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?

a. Optimal value of alpha: For Ridge regression :1.0 For Lasso Regression :10.
b. R2score on training data has decreased but it has increased on testing data in Ridge regression and Lasso regression.
c. The following will be the most important predictor variables after the change is implemented.

Feature	Coef	mod
1	LotFrontage	9.016476
2	Exterior1st_AsphShn	1.284113
3	Condition2_RRAe	-1.235881
4	BsmtFullBath	1.174806
5	RoofMatl_Membran	1.160288
6	RoofMatl_Metal	1.114697
7	RoofMatl_WdShake	1.110235
8	RoofMatl_Tar&Grv	1.106411
9	RoofMatl_WdShngl	1.060673
10	RoofMatl_Roll	1.018627

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?

We will make use of Lasso Regression model because it is using less numbers of variables and giving almost the same accurate. Its more efficient model than Ridge regression model. The r2_score of lasso is slightly higher than lasso for the test dataset so we will choose lasso regression to solve this problem.

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?

After dropping five most important predictor variables in the lasso model and new model is created , the following 5 will be the new important variables.

Feature	Coef
1	LotArea
2	Exterior1st_AsphShn
3	RoofMatl_Metal
4	RoofMatl_Membran
5	RoofMatl_WdShngl

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?

There is a trade-off between the accuracy of a machine learning model and its robustness and generalizability. A model that is too accurate on the training data may be overfit and perform poorly on new data. On the other hand, a model that is too general may not be accurate enough for any particular task. The goal is to find a balance between accuracy and generalizability that is appropriate for the specific problem being solved.

The following features will make a model robust and generalisable:-

1. **Data quality and diversity:**. The data should be of high quality, meaning that it is accurate, complete, and consistent. It should also be diverse, meaning that it represents the full range of possible inputs that the model may encounter in the real world.
2. **Model selection and evaluation:** Since machine learning models are diverse , and the best choice for a particular task will depend on the specific problem being solved. Using a separate dataset called a validation set to test the model's performance on unseen data is a good practise.
3. Model optimization and tuning: Finding the optimal values for the hyperparameters that can be adjusted to improve Model performance.
4. Model testing and deployment: Real-world data can often be different from the data that the model was trained on, and the model's performance may degrade over time. To avoid the same continuous optimization of the model is needed for better accuracy.