

Answer 1:

The optimal value of alpha is 1000 for both ridge and lasso regression. If we double the value of the alpha the R-squared value will decrease. SBrkr is the variable that change.

Answer 2:

Ridge Regression .It is most suitable when a data set contains a higher number of predictor variables than the number of observations.

Answer 3:

SBrkr - Standard Circuit Breakers & Romex

FuseA - Fuse Box over 60 AMP and all Romex wiring (Average)

Lvl - Near Flat/Level Land Contour

PConc - Poured Concrete Foundation

NAmes - North Ames Neighbourhood

Answer 4:

The model should be accurate for datasets other than the ones which were used during training. Too much weightage should not be given to the outliers so that the accuracy predicted by the model is high. To ensure that this is not the case, the outlier analysis needs to be done and only those which are relevant to the dataset need to be retained. Those outliers which it does not make sense to keep must be removed from the dataset. This would help increase the accuracy of the predictions made by the model. This would help standardize the predictions made by the model. If the model is not robust , it cannot be trusted for predictive analysis.