UK Traffic

Introduction

The UK traffic dataset is imported from kaggle. Using AADF(Annual average daily flows) describes the road traffic flow caused by each vehicles. It is categorized by the Road type.It is particularly in the Region Yokshire and Hunger.

Aim:

- To perform descriptive and explorative analysis using structure and summary measure.
- Perform histogram analysis to analyse the traffic high by vehicles.
- Perform boxplot analysis to predict under which authority the traffic is high.
- To predict which vehicle cause more traffic in particular road category.
- To predict best model based on correlation using regression model.

```
#load Uktraffic from kaggle website.

library(readxl)
traffic <- read_excel("~/Desktop/traffic.xlsx")View(traffic)



(i).Structure of traffic
str(traffic)

## Classes 'tbl_df', 'tbl' and 'data.frame':    1009 obs. of  21 variables:
##  $ AADFYear                 : num  2000 2000 2000 2000 2000 2000 2000
2000 2000 2000 ...
##  $ CP                       : num  6007 6009 6035 6054 6055 ...
##  $ Estimation_method        : chr  "Counted" "Counted" "Counted"
"Counted" ...
##  $ Estimation_method_detailed: chr  "Manual count" "Manual count" "Manual
count" "Manual count" ...
##  $ Region                   : chr  "Yorkshire and the Humber" "Yorkshire
and the Humber" "Yorkshire and the Humber" "Yorkshire and the Humber" ...
##  $ LocalAuthority           : chr  "Rotherham" "Leeds" "Doncaster"
"Calderdale" ...
##  $ Road                     : chr  "M1" "M621" "M18" "M62" ...
##  $ RoadCategory             : chr  "TM" "TM" "TM" "TM" ...
##  $ Easting                  : num  446000 432150 466400 404000 426000 ...
##  $ Northing                 : num  389300 429500 407900 416600 426200 ...
```

```
##  $ StartJunction             : chr  "M18 spur" "7" "4" "22" ...
##  $ EndJunction               : chr  "33" "M1" "5" "LA Boundary" ...
##  $ LinkLength_km             : num  3.8 1.9 6.5 6 4.9 5.2 0.7 4 5.6 8.2
...
##  $ LinkLength_miles          : num  2.36 1.18 4.04 3.73 3.04 3.23 0.43
2.49 3.48 5.1 ...
##  $ PedalCycles               : num  0 0 0 0 0 0 1 0 0 0 ...
##  $ Motorcycles               : num  168 196 25 130 229 138 25 242 190 148
...
##  $ CarsTaxis                 : num  81418 52479 26170 58960 68722 ...
##  $ BusesCoaches              : num  440 291 129 319 282 180 49 151 330 337
...
##  $ LightGoodsVehicles        : num  13950 8078 5017 11241 11818 ...
##  $ V2AxleRigidHGV            : num  5351 1293 1952 3121 4030 ...
##  $ V3AxleRigidHGV            : num  656 272 257 508 624 493 90 183 410 393
...
```

#inference:

The structure helps to predict that the it has 8 character variables and 13 numeric variables.

(ii).Summary of traffic

```
summary(traffic)
```

```
##     AADFYear          CP        Estimation_method
##  Min.   :2000   Min.   : 6007   Length:1009
##  1st Qu.:2000   1st Qu.:17990   Class :character
##  Median :2000   Median :37455   Mode  :character
##  Mean   :2000   Mean   :37114
##  3rd Qu.:2000   3rd Qu.:56223
##  Max.   :2000   Max.   :73698
##  Estimation_method_detailed    Region          LocalAuthority
##  Length:1009                Length:1009        Length:1009
##  Class :character           Class :character   Class :character
##  Mode  :character           Mode  :character   Mode  :character
##
##
##
##      Road           RoadCategory          Easting          Northing
##  Length:1009        Length:1009        Min.   :365000   Min.   :380860
##  Class :character   Class :character   1st Qu.:421000   1st Qu.:408850
##  Mode  :character   Mode  :character   Median :435530   Median :426010
##                                        Mean   :444250   Mean   :426894
##                                        3rd Qu.:459800   3rd Qu.:437800
##                                        Max.   :530000   Max.   :515400
##  StartJunction      EndJunction        LinkLength_km    LinkLength_miles
##  Length:1009        Length:1009        Min.   : 0.100   Min.   : 0.06
##  Class :character   Class :character   1st Qu.: 0.800   1st Qu.: 0.50
##  Mode  :character   Mode  :character   Median : 2.000   Median : 1.24
```

```
##                                          Mean   : 3.123   Mean   : 1.94
##                                          3rd Qu.: 3.900   3rd Qu.: 2.42
##                                          Max.   :35.500   Max.   :22.06
##   PedalCycles       Motorcycles      CarsTaxis      BusesCoaches
##  Min.   :   0.00   Min.   :  3.0   Min.   :  851   Min.   :   2
##  1st Qu.:   9.00   1st Qu.: 71.0   1st Qu.: 7967   1st Qu.:  77
##  Median :  31.00   Median :122.0   Median :13043   Median : 150
##  Mean   :  75.22   Mean   :144.4   Mean   :17457   Mean   : 247
##  3rd Qu.:  66.00   3rd Qu.:192.0   3rd Qu.:20969   3rd Qu.: 309
##  Max.   :2200.00   Max.   :797.0   Max.   :91616   Max.   :2960
##  LightGoodsVehicles V2AxleRigidHGV   V3AxleRigidHGV
##  Min.   :    0      Min.   :   0.0   Min.   :   0.00
##  1st Qu.: 1004      1st Qu.: 190.0   1st Qu.:  28.00
##  Median : 1649      Median : 349.0   Median :  52.00
##  Mean   : 2438      Mean   : 654.1   Mean   :  95.88
##  3rd Qu.: 2826      3rd Qu.: 693.0   3rd Qu.: 103.00
##  Max.   :14937      Max.   :6872.0   Max.   :1600.00

#inference

      The summary measures helps to predict that the maximum traffic is
caused by cars in range of 851 to 91616 and lightgoodsvehicles as 14937.

(iii)Using dplyr function filter

f1<-filter(traffic,Estimation_method=="Counted")
View(f1)
summary(f1)

##     AADFYear          CP        Estimation_method
##  Min.   :2000   Min.   : 6007   Length:485
##  1st Qu.:2000   1st Qu.:26091   Class :character
##  Median :2000   Median :37579   Mode  :character
##  Mean   :2000   Mean   :38991
##  3rd Qu.:2000   3rd Qu.:56609
##  Max.   :2000   Max.   :73697
##  Estimation_method_detailed    Region          LocalAuthority
##  Length:485                 Length:485         Length:485
##  Class :character           Class :character   Class :character
##  Mode  :character           Mode  :character   Mode  :character
##
##
##
##      Road           RoadCategory          Easting          Northing
##  Length:485         Length:485         Min.   :378000   Min.   :381400
##  Class :character   Class :character   1st Qu.:425000   1st Qu.:410000
##  Mode  :character   Mode  :character   Median :437090   Median :426820
##                                        Mean   :444547   Mean   :429575
##                                        3rd Qu.:459800   3rd Qu.:441000
##                                        Max.   :530000   Max.   :515400
##  StartJunction      EndJunction        LinkLength_km    LinkLength_miles
```

```
##  Length:485         Length:485         Min.   : 0.100   Min.   : 0.060
##  Class :character   Class :character   1st Qu.: 1.600   1st Qu.: 0.990
##  Mode  :character   Mode  :character   Median : 3.000   Median : 1.860
##                                        Mean   : 4.515   Mean   : 2.805
##                                        3rd Qu.: 5.300   3rd Qu.: 3.290
##                                        Max.   :35.500   Max.   :22.060
##   PedalCycles       Motorcycles      CarsTaxis      BusesCoaches
##  Min.   :   0.00   Min.   :  9.0   Min.   :  882   Min.   :   3.0
##  1st Qu.:   3.00   1st Qu.: 76.0   1st Qu.: 8276   1st Qu.:  84.0
##  Median :  22.00   Median :130.0   Median :14215   Median : 156.0
##  Mean   :  60.11   Mean   :152.8   Mean   :20468   Mean   : 233.6
##  3rd Qu.:  60.00   3rd Qu.:201.0   3rd Qu.:24768   3rd Qu.: 304.0
##  Max.   :2200.00   Max.   :797.0   Max.   :91616   Max.   :1440.0
##  LightGoodsVehicles V2AxleRigidHGV   V3AxleRigidHGV
##  Min.   :  121      Min.   :   1.0   Min.   :  0.0
##  1st Qu.: 1044      1st Qu.: 203.0   1st Qu.: 33.0
##  Median : 1853      Median : 376.0   Median : 62.0
##  Mean   : 2932      Mean   : 861.3   Mean   :124.2
##  3rd Qu.: 3646      3rd Qu.:1010.0   3rd Qu.:137.0
##  Max.   :14937      Max.   :6872.0   Max.   :725.0
```

(iii) load the packages lattice and dplyr

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(lattice)
```

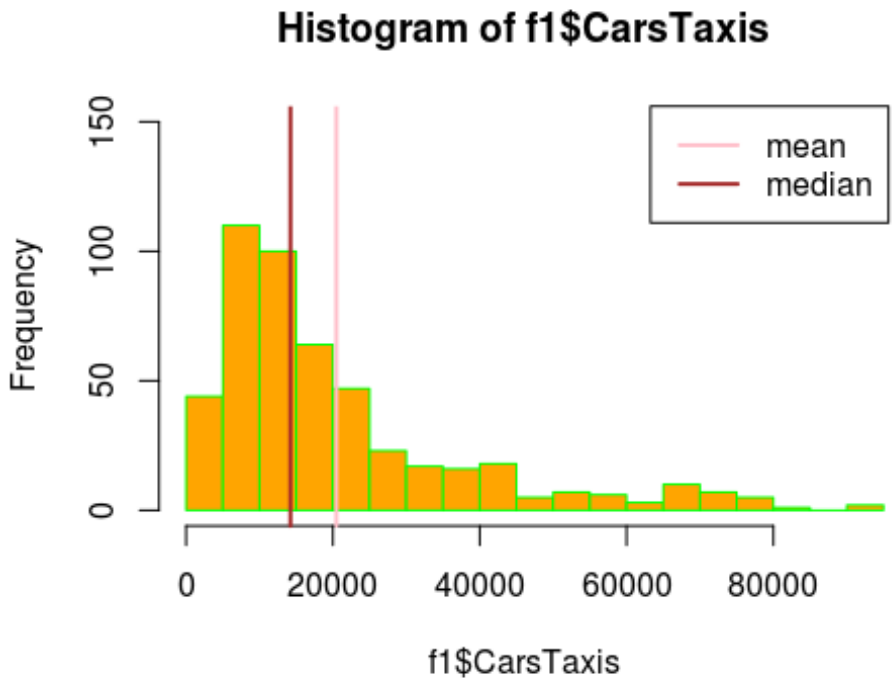
(iv)Using dplyr function filter

```
f1<-filter(traffic,Estimation_method=="Counted")
View(f1)
summary(f1)
```

```
##     AADFYear          CP        Estimation_method
##  Min.   :2000   Min.   : 6007   Length:485
##  1st Qu.:2000   1st Qu.:26091   Class :character
##  Median :2000   Median :37579   Mode  :character
##  Mean   :2000   Mean   :38991
##  3rd Qu.:2000   3rd Qu.:56609
##  Max.   :2000   Max.   :73697
```

```
##  Estimation_method_detailed    Region          LocalAuthority
##  Length:485                 Length:485         Length:485
##  Class :character           Class :character   Class :character
##  Mode  :character           Mode  :character   Mode  :character
##
##
##
##      Road           RoadCategory          Easting          Northing
##  Length:485         Length:485         Min.   :378000   Min.   :381400
##  Class :character   Class :character   1st Qu.:425000   1st Qu.:410000
##  Mode  :character   Mode  :character   Median :437090   Median :426820
##                                        Mean   :444547   Mean   :429575
##                                        3rd Qu.:459800   3rd Qu.:441000
##                                        Max.   :530000   Max.   :515400
##  StartJunction      EndJunction        LinkLength_km    LinkLength_miles
##  Length:485         Length:485         Min.   : 0.100   Min.   : 0.060
##  Class :character   Class :character   1st Qu.: 1.600   1st Qu.: 0.990
##  Mode  :character   Mode  :character   Median : 3.000   Median : 1.860
##                                        Mean   : 4.515   Mean   : 2.805
##                                        3rd Qu.: 5.300   3rd Qu.: 3.290
##                                        Max.   :35.500   Max.   :22.060
##   PedalCycles       Motorcycles      CarsTaxis      BusesCoaches
##  Min.   :   0.00   Min.   :  9.0   Min.   :  882   Min.   :   3.0
##  1st Qu.:   3.00   1st Qu.: 76.0   1st Qu.: 8276   1st Qu.:  84.0
##  Median :  22.00   Median :130.0   Median :14215   Median : 156.0
##  Mean   :  60.11   Mean   :152.8   Mean   :20468   Mean   : 233.6
##  3rd Qu.:  60.00   3rd Qu.:201.0   3rd Qu.:24768   3rd Qu.: 304.0
##  Max.   :2200.00   Max.   :797.0   Max.   :91616   Max.   :1440.0
##  LightGoodsVehicles V2AxleRigidHGV   V3AxleRigidHGV
##  Min.   :  121      Min.   :   1.0   Min.   :  0.0
##  1st Qu.: 1044      1st Qu.: 203.0   1st Qu.: 33.0
##  Median : 1853      Median : 376.0   Median : 62.0
##  Mean   : 2932      Mean   : 861.3   Mean   :124.2
##  3rd Qu.: 3646      3rd Qu.:1010.0   3rd Qu.:137.0
##  Max.   :14937      Max.   :6872.0   Max.   :725.0
```

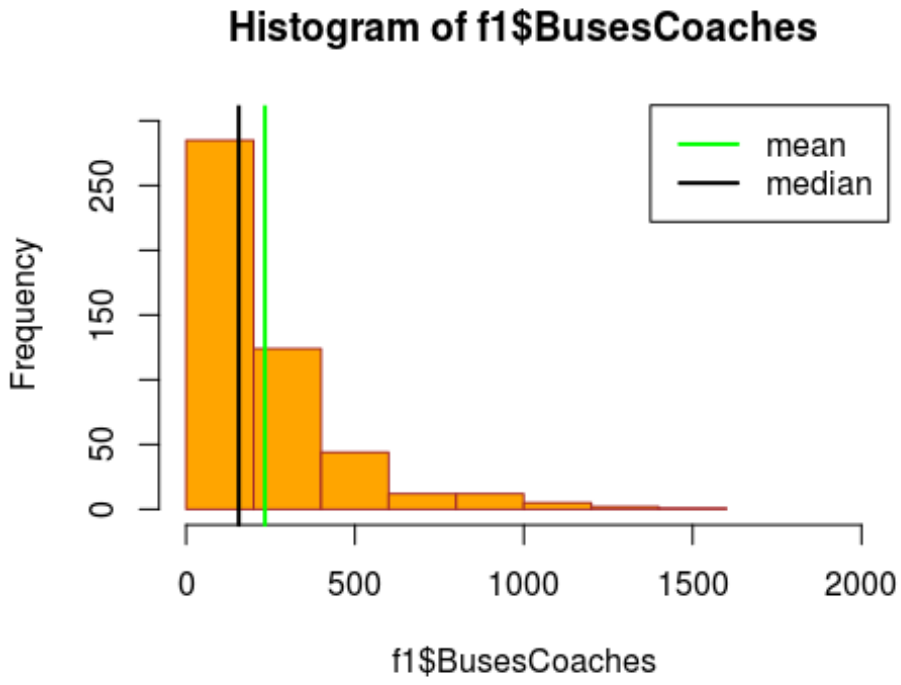
```
hist(f1$CarsTaxis,xlim=c(0,92000),border="green",breaks=30,col="Orange",ylim=
c(0,150))
abline(v=mean(f1$CarsTaxis),col="pink",lwd=2)
abline(v=median(f1$CarsTaxis),col="brown",lwd=2)
legend(x='topright',c('mean','median'),col=c('pink','brown'),lwd=c(2,2))
```




```
#inference
```

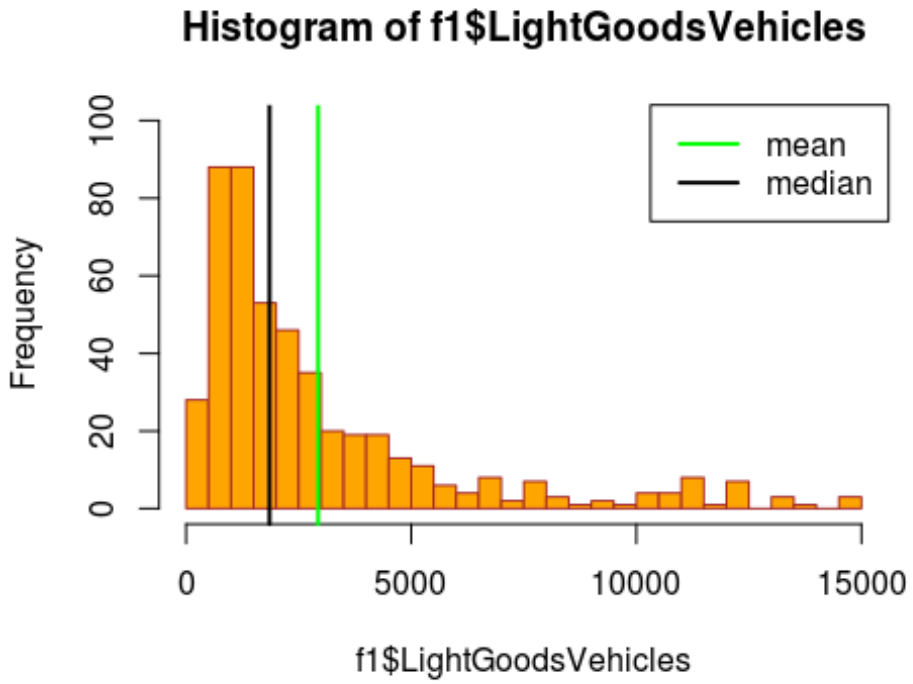
This shows that the distribution is right skewed and the traffic is high in the range of 5000 to 20000.The average of car traffic is 20000 and the traffic frequently occurs in the range of 14900

```r
hist(f1$BusesCoaches,xlim=c(0,2000),border="brown",breaks=10,col="Orange",yli
m=c(0,300))
abline(v=mean(f1$BusesCoaches),col="green",lwd=2)
abline(v=median(f1$BusesCoaches),col="black",lwd=2)
legend(x='topright',c('mean','median'),col=c('green','black'),lwd=c(2,2))
```



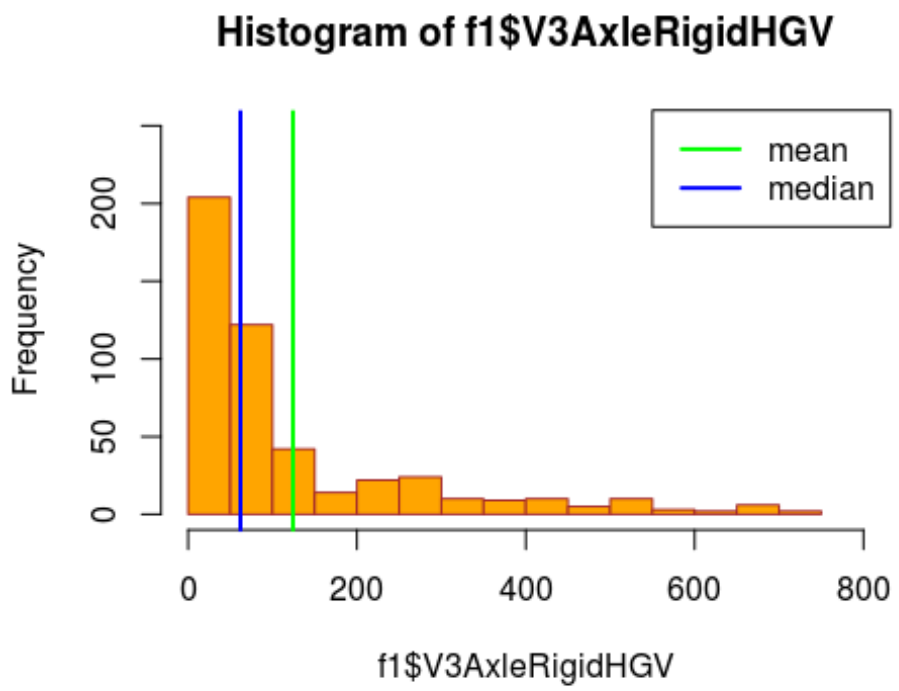

```r
#inference

      This clearly shows that the distribution is right skewed and the
traffic is high in the range of 0 to 600.
```

```
hist(f1$LightGoodsVehicles,xlim=c(0,15000),border="brown",breaks=30,col="Orange",ylim=c(0,100))
abline(v=mean(f1$LightGoodsVehicles),col="green",lwd=2)
abline(v=median(f1$LightGoodsVehicles),col="black",lwd=2)
legend(x='topright',c('mean','median'),col=c('green','black'),lwd=c(2,2))
```




```
#inference

    It helps to predict that the traffic caused by lightweightgood is high in the range of 500 to 5000.then the traffic reduced by this vehicle.
```

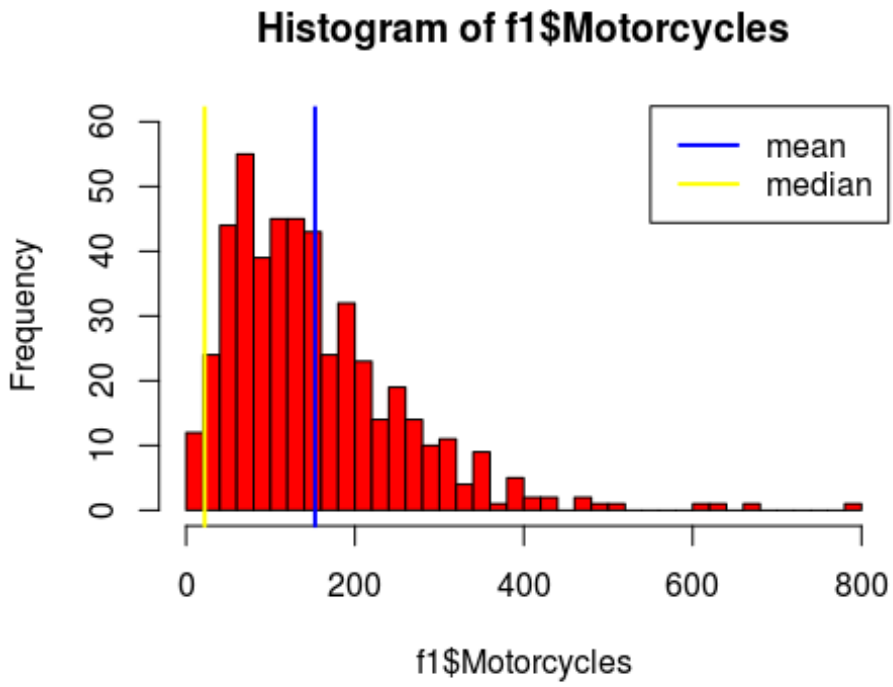
```
hist(f1$V3AxleRigidHGV,xlim=c(0,800),border="brown",breaks=20,col="Orange",yl
im=c(0,250))
abline(v=mean(f1$V3AxleRigidHGV),col="green",lwd=2)
abline(v=median(f1$V3AxleRigidHGV),col="blue",lwd=2)
legend(x='topright',c('mean','median'),col=c('green','blue'),lwd=c(2,2))
```



#inference

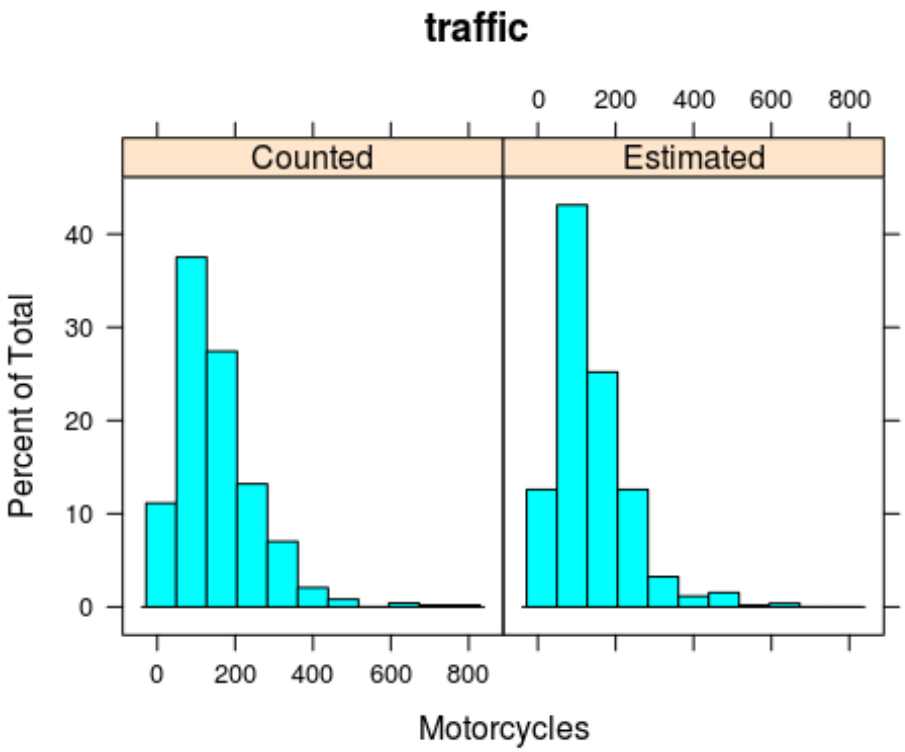
This helps to analyse that the distribution is unimodel as right skewed And the traffic is more in the range of 0 to 300.

```r
hist(f1$Motorcycles,xlim=c(0,800),breaks=30,col="red",ylim=c(0,60))
abline(v=mean(f1$Motorcycles),col="blue",lwd=2)
abline(v=median(f1$PedalCycles),col="yellow",lwd=2)
legend(x='topright',c('mean','median'),col=c('blue','yellow'),lwd=c(2,2))
```

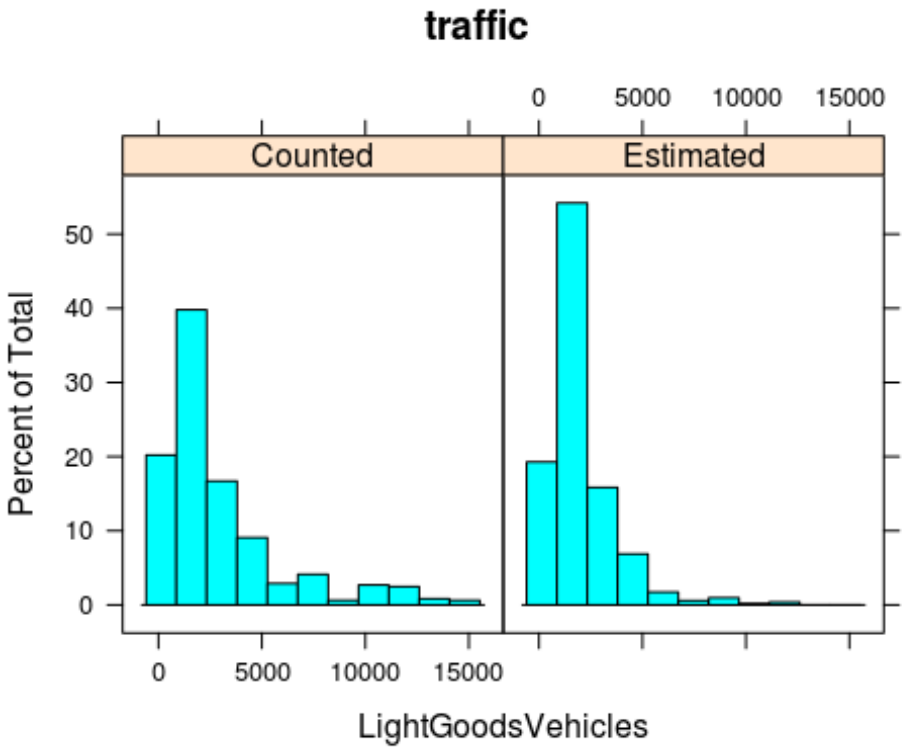



```r
#inference for overall histogram

    It helps me to predict that the traffic is more caused by the carstaxis
and lightweightgoods vehicles  based  of maximum range of vehicles flow it is
predicted.
```

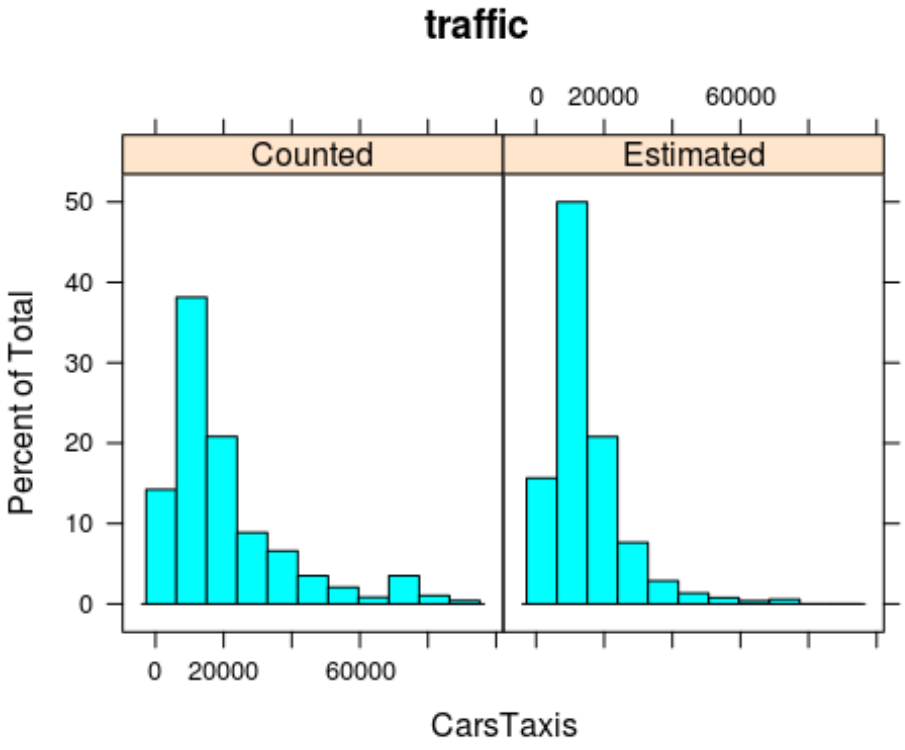
```r
attach(traffic)
histogram(~Motorcycles|Estimation_method,main="traffic")
```



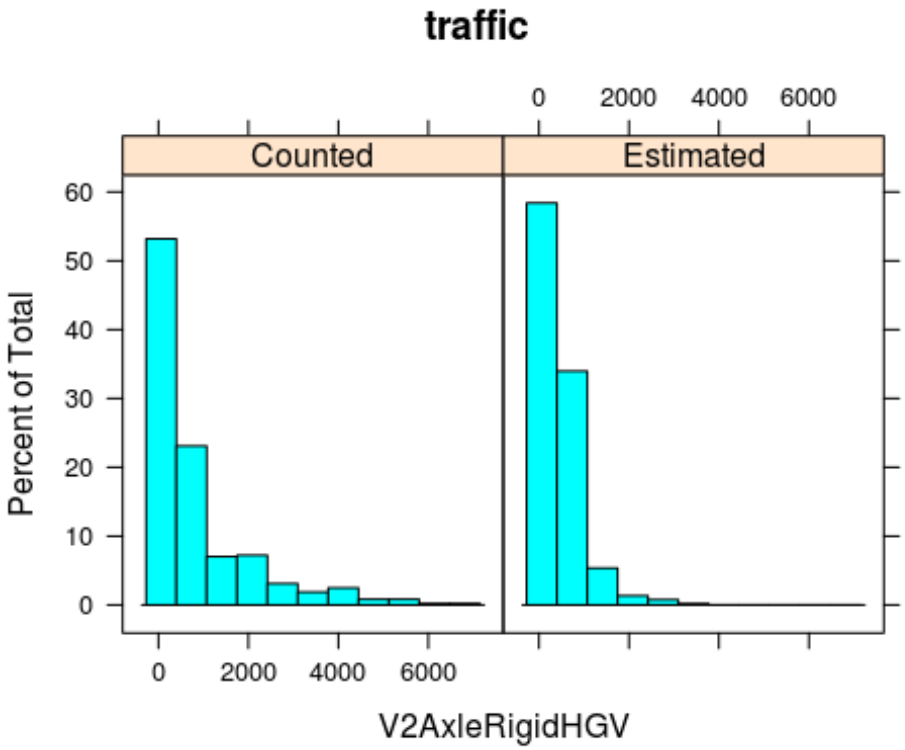

```r
histogram(~LightGoodsVehicles|Estimation_method,main="traffic")
```



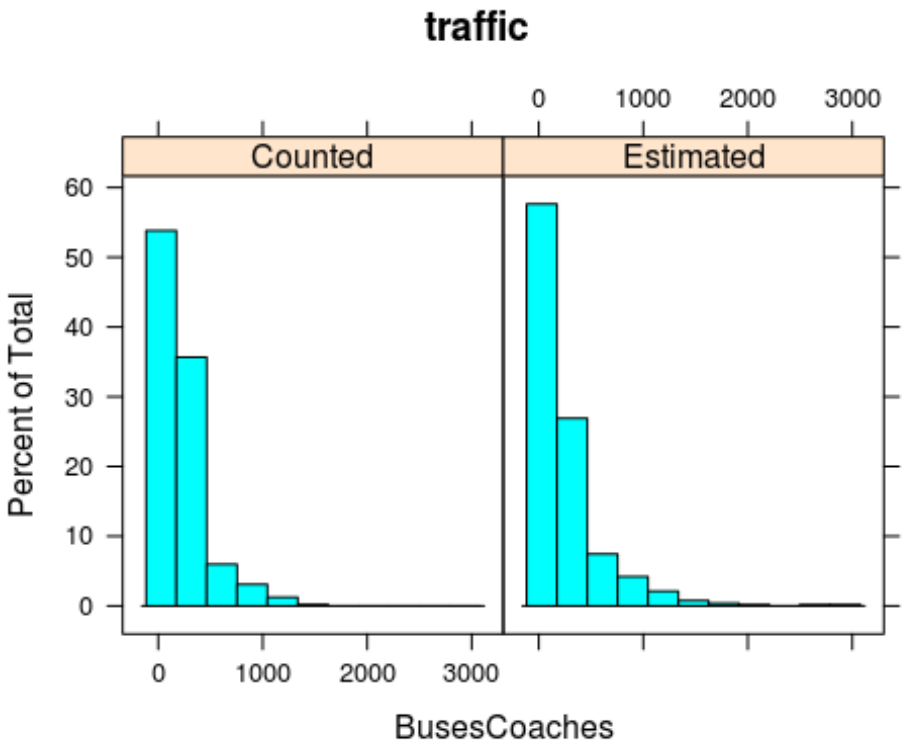

```
histogram(~CarsTaxis|Estimation_method,main="traffic")
```



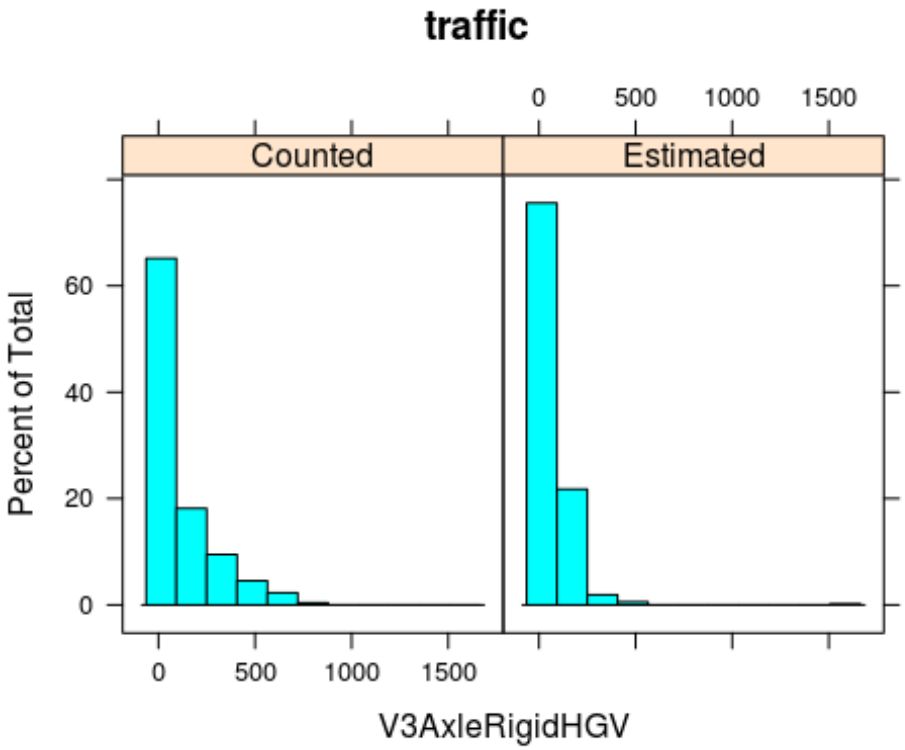

```
histogram(~V2AxleRigidHGV|Estimation_method,main="traffic")
```




```
histogram(~BusesCoaches|Estimation_method,main="traffic")
```




```
histogram(~V3AxleRigidHGV|Estimation_method,main="traffic")
```



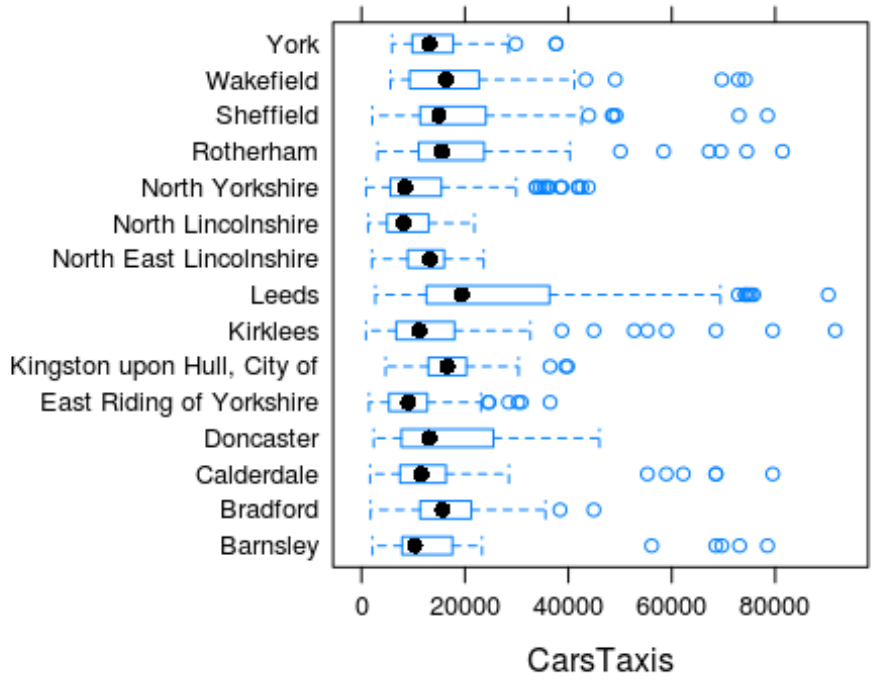

```
#inference
```

From this I have predicted that estimated method gives high range

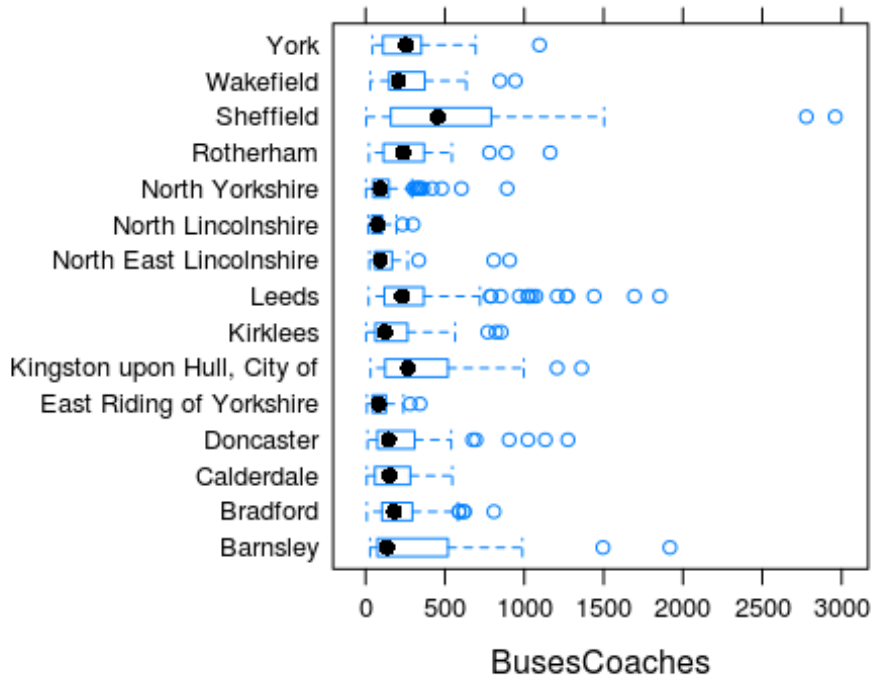
```
attach(traffic)
```

```
## The following objects are masked from traffic (pos = 3):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV
```

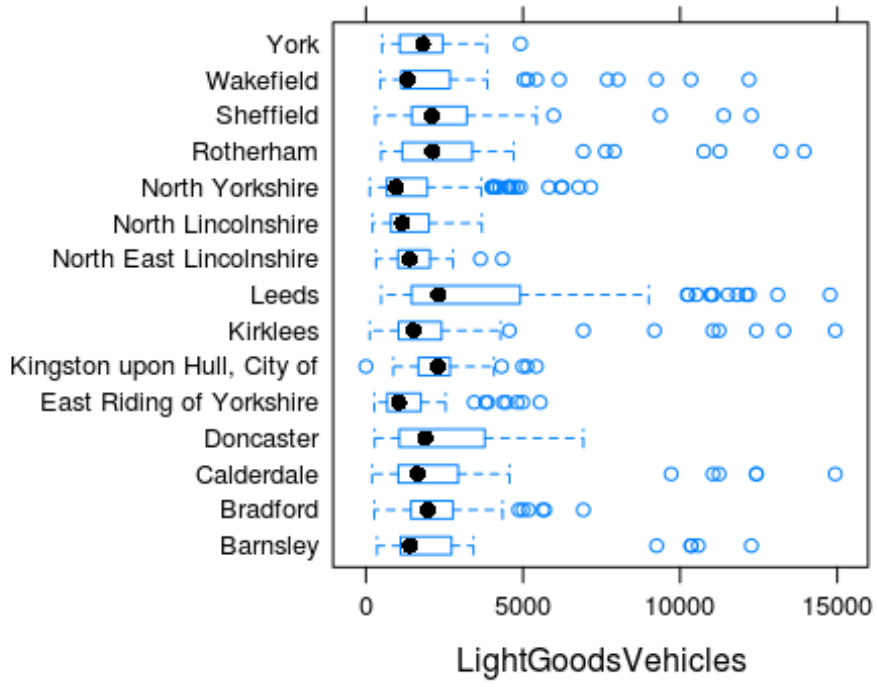
```
bwplot(LocalAuthority~CarsTaxis)
```



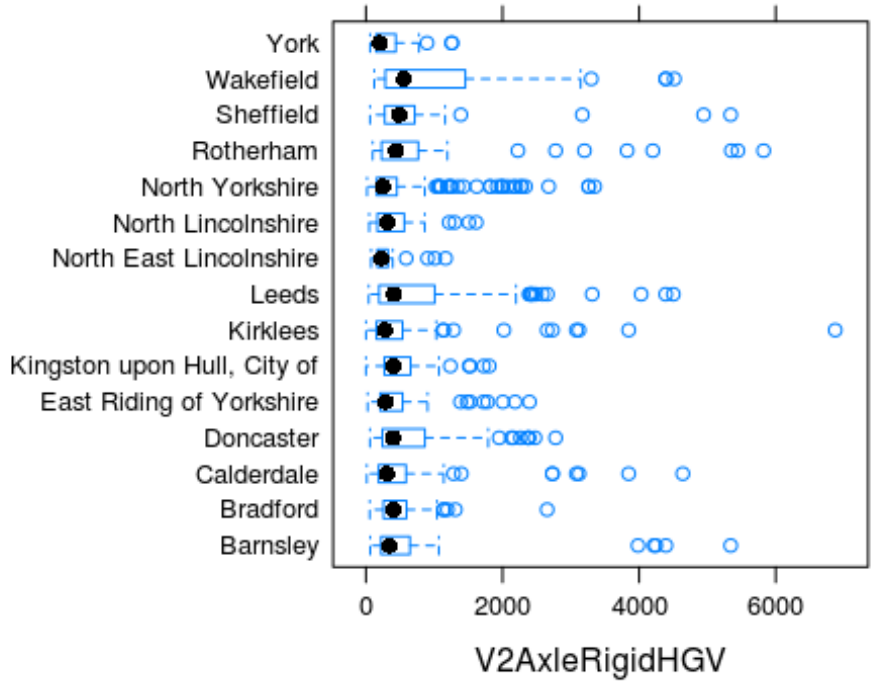

```
bwplot(LocalAuthority~BusesCoaches)
```



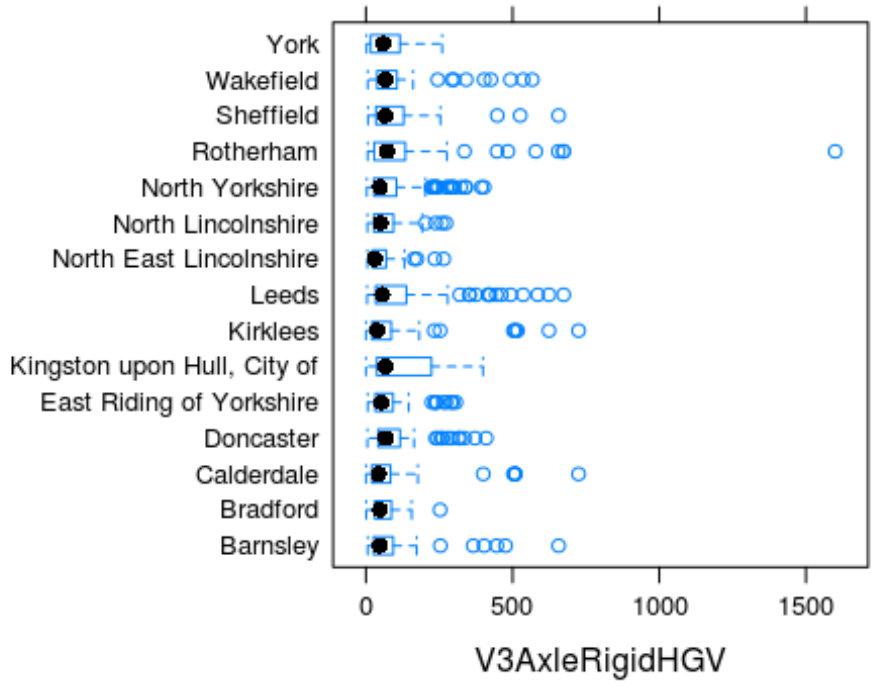
```
bwplot(LocalAuthority~LightGoodsVehicles)
```




```
bwplot(LocalAuthority~V2AxleRigidHGV)
```



```
bwplot(LocalAuthority~V3AxleRigidHGV)
```




```
#inference

        Using boxplot I predicted that the traffic flow is increased mainly
under the control of 5 local authority. It helps me to predict which
authority control the traffic is more.
```

```
# Using rbind joined the data of five local authority

c1<-filter(traffic,LocalAuthority=="Kingston upon Hull, City of")
c1
```

```
## # A tibble: 31 x 21
##    AADFYear     CP Estimation_meth… Estimation_meth… Region LocalAuthority
##       <dbl>  <dbl> <chr>            <chr>            <chr>  <chr>
##  1    2000.  6713. Estimated        Estimated using… Yorks… Kingston upon…
##  2    2000.  7482. Counted          Manual count     Yorks… Kingston upon…
##  3    2000.  8413. Estimated        Estimated using… Yorks… Kingston upon…
##  4    2000. 17889. Estimated        Estimated using… Yorks… Kingston upon…
##  5    2000. 17892. Estimated        Estimated using… Yorks… Kingston upon…
##  6    2000. 18318. Estimated        Estimated using… Yorks… Kingston upon…
##  7    2000. 18583. Counted          Manual count     Yorks… Kingston upon…
##  8    2000. 26732. Counted          Manual count     Yorks… Kingston upon…
##  9    2000. 27510. Estimated        Estimated using… Yorks… Kingston upon…
## 10    2000. 27932. Counted          Manual count     Yorks… Kingston upon…
## # ... with 21 more rows, and 15 more variables: Road <chr>,
## #   RoadCategory <chr>, Easting <dbl>, Northing <dbl>,
## #   StartJunction <chr>, EndJunction <chr>, LinkLength_km <dbl>,
## #   LinkLength_miles <dbl>, PedalCycles <dbl>, Motorcycles <dbl>,
## #   CarsTaxis <dbl>, BusesCoaches <dbl>, LightGoodsVehicles <dbl>,
## #   V2AxleRigidHGV <dbl>, V3AxleRigidHGV <dbl>
```

```
c2<-filter(traffic,LocalAuthority=="Leeds")
c2
```

```
## # A tibble: 149 x 21
##    AADFYear    CP Estimation_method Estimation_meth… Region LocalAuthority
##       <dbl> <dbl> <chr>             <chr>            <chr>  <chr>
##  1    2000. 6009. Counted           Manual count     Yorks… Leeds
##  2    2000. 6055. Counted           Manual count     Yorks… Leeds
##  3    2000. 6063. Counted           Manual count     Yorks… Leeds
##  4    2000. 6578. Counted           Manual count     Yorks… Leeds
##  5    2000. 6595. Estimated         Estimated using… Yorks… Leeds
##  6    2000. 6607. Estimated         Estimated using… Yorks… Leeds
##  7    2000. 6613. Counted           Manual count     Yorks… Leeds
##  8    2000. 7398. Estimated         Estimated using… Yorks… Leeds
##  9    2000. 7402. Counted           Manual count     Yorks… Leeds
## 10    2000. 7403. Counted           Manual count     Yorks… Leeds
## # ... with 139 more rows, and 15 more variables: Road <chr>,
## #   RoadCategory <chr>, Easting <dbl>, Northing <dbl>,
## #   StartJunction <chr>, EndJunction <chr>, LinkLength_km <dbl>,
## #   LinkLength_miles <dbl>, PedalCycles <dbl>, Motorcycles <dbl>,
## #   CarsTaxis <dbl>, BusesCoaches <dbl>, LightGoodsVehicles <dbl>,
## #   V2AxleRigidHGV <dbl>, V3AxleRigidHGV <dbl>
```

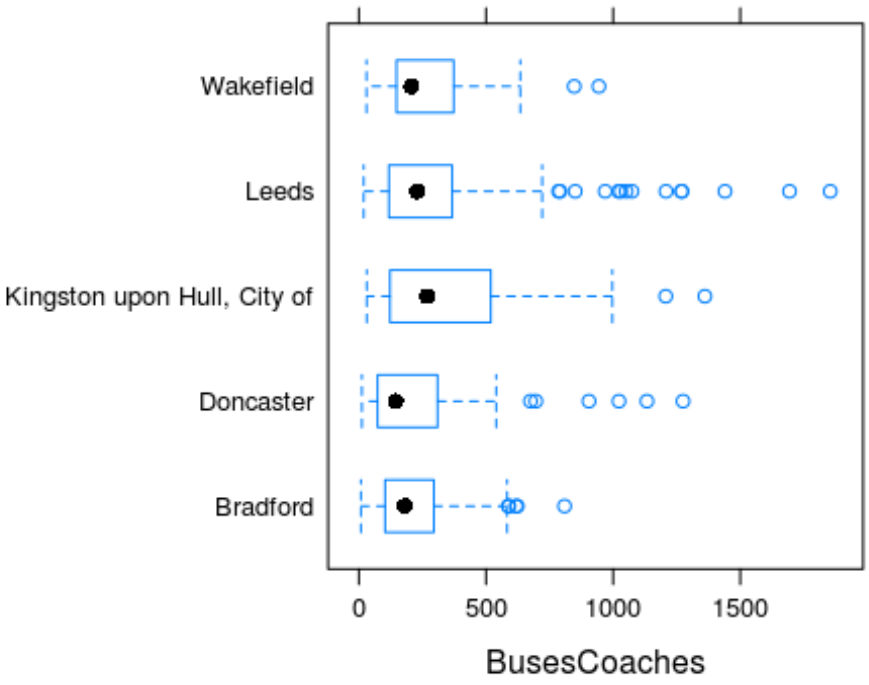
```
c3<-filter(traffic,LocalAuthority=="Wakefield")
c3
```

```
## # A tibble: 54 x 21
##    AADFYear     CP Estimation_meth… Estimation_meth… Region LocalAuthority
##       <dbl>  <dbl> <chr>            <chr>            <chr>  <chr>
##  1    2000.  6056. Counted          Manual count     Yorks… Wakefield
##  2    2000.  7375. Estimated        Estimated using… Yorks… Wakefield
##  3    2000.  7396. Counted          Manual count     Yorks… Wakefield
##  4    2000.  7399. Estimated        Estimated using… Yorks… Wakefield
##  5    2000.  7407. Counted          Manual count     Yorks… Wakefield
##  6    2000.  7419. Estimated        Estimated using… Yorks… Wakefield
##  7    2000. 16008. Counted          Manual count     Yorks… Wakefield
##  8    2000. 16578. Estimated        Estimated using… Yorks… Wakefield
##  9    2000. 17345. Estimated        Estimated using… Yorks… Wakefield
## 10    2000. 17347. Estimated        Estimated using… Yorks… Wakefield
## # ... with 44 more rows, and 15 more variables: Road <chr>,
## #   RoadCategory <chr>, Easting <dbl>, Northing <dbl>,
## #   StartJunction <chr>, EndJunction <chr>, LinkLength_km <dbl>,
## #   LinkLength_miles <dbl>, PedalCycles <dbl>, Motorcycles <dbl>,
## #   CarsTaxis <dbl>, BusesCoaches <dbl>, LightGoodsVehicles <dbl>,
## #   V2AxleRigidHGV <dbl>, V3AxleRigidHGV <dbl>

c4<-filter(traffic,LocalAuthority=="Doncaster")
c4

## # A tibble: 60 x 21
##    AADFYear    CP Estimation_method Estimation_meth… Region LocalAuthority
##       <dbl> <dbl> <chr>             <chr>            <chr>  <chr>
##  1    2000. 6035. Counted           Manual count     Yorks… Doncaster
##  2    2000. 6066. Counted           Manual count     Yorks… Doncaster
##  3    2000. 6587. Counted           Manual count     Yorks… Doncaster
##  4    2000. 7347. Estimated         Estimated using… Yorks… Doncaster
##  5    2000. 7397. Estimated         Estimated using… Yorks… Doncaster
##  6    2000. 7811. Estimated         Estimated using… Yorks… Doncaster
##  7    2000. 7815. Estimated         Estimated using… Yorks… Doncaster
##  8    2000. 7974. Counted           Manual count     Yorks… Doncaster
##  9    2000. 8156. Estimated         Estimated using… Yorks… Doncaster
## 10    2000. 8443. Counted           Manual count     Yorks… Doncaster
## # ... with 50 more rows, and 15 more variables: Road <chr>,
## #   RoadCategory <chr>, Easting <dbl>, Northing <dbl>,
## #   StartJunction <chr>, EndJunction <chr>, LinkLength_km <dbl>,
## #   LinkLength_miles <dbl>, PedalCycles <dbl>, Motorcycles <dbl>,
## #   CarsTaxis <dbl>, BusesCoaches <dbl>, LightGoodsVehicles <dbl>,
## #   V2AxleRigidHGV <dbl>, V3AxleRigidHGV <dbl>

c5<-filter(traffic,LocalAuthority=="Bradford")
c5

## # A tibble: 89 x 21
##    AADFYear    CP Estimation_method Estimation_meth… Region LocalAuthority
##       <dbl> <dbl> <chr>             <chr>            <chr>  <chr>
##  1    2000. 7377. Estimated         Estimated using… Yorks… Bradford
##  2    2000. 7401. Estimated         Estimated using… Yorks… Bradford
```

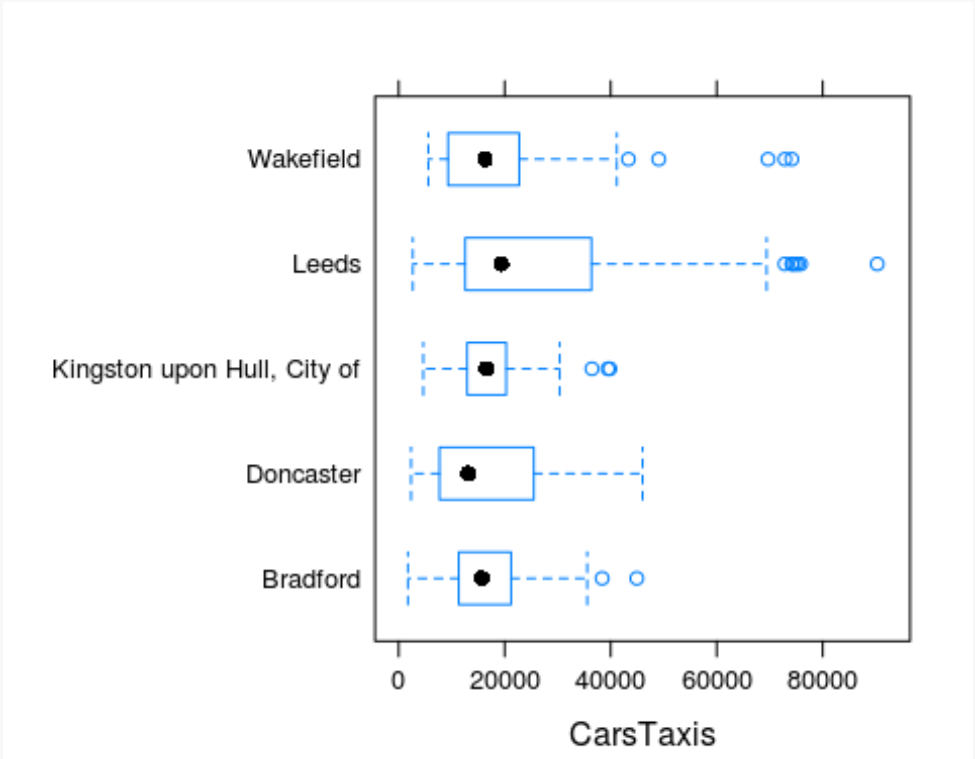
```
##  3    2000. 7405. Estimated         Estimated using… Yorks… Bradford
##  4    2000. 7409. Estimated         Estimated using… Yorks… Bradford
##  5    2000. 7413. Estimated         Estimated using… Yorks… Bradford
##  6    2000. 7421. Estimated         Estimated using… Yorks… Bradford
##  7    2000. 7734. Estimated         Estimated using… Yorks… Bradford
##  8    2000. 7853. Estimated         Estimated using… Yorks… Bradford
##  9    2000. 8524. Estimated         Estimated using… Yorks… Bradford
## 10    2000. 8579. Counted           Manual count     Yorks… Bradford
## # ... with 79 more rows, and 15 more variables: Road <chr>,
## #   RoadCategory <chr>, Easting <dbl>, Northing <dbl>,
## #   StartJunction <chr>, EndJunction <chr>, LinkLength_km <dbl>,
## #   LinkLength_miles <dbl>, PedalCycles <dbl>, Motorcycles <dbl>,
## #   CarsTaxis <dbl>, BusesCoaches <dbl>, LightGoodsVehicles <dbl>,
## #   V2AxleRigidHGV <dbl>, V3AxleRigidHGV <dbl>

cc<-rbind(c1,c2,c3,c4,c5)
cc

## # A tibble: 383 x 21
##    AADFYear     CP Estimation_meth… Estimation_meth… Region LocalAuthority
##       <dbl>  <dbl> <chr>            <chr>            <chr>  <chr>
##  1    2000.  6713. Estimated        Estimated using… Yorks… Kingston upon…
##  2    2000.  7482. Counted          Manual count     Yorks… Kingston upon…
##  3    2000.  8413. Estimated        Estimated using… Yorks… Kingston upon…
##  4    2000. 17889. Estimated        Estimated using… Yorks… Kingston upon…
##  5    2000. 17892. Estimated        Estimated using… Yorks… Kingston upon…
##  6    2000. 18318. Estimated        Estimated using… Yorks… Kingston upon…
##  7    2000. 18583. Counted          Manual count     Yorks… Kingston upon…
##  8    2000. 26732. Counted          Manual count     Yorks… Kingston upon…
##  9    2000. 27510. Estimated        Estimated using… Yorks… Kingston upon…
## 10    2000. 27932. Counted          Manual count     Yorks… Kingston upon…
## # ... with 373 more rows, and 15 more variables: Road <chr>,
## #   RoadCategory <chr>, Easting <dbl>, Northing <dbl>,
## #   StartJunction <chr>, EndJunction <chr>, LinkLength_km <dbl>,
## #   LinkLength_miles <dbl>, PedalCycles <dbl>, Motorcycles <dbl>,
## #   CarsTaxis <dbl>, BusesCoaches <dbl>, LightGoodsVehicles <dbl>,
## #   V2AxleRigidHGV <dbl>, V3AxleRigidHGV <dbl>

attach(cc)

## The following objects are masked from traffic (pos = 3):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from traffic (pos = 4):
##
```

```
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV
```

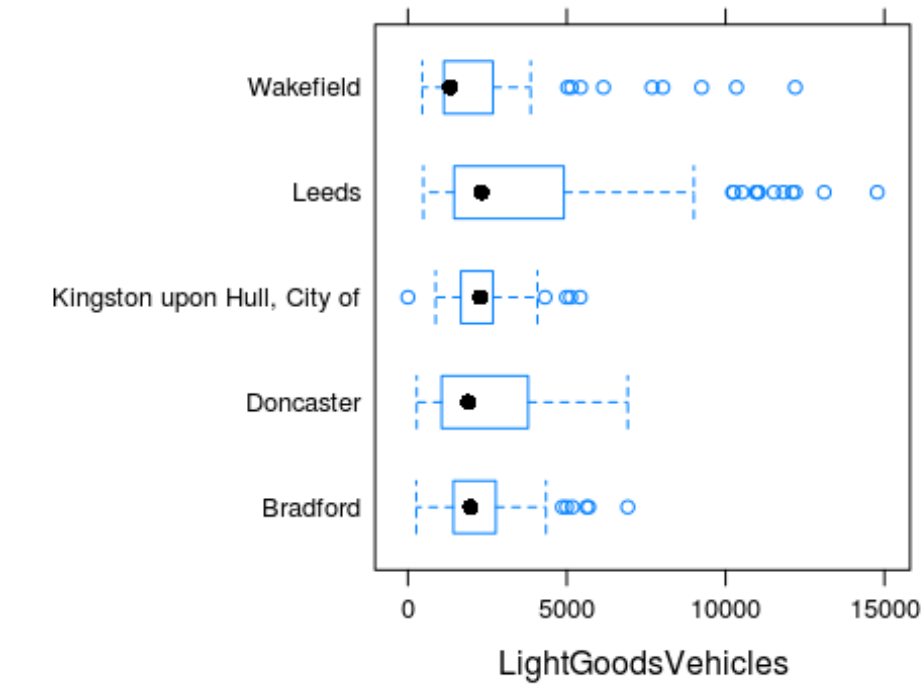
```
bwplot(LocalAuthority~BusesCoaches)
```



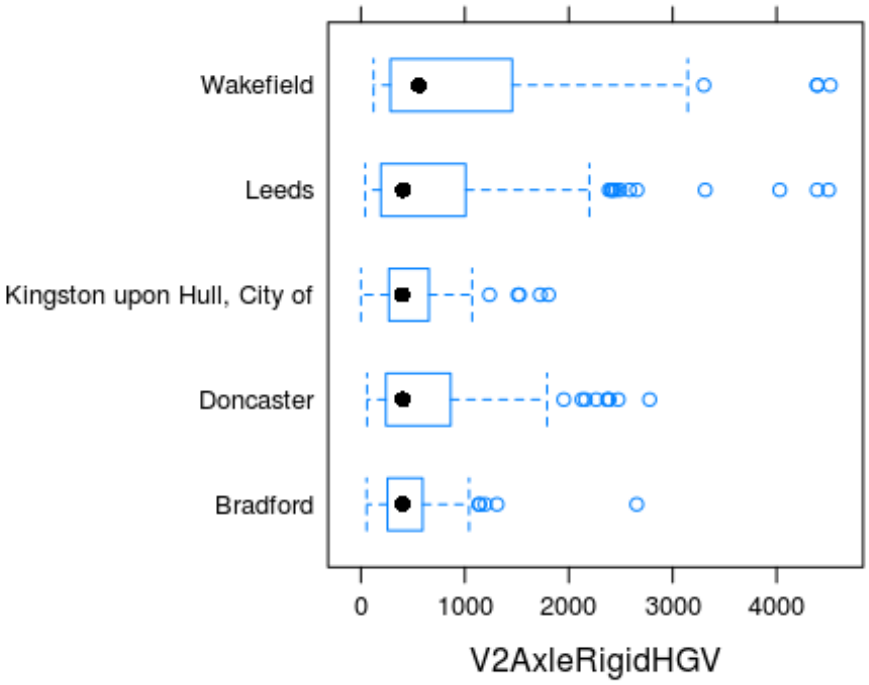
```
bwplot(LocalAuthority~CarsTaxis)
```



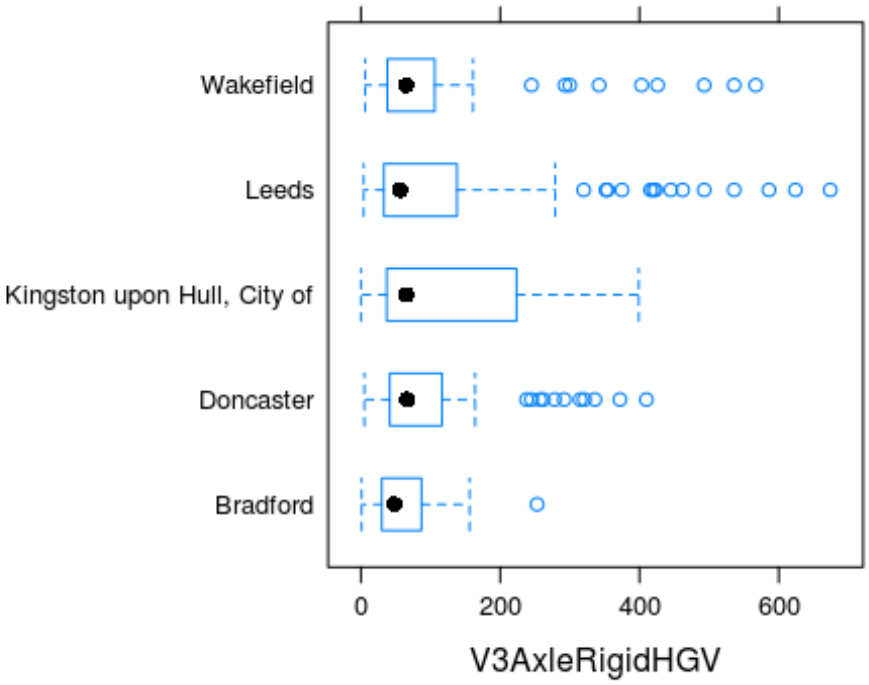

```
bwplot(LocalAuthority~LightGoodsVehicles)
```



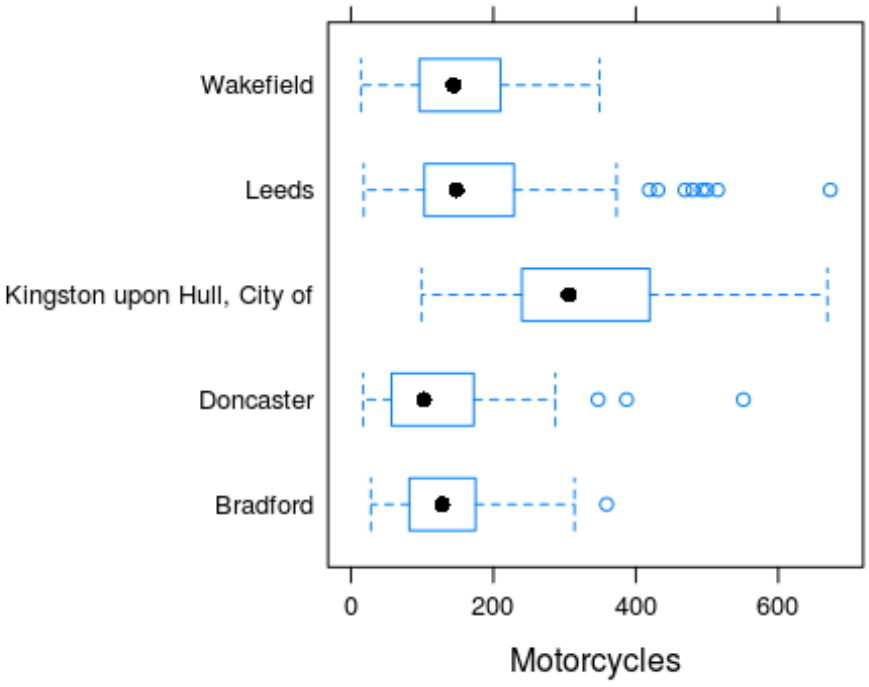
```
bwplot(LocalAuthority~V2AxleRigidHGV)
```



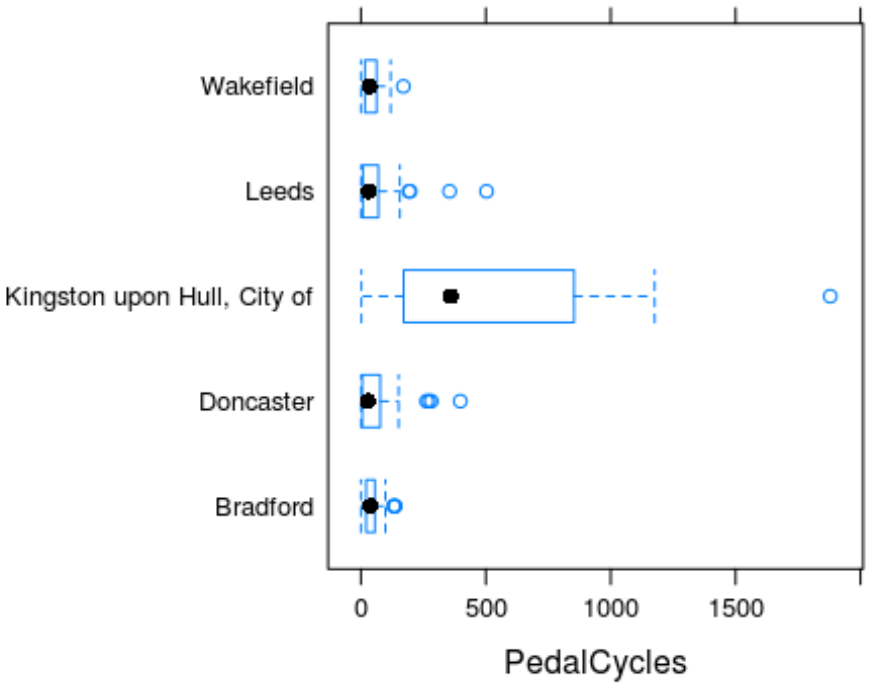
```
bwplot(LocalAuthority~V3AxleRigidHGV)
```



```
bwplot(LocalAuthority~Motorcycles)
```




```
bwplot(LocalAuthority~PedalCycles)
```



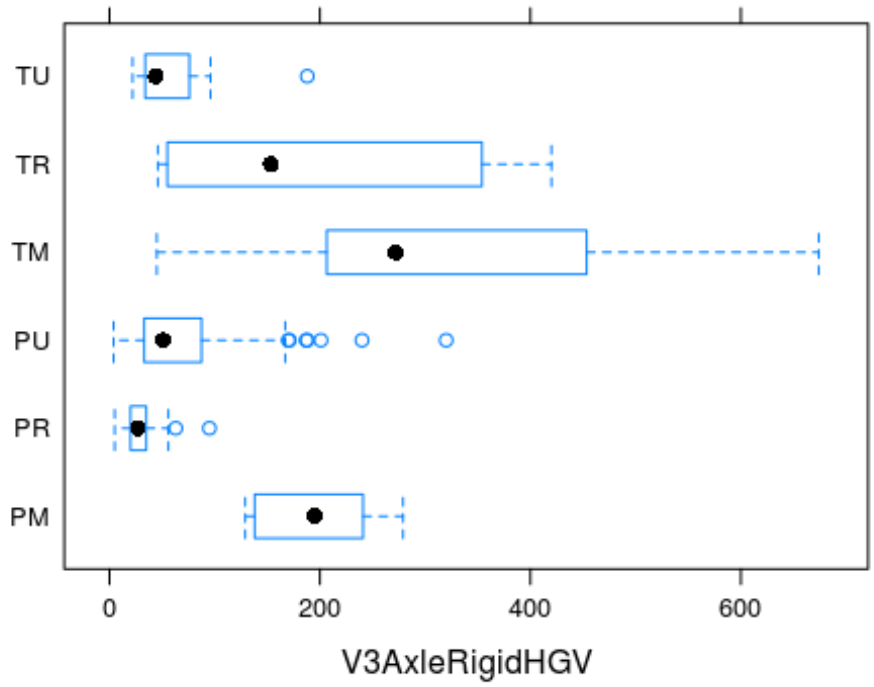
#infernce

Analysed the vehicles traffic under the 5 authority,predicted that the traffic is heavy under the control of Leeds and Kingston.But previously predicted report shows that the most of the traffic is caused by cars and lightweight goods vehicles and it is high under Leeds.I concluded that the traffic is high under Leeds

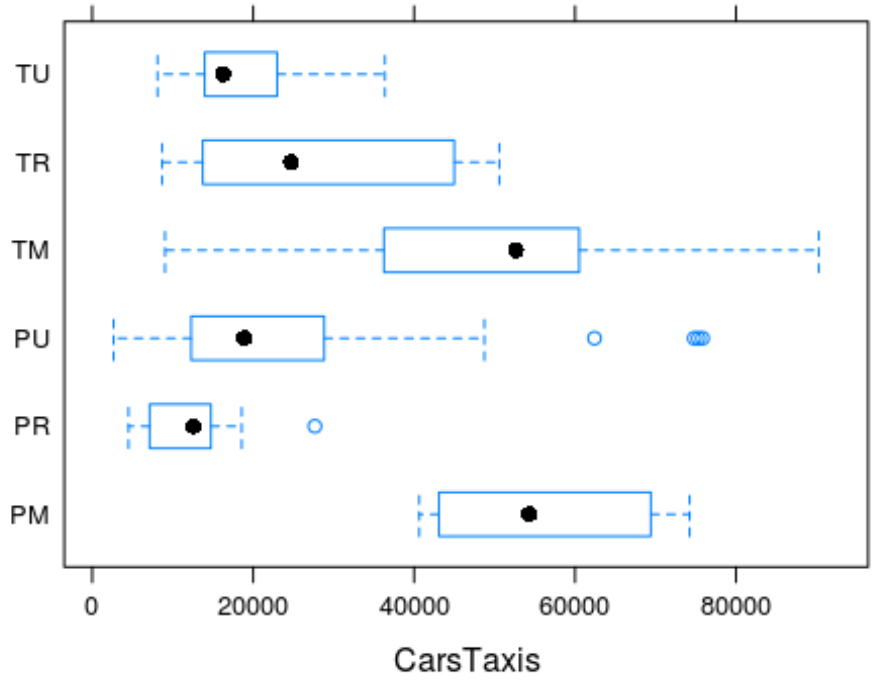
```
attach(c2)
```

```
## The following objects are masked from cc:
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV
##
## The following objects are masked from traffic (pos = 4):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from traffic (pos = 5):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV
```

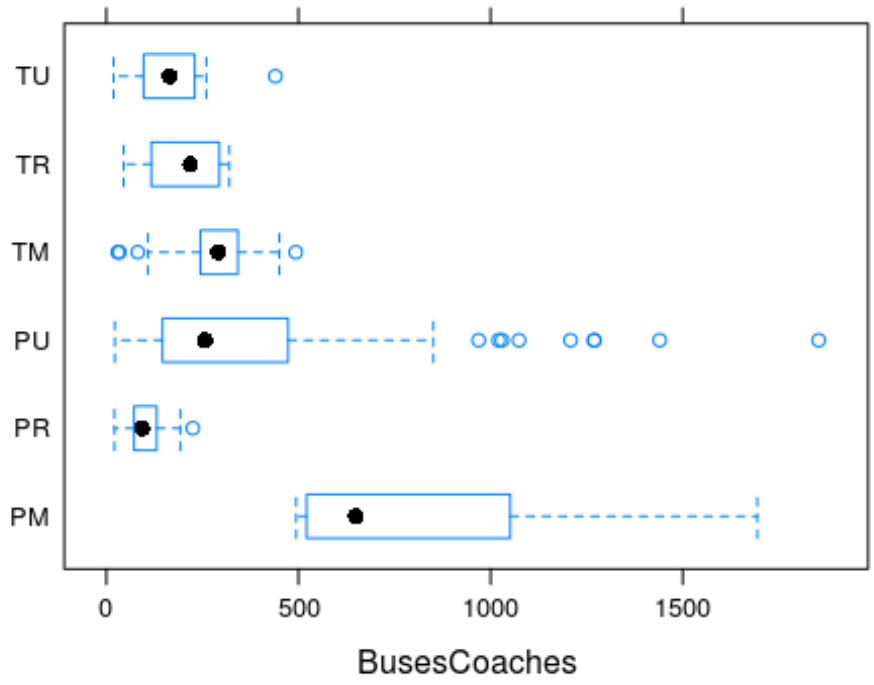
```
bwplot(RoadCategory~V3AxleRigidHGV)
```



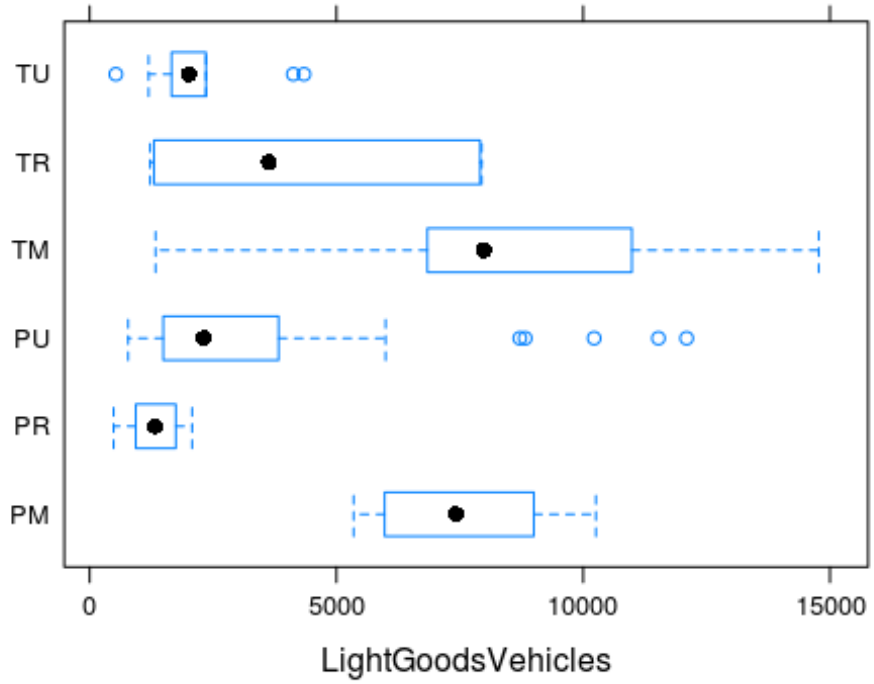

```
bwplot(RoadCategory~CarsTaxis)
```



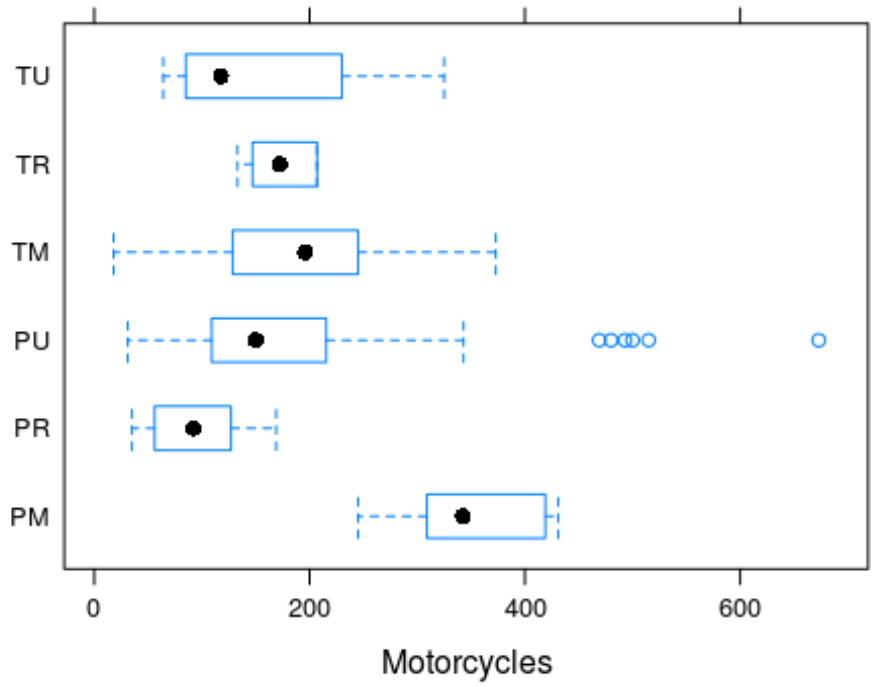

```
bwplot(RoadCategory~BusesCoaches)
```




```
bwplot(RoadCategory~LightGoodsVehicles)
```

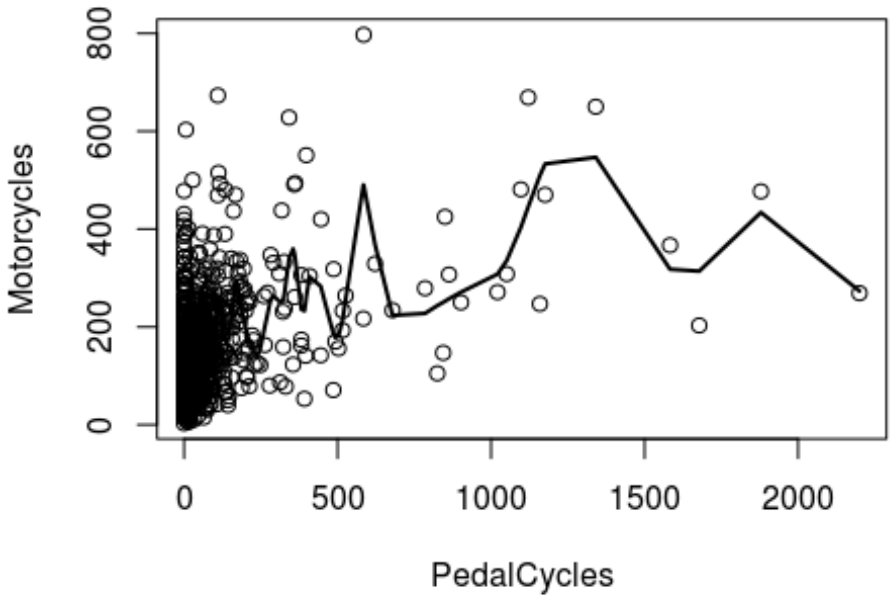



```
bwplot(RoadCategory~Motorcycles)
```




```
attach(traffic)
```

```
## The following objects are masked from c2:
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from cc:
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from traffic (pos = 5):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from traffic (pos = 6):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

plot(PedalCycles,Motorcycles)
lines(smooth.spline(PedalCycles,Motorcycles),lty=1,lwd=2)
```




```
cor(PedalCycles,Motorcycles)
## [1] 0.3939948
cor(PedalCycles,CarsTaxis)
## [1] -0.0481127
cor(PedalCycles,BusesCoaches)
## [1] 0.2661657
cor(PedalCycles,LightGoodsVehicles)
## [1] -0.08327542
cor(BusesCoaches,V2AxleRigidHGV)
## [1] 0.1652739
cor(Easting,Northing)
## [1] -0.02200243
cor(Motorcycles,CarsTaxis)
## [1] 0.5174484
cor(Motorcycles,BusesCoaches)
```

```
## [1] 0.3986205
```

```
cor(Motorcycles,LightGoodsVehicles)
```

```
## [1] 0.4426423
```

```
cor(CarsTaxis,BusesCoaches)
```

```
## [1] 0.3609451
```

```
cor(CarsTaxis,LightGoodsVehicles)
```

```
## [1] 0.9594348
```

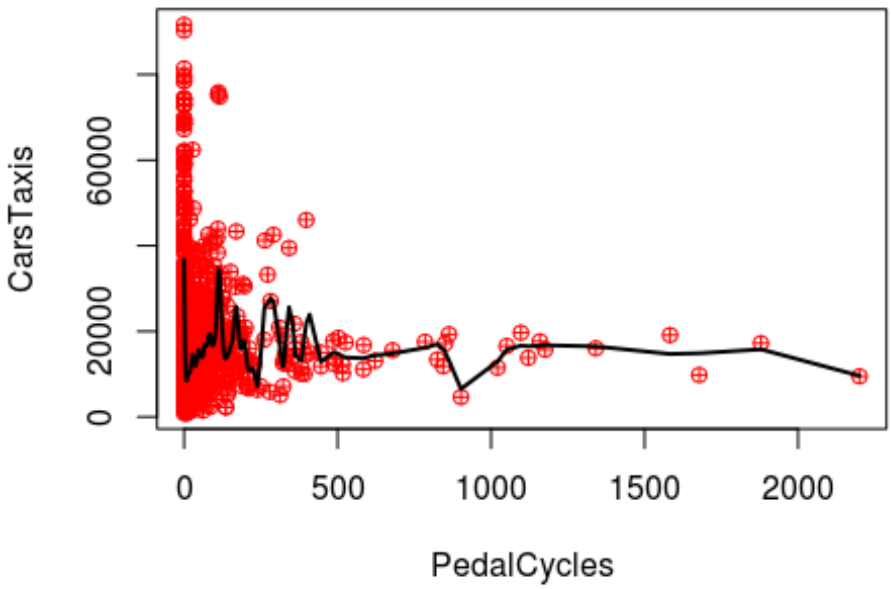
```
cor(V2AxleRigidHGV,CarsTaxis)
```

```
## [1] 0.8493717
```

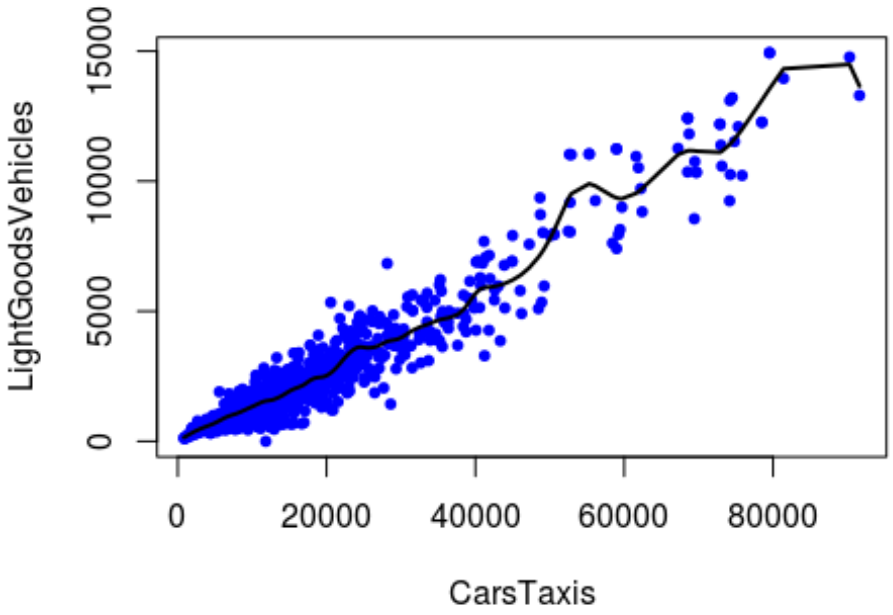
```
cor(BusesCoaches,LightGoodsVehicles)
```

```
## [1] 0.2905621
```

```r
plot(PedalCycles,CarsTaxis,pch=10,col="red")
lines(smooth.spline(PedalCycles,CarsTaxis),lty=1,lwd=2)
```



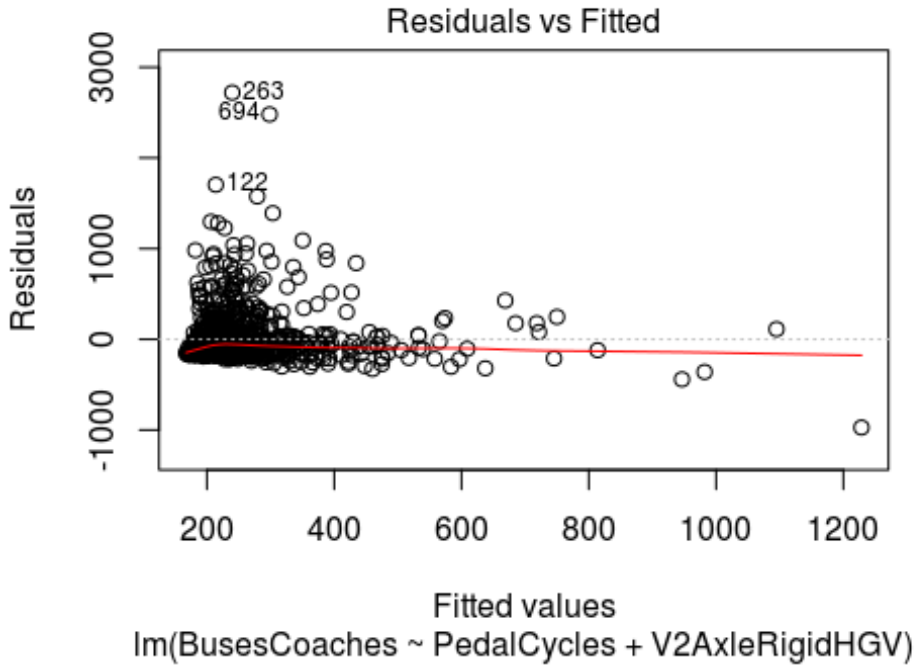
```
plot(CarsTaxis,LightGoodsVehicles,pch=20,col="blue")
lines(smooth.spline(CarsTaxis,LightGoodsVehicles),lty=1,lwd=2)
```

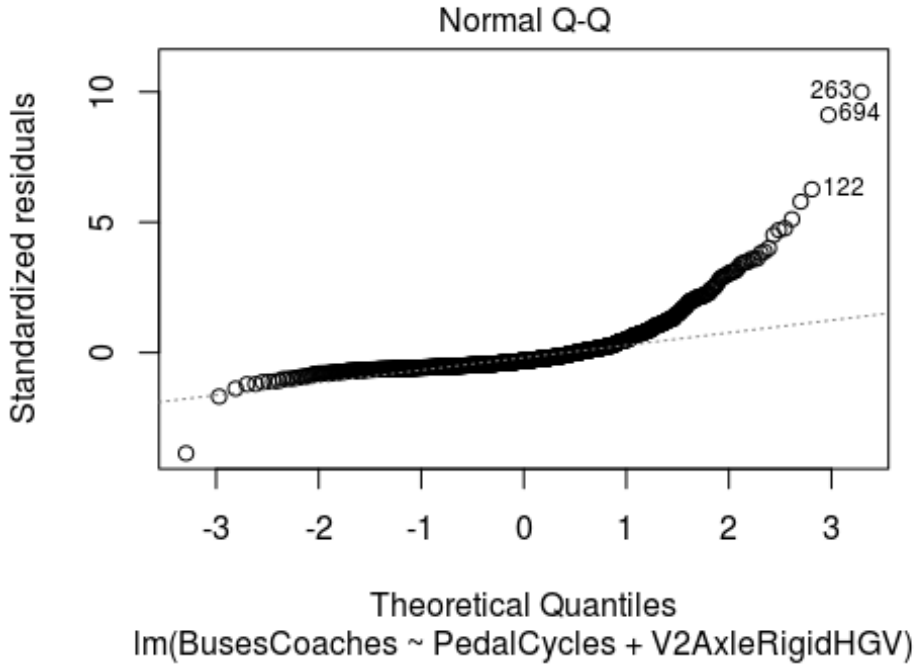


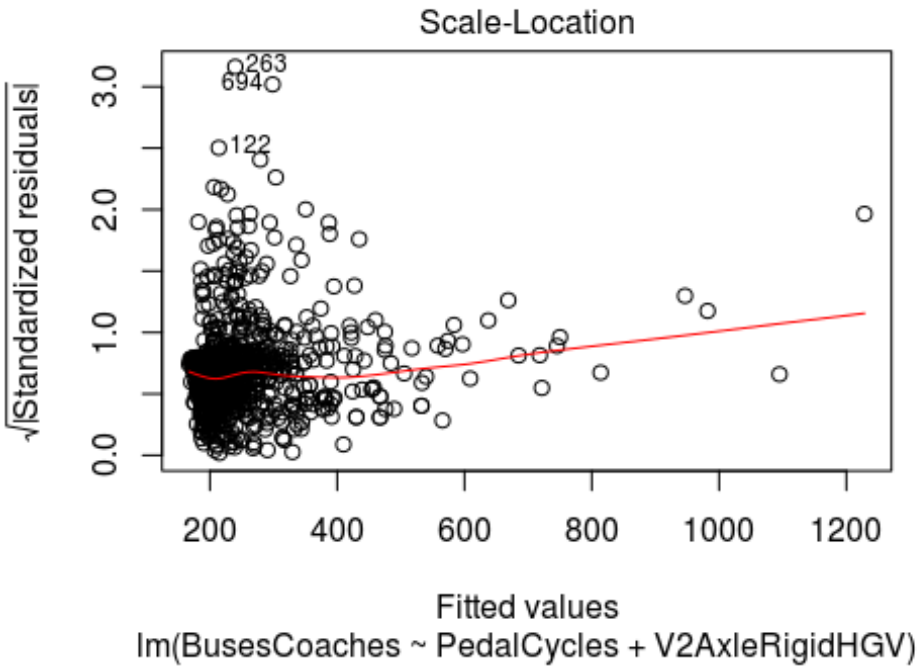

```
#inference the correlation negative for pedalcycle and carTaxis.but it is
positive between the cars and lightweightgoods


attach(traffic)

## The following objects are masked from traffic (pos = 3):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from c2:
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
```

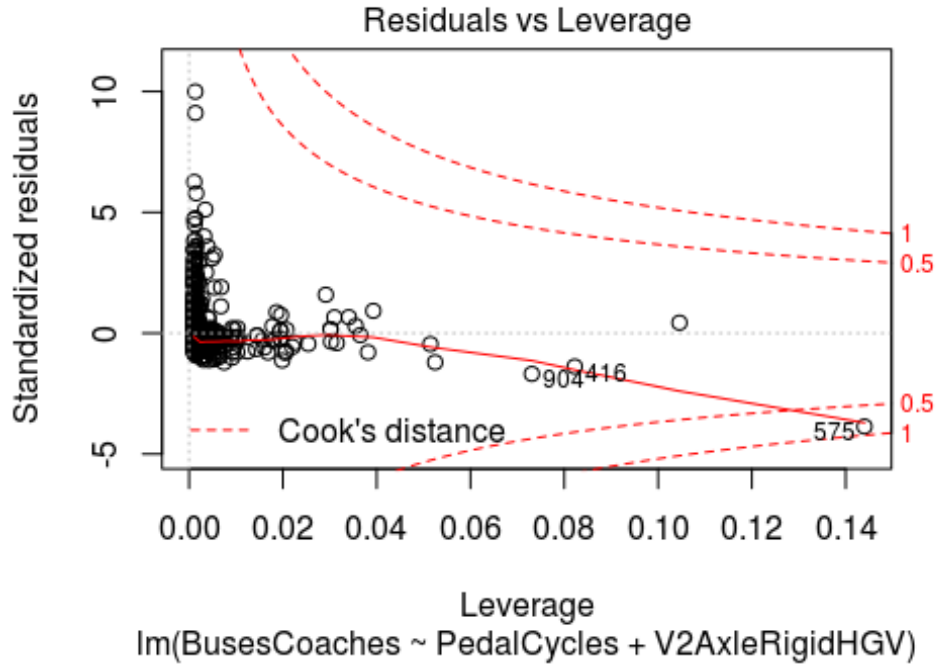
```
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from cc:
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from traffic (pos = 6):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

## The following objects are masked from traffic (pos = 7):
##
##     AADFYear, BusesCoaches, CarsTaxis, CP, Easting, EndJunction,
##     Estimation_method, Estimation_method_detailed,
##     LightGoodsVehicles, LinkLength_km, LinkLength_miles,
##     LocalAuthority, Motorcycles, Northing, PedalCycles, Region,
##     Road, RoadCategory, StartJunction, V2AxleRigidHGV,
##     V3AxleRigidHGV

l1<-lm(BusesCoaches~PedalCycles+V2AxleRigidHGV)
l1

##
## Call:
## lm(formula = BusesCoaches ~ PedalCycles + V2AxleRigidHGV)
##
## Coefficients:
##    (Intercept)     PedalCycles  V2AxleRigidHGV
##      166.16615         0.47847         0.06858
```

```
plot(l1)
```


Residuals vs Fitted
Residuals
Fitted values
lm(BusesCoaches ~ PedalCycles + V2AxleRigidHGV)
263
694
122
3000
1000
0
-1000
200
400
600
800
1000
1200


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(BusesCoaches ~ PedalCycles + V2AxleRigidHGV)
263
694
122
10
5
0
-3
-2
-1
0
1
2
3


Scale-Location
√|Standardized residuals|
263
694
122
0.0
1.0
2.0
3.0
200
400
600
800
1000
1200
Fitted values
lm(BusesCoaches ~ PedalCycles + V2AxleRigidHGV)


Residuals vs Leverage
Standardized residuals
10
5
0
-5
1
0.5
0.5
1
Cook's distance
904
416
575
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Leverage
lm(BusesCoaches ~ PedalCycles + V2AxleRigidHGV)

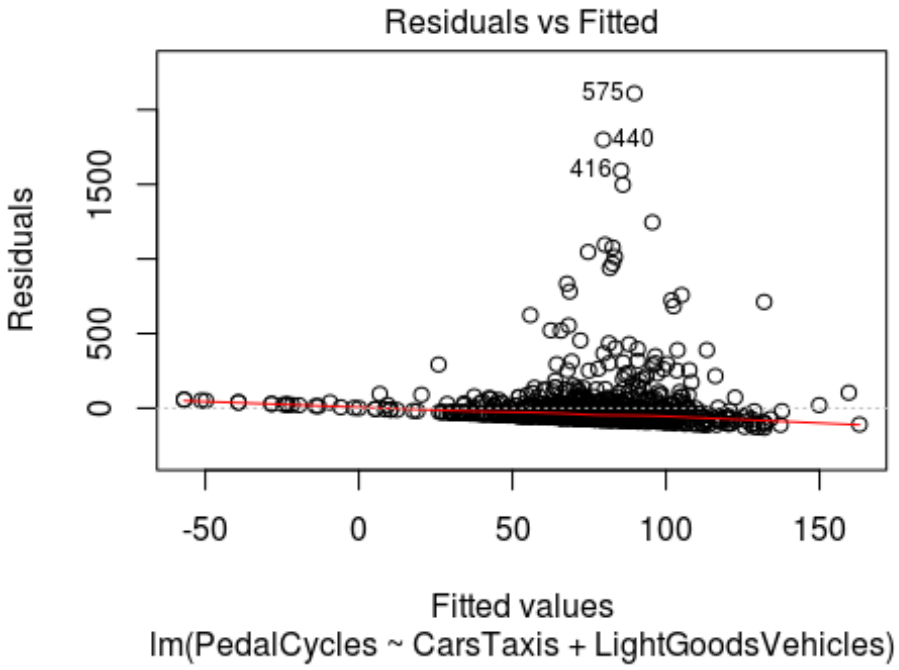
```r
summary(l1)
```

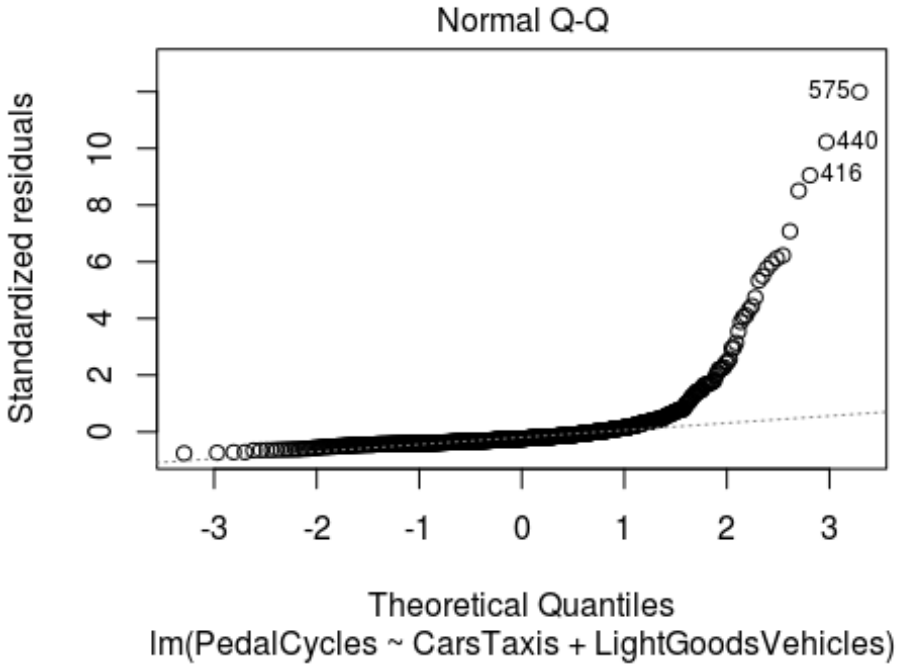
```
##
## Call:
## lm(formula = BusesCoaches ~ PedalCycles + V2AxleRigidHGV)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -972.54 -140.68  -86.81   34.98 2720.11
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.662e+02  1.167e+01  14.234  < 2e-16 ***
## PedalCycles    4.785e-01  4.874e-02   9.818  < 2e-16 ***
## V2AxleRigidHGV 6.858e-02  9.998e-03   6.859 1.21e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 272.3 on 1006 degrees of freedom
## Multiple R-squared:  0.1124, Adjusted R-squared:  0.1106
## F-statistic: 63.67 on 2 and 1006 DF,  p-value: < 2.2e-16
```

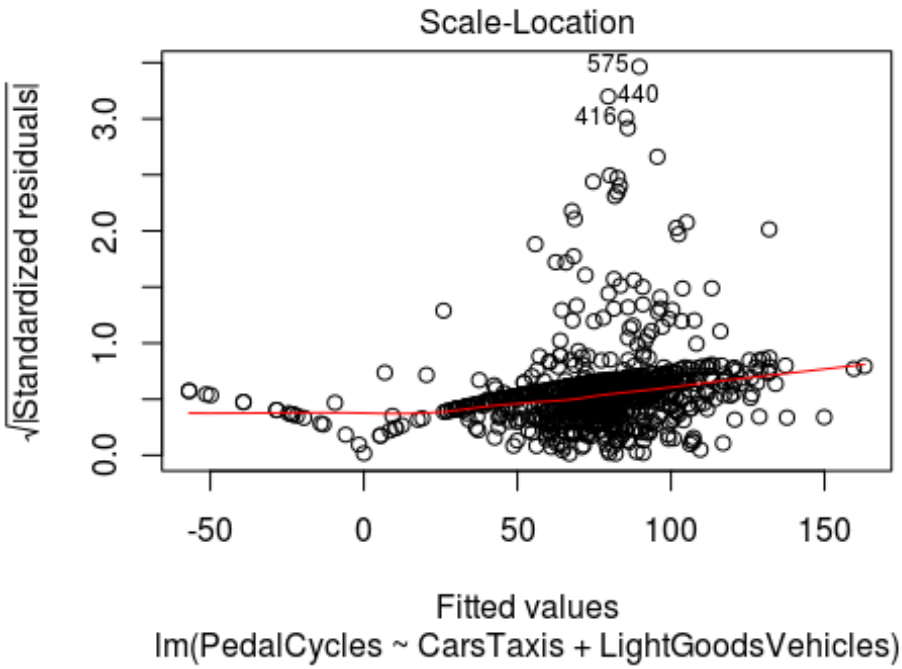
```r
l2<-lm(PedalCycles~CarsTaxis+LightGoodsVehicles)
l2
```

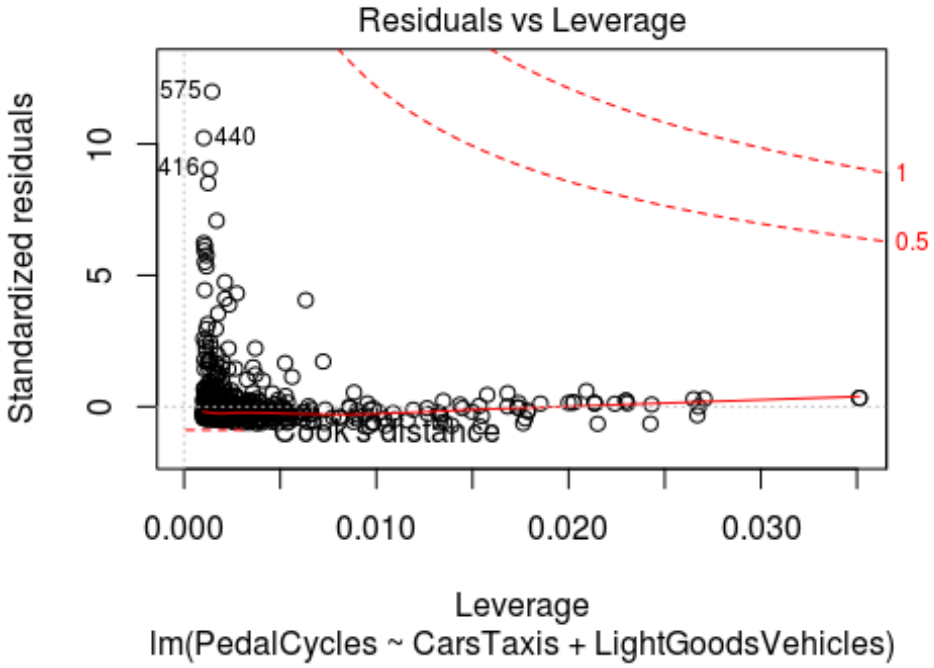
```
##
## Call:
## lm(formula = PedalCycles ~ CarsTaxis + LightGoodsVehicles)
##
## Coefficients:
##        (Intercept)           CarsTaxis  LightGoodsVehicles
##          75.479647            0.004773           -0.034277
```

```r
plot(l2)
```


Residuals vs Fitted
575
440
416
Residuals
0
500
1500
-50
0
50
100
150
Fitted values
lm(PedalCycles ~ CarsTaxis + LightGoodsVehicles)


Normal Q-Q
575
440
416
Standardized residuals
0
2
4
6
8
10
-3
-2
-1
0
1
2
3
Theoretical Quantiles
lm(PedalCycles ~ CarsTaxis + LightGoodsVehicles)


Scale-Location
575
440
416
√|Standardized residuals|
0.0
1.0
2.0
3.0
-50
0
50
100
150
Fitted values
lm(PedalCycles ~ CarsTaxis + LightGoodsVehicles)


Residuals vs Leverage
575
440
416
1
0.5
Standardized residuals
0
5
10
Cook's distance
0.000
0.010
0.020
0.030
Leverage
lm(PedalCycles ~ CarsTaxis + LightGoodsVehicles)

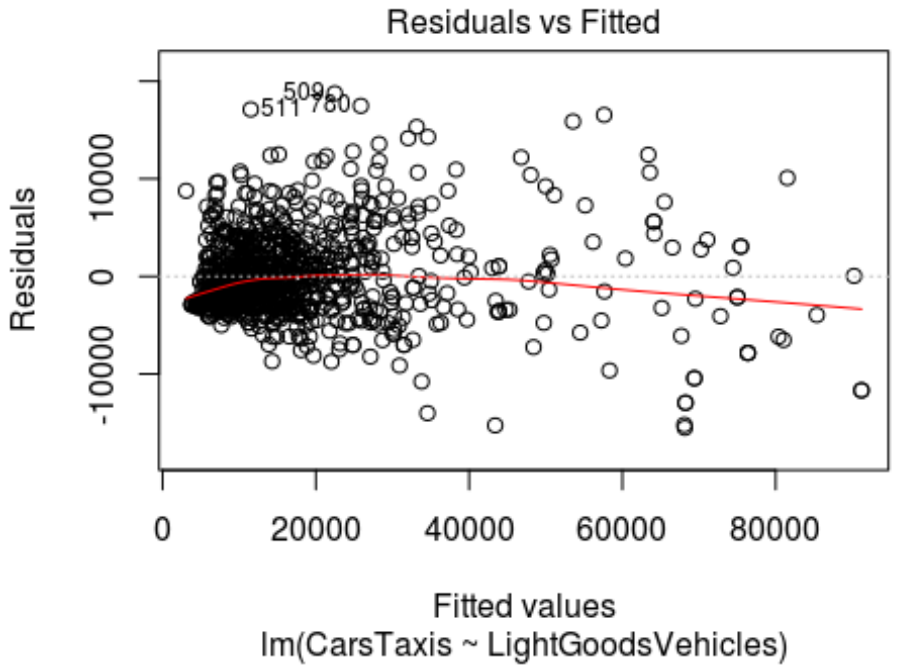
```r
summary(l2)
```

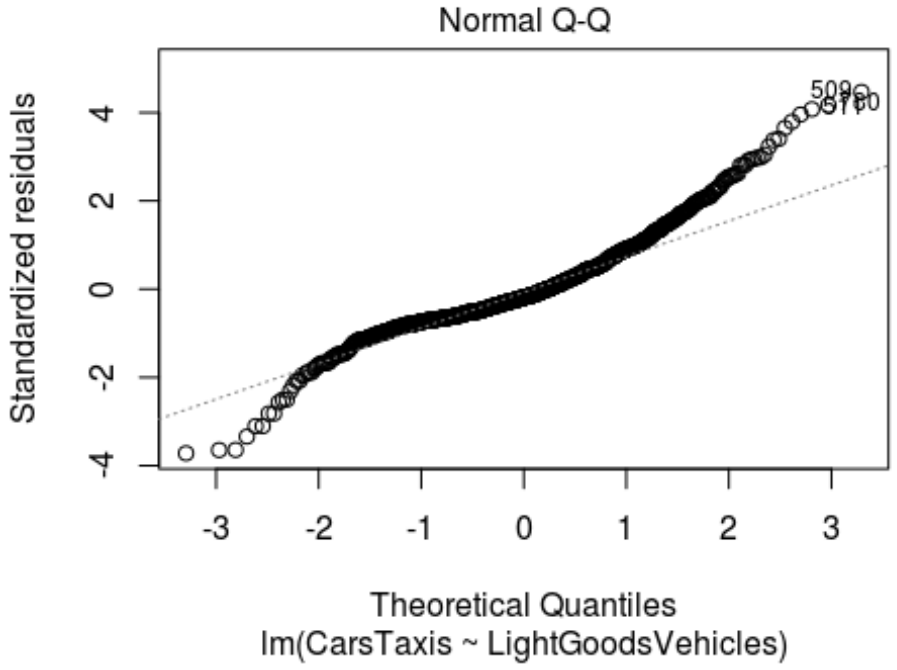
```
##
## Call:
## lm(formula = PedalCycles ~ CarsTaxis + LightGoodsVehicles)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -132.05  -63.65  -41.80   -3.91 2110.28
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        75.479647   8.854948   8.524  < 2e-16 ***
## CarsTaxis           0.004773   0.001322   3.611  0.00032 ***
## LightGoodsVehicles -0.034277   0.008128  -4.217  2.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 176.1 on 1006 degrees of freedom
## Multiple R-squared:  0.01964,    Adjusted R-squared:  0.0177
## F-statistic: 10.08 on 2 and 1006 DF,  p-value: 4.633e-05
```

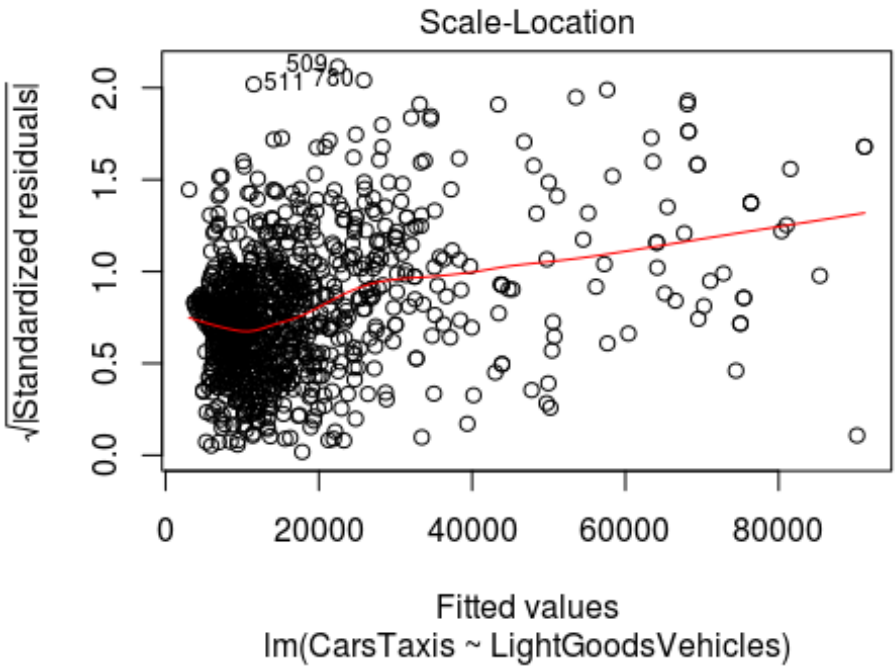
```r
l3<-lm(CarsTaxis~LightGoodsVehicles)
l3
```

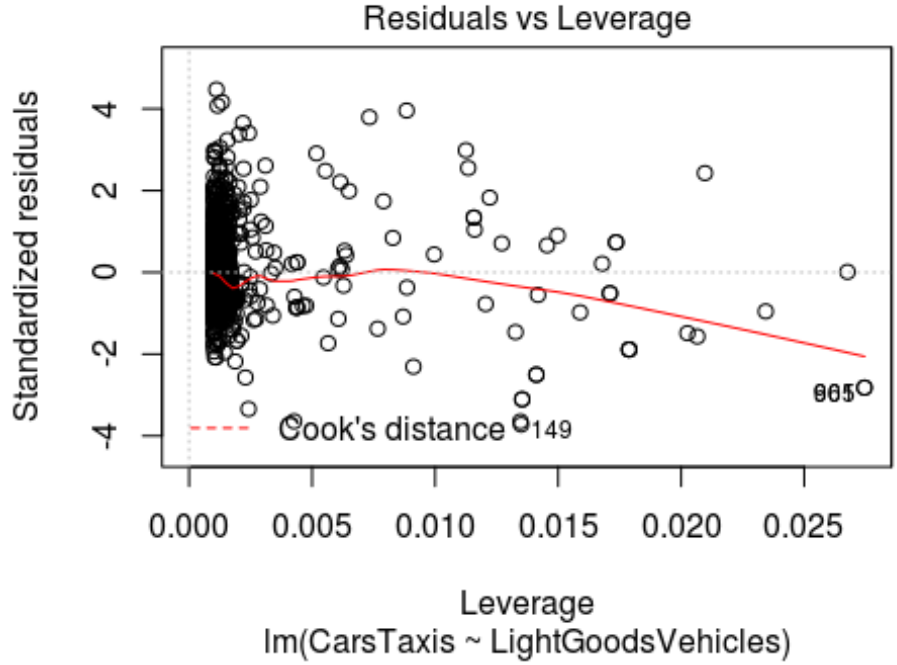
```
##
## Call:
## lm(formula = CarsTaxis ~ LightGoodsVehicles)
##
## Coefficients:
##        (Intercept)  LightGoodsVehicles
##             3069.3                 5.9
```

```
plot(13)
```


Residuals vs Fitted
Residuals
Fitted values
lm(CarsTaxis ~ LightGoodsVehicles)


Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(CarsTaxis ~ LightGoodsVehicles)


Scale-Location
509
511
780
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
0
20000
40000
60000
80000
Fitted values
lm(CarsTaxis ~ LightGoodsVehicles)


Residuals vs Leverage
Standardized residuals
-4
-2
0
2
4
Cook's distance
149
0.000
0.005
0.010
0.015
0.020
0.025
Leverage
lm(CarsTaxis ~ LightGoodsVehicles)

```r
summary(l3)
```

```
##
## Call:
## lm(formula = CarsTaxis ~ LightGoodsVehicles)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -15514.9  -2588.2   -778.7   1982.9  18750.8
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.069e+03  1.877e+02   16.36   <2e-16 ***
## LightGoodsVehicles 5.900e+00  5.464e-02  107.99   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4198 on 1007 degrees of freedom
## Multiple R-squared:  0.9205, Adjusted R-squared:  0.9204
## F-statistic: 1.166e+04 on 1 and 1007 DF,  p-value: < 2.2e-16
```

```
#inference

      From the regression model predicted that the carsTaxis and
LightweightVehicles have more correlation and form the best model.
```