** Preliminary Midterm Exam II, Spring - 2024**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **№1**  **(1)** | **№2**  **(2)** | **№3**  **(2)** | **№4**  **(2)** | **№5**  **(3)** | **Total (10)** |
|  |  |  |  |  |  |

**Анализ и визуализация данных 4**

**ВАЖНО!** Названия столбцов в данном задании и в самих данных могут отличаться. На E-COURSE необходимо загрузить 2 файла: Jupyter Notebook (.ipynb) и отчет(.docx) по проведенному анализу.

**№1** (1б)

1. Загрузить необходимые библиотеки
2. Загрузить датасет. Необходимо сразу считать столбец с датами и использовать его в качестве столбца индексации.
3. Вывести первые 10 строк датасета
4. Вывести информацию по каждому столбцу.
5. Вывести описательную статистику датасета
6. Создать столбцы Day, Month, Year взяв для них данные из индексного столбца.
7. Создать столбец Total, для которого необходимо умножить столбец Unit\_price на Quantity. Необходимо предварительно преобразовать столбец Unit\_price в числовой, убрав из него обозначения валюты.

**№2** (2б)

1. Агрегировать данные по месяцам. В качестве функции агрегирования взять сумму. Отдельно построить линейный график по столбцу «Quantity».
2. Агрегировать данные по неделям, взяв среднее. Отдельно построить линейный график по столбцу Total.
3. Агрегировать данные по неделям, взяв среднее. На агрегированных данных применить скользящее среднее, задав окно, равное 4. Агрегировать по среднему. Отдельно построить линейный график по столбцу Unit\_price.
4. Агрегировать данные по месяцам, взяв сумму. На агрегированных данных применить скользящее среднее, задав окно, равное 2. Агрегировать по среднему значению. Отдельно построить линейный график к 13 заданию, отобразив столбец Total.
5. Сгруппируйте данные по дням недели, взяв среднее. Отдельно отобразите линейный график по столбцу Quantity. Дни недели выразить буквенными обозначениями.
6. Сгруппируйте данные по месяцам, взяв сумму. Отдельно отобразите линейный график по столбцу Total. Названия месяцев выразите буквенными обозначениями.

**№ 3** (2б)

1. Создайте новый датафрейм, в котором будут храниться данные за 2021 год.
2. Покажите на графике, есть ли разница в количестве продаж(Quantity) в зависимости от деления на будние и выходные дни.
3. Покажите графически разницу в сумме продаж (Total) в разное время суток в выходные и будние дни.

**№4** (2б)

1. Агрегируйте данные по неделям, взяв среднее. Поместите результат в отдельный датафрейм.
2. Создайте столбец «Previous week average sales», для которого необходимо сдвинуть столбец Total на 1 неделю назад.
3. Создать столбец 1 week change, который отражает разницу в цене между Total и Previous week average sales.

**№5** (3б)

Написать подробный отчет, который в обязательном порядке должен содержать:

1. Описание каждого проделанного задания
2. Анализ полученных результатов в каждом задании
3. В конце отчета также необходимо сформировать общий вывод на основании всех полученных результатов.