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?

Optimal values are:
1. Ridge - 2.0
2. Lasso - 0.001

If we double the alpha values for both models, we are getting following results from old to new alpha metrics

LinearRegression()	Ridge(alpha=2.0)	Lasso(alpha=0.001)	
R2 Score (Train)	9.547200e-01	0.93727	0.91020
R2 Score (Test)	-3.084162e+20	0.89662	0.89268
RSS (Train)	6.573600e+00	9.10616	13.03574
RSS (Test)	2.702579e+22	9.05861	9.40408
MSE (Train)	6.720000e-03	0.00931	0.01333
MSE (Test)	5.607011e+19	0.01879	0.01951

LinearRegression()	Ridge(alpha=4.0)	Lasso(alpha=0.002)	
R2 Score (Train)	9.547200e-01	0.93143	0.88996
R2 Score (Test)	-3.084162e+20	0.89537	0.87326
RSS (Train)	6.573600e+00	9.95410	15.97446
RSS (Test)	2.702579e+22	9.16813	11.10584
MSE (Train)	6.720000e-03	0.01018	0.01633
MSE (Test)	5.607011e+19	0.01902	0.02304

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?

Lasso has outperformed the other models. And lasso is simpler along with it also has feature selection

Question 3

After building the model, you realized 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?

Current Lasso top 5 features are:
1. GrLivArea
2. building_remod_age
3. OverallQual
4. 1stFlrSF
5. LotArea

After dropping these top 5 we have following as Lasso's top 5 features:
1. TotRmsAbvGrd
2. GarageCars
3. FullBath
4. building_remod_age
5. Fireplaces

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?

To make the model more generalisable, we need to avoid overfitting and multiple complex transformations on the data. A simpler model is always generalizable and complex model are prone to overfit and to regularize the overfitting we can leverage regularized models like Ridge and Lasso.