



Kurtis Pykes Machine Learning Engineer

pandas is a widely used Python library for data science, analysis, and machine learning. It offers a flexible and intuitive way to handle data sets of all sizes. One of the most important functionalities of pandas is the tools it provides for reading and writing data.

For data available in a tabular format and stored as a CSV file, you can use pandas to read it into memory using the read_csv() function, which returns a pandas dataframe.

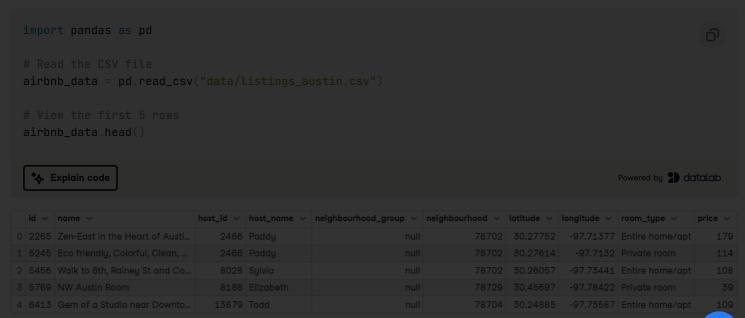
In this article, you will learn all about the read_csv() function and how to alter the parameters to customize the output. We will also cover how to write pandas dataframe to a CSV file.

Note: Check out this DataLab workbook to follow along with the code.

Importing a CSV file using the read_csv() function

Before reading a CSV file into a pandas dataframe, you should have some insight into what the data contains. Thus, it's recommended you skim the file before attempting to load it into memory: this will give you more insight into what columns are required and which ones can be discarded.

Now, let's write some code to import a file using read_csv() . Then, we can talk about what's going on and how we can customize the output we receive while reading the data into memory.



All that has gone on in the code above is we have:

