Problem Statement - FoodHub

**Submission type**

:

File Upload

**Due Date**

:

Feb 28, 12:00 AM

**Total Score**

:

40

**Available from**

:

Feb 11, 7:00 AM

**Description**

**A hand holding a phone

Description automatically generated with medium confidence**

**Context**

The number of restaurants in New York 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 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.

**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 to improve the 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

**Note: The required data has been provided in a web-page format (.html). You will find the file 'foodhub\_otder.html' in the next section. Please download the file before getting started with the project.**

**Submission Guidelines**

1. A learner notebook has been shared to guide you through the project, perform the analysis in the right manner, and generate insights from the data. **It is mandatory to download the same and follow the instructions mentioned in it to complete the project.**
2. There are two parts to the submission:
   1. A well commented solution notebook prepared from the learner notebook [format - .html]
   2. A report as you would present to the Data Science team [format - .pdf]
3. Any assignment found copied/plagiarized with other groups will not be graded and awarded zero marks.
4. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.
5. Submission will not be evaluated if
   1. it is submitted post-deadline, or,
   2. more than 2 files are submitted.

**Best Practices for Notebook**

* The final notebook should be well-documented, with inline comments explaining the functionality of code and markdown cells containing comments on the observations and insights.
* The notebook should be run from start to finish sequentially before submission.
* It is important to remove all warnings and errors before submission.
* The notebook should be submitted as an HTML file (.html) and NOT as a notebook file (.ipynb).
* It is recommended that you read the problem statement and go through the criteria and description mentioned in the rubric before starting the project.
* Please refer to the FAQ page for common project-related queries.

**Best Practices for Report**

* The report should be made keeping in mind that the audience will be the Data Science team of a company.
* The key points in the report should be the following:
  + Business Overview of the problem and solution approach
  + Key findings and insights which can drive business decisions
  + Business recommendations
* Focus on explaining the key takeaways in an easy-to-understand manner.
* The inclusion of the potential benefits of implementing the solution will give you the edge.
* Copying and pasting codes from the notebook is not a good idea, and it is better to avoid showing codes unless they are the focal point of your report.
* The report should be submitted as a PDF file (.pdf).
* A business report template has been provided for reference. It is not mandatory to use it for submission.

Happy Learning!

**Scoring guide (Rubric) - FoodHub Rubric**

| **Criteria** | **Points** |
| --- | --- |
| **Understanding the structure of the data**  - Overview of the dataset shape, datatypes - Statistical summary and check for missing values - Answer all the key questions asked in this section | 3 |
| **Univariate Data Analysis**  - Explore all the variables and provide observations on the distributions of all the relevant variables in the dataset - Answer all the key questions asked in this section | 10 |
| **Multivariate Data Analysis**  - Perform bivariate/multivariate analysis to explore relationships between the important variables in the dataset - Answer all the key questions asked in this section | 15 |
| **Quality & Use of visualisations**  - Use proper visualizations for the analysis and provide observations on the plots | 3 |
| **Conclusion and Recommendations**  - Conclude with the key insights/observations | 3 |
| **Presentation - Overall quality**  - Structure and flow - Crispness - Visual appeal - All key insights and recommendations covered | 3 |
| **Well commented Python Code**  - Structure and flow - Well commented code | 3 |
| Points | 40 |