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?

Ans-
Optimal value of alpha is 100 for ridge and 0.01 is for lasso regression models.

If we double the alpha value means doubling the penalties the models will start underfitting.

"LotArea" is the most important predictor variable.

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?

Ans-
Both ridge and lasso are giving similar results, but lasso is making some features as zero. So I will choose lasso for better clarity with less features and will be easy to interpret.

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?

Ans-
MSZoning, Neighborhood, OverallQual, Foundation, GrLivArea are five most important predictor variables now.

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?

Ans-

If the model is predicting pretty well for unseen data comparing to the trained data. We can say the model is robust and can behave well for different future data sets. Combining linear regression and ridge or lasso for regularise the features will provide a better model for prediction.