**Challenge #4**

**Mid Term Capture the Flag**

**Cyber 262**

Please write a Python program to analyze all of the three given log files and print to console the following outputs:

* We define the ***“read from a file”*** ***similarity score*** between two log files as follows:
  + Assume one log file in total reads from N1 different files; assume the other log file in total reads from N2 different files; assume N2 is greater than N1;
  + Assume the intersection of the set of N1 files and the set of N2 files contains N3 files;
  + The “read from a file” similarity score is N3/N2, i.e., N3 divided by N2;
* We define the ***“write to a file”*** ***similarity score*** between two log files as follows:
  + Assume one log file in total writes to N1 different files; assume the other log file in total writes to N2 different files; assume N2 is greater than N1;
  + Assume the intersection of the set of N1 files and the set of N2 files contains N3 files;
  + The “write to a file” similarity score is N3/N2, i.e., N3 divided by N2;
* We define the ***“program execution”*** ***similarity score*** between two log files as follows:
  + Assume one log file in total executes N1 different programs; assume the other log file in total executes N2 different programs; assume N2 is greater than N1;
  + Assume the intersection of the set of N1 programs and the set of N2 programs contains N3 files;
  + The “program execution” similarity score is N3/N2, i.e., N3 divided by N2;
* Please output the three similarity scores between Log A and Log B;
* Please output the three similarity scores between Log B and Log C;
* Please output the three similarity scores between Log C and Log A;

Note: A read/write event whose target is pipe/tty/socket/inode is not considered a file read/write. Thus they should be filtered before counting N1, N2, and N3.

**What to submit**:

* Your Python program
* The outputs you get (in a txt file)

Please put all these items in a folder and compress the folder into a single ZIP archive file. Include your team name in the file name of the ZIP archive.