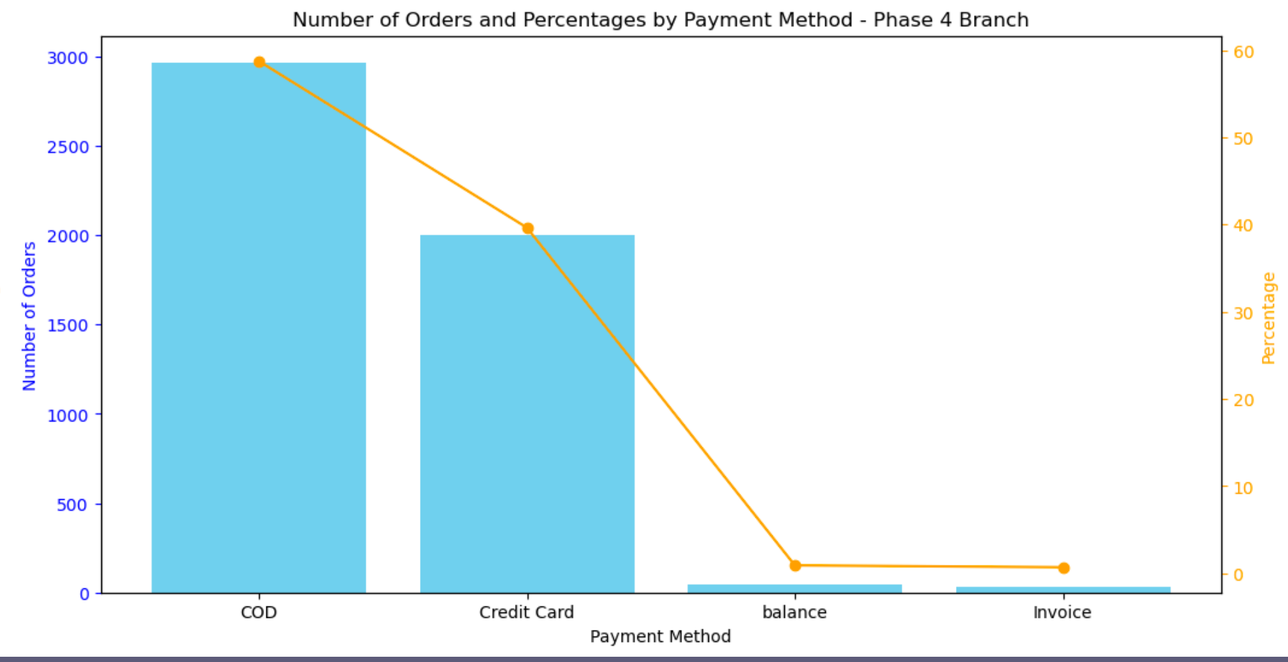
CREDIT CARD
6170 Orders

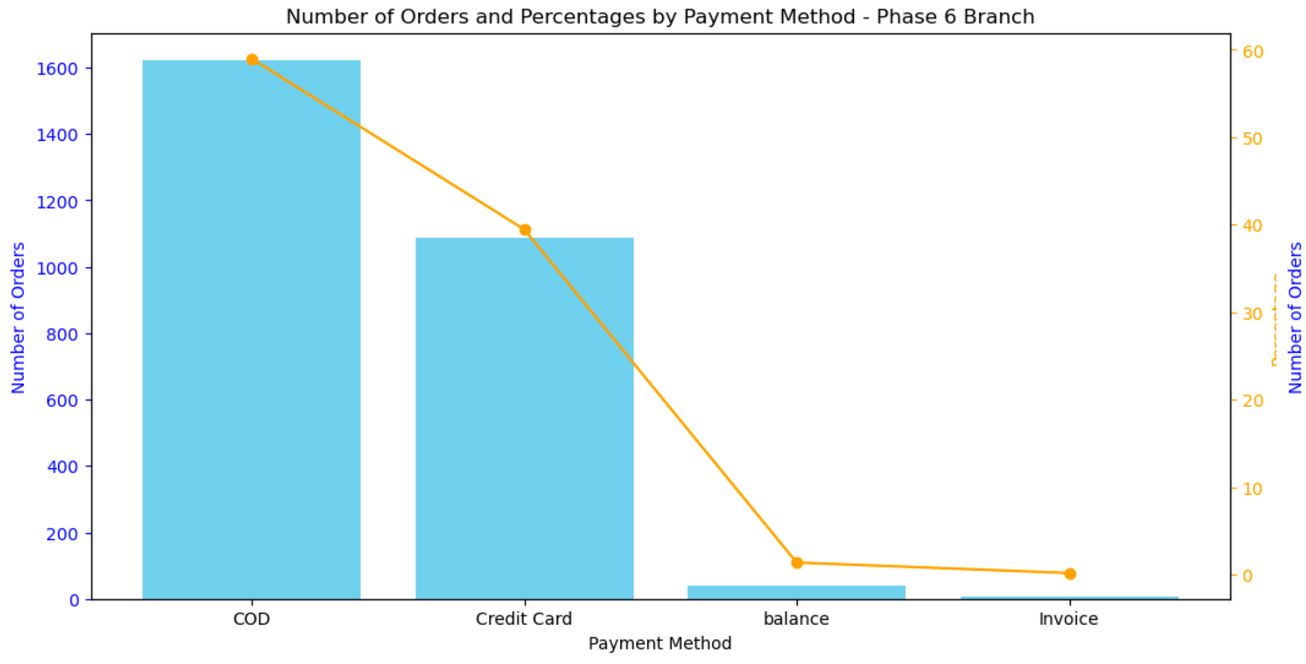
38.2% of Orders

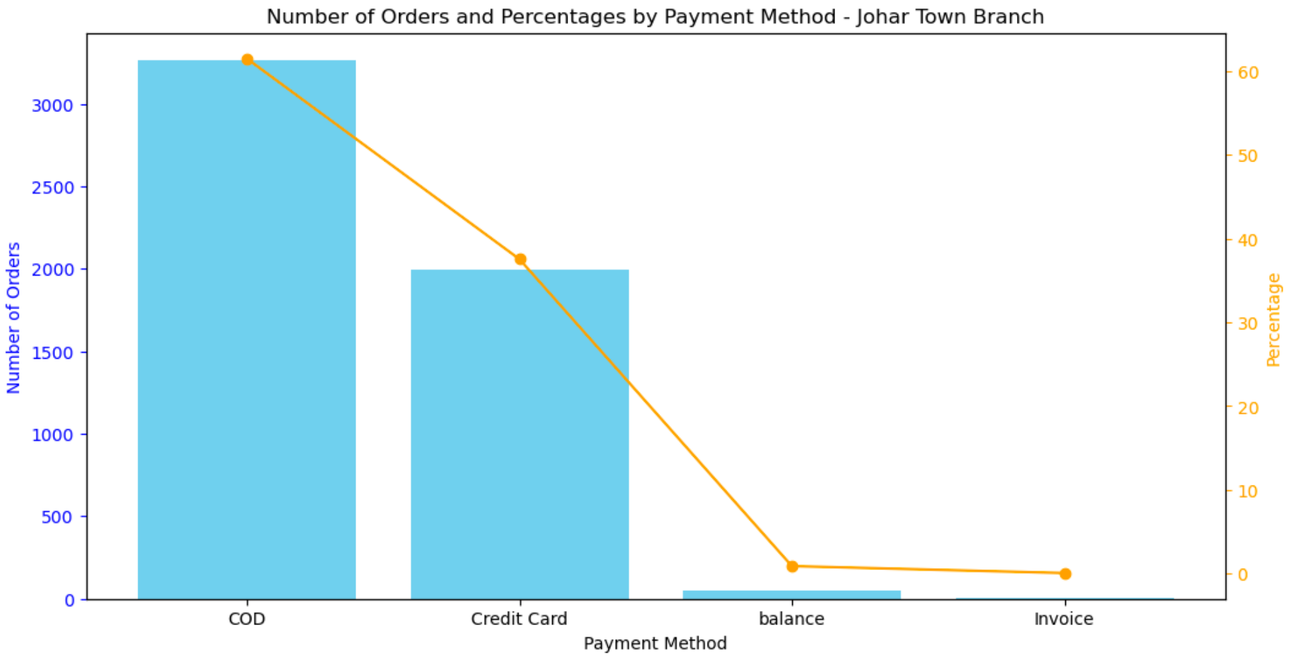
INVOICE
45
Orders

0.27% of Orders

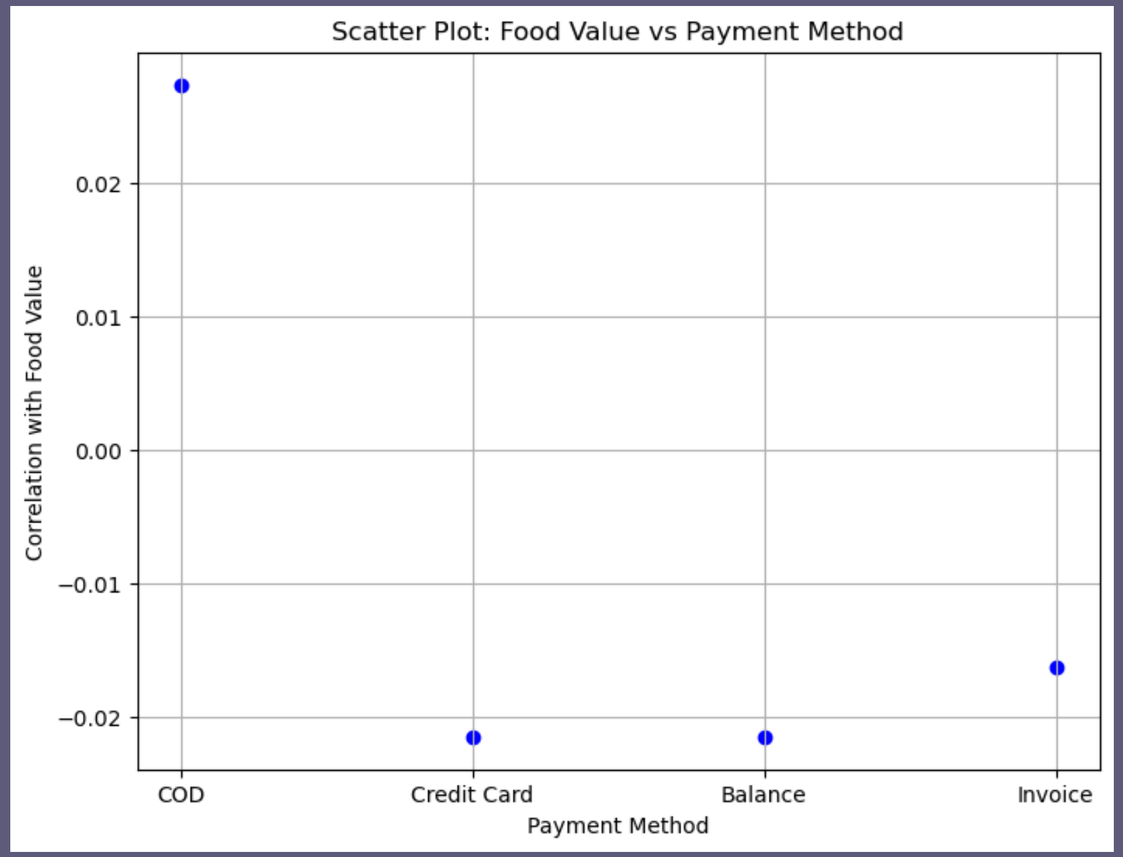

Number of Orders and Percentages by Payment Method - All Branches
Number of Orders
Percentage
Payment Method
COD
Credit Card
balance
Invoice


Number of Orders and Percentages by Payment Method - Phase 4 Branch
Number of Orders
Percentage
Payment Method
COD
Credit Card
balance
Invoice


Number of Orders and Percentages by Payment Method - Phase 6 Branch
Number of Orders
Payment Method
COD
Credit Card
balance
Invoice

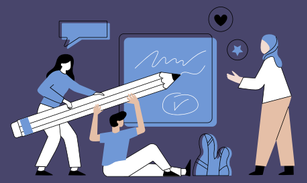

Number of Orders and Percentages by Payment Method - Johar Town Branch
Number of Orders
Percentage
Payment Method
COD
Credit Card
balance
Invoice

PAYMENT METHODS VS FOOD VALUE



- **COD (Cash on Delivery):**
  - **Positive correlation (0.0274): Indicates a slight positive relationship with food value.**
  - **On average, higher food values are associated with COD payments.**
- **Credit Card:**
  - **Negative correlation (-0.0214): Weak negative relationship with food value.**
  - **Suggests a minor tendency for lower food values with Credit Card payments.**
- **Balance:**
  - **Negative correlation (-0.0214): Similar weak negative relationship as Credit Card.**
  - **Implies a slight inclination for lower food values with Balance payments.**
- **Invoice:**
  - **Negative correlation (-0.0162): Weakest negative relationship among all methods.**
  - **Indicates a minimal tendency for lower food values with Invoice payments.**

CAUSAL INFERENCES FROM RESTAURANT DATA



EFFECT OF FOOD VALUE ON TOTAL ORDER TIME:

- **METHODOLOGY**: REGRESSION ANALYSIS CONTROLLING FOR HOUR OF THE DAY AND DAY OF THE WEEK.
- **FINDINGS:**
  - SIGNIFICANT POSITIVE CORRELATION BETWEEN FOOD VALUE AND TOTAL ORDER TIME.
  - HIGHER FOOD VALUE IMPLIES LONGER PREPARATION AND DELIVERY TIMES.
- **BUSINESS IMPLICATION:**
  - OPTIMIZE ORDER PROCESSING AND DELIVERY LOGISTICS BASED ON FOOD VALUE.
  - ADJUST MENU PRICING AND COMPOSITION FOR BALANCED DELIVERY TIMES.
  - IMPLEMENT DYNAMIC DELIVERY TIME ESTIMATES FOR IMPROVED CUSTOMER SATISFACTION.

HOUR OF THE DAY AND AVERAGE FOOD VALUE:

- **METHODOLOGY**: REGRESSION MODEL CONTROLLING FOR DAY OF THE WEEK AND SPECIAL PROMOTIONS.
- **FINDINGS**:
  - SIGNIFICANT IMPACT OF CERTAIN HOURS (LATE NIGHT, EARLY MORNING) ON AVERAGE FOOD ORDER VALUE.
  - INDICATES VARIATIONS IN CUSTOMER SPENDING HABITS DURING SPECIFIC TIME PERIODS.
- **BUSINESS IMPLICATION**:
  - INFORM TARGETED MARKETING AND MENU PLANNING FOR SPECIFIC HOURS.
  - TIME PROMOTIONAL ACTIVITIES TO COINCIDE WITH PERIODS OF HIGHER SPENDING.
  - AID IN INVENTORY MANAGEMENT AND STAFFING FOR DIFFERENT CUSTOMER DEMANDS AT VARIOUS TIMES.

Restaurant Data Dynamics

STATISTICAL TESTING

1. **Food Value and Delivery Time:**

- **Hypotheses:**
  H0: No correlation between food value and total order time.
  H1: Correlation exists between food value and total order time.

- **Analysis and Results:**
  Pearson correlation analysis: Significant correlation
  (Correlation Coefficient: 0.1797, p-value: 1.212e-114).
  Higher food value implies longer delivery times.

- **Implications**:
  Emphasizes the role of menu pricing and operational efficiency for larger orders

Food Value and Delivery Time:

```
H0: There is no correlation between the food value of an order and the total order time.
H1: There is a correlation between the food value of an order and the total order time.
```

Peak Hours Effect:

```
H0: The hour of the day does not affect the average food value of orders.
H1: The hour of the day affects the average food value of orders.
```

STATISTICAL TESTING

2. Peak Hours Effect on Average Food Value:

- **Hypotheses:**

H0: Hour of the day does not affect average food value.

H1: Hour of the day affects average food value.

- **Analysis and Implications:**

ANOVA test: Significant impact of hour of the day (F-value: 11.7438, p-value: 1.312e-27).

Customer spending patterns vary at different times, providing opportunities for targeted strategies.

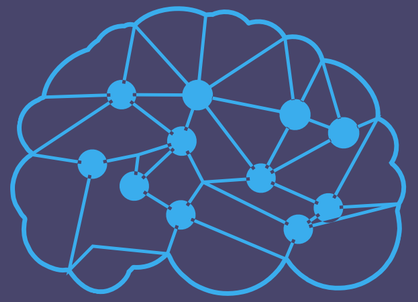

```
Pearson Correlation between Food Value and Total Order Time
Correlation Coefficient: 0.17975084965052468
p-value: 1.2126208411216282e-114
Hypothesis 2 is significant: There is a correlation between the food value of an order and the total order time.

ANOVA Test for Effect of Hour of the Day on Average Food Value
                sum_sq       df          F        PR(>F)
C(Hour)   7.707928e+07     14.0  11.743811  1.312644e-27
Residual  7.560094e+09  16126.0        NaN           NaN
Hypothesis 3 is significant: The hour of the day affects the average food value of orders.
```

ML MODEL: GRADIENT BOOSTING REGRESSOR

Objective: predicting total order delivery times

1. **Dataset preparation and features selection** such as food value, discount, and the number of orders in the queue. We employ techniques like imputation to replace missing values.
2. **Model Selection:** Gradient Boosting Regressor
3. **Loss Function:** We designed a custom MSE, which is a specialized version of the Mean Squared Error (MSE) in which predictions within a specified tolerance range are considered accurate.
4. **Model Performance:** 28% Accuracy with an MSE of 7.4.



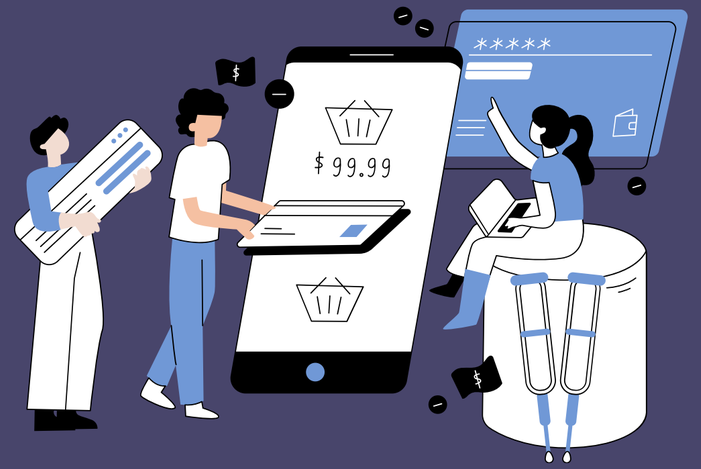
LIMITATIONS

Low Accuracy:

The low accuracy of our model could be attributed to several factors:

- complexity of the underlying patterns in the data
- insufficient feature representation
- The chosen features, while insightful, may not comprehensively capture the diverse factors influencing delivery times.
- potential missing features pose a gap in our research, possibly leading to low accuracy.

Our exploration into total order delivery time prediction uncovers not only the strengths of our model but also the inherent challenges in the dataset.



FUTURE DIRECTIONS

A. REAL-TIME DATA ANALYSIS
B. INTEGRATION OF CUSTOMER FEEDBACK DATA
C. CONTINUOUS LEVERAGING OF STATISTICAL ANALYSES FOR STRATEGIC PLANNING