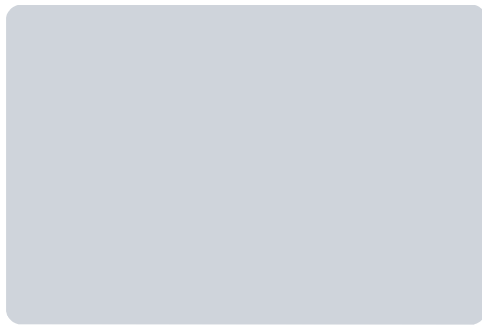Java Tutorial   Index Posts   Interview Questions   Resources

Pandas read_excel() – Reading Excel File in Python

Filed Under: Pandas

11 Comments

Home » Python » Pandas read_excel() – Reading Excel File in Python

We can use the pandas module read_excel() function to read the excel file data into a DataFrame object.

If you look at an excel sheet, it's a two-dimensional table. The DataFrame object also represents a two-dimensional tabular data structure.
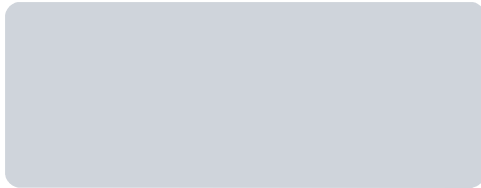
1. Pandas

‹ Previous   Next ›

Let's say we have an excel file
with two sheets – Employees and
Cars. The top row contains the
header of the table.

Excel File Sheets Data

Here is the example to read the
"Employees" sheet data and
printing it.

```python
import pandas

excel_data_df =
pandas.read_excel('r
ecords.xlsx',
sheet_name='Employee
s')

# print whole sheet
data
print(excel_data_df)
```

Output:

‹ Previous   Next ›   🔍

```
        EmpID     EmpName
EmpRole
0         1      Pankaj
CEO
1         2   David Lee
Editor
2         3    Lisa Ray
Author
```

The first parameter is the name of the excel file.

The sheet_name parameter defines the sheet to be read from the excel file.

When we print the DataFrame object, the output is a two-dimensional table. It looks similar to an excel sheet records.

2. List of Columns Headers of the Excel Sheet

We can get the list of column headers using the `columns` property of the dataframe object.

```
print(excel_data_df.columns.ravel())
```

Output:

< Previous   Next >   Q

```
['EmpID' 'EmpName'
 'EmpRole']
```

3. Printing a Column Data

We can get the column data and convert it into a list of values.

```
print(excel_data_df[
'EmpName'].tolist())
```

Output:

```
['Pankaj', 'David
Lee', 'Lisa Ray']
```

4. Pandas read_excel() usecols example

We can specify the column names to be read from the excel file. It's useful when you are interested in only a few of the columns of the excel sheet.

```python
import pandas

excel_data_df = pandas.read_excel('records.xlsx', sheet_name='Cars', usecols=['Car Name', 'Car Price'])
print(excel_data_df)
```

Output:

```
        Car Name
Car Price
0     Honda City
20,000 USD
1  Bugatti Chiron  3
Million USD
2      Ferrari 458
2,30,000 USD
```

5. Reading Excel File without Header Row

If the excel sheet doesn't have any header row, pass the header parameter value as None.

<Previous   Next >   🔍

```
excel_data_df =
pandas.read_excel('r
ecords.xlsx',
sheet_name='Numbers'
, header=None)
```

If you pass the header value as an integer, let's say 3. Then the third row will be treated as the header row and the values will be read from the next row onwards. Any data before the header row will be discarded.

6. Excel Sheet to Dict, CSV and JSON

The DataFrame object has various utility methods to convert the tabular data into Dict, CSV, or JSON format.

```
excel_data_df =
pandas.read_excel('r
ecords.xlsx',
sheet_name='Cars',
usecols=['Car Name',
'Car Price'])

print('Excel Sheet
to Dict:',
excel_data_df.to_dic
t(orient='record'))
print('Excel Sheet
to JSON:',
excel_data_df.to_jso
n(orient='records'))
```

< Previous   Next >

```
excel_data_df.to_csv
(index=False))
```

Output:

```
Excel Sheet to
Dict: [{'Car Name':
'Honda City', 'Car
Price': '20,000
USD'}, {'Car Name':
'Bugatti Chiron',
'Car Price': '3
Million USD'},
{'Car Name':
'Ferrari 458', 'Car
Price': '2,30,000
USD'}]
Excel Sheet to
JSON: [{"Car
Name":"Honda
City","Car
Price":"20,000
USD"},{"Car
Name":"Bugatti
Chiron","Car
Price":"3 Million
USD"},{"Car
Name":"Ferrari
458","Car
Price":"2,30,000
```

‹ Previous   Next ›   🔍

```
Honda City,"20,000
```

7. References

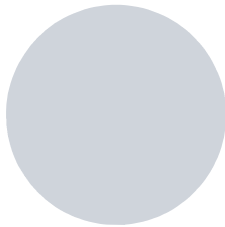[pandas read_excel() API Docs](#)

Python Pandas Module
Tutorial

◇ PREV

Pandas read_csv() - Reading
CSV File to DataFrame

NEXT ❯❯

Pankaj

I love Open
Source
technologies
and writing
about my
experience
about them
is my
passion.

Follow Author

f 🐦 in

Comments

Rengifo says:
January 13, 2022 at 3:41 am

Josh Rodriguez says:
January 5, 2022 at 12:11 pm

how can we remove a specific row? I have a 'Total' row at the end of my Excel file that I would want to remove. Thanks for the helpful info in this article!

Rajjan says:
June 16, 2021 at 1:12 am

Thanks pankaj… It saved my data. I wanted to read from a specific folder so I tweaked it for the location.

pd.read_excel('C:/Source/Datafile.xlsx', sheet_name='sheet_name')

Rahul Misra says:
January 21, 2021 at 7:20 am

Thanks, useful post.

excel_data_df = pandas.read_excel('records.xlsx', sheet_name='Cars', usecols=['Car Name', 'Car Price']) results in an empty dataframe for me. Do we need to do some setup?

‹ Previous   Next ›   🔍

usecols=[]
expects column
numbers,
indexed 0

Prabhat Kumar says:
December 17, 2020 at 8:58 pm

I must say this is
'awesome'

Havi says:
September 8, 2020 at 8:52 pm

Thanks, it was very
helpful.

mila says:
June 24, 2020 at 12:04 pm

Thank you, it was very
helpful

byli says:
June 19, 2020 at 3:31 am

thanks this was very
helpful!

firozsahib says:
May 11, 2020 at 11:37 pm

Thanks it helped me

Shilpa Bindra says:
January 7, 2020 at 3:24 am

What is we have an
xlsb file instead of
xlsx?

Comments are closed.

‹ Previous   Next ›

JournalDev is one of the most popular websites for Java, Python, Android, and related technical articles. Our tutorials are regularly updated, error-free, and complete. Every month millions of developers like you visit JournalDev to read our tutorials.

JournalDev was founded by Pankaj Kumar in 2010 to share his experience and learnings with the whole world. He loves Open source technologies and writing on JournalDev has become his passion.

Most Popular

Java / Java EE Tutorials

Core Java Tutorial

Python Tutorials

Java Interview Questions

Core Java Interview Questions

Java Design Patterns

Spring Tutorial

Favorite Sites

AskPython

CodeForGeek

GoLangDocs

LinuxForDevices

Mkyong

VM-Help

WP-Design