

# FoodHub Project

## PGP AIML - Ashley Campbell

Due 11/18/23

# Contents / Agenda

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

# Executive Summary

1. *The company should focus on restaurants that have a **high number of total ratings**, as these restaurants get the highest number of orders and create the most revenue - more so than the average rating and average order cost*
2. ***Offer discounts on weekday service in return for rating an order** to both increase the number of restaurant ratings and increase the number of weekday orders.*
3. *Top 5 revenue generating restaurants are: **Shake Shack, The Meatball Shop, Blue Ribbon Sushi, Blue Ribbon Fried Chicken, and Parm***
4. *Most popular cuisine types include: **American, Japanese, Italian, Chinese, Mexican and Indian***
5. *Vietnamese, Thai, Korean, Southern, French and Spanish restaurants have generated the least revenue*

# Business Problem Overview and Solution Approach

- **Objective**

- Understand the demand for different restaurants and perform statistical analysis to find other related factors that will enhance customer experience
- 

- **Methodology**

- Extract insights using Exploratory Data Analysis.
- Perform univariate analysis to examine each variable individually, and multivariate analysis to determine relationships between the variables
- Determine which food types are most popular
- Determine which restaurants have the most orders and generate the most income
- Determine when most orders are placed
- Determine how ratings relate to restaurant demand

# Data Overview

Variable	Description
Customer ID	This is the customer's identifier
Restaurant Name	Name of the restaurant where order was placed
Cuisine Type	Type of food ordered (American, Chinese, etc)
Cost of Order	The price of each order
Day of the week	Order placed on weekday or weekend
Rating	Customer's rating of the meal from 1 being the lowest and 5 being the highest
Food Preparation Time	Time required to prepare each meal
Delivery Time	Time required to deliver each meal

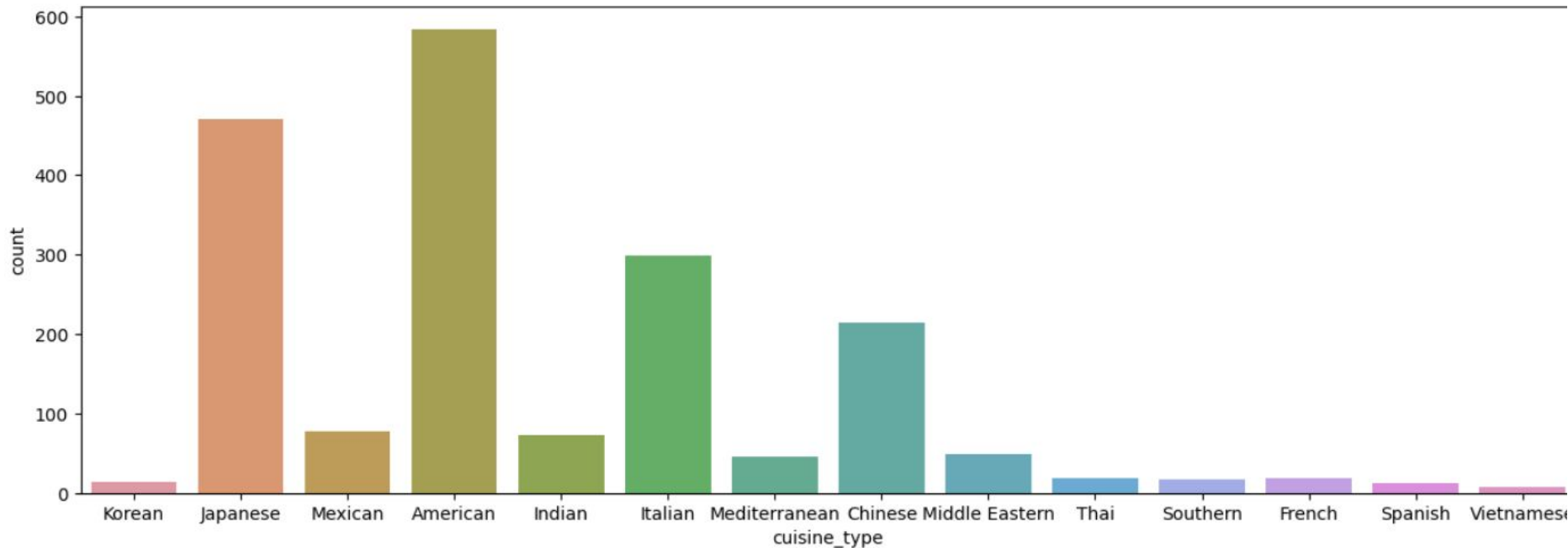
Observations	Variables
1898	9

## Notes

- Variable **data types** include **integers (4)**, **objects (4)** and **floats (1)**
- There are **no missing values** in the data set
- **Ratings were not given for 736 of 1898 orders placed**
- **Min food prep time was 20 minutes, max was 35 minutes, with an avg of 27.4 minutes per order**

\*\*\*Questions 1-5

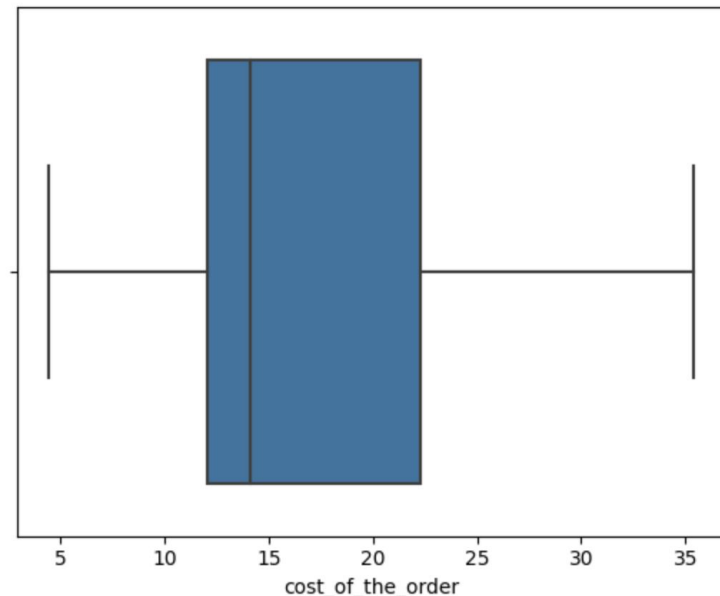
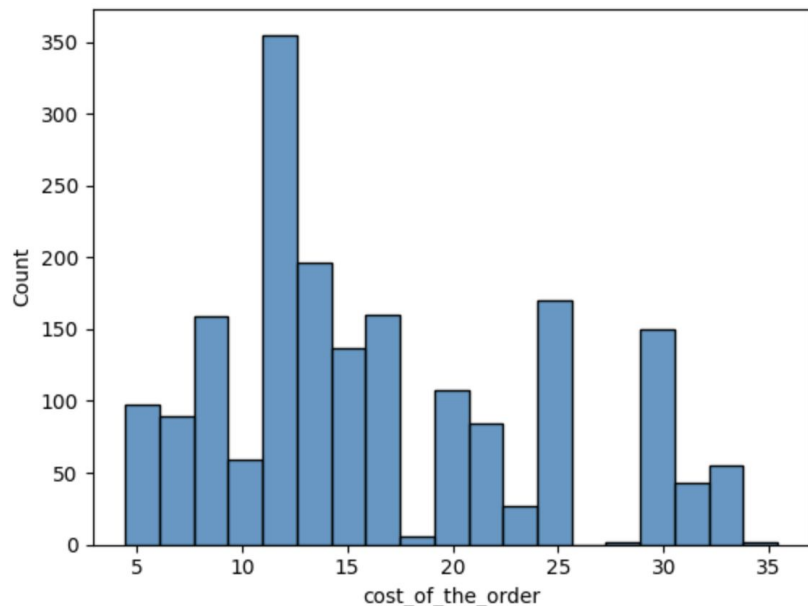
# Univariate Analysis - Cuisine Types



- 1898 orders were placed by 1200 different customers at 178 different restaurants
- 14 different types of food included in the data set
- Top cuisine types represented include American, Japanese, Italian, Chinese

\*\*\*Question 6

# Univariate Analysis - Cost of the Order

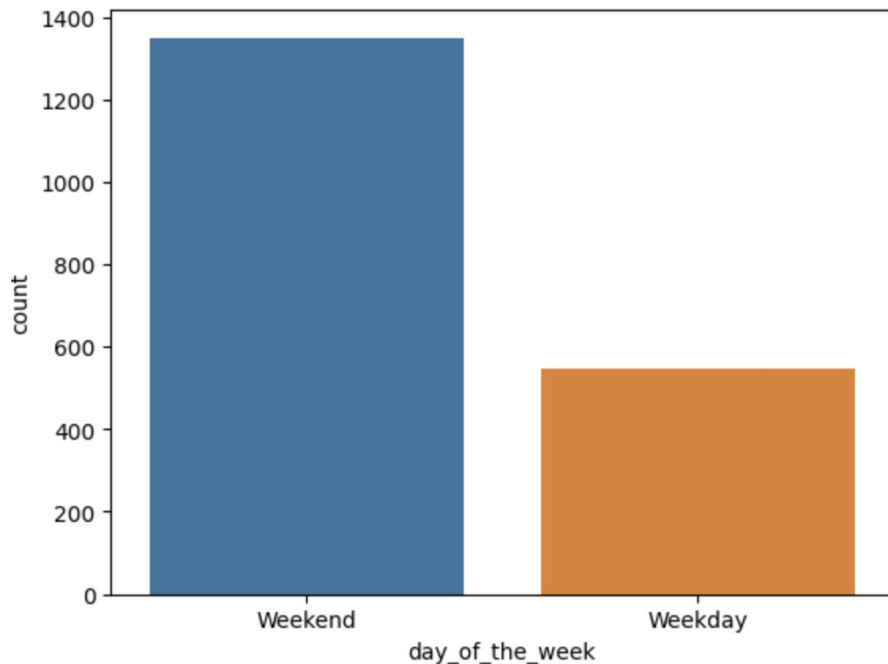


- Most orders range between \$10-17
- Right skewed distribution shows there are more orders at lower costs than higher cost
- Small or sale items are likely most popular

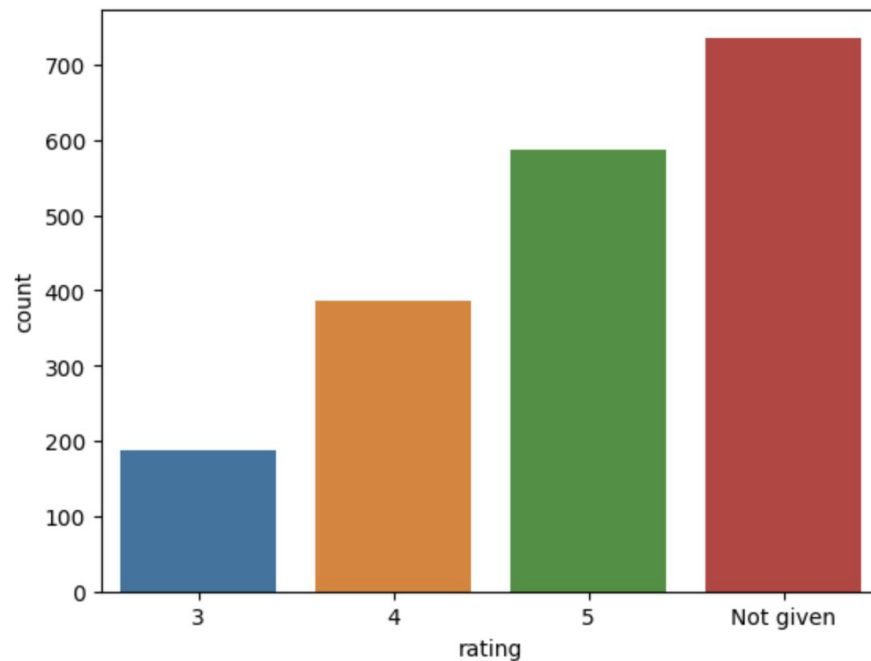
- The median order cost is \$14
- 50% of orders are between \$12 and \$22.50

\*\*\*Question 6

# Univariate Analysis - Day of the Week & Ratings



- More than twice as many orders placed on the weekend versus on weekdays.

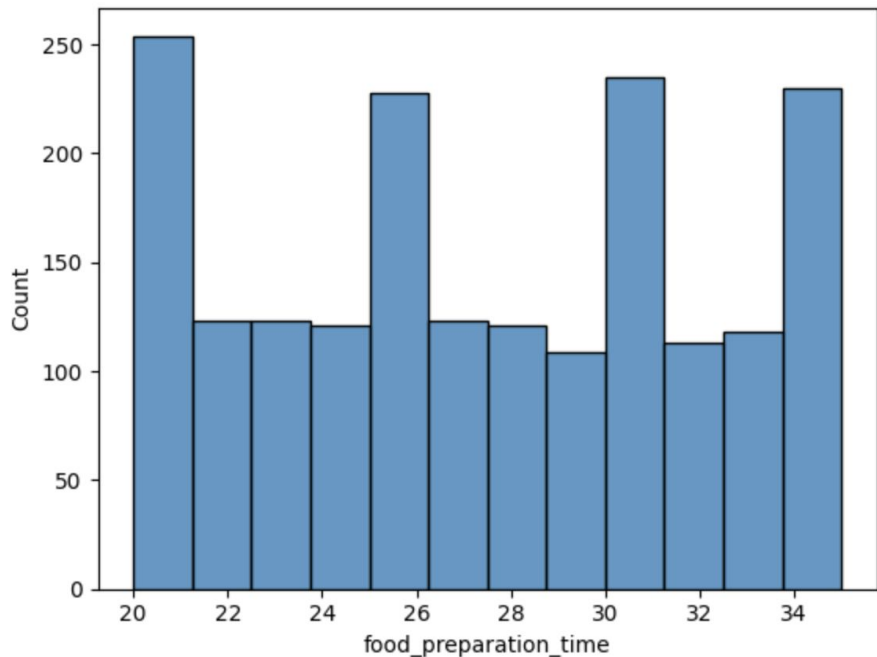


- Ratings not given for almost 40% of total orders

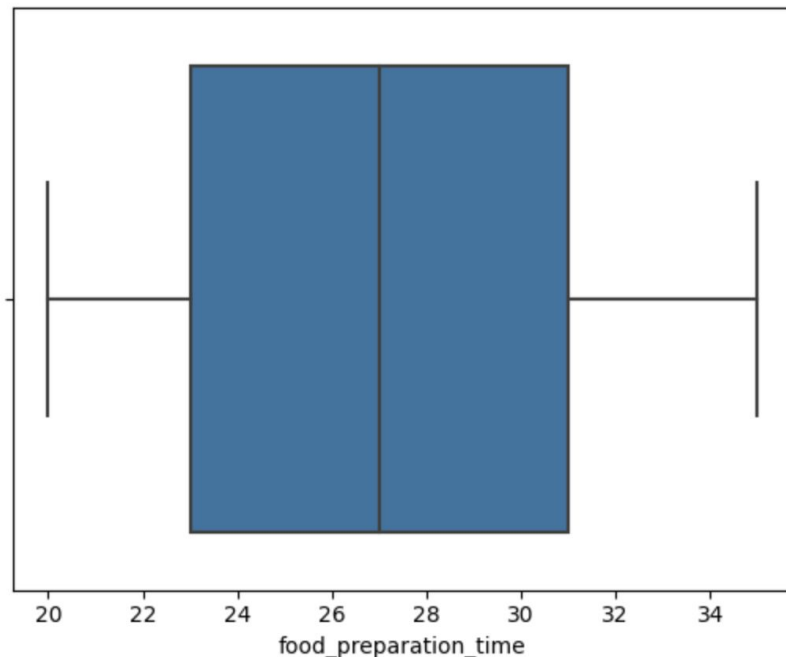
\*\*\*Question 6



# Univariate Analysis - Food Preparation Time



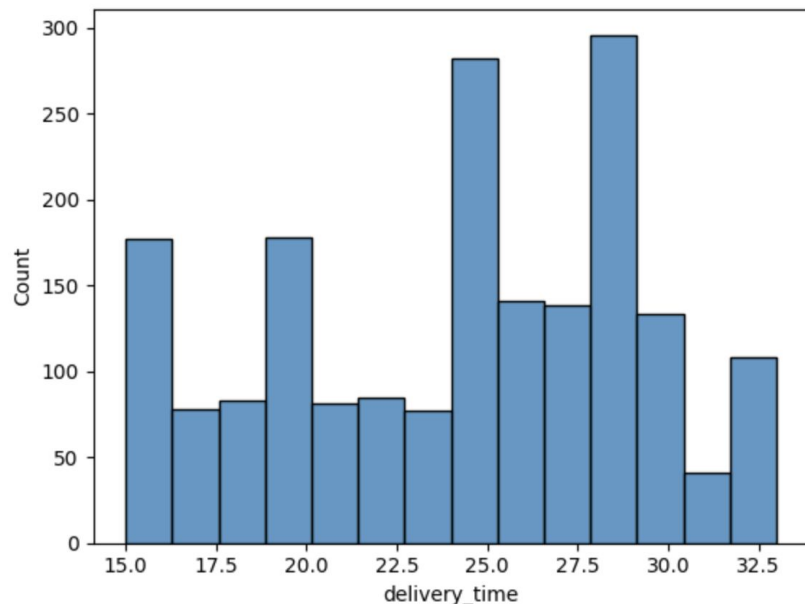
- Uniform distribution, ranging from 20-35 minutes



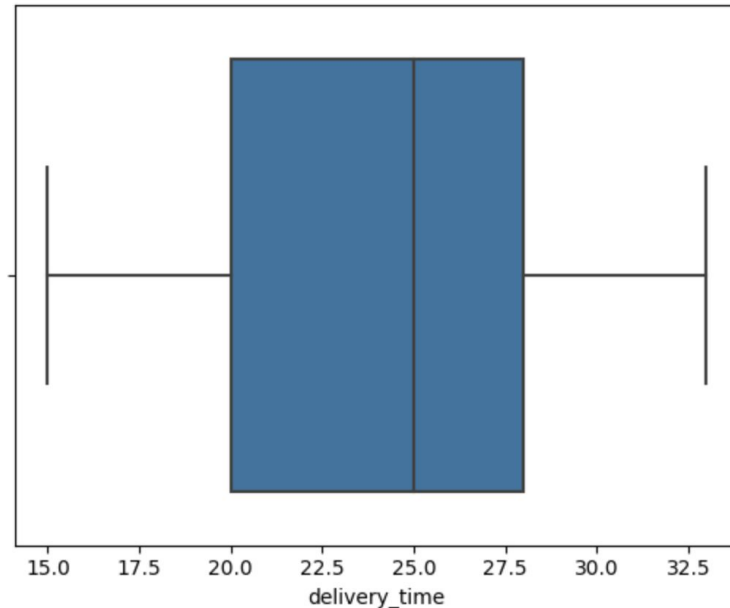
- Median food preparation time is approximately 27 minutes
- 50% of orders take between 23 and 31 minutes

\*\*\*Question 6

# Univariate Analysis - Food Delivery Time



- Delivery time ranges from 15-33 minutes
- Left skewed - more orders with shorter delivery time than longer delivery time



- Median delivery time is 25 minutes
- 50% of deliveries fall between 20 and 28 minutes
- Left skewed - median is closer to upper quartile than lower quartile

# Univariate Analysis - Most Popular Options

## Top Weekend Choices

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

## Top 5 Restaurants by Total Order

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

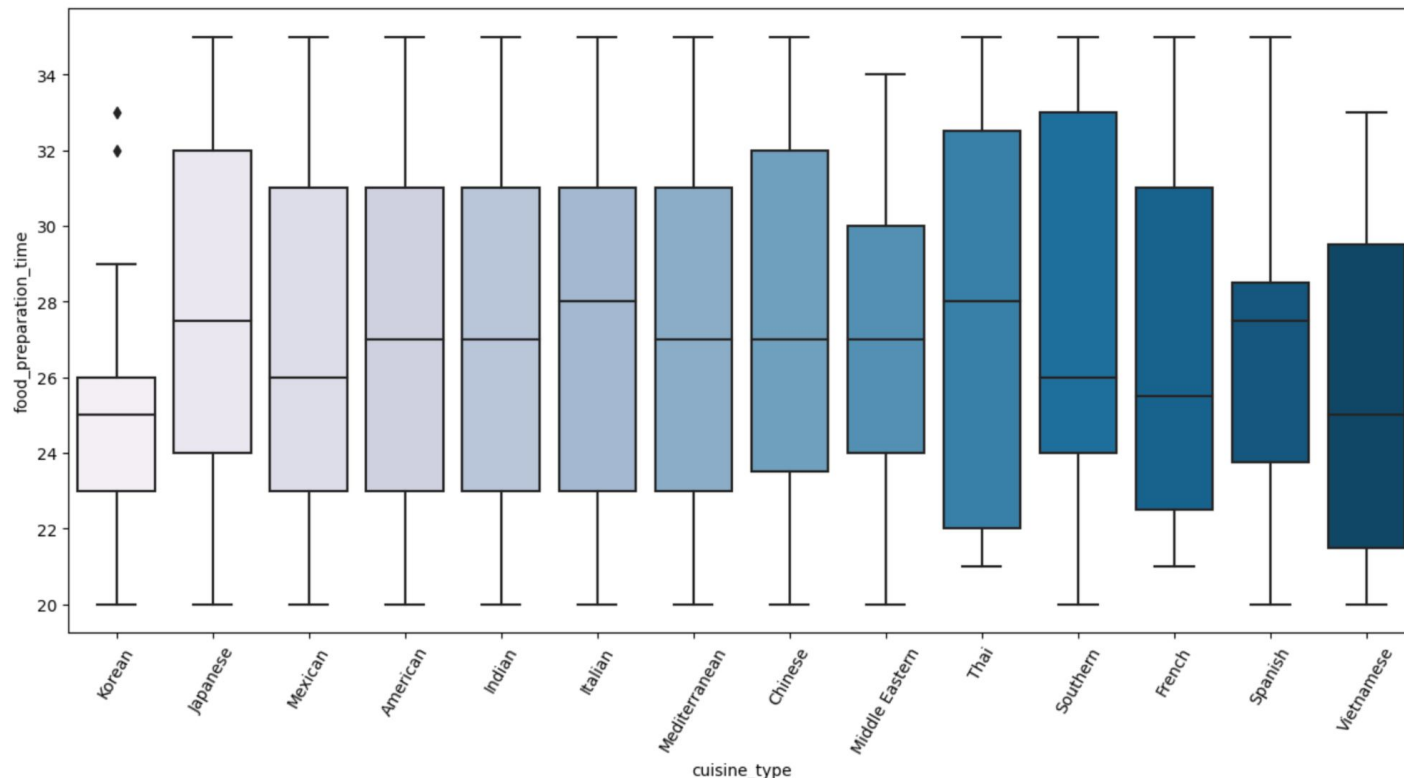
## Frequent Customer Voucher Winners

Customer ID	Orders
52832	13
47440	10
83287	9

- The mean delivery time for this data set is 24.16 minutes
- Mean delivery time on the weekend is 22 minutes vs 28 minutes during the week

\*\*\*Questions 7-11

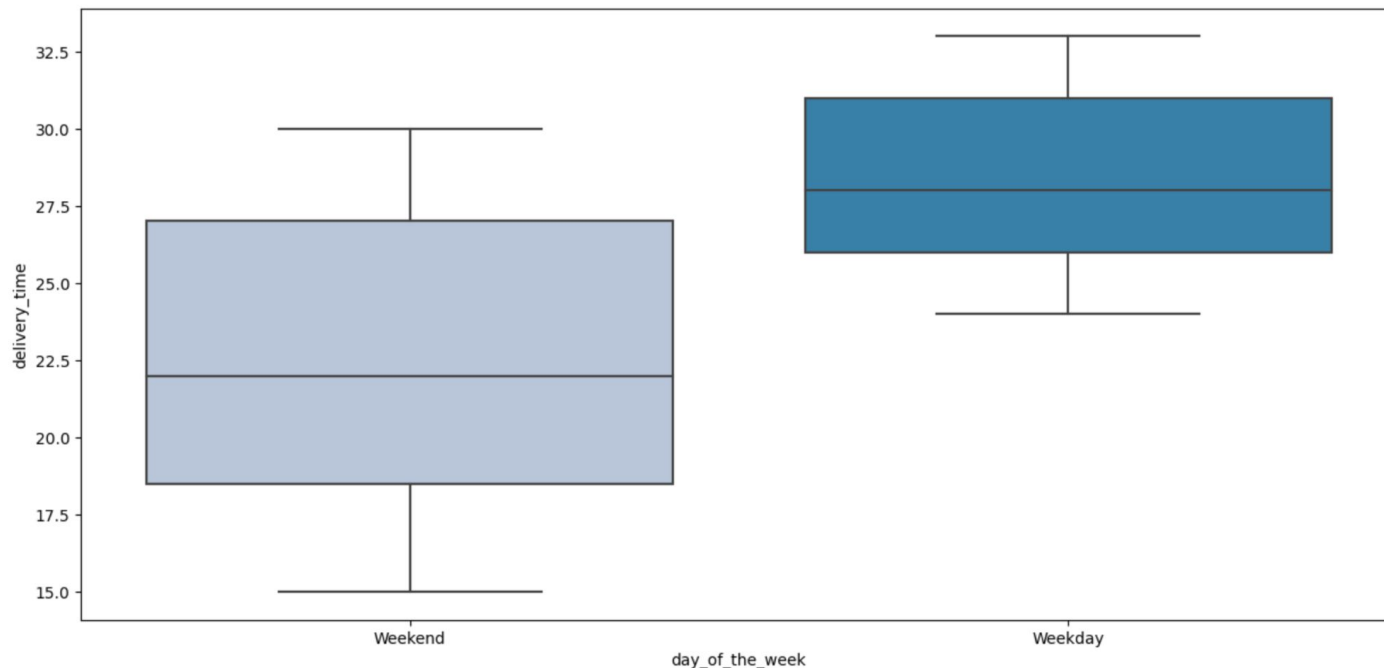
# Multivariate Analysis - Food Prep Time & Cuisine Type



- All food types have a minimum prep time of at least 20 min and max of 35 min
- IQR smallest for Korean food & largest for Thai food
- IQR for most food types < 10 with most common range from 23-31 minutes

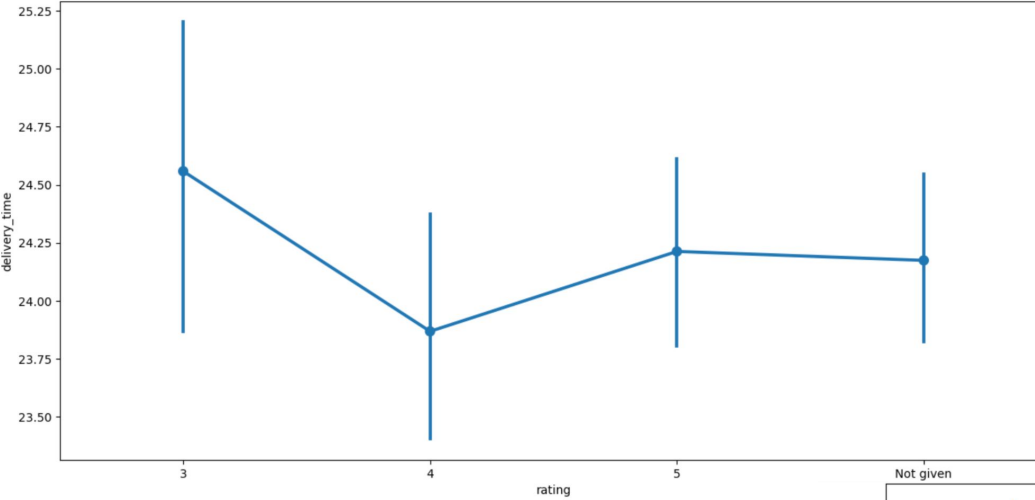
\*\*\*Question 12

# Multivariate Analysis - Day of the Week & Delivery Time



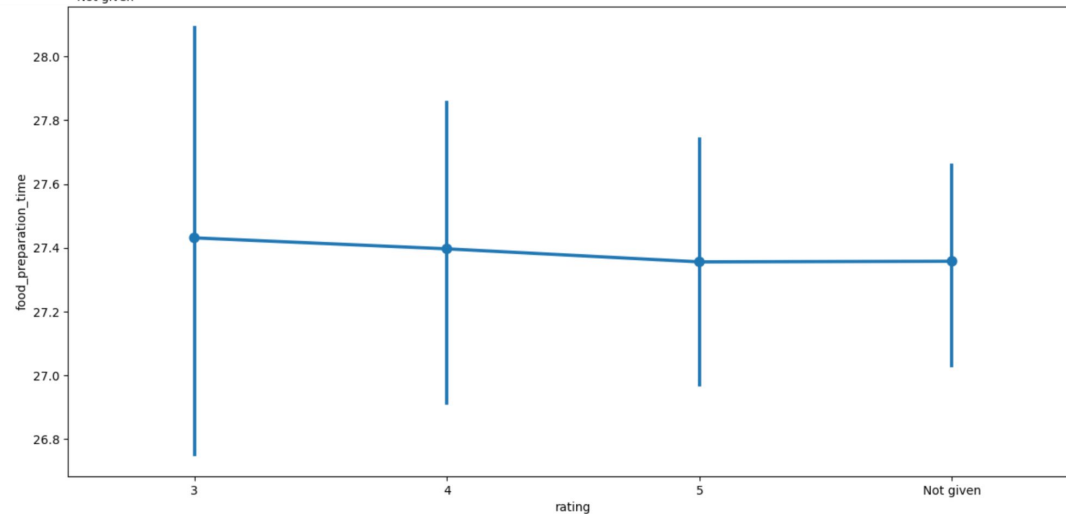
- Delivery time is **significantly faster on the weekend**, with a median delivery time of about 22 minutes (range of 15-30 minutes), vs about 28 minutes during weekdays (range of 26 - 33 minutes)

# Multivariate Analysis - Rating & Delivery / Food Prep Time

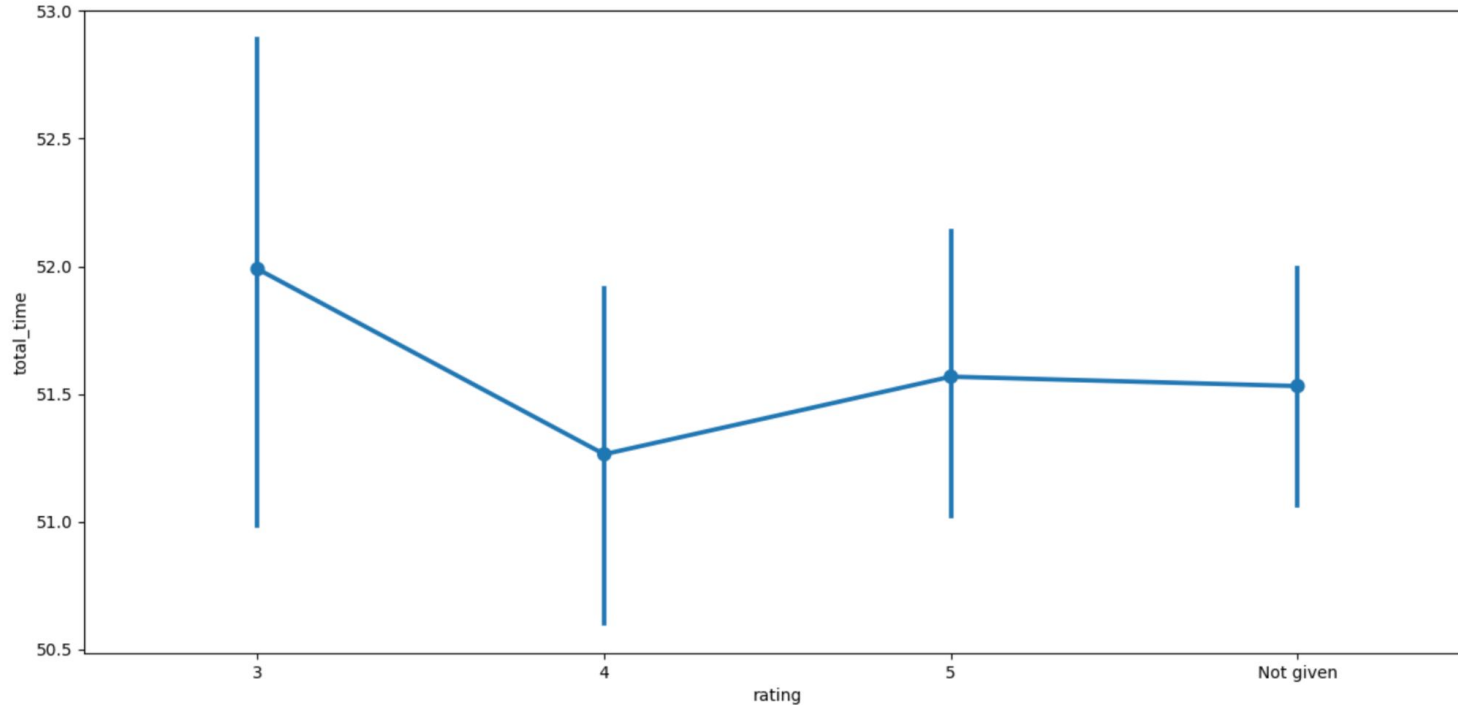


- Slightly negative correlation between delivery time and rating - faster deliveries get a higher rating

- No correlation between food prep time and rating



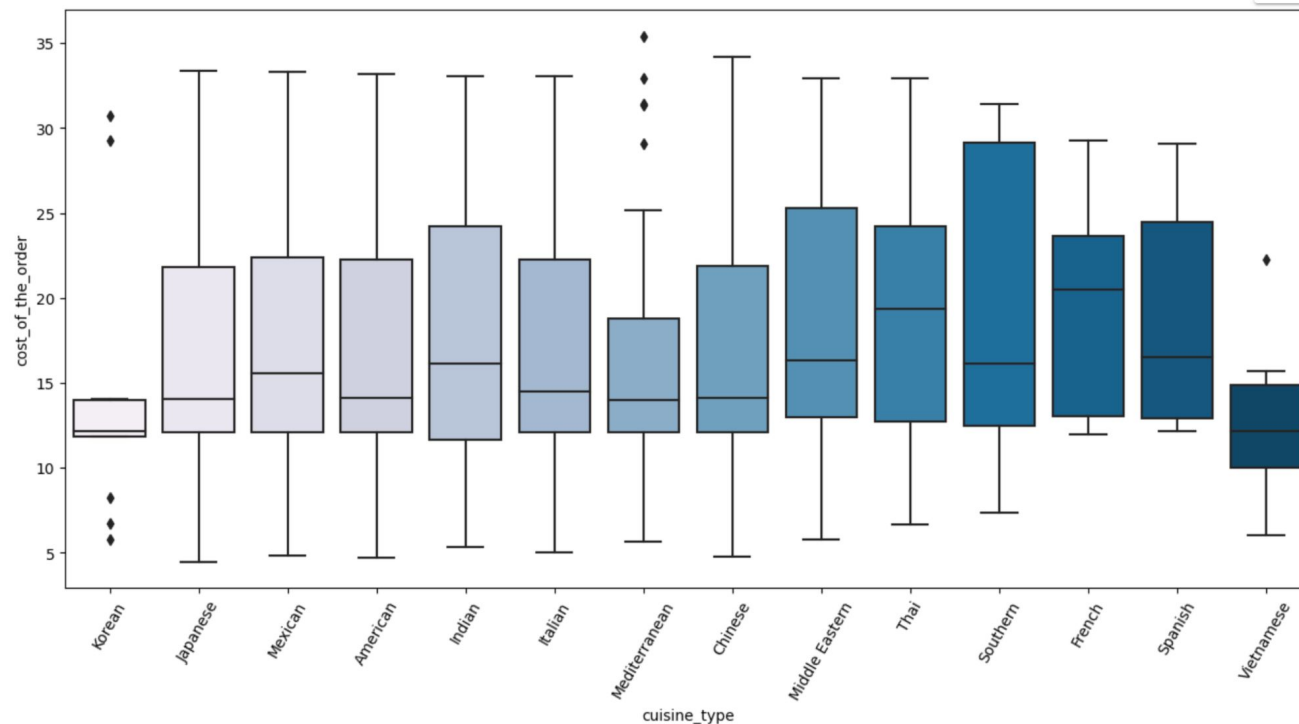
# Multivariate Analysis - Total Time & Rating



- Of 1898 orders, 200 or **10.54% of orders took > 60 minutes total**
- Very weakly negative correlation between total order time and order rating

\*\*\*Question 15

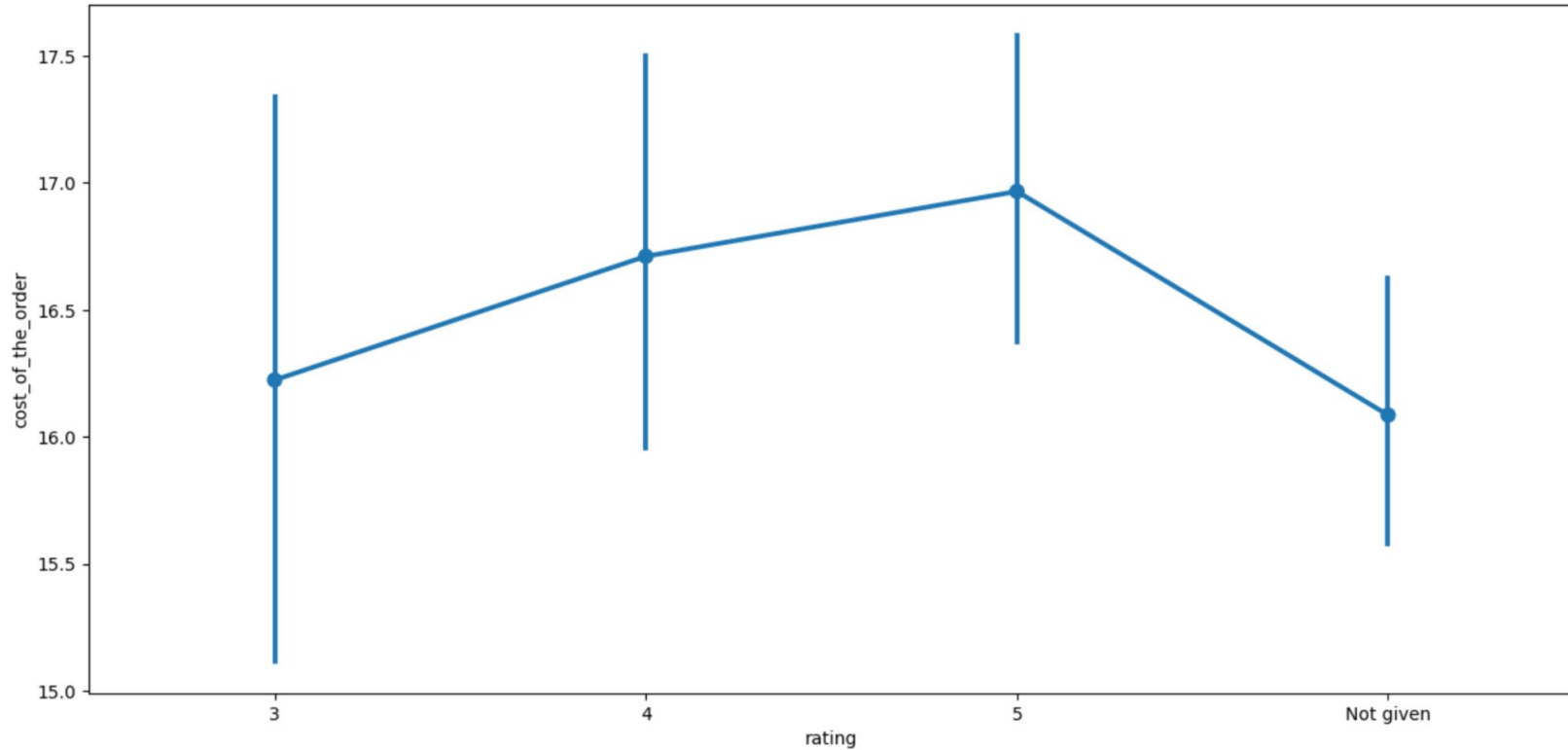
# Multivariate Analysis - Cuisine Type and Cost



- The median cost of orders is highest for French cuisine and lowest for Vietnamese cuisine
- The median cost of orders for all types of cuisine is between about \$12-\$20
- The IQR is largest for Spanish cuisine and smallest for Korean cuisine
- The middle 50% of all orders cost between \$10 and \$30

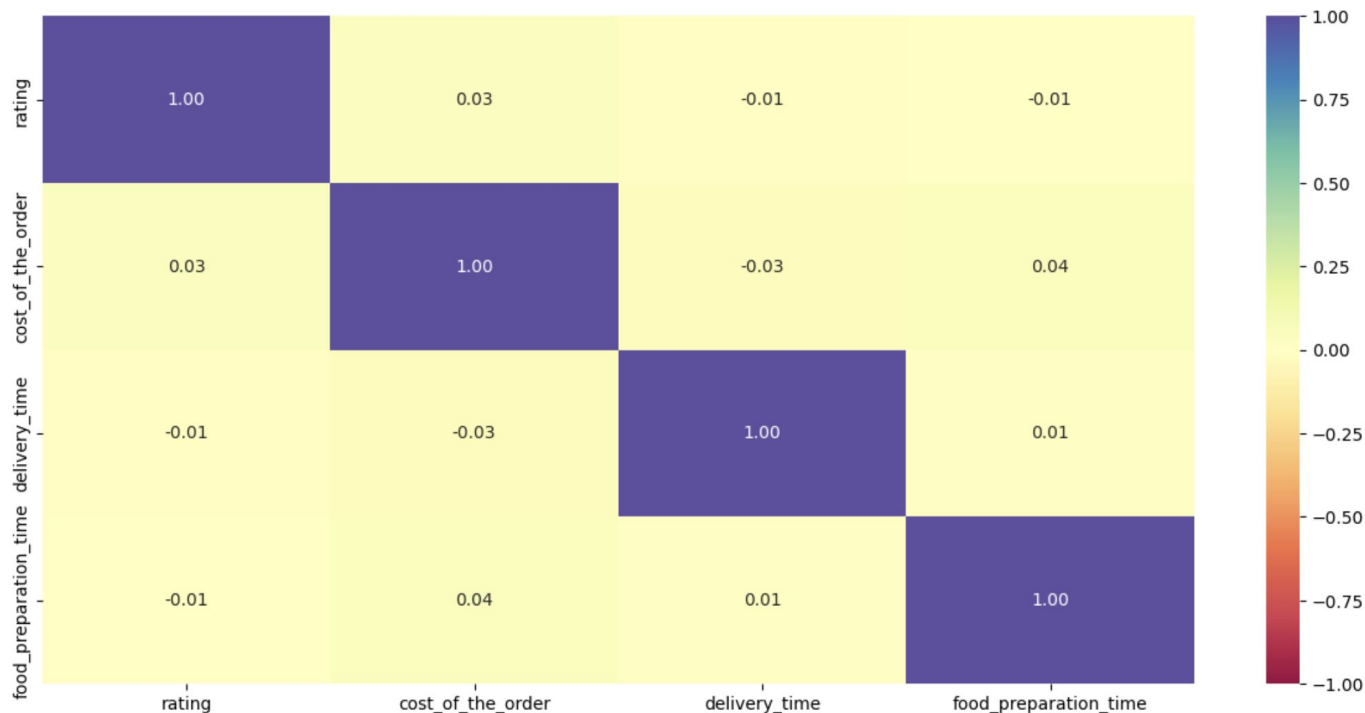


# Multivariate Analysis - Rating & Cost of Order



- Weak positive correlation between rating and cost of order
- Expensive orders have higher ratings, but within a range of \$1 (\$16-\$17)

# Correlation - Cost, Food Prep, Delivery Time & Rating



- No significant correlations between cost of the order, delivery time, food prep time and/or rating

# Multivariate Analysis - Revenue Generation

Shake Shack	3579.53
The Meatball Shop	2145.21
Blue Ribbon Sushi	1903.05
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 Bari 46	640.87
Nobu Next Door	623.67
Five Guys Burgers and Fries	506.47

## Total Revenue By Restaurant

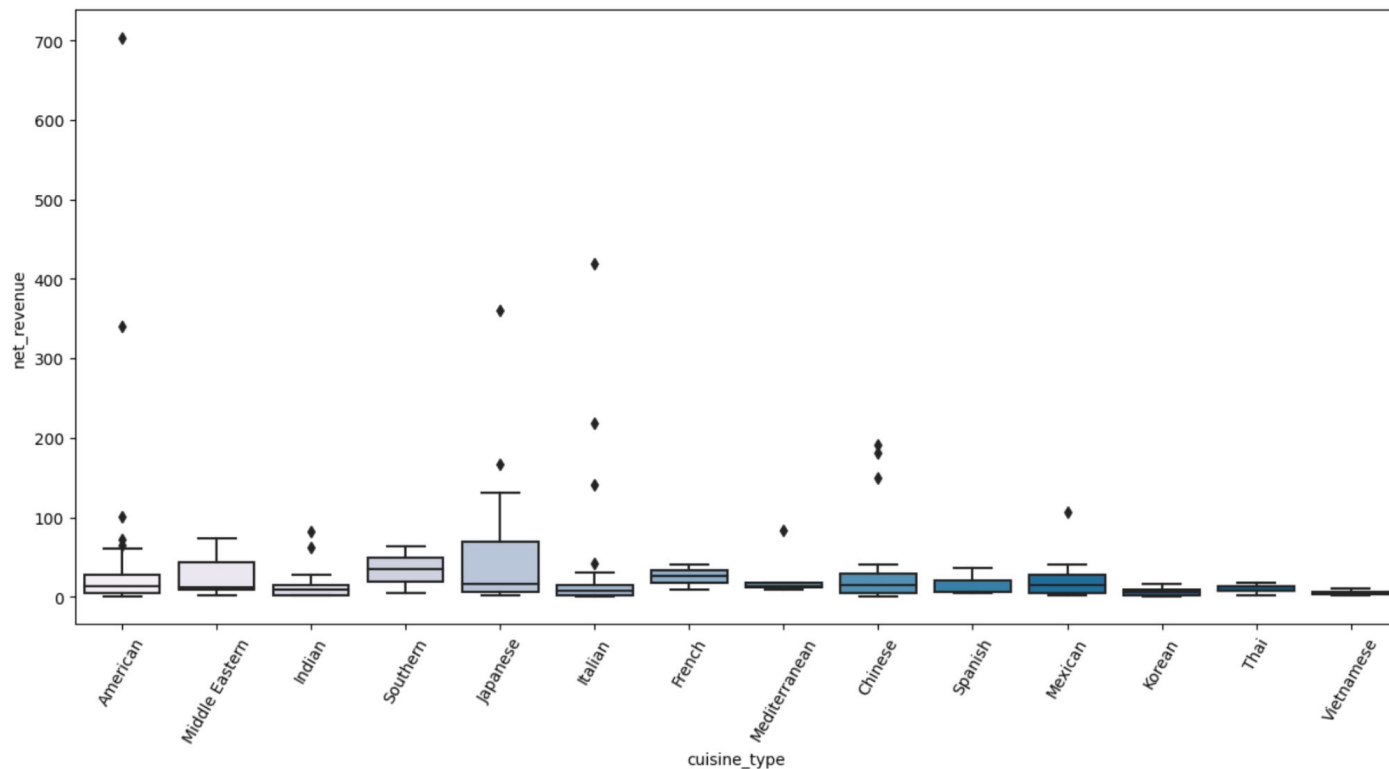
- Minor differences in top total revenue earners vs top total revenue generators for FoodHub

Shake Shack	703.61
The Meatball Shop	419.83
Blue Ribbon Sushi	360.46
Blue Ribbon Fried Chicken	340.20
Parm	218.56
RedFarm Broadway	191.47
RedFarm Hudson	180.93
TAO	167.36
Han Dynasty	149.40
Rubirosa	140.81
Sushi of Bari 46	130.50
Nobu Next Door	115.83
Blue Ribbon Sushi Bar & Grill	114.95
Chipotle Mexican Grill	106.66

## Net Revenue Earned by FoodHub

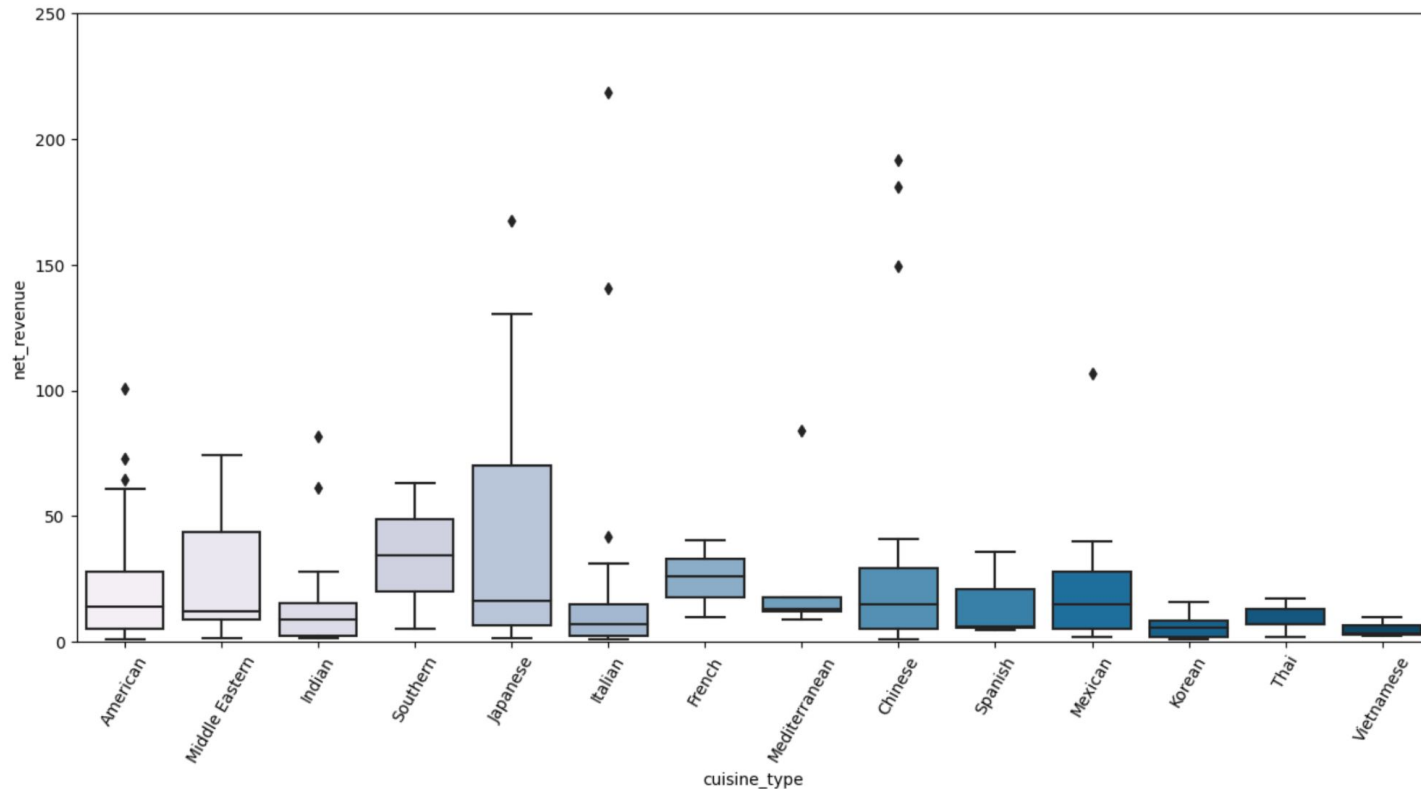
- Net earnings of 25% on orders over \$20 and 15% on orders over \$5

# Multivariate Analysis - Net Revenue & Cuisine Type



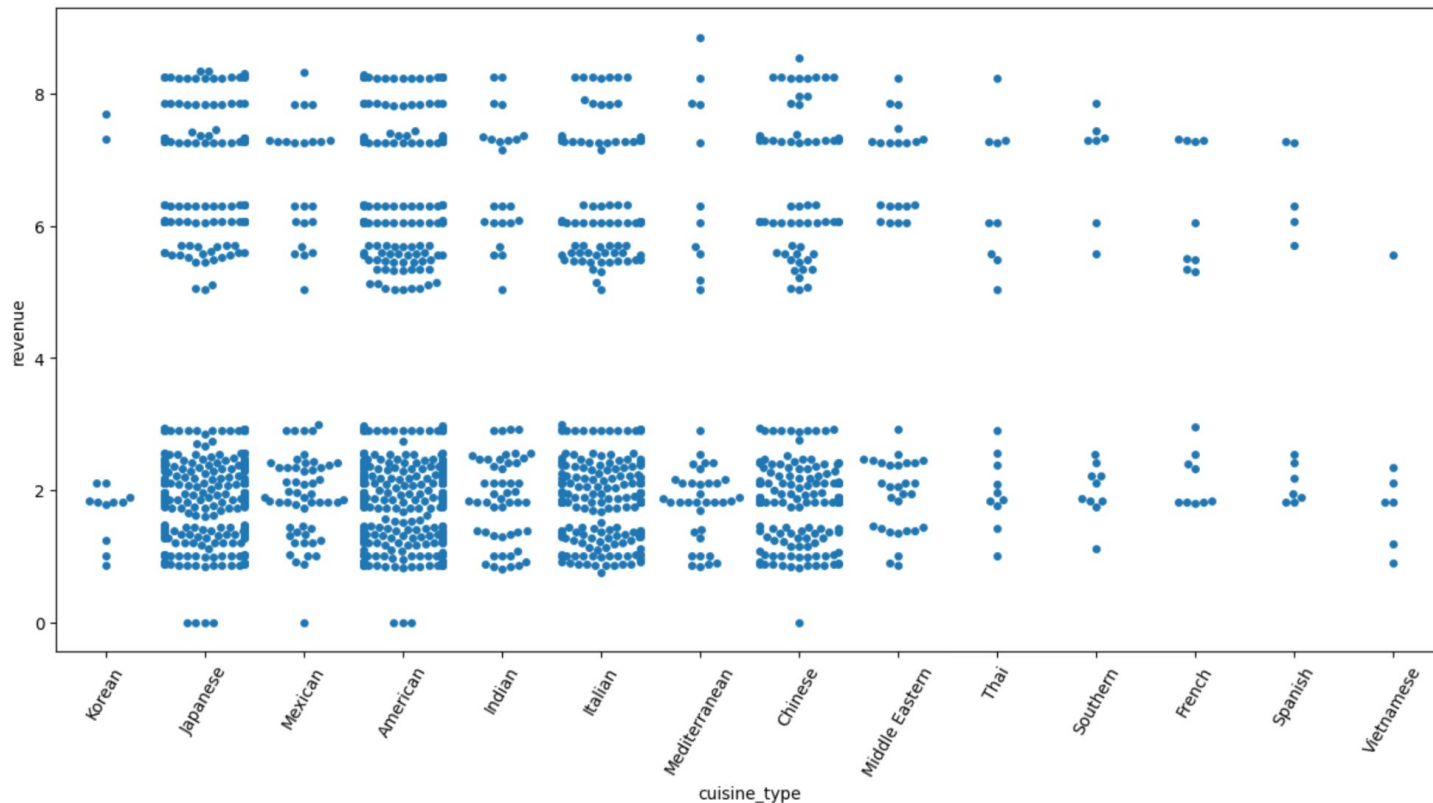
- Significant outliers for American, Italian and Japanese cuisine
- A small number of restaurants appear to be bringing in most of the money

# Closer Look - Net Revenue & Cuisine Type



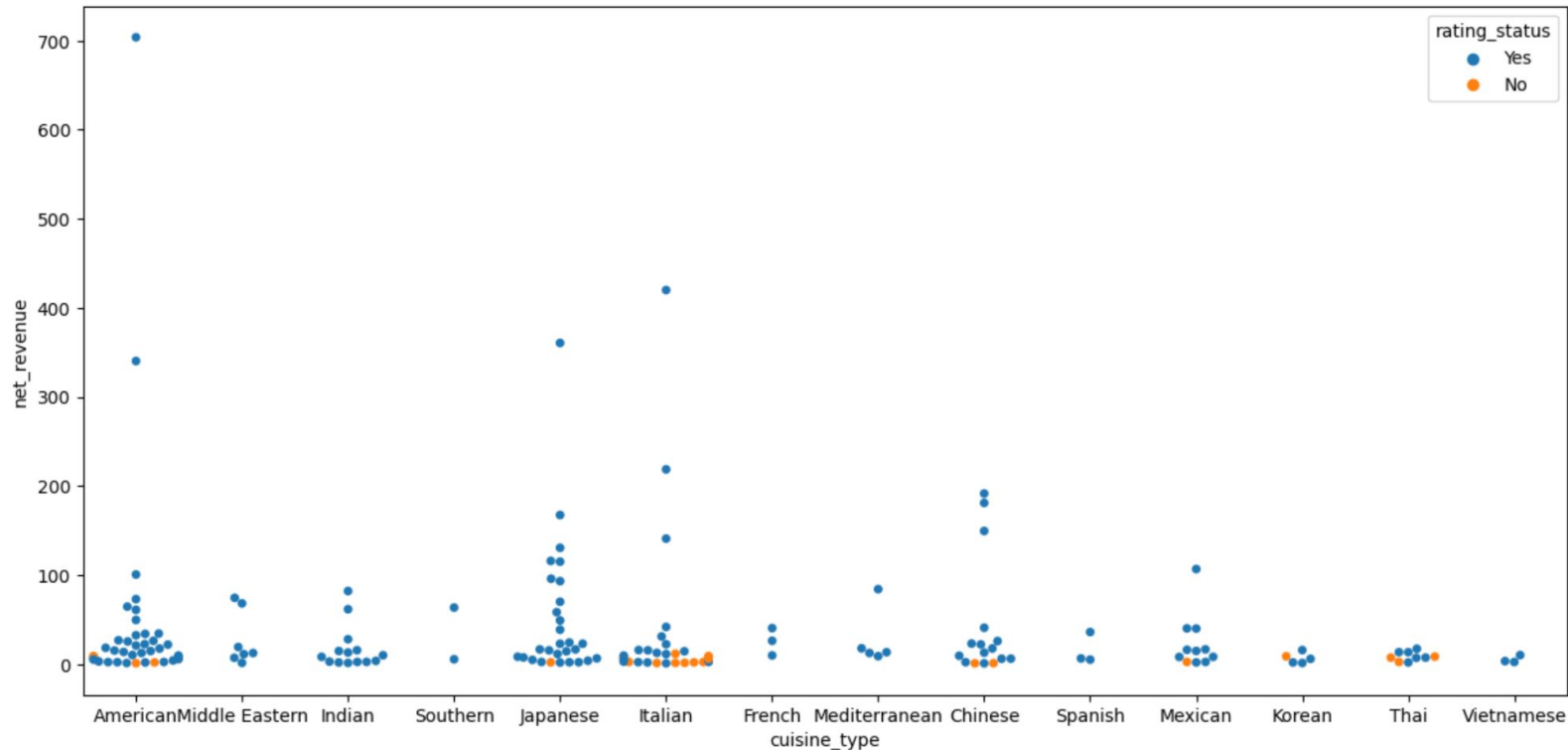
- Japanese restaurants show the greatest variability in net revenue per restaurant
- Korean, Thai and Vietnamese bring in the least amount of revenue per restaurant

# Multivariate Analysis - Revenue Per Order & Cuisine Type



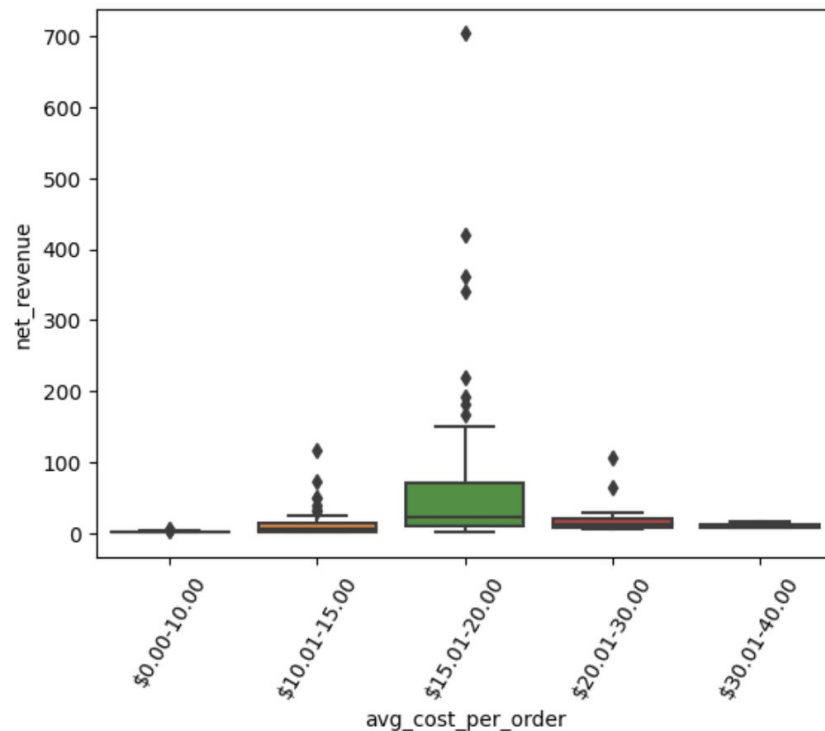
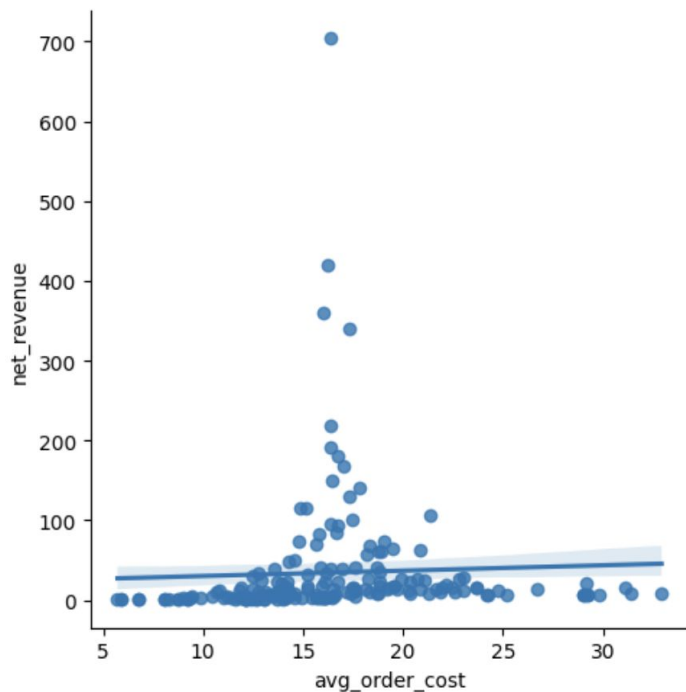
- Japanese, American, Italian and Chinese cuisines bring in the most money
- Thai, Southern, French, Spanish and Vietnamese cuisines bringing in significantly less revenue

# Multivariate Analysis -Net Revenue, Cuisine Type & Rating Status



- Net revenue from each food type is near zero for restaurants that do not have at least one rating, regardless of cuisine type

# Multivariate Analysis - Net Revenue vs Average Order Cost



- Of 1898 orders:
  - **only 555 or 29% of orders were > \$20,**
  - 890 or 47% of orders were > \$15
  - 1003 or 53% of orders were < \$15
- There does not appear to be a strong correlation between average order cost and total revenue per restaurant, most revenue coming from orders in the \$15-20 range



# Multivariate Analysis - Rating & Revenue

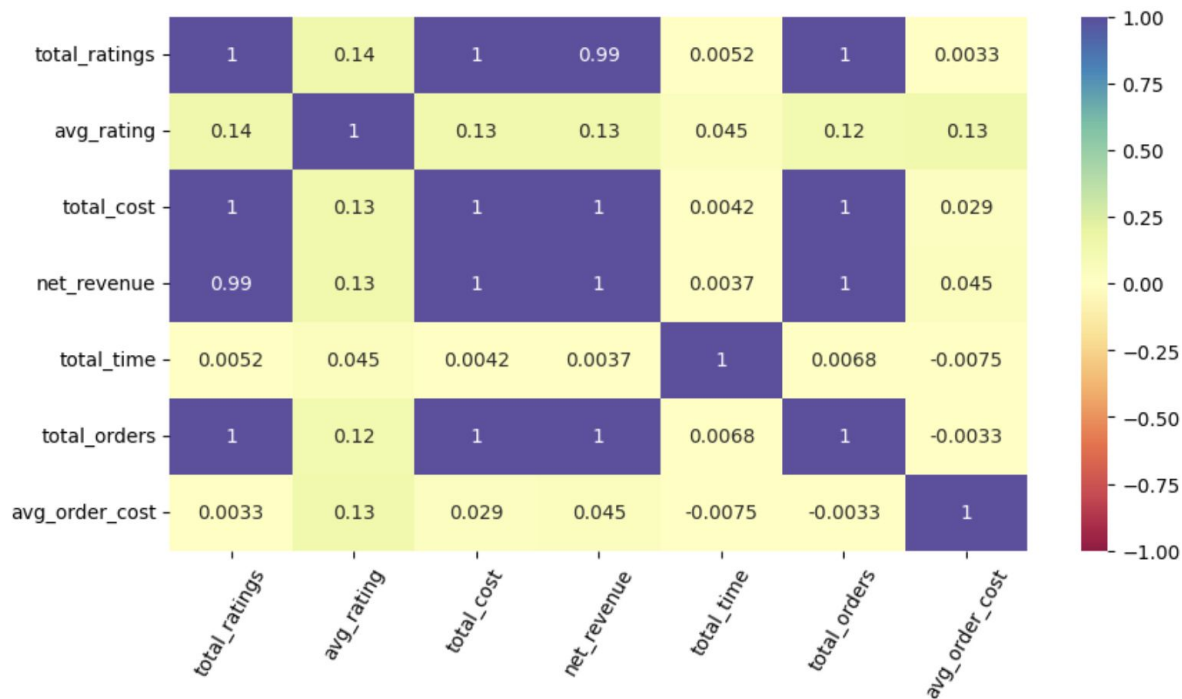
## Advertising Promotional Offer

Restaurant	Rating
The Meatball Shop	4.511905
Blue Ribbon Fried Chicken	4.328125
Shake Shack	4.278195
Blue Ribbon Sushi	4.219178

- **The total net revenue is around \$6166.3 dollars.**
  - **Approximately 12% of restaurants have no ratings on the app.**
  - **Net revenue from restaurants that do not have a rating is \$100.49, or just 1.6% of total revenue.**
- 
- **Restaurants with > 50 ratings and an average rating > 4.2**

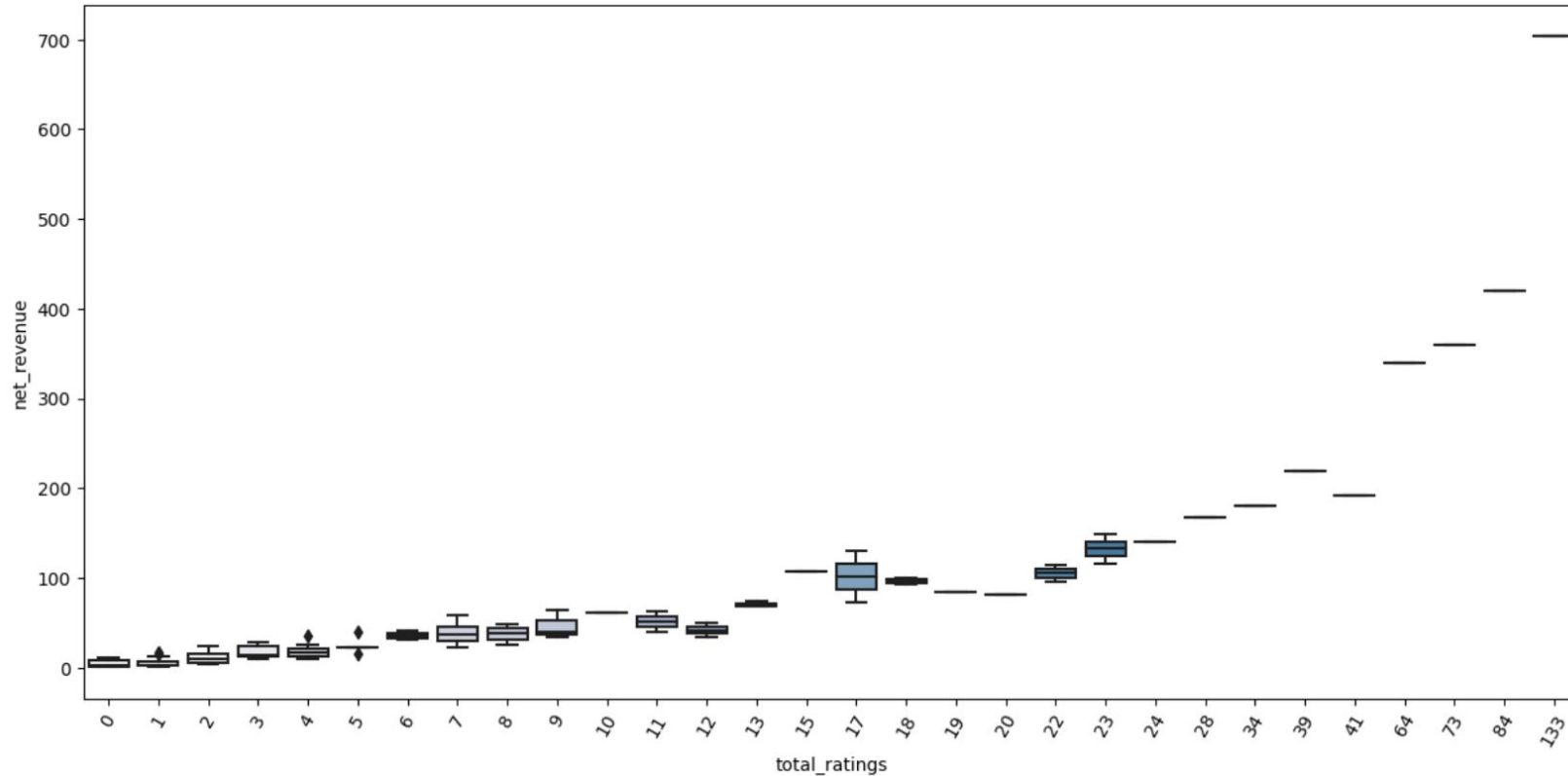
\*\*\*Question 13, 14

# Correlations: Ratings, Cost, Revenue and Time



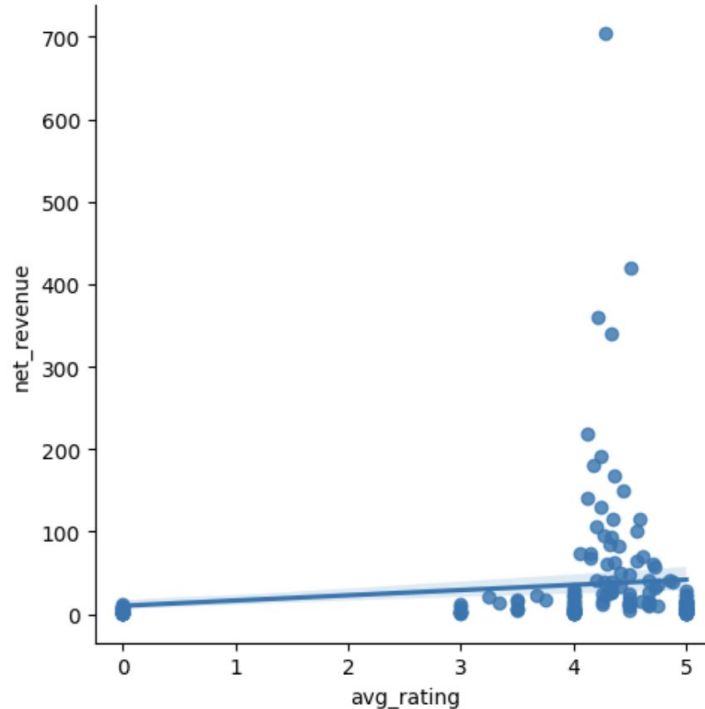
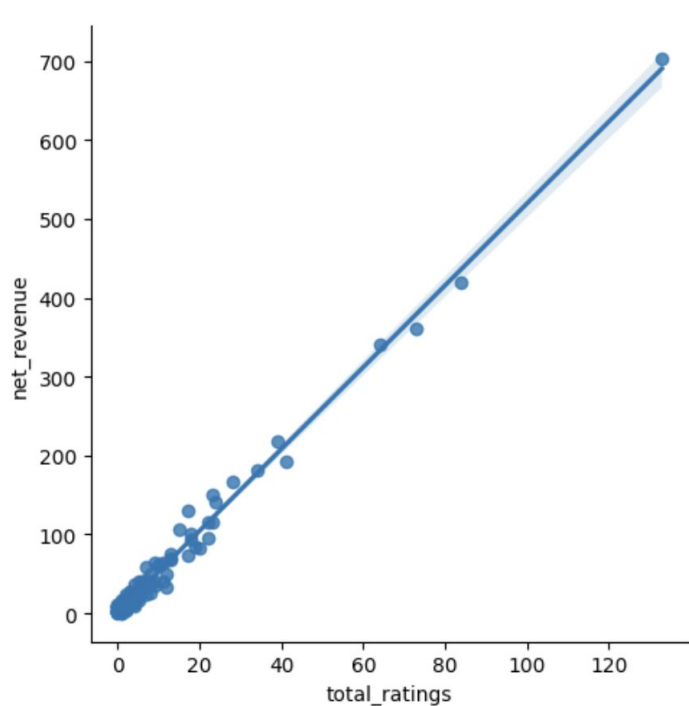
- Extremely high correlation between total ratings, total cost, total orders, and total revenue
- Very weak correlation between average rating, total revenue, total cost, total ratings and total orders
- No correlation between total time and any other variable

# Multivariate Analysis - Net Revenue & Total Ratings



- Clear positive relationship between net revenue a restaurant brings in and the total number of ratings a restaurant has received

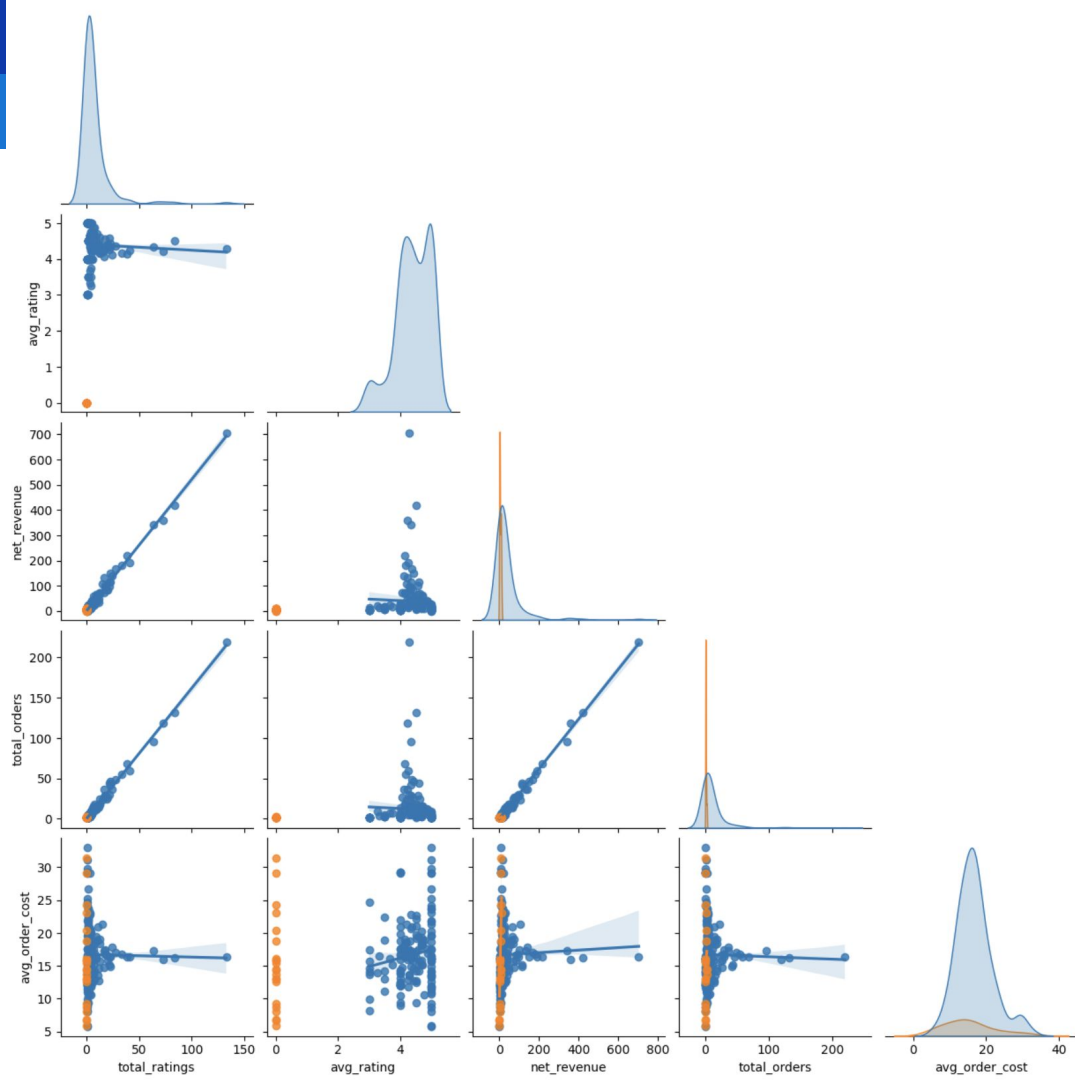
# Multivariate Analysis - Ratings and Revenue



- Very strong positive correlation between total ratings and net revenue
- Weakly positive correlation between average rating and net revenue

# Multivariate Analysis

- Strong positive correlation between net revenue, total orders and total ratings
- Very weak negative correlation between total ratings and average rating
- No significant correlation between average rating and any other factors



distribution prohibited.



**Happy Learning !**

