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: Alpha (α) is the penalty term that denotes the amount of shrinkage (or constraint) that will be implemented in the equation. The optimal value for Ridge and Lasso Regression is 1 and 10 respectively. If we double the alpha value that is from 1 to 2 and from 10 to 20. After using the alpha value and re-evaluate with sklearn we can see that predictors are the same but the coefficient has changed. The code is written in the notebook.

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. The r^2 score of Lasso Regression is slightly higher than the Ridge Regression so we will prefer Lasso Regression.
Ridge and Lasso Regression are the forms of regularised linear regressions. The regularisation can also be interpreted as prior in a maximum a posteriori estimation method. Under this interpretation, the ridge and the Lasso can make different assumptions on the class of linear transformation as they relate input and output data. In the Ridge, the coefficients of the linear transformation are normally distributed and in the Lasso they are Laplace distributed. In the Lasso, it makes it easier for the coefficients to be zero and therefore easier to eliminate some of your input variables as not contributing to the output.

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. So after removing the 5 important predictor variables we can clearly see that the r^2 score has decreased. The core is written in the Jupyter NoteBook
So the next 5 important variable we will select is

   a. 11stFlrSF - First Floor Square Feet
   b. GrLivArea - Above grade ground living area square feet
   c. Street_Pave - Pave road access to property

d. RoofMatl_metal - Roof material Metal
e. RoofStyle_Shed - Type Of Roof (Shed)

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. The model should be generalised so that the test accuracy is good enough to the training score. The model should be accurate for datasets even for the data that are not in the training set and also should take less time to train for larger datasets. We should not give too much work in our model to predict even for the outliers so that the accuracy of the model is high and it's not over-fitted with the training dataset. For processing the outliers, the analysis should be done and only those that are relevant should be kept and others should be removed. If the model is not fast and robust it cannot be trusted and also cannot be used for production for predicting in real time.