Stat631 Extra Credit

Brandon Keck

```r
library(readr)
library(dplyr)
library(ggplot2)
library(phia)
library(emmeans)
library(car)
```

```r
coffee <- read.table("coffee.txt")
head(coffee)
```

```
   Coffee Setting ReactionTime
1 Regular Library          207
2   Decaf Library          242
3 Regular    Home          234
4   Decaf    Home          236
5 Regular    Café          213
6   Decaf    Café          252
```

1. I would like this extra credit to be added to option 1: Midterm 2

2 (a)

2 factors: Coffee and Setting

Treatments: 2 (Coffee) x 3 (Setting) = 6 treatments.

```r
coffee %>%
  group_by(Coffee, Setting) %>%
  summarise(n = n())
```

```
`summarise()` has grouped output by 'Coffee'. You can override using the
`.groups` argument.
```

```
# A tibble: 6 x 3
# Groups:   Coffee [2]
  Coffee  Setting      n
  <chr>   <chr>    <int>
1 Decaf   Café         5
2 Decaf   Home         5
3 Decaf   Library      5
4 Regular Café         5
5 Regular Home         5
6 Regular Library      5
```

Replications per treatment = 5

There are 6 treatment groups, each with 5 replications, totaling 30 observations. Therefore, the design is balanced.

2 (b)

```
coffee <- within(coffee, {fCoffee = factor(Coffee); fSetting = factor(Setting)})
head(coffee)
```

```
   Coffee Setting ReactionTime fSetting fCoffee
1 Regular Library          207  Library Regular
2   Decaf Library          242  Library   Decaf
3 Regular    Home          234     Home Regular
4   Decaf    Home          236     Home   Decaf
5 Regular    Café          213     Café Regular
6   Decaf    Café          252     Café   Decaf
```
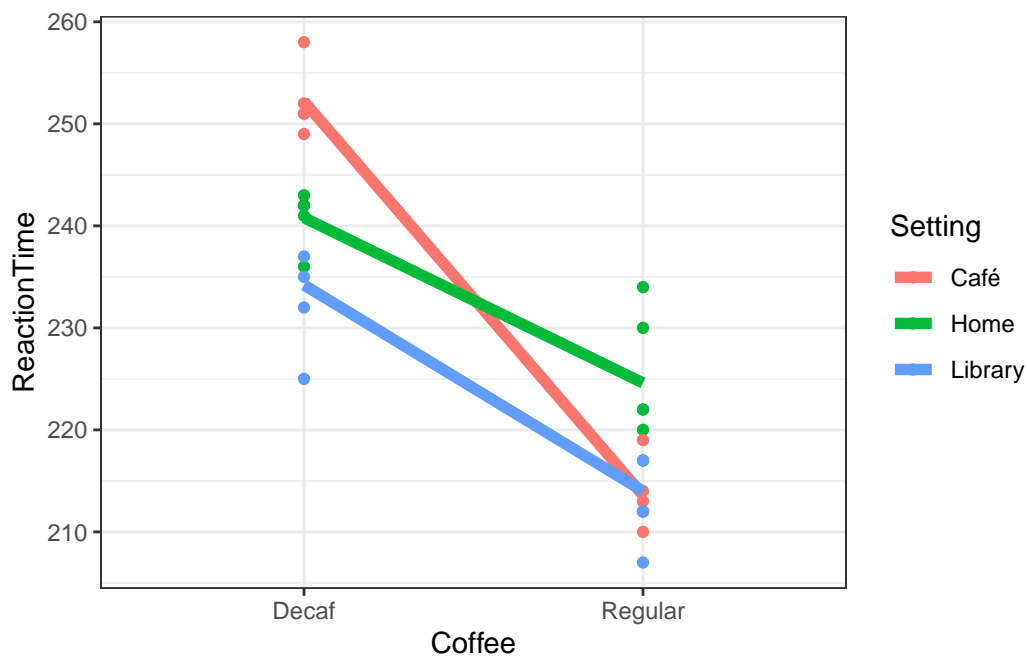
```
model1 <- lm(ReactionTime ~ fCoffee*fSetting, data = coffee)
lsmFR <- lsmeans(model1, ~ fCoffee*fSetting)
lsmFR
```

```
 fCoffee fSetting lsmean   SE df lower.CL upper.CL
 Decaf   Café        252 2.26 24      248      257
 Regular Café        214 2.26 24      209      218
```

```
Decaf    Home         241 2.26 24       236      245
Regular  Home         225 2.26 24       220      229
Decaf    Library      234 2.26 24       230      239
Regular  Library      214 2.26 24       209      219
```

```
Confidence level used: 0.95
```

```
ggplot(coffee, aes(x =Coffee, y = ReactionTime, color = Setting)) +
  geom_point() +
  stat_summary(fun = mean, geom = "line",
               aes(group = Setting), linewidth = 2) +
  theme_bw()
```


The interaction plot suggests there may be an interaction between the type of coffee type and study location. Specifically, the effect of coffee type on reaction time varies across different settings. This effect appears to be strongest in the Cafe, where the difference in reaction times between Decaf and Regular coffee is most pronounced. The differences are smaller in the Home and Library setting. Since the lines in the plot are not parallel, this indicates that the effect of coffee type is not consistent across the different study settings.

2 (c)

```
anova(aov(ReactionTime ~ fCoffee * fSetting, data = coffee))
```

```
Analysis of Variance Table

Response: ReactionTime
                Df Sum Sq Mean Sq F value    Pr(>F)
fCoffee          1 4687.5  4687.5 182.986 1.007e-12 ***
fSetting         2  504.8   252.4   9.853 0.0007517 ***
fCoffee:fSetting 2  713.6   356.8  13.928 9.657e-05 ***
Residuals       24  614.8    25.6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
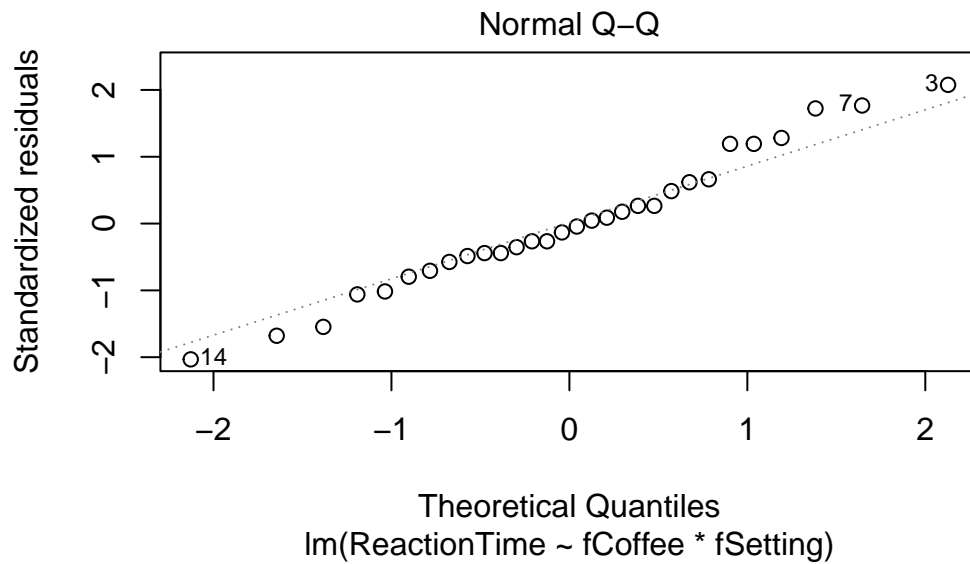
From the ANOVA table all terms in the model are statistically significant. The main effect of coffee is highly significant which indicates that the reaction time is significantly different between Decaf and Regular coffee. We also can see that the main effect of setting is also significant. This shows us that study setting has an impact on reaction time as well. Finally, the interaction between Coffee and Setting is also significant. This confirms the effect of Coffee type and it depends on the study setting which does align with what we saw in the interaction plot. This matches our interaction plot above.
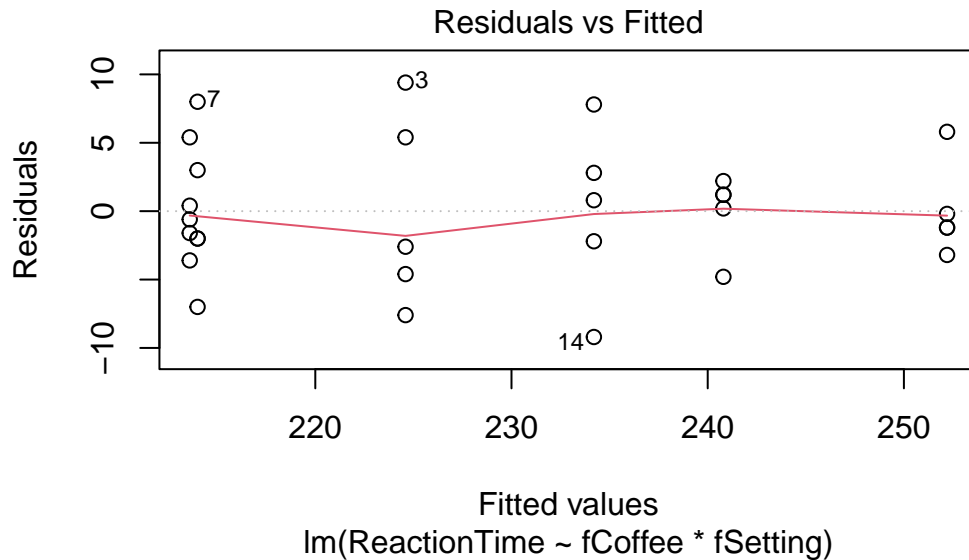
2 (d)

```
# Q-Q plot
plot(model1, 2)
```

Normal Q–Q

Theoretical Quantiles
lm(ReactionTime ~ fCoffee * fSetting)

```r
# Shapiro test for Normality
shapiro.test(model1$residuals) # p-value > 0.05 means that Normality satisfied
```

```
	Shapiro-Wilk normality test

data:  model1$residuals
W = 0.97976, p-value = 0.8192
```

```r
# Residuals vs Fitted values
plot(model1, 1)
```

Residuals vs Fitted

Fitted values
lm(ReactionTime ~ fCoffee * fSetting)

```
# Levene Test
leveneTest(ReactionTime ~ fCoffee*fSetting, data = coffee) # p-value > 0.05 Equal Variance
```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  5  0.9256 0.4817
      24
```

The Normal Q-Q plot shows that most of the residuals follow the diagonal line which indicates the normality assumption is satisfied. While there is some slight deviation at the upper end with a few outliers, the pattern is not too concerning. The Shapiro-Wilk test confirms this with a p-value of 0.8192, which is greater than our significance level of $\alpha = 0.05$ which suggest that we fail to reject the null hypothesis of normality. The residuals vs. Fitted plot shows no clear pattern or funnel shape; residuals appear to be randomly scattered with constant spread across fitted values. The levene's test also confirms this with a p-value of 0.4817, which again is greater than our significance level of $\alpha = 0.05$ Equal variance assumption is satisfied. Thus the ANVOA model and results from part c are valid and reliable.
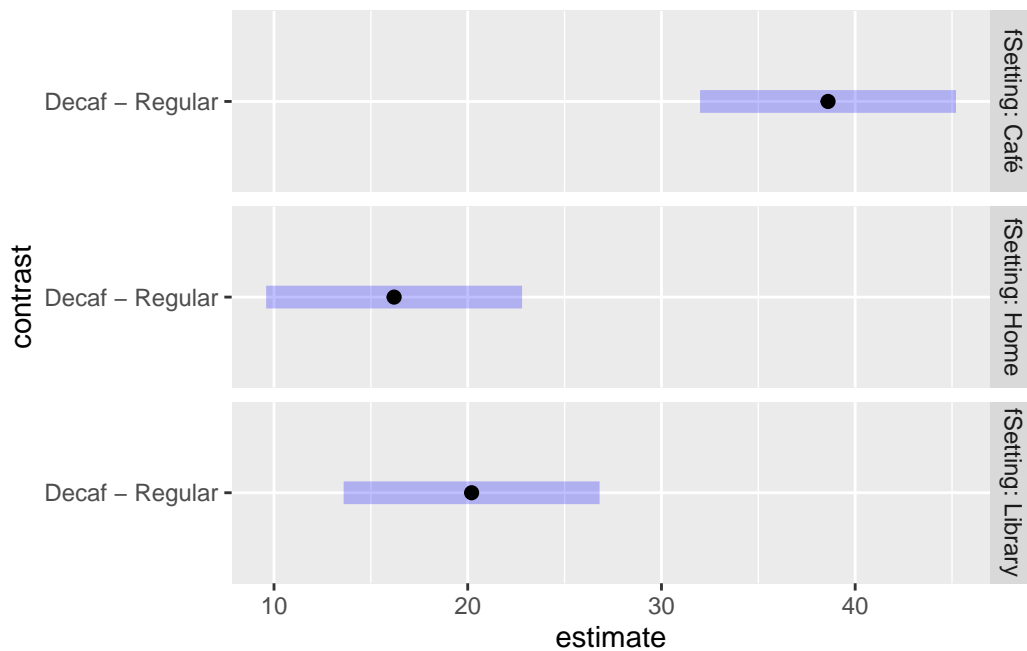
2 (e)

```
em <- emmeans(model1, pairwise ~ fCoffee | fSetting)
em$contrasts
```

```
fSetting = Café:
 contrast        estimate  SE df t.ratio p.value
 Decaf - Regular     38.6 3.2 24  12.059  <.0001

fSetting = Home:
 contrast        estimate  SE df t.ratio p.value
 Decaf - Regular     16.2 3.2 24   5.061  <.0001

fSetting = Library:
 contrast        estimate  SE df t.ratio p.value
 Decaf - Regular     20.2 3.2 24   6.310  <.0001
```

```
plot(em$contrasts)
```


From the pairwise comparison plot, we observe that in all three setting,s students who consumed decaf coffee had significantly slower reaction times compared to those who drank regular

coffee. The largest difference occurred in the Cafe setting, with a difference of 38.6 milliseconds. These results align with the earlier findings from both the interaction plot and the ANOVA table, confirming a significant interaction between coffee type and study setting on reaction time.