

# 03-Missing-Data

August 5, 2020

## 1 Missing Data

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

```
[1]: import numpy as np
import pandas as pd
```

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

```
[10]: df
```

```
[10]:
```

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

```
[12]: df.dropna()
```

```
[12]:
```

	A	B	C
0	1.0	5.0	1

```
[13]: df.dropna(axis=1)
```

```
[13]:
```

	C
0	1
1	2
2	3

```
[14]: df.dropna(thresh=2)
```

```
[14]:
```

	A	B	C
0	1.0	5.0	1
1	2.0	NaN	2

```
[15]: df.fillna(value='FILL VALUE')
```

```
[15]:
```

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

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

```
[17]: 0    1.0  
      1    2.0  
      2    1.5  
      Name: A, dtype: float64
```