FoodHub Analysis

By: Joshua Willis
Contact: jkw910@gmail.com

Contents / Agenda

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

Executive Summary

- This presentation consists of the findings for a data science analysis of the order data from the food delivery company FoodHub. The analysis conducted was done so to identify scenarios where FoodHub can enhance the customer experience.

Findings and Recommendations:

- The most popular cuisines are American, Japanese, Italian, and Chinese. This could suggest adding more restaurants in these cuisine types would enhance the customers experience.
- The most common order cost is between 10-15$, this information could be used to target more restaurants in that price range or suggest current ones have more options in that range.
- The majority of orders, by almost double, were weekend orders as opposed to weekdays. The average time of the orders were also 5 minutes slower on weekdays. This could reveal an opportunity to run specials or flag delivery times due to weekday traffic.
- The ratings show that there is a much higher variance on food preparation and delivery times for the lower rated food, perhaps indicating a restaurants consistency. The rating is higher for the more expensive cost to order findings though, indicating food quality could be a factor for higher ratings.
- 10% of the orders have over 60 minutes from ofeder till delivery, this should be flagged and analyzed further.

Business Problem Overview and Solution Approach

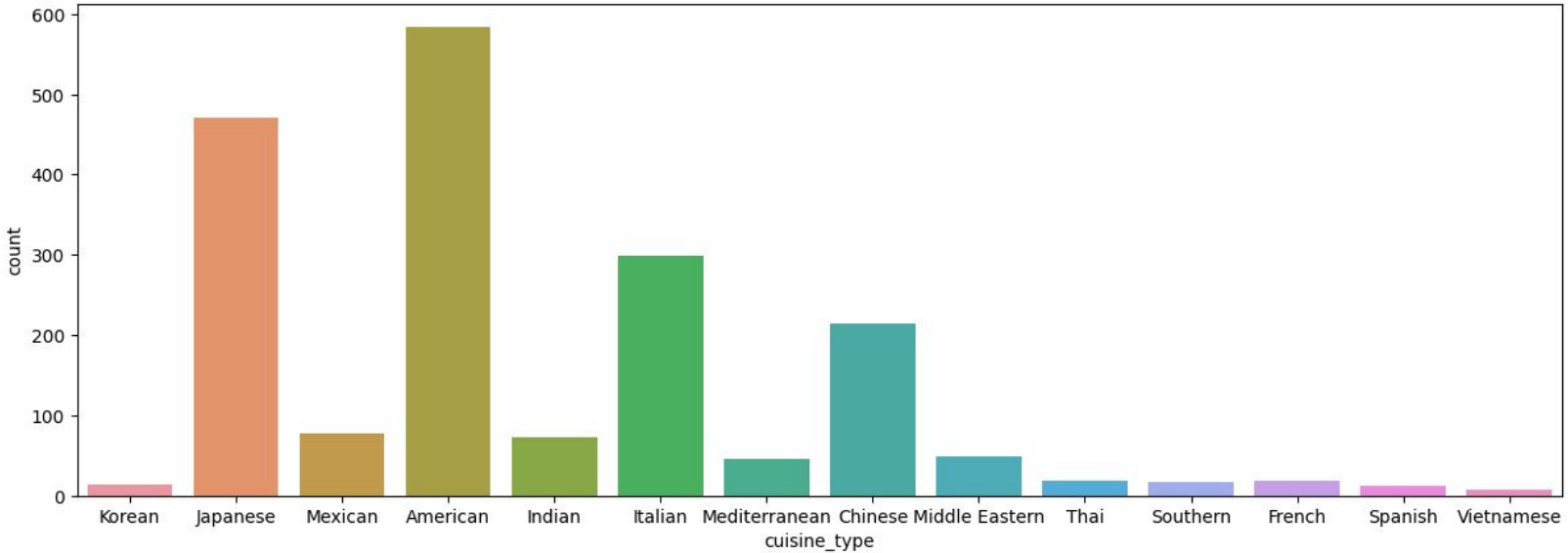
- Business Problem: FoodHub has the need to analyze orders and the associated information that has been gathered from the companies users to get an idea of demand of cuisines and restaurants, time to prep and order the food to delivery, where the majority of net revenue is coming from and customers restaurant experience ratings with these factors present. This information can ultimately lead to the goal of a better customer experience for the FoodHub user.
- Solution Approach: The approach to this dataset and business problem is to initially clean and prep the data, after taking a look at the columns and types, by removing duplicates or missing values and checking for any nulls. Then, complete a statistical summary of the data to ensure the clean up was correct and nothing appears out of the ordinary. Then complete analysis of the individual fields (cuisine types, rating, cost, timing etc) and draw some initial conclusions before moving to an analysis of the data by analyzing the relationships between the different fields. By comparing the relationships from the different fields, such as order number to cuisine type or restaurant, we can determine popularity of restaurants and cross reference that with rating, revenue, or delivery times to come to some confounded business insights.

Data Overview

- There are **1898** rows of data.
- The datatypes are as follows for the columns - o**rder_id: int, customer_id: int, restaurant_name: str, cuisine_type: str, cost_of_the_order: float, day_of_the_week: str, rating: int, food_preparation_time: int, delivery_time: int.**
- There are **no NULLS**.
- For food to be prepared the **minimum** time is **20 min**, the **maximum** time is **35 min**, and the **average** time is **27 min** (27.37).
- **736** of the orders have the rating **Not Given**.

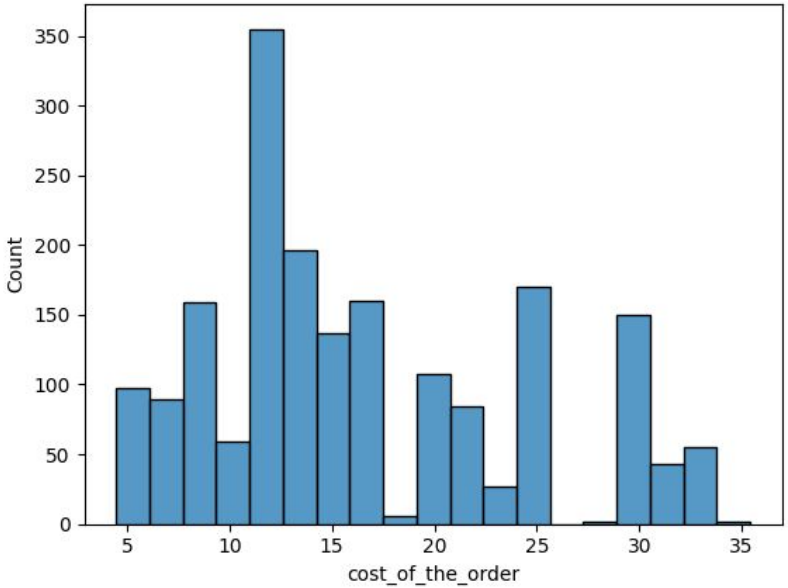
Univariate Analysis

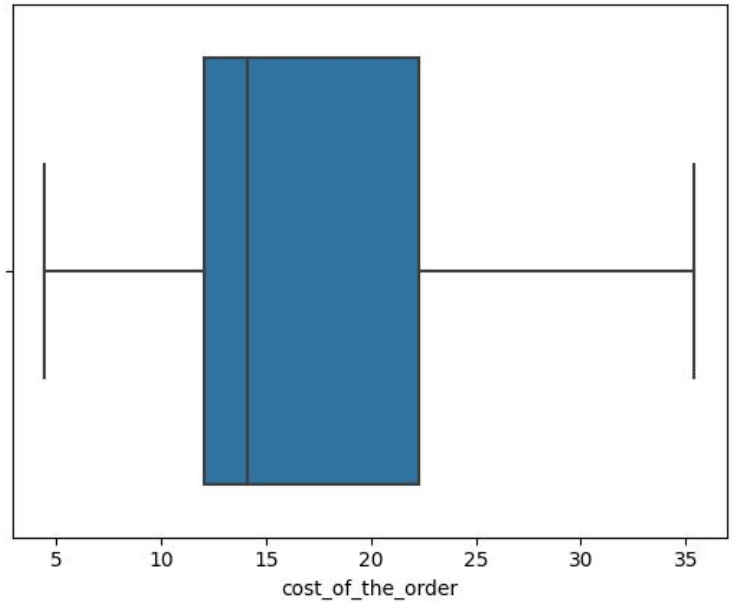
Order ID has **1898** unique findings, one for every line in the entries. Customer ID shows that there are **1200** different customers. Restaurant Name shows that there have been **178** different restaurants ordered from. Cuisine type shows that there have been 14 different types ordered.



The plot for cuisine type shows that the most popular are American, Japanese, Italian and Chinese. While the least popular are Vietnamese and French.

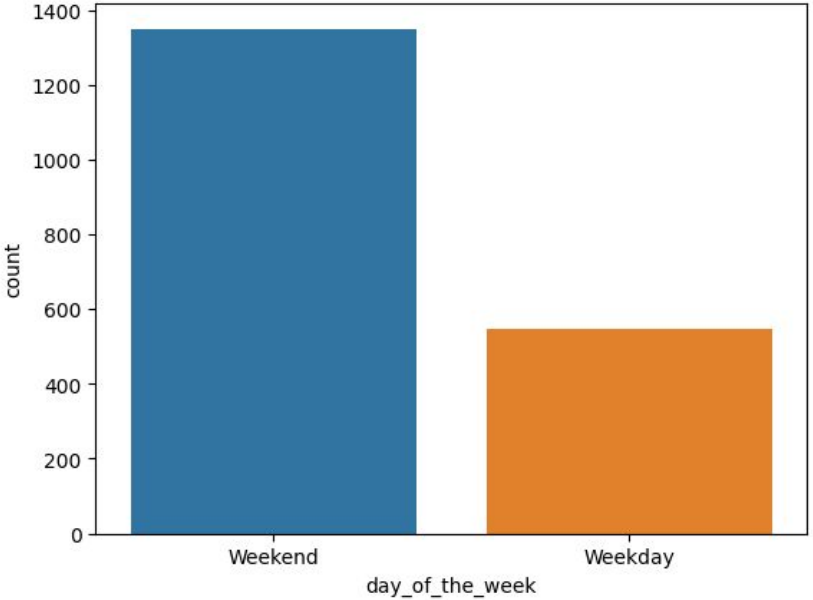
Univariate Analysis

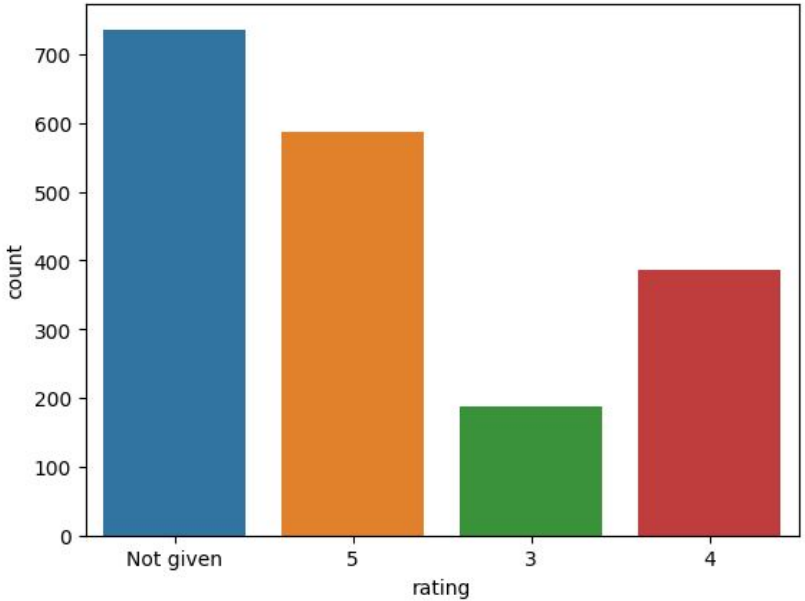




The plots for cost of order show a surge in pricing between the 10 and 15 dollar markers then a decline and a bump back up to orders 25 dollars all the way to 35 being the most expensive. The average being shown on the right to be just below 15 dollars.

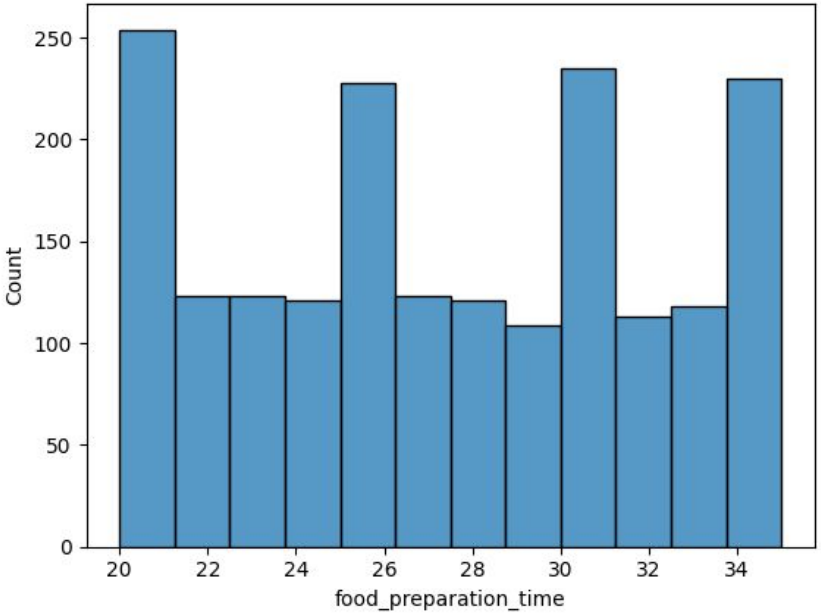
Univariate Analysis

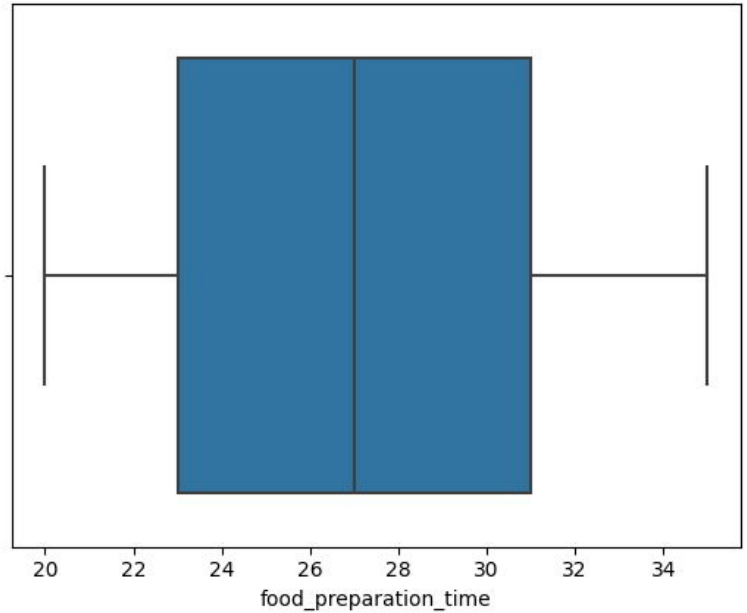




The plot for orders coming in as a weekday versus a weekend show that there are over double the amount of orders on the weekends. That ratings plot on the right shows us that the majority of establishments are rated a 5 or not given a rating at all with the remaining ratings (3 and 4) making up about 30% of the count.

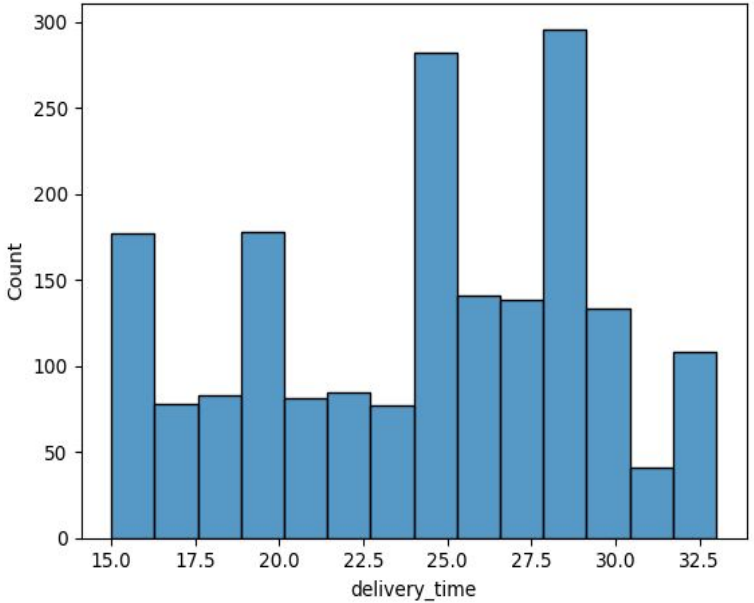
Univariate Analysis

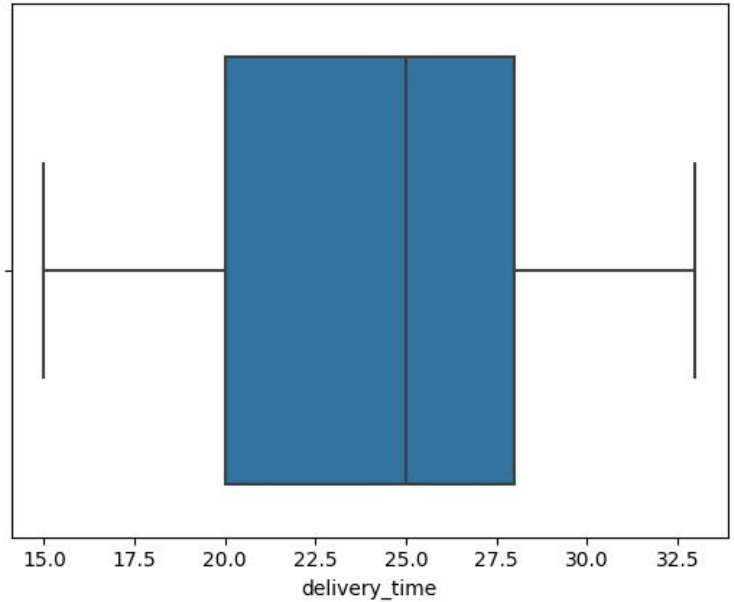




The plots for food preparation time show us a range of 20 to 35 minutes with an average right at 27 minutes. The histogram shows a unique wave pattern not seen in the other variables that may show some structured prep time around the numbers 20, 25, 30, and 34.

Univariate Analysis



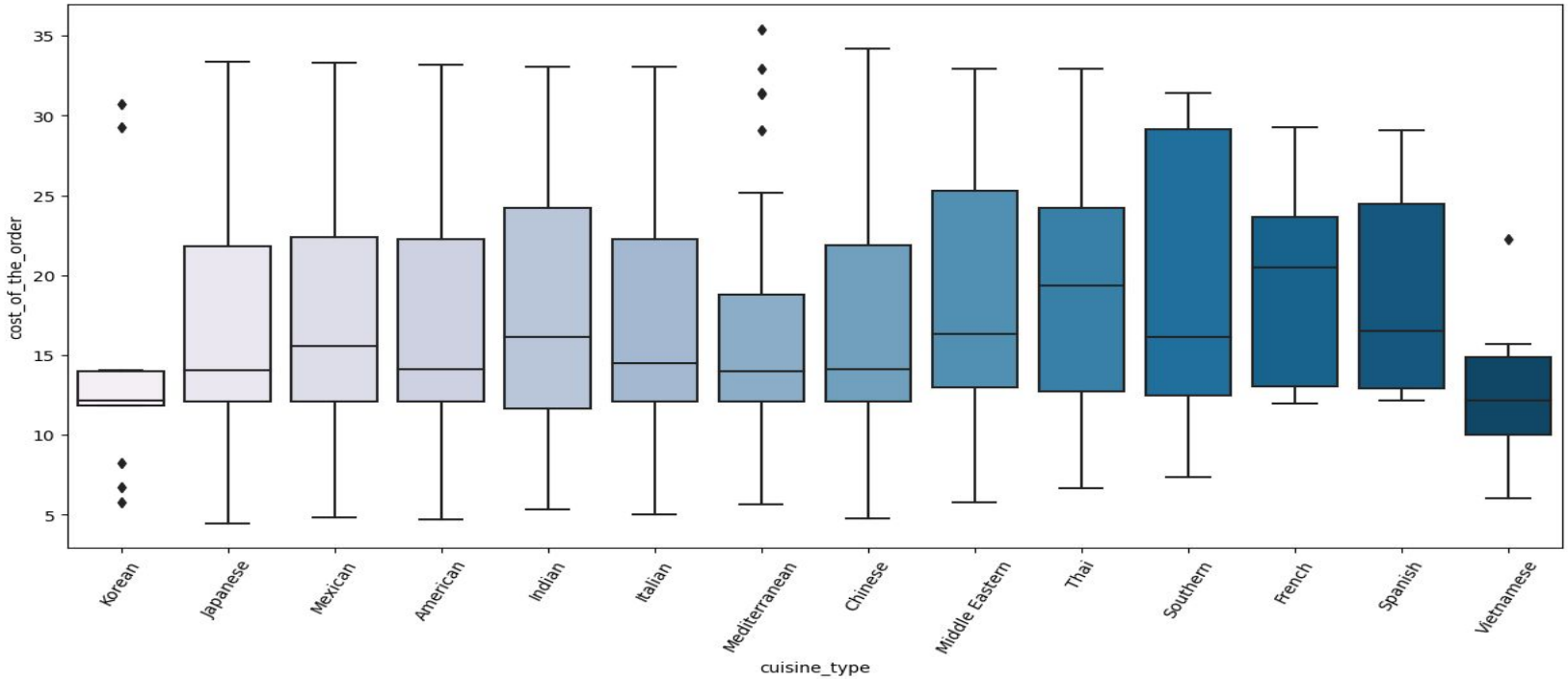


The plots for delivery time show us a range of 15 to 33 minutes with an average right at 25 minutes. The histogram shows is a wave pattern similar to the food preparation graph showing spikes in delivery time around round and digestible numbers like 15, 20, 25, and just before 30.

Univariate Analysis

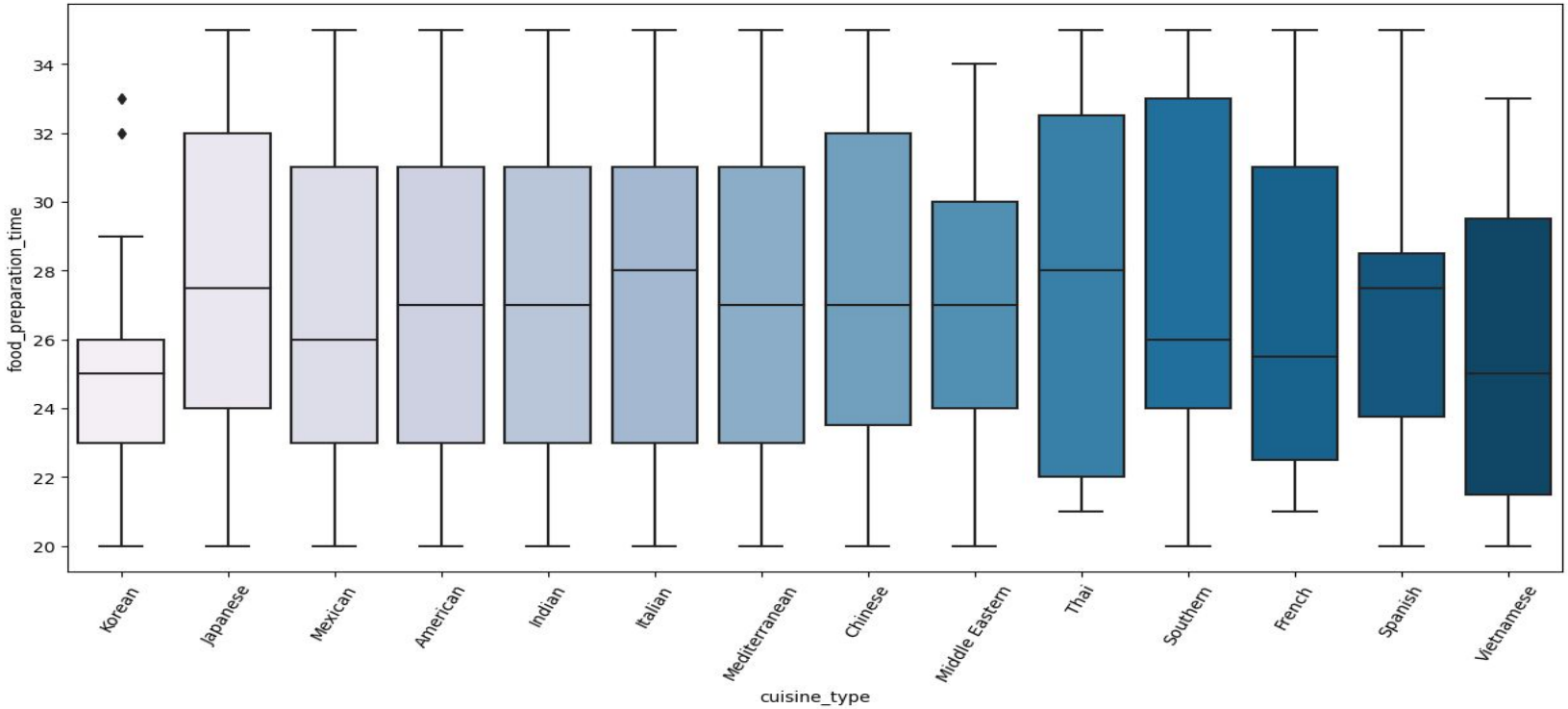
- Shake Shack, The Meatball Shop, Blue Ribbon Sushi, Blue Ribbon Fried Chicken, and Parm are the top 5 restaurants in order of number received.
- The most popular cuisine on the weekend is **American** (415 orders).
- **29.24%** of the orders cost over 20$, that is a count of 555 orders.
- The mean delivery time is **24.16 minutes**.
- The top 3 customers to give a discount to are **52832, 47440,** and **83287**.

Multivariate Analysis



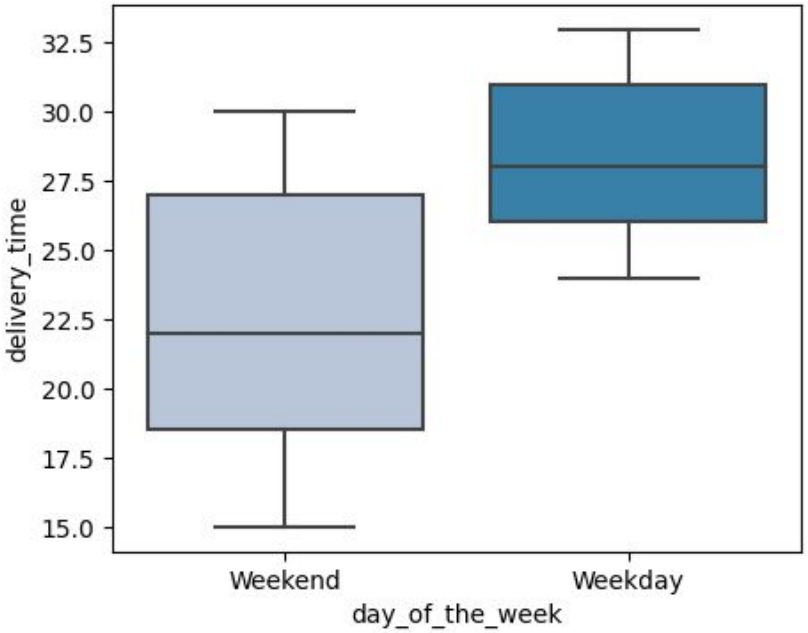
This plot shows a comparison of the cuisine type by the cost of the order, resulting in a graph that shows that the cost orders min and max are not correlated with most cuisine types besides the ones of smaller sample size mapped out earlier (Vietnamese and Korean).

Multivariate Analysis



This plot shows a comparison of the cuisine type by the food preparation time, resulting in a graph that shows that the food prep time min and max are not correlated with most cuisine types. The Average preparation times are highest for Italian food and lowest for Vietnamese.

Multivariate Analysis

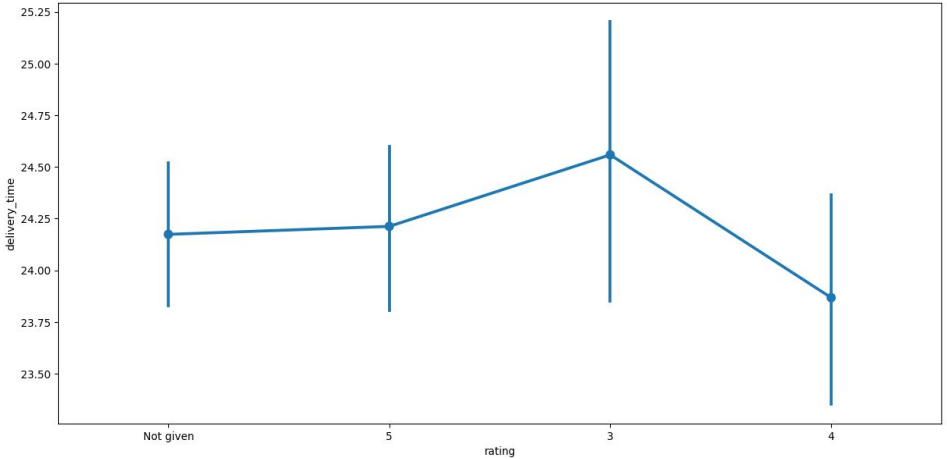


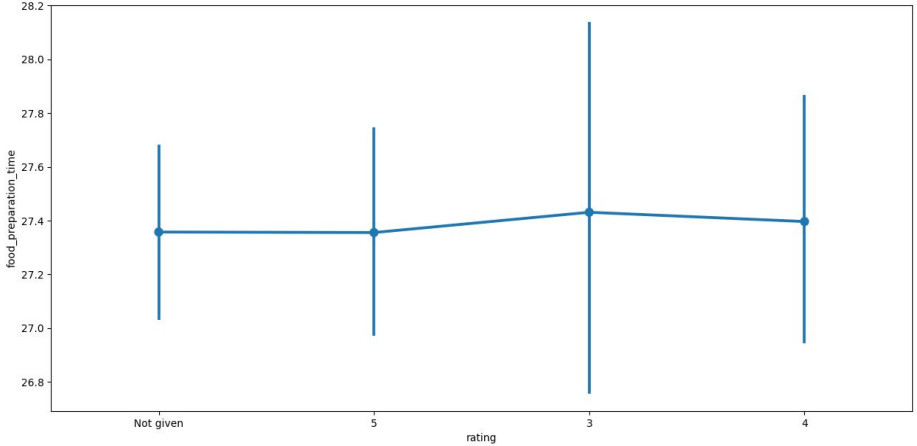
restaurant_name	
Shake Shack	3579.53
The Meatball Shop	2145.21
Blue Ribbon Sushi	1903.95
Blue Ribbon Fried Chicken	1662.29
Parm	1112.76
RedFarm Broadway	965.13
RedFarm Hudson	921.21
TAO	834.50
Han Dynasty	755.29
Blue Ribbon Sushi Bar & Grill	666.62
Rubirosa	660.45
Sushi of Gari 46	640.87
Nobu Next Door	623.67
Five Guys Burgers and Fries	506.47

Name: cost_of_the_order, dtype: float64

The plot on the left indicates that the delivery time is significantly longer for Weekday orders over weekend with an almost 5 minute differential for their averages. On the right is a breakdown of the restaurants revenue showing the most profitable chain is Shake Shack and the lowest grossing is Five Guys.

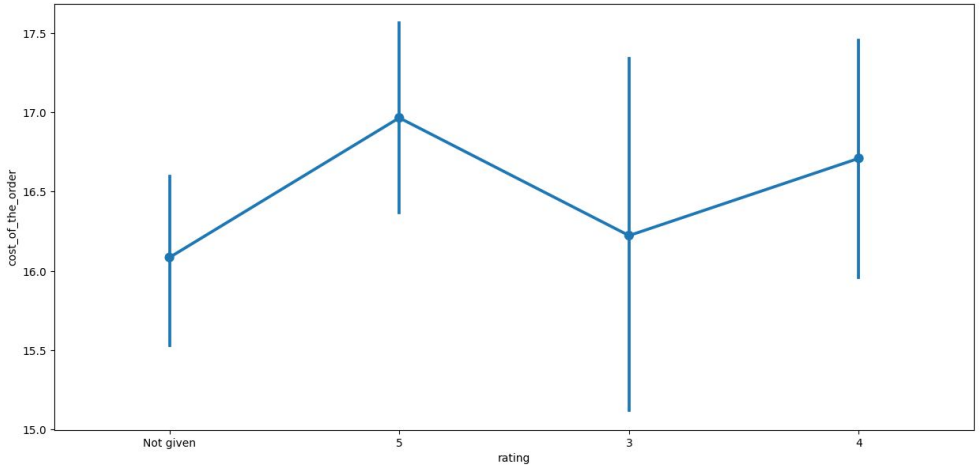
Multivariate Analysis

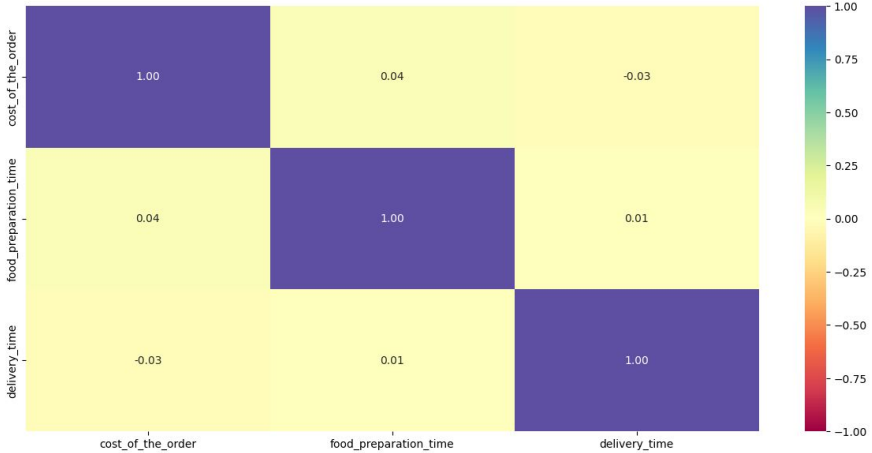




The plot on the left indicates that the delivery time and rating relationship showing that the lowest rating (3) has the most varying delivery times. Not rated and the highest rating (5) shows a major similarity. The graph on the right shows the food prep time and rating relationship, indicating that the largest variety in preparation time is the lowest rating (3) again, possibly indicating that rating and time inconsistencies are related.

Multivariate Analysis





The plot on the left indicates that the cost of the order and rating relationship showing that the lowest rating (3) has the most varying cost. Not rated and the highest rating (5) shows a major variance. The heat map on the right shows the cost of the order, food prep time, and delivery time relationships. It appears to indicate that there is a stronger relationship between prep time and cost perhaps indicating the more expensive meals costing more due to the time put into them.

Multivariate Analysis

- The restaurants with ratings higher than 4 and more than 50 reviews will be provided a promotional offer, the restaurants are **The Meatball Shop, Blue Ribbon Fried Chicken, Shake Shack, and Blue Ribbon Sushi.**
- With FoodHub charging 25% of orders over 20$ and 15% on orders over 5$ the net revenue is **$6166.30**.
- With an analysis of total time required to deliver food there are **10.5%** of orders that take more than 60 minutes to be delivered once an order is placed.
- The average delivery time on weekdays is **28 minutes**, where the average on weekends is **22 minutes.**

Thanks You

By: Joshua Willis
Contact: jkw910@gmail.com