Question 1

Alpha = 500 for Ridge and Alpha=1000 for lasso regression

Doubling up of alpha to 1000 with ridge regression, made only slight changes to the model and the top features stayed the same.

Doubling up of alpha to 2000 with lasso regression, reduced the training and test scores of the model as it reduced the number of features used by 7

Question 2

I will use ridge regression at alpha = 500 as the test and train scores were high with less difference between train and test scores

Question 3

After removing the top 5 features from the dataset, the new top 5 variables are

1stFlrSF, 2stFlrSF, YearBuilt, Functional, LotArea

Question 4

1) Removal of outliers to make sure the model error parameters are less and accurate
2) Not overfitting or underfitting by carefully choosing/removing features by ridge/lasso regression
3) Using/evaluating multiple regression algorithms on train and test data to find the best fit
4) Running the regression algorithm on the train data by creating as many possible variations to find the best fit

If we keep the model general and robust, the accuracy will automatically increase