motartrend_report

zengzengyang

11/27/2020

Motor Trend, a magazine about the automobile industry. Looking at a data set of a collection of cars, they are interested in exploring the relationship between a set of variables and miles per gallon (MPG) (outcome). They are particularly interested in the following two questions:

1.   Is an automatic or manual transmission better for MPG?
2.   Quantify the MPG difference between automatic and manual transmissions.

```
library(ggplot2)

# Lets begin by checking the data first.
summary(mtcars)
```

```
##       mpg             cyl             disp             hp
##  Min.   :10.40   Min.   :4.000   Min.   : 71.1   Min.   : 52.0
##  1st Qu.:15.43   1st Qu.:4.000   1st Qu.:120.8   1st Qu.: 96.5
##  Median :19.20   Median :6.000   Median :196.3   Median :123.0
##  Mean   :20.09   Mean   :6.188   Mean   :230.7   Mean   :146.7
##  3rd Qu.:22.80   3rd Qu.:8.000   3rd Qu.:326.0   3rd Qu.:180.0
##  Max.   :33.90   Max.   :8.000   Max.   :472.0   Max.   :335.0
##       drat             wt             qsec             vs
##  Min.   :2.760   Min.   :1.513   Min.   :14.50   Min.   :0.0000
##  1st Qu.:3.080   1st Qu.:2.581   1st Qu.:16.89   1st Qu.:0.0000
##  Median :3.695   Median :3.325   Median :17.71   Median :0.0000
##  Mean   :3.597   Mean   :3.217   Mean   :17.85   Mean   :0.4375
##  3rd Qu.:3.920   3rd Qu.:3.610   3rd Qu.:18.90   3rd Qu.:1.0000
##  Max.   :4.930   Max.   :5.424   Max.   :22.90   Max.   :1.0000
##       am             gear             carb
##  Min.   :0.0000   Min.   :3.000   Min.   :1.000
##  1st Qu.:0.0000   1st Qu.:3.000   1st Qu.:2.000
##  Median :0.0000   Median :4.000   Median :2.000
##  Mean   :0.4062   Mean   :3.688   Mean   :2.812
##  3rd Qu.:1.0000   3rd Qu.:4.000   3rd Qu.:4.000
##  Max.   :1.0000   Max.   :5.000   Max.   :8.000
```

```
data(mtcars)
head(mtcars)
```

```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
```

```
## Hornet Sportabout 18.7   8   360 175 3.15 3.440 17.02  0  0    3    2
## Valiant            18.1   6   225 105 2.76 3.460 20.22  1  0    3    1
```

#Comparing means of mpg for automatic and manual transmission

```r
mtcars$am <- factor(mtcars$am, labels=c("Automatic","Manual"))

mpgmean <- aggregate(mtcars$mpg, by=list(mtcars$am), FUN=mean)

colnames(mpgmean) <- c("am", "mpg")

mpgmean
```

```
##          am      mpg
## 1 Automatic 17.14737
## 2    Manual 24.39231
```

```r
mpgmean$mpg[2] - mpgmean$mpg[1]
```

```
## [1] 7.244939
```

```r
fit <- lm(mpg ~ am, mtcars)

summary(fit)
```

```
##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## amManual       7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285
```

#linear regression

```r
fit <- lm(mpg ~ am, mtcars)
summary(fit)
```

```
##
## Call:
```

```
## lm(formula = mpg ~ am, data = mtcars)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.3923 -3.0923 -0.2974  3.2439  9.5077
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   17.147      1.125  15.247 1.13e-15 ***
## amManual       7.245      1.764   4.106 0.000285 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 30 degrees of freedom
## Multiple R-squared:  0.3598, Adjusted R-squared:  0.3385
## F-statistic: 16.86 on 1 and 30 DF,  p-value: 0.000285

fit2 <- lm(mpg ~ am + hp, mtcars)
fit3 <- lm(mpg ~ am + hp + cyl, mtcars)
fit4 <- lm(mpg ~ am + hp + cyl + disp, mtcars)
fit5 <- lm(mpg ~ am + hp + cyl + disp + wt, mtcars)

anova(fit, fit2, fit3, fit4, fit5)

## Analysis of Variance Table
##
## Model 1: mpg ~ am
## Model 2: mpg ~ am + hp
## Model 3: mpg ~ am + hp + cyl
## Model 4: mpg ~ am + hp + cyl + disp
## Model 5: mpg ~ am + hp + cyl + disp + wt
##   Res.Df    RSS Df Sum of Sq       F    Pr(>F)
## 1     30 720.90
## 2     29 245.44  1    475.46 75.7841 3.499e-09 ***
## 3     28 220.55  1     24.89  3.9667  0.057011 .
## 4     27 216.37  1      4.19  0.6672  0.421464
## 5     26 163.12  1     53.25  8.4872  0.007257 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#Analyzing the p-values of our anova-analysis。 We need our final regression
model containing the transmission type, horsepower and weight.
library(ggplot2)

data(mtcars)
mtcars$am <- factor(mtcars$am)
ggplot(fit, mapping = aes(colour = am))
```
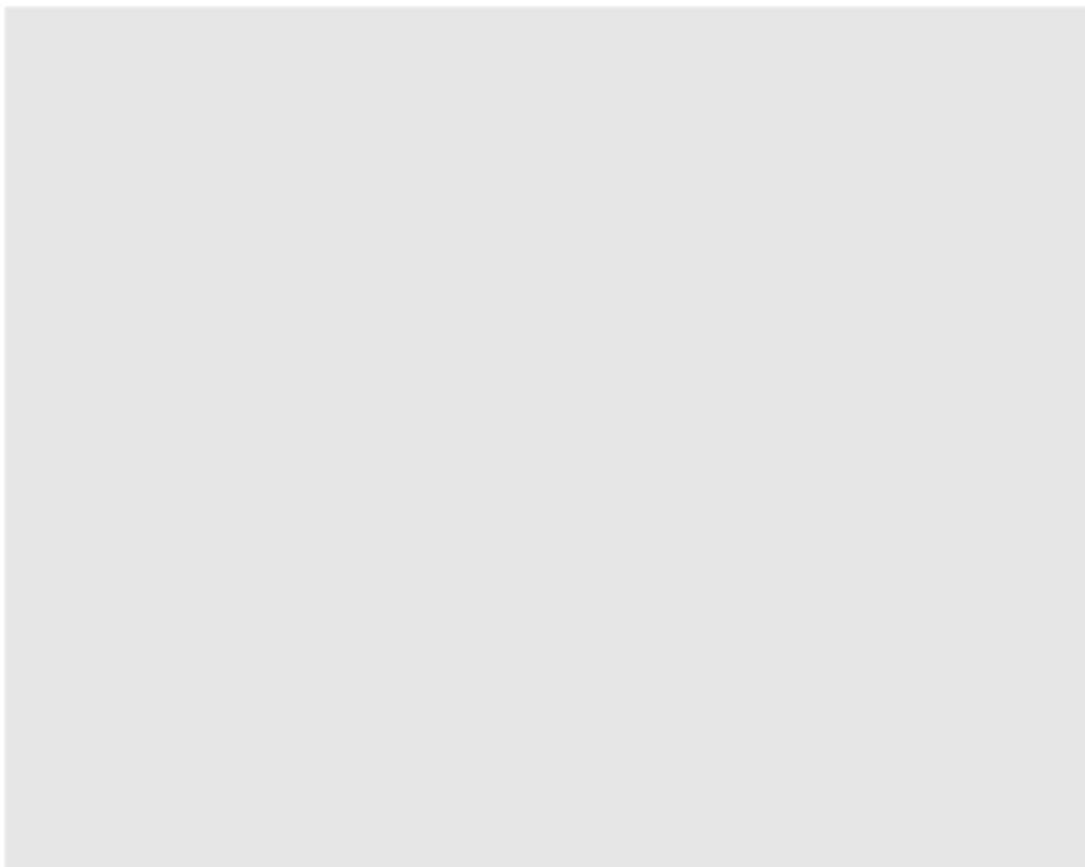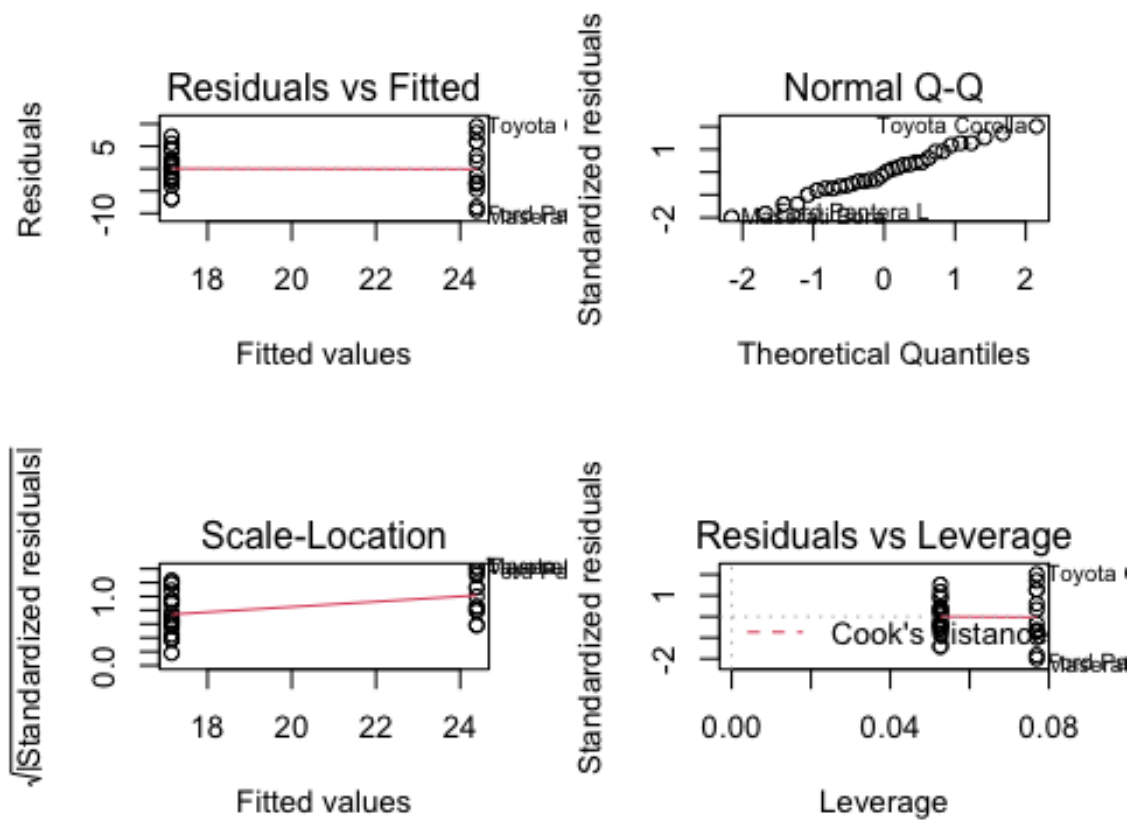
```
par(mfrow = c(2,2))
plot(fit)
```

```
fit

##
## Call:
## lm(formula = mpg ~ am, data = mtcars)
##
## Coefficients:
## (Intercept)      amManual
##      17.147         7.245

plot(fit2)
```
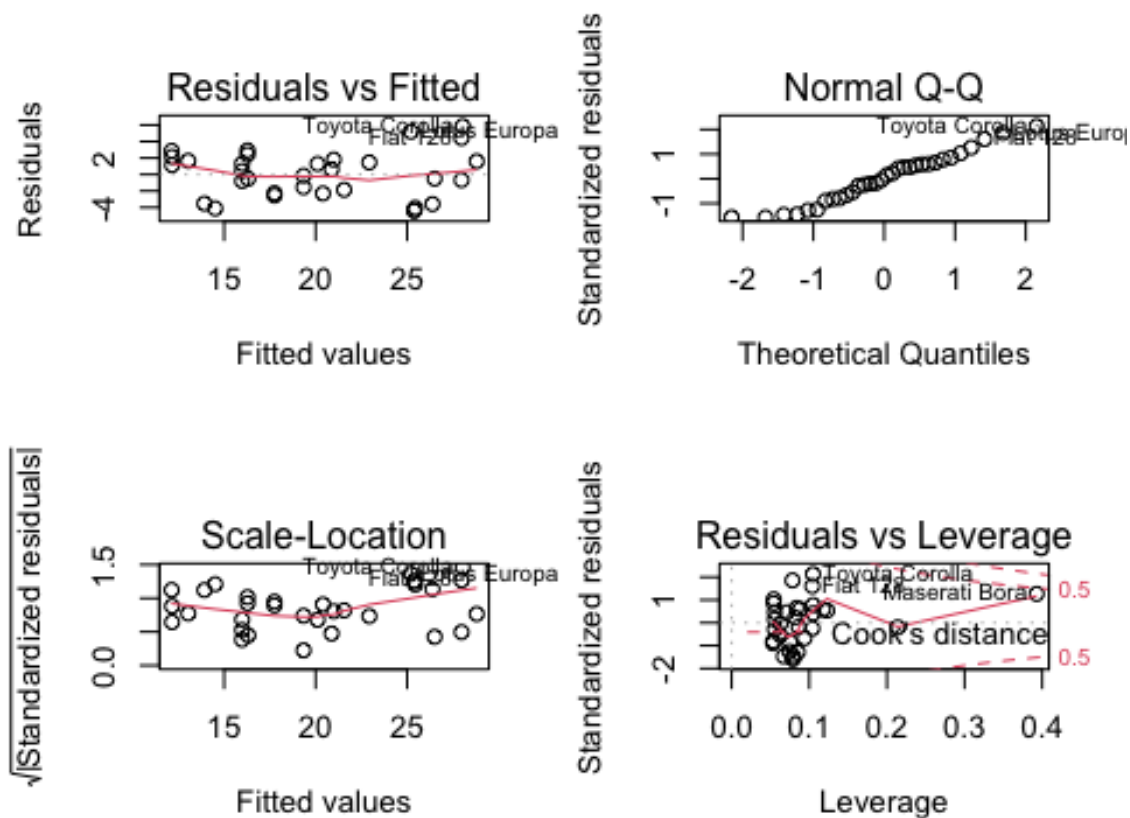
```
fit3

##
## Call:
## lm(formula = mpg ~ am + hp + cyl, data = mtcars)
##
## Coefficients:
## (Intercept)     amManual             hp           cyl
##    30.88834      3.90428       -0.03688      -1.12721

plot(fit3)
```

Residuals vs Fitted

Normal Q-Q

Scale-Location

Residuals vs Leverage
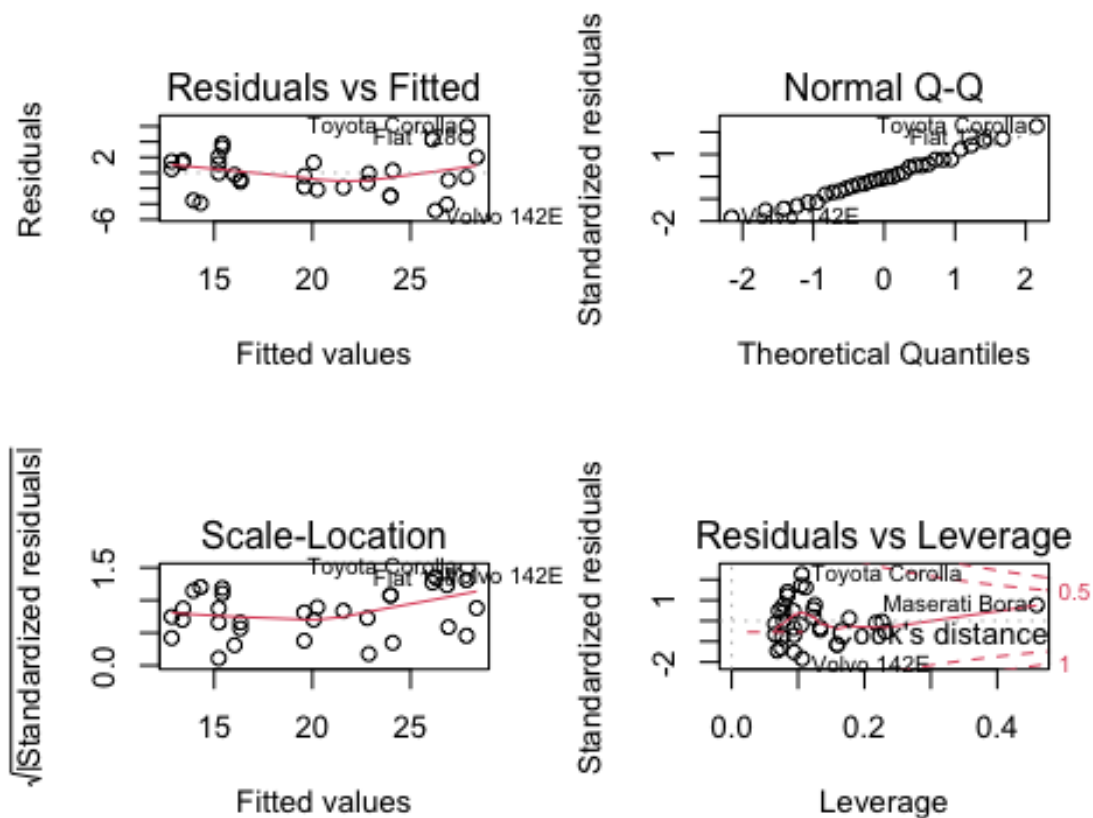

```
fit4

##
## Call:
## lm(formula = mpg ~ am + hp + cyl + disp, data = mtcars)
##
## Coefficients:
## (Intercept)      amManual           hp          cyl         disp
##    30.476440      3.445269    -0.032962    -0.834497    -0.007745

plot(fit4)
```

Residuals vs Fitted

Residuals

Toyota Corolla
Fiat 128
Volvo 142E

2
-6

15   20   25

Fitted values

Normal Q-Q

Standardized residuals

Toyota Corolla
Fiat 128

Volvo 142E

2
1
-1

-2 -1 0 1 2

Theoretical Quantiles

Scale-Location

√|Standardized residuals|

Toyota Corolla
Fiat 128   Volvo 142E

1.5
0.0

15   20   25

Fitted values

Residuals vs Leverage

Standardized residuals

Toyota Corolla
Pontiac Firebird
Cook's distance
Cadillac Fleetwood

1
-2

0.5
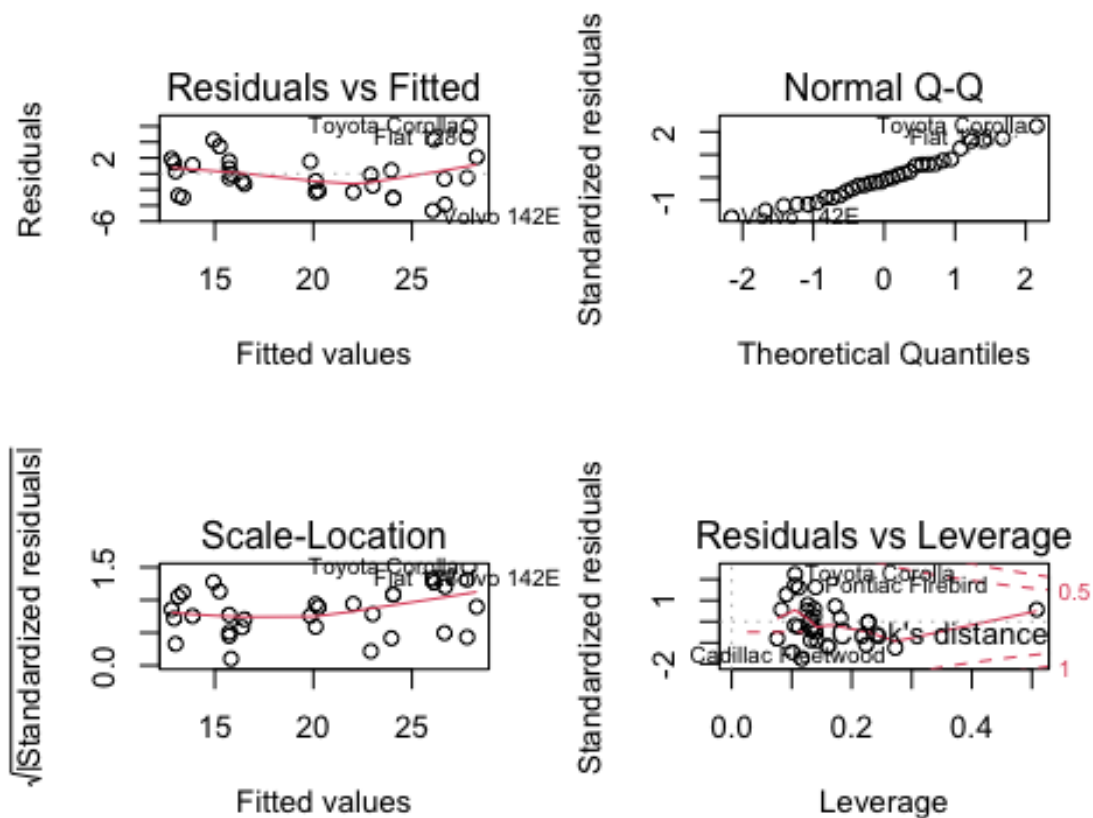
1

0.0   0.2   0.4

Leverage

```
fit5

##
## Call:
## lm(formula = mpg ~ am + hp + cyl + disp + wt, data = mtcars)
##
## Coefficients:
## (Intercept)      amManual            hp           cyl          disp
wt
##    38.20280       1.55649      -0.02796      -1.10638       0.01226      -
3.30262

plot(fit5)
```

Residuals vs Fitted

Chrysler Imperial ● Toyota Corolla ● Fiat 128 ●

Residuals: 2, -4
Fitted values: 10, 15, 20, 25, 30

Normal Q-Q

Toyota Corolla Fiat 128
Chrysler Imp

Standardized residuals: 2, -1
Theoretical Quantiles: -2, -1, 0, 1, 2

Scale-Location

Chrysler Imperial Toyota Corolla Fiat 128

√|Standardized residuals|: 0.0, 1.5
Fitted values: 10, 15, 20, 25, 30

Residuals vs Leverage

Chrysler Imperial
Maserati Bora
Cook's distance
Toyota Corona
0.5
0.5

Standardized residuals: 1, -2
Leverage: 0.0, 0.2, 0.4