# Context

Housing is one of human life's most essential needs, along with other fundamental needs such as food, water, and much more. Demand for housing has grown rapidly over the years as people's living standards improved. Housing markets have a positive impact on a country's currency, which is an important national economy scale. Housing sales prices are determined by numerous factors such as the area of the property, location of the house, the material used for construction, age of the property, number of bedrooms and garages, and so on.

There are many benefits that home buyers, property investors, and housebuilders can reap from a house-price prediction model. This model may provide a lot of information and knowledge to home buyers, property investors, and housebuilders, such as the valuation of house prices in the present market, which will help them determine house prices. Meanwhile, this model can help potential buyers decide the characteristics of a house they want according to their budget.

In this project, we will use machine learning algorithms to build the prediction model for house prices.

# Objective

The problem at hand is to predict the housing prices of a town or a suburb based on the features of the locality provided to us. In the process, we need to identify the most important features affecting the price of the house. We need to employ techniques of data preprocessing and build a linear regression model that predicts the prices for the unseen data.

# Data Dictionary

Each record in the database describes a suburb or town in a Boston. The data was drawn from the Boston Standard Metropolitan Statistical Area (SMSA) in 1970. Detailed attribute information can be found below:

**Attribute Information:**

- CRIM: Per capita crime rate by town
- ZN: Proportion of residential land zoned for lots over 25,000 sq.ft.
- INDUS: Proportion of non-retail business acres per town
- CHAS: Charles River dummy variable (= 1 if tract bounds river; 0 otherwise)
- NOX: Nitric Oxide concentration (parts per 10 million)
- RM: The average number of rooms per dwelling

- LSTAT: % lower status of the population
- MEDV: Median value of owner-occupied homes in 1000 dollars