**House price prediction**

**-Linear Regression**

**Problem Statement**

In this escalating inflation period, people have three basic needs food, water and housing. It becomes a necessity to navigate through housing prices such that buyer can have a comfortable yet affordable living space. Thus, investigating the factors that impact housing prices enables buyers to make informed decisions, optimizing costs without compromising on quality of life.

**Research Question**

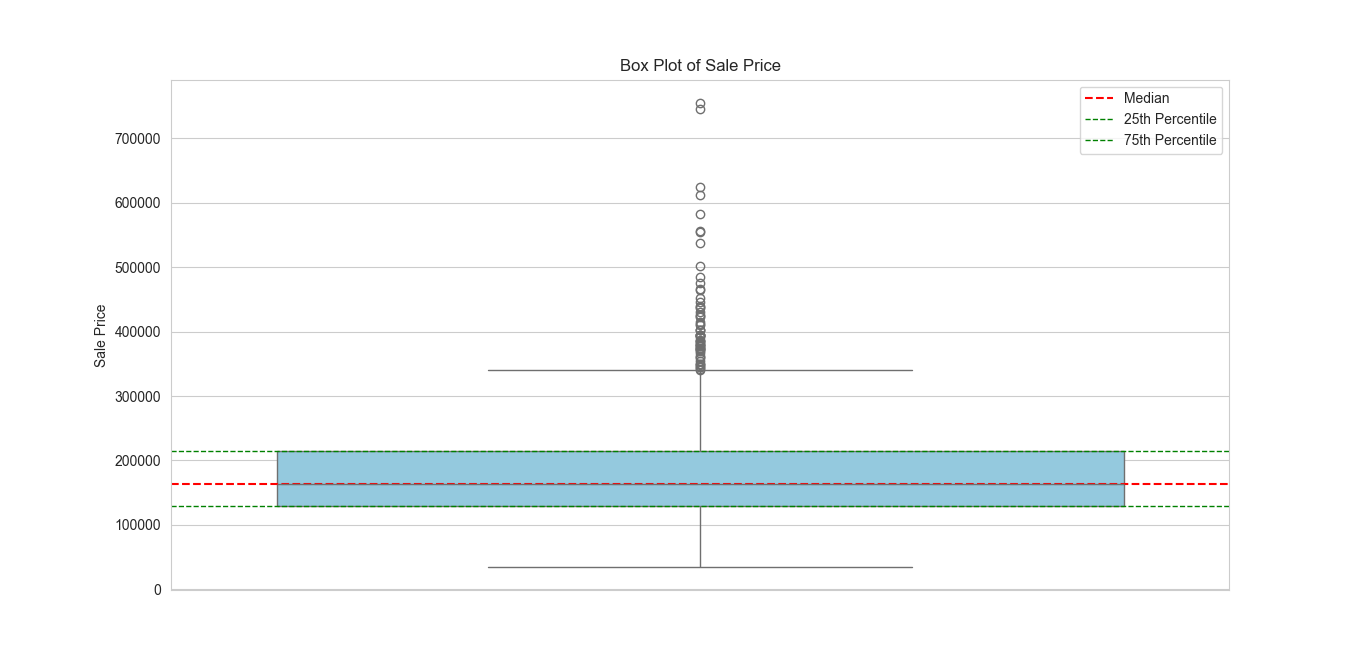
What are primary determinants influencing the housing price in the current inflationary environment.

**Assumptions**

For the scope of this project, there are handful of assumptions has been made for analyzing purposes:

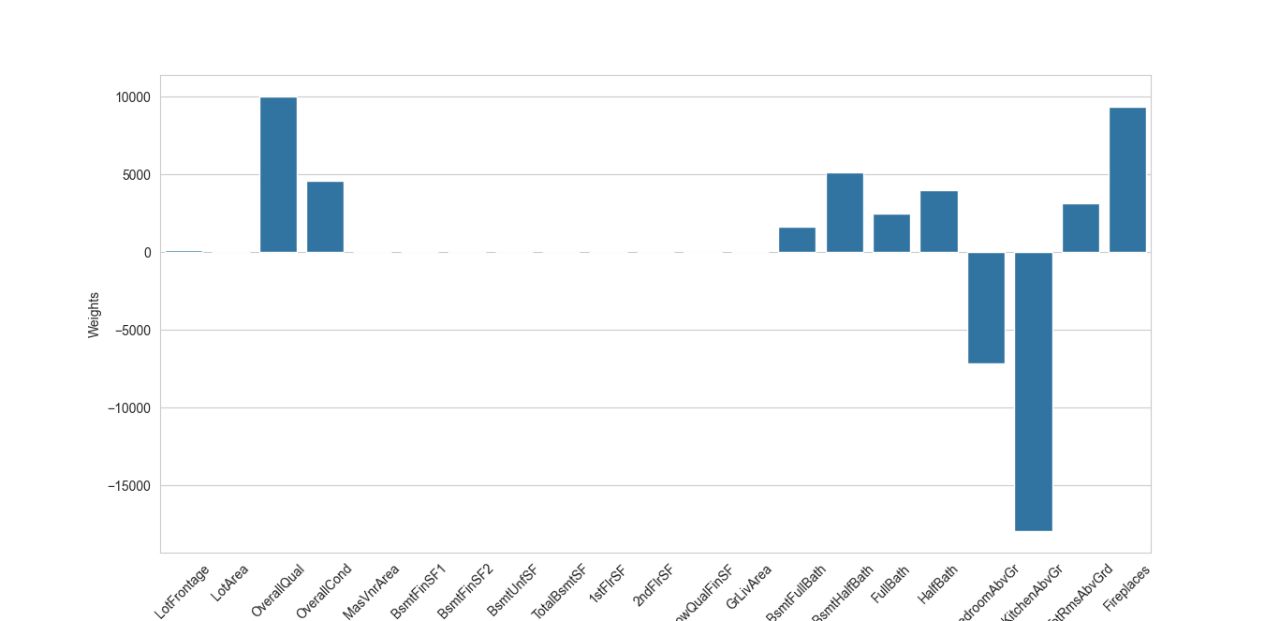
* The data has been assumed to not have many outliers which will affect the analysis drastically
* The data is valid in the current status quo.

**Distribution of the house pricing**

****

**Linear Modeling and Prediction**

For achieving the objective, linear regression model is chosen.

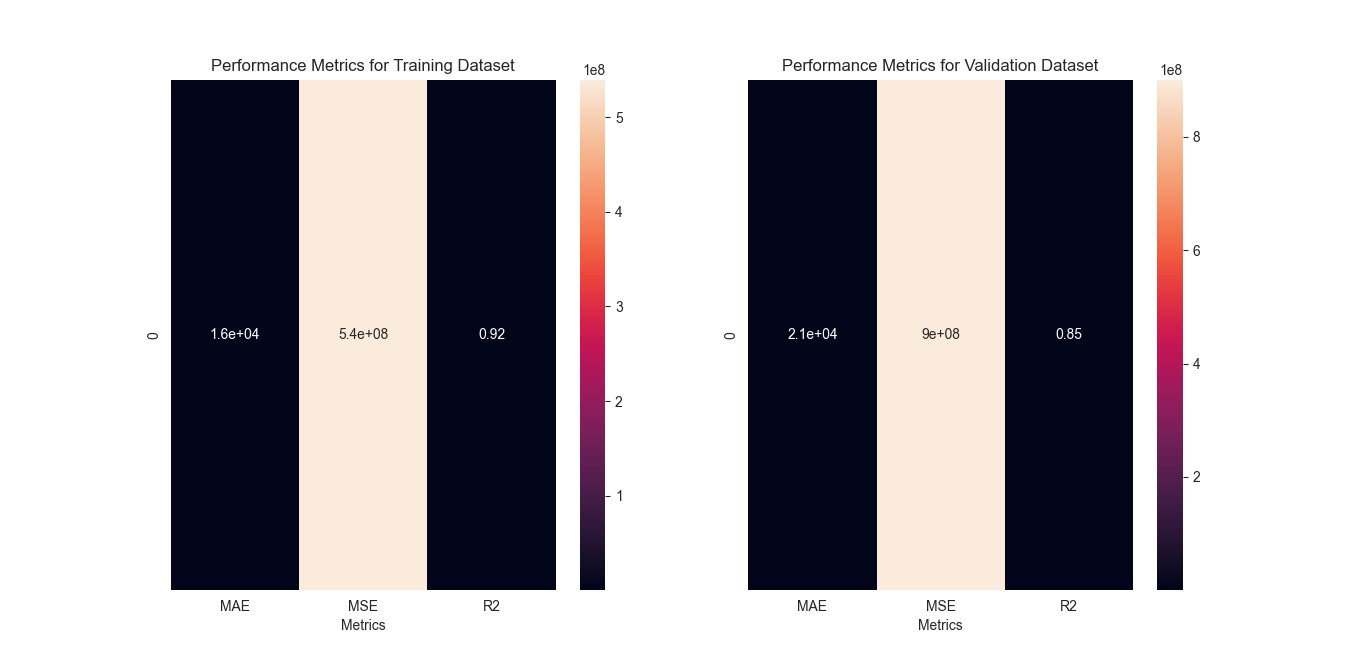


The graph showing weights assigned to first 20 features by linear regression model.

**Prediction V/S original**

Scatterplot comparing the values of prediction and actual pricing.

Approximately the majority of data points lie on the line y=x, which boosts our confidence on the regression model.

**Evaluating performance metrics for the model**

Where,

MAE is Mean Absolute error = **(1.6e+04,2.1e+04)**

MSE is Mean Squared Error= **(5.4e+08,9e+08)**

R2 is R2\_score= **(0.92,0.85)**

**Limitation of the model**

* The model does not deal with outliers, assumes the ideal case
* R2\_score has been improved but MAE and MSE is still quite large value if precision is focused