

A FOODHUB COMPANY CUSTOMER DEMAND ANALYSIS FOR BUSINESSES

GROWTH: STUDY OF MOST FAVOURITE RESTAURANTS &

CUISINES

Project Foundations for Data Science:
Food Hub Data Analysis

Professional Development Program for Applied Data Sciences
By,

Dr. Uzma Abdullah

11.AUG, 2025

CONTENTS / AGENDA

- Executive Summary
- Business Problem Overview and Solution Approach
- Data Overview
- EDA - Univariate Analysis
- EDA - Multivariate Analysis
- Appendix

EXECUTIVE SUMMARY

The growing demand for restaurant's food and online food delivery service has increased extensively recently, specifically in the city of New York due to hectic and busy daily routine of students and professionals. The increasing online food demand pushed the businesses to do a deep analysis to improve their online customer satisfaction and boost their profits. This exploratory data analysis of 1,898 FoodHub orders across 178 restaurants and 14 cuisines reveals that the platform generated a net revenue of **\$6,166.30**, with **Shake Shack** leading as the top-earning restaurant (\$703.61 from 219 orders) and **American cuisine** contributing the highest total revenue (\$1,878.27). The objective of the analysis by the food aggregator company is to improve the businesses by evaluating the data of food Hub company and online delivery data of restaurants through exploratory data analysis method. The findings are the most revenue generating restaurant is Shake Shack, producing \$703.61 revenue from 219 orders and the top revenue producing cuisine is American making about \$1878.27 revenue. The order delivery performance highlights that the mean delivery time during weekday is 28.34 minutes, whereas mean delivery time during the weekend is 22.47 minutes. The order taking time above the 60 minutes is 10.54% which includes preparation plus delivery time. Data analysis is done by Python programming, the results shows that the overall highest revenue is produced by American Cuisine, while Shake Shack is the highest revenue-generating restaurant having high order count. It is seen that almost 89% of the orders are delivered within the time period of 60 minutes showing company efficient delivery during weekends. The conclusions of the analysis: primarily food delivery businesses should promote high revenue restaurants and cuisines specifically American, Japanese and Italian. Moreover, they should improve weekday delivery swiftness either by hiring more staff or improved delivery routes. The restaurants receiving low-rated vendor reviews should build credibility to attract customers and boost their contentment and engagement. Next, in order to maintain customer interest and orders frequency they should focus on marketing of popular cuisines. Finally, delivery time for orders exceeding than an hour should be targeted controlled and needs further optimization.

BUSINESS PROBLEM SUMMARY AND SOLUTION METHODOLOGY

Problem Definition:

The company Food Hub is an online food service provider, it is a platform which gets online food orders from customers to get from their multiple partner restaurants and cuisines, where their hired staff deliver food items to the customers. Although the company produces significant revenue, but still there are concerns about delivery efficiency, order preparation times, and customer satisfaction. The revenue distribution across the restaurants and cuisine is not smooth as some orders exceed the acceptable delivery threshold of 60 minutes. The FoodHub company necessities actionable steps to raise organisations, expand delivery performance, hence increases overall profit and adopts cost-effectiveness policies for the growth of the company.

Solution Approach / Methodology:

To get the solution to the problem, the methodology adopted is to collect the data for the key independent variables which are involved in marketing strategy. It includes Unique ID of the order, ID of the customer who ordered the food, name of the restaurant, Cuisine ordered by the customer, Cost of the order, day of the week: Indicates whether the order is placed on a weekday or weekend (The weekday is from Monday to Friday and the weekend is Saturday and Sunday), Rating given by the customer out of 5, food preparation time: Time (in minutes) taken by the restaurant to prepare the food. This is calculated by taking the difference between the timestamps of the restaurant's order confirmation and the delivery person's pick-up confirmation, delivery time (time in minutes) taken by the delivery person to deliver the food package. This is calculated by taking the difference between the timestamps of the delivery person's pick-up confirmation and drop-off information

The following step is to clean the data and prepare for the analysis. Missing values should be handled, the column formats must be standard and correct data is to be picked for the analysis which gives more comprehensive information to explore the study. Further, exploratory data analysis is performed to observe the trends in revenue, delivery time and customer responses/views about delivery and food are taken. It also includes to recognize the top rating restaurants and cuisines through customer rating and their profit generation. Additionally, the net revenue is computed as 25% commission is given for the orders which are greater than \$20 and 15% for the orders which varies between \$5-\$20 respectively. In order to check the metrics of delivery performance, the data is visualized for the orders which takes more than 60 minutes including food preparation and time taken for delivery. The purpose of the study is to emphasize the high-performance restaurants and cuisines and cover the grey areas in relation to delivery time and customer ratings to improve the business performance. The recommendations would be to enhance business functioning, improved marketing strategies and customer contentment.

DATA OVERVIEW

The dataset contains 1,898 rows and 9 columns, capturing detailed information about customer orders placed through FoodHub. It includes order IDs, customer IDs, restaurant details, cuisine

types, order costs, the day of the week, customer ratings, food preparation times, and delivery times. No missing values are present in the dataset after data validation.

The dataset has 1,898 rows and 9 columns.

Columns and Data Types:

- order_id — Integer (Unique identifier for each order)
- customer_id — Integer (Unique identifier for each customer)
- restaurant_name — String (Name of the restaurant)
- cuisine_type — String (Type of cuisine served)
- cost_of_the_order — Float (Cost in dollars)
- day_of_the_week — String (Day when the order was placed)
- rating — String (Customer's rating for the order)
- food_preparation_time — Integer (Time taken to prepare the food, in minutes)
- delivery_time — Integer (Time taken for delivery, in minutes)

Datatypes of the different columns in the dataset.

- order_id → int64
- customer_id → int64
- restaurant_name → object
- cuisine_type → object
- cost_of_the_order → float64
- day_of_the_week → object
- rating → object
- food_preparation_time → int64
- delivery_time → int64

There are no missing values in the dataset. If missing values had been found, numeric columns would have been filled with the median, and categorical columns with the mode.

STATISTICAL SUMMARY OF FOOD PREPARATION TIME

- Minimum time: 20 minutes
- Average time: 27.37 minutes
- Maximum time: 35 minutes

Zero orders are missing a rating. None missing values are present in the dataset after data validation. The dataset has 1,898 rows and 9 columns.

STATISTICAL SUMMARY TABLE

order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of_the_week	rating	food_preparation_time	delivery_time	
0	1477147	337525	Hangawi	Korean	30.75	Weekend	Not given	25	20
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend	Not given	25	23

2	1477070	66393	Cafe Habana	Mexican	12.23	Weekday	5	23	28
3	1477334	106968	Blue Ribbon Fried Chicken	American	29.20	Weekend	3	25	15
4	1478249	76942	Dirty Bird to Go	American	11.59	Weekday	4	25	24

The statistical summary shows the minimum time taken to prepare food is 20 min, the maximum time is 35 min, the average time is 27.371. To find the number of orders not rated, used 'value_counts()' on the rating column with 'dropna = False'. This make sure all values which are missing or not given are counted. The output shows 736 orders were not rated, 588 orders were given No.5 rating, 386 were given No.4 ratings, 188 orders were given No.3 rating.

EXPLORATORY DATA ANALYSIS (EDA)

Univariate Analysis

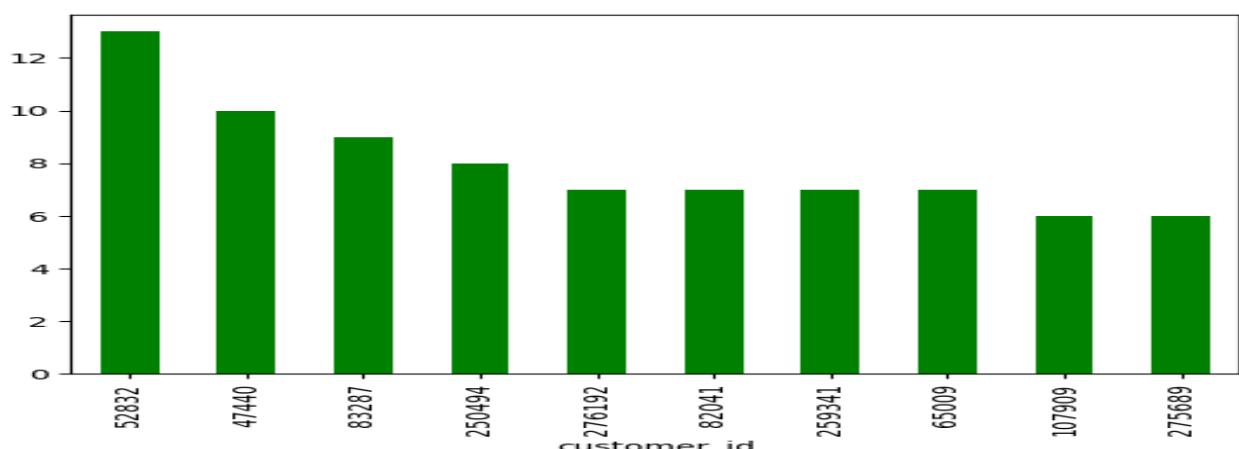
By using Python program, the data set have been checked having no missing values so no data cleaning or imputation is required. The statistical summary shows the minimum time taken to prepare food is 20 min, the maximum time is 35 min, the average time is 27.371

The unique number of IDs is 1898 which shows no duplicate orders. The unique customers ID is 1200, unique restaurants are 178 and the unique cuisine type are 14. The day of week are Weekend and Weekdays with the rating is categorised as “Not given”, 5, 4,3.

To find the number of orders not rated, the command 'value_counts ()' on the rating column with 'dropna = False' is used. This make sure all values which are missing or not given are counted. The output shows 736 orders were not rated, 588 orders were given No.5 rating, 386 were given No.4 ratings, 188 orders were given No.3 rating.

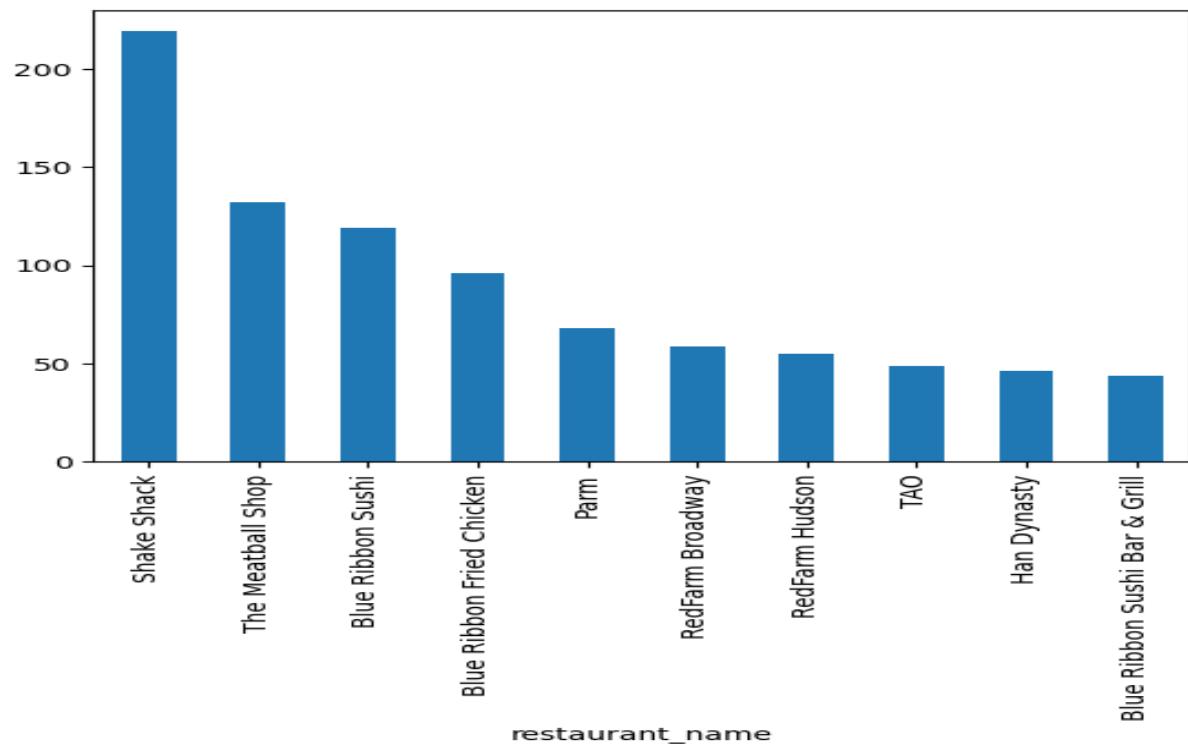
The number of individual orders which are placed are 1898 which helps to know about business metrics. The number of most frequently ordered customers which are 1200 in this data set. The above command shows frequently ordered customers which are 1200 in this data set.

TOP TEN ACTIVE CUSTOMERS:



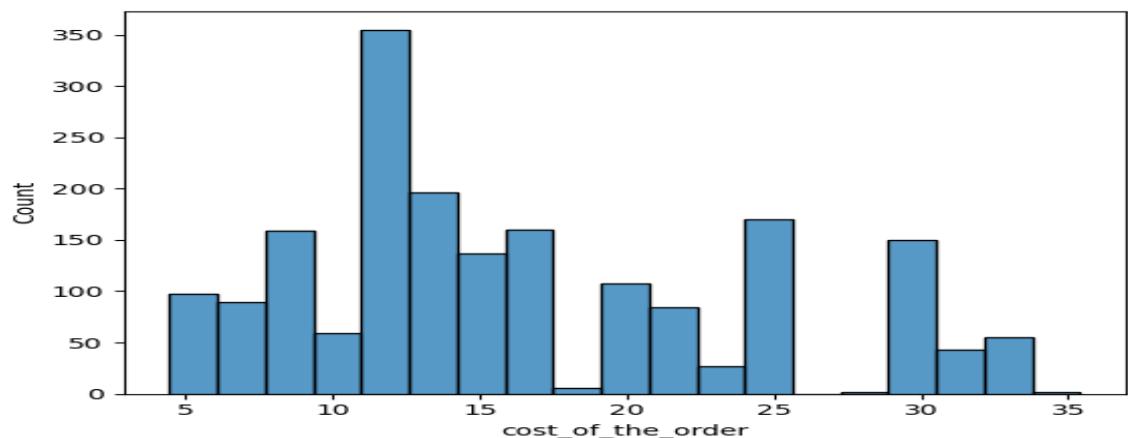
The above graph shows the Top Ten Active Customers by order count. The x-axis shows the number of orders by customers and y-axis shows the customers ID. The customer ID 52832 shows 13 orders making him most active customer. Similarly moving towards right in the graph shows the decline in orders by respective customers ID. The EDA suggests marketing strategies for the company by looking into customer purchasing behaviour.

BAR CHART FOR THE TOP 10 RESTAURANTS



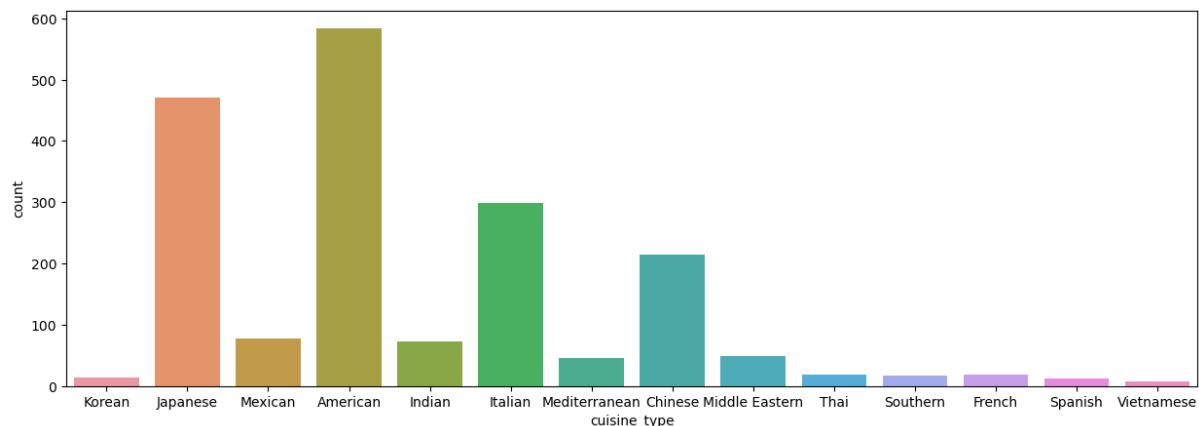
The x-axis shows the restaurant name and the y-axis shows the number of orders received by the restaurants. The highest number of orders are received by Shake Shack, followed by the meatball shop and Blue Ribbon Sushi.

HIGHEST CUSTOMER DEMANDING RESTAURANT



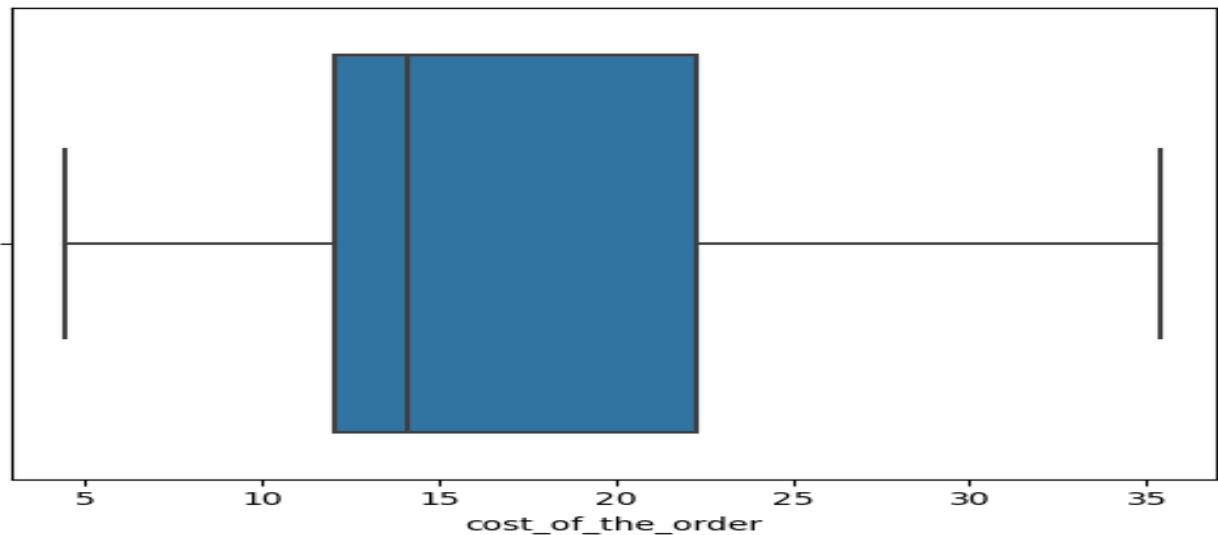
The x-axis shows the restaurant name the y-axis shows the number of orders received by the restaurants. The highest number of orders are received by Shake Shack, followed by the meatball shop and Blue Ribbon Sushi.

HIGHEST CUSTOMER DEMANDING CUISINE



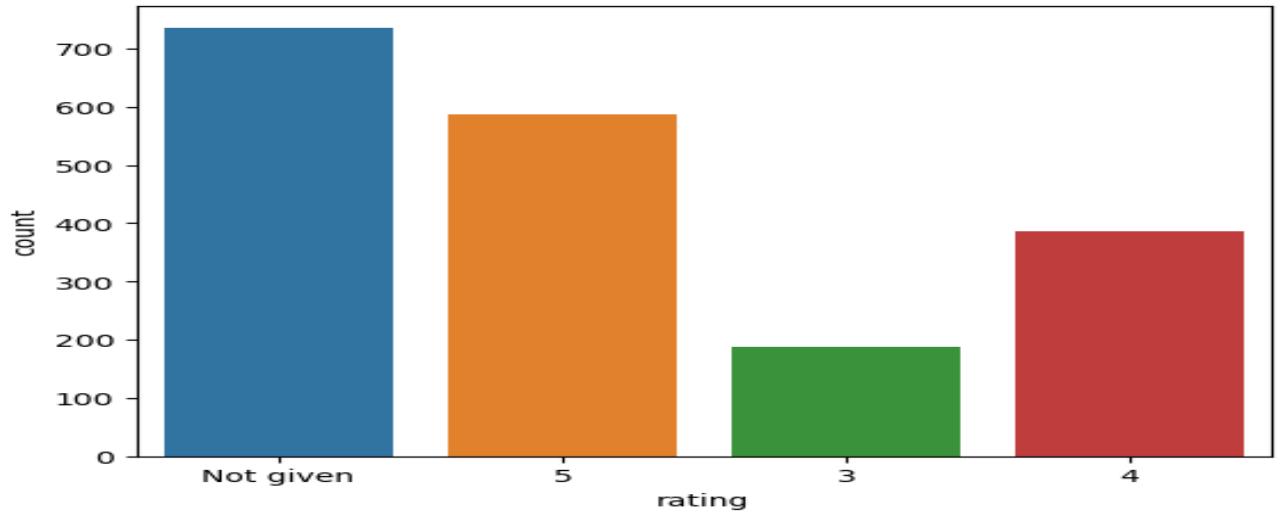
The most demanding food order is by American Cuisine. The height of cuisine bar shows the demand by customers, the one with tallest height has more demand as compared to smaller. This will assist the business to define marketing schemes and menu classification.

BOX PLOT FOR THE COST OF ORDER



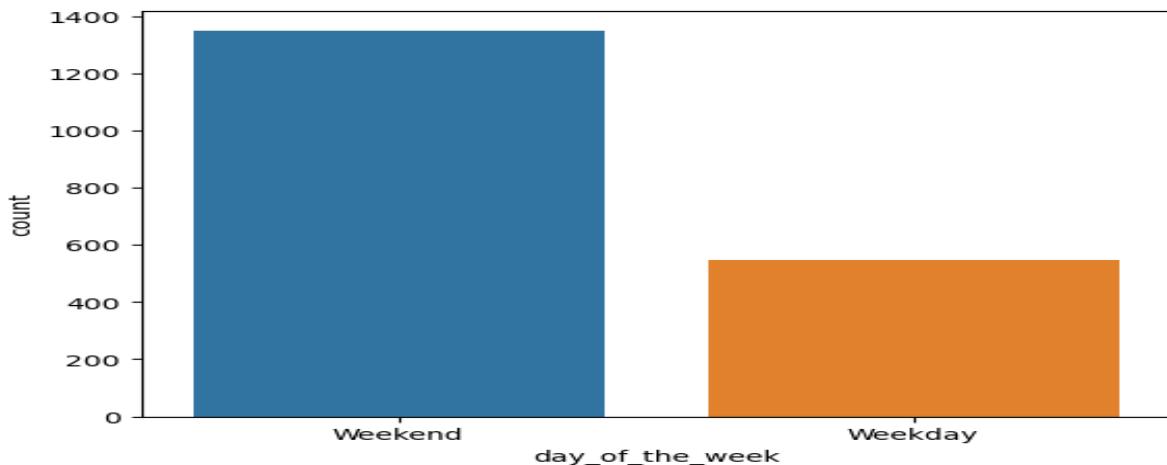
Most of the orders are clustered between \$15-\$20, the promotional offer for free delivery for order above \$ 20 can increase the sales. There is less variation in the data set reflecting the customer friendly pricing strategy by the restaurants.

HISTOGRAM FOR THE COST OF ORDER



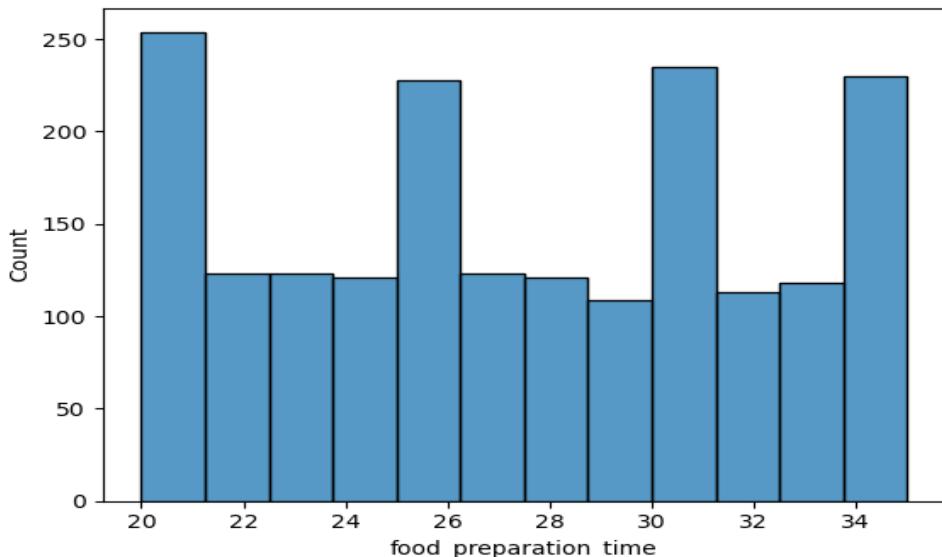
The highest bar shows rating not given, the other ratings show the business to improve service quality and other areas of improvement. The most demanding food order is by American Cuisine.

BAR CHART FOR THE WEEKDAYS AND WEEKEND COMPARISON



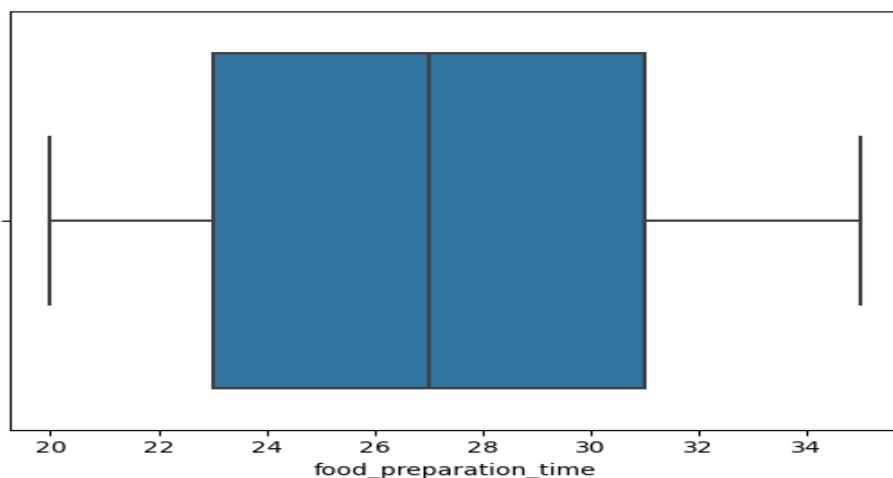
The day set is split into only weekend and weekdays, by looking into bar chart the weekends orders are significantly high as compared to weekdays. This could be due to holiday, social time gathering or leisure dining. More customer demand asks restaurants and delivery services to hire more staff with increased inventory in order to provide speedy delivery to the customers. Also, promotional offers during weekdays can sustain the businesses during working days.

HISTOGRAM FOR THE COST OF ORDER VS FOOD PREPARATION TIME



The graph shows most of the orders take average time between 20-35 minutes. There is variation in cost of order vs the food preparation time, few high-cost items are taking less time, where as few high-cost items are taking more time, similar is the case with low-cost food item. There is no strong linear relationship between cost of order and food preparation time. For the businesses standard cooking time is ideal to meet customer satisfaction and ensure delivery consistency.

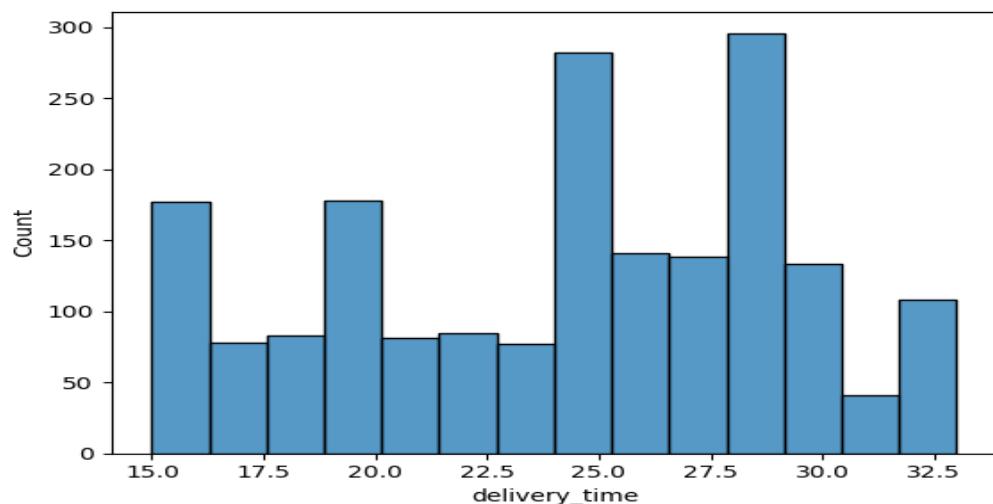
BOX PLOT FOR THE COST OF ORDER VS FOOD PREPARATION TIME



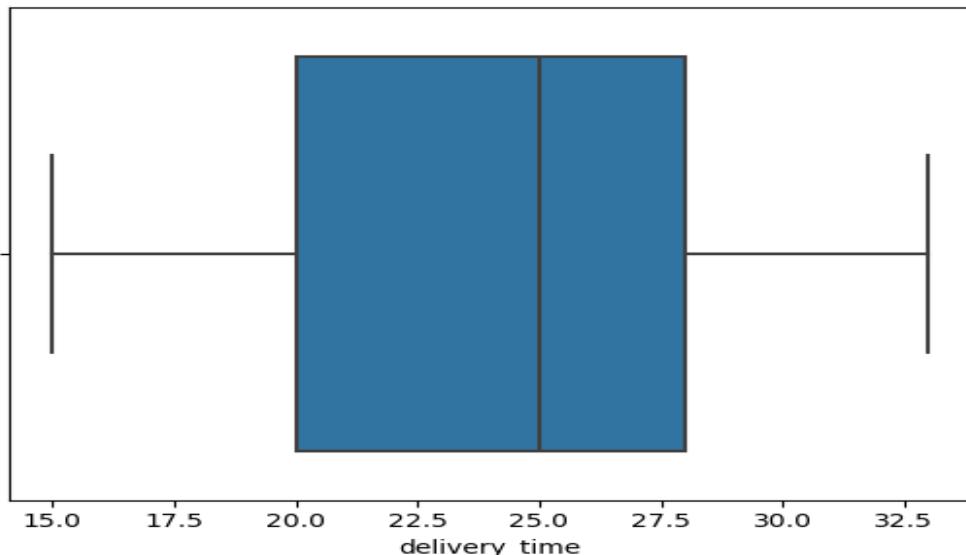
The boxplot shows the time taken for the food is between 23-31 minutes, showing the kitchen working time is adhere to business policy and graph shows there are no outliers. If high cost orders are taking more time to prepare the company should balance the working speed and

revenue, but if low cost items are taking more time there is need to increase operational efficacy to save the time and labour cost.

HISTOGRAM FOR DELIVERY TIME AND NUMBER OF ORDERS



BOXPLOT FOR DELIVERY TIME AND NUMBER OF ORDERS



TOP 5 RESTAURANTS WITH HIGHEST NUMBER OF ORDERS

restaurant_name	
Shake Shack	219
The Meatball Shop	132
Blue Ribbon Sushi	119
Blue Ribbon Fried Chicken	96
Parm	68
Name: count, dtype: int64	

The data shows Shake Shack restaurant is getting highest number of orders. The table shows the meatball shop has the second highest orders, Blue Ribbon Sushi have 119 orders and Blue ribbon fried chicken has 96 orders. The least number of orders is for Parm's.

MOST POPULAR CUISINES ON WEEKENDS

The most popular cuisines on weekends are American cuisine. The number of total orders that cost above 20 dollars is 555. Percentage of orders above 20 dollars is 29.24 %. The mean delivery time for this dataset is 24.16 minutes

NUMBER OF ORDERS PLACED BY EACH CUSTOMER

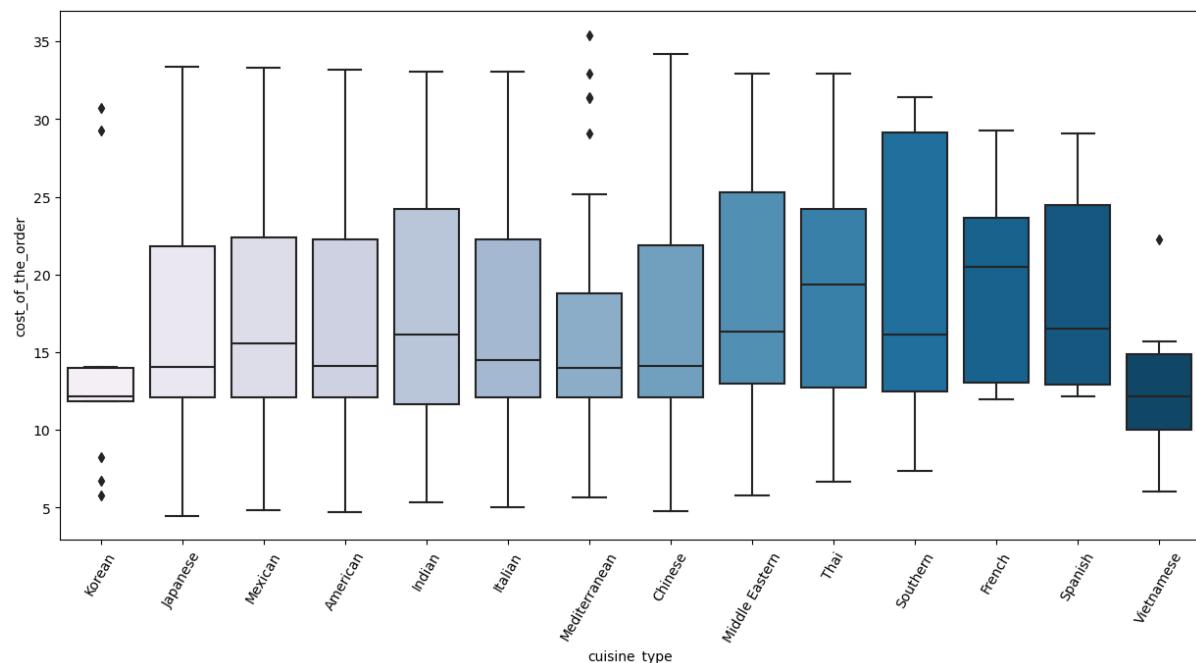
This output shows the number of orders placed by each customer in your dataset.

The result means:

- Customer 52832 placed 13 orders
- Customer 47440 placed 10 orders
- Customer 83287 placed 9 orders

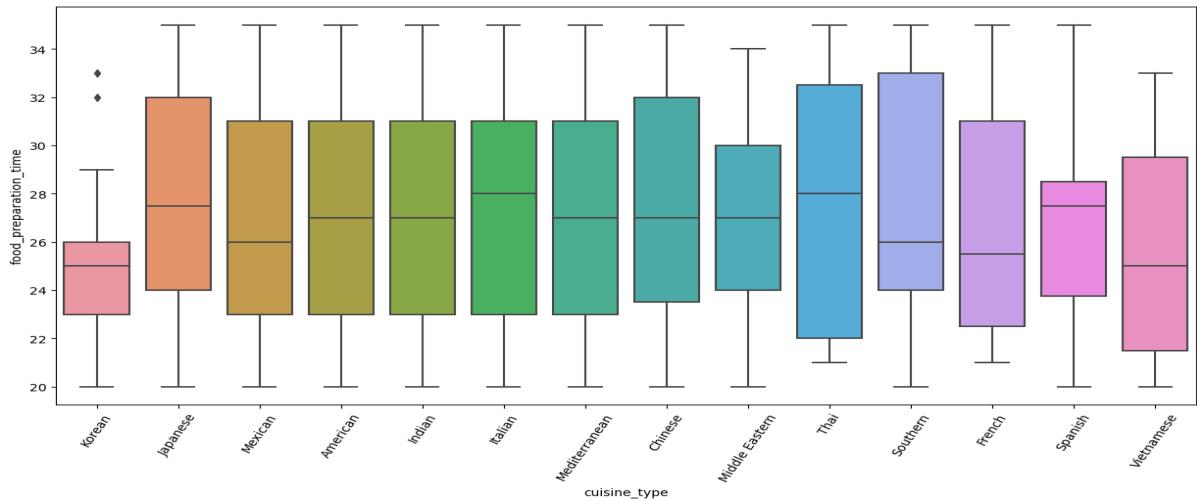
This is useful for identifying repeat customers and analysing customer loyalty.

MULTIVARIATE ANALYSIS



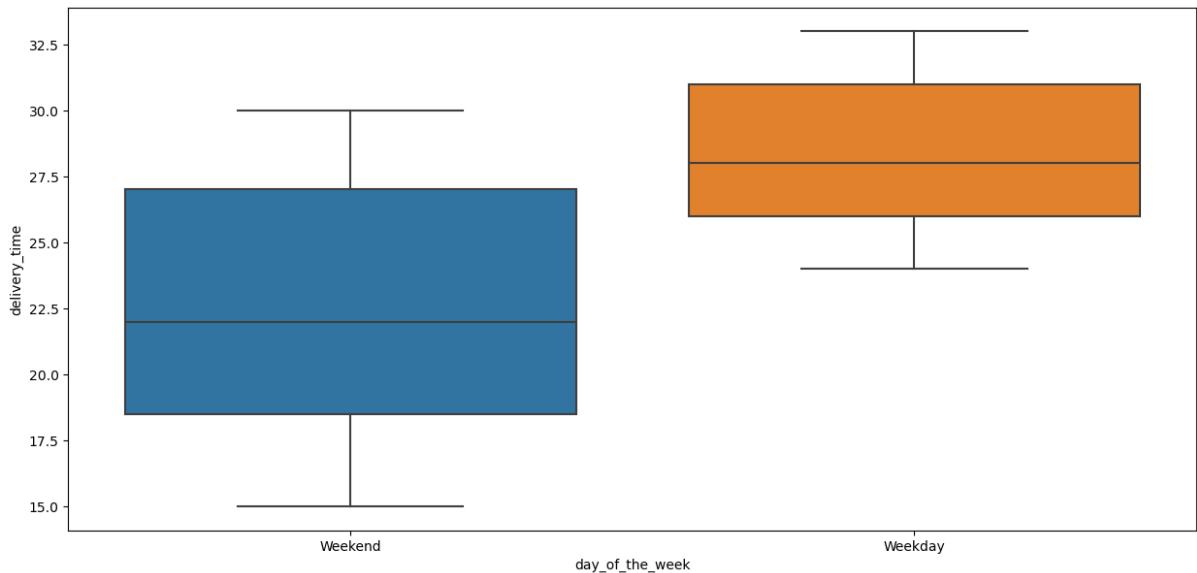
In the boxplot there is notable difference in pricing across cuisines. Southern and French cuisines have the highest median order costs, whereas Korean and Vietnamese cuisines have the lowest median prices. Thai, Southern, and Spanish cuisines have more cost ranges, their menu price is diverse with possible premium offers. In order to achieve maximum revenue potential, the menu pricing strategies with promotional offers can bring profits to businesses.

RELATIONSHIP BETWEEN FOOD PREPARATION TIME AND CUISINE TYPE



The graph shows the cuisines with short and consistent preparation time as Korean are efficient during peak-hour efficiency, while cuisines with longer prep time may require optimized kitchen work flow.

RELATIONSHIP BETWEEN DAY OF THE WEEK AND DELIVERY TIME



The boxplot shows that the weekends have slightly longer and more variable delivery times, while weekdays are more consistent.

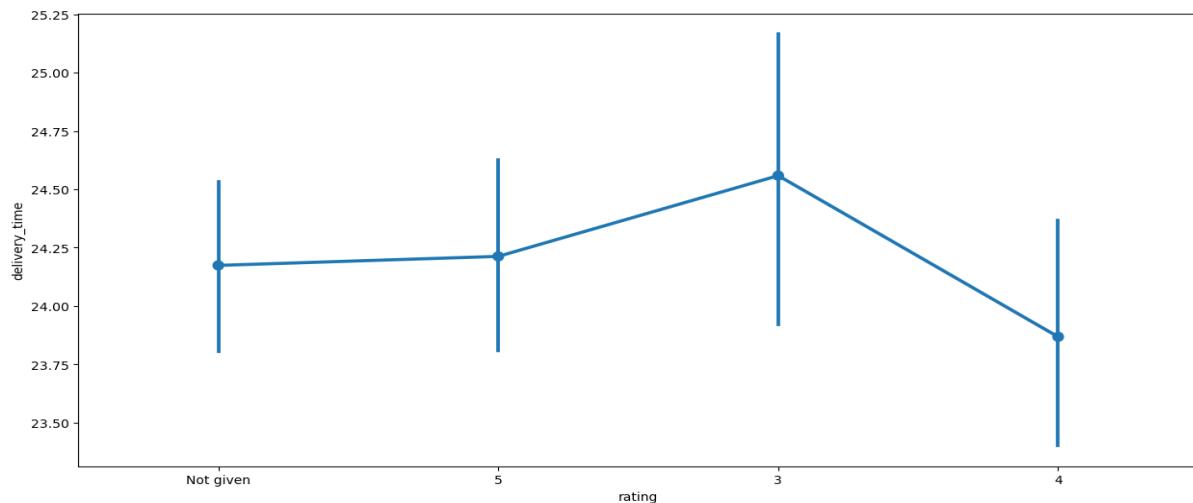
RESTAURANT AND THEIR REVENUE

		Shake Shack	3579.53
The Meatball Shop	2145.21	Blue Ribbon Sushi	1903.95
Blue Ribbon Fried Chicken	1662.29	Parm	1112.76

RedFarm Broadway	965.13	RedFarm Hudson	921.21
TAO	834.50	Han Dynasty	755.29
Blue Ribbon Sushi Bar & Grill	666.62	Rubirosa	660.45
Sushi of Gari 46	640.87	Nobu Next Door	623.67
Five Guys Burgers and Fries	506.47	Name: cost_of_the_order, dtype: float64	

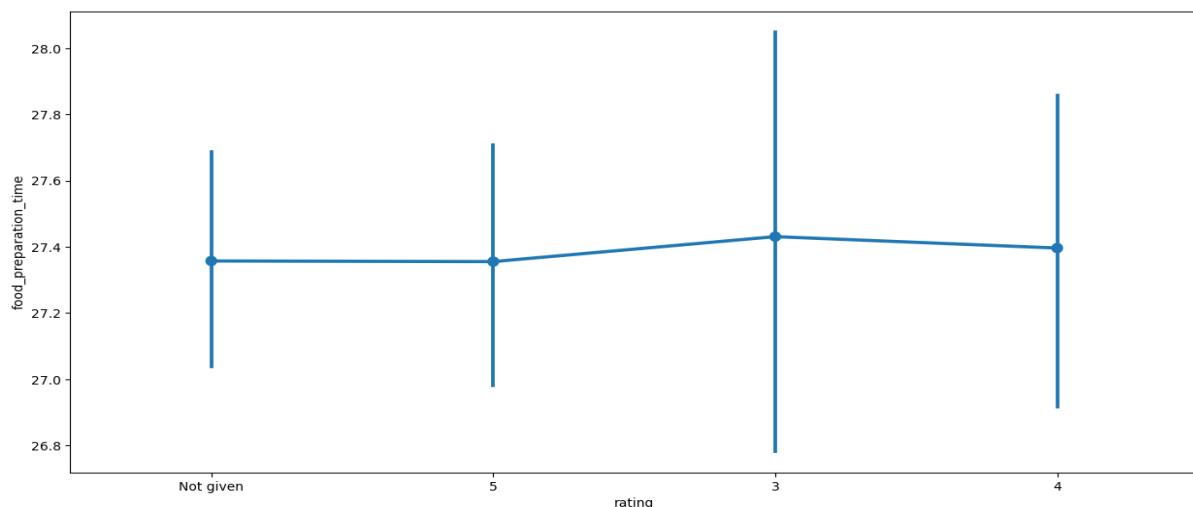
A small group of restaurants generates the majority of total revenue, with Shake Shack leading significantly ahead of competitors. Premium dining spots and well-known chains, such as The Meatball Shop and Blue Ribbon Sushi, show strong revenue performance, while most other restaurants earn moderately in comparison.

RELATIONSHIP BETWEEN RATING AND DELIVERY TIME



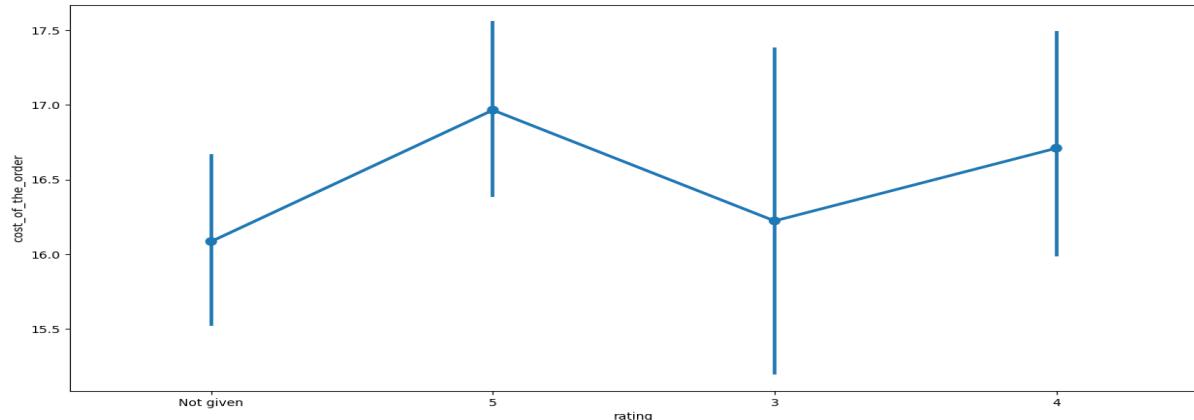
The point plot shows that the average delivery time remains almost consistent across different ratings, having minor variations. By looking at the graph, the delivery time does not seem to be a strong factor to effect customer rating.

RELATIONSHIP BETWEEN RATING AND FOOD PREPARATION



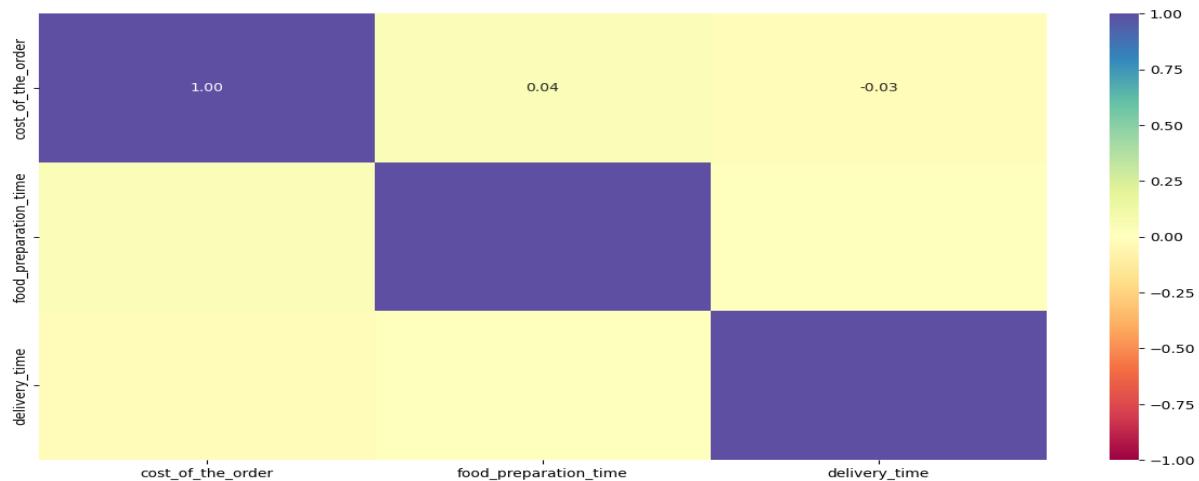
The graph shows the minor variations, the small vertical lines show the confidence interval reflecting the difference is not significant statistically. The food preparation time does not impact customer ratings.

RELATIONSHIP BETWEEN RATING AND COST OF THE ORDER



Again, the statistical difference is not much to impact the cost of order, hence the customer rating is not impacting the cost of order much.

HEATMAP FOR DELIVERY TIME, COST OF ORDER AND FOOD PREPARATION



The heat map shows there is not a strong correlation between three variables, cost, food preparation time and delivery time. Each variable appears to be independent of other in the data set.

RESTAURANTS RATING

restaurant_name	rating
0 Shake Shack	133
1 The Meatball Shop	84
2 Blue Ribbon Sushi	73
3 Blue Ribbon Fried Chicken	64
4 RedFarm Broadway	41

All restaurants in the list shows the rating above 50 except RedFarm Broadway.

RESTAURANT RATING MORE THAN 50

restaurant_name	rating
0 The Meatball Shop	4.511905
1 Blue Ribbon Fried Chicken	4.328125
2 Shake Shack	4.278195
3 Blue Ribbon Sushi	4.219178

REVENUE GENERATION

	order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of
0	1477147	337525	Hangawi	Korean	30.75	Weekend
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend
2	1477070	66393	Cafe Habana	Mexican	12.23	Weekday
3	1477334	106968	Blue Ribbon Fried Chicken	American	29.20	Weekend
4	1478249	76942	Dirty Bird to Go	American	11.59	Weekday

The net revenue is around 6166.3 dollars. Percentage of orders taking more than 60 minutes: The mean delivery time on weekdays is around 28.34 minute. The mean delivery time on weekend is around 22.47 minute

TOP CUSISINE AND TOTAL REVENUE

American	1880.4130	Japanese	1481.0765
Italian	966.8845	Chinese	690.7530
Mexican	257.4205	Indian	246.9940
Middle Eastern	193.5535	Mediterranean	136.3280
Thai	78.0525	French	76.1925
Southern	68.7535	Spanish	47.2365
Korean	33.3030	Vietnamese	15.7530
Name: Revenue, dtype: float64			

CONCLUSION

The study shows from the data that renowned Shake Shack restaurant is the highest revenue-producing restaurant, earning **\$6,166.30**, with the **\$703.61** as the maximum standard order value. Among cuisines American, Japanese, and Italian arose as the prominent classes in terms of both order volume and total revenue gained from the business. The FoodHub delivery efficiency remains robust, with only **10.54%** of orders exceeding the 60-minute threshold level. Nevertheless, weekday supplies average **28.34 minutes**, which is slower than weekend deliveries (**22.47 minutes**) when comparatively studied.

Effective analysis using a box plot of food preparation time shown that the mainstream of orders is arranged within **24 to 31 minutes**, with a median of **27 minutes**. The preparation process shows consistency with no intense outliers, suggesting well-managed kitchen systems. Nonetheless, the spread up to **35 minutes** for certain orders indicates potential for further efficiency gains, particularly for items that may interrupt order completion and subsequently increase delivery times. Among all typically four restaurants recorded over 50 customer ratings, each retaining an average rating above 4, demonstrating strong customer satisfaction within this sector.

RECOMMENDATIONS

The study recommends the deliberate advertising and promotions announcement to boost the revenue for the highly rated restaurants and increase their market share. Efficient working on weekdays deliveries should be focused either through better route planning for the quick deliveries, if needed increase the kitchen and delivery staff, also strong communication among them to improve performance, thus eventually improving the operational efficiency. Consumer demand aligning with high-demand cuisines and restaurants should be expanded through different offers of menu. The restaurant not receiving online rating should adopt such like marketing strategies in which consumer willingly give rating, it could be like giving some points on rating. Late delivery risks should be reduced along with reduced food preparation time. To promote more business 60 min service time management should be brought down by adopting the monitoring tools to achieve customer satisfaction.

