

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 the prequel to this course, you have already learnt 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. In this course, you'll extend this knowledge base by learning to import data (i) from the web and (ii) a special and essential case of this: pulling data from Application Programming Interfaces, also known as APIs, such as the Twitter streaming API, which allows us to stream real-time tweets.

1 Importing data from the Internet **FREE**

0%

The web is a rich source of data from which you can extract various types of insights and findings. In this chapter, you will learn how to get data from the web, whether it be stored in files or in HTML. You'll also learn the basics of scraping and parsing web data.

[VIEW CHAPTER DETAILS](#)

[Continue Chapter](#)

2 Interacting with APIs to import data from the web

0%

In this chapter, you will push further on your knowledge of importing data from the web. You will learn the basics of extracting data from APIs, gain insight on the importance of APIs and practice getting data from them with dives into the OMDB and Library of Congress APIs.

[VIEW CHAPTER DETAILS](#)

[Continue Chapter](#)

3 Diving deep into the Twitter API

0%

In this chapter, you will consolidate your knowledge of interacting with APIs in a deep dive into the Twitter streaming API. You'll learn how to stream real-time Twitter data and to

analyze and visualize it!

[VIEW CHAPTER DETAILS](#)