



(<http://www.pieriandata.com>)

Missing Data

Let's show a few convenient methods to deal with Missing Data in pandas:

In [1]:

```
import numpy as np
import pandas as pd
```

In [9]:

```
df = pd.DataFrame({'A': [1, 2, np.nan],
                   'B': [5, np.nan, np.nan],
                   'C': [1, 2, 3]})
```

In [10]:

```
df
```

Out[10]:

	A	B	C
0	1.0	5.0	1
1	2.0	NaN	2
2	NaN	NaN	3

In [12]:

```
df.dropna()
```

Out[12]:

	A	B	C
0	1.0	5.0	1

In [13]:

```
df.dropna(axis=1)
```

Out[13]:

	C
0	1
1	2
2	3

In [14]:

```
df.dropna(thresh=2)
```

Out[14]:

	A	B	C
0	1.0	5.0	1
1	2.0	NaN	2

In [15]:

```
df.fillna(value='FILL VALUE')
```

Out[15]:

	A	B	C
0	1	5	1
1	2	FILL VALUE	2
2	FILL VALUE	FILL VALUE	3

In [17]:

```
df['A'].fillna(value=df['A'].mean())
```

Out[17]:

0	1.0
1	2.0
2	1.5

Name: A, dtype: float64

Great Job!