

Problem 1. From the files in logs.tar.xz output the top 3 most frequent IP addresses that performed the GET method from 10 to 17 hours 2015-01-13.

The answer for this task will be the top 3 most frequent IP addresses and their frequencies. And commands that you've used to obtain it

Problem 2. In the same archive find a list of all files with the .tsv extension that are larger than 5 MB and start archiving in the background.

The answer for this task will be the number of items (.tsv files) an archive contains. And commands that you've used to obtain it

Problem 3

The data_for_science.tar.xz directory contains files of the following format: target class, tab, comma-separated list of keywords. Your task is to find unique words for the *bad* class, which are contained in the three largest files. Remember, *DOG* and *dog* are the same word.

The answer for this task will be 10,11 and 12 entries in a sorted set of unique words. And commands that you've used to obtain it