**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?

**Ans :**

Obtained value from existing models

Optimal value of lambda for ridge - 40

Optimal value of lambda for lasso – 0.002

Top 5 features : Neighborhood\_NoRidge, Neighborhood\_NridgHt, OverallQual, GrLivArea, Neighborhood\_Crawfor

If we multiply the values by 2

Optimal value of lambda for ridge - 80

Optimal value of lambda for lasso – 0.004

Ridge :

1. Mean square error slightly increases with the new values
2. R2 value decreased by 1% on training data
3. Increase in coefficient values observed

Lasso :

1. Mean square error slightly increases with the new values
2. R2 value decreased by 1% on training data
3. Change is coefficient values observed, some coefficients are now zero
4. Top Features : Neighborhood\_NoRidge, Neighborhood\_NridgHt, OverallQual, GrLivArea, ExterQual

**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?

**Ans :**

Optimal value of lambda for ridge - 40

Optimal value of lambda for lasso – 0.002

Lasso would be a better choice as lasso will penalize more on the dataset and can also help in feature elimination.

**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?

**Ans :**

Top 5 features are Neighborhood\_NoRidge, Neighborhood\_NridgHt, OverallQual, GrLivArea and Neighborhood\_Crawfor

Current Model accuracy is 82%

New model accuracy will be 80%

New Top 5 features will be : 2ndFlrSF, 1stFlrSF, Exterior2nd\_ImStucc, ExterQual, KitchenQual

**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?

**Ans :**

To make model robust and generalisable

1. Model accuracy should be > 70-75%: The created Lasso model has accuracy of 83% on training data and 82% on test data
2. Mean square error should be closer to zero, model has 0.18(for training data) and 0.17(for test data) value

which is correct.

Thus we are sure that model is robust and generalisable