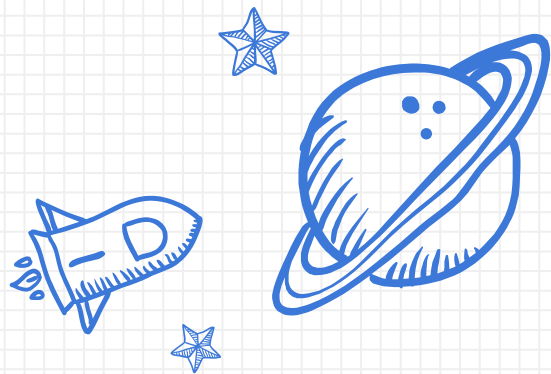# Data Analysis with Pandas and Python

Nice to meet you!

Boris Paskhaver

- Software Engineer/Consultant
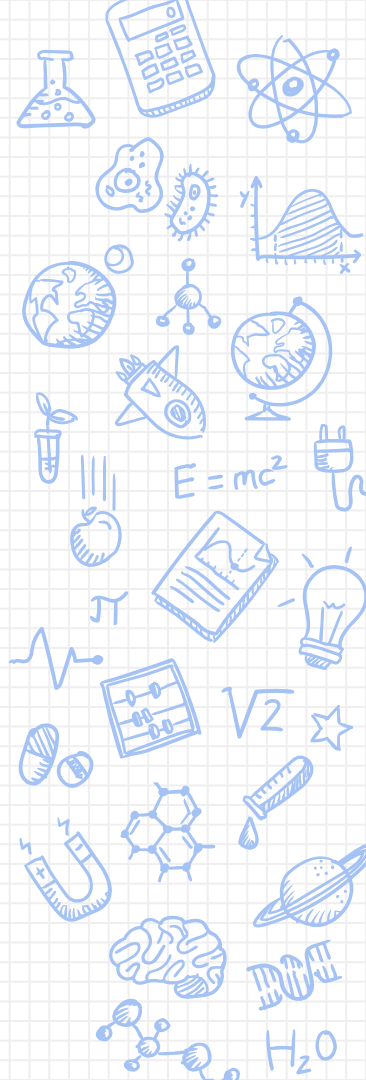- Teacher on Udemy.com
- Author of Pandas in Action

Data Analysis

The process of deriving an insight
from a dataset

What is pandas?

✘ Data analysis library built on top of the Python programming language

✘ A robust toolkit for analyzing, filtering, manipulating, aggregating, merging, pivoting, and cleaning data

What is pandas?

✘   Pandas is not named after the furry animal.
Its name is short after "panel data structures",
a term that refers to "data sets with
observations over time".

✘   "Excel for Python" or "Excel on steroids".

Prerequisites

Required

✘ Basic knowledge of spreadsheet software

✘ Basic knowledge of data types (text/strings, integers, floats, Booleans)

Recommended

✘ Intermediate knowledge of spreadsheet software (pivot tables, VLOOKUP, functions)

✘ Basic understanding of Python (variables, functions, lists)

About this Course

✘ Split up into sections, each focusing on a specific feature or data structure

✘ Bundled with datasets

Installing Pandas and Python

✘ The easiest setup process is installing the **Anaconda distribution**, a bundle of Python, pandas, and 100+ data analysis libraries