# Subjective Questions

**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?

**Ans –** Optimal Value of alpha for Lasso Regression is **0.001**

Optimal Value of alpha for Ridge Regression is **1**

**When double the value of alpha for lasso –**

We get less r2 score of train (from 0.893 to 0.88) and slightly better r2 score for test (from 0.845 to 0.847) but RSMLE (as we have log the dependent variable) value is increasing (from 0.155 to 0.156)

**When double the value of alpha for ridge –**

We get less r2 score of train (from 0.922 to 0.918) and slightly better r2 score for test (from 0.836 to 0.837) but RSMLE value is increasing (from 0.161 to 0.162)

Most important predictor variable for both regressions are as follows -

| Regression | Most Important Variables |
|---|---|
| **Lasso Regression** | OverallQual, Neighborhood, TotalBsmtSF, YearBuilt, 2ndFlrSF |
| **Ridge Regression** | RoofMatl, Neighborhood, PoolQC, MSZoning, Heating |

**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?

**Ans –** We will use lasso regression to apply as it is giving as better rmsle (root mean squared log error) and the difference in r2 score of test and train is also very close compared to ridge regression.

Lasso regression (L1 regularization), uses a tuning parameter called lambda as the **penalty is absolute value of magnitude of coefficients** which is identified by cross validation. As the **lambda value increases Lasso shrinks the coefficient towards zero** and it make the **variables exactly equal to 0**. Lasso also does variable selection. When lambda value is small it performs simple linear regression and as lambda value increases, shrinkage takes place and variables with 0 value are neglected by the model.

Ridge regression (L2 regularization), uses a tuning parameter called lambda as the **penalty is square of magnitude of coefficients** which is identified by cross validation. Residual sum of squares should be small by using the penalty. Hence the coefficients that have **greater values gets penalized**. As we increase the value of lambda the variance in model is dropped and bias remains constant. **Ridge regression includes all variables in final model unlike Lasso Regression.**

Because above reasons, **lasso will make some coefficients zero** and **model will be less complex** than ridge. Therefore, we will go for **lasso regression** compare to ridge regression.

**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?

**Ans –**

Those 5 most important predictor variables that will be excluded are :-

| Variables | Description |
|---|---|
| **TotalBsmtSF** | Total square feet of basement area |
| **2ndFlrSF** | Second floor square feet |
| **YearBuilt** | Original construction date |
| **OverallCond** | Rates the overall condition of the house |
| **LandContour_HLS** | Flatness of the property - Hillside - Significant slope from side to side |

**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?

**Ans –** The simplification of model can be done at several levels like –

1. Choice of simpler functions
2. Keeping the number of model parameters small,
3. If the family is that of polynomials, then keeping the degree of the polynomial low, etc

According to Occam's Razor, model is robust and generalisable when it makes fewer assumptions about data yet to be seen. If we have 2 models which shows similar performance then we will choose the model which makes fewer assumptions.

Simple models are robust and easy to explain but when it comes to accuracy it makes more error in training set as it goes for more generalise approach compare to complex model which leads to overfitting.