Mach 1 Regression: Consumer Confidence and NJ Unemploymetn

```
library(readxl)
library(data.table)
library(fpp3)
```

```
## -- Attaching packages --------------------------------------- fpp3 0.4.0 --

## v tibble      3.1.5     v tsibble     1.1.0
## v dplyr       1.0.7     v tsibbledata 0.3.0
## v tidyr       1.1.4     v feasts      0.2.2
## v lubridate   1.8.0     v fable       0.3.1
## v ggplot2     3.3.5

## -- Conflicts ------------------------------------------------ fpp3_conflicts --
## x dplyr::between()     masks data.table::between()
## x lubridate::date()    masks base::date()
## x dplyr::filter()      masks stats::filter()
## x dplyr::first()       masks data.table::first()
## x lubridate::hour()    masks data.table::hour()
## x tsibble::intersect() masks base::intersect()
## x tsibble::interval()  masks lubridate::interval()
## x lubridate::isoweek() masks data.table::isoweek()
## x tsibble::key()       masks data.table::key()
## x dplyr::lag()         masks stats::lag()
## x dplyr::last()        masks data.table::last()
## x lubridate::mday()    masks data.table::mday()
## x lubridate::minute()  masks data.table::minute()
## x lubridate::month()   masks data.table::month()
## x lubridate::quarter() masks data.table::quarter()
## x lubridate::second()  masks data.table::second()
## x tsibble::setdiff()   masks base::setdiff()
## x tsibble::union()     masks base::union()
## x lubridate::wday()    masks data.table::wday()
## x lubridate::week()    masks data.table::week()
## x lubridate::yday()    masks data.table::yday()
## x lubridate::year()    masks data.table::year()
```

```
library(car)
```

```
## Loading required package: carData

##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(corrplot)
```

```
## corrplot 0.90 loaded
```

```
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
library(ggplot2)
library(readxl)
Tng_Ctr_Hour2 <- read_excel("C:/RBS/Business Forecasting/Group Project/Tng_Ctr_Hour2.xlsx")
View(Tng_Ctr_Hour2)
```

```
setDT(Tng_Ctr_Hour2)
class(Tng_Ctr_Hour2)
```

```
## [1] "data.table" "data.frame"
```

```
Tng_Ctr_Hour2[, Quarter := factor(Quarter, ordered = T)]
Tng_Ctr_Hour2[, Month := factor(Month, ordered = T)]
summary(Tng_Ctr_Hour2)
```

```
##      Year           Quarter     Month      Device_Hrs     DH_Prev_Year
##  Length:81          Q1:21   Apr    : 7   Min.   : 222.8   Length:81
##  Class :character   Q2:21   Aug    : 7   1st Qu.: 899.0   Class :character
##  Mode  :character   Q3:21   Feb    : 7   Median :1008.0   Mode  :character
##                     Q4:18   Jan    : 7   Mean   : 990.1
##                             Jul    : 7   3rd Qu.:1101.7
##                             Jun    : 7   Max.   :1519.9
##                             (Other):39
##  DH_YoY_Change      DH_YoY_Ch_Per      Total_Inst_Hrs
##  Length:81          Length:81          Min.   : 504.6
##  Class :character   Class :character   1st Qu.:1937.3
##  Mode  :character   Mode  :character   Median :2203.2
```

```
##                                     Mean   :2165.7
##                                     3rd Qu.:2446.8
##                                     Max.   :3084.1
##
##  Total_Inst_Hrs_Prev_Year Inst_Hrs_YoY_Change Total_Inst_Hrs_YoY_Change_Per2
##  Length:81                Length:81           Length:81
##  Class :character         Class :character    Class :character
##  Mode  :character         Mode  :character    Mode  :character
##
##
##
##
##    Cons_Sent          NJURN
##  Min.   : 70.30   Min.   : 2.900
##  1st Qu.: 89.00   1st Qu.: 4.100
##  Median : 93.80   Median : 4.900
##  Mean   : 91.49   Mean   : 5.615
##  3rd Qu.: 97.90   3rd Qu.: 6.200
##  Max.   :101.40   Max.   :16.600
##
```

```
str(Tng_Ctr_Hour2)
```

```
## Classes 'data.table' and 'data.frame':   81 obs. of  13 variables:
##  $ Year                          : chr  "2015-01" "2015-02" "2015-03" "2015-04" ...
##  $ Quarter                       : Ord.factor w/ 4 levels "Q1"<"Q2"<"Q3"<..: 1 1 1 2 2 2 3 3 3 4 ...
##  $ Month                         : Ord.factor w/ 12 levels "Apr"<"Aug"<"Dec"<..: 5 4 8 1 9 7 6 2 12 3 ...
##  $ Device_Hrs                    : num  960 944 1429 1097 916 ...
##  $ DH_Prev_Year                  : chr  "NA" "NA" "NA" "NA" ...
##  $ DH_YoY_Change                 : chr  "NA" "NA" "NA" "NA" ...
##  $ DH_YoY_Ch_Per                 : chr  "NA" "NA" "NA" "NA" ...
##  $ Total_Inst_Hrs                : num  1701 1614 2533 2152 1695 ...
##  $ Total_Inst_Hrs_Prev_Year      : chr  "NA" "NA" "NA" "NA" ...
##  $ Inst_Hrs_YoY_Change           : chr  "NA" "NA" "NA" "NA" ...
##  $ Total_Inst_Hrs_YoY_Change_Per2: chr  "NA" "NA" "NA" "NA" ...
##  $ Cons_Sent                     : num  98.1 95.4 93 95.9 90.7 96.1 93.1 91.9 87.2 90 ...
##  $ NJURN                         : num  6.8 6.7 6.3 5.8 6 5.9 6.2 5.5 5.1 4.9 ...
##  - attr(*, ".internal.selfref")=<externalptr>
```

```
TCH2 = subset(Tng_Ctr_Hour2, select = c(Month, Device_Hrs, Cons_Sent, NJURN ))
```

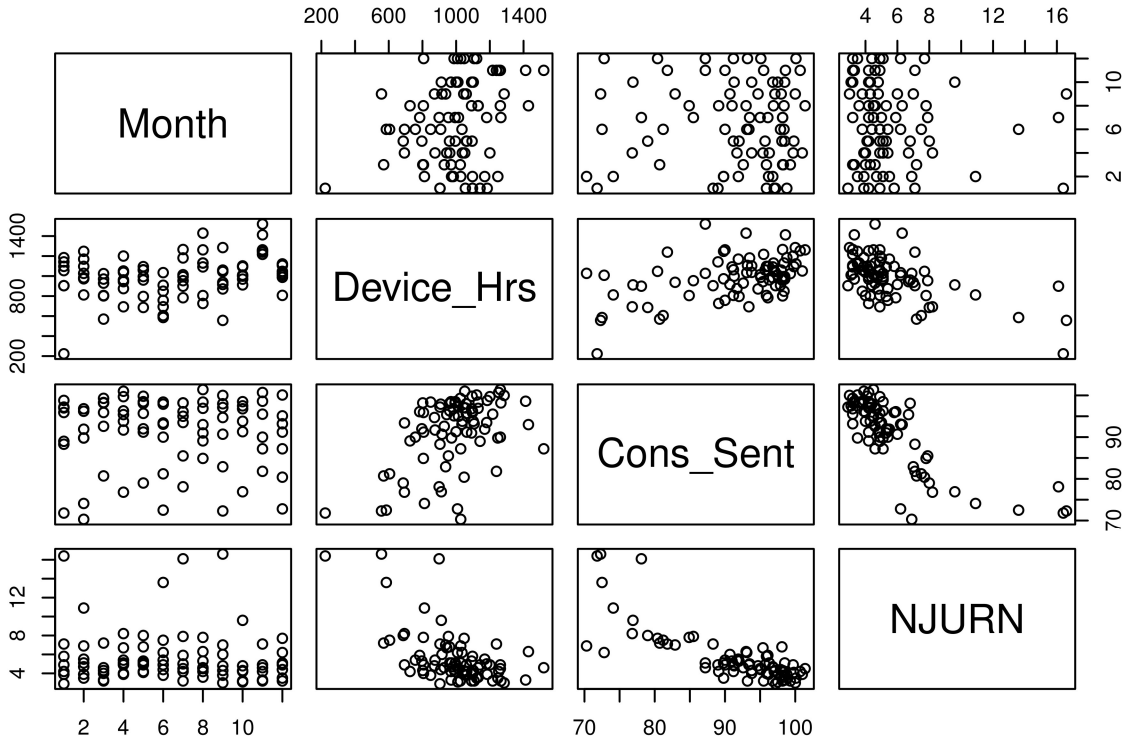
```
LM_TCH = lm(Device_Hrs ~ ., TCH2)
summary(LM_TCH)
```

```
##
## Call:
## lm(formula = Device_Hrs ~ ., data = TCH2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -358.27  -81.39    4.97   89.57  380.81
##
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  628.036    368.350   1.705 0.092828 .
## Month.L      191.566     54.791   3.496 0.000842 ***
## Month.Q      123.589     55.897   2.211 0.030454 *
## Month.C     -101.173     54.663  -1.851 0.068598 .
## Month^4      -78.516     55.482  -1.415 0.161650
## Month^5      -23.754     54.505  -0.436 0.664370
## Month^6     -186.716     54.559  -3.422 0.001063 **
## Month^7      -84.275     55.469  -1.519 0.133390
## Month^8     -206.045     54.858  -3.756 0.000364 ***
## Month^9      120.027     53.920   2.226 0.029380 *
## Month^10       1.480     53.683   0.028 0.978094
## Month^11     -72.850     53.229  -1.369 0.175693
## Cons_Sent      5.674      3.494   1.624 0.109096
## NJURN        -27.755     10.178  -2.727 0.008151 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 139.5 on 67 degrees of freedom
## Multiple R-squared:  0.6347, Adjusted R-squared:  0.5638
## F-statistic: 8.954 on 13 and 67 DF,  p-value: 2.83e-10
```

We show a relatively strong correlation between the model and Device Hours

```
pairs(TCH2)
```

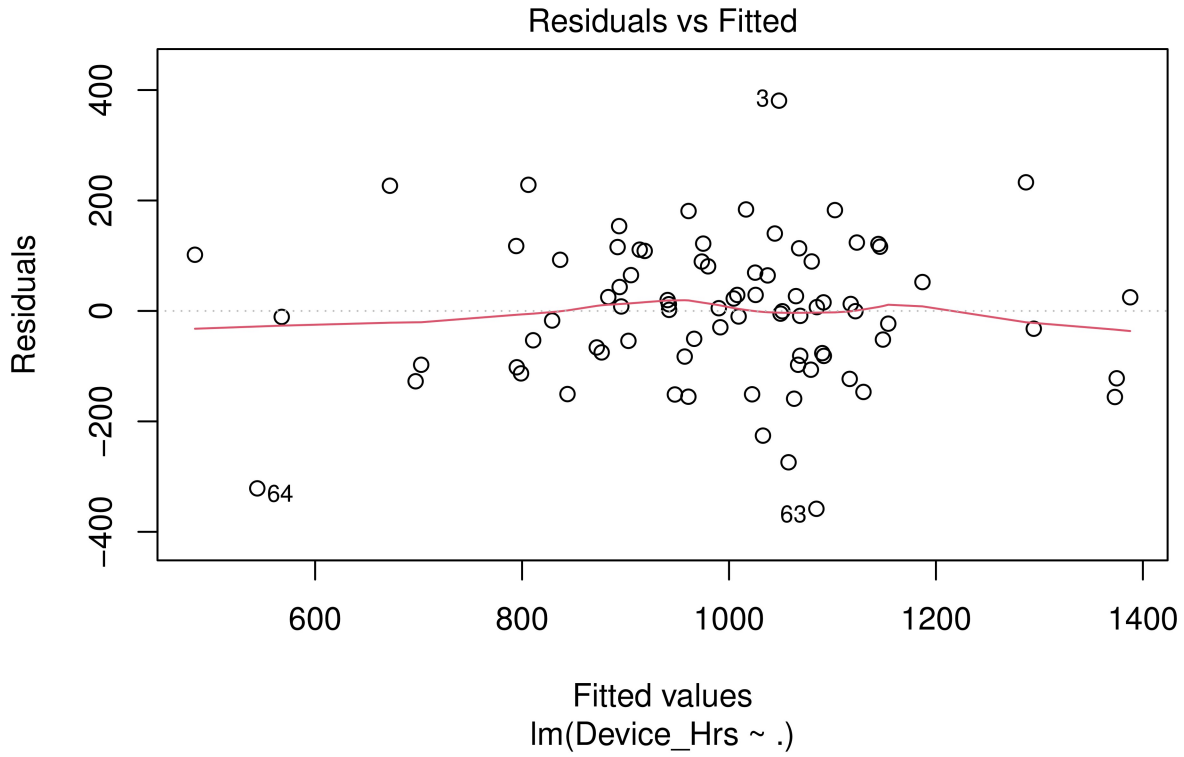


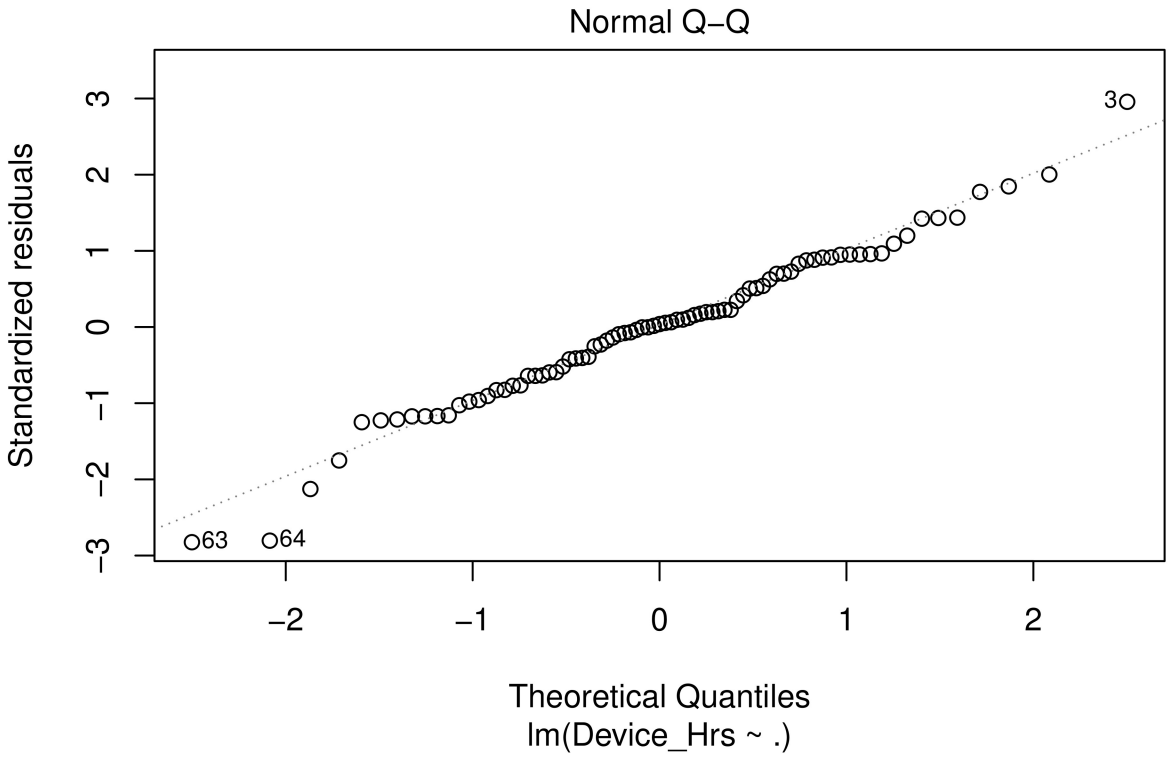
We can see some positive correlation between Device Hours and Consumer Sentiment and some negative correlation between Device Hours and NJ URN

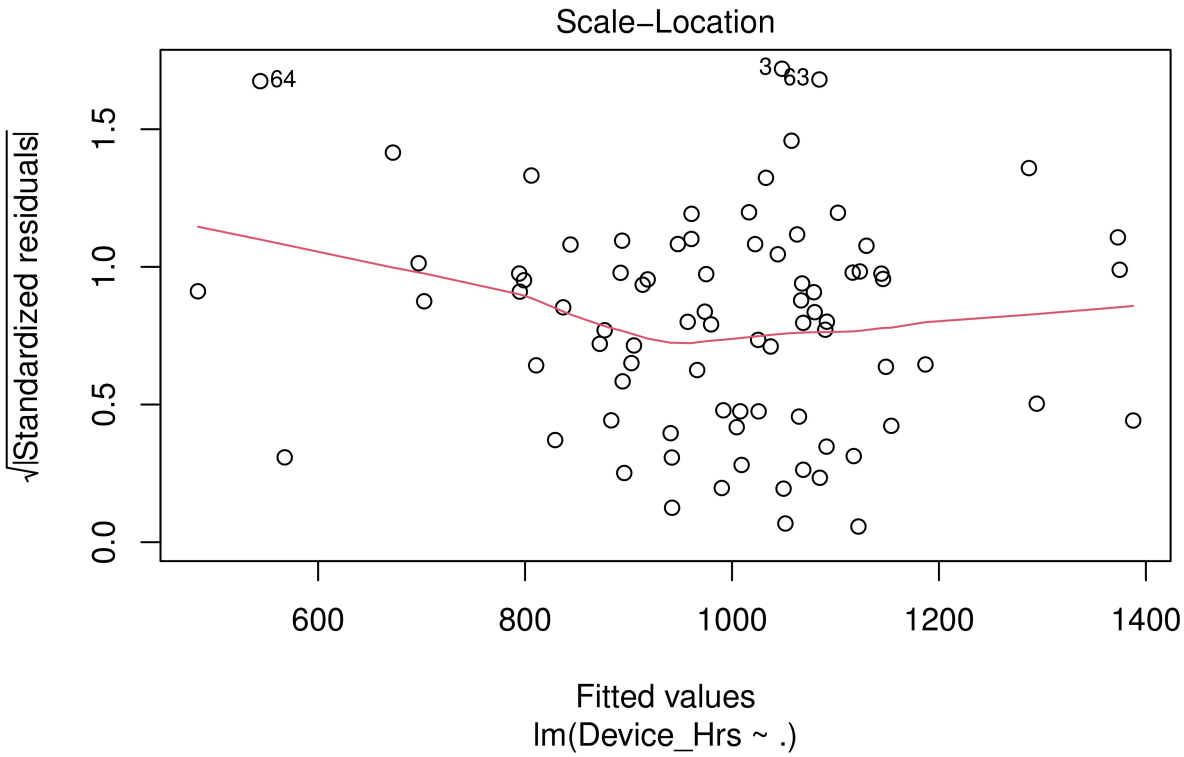
```
str(TCH2)
```

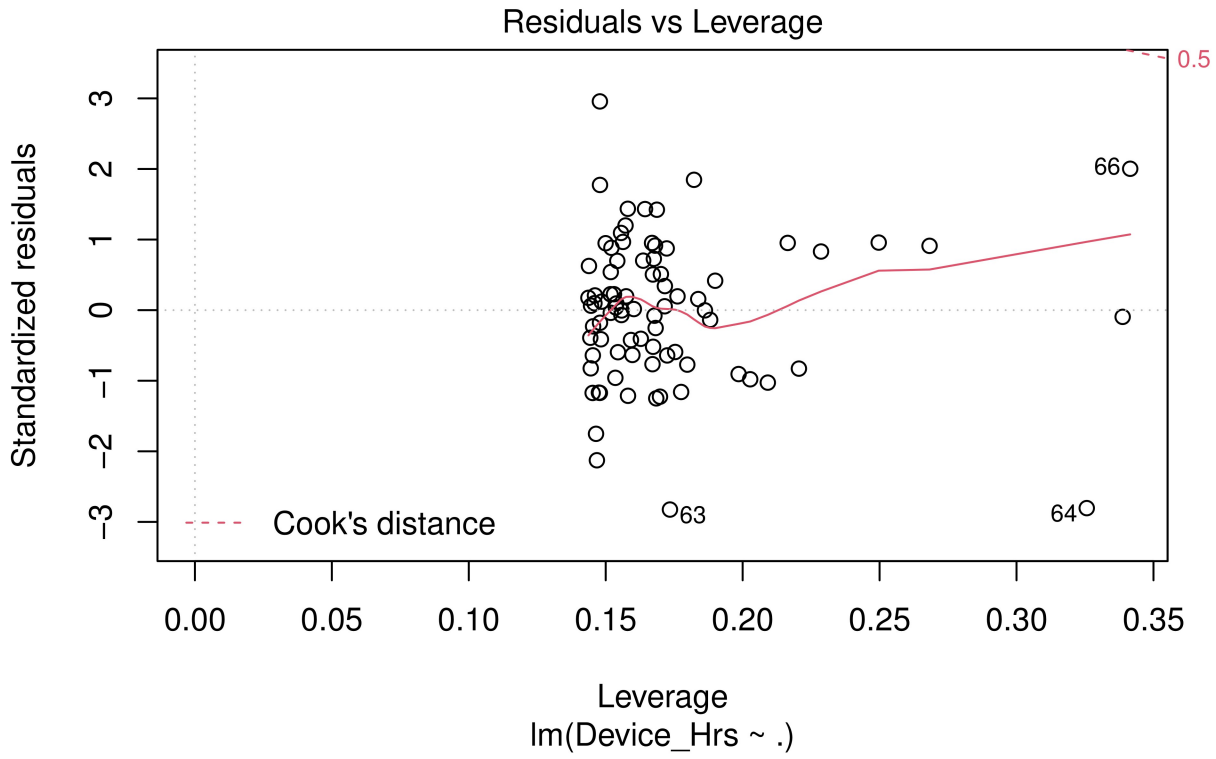
```
## Classes 'data.table' and 'data.frame':   81 obs. of  4 variables:
##  $ Month     : Ord.factor w/ 12 levels "Apr"<"Aug"<"Dec"<..: 5 4 8 1 9 7 6 2 12 11 ...
##  $ Device_Hrs: num  960 944 1429 1097 916 ...
##  $ Cons_Sent : num  98.1 95.4 93 95.9 90.7 96.1 93.1 91.9 87.2 90 ...
##  $ NJURN     : num  6.8 6.7 6.3 5.8 6 5.9 6.2 5.5 5.1 4.9 ...
##  - attr(*, ".internal.selfref")=<externalptr>
```

```
plot(LM_TCH)
```

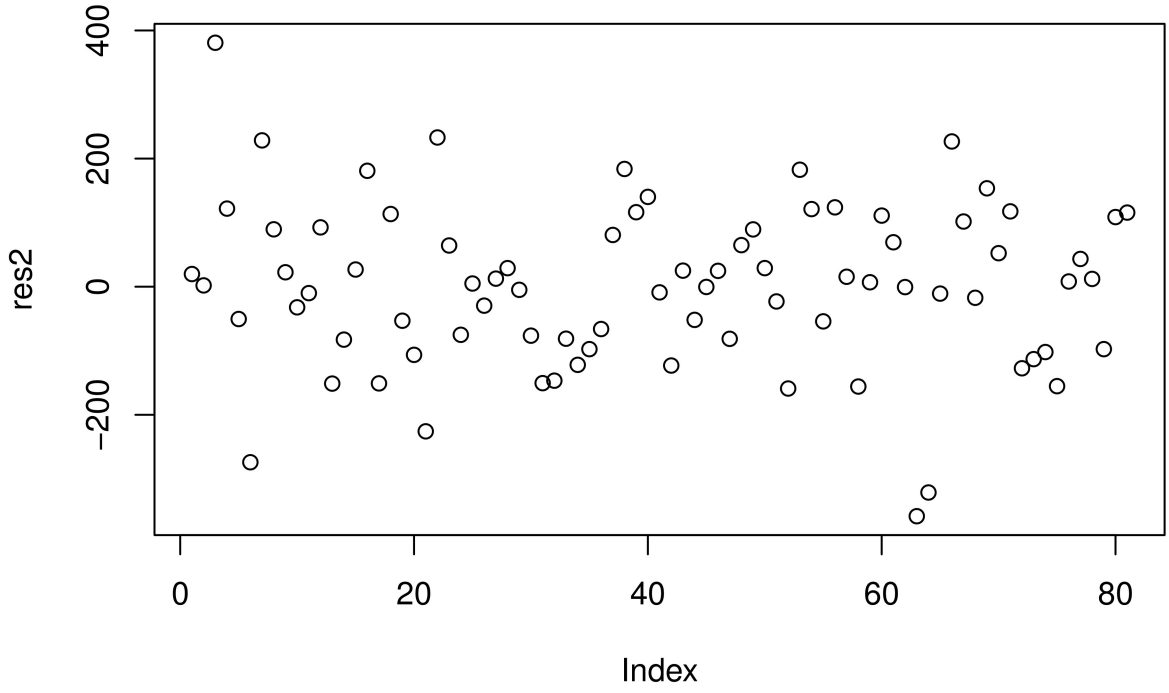



Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Device_Hrs ~ .)
3
63
64
−3
−2
−1
0
1
2
3

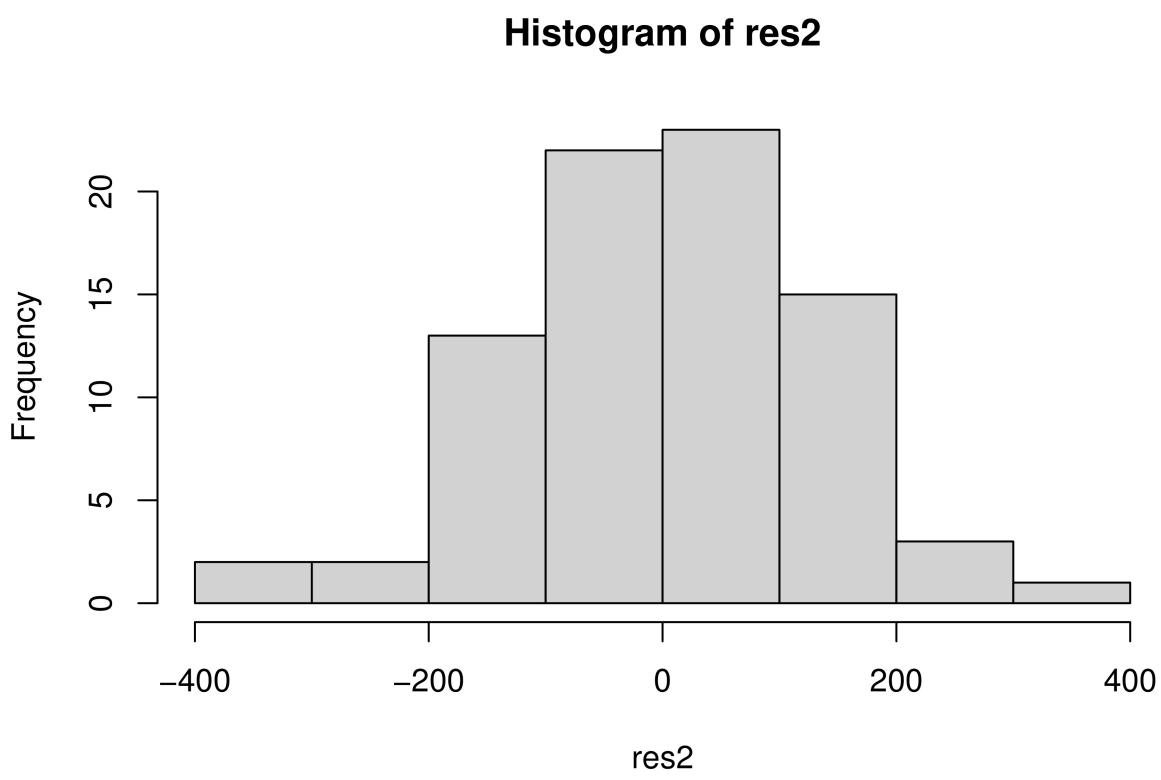

Scale−Location
64
3
63
√|Standardized residuals|
0.0
0.5
1.0
1.5
600
800
1000
1200
1400
Fitted values
lm(Device_Hrs ~ .)




```
res2= residuals(LM_TCH)
plot(res2)
```




```
hist(res2)
```



Step of our formula

```
TCH.Step = step(LM_TCH)
```

```
## Start:  AIC=812.61
## Device_Hrs ~ Month + Cons_Sent + NJURN
##
##             Df Sum of Sq     RSS    AIC
## <none>                   1304036 812.61
## - Cons_Sent  1     51325 1355361 813.74
## - NJURN      1    144740 1448776 819.13
## - Month     11   1067021 2371057 839.04
```

```
TCH.Step
```

```
##
## Call:
## lm(formula = Device_Hrs ~ Month + Cons_Sent + NJURN, data = TCH2)
##
## Coefficients:
## (Intercept)      Month.L      Month.Q      Month.C      Month^4      Month^5
##     628.036      191.566      123.589     -101.173      -78.516      -23.754
##     Month^6      Month^7      Month^8      Month^9     Month^10     Month^11
##    -186.716      -84.275     -206.045      120.027        1.480      -72.850
##   Cons_Sent         NJURN
##       5.674      -27.755
```

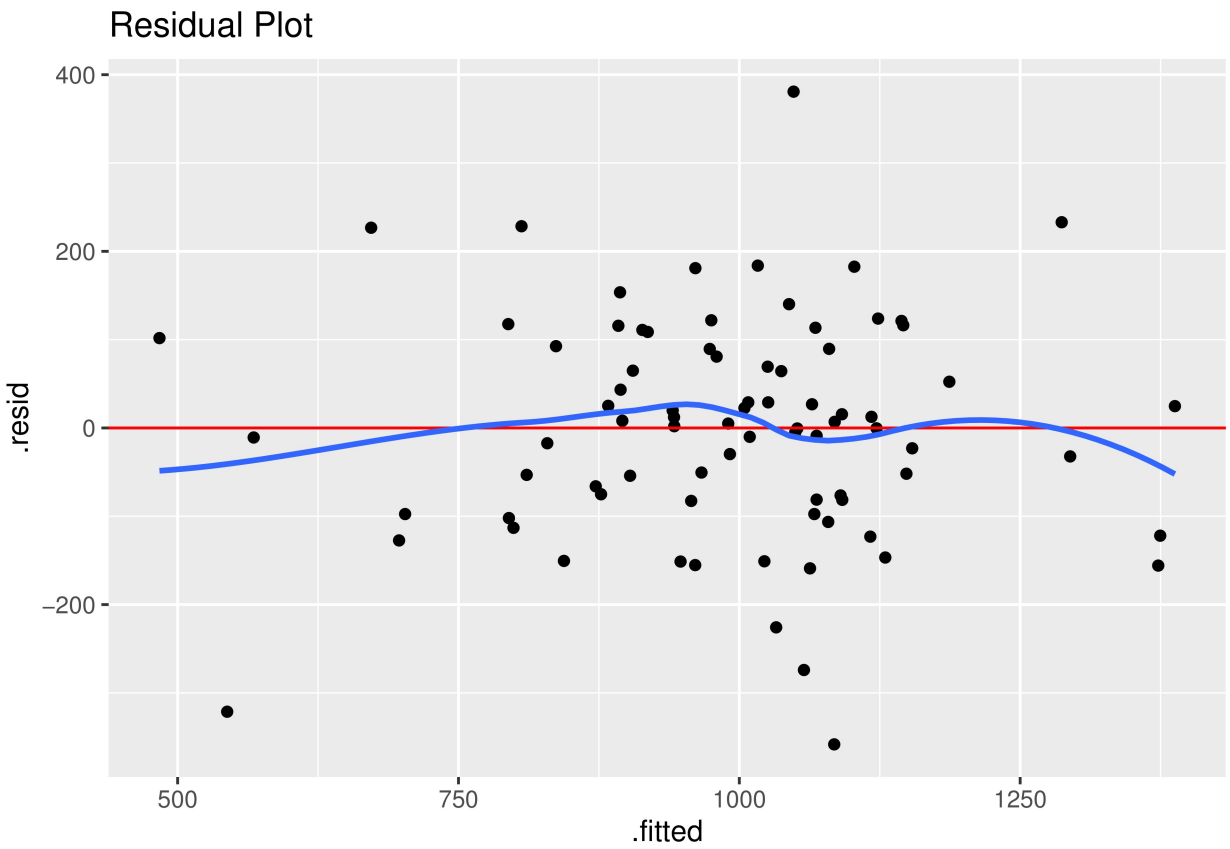
Summary of the Step

```
summary(TCH.Step)
```

```
##
## Call:
## lm(formula = Device_Hrs ~ Month + Cons_Sent + NJURN, data = TCH2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -358.27  -81.39    4.97   89.57  380.81
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  628.036    368.350   1.705 0.092828 .
## Month.L      191.566     54.791   3.496 0.000842 ***
## Month.Q      123.589     55.897   2.211 0.030454 *
## Month.C     -101.173     54.663  -1.851 0.068598 .
## Month^4      -78.516     55.482  -1.415 0.161650
## Month^5      -23.754     54.505  -0.436 0.664370
## Month^6     -186.716     54.559  -3.422 0.001063 **
## Month^7      -84.275     55.469  -1.519 0.133390
## Month^8     -206.045     54.858  -3.756 0.000364 ***
## Month^9      120.027     53.920   2.226 0.029380 *
## Month^10       1.480     53.683   0.028 0.978094
## Month^11     -72.850     53.229  -1.369 0.175693
## Cons_Sent      5.674      3.494   1.624 0.109096
## NJURN        -27.755     10.178  -2.727 0.008151 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 139.5 on 67 degrees of freedom
## Multiple R-squared:  0.6347, Adjusted R-squared:  0.5638
## F-statistic: 8.954 on 13 and 67 DF,  p-value: 2.83e-10
```

```
p1= ggplot(TCH.Step) +
  aes(x=.fitted, y=.resid)+
  geom_point() + geom_abline(intercept = 0, slope = 0, color="red") + geom_smooth(method = "loess", se =
  labs(title= "Residual Plot")
p1
```

```
## `geom_smooth()` using formula 'y ~ x'
```


Residual Plot
400
200
0
−200
.resid
500
750
1000
1250
.fitted

“‘