

# Foodhub Case Study

## Foodhub Data Analysis and Python for Data Science

February 12, 2024

# Contents / Agenda

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

# Data description

- The data contains the different data related to a food order. The detailed data dictionary is given below

## Data Dictionary

- Order\_id: Unique ID of the order
- Customer\_id: ID of the customer who orders the food
- Restaurant\_name: Name of the restaurant
- Cuisine\_type: Cuisine ordered by the customer
- Cost\_of\_the\_order: 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: rating given by the customer out of 5
- Food\_preparation\_time: Time (in minutes) the restaurant takes 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: (in minutes) the delivery person takes 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.

# Recommendations

- FoodHub should prioritize partnerships with restaurants offering American, Japanese, Italian, and Chinese cuisines, given their high popularity among our customers.
- To incentivize top-performing establishments, like Shake Shack, which fulfills a significant portion of orders, FoodHub should introduce promotional offers. This strategy could enhance customer satisfaction and encourage repeat business.
- Given the surge in order volumes over weekends, recruiting additional delivery personnel for these periods is advisable to guarantee prompt order fulfillment. Launching weekend-specific promotional offers could further elevate order volumes, maximizing revenue during these peak times.
- Customer ratings are indispensable for assessing satisfaction levels. An analysis to uncover the factors contributing to the low frequency of ratings is necessary. Enhancing the app's rating interface to be more engaging may encourage more customers to provide feedback on their orders.
- With approximately 11% of orders experiencing delivery times exceeding 60 minutes, FoodHub must address this issue to mitigate customer dissatisfaction. Offering incentives to delivery staff for consistently timely deliveries might be an effective strategy to improve service efficiency and customer contentment.

# Business Problem Overview and Solution Approach

- FoodHub, a food aggregator company, has compiled data from various orders placed by registered customers through their online platform. They aim to analyze this data to gain insights into the popularity of different restaurants. This analysis will enable them to improve the customer experience by tailoring their services to meet demand more effectively.
- Conclusions in analyzing the raw data obtained from the FoodHub portal will be carried out by ensuring that the data is clean and as complete as possible. The next step in data analysis will be to statistically analyze the data to look for patterns and trends that can inform business decisions. This will involve using various statistical techniques, such as descriptive statistics, to summarize the data and correlation analysis to identify relationships between different variables.
- The insights gained from this analysis will then be used to formulate recommendations for business strategy, such as targeting specific customer segments, optimizing product offerings, or improving operational efficiencies. Finally, the analysis will conclude with a review of the implications of these findings on the business and suggest actionable steps to leverage opportunities or mitigate challenges identified during the analysis.

# Data Overview

- The data under review includes:
- The minimum, average, and maximum time for food to be prepared once an order is placed.
- The number of orders customers have not rated.
- A ranking of cuisine types and their popularity by
  - Preference
  - Cost range
- Customer satisfaction with delivery time.
- Customer satisfaction with preparation time.
- Delivery time as it relates to weekday and weekend order delivery times.

# Understanding the Structure of the Data

1. How many rows and columns are present in the dataframe?

There are 1898 rows in the dataframe.

There are 9 columns in the dataframe.

2. What are the datatypes of the different columns in the dataset?

Column	Dtype
<b>Order_id</b>	Int64
<b>Customer_id</b>	Int64
<b>Restaurant_name</b>	Object
<b>Cuisine_type</b>	Object
<b>Cost_of_the_order</b>	Float64
<b>Day_of_the_week</b>	Object
<b>Rating</b>	Object
<b>Food_preparation_time</b>	Int64
<b>Delivery time</b>	Int64

# Understanding the Structure of the Data

3. Are there any missing values in the dataframe?

There are no missing values in the data frame.

4. What is the minimum, average, and maximum time for food to be prepared once an order is placed?

	order_id	customer_id	cost_of_the_order	food_preparation_time	delivery_time
count	1.898000e+03	1898.000000	1898.000000	1898.000000	1898.000000
mean	1.477496e+06	171168.478398	16.498851	27.371970	24.161749
std	5.480497e+02	113698.139743	7.483812	4.632481	4.972637
min	1.476547e+06	1311.000000	4.470000	20.000000	15.000000
25%	1.477021e+06	77787.750000	12.080000	23.000000	20.000000
50%	1.477496e+06	128600.000000	14.140000	27.000000	25.000000
75%	1.477970e+06	270525.000000	22.297500	31.000000	28.000000
max	1.478444e+06	405334.000000	35.410000	35.000000	33.000000

The minimum time it takes for food to be prepared once an order is placed is: 20 minutes.  
 The average time it takes for food to be prepared once an order is placed is: 27.37 minutes.  
 the Maximum time it takes for food to be prepared once an order is placed is: 35 minutes.

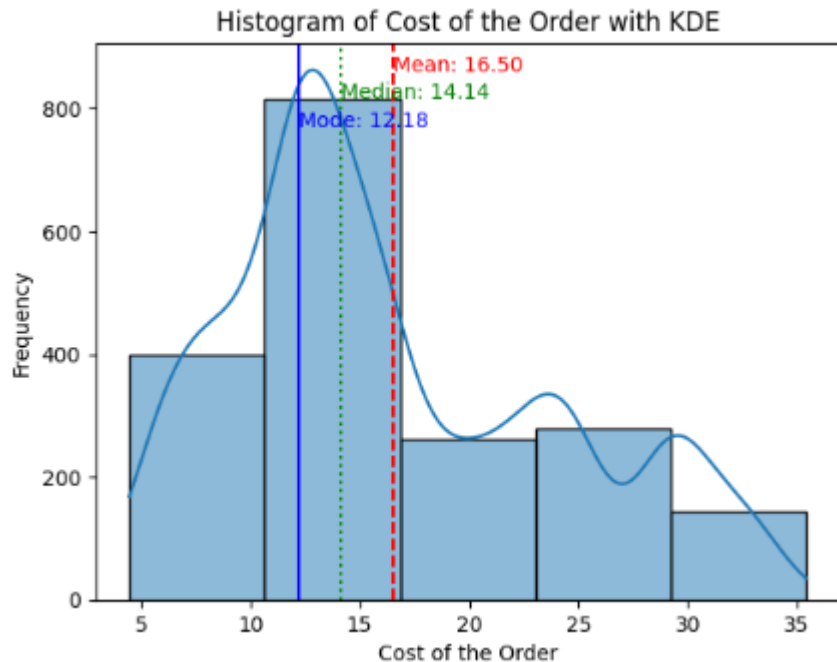


# EDA – Univariate Analysis Questions

6. Explore all the variables and provide observations on their distributions.
  - a. There are 1898 unique order IDs.
  - b. There are 1200 unique customer IDs.
  - c. There are 178 unique restaurant names.
  - d. There are 19 unique cuisine types.

# EDA – Univariate Analysis Questions

- Cost of the Order



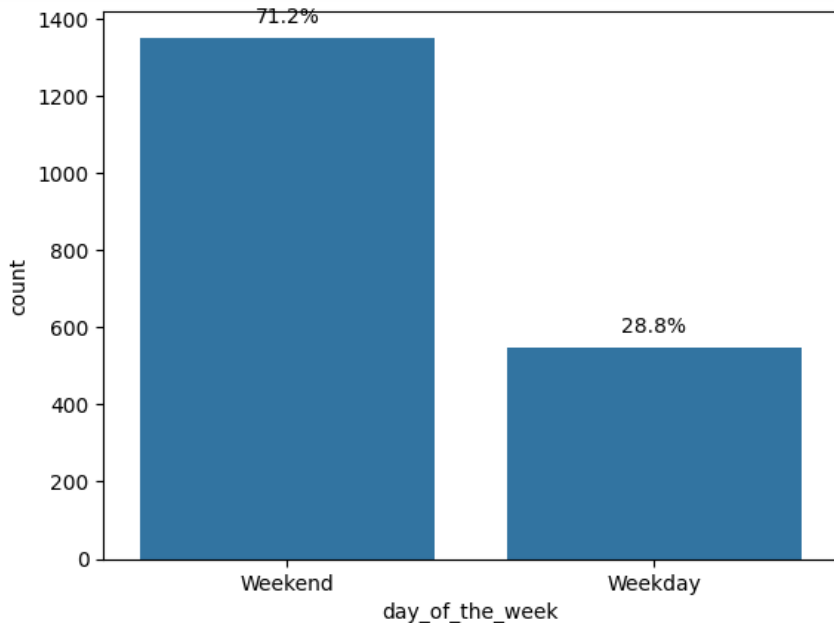
The average cost of the order (Mean) is greater than the median cost indicating that the distribution for the cost of the order is right-skewed.

The mode of the distribution indicates that a large chunk of customers prefers to order food that costs around 10 – 12 dollars.

There are few orders that cost greater than \$30. These orders might be for more expensive meals.

# EDA – Univariate Analysis Questions

- Day of the Week



During the week there are approximately 547 cuisine orders.

During the weekend there are about 1351 cuisine orders.

Observations:

Customers seem to prefer to order food online during the weekend.

Cuisine was ordered approximately, on average, 590 times on weekdays.

- Number of unique days of the week 2

# EDA – Univariate Analysis Questions

5. How many orders have customers not rated?

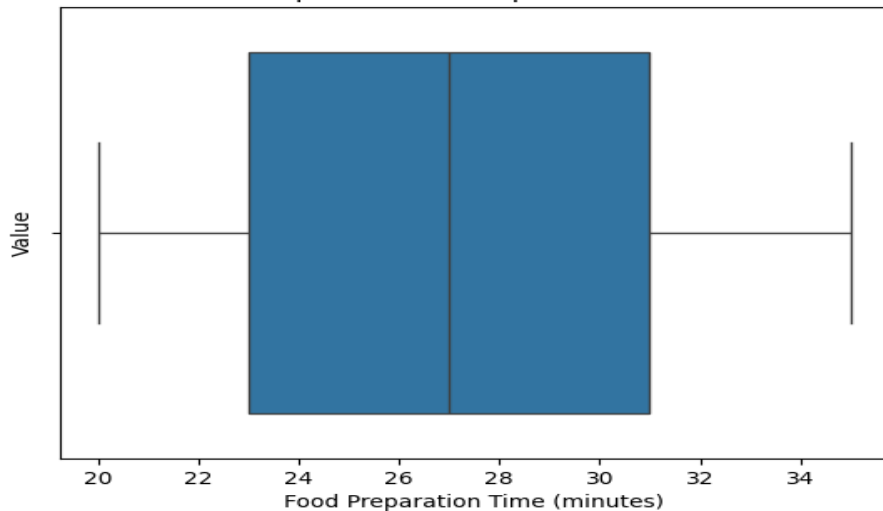
Food Order Ratings	Count
Not given	736
5	588
4	386
3	188

There were 736 orders not rated

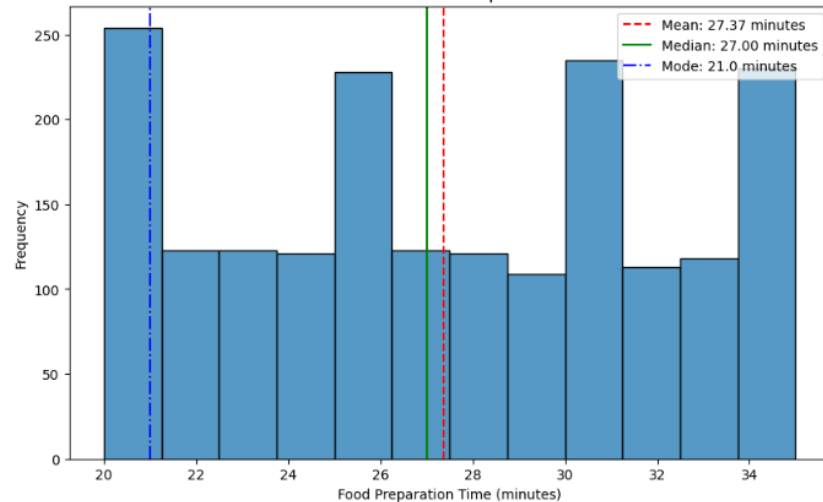
# EDA – Univariate Analysis Questions

- Distribution of food preparation time

Boxplot of Food Preparation Time



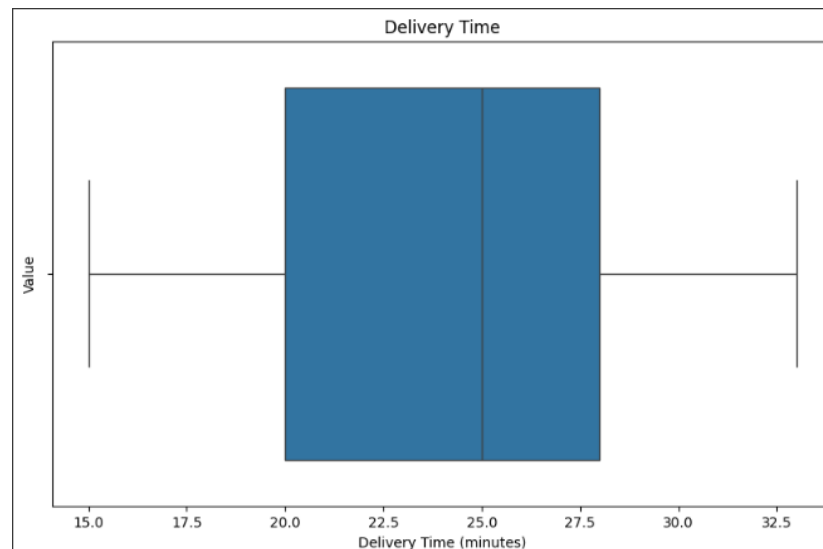
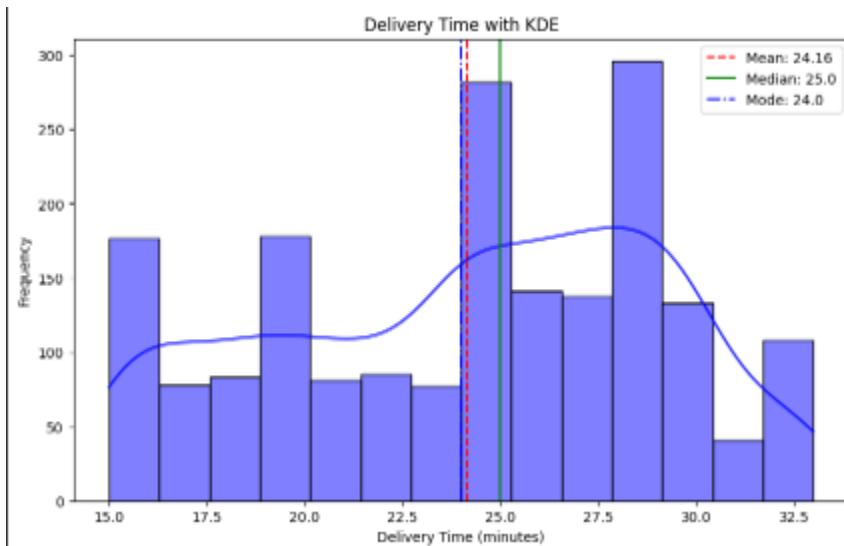
Distribution of Food Preparation Time



The boxplot mean, and the histogram median suggests that the average time to prepare any order was 27 minutes.

The mode in the bar chart suggests that most of the orders were prepared in 21 minutes.

# EDA – Univariate Analysis Questions



The boxplot shows a minimum delivery time is 15 minutes.

The (Mean) average time in the histplot and boxplot for order delivery is 24.16 minutes.

The boxplot shows the Maximum delivery time is 33 minutes.

# EDA – Univariate Analysis Questions

7. Which are the top 5 restaurants regarding the number of orders received?

Top Five Restaurants (in descending order)	
Restaurant	Number of orders
Shake Shack	219
The Meatball Shop	132
Blue Ribbon Sushi	119
Blue Ribbon Chicken	96
Parm	68

3. Which is the most popular?
4. What percentage of the orders cost more than 20 dollars?
5. What is the mean order delivery time?

## EDA – Univariate Analysis Questions

8. Which is the most popular cuisine on weekends?

The most popular cuisine on weekends is American.

9. What percentage of the orders cost more than 20 dollars?

The percentage of orders above 20 dollars is 28.03%

10. What is the mean order delivery time?

The Mean delivery time for this dataset is 24.16 minutes.



## EDA – Univariate Analysis Questions

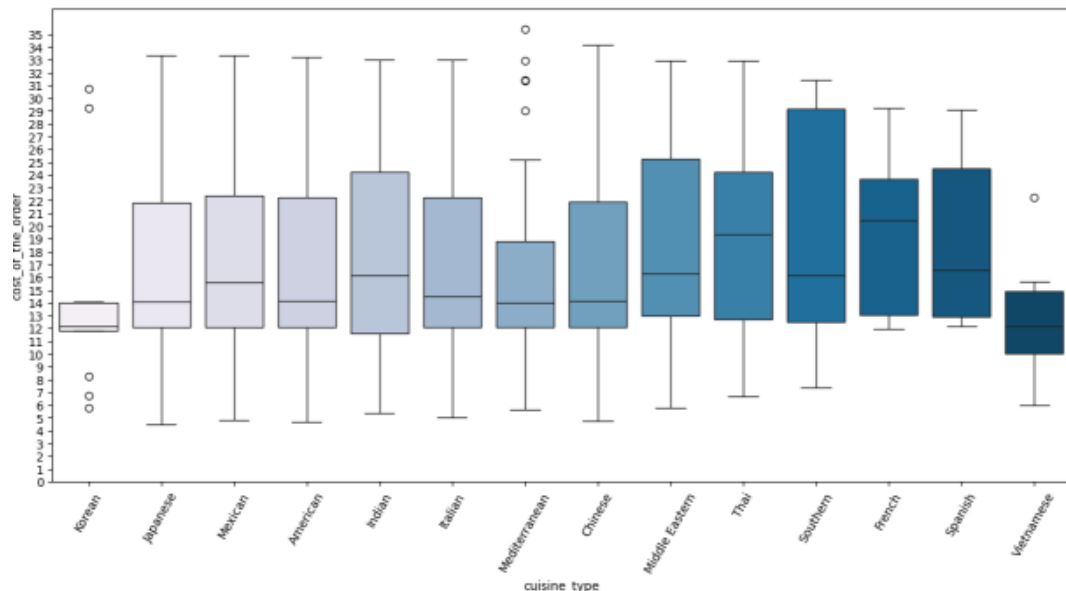
11. the company has decided to give 20% discount vouchers to the top 3 most frequent customers. Find the IDs of these customers and the number of orders they placed.

Top three most frequent customers	
Customer number	Number of orders
52832	13
47440	10
83287	9

# EDA – Multivariate Analysis Questions

Question 12. Perform a multivariate analysis to explore relationships between the critical variables in the data set.

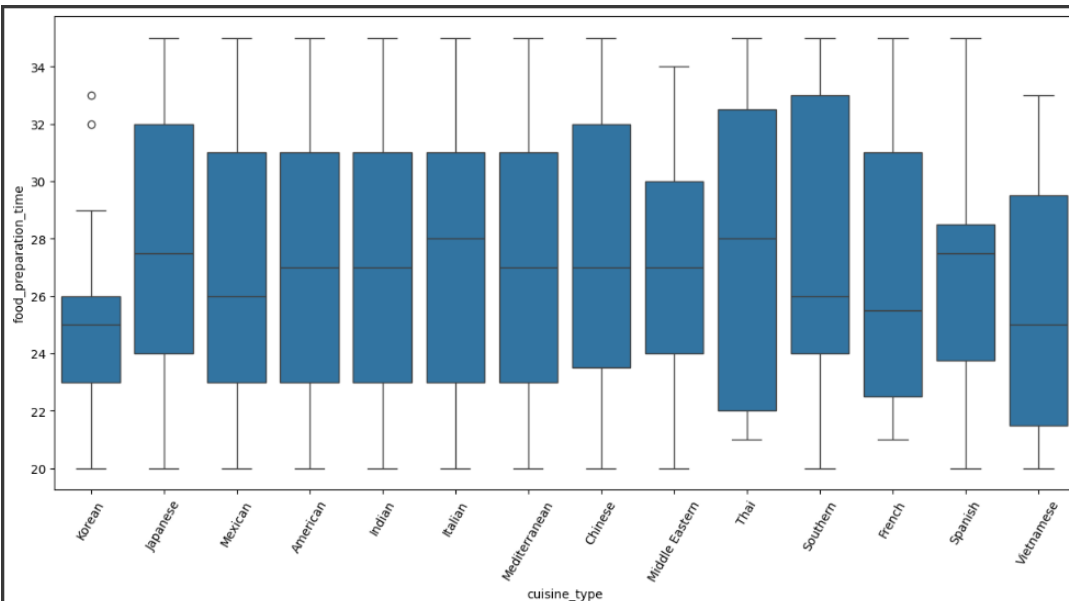
## Cost of the order



- Vietnamese and Korean cuisines are less expensive than the other 14 cuisines.
- the costs for Italian, American, Chinese, and Japanese cuisines are similar at \$22.
- Outliers are present for Korean, Mediterranean and Vietnamese cuisines.
- The most expensive cuisines are Southern at \$29 and Middle Eastern cuisines at \$25.
- \$25.

# EDA – Multivariate Analysis Questions

- Cuisine vs Food Preparation time



Food preparation time is very consistent for most of the cuisines the median food preparation time lies between 24 and 30 minutes for all the cuisines.

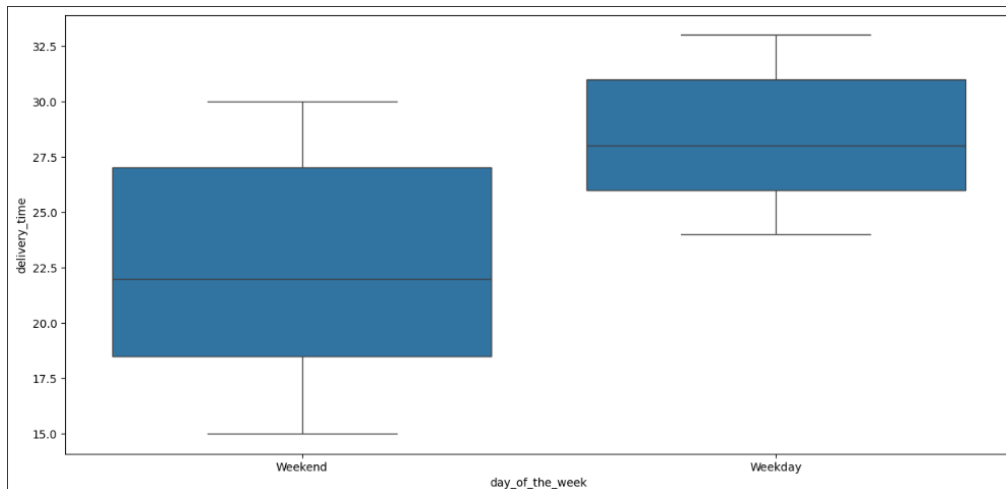
Outliers are present for the food preparation time of Korean cuisine.

Korean cuisine is less preparation time intensive at 26 minutes.

Southern cuisine is the most preparation time intensive 33 minutes.

The next time-intensive cuisine is Thai at 32 minutes, followed by Chinese cuisine at 31 minutes. The median food preparation time

# EDA – Multivariate Analysis Questions



Overall, delivery times during the weekend are between 15 and 30 minutes. The median delivery time for weekends is 22 minutes, with 50% of the deliveries being made under 22 minutes.

Overall delivery times during the weekday are between 24 and 33 minutes. The median delivery time for weekdays is 28 minutes, with 50% of the deliveries being made under 28 minutes.

# EDA – Multivariate Analysis Questions

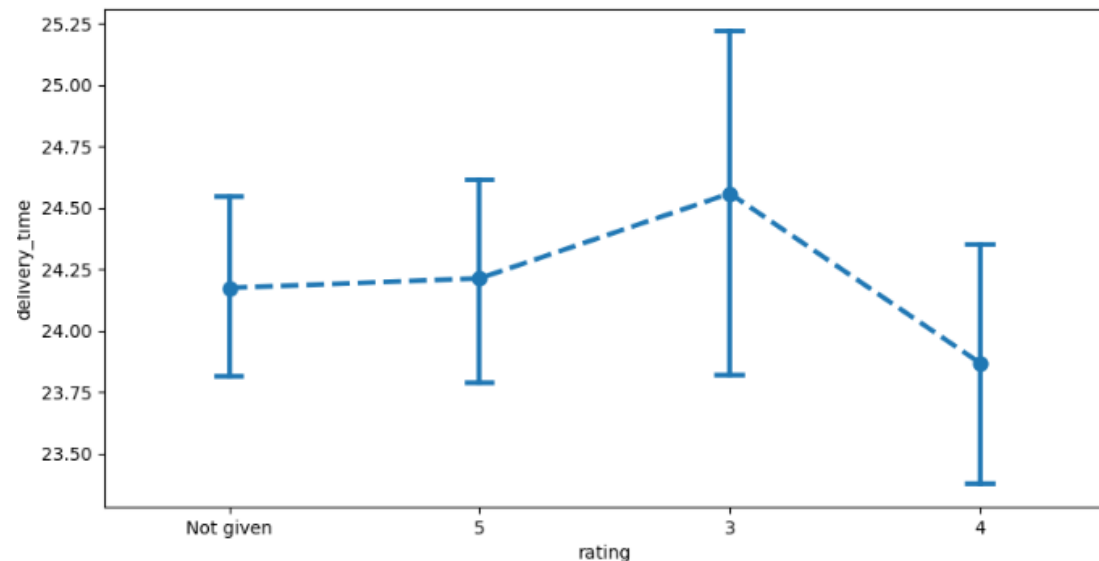
Revenue generated by the restaurants

Restaurant Name	Revenue (in U.S. dollars)
Shake Shack	3493
The Meatball Shop	2092
Blue Ribbon Sushi	1857
Blue Ribbon Fried Chicken	1628
Parm	1084
RedFarm Broadway	940
RedFarm Hudson	903
TAO	816
Han Dynasty	739
Blue Ribbon Sushi Bar & Grill	648
Rubirosa	645
Sushi of Gari 46	630
Nobu Next Door	606
Five Guys Burgers and Fries	500

These 14 restaurants are generating more than \$500 in revenue.

# EDA – Multivariate Analysis Questions

- Rating vs. Delivery time



The average delivery time for orders not rated is approximately 24.15 minutes, meaning that deliveries not rated took an approximate average of 24.15 minutes.

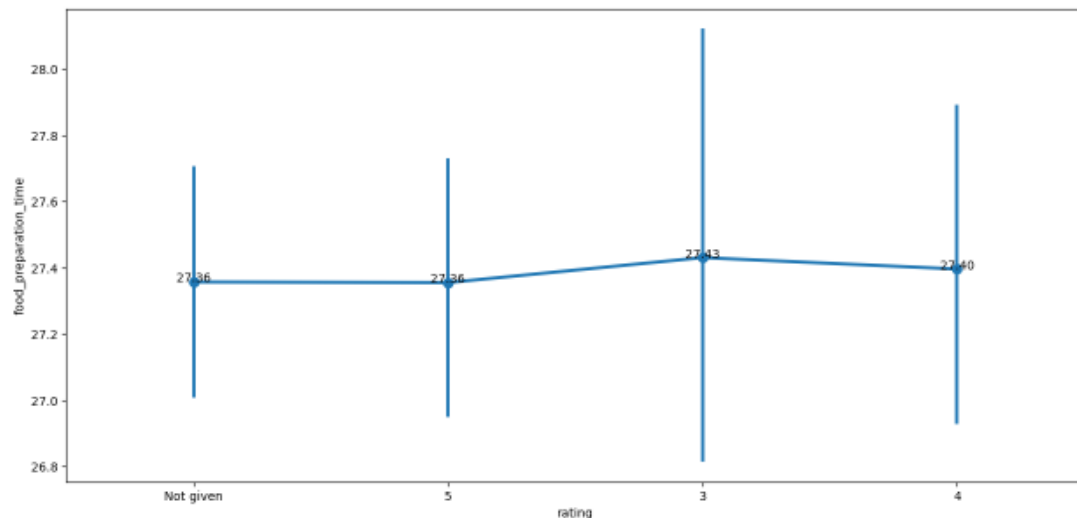
The average delivery time for orders rated as 5 is 24.25 minutes, meaning that deliveries rated as 5 took approximately 24.25 minutes.

The average delivery time for orders rated as 3 is approximately 24.60 minutes, meaning that deliveries rated as 3 took approximately 24.60 minutes.

The average delivery time for orders rated as 4 is approximately 23.90 minutes, meaning that deliveries rated as 4 took approximately 23.90 minutes.

# EDA – Multivariate Analysis Questions

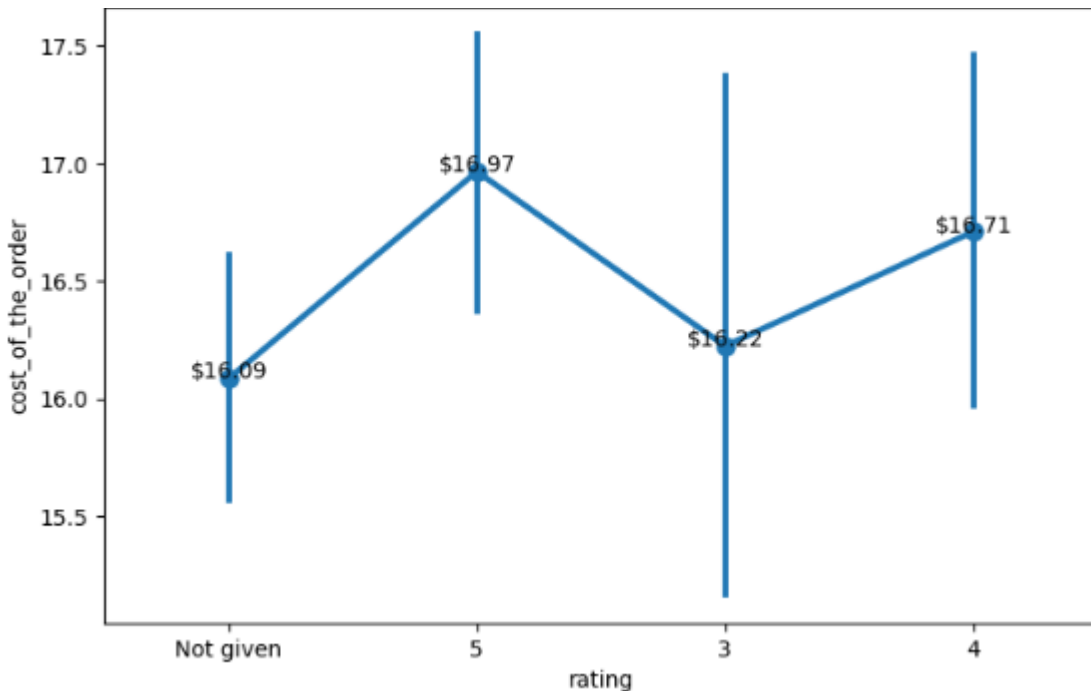
- Rating vs. Food preparation time



The very close proximity of data points around 27.4 on the y-axis suggests that customers do not significantly factor in food preparation time when rating cuisine, or that the variation is not enough to affect ratings significantly.

# EDA – Multivariate Analysis Questions

- Rating vs. Cost of the order



High-cost orders are being rated well and low-cost orders are not being rated. cuisine costing, on average, \$16.90 was not rated.

cuisine costing on average \$16.97 was rated at 5

cuisine costing on average \$16.22 was rated at 3

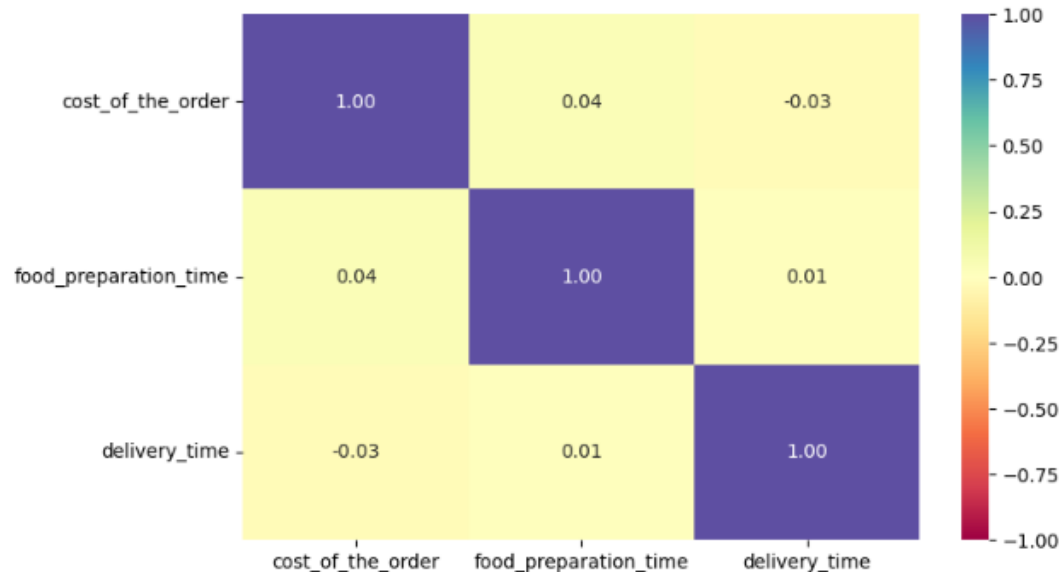
Cuisine costing on average \$16.71 was rated at 4

It appears that with the average cost of the order being closely clustered around \$16 the rating for the cost of the order may be influenced by the quality of the meal rather than anything else,



# EDA – Multivariate Analysis Questions

- Correlation among variables



In a heatmap, values approaching one in the cells where columns and rows intersect indicate a stronger correlation between the respective column and row variables.

- With a value of -0.03, the map suggests that there is no correlation between the cost of the order and delivery time.
- With a value of 0.04, between the cost of the order and the food preparation time, the map suggests there is no correlation between the cost of the order and food preparation time

# EDA – Multivariate Analysis Questions

## Question 13

The company wants to provide a promotional offer in the restaurant's advertisement. The condition to get the offer is that the restaurants must have a rating count of more than 50, and the average rating should be greater than 4. Find the restaurants fulfilling the criteria to get the promotional offer.

Restaurants with rating count over 50	
Restaurant name	Rating
Shake Shack	133
The Meatball Shop	84
Blue Ribbon Sushi	73
Blue Ribbon Fried Chicken	64

# EDA – Multivariate Analysis Questions

## Question 14

The company charges the restaurant 25% on orders greater than 20 dollars and 15% on orders greater than 5 dollars. Find the net revenue generated by the company across all orders.

	order_id	customer_id	restaurant_name	cuisine_type	cost_of_the_order	day_of_the_week	rating	food_preparation_time	delivery_time	Revenue
0	1477147	337525	Hangawi	Korean	30.75	Weekend	Not given	25	20	7.6875
1	1477685	358141	Blue Ribbon Sushi Izakaya	Japanese	12.08	Weekend	Not given	25	23	1.8120
2	1477070	66393	Cafe Habana	Mexican	12.23	Weekday	5	23	28	1.8345
3	1477334	106968	Blue Ribbon Fried Chicken	American	29.20	Weekend	3	25	15	7.3000
4	1478249	76942	Dirty Bird to Go	American	11.59	Weekday	4	25	24	1.7385

The net revenue generated on all the orders given in the dataset is approximately 61.66.3 dollars.

# EDA – Multivariate Analysis Questions

## Question 15

The company wants to analyze the total time required to deliver the food. What percentage of orders take more than 60 minutes to get delivered from the time the order is placed?

Approximately 10.54% of the total orders have more than 60 minutes of total delivery time.

# EDA – Multivariate Analysis Questions

## Question 16

The company wants to analyze the delivery time of the orders on weekdays and weekends. How does the mean delivery time vary during weekdays and weekends?

The mean delivery time on weekdays is around 28 minutes

# Analysis Questions

15. The company wants to analyze the total time required to deliver the food. What percentage of orders take more than 60 minutes to get delivered from when the order is placed?

- Approximately 10.54% of the orders have more than 60 minutes of total delivery time.

16. The company wants to analyze the delivery time of the orders on weekdays and weekends. How does the delivery time vary during weekdays and weekends?

- The mean delivery time on weekdays is around 26 minutes whereas the mean delivery time on weekends is around 22 minutes.
- This could be due to the reduction in traffic volume on the weekends

# Executive Summary – Question 17

- **Conclusions**

- 1. Approximately 80% of the orders are for American, Japanese, Italian, and Chinese cuisines, making them the most popular choices of FoodHub customers.**
- 2. Shake Shack is the most popular restaurant that has received the highest number of orders.**
- 3. Order volumes increase on the weekends compared to the weekdays.**
- 4. Delivery time over the weekends is less compared to the weekdays. The reduction in deliveries could result from decreased road traffic during the weekends.**
- 5. 38.8% of orders were not rated**