

1 Word frequency

For our activity this week we will be finding the most common words in a corpus. This will give you a chance to work with a few different data types, and will also give you practice searching for help in accomplishing a programming task.

1. Download a text file (probably ending with `.txt`) that holds the works of Shakespeare, or any other voluminous text.
2. Write a python program that reads this file and prints 15 most frequent words, along with their frequency (as a percentage of all words).

Extra fun Refine your definition of "word", e.g., to exclude punctuation, and to count variants of the same word as identical.

Median fun Print out enough of the most common words to account for 50% of the words that occur. How many words need be included?

"Spelling bee" fun Count the number of words that have exactly seven unique letters in them.

Palindrome fun Print all the palindrome words that occur.