



Даша: фильм или сериал



Дедлайн сдачи: 14.03.24 23:59

Формат сдачи:

1. тетрадка с расширением

`.ipynb` **или** ссылка на тетрадку в Google Colab с разрешением на комментирование;

2. полученный JSON-файл с отзывами и оценками, названный

`reviews.json`.



Требования к коду:

1. Сохранить HTML-разметку со

всеми отзывами со всех страниц в файл `reviews.html`.

2. Оформить часть со скачиванием и сохранением в файл HTML-разметки как функцию

`get_html()`, которая принимает на вход ту часть ссылки, которая меняется при обращении к разным фильмам (например, сравните ссылки на отзывы к “Oppenheimer”: <https://letterboxd.com/film/oppenheimer-2023/reviews/by/added/>, — и к “The Grand Budapest Hotel”: <https://letterboxd.com/film/the-grand-budapest-hotel/reviews/by/added/>), и количество страниц, с которых нужно выгрузить отзывы.

3. Оформить часть с выделением отзывов и оценок и сохранение их в JSON-файл под названием

`reviews.json` как функцию `get_reviews_stars()`, которая принимает на вход название файла, где находится вся HTML-разметка (в нашем случае `reviews.html`).

4. Следить за PEP-8 и приличным названием переменных 😊

У Вас есть сайт letterboxd.com с отзывами на фильмы и сериалы. Ваши задачи:

1. Выбрать один любой фильм (или сериал), у которого есть побольше отзывов на сайте (например, “[Oppenheimer](#)”).
2. При помощи библиотеки `requests` скачать HTML-разметку как минимум первой страницы с отзывами.

NB! Особенно поощряться будет наличие всех отзывов. В этом Вам может помочь посмотреть на то, как, например, задается ссылка на вторую страницу:

<https://letterboxd.com/film/oppenheimer-2023/reviews/by/added/page/2/>

3. При помощи библиотеки `beautifulsoup` научиться автоматически доставать со страницы **отзыв** и **оценку** (от 0 до 10) каждого пользователя.

NB! В этом поможет класс, начинающийся с `rating -green rated`, например `class="rating -green rated-9"`.

4. Сохранить все отзывы и оценки как список словарей с ключами: `'id'` (номер отзыва в списке), `'text_review'` (полный текст отзыва) и `'rate_review'` (рейтинг фильма или сериала от пользователя, оставившего отзыв). Например:

```
[
    {'id': 1, 'text_review': 'Im Kino mit Freunden...', 'rate_
    {'id': 2, 'text_review': 'Great. Really like...', 'rate_
    ...
]
```

Полученный список сохраните как JSON-файл под названием `reviews.json` (не забудьте про параметры `ensure_ascii` и `indent`).

5. Сдайте в [гугл-форму](#) Ваш код целиком, а также полученный файл `reviews.json`.