

ДОПОЛНИТЕЛЬНЫЕ МАТЕРИАЛЫ К ЗАНЯТИЮ 0.2.

ЧТО ДЕЛАТЬ ДАЛЬШЕ?

Теория

Разберитесь в основах Книга [Python Data Science Handbook](#) от O'Reilly. Там нужны главы:

- [Introduction to NumPy](#);
- [Data Manipulation with Pandas](#).

Обращайтесь к первоисточнику У библиотек Питона хорошая документация, в которой в случае необходимости можно найти нужные ответы:

- [Numpy](#);
- [Scipy](#);
- [pandas](#).

Быстрые ответы Иногда возникает вопрос, искать ответ на который по всей документации неудобно. Поэтому вот видео-ответы на некоторые (в плейлисте на Youtube есть ещё):

- [How do I find and remove duplicate rows in pandas?](#)
- [How do I create a pandas DataFrame from another object?](#)
- [How do I apply a function to a pandas Series or DataFrame?](#)

Быстрое напоминание Чтобы быстро вспомнить нужную конструкцию, вот несколько шпаргалок:

- [NumPy + SciPy + Pandas](#);
- [Numpy](#);
- [Pandas](#);
- [SciPy](#).

Практика

Эти задания нужны, чтобы вы руками написали достаточно много кода, зная при этом, какой ответ ожидать.

Тренируйте навык Создайте новый Jupyter Notebook и пошагово повторите там действия из [10 Minutes to pandas](#) — официального введения в библиотеку.

Тренируйте навык

[Номер 1](#). Импортируйте библиотеки, создайте датасет, датафрейм, прочитайте из CSV, найдите максимумы, визуализируйте.

[Номер 2](#). Прочитайте из TXT и экспортируйте туда, выберете нижние/верхние записи, сгруппируйте, отсортируйте данные.

[Номер 3](#). Создайте функции. Прочитайте из Excel и экспортируйте туда,

Outliers, Лямбда-функции, нарежьте данные.

Номер 4. Добавьте/удалите столбцы, делайте операции с индексами.

Тренируйте навык Большой ноутбук [100-pandas-puzzles](#). Хотя заданий там 60.

Тренируйте в контексте Чтобы тема лучше усвоилась, нужно не только много раз повторить действие, но и понимать зачем — в каком контексте и для чего это можно применить. Поэтому советуем поработать немного с ноутбуками из этого [репозитория](#). Внутри папок «Гитхаба» лежат задания на такие навыки:

- [Getting & Knowing Your Data](#);
- [Filtering & Sorting](#);
- [Grouping](#);
- [Apply](#);

- Merge;
- Stats;
- Visualization;
- Creating Series and DataFrames;
- Time Series.

Надеемся, файл был полезным. Будем рады обратной связи. Какие материалы вам нужны и интересно получать?

Если не работают ссылки, есть идеи, или вы нашли материал, который мы упустили, напишите координатору программы Андрею.