EDYODA DIGITAL UNIVERSITY

Data Science Certification Exam - Test 2

Question 1 [Marks 40]

You work in a news agency & are responsible for identifying the most important topic for writing articles by your content team. Develop a Jupyter Notebook which on executing every time shows top 10 Twitter trending topics in India in the last 24 hours. Hints: Connecting with Twitter API. Bringing Dashboard-like behaviour in Jupyter notebook. *Please note: Visualisation is mandatory. You will receive 0 marks if you do not add visualisation.*

Question 2 [Marks 40]

Epilepsy is a nervous system disorder that affects movement. The dataset contains 195 records of various people with 23 features that contain biomedical measurements. Your model will be used to differentiate healthy people from people having the disease. Target Column is 'status'. Identify the model with the best params. *Please note: Visualisation is mandatory. You will receive 0 marks if you do not add visualisation.* Data Link –

https://github.com/edyoda/data-science-complete-tutorial/blob/master/Data/epilepsy.data

Question 3 [Marks 40]

Predicting Shopping Mall Sales. You will have to create a model to predict revenue. Identify the model with the best params. Target Column – Revenue. *Please note: Visualisation is mandatory. You will receive 0 marks if you do not add visualisation.* Data Link –

https://github.com/edyoda/data-science-complete-tutorial/blob/master/Data/Shopping_Revenue.csv

Question 4 [Marks 80]

Student Attendance Dashboard for Decision Making. You will be given a Google Sheet with different sheets holding data of different classes. Create a dashboard showing a summary of what's happening across different classes. Which students need attention for missing sessions? Which

students are unsatisfied? Which students are very satisfied? <Date> Column shows attendance. R-<Date> Column shows the rating by the candidate. Extra marks for identifying more useful insights. Missing values of the Type column can be treated as STUDENT. This data gets updated on a daily basis. The latest data gets injected in early columns. Your code shouldn't need any changes to deal with the latest data. *Please note: Visualisation is mandatory. You will receive 0 marks if you do not add visualisation.* Data Link –

https://docs.google.com/spreadsheets/d/1Y_g6LQI2GLqpKf-fj-rH7D9gfnP2MJzpFA6t0AsBMi8/edit?usp=sharing

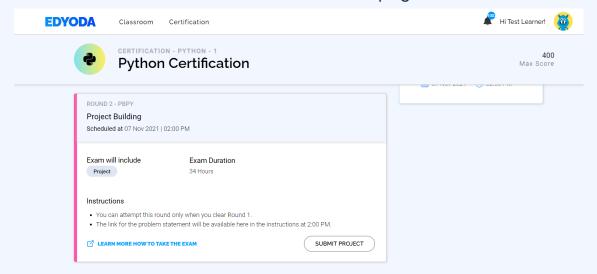## Grading Details

- **Total Grades:** 200
- **Passing Grades:** 150

Grading Guide

Please ensure that you cover all the required pointers for all questions. Missing any of the pointers will result in losing marks.

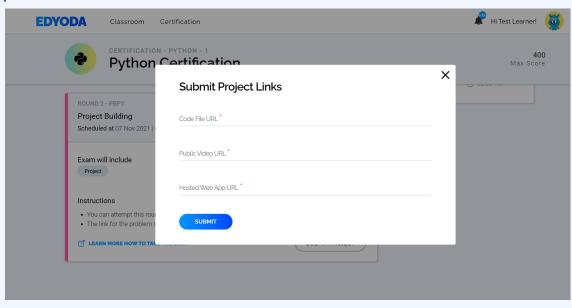| Question	Grading Objective	Total Marks
Question 1	Connecting Twitter API	10
	Getting Top 10 Trends	20
	Dashboard Visualization	10
Question 2	Data Analysis: 1. Data Cleaning 2. Data Visualization	15
	Building Predictive Model 1. Organizing Data into Training and Test 2. Splitting data into dependent and independent variable 3. Building multiple models(at least 4) and finding out best out of them	20
	Feature Selection	5
Question 3	Data Analysis 1. Data Cleaning 2. Data Visualization	15
	Building Predictive Model 1. Organizing Data into Training and Test 2. Splitting data into dependent and independent variable 3. Building multiple models(at least 4) and finding out best out of them	20
	Feature Selection	5
Question 4	Which students need attention for missing sessions? Answer with visualization	20
	Which students are unsatisfied? Answer with visualization	20
	Which students are very satisfied? Answer with visualization	20
	Creating your own insights	20

Submission Details

- ***You have to submit the Jupyter notebooks for all 4 questions***. You don't have to use PowerBI. Please upload all your notebooks in a folder on Google drive. Please ensure that the folder is available publicly or else the grading team will give you zero marks. To check that please try to open the URL in incognito mode. You have to share the public URL of this google drive folder.

- ***Please ensure that you DO NOT COPY the code from your batch mates or the internet. Our proctoring tool will detect plagiarism and you will receive zero marks in the exam.***

- ***Please DO NOT share your code publicly. If anyone copies your code then both you and the other person will receive zero marks.***

- You have to submit the code URL on the certification details page on the classroom portal. Follow the steps below:
  - Goto Round 2 card on the certification details page

  - Click on the "Submit Project" button
  - Fill the form in the popup and submit. Your project does not require a video URL and hosted web app URL so you can simply

paste the code URL itself.

- Please ensure that you make your submissions on time. **No late submissions will be accepted.**