**1 задание**

*Какую оплату вы считаете справедливой для выполнения асессором одного микрозадания из этого файла?*

Согласно полученным расчётам цена одного микрозадания должна составлять 17N условных единиц.

*Опишите подробно все этапы вашего решения.*

Первым этапом парсим файл с исходными данными. В процессе написания кода были выявлены аномалии в данных, такие как дата завершения раньше чем дата начала задания. Поэтому на первому этапе мы их отсеяли (как статистическую погрешность).

На втором этапе мы группируем исходные данные по пользователям. В каждой группе производим сортировку данных по возрастанию даты завершения задания.

На третьем этапе для каждой группы делаем разбивку на подгруппы по времени. Делается для того чтобы вычислить периоды работы ассессора с системой. Например, ассессор может взять одновременно n заданий на выполнение, но выполнять он их будет все по очереди. Поэтому эти n заданий будут объединены в подгруппу с минимальным временем начала и максимальным временем выполнения. За указанный период посчитаем количество выполненных ассессором микрозаданий.

С полученными подгруппами переходим на заключительный этап, где рассчитываем среднюю стоимость одного микрозадания. В каждой подгруппе делаем расчёт стоимости одного микрозадания по формуле:

Затем полученные результаты усредняем и получаем справедливую стоимость выполнения одного микрозадания.

*Приложите, пожалуйста, исходный код решения задачи или приведите примеры частей кода в описании решения.*

Весь исходный код представлен в проекте «TaskOne».

**2 задание**

*Используя данные об оценках, установите, какие асессоры хуже всего справились с заданием.*

Следующие 5 ассессоров хуже справились с заданием: 56(uid), 3(uid), 390(uid), 18(uid), 234(uid)

*На какие показатели вы ориентировались и какие метрики вы использовали для ответа на этот вопрос?*

Для ответа на данный вопрос используются метрики корректность ответа и идентификатор документа.

*Можно ли предложить какие-то новые метрики для подсчета качества асессоров с учетом природы оценок у этого бинарного задания?*

Метрик можно собрать много различных, но важно их собирать с условием того на какие вопросы мы хотим получить ответы.

*Опишите подробно все этапы вашего решения.*

Первым этапом парсим файл с исходными данными.

На втором этапе мы группируем исходные данные по идентификаторам документов. Для каждой документа высчитываем процент правильных ответов ассессоров. Исходя из полученного процента выставляем коэффициент сложности документа: [0-30) – коэф. 3, [30-60) – коэф. 2, [60-100] – коэф. 1. Чем меньше людей справилось с выполнением задания, тем сложнее считаем оцениваемый документ.

На третьем этапе рассчитываем для каждого оцениваемого документа его стоимость с учётом коэффициента сложности. После чего группируем результаты по идентификаторам пользователей. Для каждого пользователя рассчитываем среднее значение по оцененным документам.

Для получения результатом сортируем полученную стоимость каждого пользователя.

Среди полученной стоимости выделяется первая пятёрка:

1. 56(uid) => 0,43066

2. 3(uid) => 0,46479

3. 390(uid) => 0,48058

4. 118(uid) => 0,48082

5. 234(uid) => 0,48485

6. 550(uid) => 0,74219

7. 9(uid) => 0,74704

8. 335(uid) => 0,75895

9. 71(uid) => 0,75982

10. 154(uid) => 0,76049