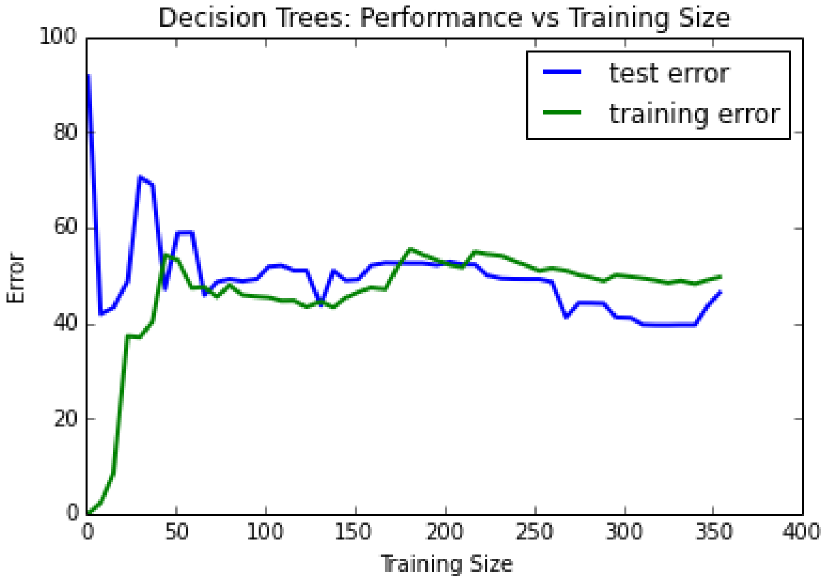
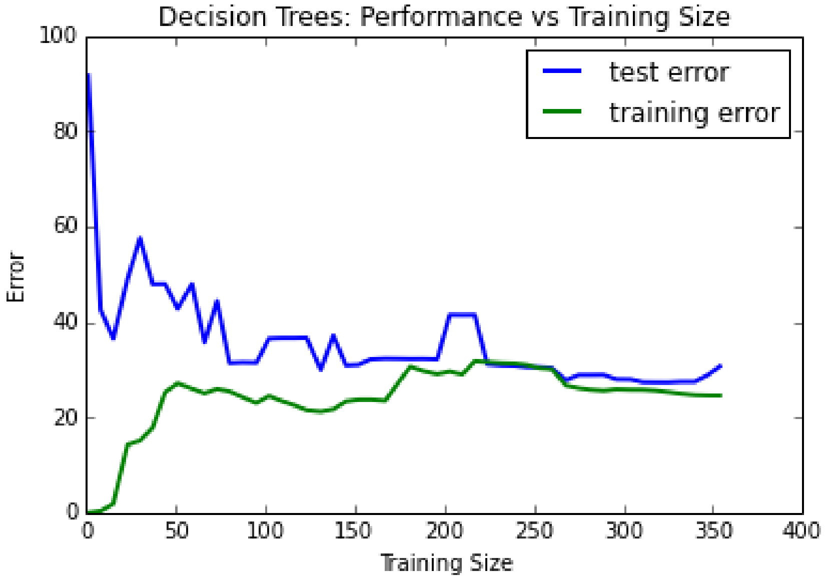
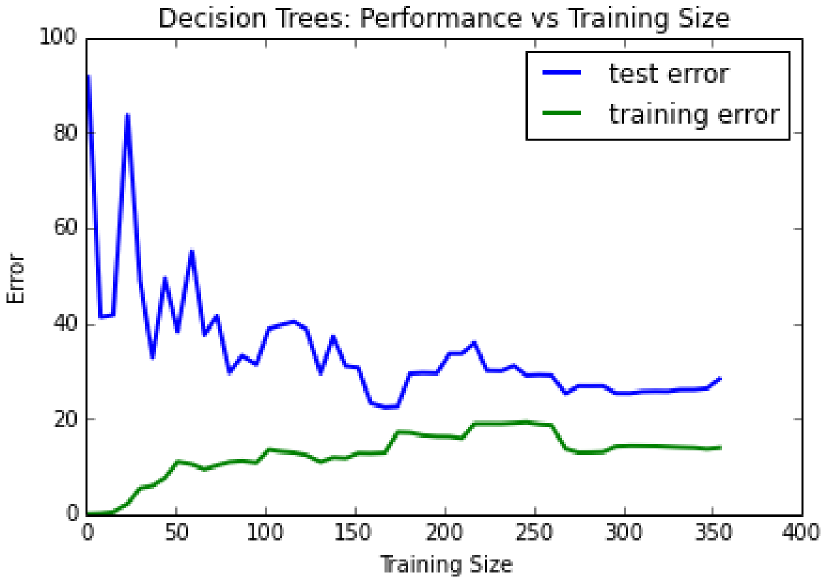
```
Number of houses 506
Number of features 13
Minimum price 5.0
Maximum price 50.0
Mean price 22.5328063241
Median price 21.2
Standard deviation of price 9.18801154528
Decision Tree with Max Depth:
1
```



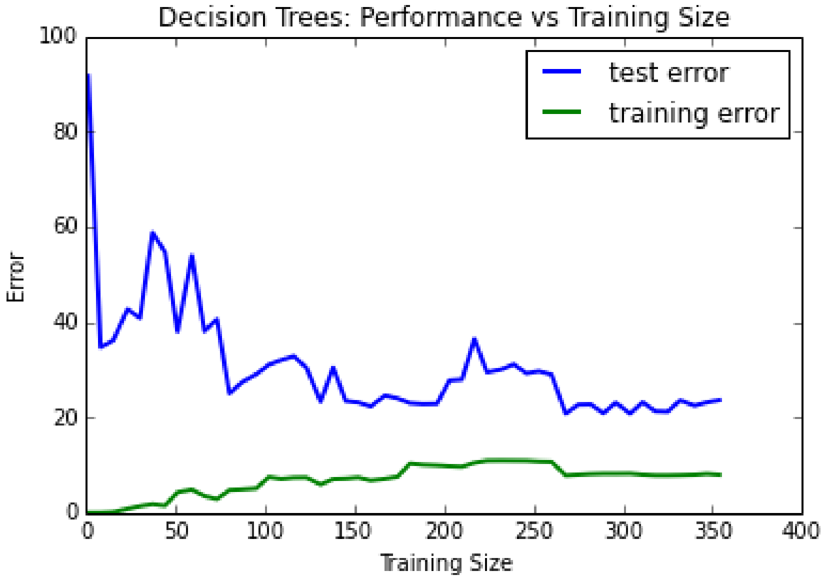
```
Decision Tree with Max Depth:
2
```



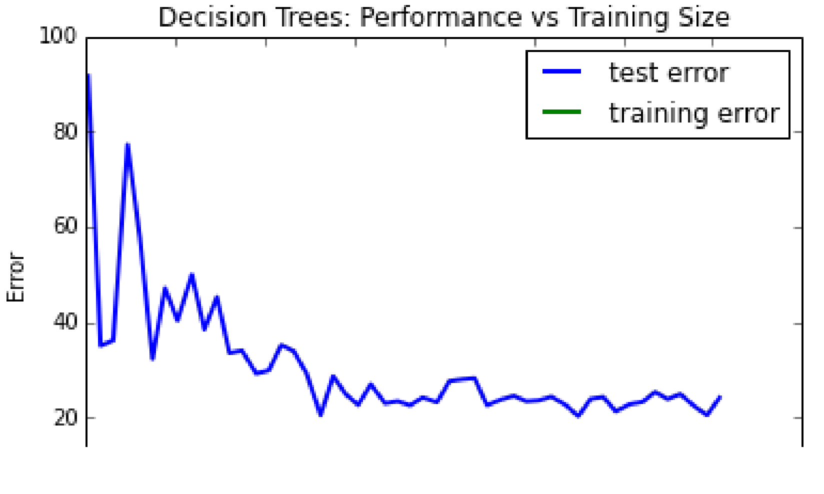
```
Decision Tree with Max Depth:
3
```



```
Decision Tree with Max Depth:
4
```

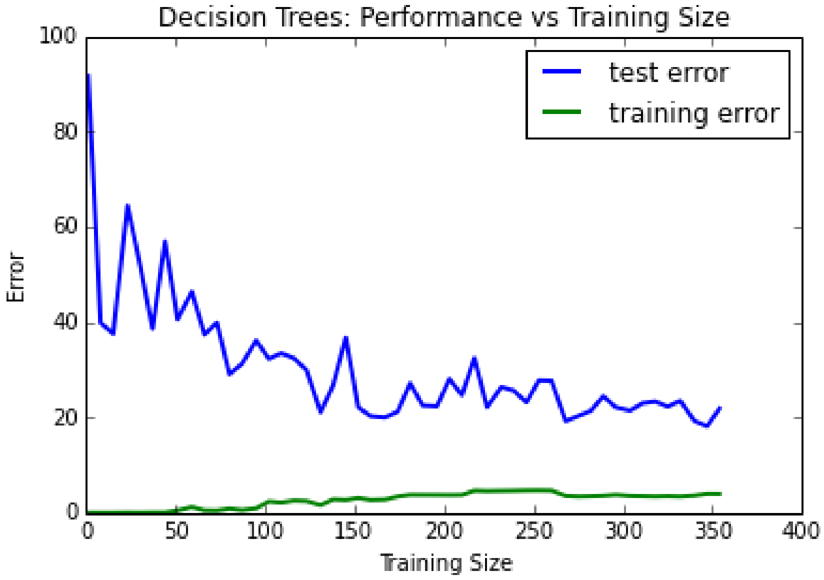


```
Decision Tree with Max Depth:
5
```

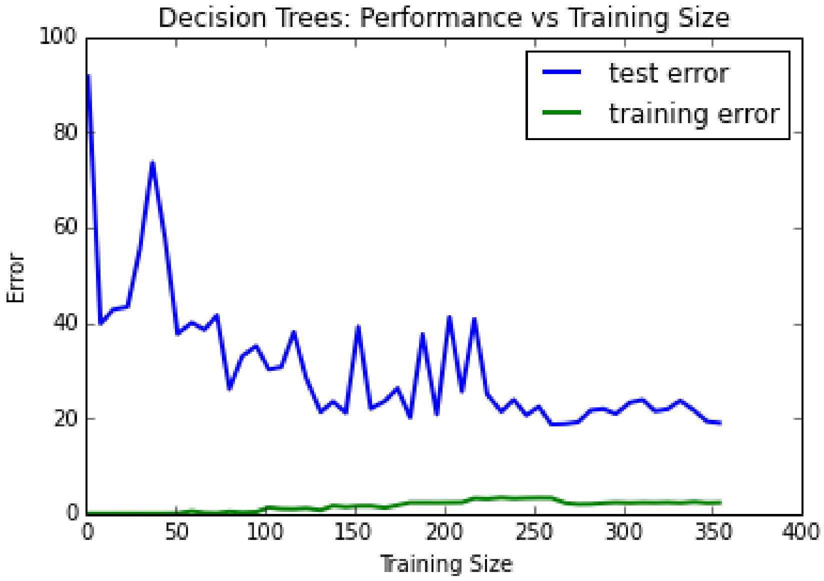


0
0
50
100
150
200
250
300
350
400
Training Size

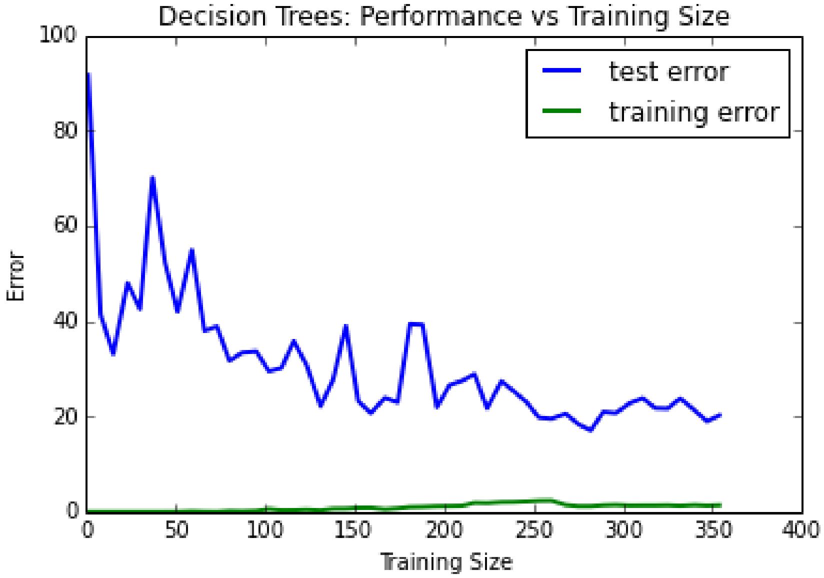
```
Decision Tree with Max Depth:
6
```



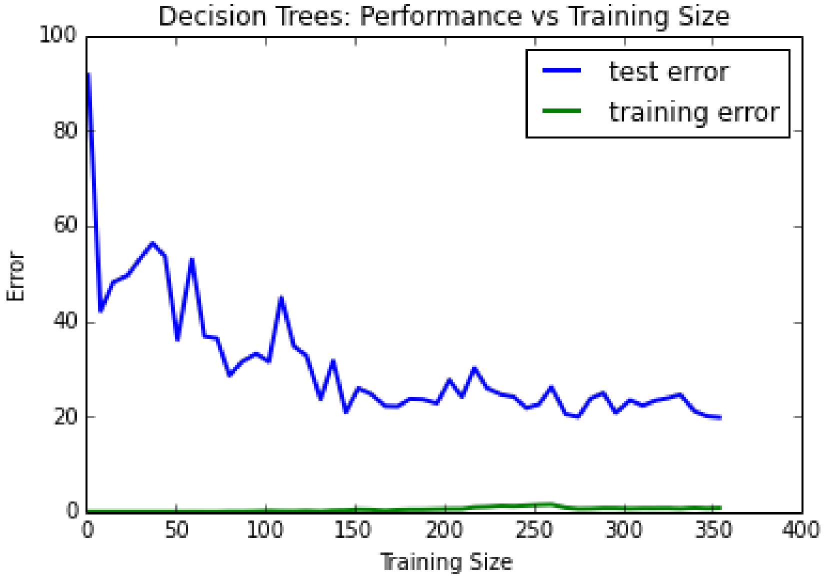
```
Decision Tree with Max Depth:
7
```



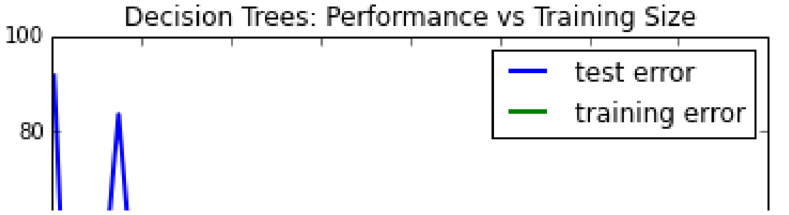
```
Decision Tree with Max Depth:
8
```

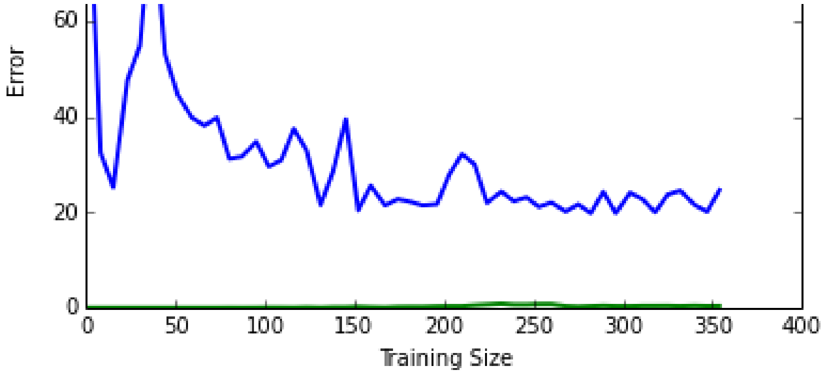


```
Decision Tree with Max Depth:
9
```

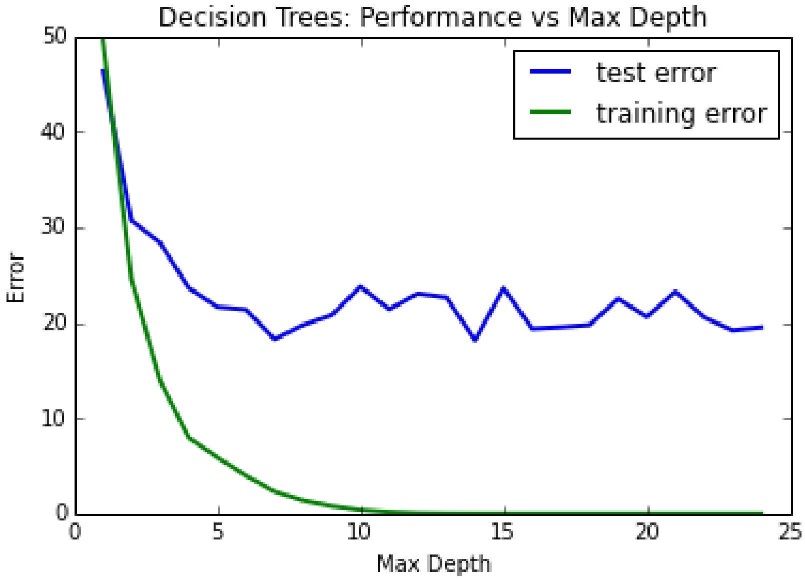


```
Decision Tree with Max Depth:
10
```





```
Model Complexity:
```



```
Final Model:
GridSearchCV(cv=None, error_score='raise',
       estimator=DecisionTreeRegressor(criterion='mse', max_depth=None, max_f
eatures=None,
           max_leaf_nodes=None, min_samples_leaf=1, min_samples_split=2,
           min_weight_fraction_leaf=0.0, presort=False, random_state=None,
           splitter='best'),
       fit_params={}, iid=True, n_jobs=1,
       param_grid={'max_depth': (1, 2, 3, 4, 5, 6, 7, 8, 9, 10)},
       pre_dispatch='2*n_jobs', refit=True,
       scoring=make_scorer(performance_metric, greater_is_better=False),
       verbose=0)
House: [[  1.19500000e+01   0.00000000e+00   1.81000000e+01   0.00000000e+00
    6.59000000e-01   5.60900000e+00   9.00000000e+01   1.38500000e+00
    2.40000000e+01   6.80000000e+02   2.02000000e+01   3.32090000e+02
    1.21300000e+01]]
Prediction: [ 19.99746835]
```

Evaluating Model Performance

Q. Which measure of model performance is best to use for predicting Boston housing data and analyzing the errors? Why do you think this measurement most appropriate?

I think squared difference between predication and actual value is most appropriate. As this error output is differentiable and non negative.

Q. Why might the other measurements not be appropriate here?

Ohter method like absolute difference is not diferentiable.

Q. Why is it important to split the Boston housing data into training and testing data?

To avoid picking up a wrong model that overfits training data.

Q. What happens if you do not do this?

With the increase of model complexity, the model will overfit training data. In that case the training error will be minimal but the model will fail to predict new cases.

Q. What does grid search do and why might you want to use it?

Grid search trains the model with different parameters and chooses a parameter that minimizes testing set error.

Q. Why is cross validation useful and why might we use it with grid search.

Cross validation uses different set of samples than training samples so not prone to overfitting. Grid search chooses parameter that minimizes cross validation/testing error.

Analyzing Model Performance

Q. Look at all learning curve graphs provided. What is the general trend of training and testing error as training size increases?

Training error increases but testing error decreases.

Q. Look at the learning curves for the decision tree regressor with max depth 1 and 10 (first and last learning curve graphs). When the model is fully trained does it suffer from either high bias/underfitting or high variance/overfitting?

The max depth 1 model suffers high bias. The max depth 10 model doesn't seem to suffer high bias or variance as its test error is not significantly higher than other models.

Q. Look at the model complexity graph. How do the training and test error relate to increasing model complexity?

Training error decreases with model complexity. Test error decreases initially with model complexity. Then the test error can increase or decrease with increase of model complexity.

Based on this relationship, which model (max depth) best generalizes the dataset and why?

Decision tree with max depth 5 seems to best generalizes the dataset as the test error is similar with training set size and the test error does not decrease significantly with increase of model complexity after 5.

```
In []:
```