ISRL Chapter 8 Lab 2 - Fitting Regression Trees

p327

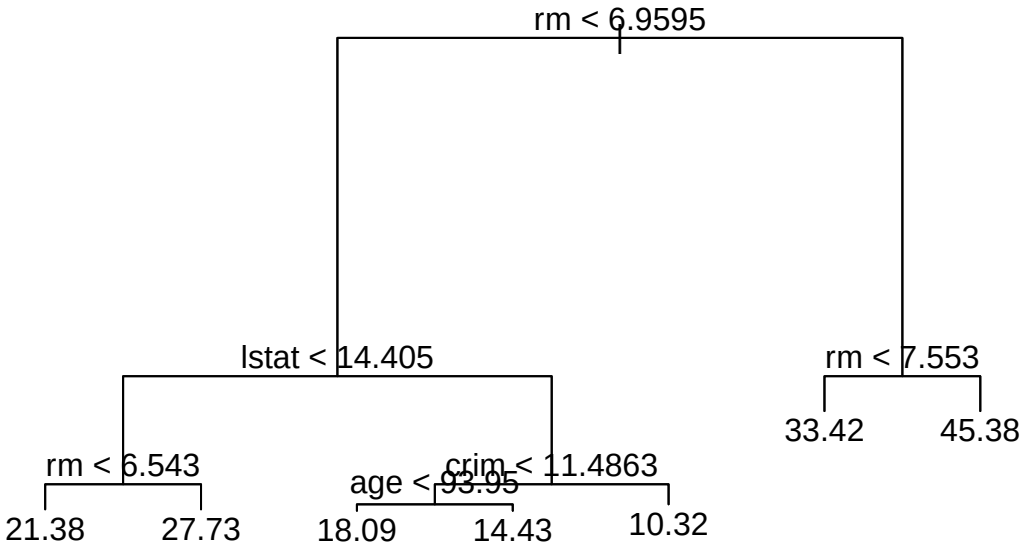
```
library(tree)
library(MASS)
set.seed(1)
train = sample(1:nrow(Boston), nrow(Boston)/2)
tree.boston=tree(medv ~ ., Boston, subset=train)
summary(tree.boston)
```

```
##
## Regression tree:
## tree(formula = medv ~ ., data = Boston, subset = train)
## Variables actually used in tree construction:
## [1] "rm"    "lstat" "crim"  "age"
## Number of terminal nodes:  7
## Residual mean deviance:  10.38 = 2555 / 246
## Distribution of residuals:
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -10.1800  -1.7770  -0.1775   0.0000   1.9230  16.5800
```

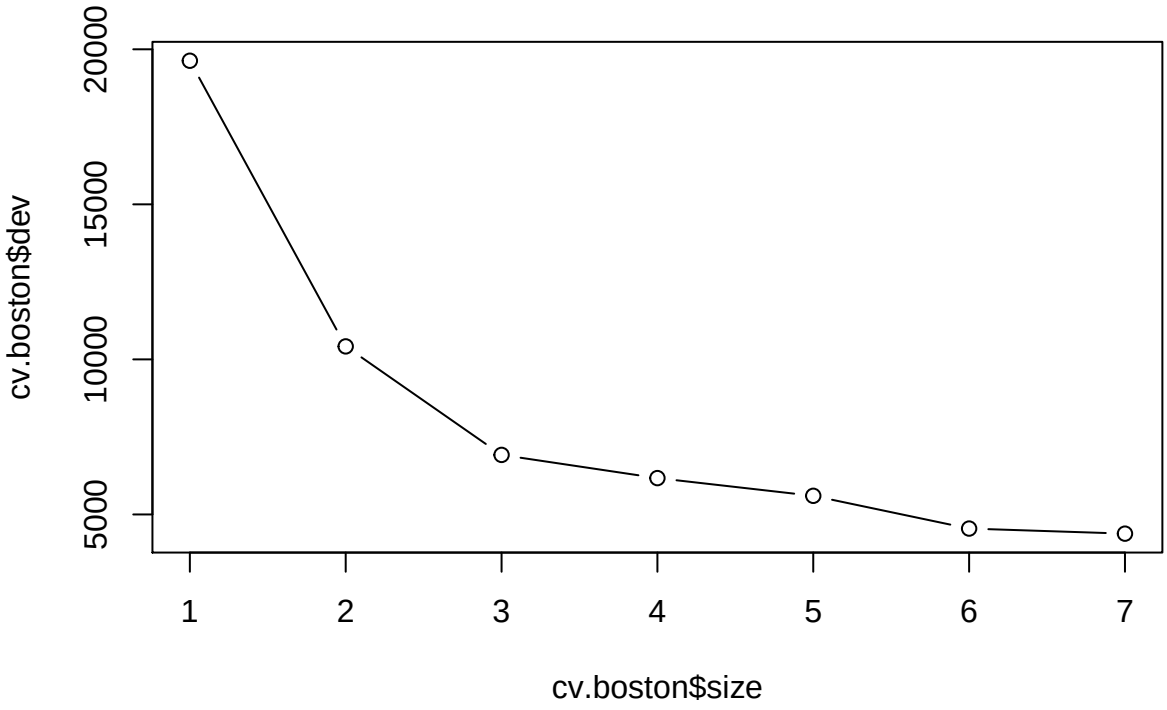
Fix the plot.new error. StackOverFlow Answer

Plot the Unpruned Tree

```
{
plot(tree.boston)
text(tree.boston,pretty=0) # plot.new error
}
```



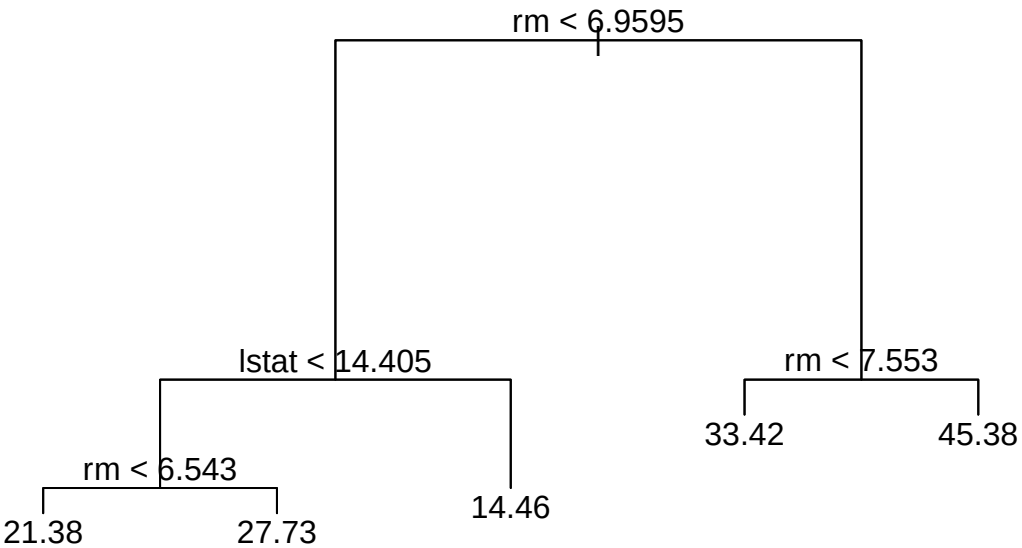
```
cv.boston=cv.tree(tree.boston)
plot(cv.boston$size, cv.boston$dev, type='b')
```



Prune the Tree

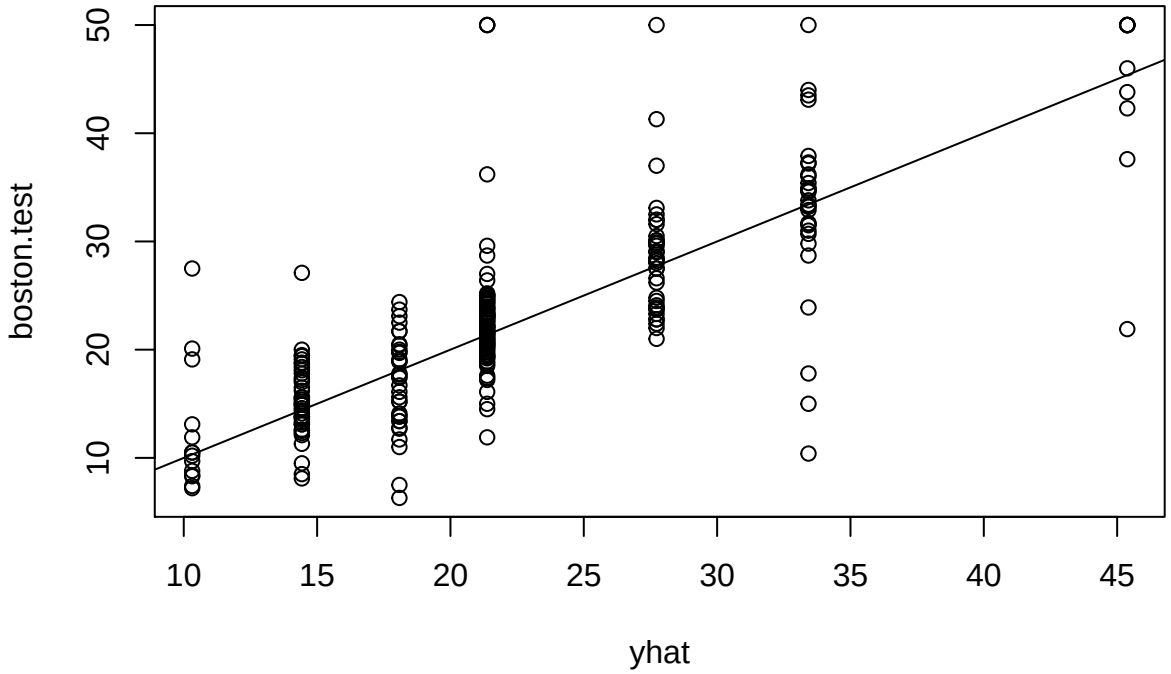
```
prune.boston=prune.tree(tree.boston, best=5)

{plot(prune.boston)
text(prune.boston, pretty=0)
}
```



Predict the Response for the Test Data with Unpruned Tree

```
yhat=predict(tree.boston, newdata=Boston[-train ,])
boston.test=Boston[-train, "medv"]
{plot(yhat,boston.test)
abline (0 ,1)
}
```




```
mean((yhat-boston.test)^2)
```

```
## [1] 35.28688
```