



Advanced Analytics Assignment 1

Team Members:

Mohammed Hussain – 14551

Revanth Guthala – 14333

Rohit Polavarapu – 14654

Sanjeev Kumar Mishra – 14558

Date of submission: 22-Feb-2018

Advanced Analytics Assignment 1 - ANOVA

Mohammed Hussain, Revanth Guthala, Rohit Polavarapu, Sanjeev Kumar Mishra

Reading the Spyder_Strength dataset

```
given_data<-  
read.table("C:/Users/Admin/Downloads/MG226/assignment/spyder_strength.data",  
header = TRUE)
```

We are initializing a function here that will perform 4 different Normality Tests on the given input and help us determine whether the given input follows Normality or not.

```
normtest<-function(x)  
{  
  s<-shapiro.test(x)  
  ad<-ad.test(x)  
  cvm<-cvm.test(x)  
  ll<-lillie.test(x)  
  sf<-sf.test(x)  
  df<-data.frame(Method=c(s$method, ad$method, cvm$method, ll$method,  
sf$method),  
                  P.Value=c(s$p.value, ad$p.value, cvm$p.value, ll$p.value,  
sf$p.value))  
  df  
}
```

We convert the numerical values into categorical values and distribute the data into different data types.

```
y<-given_data$Torque  
hole_type <- as.factor(given_data$Hole)  
assem_type <- as.factor(given_data$Assembly)  
hub_type <- as.factor(given_data$Hub)  
#torque_reqd <- as.factor(given_data$Torque)  
treatment<-NULL  
for(i in 1:64)  
{  
  if ((hole_type[i]==-1) & (assem_type[i]==-1) & (hub_type[i]==-1))  
treatment[i]<-"combi_111"  
  if ((hole_type[i]==-1) & (assem_type[i]==-1) & (hub_type[i]==1))  
treatment[i]<-"combi_112"  
  if ((hole_type[i]==-1) & (assem_type[i]==1) & (hub_type[i]==-1))  
treatment[i]<-"combi_121"  
  if ((hole_type[i]==-1) & (assem_type[i]==1) & (hub_type[i]==1))  
treatment[i]<-"combi_122"  
  if ((hole_type[i]==1) & (assem_type[i]==-1) & (hub_type[i]==-1))  
treatment[i]<-"combi_211"  
  if ((hole_type[i]==1) & (assem_type[i]==-1) & (hub_type[i]==1))
```

```

treatment[i]<-"combi_212"
  if ((hole_type[i]==1) & (assem_type[i]==1) & (hub_type[i]==-1))
treatment[i]<-"combi_221"
  if ((hole_type[i]==1) & (assem_type[i]==1) & (hub_type[i]==1))
treatment[i]<-"combi_222"
}
treatment<-as.factor(treatment)

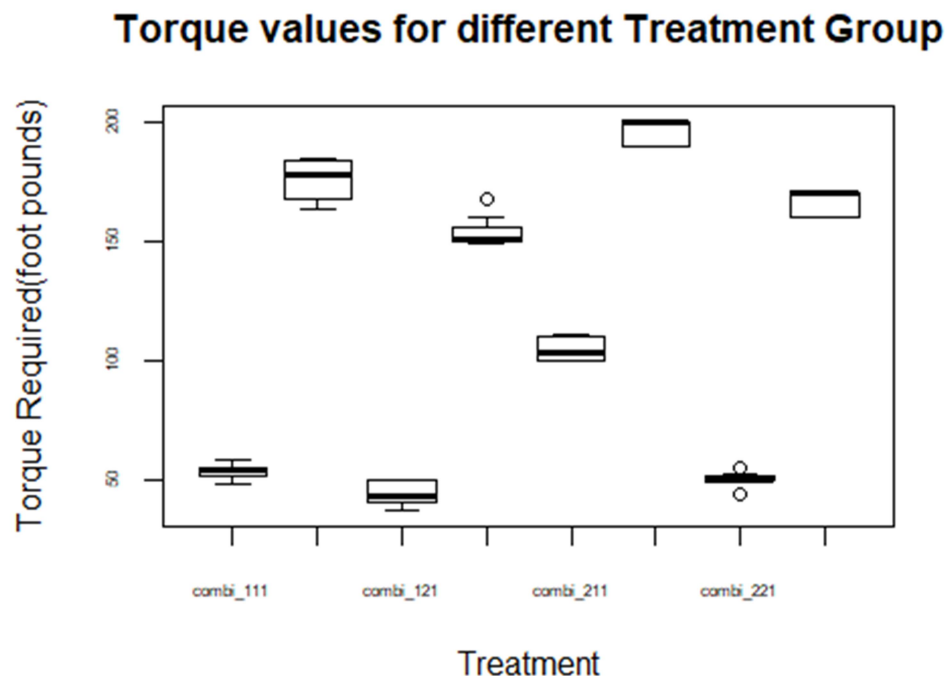
```

Creating a Boxplot to visualize the values of torque for various treatments.

```

boxplot(y~treatment,font.main=1,xlab="Treatment", ylab="Torque Required(foot
pounds)", main = "Torque values for different Treatment Group", cex.axis =
0.50)

```



It can be observed from the boxplot that various combinations of Hole type, Assembly Type and Hub type require significantly different ranges of torque for destructive testing.

One Way ANOVA

H₀: No difference in means of the torque values required for the various treatments

H_a: There is a significant difference in the torque values required over the various treatments

```

anova(aov(y~treatment))

```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## treatment  7 219653    31379  1013.3 < 2.2e-16 ***
## Residuals 56   1734         31
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The high F value obtained from the One Way ANOVA test signifies variation in the Mean Torques for various treatments.

Three Factor ANOVA

```
model=aov(y~hole_type*assem_type*hub_type)
anova(model)
```

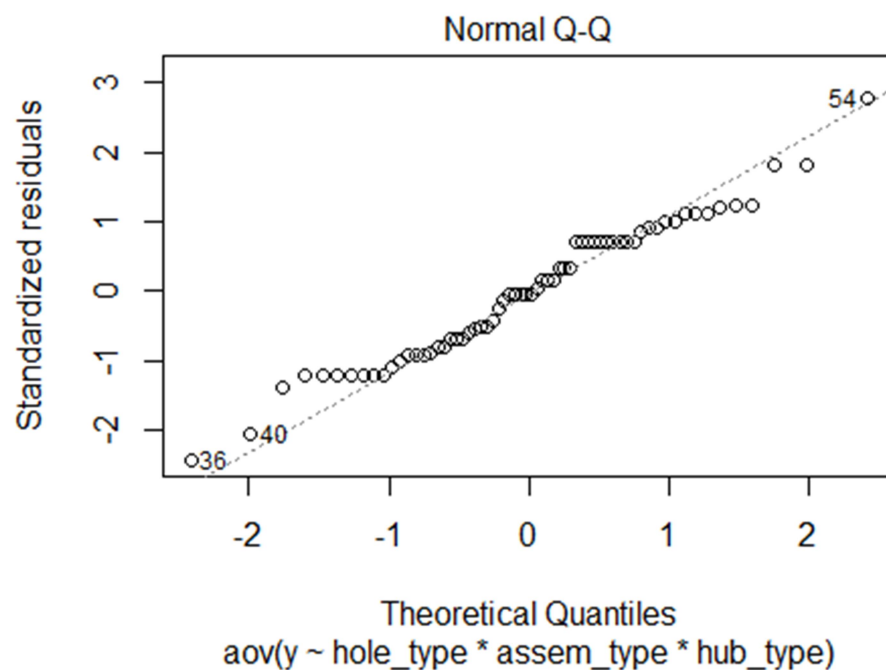
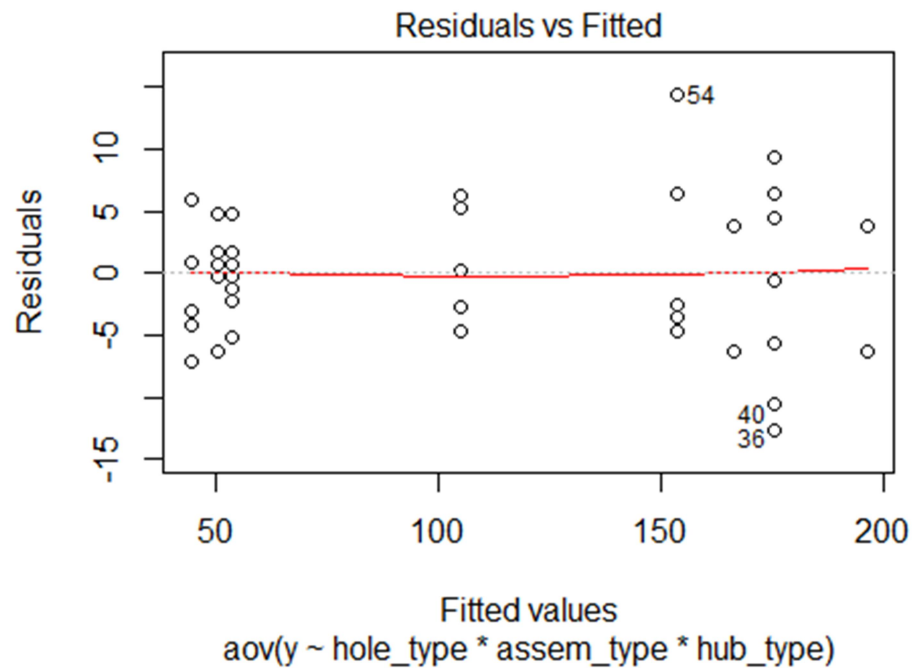
```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq    F value    Pr(>F)
## hole_type    1    8258     8258  266.6837 < 2.2e-16 ***
## assem_type    1   13369    13369  431.7289 < 2.2e-16 ***
## hub_type      1  193050   193050  6234.1653 < 2.2e-16 ***
## hole_type:assem_type  1    2849     2849   91.9991 2.026e-13 ***
## hole_type:hub_type    1     594      594   19.1865 5.250e-05 ***
## assem_type:hub_type    1     135      135    4.3641 0.04126 *
## hole_type:assem_type:hub_type  1    1397     1397   45.1097 1.017e-08 ***
## Residuals      56    1734         31
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The F-values and the corresponding p-values for each of the above term indicates the following :

1. There are significant Main Effects of the 3 Factors (Hub, Assembly, Hole) on mean Torque values.
2. There are significant 2-way Interactions between any 2 of the given Factors, barring the interactions between assembly type and hub type, between which the interactions are mild(p-value = 0.04).
3. There is a significant 3-way Interaction Effect in the Model between Hub, Assembly, and Hole.

Residual Analysis

```
res<-resid(model)
plot(model,which=c(1,2))
```



It can be observed from the plot of residuals that they are normally distributed. To check for this assumption, we perform the normality test. The normtest function is already defined above.

Normality Test for Residuals

```
library(nortest)
normtest(res)
```

```
##                               Method      P.Value
## 1                Shapiro-Wilk normality test 0.172357678
## 2            Anderson-Darling normality test 0.043735579
## 3            Cramer-von Mises normality test 0.051058803
## 4 Lilliefors (Kolmogorov-Smirnov) normality test 0.004193211
## 5            Shapiro-Francia normality test 0.121477563
```

The p-values indicate a normal distribution for the residuals.

Homogeneity Test for Residuals

```
bartlett.test(res~hole_type)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  res by hole_type
## Bartlett's K-squared = 2.8068, df = 1, p-value = 0.09386
```

```
bartlett.test(res~assem_type)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  res by assem_type
## Bartlett's K-squared = 0.3554, df = 1, p-value = 0.5511
```

```
bartlett.test(res~hub_type)
```

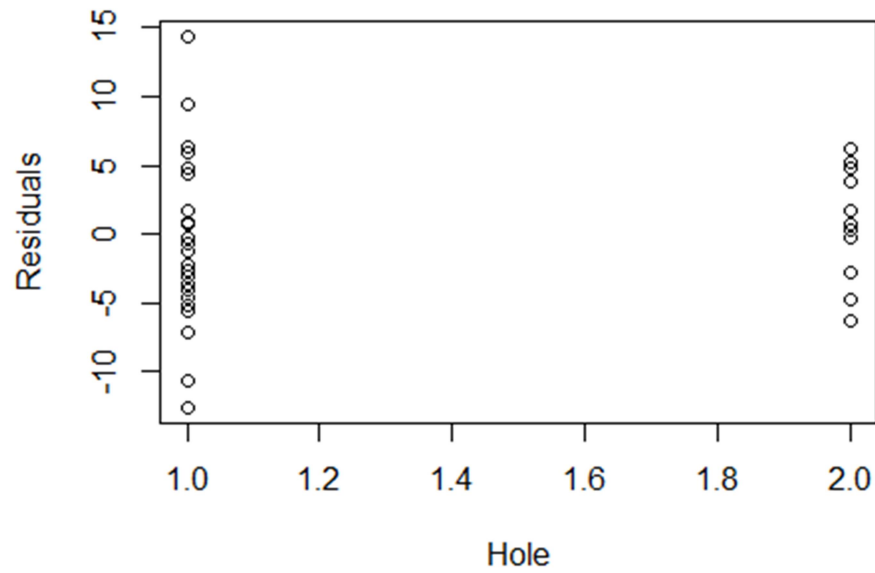
```
##
## Bartlett test of homogeneity of variances
##
## data:  res by hub_type
## Bartlett's K-squared = 6.121, df = 1, p-value = 0.01336
```

The p-values thus obtained suggest that the variance for residuals across various factor combinations are equal i.e. the residuals are homoscedastic except for the Assembly Type – for which the data is not homoscedastic.

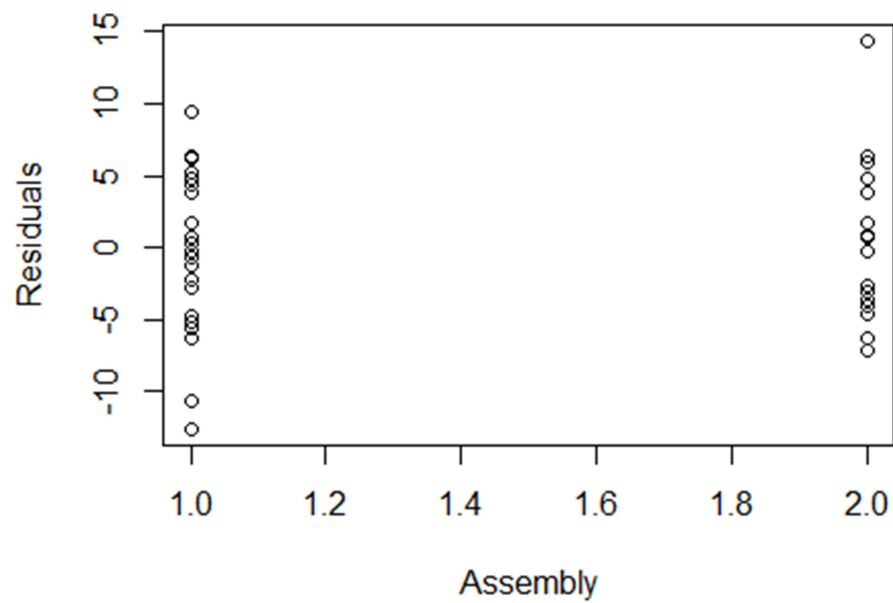
We can visualise the above results for individual, two way interactions and three way interactions as follows.

Individual Residual Analysis

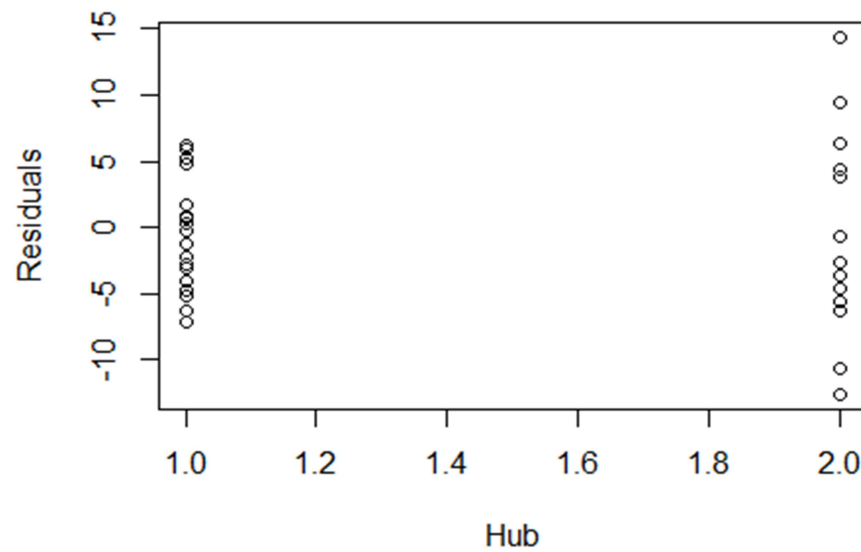
```
plot(as.numeric(hole_type),res,xlab="Hole",ylab="Residuals")
```



```
plot(as.numeric(assem_type),res,xlab="Assembly",ylab="Residuals")
```

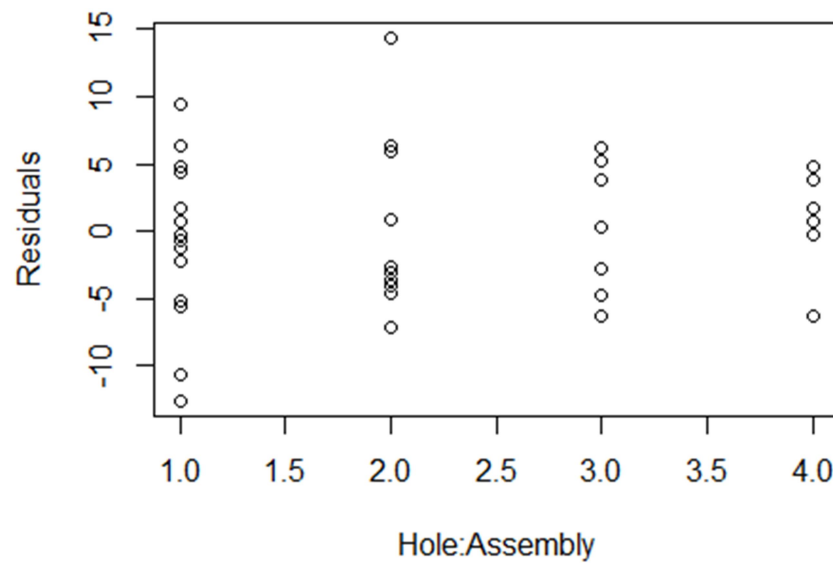


```
plot(as.numeric(hub_type),res,xlab="Hub",ylab="Residuals")
```

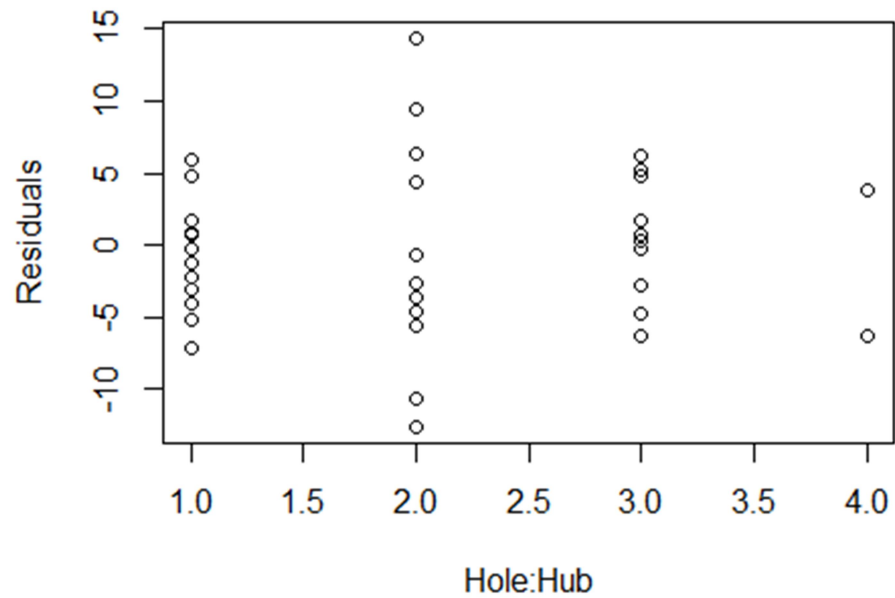


Residual Analysis for Two way Interaction

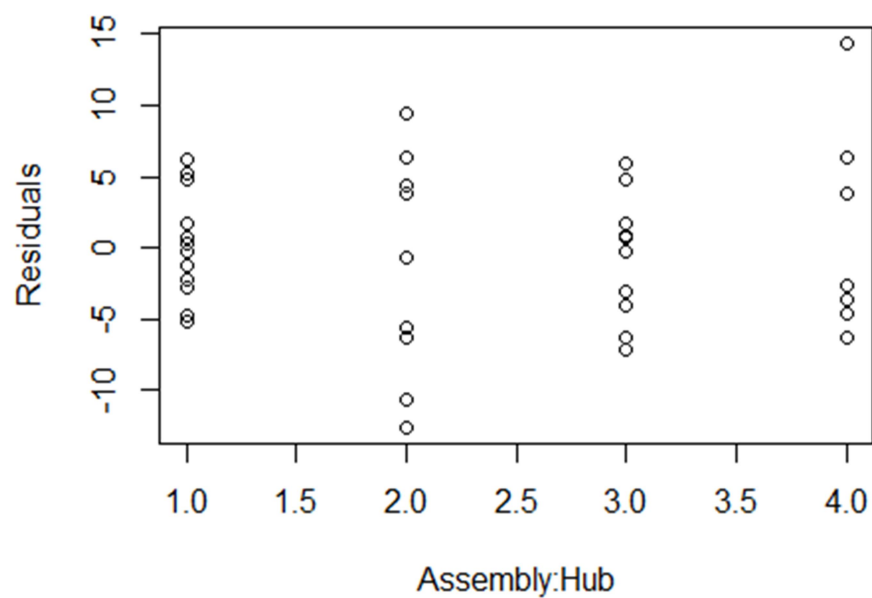
```
plot(as.numeric(hole_type:assem_type),res,xlab="Hole:Assembly",ylab="Residuals")
```




```
plot(as.numeric(hole_type:hub_type),res,xlab="Hole:Hub",ylab="Residuals")
```

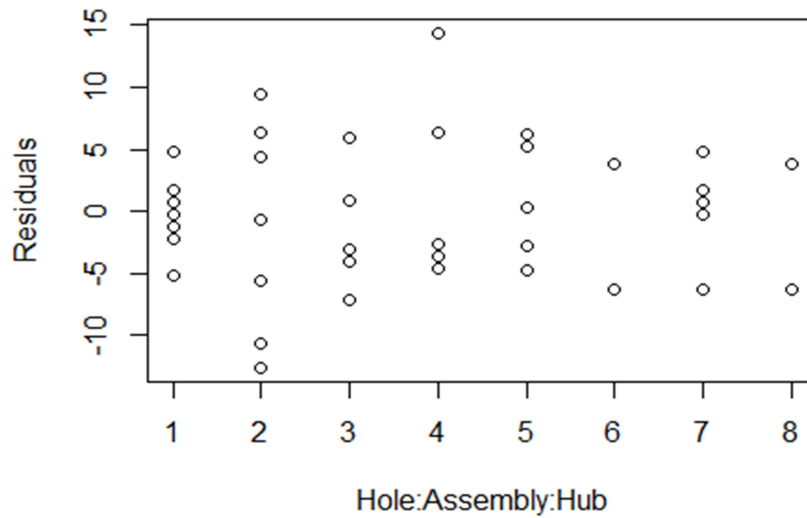


```
plot(as.numeric(assem_type:hub_type),res,xlab="Assembly:Hub",ylab="Residuals")
```



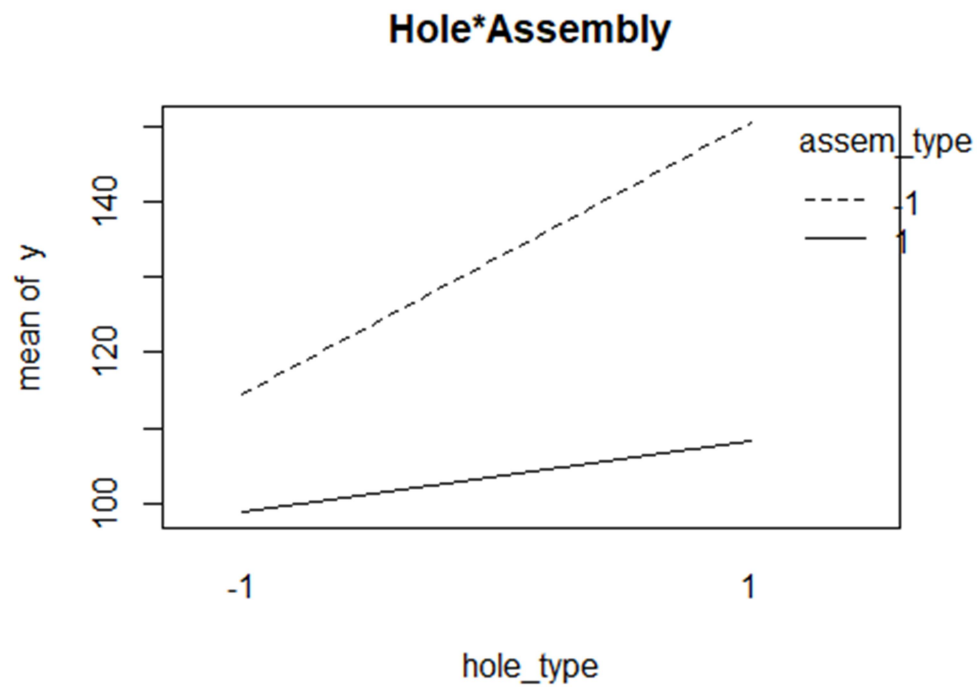
Residual Analysis for Three way Interaction

```
plot(as.numeric(hole_type:assem_type:hub_type),res,xlab="Hole:Assembly:Hub",y  
lab="Residuals")
```

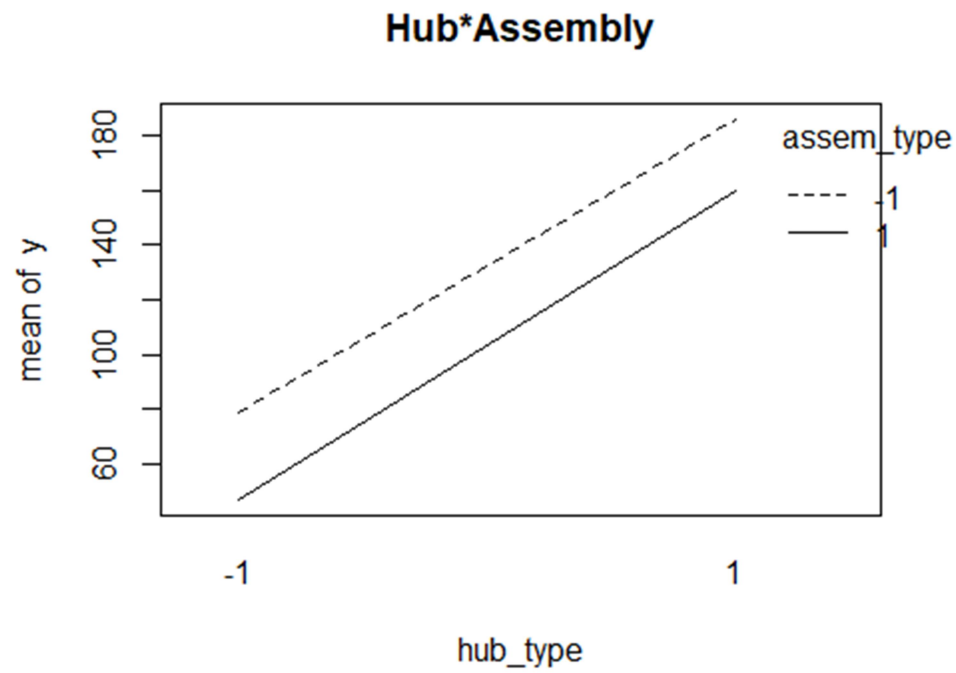


Interaction Plots

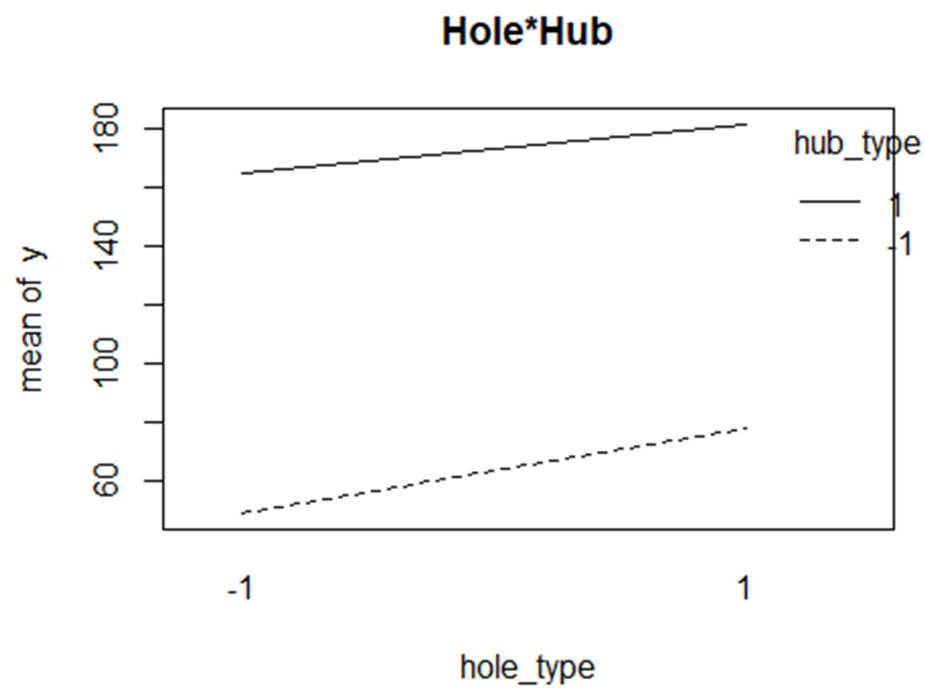
```
interaction.plot(hole_type,assem_type,y,main="Hole*Assembly")
```



```
interaction.plot(hub_type,assem_type,y,main="Hub*Assembly")
```



```
interaction.plot(hole_type,hub_type,y,main="Hole*Hub")
```



Tukey's HSD Test

The Tukey's Honest Significant Difference test is to find out which specific groups's means (compared with each other) are different. The test compares all possible pairs of means.

```
TukeyHSD(model)
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = y ~ hole_type * assem_type * hub_type)
##
## $hole_type
##      diff      lwr      upr p adj
## 1--1 22.71875 19.93186 25.50564    0
##
## $assem_type
##      diff      lwr      upr p adj
## 1--1 -28.90625 -31.69314 -26.11936    0
##
## $hub_type
##      diff      lwr      upr p adj
## 1--1 109.8437 107.0569 112.6306    0
##
## $`hole_type:assem_type`
##      diff      lwr      upr      p adj
## 1:-1--1:-1 36.0625 30.852944 41.2720558 0.0000000
## -1:1--1:-1 -15.5625 -20.772056 -10.3529442 0.0000000
## 1:1--1:-1 -6.1875 -11.397056 -0.9779442 0.0137571
## -1:1-1:-1 -51.6250 -56.834556 -46.4154442 0.0000000
## 1:1-1:-1 -42.2500 -47.459556 -37.0404442 0.0000000
## 1:1--1:1 9.3750 4.165444 14.5845558 0.0000799
##
## $`hole_type:hub_type`
##      diff      lwr      upr p adj
## 1:-1--1:-1 28.8125 23.60294 34.02206    0
## -1:1--1:-1 115.9375 110.72794 121.14706    0
## 1:1--1:-1 132.5625 127.35294 137.77206    0
## -1:1-1:-1 87.1250 81.91544 92.33456    0
## 1:1-1:-1 103.7500 98.54044 108.95956    0
## 1:1--1:1 16.6250 11.41544 21.83456    0
##
## $`assem_type:hub_type`
##      diff      lwr      upr p adj
## 1:-1--1:-1 -31.8125 -37.02206 -26.60294    0
## -1:1--1:-1 106.9375 101.72794 112.14706    0
## 1:1--1:-1 80.9375 75.72794 86.14706    0
## -1:1-1:-1 138.7500 133.54044 143.95956    0
## 1:1-1:-1 112.7500 107.54044 117.95956    0
```

```
## 1:1--1:1    -26.0000 -31.20956 -20.79044      0
##
## `$hole_type:assem_type:hub_type`
##           diff          lwr          upr          p adj
## 1:-1:-1--1:-1:-1  51.500  42.740286  60.2597138 0.0000000
## -1:1:-1--1:-1:-1  -9.125 -17.884714  -0.3652862 0.0354797
## 1:1:-1--1:-1:-1   -3.000 -11.759714   5.7597138 0.9587675
## -1:-1:1--1:-1:-1 122.375 113.615286 131.1347138 0.0000000
## 1:-1:1--1:-1:-1 143.000 134.240286 151.7597138 0.0000000
## -1:1:1--1:-1:-1 100.375  91.615286 109.1347138 0.0000000
## 1:1:1--1:-1:-1 113.000 104.240286 121.7597138 0.0000000
## -1:1:-1-1:-1:-1  -60.625 -69.384714 -51.8652862 0.0000000
## 1:1:-1-1:-1:-1  -54.500 -63.259714 -45.7402862 0.0000000
## -1:-1:1-1:-1:-1   70.875  62.115286  79.6347138 0.0000000
## 1:-1:1-1:-1:-1   91.500  82.740286 100.2597138 0.0000000
## -1:1:1-1:-1:-1   48.875  40.115286  57.6347138 0.0000000
## 1:1:1-1:-1:-1   61.500  52.740286  70.2597138 0.0000000
## 1:1:-1--1:1:-1    6.125  -2.634714  14.8847138 0.3667120
## -1:-1:1--1:1:-1 131.500 122.740286 140.2597138 0.0000000
## 1:-1:1--1:1:-1 152.125 143.365286 160.8847138 0.0000000
## -1:1:1--1:1:-1 109.500 100.740286 118.2597138 0.0000000
## 1:1:1--1:1:-1 122.125 113.365286 130.8847138 0.0000000
## -1:-1:1-1:1:-1 125.375 116.615286 134.1347138 0.0000000
## 1:-1:1-1:1:-1 146.000 137.240286 154.7597138 0.0000000
## -1:1:1-1:1:-1 103.375  94.615286 112.1347138 0.0000000
## 1:1:1-1:1:-1 116.000 107.240286 124.7597138 0.0000000
## 1:-1:1--1:-1:1   20.625  11.865286  29.3847138 0.0000000
## -1:1:1--1:-1:1  -22.000 -30.759714 -13.2402862 0.0000000
## 1:1:1--1:-1:1   -9.375 -18.134714  -0.6152862 0.0278403
## -1:1:1-1:-1:1  -42.625 -51.384714 -33.8652862 0.0000000
## 1:1:1-1:-1:1  -30.000 -38.759714 -21.2402862 0.0000000
## 1:1:1--1:1:1    12.625   3.865286  21.3847138 0.0007601
```

It can be observed that the p-values in most of the cases are low, suggesting a significant effect of different combinations on the torque required for destructive testing.

CONCLUSION

There is significant difference in the torque value for various combinations of the given factors of hub type, assembly type and hole type. From plotting and comparing the torque required for various groups, a combination of Hole type 2, Assembly Method 1 and Hub Shape 2 seem to be the strongest combination for manufacturing Spyders.