

# STATS 419 Final Project

Tyler Chun, Jaehee Jeong, Yolanda Jin, Harrison DiStefano

11/24/2021

## Library

```
library(DataExplorer)
library(MASS)
```

## Approach: Remove point 8 and Treat Speed as numeric and Box-cox

- Treat Speed as numeric
- Transformation of response

## 1. Import the data and clean the data set to fit into the model

```
df = read.csv("df.csv")

# Do initial Analysis and fit model
Response = c(df$Replicate.1,df$Replicate.2,df$Replicate.3)

# Change Speed to original speed numbers
df$Speed_f=as.factor(df$Speed)

df$Speed[df$Speed_f==1]=50
df$Speed[df$Speed_f==2]=30
df$Speed[df$Speed_f==3]=10
df$Speed[df$Speed_f==4]=2.5

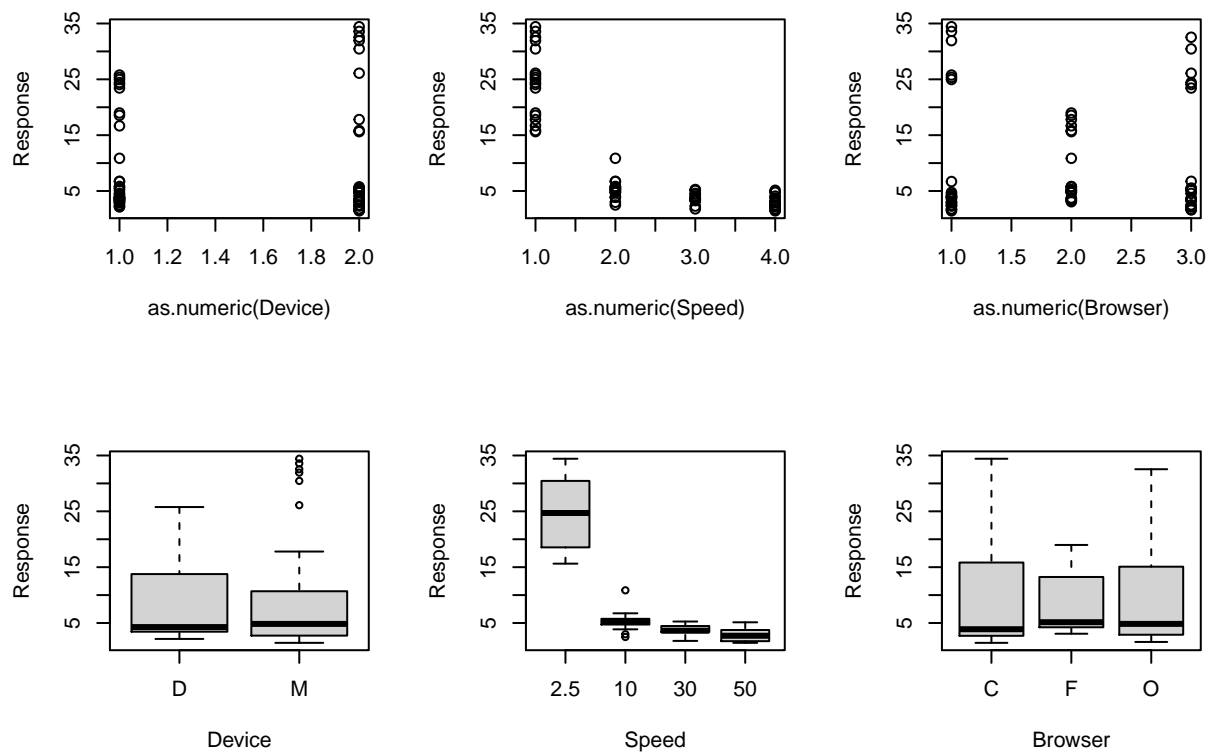
# Save variables as a factor
df$Speed = as.factor(df$Speed)
df$Device = as.factor(df$Device)
df$Browser = as.factor(df$Browser)

# Reshape the df
new_df =
data.frame(Device = rep(df$Device,3),
Speed = rep(df$Speed,3),
Browser = rep(df$Browser,3),
Response,
replicate = as.factor(c(rep(1,24),rep(2,24),rep(3,24))))
```

## 1.1 EDA

```
# Create eda html
# create_report(df)

# Create plots
par(mfrow=c(2,3))
plot(data= new_df, Response~ as.numeric(Device) + as.numeric(Speed)+as.numeric(Browser))
plot(data= new_df, Response~ (Device) + (Speed)+(Browser))
```



## 2. Fit into the model and Check the model with the model Assumptions and Tukey

```
# Create a linear model with two way interaction effects
m1 <-lm(data=new_df, Response~(Device + Speed + Browser)^2 + replicate)
summary(m1)
```

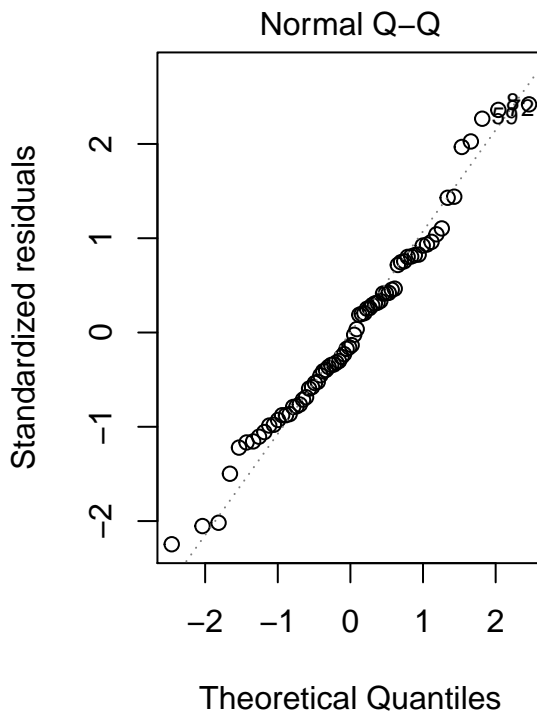
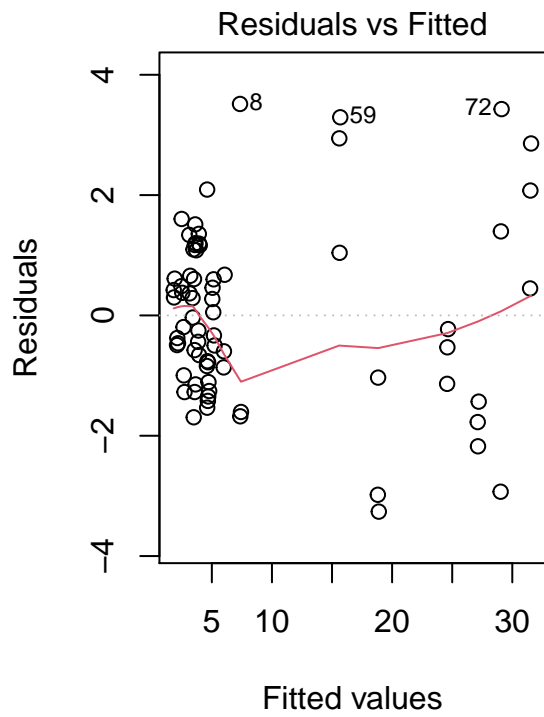
```
##
## Call:
## lm(formula = Response ~ (Device + Speed + Browser)^2 + replicate,
##     data = new_df)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.2615 -1.0528 -0.2109  1.0512  3.5149
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    27.12306    0.90071   30.113 < 2e-16 ***
## DeviceM         4.35889    0.98668    4.418 5.08e-05 ***
## Speed10        -22.51500    1.13932  -19.762 < 2e-16 ***
## Speed30        -24.00167    1.13932  -21.067 < 2e-16 ***
## Speed50        -24.65889    1.13932  -21.644 < 2e-16 ***
## BrowserF       -11.51625    1.10314  -10.440 2.30e-14 ***
## Browser0        -2.54542    1.10314   -2.307 0.025044 *
## replicate2      0.02250    0.49334    0.046 0.963798
## replicate3      0.08000    0.49334    0.162 0.871808
## DeviceM:Speed10 -5.40000    1.13932   -4.740 1.70e-05 ***
## DeviceM:Speed30 -3.98667    1.13932   -3.499 0.000966 ***
## DeviceM:Speed50 -4.17889    1.13932   -3.668 0.000576 ***
## DeviceM:BrowserF -1.15417    0.98668   -1.170 0.247436
## DeviceM:Browser0 0.09417    0.98668    0.095 0.924334
## Speed10:BrowserF 14.25333    1.39538   10.215 4.97e-14 ***
## Speed30:BrowserF 13.09167    1.39538    9.382 8.97e-13 ***
## Speed50:BrowserF 13.64333    1.39538    9.778 2.25e-13 ***
## Speed10:Browser0 3.92167    1.39538    2.810 0.006958 **
## Speed30:Browser0 2.81667    1.39538    2.019 0.048703 *
## Speed50:Browser0 1.91000    1.39538    1.369 0.176944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.709 on 52 degrees of freedom
## Multiple R-squared:  0.9765, Adjusted R-squared:  0.968
## F-statistic: 113.9 on 19 and 52 DF, p-value: < 2.2e-16
```

```
anova(m1)
```

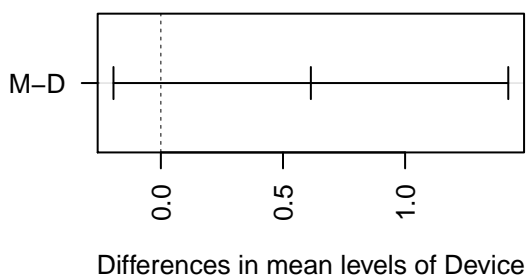
```
## Analysis of Variance Table
##
## Response: Response
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Device         1    6.8    6.79    2.3247 0.1333922
## Speed          3 5710.8 1903.59 651.7780 < 2.2e-16 ***
## Browser        2   46.4   23.21   7.9460 0.0009745 ***
## replicate      2    0.1    0.04   0.0140 0.9861143
## Device:Speed   3   74.3   24.77   8.4797 0.0001104 ***
## Device:Browser 2    5.8    2.90   0.9927 0.3774856
## Speed:Browser  6  474.9   79.15  27.1020 2.187e-14 ***
## Residuals     52  151.9    2.92
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,2))
plot(m1,c(1,2))
```

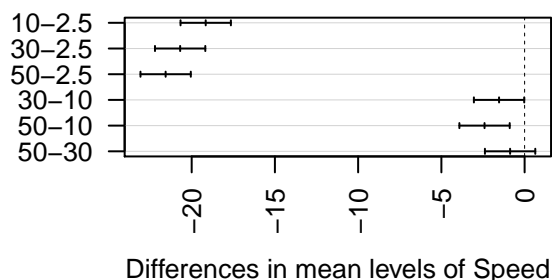


```
par(mfrow=c(2,2))
plot(TukeyHSD(aov(m1)),las=2)
```

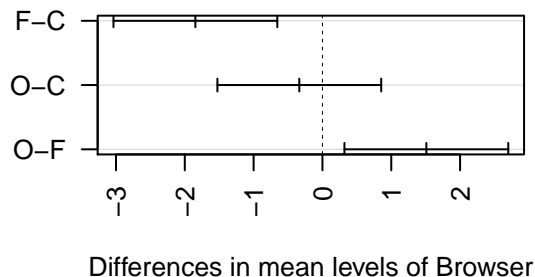
### 95% family-wise confidence level



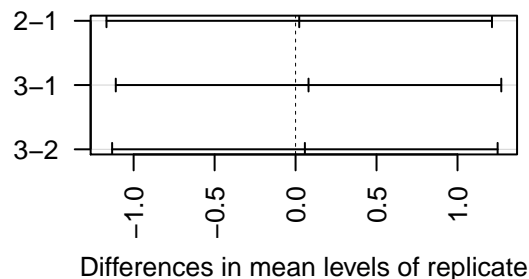
### 95% family-wise confidence level



### 95% family-wise confidence level



### 95% family-wise confidence level



```
TukeyHSD(aov(m1))
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = m1)
##
## $Device
##      diff      lwr      upr    p adj
## M-D 0.6141667 -0.1941313 1.422465 0.1333922
##
## $Speed
##      diff      lwr      upr    p adj
## 10-2.5 -19.1566667 -20.668600 -17.64473339 0.0000000
## 30-2.5 -20.6922222 -22.204155 -19.18028895 0.0000000
## 50-2.5 -21.5638889 -23.075822 -20.05195561 0.0000000
## 30-10 -1.5355556 -3.047489 -0.02362228 0.0452395
## 50-10 -2.4072222 -3.919155 -0.89528895 0.0005442
## 50-30 -0.8716667 -2.383600 0.64026661 0.4272319
##
## $Browser
##      diff      lwr      upr    p adj
## F-C -1.84625 -3.0364793 -0.6560207 0.0013069
## O-C -0.33625 -1.5264793 0.8539793 0.7752172
## O-F 1.51000 0.3197707 2.7002293 0.0096095
```

```

##
## $replicate
##      diff      lwr      upr      p adj
## 2-1 0.0225 -1.167729 1.212729 0.9988539
## 3-1 0.0800 -1.110229 1.270229 0.9856117
## 3-2 0.0575 -1.132729 1.247729 0.9925397
##
## $'Device:Speed'
##      diff      lwr      upr      p adj
## M:2.5-D:2.5  4.00555556  1.461810  6.5493015 0.0001942
## D:10-D:2.5 -16.45666667 -19.000413 -13.9129208 0.0000000
## M:10-D:2.5 -17.85111111 -20.394857 -15.3073652 0.0000000
## D:30-D:2.5 -18.69888889 -21.242635 -16.1551430 0.0000000
## M:30-D:2.5 -18.68000000 -21.223746 -16.1362541 0.0000000
## D:50-D:2.5 -19.47444444 -22.018190 -16.9306985 0.0000000
## M:50-D:2.5 -19.64777778 -22.191524 -17.1040319 0.0000000
## D:10-M:2.5 -20.46222222 -23.005968 -17.9184763 0.0000000
## M:10-M:2.5 -21.85666667 -24.400413 -19.3129208 0.0000000
## D:30-M:2.5 -22.70444444 -25.248190 -20.1606985 0.0000000
## M:30-M:2.5 -22.68555556 -25.229301 -20.1418096 0.0000000
## D:50-M:2.5 -23.48000000 -26.023746 -20.9362541 0.0000000
## M:50-M:2.5 -23.65333333 -26.197079 -21.1095874 0.0000000
## M:10-D:10 -1.39444444 -3.938190  1.1493015 0.6675788
## D:30-D:10 -2.24222222 -4.785968  0.3015237 0.1216793
## M:30-D:10 -2.22333333 -4.767079  0.3204126 0.1281090
## D:50-D:10 -3.01777778 -5.561524 -0.4740319 0.0099469
## M:50-D:10 -3.19111111 -5.734857 -0.6473652 0.0052260
## D:30-M:10 -0.84777778 -3.391524  1.6959681 0.9636007
## M:30-M:10 -0.82888889 -3.372635  1.7148570 0.9677688
## D:50-M:10 -1.62333333 -4.167079  0.9204126 0.4821241
## M:50-M:10 -1.79666667 -4.340413  0.7470792 0.3514285
## M:30-D:30  0.01888889 -2.524857  2.5626348 1.0000000
## D:50-D:30 -0.77555556 -3.319301  1.7681904 0.9776717
## M:50-D:30 -0.94888889 -3.492635  1.5948570 0.9346661
## D:50-M:30 -0.79444444 -3.338190  1.7493015 0.9744656
## M:50-M:30 -0.96777778 -3.511524  1.5759681 0.9279356
## M:50-D:50 -0.17333333 -2.717079  2.3704126 0.9999988
##
## $'Device:Browser'
##      diff      lwr      upr      p adj
## M:C-D:C  0.9675000 -1.09667912  3.0316791 0.7346401
## D:F-D:C -1.2691667 -3.33334578  0.7950124 0.4628240
## M:F-D:C -1.4558333 -3.52001245  0.6083458 0.3101399
## D:O-D:C -0.3833333 -2.44751245  1.6808458 0.9937241
## M:O-D:C  0.6783333 -1.38584578  2.7425124 0.9246500
## D:F-M:C -2.2366667 -4.30084578 -0.1724876 0.0264498
## M:F-M:C -2.4233333 -4.48751245 -0.3591542 0.0126633
## D:O-M:C -1.3508333 -3.41501245  0.7133458 0.3925184
## M:O-M:C -0.2891667 -2.35334578  1.7750124 0.9983437
## M:F-D:F -0.1866667 -2.25084578  1.8775124 0.9998028
## D:O-D:F  0.8858333 -1.17834578  2.9500124 0.7998618
## M:O-D:F  1.9475000 -0.11667912  4.0116791 0.0748945
## D:O-M:F  1.0725000 -0.99167912  3.1366791 0.6423285
## M:O-M:F  2.1341667  0.06998755  4.1983458 0.0388256

```

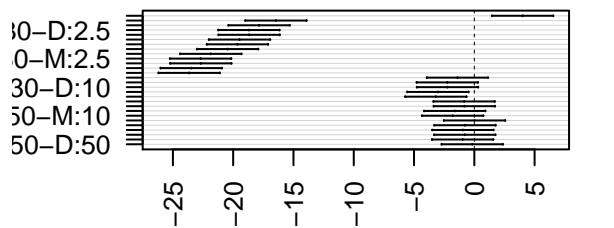
```

## M:0-D:0 1.0616667 -1.00251245 3.1258458 0.6521451
##
## $'Speed:Browser'
##          diff          lwr          upr          p adj
## 10:C-2.5:C -25.21500000 -28.5902039 -21.8397961 0.0000000
## 30:C-2.5:C -25.99500000 -29.3702039 -22.6197961 0.0000000
## 50:C-2.5:C -26.74833333 -30.1235372 -23.3731295 0.0000000
## 2.5:F-2.5:C -12.09333333 -15.4685372 -8.7181295 0.0000000
## 10:F-2.5:C -23.05500000 -26.4302039 -19.6797961 0.0000000
## 30:F-2.5:C -24.99666667 -28.3718705 -21.6214628 0.0000000
## 50:F-2.5:C -25.19833333 -28.5735372 -21.8231295 0.0000000
## 2.5:0-2.5:C -2.49833333 -5.8735372 0.8768705 0.3440631
## 10:0-2.5:C -23.79166667 -27.1668705 -20.4164628 0.0000000
## 30:0-2.5:C -25.67666667 -29.0518705 -22.3014628 0.0000000
## 50:0-2.5:C -27.33666667 -30.7118705 -23.9614628 0.0000000
## 30:C-10:C -0.78000000 -4.1552039 2.5952039 0.9996548
## 50:C-10:C -1.53333333 -4.9085372 1.8418705 0.9173533
## 2.5:F-10:C 13.12166667 9.7464628 16.4968705 0.0000000
## 10:F-10:C 2.16000000 -1.2152039 5.5352039 0.5644195
## 30:F-10:C 0.21833333 -3.1568705 3.5935372 1.0000000
## 50:F-10:C 0.01666667 -3.3585372 3.3918705 1.0000000
## 2.5:0-10:C 22.71666667 19.3414628 26.0918705 0.0000000
## 10:0-10:C 1.42333333 -1.9518705 4.7985372 0.9489270
## 30:0-10:C -0.46166667 -3.8368705 2.9135372 0.9999982
## 50:0-10:C -2.12166667 -5.4968705 1.2535372 0.5908681
## 50:C-30:C -0.75333333 -4.1285372 2.6218705 0.9997523
## 2.5:F-30:C 13.90166667 10.5264628 17.2768705 0.0000000
## 10:F-30:C 2.94000000 -0.4352039 6.3152039 0.1442546
## 30:F-30:C 0.99833333 -2.3768705 4.3735372 0.9967405
## 50:F-30:C 0.79666667 -2.5785372 4.1718705 0.9995784
## 2.5:0-30:C 23.49666667 20.1214628 26.8718705 0.0000000
## 10:0-30:C 2.20333333 -1.1718705 5.5785372 0.5345643
## 30:0-30:C 0.31833333 -3.0568705 3.6935372 1.0000000
## 50:0-30:C -1.34166667 -4.7168705 2.0335372 0.9660606
## 2.5:F-50:C 14.65500000 11.2797961 18.0302039 0.0000000
## 10:F-50:C 3.69333333 0.3181295 7.0685372 0.0208243
## 30:F-50:C 1.75166667 -1.6235372 5.1268705 0.8234652
## 50:F-50:C 1.55000000 -1.8252039 4.9252039 0.9116590
## 2.5:0-50:C 24.25000000 20.8747961 27.6252039 0.0000000
## 10:0-50:C 2.95666667 -0.4185372 6.3318705 0.1389905
## 30:0-50:C 1.07166667 -2.3035372 4.4468705 0.9940501
## 50:0-50:C -0.58833333 -3.9635372 2.7868705 0.9999784
## 10:F-2.5:F -10.96166667 -14.3368705 -7.5864628 0.0000000
## 30:F-2.5:F -12.90333333 -16.2785372 -9.5281295 0.0000000
## 50:F-2.5:F -13.10500000 -16.4802039 -9.7297961 0.0000000
## 2.5:0-2.5:F 9.59500000 6.2197961 12.9702039 0.0000000
## 10:0-2.5:F -11.69833333 -15.0735372 -8.3231295 0.0000000
## 30:0-2.5:F -13.58333333 -16.9585372 -10.2081295 0.0000000
## 50:0-2.5:F -15.24333333 -18.6185372 -11.8681295 0.0000000
## 30:F-10:F -1.94166667 -5.3168705 1.4335372 0.7118160
## 50:F-10:F -2.14333333 -5.5185372 1.2318705 0.5759221
## 2.5:0-10:F 20.55666667 17.1814628 23.9318705 0.0000000
## 10:0-10:F -0.73666667 -4.1118705 2.6385372 0.9998003
## 30:0-10:F -2.62166667 -5.9968705 0.7535372 0.2764044

```

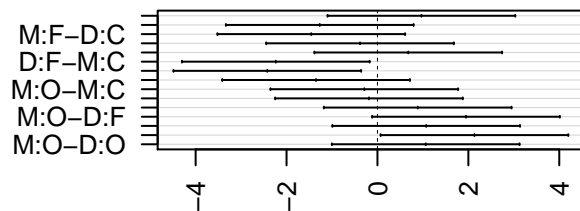
## 50:0-10:F	-4.28166667	-7.6568705	-0.9064628	0.0034716
## 50:F-30:F	-0.20166667	-3.5768705	3.1735372	1.0000000
## 2.5:0-30:F	22.49833333	19.1231295	25.8735372	0.0000000
## 10:0-30:F	1.20500000	-2.1702039	4.5802039	0.9847214
## 30:0-30:F	-0.68000000	-4.0552039	2.6952039	0.9999084
## 50:0-30:F	-2.34000000	-5.7152039	1.0352039	0.4424759
## 2.5:0-50:F	22.70000000	19.3247961	26.0752039	0.0000000
## 10:0-50:F	1.40666667	-1.9685372	4.7818705	0.9528368
## 30:0-50:F	-0.47833333	-3.8535372	2.8968705	0.9999974
## 50:0-50:F	-2.13833333	-5.5135372	1.2368705	0.5793725
## 10:0-2.5:0	-21.29333333	-24.6685372	-17.9181295	0.0000000
## 30:0-2.5:0	-23.17833333	-26.5535372	-19.8031295	0.0000000
## 50:0-2.5:0	-24.83833333	-28.2135372	-21.4631295	0.0000000
## 30:0-10:0	-1.88500000	-5.2602039	1.4902039	0.7474358
## 50:0-10:0	-3.54500000	-6.9202039	-0.1697961	0.0316242
## 50:0-30:0	-1.66000000	-5.0352039	1.7152039	0.8679064

### 95% family-wise confidence level



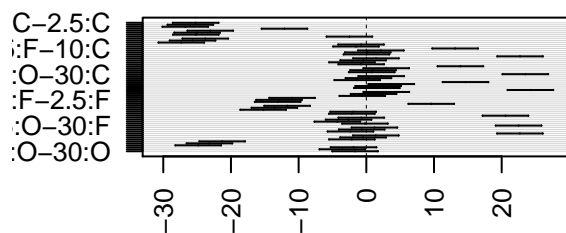
Differences in mean levels of Device:Speed

### 95% family-wise confidence level



Differences in mean levels of Device:Browser

### 95% family-wise confidence level



Differences in mean levels of Speed:Browser

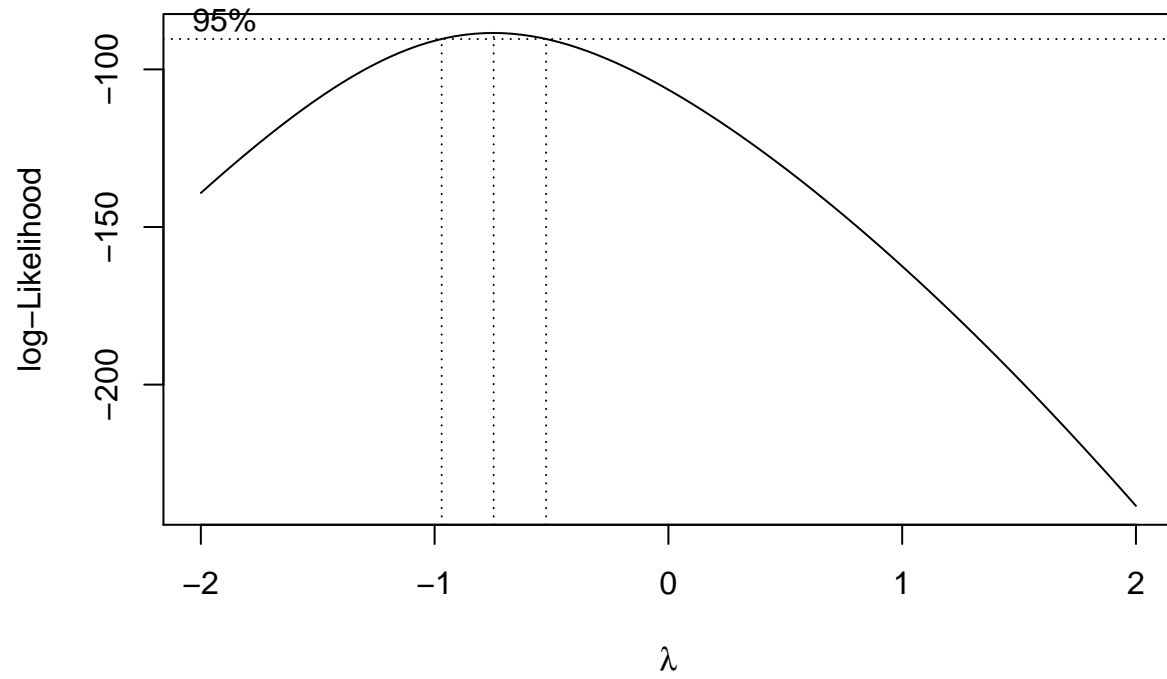
## 3. Boxcox Transformation

```
# Delete outlier
new_df_8_deleted <- new_df[-8,]

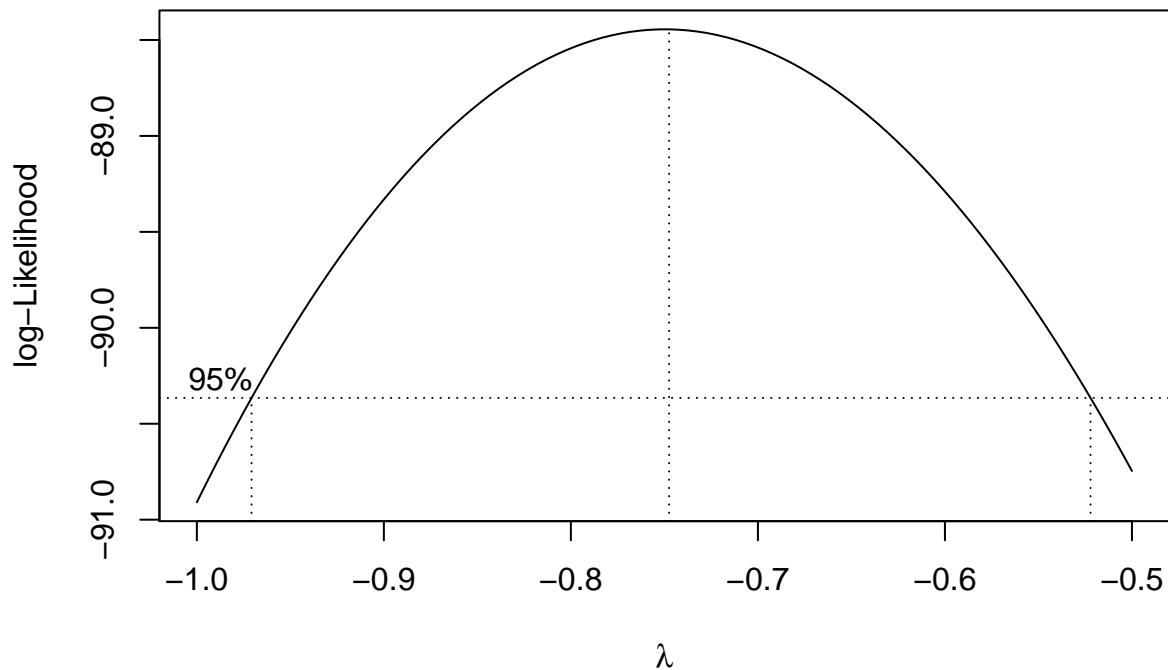
new_df_8_deleted$Speed <- as.numeric(as.character(new_df_8_deleted$Speed))
m1.1 <- lm(data=new_df_8_deleted, Response~(Device + Speed + Browser)^2 + replicate)
```



```
# To see the exact best lambda:  
lmod1_bc <- boxcox(m1.1)
```



```
# Zoom in to select an easier to interpret lambda  
boxcox(m1.1, plotit=T, lambda = seq(-1, -0.5, by=0.1))
```



```
lambda <- lmod1_bc$x[which.max(lmod1_bc$y)]
```

```
# lambda is -0.747474747474747  
paste0("lambda is ",lambda)
```

```
## [1] "lambda is -0.747474747474747"
```

```
# We will select a cleaner lambda for inteprétation as -0.7
```

```
lambda <- -0.7  
new_df_8_deleted['Response_t'] <- (new_df_8_deleted$Response^lambda - 1) / lambda
```

### 3.1 Fit the model after transformation and

```
m2 <-lm(data=new_df_8_deleted, Response_t~(Device + Speed + Browser)^2 + replicate)  
summary(m2)
```

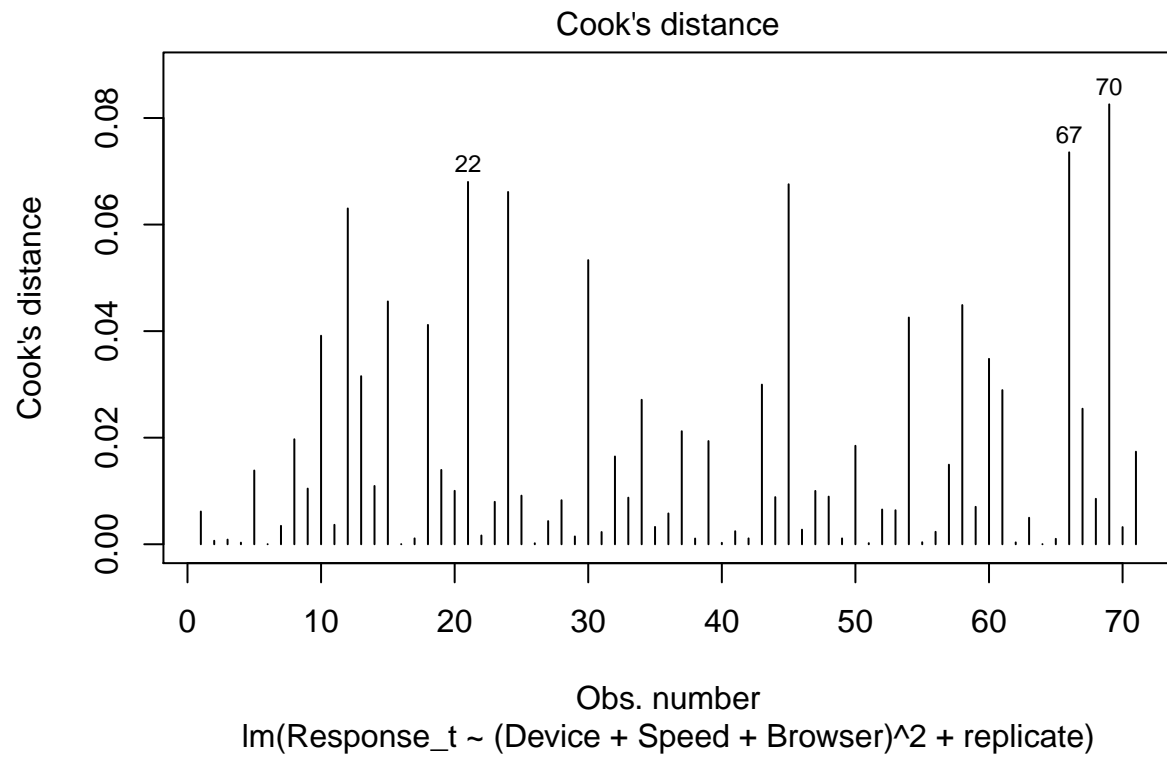
```
##  
## Call:  
## lm(formula = Response_t ~ (Device + Speed + Browser)^2 + replicate,  
##     data = new_df_8_deleted)  
##  
## Residuals:
```

```
##           Min           1Q       Median           3Q           Max
## -0.285818 -0.079701 -0.004034  0.085380  0.253185
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.216009   0.059070  20.586 < 2e-16 ***
## DeviceM        -0.112954   0.066623  -1.695  0.09527 .
## Speed          -0.010304   0.001686  -6.111 8.45e-08 ***
## BrowserF       -0.116709   0.074017  -1.577  0.12019
## Browser0        0.035285   0.072006   0.490  0.62593
## replicate2      0.005479   0.038598   0.142  0.88761
## replicate3     -0.016579   0.038598  -0.430  0.66909
## DeviceM:Speed   -0.002127   0.001691  -1.257  0.21355
## DeviceM:BrowserF 0.218293   0.077195   2.828  0.00639 **
## DeviceM:Browser0 0.103918   0.076278   1.362  0.17826
## Speed:BrowserF   0.005190   0.002075   2.500  0.01520 *
## Speed:Browser0  -0.002781   0.002063  -1.348  0.18275
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1321 on 59 degrees of freedom
## Multiple R-squared:  0.7677, Adjusted R-squared:  0.7244
## F-statistic: 17.73 on 11 and 59 DF, p-value: 7.403e-15
```

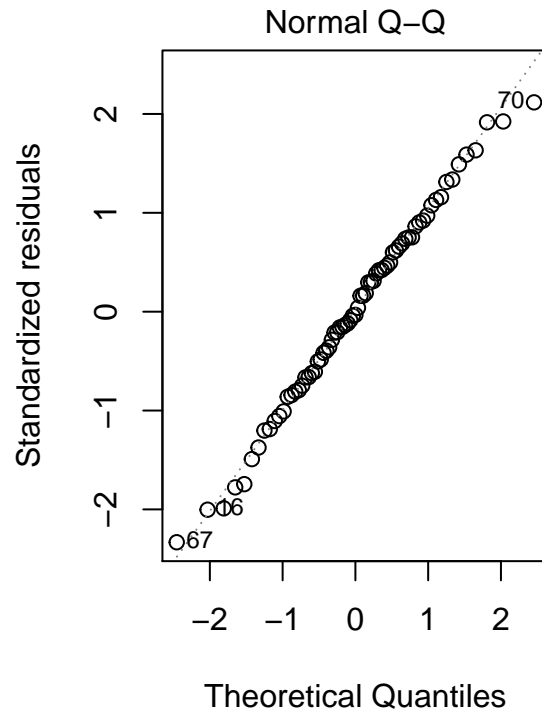
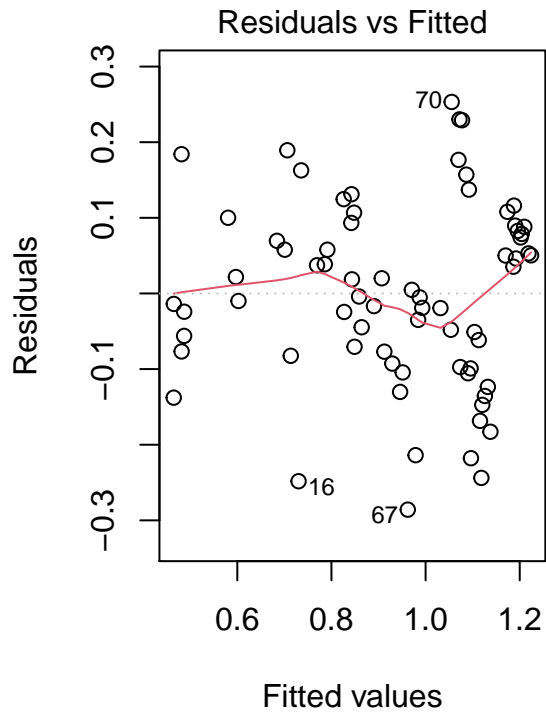
```
anova(m2)
```

```
## Analysis of Variance Table
##
## Response: Response_t
##           Df Sum Sq Mean Sq F value    Pr(>F)
## Device      1 0.04866  0.04866    2.7880  0.100273
## Speed       1 2.74277  2.74277  157.1355 < 2.2e-16 ***
## Browser     2 0.18243  0.09121    5.2257  0.008138 **
## replicate   2 0.00784  0.00392    0.2247  0.799430
## Device:Speed 1 0.02425  0.02425    1.3893  0.243250
## Device:Browser 2 0.13270  0.06635    3.8013  0.027999 *
## Speed:Browser 2 0.26472  0.13236    7.5831  0.001172 **
## Residuals   59 1.02983  0.01745
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

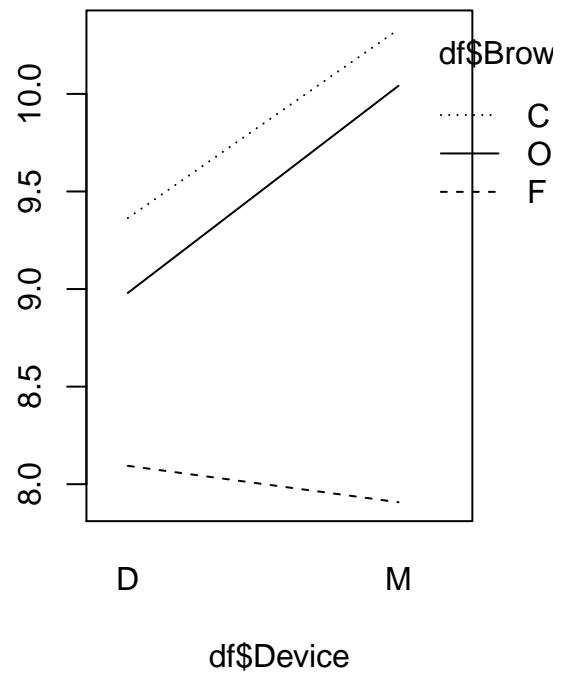
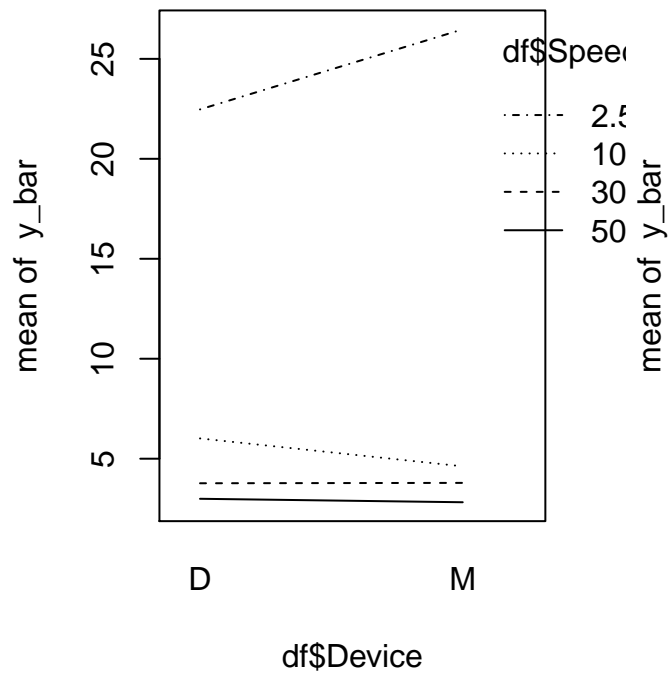
```
# Cook's distance
par(mfrow=c(1,1))
plot(m2,4)
```



```
# Residual vs Fitted  
par(mfrow=c(1,2))  
plot(m2,c(1,2))
```



```
# Interaction plots
y_bar = (df$Replicate.1 + df$Replicate.2 + df$Replicate.3)/3
interaction.plot(df$Device, df$Speed, y_bar)
interaction.plot(df$Device, df$Browser, y_bar)
```



```
interaction.plot(df$Browser, df$Speed, y_bar)
```

