



## Dear Participant,

Welcome to the second stage of **Data Science and Machine Learning** track of Spectrum Internship Drive. As you have completed the Task 1, hope you must be comfortable and have a good knowledge of Basic python, pandas, numpy and file handling. So now it's time to dive a bit more deeper into the field and explore. So with the knowledge from Stage 1, here we go!!

Welcome to the **Stage 2.**

This stage will introduce you to visualize and plot data using **Matplotlib** and **Seaborn** and also make you used to working with data, cleaning data, various types of plotting and visualization you can do with data, and how numpy and matplotlib libraries help you with the same, this is what we call as **EDA(Exploratory Data Analysis)**

## TASK 2 :

### Prerequisites::

- Numpy                      -Pandas                      -Google Colaboratory

### Technology Stack to be used:

- Numpy    - Pandas    -Matplotlib                      -Seaborn

### Tasks:

This time your task will be covered in 2 parts, i.e there will be 2 separate notebook for part 1 and 2. Go through part 1 first and complete it.

At the end of part 1 notebook, you will find a link for part 2.

So click on the link below and open the colaboratory notebook.

Below link is for part 1

<https://colab.research.google.com/drive/10xr5VWec7wiibNbpXwy4BLhsjrPSp9ri?usp=sharing>

Link for part 2(also given at the end of part 1 notebook):-

[https://colab.research.google.com/drive/1PMjpgyFk\\_-4Cqw2C49\\_3znhSBeTJCoL#scrollTo=SdQ2oLeORw6C](https://colab.research.google.com/drive/1PMjpgyFk_-4Cqw2C49_3znhSBeTJCoL#scrollTo=SdQ2oLeORw6C)



## Important Instructions before attempting the task:

- After opening the link above, go to top left and click on File->Save a Copy in Drive. Wait for a few seconds, and a new window will open. This will be the copy of the Task 2(part 1) colaboratory notebook. You have to attempt your task in this copy. Similarly you will be doing for Task 2(part 2).
- Your colaboratory notebook is saved, but everytime you close and reopen it again, you have to run all the cells, till the last cell you have coded, to download all the datasets and give the notebook a fresh run.

## Submission Guidelines:

- You can submit your task 2 i.e both part 1 and 2 by
  - Uploading your task to google drive or a public github repository.

Upon completion of your task, go to top left and click on File->Download->Download .ipynb. In this way download both part 1 and part 2. Then, Create a folder as Task-2 and keep both the .ipynb files(part 1 and 2) in it. Then upload the folder to either Github, or drive. Copy the link for github, or for google drive, right click on the uploaded folder, click Share, then click on "Change to anyone with the link", then Copy Link, and submit that link in the Dashboard.

**Note:- Read the Submission instruction properly and share the drive link which can be accessed.**

In case of any confusion or queries, you can drop a text in the data science group in Telegram which you would have joined by now. Or if you haven't joined, log in to your dashboard at <https://internship.spectrumcet.com> , and click on the bottom right telegram button.

## LAST DATE OF SUBMISSION:

11/07/2021 , 11:59 pm.

Warm regards,

**SPECTRUM, CET-B**