marks_proj_work

Mark Austin

2022-11-14

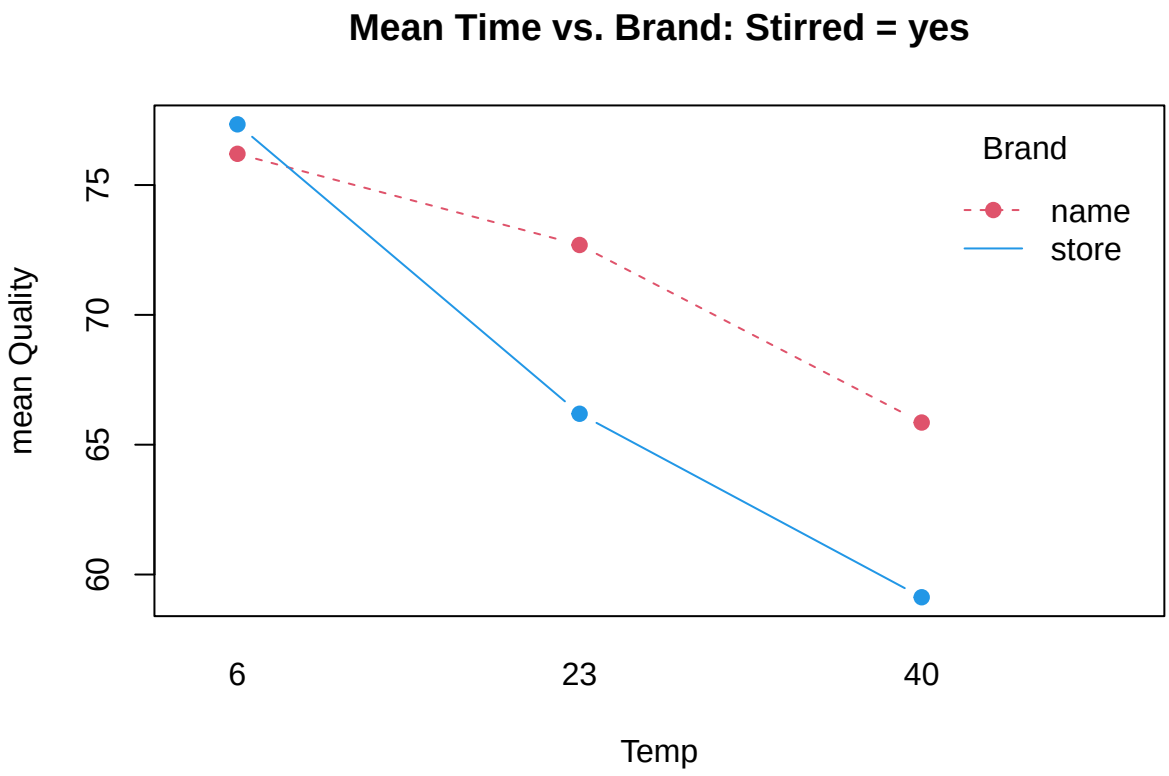
Exploratory Analysis

```
#using Halid's code to keep data frame consistent
library(tidyverse)
df_eff <- read_csv('effervescence.csv', col_types = 'fffnnn')
df_eff %>% head()
```

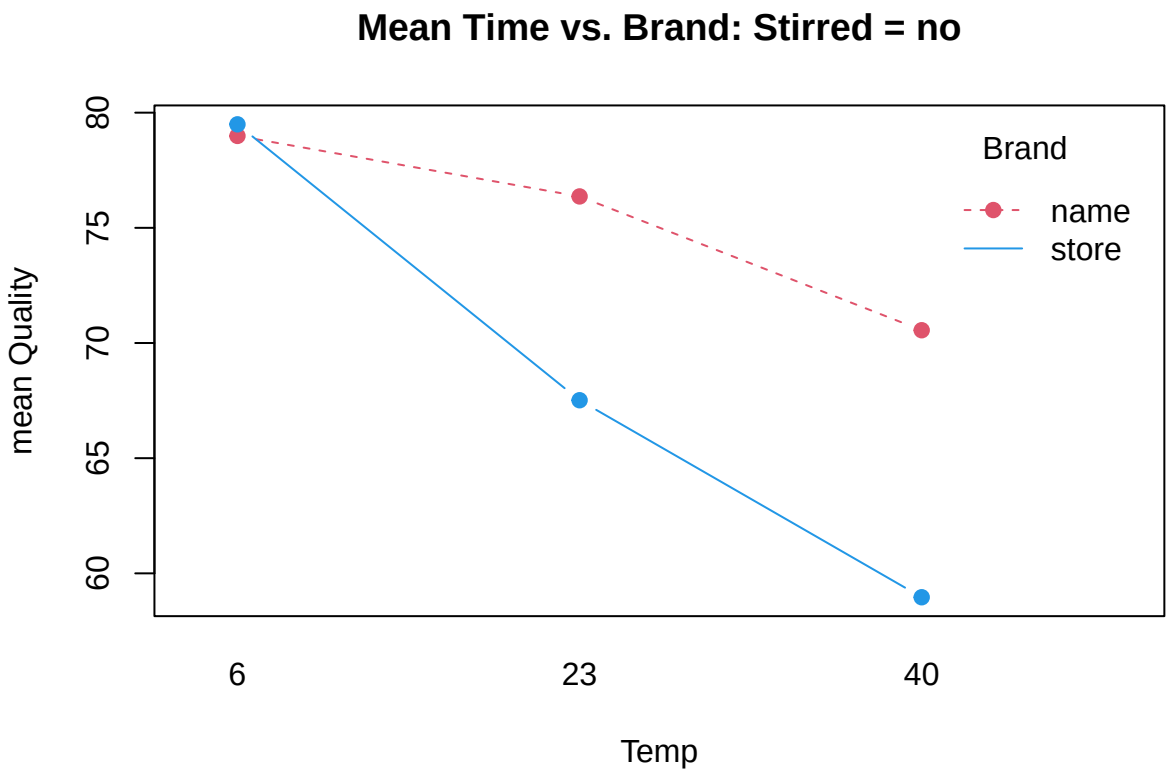
```
## # A tibble: 6 x 6
##   Brand Temp  Stirred Order  Time OrgTime
##   <fct> <fct> <fct>   <dbl> <dbl>   <dbl>
## 1 name  6     yes         8  77.2    75.5
## 2 name  23    yes         3  75.4    68.1
## 3 name  40    yes         7  68.1    44.8
## 4 store 6     yes         1  77.9    78.4
## 5 store 23    yes         2  66.4    40.6
## 6 store 40    yes        18  59.8    27.4
```

```
##3 factor interaction plot based on HW7 code

with(df_eff%>%filter(Stirred=="yes"),interaction.plot(Temp,Brand,Time,
            type="b", pch=19, col=c(2,4), ylab="mean Quality",
           main="Mean Time vs. Brand: Stirred = yes"))
```




```
with(df_eff%>%filter(Stirred=="no"),interaction.plot(Temp,Brand,Time,
          type="b", pch=19, col=c(2,4), ylab="mean Quality",
          main="Mean Time vs. Brand: Stirred = no"))
```



From the three factor interaction plots, does the 3 factor interaction look obvious here? The brand by temp interaction is clear.

Analysis and Results

```
#model with stirred as block effect without interaction
aov_block_eff <- aov(lm_block_eff <- lm(Time ~ Brand * Temp + Stirred, data = df_eff))
summary(lm_block_eff)
```

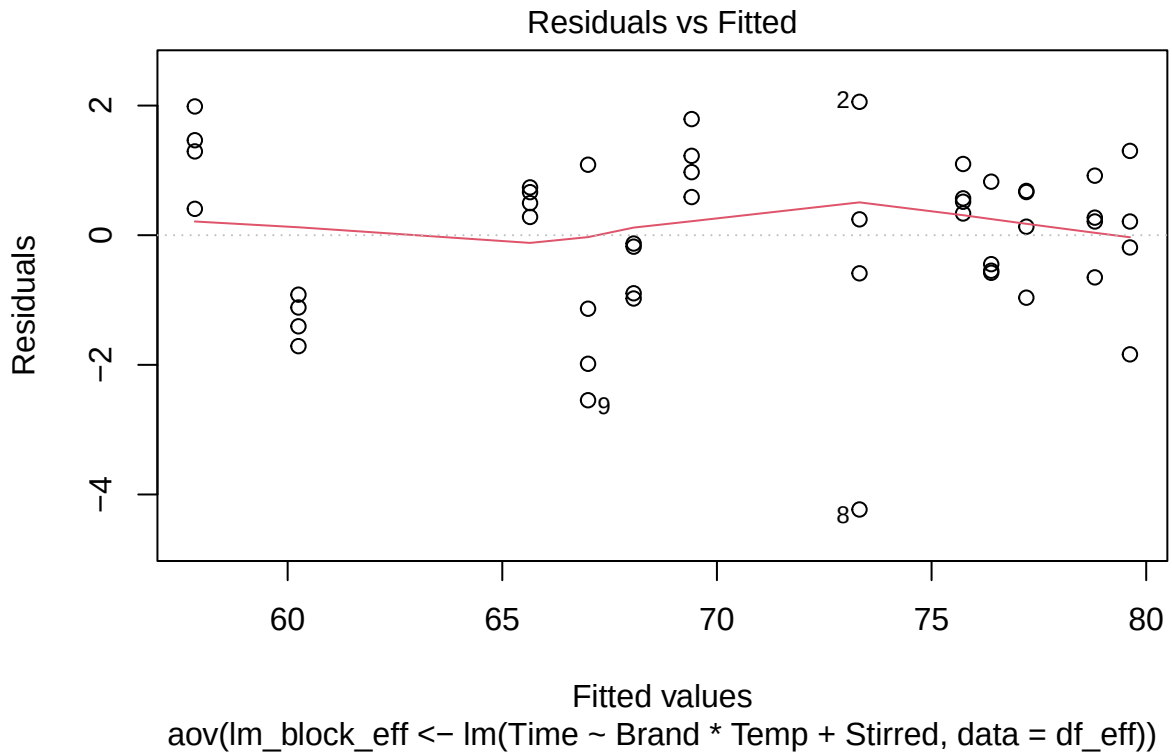
```
##
## Call:
## lm(formula = Time ~ Brand * Temp + Stirred, data = df_eff)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.2315 -0.7120  0.2577  0.7596  2.0577
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        76.3899     0.5035 151.710  < 2e-16 ***
## Brandstore          0.8182     0.6593   1.241    0.222
## Temp23             -3.0690     0.6593  -4.655 3.38e-05 ***
## Temp40             -9.3922     0.6593 -14.246  < 2e-16 ***
## Stirredno           2.4133     0.3806   6.340 1.42e-07 ***
## Brandstore:Temp23  -8.4923     0.9323  -9.108 2.11e-11 ***
```

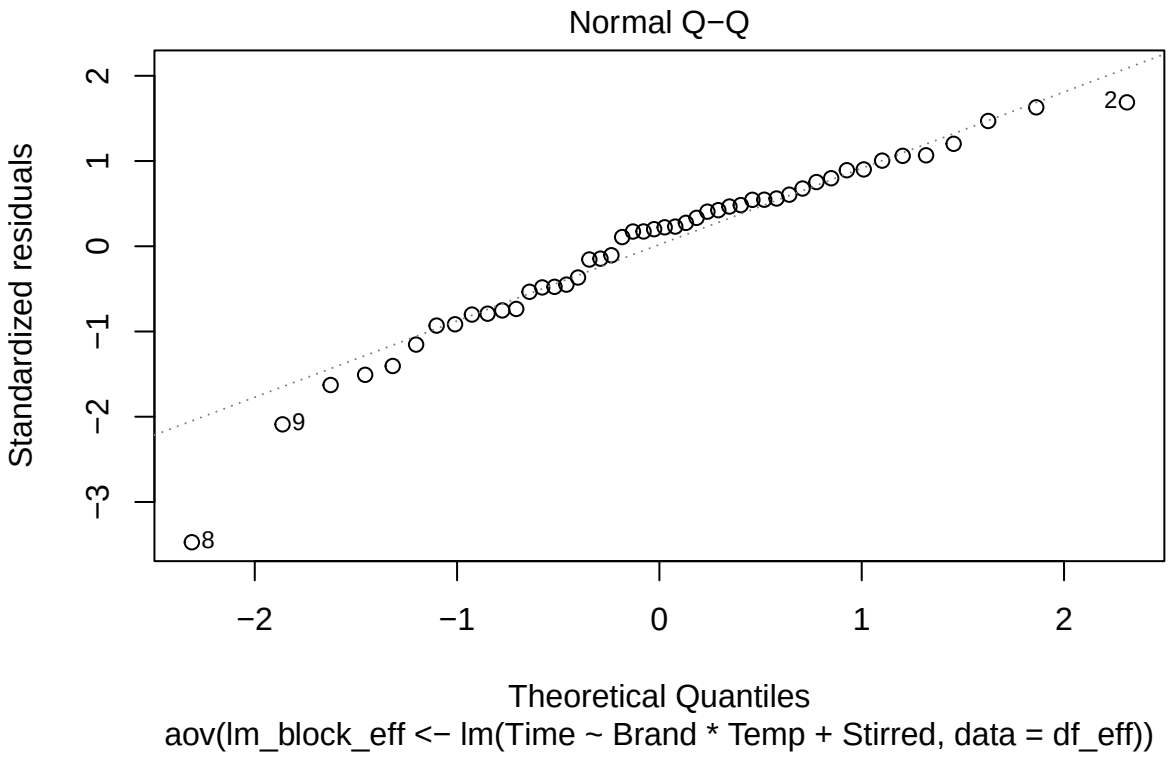
```
## Brandstore:Temp40  -9.9781     0.9323 -10.702 1.93e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.319 on 41 degrees of freedom
## Multiple R-squared:  0.9699, Adjusted R-squared:  0.9655
## F-statistic: 220.3 on 6 and 41 DF,  p-value: < 2.2e-16
```

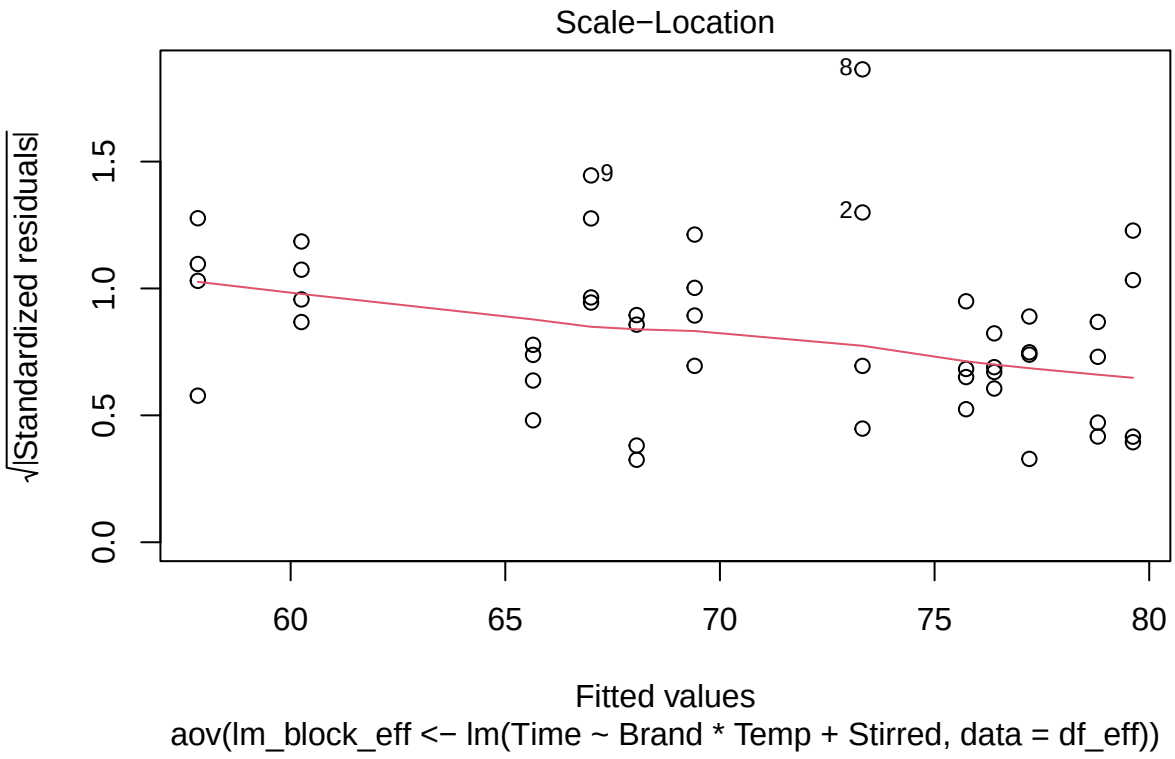
```
summary(aov_block_eff)
```

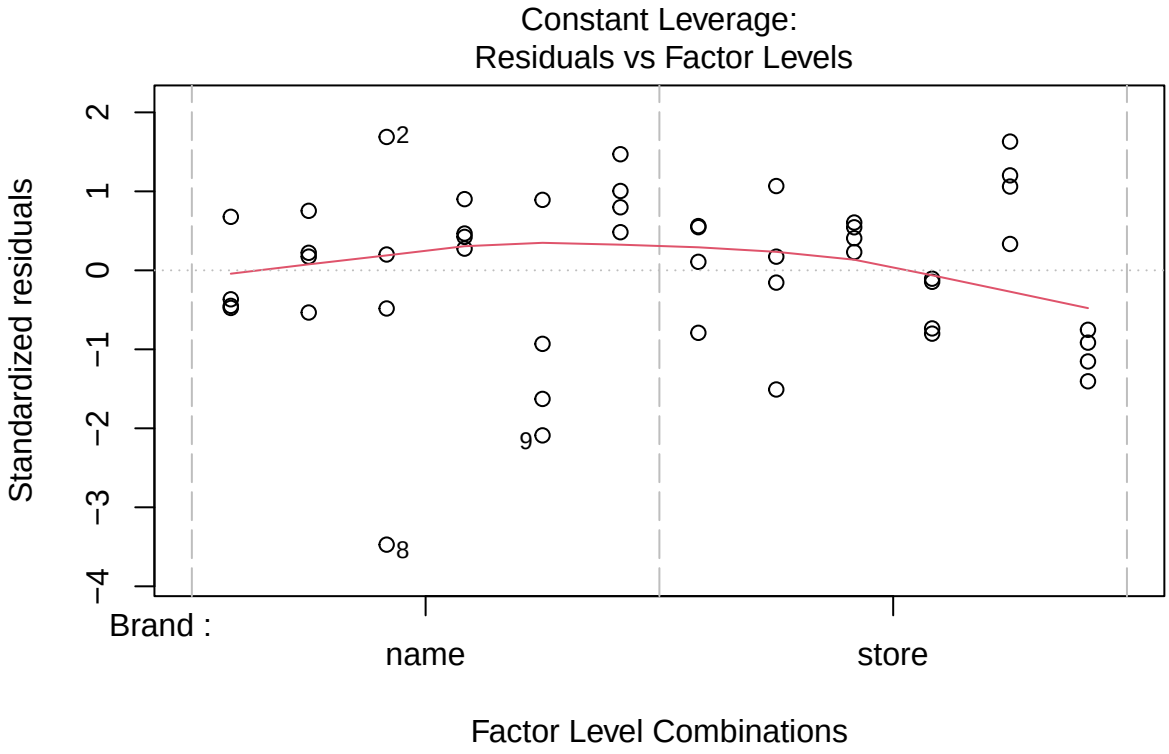
```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Brand        1  342.0   342.0  196.72  < 2e-16 ***
## Temp         2 1654.7   827.4  475.89  < 2e-16 ***
## Stirred      1   69.9    69.9   40.20 1.42e-07 ***
## Brand:Temp   2  231.9   115.9   66.68 1.30e-13 ***
## Residuals   41   71.3     1.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(aov_block_eff)
```




Normal Q−Q
Standardized residuals
Theoretical Quantiles
aov(lm_block_eff <− lm(Time ~ Brand * Temp + Stirred, data = df_eff))
2
9
8
−3
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
8
9
2
60
65
70
75
80
Fitted values
aov(lm_block_eff <− lm(Time ~ Brand * Temp + Stirred, data = df_eff))



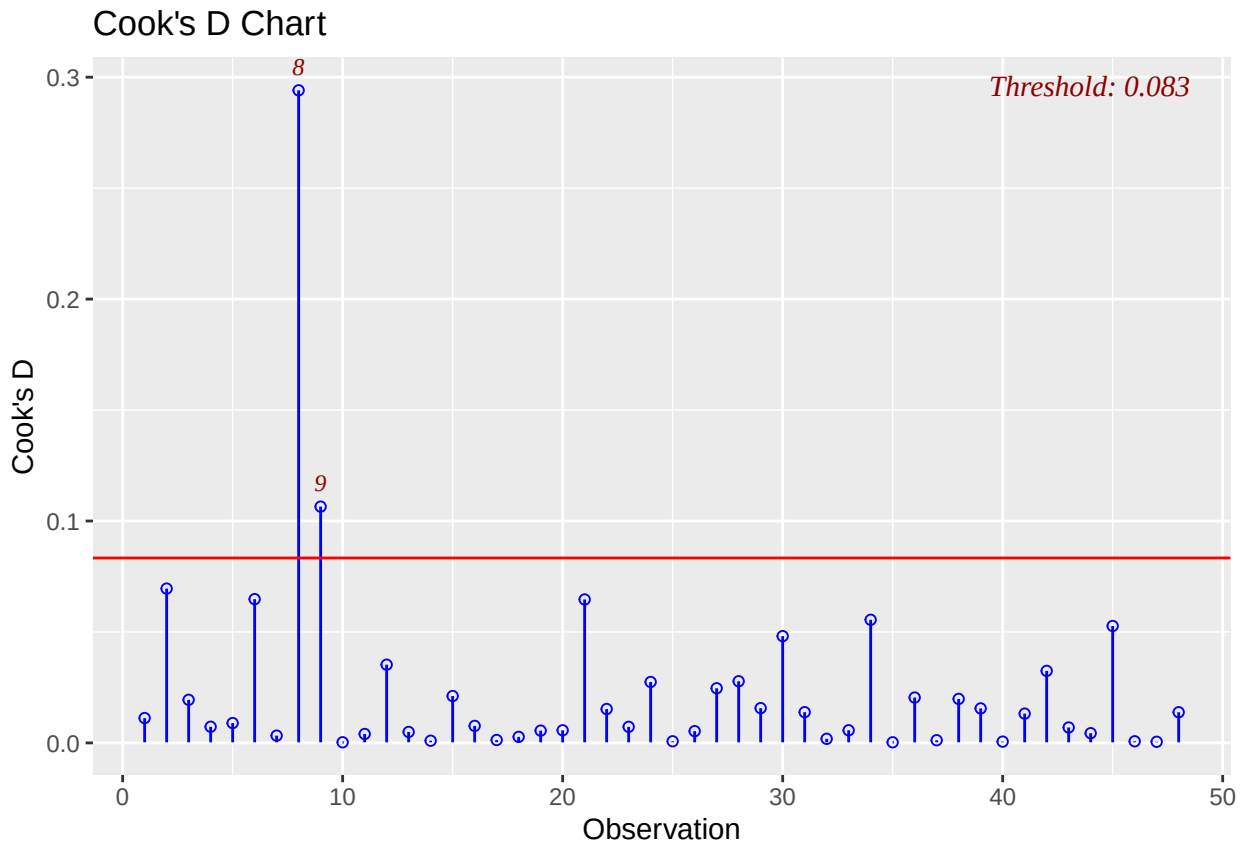

```
library(olsrr)
```

```
## Warning: package 'olsrr' was built under R version 4.1.3
```

```
##
## Attaching package: 'olsrr'
```

```
## The following object is masked from 'package:datasets':
##
##     rivers
```

```
ols_plot_cooksd_chart(lm_block_eff)
```




```
#added covariate Order model with stirred as block effect without interaction
aov_block_order_eff <- aov(lm_block_order_eff <- lm(Time ~ Brand * Temp + Stirred + Order, data = df_ef
summary(lm_block_order_eff)
```

```
##
## Call:
## lm(formula = Time ~ Brand * Temp + Stirred + Order, data = df_eff)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9257 -0.7714  0.0969  0.8613  2.4030
##
## Coefficients:
##                    Estimate Std. Error t value Pr(>|t|)
## (Intercept)        75.76840    0.66484 113.965  < 2e-16 ***
## Brandstore          0.95630    0.65881   1.452   0.1544
## Temp23             -2.91120    0.66104  -4.404 7.73e-05 ***
## Temp40             -9.17522    0.66943 -13.706  < 2e-16 ***
## Stirredno           1.46630    0.77003   1.904   0.0641 .
## Order               0.03946    0.02800   1.409   0.1664
## Brandstore:Temp23  -8.64026    0.92729  -9.318 1.42e-11 ***
## Brandstore:Temp40 -10.22471    0.93780 -10.903 1.51e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.303 on 40 degrees of freedom
```

```
## Multiple R-squared:  0.9713, Adjusted R-squared:  0.9663
## F-statistic: 193.7 on 7 and 40 DF,  p-value: < 2.2e-16
```

```
summary(aov_block_order_eff)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Brand        1  342.0   342.0 201.452  < 2e-16 ***
## Temp         2 1654.7   827.4 487.344  < 2e-16 ***
## Stirred      1   69.9    69.9  41.166 1.23e-07 ***
## Order        1    0.9     0.9   0.534    0.469
## Brand:Temp   2  234.3   117.2  69.010 1.08e-13 ***
## Residuals   40   67.9     1.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
library(car)
```

```
## Warning: package 'car' was built under R version 4.1.3
```

```
## Loading required package: carData
```

```
## Warning: package 'carData' was built under R version 4.1.3
```

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

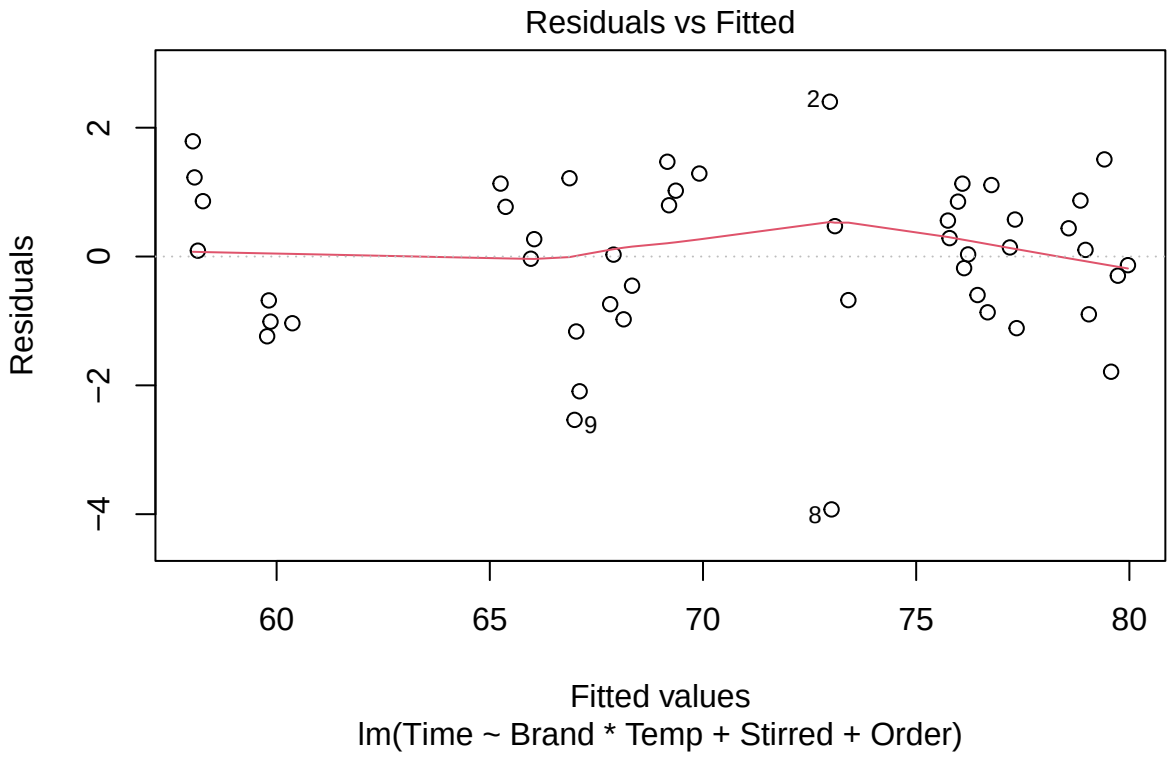
```
## The following object is masked from 'package:dplyr':
##
##     recode
```

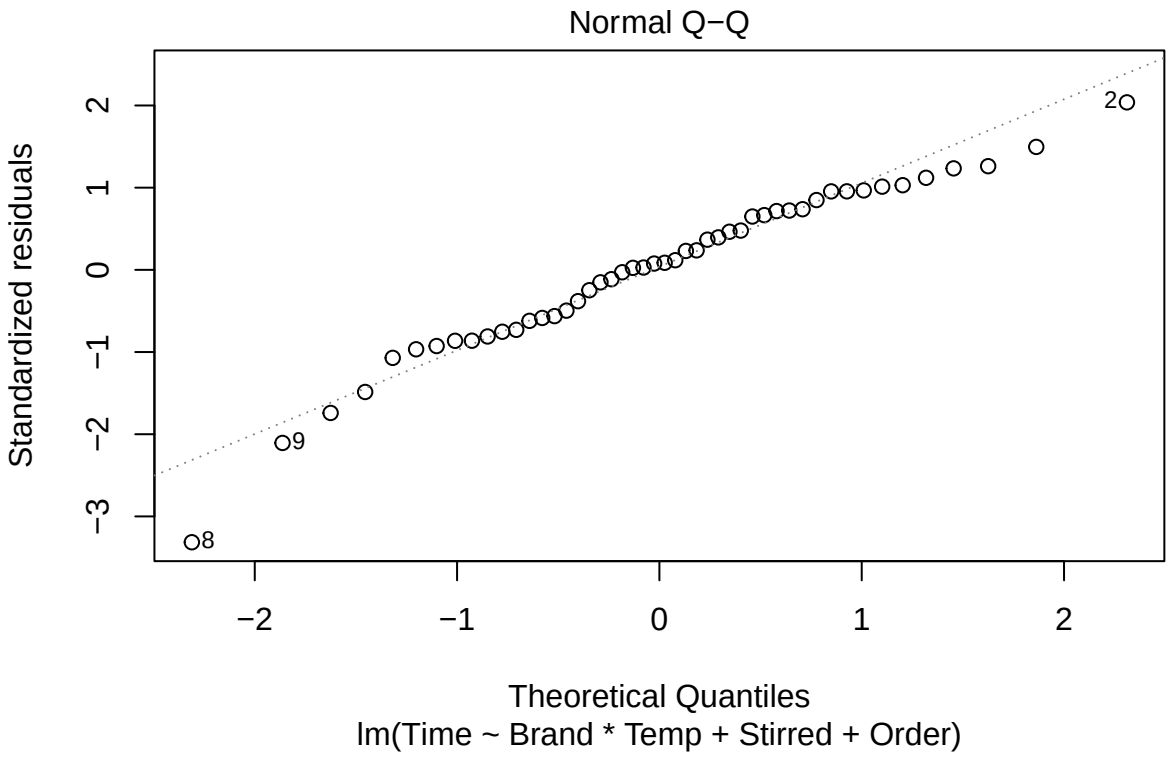
```
## The following object is masked from 'package:purrr':
##
##     some
```

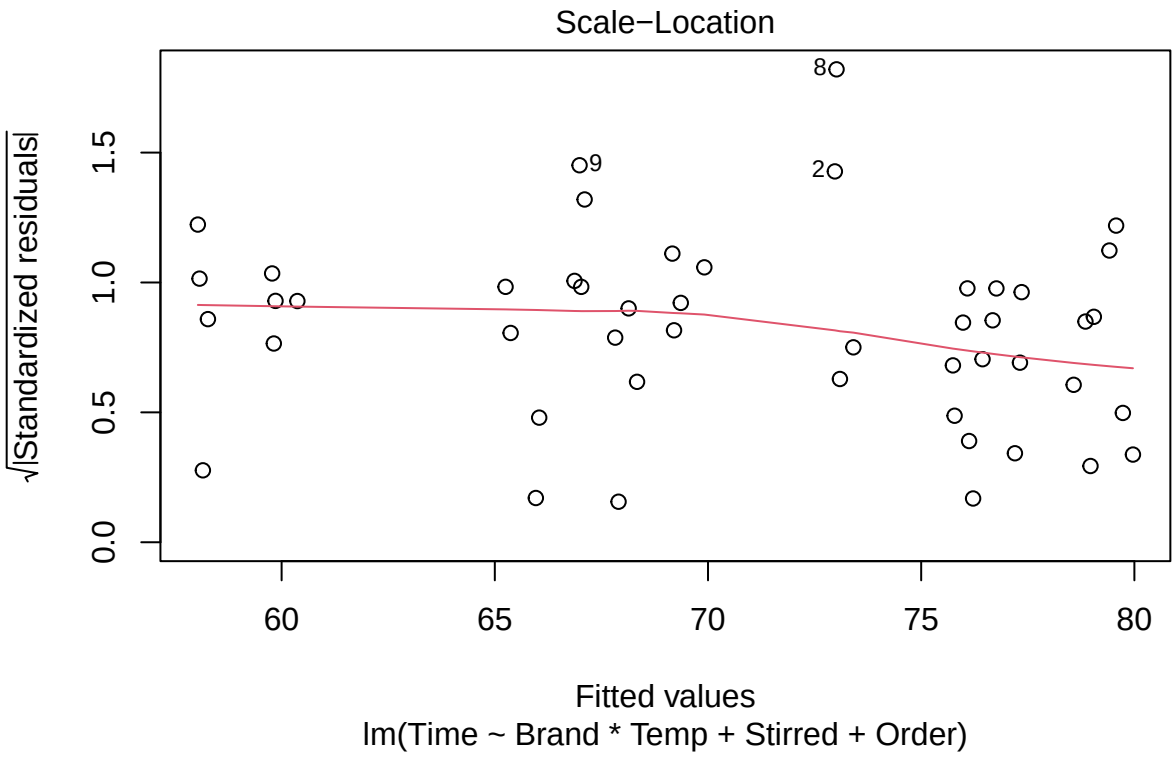
```
Anova(aov_block_order_eff, type=3) # type 3 SS
```

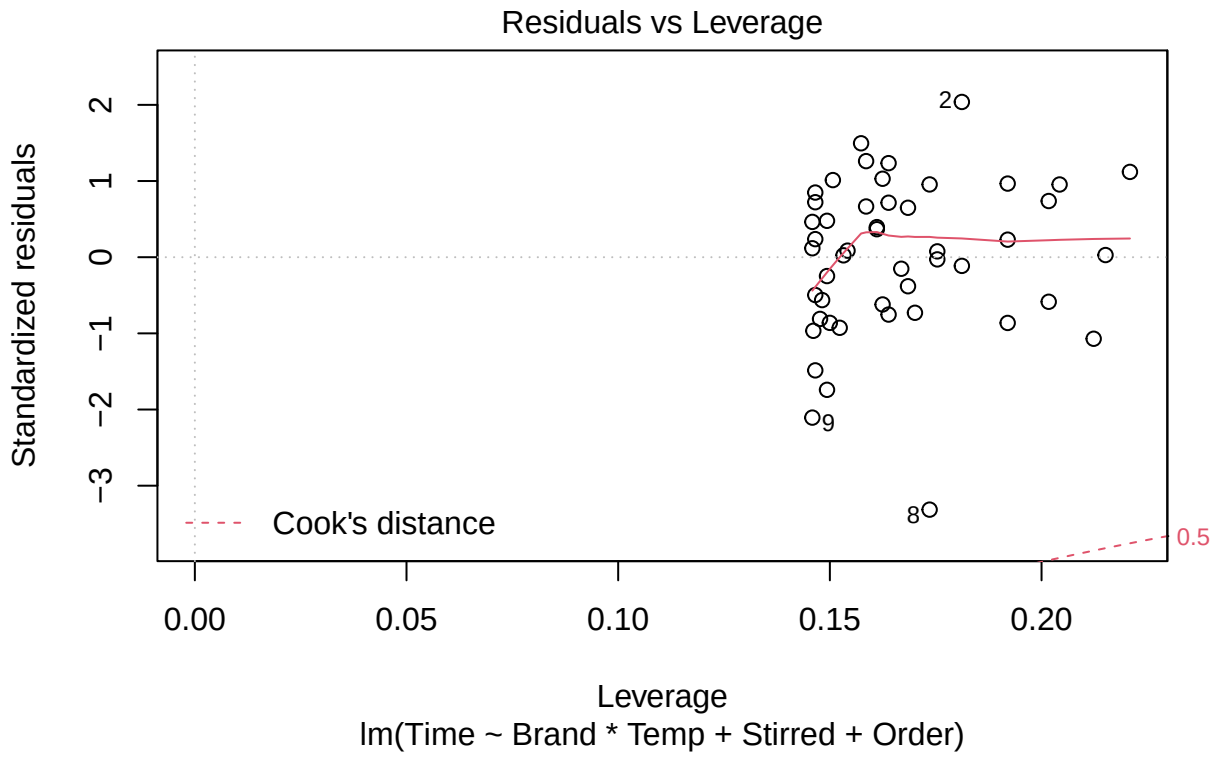
```
## Anova Table (Type III tests)
##
## Response: Time
##              Sum Sq Df    F value    Pr(>F)
## (Intercept) 22049.8  1 12987.9591 < 2.2e-16 ***
## Brand           3.6  1     2.1070   0.15442
## Temp          335.9  2    98.9237 3.275e-16 ***
## Stirred         6.2  1     3.6261   0.06409 .
## Order           3.4  1     1.9864   0.16645
## Brand:Temp    234.3  2    69.0102 1.076e-13 ***
## Residuals      67.9 40
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(lm_block_order_eff)
```

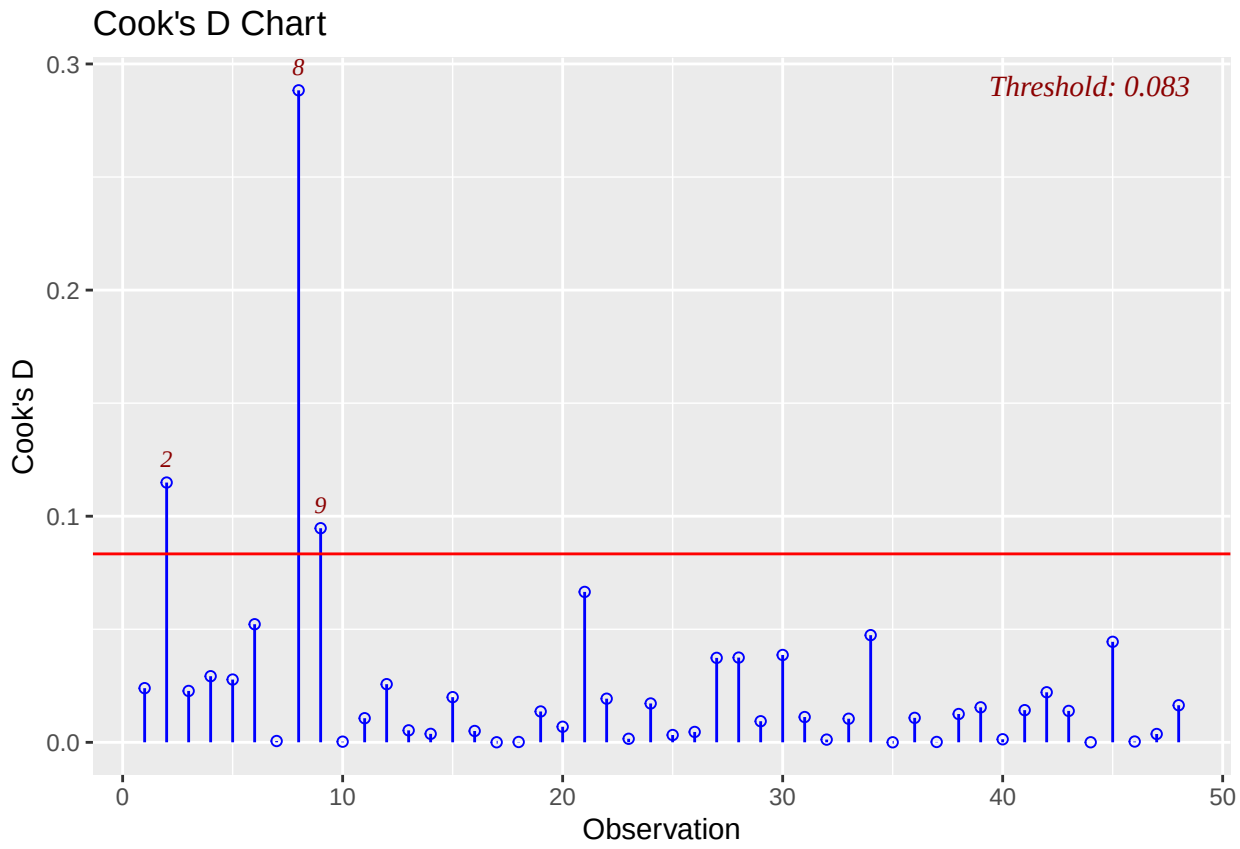



Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(Time ~ Brand * Temp + Stirred + Order)
2
9
8
−3
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
8
9
2
60
65
70
75
80
Fitted values
lm(Time ~ Brand * Temp + Stirred + Order)




```
ols_plot_cooksd_chart(lm_block_order_eff)
```




```
#added covariate Order to model with 3 factor interaction
aov_three_order_eff <- aov(lm_three_order_eff <- lm(Time ~ Brand * Temp*Stirred + Order, data = df_eff))
summary(lm_three_order_eff)
```

```
##
## Call:
## lm(formula = Time ~ Brand * Temp * Stirred + Order, data = df_eff)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6373 -0.3259  0.0077  0.4420  2.6391
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   76.3849     0.6759 113.009  < 2e-16 ***
## Brandstore                     1.0930     0.7740   1.412  0.16673
## Temp23                        -3.6070     0.7971  -4.525 6.67e-05 ***
## Temp40                       -10.4002     0.7768 -13.389 2.47e-15 ***
## Stirredno                      3.1339     1.0819   2.897  0.00646 **
## Order                         -0.0128     0.0282  -0.454  0.65270
## Brandstore:Temp23             -7.5228     1.1145  -6.750 8.06e-08 ***
## Brandstore:Temp40             -7.6899     1.1517  -6.677 1.00e-07 ***
## Brandstore:Stirredno          -0.6392     1.0870  -0.588  0.56026
## Temp23:Stirredno               0.9735     1.1047   0.881  0.38420
## Temp40:Stirredno               1.8751     1.0902   1.720  0.09427 .
## Brandstore:Temp23:Stirredno   -1.8430     1.5628  -1.179  0.24622
```

```
## Brandstore:Temp40:Stirredno  -4.4163     1.5906  -2.777  0.00876 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.087 on 35 degrees of freedom
## Multiple R-squared:  0.9826, Adjusted R-squared:  0.9766
## F-statistic: 164.3 on 12 and 35 DF,  p-value: < 2.2e-16
```

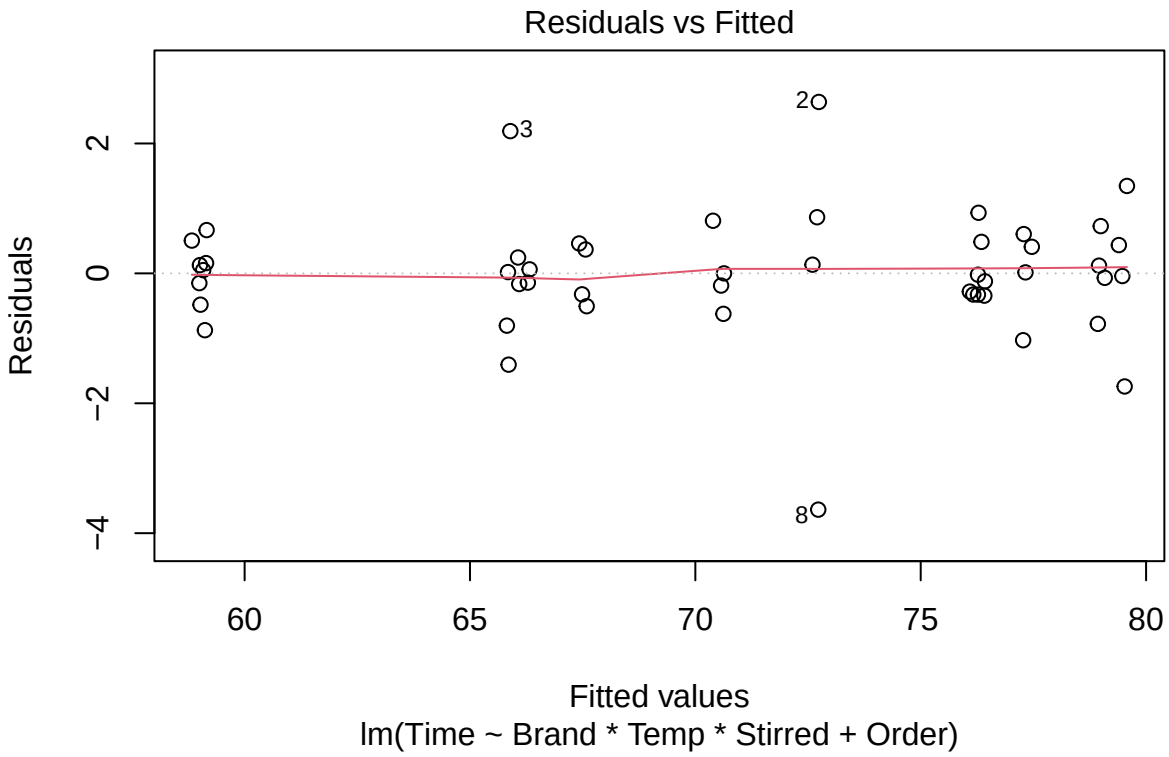
```
summary(aov_three_order_eff)
```

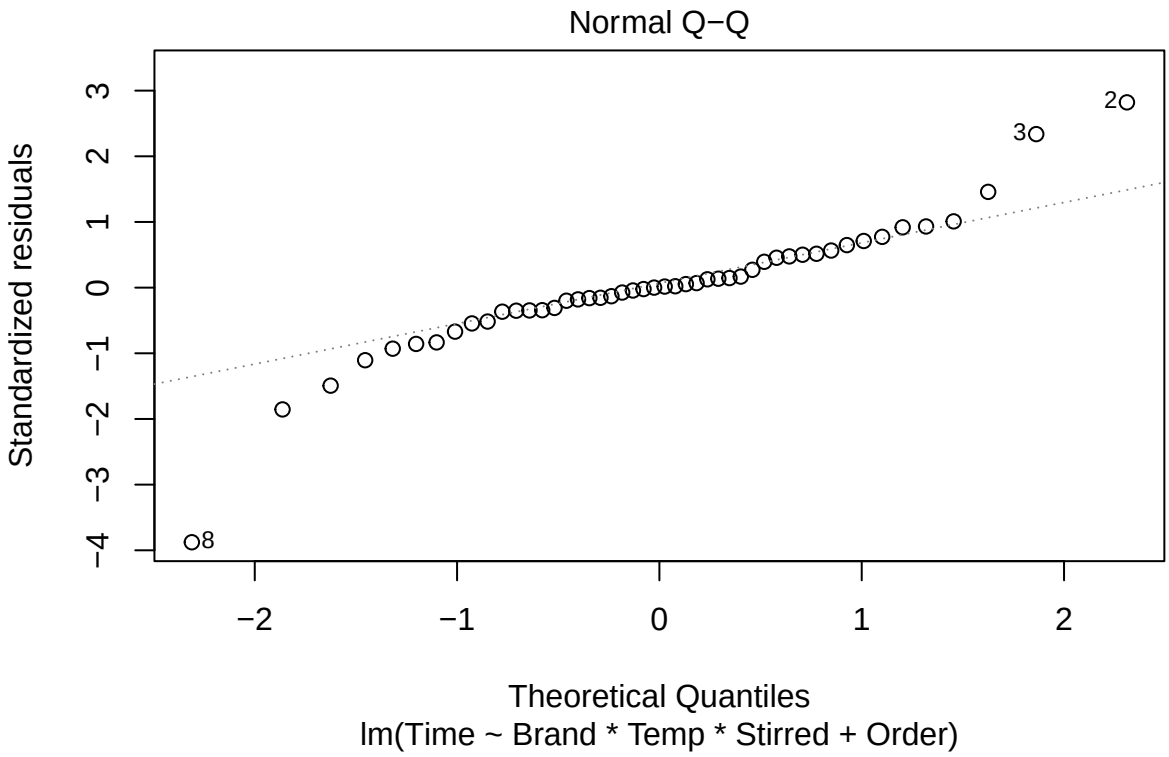
```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## Brand               1  342.0   342.0 289.512  < 2e-16 ***
## Temp                2 1654.7   827.4 700.374  < 2e-16 ***
## Stirred             1   69.9    69.9  59.161 5.01e-09 ***
## Order               1    0.9     0.9   0.767 0.387161
## Brand:Temp          2  234.3   117.2  99.176 3.81e-15 ***
## Brand:Stirred       1   17.3    17.3  14.641 0.000514 ***
## Temp:Stirred        2    0.0     0.0   0.018 0.982371
## Brand:Temp:Stirred  2    9.2     4.6   3.904 0.029469 *
## Residuals          35   41.3     1.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

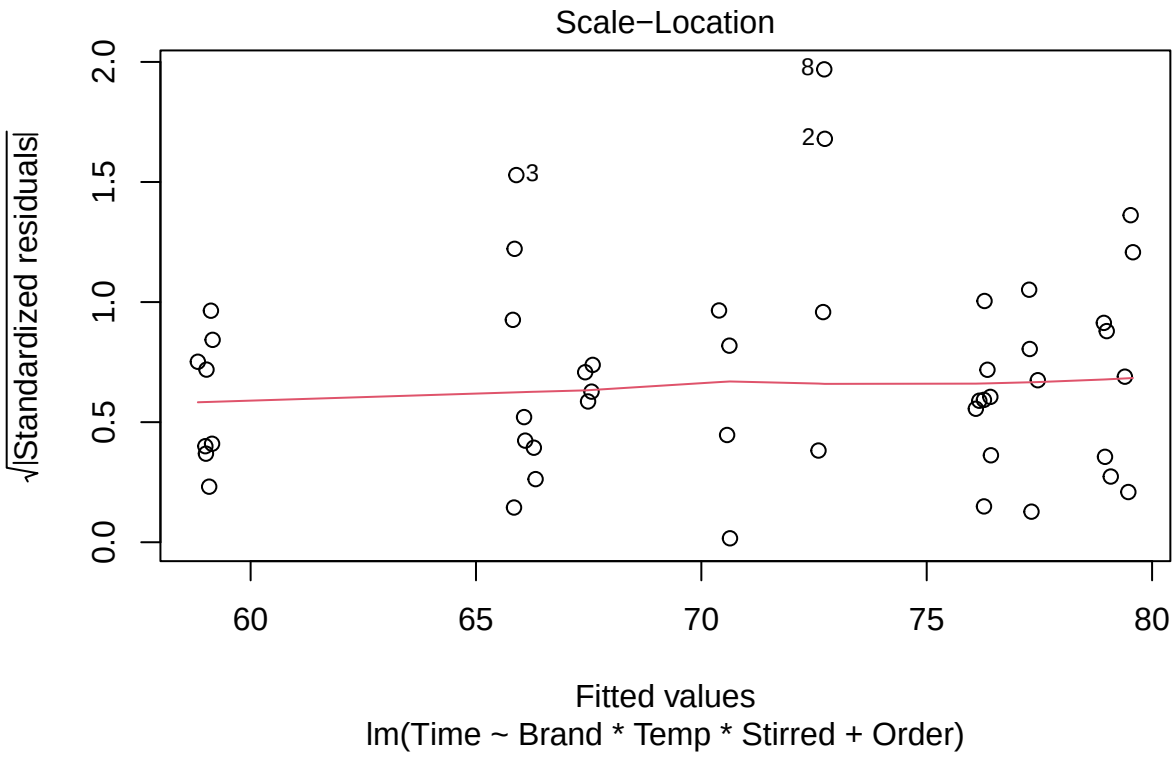
```
library(car)
Anova(aov_three_order_eff, type=3) # type 3 SS
```

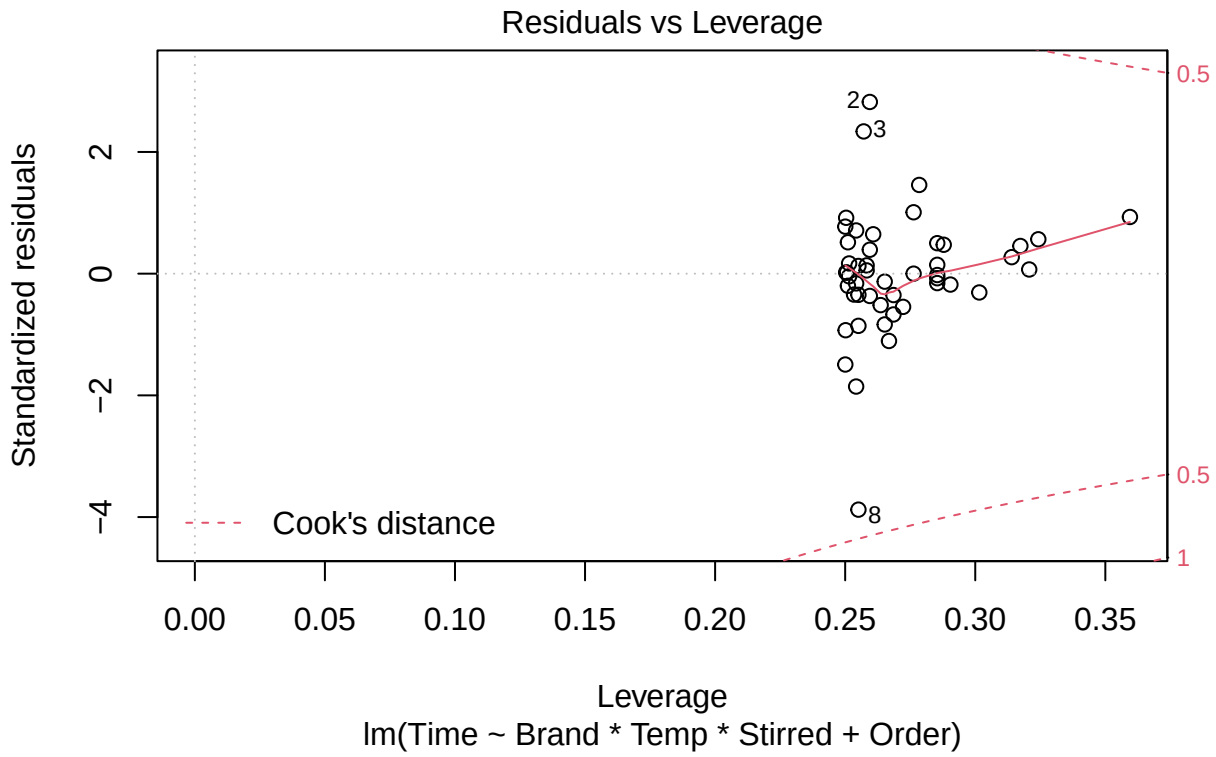
```
## Anova Table (Type III tests)
##
## Response: Time
##                      Sum Sq Df    F value    Pr(>F)
## (Intercept)        15086.6  1 12770.9606 < 2.2e-16 ***
## Brand                  2.4  1     1.9942  0.166731
## Temp                 220.9  2    93.4777 9.153e-15 ***
## Stirred                9.9  1     8.3899  0.006465 **
## Order                  0.2  1     0.2060  0.652697
## Brand:Temp            69.5  2    29.3957 3.224e-08 ***
## Brand:Stirred          0.4  1     0.3458  0.560255
## Temp:Stirred           3.5  2     1.4816  0.241180
## Brand:Temp:Stirred     9.2  2     3.9044  0.029469 *
## Residuals             41.3 35
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(lm_three_order_eff)
```


Residuals vs Fitted
Residuals
Fitted values
lm(Time ~ Brand * Temp * Stirred + Order)
2
3
8
60
65
70
75
80
−4
−2
0
2


Normal Q−Q
Standardized residuals
3
2
1
0
−1
−2
−3
−4
2
3
8
−2
−1
0
1
2
Theoretical Quantiles
lm(Time ~ Brand * Temp * Stirred + Order)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
8
2
3
60
65
70
75
80
Fitted values
lm(Time ~ Brand * Temp * Stirred + Order)




```
#adding Halid's code to see how full 3 factor interaction compares
aov_eff <- aov(lm_eff <- lm(Time ~ Brand * Temp * Stirred, data = df_eff))
summary(lm_eff)
```

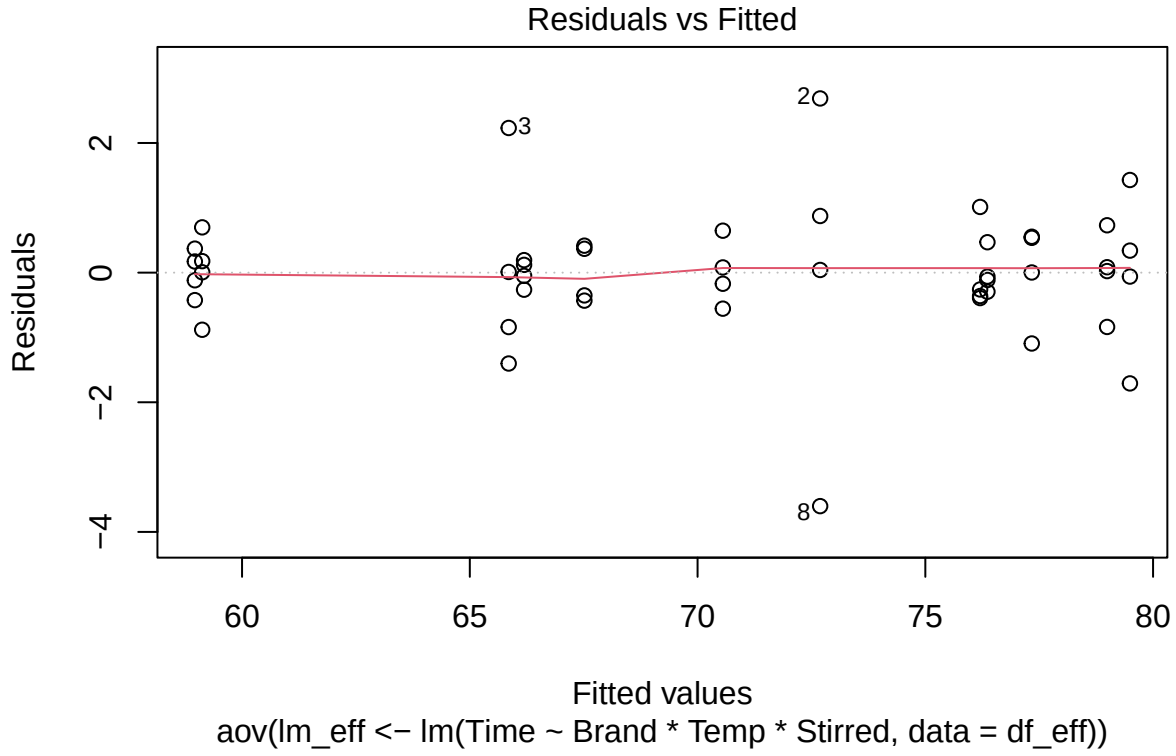
```
##
## Call:
## lm(formula = Time ~ Brand * Temp * Stirred, data = df_eff)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.6021 -0.3538  0.0077  0.3816  2.6871
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                   76.2024     0.5374 141.794  < 2e-16 ***
## Brandstore                     1.1346     0.7600   1.493 0.144185
## Temp23                        -3.5110     0.7600  -4.620 4.78e-05 ***
## Temp40                       -10.3490     0.7600 -13.617 9.05e-16 ***
## Stirredno                      2.7882     0.7600   3.669 0.000783 ***
## Brandstore:Temp23             -7.6348     1.0748  -7.103 2.40e-08 ***
## Brandstore:Temp40             -7.8628     1.0748  -7.315 1.27e-08 ***
## Brandstore:Stirredno          -0.6328     1.0748  -0.589 0.559694
## Temp23:Stirredno               0.8839     1.0748   0.822 0.416308
## Temp40:Stirredno               1.9135     1.0748   1.780 0.083476 .
## Brandstore:Temp23:Stirredno   -1.7150     1.5200  -1.128 0.266685
```

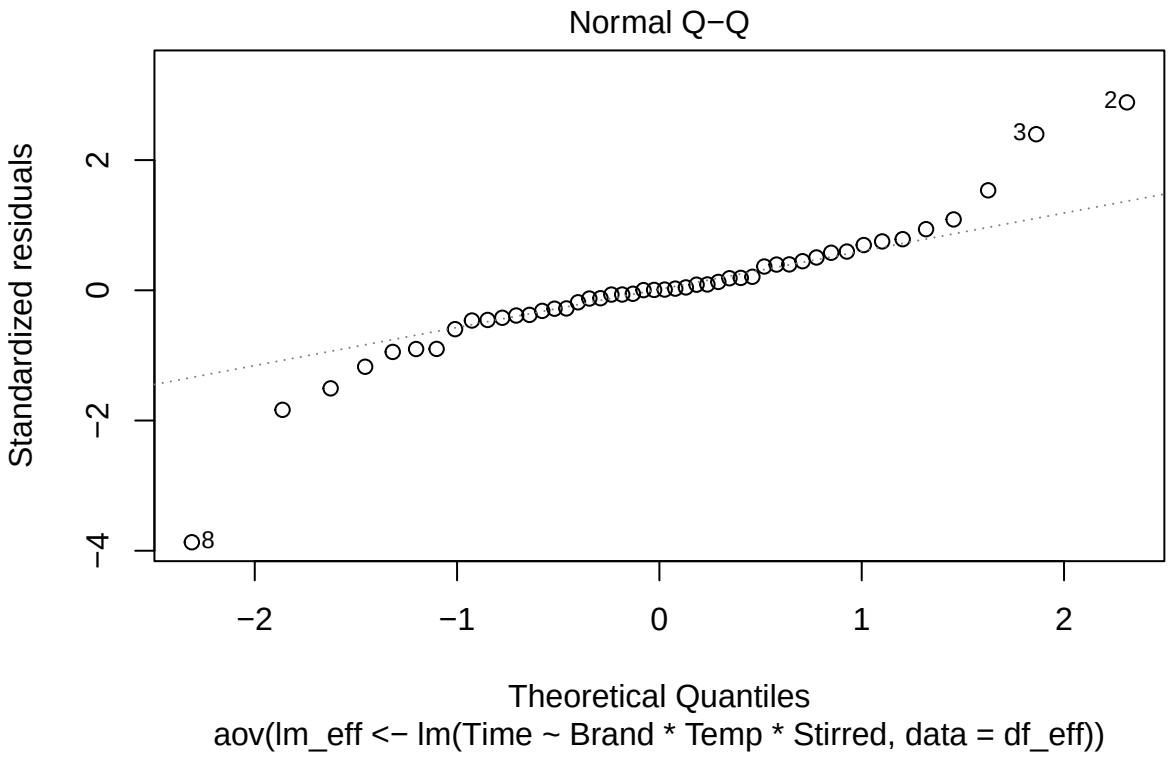
```
## Brandstore:Temp40:Stirredno  -4.2307     1.5200  -2.783 0.008519 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.075 on 36 degrees of freedom
## Multiple R-squared:  0.9824, Adjusted R-squared:  0.9771
## F-statistic: 183.2 on 11 and 36 DF,  p-value: < 2.2e-16
```

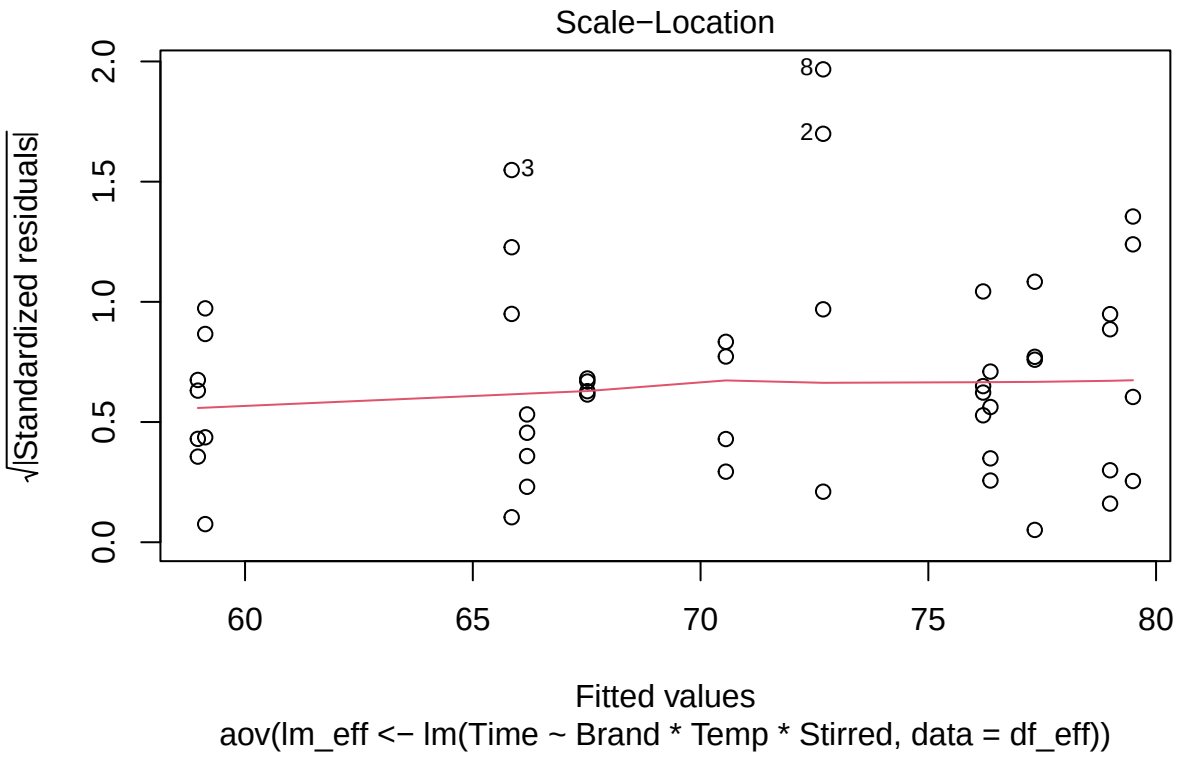
```
summary(aov_eff)
```

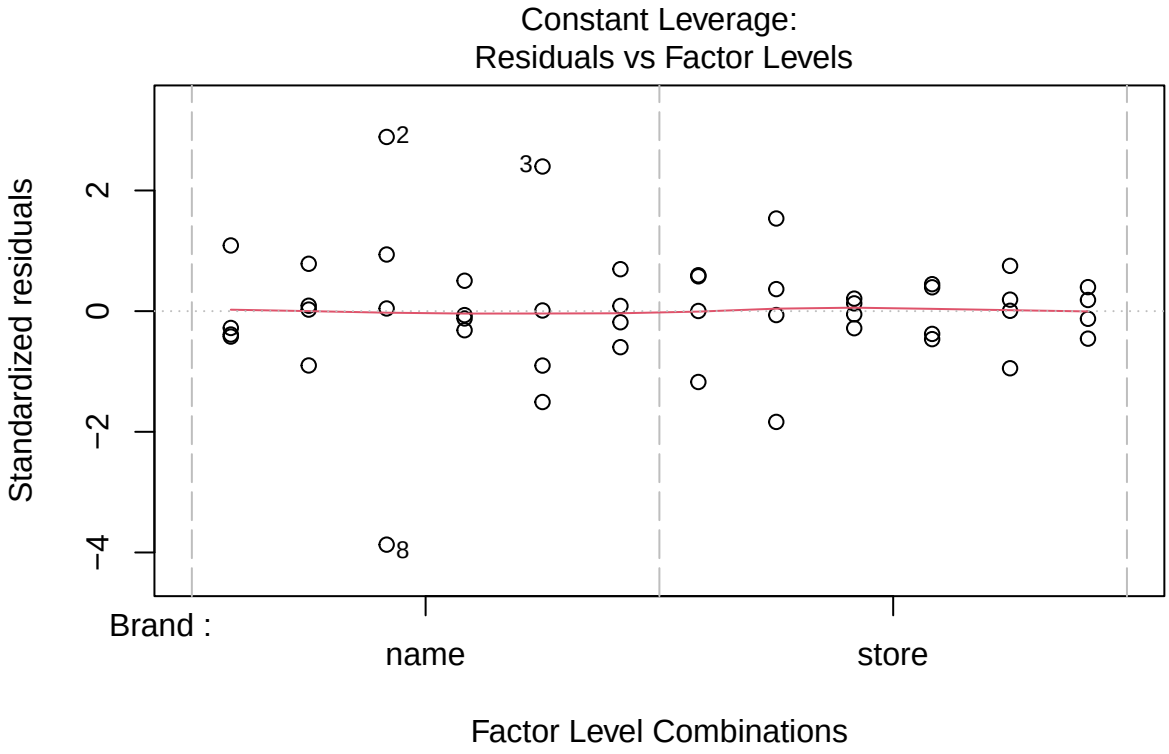
```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## Brand               1  342.0   342.0 296.041  < 2e-16 ***
## Temp                2 1654.7   827.4 716.169  < 2e-16 ***
## Stirred             1   69.9    69.9  60.495 3.22e-09 ***
## Brand:Temp          2  231.9   115.9 100.345 1.90e-15 ***
## Brand:Stirred       1   20.5    20.5  17.753 0.000161 ***
## Temp:Stirred        2    0.1     0.1   0.054 0.947535
## Brand:Temp:Stirred  2    9.1     4.5   3.919 0.028838 *
## Residuals          36   41.6     1.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
plot(aov_eff)
```

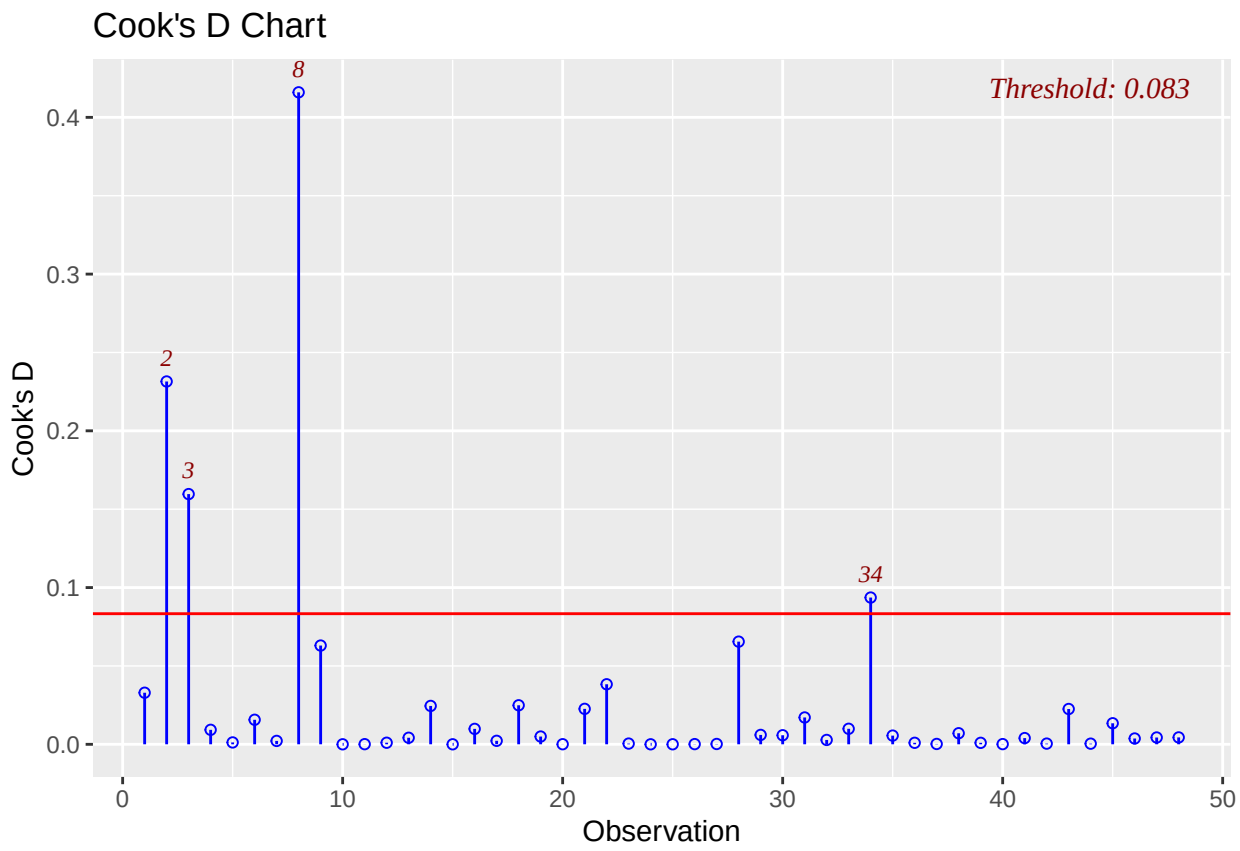



Normal Q−Q
Standardized residuals
Theoretical Quantiles
aov(lm_eff <− lm(Time ~ Brand * Temp * Stirred, data = df_eff))
2
3
8
−4
−2
0
2
−2
−1
0
1
2


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
8
2
3
60
65
70
75
80
Fitted values
aov(lm_eff <− lm(Time ~ Brand * Temp * Stirred, data = df_eff))




```
#adding Cook's D here
ols_plot_cooksd_chart(lm_eff)
```


Cook's D Chart
Threshold: 0.083
8
2
3
34
0.4
0.3
0.2
0.1
0.0
Cook's D
0
10
20
30
40
50
Observation