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 to double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?

Ans:

The purpose of regularization is to ensure that model is not overtly complex. For ridge and lasso regression we penalize the model for its complexity.

Lambda is the coefficient for the regularization term R(w).
- Ridge uses sum of squared coefficients
- Lasso uses sum of absolute value of coefficients

The optimal value of alpha for Ridge is 2 and for Lasso it is 0.0001. With these alphas the R2 of the model was approximately values nearing to 0.83.

After doubling the alpha values in the Ridge and Lasso, the prediction accuracy remains around 0.82 but there is a small change in the co-efficient values. The new model is created and demonstrated in the Jupiter notebook.

Below are the changes in the co-efficient.

	Ridge Co-Efficient
Total_sqr_footage	0.168551
GarageArea	0.101170
TotRmsAbvGrd	0.066366
OverallCond	0.046645
LotArea	0.044408
LotFrontage	0.032310
CentralAir_Y	0.030935
Total_porch_sf	0.030378
Neighborhood_StoneBr	0.028599
Alley_Pave	0.024118
MSSubClass_70	0.023206
RoofMatl_WdShngl	0.022779
SaleType_Con	0.022595
Neighborhood_Veenker	0.022175
OpenPorchSF	0.021912
HouseStyle_2.5Unf	0.021474
KitchenQual_Ex	0.019456
PavedDrive_P	0.018754

	Ridge Doubled Alpha Co-Efficient
Total_sqr_footage	0.148479
GarageArea	0.091536
TotRmsAbvGrd	0.067673
OverallCond	0.042534
LotArea	0.039162
Total_porch_sf	0.032932
CentralAir_Y	0.030906
LotFrontage	0.027828
Neighborhood_StoneBr	0.026176
MSSubClass_70	0.022350
OpenPorchSF	0.021735
Alley_Pave	0.021469
Neighborhood_Veenker	0.019891
KitchenQual_Ex	0.019890
BsmtQual_Ex	0.019874
HouseStyle_2.5Unf	0.018605
MasVnrType_Stone	0.018520
RoofMatl_WdShngl	0.017973

Lasso Regression Model

	Lasso Co-Efficient
Total_sqr_footage	0.201620
GarageArea	0.110891
TotRmsAbvGrd	0.061496
LotArea	0.045611
OverallCond	0.045524
CentralAir_Y	0.032341
Total_porch_sf	0.028519
Neighborhood_StoneBr	0.022867
Alley_Pave	0.020339
OpenPorchSF	0.019049
MSSubClass_70	0.018729
KitchenQual_Ex	0.016971
LandContour_HLS	0.016961
BsmtQual_Ex	0.016644
Condition1_Norm	0.016300
MasVnrType_Stone	0.014740
Neighborhood_Veenker	0.014612
LotFrontage	0.013990

	Lasso Doubled Alpha Co-Efficient
Total_sqr_footage	0.204096
GarageArea	0.104112
TotRmsAbvGrd	0.064101
OverallCond	0.041420
CentralAir_Y	0.032421
Total_porch_sf	0.030378
LotArea	0.026931
BsmtQual_Ex	0.018090
KitchenQual_Ex	0.016499
Neighborhood_StoneBr	0.016400
Alley_Pave	0.016017
OpenPorchSF	0.015051
LandContour_HLS	0.014552
MSSubClass_70	0.014316
MasVnrType_Stone	0.013506
Condition1_Norm	0.013193
SaleCondition_Partial	0.010818
LotConfig_CulDSac	0.009039

Overall, since the alpha values is very small, we don’t see much change in the model after doubling the alpha and building the model.

Question 2
You have determined the optimal value of lambda for ridge and lasso regression during the assignmen t. Now, which one will you choose to apply and why?

Ans: The metrics for Advanced Regression Model building is recorded as below: -

Metrics Recorded	**Ridge Regression**	**Lasso Regression**
Optimal Alpha Value	2	0.0001
Mean Square Error	0.0018018986744686388	0.0018277328764762812
Mean Absolute Error	0.02914698264772252	0.028565953511134015

- The Mean Squared Error and Mean Absolute Error of both the models are almost same.
- Since Lasso helps in feature reduction (as the coefficient value of some of the features become zero), Lasso has a better edge over Ridge and should be used as the final model.

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?

Ans: The five most important predictor variables in the current lasso model is: -

Total_sqr_footage	0.201620
GarageArea	0.110891
TotRmsAbvGrd	0.061496
LotArea	0.045611
OverallCond	0.045524

We build a Lasso model in the Jupiter notebook after removing these attributes from the dataset.
The R2 of the new model without the top 5 predictors drops to .73
The Mean Squared Error changes to 0.002831125450373643
The Mean Absolute Error changes to 0.03945336948638169

Upon deleting them the new five predictors are

LotFrontage	0.146604
Total_porch_sf	0.070877
HouseStyle_2.5Unf	0.061791
HouseStyle_2.5Fin	0.052376
Neighborhood_Veenker	0.041474

Question 4

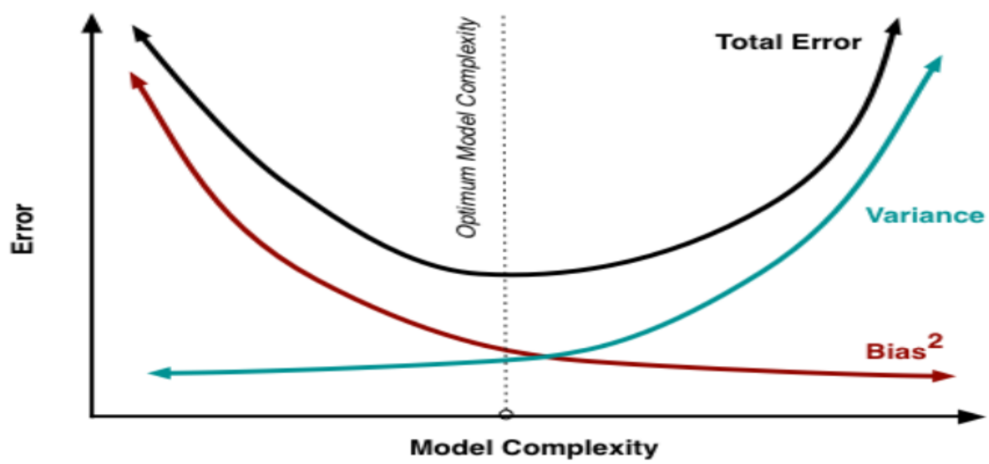
How can you make sure that a model is robust and generalizable? What are the implications of the same for the accuracy of the model and why?

As Per Occam's razor – model should be as simple as necessary.
So according to above expectation simple model have an edge over other complex models. The advantages of simple model are as below:
• Generalizability
• Robustness

• Making few assumptions
• Less data is required for learning

Robust model is not sensitive to training data. Robust models have low variance and high bias.
• Variance = How sensitive is model to the training data. This refers to consistency of the model.
• Bias = Accuracy of the data on unseen future data.

Making a model simple lead to Bias-Variance Trade-off:
• A complex model will need to change for every little change in the dataset and hence is very unstable and extremely sensitive to any changes in the training data.
• A simpler model that abstracts out some pattern followed by the data points given is unlikely to change wildly even if more points are added or removed.



Model is always trained on training set and evaluated on unseen data (test set). Adding to many predictor variables in the model may lead to complex model. Complex model deteriorates the performance of the model (r2 score). Complex model introduces problem of overfitting where model memorized the data and is not generalized. When such model is evaluated against the unseen data the performance is very poor.