

# Data To Fish

Data Science Tutorials

LAST UPDATED ON APRIL 3RD, 2019,

## How to Create Pandas DataFrame in Python

Need to create pandas DataFrame in Python?

If so, I'll show you two different methods to create pandas DataFrame:

- By typing the values in Python itself to create the DataFrame
- By importing the values from a file (such as an Excel file), and then creating the DataFrame in Python based on the values imported

### Method 1: typing values in Python to create pandas DataFrame

To create pandas DataFrame in Python, you can follow this generic template:

```
from pandas import DataFrame

Data = {'First Field Name': ['First value', 'Second value',...],
        'Second Field Name': ['First value', 'Second value',...],
        ....}
```

```
}  
  
df = DataFrame (Data, columns = ['First Field Name','Second Field
```

Note that there is no need to use quotes around *numeric* values (unless you wish to capture those values as strings).

Now let's see how to apply the above template using a simple example.

To start, let's say that you have the following data about Cars, and that you want to capture that data in Python using pandas DataFrame:

Brand	Price
Honda Civic	22,000
Toyota Corolla	25,000
Ford Focus	27,000
Audi A4	35,000

This is how the Python code would look like for our Cars example:

```
from pandas import DataFrame  
  
Cars = {'Brand': ['Honda Civic','Toyota Corolla','Ford Focus','Audi A4'],  
        'Price': [22000,25000,27000,35000]}  
  
df = DataFrame(Cars,columns= ['Brand', 'Price'])  
  
print (df)
```

Run the Python code, and you'll get the following DataFrame:

```
      Brand  Price
0   Honda Civic  22000
1  Toyota Corolla  25000
2   Ford Focus  27000
3    Audi A4   35000
```

Let's now review the second method of importing the values into Python to create the DataFrame.

## Method 2: importing values from an Excel file to create pandas DataFrame

You can use the syntax `read_excel` to import the Excel file into Python in order to create your DataFrame:

```
import pandas as pd

Data = pd.read_excel(r'Path where the Excel file is stored\File name.xlsx')
df = pd.DataFrame(Data, columns= ['First Field Name', 'Second Field Name'])

print (df)
```

Let's say that you have the following table stored in an Excel file (where the Excel *file name* is 'Cars'):

Brand	Price
Honda Civic	22,000
Toyota Corolla	25,000
Ford Focus	27,000
Audi A4	35,000

In the Python code below, you'll need to change the *path name* to the location where the Excel file is stored on *your* computer.

In my case, the Excel file is saved on my desktop, under the following path:

**'C:\Users\Ron\Desktop\Cars.xlsx'**

Once you imported the data into Python, you'll be able to assign it to the DataFrame. Here is the full Python code for our example:

```
import pandas as pd

Cars = pd.read_excel(r'C:\Users\Ron\Desktop\Cars.xlsx')
df = pd.DataFrame(Cars, columns= ['Brand', 'Price'])

print (df)
```

As before, you'll get the same pandas DataFrame in Python:

**Note:** you will have to install *xlrd* if you get the following error when running the code:

**ImportError: Install xlrd >= 1.0.0 for Excel support**

You may then use the PIP install method to install *xlrd* as follows:

```
pip install xlrd
```

You can also create the same DataFrame if you need to import a CSV file into Python, rather than using an Excel file.

## Get the maximum value from the DataFrame

Once you have your values in the DataFrame, you can perform a large variety of operations. For example, you may calculate stats using pandas.

For instance, let's say that you want to find the maximum price among all the Cars within the DataFrame.

Obviously, you can derive this value just by looking at the data-set, but the method presented below would work for much larger data-sets.

To get the maximum price in our Cars example, you'll need to add the following portion to the Python code (and then print the results):

```
max1 = df['Price'].max()
```

And the complete Python code is:

```
from pandas import DataFrame

Cars = {'Brand': ['Honda Civic', 'Toyota Corolla', 'Ford Focus', 'Audi A4'],
        'Price': [22000, 25000, 27000, 35000]}

df = DataFrame(Cars, columns= ['Brand', 'Price'])
```

```
max1 = df['Price'].max()  
print (max1)
```

Once you run the code, you'll get the value of *35,000*, which is indeed the maximum price!

You can check the pandas documentation to learn more about creating pandas DataFrame.

 **PYTHON**

[Home](#) » [Blog](#) » [How to Create Pandas DataFrame in Python](#)