

CSCE 4273-001: Assignment 0

Due 11:59pm Monday, January 27, 2025

1 Word Count

File “pg100.txt” contains a complete collection of works by William Shakespeare. In this assignment, you will count the number of times each word appears in Shakespeare’s all works. You will run a Spark program on the cluster to do that. Please follow the instructions provided in “instructions.pdf” to access the cluster at AHPCC and run the spark program in “pywordcount.py” with the input data “pg100.txt”.

This is assignment will not be graded.

What to submit:

Submit the output file.