Python Foundations: FoodHub Project

Data Science & Business Analytics

August 9, 2024

Contents/Agenda

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

Executive Summary (1)

Conclusions

- Shake Shack is the most popular restaurant based on the data.
- Japanese, American, Italian, and Chinese cuisine are the most popular.
- There are no missing values in the data set.
- 736 orders have not been rated by customers .
- Order volume is higher during the weekend, potentially due to customers opting for convenience during their time off of work or school.
- Delivery time is longer on weekdays, most likely due to more traffic.

Executive Summary (2)

Recommendations

- FoodHub should focus on gathering more ratings to get a better idea of customer satisfaction. When customers give high ratings to a restaurant, orders may increase, which would create more revenue for Foodhub.
- To increase the number of customers who rate their orders, FoodHub should provide incentives to customers, such as free delivery or a discount at checkout.
- FoodHub should add more restaurants that serve the most popular cuisine types.
- FoodHub should provide incentives for repeat customers so these customers continue to place orders consistently.

Business Problem Overview & Solution Approach (1)

Business Problem Overview:

- The food aggregator company, FoodHub, wants to analyze the data they have collected from different orders to determine the demand of different restaurants.
- Foodhub wants to enhance their customer experience by analyzing the data they have collected which includes variables such as, rating, cuisine type, cost, etc.

Business Problem Overview & Solution Approach (2)

The Solution Approach

- The rating provided by customers can be used to determine customer satisfaction.
- When customers are satisfied with their restaurant orders, they are more likely to keep ordering from said restaurant. Moreover, better rated restaurants will attractive more customers.
- Therefore, FoodHub should focus on the rating given by customers. This will increase orders and as result increase revenue for FoodHub.
- To increase the amount of ratings, FoodHub should provide incentives to customers when they rate their order. The promotions could include free delivery or a discount at checkout.

Data Overview (1)

The dataset contains data relating to customers food orders.

The dataset includes:

- **order id**: unique ID of the order
- **customer id**: ID of the customer
- **restaurant name**: name of the restaurant
- **cuisine type:** type of food ordered by customer
- **cost**: cost of the order
- **day of the week**: weekday or weekend
- **Rating:** rating given by customer out of 5
- **food preparation time**: time taken by restaurant to prepare the food
- **delivery time**: time taken for the order to be delivered

Data Overview (2)

1. There are 1898 rows and 9 columns present in the data.
2. The data types present in the dataset include float (cost of the order), integer (order id, customer id, food preparation time, & delivery time), and object (restaurant name, cuisine type, day of the week, and rating).
3. There are no missing values in the data
4. The statistical summary for the food preparation time:
    - It takes a minimum of 20 minutes to prepare the food once an order is placed
    - It takes on average 27.37 minutes to prepare the food once an order is placed
    - It takes a maximum of 35 minutes to prepare the food once an order is placed
5. There are 736 unrated orders

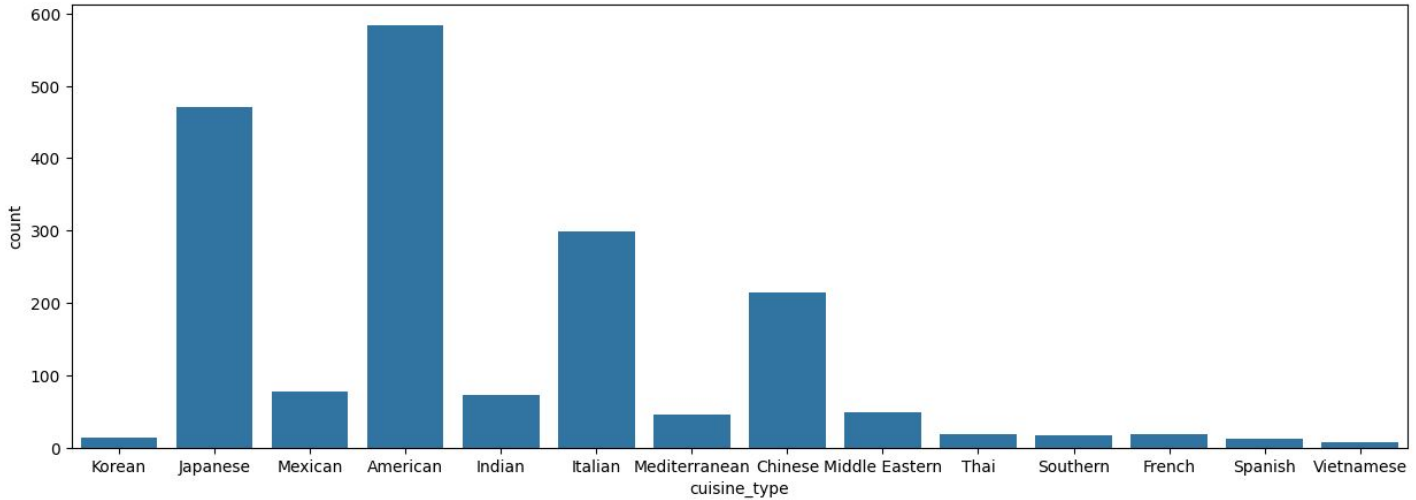
Univariate Analysis

The trends present in each variable were explored using histograms, box plots, count plots, bar plots, and bar graphs.

6.

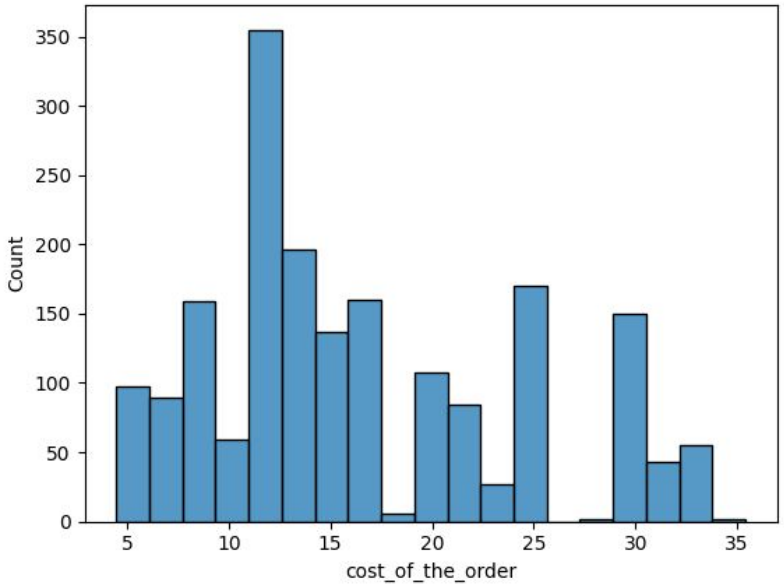
- There are 1898 unique orders
- There are 1200 unique customers
- There are 178 unique restaurants listed in the dataset
- There are 14 unique cuisines in the dataset

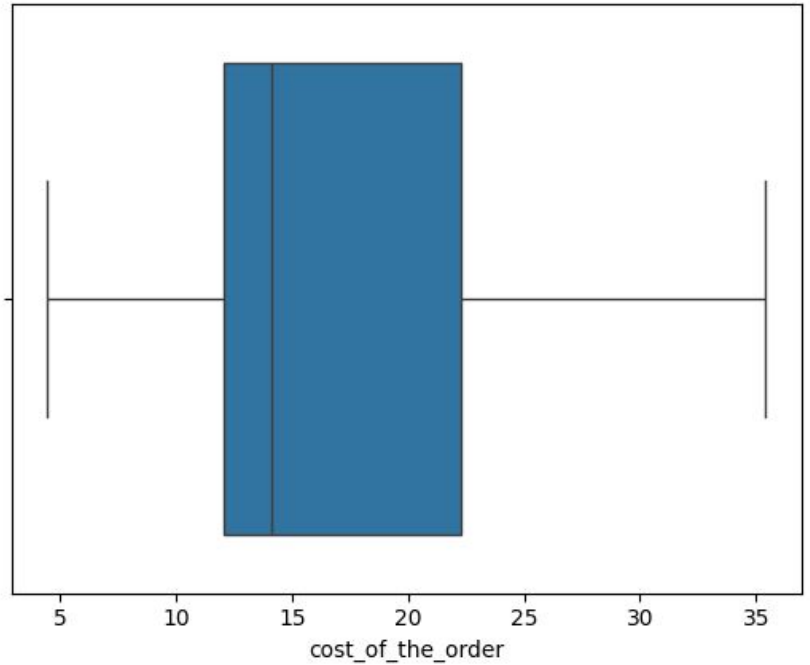
Univariate Analysis - Cuisine Type



- American cuisine was ordered 590 times, making it the most popular cuisine type,
- Japanese, Italian, and Chinese food are among the most popular cuisine type.
- Vietnamese, Spanish, and Korean cuisine seem to be the least popular.

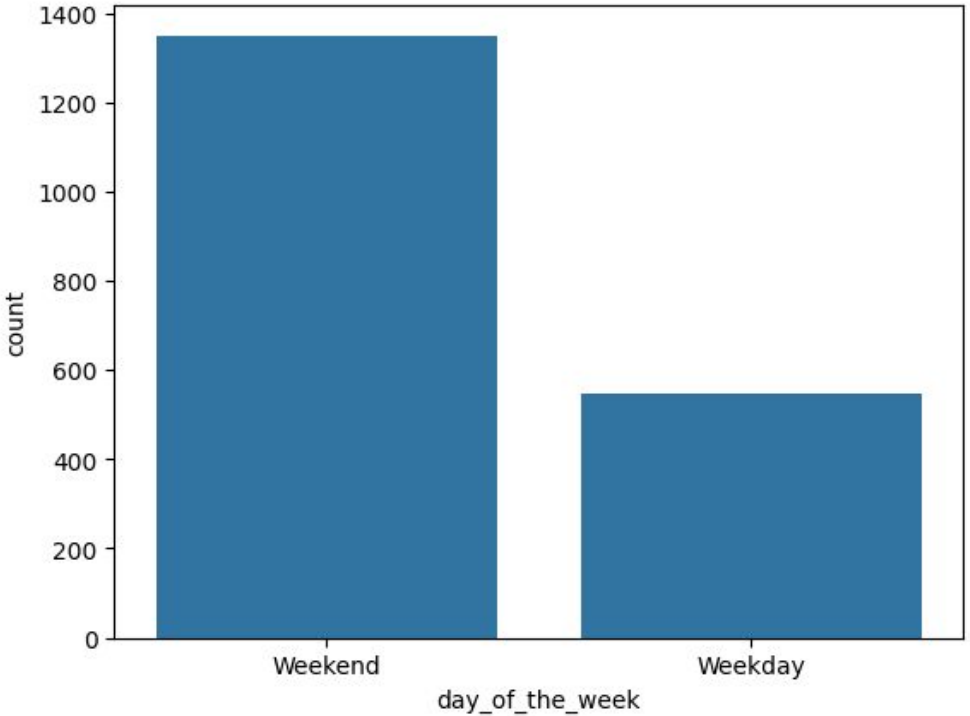
Univariate Analysis - Cost of the Order





- The distribution is slightly right-skewed.
- The mode of the distribution shows that most people order meals that cost around $10 to $13
- The median is about $14
- There are no outliers in this distribution

Univariate Analysis - Day of the Week

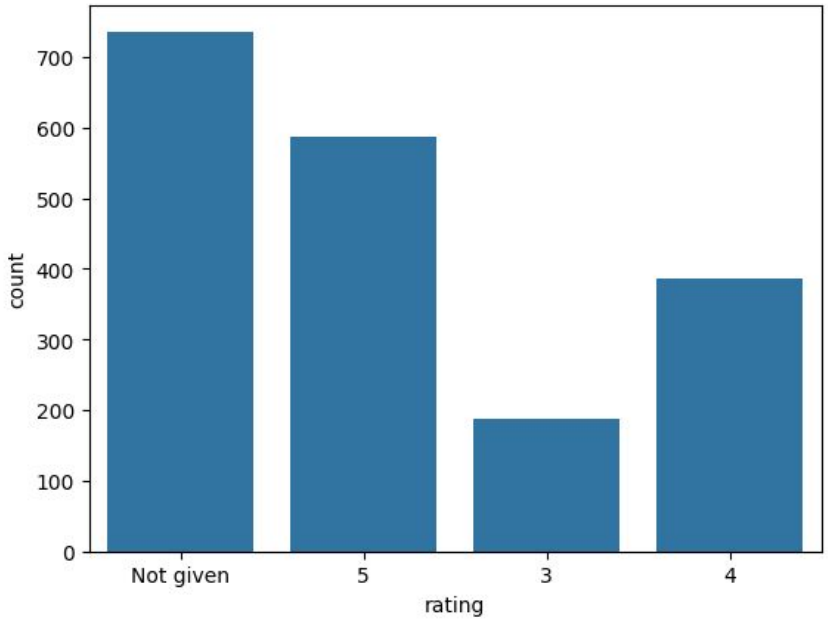


day_of_the_week	count
Weekend	1351
Weekday	547

dtype: int64

- Food orders are higher during the weekend.

Univariate Analysis - Rating

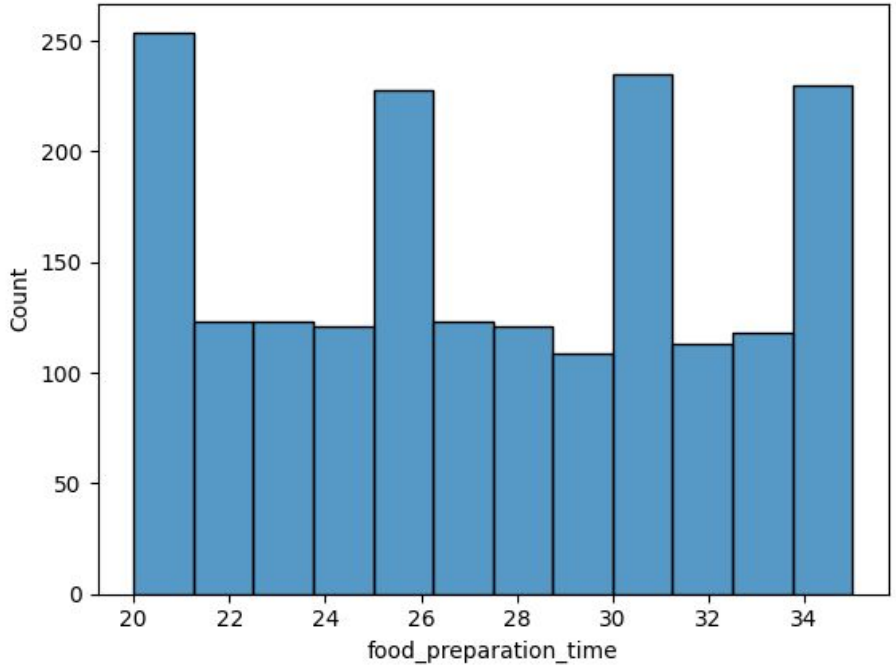


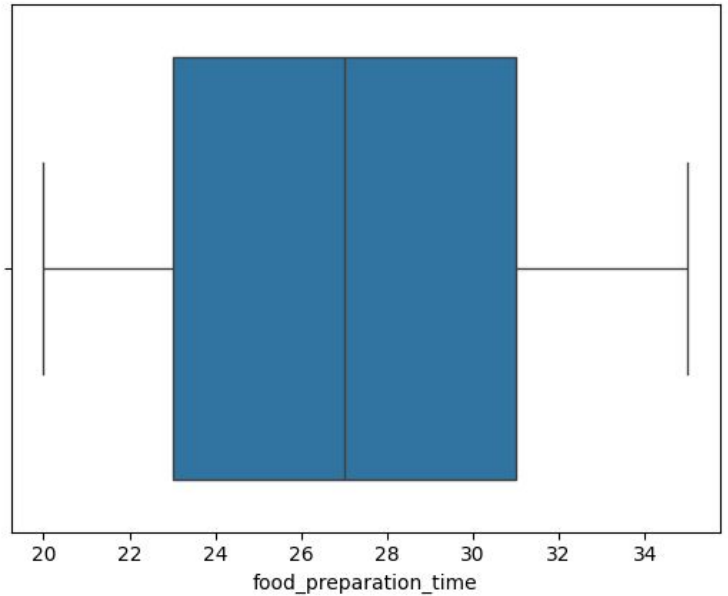
rating	count
Not given	736
5	588
4	386
3	188

dtype: int64

- Majority of FoodHub's customers do not rate their orders
- When customers do rate their orders, a rating of 5/5 is given.

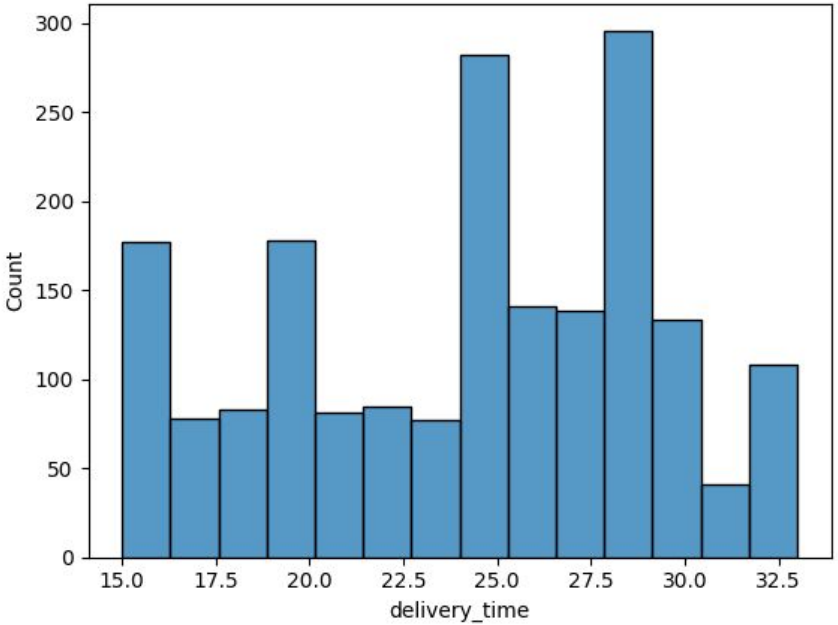
Univariate Analysis- Food Preparation Time

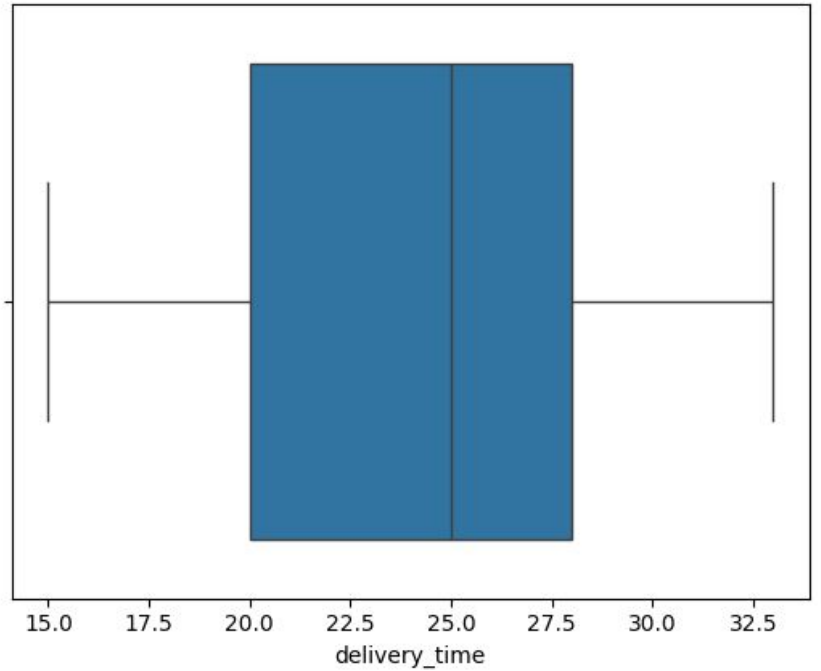




- There are no outliers in this distribution
- The distribution is uniform with a median of 27.

Univariate Analysis - Delivery Time





- There are no outliers in this distribution
- The distribution is left skewed, meaning the average delivery time is smaller than the median

Univariate Analysis

7. Shake Shack, The Meatball Shop, Blue Ribbon Sushi, Blue Ribbon Fried Chicken, and Parm were the restaurants that received the highest number of c

restaurant_name	count
Shake Shack	219
The Meatball Shop	132
Blue Ribbon Sushi	119
Blue Ribbon Fried Chicken	96
Parm	68

8. American food is the most popular cuisine on the weekend.

9. A total of 555 orders cost above $20. The percentage of the orders that cost above $20 is 29.24%

Univariate Analysis

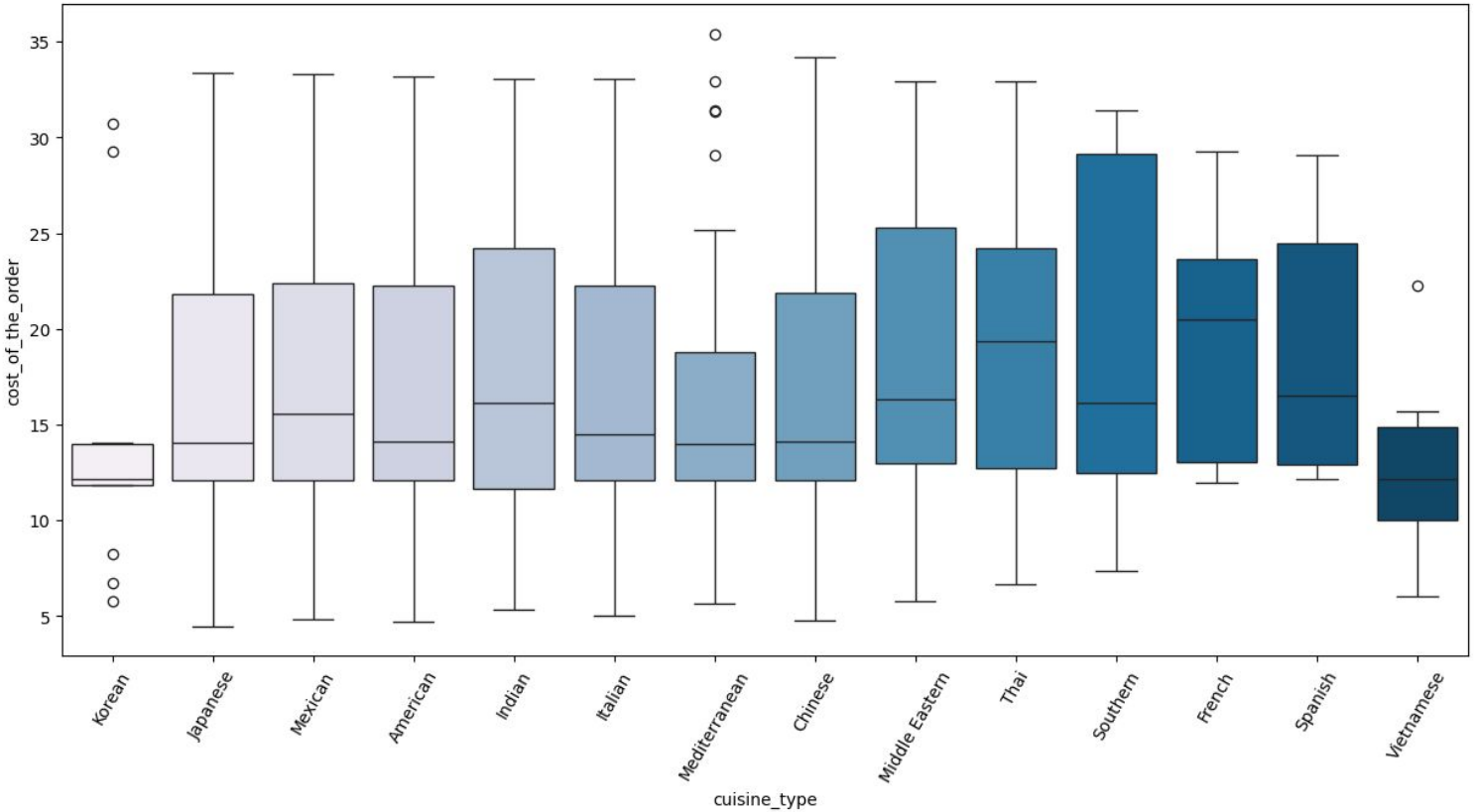
10. The mean delivery time is 24.16 minutes

11. The top 5 frequent customers are customers with ID 52832, 47440, 83287, 250494, & 259341.

customer_id	
52832	13
47440	10
83287	9
250494	8
259341	7

dtype: int64

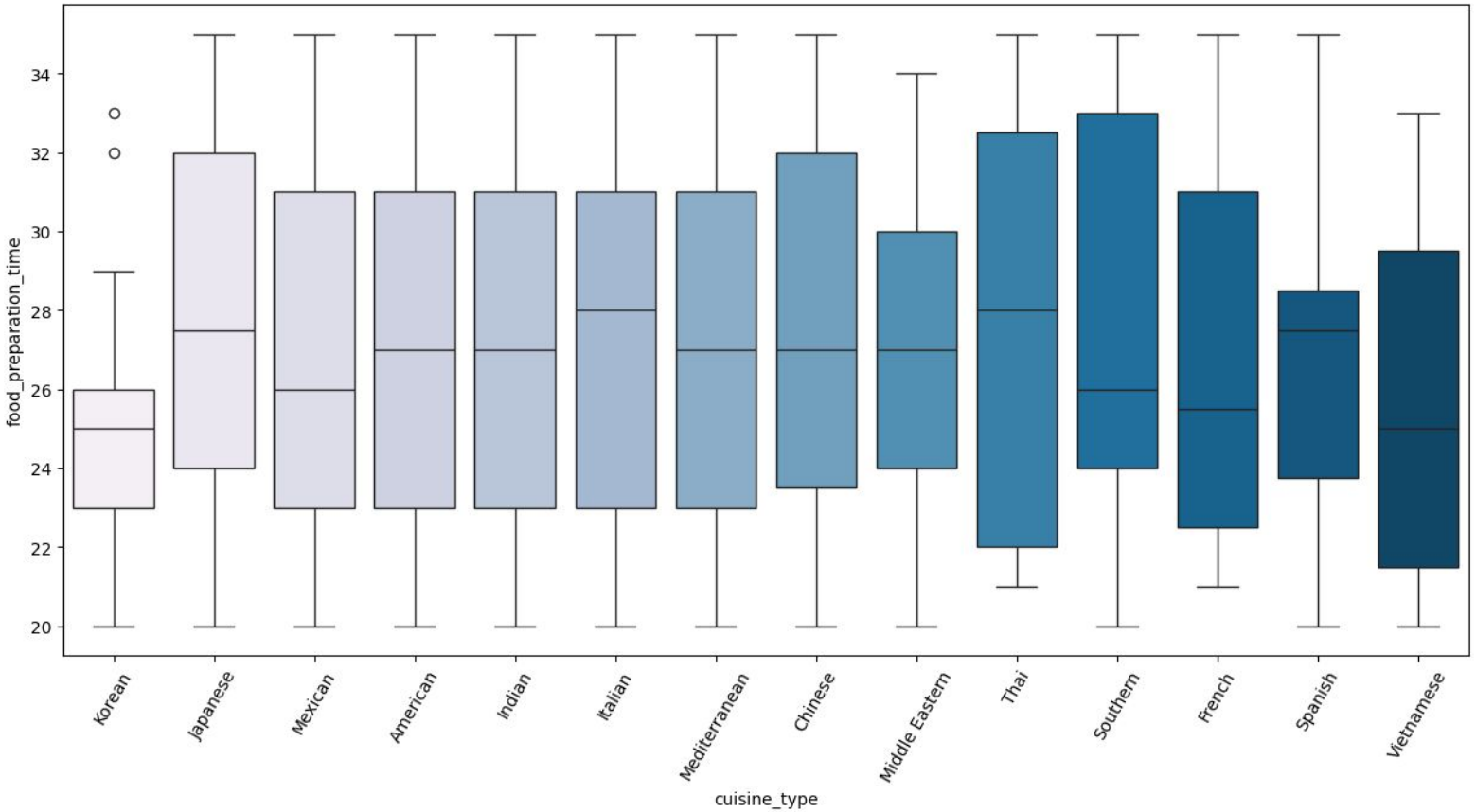
Multivariate Analysis - Cuisine Vs Cost of the order (1)



Multivariate Analysis - Cuisine Vs Cost of the order (2)

- Vietnamese and Korean cuisines cost less compared to other cuisines.
- There are outliers for the cost of Korean, Mediterranean and Vietnamese cuisines.
- French and Spanish cuisines are more expensive compared to other cuisines.

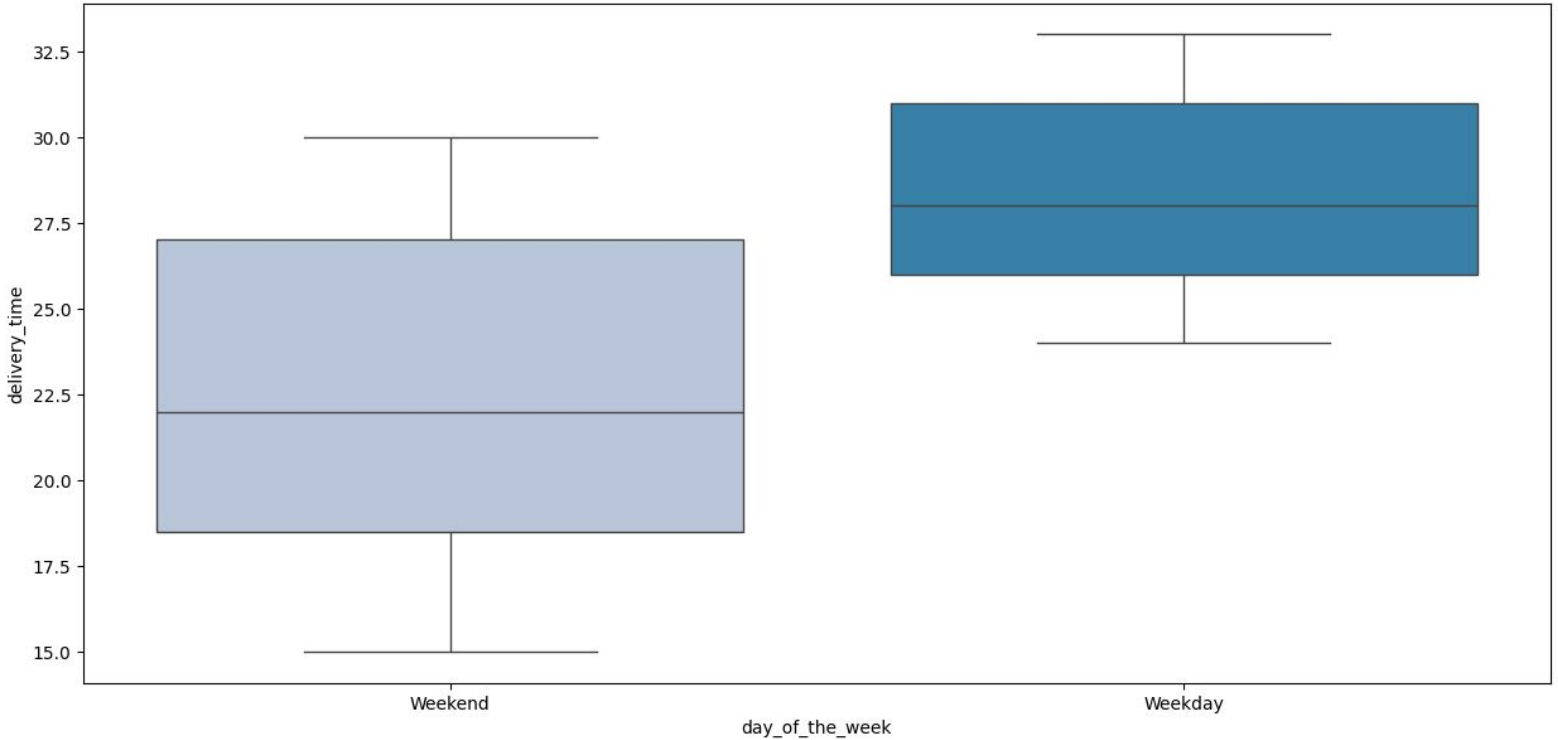
Multivariate Analysis – Cuisine Vs Food Preparation Time



Multivariate Analysis - Cuisine Vs Food Preparation Time (2)

- The median food preparation time is between 24 and 30 minutes for all the cuisines.
- There are outliers present for the food preparation time of Korean cuisine.
- Korean cuisine takes less time compared to the other cuisines.
- Southern cuisine takes more time compared to the other cuisines.

Multivariate Analysis - Day of the week vs Delivery time (1)



Multivariate Analysis - Day of the week vs Delivery time (2)

- The delivery time for all the orders over the weekends is less compared to weekdays

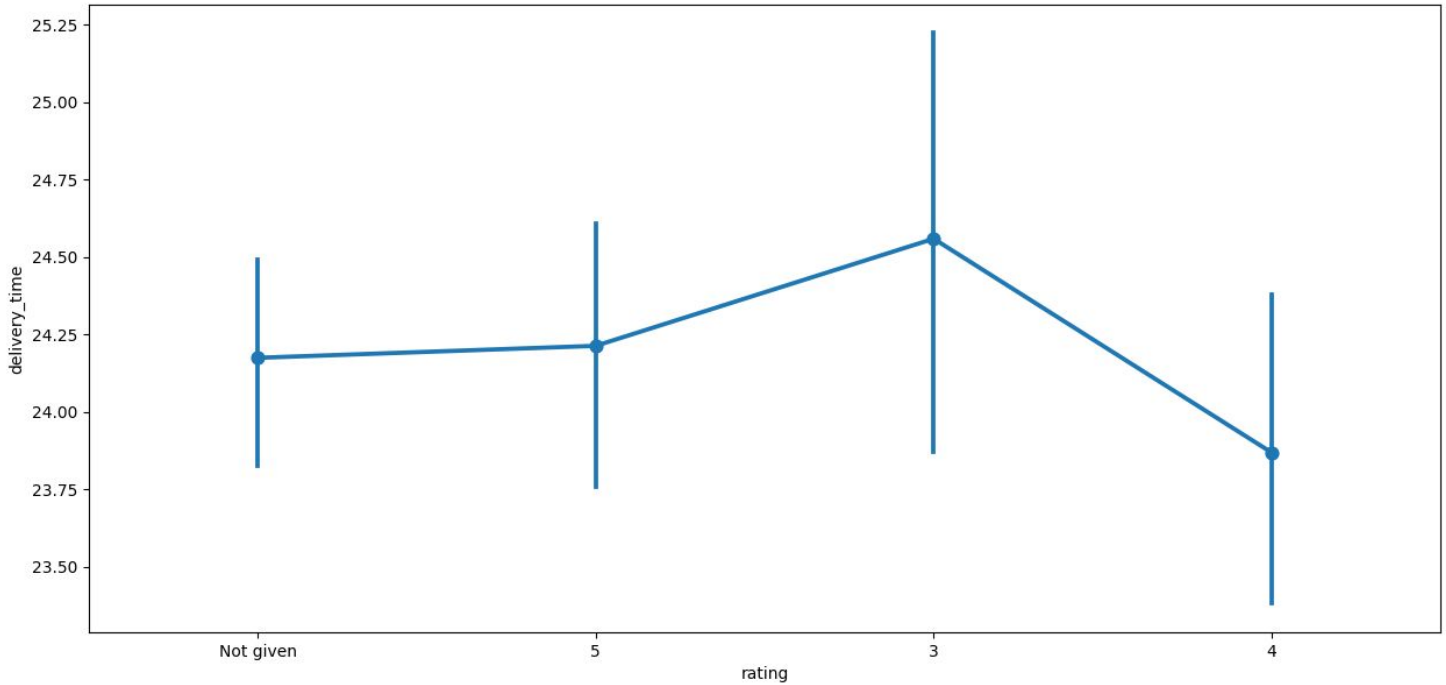
Revenue Generated by the Restaurants

restaurant_name	cost_of_the_order
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

dtype: float64

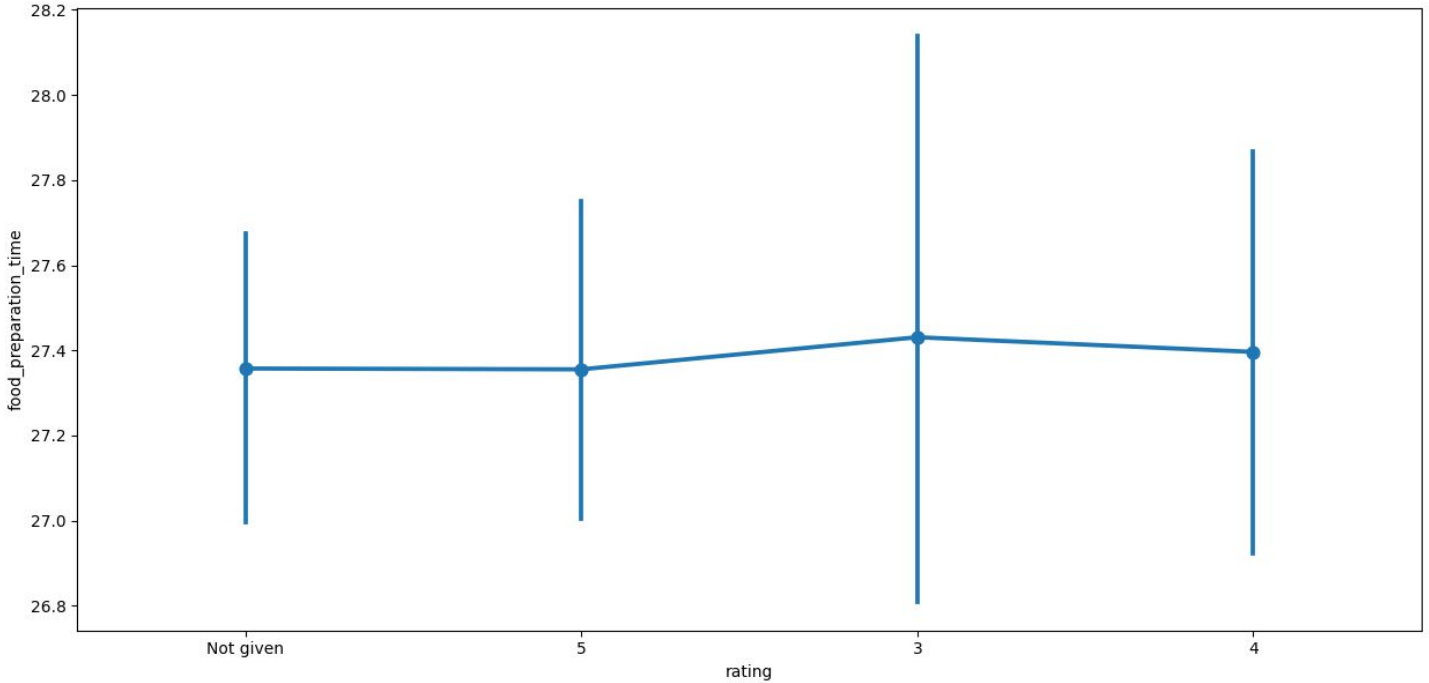
The 14 restaurants listed here are generating more than 500 dollars in revenue.

Multivariate Analysis - Rating Vs Delivery Time



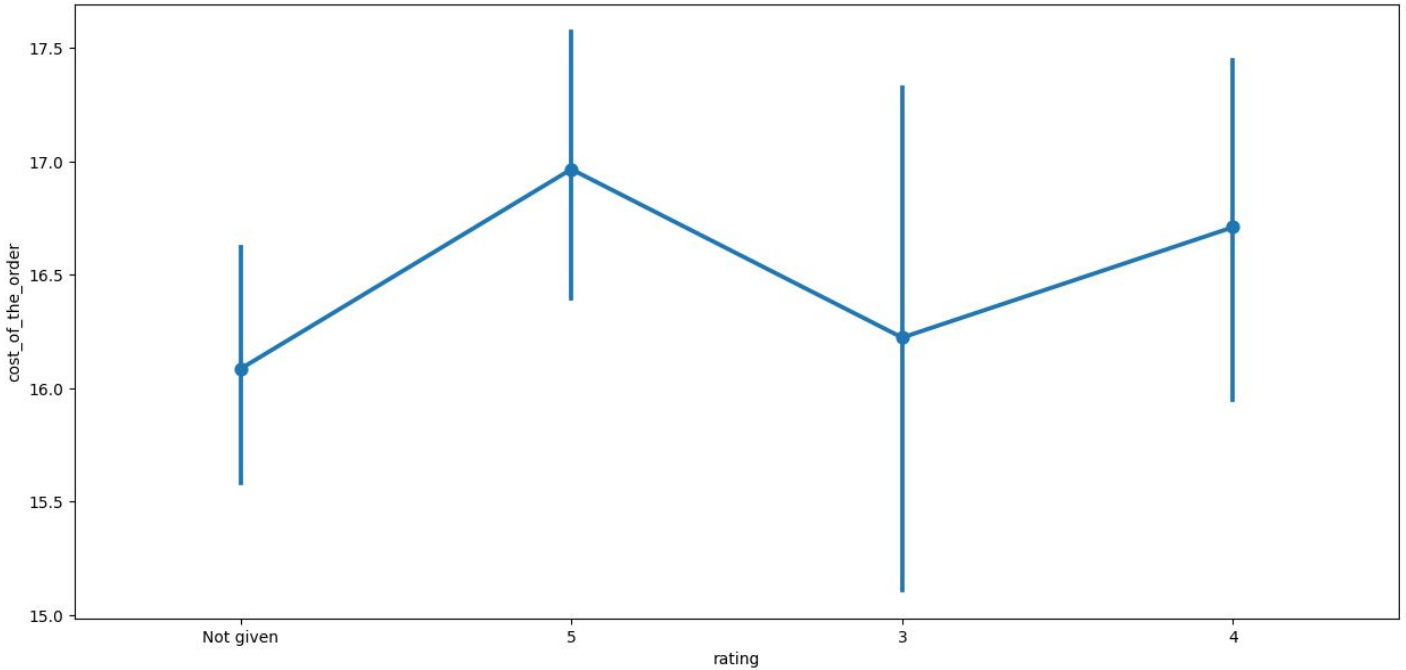
- Based on the figure, longer delivery time plays a role in lower rated order.

Multivariate Analysis - Rating Vs Food Preparation Time



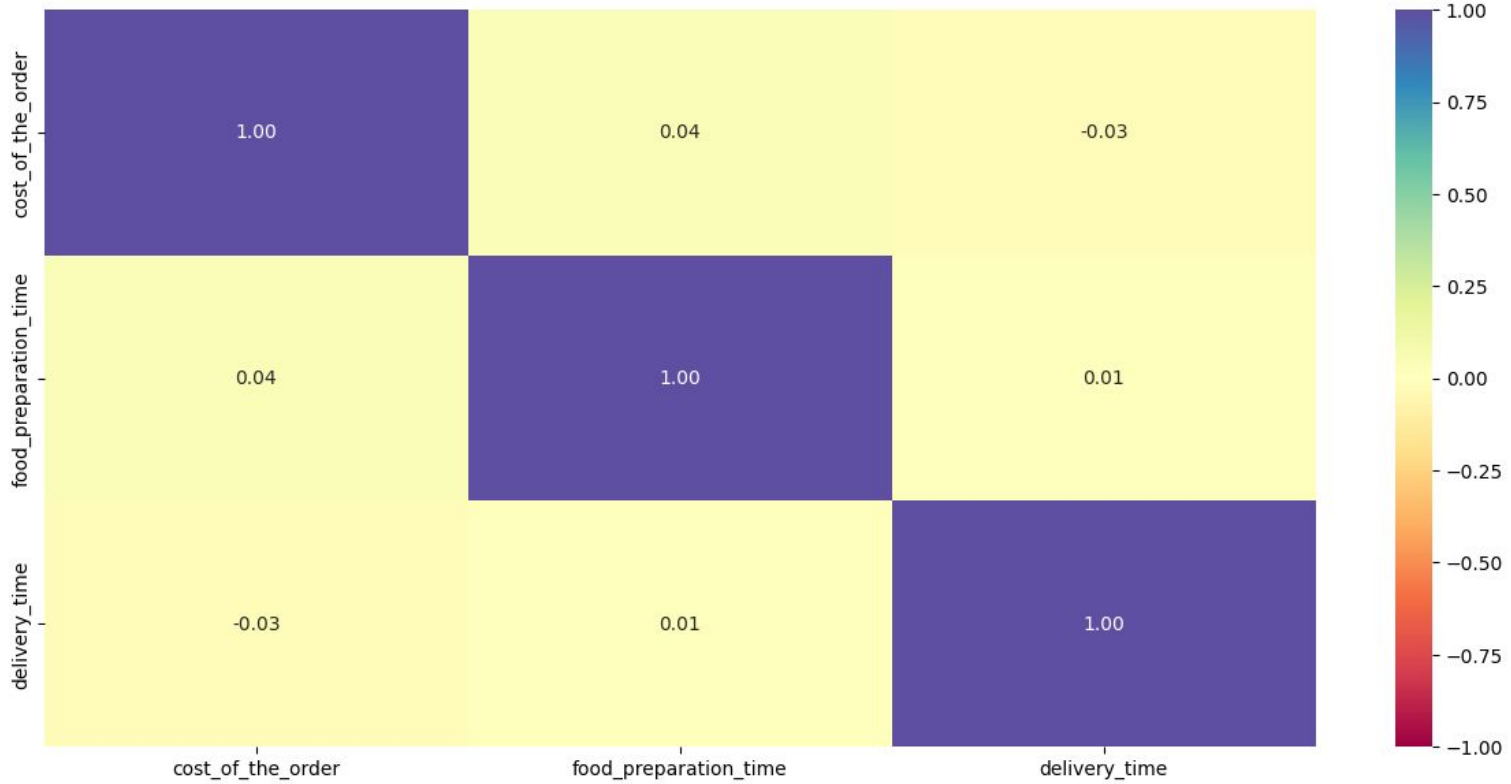
There does not seem to be a connection between food preparation time and rating.

Multivariate Analysis - Rating Vs Cost of the Order



Orders with a higher cost have received better ratings, while lower cost orders have not been rated.

Multivariate Analysis - Correlation among variables



There seems to be very little correlation between delivery time, food preparation time, and cost of the order.

Multivariate Analysis

13. The below restaurants fulfill the criteria for the promotion

	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

Multivariate Analysis

14. The net revenue generated by the company across all orders is around 6166.3 dollars.

15. The percentage of orders that have more than 60 minutes of total delivery time is 10.54 %.

16. The mean delivery time on weekdays is around 28 minutes. The mean delivery time on the weekend is around 22 minutes. Therefore, delivery is faster on the weekends.