**Objective**:

The objective of the project is predict the house prices based on historical data.

**Outcome**:

Found that Random Forest works best in predicting the house prices when compared to Linear Regression.

**Tools Used**:

1. Python for data analysis

2. Scikit-learn – Linear Regression, Random Forest algorithm

**Execution**:

The data is read from a csv file.

It is analysed using histogram, heatmap to check correlation between variables.

The categorical data is converted to numerical data specifically ‘Ocean proximity’

Logarithmic transformation is applied to overcome skewness in the data.

Created two more columns 'bedroom\_ratio', 'household\_rooms' that helps in analyzing the dataset.

Linear regression and Random Forest Regressor algorithms are is applied to predict the house prices.

It is found that Random Forest Regressor works best in this scenario.