Question 1

Ques: What is the optimal value of alpha for ridge and lasso regression?

Ans: The optimal value of alpha for Ridge is 15 and for Lasso is 0.001

Ques: What will be the changes in the model if you choose double the value of alpha for both ridge and lasso?

Ans: R2 Score decreases, RSS increases, MSE increases and RMSE increases

Ques: What will be the most important predictor variables after the change is implemented?

Ans:

For Ridge

```
OverallQual             0.085266
Neighborhood_Crawfor    0.066386
GrLivArea               0.056770
Neighborhood_Somerst    0.053916
Neighborhood_NridgHt    0.051513
Condition1_Norm         0.049232
OverallCond             0.041963
GarageCars              0.039767
BsmtExposure_Gd         0.038566
Exterior1st_BrkFace     0.037650
Name: Ridge, dtype: float64
```

For Lasso

```
Neighborhood_Crawfor    0.160802
ExterCond_Other         0.142087
Neighborhood_StoneBr    0.132370
Neighborhood_NridgHt    0.130550
GarageQual_Other        0.115046
Neighborhood_Somerst    0.107802
LandContour_Low         0.100008
SaleType_Other          0.097972
Neighborhood_ClearCr    0.097446
Exterior1st_BrkFace     0.095600
Name: Lasso, dtype: float64
```

Question 2

Ques: 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: Both Ridge and Lasso has similar scores in the metrics but I will choose Lasso because we have too many features and we need feature selection as well.

Question 3

Ques: 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?

2ndFlrSF	0.106822
1stFlrSF	0.095199
Condition1_Norm	0.069749
BsmtExposure_Gd	0.056734
Exterior1st_BrkFace	0.052738

Question 4

Ques: 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: The model is robust and generalised when it performs good on unseen data. R2 score on test data will tell whether the model underfit or overfit. Regularization helps to build the model robust and generalised.