Stat453_Assignment03

Belina Jang V00924202

2023-03-10

Questions

Question 1. The yield of a chemical process is being studied. The two most important variables are thought to be the pressure and the temperature. Three levels of each factor are selected, and a factorial experiment with two replicates is performed. The yield data follow:

	Pressure		
Temperature	200	215	230
150	90.4	90.7	90.2
	90.2	90.6	90.4
160	90.1	90.5	89.9
	90.3	90.6	90.1
170	90.5	90.8	90.4
	90.7	90.9	90.1

```
rm(list = ls())
Temperature=rep(c(rep(150,2),rep(160,2),rep(170,2)),3)
Pressure<-c(rep(200,6), rep(215,6),rep(230,6))
Yield = c(90.4, 90.2, 90.1, 90.3, 90.5, 90.7,
          90.7, 90.6, 90.5, 90.6, 90.8, 90.9,
          90.2, 90.4, 89.9, 90.1, 90.4, 90.1)
q1_data <- data.frame(Yield,Temperature, Pressure)

#Using ANOVA

res.aov <-aov(Yield~ factor(Temperature)*factor(Pressure),data=q1_data)
summary(res.aov)
```

```
##                                    Df Sum Sq Mean Sq F value   Pr(>F)
## factor(Temperature)                 2 0.3011  0.1506   8.469 0.008539 **
## factor(Pressure)                    2 0.7678  0.3839  21.594 0.000367 ***
## factor(Temperature):factor(Pressure)  4 0.0689  0.0172   0.969 0.470006
## Residuals                           9 0.1600  0.0178
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
res.aov$coefficients
```

```
##                                (Intercept)
##                               9.030000e+01
##                       factor(Temperature)160
##                                -1.000000e-01
##                       factor(Temperature)170
##                                 3.000000e-01
##                          factor(Pressure)215
##                                 3.500000e-01
##                          factor(Pressure)230
##                                -8.888948e-15
## factor(Temperature)160:factor(Pressure)215
##                                 1.573213e-14
## factor(Temperature)170:factor(Pressure)215
##                                -1.000000e-01
## factor(Temperature)160:factor(Pressure)230
##                                -2.000000e-01
## factor(Temperature)170:factor(Pressure)230
##                                -3.500000e-01
```

```
#Multiple comparisons

TUKEY <- TukeyHSD(x=res.aov, conf.level=0.95)
TUKEY
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Yield ~ factor(Temperature) * factor(Pressure), data = q1_data)
##
## $`factor(Temperature)`
##               diff         lwr        upr     p adj
## 160-150 -0.1666667 -0.38159536 0.04826203 0.1313152
## 170-150  0.1500000 -0.06492869 0.36492869 0.1809078
## 170-160  0.3166667  0.10173797 0.53159536 0.0066518
##
## $`factor(Pressure)`
##               diff        lwr         upr     p adj
## 215-200  0.3166667  0.1017380  0.53159536 0.0066518
## 230-200 -0.1833333 -0.3982620  0.03159536 0.0944905
## 230-215 -0.5000000 -0.7149287 -0.28507131 0.0002951
##
## $`factor(Temperature):factor(Pressure)`
##                             diff         lwr         upr     p adj
## 160:200-150:200 -1.000000e-01 -0.62747453  0.42747453 0.9959112
## 170:200-150:200  3.000000e-01 -0.22747453  0.82747453 0.4489114
## 150:215-150:200  3.500000e-01 -0.17747453  0.87747453 0.2916707
## 160:215-150:200  2.500000e-01 -0.27747453  0.77747453 0.6429825
## 170:215-150:200  5.500000e-01  0.02252547  1.07747453 0.0397898
## 150:230-150:200 -1.421085e-14 -0.52747453  0.52747453 1.0000000
## 160:230-150:200 -3.000000e-01 -0.82747453  0.22747453 0.4489114
## 170:230-150:200 -5.000000e-02 -0.57747453  0.47747453 0.9999713
## 170:200-160:200  4.000000e-01 -0.12747453  0.92747453 0.1812587
## 150:215-160:200  4.500000e-01 -0.07747453  0.97747453 0.1099996
## 160:215-160:200  3.500000e-01 -0.17747453  0.87747453 0.2916707
```

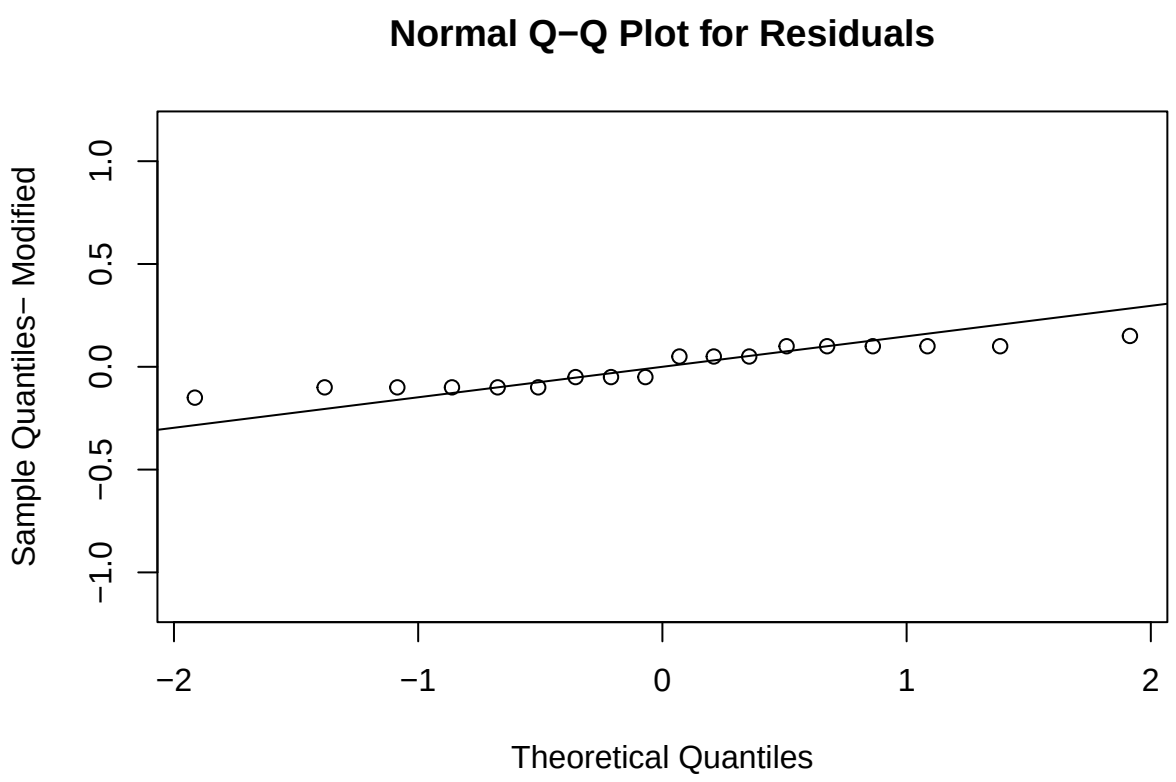
```
## 170:215-160:200  6.500000e-01  0.12252547  1.17747453 0.0147551
## 150:230-160:200  1.000000e-01 -0.42747453  0.62747453 0.9959112
## 160:230-160:200 -2.000000e-01 -0.72747453  0.32747453 0.8316762
## 170:230-160:200  5.000000e-02 -0.47747453  0.57747453 0.9999713
## 150:215-170:200  5.000000e-02 -0.47747453  0.57747453 0.9999713
## 160:215-170:200 -5.000000e-02 -0.57747453  0.47747453 0.9999713
## 170:215-170:200  2.500000e-01 -0.27747453  0.77747453 0.6429825
## 150:230-170:200 -3.000000e-01 -0.82747453  0.22747453 0.4489114
## 160:230-170:200 -6.000000e-01 -1.12747453 -0.07252547 0.0241003
## 170:230-170:200 -3.500000e-01 -0.87747453  0.17747453 0.2916707
## 160:215-150:215 -1.000000e-01 -0.62747453  0.42747453 0.9959112
## 170:215-150:215  2.000000e-01 -0.32747453  0.72747453 0.8316762
## 150:230-150:215 -3.500000e-01 -0.87747453  0.17747453 0.2916707
## 160:230-150:215 -6.500000e-01 -1.17747453 -0.12252547 0.0147551
## 170:230-150:215 -4.000000e-01 -0.92747453  0.12747453 0.1812587
## 170:215-160:215  3.000000e-01 -0.22747453  0.82747453 0.4489114
## 150:230-160:215 -2.500000e-01 -0.77747453  0.27747453 0.6429825
## 160:230-160:215 -5.500000e-01 -1.07747453 -0.02252547 0.0397898
## 170:230-160:215 -3.000000e-01 -0.82747453  0.22747453 0.4489114
## 150:230-170:215 -5.500000e-01 -1.07747453 -0.02252547 0.0397898
## 160:230-170:215 -8.500000e-01 -1.37747453 -0.32252547 0.0023855
## 170:230-170:215 -6.000000e-01 -1.12747453 -0.07252547 0.0241003
## 160:230-150:230 -3.000000e-01 -0.82747453  0.22747453 0.4489114
## 170:230-150:230 -5.000000e-02 -0.57747453  0.47747453 0.9999713
## 170:230-160:230  2.500000e-01 -0.27747453  0.77747453 0.6429825
```

(a) Analyze the data and draw conclusions. Use a = 0.05.

∴ The p-value (0.470006) for the interaction term is bigger than 0.05, therefore the effect of interaction is not significant. So we can just look at the Tukey test result for the main effects. Temperature: p-value (0.0066518) for the difference between levels 170 and 160 < a=0.05, therefore the effect of temperature is significant. Pressure: p-values fore the difference between levels are smaller than a=0.05 except for the difference between 230 and 200 (p-value: 0.0944905), therefore the effect of pressure is significant. So both temperature and pressure are significant factors, but interaction is not significant.

(b) Prepare appropriate residual plots and comment on the model's adequacy.

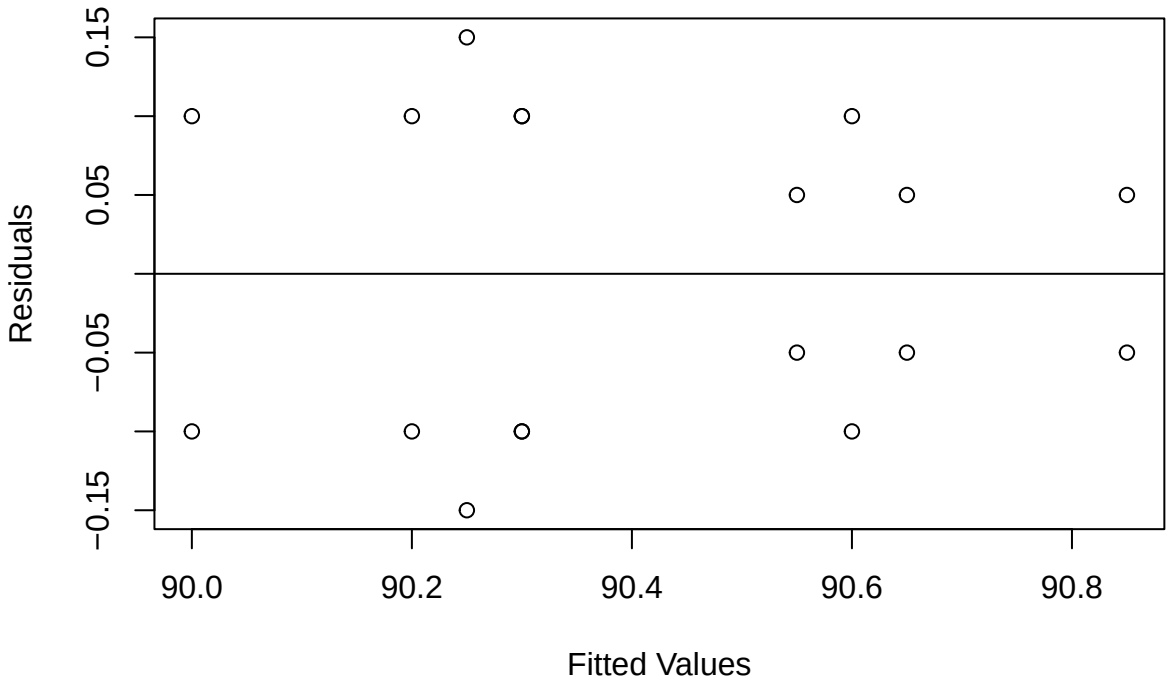
```r
#Normality
residuals=res.aov$residuals
qqnorm(residuals, ylim=c(min(residuals)-1,max(residuals)+1), main = "Normal Q-Q Plot for Residuals",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles- Modified",
       plot.it = TRUE, datax = FALSE)

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




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

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.87366, p-value = 0.02046
```

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



∴ Normality plot does not seem to violate normality assumption but the p-value of the shapiro test is smaller than a=0.05, so it might not be normally distributed. Therefore there's an indication of violation of the normality assumption. Since the variance plots seem spread out, there's no indication of violation of the constant variance assumption.

(c) Under what conditions would you operate this process (i.e., which conditions maximize the yield)?

```
q1_data["Fitted"] = res.aov$fitted.values
q1_data[res.aov$fitted.values == max(res.aov$fitted.values),]
```

```
##    Yield Temperature Pressure Fitted
## 11  90.8         170      215  90.85
## 12  90.9         170      215  90.85
```

∴ Since the interaction term is not significant, we don't need to worry about the effect of interaction, so temperature at 170(since the diff of 170-160 is positive and significant) and pressure at 215(since the diff of 215-200 is positive and significant) would maximize the yield.

Question 2. The C. F. Eye Care company manufactures lenses for transplantation into the eye following cataract surgery. An engineering group has conducted an experiment involving two factors to determine their effect on the lens polishing process. The results of this experiment are summarized in the following ANOVA display:

Source	*DF*	*SS*	*MS*	*F*	*P*-value
Factor *A*	?	?	0.0833	0.05	0.952
Factor *B*	?	96.333	96.333	57.80	<0.001
Interaction	2	12.167	6.0833	3.65	?
Error	6	10.000	?		
Total	11	118.667			

(a) Complete the table

```
dfB = floor(96.333/96.333) #1
dfA = 2/dfB #2
SSA = 0.0833*dfA #0.1666
MSE = 10/6 #1.666667
PInteraction = pf(3.65, 2, 6, lower.tail = FALSE) #0.09181187
```

Source	*DF*	*SS*	*MS*	*F*	*P*-value
Factor *A*	2	0.1666	0.0833	0.05	0.952
Factor *B*	1	96.333	96.333	57.80	<0.001
Interaction	2	12.167	6.0833	3.65	0.0918
Error	6	10.000	1.6667		
Total	11	118.667			

(b) How many replicates are in this experiment?

∴ Since a-1 = 2 and b-1 = 1, a = 3 and b = 2. Since df(error) = 6 = ab(n-1) and df(Total)= abn-1 = 11, n = 2. Therefore there are 2 replicates in this experiment.

(c) Does the effect of factor B depend on the level of factor A? Justify.

∴ Since the p-value of the interaction term is bigger than a=0.05, the effect of the interaction term is not significant which shows that the effect of factor B does not depend on the level of factor A.

Question 3. An experiment was performed to improve the yield of a chemical process. Four factors were selected, and two replicates of a completely randomized experiment were run. The results are shown in the following table:

Treatment Combination	Replicate I	Replicate II	Treatment Combination	Replicate I	Replicate II
(1)	90	93	*d*	98	95
a	74	78	*ad*	72	76
b	81	85	*bd*	87	83
ab	83	80	*abd*	85	86
c	77	78	*cd*	99	90
ac	81	80	*acd*	79	75
bc	88	82	*bcd*	87	84
abc	73	70	*abcd*	80	80

```
A=rep(c(-1,1),8)
B=rep(c(rep(-1,2),rep(1,2)),4)
C=rep(c(rep(-1,4),rep(1,4)),2)
D=c(rep(-1,8),rep(1,8))
rep1 = c(90, 74, 81, 83, 77, 81, 88, 73,
         98, 72, 87, 85, 99, 79, 87, 80)
rep2 = c(93, 78, 85, 80, 78, 80, 82, 70,
         95, 76, 83, 86, 90, 75, 84, 80)
A = rep(A, rep(2, 16))
B = rep(B, rep(2, 16))
C = rep(C, rep(2, 16))
D = rep(D, rep(2, 16))
yield = cbind(rep1,rep2)
yield.vec = c(t(yield))
yield_data = data.frame(A, B, C, D, yield.vec)
```

(a) Estimate the factor effects.

```
#Using ANOVA
res.aov <-aov(yield.vec~A*B*C*D,data=yield_data)
res.aov$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
```

```
res.aov$coefficients[-1]*2 #factor effects
```

```
##       A       B       C       D     A:B     A:C     B:C     A:D     B:D     C:D
```

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

Factor effects are as follows: A = -9.0625, B = -1.3125, C = -2.6875, D = 3.9375, AB = 4.0625, AC = 0.6875, BC = -0.5625, AD = -2.1875, BD = -0.1875, CD = 1.6875, ABC = -5.1875, ABD = 4.6875, ACD = -0.9375, BCD = -0.9375, ABCD = 2.4375.

(b) Prepare an analysis of variance table, and determine which factors are important in explaining yield.

```
res.aov = aov(yield.vec~A*B*C*D, data=yield_data)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## A            1  657.0   657.0  85.816 7.87e-08 ***
## B            1   13.8    13.8   1.800 0.198445
## C            1   57.8    57.8   7.547 0.014317 *
## D            1  124.0   124.0  16.200 0.000979 ***
## A:B          1  132.0   132.0  17.245 0.000749 ***
## A:C          1    3.8     3.8   0.494 0.492302
## B:C          1    2.5     2.5   0.331 0.573296
## A:D          1   38.3    38.3   5.000 0.039945 *
## B:D          1    0.3     0.3   0.037 0.850417
## C:D          1   22.8    22.8   2.976 0.103793
## A:B:C        1  215.3   215.3  28.118 7.15e-05 ***
## A:B:D        1  175.8   175.8  22.959 0.000200 ***
## A:C:D        1    7.0     7.0   0.918 0.352162
## B:C:D        1    7.0     7.0   0.918 0.352162
## A:B:C:D      1   47.5    47.5   6.208 0.024077 *
## Residuals   16  122.5     7.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

∴ Factors and interactions A, C, D, AB, AD, ABC, ABD, ABCD are significant.

(c) Write down a regression model for predicting yield, assuming that all four factors were varied over the range from −1 to +1 (in coded units).

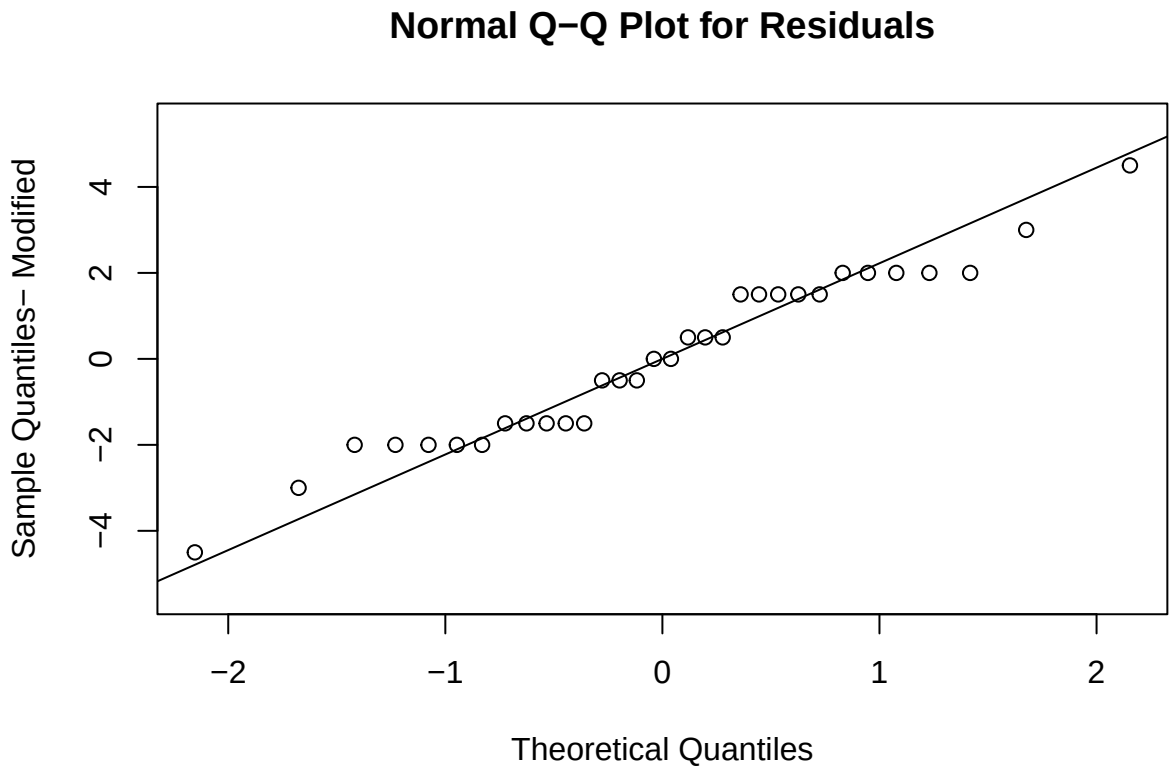
```
res.aov$coefficients
```

```
## (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 = 82.78125 -4.53125A -1.34375C + 1.96875D + 2.03125AB -1.09375AD -2.59375ABC + 2.34375ABD + 1.21875ABCD

(d) Does the residual analysis appear satisfactory?

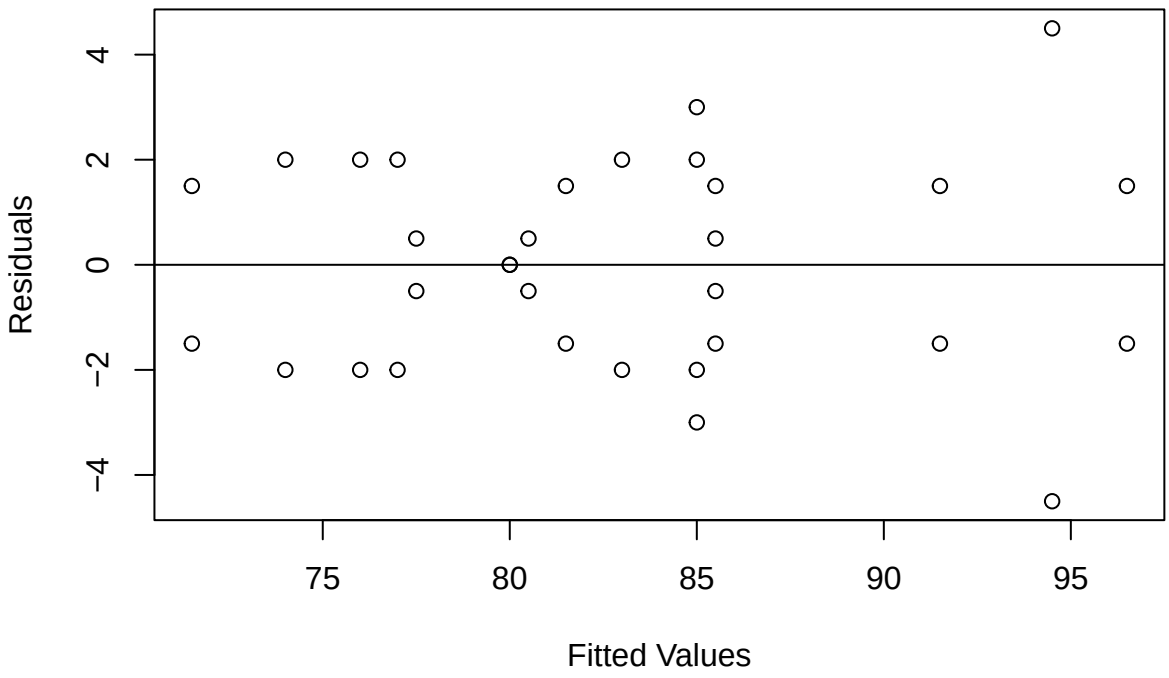
```
#residual analysis
#Normality
res.aov = aov(yield.vec~A*B*C*D, data=yield_data)
residuals=res.aov$residuals
qqnorm(residuals, ylim=c(min(residuals)-1,max(residuals)+1), main = "Normal Q-Q Plot for Residuals",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles- Modified",
       plot.it = TRUE, datax = FALSE)

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




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

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.96135, p-value = 0.2989
```

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



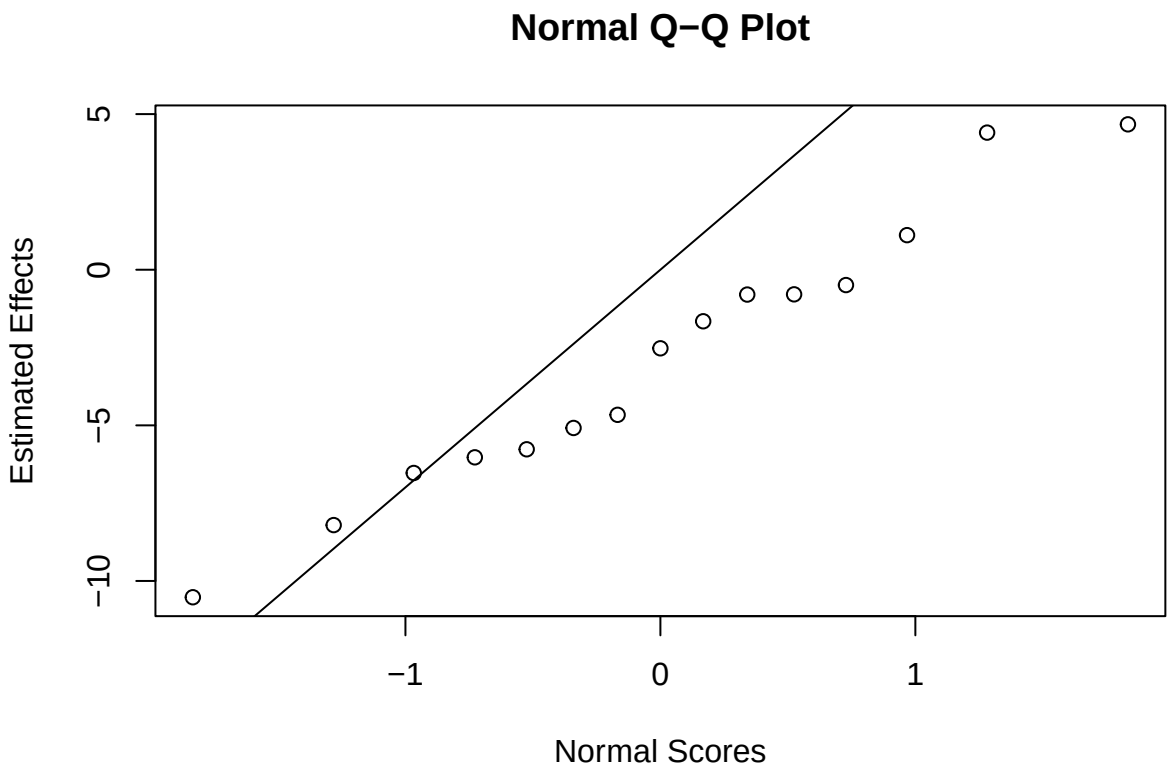
∴ Since the p-value from Shapiro Wilks test is $0.2989 > 0.05$, we fail to reject the null hypothesis that the residuals are normally distributed. Therefore it is normally distributed. There's an indication of violation of the constant variance assumption since there seems to be a pattern in Fitted values vs. residuals plot.

Question 4. The effect estimates from a 2^4 factorial experiment are listed here.

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		

(a) Are any of the effects significant?

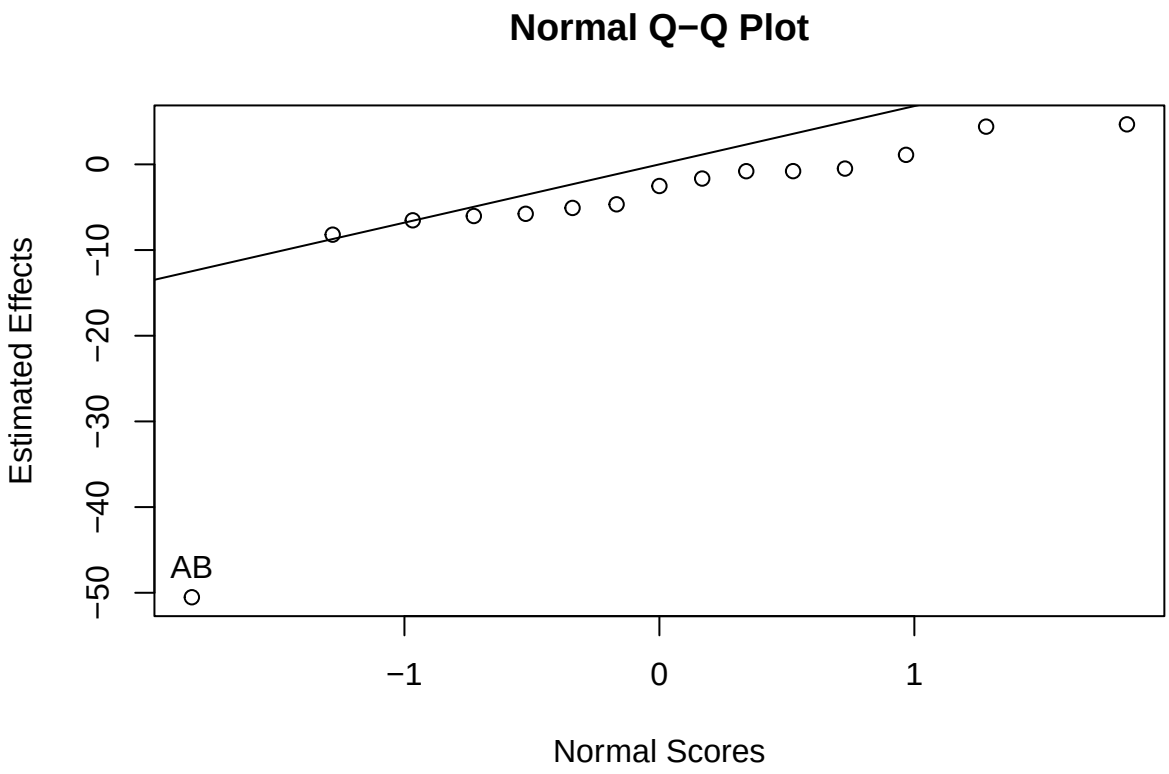
```
effects = c(-2.5251,-1.6564,4.4054,1.1109,-0.4932,-10.5229,-5.0842,-6.0275,-5.7696,-8.2045,
            4.6707,-6.5304,-4.6620,-0.7914,-0.7982)
names(effects) = c("ABCD","AD","BCD","AC","ACD","AB","ABD","D","ABC","C","CD","B","BD","A","BC")

fullnormal(effects,names(effects),alpha=0.025)
```



∴ No, all the effects doesn't seem to be significant.

(b) What happens if the effect of the interaction AB was -50.5229 instead of -10.5229?

```
effects[6] = -50.5229
fullnormal(effects,names(labels),alpha=.025)
```



∴ The effect of the interaction AB is significant when the effect of AB is -50.5229.

Question 5. An article in Quality and Reliability Engineering International (2010, Vol. 26, pp. 223-233) presents a 2^5 factorial design. The experiment is shown in the following table:

A	B	C	D	E	y
-1	-1	-1	-1	-1	8.11
1	-1	-1	-1	-1	5.56
-1	1	-1	-1	-1	5.77
1	1	-1	-1	-1	5.82
-1	-1	1	-1	-1	9.17
1	-1	1	-1	-1	7.8
-1	1	1	-1	-1	3.23
1	1	1	-1	-1	5.69
-1	-1	-1	1	-1	8.82
1	-1	-1	1	-1	14.23
-1	1	-1	1	-1	9.2
1	1	-1	1	-1	8.94
-1	-1	1	1	-1	8.68
1	-1	1	1	-1	11.49
-1	1	1	1	-1	6.25
1	1	1	1	-1	9.12
-1	-1	-1	-1	1	7.93
1	-1	-1	-1	1	5
-1	1	-1	-1	1	7.47
1	1	-1	-1	1	12
-1	-1	1	-1	1	9.86
1	-1	1	-1	1	3.65
-1	1	1	-1	1	6.4
1	1	1	-1	1	11.61
-1	-1	-1	1	1	12.43
1	-1	-1	1	1	17.55
-1	1	-1	1	1	8.87
1	1	-1	1	1	25.38
-1	-1	1	1	1	13.06
1	-1	1	1	1	18.85
-1	1	1	1	1	11.78
1	1	1	1	1	26.05

```
A <- rep(c(-1,1), 16)
B <- rep(c(-1,-1,1,1),8)
```

```
C <- rep(c(rep(-1,4),rep(1,4)),4)
D <- rep(c(rep(-1,8),rep(1,8)),2)
E <- c(rep(-1,16),rep(1,16))
y <- 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
q5_data <- data.frame(A,B,C,D,E,y)
```

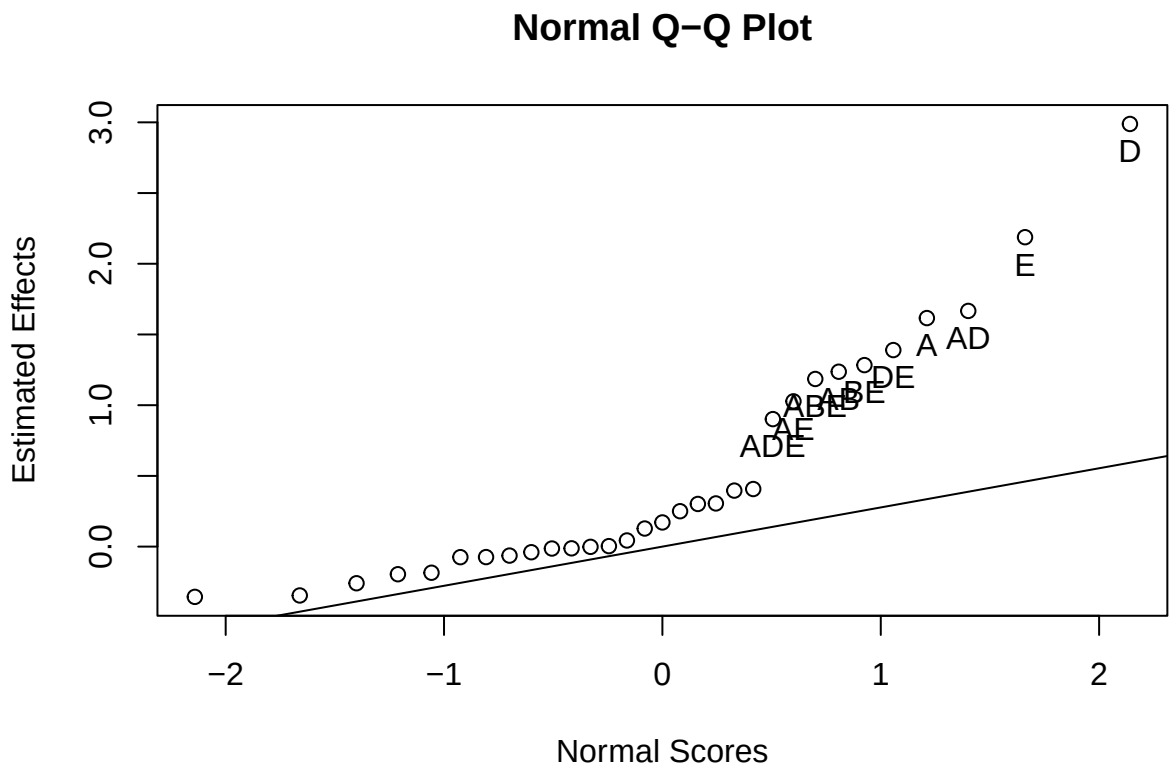
(a) Analyze the data from this experiment. Identify the significant factors and interactions.

```
#Using ANOVA
res.lm<-lm(y~A*B*C*D*E, data=q5_data)
summary(res.lm)
```

```
##
## Call:
## lm(formula = y ~ A * B * C * D * E, data = q5_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            2.988438        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          1.666562        NaN     NaN      NaN
## B:D         -0.013438        NaN     NaN      NaN
## C:D          0.003437        NaN     NaN      NaN
## A:E          1.027187        NaN     NaN      NaN
## B:E          1.283437        NaN     NaN      NaN
## C:E          0.301562        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        0.305313        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
```

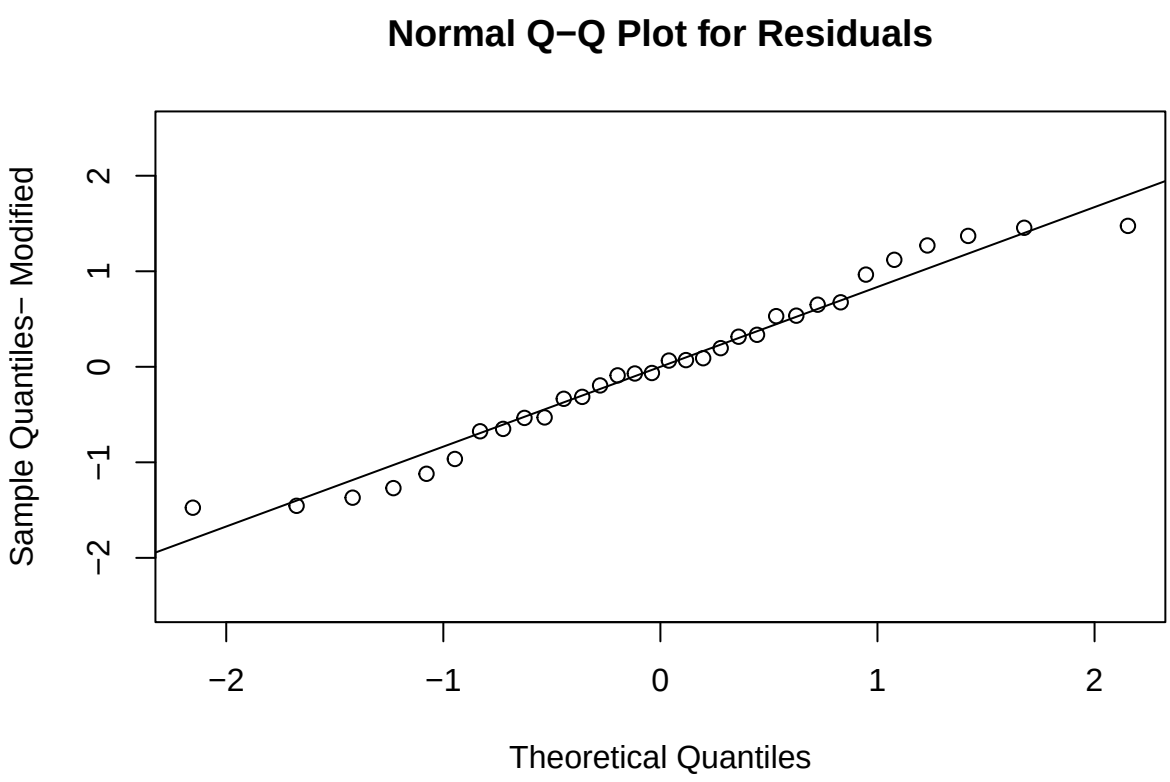
```
fullnormal(coef(res.lm)[-1],alpha=.05)
```



Factors and Interactions A, D, E, AB, AD, AE, BE, DE, ABE and ADE are significant.

(b) Analyze the residuals from this experiment. Are there any indications of model inadequacy or violations of the assumptions?

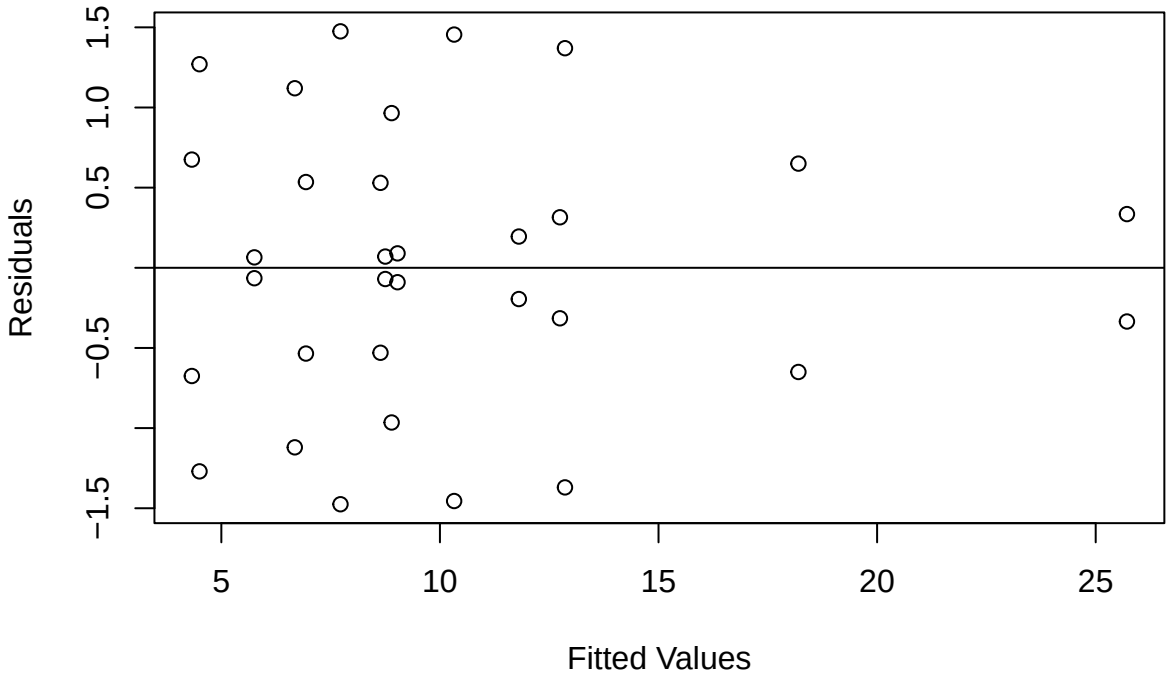
```
#residual analysis
#Normality: drop C (the reduced model)
res.aov = aov(y~A*B*D*E, data=q5_data)
residuals=res.aov$residuals
qqnorm(residuals, ylim=c(min(residuals)-1,max(residuals)+1), main = "Normal Q-Q Plot for Residuals",
       xlab = "Theoretical Quantiles", ylab = "Sample Quantiles- Modified",
       plot.it = TRUE, datax = FALSE)

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




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

```
##
##  Shapiro-Wilk normality test
##
## data:  residuals
## W = 0.96505, p-value = 0.3751
```

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



∴ Since the p-value from Shapiro Wilks test is $0.3751 > 0.05$, we fail to reject the null hypothesis that the residuals are normally distributed. So there's no indication of violation of the normality assumption. There might be a violation of the constant variance assumption since higher fitted values seem to have smaller variance.

(c) 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. Compare your results with those obtained in part (a).

```
# drop factor C
res.aov<-aov(y~A*B*D*E, data=q5_data)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## A            1  83.56   83.56  57.233 1.14e-06 ***
## B            1   0.06    0.06   0.041 0.841418
## D            1 285.78  285.78 195.742 2.16e-10 ***
## E            1 153.17  153.17 104.910 1.97e-08 ***
## A:B          1  48.93   48.93  33.514 2.77e-05 ***
## A:D          1  88.88   88.88  60.875 7.66e-07 ***
## B:D          1   0.01    0.01   0.004 0.950618
## A:E          1  33.76   33.76  23.126 0.000193 ***
## B:E          1  52.71   52.71  36.103 1.82e-05 ***
## D:E          1  61.80   61.80  42.328 7.24e-06 ***
## A:B:D        1   3.82    3.82   2.613 0.125501
```

```
## A:B:E        1  44.96   44.96  30.794 4.40e-05 ***
## A:D:E        1  26.01   26.01  17.815 0.000650 ***
## B:D:E        1   0.05    0.05   0.035 0.854935
## A:B:D:E      1   5.31    5.31   3.634 0.074735 .
## Residuals   16  23.36    1.46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

∴ Factor C is not important. The remaining design is a 2^4 factorial design with 2 replicates. And the same factors and interactions (A, D, E, AB, AD, AE, BE, DE, ABE and ADE) are significant as part (a).

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

```r
# Add fitted as a column to the data frame
q5_data$Fitted = res.aov$fitted.values
q5_data[res.aov$fitted.values==max(res.aov$fitted.values),] #also can use "which"
```

```
##    A B  C D E     y Fitted
## 28 1 1 -1 1 1 25.38 25.715
## 32 1 1  1 1 1 26.05 25.715
```

The maximum fitted value is 25.715 and the settings of the active factors that maximize the predicted response are (A=1, B=1, C=-1, D=1, E=1) and (A=1, B=1, C=1, D=1, E=1).