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 double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer:

In case of Ridge optimal value for alpha : 5
In case of Lasso optimal value for alpha : 100

After doubling the alpha following metrics change and the new coefficients are as follows:

Ridge:

```
Train R2 : 0.8754766612309398
Test R2  : 0.8415845938340395
```

Doubled Alpha R2

```
Train R2 : 0.8627996682804974
Test R2  : 0.8335417800135897
```

Lasso:

```
Train R2 : 0.8741957138330694
Test R2  : 0.8501297369349572
```

Doubled Alpha R2

```
Train R2 : 0.8609057091768293
Test R2  : 0.8403681385501198
```

In both Ridge and Lasso, the **alpha** parameter controls the "strength" of regularization. Increasing alpha adds more penalty to the complexity of the model.
When doubled alpha, we observed a consistent trend across both models:

- **Train R2 decreased:** This happens because the model is now "constrained" and cannot fit the training data as tightly as before.
- **Test R2 decreased:** While regularization is meant to improve generalization, doubling it in this specific case seems to have pushed the models towards "simplicity" (increased bias) and less flexible.

Original alpha was a better fit in terms of keeping bias variance trade off at optimal level. Doubling alpha is destabilising the model.
Most important number of Predictor variables remained the same for both the models:

```
['GrLivArea', 'Neighborhood_NoRidge', 'GarageArea',
'Neighborhood_NridgHt', 'KitchenQual_TA']
```

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?

Answer : Looking at the results , Lasso model seems to be best fit because of multiple reasons as below:

- Effective Feature Selection and much more interpretable.
- It provides the best balance—it is complex enough to capture the patterns in your data (low bias) but regulated enough to ignore the noise (low variance)
- With a Test R2: **0.850**, the Lasso model explains 85% of the variance in the test data, which is higher than that of the Ridge model.
- With Ridge model which retains all the features, although shrinking the coefficients to 0 ,still it is less interpretable from the business point of view.

Question 3

After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?

Answer:

After removing top 5 most predictors from the dataset below are new predictors in the new lasso fit model:

```
1stFlrSF        244837.972443
2ndFlrSF        132385.160749
BsmtQual_Gd     -35107.867281
MasVnrArea       34866.612652
BsmtQual_None   -34271.106271
```

Question 4

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

Answer:

To ensure model is robust and generalised , the model needs to learn the p
atterns in the data rather than memorising the data or noises in data itse
lf. This can ensured by the making the simpler model which has optimal bia

s variance tradeoff. This can be implemented by techniques like train-test split , cross validation, regularisation.

Implication on accuracy will be that we will get a slightly less accurate model but this is much better rather than overfitting the model which fails to perform well on unseen data. Stable model performs better compared to overfitted model or very simple model.