

PREDICTED HOUSE PRICES



Participated in
KAGGLE competition.
Predicted prices of
houses with
numerous factors
using supervised
learning.

BRIEF OF DATA

- ~Training set containing 80 columns and 1461 rows.
- ~Containing 6965 NA values in the dataset.
- ~Changing columns according to the description file given.

STEPS

- Changing NA values to the relevant values in the factor columns.
- Plotting graphs of relevant columns to get useful insights.
- Removing irrelevant columns using the correlation matrix.
- Removing the NA values.
- Applying the algorithms of supervised learning.

CHANGING NA VALUES IN FACTOR COLUMNS

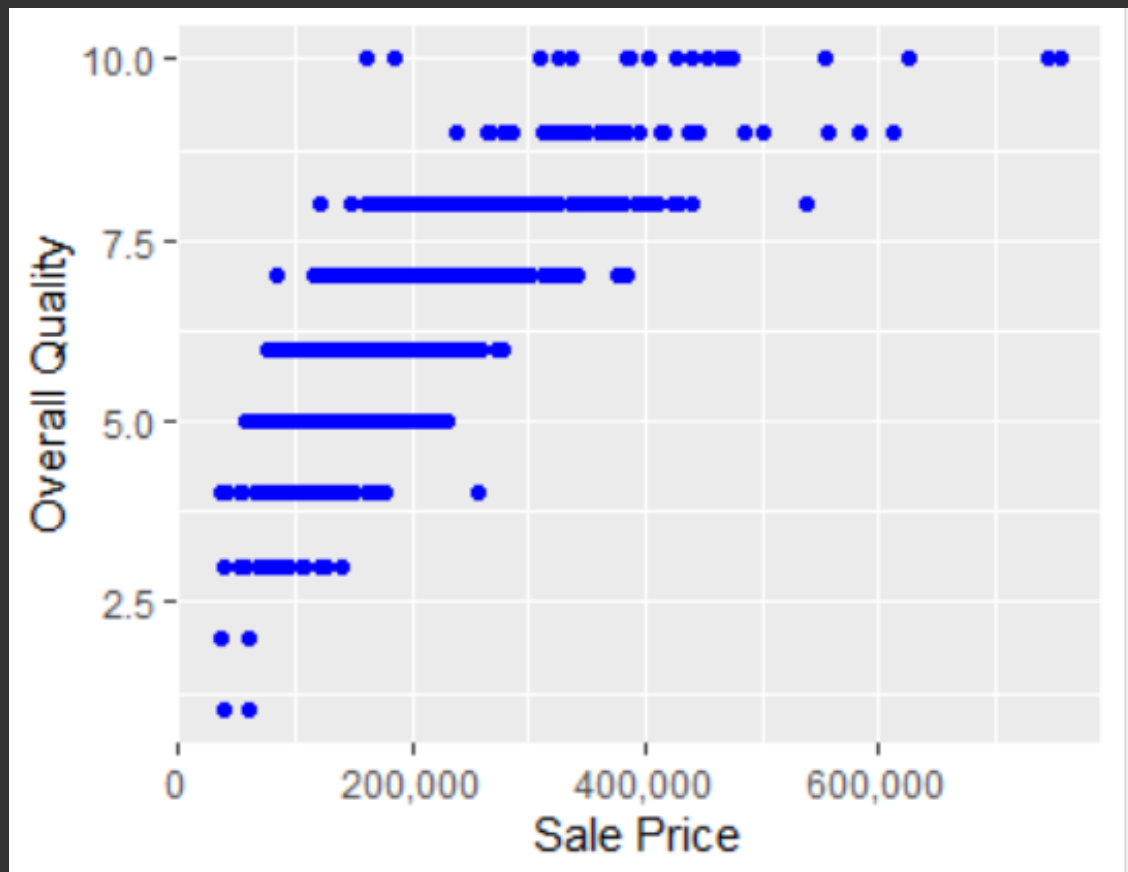
- Alley-"No Access"
- FireplaceQu-"No Fireplace"
- BsmtQual-"No~Basement"
- BsmtCond-"NoBasement"
- BsmtExposure-"NoBasement"
- BsmtFinType1-"No-Basement"
- BsmtFinType2-"No_Basement"

GarageType-"NoGarage"
GarageFinish-"No-Garage"
GarageQual-"No_Garage"
GarageCond-"No~Garage"
PoolQC-"No Pool"
Fence-"No Fence"
MiscFeature-"None"

PLOTTING SALE PRICE

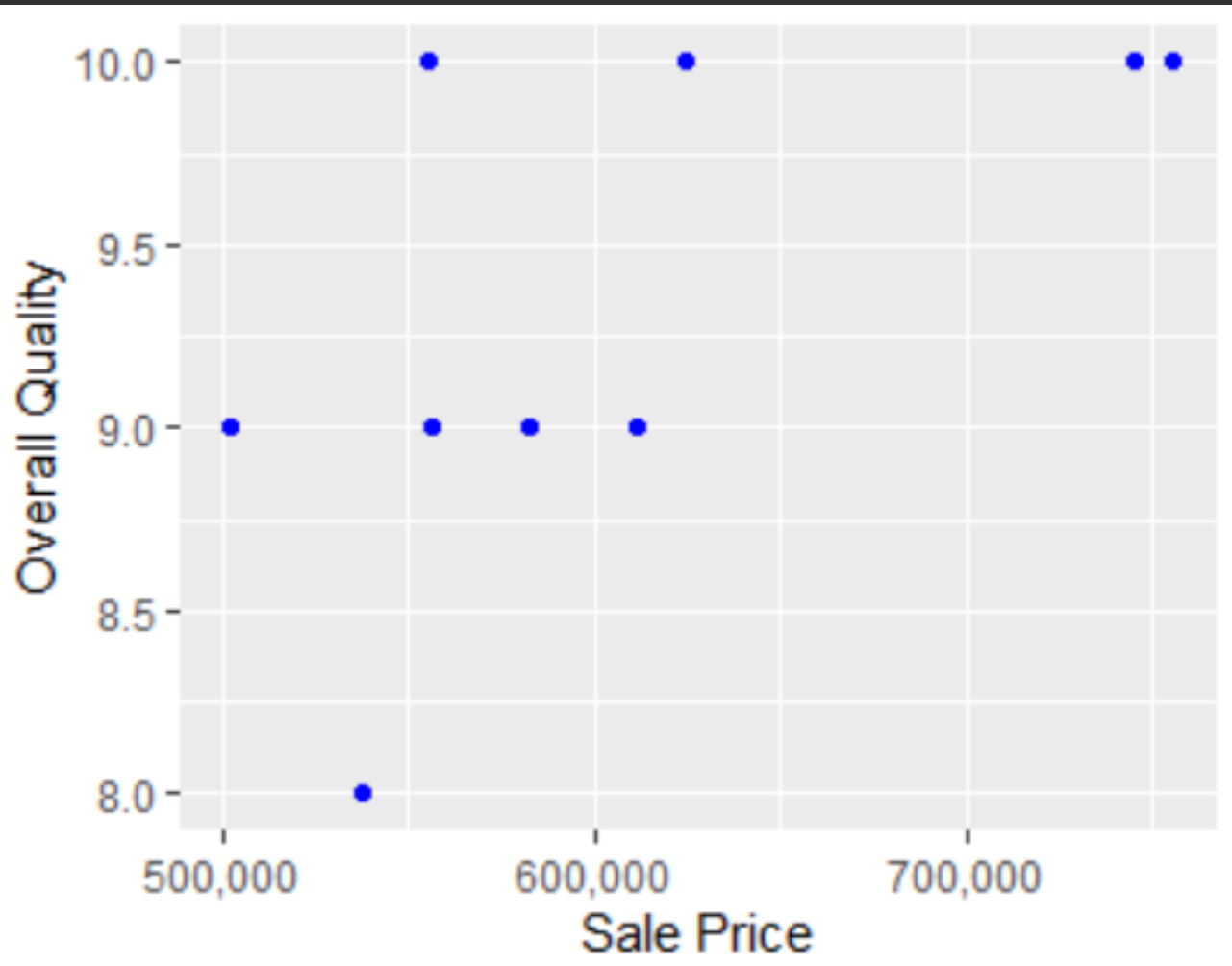


PLOTTING OVERALL QUAL AND SALES



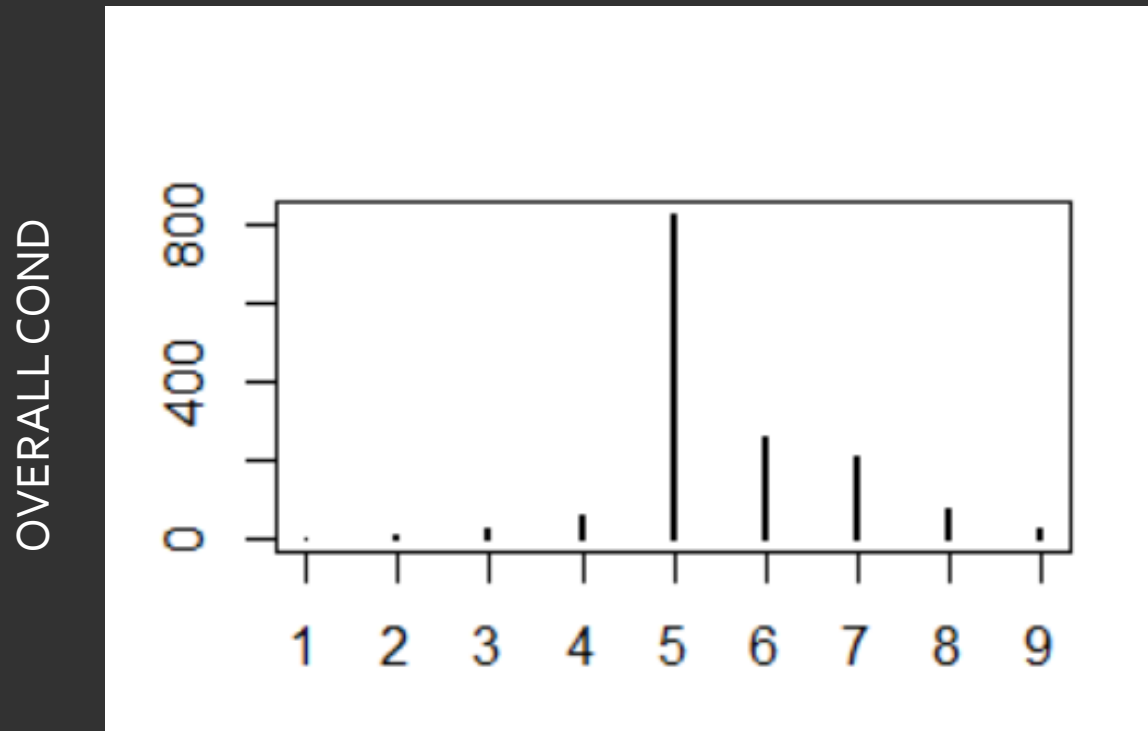
This shows quality above 8 houses were sold at very high prices.

CHECKING THE QUALITY OF HOUSES THAT ARE SOLD AT HIGHEST PRICES



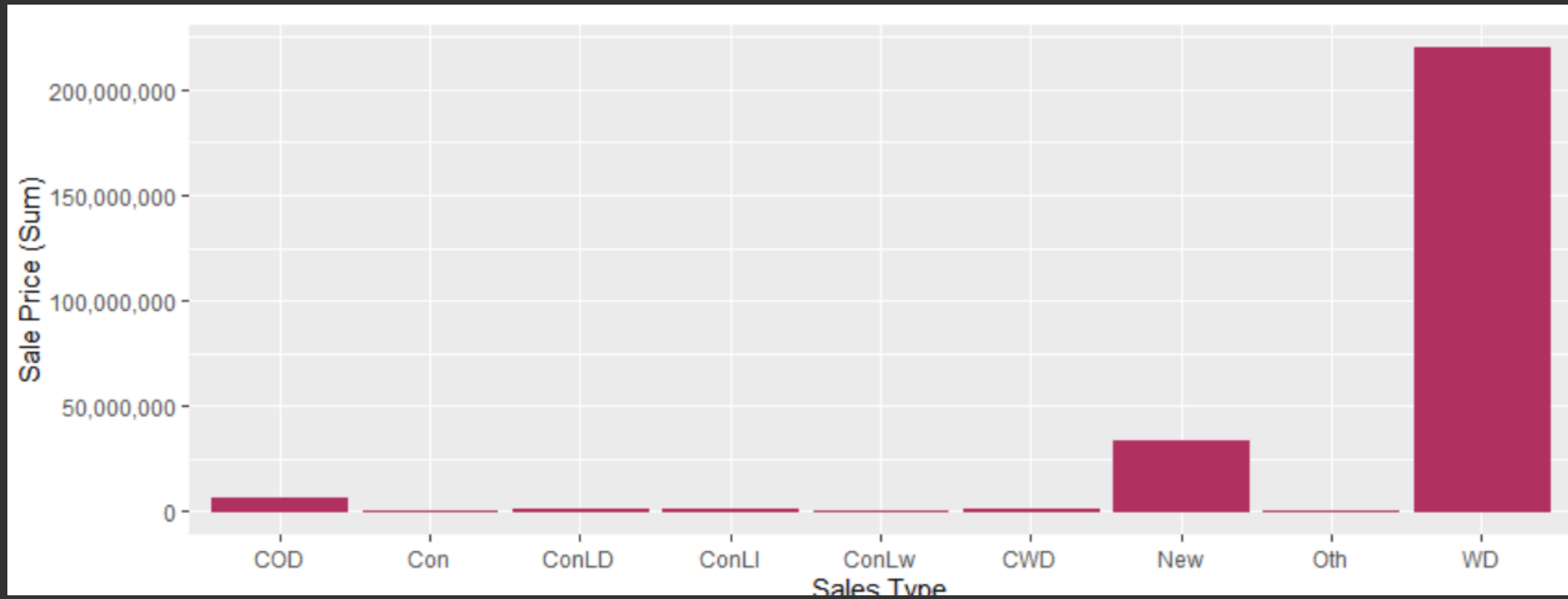
- THIS SHOWS HOUSES HAVING 9 OR ABOVE WERE SOLD AT HIGHEST PRICES.
- IN PREVIOUS VIZUALTISATION, THE CONCLUSION WAS HOUSES ABOVE 8 WERE HAVING HIGH PRICES.
- BUT THIS IS NOT THE ACTUAL CASE

PLOTTING OVERALL COND (CONDITION)



- In the data set no of houses having overall condition 5 were abundant.

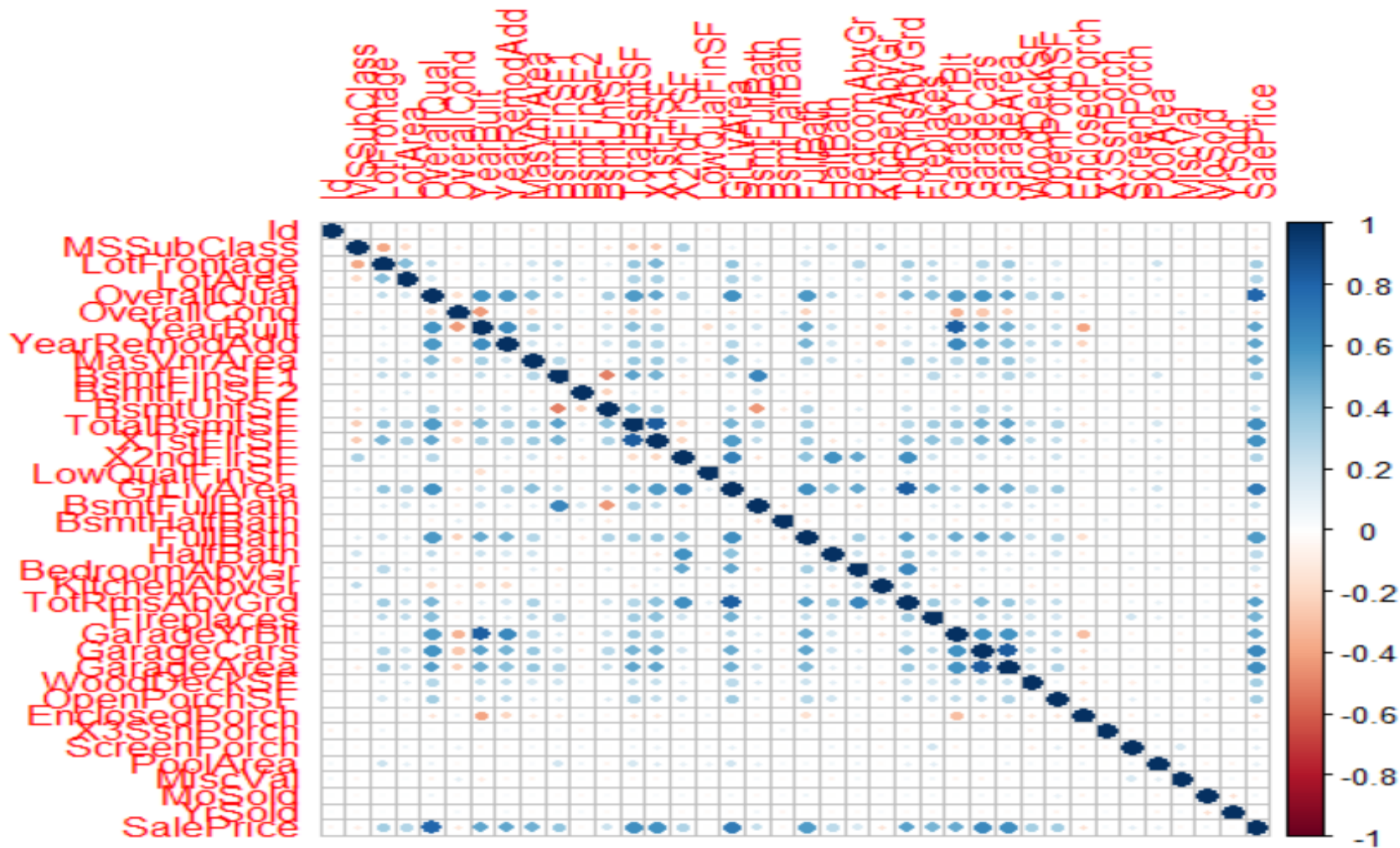
PLOTTING SALES TYPE OF HOUSES



- WD sale type contributes highest significance in price determination.

REMOVING THE NA VALUES

- USING THE OMIT FUNCTION IN R.
- TO PLOT THE CORRELATION MATRIX OF NUMERIC COLUMNS IN THE DATASET.



REMOVING COLUMNS DOES NOT SHOW
ANY MAJOR SIGNIFICANCE TO THE SALES
PRICE OF THE HOUSES.

NOW APPLYING ALOGRITHMS

LINEAR REGRESSION

I GOT A SCORE OF 0.24670
USING LINEAR RERESSION


DECISION TREES

I GOT A SCORE OF 0.24403
USING DECISION TRESS

NOTE: THESE SCORES ARE PUBLIC SCORE WHICH ARE BASED ON SOME PART
OF THE PREDICTION

PROFILE LINK(KAGGLE)

<https://www.kaggle.com/gauravss>



Gaurav

Joined 9 months ago · last seen in the past day

Following 2

Competitions Novice

[Home](#) [Competitions \(1\)](#) [Datasets](#) [Notebooks](#) [Discussion](#) ... [Edit Profile](#)

Competitions Novice	Datasets Novice	Notebooks Novice	Discussion Novice
Unranked	Unranked	Unranked	Unranked
<div><div>0</div><div>0</div><div>0</div></div>	<div><div>0</div><div>0</div><div>0</div></div>	<div><div>0</div><div>0</div><div>0</div></div>	<div><div>0</div><div>0</div><div>0</div></div>
<div>House Prices: ... Ongoing Top 84%</div>			