Problem Statement - FoodHub

**Submission type**: File Upload

**Due Date** : Oct 19, 5:30 AM CDT

**Total Marks** : 60

**Available from** Sep 26, 9:30 AM

**Description**

**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\_of\_the\_order: 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

**Submission Guidelines**

1. There are two ways to work on this project:

**i. Full-code way:**The full code way is to write the solution code from scratch and only submit a final Jupyter notebook with all the insights and observations.

**ii. Low-code way**. The low-code way is to use an existing solution notebook template to build the solution and then submit a business presentation with insights and recommendations.

The primary purpose of providing these two options is to allow learners to opt for the approach that aligns with their individual learning aspirations and outcomes. The below table elaborates on these two options.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Submission type | Who should choose | What is the same across the two | What is different across the two | Final submission file [IMP] | Submission Format |
| Full-code | Learners who aspire to be in hands-on coding roles in the future focussed on building solution codes from scratch | Perform exploratory data analysis to identify insights and recommendations for the problem | Focus on code writing: 10 - 20% grading on the quality of the final code submitted | Solution notebook from the full-code template submitted in .html format | .html |
| Low-code | Learners who aspire to be in managerial roles in the future-focussed on solution review, interpretation, recommendations, and communicating with business |  | Focus on business presentation: 10 - 20% grading on the quality of the final business presentation submitted | Business presentation in .pdf format with problem definition, insights, and recommendations | .pdf |

Please follow the below steps to complete the assessment. Kindly note that if you submit a presentation, ONLY the presentation will be evaluated. Please make sure that all the sections mentioned in the rubric have been covered in your submission.

**i. Full-code version**

* Download the full-code version of the learner notebook.
* Follow the instructions provided in the notebook to complete the project.
* Clearly write down insights and recommendations for the business problems in the comments.
* Submit only the solution notebook prepared from the learner notebook [format: .html]

**ii.** **Low-code version**

* Download the low-code version of the learner notebook.
* Follow the instructions provided in the notebook to complete the project.
* Prepare a business presentation with insights and recommendations to the business problem.
* Submit only the presentation [format: .pdf]

2. Any assignment found copied/plagiarized with other submissions will not be graded and awarded zero marks.

3. Please ensure timely submission as any submission post-deadline will not be accepted for evaluation.

4. Submission will not be evaluated if

* it is submitted post-deadline, or,
* more than 1 file is submitted.

**Best Practices for Full-code submissions**

* 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 in a sequential manner 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).
* Please refer to the FAQ page for common project-related queries.

**Best Practices for Low-code submissions**

* The presentation should be made keeping in mind that the audience will be the Data Science lead of a company.
* The key points in the presentation 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 from the notebook is not a good idea, and it is better to avoid showing codes unless they are the focal point of your presentation.
* The presentation should be submitted as a PDF file (.pdf) and NOT as a .pptx file.
* Please refer to the FAQ page for common project-related queries.

Happy Learning!

**Scoring guide (Rubric) - FoodHub Project 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 | 5 |
| **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 | 15 |
| **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 | 20 |
| **Quality & Use of visualisations**  - Use proper visualizations for the analysis and provide observations on the plots | 6 |
| **Conclusion and Recommendations**  - Conclude with the key insights/observations | 6 |
| **Presentation/Notebook - Overall Quality**  - Structure and flow - Crispness - Visual appeal - All key insights and recommendations covered? OR - Structure and flow - Well commented code - All key insights and recommendations covered? | 8 |
| Points | 60 |

FAQ - FoodHub

**1. How should one approach the FoodHub project?**

* Before starting the project, please read the problem statement carefully and go through the criteria and descriptions mentioned in the rubric.
* Once you understand the task, download the learner notebook and the dataset to get started with the project.
* Many parts of the learner notebook are omitted and replaced with questions. You are expected to fill in the gaps as per the instructions/questions. Try to use descriptive statistics and visualization to understand the data.
* Once the EDA is completed, it is important to close the analysis with key findings and recommendations to the business.

**2. Is there a way to transfer the graphs from Colab to the presentation without it looking blurry?**

There are multiple ways to transfer the graphs.

1. Use the following line of code just after the visualization code:

plt.savefig("output.jpg", bbox\_inches='tight')

For example:

sns.histplot(data=data, x='column')  
plt.savefig("output.jpg", bbox\_inches='tight')

2. Use the snipping tool to snip the visual plot from the Jupyter notebook and paste the snip in ppts.

3. Right-click on the image and click on copy and paste the copied plot in the ppt or document.

**3. How should the presentation look like?**

The presentation should be made keeping in mind that the audience will be the Data Science team of an organization and should focus on explaining the key takeaways in an easy-to-understand manner. The visualizations used should be readable and it is better to avoid showing codes unless they are the focal point of your presentation.

**4. In Question 14, do we need to add the cost of the orders while calculating total revenue?**

The **total revenue** should be calculated based on the applied surcharges only.

Project Discussion Forum

**Due Date**

:

Oct 18, 2:30 AM CDT

**Instructions**

This is the Project discussion thread for the **FoodHub** project of **Python Foundations**.

This discussion forum is open from **26th September** to**17th October**to help you with your project-related queries.

You can also use this forum, to discuss your approaches, and share roadblocks you face and discuss them amongst your peers.

Points to note :

Please feel free to raise any blockers or doubts that are preventing you from moving ahead with the project.

1. Please do **not** share your codes here as we do not want anyone to be biased by someone else's work.
2. For faster resolution, please make sure you mention the following points:  
   a. Error screenshot  
   b. Python version or any other technical details you want to mention  
   c. The steps that you have taken to resolve the issue and why you think it might still be occurring.
3. Please do not post general questions outside the scope of the project.
4. Please maintain the professional decorum of the forum.
5. The response time is 12-24 hours on weekdays.
6. Queries hinting towards or giving away the solution will be removed from the forum.

Please refer to the following links for frequently asked questions to make sure that your question is not already answered: