

1. See notebook

2. Top ten features with scores and RMSE:

Feature	RMSE Test	Score Test
OverallQual	49018.435788	0.645463
ExterQual_Coded	58260.148655	0.499175
GrLivArea	61369.555159	0.444290
KitchenQual_Coded	61622.301138	0.439703
TotalBsmtSF	62430.331676	0.424913
1stFlrSF	62876.256239	0.416668
GarageCars	63131.022101	0.411931
GarageArea	63377.710181	0.407326
BsmtQual_Coded	66831.518131	0.340970
GarageFinish_Coded	68071.654684	0.316285

3. Top ten two-pair predictors with scores and RMSE:

Feature_1	Feature_2	RMSE Test	Score Test
1stFlrSF	OverallQual	43764.688318	0.717388
OverallQual	1stFlrSF	43764.688318	0.717388
TotalBsmtSF	OverallQual	44577.664096	0.706791
OverallQual	TotalBsmtSF	44577.664096	0.706791
OverallQual	GrLivArea	45673.467009	0.692199
GrLivArea	OverallQual	45673.467009	0.692199
GarageArea	OverallQual	46468.134901	0.681395
OverallQual	GarageArea	46468.134901	0.681395
KitchenQual_Coded	OverallQual	47070.072304	0.673087
OverallQual	KitchenQual_Coded	47070.072304	0.673087
OverallQual	GarageCars	47124.910052	0.672325
GarageCars	OverallQual	47124.910052	0.672325
OverallQual	ExterQual_Coded	47433.824631	0.668015
ExterQual_Coded	OverallQual	47433.824631	0.668015
OverallQual	BsmtQual_Coded	48237.754631	0.656666
BsmtQual_Coded	OverallQual	48237.754631	0.656666
GarageFinish_Coded	OverallQual	48571.271663	0.651902
OverallQual	GarageFinish_Coded	48571.271663	0.651902
OverallQual	OverallQual	49018.435788	0.645463
1stFlrSF	ExterQual_Coded	49640.643171	0.636405

4. Single model with all features:

```
The model performance for testing set
-----
RMSE is 35623.30345000588
Score is 0.8127547098528062
```

**Observation: the 2 pair prediction was, at best, 0.72, but using all features resulted in a score of 0.81.**

5. 5NN and 10NN

5NN:

```
The model performance for testing set
-----
RMSE is 51966.07983368139
Score is 0.6015421069563245
```

10NN

```
The model performance for testing set
-----
RMSE is 52392.16317256475
Score is 0.5949812006956909
```

6.

**Since the score is better with linear regression, it is a better tool for inference**