FoodHub Orders - Analysis

ML & AI

Azin Faghihi
November 2024

Contents / Agenda

- Executive Summary

- Business Problem Overview and Solution Approach

- Data Overview

- EDA - Univariate Analysis

- EDA - Multivariate Analysis

- Appendix

Executive Summary

Cuisines:

- The median food preparation time is highest for **Italian** cuisine and lowest for **Korean** cuisine.
- **Thai** cuisine has the widest spread (IQR) in food preparation time, while Korean cuisine has the narrowest.
- The median cost of the order is highest for **French** cuisine and lowest for **Vietnamese** cuisine.
- **Southern** cuisine has the widest spread (IQR) in cost of the order, while **Korean** cuisine has the narrowest.
- The most popular cuisine on weekends and weekdays is the **American** cuisine.

Delivery Time:

- Food preparation takes a minimum of 20 minutes, an average of 27 minutes, and a maximum of 35 minutes.
- Average delivery times are longer on weekdays compared to weekends.

Preparation Time:

- Food preparation takes a minimum of 20 minutes, an average of 27 minutes, and a maximum of 35 minutes.

Recommendations

Customers:

- Encourage customers to leave reviews to increase their chances of receiving more discount vouchers.
  - This is particularly important for Japanese and Italian cuisines, which currently have the highest proportion of unrated orders.

Restaurants:

- Reduce the total time and consider weekend/weekday deliveries more carefully
  - If possible, reduce the preparation time for the Italian and Thai restaurants.
  - If possible, reduce the delivery time on weekdays.
  - Hire more help during the weekends for the higher number of orders

Business Problem Overview and Solution Approach

Problem:

    FoodHub is a food aggregator company and would like to learn about the demand of the restaurants and also enhance their customer experience using data science.

Approach:

    We conduct exploratory data analysis on the available data of previous orders.

Data Overview

- There are 1,898 rows and 9 columns in the dataset.
- There are no null values however some ratings are tabulated as 'Not given'.
- Here's the data dictionary:

Column	dtype	Description
order_id	int64	Unique ID of the order
customer_id	int64	ID of the customer who ordered the food
restaurant_name	object	Name of the restaurant
cuisine_type	object	Cuisine ordered by the customer
cost_of_the_order	float64	Cost of the order
day_of_the_week	object	Indicates whether the order is placed on a weekday or weekend
rating	object	Rating given by the customer out of 5
food_preparation_time	int64	Time (in minutes) taken by the restaurant to prepare the food.
delivery_time	int64	Time (in minutes) taken by the delivery person to deliver the food package.

Data Overview

- Food preparation takes a minimum of 20 minutes, an average of 27 minutes, and a maximum of 35 minutes.
- About **40%** of the orders are **not rated**.

Column	cost_of_the_order	food_preparation_time	delivery_time
count	1898.000000	1898.000000	1898.000000
unique	NaN	NaN	NaN
top	NaN	NaN	NaN
freq	NaN	NaN	NaN
mean	16.498851	27.371970	24.161749
std	7.483812	4.632481	4.972637
min	4.470000	20.000000	15.000000
25%	12.080000	23.000000	20.000000
50%	14.140000	27.000000	25.000000
75%	22.297500	31.000000	28.000000
max	35.410000	35.000000	33.000000

Univariate Analysis

There are **~2K orders** in the dataset for **1,200 customers** for **178 restaurants** in **14 cuisine types.**

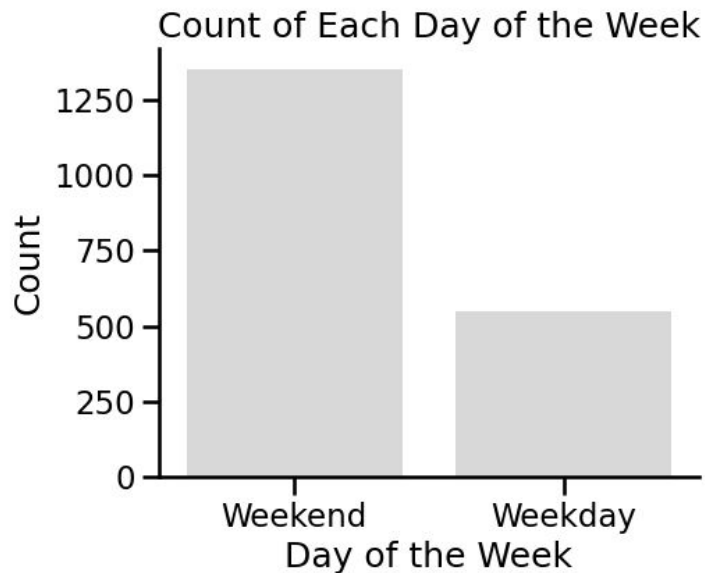Following slides show the univariate analysis of the data. The variables are:

Day of the week
Rating
Food preparation time
Delivery time
Restaurant
Cuisine types
Cost of the order
Customers

Univariate Analysis
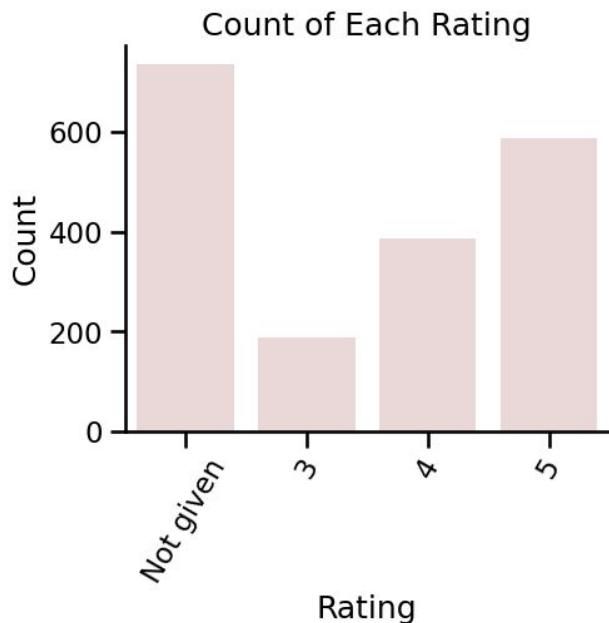
Day of the week (weekends/weekdays):

- Around **30%** of orders are placed on **weekdays** and **70%** on **weekends.**

Univariate Analysis

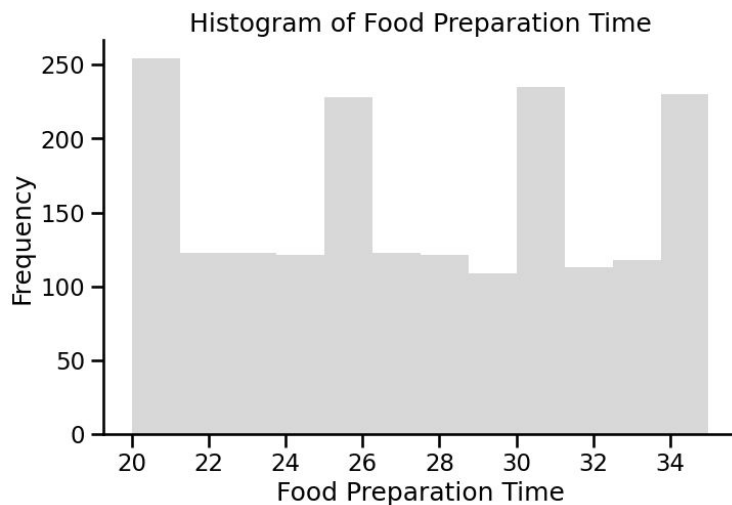## Ratings:

- About **40%** of the orders are **not rated**.

Count of Each Rating

rating	count
3	188
4	386
5	588
Not given	736

Univariate Analysis

Food Preparation Time:

- Food preparation takes a minimum of 20 minutes, an average of 27 minutes, and a maximum of 35 minutes.
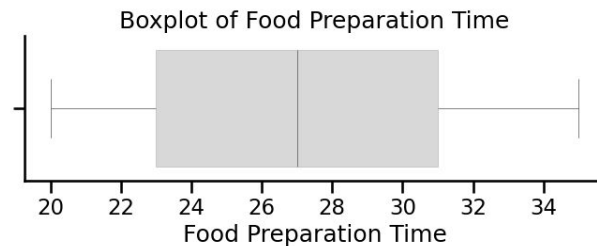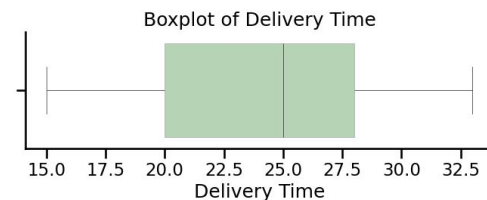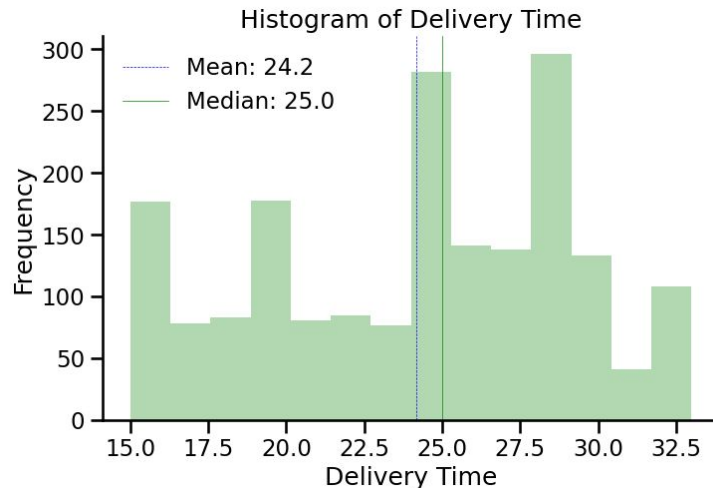
Univariate Analysis

Delivery Time:

- Delivery takes a minimum of 15 minutes, **_an average of 24 minutes_**, and a maximum of 33 minutes.

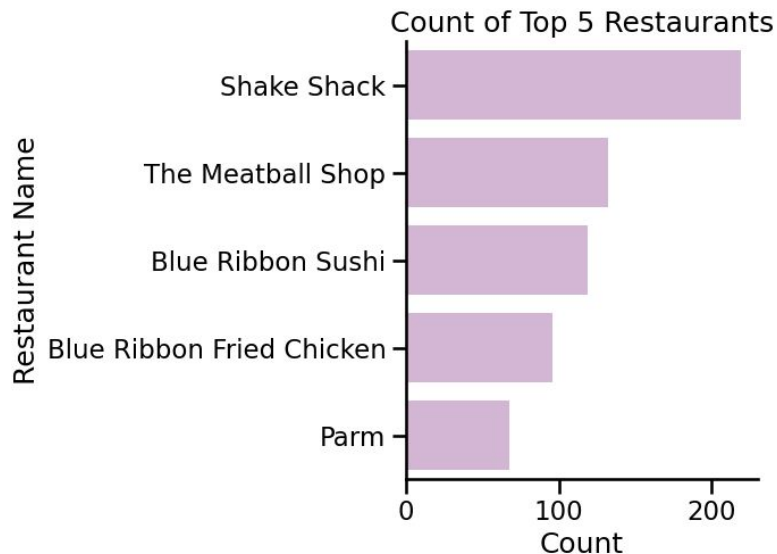- The distribution of the delivery time is slightly left-skewed.

Delivery Time (min)	
min	15
mean	24
max	33

Histogram of Delivery Time

Mean: 24.2
Median: 25.0

Boxplot of Delivery Time

Univariate Analysis

Restaurants:

- Here are the top 5 restaurants with the highest *number of orders* placed.

Count of Top 5 Restaurants

Univariate Analysis

Cuisine Types:

- The *most popular* cuisine on weekends and weekdays is the **American** cuisine.

Count of Each Cuisine Type
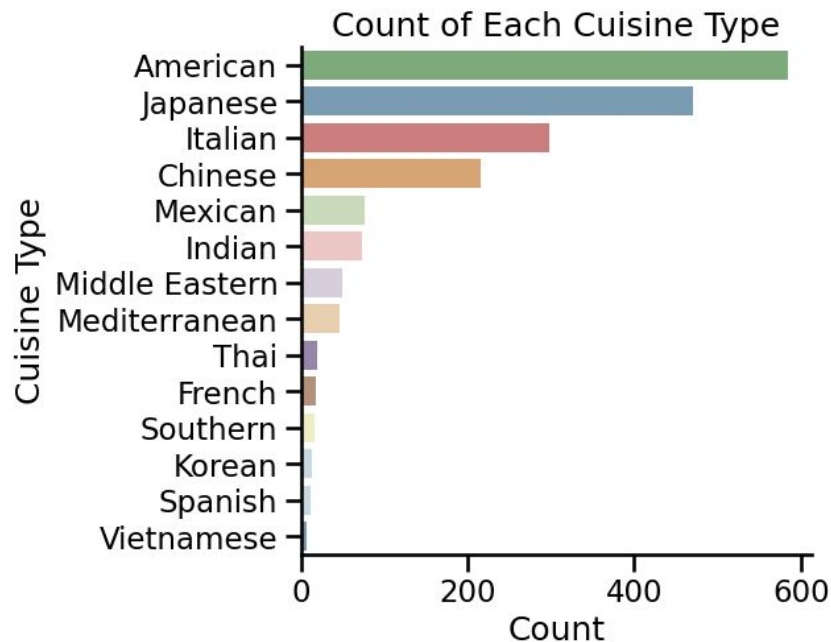
Univariate Analysis

Cost of the order:

- The distribution of order costs is right-skewed.
- The average cost of an order is around $16.5.
    - Around 29% of the orders *cost more that $20*.
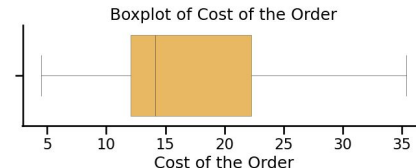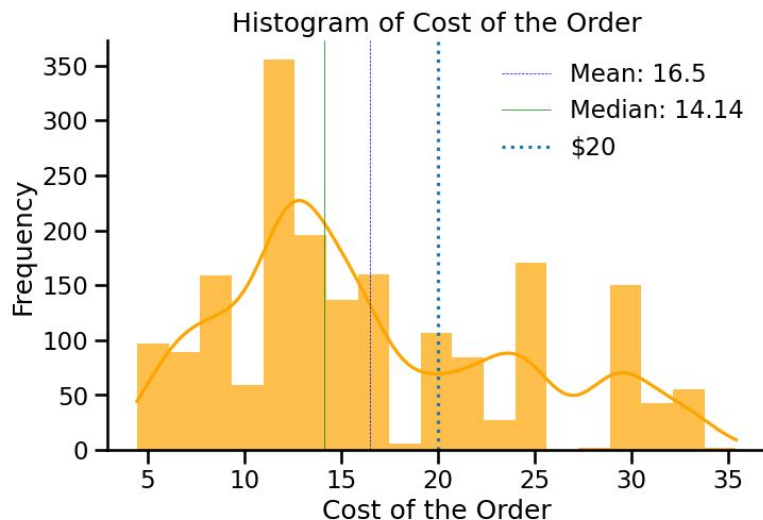
Univariate Analysis

Customers:

- Here are the 5 *most frequent customers* and are eligible to get 20% discount vouchers.

customer_id	count
52832	13
47440	10
83287	9
250494	8
259341	7

Multivariate Analysis

Following slides show the multivariate analysis of the data. The variables are:

Cuisine type *vs* cost of the order

Cuisine type *vs* food preparation time

Day of the week *vs delivery time*

Restaurants and cost of the order

Day of the week and cost of the order

Rating *vs* delivery time

Rating *vs* preparation time

Rating *vs* cost of the order

Orders, restaurants, ratings

Cost, rating, delivery time, preparation time

Multivariate Analysis

Cuisine type *vs* cost of the order:

- The median cost of the order is highest for French cuisine and lowest for Vietnamese cuisine.
- Southern cuisine has the widest spread (IQR) in cost of the order, while Korean cuisine has the narrowest.

Boxplot of Cost of the Order for Various Cuisine Types
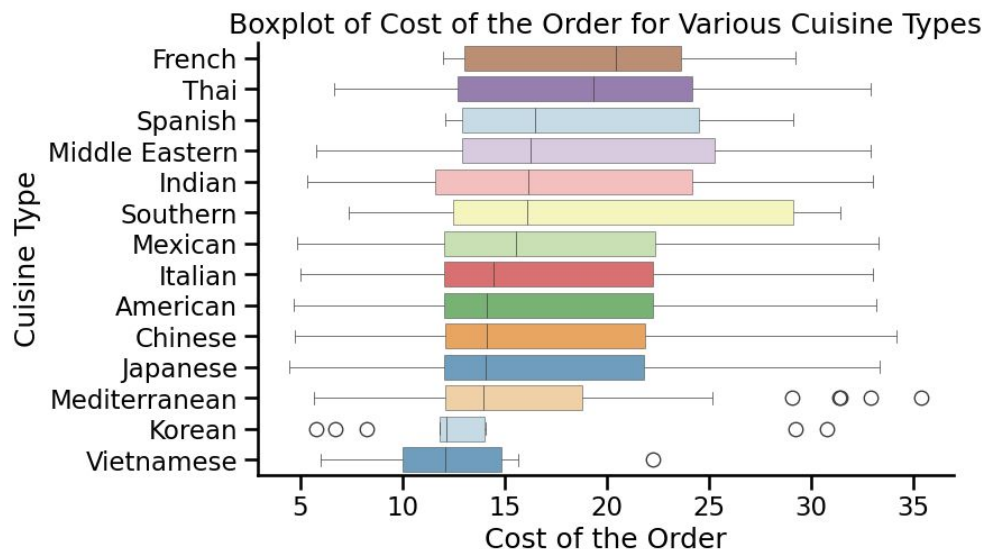
Multivariate Analysis

Cuisine type *vs* preparation time:

- The median food preparation time is highest for Italian cuisine and lowest for Korean cuisine.
- Thai cuisine has the widest spread (IQR) in food preparation time, while Korean cuisine has the narrowest.

Box plot of Food Preparation Time for Various Cuisine Types
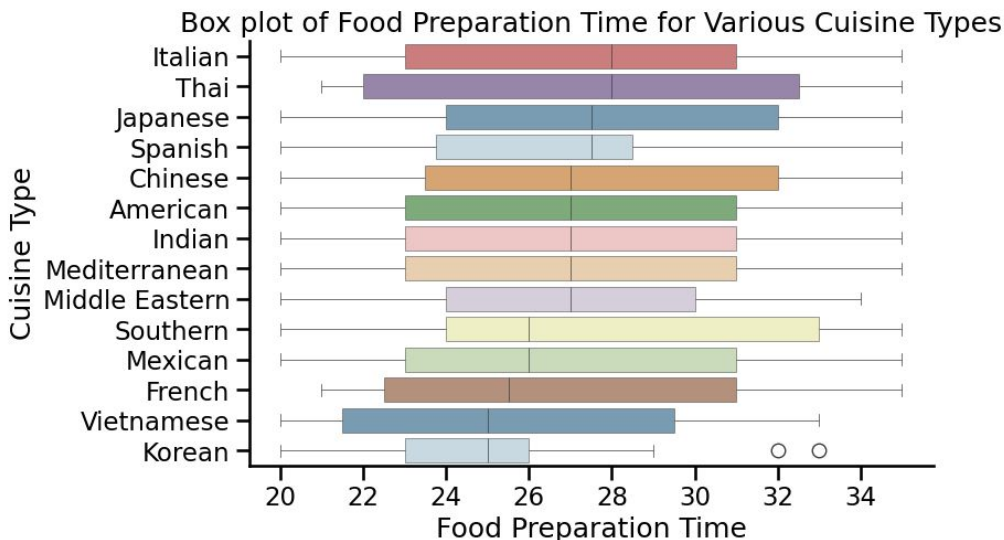
Multivariate Analysis

Day of the week *vs* delivery time:

- The median delivery time is longer on weekdays compared to the weekends.
- The average delivery time is around 28 minutes on weekdays and 22 minutes on weekends.
  - The average delivery time is 6 minutes longer on the weekdays compared to the weekend.

Boxplot of Delivery Time for Various Day of the Weeks
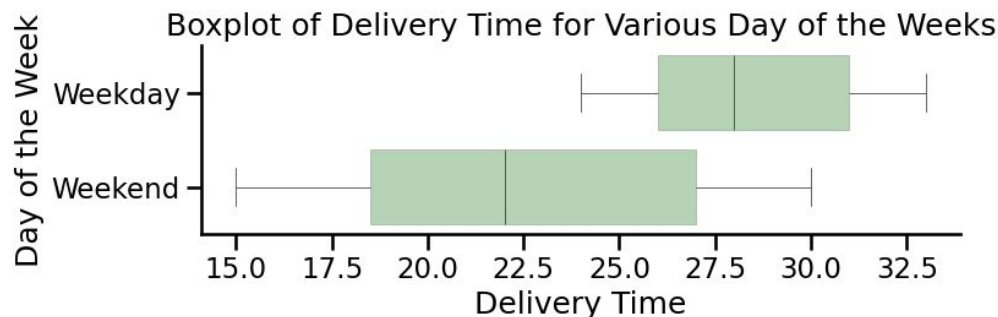
Delivery Time (min)

day_of_the_week	
Weekday	28
Weekend	22

Multivariate Analysis

Restaurants/day of the week *and* the cost of the order

- The table below shows the top 10 restaurants with the highest revenues.
- The total revenue is $31,315 and 72% of it is generated on weekends.

day_of_the_week	Revenue
Weekday	$8,922
Weekend	$22,393

restaurant_name	Revenue
Shake Shack	$3580
The Meatball Shop	$2145
Blue Ribbon Sushi	$1904
Blue Ribbon Fried Chicken	$1662
Parm	$1113
RedFarm Broadway	$965
RedFarm Hudson	$921
TAO	$834
Han Dynasty	$755
Blue Ribbon Sushi Bar & Grill	$667

Multivariate Analysis
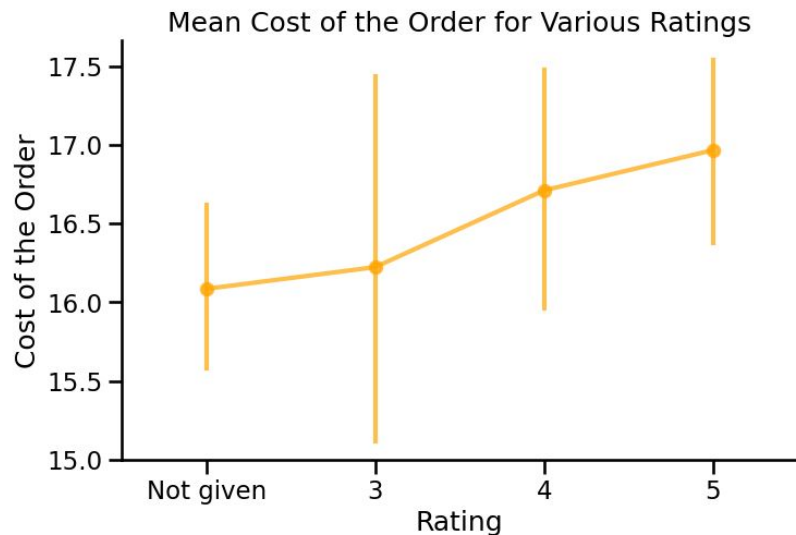
Rating *vs* delivery time/preparation time

- The ratings improve as the delivery and meal preparation time decrease.
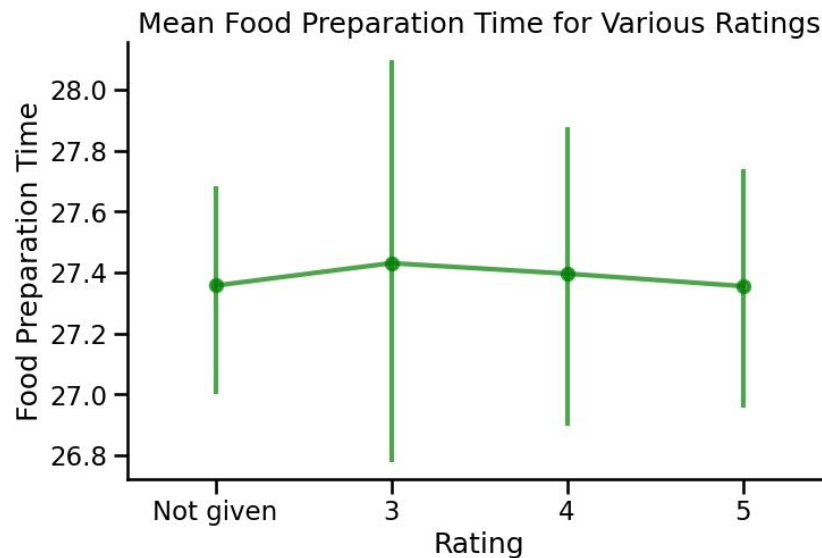
# Multivariate Analysis

Rating *vs* cost of the order

- Higher ratings are given to the higher-priced orders.

Mean Cost of the Order for Various Ratings

Multivariate Analysis

Correlation among variables (rating, cost, delivery time, preparation time)

- Negative correlation between the delivery/preparation time and the rating.
- Positive correlation between the cost and the rating.
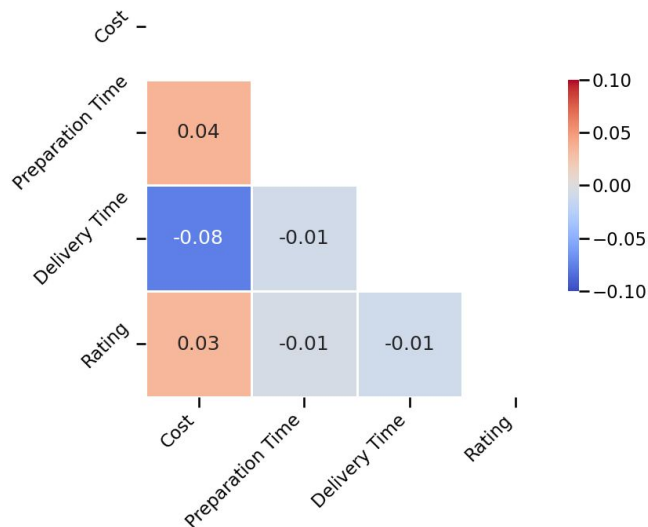
Promotional Offer

Condition to get the offer:

- Rating count > 50
- Average rating > 4

The following restaurants would receive the offer.

	count	mean
Restaurant		
Shake Shack	133	4.3
The Meatball Shop	84	4.5
Blue Ribbon Sushi	73	4.2
Blue Ribbon Fried Chicken	64	4.3

Revenue
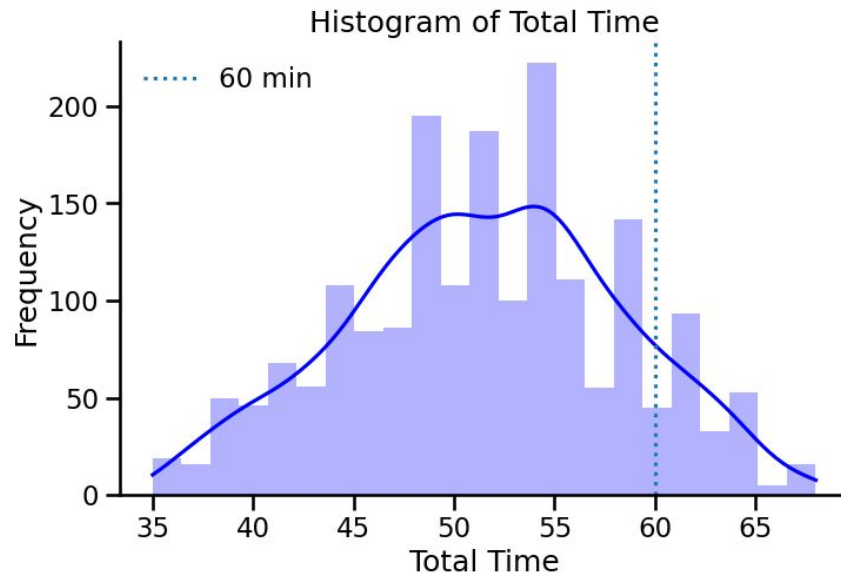
The company charges the restaurants

- 25% on the orders having cost greater than 20 dollars
- 15% on the orders having cost greater than 5 dollars

The net revenue is $6,166.30.

Total Time: preparation + delivery

Preparation time + delivery time

- Around 10.5% of orders take more than 60 min to get delivered from the time the order is placed.
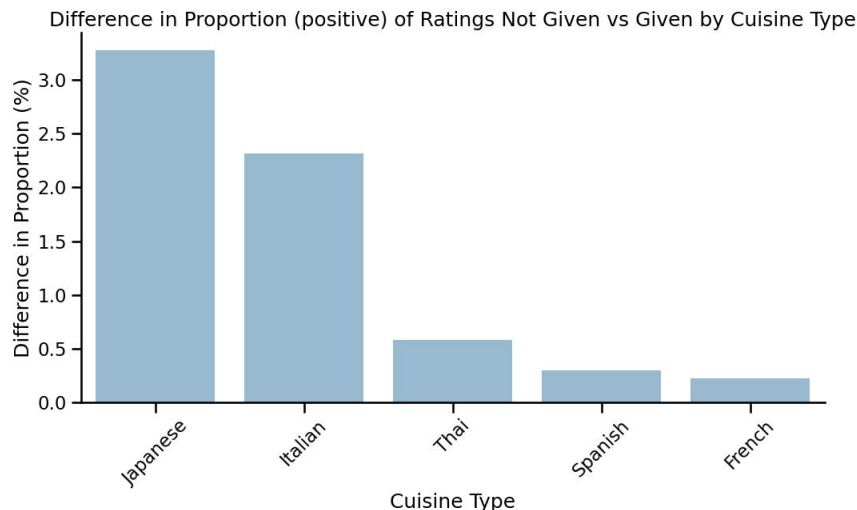

Histogram of Total Time

APPENDIX

Cuisine type vs rating (given/not given)

Here are the proportions of the cuisine types for orders when they are rated and not.

- Looking to see if in a particular cuisine, there are more cases of non-given ratings.

Difference in Proportion (positive) of Ratings Not Given vs Given by Cuisine Type

cuisine_type	Not Given	Given	Difference
Japanese	0.267663	0.234940	0.032723
Italian	0.171196	0.148021	0.023175
Thai	0.013587	0.007745	0.005842
Spanish	0.008152	0.005164	0.002989
French	0.010870	0.008606	0.002264
Vietnamese	0.002717	0.004303	-0.001586
Mexican	0.039402	0.041308	-0.001906
Korean	0.005435	0.007745	-0.002310
Chinese	0.111413	0.114458	-0.003045
Southern	0.005435	0.011188	-0.005753
Mediterranean	0.019022	0.027539	-0.008517
Middle Eastern	0.020380	0.029260	-0.008879
Indian	0.031250	0.043029	-0.011779
American	0.293478	0.316695	-0.023217

Happy Learning !