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?

Ridge: 100
Lasso: 0.001
In all models, the training scores have slightly decreased and the testing scores have slightly increased. The change is more obvious for Ridge after RFE. Here is the change
was largest, minimizing the difference between the train and test data.
'GrLivArea', 'OverallQual', 'TotalBsmtSF', 'OverallCond', 'BuiltYear' is important
changing forecasts

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?

The r2 score for lasso is slightly higher and the difference between training and testing is smaller
lower. So I chose Lasso. Lasso helps to reduce the objects in the image, .
helps create a simple 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?

The top predictor variables are: 'GrLivArea', 'OverallQual', 'TotalBsmtSF', 'OverallCond',
'YearBuilt'

GrLivArea	0.128692
OverallQual	0.065629
TotalBsmtSF	0.048505
OverallCond	0.043846
YearBuilt	0.043804

After removing them, the predictor variables are:

BsmtFinSF1	0.143149
BsmtUnfSF	0.119984
MSZoning_RL	0.086269
2ndFlrSF	0.085585
MSZoning_RM	0.072373

The predictor variables remain the same, with a slightly different order for Lasso applied to
the RFE created variables.

GrLivArea	0.136432
MSZoning_RL	0.083509
TotalBsmtSF	0.073250

MSZoning_RM 0.069320
GarageCars 0.041585

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 make sure a regression model is robust and generalizable, numerous techniques need to be employed. Feature selection, regularization, cross-validation, and hyperparameter tuning assist optimize model complexity and prevent overfitting. A proper educate-test break up and coping with outliers are important for evaluating how properly the version generalizes to unseen statistics. Ensemble techniques and suitable information preprocessing in addition beautify generalization. Monitoring and retraining the model as facts evolves is also crucial.
A robust, generalizable model minimizes overfitting and enhances predictions on new data. This may reduce training data accuracy but ensures more reliable performance across diverse datasets. Sometimes, trading off some training data accuracy for better generalization leads to a more effective real-world model.