assignment3_stat453

Koki Itagaki

2024-03-05

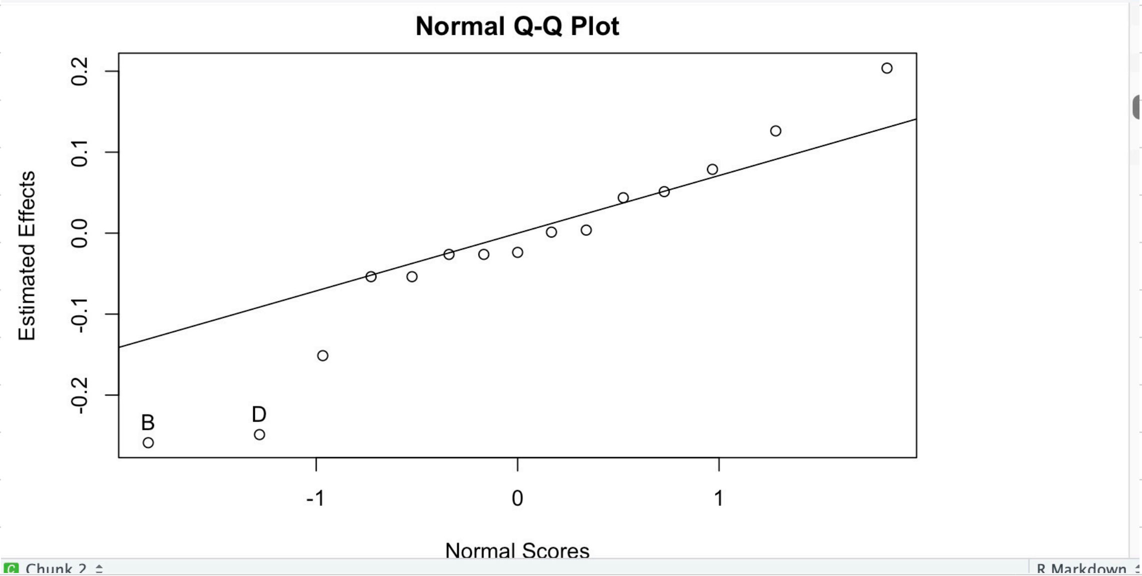
###Question1

```
#a) Analyze the data and determine which factor is not significant.
A <- c(-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1)
B <- c(-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1)
C <- c(-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1)
D <- c(-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1)
y <- c(2.45,3.36,2.16,2.29,2.49,3.39,2.32,2.44,1.84,2.24,1.69,1.87,2.29,2.92,2.04,2.03)
data <- data.frame(A,B,C,D,y)


model <- lm(y ~ A*B*C*D , data = data)
#fullnormal(coef(model)[-1],alpha=.025)

#Prepare an ANOVA table
anova_table <- anova(model)
```

```
## Warning in anova.lm(model): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
# Display the ANOVA table
print(anova_table)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df  Sum Sq Mean Sq F value Pr(>F)
## A          1 0.66423 0.66423     NaN    NaN
## B          1 1.07122 1.07122     NaN    NaN
## C          1 0.25502 0.25502     NaN    NaN
## D          1 0.99002 0.99002     NaN    NaN
## A:B        1 0.36603 0.36603     NaN    NaN
## A:C        1 0.00002 0.00002     NaN    NaN
## B:C        1 0.00903 0.00903     NaN    NaN
## A:D        1 0.04623 0.04623     NaN    NaN
## B:D        1 0.04202 0.04202     NaN    NaN
## C:D        1 0.09922 0.09922     NaN    NaN
## A:B:C      1 0.01102 0.01102     NaN    NaN
```


Normal Q-Q Plot
Estimated Effects
0.2
0.1
0.0
-0.1
-0.2
B
D
-1
0
1
Normal Scores
Chunk 2
R Markdown

```r
#e) Confound the projected design with blocks using the highest order interaction
#as a confounding. Write down the runs for both blocks and estimate the block
#effect. What is the block effect really estimating in this case?

# Create a new factor "Block"
data$Block <- ifelse(data$A * data$B * data$D > 0, "1", "2")

# Fit the model with the block factor
lm_block <- lm(y ~ A * B * D + Block, data = data)

# Summary of the model
summary(lm_block)
```

```
##
## Call:
## lm(formula = y ~ A * B * D + Block, data = data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -0.34  -0.08   0.00   0.08   0.34
##
## Coefficients: (1 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.40750    0.08214  29.310 1.99e-09 ***
## A            0.20375    0.05808   3.508  0.00798 **
## B           -0.25875    0.05808  -4.455  0.00213 **
## D           -0.24875    0.05808  -4.283  0.00268 **
## Block2      -0.08750    0.11616  -0.753  0.47288
## A:B         -0.15125    0.05808  -2.604  0.03142 *
## A:D         -0.05375    0.05808  -0.925  0.38181
## B:D          0.05125    0.05808   0.882  0.40330
## A:B:D             NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2323 on 8 degrees of freedom
## Multiple R-squared:  0.8814, Adjusted R-squared:  0.7777
## F-statistic: 8.497 on 7 and 8 DF,  p-value: 0.003616
```

```r
#So i got -0.08750 as the block effect
```

###Question2

```r
# Import the data:
data = read.csv("yield.csv")
```

```
# Import the data:
data = read.csv("yield.csv")
data
```

```
##     A  B  C  D Rep Yield
## 1  -1 -1 -1 -1   1    90
## 2   1 -1 -1 -1   1    74
## 3  -1  1 -1 -1   1    81
## 4   1  1 -1 -1   1    83
## 5  -1 -1  1 -1   1    77
## 6   1 -1  1 -1   1    81
## 7  -1  1  1 -1   1    88
## 8   1  1  1 -1   1    73
## 9  -1 -1 -1  1   1    98
## 10  1 -1 -1  1   1    72
## 11 -1  1 -1  1   1    87
## 12  1  1 -1  1   1    85
## 13 -1 -1  1  1   1    99
## 14  1 -1  1  1   1    79
## 15 -1  1  1  1   1    87
## 16  1  1  1  1   1    80
## 17 -1 -1 -1 -1   2    93
## 18  1 -1 -1 -1   2    78
## 19 -1  1 -1 -1   2    85
## 20  1  1 -1 -1   2    80
## 21 -1 -1  1 -1   2    78
## 22  1 -1  1 -1   2    80
## 23 -1  1  1 -1   2    82
## 24  1  1  1 -1   2    70
## 25 -1 -1 -1  1   2    95
## 26  1 -1 -1  1   2    76
## 27 -1  1 -1  1   2    83
## 28  1  1 -1  1   2    86
## 29 -1 -1  1  1   2    90
## 30  1 -1  1  1   2    75
## 31 -1  1  1  1   2    84
## 32  1  1  1  1   2    80
```

```
# Fit a linear model
model <- lm(Yield ~ A*B*C*D , data = data)
# (a) Estimate the factor effects
effects <- 2* coef(model)



# Display the estimated factor effects
```

```
print(effects)
```

```
## (Intercept)           A           B           C           D         A:B 
##    165.5625     -9.0625     -1.3125     -2.6875      3.9375      4.0625 
##         A:C         B:C         A:D         B:D         C:D       A:B:C 
##      0.6875     -0.5625     -2.1875     -0.1875      1.6875     -5.1875 
##       A:B:D       A:C:D       B:C:D     A:B:C:D 
##      4.6875     -0.9375     -0.9375      2.4375
```

```
# Based on the R output the  estimated factor effects for the  2^4 factorial design is:

#Intercept: 82.78125, A: -4.53125, B: -0.65625, C: -1.34375, D: 1.96875, A:B: 2.03125, A:C: 0.34375, B:C: -0.28125, A:D: -1.09375, B:D: -0.09375, C:D: 0.84375, A:B:C: -2.59375, A:B:D: 2.34375, A:C:D: -0.46875, B:C:D: -0.46875, A:B:C:D: 1.21875


# (b) Prepare an ANOVA table
anova_table <- anova(model)

# Display the ANOVA table
print(anova_table)
```

```
## Analysis of Variance Table
## 
## Response: Yield
##           Df Sum Sq Mean Sq F value    Pr(>F)    
## A          1 657.03  657.03 85.8163 7.875e-08 ***
## B          1  13.78   13.78  1.8000 0.1984451    
## C          1  57.78   57.78  7.5469 0.0143171 *  
## D          1 124.03  124.03 16.2000 0.0009794 ***
## A:B        1 132.03  132.03 17.2449 0.0007491 ***
## A:C        1   3.78    3.78  0.4939 0.4923019    
## B:C        1   2.53    2.53  0.3306 0.5732962    
## A:D        1  38.28   38.28  5.0000 0.0399447 *  
## B:D        1   0.28    0.28  0.0367 0.8504174    
## C:D        1  22.78   22.78  2.9755 0.1037933    
## A:B:C      1 215.28  215.28 28.1184 7.146e-05 ***
## A:B:D      1 175.78  175.78 22.9592 0.0001997 ***
## A:C:D      1   7.03    7.03  0.9184 0.3521621    
## B:C:D      1   7.03    7.03  0.9184 0.3521621    
## A:B:C:D    1  47.53   47.53  6.2082 0.0240766 *  
## Residuals 16 122.50    7.66                      
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

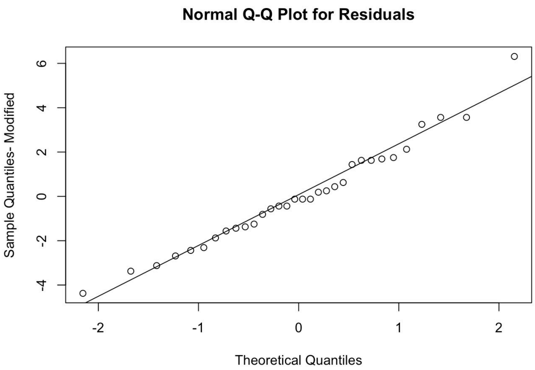
```
#From the anova table, the p-value of A,C,D,A:B,A:D,A:B:C,A:B:D,A:B:C:D are less than 0.05 = a. So these factors
are important factors in this experiment.


#(c)ch
model<- lm(Yield ~ A*B*C*D , data = data)
model
```

```
##
## Call:
## lm(formula = Yield ~ A * B * C * D, data = data)
##
## Coefficients:
## (Intercept)            A            B            C            D          A:B
##    82.78125     -4.53125     -0.65625     -1.34375      1.96875      2.03125
##         A:C          B:C          A:D          B:D          C:D        A:B:C
##     0.34375     -0.28125     -1.09375     -0.09375      0.84375     -2.59375
##       A:B:D        A:C:D        B:C:D      A:B:C:D
##     2.34375     -0.46875     -0.46875      1.21875
```

```
#Yield=β0+β1A+β3C+β4D+β5AB+β6AC+β7AD+β8BC+β9BD+β10CD+β11ABC+β12ABD+β13ACD+β14BCD+β15,ABCD+ϵ
#So, Yield=82.78125-4.53125A-0.65625B-1.34375C+1.96875 D+2.03125AB-1.09375AD-2.59375ABC+2.34375ABD+1.21875ABCD+ϵ




#(d)

#Final model - remove non-significant terms
res.aov<-aov(Yield~A*B*C*D-A:C-B:C-B:D-C:D-A:C:D-B:C:D,data=data)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## A            1  657.0   657.0  87.109 4.17e-09 ***
## B            1   13.8    13.8   1.827 0.190204
## C            1   57.8    57.8   7.661 0.011226 *
## D            1  124.0   124.0  16.444 0.000527 ***
## A:B          1  132.0   132.0  17.505 0.000385 ***
## A:D          1   38.3    38.3   5.075 0.034576 *
## A:B:C        1  215.3   215.3  28.542 2.31e-05 ***
## A:B:D        1  175.8   175.8  23.305 8.01e-05 ***
## A:B:C:D      1   47.5    47.5   6.302 0.019913 *
## Residuals   22  165.9     7.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

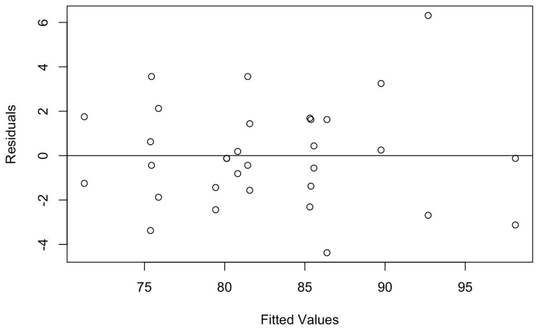
```r
#Residual Analysis
#Normality
yield_residuals=res.aov$residuals
qqnorm(yield_residuals, ylim=c(min(yield_residuals),max(yield_residuals)), main = "Normal Q-Q Plot for Residual
s",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles- Modified",
       plot.it = TRUE, datax = FALSE)

qqline(yield_residuals, datax = FALSE, distribution = qnorm)
```



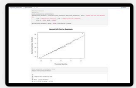

```r
#Test normality using Shapiro Wilks
shapiro.test(yield_residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  yield_residuals
## W = 0.97742, p-value = 0.7219
```

```
#Check Variance
Fitted_values=res.aov$fitted.values
plot(Fitted_values,yield_residuals,ylab="Residuals",xlab="Fitted Values")
abline(h=0)
```




```
#From the q-q plot for residuals, the almost of all data are near the streight line. Also, by the Shapiro-Wilk te
st, Since the p-value is greater than 0.05,we fail to reject the null hypothesis that the data is normally distri
buted.
#it means the distribution is normally distributed.


#Moreover,The   residual    vs  fitted  value   plot  shows
# that the pattern of scatter is the almost same.
#so the variance is adequate
#So this is an adequate model to test.
```

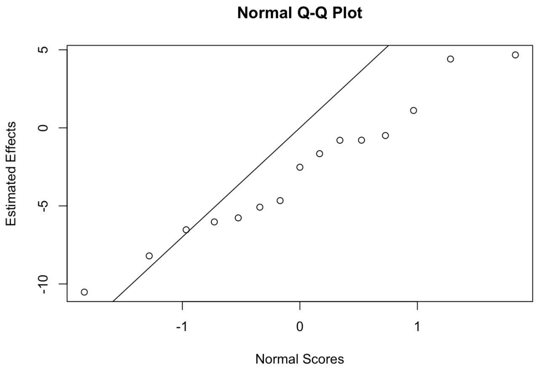


###Question3

```
#a) Are any of the effects significant?
```

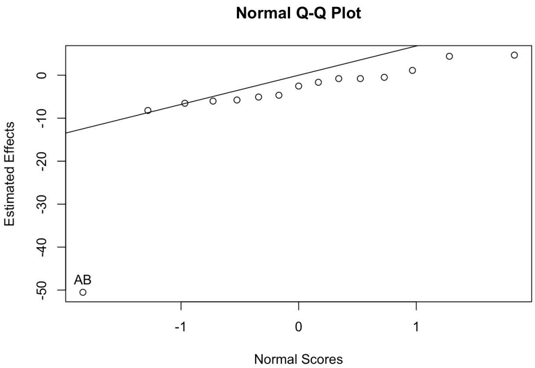
###Question3

```
#a) Are any of the effects significant?
# Define the effect estimates
effects <- c(ABCD = -2.5251, AD = -1.6564, BCD = 4.4054, AC = 1.1109,
             ACD = -0.4932, AB = -10.5229, ABD = -5.0842, D = -6.0275,
             ABC = -5.7696, C = -8.2045, CD = 4.6707, B = -6.5304,
             BD = -4.6620, A = -0.7914, BC = -0.7982)

#According to the fullnormal graph, there is no labels.
#It means that thereare no significant effects.

library(daewr)
fullnormal(effects,alpha=.025)
```




```
#b) What happens if you the effect of the interaction AB was -50.5229 instead of
#-10.5229?
effects2 <- c(ABCD = -2.5251, AD = -1.6564, BCD = 4.4054, AC = 1.1109,
              ACD = -0.4932, AB = -50.5229, ABD = -5.0842, D = -6.0275,
              ABC = -5.7696, C = -8.2045, CD = 4.6707, B = -6.5304,
```

```
#b) What happens if you the effect of the interaction AB was -50.5229 instead of
#-10.5229?
effects2 <- c(ABCD = -2.5251, AD = -1.6564, BCD = 4.4054, AC = 1.1109,
              ACD = -0.4932, AB = -50.5229, ABD = -5.0842, D = -6.0275,
              ABC = -5.7696, C = -8.2045, CD = 4.6707, B = -6.5304,
              BD = -4.6620, A = -0.7914, BC = -0.7982)

fullnormal(effects2,alpha=.025)
```




```
#According to the fullnormal graph, there is one label which is AB.
#So, only AB is significant effect.
```

###Question4

```
#(a) Analyze the data from this experiment. Identify the significant factors
#and interactions and removing the non-significant terms, when
#appropriate.
A <- c(-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1)
B <- c(-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1)
```

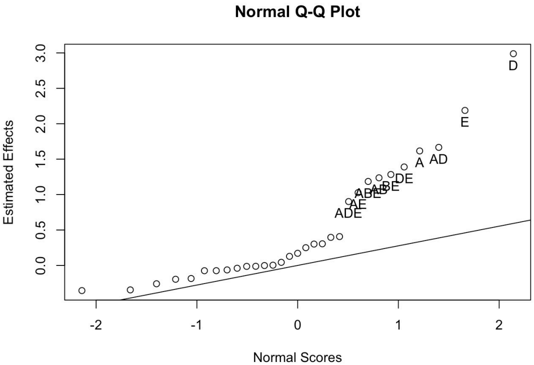
```
C <- c(-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1)
D <- c(-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1)
E <- c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
Obs <- c(8.11,5.56,5.77,5.82,9.17,7.8,3.23,5.69,8.82,14.23,9.2,8.94,8.68,11.49,6.25,9.12,7.93,5,7.47,12,9.86,3.6
5,6.4,11.61,12.43,17.55,8.87,25.38,13.06,18.85,11.78,26.05)
Data <- data.frame(A,B,C,D,E,Obs)
Data
```

```
##     A  B  C  D  E   Obs
## 1  -1 -1 -1 -1 -1  8.11
## 2   1 -1 -1 -1 -1  5.56
## 3  -1  1 -1 -1 -1  5.77
## 4   1  1 -1 -1 -1  5.82
## 5  -1 -1  1 -1 -1  9.17
## 6   1 -1  1 -1 -1  7.80
## 7  -1  1  1 -1 -1  3.23
## 8   1  1  1 -1 -1  5.69
## 9  -1 -1 -1  1 -1  8.82
## 10  1 -1 -1  1 -1 14.23
## 11 -1  1 -1  1 -1  9.20
## 12  1  1 -1  1 -1  8.94
## 13 -1 -1  1  1 -1  8.68
## 14  1 -1  1  1 -1 11.49
## 15 -1  1  1  1 -1  6.25
## 16  1  1  1  1 -1  9.12
## 17 -1 -1 -1 -1  1  7.93
## 18  1 -1 -1 -1  1  5.00
## 19 -1  1 -1 -1  1  7.47
## 20  1  1 -1 -1  1 12.00
## 21 -1 -1  1 -1  1  9.86
## 22  1 -1  1 -1  1  3.65
## 23 -1  1  1 -1  1  6.40
## 24  1  1  1 -1  1 11.61
## 25 -1 -1 -1  1  1 12.43
## 26  1 -1 -1  1  1 17.55
## 27 -1  1 -1  1  1  8.87
## 28  1  1 -1  1  1 25.38
## 29 -1 -1  1  1  1 13.06
## 30  1 -1  1  1  1 18.85
## 31 -1  1  1  1  1 11.78
## 32  1  1  1  1  1 26.05
```

```
Model <- lm(Obs~A*B*C*D*E,data = Data)
coef(Model)
```

```
## (Intercept)           A           B           C           D           E 
##  10.1803125   1.6159375   0.0434375  -0.0121875   2.9884375   2.1878125 
##         A:B         A:C         B:C         A:D         B:D         C:D 
##   1.2365625  -0.0015625  -0.1953125   1.6665625  -0.0134375   0.0034375 
##         A:E         B:E         C:E         D:E       A:B:C       A:B:D 
##   1.0271875   1.2834375   0.3015625   1.3896875   0.2503125  -0.3453125 
##       A:C:D       B:C:D       A:B:E       A:C:E       B:C:E       A:D:E 
##  -0.0634375   0.3053125   1.1853125  -0.2590625   0.1709375   0.9015625 
##       B:D:E       C:D:E     A:B:C:D     A:B:C:E     A:B:D:E     A:C:D:E 
##  -0.0396875   0.3959375  -0.0740625  -0.1846875   0.4071875   0.1278125 
##     B:C:D:E   A:B:C:D:E 
##  -0.0746875  -0.3553125
```

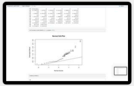
```
fullnormal(coef(Model)[-1],alpha=.025)
```




```
summary(Model)
```

```
##
```

```
##
## Call:
## lm(formula = Obs ~ A * B * C * D * E, data = Data)
##
## Residuals:
## ALL 32 residuals are 0: no residual degrees of freedom!
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.180312        NaN     NaN      NaN
## A            1.615938        NaN     NaN      NaN
## B            0.043438        NaN     NaN      NaN
## C           -0.012187        NaN     NaN      NaN
## D            2.988437        NaN     NaN      NaN
## E            2.187813        NaN     NaN      NaN
## A:B          1.236562        NaN     NaN      NaN
## A:C         -0.001563        NaN     NaN      NaN
## B:C         -0.195313        NaN     NaN      NaN
## A:D          1.666563        NaN     NaN      NaN
## B:D         -0.013438        NaN     NaN      NaN
## C:D          0.003437        NaN     NaN      NaN
## A:E          1.027188        NaN     NaN      NaN
## B:E          1.283437        NaN     NaN      NaN
## C:E          0.301563        NaN     NaN      NaN
## D:E          1.389687        NaN     NaN      NaN
## A:B:C        0.250313        NaN     NaN      NaN
## A:B:D       -0.345312        NaN     NaN      NaN
## A:C:D       -0.063437        NaN     NaN      NaN
## B:C:D        0.305312        NaN     NaN      NaN
## A:B:E        1.185313        NaN     NaN      NaN
## A:C:E       -0.259062        NaN     NaN      NaN
## B:C:E        0.170938        NaN     NaN      NaN
## A:D:E        0.901563        NaN     NaN      NaN
## B:D:E       -0.039687        NaN     NaN      NaN
## C:D:E        0.395938        NaN     NaN      NaN
## A:B:C:D     -0.074063        NaN     NaN      NaN
## A:B:C:E     -0.184688        NaN     NaN      NaN
## A:B:D:E      0.407187        NaN     NaN      NaN
## A:C:D:E      0.127812        NaN     NaN      NaN
## B:C:D:E     -0.074688        NaN     NaN      NaN
## A:B:C:D:E   -0.355312        NaN     NaN      NaN
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 31 and 0 DF,  p-value: NA
```



```
Model2 <- aov(Obs~A+B+D+E+A*B+A*D+A*E+B*E+D*E+A*B*E+A*D*E,data = Data)
summary(Model2)
```

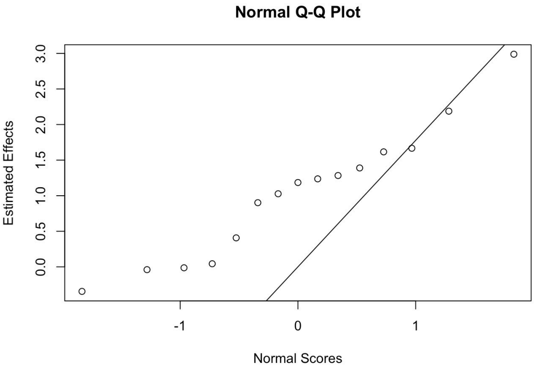
```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## A            1  83.56   83.56  51.362 6.10e-07 ***
## B            1   0.06    0.06   0.037 0.849178    
## D            1 285.78  285.78 175.664 2.30e-11 ***
## E            1 153.17  153.17  94.149 5.24e-09 ***
## A:B          1  48.93   48.93  30.076 2.28e-05 ***
## A:D          1  88.88   88.88  54.631 3.87e-07 ***
## A:E          1  33.76   33.76  20.754 0.000192 ***
## B:E          1  52.71   52.71  32.400 1.43e-05 ***
## D:E          1  61.80   61.80  37.986 5.07e-06 ***
## A:B:E        1  44.96   44.96  27.635 3.82e-05 ***
## A:D:E        1  26.01   26.01  15.988 0.000706 ***
## Residuals   20  32.54    1.63                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#From the full normal plot, it shows that  factors A,D,E,A:D,D:E,B:E,A:B,A:E,A:B:E,A:D:E are significant.
#Also,ANOVA analysis shows the factors A,D,E,AB,AD,AE,BE,DE,ABE,ADE as significant.

#(b)One of the factors from this experiment does not seem to be important.
#If you drop this factor, what type of design remains? Analyze the data
#using the full factorial model for only the four active factors, including
#model adequacy checking.

A <- c(-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1)
B <- c(-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1)
D <- c(-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1)
E <- c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
Obs <- c(8.11,5.56,5.77,5.82,9.17,7.8,3.23,5.69,8.82,14.23,9.2,8.94,8.68,11.49,6.25,9.12,7.93,5,7.47,12,9.86,3.6
5,6.4,11.61,12.43,17.55,8.87,25.38,13.06,18.85,11.78,26.05)
Data <- data.frame(A,B,D,E,Obs)

Model2 <- lm(Obs~A*B*D*E,data = Data)
coef(Model)
```

```
##  (Intercept)            A            B            C            D            E 
##   10.1803125    1.6159375    0.0434375   -0.0121875    2.9884375    2.1878125 
##          A:B          A:C          B:C          A:D          B:D          C:D 
##    1.2365625   -0.0015625   -0.1953125    1.6665625   -0.0134375    0.0034375 
##          A:E          B:E          C:E          D:E        A:B:C        A:B:D 
##    1.0271875    1.2834375    0.3015625    1.3896875    0.2503125   -0.3453125 
##        A:C:D        B:C:D        A:B:E        A:C:E        B:C:E        A:D:E 
```

```
##       A:C:D       B:C:D       A:B:E       A:C:E       B:C:E       A:D:E
##  -0.0634375   0.3053125   1.1853125  -0.2590625   0.1709375   0.9015625
##       B:D:E       C:D:E     A:B:C:D     A:B:C:E     A:B:D:E     A:C:D:E
##  -0.0396875   0.3959375  -0.0740625  -0.1846875   0.4071875   0.1278125
##     B:C:D:E   A:B:C:D:E
##  -0.0746875  -0.3553125
```

```
fullnormal(coef(Model2)[-1],alpha=.025)
```



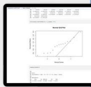

```
summary(Model2)
```

```
##
## Call:
## lm(formula = Obs ~ A * B * D * E, data = Data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.4750 -0.5637  0.0000  0.5637  1.4750
##
```

```
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 10.18031    0.21360  47.661  < 2e-16 ***
## A            1.61594    0.21360   7.565 1.14e-06 ***
## B            0.04344    0.21360   0.203 0.841418
## D            2.98844    0.21360  13.991 2.16e-10 ***
## E            2.18781    0.21360  10.243 1.97e-08 ***
## A:B          1.23656    0.21360   5.789 2.77e-05 ***
## A:D          1.66656    0.21360   7.802 7.66e-07 ***
## B:D         -0.01344    0.21360  -0.063 0.950618
## A:E          1.02719    0.21360   4.809 0.000193 ***
## B:E          1.28344    0.21360   6.009 1.82e-05 ***
## D:E          1.38969    0.21360   6.506 7.24e-06 ***
## A:B:D       -0.34531    0.21360  -1.617 0.125501
## A:B:E        1.18531    0.21360   5.549 4.40e-05 ***
## A:D:E        0.90156    0.21360   4.221 0.000650 ***
## B:D:E       -0.03969    0.21360  -0.186 0.854935
## A:B:D:E      0.40719    0.21360   1.906 0.074735 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.208 on 16 degrees of freedom
## Multiple R-squared:  0.9744, Adjusted R-squared:  0.9504
## F-statistic: 40.58 on 15 and 16 DF,  p-value: 7.07e-10
```

```
Model3 <- aov(Obs~A+B+D+E+A*B+A*D+A*E+B*E+D*E+A*B*E+A*D*E,data = Data)
summary(Model3)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## A            1  83.56   83.56  51.362 6.10e-07 ***
## B            1   0.06    0.06   0.037 0.849178
## D            1 285.78  285.78 175.664 2.30e-11 ***
## E            1 153.17  153.17  94.149 5.24e-09 ***
## A:B          1  48.93   48.93  30.076 2.28e-05 ***
## A:D          1  88.88   88.88  54.631 3.87e-07 ***
## A:E          1  33.76   33.76  20.754 0.000192 ***
## B:E          1  52.71   52.71  32.400 1.43e-05 ***
## D:E          1  61.80   61.80  37.986 5.07e-06 ***
## A:B:E        1  44.96   44.96  27.635 3.82e-05 ***
## A:D:E        1  26.01   26.01  15.988 0.000706 ***
## Residuals   20  32.54    1.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



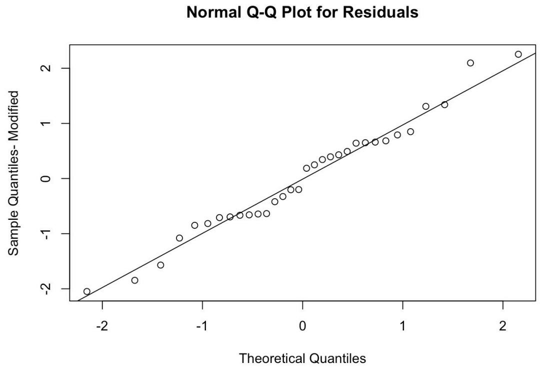
```
#From (a), the factor c seems to be not important, also there are no interactions which are significant with c. S
```

```
#From (a), the factor c seems to be not important, also there are no interactions which are significant with c, S
o, I dropped the
#factor c. This is  2^4 factorial design.
#Even though I drop the factor c, the remaining 4 factors still
#have the same results as the (a).

#For the model checking,




#model checking
#Final model - remove non-significant terms
res.aov<-aov(Model3)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## A            1  83.56   83.56  51.362 6.10e-07 ***
## B            1   0.06    0.06   0.037 0.849178    
## D            1 285.78  285.78 175.664 2.30e-11 ***
## E            1 153.17  153.17  94.149 5.24e-09 ***
## A:B          1  48.93   48.93  30.076 2.28e-05 ***
## A:D          1  88.88   88.88  54.631 3.87e-07 ***
## A:E          1  33.76   33.76  20.754 0.000192 ***
## B:E          1  52.71   52.71  32.400 1.43e-05 ***
## D:E          1  61.80   61.80  37.986 5.07e-06 ***
## A:B:E        1  44.96   44.96  27.635 3.82e-05 ***
## A:D:E        1  26.01   26.01  15.988 0.000706 ***
## Residuals   20  32.54    1.63                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Residual Analysis
#Normality
engineer_residuals=res.aov$residuals
qqnorm(engineer_residuals, ylim=c(min(engineer_residuals),max(engineer_residuals)), main = "Normal Q-Q Plot for R
esiduals",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles- Modified",
       plot.it = TRUE, datax = FALSE)

qqline(engineer_residuals, datax = FALSE, distribution = qnorm)
```

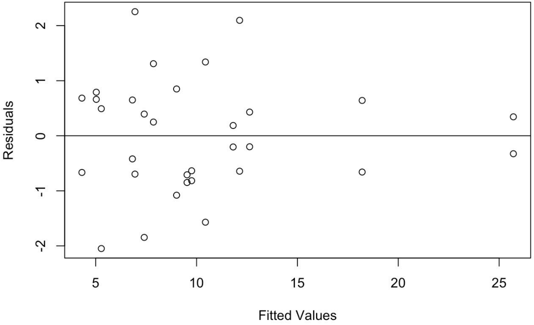
Normal Q-Q Plot for Residuals




```r
#Test normality using Shapiro Wilks
shapiro.test(engineer_residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  engineer_residuals
## W = 0.97507, p-value = 0.6491
```

```r
#Check Variance
Fitted_values=res.aov$fitted.values
plot(Fitted_values,engineer_residuals,ylab="Residuals",xlab="Fitted Values")
abline(h=0)
```




```
#From the q-q plot for residuals, the almost of all data are near the streight line. Also, by the Shapiro-Wilk te
st, Since the p-value(0.6491) is greater than 0.05,we fail to reject the null hypothesis that the data is normall
y distributed.
#it means the distribution is normally distributed.


#Moreover,The   residual    vs  fitted  value   plot  shows
# that the pattern of scatter is the almost same.
#so the variance is adequate
#So this is an adequate model to test.

#(c)Find settings of the active factors that maximize the predicted response.

#Based on the transformed data above we got
#Yi,j,k,l = 10.1803125 + 1.6159375αi+0.0434375βj + 2.9884375γk
# + 2.1878125δl + ϵi,j,k,l
```

###Question5

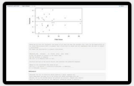
```
#Consider the full 25 factorial design Question 4 above. Suppose that
#this experiment had been run in two blocks with ABCDE confounded with the
#blocks. Set up the blocked design and perform the analysis. Compare your results
#with the results obtained for the completely randomized design in Question 4.
```

```
A <- c(-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1)
B <- c(-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1)
C <- c(-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1)
D <- c(-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1)
E <- c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
Obs <- c(8.11,5.56,5.77,5.82,9.17,7.8,3.23,5.69,8.82,14.23,9.2,8.94,8.68,11.49,6.25,9.12,7.93,5,7.47,12,9.86,3.6
5,6.4,11.61,12.43,17.55,8.87,25.38,13.06,18.85,11.78,26.05)
# Define the block variable
#positive is 2 negative 1 is 1
block <- c(1,2,2,1,2,1,1,2,2,1,1,2,1,2,2,1,2,1,1,2,1,2,2,1,1,2,2,1,2,1,1,2)

Data <- data.frame(A,B,C,D,E,Obs,block)


# Fit the model with the block effect and confounding ABCDE with blocks
Model_Mod <- lm(Obs ~ block + A*B*C*D*E, data = Data)

# Display the ANOVA table for the modified design
anova_table_Mod <- anova(Model_Mod)
```

```
## Warning in anova.lm(Model_Mod): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
print(anova_table_Mod)
```

```
## Analysis of Variance Table
##
## Response: Obs
##            Df  Sum Sq Mean Sq F value Pr(>F)
## block       1   4.040   4.040     NaN    NaN
## A           1  83.560  83.560     NaN    NaN
## B           1   0.060   0.060     NaN    NaN
## C           1   0.005   0.005     NaN    NaN
## D           1 285.784 285.784     NaN    NaN
## E           1 153.169 153.169     NaN    NaN
## A:B         1  48.931  48.931     NaN    NaN
## A:C         1   0.000   0.000     NaN    NaN
## B:C         1   1.221   1.221     NaN    NaN
## A:D         1  88.878  88.878     NaN    NaN
## B:D         1   0.006   0.006     NaN    NaN
## C:D         1   0.000   0.000     NaN    NaN
## A:E         1  33.764  33.764     NaN    NaN
## B:E         1  52.711  52.711     NaN    NaN
## C:E         1   2.910   2.910     NaN    NaN
## D:E         1  61.799  61.799     NaN    NaN
```



```
## C:D         1   0.000   0.000     NaN     NaN
## A:E         1  33.764  33.764     NaN     NaN
## B:E         1  52.711  52.711     NaN     NaN
## C:E         1   2.910   2.910     NaN     NaN
## D:E         1  61.799  61.799     NaN     NaN
## A:B:C       1   2.005   2.005     NaN     NaN
## A:B:D       1   3.816   3.816     NaN     NaN
## A:C:D       1   0.129   0.129     NaN     NaN
## B:C:D       1   2.983   2.983     NaN     NaN
## A:B:E       1  44.959  44.959     NaN     NaN
## A:C:E       1   2.148   2.148     NaN     NaN
## B:C:E       1   0.935   0.935     NaN     NaN
## A:D:E       1  26.010  26.010     NaN     NaN
## B:D:E       1   0.050   0.050     NaN     NaN
## C:D:E       1   5.017   5.017     NaN     NaN
## A:B:C:D     1   0.176   0.176     NaN     NaN
## A:B:C:E     1   1.092   1.092     NaN     NaN
## A:B:D:E     1   5.306   5.306     NaN     NaN
## A:C:D:E     1   0.523   0.523     NaN     NaN
## B:C:D:E     1   0.179   0.179     NaN     NaN
## Residuals   0   0.000     NaN
```

```r
#This is from question4
A <- c(-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1,-1,1)
B <- c(-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1,-1,-1,1,1)
C <- c(-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1,-1,-1,-1,-1,1,1,1,1)
D <- c(-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1)
E <- c(-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,-1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1)
Obs <- c(8.11,5.56,5.77,5.82,9.17,7.8,3.23,5.69,8.82,14.23,9.2,8.94,8.68,11.49,6.25,9.12,7.93,5,7.47,12,9.86,3.6
5,6.4,11.61,12.43,17.55,8.87,25.38,13.06,18.85,11.78,26.05)
Data <- data.frame(A,B,C,D,E,Obs)
Data
```

```
##      A  B  C  D  E   Obs
## 1   -1 -1 -1 -1 -1  8.11
## 2    1 -1 -1 -1 -1  5.56
## 3   -1  1 -1 -1 -1  5.77
## 4    1  1 -1 -1 -1  5.82
## 5   -1 -1  1 -1 -1  9.17
## 6    1 -1  1 -1 -1  7.80
## 7   -1  1  1 -1 -1  3.23
## 8    1  1  1 -1 -1  5.69
## 9   -1 -1 -1  1 -1  8.82
## 10   1 -1 -1  1 -1 14.23
## 11  -1  1 -1  1 -1  9.20
## 12   1  1 -1  1 -1  8.94
```

assignment3_stat453.html Open in Browser Find Publish

```
## 13 -1 -1  1  1 -1  8.68
## 14  1 -1  1  1 -1 11.49
## 15 -1  1  1  1 -1  6.25
## 16  1  1  1  1 -1  9.12
## 17 -1 -1 -1 -1  1  7.93
## 18  1 -1 -1 -1  1  5.00
## 19 -1  1 -1 -1  1  7.47
## 20  1  1 -1 -1  1 12.00
## 21 -1 -1  1 -1  1  9.86
## 22  1 -1  1 -1  1  3.65
## 23 -1  1  1 -1  1  6.40
## 24  1  1  1 -1  1 11.61
## 25 -1 -1 -1  1  1 12.43
## 26  1 -1 -1  1  1 17.55
## 27 -1  1 -1  1  1  8.87
## 28  1  1 -1  1  1 25.38
## 29 -1 -1  1  1  1 13.06
## 30  1 -1  1  1  1 18.85
## 31 -1  1  1  1  1 11.78
## 32  1  1  1  1  1 26.05
```

```
Model <- lm(Obs~A*B*C*D*E,data = Data)
aov.model = aov(Model)
summary(aov.model)
```

```
##             Df Sum Sq Mean Sq
## A            1  83.56   83.56
## B            1   0.06    0.06
## C            1   0.00    0.00
## D            1 285.78  285.78
## E            1 153.17  153.17
## A:B          1  48.93   48.93
## A:C          1   0.00    0.00
## B:C          1   1.22    1.22
## A:D          1  88.88   88.88
## B:D          1   0.01    0.01
## C:D          1   0.00    0.00
## A:E          1  33.76   33.76
## B:E          1  52.71   52.71
## C:E          1   2.91    2.91
## D:E          1  61.80   61.80
## A:B:C        1   2.01    2.01
## A:B:D        1   3.82    3.82
## A:C:D        1   0.13    0.13
## B:C:D        1   2.98    2.98
## A:B:E        1  44.96   44.96
```

```
## 29 -1 -1  1  1  1 13.06
## 30  1 -1  1  1  1 18.85
## 31 -1  1  1  1  1 11.78
## 32  1  1  1  1  1 26.05
```

```
Model <- lm(Obs~A*B*C*D*E,data = Data)
aov.model = aov(Model)
summary(aov.model)
```

```
##              Df Sum Sq Mean Sq
## A             1  83.56   83.56
## B             1   0.06    0.06
## C             1   0.00    0.00
## D             1 285.78  285.78
## E             1 153.17  153.17
## A:B           1  48.93   48.93
## A:C           1   0.00    0.00
## B:C           1   1.22    1.22
## A:D           1  88.88   88.88
## B:D           1   0.01    0.01
## C:D           1   0.00    0.00
## A:E           1  33.76   33.76
## B:E           1  52.71   52.71
## C:E           1   2.91    2.91
## D:E           1  61.80   61.80
## A:B:C         1   2.01    2.01
## A:B:D         1   3.82    3.82
## A:C:D         1   0.13    0.13
## B:C:D         1   2.98    2.98
## A:B:E         1  44.96   44.96
## A:C:E         1   2.15    2.15
## B:C:E         1   0.94    0.94
## A:D:E         1  26.01   26.01
## B:D:E         1   0.05    0.05
## C:D:E         1   5.02    5.02
## A:B:C:D       1   0.18    0.18
## A:B:C:E       1   1.09    1.09
## A:B:D:E       1   5.31    5.31
## A:C:D:E       1   0.52    0.52
## B:C:D:E       1   0.18    0.18
## A:B:C:D:E     1   4.04    4.04
```

```
#Compared to the design from question 4, the value of question 5 has less
#values.
```