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:
a. Optimum value of alpha in ridge = 1.0
   Optimum value of alpha in lasso = 0.001

b. Lasso :
   - Train and Test score of Lasso decreased
   - model gets penalised with more coefficients made zero
   - model becomes more simpler

   Ridge :
   - Ridge Train score decreased , Test score increased
   - some of coefficients moving toward zero , penalty applied
   - model becomes simpler

c. Lasso :
   - GarrageCars_3 , FullBath_3 ,OverallQual_8 ,OverallQual_9 ,2ndFlrSF

   Ridge :
   - BsmtFinSF1 , TotRmsAbvGrd_10 , FullBath_3 ,OverallQual_9, LotArea

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:
Would be choosing lasso coefficient and alpha of lasso regression which would make model simple and feature selection

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:
MasVnrArea , 2ndFlrSF , OverallQual_8 , Neighborhood_NridgHt , BsmtExposure_Gd

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:

- Even after removing the variables the model accuracy hasn't decreased drastically (it is in >65%)
- With penalty assigned to the coefficients the model became simpler and few coefficients were removed
- With cross validation we eliminated the overfitting