

Лабораторная работа № 1

Современные подходы к обработке и хранению больших данных. Формы представления данных, типы и виды данных.

Задание:

1. Установить программный комплекс Anaconda 3 и браузер БД SQLite3:
 - a. <https://www.anaconda.com/products/individual>
 - b. <https://sqlitestudio.pl/>
2. Изучить следующий сайт для освоения основ языка программирования Python:
 - a. <https://learnxinyminutes.com/docs/ru-ru/python-ru/>
3. Создав «блокнот» в Anaconda 3, произвести получение и обработку файлов из папки с вариантом, используя библиотеку **pandas**.
4. Ответить на вопросы при защите лабораторной работе.
Тематики:
 - a. Общий принцип работы составленного блокнота
 - b. Вопрос по исходному тексту
 - c. Теоретический вопрос по языку Python, одному из форматов, с которым работали в Л/р, или вопрос из первой лекции.

Группа № 1

- 1) CSV: Подсчёт числа погибших в каждой стране за март.
- 2) SQLite: Подсчитать общее число отзывов, выяснить, какого типа отзывов больше: положительных, негативных или нейтральных.
- 3) JSON: Показать тенденцию изменения полового состава населения Албании.

Группа № 2

- 1) CSV: Подсчёт числа тех сотрудников, у кого оплата переработки (Overtime Pay) превысила базовую ставку (Base Pay).
- 2) SQLite: Вывести число «шедевров» (балл в review более 9.5) и «недоразумений» (балл более 0 и менее 1).
- 3) JSON: Показать тенденцию изменения детского населения в Дании

Группа № 3

- 1) CSV: Показать тексты публикаций от 23 февраля 2015 г.
- 2) SQLite: Показать имена и должность тех, у кого оплата составила более 300.000 долларов
- 3) JSON: Показать график показателей импорта и экспорта стран-экспортёров (считать таковой, если в 2010 г. процент экспорта превысил процент импорта в долях ВВП), страны называются на букву “В”.

Группа № 4

- 1) CSV: Подсчёт числа выздоровевших в каждой стране за февраль. Вывод в формате «страна-число».
- 2) SQLite: Показать уровень продаж за ноябрь, число проданных билетов и среднюю цену за билет
- 3) JSON: Показать уровень фертильности стран, названия которых начинаются на “Г”.

Группа № 5

- 4) CSV: Показать уникальные названия авиакомпаний.
- 5) SQLite: Показать среднюю ставку административного аналитика.
- 6) JSON: Показать тенденцию изменения населения Сербии.

Группа № 6

- 7) CSV: Вывод числа подтверждённых случаев 10 мая 2020 г. во всех странах. Выделить те страны, где обнаружено менее 10000 случаев.
- 8) SQLite: Кто зарабатывает больше: те, кого зовут Yvonne или Yvette?
- 9) JSON: Показать изменение погибших и пострадавших в ДТП в США.