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 to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Answer: As per the changes, we got the optimal values of alpha for ridge is 10 and for lasso its 0.001. With these alphas the adjusted R^2 is 0.90.

As mentioned after doubling the value of alpha for both ridge and lasso, there is no noticeable difference in the adjust R^2 or RMSE, and the import predictor variables after this change is implemented are as follows:

	Ridge (alpha=20.0)	Lasso (alpha=0.001)	Ridge (alpha = 40.0)	Lasso (alpha = 0.002)
LotArea	0.040234	0.038282	0.040861	0.038249
OverallQual	0.094401	0.097098	0.093708	0.098803
OverallCond	0.044810	0.045685	0.042980	0.043678
YearBuilt	0.024054	0.026235	0.021814	0.026624
YearRemodAdd	0.014138	0.012993	0.015512	0.014334
BsmtFinType1	0.011778	0.011132	0.011810	0.010611
BsmtFinSF1	0.028071	0.028217	0.028352	0.028596
BsmtFinSF2	0.000000	0.000000	0.000000	0.000000
HeatingQC	0.012877	0.012072	0.013472	0.012730
CentralAir	0.016852	0.016988	0.016949	0.016841
1stFlrSF	0.130980	0.136397	0.124790	0.134151
2ndFlrSF	0.107580	0.110587	0.101843	0.108058
BsmtFullBath	0.016605	0.016636	0.015710	0.015439
BsmtHalfBath	0.000000	0.000000	0.000000	0.000000
HalfBath	0.011801	0.010056	0.013301	0.009812

Few of the predictor variables have been shown based on doubling the alpha for Ridge and Lasso, only there is a slight increase in model coefficients based on this analysis.

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: The optimum lambda value in case of Ridge and Lasso is as follows:

Ridge: 20

Lambda: 0.001

After comparing the model adjusted R^2 score, it sounds its almost same which is around 0.90.

We have chosen Lasso, as it gives lesser number of predictor variables, as some of the model coefficients becomes zero, and hence Lasso has a additional advantage over Ridge and should be used as the final model.

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?

Answer:

The top 5 features in the current model are:

- 1stFlrSF
- 2ndFlrSF
- OverallQual
- OverallCond
- SaleCondtion_Partial

After dropping these 5 features, and then we build a Lasso regression model after dropping these 5 features, we noticed that the adjusted R^2 is dropped to 0.80 from 0.90 and RMSE is increased to 0.17 from 0.124, and hence now the new top 5 predictor are as follows:

	Lasso
GarageArea	0.080422
LotArea	0.079027
FireplaceQu	0.071422
KitchenQual	0.062614
BsmtFinSF1	0.052331

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?

Answer:

According to Occam's Razor making a model as simple as possible but not making it any sort of simpler. For making model simple, we can use some regularization techniques like Ridge and Lasso along with some feature reduction techniques like RFE. Some of the points are as under:

- Simpler models are usually more generic which means they relies on learning the inherent trends/dependencies present in the model instead of learning the training data very strongly.
- Simpler models require fewer training samples for doing effective training than the more complex models, and hence are easier to train.
- Simpler models are more robust which means simpler models have low variance and if we compare it with a more complex model with many features and extreme coefficients can learn the training data completely and performs poorly on the test data, and any variation in the training dataset will change the model parameters completely.

Therefore, to make a model more robust and generalizable, we should always make the more simple but not simpler.

Regularization can be used to make model simpler by introducing a penalty term which makes the overall predicted value tends to be near to zero, and we have two techniques for this like Ridge and Lasso.

Also, we can make model simple according to Bias-Variance Trade-off:

- A complex model will need to change for every little change in the dataset and hence is very unstable and extremely sensitive to any changes in the training data.
- A simpler model that abstracts out some pattern followed by the data points given is unlikely to change wildly even if more points are added or removed.

Thus accuracy of the model can be maintained by keeping the balance between Bias and Variance as it minimizes the total error as shown in the below graph: