

COSC 311 Lab 1 (10 points)

Due Wednesday, 22 February 2023 at 11:59 PM

Please finish the following tasks and submit your lab report via MyClasses. Your submission must contain your source code file (.py file) and a PDF document. For each task, please include the output/results of your program (may use screenshot).

The attached is a Science paper (.txt file) published in Feb. 9, 2023. Write a Python program to analyze this paper and answer the following questions.

Task 1: How many different words appeared in this paper?

Task 2: What are the 10 words that appear most frequently (from high to low)?

Task 3: What are the appearance frequencies for the following words?

- Summerfelt
- wastewater
- greenhouse
- salmon

Task 4: What are the words appear exactly 1 time, 2 times, 5 times, and 10 times, respectively?

Task 5: Draw a bar figure to show the average length of words for each appearance frequency.

Policy

1. Each student **MUST** finish this lab independently. **NO TEAM WORK** and **DISCUSSION** are allowed. If you need any help, please feel free to contact the instructor.

2. You need to write your whole program in an editor and save your source code as a ".py" file, which will be submitted to MyClasses together with your PDF report.