In [1]:
```
import numpy as np
import pandas as pd
from pandas import Series,DataFrame
```

In [7]:
```
#Lets get a dataframe with duplicates

dframe = DataFrame({'key1': ['A'] * 2 + ['B'] * 3,
                    'key2': [2, 2, 2, 3, 3]})

#Show
dframe
```

Out[7]:

	key1	key2
0	A	2
1	A	2
2	B	2
3	B	3
4	B	3

In [8]:
```
#We can use duplicated to find duplicates
dframe.duplicated()
```

Out[8]:
```
0    False
1     True
2    False
3    False
4     True
dtype: bool
```

In [9]:
```
# We can also drop duplicates like this:
dframe.drop_duplicates()
```

Out[9]:

	key1	key2
0	A	2
2	B	2
3	B	3

In [10]:
```
#You can filter which duplicates to drop by a single column
dframe.drop_duplicates(['key1'])
```

Out[10]:

	key1	key2
0	A	2
2	B	2

In [11]:
```
#Show original
dframe
```

Out[11]:

	key1	key2
0	A	2
1	A	2
2	B	2
3	B	3
4	B	3

In [14]:
```
#By default the first value was taken for the duplicates, we can also take the la
dframe.drop_duplicates(['key1'],take_last=True)
```

Out[14]:

	key1	key2
1	A	2
4	B	3

In [ ]:
```
# That's about it for finding duplicates, pretty easy huh?

#Up next: Mapping
```