## LogAnalyzer (10 баллов)

### Легенда

Веб-программист Катя обнаружила очень странные дела, происходящие с сайтом. Благо, всё, происходящее на сервере, не проходит мимо и на всякий случай пишется в текстовый файл. Привело это, правда, к тому, что файл с логами достиг небывалых размеров и работать с ним даже в Sublime Text стало совсем некомфортно, что уж говорить об обычном виндовом notepad.exe, который зависал при попытке просто посмотреть логи… Да еще и менеджер проекта с продажниками периодически просят узнать, насколько популярна та или иная страница сайта. Упростим Кате жизнь?

### Постановка задачи

Дан лог некоторого сервера. Необходимо его проанализировать и сделать новый лог-файл.

Формат файла:

level | date | ip | source | exec\_time

где

* **level** – уровень логирования. Возможные варианты: INFO, WARN, ERROR, FATAL, DEBUG. Это поле всегда содержит 6 символов. Если level меньше 6, то оставшееся место добивается пробелами
* **date** – время запроса в формате yyyy-mm-dd hh:mm:ss. Например, 2017-04-08 16:23:42
* **ip** – ip address, с которого пришел запрос. Это поле всегда содержит 15 символов. Если ip меньше 15, то оставшееся место добивается пробелами
* **source** – запрашиваемый ресурс
* **exec\_time** – время обработки запроса

Пример файла:

WARN | 2016-11-06 18:42:27 | 148.183.254.223 | /dml/gnxdvthpm/imnwqjslc/gvnlmro/uyn | 2667  
INFO | 2016-11-06 18:42:29 | 158.167.157.188 | /ycmmdrvj | 2544  
WARN | 2016-11-06 18:42:32 | 206.68.20.151 | /rlfe/upqyh/wdwrfaiasp | 2478  
DEBUG | 2016-11-06 18:42:39 | 143.158.240.156 | /iaxdt/xojjiqc | 2447  
ERROR | 2016-11-06 18:42:45 | 243.185.193.212 | /wqnzup/vuyufwaj/uwy/ngjb | 2350  
FATAL | 2016-11-06 18:42:50 | 158.167.157.188 | /wkf | 2465

Для выполнения следует выбрать вариант под номером, который получается от остатка деления Вашего номера в списке группы на 4. В результате выполнения задания нужно будет создать файл с логами, в который надо будет включить только требуемые записи. Также данные записи должны быть отсортированы по двум ключам: первый - по времени выполнения, второй — по названию ресурса (оба по возрастанию).

0. Найти медиану по времени обработки запроса для указанного в параметрах source (по умолчанию по всем запросам). В новый файл с логами включить только те запросы, время обработки которых превышает медианное.

1. Найти 99 перцентиль по времени обработки запроса. В новый файл с логами включить только те запросы, время обработки которых меньше найденного.

2. Вывести все записи за последние N минут к заданному моменту времени, включить все их в новый файл с логами.

3. Найти три самых популярных ресурса на сервере. Подсчет популярности ведется по числу поступивших запросов. Ресурсы, к которым было произведено одинаковое количество запросов, считаются равнозначными, и должны учитываться при выводе статистики. Например, если к ресурсам A,B,C,D,E было совершено 5,3,4,5,6 запросов, то вывести надо все записи, соответствующие ресурсам E,A,D,C. В новый файл с логами включить все записи, соответствующие данным ресурсам.

**Требования:**

* Справка по -h или --help
* Проверка входных параметров, если таковые имеются
* Во всех вариантах на входе подается имя файла с логами (имеется в папке с заданием)
* Отсутствие ошибок при проверке кода на [http://shellcheck.net](http://shellcheck.net/)

Срок сдачи: 21.11.24, 23:59, GMT+5

Срок защиты: 29.11.24