STAT 212- Homework #5

Part A.

2.a

```
cellphone <- read.table("CellPhoneRL.txt", header = TRUE)
Y=cellphone$y
S=cellphone$S
C=cellphone$C
fit=aov(Y~S*C)
anova(fit)
```

Analysis of Variance Table

Response: Y

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
S	2	2.42453	1.21227	146.1735	< 2.2e-16	***
C	4	0.76724	0.19181	23.1282	1.345e-11	***
S:C	8	0.12070	0.01509	1.8192	0.09103	.
Residuals	60	0.49760	0.00829			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Test-statistic for interaction: $F = 1.8192$
p-value: 0.0910
p-value is greater than 0.01 level of significance, so we fail to reject the null hypothesis ##that there is no interaction

Test-statistic for S: $F = 146.1735$
p-value: $< 2.2e^{-16}$
p-value is less than the .01 level of significance, so we reject the null
hypothesis for effect S.

Test-statistic for C: $F = 23.1282$
p-value: $1.345e^{-11}$
p-value is less than the .01 level of significance, so we reject the null
hypothesis for effect C

```
##2.b
TukeyHSD(fit)


   Tukey multiple comparisons of means
     95% family-wise confidence level

Fit: aov(formula = Y ~ S * C)

$S
        diff        lwr         upr p adj
L-H -0.4388 -0.5007017 -0.3768983     0
M-H -0.2520 -0.3139017 -0.1900983     0
M-L  0.1868  0.1248983  0.2487017     0

$C
               diff          lwr          upr        p adj
B2-B1  0.27533333  0.181809954  0.36885671 0.0000000
B3-B1  0.15800000  0.064476621  0.25152338 0.0001238
B4-B1  0.24533333  0.151809954  0.33885671 0.0000000
B5-B1  0.08800000 -0.005523379  0.18152338 0.0746708
B3-B2 -0.11733333 -0.210856713 -0.02380995 0.0069794
B4-B2 -0.03000000 -0.123523379  0.06352338 0.8949578
B5-B2 -0.18733333 -0.280856713 -0.09380995 0.0000049
B4-B3  0.08733333 -0.006190046  0.18085671 0.0782490
B5-B3 -0.07000000 -0.163523379  0.02352338 0.2314578
B5-B4 -0.15733333 -0.250856713 -0.06380995 0.0001330
```

From the output and resulting p-values, we can see that pairs of cell phones that are
significantly different at a .01 level of significance based on their respective main effects ## are:
B2-B1, B3-B1, B4-B1, B3-B2, B5-B2, and B5-B4

Furthermore, we see that the signal pairs that are significantly different at a .01 level of
significance based on their main effects are: L-H, M-H, and M-L

```
##2.d
interaction.plot(C, S, Y, xlab = "Type of Cellphone", ylab =
"SAR Level (W/kg)", trace.label = "Signal Strength", main =
"Interaction Plot for Cellphone Radiation Data")
```

Interaction Plot for Cellphone Radiation Data
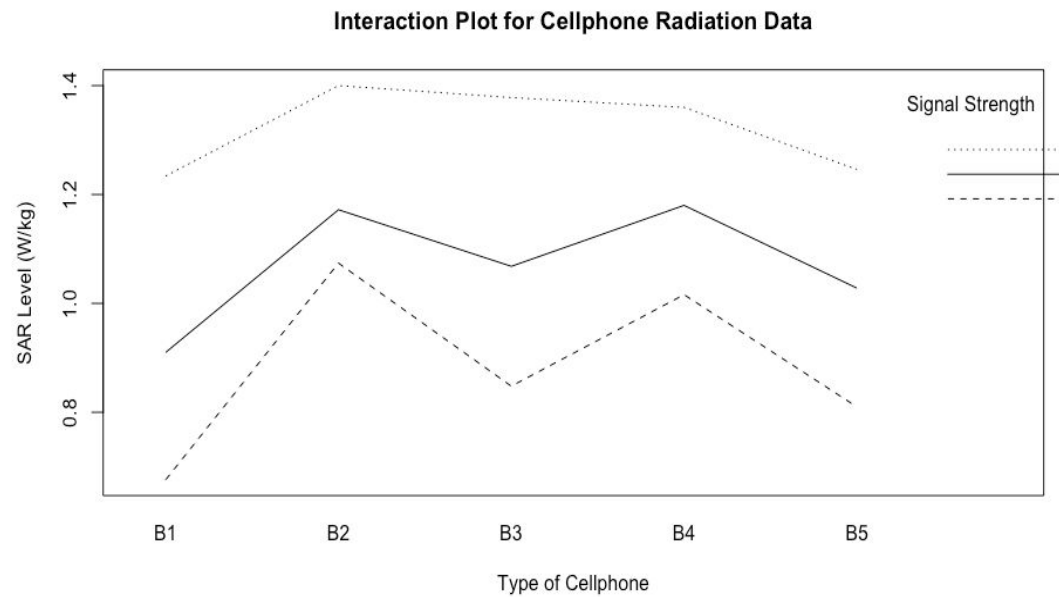
## The signals are High (.......), Medium (_____) and Low(---------)
None of the line segments intersect, but there are indications of a weak interaction between phone type B1 and the Low signal strength, as well as a very weak interaction between B1 and High signal strength. Furthermore, we can see an interaction between B3 and High signal strength, as well as interactions with B2 and both High and Low signals. The rest of the pairing indicate a weak, if any, interaction.


```
interaction.plot(S, C, Y, xlab = "Signal Strength", ylab = "SAR
Level (W/kg)", trace.label = "Cellphone Type", main =
"Interaction Plot for Cellphone Radiation Data")
```

**Interaction Plot for Cellphone Radiation Data**
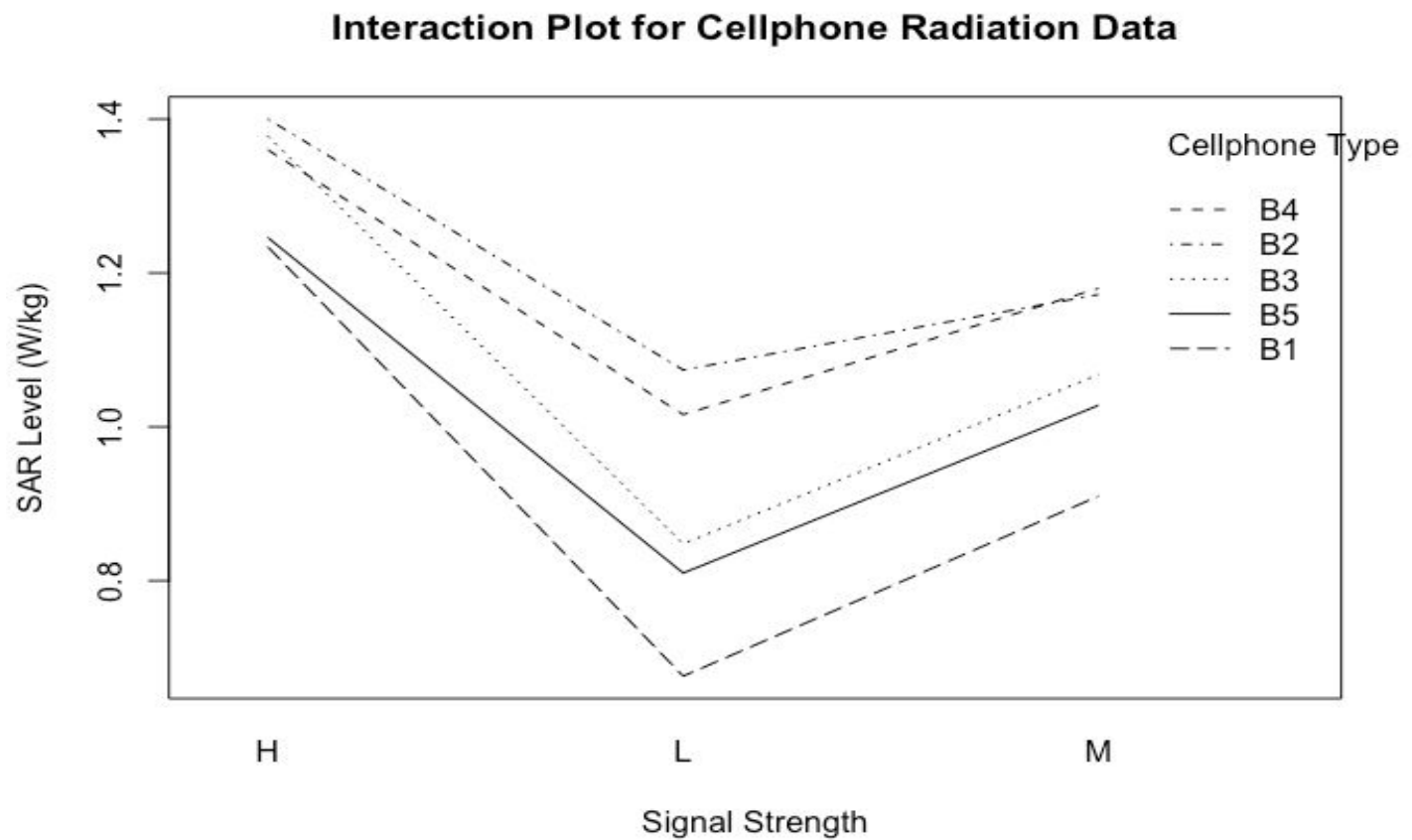
From this version of the interaction plot, we see that there is a slight interaction between High signal strength and the B1 phone as well as a stronger interaction between the High signal strength and the B3 phone. We can also see an interaction between the low signal strength and the B2 phone. And it appears that the rest of the pairing have a weak, if any, interaction.

```
## residuals plot
plot(fit, which = 1)
```

Residuals vs Fitted

aov(Y ~ S * C)
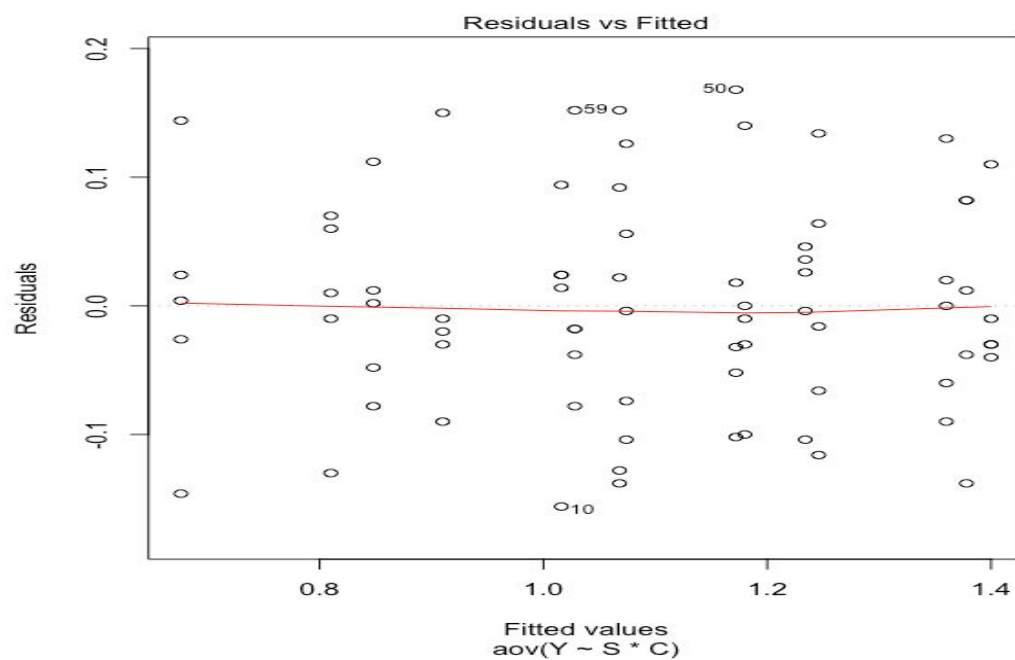
Our residuals plot suggests that our assumption of homoscedasticity is likely correct, because there does not appear to be any kind of pattern amongst the residuals.

```
#Normal Q-Q Plot
plot (fit, which = 2)
qqline(rstandard(fit), col = "RED", lwd = 2)
```


Normal Q-Q

aov(Y ~ S * C)
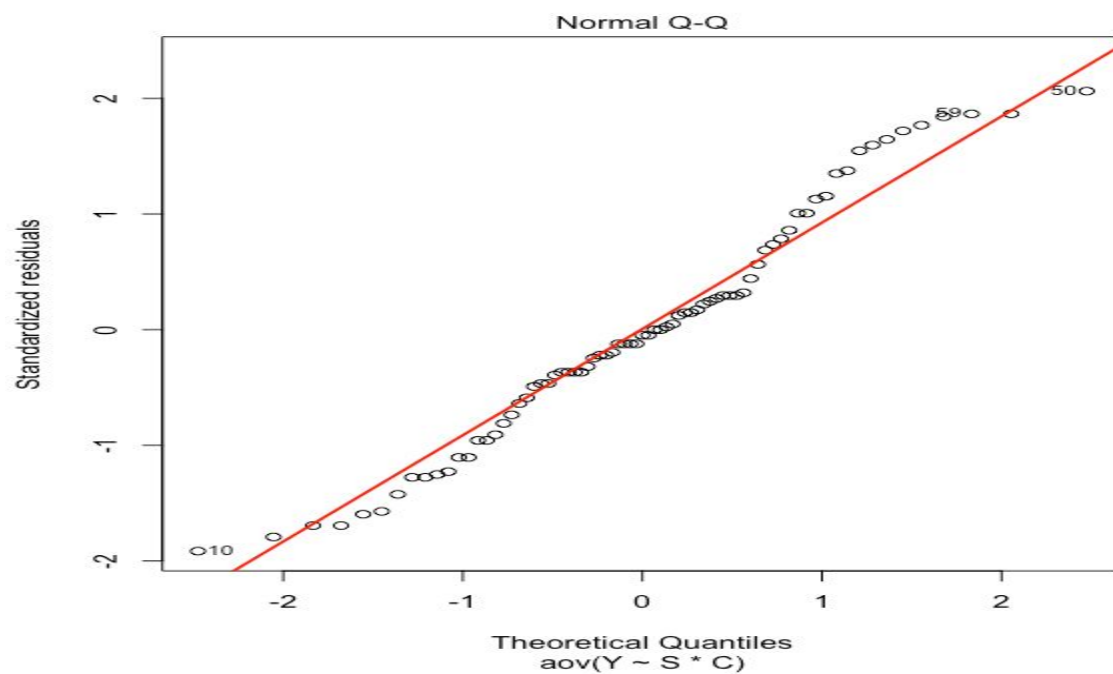
Our Q-Q plot suggests that our assumption of normality is likely correct, because there does not appear to be any specific kind of trend happening with the data points, and they are distributed pretty evenly along the qq-line.

Part B.

##3.a and 3.b

$$\text{model}: X_{ij} = \mu + a_i + b_j + e_{ij}, \quad i = 1, 2, 3, 4, \quad j = 1, 2, 3, \cdots, 8$$

$$\sum a_i = 0 \quad \text{and} \quad var(b_j) = \sigma_b^2$$

↓ Treatment effects are represented by a_i
- Blocking effects are represented by b_j

Null - $H_0: a_1 = a_2 = a_3 = a_4$

Alternative - H_a: at least one a_i is different

##3.d
```
reactions = read.table("PilotReacTimes.txt", header = TRUE)
attach(reactions)
fit = aov(times ~ design + pilot, data = reactions)
anova(fit)
```
```
Analysis of Variance Table

Response: times
          Df  Sum Sq Mean Sq F value    Pr(>F)
design     3  8.3729  2.7910  9.1936 0.0004409 ***
pilot      7 29.1090  4.1584 13.6982 1.495e-06 ***
Residuals 21  6.3751  0.3036

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

##
 The resulting p-values for design and pilot are 4.409e^-4 and 1.495e^-6, respectively. Both values are less than the .01 level of significance, thus we reject our null hypothesis that all the treatment effects are the same.

Part C
Problem 4
```
TukeyHSD(fit, "design", conf.level=0.99)
    Tukey multiple comparisons of means
      99% family-wise confidence level

Fit: aov(formula = times ~ design + pilot, data = reactions)

$design
         diff          lwr         upr      p adj
B-A -0.83875 -1.8099955   0.1324955 0.0290281
C-A -0.00875 -0.9799955   0.9624955 0.9999881
D-A -1.16375 -2.1349955  -0.1925045 0.0019931
C-B  0.83000 -0.1412455   1.8012455 0.0310802
D-B -0.32500 -1.2962455   0.6462455 0.6458100
D-C -1.15500 -2.1262455  -0.1837545 0.0021465
```
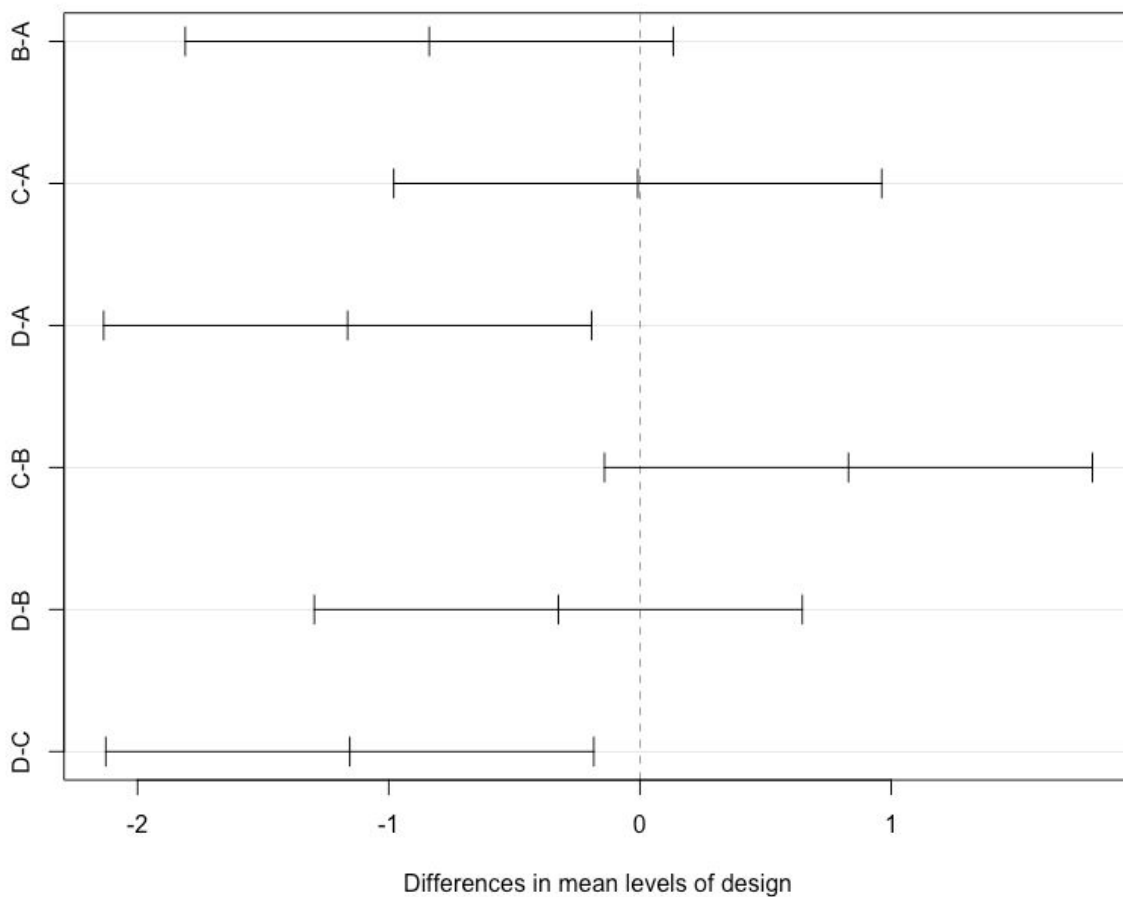
```
plot(TukeyHSD(fit, "design", conf.level = 0.99))
```

99% family-wise confidence level

Differences in mean levels of design

Based on both the `TukeyHSD` output and the confidence level plot, we see that the only design pairs that do not contain zero and thus differ significantly are D-A and D-C...the rest of the pairs are not significantly different.

##5.a
Fabric is the blocking factor because, given a random block design, the blocking factor is the factor that observations are being made on.

##5.b
H0 = None of the chemicals are different in terms of the fabrics' mean strength.
Ha = at least one of the chemicals is different in terms of the fabrics' mean strength.

	DF	Sum of Squares	Mean squares	F - Value	P
Treatment	3	2.4815	0.827167	19.424681	6.717016e-05
Block	4	5.4530	1.36325	32.013724	2.566073e-06
Residuals	12	0.5110	0.0425833		
		8.4455			

n = 20, because for each of the 4 chemicals being tested, there are 5 samples of fabric.
Treatment D.O.F = k - 1 = 4 - 1 = 3
Blocking D.O.F. = b - 1 = 5 - 1 = 4
SSB = Blocking Sum of Squares
Treatment Sum of Squares (SSTr) = SST - (SSB + SSR) = 8.4455 - (0.511 + 5.4530) = 2.4815
MSTr = SSTr/(k-1) = 2.4815/3 = 0.827167
MSB = SSB / (b - 1) = 5.4530/4 = 1.36325
MSR = SSR / (k - 1)(b - 1) = 0.511/12 = 0.0425833
Treatment F-Value = MSTr / MSR = 0.827167 / 0.0425833 = 19.424681
1 - pf(19.424681, 3, 12)

```
[1] 6.717016e-05
```

Blocking F-Value = MSB / MSR = 1.36325 / 0.0425833 = 32.013724
1 - pf(32.013724, 4, 12)

```
[1] 2.566073e-06
```

The resulting p-values are both less than the 0.05 level of significance, thus we reject our null hypothesis that none of the chemicals are different in terms of the fabrics' mean strengths.