JakeDineen_5.2

August 8, 2018

In [1]:
```
import pandas as pd
import numpy as np
#Store data
state_data = {'State':['Alabama','Alaska','Arizona','Arkansas'],'PostCode':['AL','AK',

#Read as dataframe + Display
state = pd.DataFrame(state_data, columns=['State','PostCode','Area','Pop'])
state
```

Out[1]:

	State	PostCode	Area	Pop
0	Alabama	AL	52,423	4,040,587
1	Alaska	AK	656,424	550,043
2	Arizona	AZ	*	3,665,228
3	Arkansas	AR	53,182	2,350,750

In [2]:
```
#Set State as index
state = state.set_index('State')
state
```

Out[2]:

State	PostCode	Area	Pop
Alabama	AL	52,423	4,040,587
Alaska	AK	656,424	550,043
Arizona	AZ	*	3,665,228
Arkansas	AR	53,182	2,350,750

In [3]:
```
#Replace the '*' with '0'.
state.replace('*','0', inplace = True)
state
```

Out[3]:

State	PostCode	Area	Pop
Alabama	AL	52,423	4,040,587
Alaska	AK	656,424	550,043
Arizona	AZ	0	3,665,228
Arkansas	AR	53,182	2,350,750

In [6]:
```
#Define a function to replace ',' with ''.
def replace_commas(x):
```

```
        try:
            return x.replace(',', '')
        except:
            print('Error')

    state['Area'] = state['Area'].map(replace_commas)
    state['Pop'] = state['Pop'].map(replace_commas)
    state
```

Out[6]:

State	PostCode	Area	Pop
Alabama	AL	52423	4040587
Alaska	AK	656424	550043
Arizona	AZ	0	3665228
Arkansas	AR	53182	2350750