

Тест начат	Понедельник, 6 ноября 2023, 14:24
Состояние	Завершены
Завершен	Понедельник, 6 ноября 2023, 16:38
Прошло времени	2 час. 13 мин.
Баллы	5,00/5,00
Оценка	20,00 из 20,00 (100%)

Вопрос 1

Верно

Баллов: 1,00 из 1,00

Метод наименьших квадратов

В этом задании вам необходимо реализовать алгоритм МНК (метод наименьших квадратов) в виде функции с использованием библиотеки numpy:

```
def mnk(x_, y_):
    ...
    return a, b
```

Входные данные:

- x_: временной ряд x длиной 30 точек, сгенерированный с помощью функции linspace;
- y_: временной ряд y длиной 30 точек, сгенерированный с помощью функции linspace с коэффициентами a и b.

Написанная функция должна возвращать коэффициенты a и b функции линейной аппроксимации ($y = ax + b$).

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 import numpy as np
2 def mnk(x_, y_):
3     # запишите далее ваш код
4     n = len(x_)
5     x_mean = np.mean(x_)
6     y_mean = np.mean(y_)
7     # Вычисляем коэффициенты наклона (a) и пересечения (b)
8     a = np.sum((x_ - x_mean) * (y_ - y_mean)) / np.sum((x_ - x_mean) ** 2)
9     b = y_mean - a * x_mean
10    return a, b
11
```

	Тест	Ожидается	Получил	
✓	<pre>import numpy as np x = np.linspace(0, 10, 30) y = np.linspace(0, 10, 30)*(10+4*np.random.rand(30)) + (10+4*np.random.rand(30)) a,b = mnk(x,y) y_pred = a*5+b res = np.array([a,b,y_pred]) y_true = 12*5+12 print(np.linalg.norm(res - np.array([12., 12., 72]))) <2.)</pre>	True	True	✓

Все тесты пройдены! ✓

Верно
 Баллы за эту попытку: 1,00/1,00.

Вопрос **2**

Верно

Баллов: 1,00 из 1,00

Анализ информации о клиентах

Вам предложены данные о клиентах, которые хранятся в следующем формате:

	name	city	age
0	Petr	Moscow	41.0
1	Anna	Novosibirsk	28.0

Выполните следующие задания применительно к данным:

- Удалите строки, где есть пропуски.
- Удалите дубликаты, если они имеются в данных.
- Замените нулевые значения средними по столбцу без учета строки нулевого значения.
- Сохраните результаты проделанных действий в датафрейм с именем df

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 import pandas as pd
2 df = pd.read_csv('users.csv', sep=',')
3
4 # далее запишите ваш код
5 df = df.dropna().drop_duplicates()
6 age_mean = df[df['age'] !=0]['age'].mean()
7 df.loc[df['age'] == 0, 'age'] = age_mean
```

	Тест	Ожидается	Получил	
✓	print(df['age'].mean()==33.5)	True	True	✓

Все тесты пройдены! ✓

Верно

Баллы за эту попытку: 1,00/1,00.

Вопрос **3**

Верно

Баллов: 1,00 из 1,00

Анализ электромобилей — 1

Вам предложены данные об электромобилях, которые хранятся в файле [Electric_Car.csv](#).

Выполните следующие действия применительно к данным:

1. Сегрегируйте данные по брендам (поле Brand).
2. Вычислите среднее значение цены (поле PriceEuro) для каждой группы.
3. Сохраните результаты проделанных действий в датафрейм с именем Carsgroupby.

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 | import pandas as pd
2 | Cars = pd.read_csv('Electric_Car.csv')
3 |
4 | # далее напишите ваш код
5 | Carsgroupby = Cars.groupby('Brand')['PriceEuro'].mean().reset_index()
```

	Тест	Ожидается	Получил	
✓	print(Carsgroupby['PriceEuro'].sum())==1706238.541025641)	True	True	✓

Все тесты пройдены! ✓

Верно
Баллы за эту попытку: 1,00/1,00.

Вопрос **4**

Верно

Баллов: 1,00 из 1,00

Анализ электромобилей — 2

Вам предложены данные об электромобилях, которые хранятся в файле [Electric_Car.csv](#).

Выполните следующие действия применительно к данным:

1. Сделайте срез по столбцу PriceEuro со значениями более 50000 и по столбцу TopSpeed_KmH со значениями более 200.
2. Переназначьте индексы таблицы начиная с нуля.
3. Сохраните результаты проделанных действий в датафрейм с именем Cars_speed.

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 | import pandas as pd
2 | Cars = pd.read_csv('Electric_Car.csv')
3 |
4 | # далее запишите ваш код
5 | Cars_speed = Cars[(Cars['PriceEuro'] > 50000) & (Cars['TopSpeed_KmH'] > 200)].reset_index()
```

	Тест	Ожидается	Пол
✔	<pre>import numpy as np Cars_speed_prove = Cars.loc[(Cars.PriceEuro > 50000) & (Cars.TopSpeed_KmH > 200)].reset_index() print(np.sum(Cars_speed.PriceEuro)+np.sum(Cars_speed.TopSpeed_KmH) == np.sum(Cars_speed_prove.PriceEuro)+np.sum(Cars_speed_prove.TopSpeed_KmH))</pre>	True	True

Все тесты пройдены! ✔

Верно

Баллы за эту попытку: 1,00/1,00.

Вопрос **5**

Верно

Баллов: 1,00 из 1,00

Анализ электромобилей — 3

Вам предложены данные, которые хранятся в датафреймах [EC.csv](#) , [EVP.csv](#).

Вам необходимо:

1. Выполнить операцию слияния таблиц по столбцу Brand и сохранить результат в переменную integral.
2. Получить описательные статистики для полученной таблицы integral с помощью функции describe().
3. Сохранить результат вывода описательных статистик в переменную describe_pd.

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 import pandas as pd
2 import numpy as np
3 EC = pd.read_csv('EC.csv', sep=',')
4 EVP = pd.read_csv('EVP.csv', sep=',')
5
6 # далее запишите ваш код
7 EC['Brand'] = EC['Brand'].str.replace(' ', '')
8 EVP['Brand'] = EVP['Brand'].str.replace(' ', '')
9 integral = EVP.merge(EC.sort_values(['Brand', 'PriceEuro']).drop_duplicates('Brand'), on
10 describe_pd = integral.describe()
```

	Тест	Ожидается	Получил	
✔	import pandas as pd import numpy as np EC = pd.read_csv('EC.csv', sep=',') EVP = pd.read_csv('EVP.csv', sep=',') integral_1 = EC.merge(EVP, how='outer', on='Brand') describe_pd_1 = integral_1.describe() print(describe_pd_1.sum().count() == describe_pd.sum().count())	True	True	✔

Все тесты пройдены! ✔

Верно
Баллы за эту попытку: 1,00/1,00.