Principles of Programming Languages & Translators

Department of Computer & Informatics Engineering

University of Patras

Spring Semester 2022

Teachers: Ioannis Garofalakis, Spyridon Sioutas, Panagiotis Hatzidoukas

Optional Laboratory Exercise

The Python ecosystem, which includes the programming language as well as a number of external software packages and libraries, is a powerful open source tool, which has great utility outside the strictly limited area of computer science, and in the research of every specialty regarding the analysis and processing of data.

The Python language is essentially a general-purpose programming language, unlike other languages that focus strictly on specific areas such as the statistical analysis. Its great advantage concerns the possibility of writing code in which the libraries relevant to the requested scientific areas of each project can be used each time.

Specifically now for this lab exercise a data set with elements such as the amount and type of waste collected for recycling in the city of Buffalo will be used. The specific data is provided by the website https://data.buffalony.gov/Quality-of-Life/Monthly-Recycling-and-Waste Collection-Statistics/2cjd-uvx7

Wanted:

First, a python script should be implemented that will collect the data from the above OpenData Buffalo website. Then a program should be built in python that would process only the necessary data regarding the requested graphs of the exercise.

Specifically, it will be necessary to export the graphs for the following cases:

- Total quantities of recyclable items per year
- Overall presentation of recyclable items and their corresponding quantities

- Presentation of the 5 months with the largest amount of recyclables items, regardless of year and type of recyclable items

At the end, the specific data should be loaded into corresponding tables of a MySQL database (or any database you wish), and exported to corresponding .csv files (the .csv will also be a deliverable of the exercise).

Attention all the necessary data transformation actions should be done exclusively in Python language. The database will be used for storage ONLY .

Hint: Use the necessary Python libraries such as pandas, xlrd , matplotlib or any others you deem necessary.

Deliverable

1. **Written Report** (in pdf or word file) that will include:
   - o The code in python language enriched with comments
   - o Screenshots of examples of the application (and its shape database)
   - o The requested graphs (with titles, legend)
   - o Comments - Assumptions made in developing the work

2. Compressed into **a zip file:**
   - o The above written reference
   - o The **FINAL** code in python.
   - o Export of the schema, database data
   - o The .csv files that will contain the final data

The zip file must be named the student's registration number (eg **3972.zip),** and upload **(MANDATORY)** to e-class.
In a separate .txt file inside the zip, mention the name, the year, the student's registration number and e-mail address.

Clarifications

1. The exercise is **individual**
2. The exercise is **optional with a bonus of up to 1 unit in the final grade, as long as her grade is >=5**
3. Final delivery date is the date of the written exam June 2022 **ONLY!** The date will be determined accordingly the oral examination of the work.
4. The exercise **WILL NOT BE RESERVED** for the next ACADEMIC YEAR!

5. For any questions or suggestions you can contact by e-mail at mvonitsanos@ceid.upatras.gr or in the "Discussions" area on his page course in eclass (https://eclass.upatras.gr/modules/forum/?course=CEID1091)