TUTORIAL

pandas read_csv() T ata

Python

Importing data is the first step in any dat a scientists prefer the pandas read_csv() function to do this.

December 2024 9 min read

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.

```
import pandas as pd

# Read the CSV file
airbnb_data = pd.read_csv("data/listings_austin.csv")

# View the first 5 rows
airbnb_data.head()
```

Explain code

Powered by datalab

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price
0	2265	Zen-East in the Heart of Austi...	2466	Paddy	null	78702	30.27752	-97.71377	Entire home/apt	179
1	5245	Eco friendly, Colorful, Clean, ...	2466	Paddy	null	78702	30.27614	-97.7132	Private room	114
2	5456	Walk to 6th, Rainey St and Co...	8028	Sylvia	null	78702	30.26057	-97.73441	Entire home/apt	108
3	5769	NW Austin Room	8186	Elizabeth	null	78729	30.45697	-97.78422	Private room	39
4	6413	Gem of a Studio near Downto...	13879	Todd	null	78704	30.24885	-97.73587	Entire home/apt	109

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

1. Imported the pandas library into our environment