

How can I read Excel files using Pandas?

Authored by
stats writer

April 18, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I read Excel files using Pandas?*. PSYCHOLOGICAL SCALES.
Retrieved from <https://scales.arabpsychology.com/?p=136777>

Pandas is a popular open-source library in Python that is widely used for data analysis and manipulation. It provides various functions and methods for efficiently working with tabular data, including reading and writing files in different formats. To read Excel files using Pandas, one can use the "read_excel()" function, which takes in the file path as an input and returns a DataFrame object containing the data from the Excel file. This function also allows for specifying specific sheets, columns, and rows to be read, making it a versatile tool for handling large and complex Excel files. Overall, using Pandas to read Excel files offers a convenient and efficient way to extract and analyze data from these commonly used spreadsheet files.

The Ultimate Guide: Read Excel Files with Pandas

Excel files are one of the most common ways to store data. Fortunately the pandas function read_excel() allows you to easily read in Excel files.

This tutorial explains several ways to read Excel files into Python using pandas.

Example 1: Read Excel File into a pandas DataFrame

Suppose we have the following Excel file:

	A	B	C	D	E	F
1	playerID	team	points			
2		1 Lakers	26			
3		2 Mavs	19			
4		3 Bucks	24			
5		4 Spurs	22			
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						

The following code shows how to use the `read_excel()` function to import this Excel file into a pandas `DataFrame`:

```
import pandas as pd
```

```
#import Excel file
```

```
df = pd.read_excel('data.xlsx')
```

```
#view DataFrame
```

```
df
```

```
playerID team points
```

0 1 Lakers 26

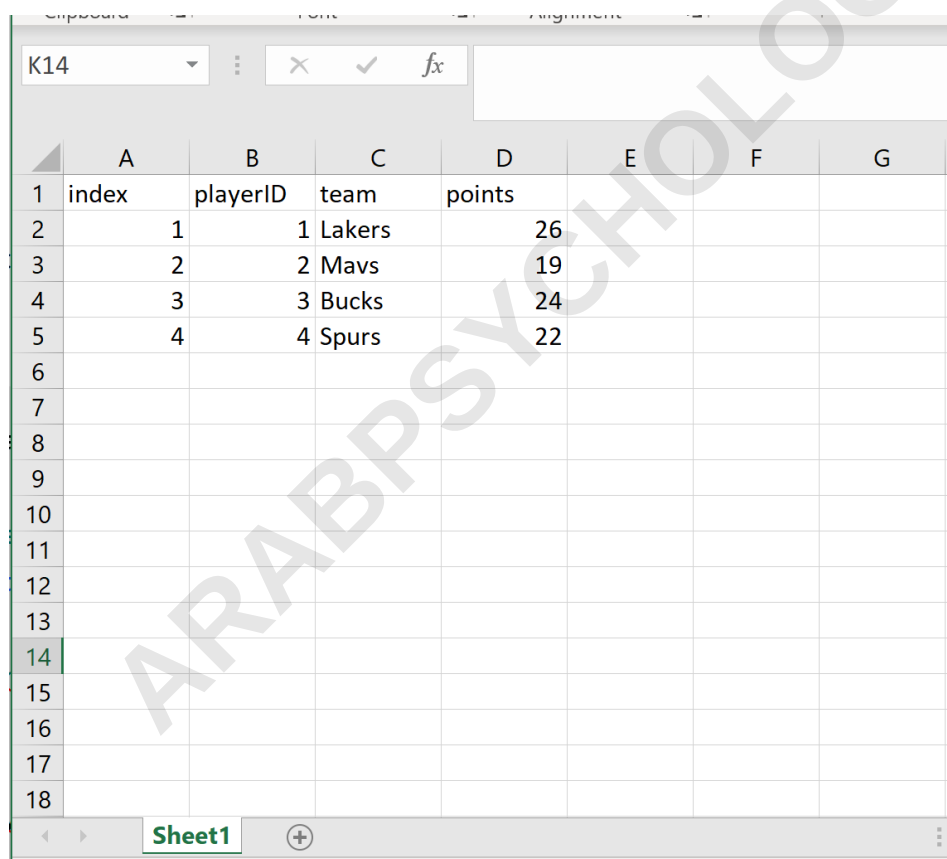
1 2 Mavs 19

2 3 Bucks 24

3 4 Spurs 22

Example 2: Read Excel File with Index Column

Sometimes you may also have an Excel file in which one of the columns is an index column:



	A	B	C	D	E	F	G
1	index	playerID	team	points			
2	1	1	Lakers	26			
3	2	2	Mavs	19			
4	3	3	Bucks	24			
5	4	4	Spurs	22			
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							

In this case you can use `index_col` to tell pandas which column to use as the index column when importing:

```
import pandas as pd
```

```
#import Excel file, specifying the index column
```

```
df = pd.read_excel('data.xlsx', index_col='index')
```

```
#view DataFrame
```

```
df
```

```
playerID team points
```

```
index
```

```
1 1 Lakers 26
```

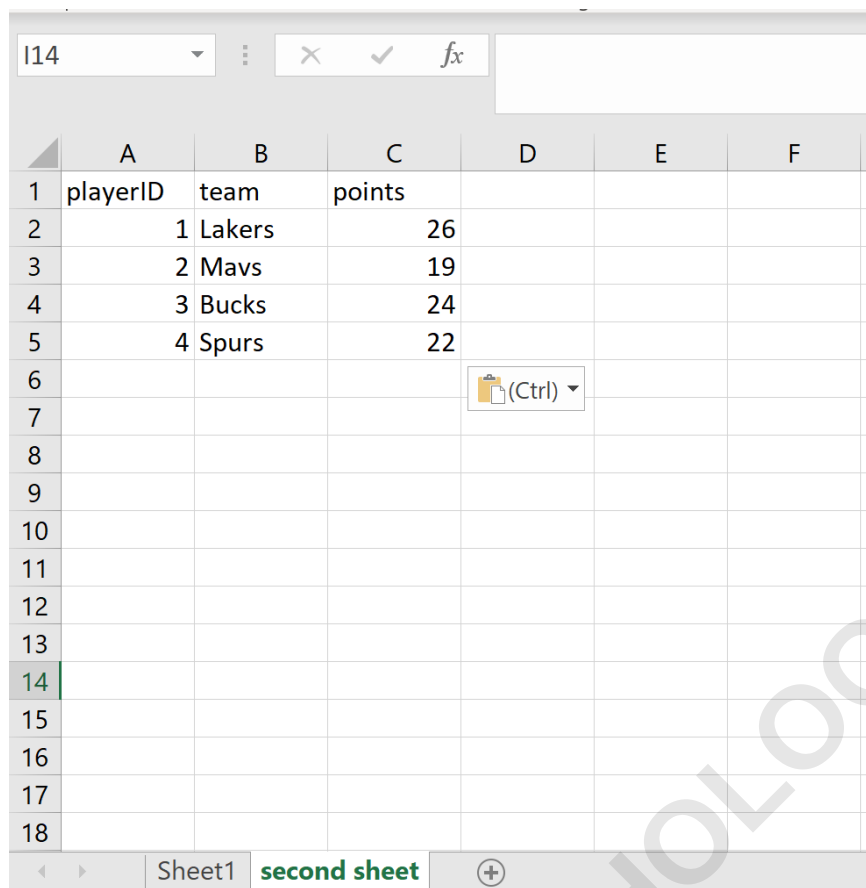
```
2 2 Mavs 19
```

```
3 3 Bucks 24
```

```
4 4 Spurs 22
```

Example 3: Read Excel File Using Sheet Name

You can also read specific sheet names from an Excel file into a pandas DataFrame. For example, consider the following Excel file:



	A	B	C	D	E	F
1	playerID	team	points			
2	1	Lakers	26			
3	2	Mavs	19			
4	3	Bucks	24			
5	4	Spurs	22			
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						

To read a specific sheet in as a pandas DataFrame, you can use the `sheet_name()` argument:

```
import pandas as pd
```

```
#import only second sheet
```

```
df = pd.read_excel('data.xlsx', sheet_name='second sheet')
```

```
#view DataFrame
```

```
df
```

playerID team points

0 1 Lakers 26

1 2 Mavs 19

2 3 Bucks 24

3 4 Spurs 22

Common Error: Install xlrd

When you attempt to use the `read_excel()` function, you may encounter the following error:

ImportError: Install xlrd >= 1.0.0 for Excel support

In this case, you need to first install xlrd:

`pip install xlrd`

Once this is installed, you may proceed to use the `read_excel()` function.

How to Read CSV Files with Pandas

How to Export a Pandas DataFrame to Excel