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: After analysis optimal value of alpha for Ridge is 1 and for Lasso is 0.001. Id we double the value of alpha for Ridge then coefficients gets decreased. In Lasso it also decreases and some time it turned out to be 0 as well.

The most important predictor variables are –

1. OverallQual
2. GrLivArea
3. TotalBsmtSF
4. GarageCars
5. FullBath

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 optimal value for Lasso regression is 0.001 because mean_train_score is 0.88 and mean_test_score is 0.84 at this value

The optimal value for Ridge is 1 because mean_train_score is 0.88 and mean_test_score is 0.84 at this value.

Since the accuracy is almost same for both Lasso and Ridge. But will choose Lasso because here we can do feature selection as well.

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?

The most important predictor variable are –

1. OverallQual
2. GrLivArea
3. TotalBsmtSF
4. GarageCars
5. FullBath

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?

The generalised model is that one which performs well on unseen data. In this case bias and variance should be low as possible. We can say a model is robust when variation doesn't affect on its performance.

When a model performs well on train dataset but fails to perform well on test dataset then it is overfit model. So, to deal with overfit we use regularization. So, Lasso and Ridge comes into picture. There we add penalty term to deal with coefficients on different alpha values.