Lab1

January 20, 2023

Lab 1 will have two parts (A) and (B). In Lab 1, you are required to implement in Python, the classes and objects. A parent class will define the methods that will be inherited within the data classes. You will do this together for the two datasets (Happiness [1] and Covid [2]) that are provided. Initially you can get the dataset from the Canvas Labs Folder.

Datasets

Dataset 1 is from the World Happiness Report presented at the United Nations. You can read more about this dataset here: [1]. Dataset 2 is from recent COVID pandemic and it gives statistics about different countries. You can read more about this here: [2].

Part A

Python Basics: Object oriented concepts like inheritance, polymorphism: You can use only 10 rows from each dataset. Create a data structure for these two datasets, list different columns, if its a numeric value, show mean, variance and stand deviation. You should be able to search for min and max values for each column. For example, one data class represents the Happiness dataset and another data class represents the COVID datset. Describe the data structures for example lists, sets, etc. that are used. **You are required to use the Draw IO[1] online tool to draw the UML diagram for your implementation and include in your report in the Appendix.

Part B

Getting the data and visualizing it: (*Part B is based on lecture 3 - you can attempt it after the lecture that week) You need to be able to fetch the data in the ways possible from 1) CSV Files, 2) Scrapinng the Web, 3) by Web APIs. Investigate if this is possible for each dataset and if so, implement it. That means, being able to fetch data in different ways. Data gathered should be stored locally in a MongoDB database for further usage and visualization. This is done via Python code. You are then required to visualize different aspects about the datasets with interesting observations using line charts, bar charts, heatmaps and scatterplots. to explore a little more than basic graphs. Describe your graphs, observations and label them well.

Work Environment

This lab project will be created using a virtual environment in Anaconda to show application level isolation.

Lab Report

The report should be written in the lab report format [3] using LaTeX in Overleaf. Download the pdf file after you finish writing and submit the pdf along with your zipped code files to Canvas. **You must document your code properly so its readable.

References

[1] "World Happiness Report Dataset." https://www.kaggle.com/unsdsn/world-happiness?select=2019.csv. Accessed: 2021-01-05

[2] "WHO Coronavirus Disease (COVID-19) Dashboard." https://covid19.who.int/table. Accessed: 2021-01-28.

[1] https://app.diagrams.net/