Question 1
What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you cho ose double the value of alpha for both ridge and lasso?
What will be the most important predictor variables after the change is implemented?

Ans: Optimal value for ridge is 0.01
Optimal value for lasso is 0.0001
Ridge:
r2-square on training data: 0.953
r2-square on test data: 0.858
MSE on training data: 38837386258.63372
MSE on test data: 39495635164.90639

Lasso:
r2-square on training data: 0.942
r2-square on test data: 0.866
MSE on training data: 0.00914
MSE on test data: 0.0220

After changes in alpha,
- variance test and train r2-square of lasso, test Mean square error are decreased. MSE of test also is decreased.
- Variance of test and train r2-square is same and Mean square error are decreased drastically
Ridge:
train r2-score: 0.952
test r2-score: 0.861
train mean squared error: 0.00743
test mean squared error: 0.0228
most important predictor variables: RoofMatl_WdShngl, RoofMatl_Membran, RoofMatl_Tar&Grv, TotalBs mtSF, RoofMatl_Roll
Lasso:
train r2-score: 0.927
test r2-score: 0.865
train mean squared error: 0.0113
test mean squared error: 0.022
most important predictor variables: TotalBsmtSF, 2ndFlrSF, GarageArea, RoofMatl_WdShngl, MSZoning_ RL

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: Lasso regression is chosen with alpha value: 0.0002, since it has a very low MSE and trying to explain most of features.

Question 3
After building the model, you realised that the five most important predictor variables in the lasso model are not avai lable 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: After removing and building the model: 5 most important features are OverallQual_9, LotArea, SaleType_Con LD, RoofMatl_WdShake, RoofMatl_CompShg

Question 4
How can you make sure that a model is robust and generalisable? What are the implications of the same for the accu

racy of the model and why?
Ans: We can generalize the model based on the test and train metrics variance. If these metrics variance is less.