Fuel Outlier HW

```
library(fpp)
```

```
## Loading required package: forecast
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

```
## Loading required package: fma
```

```
## Loading required package: expsmooth
```

```
## Loading required package: lmtest
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
## Loading required package: tseries
```

```
library(fpp2)
```

```
## -- Attaching packages --------------------------------------------- fpp2 2.4 --
```

```
## v ggplot2 3.3.5
```

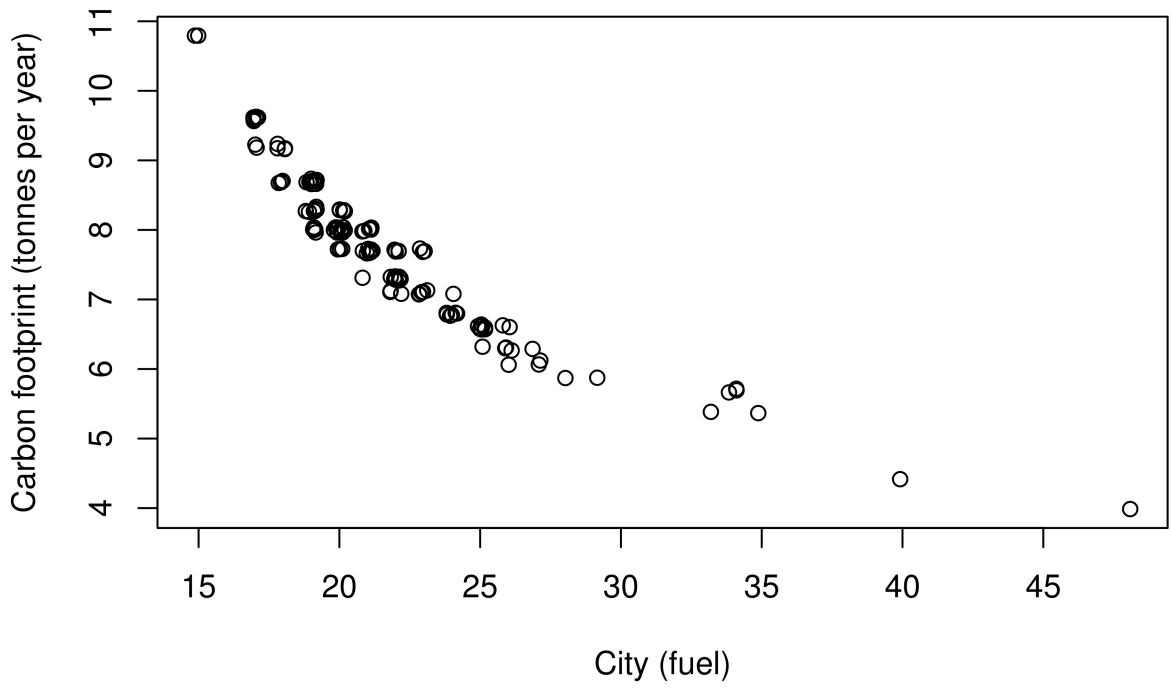
```
##
```

```
##
## Attaching package: 'fpp2'
```

```
## The following objects are masked from 'package:fpp':
##
##     ausair, ausbeer, austa, austourists, debitcards, departures,
##     elecequip, euretail, guinearice, oil, sunspotarea, usmelec
```

```
View(fuel)
```

```
# Get and load library FPP
# Lets do a plot of two variables
plot(jitter(Carbon) ~ jitter(City), xlab="City (fuel)",
ylab="Carbon footprint (tonnes per year)", data=fuel)
```



This is a non-linear relationship.

```
# Simple Linear Regression
fit <- lm(Carbon ~ City, data=fuel)
```

```
res <- residuals(fit)
```

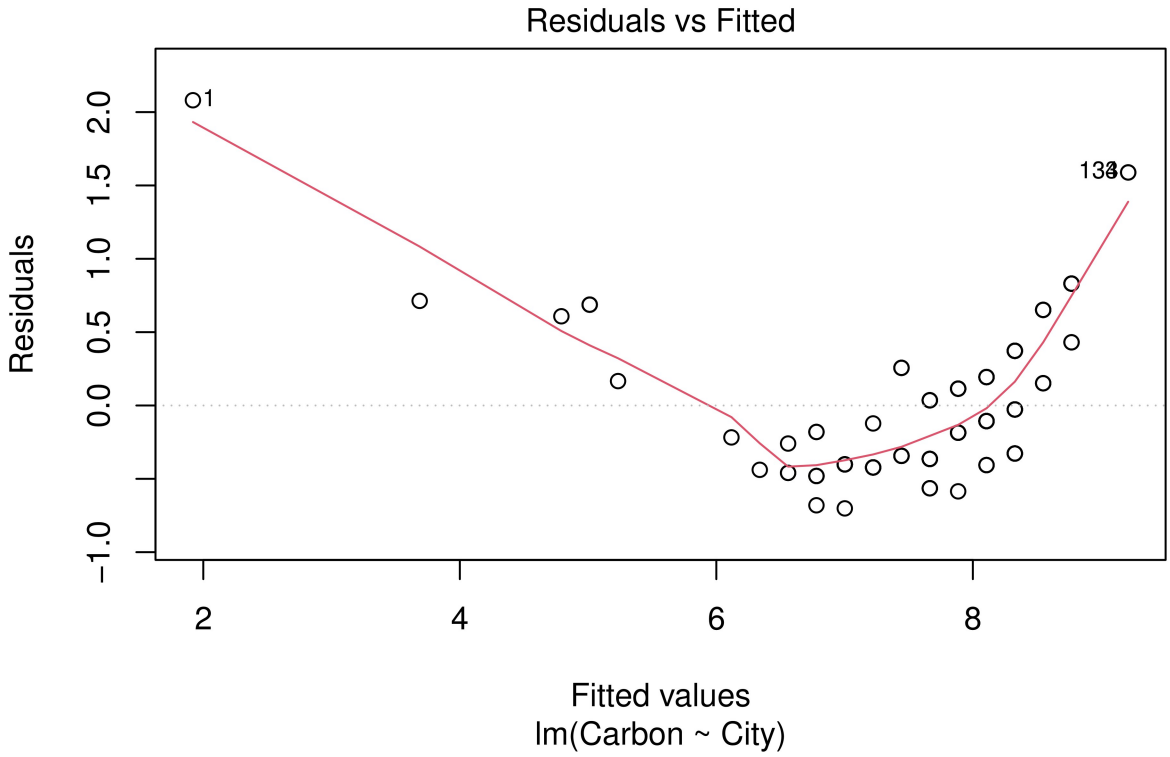
On with Forecasting

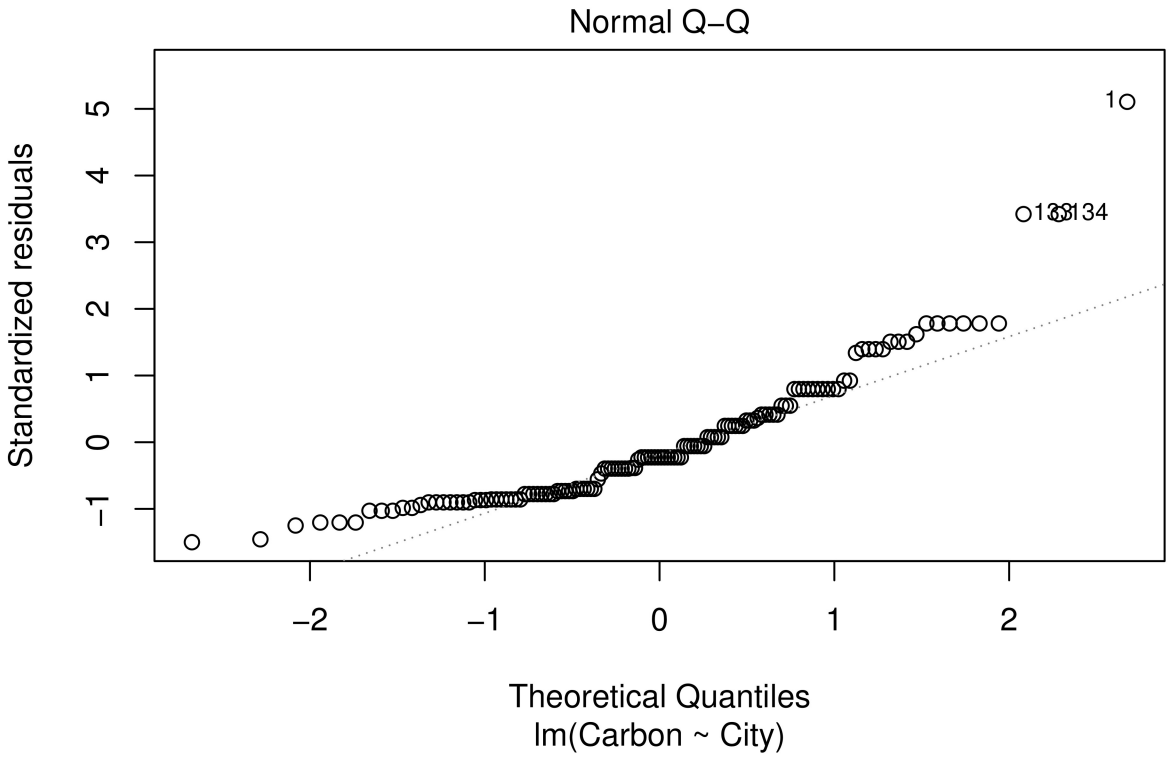
```
# to see the output from regression
fitted(fit)[1]
```

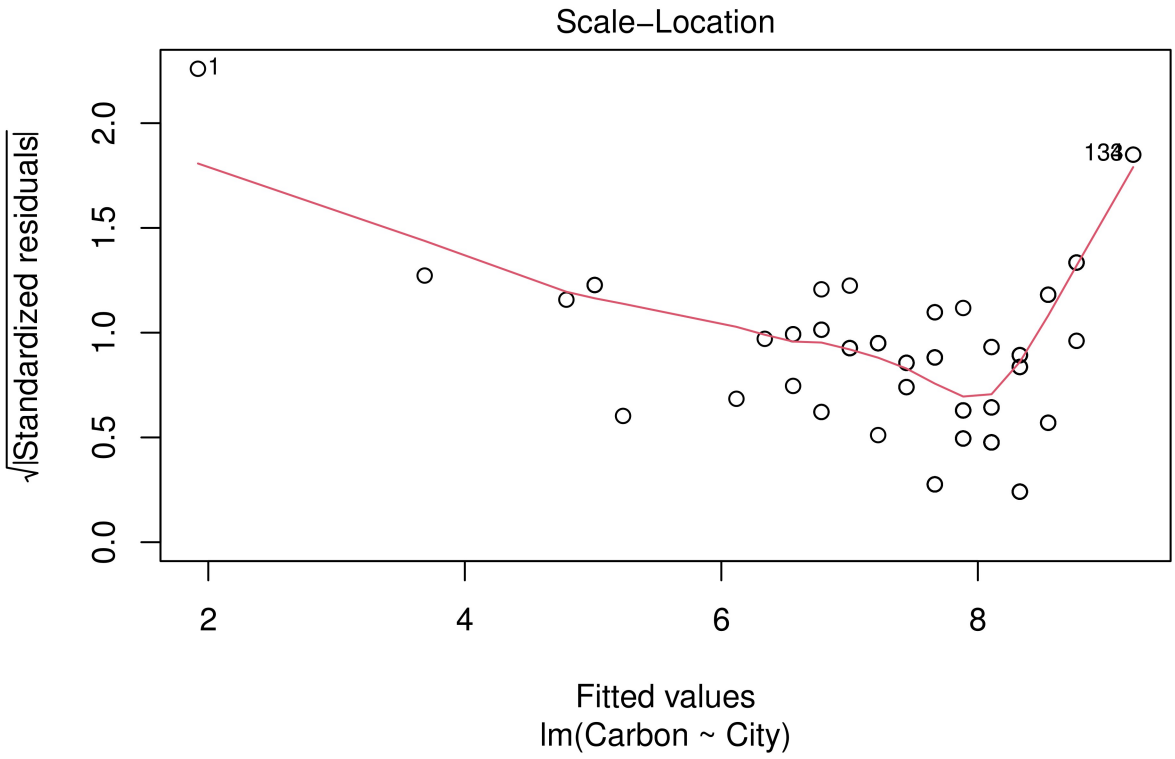
```
##       20
## 7.001388
```

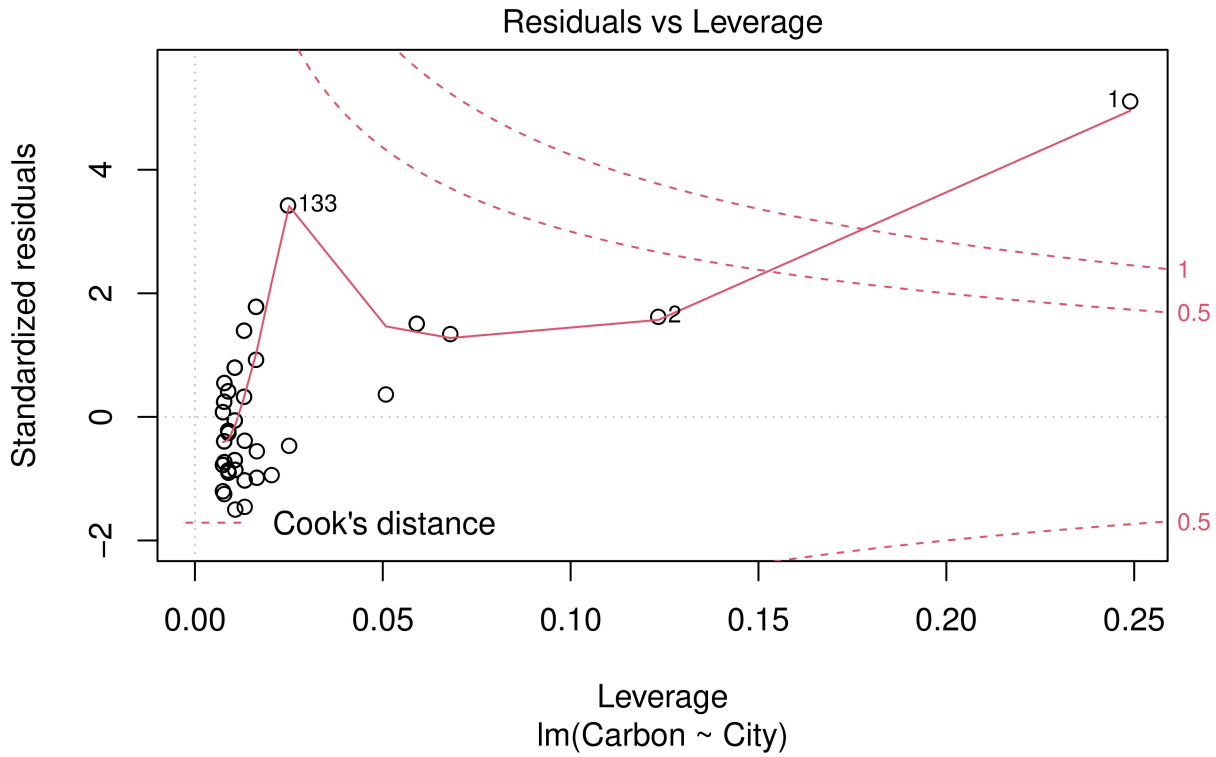
```r
# ask to do a forecast for predictor = 30
fcast <- forecast(fit, newdata=data.frame(City=30))
```

```r
plot(fit)
```




Normal Q−Q
Standardized residuals
−1
0
1
2
3
4
5
1
133
134
−2
−1
0
1
2
Theoretical Quantiles
lm(Carbon ~ City)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
1
2
4
6
8
Fitted values
lm(Carbon ~ City)



Outliers: 1: Toyota Prius 2: Honda Civic Hybrid 133: Chevrolet Colorado Cab Chassis inc 2WD

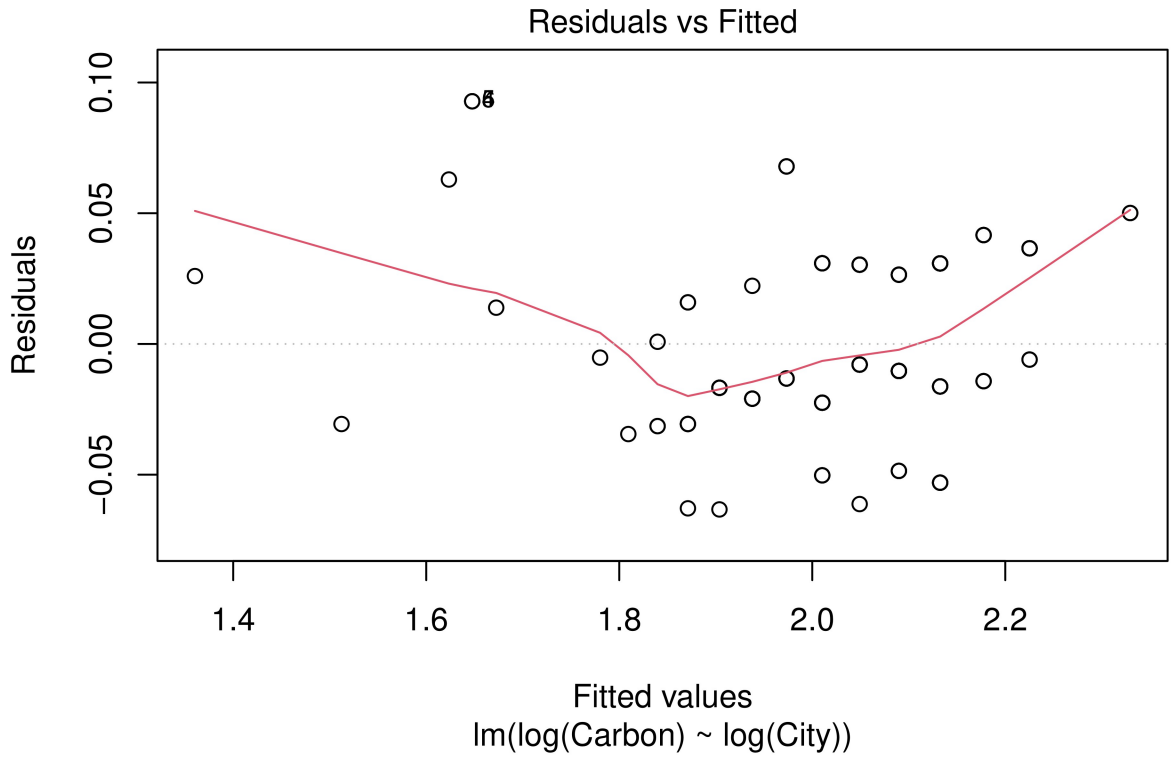
```
# Lets do a transformation of response and predictors to get a better fit.
# Split the plotting output area into 2 columns and 1 row.
par(mfrow=c(1,2))
```

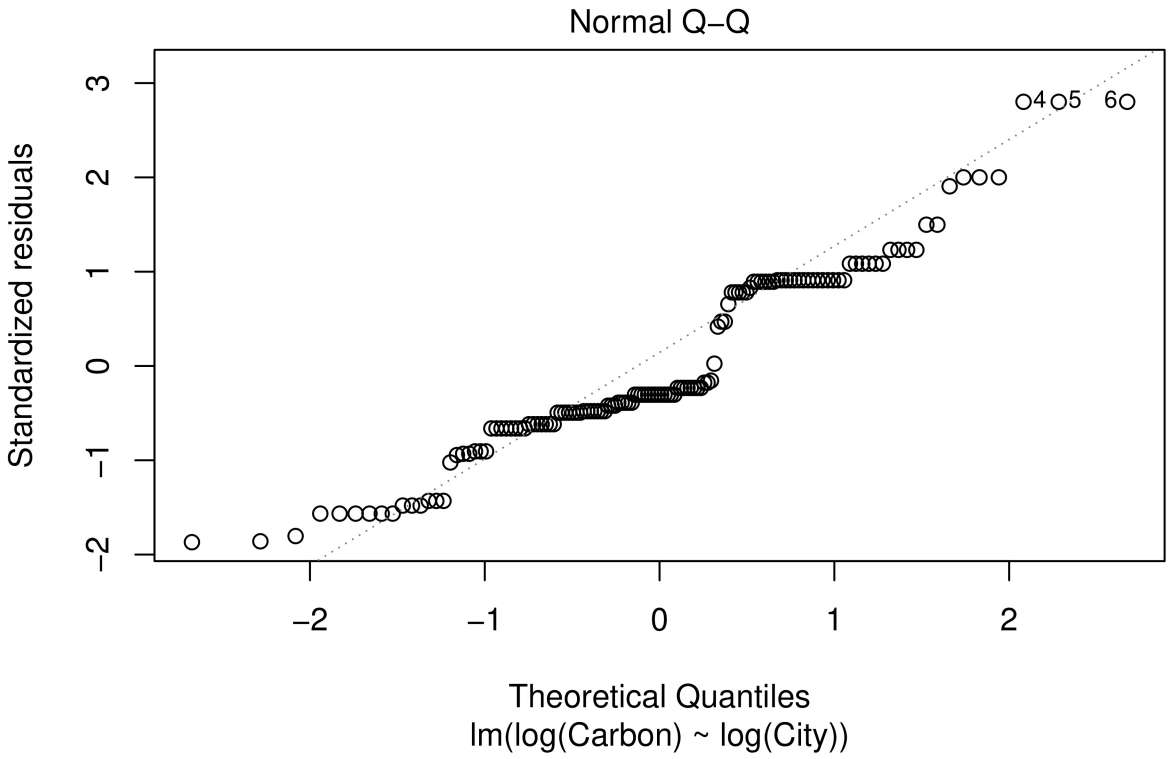
```
# Do a Regression on log of the data.
fit2 <- lm(log(Carbon) ~ log(City), data=fuel)
```

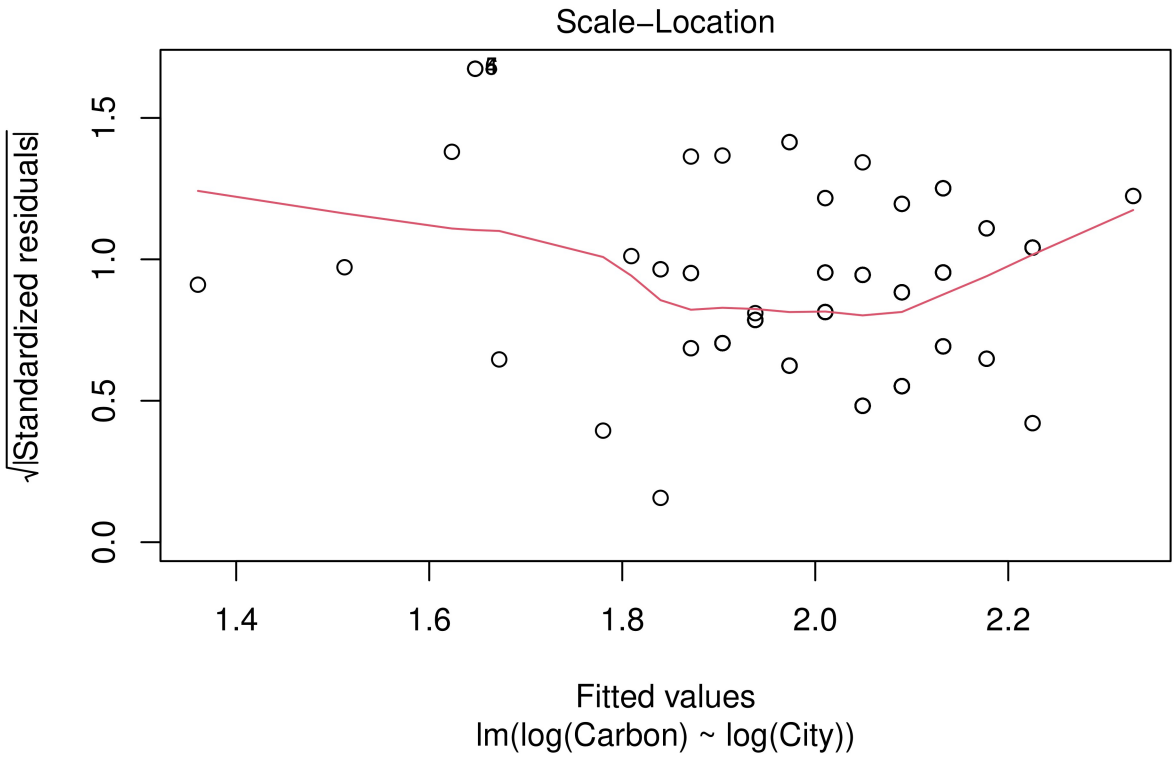
```
lm(fit2)
```

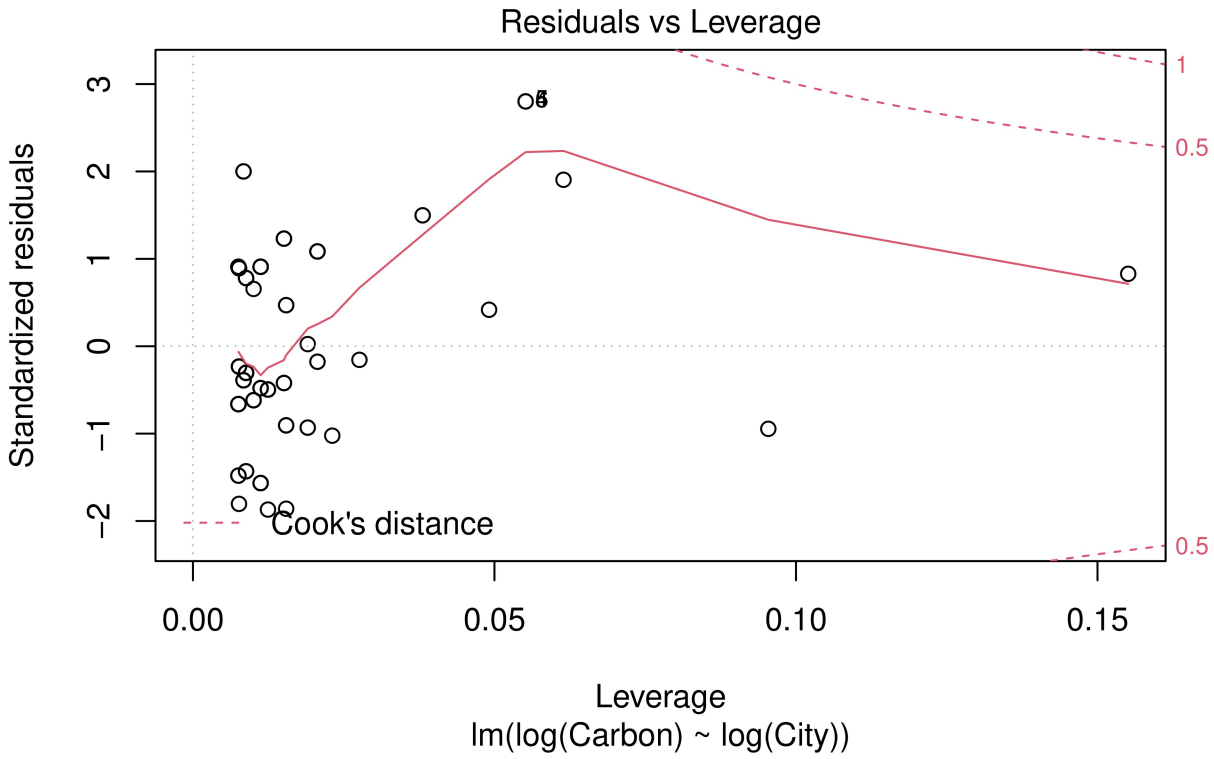
```
##
## Call:
## lm(formula = fit2)
##
## Coefficients:
## (Intercept)    log(City)
##      4.5858      -0.8332
```

```
plot(fit2)
```


Residuals vs Fitted
Residuals
0.10
0.05
0.00
−0.05
1.4
1.6
1.8
2.0
2.2
Fitted values
lm(log(Carbon) ~ log(City))


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(log(Carbon) ~ log(City))
4
5
6
−2
−1
0
1
2
3


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
1.4
1.6
1.8
2.0
2.2
Fitted values
lm(log(Carbon) ~ log(City))



Outliers: 4. Ford Escape Hybrid FWD 5. Mazda Tribute Hybrid 2WD