

How can I skip rows when reading an Excel file using Pandas?

Authored by
stats writer

June 25, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I skip rows when reading an Excel file using Pandas?*.
PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=152290>

Pandas is a popular Python library used for data analysis and manipulation. It provides various functions and methods for reading and writing data files, including Excel files. When reading an Excel file using Pandas, it is possible to skip rows that are not needed for analysis. This can be achieved by specifying the desired number of rows to be skipped in the "skiprows" parameter of the "read_excel()" function. This allows for efficient data processing and ensures that only relevant data is included in the analysis. By using this feature, users can customize their data reading process and extract only the necessary information from an Excel file.

Pandas: Skip Rows when Reading Excel File

You can use the following methods to skip rows when reading an Excel file into a pandas DataFrame:

Method 1: Skip One Specific Row

```
#import DataFrame and skip row in index position 2  
df = pd.read_excel('my_data.xlsx', skiprows=)
```

Method 2: Skip Several Specific Rows

```
#import DataFrame and skip rows in index positions 2  
and 4  
df = pd.read_excel('my_data.xlsx', skiprows=)
```

Method 3: Skip First N Rows

```
#import DataFrame and skip first 2 rows  
df = pd.read_excel('my_data.xlsx', skiprows=2)
```

The following examples show how to use each method in practice with the following Excel file called `player_data.xlsx`:

	A	B	C	D	E	F
1	team	points	rebounds	assists		
2	A	24	8	5		
3	B	20	12	3		
4	C	15	4	7		
5	D	19	4	8		
6	E	32	6	8		
7	F	13	7	9		
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						

Example 1: Skip One Specific Row

We can use the following code to import the Excel file and skip the row in index position 2:

```
import pandas as pd
```

```
#import DataFrame and skip row in index position 2
```

```
df = pd.read_excel('player_data.xlsx', skiprows=)
```

```
#view DataFrame
```

```
print(df)
```

```
team points rebounds assists
```

```
0 A 24 8 5
```

```
1 C 15 4 7
```

```
2 D 19 4 8
```

```
3 E 32 6 8
```

```
4 F 13 7 9
```

Notice that row in index position 2 (with team 'B') was skipped when importing the Excel file into the pandas DataFrame.

Note: The first row in the Excel file is considered to be row 0.

Example 2: Skip Several Specific Rows

We can use the following code to import the Excel file and skip the rows in index positions 2 and 4:

```
import pandas as pd
```

#import DataFrame and skip rows in index positions 2 and 4

df = pd.read_excel('player_data.xlsx', skiprows=)

#view DataFrame

print(df)

team points rebounds assists

0 A 24 8 5

1 C 15 4 7

2 E 32 6 8

3 F 13 7 9

Example 3: Skip First N Rows

We can use the following code to import the Excel file and skip the first two rows:

import pandas as pd

#import DataFrame and skip first 2 rows

df = pd.read_excel('player_data.xlsx', skiprows=2)

#view DataFrame

print(df)

B 20 12 3
0 C 15 4 7
1 D 19 4 8
2 E 32 6 8
3 F 13 7 9

Notice that the first two rows in the Excel file were skipped and the next available row (with team 'B') became the header row for the DataFrame.

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

How to Export NumPy Array to CSV File