Q1) The optimal value for Ridge was 10 and for Lasso it was 100. On doubling the value of the alpha we noticed that the r2 accuracy dropped for Ridge but for Lasso it improved. The most important predictor was Neighbourhood: (NWAmes).

Q2)Among Ridge and Lasso I would be going ahead with the Ridge with an alpha of 10 since the r2 score of Ridge is much better than that of the Lasso.

Q3)

Q4) In order to ensure the model is robust and generalizable we have to ensure that the model does not overfit or underfit on the data. This is where the choice of the alpha matters during this hyper parameter tuning we try to ensure that this value does not overfit or underfit and choose an optimal value. The outcome of this would actually mean a decent performance not just on the train data but also on the test. If the model built gives us a very good number for accuracy on train but not so good value for test then it implies that we have overfit.