1. Задан двухмерный массив ar1 размерности (10, 10), состоящий из случайных целых чисел в пределах от 0 до 15. Вычислить разность s\_odd - s\_even, где s\_odd - сумма элементов, стоящих на позиции (x, y), где (x + y) является нечетным числом; s\_even - сумма элементов, стоящих на позиции (x, y), где (x + y) является четным числом.

Решить задачу средствами numpy и/или pandas. Не использовать циклы и конструкции стандартного Python там, где можно использовать возможности данных библиотек

+

2. Датасет: Chinook\_Sqlite.sqlite +

С помощью кода на Python с использованием sqlite3 и SQL решить задачу. Реализовать функции на Python:

1. Которая возвращает все имеющиеся жанры.

2. Которая возвращает ID жанров, в которых написано более 100 треков, и их (жанров) название.

3. Датасет: album\_artist.xlsx +

С помощью кода на Python с использованием xlwings решить задачу. Вынести названия артистов на отдельный лист "Артисты" и присвоить каждому артисту уникальный идентификатор. Заменить названия артистов на исходном листе на идентификаторы с листа "Артисты".

4. Датасет: us-countries.csv +

Создайте таблицу, где по строкам располагаются названия штатов, по столбцам - каждый из 12 месяцев 2020 года, а в ячейках таблицы хранится суммарное количество смертей в данном штате в этот месяц. Если информация за какой-то месяц отсутствует, укажите в этой ячейке значение 0.

5. По данным из файла data/meals.json сформировать словарь, в котором по идентификатору блюда можно получить список ингредиентов. +

6. Датасет: Womens Clothing E-Commerce Reviews.csv +

Для каждого уникального значения в столбце Division Name найти топ-5 самых часто используемых слов в описании отзыва. Исключить из рассмотрения стоп-слова.

7. Датасет: people.\*.csv ------

В people.\*.csv найти номера карт, сумма цифр в которых кратна 7. Выполнить задание с использованием Dask, распараллелив процесс обработки данных. Выполнить задание с использованием Dask (корректным!), распараллелив процесс обработки данных (использование Dask должно приводить к истинной параллельной обработке данных).

8. Подсчитать, сколько раз во всех текстовых файлах, лежащих в каталоге fish, встречаются слова, начинающиеся с прописной буквы. Выполнить задание с использованием Dask, распараллелив процесс обработки данных. Выполнить задание с использованием Dask (корректным!), распараллелив процесс обработки данных (использование Dask должно приводить к истинной параллельной обработке данных).

9. Сгенерируйте 10 строк длиной в 10000 символов, состоящих из маленьких английских букв. Посчитайте, сколько раз в этих 10 строках встречается шаблон "x?y?z", где знак вопроса обозначает один любой символ. Решение этой задачи распараллелить, используя multiprocessing Pool. Сравнить продолжительность последовательного и параллельного решения задачи.

10. Создайте dask.array размерности 10 тыс на 5, заполненный случайными целыми числами на отрезке [0, 10]. Создайте версию этого массива, оставив только строки, в которых есть число 7