Lab07 - War and Peace

Due: Fri May 16, 2014

This lab focuses on:
- ❏ Chapter 18 Collections

First step - Copy all the files from my k: drive, notably including `war_and_peace.txt`.
The Gutenberg Project is cool. They put free ebooks online: www.gutenberg.org. Click, click, click… I found War and Peace by Tolstoy. It's just a text file, so we can play with it.

Part 1 - Total words

Question: How many words are in War and Peace?

Here's the EZ pseudo-code:
```
foreach( token in file) {
    increment word count
}
```

Remember to start small! Don't start debugging with War and Peace. Create a small text file to use first. Get your program working. And then try War and Peace.

We have used the `Scanner` class before. In Lab 07, we want to scan for words separated by whitespace AND punctuation. And what about apostrophes?
I discuss this in my README file, so please check that out.

I only count words with 3 characters or more. I wasn't interested in "a" or "an" or...

Part 2 - Unique words

Question: How many **different** words are in War and Peace?

Here's the pseudo-code:
```
create new Collection
foreach( token in file) {
    if( collection does NOT contain token) {
        add token to Collection
        increment word count
    }
}
```

Remember:
- Copy-paste your basic token loop from part 1… it should work here as well
- Make sure that your String comparison is **case-insensitive**
- After processing all tokens, the size of your collection is your unique word count

Please do this 4 times… for 4 different Collection classes: `ArrayList`, `LinkedList`, `HashSet`, and `TreeSet`. Time and report the execution of each.

Part 3 - Rank words

Question: What are the **top 100** most used words in War and Peace?

This one's a little tougher. You need to use a `Map`.

```
create new Map
foreach( token in file) {
   current tally = Map get for token
   if( tally == null) {
      create a new tally
      put it in Map
   }
   else {
      increment the tally
   }
}
convert your Map to List
sort List by tally
report top 100
```

There's a lot of detail here.
Consult my README.

Good luck!

QOTD
We can know only that we know nothing.
And that is the highest degree of human wisdom.
- L Tolstoy, War and Peace (www.goodreads.com/work/quotes/4912783)