



Introduction and Basic Dataframe

Profile



 [linkedin.com/in/romansyasetyo/](https://www.linkedin.com/in/romansyasetyo/)



Table of Content

What will We Learn Today?

1. About Pandas
2. About DataFrame
3. Creating DataFrame
4. Read DataFrame
5. Export DataFrame

Hands on using Python





Google Colab

<https://colab.research.google.com/>





About Pandas





About Pandas

- Pandas is a fast, powerful, flexible and easy to use open source **data analysis and manipulation tool**, built on top of the Python programming language
- Tools for **reading and writing data** between in-memory data structures and different formats: CSV and text files, Microsoft Excel, SQL databases, and the fast HDF5 format





About DataFrame



About DataFrame

- Main object in data analysis with Python
- Represent data with **rows & columns**

	Column names								
	Name	Team	Number	Position	Age	Height	Weight	College	Salary
0	Avery Bradley	Boston Celtics	0.0	PG	25.0	6-2	180.0	Texas	7730337.0
1	John Holland	Boston Celtics	30.0	SG	27.0	6-5	205.0	Boston University	NaN
2	Jonas Jerebko	Boston Celtics	8.0	PF	29.0	6-10	231.0	NaN	5000000.0
3	Jordan Mickey	Boston Celtics	NaN	PF	21.0	6-8	235.0	LSU	1170960.0
4	Terry Rozier	Boston Celtics	12.0	PG	22.0	6-2	190.0	Louisville	1824360.0
5	Jared Sullinger	Boston Celtics	7.0	C	NaN	6-9	260.0	Ohio State	2569260.0
6	Evan Turner	Boston Celtics	11.0	SG	27.0	6-7	220.0	Ohio State	3425510.0

Source : GeeksforGeeks



Creating DataFrame





Creating DataFrame

```
# Method 1: Creating Pandas DataFrame from lists of lists.  
  
# initialize list of lists  
data_1 = [['tom', 10], ['nick', 15], ['juli', 14]]  
  
# Create the pandas DataFrame  
df_1 = pd.DataFrame(data_1, columns = ['Name', 'Age'])  
  
# print dataframe.  
df_1
```

	Name	Age
0	tom	10
1	nick	15
2	juli	14





Creating DataFrame

```
# Method 2: Creating DataFrame from dict of array/lists

# initialise data of lists.
data_2 = {'Name': ['Tom', 'nick', 'krish', 'jack'],
          'Age': [20, 21, 19, 18]}

# Create DataFrame
df_2 = pd.DataFrame(data_2)

# Print the output.
df_2
```

	Name	Age
0	Tom	20
1	nick	21
2	krish	19
3	jack	18



Creating DataFrame

```
# Method 3: Creating Dataframe from list of dicts
```

```
# Initialise data to lists.
```

```
data_3 = [{'a': 1, 'b': 2, 'c': 3},  
          {'a': 10, 'b': 20, 'c': 30}]
```

```
# Creates DataFrame.
```

```
df_3 = pd.DataFrame(data_3)
```

```
# Print the data
```

```
df_3
```

	a	b	c
0	1	2	3
1	10	20	30



Creating DataFrame

```
# Method 4: Creating DataFrame using zip() function.
```

```
# Python program to demonstrate creating  
# pandas Datadaframe from lists using zip.
```

```
# List1
```

```
Name = ['tom', 'krish', 'nick', 'juli']
```

```
# List2
```

```
Age = [25, 30, 26, 22]
```

```
# get the list of tuples from two lists.
```

```
# and merge them by using zip().
```

```
data_4 = list(zip(Name, Age))
```

```
# Converting lists of tuples into
```

```
# pandas Dataframe.
```

```
df_4 = pd.DataFrame(data_4,  
                     columns = ['Name', 'Age'])
```

```
# Print data.
```

```
df_4
```

	Name	Age
0	tom	25
1	krish	30
2	nick	26
3	juli	22



Read DataFrame





Read DataFrame

- Pandas is a strong library since it can get data from many sources, for example CSV, TSV, Excel, JSON, HTML, etc.

Getting data from external is pretty straightforward, the example can be seen at following code:

```
df = pd.read_csv('filepath.type', delimiter=',')  
df = pd.read_excel('filepath.type', sheet_name=0, header=0, names=None, index_col=None)
```





Export DataFrame





Export DataFrame

- When we have processed the DataFrame and want to export it to external file, pandas have covered the functions for us:

```
# Export to CSV
df.to_csv('filepath.csv', index=False, delimiter=';')

# Export to Excel
df.to_excel('filepath.xlsx', sheet_name='Sheet1', index=False)
```



Thank You

