Individual Python Project Project Deliverables Due Dec 2nd 2021

Must Meet 4 Requirements

1. Must involve reading in Files(CSV, JSON) OR interacting with Databases (Access, SQL, PostGRESQL or MongoDB). Will involve reading in files

2. Must Involve NumPY OR Pandas OR SciPY

Will involve Pandas

3. Must have a Visualization component – either Data

Visualization (Plots) or a GUI

Will involve Data Visualization (Plots)

4. A learning component beyond that is covered in class. Additional libraries that we haven’t used in class or going deep into an existing library introduced in class.

Going pretty deep into Pandas

Deliverables

• A Less than 4-min video describing your Project

• Explain the problem and/or Dataset

• What you want the program to do

• Show the output of the program

• Explain any unique features that you incorporated.

• GITHUB CODE SUBMISSION

• Upload all necessary files to the Github Project Location

• Submit link to Github Code Repository

• Upload Video to Github or Link to access the video on

youtube/vimeo or any cloud storage.

Grading Criteria

• Grading Criteria

• Code Construction (40points)

• File Organization and submission to Github

• Commenting of Code

• Appropriate use of Functions and Variable Names

• Use of Object-Oriented Programming (BONUS)

• Technical Complexity of Program (35points)

• Video Presentation (25points)

• To Note: We do not grade on whether the results are correct or the model you used is appropriate or not. We check on how the python code is designed, constructed and organized.

Criteria that I would consider a Bad Project

• This project is a direct copy from the web. Please cite all data sets or data sources used. Or taken an existing tutorial available online and simply replicated it.

• This obviously takes 5 days to complete.

• Did not take any risk. Nothing new learned by the student.

• Really bad code organization and explanation.

• It clearly does not work and the solution is mocked up…or the solution was hard-coded in.

TA and Myself

• TA’s will not have the time to debug the code for you. Do not send code over to them to fix.

• We can perhaps help to fix an error printed by the console screen or pointing to a possible error in the

code based on the code error.

• Instructors + TA can help point to a solution pathway for the project.

Sample New Learning Component

• Graph Based Models

• Network Analysis

• Advanced Data Visualization with Large Data Sets

• Image and/or Video Analysis Processing Libraries

• Building GUI (Plotly + DASH) + Databases

• Machine Learning (NLP, DeepML)

• Computational Geometry

• Finite Element Analysis, Computational Fluid Dynamics

• Digital Signal Processing

• Economic Model, combining Simulation (extensive use of numpy, scipy libraries)

Sample Project Themes

• A 3D Game

• Data Entry, Data Recording, Report Dashboard

• A Recommendation and Prediction System

• Simulation (Discrete, Continuous)

• Munging through Large Datasets and Analyzing/Predicting outcomes

• Network Analysis (Traffic, Social, Transport)

• Machine Learning (NLP, DeepML)

• Automation of Tasks

• Web Scraping (Legal ofcourse)

• Web Design (Django or Flask)

To Do

• Submit 1 page proposal of Project Idea with Instructor.

• Deadline : Nov 2nd 2021.

• Even if you have got an email approval from Dr Starly, must still submit it so that I am sure that

everyone’s project idea is approved.

SAMPLE DATA SOURCES

• https://catalog.data.gov/dataset

• Kaggle

• https://earthdata.nasa.gov/

• https://pds.nasa.gov/datasearch/data-search/

• https://registry.opendata.aws/

• https://cloud.google.com/bigquery/public-data/

• https://archive.ics.uci.edu/ml/datasets.php

• Data from other NC State courses

• Your own company’s data