

How can I combine multiple Excel sheets in Pandas?

Authored by
stats writer

April 18, 2024

RECOMMENDED CITATION

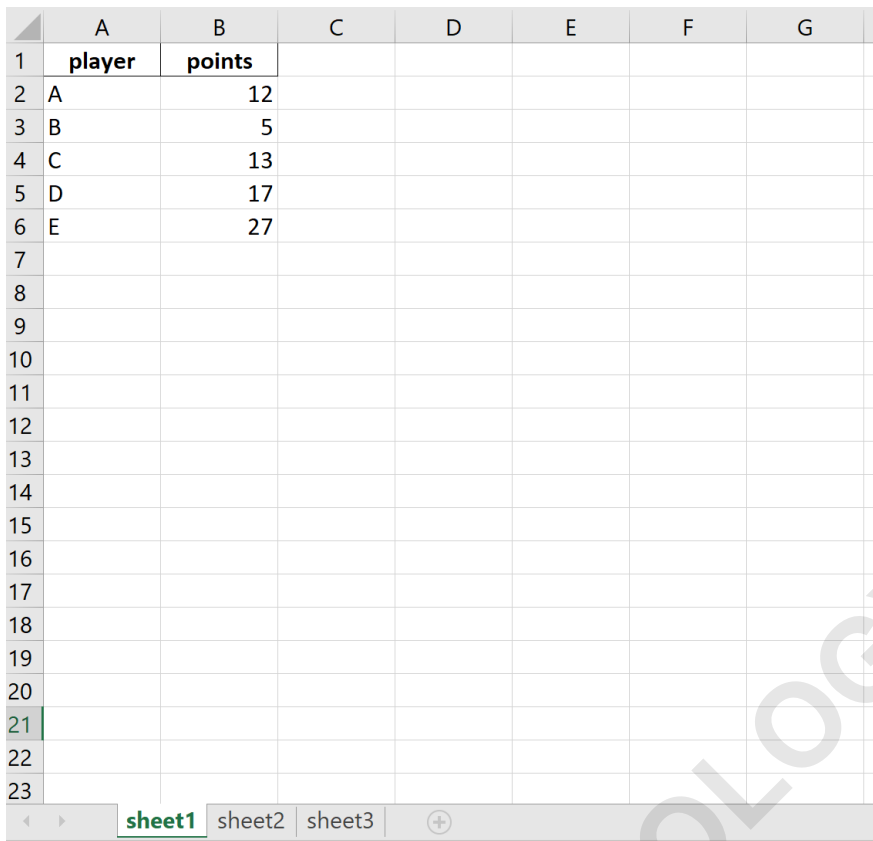
stats writer (2024). *How can I combine multiple Excel sheets in Pandas?*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=136793>

Combining multiple Excel sheets in Pandas refers to the process of merging or joining several Excel sheets into one consolidated dataset using the Python library, Pandas. This can be achieved by using the "concat" function, which combines the data from multiple sheets based on specified columns or indexes. It is a useful technique for organizing and analyzing large amounts of data from different sources. By combining multiple sheets, users can easily manipulate and perform various operations on the data, making it a valuable tool for data analysis and management.

Combine Multiple Excel Sheets in Pandas

Often you may want to import and combine multiple Excel sheets into a single pandas DataFrame.

For example, suppose you have the following Excel workbook called data.xlsx with three different sheets that all contain two columns of data about basketball players:



The image shows a screenshot of an Excel spreadsheet. The spreadsheet has columns labeled A through G and rows numbered 1 through 23. The data is as follows:

	A	B	C	D	E	F	G
1	player	points					
2	A	12					
3	B	5					
4	C	13					
5	D	17					
6	E	27					
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							

The spreadsheet interface shows three sheets: 'sheet1', 'sheet2', and 'sheet3'. The 'sheet1' tab is selected.

We can easily import and combine each sheet into a single pandas DataFrame using the pandas functions `concat()` and `read_excel()`, but first we need to make sure that `xlrd` is installed:

```
pip install xlrd
```

Once this is installed, we can use the following code to import and combine these three sheets into one pandas DataFrame:

```
#load pandas library
```

```
import pandas as pd
```

```
#import and combine the three sheets into one pandas  
DataFrame
```

```
df = pd.concat(pd.read_excel('data.xlsx',  
sheet_name=None), ignore_index=True)
```

```
#view DataFrame
```

```
df
```

```
player points
```

```
0 A 12
```

```
1 B 5
```

```
2 C 13
```

```
3 D 17
```

```
4 E 27
```

```
5 F 24
```

```
6 G 26
```

```
7 H 27
```

```
8 I 27
```

```
9 J 12
```

```
10 K 9
```

```
11 L 5
```

```
12 M 5
```

```
13 N 13
```

14 O 17

How This Code Works

There are only two pieces to understanding how this single line of code is able to import and combine multiple Excel sheets:

1. Read in all sheets.

```
pd.read_excel('data.xlsx', sheet_name=None)
```

This chunk of code reads in all sheets of an Excel workbook. By default, the `read_excel()` function only reads in the first sheet, but through specifying `sheet_name=None` we are able to read in every single sheet in the Excel workbook.

2. Concatenate all sheets.

```
pd.concat(DataFrames to concatenate,  
ignore_index=True)
```

This chunk of code simply concatenates all of the `DataFrames` from each Excel sheet into one single

pandas DataFrame. By specifying `ignore_index=True`, we're telling pandas that the names of the individual sheets are not important.

Note that this code only works if each of the Excel sheets has the same format. In this example, each sheet had two columns of data and each column had the same name, which is why this single line of code worked so easily to combine each of the Excel sheets into one pandas DataFrame.

[The Ultimate Guide: How to Read Excel Files with Pandas](#)
[How to Write Pandas DataFrames to Multiple Excel Sheets](#)