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?

Answer:
Optimal Value of alpha are as below:

- For Ridge: 100
- For Lasso: 0.001

Below are the observations when we doubled the value of alpha:

Original:

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.903850	0.905287
1	R2 Score (Test)	0.863691	0.861922
2	RSS (Train)	15.231463	15.003830
3	RSS (Test)	10.128449	10.259895
4	MSE (Train)	0.122140	0.121224
5	MSE (Test)	0.152067	0.153050

When alpha doubled:

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.899197	0.901286
1	R2 Score (Test)	0.865545	0.866069
2	RSS (Train)	15.968639	15.637783
3	RSS (Test)	9.990682	9.951741
4	MSE (Train)	0.125061	0.123758
5	MSE (Test)	0.151029	0.150734

As we can see from above, R2 for Test got slightly increased. Also, RSS for train got increased. Coefficients also got slightly decreased. Since alpha values are small so variation is small. BedroomAbvGr and Fireplaces came under the most important factors. Also, OverallCond was removed from Ridge and Neighborhood_StoneBr was removed from Lasso.

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?

Answer:

	Metric	Ridge Regression	Lasso Regression
0	R2 Score (Train)	0.903850	0.905287
1	R2 Score (Test)	0.863691	0.861922
2	RSS (Train)	15.231463	15.003830
3	RSS (Test)	10.128449	10.259895
4	MSE (Train)	0.122140	0.121224
5	MSE (Test)	0.152067	0.153050

As we can see Mean squared error and R2 is almost same in both the models, we will go ahead with Lasso since it would give option to penalize or feature elimination also.

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?

Answer: Coding part in python note book.
After the new model was build the five most important predictor variables are as shown in below screenshot:

Top coeffcients of Ridge

```
In [124]: betas.sort_values(by=['Ridge'],ascending=False, inplace=True)
          betas['Ridge'].head(5)

Out[124]: BsmtFinSF1      0.055823
          YearBuilt       0.053482
          BsmtUnfSF       0.051150
          BedroomAbvGr    0.046803
          OverallCond     0.043695
          Name: Ridge, dtype: float64
```

Top coeffcients of Lasso

```
In [125]: betas.sort_values(by=['Lasso'],ascending=False, inplace=True)
          betas['Lasso'].head(5)

Out[125]: YearBuilt       0.081480
          BsmtFinSF1      0.076841
          BsmtUnfSF       0.072146
          OverallCond     0.056176
          BedroomAbvGr    0.051612
          Name: Lasso, dtype: float64
```

Question 4

How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

Answer: The model is robust and generalized when it is stable and behaves accurately on unseen data also. Model can be made more robust and generalizable by trading off between bias and Variance. As per Occam's Razor, the model should be complex enough to explain the properties and simple enough so that it is easily explainable. So basically, we tradeoff between variance and bias by choosing model complexity. While creating model all the outliers in the dataset should be taken care so that it does not gives abrupt results on unseen data. Regularization should be introduced to reduce the error and avoid overfitting.

As explained above sometimes to make model robust and generalizable we have to compromise with accuracy. A model which is perfectly accurate is usually over fitted model.