CTA 200: Homework 2

Problem 1: Write a program that will read a file and determine that 5 most commonly used words with at least n characters. The program should be called with python name.py filename -n, where n is the minimum number of charters. The program should write the 5 most common words to a txt file with the number of occurrences.

Problem 2: Write a recursive function that takes input N and s such that it computes

$$\sum_{i=1}^{N} \frac{1}{i^s} \tag{0.1}$$

What happens when you take N to be very large? You can fix this problem by using a function in the "sys" library (look up the error message you get to find the solution).

Now write a non-recursive function that performs the same calculation. Which one runs faster? (use the "time" library to help you measure the runtime)

In the limit $N \to \infty$ this sum is the Riemann zeta function. This function is part of the scipy.special library. Make a log log plot show the difference between you function and the zeta function as you increase N from 10 to 100000. How much slower/faster is the calculation of the zeta function compared to your functions for 2,3 or 4 significant figure accuracy?

Send solutions to cta200@cita.utoronto.ca by Friday, May 8 at 5 pm
Solutions should be given as working (commented!) python scripts