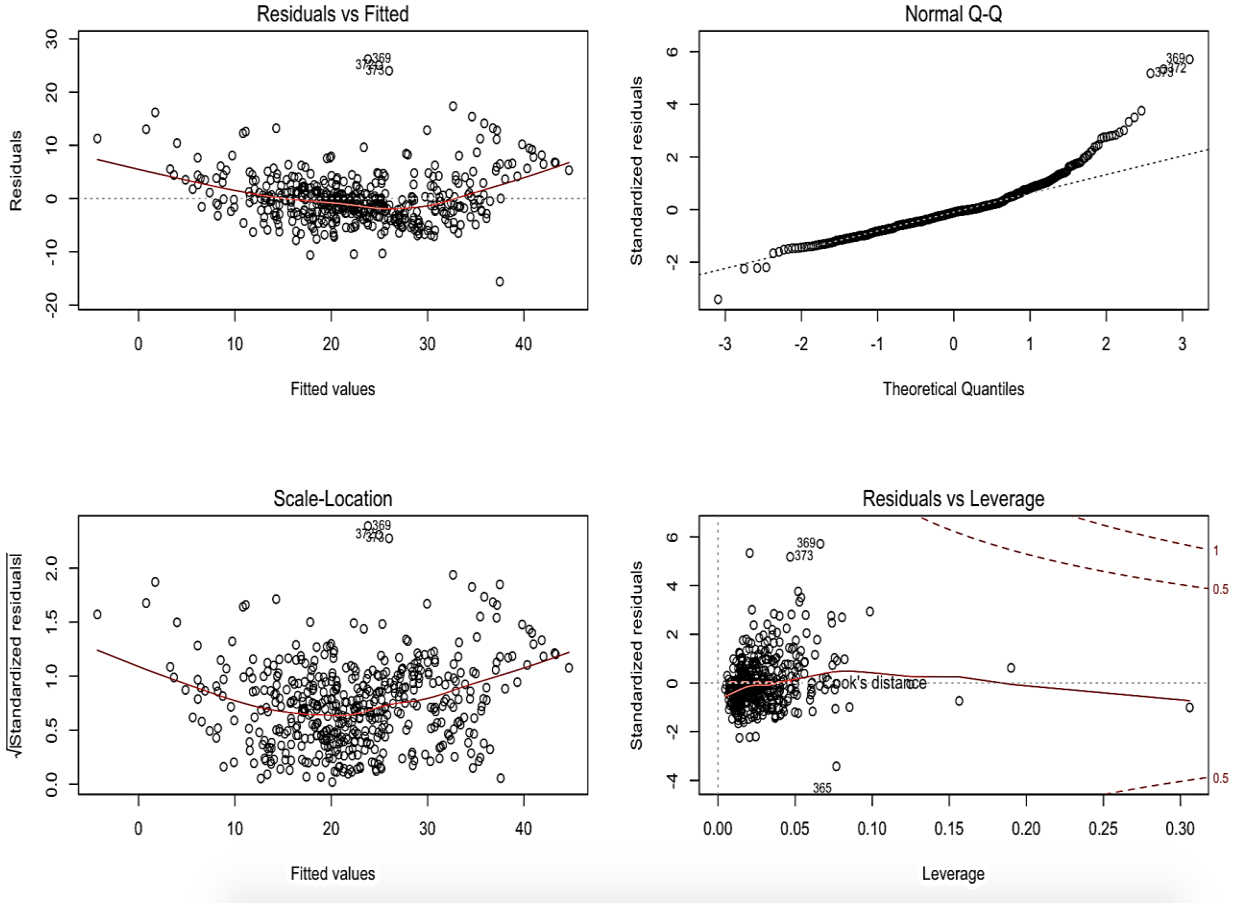
Page 1: code for regression and resulting model.

```
library(MASS)
library(mlbench)

data(BostonHousing)
summary(BostonHousing)
fit=lm(medv~.,BostonHousing)
par(mfrow=c(2,2))
plot(fit)
```

Page 2: a screenshot of your diagnostic plot and a few sentences of your explanation.

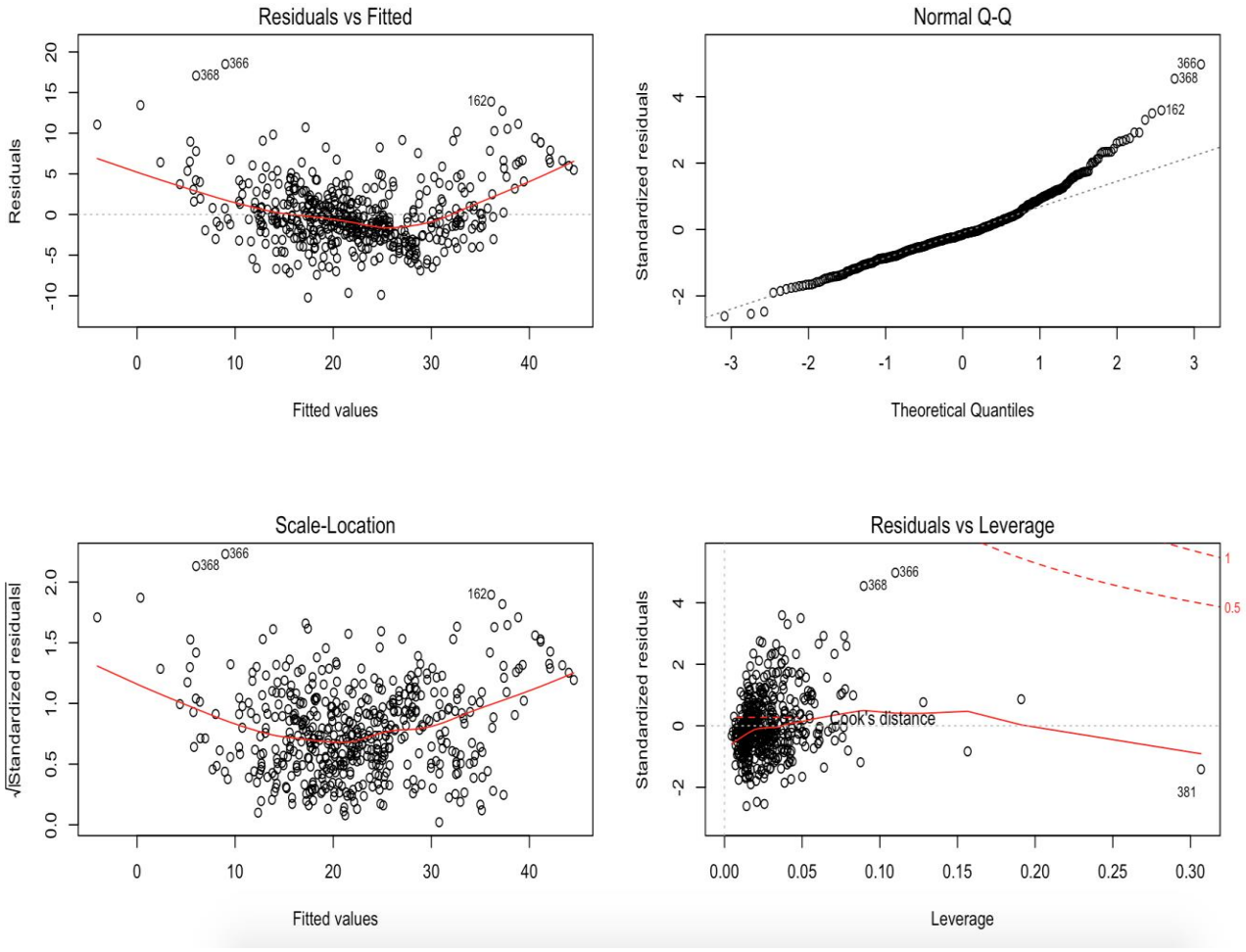


Check the points at the right of the Q-Q plot. Some points have quantiles larger than 2 and the point have quantile of -3, they are obviously deviating from the line. Therefore, they are possibly the outliers.

Check the Residuals vs Leverage plot. We can see there are points with large residuals and leverage, meaning that they have large impact on regression.

Combining these two observations, we decided to remove points that have standardized residuals larger than 3 and points have cook's distance larger than 4/n. It results in removing 8 points.

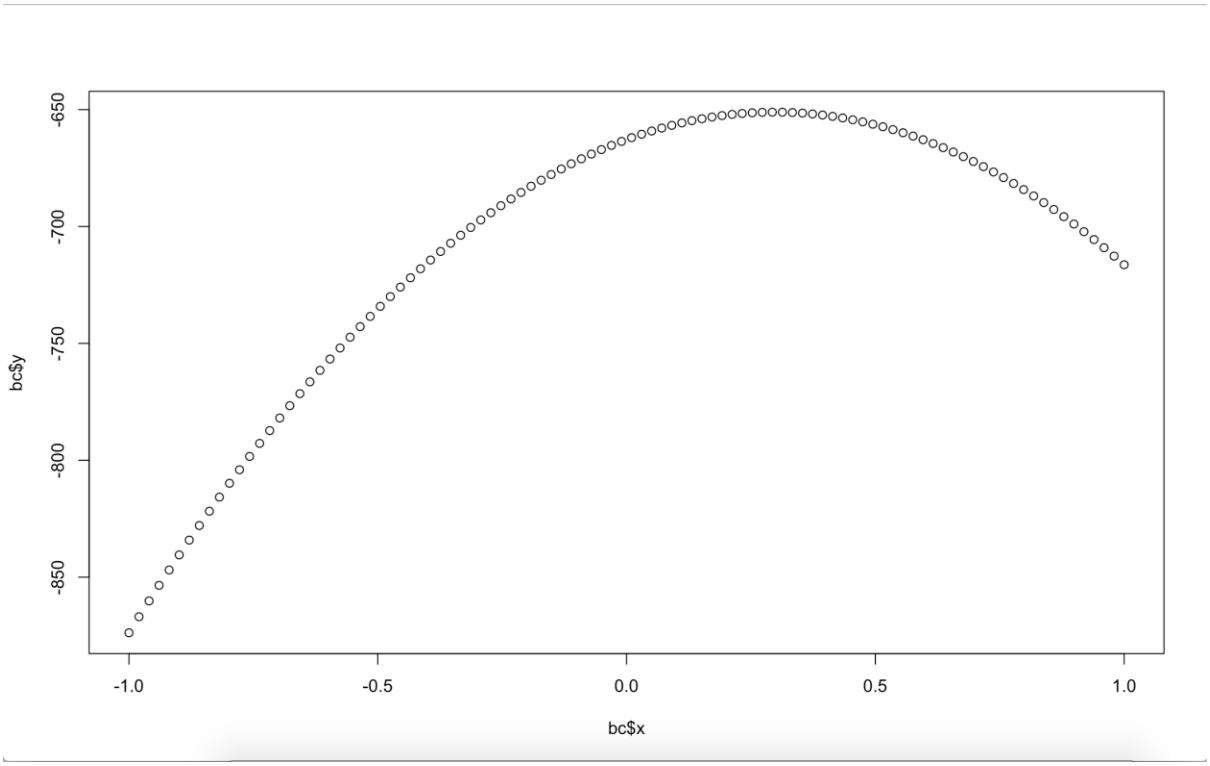
Page 3: a screenshot of your new diagnostic plot.



Page 4: a screenshot of your code for subproblem 2.

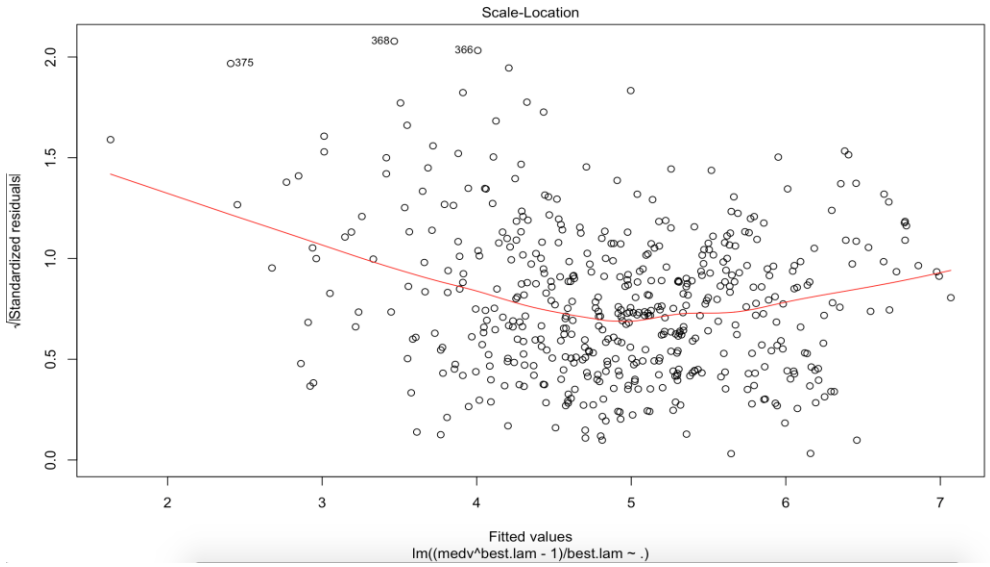
```
cooksd = cooks.distance(fit)
index.cook<-which(cooksd>4/nrow(BostonHousing))
stdres = abs(rstandard(fit))
index.stdres<-which(stdres>3)
removed<-intersect(index.stdres,index.cook)

# remove certain rows
newData <- BostonHousing[-removed,]
newfit=lm(medv~.,newData)
```

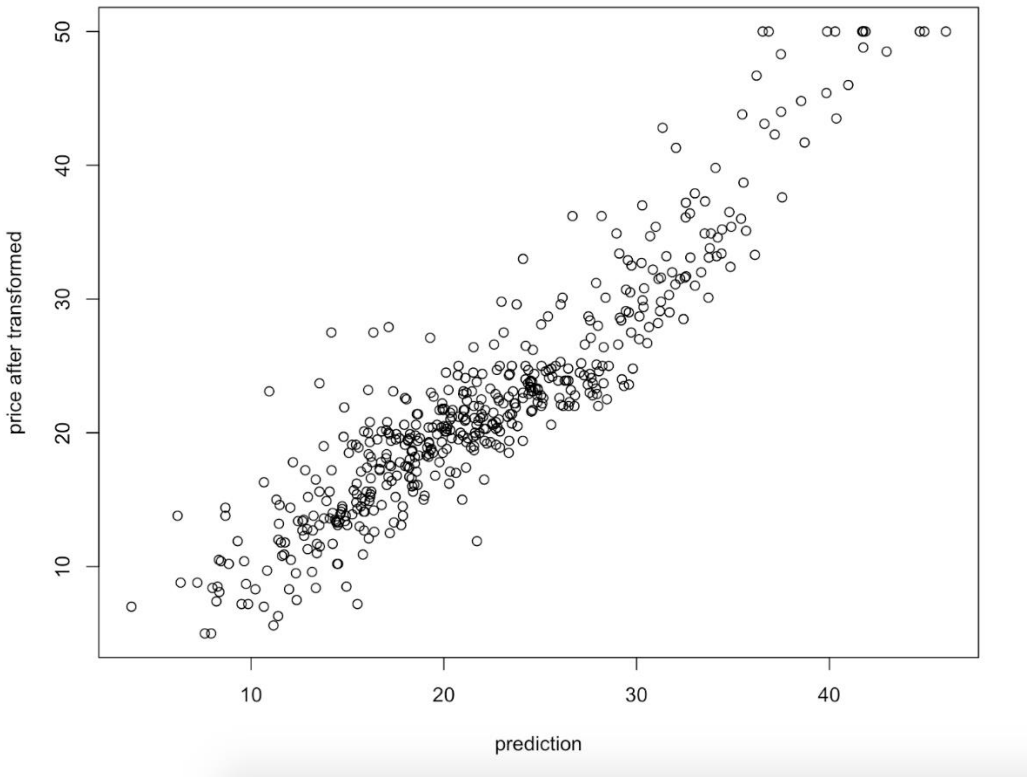
Page 5: a screenshot of Box-Cox transformation plot and the best value you chose.



I choose lambda = 0.30

Page 6: result of the standardized residuals of the regression after Box-Cox transformation and a plot of fitted house price against true house price.





Page 7: code for subproblems 3 and 4.

```
# problem 3
bc = boxcox(newfit, lambda = seq(-1,1))
plot(bc)
best.lam = bc$x[which(bc$y==max(bc$y))]

#problem 4
new_y=(newData$medv^best.lam-1)/best.lam
newfit.inv = lm((medv^best.lam-1)/best.lam~.,newData)
par(mfrow=c(2,2)) # Change the panel layout to 2 x 2
plot(newfit.inv)
par(mfrow=c(1,1))

# true vs fitted
plot((predict(newfit.inv,newData[,1:13])*best.lam+1)^(1/best.lam),newData$medv,
     xlab = "prediction",ylab="price after transformed")
```