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?

The optimal value of alpha for Lasso Regression is 0.01.
The optimal value of alpha for Ridge Regression is 2.

If we double the value of alpha in Ridge Regression, then the model will apply the penalty to the curve and try to make the model more generalize and simple which in turn generates more error for the test and train dataset.

If we double the value of alpha in Lasso Regression, then the model gets penalized which causes more coefficients to become 0, which causes the r-squared error to go up.

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?

The Mean Squared Error and Mean Absolute Error of both the models are almost same. MSE and MAE values for Ridge & Lasso regression are 0.0018 but since Lasso helps in feature reduction (as the coefficient value of some of the features become zero), Lasso has a better edge over Ridge and should be used as the final model.

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?

Five important predictor variables are: **GrLiveArea, OverallQual, OverallCond, TotalBsmtSF, GarageArea**

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?

As Per Occam's razor model should be as simple as necessary. So according to above expectation simple model have an edge over other complex models. The advantages of simple model are as below:

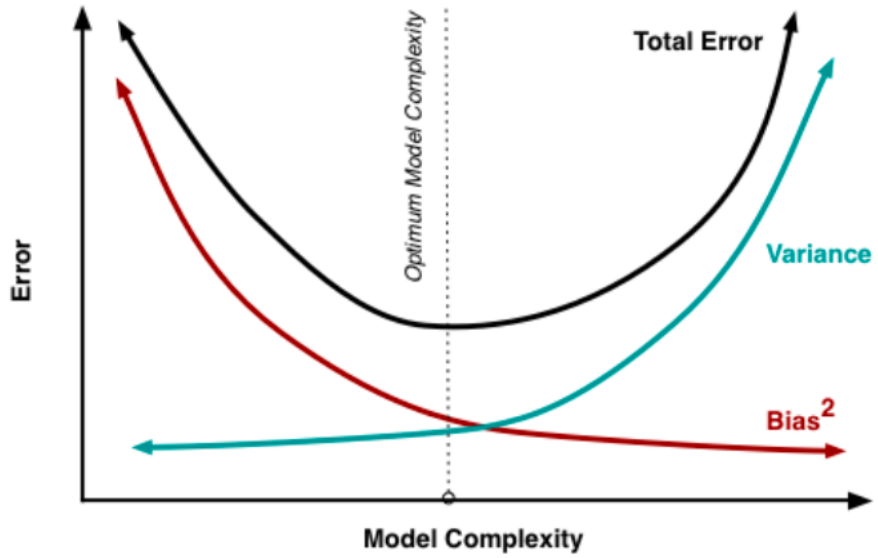
- Generalizability
- Robustness
- Making few assumptions
- Less data is required for learning

Definition:

Variance = How sensitive is model to the training data. This refers to consistency of the model.

Bias = Accuracy of the data on unseen future data.

Making a model simple lead to Bias-Variance Trade-off:



Adding to many predictor variables in the model may lead to complex model. Complex model deteriorates the performance of the model (r2 score). Complex model introduces problem of overfitting where model memorized the data and is not generalized. When such model is evaluated against the unseen data the performance is very poor.