M.Sc. In Data Science
Programming for Data Scientists
2020/2021 Academic Year

Assignment 4 – Dynamic Web Visualisations.

Motivation for Assignment

Cleaning, rearranging, mutating, and processing data is all well and good, but – at the end of the day – you need to be able to produce and publish outcomes. Creating visualisations is an excellent example of this output, and arranging for the publication of visualisations is key, too. Targetting the WWW as a visualisation delivery mechanism ensures the widest audience for your output. Enabling the dynamic creation of visualisations via a web-based application can also enhance interactions with published outcomes. This assignment explores the technologies and techniques used to enable such publications.

Details of Assignment

Study the included CSV dataset which was downloaded from the Irish Government's Covid-19 dataset site.

Your tasks are as follows (with individual mark allocations shown within square brackets):

1. Produce Python code (in a Jupyter Notebook) which tidies the data from the above link and makes it suitable for analysis using pandas/plotly. Be sure to annotate this notebook, detailing the processing steps you devised to tidy the data. Your tidied data, saved as a new CSV, is to be used in Part 2 of this assignment. [10]

2. Create a *second* Jupyter Notebook which creates three *individual* visualisations which each highlight an analysis outcome for this data. Use the plotly visualisation library. [15]

3. With your visualisations complete, you are to create an interactive, dynamic web application which showcases the visualisations from Part 2 of this assignment. Your webapp publishes the visualisations (so that they are visible from a web browser, as served up by your webapp's web server). You are to use the Flask micro-web framework to build your web application (see: https://palletsprojects.com/p/flask/ and https://www.fullstackpython.com/flask.html for more on Flask, as well as the relevant resources discussed recently in class). [10]

Marks Allocation, Submission, and Deadline

- This assignment is worth 35% of your total grade, with each part allocated an individual mark (as above).

- You are to submit a Jupyter Notebook which details the code and annotations which tidies the original data into a format that supports analysis.

- You are also to submit a *second* Jupyter Notebook (which has been through *Kernel… Restart & Clear Output* before saving) showing the work product which produces your visualisations.

- Finally, you are also required to submit a ZIP archive of your dynamic web application (containing all of the Python code to your web application). Share this ZIP file with paul.barry@itcarlow.ie via OneDrive.

- The due date/time for this assignment is: **5:00pm on Friday January 15th 2021 –** your email to paul.barry@itcarlow.ie must arrive prior to this deadline.

This is an individual assignment: you are expected to work on your own, and that the work you submit is written by you. **You must declare if this is not the case.**