**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:
For Ridge the alpha value is 10
For Lasso, the alpha value is .001
The most important predictor variable after change is implemented is: **1stFlrSF**

**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:
Although both the models performances are almost the same, I will choose lasso, as its test R2 score is : 0.8778 as compared to 0.8774 for ridge

**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:
Top 5 predictor variables after removing previous ones are:
FullBath
GarageArea
Fireplaces
LotArea
ExterQual

**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?

Answer:
Occam's Razor: Make a model as simple as possible but not any simpler.
Model can be made simple by reducing the number of features,regularization techniques (Ridge and lasso) etc.

**Simpler models are more generic:** Instead of learning the training data very strongly, simpler models learn the inherent trend/dependencies present in the model.
**Robustness:** A more complex model with many features and extreme coefficients learn training data completely and perform poorly on test data. A slight change in training data will change the model parameters completely. Simpler models have **low variance.**

**Disadvantages of simple models:**
Very simple models can be too naive and might give a low R2 score or RMSE and may not be able to explain all variance in data. Too simple models might underfit the data. Simpler models have **high bias.**

**Regularization techniques (Ridge and Lasso) can be used to optimize this bias and variance trade off. We can choose value of alpha to adjust model complexity**