Problem Statement:

Which features are likely to impact the price of houses in Ames, Iowa and where are they located?
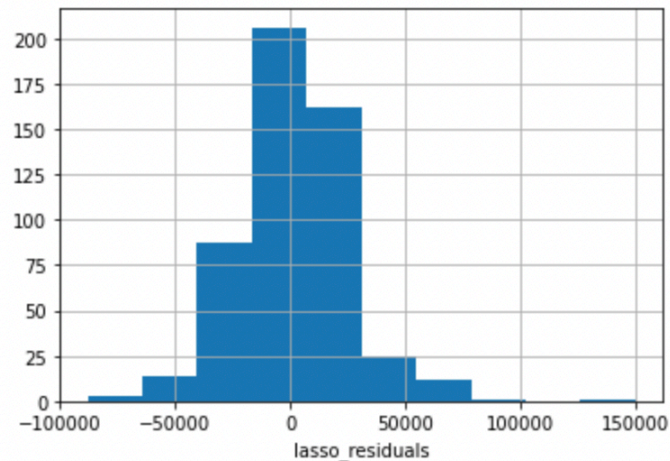
Contents

Modelling and Prediction
Key Features
Conclusion and Recommendation

Modelling and Prediction

Comparing the Lasso and Ridge models , I would prefer to use the Lasso Model.
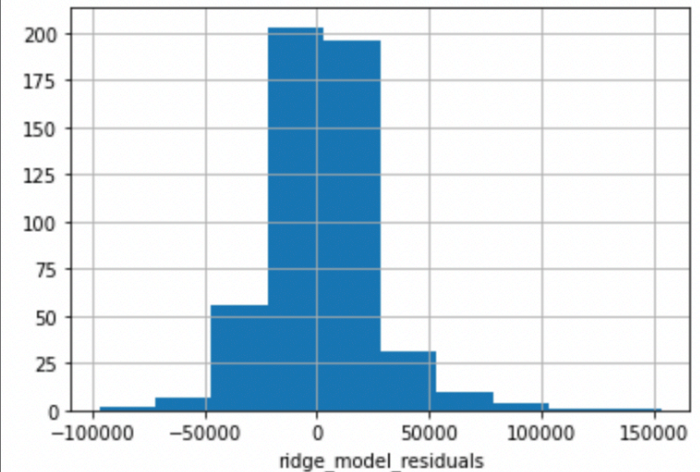


```
lasso cross_val_score 0.8836230252908578
lasso train r2 = 0.9174563356200337
lasso test r2 = 0.9055230240099783

Text(0.5, 0, 'lasso_residuals')
```



```
ridge_model cross_val_score 0.8820418938868252
ridge_model train r2 = 0.9159015377389182
ridge_model test r2 = 0.900507077461408

Text(0.5, 0, 'ridge_model_residuals')
```

Actual vs. Predicted Sale Price

- The predicted price that is coming up for the Test Dataset once the Lasso Model was applied on it.

	hicoef_feat	coef	abs_coef
0	overall_qual	28925.662296	28925.662296
2	kitchen_qual	14008.614555	14008.614555
1	total_bsmt_sf	13871.897538	13871.897538
7	lot_area	10886.930733	10886.930733
4	neighborhood_NridgHt	9931.718482	9931.718482
6	garage_area	9927.083796	9927.083796
11	neighborhood_NoRidge	9689.745146	9689.745146
3	neighborhood_StoneBr	8571.608779	8571.608779
5	sale_type_New	5690.419318	5690.419318
8	bsmt_exposure	5356.730456	5356.730456
9	bldg_type_TwnhsE	-5321.998985	5321.998985
13	neighborhood_Crawfor	4810.512051	4810.512051
12	screen_porch	4682.486610	4682.486610
10	neighborhood_GrnHill	2496.595635	2496.595635

Key Features

- These features are the key features in the model that cause for 90% ($R^2 = 0.90$) of variance in our data

Conclusion and Recommendations

- ***Most Valuable Features***: overall quality of the house, total basement square feet, kitchen quality and lot area

- ***Least Valuable Features***: misc. value, fireplaces, id(which seems like postal id), the time when property was sold (as in older property has lesser price)

- ***Fireplaces:*** The more quality fireplaces a house has, the higher the value, as Ames has low temperature range that means fireplace is essential.