Final Project

Deadline : Sat Dec 10 11:59 pm

Topic

- Perform 5 Analysis on a dataset.

Dataset (20 points)

- Select a dataset of your choice.

- The dataset should have sufficient dimensions that you can perform good analysis. (Think Rationally)

- You can download multiple datasets and use all of them to perform your analysis.

- You can use a CSV downloaded file or using REST API you can gather your own data.

- Your final submission should have a folder called 'data'.

- For people having a single file of more than 200 MB, you would have to think of a way to store your data since Github has a size limitation on a individual file.

- A single high level script + ipynb should be there called "collect_data.py"

Analysis (50 points)

- Have a question that you want to ask to the data.

- Each of the analysis needs to have a plot which helps to easily deduce the answer to the question.

- The output of all the analysis should be in a single readme.md file (No PDF, no ppt, no txt) at the root level of the folder.

- Each of the analysis should be in a single file ("analysis_N.py"). (One file for one analysis.)

- Try not to download any data in your analysis script.

Bonus Points (30 points)

- Integrate more than one data source.

- Analysis has pre-exploration steps.

- Analysis are dynamic.

- Create a dashboard.

- Integrate an API that would help user perform analysis. Eg. API.ai, Twilio.