KJC4 4CREDITS

NARRATIVE FOR HOMEWORK9

I imported the string module .I then defined function byFreq which will return the second item in a pair. I opened the Dracula file for reading and read it into a variable “Dracula” using the read method. I then created 2 variables “chapters\_beg” and “chapters\_end” to store the index of where I wanted my slicing to start and end. This was done to extract only the chapters. I then stored the sliced text in a variable ”text”. I created an empty list “stripped\_punc”. I split “text” and used two for loops to loop though each element in this list, remove the punctuation and convert them to lower case. I then appended each of them to my list “stripped\_punc”. I then read in a text file holding the stop words I had picked using the read method. Each word was separated by comma so I split based on (“,”) and stored them in a variable “stop\_words”. I created an empty list “stripped\_stops”. I then looped through each element in “stripped\_punc” and if was not present in “stop\_words”, I appended it to my list “stripped\_stops” . I created an empty dictionary “dict”. I looped through each element in stripped\_stops, if the element was in dict already,I added one to its count, if it was not in dict, I set its count to 1. I used list(dict.items()) to create a list containing (key,value) pairs and stored it in a variable “items”. I then sorted items twice to get the pairs in a form such that the ones with highest frequencies came first but ones with the same frequencies came in alphabetical order. I then used a for loop to get the 100 most common words and their counts and printed them out.