ПМ

Домашнее задание «Библиотека Pandas»

 \Box

Задание

🗎 Вопросы по заданию

Преподаватель: Константин Башевой

Домашнее задание

Задание 1

Скачайте с сайта grouplens.org...movielens/ датасет любого размера. Определите какому фильму было выставлено больше всего оценок 5.0.

Задание 2

По данным файла power.csv посчитайте суммарное потребление стран Прибалтики (Латвия, Литва и Эстония) категорий 4, 12 и 21 за период с 2005 по 2010 года. Не учитывайте в расчетах отрицательные значения quantity.

Задание 3

Выберите страницу любого сайта с табличными данными. Импортируйте таблицы в pandas dataframe.

Примеры страниц (необязательно брать именно эти):

https://fortrader.org/quotes www.finanz.ru...om-vremeni

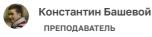
Решение Зачет

LINK

 $https://github.com/Peter M-lab/PYDA1/blob/main/09_pandas/Untitled.ipynb ~~11 \ \text{anp.}~2021$

Вы загрузили решение 11 апр. в 21:00

Константин Башевой поставил(а) зачет 12 апр. в 10:59



12 апреля 2021 10:59

Добрый день, спасибо за решения.

Все верно, здорово сделано.

В задании 2 можно немного сократить код, используя для дат метод between:

data_['year'].between(2005, 2010) .



Вам понравилось?

Вопросы по заданию

Задайте вопрос — вам помогут одногруппники и эксперты

пм

На какой вопрос вы хотите получить ответ?

ПМ

Спросить

Все вопросы Помочь с ответом



Ярослав Гребюк

data[data['rating'] == 5.0] сделал, только не пойму, файл inplace нужно делать, перед дальнейшим join?

Написать ответ

🗎 Ответов пока нет



Г Ярослав Гребюк

фильтр

Через одно место, но вроде получилось?

data = pd.read_csv('ratings.csv')

data.head()

2 rating 100836

sort = data[data['rating'] == 5.0]

sort.to_csv("rating5.csv")

2 rating 13211

join = sort.merge(movies, on='movield', how='left')

join.head()

best_movies = join[join['rating'] == 5]

best_movies['title'].value_counts()

Написать ответ

🖹 Ответов пока нет

СЖ Семетей Жоробеков

метод isin

groupby('item id').count().sort_values Смотрится слишком SQLeвo. value_counts() гораздо приятней. Не понял как приспособить метод isin.

В случае обычного фильтра я получаю понятную таблицу:

country year quantity category

373 Lithuania 2014 39.0 1

374 Lithuania 2013 192.0 1

В случае с isin

0 False

1 False

2 False

3 False

4 False

ПМ

