Numerical dataset

dataset name : Housing
number of classes: 10

we want to predict median_house_value
data rows : 20640
The number of samples used in training : 16512
The number of samples used in testing : 4128

dataset classes :

```
Data columns (total 10 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   longitude           20640 non-null  float64
 1   latitude            20640 non-null  float64
 2   housing_median_age  20640 non-null  float64
 3   total_rooms         20640 non-null  float64
 4   total_bedrooms      20640 non-null  float64
 5   population          20640 non-null  float64
 6   households          20640 non-null  float64
 7   median_income       20640 non-null  float64
 8   median_house_value  20640 non-null  float64
 9   ocean_proximity     20640 non-null  object
dtypes: float64(9), object(1)
```

We made 2 models

1. Linear Regression model
2. K Neighbors Regressor

Linear Regression model

Here is the model evaluation

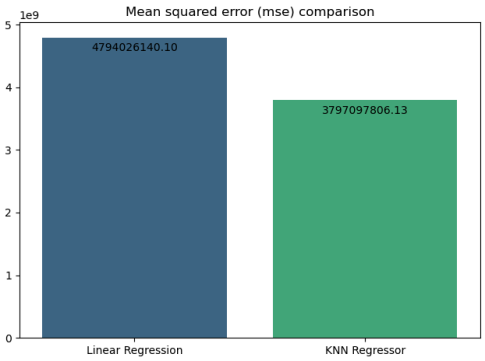
Mean Squared Error: 4794026140.10
Mean Absolute Error: 50990.17
R-Squared: 0.64

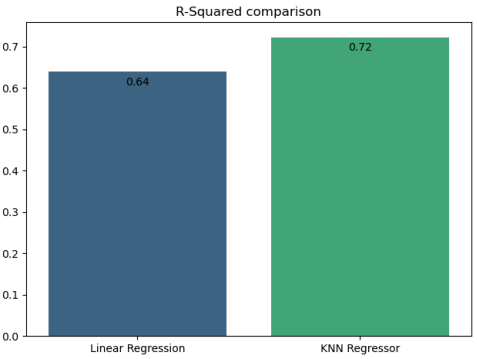
K Neighbors Regressor

Here is the model evaluation

Mean Squared Error: 3797097806.127921
Mean Absolute Error: 41750.77063953488
R-Squared: 0.722263437057932

comparison





As figure show us KNN is better