1. Housing train shows 405 observations of 14 variables compared to the 506 observations of 14 variables present in the original data set. It seems like there are the same number of X variables.

2. With LASSO the optimal value of lambda is 0.00355648. The in-sample RMSE is 0.165. The out-of-sample RMSE is 0.00356.

3. With ridge regression the optimal value of lambda is 0.03727594. The out-of-sample RMSE is 0.233.

4. I don't think that you would be able to estimate a simple linear regression model on a p¿¿n data set since most algorithms assume that this is not the case. As always we are wanting to balance a model that has enough complexity to generalize to new data sets, but not so much that misinterprets random noise as being significant. I believe that the model is doing well on this trade-off because the in sample RMSE is not so high that it would seem overfit (I think).