



| | |
|--------------------------------|---|
| Timeframe | 1 week |
| Submission Deadline | Thursday 12th October 11:59pm |
| Submission Link | Submit a link to your solutions here: https://forms.gle/YzD6Zv6GL7u7i8bU7 |
| Additional Instructions | Use any software/tools that you are comfortable with. |

Context

The number of restaurants is increasing day by day. Lots of students and busy professionals rely on those restaurants due to their hectic lifestyles. Online food delivery service is a great option for them. It provides them with good food from their favorite restaurants. A food aggregator company FoodHub offers access to multiple restaurants through a single smartphone app.

The app allows 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 on the delivery order from the restaurants.

Objective

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. Suppose you are hired as a Data Scientist in this company and the Data Science team has shared some of the key questions that need to be answered. Perform the data analysis to find answers to these questions that will help the company improve its business.

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 ordered the food
- restaurant_name: Name of the restaurant
- cuisine_type: Cuisine ordered by the customer
- cost: 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) taken by the restaurant 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: Time (in minutes) taken by the delivery person 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

Questions:

1. What are your general thoughts on the dataset? How would you describe the data? (1 mark)
2. Find the summary statistics and write your observations based on that. (3 marks)
3. What are your thoughts on the categorical variables? (2 marks)
4. How many orders are not rated? Please show how you arrived at this. (1 mark)

5. Explore all the variables and provide observations on the distributions of all the relevant variables in the dataset. (5 marks)
6. Which are the top 5 restaurants that have received the highest number of orders? (1 mark)
7. Which is the most popular cuisine on weekends. Explain how you arrived at this. (1 mark)
8. How many orders had a total cost of above 20 dollars? What is the percentage of such orders in the dataset? (1 mark)
9. Find the mean delivery time based on this dataset. (1 mark)
10. Suppose the company has decided to give a free coupon of 15 dollars to the customer who has spent the maximum amount on a single order. Find the ID of the customer along with the order details. (2 marks)
11. Perform bivariate analysis to explore relationships between the important variables in the dataset. (8 marks)
12. Suppose 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. Help us identify the restaurants fulfilling the criteria to get the promotional offer. (2 marks)
13. Suppose the company charges the restaurant 25% on the orders having a cost greater than 20 dollars and 15% on the orders having a cost greater than 5 dollars. What is the net revenue generated on all the orders given in the dataset? (2 marks)
14. Suppose the company wants to analyze the total time required to deliver the food. Help us identify the percentage of orders that have more than 60 minutes of total delivery time. (2 marks)
15. Suppose the company wants to analyze the delivery time of the orders on weekdays and weekends. Help us find the mean delivery time on weekdays and weekends. Analyze the results. (2 marks)

16. What other data explorations techniques have you not captured in the above questions? These are additional methods to help us understand our data better.(3 Marks)
17. Write the conclusions and business recommendations derived from the analysis.
(3 marks)

Partner Organizations

