

Frame5

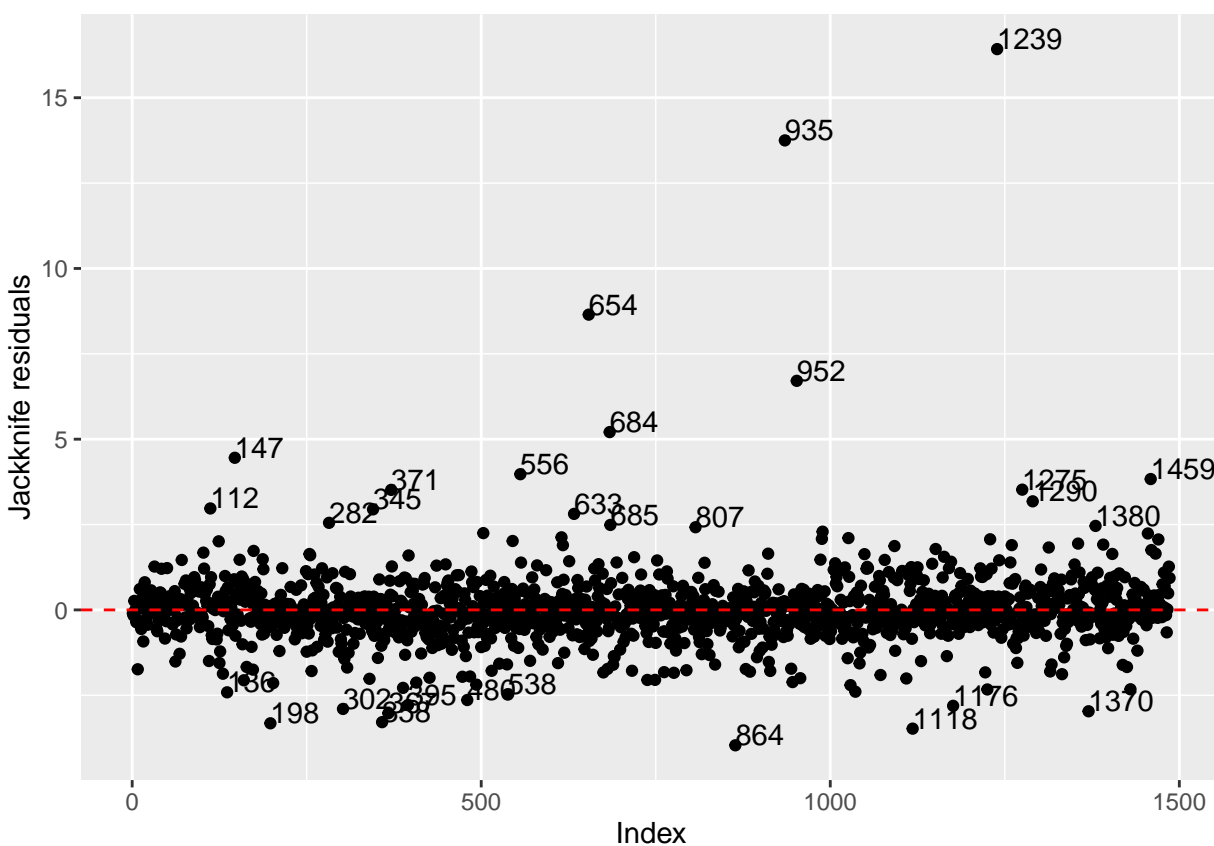
Ottó Hólm Reynisson

15.september 2016

Residuals

Begin by looking at the residuals from this model

```
indexPlotJackResiduals(lm.all)
```

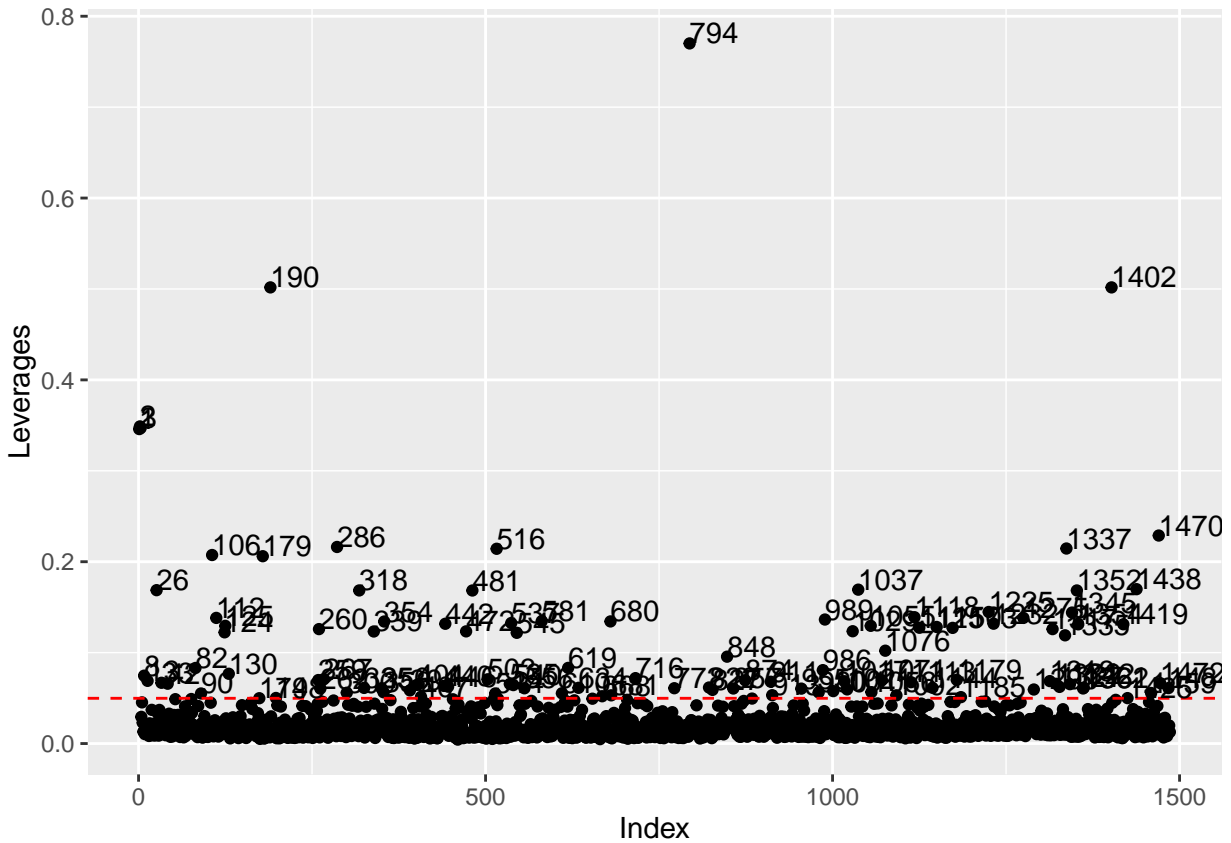


Here the blue line represent 2 standard deviation from 0 and green line 3 standard deviation away from 0. To help us identify the possible outliers we mark points 2 sd from 0 with their index.

Leverages

The next thing to do is looking at the leverages, that is the measure of how far independent variable values of an observation are from those of the other observation.

```
indexPlotLeverage(lm.all)
```



Studentized residuals

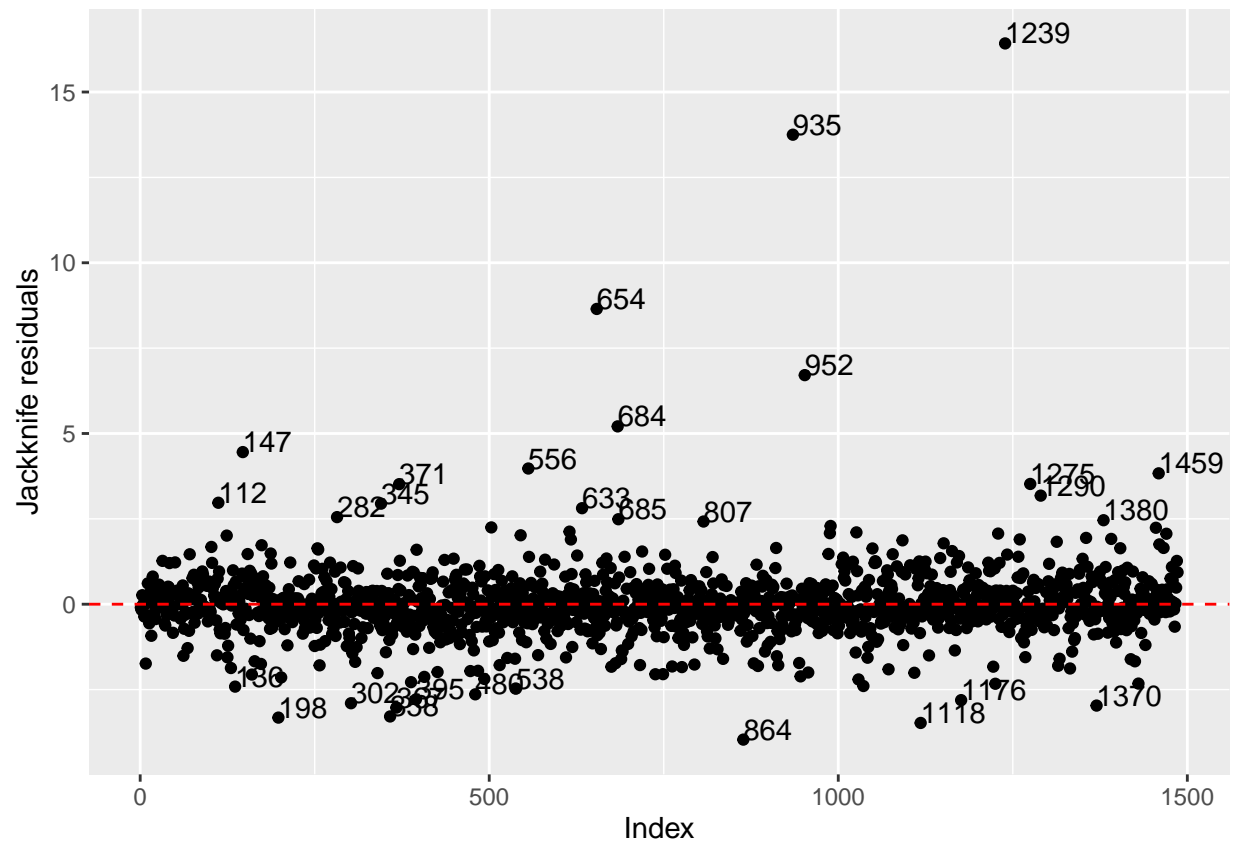
Studentized residuals are sometimes preferred in residual plots as they have been standardized to have equal variance. They are also a big part in the Jackknife residuals that follows

```
indexPlotStResiduals(lm.all)
```

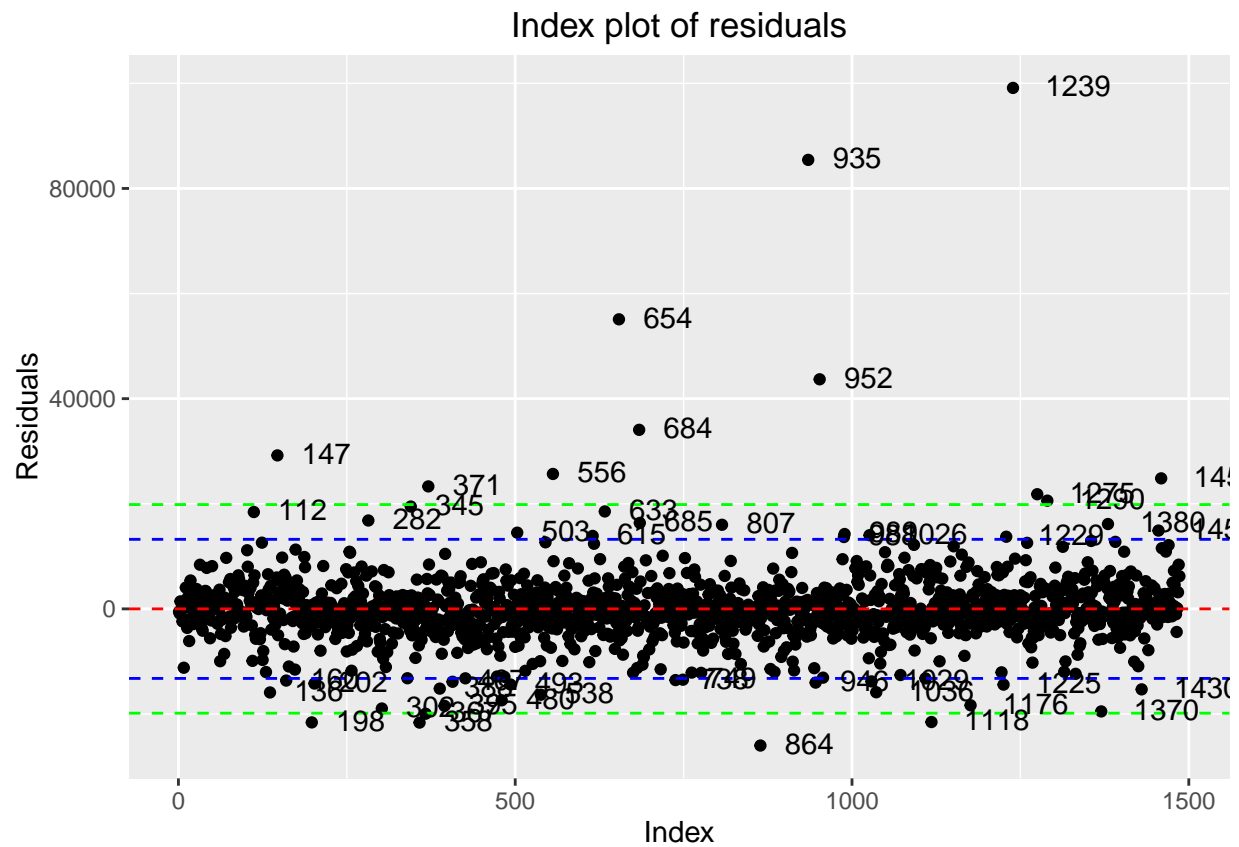


Blue and green line represent as before 2 and 3 sd from zero.

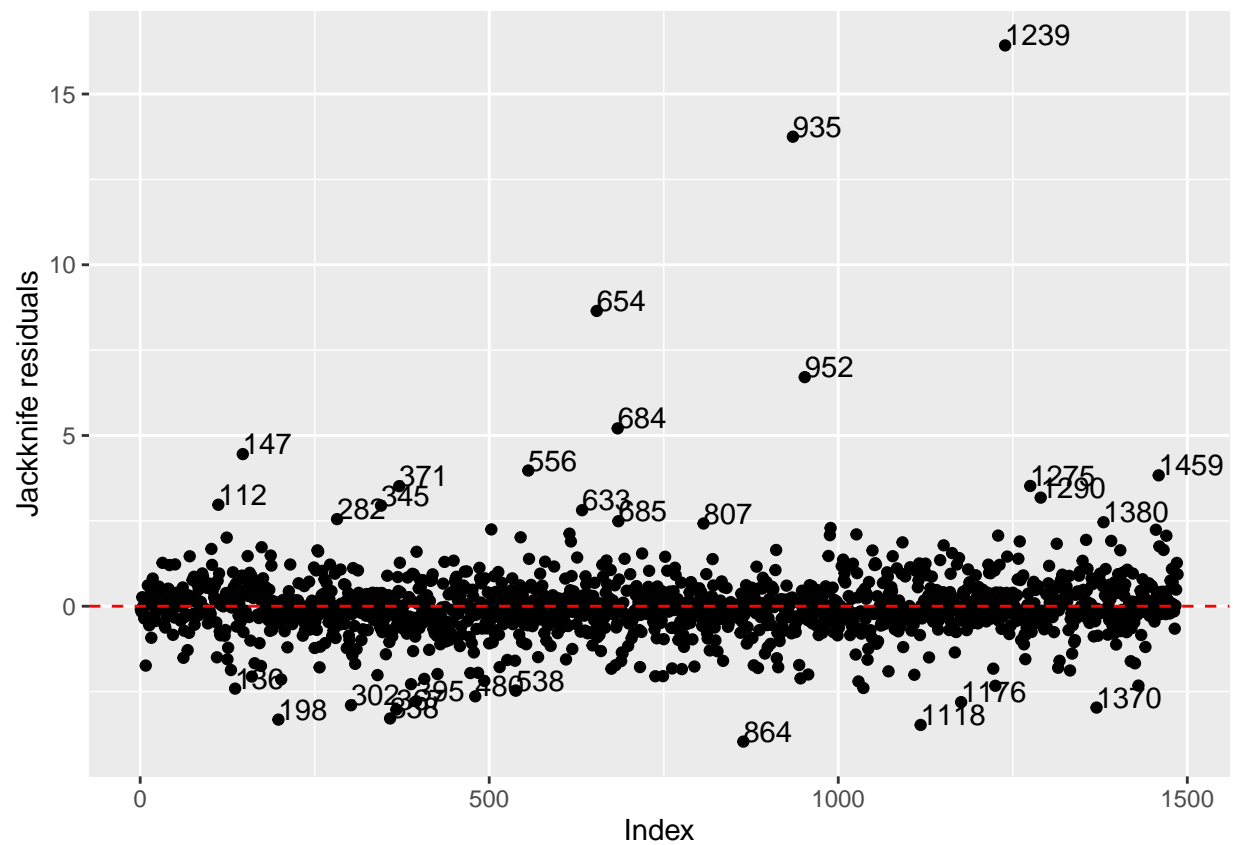
```
indexPlotJackResiduals(lm.all)
```



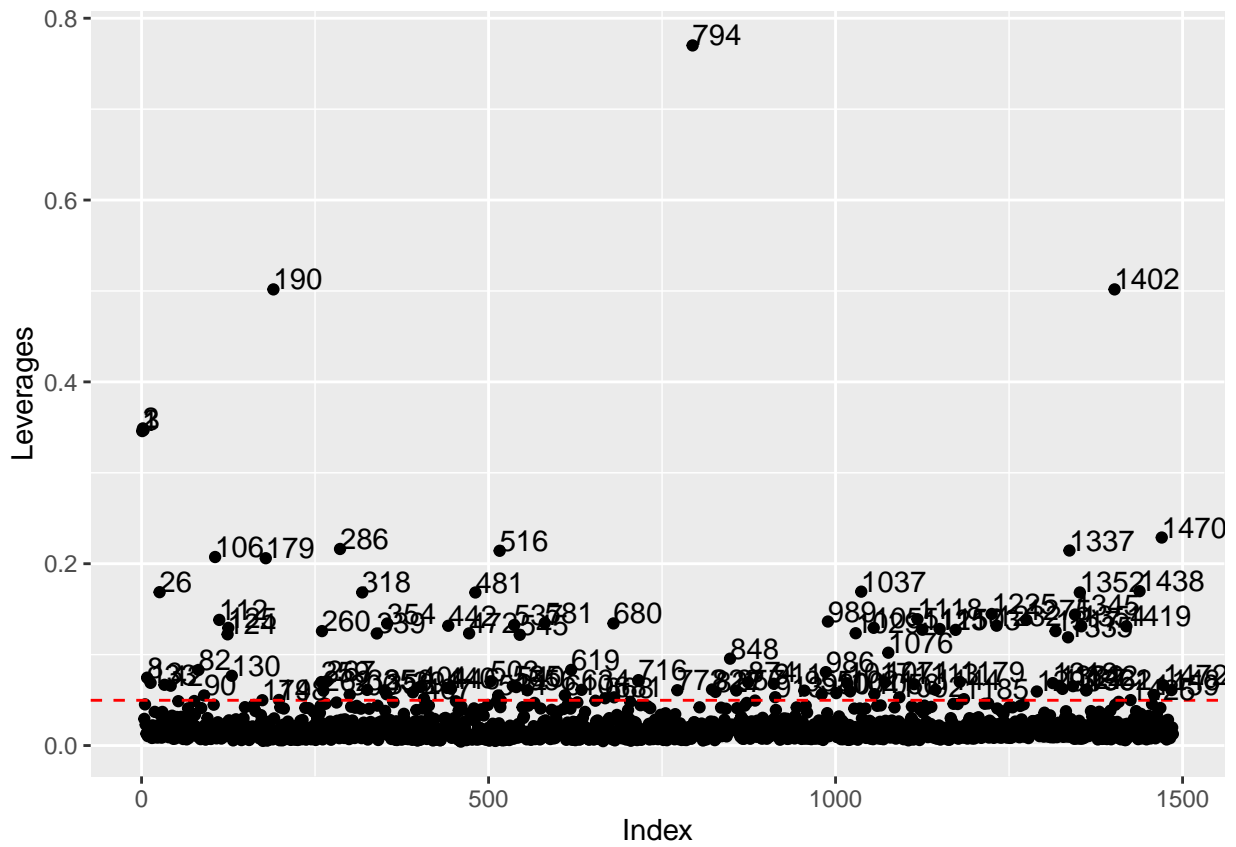
```
##### Föll fyrir residuals og annað er í outliers.R skránni #####
# functions to plot residuals and leverage
indexPlotResiduals(lm.all)
```



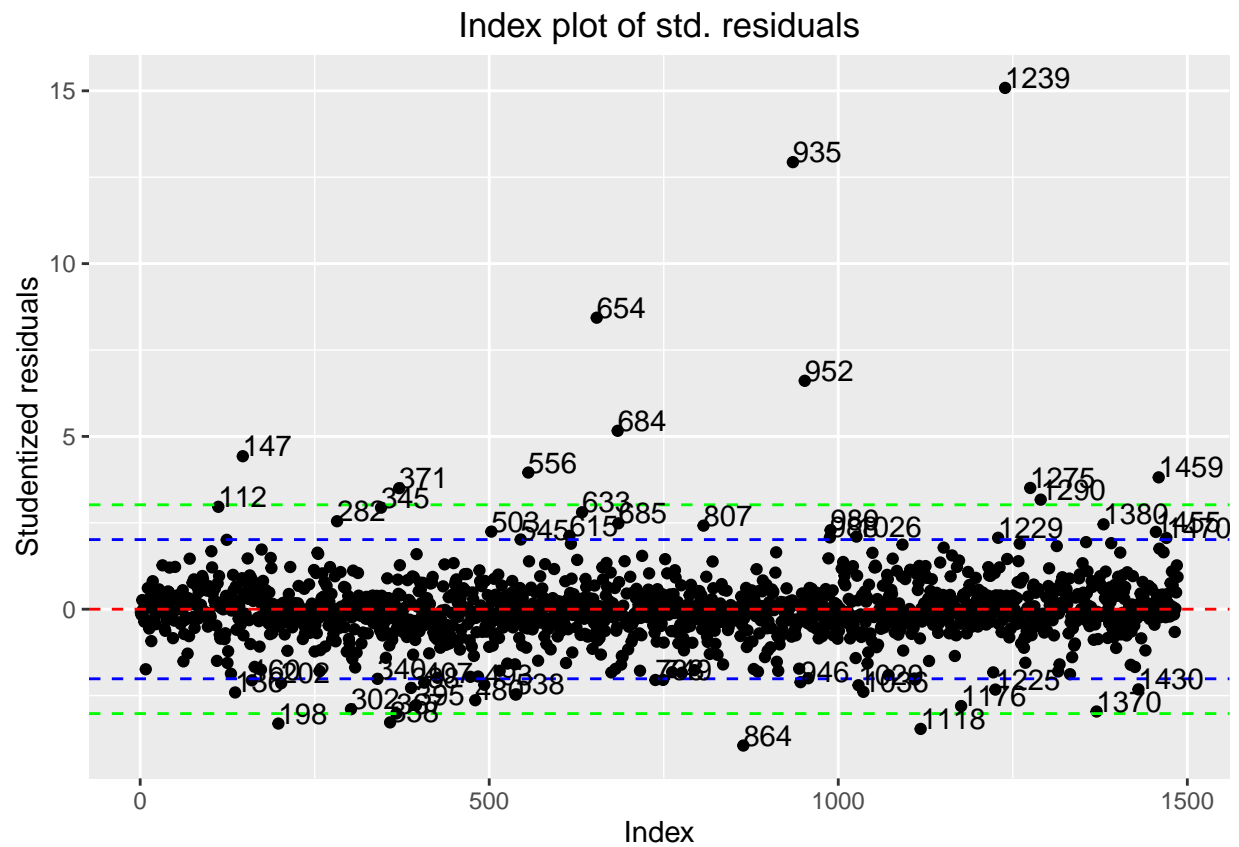
```
indexPlotJackResiduals(lm.all)
```



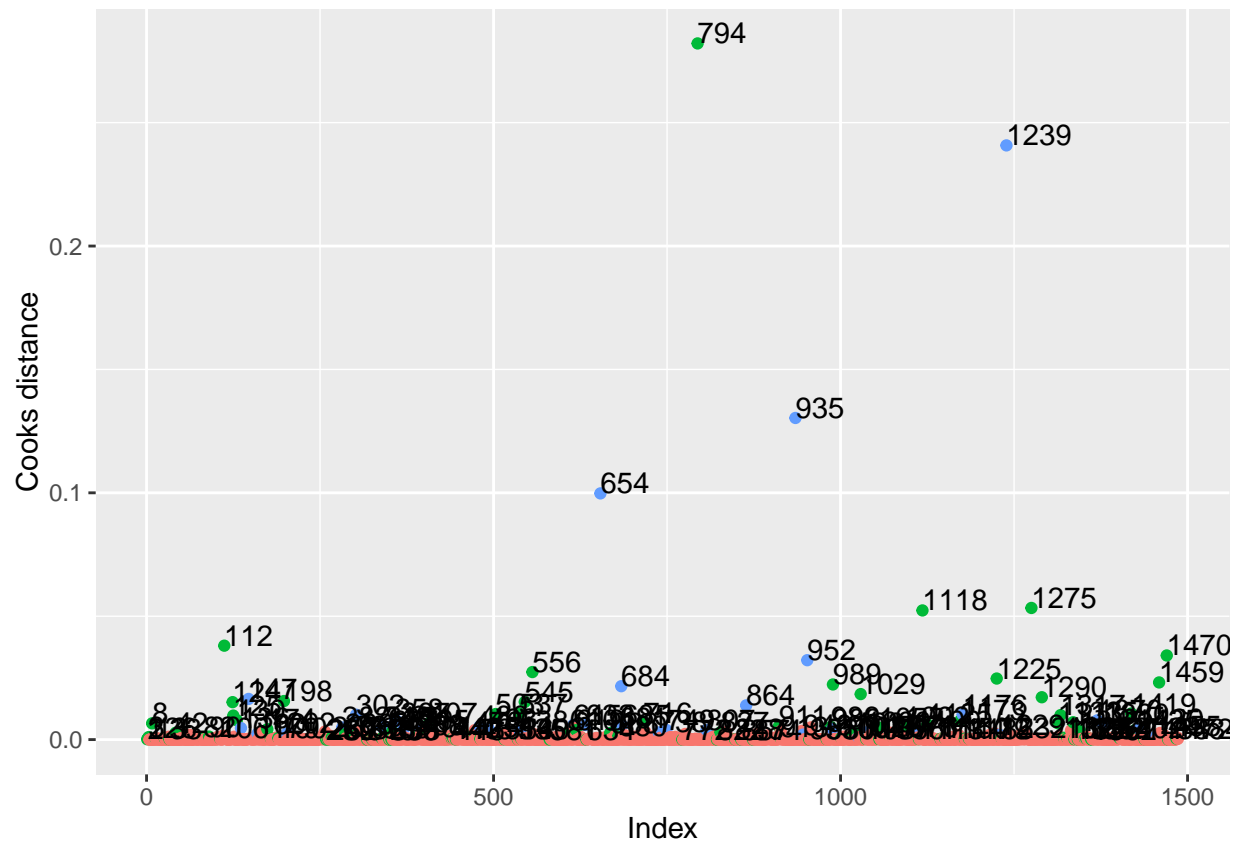
```
indexPlotLeverage(lm.all)
```



```
indexPlotStResiduals(lm.all)
```



blue dots have high residuals and green have high leverage
`indexPlotCookdistance(lm.all)`



```
##### R squared adjusted fyir lm.all
Radj.all <- CalculateRadjusted(lm.all, test) # 0.7990392
```

```
##### lm.allNoOutliers Here I take out all outliers and get a new model
lm.allNoOutlier <- removeOutliersWStdResMoreThanThree(lm.all)
Radj.allNoOutliers <- CalculateRadjusted(lm.allNoOutlier, test) # 0.7328961
```

```
##### lm.allNoInfluential Here I take out all influential points and get
lm.allNoInfluential <- removeInfluential(lm.all,maxCookDistance = 4/n)
Radj.allNoInfluential <- CalculateRadjusted(lm.allNoInfluential, test) # 0.7327364
```

Transformation