

Import necessary python packages

```
In [1]: import numpy as np
import pandas as pd
import scipy.stats as stats
import matplotlib.pyplot as plt
%matplotlib inline
```

Import Boston data set and store it in a variable called boston.

```
In [2]: import sklearn
from sklearn.datasets import load_boston
boston_data = load_boston()
```

Boston is a dictionary, understanding the keys of this dictionary.

```
In [3]: boston_data.keys()
```

```
Out[3]: dict_keys(['feature_names', 'data', 'target', 'DESCR'])
```

```
In [4]: boston_data.data.shape
```

```
Out[4]: (506, 13)
```

Print the feature names of boston data set.

```
In [5]: print(boston_data.feature_names)

['CRIM' 'ZN' 'INDUS' 'CHAS' 'NOX' 'RM' 'AGE' 'DIS' 'RAD' 'TAX' 'PTRATIO'
 'B' 'LSTAT']
```

This dataset have 506 instances(rows) and 13 attributes or parameters(columns). We will predict the housing prices in boston region using the features given. Understand the characteristics of Boston Data set Features