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?

- Optimal Value for alpha is 15.
- If double the value of alpha – The variable rank and the coefficients remain unaffected
- Most important predictor variable after the change is implemented will be – BsmtHalfBath

```
alpha = 15
ridge = Ridge(alpha=alpha)

ridge.fit(X_train, y_train)
len(sorted(ridge.coef_))
```

```
12
```

```
ridge_coeff = list(zip(X_trai
pd.DataFrame(ridge_coeff, col
```

	Coefficient	Feature
0	BsmtHalfBath	18510.245008
1	WoodDeckSF	8009.290701
2	3SsnPorch	7141.974696
3	ScreenPorch	5515.502033
4	MiscVal	5031.990633
5	LotShape_IR2	3548.865426
6	HeatingQC_Po	3395.322230
7	SaleType_CWD	2248.683676
8	SaleType_Con	222.535779
9	SaleType_ConLI	190.990552

```
alpha = 30
ridge = Ridge(alpha=alpha)

ridge.fit(X_train, y_train)
len(sorted(ridge.coef_))
```

```
12
```

```
ridge_coeff = list(zip(X_train
pd.DataFrame(ridge_coeff, colu
```

	Coefficient	Feature
0	BsmtHalfBath	18510.245008
1	WoodDeckSF	8009.290701
2	3SsnPorch	7141.974696
3	ScreenPorch	5515.502033
4	MiscVal	5031.990633
5	LotShape_IR2	3548.865426
6	HeatingQC_Po	3395.322230
7	SaleType_CWD	2248.683676
8	SaleType_Con	222.535779
9	SaleType_ConLI	190.990552

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?

I will choose lasso over ridge

- Alpha value is same for both Ridge and Lasso
- Lasso will be effective with large number of variables.

Question 3

After building the model, you realized 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?

LotShape_IR2, HeatingQC_Po, SaleType_CWD, SaleType_Con, SaleType_ConLI

Question 4

How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

Model can be robust and generalizable with optimum variance and bias.

Optimum value of alpha should be selected. Penalty should be less.

Ridge and Lasso , r2 score can be used for the same.