

Тест начат	Понедельник, 4 декабря 2023, 17:07
Состояние	Завершены
Завершен	Понедельник, 4 декабря 2023, 17:10
Прошло времени	3 мин.
Баллы	16,00/16,00
Оценка	30,00 из 30,00 (100%)

Вопрос
 Инфо

Итоговые задания модуля выполняются на основе датасета [Supermarket sales](#), который содержит о продажах в трех разных супермаркетах.

Мы используем в задании [файл supermarket.csv](#) , который уже подгружен в ЛМС.

В некоторых заданиях мы попросим вас привести код скрипта, а в некоторых — ответить на вопрос датасету.

При выполнении заданий рекомендуем:

- скачать датасет и изучить представленные в нем данные;
- все задания выполнять в самостоятельно созданном Python-ноутбуке, а в LMS вносить требуемый результат;
- при внесении кода в LMS, не забывайте **импортировать** для его выполнения **библиотеки** (если это не сделано в окне ввода кода);
- если в задании требуется вывести на экран результат работы, используйте команду `print`

Удачи в выполнении заданий!

Вопрос **1**

Верно

Баллов: 1,00 из 1,00

Первоначальная работа с датафреймом Supermarket s

Это задание выполняется на основе датасета [Supermarket sales](#), который содержит данные о пр трех разных супермаркетах.

В этом задании вам необходимо написать скрипт, который сделает следующее:

- Поля Date и Time в основном датафрейме df будут приведены к типу datetime;
- Выведет на экран True, если в датасете есть пропуски, и False, если пропусков нет.

В качестве ответа загрузите ваш код программы на Python.

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 import pandas as pd
2 df = pd.read_csv('supermarket.csv', sep=',')
3
4 # Далее запишите ваш код
5 df["Date"] = pd.to_datetime(df['Date'])
6 df["Time"] = pd.to_datetime(df['Time'])
7 if df.isna().sum().sum() == 0:
8     print(False)
9 else:
10    print(True)
```

	Тест	Ожидается	Получил
✓	print()	False	False

	Тест	Ожидается	Получил
✔	print(df.info())	<class 'pandas.core.frame.DataFrame'> RangeIndex: 1000 entries, 0 to 999 Data columns (total 17 columns): # Column Non-Null Count Dtype --- - 0 Invoice ID 1000 non-null object 1 Branch 1000 non-null object 2 City 1000 non-null object 3 Customer type 1000 non-null object 4 Gender 1000 non-null object 5 Product line 1000 non-null object 6 Unit price 1000 non-null float64 7 Quantity 1000 non-null int64 8 Tax 5% 1000 non-null float64 9 Total 1000 non-null float64 10 Date 1000 non-null datetime64[ns] 11 Time 1000 non-null datetime64[ns] 12 Payment 1000 non-null object 13 cogs 1000 non-null float64 14 gross margin percentage 1000 non-null float64 15 gross income 1000 non-null float64 16 Rating 1000 non-null float64 dtypes: datetime64[ns](2), float64(7), int64(1), object(7) memory usage: 132.9+ KB None	<class 'pandas.core.frame.DataFrame' RangeIndex: 1000 entries, 0 to 999 Data columns (total 17 columns): # Column Non-Null Count Dtype --- - 0 Invoice ID 1000 non-null object 1 Branch 1000 non-null object 2 City 1000 non-null object 3 Customer type 1000 non-null object 4 Gender 1000 non-null object 5 Product line 1000 non-null object 6 Unit price 1000 non-null float64 7 Quantity 1000 non-null int64 8 Tax 5% 1000 non-null float64 9 Total 1000 non-null float64 10 Date 1000 non-null datetime64[n 11 Time 1000 non-null datetime64[n 12 Payment 1000 non-null object 13 cogs 1000 non-null float64 14 gross margin percentage 1000 non-null float64 15 gross income 1000 non-null float64 16 Rating 1000 non-null float64 dtypes: datetime64[ns](2), float64(7), int64(1), object memory usage: 132.9+ KB None

Все тесты пройдены! ✔

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

Вопрос 2

Верно

Баллов: 5,00 из 5,00

Вычисление общего дохода по дням и неделям

Задание 1. Напишите скрипт, который будет вычислять общий доход в разбивке по местополож супермаркета (City) и в динамике. Чтобы отследить динамику изменения дохода (gross income) и график, отражающий суммарный доход в разрезе каждого магазина по датам.

Дайте ответы на вопросы, приведенные ниже.

Важно: для ввода ответа используйте строчные символы.

Вопрос 1. Запишите в поля ввода суммы дохода для каждого местоположения магазина. В качестве разделителя дробной и целой части используйте символ точки «.».

Mandalay	<input type="text" value="5057.0320"/>	✓
Naupytaw	<input type="text" value="5265.1765"/>	✓
Yangon	<input type="text" value="5057.1605"/>	✓

Вопрос 2. Проанализируйте динамику суммарного изменения дохода (сумму поля gross income на графике. Почему мы не можем использовать графическое представление изменения дохода для анализа данных?

✓

Задание 2. Используя ранее выполненные задания, проанализируйте динамику изменения сум по неделям для каждого расположения магазина.

Вопрос. Какой город имел максимальную недельную прибыль за рассмотренный в датасете период?

✓

Вопрос 3

Верно

Баллов: 7,00 из 7,00

Исследование характеристик и зависимостей датасета

Используя ранее выполненные задания, исследуйте датасет и дайте ответы на вопросы.

Важно:

- для ввода ответа используйте строчные символы;
- если в ответе требуется указать название колонки или значение поля, запишите его точно так, как название записано в исходных данных.

Вопрос 1. Изучите количество купленных товаров по типу. Запишите в поле тип товара, который вы всего купили.

✓

Вопрос 2. Проанализируйте нормализованную долю покупателей-членов бонусной программы. Результат округлите до трех знаков после запятой. Запишите в поля ввода долю каждой группы. В качестве разделителя дробной и целой части используйте символ точки «.».

Member	<input type="text" value="0.501"/>	✓
Normal	<input type="text" value="0.499"/>	✓

Вопрос 3. Определите и запишите в поля для ввода средний чек по каждому платежному методу. Результат округлите до двух знаков после запятой. В качестве разделителя дробной и целой части используйте символ точки «.».

Cash	<input type="text" value="326.18"/>	✓
Credit card	<input type="text" value="324.01"/>	✓
Ewallet	<input type="text" value="318.82"/>	✓

Вопрос 4. Проанализируйте динамику среднего чека по каждому платежному методу по неделям. Запишите в поле ввода название платежного метода, имеющего наименьший средний чек на 3-й неделе.

✓

Вопрос **4**
Верно
Баллов: 2,00 из 2,00

Анализ признаков и геометрические вычисления для датасета Electric cars

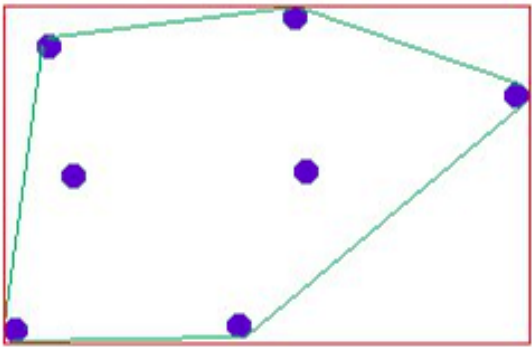
В этом задании вы будете работать с датасетом, содержащим информацию об электромобилях [cars.csv](#).

Если вы занимаетесь разработкой модели анализа перспективности сегментов рынка или оценки страховки автомобилей, то вам может пригодиться проведение парного анализа признаков и возможности создания новых удобных признаков на основе анализа. В этом случае полезным будет предварительный анализ возможных комбинированных производных признаков, таких как площадь и другие геометрические фигуры.

Что необходимо сделать:

1. Загрузить датасет и проанализировать его структуру.
2. Построить график зависимости значений столбца PriceEuro от TopSpeed_KmH с помощью функции scatter библиотеки matplotlib.
3. Подсчитать визуально количество вершин описывающего *многоугольника*.
4. Определить, какую площадь имеет описывающий *прямоугольник*.

Описывающий многоугольник — выпуклый многоугольник, вершинами которого являются диаграммы, при этом все остальные вершины находятся внутри многоугольника. На рисунке вы видите пример построения описывающего многоугольника (показан синим цветом) и описывающего прямоугольника (показан красным цветом).



После выполнения задания введите полученные результаты в поле ввода и ответьте на вопрос

Вопрос 1. Какое количество вершин имеет описывающий многоугольник?

7 ✓

Вопрос 2. Какую площадь имеет описывающий прямоугольник?

Площадь описывающего многоугольника лежит в диапазоне 70 000 000 – 90 000 000 ✓

Вопрос **5**
Верно
Баллов: 1,00 из 1,00

Категоризация данных датасета Electric cars

Напишите программу, выполняющую категоризацию данных датасета [Electric cars.csv](#).

- 1.Используя модуль LabelEncoder из библиотеки sklearn, выполните кодирование категория признака BodyStyle.
- 2.Найдите сумму полученных значений категориального признака и сохраните результат в переменную sum.
- 3.Найдите количество уникальных категорий полученного признака и сохраните результат в переменную cat_count.

Ответ: (штрафной режим: 0%)

Сбросить ответ

```
1 import numpy as np
2 import pandas as pd
3
4 df = pd.read_csv('Electric_Car.csv')
5
6 # Далее запишите ваш код
7 from sklearn.preprocessing import LabelEncoder
8 label_encoder = LabelEncoder()
9 df['BodyStyle'] = label_encoder.fit_transform(df['BodyStyle'])
10 sum = df['BodyStyle'].sum()
11 cat_count = df['BodyStyle'].nunique()
```

	Тест	Ожидается	Получил	
✓	print(sum+cat_count==429)	True	True	✓

Все тесты пройдены! ✓

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