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

0%

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.

VIEW CHAPTER DETAILS

Continue Chapter

2 Importing data from other file types

0%

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.

VIEW CHAPTER DETAILS

Continue Chapter

3 Working with relational databases in Python

0%

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.

VIEW CHAPTER DETAILS