

## Python Project – Boston Housing Data Exploration

Each record in the database describes a Boston suburb or town. The data was drawn from the Boston Standard Metropolitan Statistical Area (SMSA) in 1970.

The attributes are defined as follows:

1. CRIM: per capita crime rate by town
2. ZN: proportion of residential land zoned for lots over 25,000 sq.ft.
3. INDUS: proportion of non-retail business acres per town
4. CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
5. NOX: nitric oxides concentration (parts per 10 million)
6. RM: average number of rooms per dwelling
7. AGE: proportion of owner-occupied units built prior to 1940
8. DIS: weighted distances to five Boston employment centres
9. RAD: index of accessibility to radial highways
10. TAX: full-value property-tax rate per \$10,000
11. PTRATIO: pupil-teacher ratio by town
12. B:  $1000(B_k - 0.63)^2$  where  $B_k$  is the proportion of blacks by town
13. LSTAT: % lower status of the population
14. MEDV: Median value of owner-occupied homes in \$1000s

To do:

- Explore the data and develop the understanding about the contents of the data.
- Document your understanding in a word document, which speaks primarily on what data contains. (Not more than 500 words and in English :P)
- Use the correct visualization to represent some key findings
- Find out 3 findings and 3 insights from the data, use adaptive visualization