Programming Lab 4 – Reading a Classic
PHYS 2511 – Prof. Matthew Newby – Spring 2020

Goal:	Find the number of times each major character is mentioned in each chapter for Mary Shelley's Frankenstein. This is a **collaborative** assignment.
Requirements:	Create a program that reads in the plain text text of *Frankenstein*, and can determine how many times each character is mentioned in each chapter throughout the book. You will then attempt to make connections (or disconnections!) between the characters (listed below) based on this information.
Inputs:	• The text of *Frankenstein.*
Outputs:	• The number of appearances of each character's name by chapter. • A graph (or graphs) of this information.
Optional:	• Mathematically correlate character appearances with each other. • Perform additional analyses on the data (your choice). • Perform the same analyses with a different book.

Background:

Project Gutenberg (http://www.gutenberg.org/) is dedicated to providing public domain books that can be downloaded for free. The plain text of Mary Shelley's *Frankenstein* will be the subject of this activity, and the full plain text can be found here: https://www.gutenberg.org/ebooks/84 (as "Plain Text UTF-8". The encoding format is "UTF-8," which is a more modern form of ASCII).

For this lab we will read in the full text of *Frankenstein* and then count the number of times each character (listed below) is mentioned by name in each chapter. From this information, we will try to answer the following questions for each of the characters:

- Do they meet Victor Frankenstein (the main character who creates the monster)?
- Do the meet the creature (the monster)?
- What is their relationship (if any) with/to Victor?
- Do they die in the story, and if so, when?

Even if you have read the book and know the answers to these questions, you will still have to hypothesize based on the data from the book. You may even find that your data leads you to contradict the what you know about the book (this is a lesson in drawing conclusions from a limited set of data, then)!

If you have never read the book, or do not know the plot well, I highly recommend waiting until *after* you come to conclusions from your analysis to read the book or a summary. It is very interesting to contrast the conclusions you draw with a limited view of the data with what the "real" answer is.

The Cast: These are the characters you must analyze, and the name you want to use when searching the book is in parentheses, if needed. Names are case-sensitive! The two main characters are listed first, then the rest follow alphabetically:

Victor Frankenstein (Victor)
The Creature (creature)
Agatha
Caroline
Old Man De Lacey (De Lacey)
Elizabeth
Ernest
Felix
Henry Clerval (Henry)
Justine
William

Notes:

1. You will want to use file I/O to open and read the file. A basic way to do this is:

```
my_file = open("file_name_as_a_string", "r")
for line in file:
        STUFF
my_file.close()
```

This will only find the file if it is in the same directory as your python script. There are other ways to do this with files in other directories, but we won't cover those here.

2. First figure out how to analyze the file, then worry about making the figure.

3. The figure should be made with the matplotlib library. This will be discussed in class, and a guide is on Canvas.

4. This is a collaborative assignment! You must both work on the project and both of your contributions must show up in the Github history. Choose one repository to work in, and leave a text file in the other one pointing me to collaborative work. You will find working with another person to be very helpful for this assignment.