

Programming Assignment 4

Due: Friday, May 15th

There are two goals for this assignment:

- To write a Python program that reads in text files.
- To practice with dictionaries.

Write a Python program that **reads** in a text file (the file name should be provided as a function parameter), and **returns** a dictionary where the keys are words and the value is the number of times the word appears in the text file (“Hello” should be treated as the same word as “hello”, “hello.”, and “-hello” - in other words, all punctuation should be removed and capitalization should not be considered). I have included the file hamlet.txt to help you practice.

Instructions:

- Create a function that reads in a file (it will accept a directory as a parameter). Practice with hamlet.txt. Be careful, larger files might cause an “out-of-memory” error.
- Assign the contents of the file to a variable as a single string.
- Remove all punctuation from the string.
- Make all characters in the string lowercase.
- Split the string into a list of words.
- Create a dictionary where the keys are the words in the string and the values are the number of times the word appears in the string.
- **Return** this dictionary.

Call your script `shakespeare.py`

I have named my main function `main`. A sample session of the program appears below.

Session 1

```
>>> import shakespeare
>>> words_with_counts = shakespeare.main("/Users/theresamiglervondollen/Desktop/hamlet.txt")
>>> words_with_counts['the']
993
>>> words_with_counts['laugh']
6
```

Here, your output is simply the dictionary.

Every program that you submit in this class should begin with a comment block resembling the model below.

```
# your name
# your userid@ucsc.edu
# programming assignment 4
# a short description of what the program does
```

Test your program thoroughly. Attach the file `shakespeare.py` to Programming Assignment 4 on eCommons.