CSE 140 Section 8 Exercise

Scenario

Suppose that you and your partner have been tasked with developing a Python program, `word_graph.py`, which visually compares the frequency of various words in a corpus. This program must be written, tested and submitted by next Monday. *(Not really! This is just an exercise!)*

Your program will take the following command-line arguments:

```
python word_graph.py [directory path] [word 1] [word 2] ... [word n]
```

The directory path is assumed to contain only text files, and no subdirectories. These text files contain raw, clean text, which can be split into words using Python's `split` function.

Your program will produce a bar graph comparing the frequency of the words in the corpus. The y-axis of this graph is frequency; the x-axis has a bar for each word specified in the command-line arguments.

Program Design

- Decide how your program will report its results. Sketch some example output. Will it display the graph on the screen, or write the results to a file?

- Specify what functions you will write to decompose this problem. Write down the declaration and docstring of each function.

- Outline the main method of your program.

Division of Labor

Decide how work will be fairly divided between you and your partner, and develop a plan for completing the program by your deadline. Write down the key details of this plan. Some factors you should consider are:

- How will you share your source code? How will changes you make be passed to your partner?

- Who will work on what parts of the program?

- Who will write tests for each function?

- Who will validate the results?

- What happens if you find a bug in your partner's code? Do you wait for them to fix it, or fix it yourself?

- As you work, your program design will likely change. How will these changes be communicated?

- What are two advantages and two disadvantages of writing software in a group?