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?

Ridge Regression: - When the curve between -ve mean absolute error and alpha is plotted, value of alpha increases from 0, the error term decreases, and the train error is minimum. Hence alpha = 2 is decided to be optimal value for ride reg.
Lasso Regression: - to begin with alpha's value was kept very small (0.01), when alpha value was increased the model was penalizing the more and made most of the Coeff values zero.
When the alphas were doubled on ride reg, model will apply more penalty on the curve and model will be more generalized and simpler. with doubled alpha, errors for both test and train were increased. When alpha was doubled in Lasso reg, the model was pushing the coefficients of the variable to "0" and R2 square was also decreased.

The most important predictor variables after the changes were implemented were:

Ridge Regression:	*Lasso Regression:*
MSZoning_RL	GrLivArea
MSZoning_RH	LotArea
MSZoning_FV	LotFrontage
MSZoning_RM	FirePlaces
SaleCondition_Partial	BsmtFinSF1
GrLivArea	OverallCond
Neigborhood_StoneBr	TotalBsmtSF
SaleCondition_Normal	OverallQual
Exterior1st_BrkFace	GarageArea
Neighborhood_Crawfor	LotArea

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?

A tuning parameter lambda, sometimes called a penalty parameter, controls the strength of the penalty term in ridge regression and lasso regression. It is basically the amount of shrinkage, where data values are shrunk towards a central point, like the mean.
As lambda increases variance in model decreases and bias will remain constant. All variables are included in Ridge Regression models.
In Lasso when the lambda increases, coefficients are shrunk to zero and makes the variables ‘0’. Lasso allows for variable selection. When Lambda value increases, variables with 0 value are neglected by the model. So for simpler and easily explainable model we can use Lasso.

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 five most important predictor variables now will be

- MSSubClass
- RoofMatl_Membran
- MSZoning_RL
- MSZoning_FV
- MSZoning_RH

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 answer in a single line, simpler models are robust and generalizable. But the downside will be accuracy will be less and bias will be more. Upside is we will have less variance and more generalizable model. So, depending on the use-case we need to do the trade-off. For eg: in determining if a particular drug will be good for curing cancer will less side effects, the model accuracy should be higher. But in a case where I am building a propensity base for a Bank's product the accuracy threshold can be lower because the False Negatives are not a do-or-die situation.