DataCamp
We're hiring!

PAID COURSE

Importing Data in Python (Part 1)

Replay Course

3 hours | 15 Videos | 54 Exercises | 19,277 Participants | 4150 XP

STATEMENT OF ACCOMPLISHMENT

Download

This course is part of these tracks:

Data Analyst with Python

Data Scientist with Python

Importing & Cleaning Data with Python

Python Developer

Hugo Bowne-Anderson
Data Scientist at DataCamp

Hugo hearts all things Pythonic and is charged with building out DataCamp's Python curriculum. He can be found at hackathons, meetups & code sprints, primarily in NYC. Before joining the ranks of

DataCamp, he worked in applied mathematics (biology) research at Yale University.

See More

COLLABORATOR(S)

 **Francisco Castro**

PREREQUISITES

- Intro to Python for Data Science
- Intermediate Python for Data Science

DATASETS

- Chinook (SQLite)
- LIGO (HDF5)
- Battledeath (XLSX)
- Extent of infectious diseases (DTA)
- Gene expressions (MATLAB)
- MNIST
- Sales (SAS7BDAT)
- Seaslugs
- Titanic

Course Description

As a Data Scientist, on a daily basis you will need to clean data, wrangle and munge it, visualize it, build predictive models and interpret these models. Before doing any of these, however, you will need to know how to get data into Python. In this course, you'll learn the many ways to import data into Python: (i) from flat files such as .txts and .csvs; (ii) from files native to other software such as Excel spreadsheets, Stata, SAS and MATLAB files; (iii) from relational databases such as SQLite & PostgreSQL.

1 Introduction and flat files FREE 100%

In this chapter, you'll learn how to import data into Python from all types of flat files, a simple and

prevalent form of data storage. You've previously learned how to use NumPy and pandas - you will learn how to use these packages to import flat files, as well as how to customize your imports.

Welcome to the course!	50 xp
Exploring your working directory	50 xp
Importing entire text files	100 xp
Importing text files line by line	100 xp
The importance of flat files in data science	50 xp
Pop quiz: examples of flat files	50 xp
Pop quiz: what exactly are flat files?	50 xp
Why we like flat files and the Zen of Python	50 xp
Importing flat files using NumPy	50 xp
Using NumPy to import flat files	100 xp
Customizing your NumPy import	100 xp
Importing different datatypes	100 xp
Working with mixed datatypes (1)	50 xp
Working with mixed datatypes (2)	100 xp
Importing flat files using pandas	50 xp
Using pandas to import flat files as DataFrames (1)	100 xp
Using pandas to import flat files as DataFrames (2)	100 xp
Customizing your pandas import	100 xp
Final thoughts on data import	50 xp

HIDE CHAPTER DETAILS Completed

2 Importing data from other file types 100%

You've learned how to import flat files, but there are many other file types you will potentially have to work with as a data scientist. In this chapter, you'll learn how to import data into Python from a wide array of important file types. You will be importing file types such as pickled files, Excel spreadsheets, SAS and Stata files, HDF5 files, a file type for storing large quantities of numerical data, and MATLAB files.

Introduction to other file types	50 xp
Not so flat any more	50 xp
Loading a pickled file	100 xp
Listing sheets in Excel files	100 xp
Importing sheets from Excel files	100 xp
Customizing your spreadsheet import	100 xp
Importing SAS/Stata files using pandas	50 xp
How to import SAS7BDAT	50 xp
Importing SAS files	100 xp
Using read_stata to import Stata files	50 xp
Importing Stata files	100 xp
Importing HDF5 files	50 xp
Using File to import HDF5 files	50 xp
Using h5py to import HDF5 files	100 xp
Extracting data from your HDF5 file	100 xp
Importing MATLAB files	50 xp
Loading .mat files	100 xp
The structure of .mat in Python	100 xp

HIDE CHAPTER DETAILS Completed

3 Working with relational databases in Python 100%

In this chapter, you'll learn how to extract meaningful data from relational databases, an essential element of any data scientist's toolkit. You will be learning about the relational model, creating SQL queries, filtering and ordering your SQL records, and advanced querying by JOINing database tables.

Introduction to relational databases	50 xp
Pop quiz: The relational model	50 xp
Creating a database engine in Python	50 xp
Creating a database engine	100 xp
What are the tables in the database?	100 xp
Querying relational databases in Python	50 xp
The Hello World of SQL Queries!	100 xp
Customizing the Hello World of SQL Queries	100 xp
Filtering your database records using SQL's WHERE	100 xp
Ordering your SQL records with ORDER BY	100 xp
Querying relational databases directly with pandas	50 xp
Pandas and The Hello World of SQL Queries!	100 xp
Pandas for more complex querying	100 xp
Advanced Querying: exploiting table relationships	50 xp
The power of SQL lies in relationships between tables: INNER JOIN	100 xp
Filtering your INNER JOIN	100 xp
Final Thoughts	50 xp

HIDE CHAPTER DETAILS Completed