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?

1. For Ridge regression, it is 10. For Lasso regression, it is 0.001
2. On doubling for both, scores for both the cases reduced.
   Following are the important predictor variables, in that order:
       i. Ridge: OverallQual, Neighborhood_Crawfor, Neighborhood_NridHt, GrLivArea, Neighborhood_Somerst
       ii. Lasso: GrLivArea, OverallQual, Neighborhood_Crawfor, Neighborhood_Somerst, Neighborhood_NridHt

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?
Comparing both, Lasso gives better scores with reduced variables to handle.

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?

Removing Neighborhood, OverallQual, GrLivArea variables from the dataset and re running the code, following are the observations:

1. Ridge: (0.106, 'MSZoning_FV'),
```
 (0.101,  '2ndFlrSF'),
 (0.093,  '1stFlrSF'),
 (0.068,  'Exterior1st_BrkFace'),
 (0.065,  'Functional_Typ'),
```
2. Lasso: (0.108, '2ndFlrSF'),
```
 (0.106,  'MSZoning_FV'),
 (0.098,  '1stFlrSF'),
 (0.064,  'YearBuilt'),
 (0.064,  'Exterior1st_BrkFace'),
```

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 robust and generalisable, following could be considered:

1. Weighing each predictor variable differently
2. Choosing methods that are less sensitive to outliers

More generalised the model is, more likely it will not handle a particular use case effectively. A balance may be needed between generalisation and accuracy.