

**Assignment 1.2** Develop a linear regression model to predict the median house value based on the California Housing details( same dataset from assignment 1.1)

Notes :

- Drop rows with missing data
- Drop ocean\_proximity column (techniques to convert values is not discussed yet)
- Drop latitude and longitude columns (these are not numerical values)
- y – value is in column median\_house\_value
- use the following six features :
  - o housing\_median\_age
  - o total\_rooms
  - o total\_bedrooms
  - o population
  - o households
  - o median\_income
  
- experiment with different learning rates and write your observations
- Compare the results with and without normalization
- Compare the results between the following scoring methods on all attributes
  - o Z-score
  - o Min-Max Scaling
  - o Robust scaling

Capture the following details from each of the experiments

Run number	Description	Most appropriate learning rate	$\frac{1}{2}$ MSE	R square	Does it converge	Number of iterations for convergence	Run specific comments
1	Without normalization						
2	With Z-score normalization						

3	With Min-max scaling						
4	With Robust scaling						