

Assessment Test

Общее описание:

В рамках данного задания мы будем анализировать тексты различных участников рэп-баттлов (преимущественно VERSUS) на предмет различных интересных факторов. Само приложение для анализа должно представлять собой консольную утилиту, которая принимает в себя различные аргументы. В зависимости от них поведение приложения меняется и выдает нам различные интересные факторы.

Данные для анализа необходимо скачать по ссылке:

<https://drive.google.com/file/d/1hz2J73ftRZilHaxLFk-eUjSMIOo1OSNH/view?usp=sharing>

Правила выполнения тестового задания:

1. Выложить выполненное задание на github и поделиться ссылкой на проект;
2. В Readme написать, как правильно запускать ваш файл, например `ruby versus.rb --top-bad-words=3`;
3. Для выполнения всех уровней можно использовать любые библиотеки или API;
4. При выполнении задания на ООП языке следует выделять логику в классы и методы, иначе в функции. Очень желательно избегать названия переменных вида `i, j, k, ans` и тд. Использовать «говорящие» названия.

Условия тестового задания:

Уровень 1

На первом уровне сложности нам нужно проанализировать все тексты и выяснить самого нецензурного участника площадки. У каждого из нас свое понятие нецензурности, поэтому алгоритм определения данного фактора оставляем за вами.

Параметр `--top-bad-words`, который передается в программу, показывает максимальное количество самых нецензурных участников, которое следует вывести на экран. Список «bad words» должен лежать в **yaml** файле.

Дополнительно необходимо будет посчитать общее количество баттлов для каждого участника и среднее количество всех слов в раунде.

Серьезно отнеситесь к форматированию данных. На данном уровне задания оно должно полностью соответствовать примеру ниже.

Пример работы программы:

```
ruby versus.rb --top-bad-words=3
```

Гнойный | 12 батлов | 127 нецензурных слов | 10.58 слова на баттл | 232 слова в раунде |

Охххуmiron | 7 батлов | 24 нецензурных слова | 3.42 слова на баттл | 317 слов в раунде |

Галат | 3 батла | 2 нецензурных слов | 0.66 слова на баттл | 207 слов в раунде |

Уровень 2

На втором уровне сложности мы будем искать для определенного участника его любимые слова в батлах. Имя участника должно передаваться как параметр `--name`. Чтобы ограничить вывод самых популярных слов, стоит использовать параметр `--top-words`. По дефолту необходимо выводить 30 слов.

Не забудьте исключить разные местоимения и прочие мелочи.

Пример работы программы:

```
ruby versus.rb --top-words=20 --name=Толик
```

Рэпер Толик не известен мне. Зато мне известны:

Гнойный

Охххуmiron

Галат

```
ruby versus.rb --top-words=5 --name=Оххуmiron
```

Факты - 5 раз

Папочку - 2 раза

Микрофоны - 1 раз

Птички - 1 раз

Пожертвую - 1 раз