FoodHub Data Analysis

Project Python Foundations: PGP in AI/ML

Saurabh Mittal

December 14, 2023

Contents / Agenda

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

Executive Summary (Conclusions)

- **Data contains ~1900 records with 1200 customers, 178 restaurants, 14 cuisines and there exist no missing values**
- Average food preparation time is around 14.5 mins, with maximum variability in Thai cuisine between 22-32 mins.
- About 39% of the orders are not rated that impact marketing plans
- Major cuisines ordered are American, Japanese, Italian and Chinese
- Average cost of the order is 14.5 dollars, with 30% of orders over 20 dollars, and majority within the range of 5 dollars to 33 dollars.
- Average delivery time is around 24 mins, with high order count on weekends and high variability (18 - 27 mins) on weekends as well. About 10 % of the orders are late by more than 60 mins. Mean delivery time on weekdays is 22 mins and on weekends is 28 mins.
- Orders are doubled on weekends
- There are 3 most frequent customers with at least 9 orders
- There are 3 restaurants with over 100 orders and 1 restaurant with over 200 orders
- The top 5 restaurants generate at least 1100 dollars revenue for themselves. The next 9 restaurants generate at least 500 dollars in revenue.
- Ratings are higher if the preparation time is lower. Counter intuitively, ratings are higher with higher costs. Delivery times, though expected to have some relationship with ratings, is not impacting ratings with current values
- **Net revenue with 178 restaurants for the Company is over 6000 dollars**

Executive Summary (Recommendations)

- Incentivize customers to provide ratings
- Incentivize restaurants with highest ratings to lower their preparation time on the weekends.
- Provide discounts on American, Japanese, Italian and Chinese cuisines
- Provide referral discounts to the most frequent customers
- Gather more data to resolve high delivery times for 10% of the orders, and understanding the impact of delivery times on the ratings. Intuitively, a lower delivery time should result in a higher rating.

Business Problem Overview

- A food aggregator company FoodHub offers access to multiple restaurants through a single smartphone app. The app allows the restaurants to receive a direct online order from a customer. The app assigns a delivery person from the company to pick up the order after it is confirmed by the restaurant. The delivery person then uses the map to reach the restaurant and waits for the food package. Once the food package is handed over to the delivery person, he/she confirms the pick-up in the app and travels to the customer's location to deliver the food. The delivery person confirms the drop-off in the app after delivering the food package to the customer. The customer can rate the order in the app. The food aggregator earns money by collecting a fixed margin of the delivery order from the restaurants.
- The food aggregator company has stored the data of the different orders made by the registered customers in their online portal. They want to analyze the data to get a fair idea about the demand of different restaurants which will help them in enhancing their customer experience.

Solution Approach / Methodology

- Employ Exploratory Data Analysis (EDA) to understand data features and correlations between various variables

- Utilize advanced Python constructs and libraries to slide/dice data to perform statistical assessments and visualizations to understand the relationship with both categorical and numerical variables.

- Utilize Python libraries such as numpy, pandas, matplotlib, and seaborn to perform the analysis.

- Provide conclusions and recommendations.

Data Overview

- Data contains ~1900 records with 9 columns
- No missing values

```
RangeIndex: 1898 entries, 0 to 1897
Data columns (total 9 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   order_id               1898 non-null   int64
 1   customer_id            1898 non-null   int64
 2   restaurant_name        1898 non-null   object
 3   cuisine_type           1898 non-null   object
 4   cost_of_the_order      1898 non-null   float64
 5   day_of_the_week        1898 non-null   object
 6   rating                 1898 non-null   object
 7   food_preparation_time  1898 non-null   int64
 8   delivery_time          1898 non-null   int64
dtypes: float64(1), int64(4), object(4)
```

Business Statistics

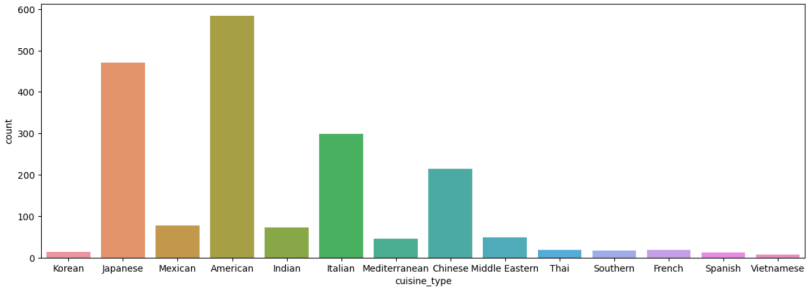
- Food preparation Time:
  - Min: 20 mins, Max: 35 mins, Avg: 27.37 mins
- 736 orders not rated
  - Ratings of 3, 4 and 5

Univariate Analysis

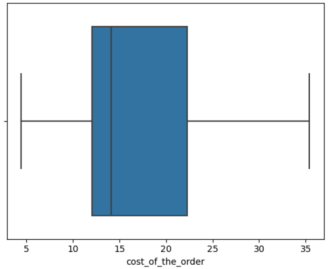
Customers: 1200

Restaurants: 178

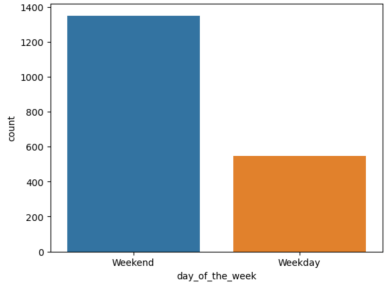
Cuisines: 14



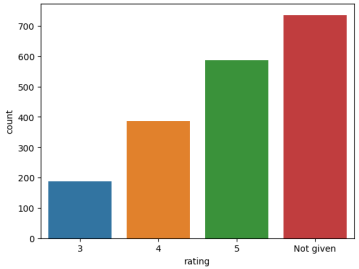
Major cuisines ordered:
American, Japanese, Italian and Chinese



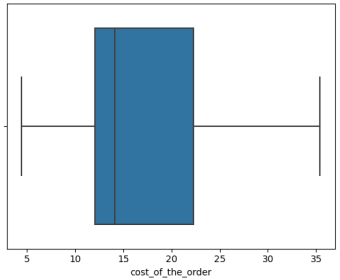
Avg order cost: $14.5,
50% orders between $12 - $22.5



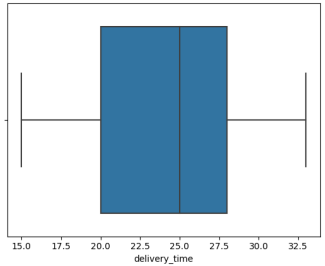
Orders on weekends are doubled of weekdays



About 39% orders not rated



Avg food preparation time around 14.5 mins and 50% orders between 12-23 mins



Avg. delivery time around 24 mins, with 50% between 20-28 mins

Univariate Analysis (Contd.)

Top 5 Restaurants

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

Weekend Cuisine Popularity: **American**

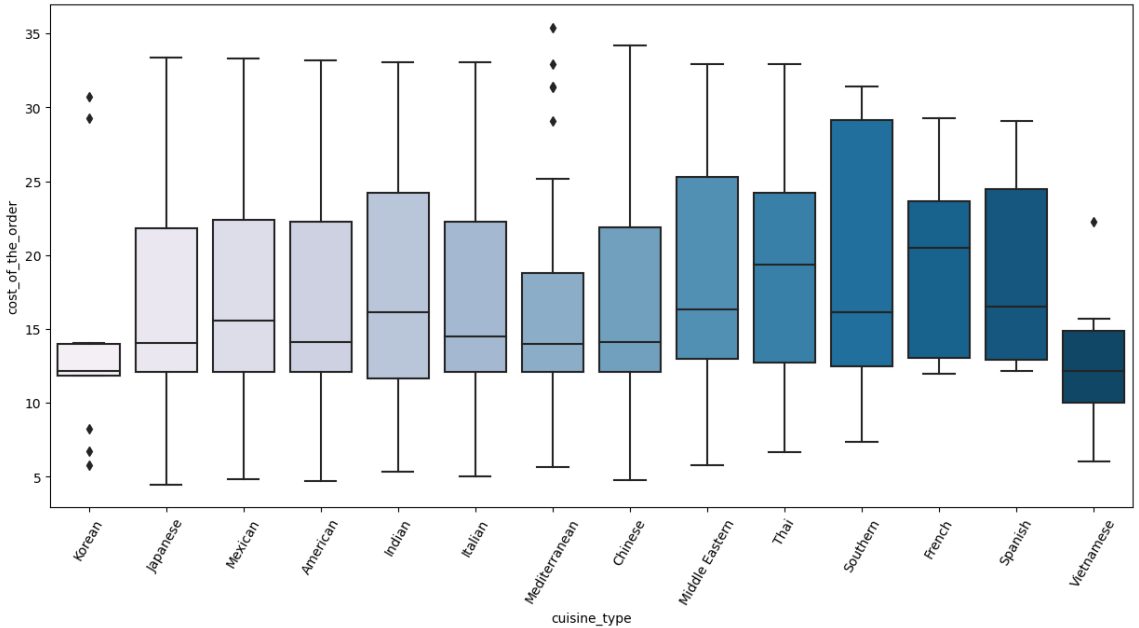
American	415
Japanese	335
Italian	207
Chinese	163
Mexican	53
Indian	49
Mediterranean	32
Middle Eastern	32
Thai	15
French	13
Korean	11
Southern	11
Spanish	11
Vietnamese	4

Percentage of Orders with cost over $20	29.24%
Mean Order Delivery Time	24.15 mins
Most Frequent Customers	3 with 13,10,9 orders respectively

Multivariate Analysis

Cuisine vs Order Cost

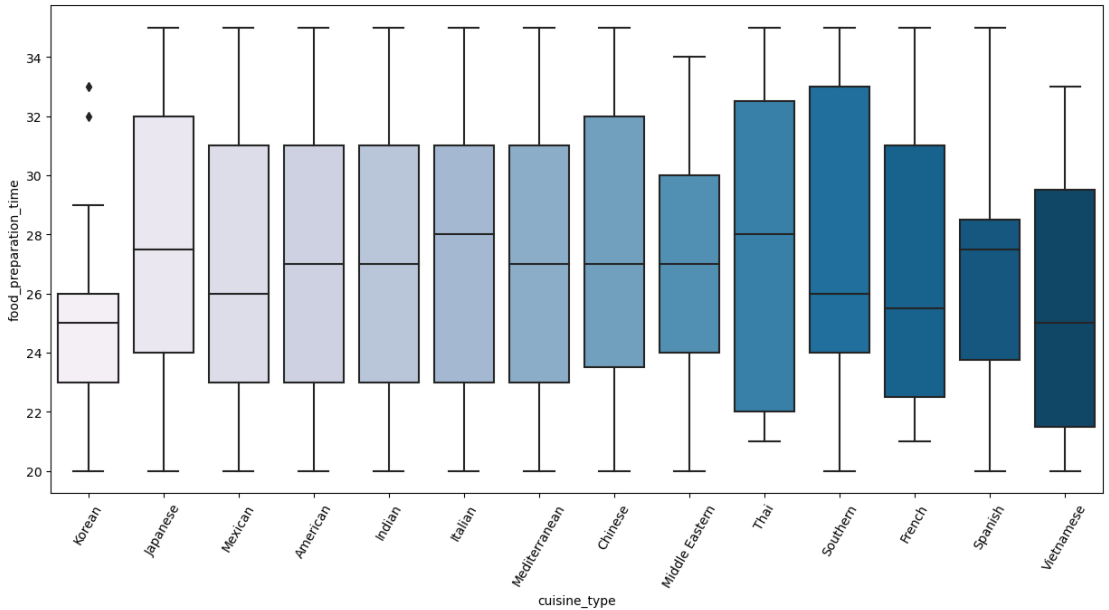
- Vietnamese is the cheapest
- Southern has the largest range in 50% of its offerings
- Some of the Mediterranean orders are the most expensive
- Majority of the cuisines in the range of $5 - $33



Multivariate Analysis (contd.)

Cuisine vs Food Preparation Time

- Max. variability across cuisines: **10 mins**
    - About 50% of Thai orders take between 22-32 mins
- Min. variability across cuisines: **3 mins**
    - About 50% of the Korean orders take 23-26 mins

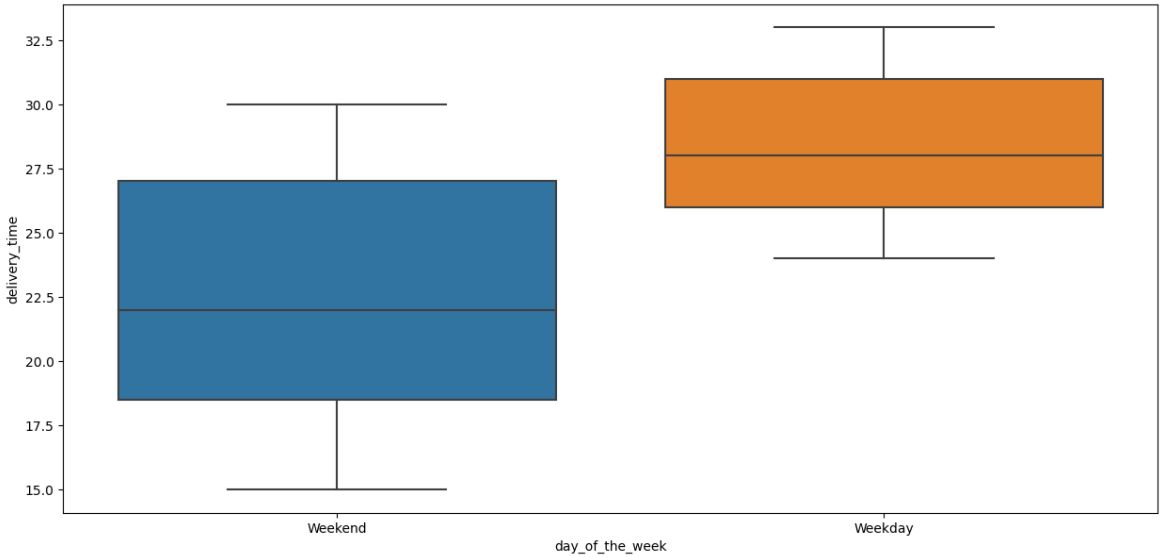


Multivariate Analysis (contd.)

Day of the Week vs Delivery Time

- Weekday delivery time higher than Weekend, perhaps due to office traffic
- Variability on weekends: 18 – 27 mins
- Variability on weekdays: 26 – 32 mins

- **More orders on weekends contribute to more data and more variability**

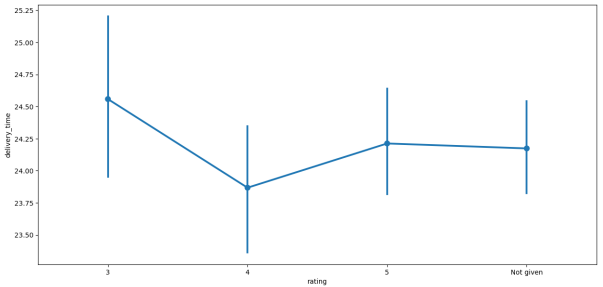


Multivariate Analysis (contd.)

- Revenue generated by Restaurants

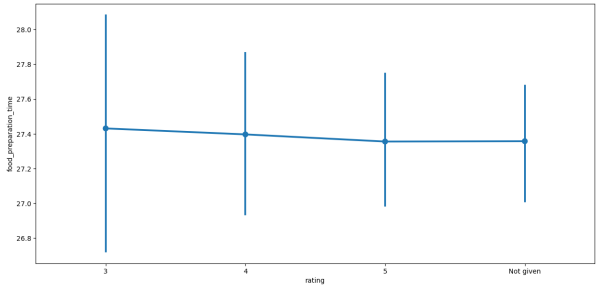
Shake Shack	3579.53
The Meatball Shop	2145.21
Blue Ribbon Sushi	1903.95
Blue Ribbon Fried Chicken	1662.29
Parm	1112.76
RedFarm Broadway	965.13
RedFarm Hudson	921.21
TAO	834.50
Han Dynasty	755.29
Blue Ribbon Sushi Bar & Grill	666.62
Rubirosa	660.45
Sushi of Gari 46	640.87
Nobu Next Door	623.67
Five Guys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

Multivariate Analysis (contd.)



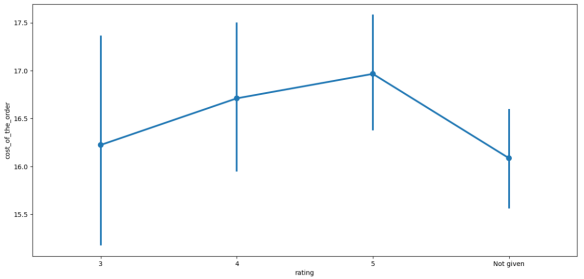
Rating vs Delivery Time
- non-linear relationship

Lower delivery times are in general related to better ratings. More data required for a conclusive trend.



Rating vs Food Preparation Time
- linear relationship

Preparation time is negatively correlated.



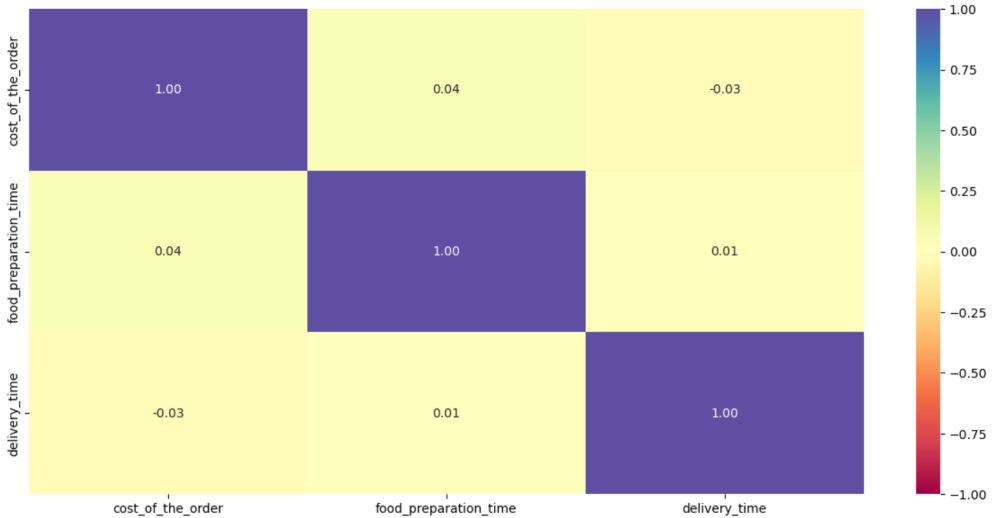
Rating vs Order Cost
- linear relationship.

Higher cost is positively correlated

Multivariate Analysis (contd.)

- No correlation between Order cost, food preparation time and delivery time

- Correlation does exist for the three variables with ratings (previous slide)



Multivariate Analysis (contd.)

- **Question 13: The company wants to provide a promotional offer in the advertisement of the restaurants. 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.**

121 restaurants with avg. ratings > 4

	restaurant_name	rating
0	Shake Shack	60
1	The Meatball Shop	53
2	Blue Ribbon Sushi	32
3	Blue Ribbon Fried Chicken	32
4	RedFarm Broadway	18
...	...	...
116	Haru Gramercy Park	1
117	Galli Restaurant	1
118	Five Leaves	1
119	El Parador Cafe	1
120	indikitch	1

121 rows × 2 columns

2 restaurants with cumulative ratings > 50

	restaurant_name	rating
0	Shake Shack	60
1	The Meatball Shop	53

Multivariate Analysis (contd.)

- **Question 14: The company charges the restaurant 25% on the orders having cost greater than 20 dollars and 15% on the orders having cost greater than 5 dollars. Find the net revenue generated by the company across all orders.**
  - The net revenue is around 6166.3 dollars

- **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? (The food has to be prepared and then delivered.**
  - Number of orders late by 60 mins: 200
  - Percentage of orders late by 60 mins: 10.54 %

- **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
  - The mean delivery time on weekends is around 22 minutes

Great Learning

Happy Learning !