**ECE 1111: Engineering Computation I**

**Homework No. 12: Histograms, Word Counts and Data Structures**

**Deposit your work in:**

/data/courses/ece\_1111/current/homework/hw\_12/<lastname\_firstname>

**Goal:** Introduce you to how data structures and algorithms can play an important role in high performance computing.

**Description:** Starting with this file:

/data/courses/ece\_1111/resources/data/text/data\_v00.txt

Compute a histogram of the word frequencies.

(1) (/p01) Use Unix command line tools. Google search “how to compute a sorted list of word frequencies using Linux command line tools.” You will find lots of solutions.

(2) (/p02) Write a C++ program that does the same thing. Your functions that open the file, read the text, parse it (split it into words), and compute frequencies should be implemented as member functions in a class named ComputeWordFrequencies. You should split each task into small functions that do one thing. Your driver program should then call these functions and should not contain any significant code other than function calls.

(3) (/p03) Write a Python program that does the same thing (hint: use dictionaries).

Compare the run-time of each program and analyze the results. Please your comments in a file AAREADME.txt in the base directory. Do all three versions produce identical results? If not, why? Demonstrate how they differ on some small test cases.

Task No. 2 will be evaluated based not only on the accuracy of your solution, but also on the quality of your code.