Data601-Homework-2

September 28, 2020

1 Overview

In this homework you will put the concepts and techniques you learned in the first 5 weeks of the course to use. You will get data from a Python API or with web scraping and then apply an exploratory data analysis with this dataset.

1.1 Deadlines:

1. Ask any dataset, API, website-related questions by COB October 2nd.
2. Submit the completed deliverables at the beginning of class on October 8th at 4:30 PM.

2 The API's you can choose from

- Spotify API
- Yelp API
- Twitter API
- Open Weather API

If you have found any other API's that might be interesting to use, please check with me before you begin.

3 Webscraping

- Make sure you check the webscraping policy of the webpage you will be scraping.

4 Deliverables

For this project there will be two deliverables:

1. A Github repo that contains your project.
2. A Jupyter Notebook (I will refer to it as `technical notebook` or `report`).

4.1 Github Guidelines

- Put an attention grabbing title to the repo.
- Github should include a license.
- ReadMe should be included.

- ReadMe should include these sections:
  - Summary
  - Data description and links to source.
  - Summary of the folders/files in the repo.
  - Check this link for a clean repo

4.2 Report Guidelines

This will a jupyter notebook and it will be uploaded to your github repo.

Should include the following two sections:

4.2.1 Getting the data

In this section, describe how you used API.

- You should include the code that you wrote
- Make sure your code is reproducible on any other machine.(outline each step one should take to replicate your results.)
- When this code runs, it should create a folder named 'data' and it should save the data in this folder as a 'csv' file.

Make sure that you don't share your account authentication if applicable.

4.2.2 Analyzing the data

In this section, you will repeat hw-1 with the dataset above. This part should include:

- Your research question: Clear statement of your question, background and motivation for the question.
- Methods of analysis, Rationale for using these methods, interpretation and conclusions.
- Limitations and further research
- References

Please use visualizations and narration to support your analysis.