How can I ignore the first column when importing a CSV file in Pandas?

Authored by

stats writer

June 25, 2024

When importing a CSV file in Pandas, it is possible to ignore the first column by using the "usecols" parameter and specifying the columns to be imported. This allows for more efficient data manipulation and analysis, as the unnecessary first column is not included in the data set. This can be particularly useful when dealing with large CSV files with many columns, as it reduces the amount of memory and processing power required.

Pandas: Ignore First Column when Importing CSV File

You can use the following basic syntax to ignore the first column when importing a CSV file into a pandas DataFrame:

```
withopen('basketball_data.csv') as x:
ncols = len(x.readline().split(','))

df = pd.read_csv('basketball_data.csv', usecols=range(1,ncols))
```

This particular example will read each column from a CSV file called basketball_data.csv into a pandas DataFrame except for the first column.

Using this code, we first find the number of columns in the CSV file and assign it to a variable called ncols.

Then we use the usecols argument to specify that we only want to import the columns in the range from 1 (i.e.

the second column) to the last column of the CSV file.

The following example shows how to use this syntax in practice.

Example: Ignore First Column when Importing CSV File in Pandas

Suppose we have the following CSV file called basketball_data.csv:

```
team,points,rebounds
A, 22, 10
B, 14, 9
C, 29, 6
D, 30, 2
```


We can use the following syntax to import the CSV file into a pandas DataFrame and ignore the first column:

```
import pandas as pd

#calculate number of columns in CSV file
```

```
with open('basketball_data.csv') as x:
ncols = len(x.readline().split(','))

#import all columns except first column into DataFrame
df = pd.read_csv('basketball_data.csv',
usecols=range(1,ncols))

#view resulting DataFrame
print(df)

points rebounds
0 22 10
1 14 9
2 29 6
3 30 2
```

Notice that the first column called team was dropped when we imported the CSV file into pandas.

Note that if you already know the total number of columns in the CSV file ahead of time, you can directly supply that value to the usecols argument.

For example, suppose we already knew that there were three columns in the CVS file.

We could use the following syntax to import the CSV file into a pandas DataFrame and ignore the first column:

```
import pandas as pd

#import all columns except first column into DataFrame
df = pd.read_csv('basketball_data.csv', usecols=range(1,3))

#view resulting DataFrame
print(df)

points rebounds
0 22 10
1 14 9
2 29 6
3 30 2
```

Notice that the first column called team was dropped when we imported the CSV file into pandas.

Note: You can find the complete documentation for the pandas read_csv() function .

The following tutorials explain how to perform other common tasks in Python: