Problem Definition: The goal of this project was to include all features (with the exception of neighborhood) in a predictive model for median Boston home value . RMSE was used as the evaluation metric to select the best model. RMSE represents the average error between predicted and expected response values. Therefore a smaller RMSE suggests a more effective model.

Research Design/Programming: A preliminary comparison between 4 algorithms including Linear Regression, ElasticNet, Random Forest Regressor and Gradient Boosting Regressor (GBR) were evaluated using cross validation. The model generated by GBR had the lowest RMSE by a considerable amount and was the front runner for final model selection.

	not normalized	normalized
GBR	3664.080390	3758.748656
RandomForest	4226.065681	4228.099475
LinearReg	4849.249199	4881.982896
ElasticNet	5335.610844	9295.204346

Following the preliminary cross validation comparison, Random Search was performed for hyperparameter tuning. GBR produced the best predictive model.

	Tuned Models	Baseline
GBR	2484.710945	3664.080390
RandomForest	4592.313726	4226.065681
Linear	4789.666822	4849.249199
ElasticNet	4871.031537	5335.610844

On average the GBR model predictions were \$2,484 off of the expected median home value.

Recommendations for Management:

      This initial exploration of model efficiency suggests that GBR should be selected as the preferred algorithm. The dataset provided is small at n = 506. The models predictive capacity would benefit from collecting more data.

      Computing power was a limiting factor during this project which is why Random Search was used for hyperparameter tuning instead of a more exhaustive option such as Grid Search. Before deploying the model into production, a more exhaustive hyperparameter tuning optimization should be conducted.

      Ultimately the model has significant room to improve. Currently the limiting factors are computing resources and dataset size. GBR clearly outperformed opposing algorithms. Similarly GBR showed the greatest reduction in RMSE following hyperparameter tuning.