Project-2:

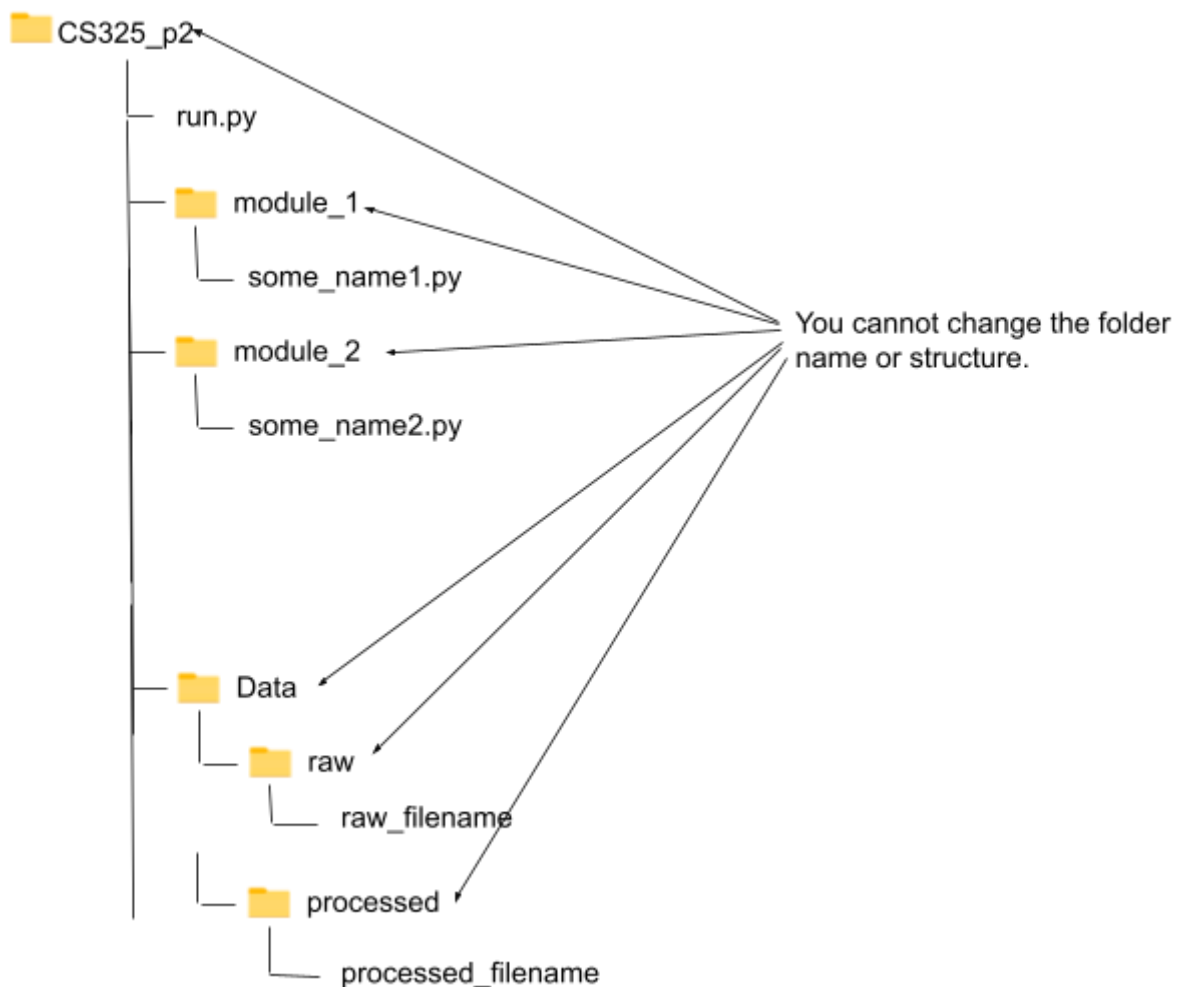**Due Date: 03/12**

Description:

1. Break down your project-1 code into logically two different files/modules.
2. Use import to call the functionality of one file into another.
3. Use the folder structure given below and name the folders as shown.

```
CS325_p2
├── run.py
├── module_1
│   └── some_name1.py
├── module_2
│   └── some_name2.py
├── Data
│   ├── raw
│   │   └── raw_filename
│   └── processed
│       └── processed_filename
```

You cannot change the folder name or structure.

4. The run.py will contain a main function and this is the main python program which is going to call the rest of the modules. You can name the rest of the *.py files as you want but the folder name and structure cannot be changed. For example to run the program it will be
   > python3 run.py URL/File/argument
5. Store the unprocessed or the raw file inside the folder Data->raw. You can name the raw or the unprocessed data files as you want. I have seen students having author

names and some other content in their output file. Puth those files in this directory. Rewrite your code to achieve it.
6. In the processed folder the files should only contain the articles and nothing else. Rewrite your code to achieve it.
7. Document your code, that is write comments on top of **every python file** explaining what this module expects as **input** and what will be the **output** of it. Also explain the **working** of the code.
8. While creating the modules and different functionalities. You have to use one of the SOLID principles. You can use any one of them. But you have to mention on top of the python file which SOLID principle you are using and what is the benefit of it.
9. Create a Class diagram of the whole software and save it as a JPEG/PNG.

Rubric:

You should be able to break down the code into two modules (containing functions and classes) which is important, that is the basis of this project. If you are not able to break down the code into two logical well defined modules 30% of your points will be gone.

1. Create a new git branch of the repository you are using for project-1. In the new branch upload the content of the folder CS325_p2(as described above in description): 30 points
2. Following the given folder structure: 10 points
3. Writing comments on top of each python file; 10 points
4. Able to implement and justify a SOLID principle in the project: 30 points
5. Able to create a class diagram (upload to github): 20 points

You have to send the public github link to your repo to the TAs, send it to any one of them. **Please be sure that the subject of the email that you are going use should be "Project-2 CS325" and in the body do not forget to mention your name with your 800 number.** They will grade randomly so there is no fixed TA to a corresponding student.
TAs email IDs: ntavlee@siue.edu and nimorga@siue.edu

ZIP AND UPLOAD TO MOODLE

IF YOU HAVE ANY QUESTION ASK ME AFTER THE CLASS OR DURING THE CLASS, I AM HAPPY TO EXPLAIN.