

ClassNum/Pan



default ▼

FINISHED ▶ ↻ 📖 ⚙️

```
%pyspark
from pandas import Series, DataFrame
import pandas as pd
obj = Series([4, 7, -5, 3])
obj
obj.values
obj.index
obj2 = Series([4, 7, -5, 3], index=['d', 'b', 'a', 'c'])
obj2
obj2.index
obj2['a']
obj2['d'] = 6
obj2[['c', 'a', 'd']]
obj2
obj2[obj2 > 0]
obj2 * 2
import numpy as np
np.exp(obj2)
'b' in obj2
'e' in obj2
sdata = {'Ohio': 35000, 'Texas': 71000, 'Oregon': 16000, 'Utah': 5000}
obj3 = Series(sdata)
obj3
states = ['California', 'Ohio', 'Oregon', 'Texas']
obj4 = Series(sdata, index=states)
obj4
pd.isnull(obj4)
pd.notnull(obj4)
obj4.isnull()
obj3
obj4
obj3 + obj4
obj4.name = 'population'
obj4.index.name = 'state'
obj4
obj.index = ['Bob', 'Steve', 'Jeff', 'Ryan']
obj
data = {'state': ['Ohio', 'Ohio', 'Ohio', 'Nevada', 'Nevada'],
        'year': [2000, 2001, 2002, 2001, 2002],
        'pop': [1.5, 1.7, 3.6, 2.4, 2.9]}
frame = DataFrame(data)
frame
DataFrame(data, columns=['year', 'state', 'pop'])
frame2 = DataFrame(data, columns=['year', 'state', 'pop', 'debt'],
                   index=['one', 'two', 'three', 'four', 'five'])
frame2
frame2.columns
frame2['state']
frame2.year
frame2.ix['three']
frame2['debt'] = 16.5
frame2
```

ClassNum/Pan

ClassNum/Pan

ClassNum/Pan

ClassNum/Pan

ClassNum/Pan



keyboard settings lock default

```

val = Series([-1.2, -1.5, -1.7], index=['two', 'four', 'five'])
frame2['debt'] = val
frame2
frame2['eastern'] = frame2.state == 'Ohio'
frame2
del frame2['eastern']
frame2.columns
pop = {'Nevada': {2001: 2.4, 2002: 2.9},
      'Ohio': {2000: 1.5, 2001: 1.7, 2002: 3.6}}
frame3 = DataFrame(pop)
frame3
frame3.T
pdata = {'Ohio': frame3['Ohio'][:-1],
        'Nevada': frame3['Nevada'][:2]}
DataFrame(pdata)
frame3.index.name = 'year'; frame3.columns.name = 'state'
frame3
frame3.values
array([[2000, 'Ohio', 1.5, nan],
       [2001, 'Ohio', 1.7, -1.2],
       [2002, 'Ohio', 3.6, nan],
       [2001, 'Nevada', 2.4, -1.5],
       [2002, 'Nevada', 2.9, -1.7]], dtype=object)

```

Took 31 sec. Last updated by anonymous at February 16 2017, 6:13:49 PM.

FINISHED ▶ 🔍 📖 ⚙️

```

%pyspark
from pandas import Series , DataFrame
import pandas as pd
obj = Series([4,7,-5,3])
obj

```

```

0    4
1    7
2   -5
3    3
dtype: int64

```

Took 0 sec. Last updated by anonymous at February 16 2017, 6:35:07 PM.

FINISHED ▶ 🔍 📖 ⚙️

```

%pyspark
obj.values

```

array([4, 7, -5, 3])



Took 0 sec. Last updated by anonymous at February 16 2017, 6:35:51 PM.

FINISHED ▶ 🔍 📖 ⚙️

```

%pyspark
obj.index

```

RangeIndex(start=0, stop=4, step=1)



Took 0 sec. Last updated by anonymous at February 16 2017, 6:36:30 PM.

ClassNum/Pan

FINISHED ▶ ↻ 📖 ⚙️

Took 0 sec. Last updated by anonymous at February 16 2017, 6:42:20 PM.

FINISHED

↓

FINISHED    

Took 0 sec. Last updated by anonymous at February 16 2017, 6:44:04 PM.

FINISHED ▶ ↻ 📖 ⚙️

Took 1 sec. Last updated by anonymous at February 16 2017, 7:04:08 PM.

FINISHED    

```
%pyspark
class Num/Pan
```

default ▼

11

```
dtype: int64
```

FINISHED ▶ ⌵ ⌶ ⚙

obj4

```
dtype: float64
```

FINISHED

```
obj4.isNull()
```

```
dtype: bool
```

FINISHED ▶ 🔍 📖 ⚙️

```
obj4.index.name = 'state'
```

FINISHED ▶ 🔍 📖 ⚙️

obj4

Texas	71000.0
-------	---------

Name: population, dtype: float64

Zeppelin

Took 0 sec. Last updated by anonymous at February 16 2017, 7:02:35 PM.

ClassNum/Pan

▶

⌵

📖

✍

📄

⬇

📄

🗑

⌛

FINISHED

⌵

⚙

🔒

📖

default

⌵

%pyspark

```
DataFrame(data, columns=['year', 'state', 'pop'])
frame2 = DataFrame(data, columns=['year', 'state', 'pop', 'debt'],
                    index=['one', 'two', 'three', 'four', 'five'])
```

Took 0 sec. Last updated by anonymous at February 16 2017, 7:10:10 PM.

%pyspark

```
pop = {'Nevada': {2001: 2.4, 2002: 2.9},
       'Ohio': {2000: 1.5, 2001: 1.7, 2002: 3.6}}
```

FINISHED

▶

⌵

📖

⚙

Took 0 sec. Last updated by anonymous at February 16 2017, 7:11:24 PM.

|

READY

▶

⌵

📖

⚙