"Subjective 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 --- The optimal value is those value which satisfy the constraints with minimum and maximum values. Optimal value function is used across different field and in regression.

Alpha is one of the imp parameters of optimal value, in general value of alpha is 0<1 (Greater than 0 but less than 1) or 0 <= Alpha <=infinity

As per assignments, Optimal alpha value for Ridge regression is = 20
Lasso regression is = 0.0001

When we choose to double the alpha for Ridge than model will be generic and the might be no fit for each data set.

In case of double of alpha for Lasso than model variable will aligns towards zero.

Imp Predictor variables are - Garage Area, Pool Area, Sale type, Full Bath , Half Bath.

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 -- I will choose the Lasso regression model based on optimal value of the lambda.
Lasso does pretty well, in terms of lambda calculation compare to Ridge.

Lasso tends to zero therefore very less variable tends to influence the model, when lambda is large enough.

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-- 5 most important variables are –

A. Garage Area.
B. BedroomAbvGR.
C. Kitchen AbvGR.
D. Pool Area.

E. TotalBSMTSF.

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 -- In order to make model robust and generalisable it should be simple and not complex. It should you less variable, so that less variable or error.