

Semester: Fall 2021
SE 412: Project
Instructor: Dr. Ziad Al-Sharif
Due Date: January 13th, 2021, at 11:55 pm sharp; no extensions whatsoever

Note: you are only allowed to use these three packages: **urllib**, **matplotlib**, and **re**. Any use of other packages such as BeautifulSoup will get you a ZERO on the project.

Watch the video about opening a webpage using the “**urllib**” package that is available at (<https://youtu.be/J6XGaXe1BCs>), the video about the “**matplotlib**” package that is available at (<https://youtu.be/AjU4-KBvb-4>), and the video about the “**re**” package that is available at (<https://youtu.be/lePyzyEN7gs>). Then write a python script that does the following:

- 1) [2 Points] Asks the user to enter any URL for a website (webpage), read the contents of that page, and prints out on the screen all of:
 - a. [0.5 Point] Total number of lower letters [a-z], the sum of occurrences for all lower letters
 - b. [0.5 Point] Total number of upper letters [A-Z], the sum of occurrences for all upper letters
 - c. [0.5 Point] Total number of digits [0-9], the sum of occurrences for all digits
 - d. [0.5 Point] Total number of occurrences for the string “href=”
- 2) [2 Points] Use **matplotlib** to draw a **pie** chart that shows the occurrences of lower letters, upper letters, and digits in the source of the opened webpage
- 3) [2 Points] Use the **re** package to identify all of the links (i.e. URLs) and emails that may be available in the source of that webpage.
- 4) [2 Points] Use the **matplotlib** package to draw a **column** chart with the frequency of all letters in that webpage (A-Z and a-z). Note in this part you should ignore the difference between lower case and uppercase letters; in other words, you should treat ‘A’ and ‘a’ as one letter and so on.
- 5) [1 Point] Use the **matplotlib** package to draw a **column** chart with the frequency of all characters on that webpage. Note that characters are NOT limited to letters and digits, it includes any key on the keyboard
- 6) [1 Point] Use the **matplotlib** package to draw a **pie** chart with the frequency of digits in that webpage (0-9)

Note:

- A. Your program should be MENU-driven, in which the script will keep asking the user to enter a choice between (1-5) until the user enters a zero (0), then the program will exit/terminate.
- B. Any program that is not menu-driven will not be graded (you will get a ZERO for it)
- C. Any two programs that may show similarities that can be considered as cheating and a ZERO will be assigned to them, further actions can be considered against cheaters.