Main Data structures:

1. case insensitive hashtable
2. case sensitive hashtable
3. heap

The insensitive hashtable is designed to hold the lower case version of each word and keep track of the number of occurrences and versions of each word. The update function to add or update words in the hashtable is expected O(1). Worst Case is O(n) where n is the sum of all the different words and their versions in the text file. At the end of its iteration it will hold lower case versions of each word, their frequencies and number of versions. Each word gets its own node to be stored in.

The sensitive hashtable is designed to hold the different versions of each word. Collisions are dealt with through linked lists. Expected case O(1) running time. Worst case O(n) where n is the sum of all the different words and their versions in the text file. At the end of its iteration it will hold the case sensitive version of each word from the input file. Each word will be stored in its own node.

The heap is only used for sorting and final output. After all the data has been gathered from the file and put into the hashtables; it is then extracted to an array which then gets put through a heapsort. Once sorted, the names and numbers get output accordingly.

Challenges: C is hard.
I thought i was done with my code about a week before it was due, but then i realized that once i stuff a book in the input file, i run out of memory. I thought a big file was considered to be a few hundred words; however, with a large number of unique words, my program failed. I did not understand the concept of free till a day before the code was due, and thus, have only managed to put a few in which did not solve the problem with huge input files.

Another issue would be that i had to learn how to use the terminal. So far, with all my experience in Java i have never needed to worry about memory or how compiling works. Eclipse has held my hand through all this mess; however, with C, eclipse refuses to be cooperative. So much of my time was put into learning and trying to use new text editors and the terminal.