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 of alpha for Ridge is 7 and for Lasso it is 100. With these alphas the R2 of the model was approximately 0.839. After doubling the alpha values in the Ridge and Lasso, the prediction accuracy remains around 0.835 but there is a small change in the co-efficient values. The new model is created and demonstrated in the Jupiter notebook. Below are the changes in the co-efficients.

For ridge :

Coeff values are deceresing as alpha will increase.r2_score of train data is also drop from .839 to 0.835

For lasso:

As alpha value increse But r2score is also dropped by .839 to .833

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 helps in feature reduction (as the coefficient value of some of the features become zero), Lasso has a better edge over Ridge and should be used as the final model.

Lasso as its giving feature selection option also. It has removed unwanted features from model without affecting the model accuracy. Which makes are model generalized and simple and accurate.

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?

ANS : Top 5 features are Neighborhood_NoRidge,
Neighborhood_NridgHt,
2ndFlrSF,
OverallQual,
Neighborhood_Veenker.

After dropping them model accuracy reduced from 84% to 72% . Now top most features are:
Next top 5 features after droping 5 main predictors
1stFlrSF,
MSSubClass_90,
MSSubClass_120,
TotalBsmtSF,
HouseStyle_1Story

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?
ANS :
To make model robust and generalisable 3 features are required:

1. Model accuracy should be > 70-75%: In our case its coming 80%(Train) and 81%(Test) which is correct.

2. P-value of all the features is < 0.05

3. VIF of all the features are < 5 Thus we are sure that model is robust and generalisable.

As Per, Occam's Razor— given two models that show similar 'performance' in the finite training or test data, we should pick the one that makes fewer on the test data due to following reasons:-

- Simpler models are usually more 'generic' and are more widely applicable
- Simpler models require fewer training samples for effective training than the more complex ones and hence are easier to train.
- Simpler models are more robust.
  - Complex models tend to change wildly with changes in the training data set
  - Simple models have low variance, high bias and complex models have low bias, high variance
  - Simpler models make more errors in the training set. Complex models lead to overfitting — they work very well for the training samples, fail miserably when applied to other test samples

Therefore, to make the model more robust and generalizable, make the model simple but not simpler which will not be of any use