**What Produced Better Results**

Multi linear regression produced significantly better results compared to the simple linear regressions. When performing the simple linear regression, I achieved the best R2 score using the medium house value feature. I got an R2 score of approximately 0.4681 with the next highest R2 score of approximately 0.202 from the average rooms feature. The multi linear regression using all the available features achieved an R2 score of approximately 0.6009. The multi linear regression was able to achieve this better result by considering multiple factors when making the prediction. If there were more features associated with the house this score likely could be increased further and the predictions could be more accurate. Something like the average number of bathrooms would be a good feature to use as it is something that is relevant to the price of houses.

**What I Learned**

I learned a lot of terminology surrounding machine learning and learned some of the machine learning methods there are. I didn’t know that there was a way of measuring the accuracy of a model. I guess it makes sense that that is a mathematical way of scoring the model since it would be very difficult to figure out if it is working correctly otherwise. I also didn’t know of any of the machine learning methods and how they worked. K-nearest neighbors seems like it would be the easiest to implement your own algorithm for since it is using existing data and seeing what the *k* closest data points are to classify the new data point.

**What Challenges I Faced**

I initially struggled with setting up the linear regression for the final project. I didn’t fully understand everything I read from the textbook and didn’t know what the fit function was doing. This made me think I was missing the part where I actually train the model and made me think the R2 score I was getting was incorrect. This was especially the case because the feature I started with when getting the linear regression set up had a very bad score and made it look like something was wrong. I also struggled some with setting up the k-nearest neighbors classification. I forgot that it takes an integer or string for the second array. I had to go through the error messages and figure out what was causing the errors to happen. Once I figured out it needed to be an integer, I rounded the values so they could be used as the classification.

**Performance Differences of k-NN and Linear Regression**

The linear regression performed better than the k-nearest neighbor classification for what I used it for. The k-nearest neighbor classification got an R2 score of approximately 0.3421 and the linear regression got an R2 score of approximately 0.4681 when used on the same feature. This discrepancy could also be caused by a combination of the data pool being too small and rounding the actual values for the classification. Overall I think k-NN could out perform linear regression if used with the correct data and a sufficient data pool.