**Project: Predicting Boston Housing Prices**

This project focuses on finding the best linear regression model to estimate the median value (in $1000's) of owner-occupied homes in the Suburbs of Boston.

**Dataset and inputs:**

This dataset concerns housing values in Boston suburbs. It's based on the "[Boston Housing Dataset](https://archive.ics.uci.edu/ml/datasets/Housing)" from University of California, Irvine.

The target is medv: median value of owner-occupied homes in terms of thousands of dollars ($1000s).

Features:

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(Bk-0.63)^2 where Bk is the proportion of black people by town.
13. lsat: percent lower status of the population.
14. medv: median value of owner-occupied homes in terms of thousands of dollars ($1000s).

**Techniques applied:**

I have applied linear regression technique to estimate the median value (in $1000's) of owner-occupied homes in the Suburbs of Boston.

**Conclusion:** With an R-squared value of 0.829 and a Mean Squared Error of 0.0284 model 4 is considered as best model.