

# ▼ Pandas

```
import pandas as pd
```

```
data={"name":["Bill","Tom","Tim","John","Alex","Vanessa","Kate"],
      "score":[90,80,85,75,95,60,65],
      "sport":["Wrestling","Football","Skiing","Swimming","Tennis",
               "Karete","Surfing"],
      "sex":["M","M","M","M","F","F","F"]}
```

```
df=pd.DataFrame(data)
```

df

	name	score	sport	sex
0	Bill	90	Wrestling	M
1	Tom	80	Football	M
2	Tim	85	Skiing	M
3	John	75	Swimming	M
4	Alex	95	Tennis	F
5	Vanessa	60	Karete	F
6	Kate	65	Surfing	F

```
df=pd.DataFrame(data,columns=["name","sport","sex","score"])
```

df

	name	sport	sex	score
0	Bill	Wrestling	M	90
1	Tom	Football	M	80
2	Tim	Skiing	M	85
3	John	Swimming	M	75
4	Alex	Tennis	F	95
5	Vanessa	Karete	F	60
6	Kate	Surfing	F	65

```
df.head()
```

	name	sport	sex	score
--	------	-------	-----	-------

0	Bill	Wrestling	M	90
1	Tom	Football	M	80
2	Tim	Skiing	M	85
3	John	Swimming	M	75
4	Alex	Tennis	F	95

```
df.tail()
```

	name	sport	sex	score
2	Tim	Skiing	M	85
3	John	Swimming	M	75
4	Alex	Tennis	F	95
5	Vanessa	Karete	F	60
6	Kate	Surfing	F	65

```
df.tail(3)
```

	name	sport	sex	score
4	Alex	Tennis	F	95
5	Vanessa	Karete	F	60
6	Kate	Surfing	F	65

```
df.head(2)
```

	name	sport	sex	score
0	Bill	Wrestling	M	90
1	Tom	Football	M	80

```
df=pd.DataFrame(data,columns=["name", "sport", "gender", "score", "age"])
df
```

	name	sport	gender	score	age
0	Bill	Wrestling	NaN	90	NaN
1	Tom	Football	NaN	80	NaN
2	Tim	Skiing	NaN	85	NaN
3	John	Swimming	NaN	75	NaN
4	Alex	Tennis	NaN	95	NaN

<b>5</b>	Vanessa	Karete	NaN	60	NaN
<b>6</b>	Kate	Surfing	NaN	65	NaN

```
df=pd.DataFrame(data,columns=["name", "sport", "gender", "score", "age"],
                index=["one","two","three","four","five","six","seven"])

df
```

	name	sport	gender	score	age
<b>one</b>	Bill	Wrestling	NaN	90	NaN
<b>two</b>	Tom	Football	NaN	80	NaN
<b>three</b>	Tim	Skiing	NaN	85	NaN
<b>four</b>	John	Swimming	NaN	75	NaN
<b>five</b>	Alex	Tennis	NaN	95	NaN
<b>six</b>	Vanessa	Karete	NaN	60	NaN
<b>seven</b>	Kate	Surfing	NaN	65	NaN

```
df["name"]
```

```
one      Bill
two      Tom
three    Tim
four     John
five     Alex
six      Vanessa
seven    Kate
Name: name, dtype: object
```

```
my_columns=["name","sport"]
x = df[my_columns]
```

```
x
```

	name	sport
<b>one</b>	Bill	Wrestling
<b>two</b>	Tom	Football
<b>three</b>	Tim	Skiing
<b>four</b>	John	Swimming
<b>five</b>	Alex	Tennis
<b>six</b>	Vanessa	Karete
<b>seven</b>	Kate	Surfing

df.sport

```
one      Wrestling
two      Football
three    Skiing
four     Swimming
five     Tennis
six      Karete
seven    Surfing
Name: sport, dtype: object
```

df.loc[["one"]]

	name	sport	gender	score	age
one	Bill	Wrestling	NaN	90	NaN

df.loc[["one","two"]]

	name	sport	gender	score	age
one	Bill	Wrestling	NaN	90	NaN
two	Tom	Football	NaN	80	NaN

df["age"]=18

```
df=pd.DataFrame(data,columns=["name", "sport", "gender", "score", "age"],
                 index=["one","two","three","four","five","six","seven"])
values=[18,19,20,18,17,17,18]
df["age"]=values
df
```

	name	sport	gender	score	age
one	Bill	Wrestling	NaN	90	18
two	Tom	Football	NaN	80	19
three	Tim	Skiing	NaN	85	20
four	John	Swimming	NaN	75	18
five	Alex	Tennis	NaN	95	17
six	Vanessa	Karete	NaN	60	17
seven	Kate	Surfing	NaN	65	18

```
df["pass"]=df.score>=70
df
```

	name	sport	gender	score	age	pass
--	------	-------	--------	-------	-----	------

one	Bill	Wrestling	NaN	90	18	True
two	Tom	Football	NaN	80	19	True
three	Tim	Skiing	NaN	85	20	True
four	John	Swimming	NaN	75	18	True
five	Alex	Tennis	NaN	95	17	True
six	Vanessa	Karete	NaN	60	17	False
seven	Kate	Surfing	NaN	65	18	False

```
del df["pass"]
df
```

	name	sport	gender	score	age
one	Bill	Wrestling	NaN	90	18
two	Tom	Football	NaN	80	19
three	Tim	Skiing	NaN	85	20
four	John	Swimming	NaN	75	18
five	Alex	Tennis	NaN	95	17
six	Vanessa	Karete	NaN	60	17
seven	Kate	Surfing	NaN	65	18

```
scores={"Math":{"A":85,"B":90,"C":95}, "Physics":{"A":90,"B":80,"C":75}}
```

```
scores_df=pd.DataFrame(scores)
scores_df
```

	Math	Physics
A	85	90
B	90	80
C	95	75

```
scores_df.index.name="name"
scores_df.columns.name="lesson"
```

```
scores_df
```

lesson	Math	Physics
name		
A	85	90

<b>B</b>	90	80
----------	----	----

<b>C</b>	95	75
----------	----	----

```
scores_df.values
```

```
array([[85, 90],  
       [90, 80],  
       [95, 75]])
```

```
scores_df.keys
```

```
<bound method NDFrame.keys of lesson  Math  Physics  
name  
A      85      90  
B      90      80  
C      95      75>
```