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?

The optimum alpha values of Ridge and Lasso regression are 0.8 and 0.0001. When the alpha values are doubled, the Lasso and Ridge prediction score became almost the same with the R2 value of 88%.

The most important predictors for Lasso are below,

Lasso (alpha=0.0002)	
GrLivArea	0.224467
OverallQual	0.187642
TotalBsmtSF	0.135337
OverallCond	0.077752
GarageCars	0.069602
LotArea	0.049616
KitchenQual	0.045449
MSZoning_FV	0.033596

MSZoning_RL	0.021558
SaleType_New	0.014374
MSZoning_RH	0.003176
GarageQual	0.001920
BsmtUnfSF	-0.047578
Functional	-0.052571
age	-0.104610

The most important predictors for Ridge are below,

Ridge (alpha=1.6)	
GrLivArea	0.221255
OverallQual	0.174458
TotalBsmtSF	0.135842
MSZoning_FV	0.090149
MSZoning_RL	0.075796
OverallCond	0.074215
GarageCars	0.073071

MSZoning_RH	0.068992
MSZoning_RM	0.056515
LotArea	0.056401
KitchenQual	0.051462
SaleType_New	0.039013
GarageQual	0.035141
SaleType_Con	0.029339
Heating_Grav	0.026478
GarageFinish_NOT APPLICABLE	0.017955
Heating_OthW	-0.021035
SaleCondition_Partial	-0.022564
RoofStyle_Shed	-0.029762
Foundation_Wood	-0.032239
SaleType_CWD	-0.041571
Exterior1st_BrkComm	-0.044418
BsmtUnfSF	-0.049158
Functional	-0.063487
age	-0.099918

** Note - The coefficient values are sorted in the descending order

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?

I will choose Lasso for the bellow reasons,

1. Lasso Performed better than Ridge and Linear Regression
2. Lasso prediction score on train(0.897) and test(0.890) is close to zero variance(0.007)
3. Ridge prediction score on train(0.901) and test(0.876) is close to 3% variance
4. RMSE value of Prediction on test is slightly lesser for lasso compared to Ridge, the smaller the better. So Lasso wins even with RMSE
5. In Lasso, some of these coefficients became 0, thus resulting in an easier interpretation with predictors

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?

The 5 most important predictor variables excluding the initial 5 predictors by the original lasso model are below,

	Lasso (alpha=0.001)
GarageCars	0.177129

KitchenQual	0.175559
LotArea	0.146529
OverallCond	0.051638
SaleType_New	0.012155

Here the dropList contains the initial 5 most important predictors,

```
dropList=['GrLivArea','OverallQual','TotalBsmtSF','MSZoning_FV','MSZoning_RL']#
initial 5 most important features
```

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?

- The model should adhere to the optimal feature selection w.r.t Bias/Variance Trade-off, we should not have more features and have a complex model or fewer features with more Bias. The optimal number of features will help us to build a robust and generalizable model
- The model has high accuracy in training but in tests, the scores are very low means the model memorised the pattern in the train data and resulted in overfitting. So the train/test prediction accuracy should not have a high variance
- Techniques like Regularization using Lasso and Ridge help to avoid overfitting