

Question 1

What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer:

- The optimal value of alpha for
 - o Ridge is 9
 - o Lasso is 100

```
# Printing the best hyperparameter alpha
print(model_cv.best_params_)
3] ✓ 0.0s
· {'alpha': 9.0}
```

```
# Printing the best hyperparameter alpha
print(model_cv.best_params_)
40] ✓ 0.0s
· {'alpha': 100}
```

For both: Changing the values did not affect the coefficients.

Top Features are:

Neighborhood_NoRidge, Neighborhood_NridgHt, OverallQual, overallQual
Neighborhood_Veenkar

Question 2

You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?

I will choose Lasso as its giving feature selection option also. It has removed unwanted features from model without affecting the model accuracy. Which makes are model generalized and simple and accurate.

Question 4

How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?

To make model robust and generalisable 3 features are required:

1. **Model accuracy** should be $> 70-75\%$: I our case its coming 80% (Train) and 81% (Test) which is correct.
2. **P-value** of all the features is < 0.05
3. VIF of all the features are < 5