



свою презентацию с рассказа истории, он привлекает наше внимание. Мы любим книги. Мы любим музыку и песни. Мы любим фильмы. Как вы можете использовать истории в науке о данных? В хороших отчетах есть элементы повествования.

Попробуйте рассказать историю с помощью своего анализа

2

## Глава II

### Инструкции

•

Используйте эту страницу в качестве единственной ссылки. Не слушайте никаких слухов и домыслов

о том, как приготовить ваше решение.

•

Здесь и далее мы используем Python 3 как единственно правильную версию Python.

•

Файлы python для упражнений на python (module01, module02, module03) должны иметь блок в конце: `if __name__ == '__main__':`.

•

Обратите внимание на права доступа к вашим файлам и каталогам.

•

Для оценки ваше решение должно находиться в вашем репозитории GIT.

•

Ваши решения будут оценены вашими товарищами-рыбами.

•

Вы не должны оставлять в своем каталоге никаких других файлов, кроме тех, которые явно указаны

в инструкциях по выполнению упражнения. Рекомендуется изменить ваш файл `.gitignore`, чтобы избежать несчастных случаев.

•

Когда вам нужно получить точный результат в ваших программах, запрещается отображать предварительно рассчитанный результат вместо правильного выполнения упражнения.

•

У вас есть вопрос? Спросите своего соседа справа. В противном случае попробуйте обратиться к своему соседу слева.

•

Ваше справочное руководство: mates / Интернет / Google.

•

вы можете задавать вопросы в Slack.

•

Внимательно прочитайте примеры. Они могут требовать вещей, которые иным образом не указаны в предмете.

•

И да пребудет с вами Сила!

3

## Глава III

# Конкретные инструкции дня

- 

Нет кода в глобальной области видимости. Используйте функции!

- 

Любое не перехваченное исключение приведет к аннулированию работы, даже в случае ошибки, которую вам было предложено протестировать

- 

Используемый интерпретатор - Python 3

- 

Разрешена любая встроенная функция

- 

Вы можете импортировать следующие библиотеки: os, sys, urllib, requests, beautifulsoup, json, pytest, collections, functools, datetime, re

- 

Используйте Jupyter Notebook для создания отчета

4

## Глава IV

### Обязательная часть

В этой спешке вы будете работать над своим собственным аналитическим отчетом. Вы будете анализировать данные из базы данных MovieLens. К концу спешки у вас будет

два файла: `movielens_analysis.py`

и `movielens_report.ipynb`. В первом файле вам

нужно будет создать свой собственный модуль с классами и методами. Во втором файле вы создадите сам отчет, используя только свой модуль.

### Модуль

Помните, что цель гонки - укрепить свои навыки.

Постарайтесь использовать как можно больше из того, что вы узнали за предыдущие дни.

- 

Используйте уменьшенную версию набора данных MovieLens,

[загрузите](#)

ее, пожалуйста

- 

Прочтите `README.txt` очень осторожно. Сосредоточьтесь на файловых структурах

- 

В вашем модуле вам нужно будет создать 4 класса, соответствующие 4 файлам из данных, и 1 класс для тестирования

- 

Приведенные ниже классы и методы являются обязательными, но вы можете добавить к ним все, что соответствует вашим потребностям

Оценки классов, теги, фильмы, ссылки можно найти во вложениях

Классные тесты:

Создавайте тесты с помощью PyTest для каждого метода из вышеперечисленных классов.

Они должны проверить:

- 

если методы возвращают правильные типы данных

•

если элементы списков имеют правильные типы данных

5

Rush 00 – Piscine Python для науки о данных

Аналитика MovieLens

•

если возвращенные данные отсортированы правильно

Проведите тесты, прежде чем переходить к следующему этапу спешки.

## Сообщить

Используя только классы и методы из `movielens_analysis.py`,  
подготовьте свой отчет.

Вы должны сделать это в записной книжке Jupyter.

Это отличный инструмент, особенно если вы специалист по обработке данных.

Это дает вам возможность работать с кодом в интерактивном  
режиме, запуская и перезапуская разные ячейки с разными значениями.

Вам не нужно перезапускать весь ваш код с самого начала.

Кроме того, вы можете помещать в ячейки не только код, но и текст,  
что является отличной функцией для создания отчетов.

Установите его в свою среду.

В этой части спешки мы дадим вам больше свободы.

Мы не собираемся определять структуру вашего отчета.

Цель отчета - рассказать нам интересную историю  
о наборе данных MovieLens.

Найдите хорошую структуру и правильную последовательность.

Единственные ограничения:

•

вы должны использовать все без исключения методы из `movielens_analysis.py` кроме классных  
тестов

•

каждая ячейка в вашей записной книжке должна содержать волшебную команду `%timeit`

•

запрещен любой другой импорт, а также использование встроенных функций. Если  
они вам нужны, заранее поместите их в свой модуль

6

## Глава V

## Бонусная часть

•

Добавьте к классам больше методов, которые могут оказаться полезными и интересными для  
вашего

отчета. Не забудьте их тоже протестировать

•

Улучшайте тесты. Проверьте также правильность ваших расчетов. Предварительно  
рассчитайте вручную некоторые результаты и показатели и проверьте, возвращают ли методы  
правильную

информацию, если вы вводите для них конкретные входные данные

7

# Глава VI

## Явка и экспертная оценка

Включите свою работу, используя свой репозиторий git, как обычно.

Во время оценки будет оцениваться только та работа, которая находится в вашем репозитории.

Во время коррекции вас будут оценивать по вашей явке (никаких функций

, которые делают всю тяжелую работу за вас), а также по вашей способности представить, объяснить и обосновать свой выбор.