HW3

Hunter Schwartz

1.

```r
question <- read.table("CH16PR08.txt", header=F)
colnames(question) <- c("Rate", "Color", "ID")

rate <- question$Rate
color <- factor(question$Color, labels=c("blue", "green", "orange"))
id <- factor(question$ID)
```
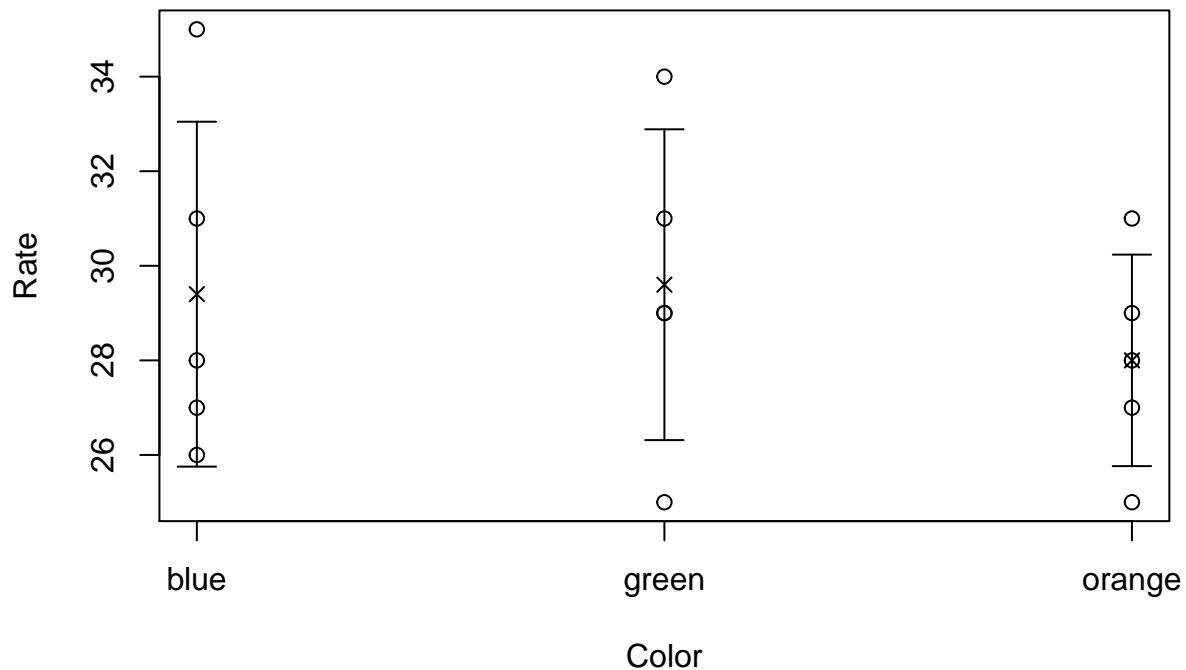
2.

```r
rate.mean <- aggregate(rate, by=list(color), mean)$x
rate.sd <- aggregate(rate, by=list(color), sd)$x

stripchart(rate ~ color, vertical=T, pch=1,
           main="Response Rate vs. Flier Color (with standard deviation)",
           ylab="Rate", xlab="Color")
points(rate.mean, pch=4)
arrows(1:3, rate.mean - rate.sd, 1:3, rate.mean + rate.sd, angle=90, code=3, length=.1)
```

Response Rate vs. Flier Color (with standard deviation)

The means (denoted 'x') appear to be similar between groups. The bars represent one standard deviation in either direction of the mean of points within each group; the standard deviation appears to vary slightly between the groups, in particular it is smaller in the 'orange' group, but perhaps not significantly.

3.

```
model <- aov(rate ~ color)
anova(model)

## Analysis of Variance Table
##
## Response: rate
##           Df Sum Sq Mean Sq F value Pr(>F)
## color      2    7.6     3.8  0.3918 0.6842
## Residuals 12  116.4     9.7
```

4.

```
print( aggregate(model$residuals, by=list(color), FUN=sum) )

##   Group.1             x
## 1    blue -1.998401e-15
## 2   green  8.881784e-16
```

```
## 3   orange  1.554312e-15
print( sum(model$residuals) )
```

```
## [1] 4.440892e-16
```

The sum of residuals within each group sum to zero (and by extension, so does the sum of all the residuals).

5.

```
alpha <- .10
qf(1 - alpha, df1=2, df2=12)
```

```
## [1] 2.806796
```

$H_0 : \mu_1 = \mu_2 = \mu_3$
$H_a : \mu_1 \neq \mu_2$ OR $\mu_1 \neq \mu_3$ OR $\mu_2 \neq \mu_3$
We reject H_0 if $F_{obs} > 2.806796$. Since by the ANOVA analysis $F_{obs} = 0.3918$, we fail to reject the null hypothesis and conclude that any difference between the group means may be due to random chance.

6.

Although no significant difference was found between the groups based on color, we cannot necessarily conclude that the executive's statement is correct, as a white color 'treatment' was not included in the experiment. Consider a hypothetical case where few people respond to surveys written on white paper, but many more respond to the colored surveys, no matter what color, just by virtue of their being colored. This would be consistent with the results of our experiment yet the executive would be wrong.

7.

```
  model.mse <- sum(model$residuals ^ 2) / model$df.residual
  model.se <- rep( sqrt( model.mse/5 ), 3 ) # n=5 for each group
  model.cc.90 <- qt(1 - alpha/2, df=12) # 90% C.I.
rate.allow <- model.se * model.cc.90

stripchart(rate ~ color, vertical=T, pch=1,
           main="Response Rate vs. Flier Color (with 90% C.I.'s)",
           ylab="Rate", xlab="Color")
points(rate.mean, pch=4)
arrows(1:3, rate.mean - rate.allow, 1:3, rate.mean + rate.allow, angle=90, code=3, length=.1)
```

Response Rate vs. Flier Color (with 90% C.I.'s)
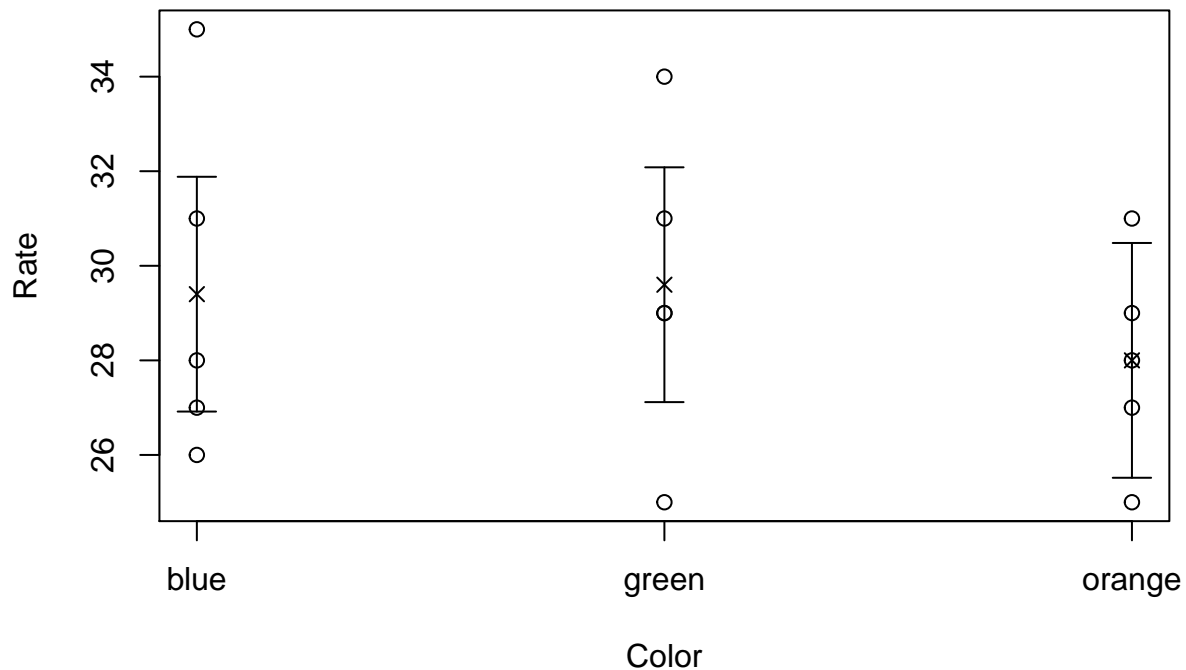
The intervals of each group seem wide enough to contain the the means of the other groups, which may suggest there is not a significant difference between the means based on color.

8.

```
blue.mean <- rate.mean[1]
orange.mean <- rate.mean[3]

diff <- orange.mean - blue.mean
diff.se <- sqrt( model.mse / (1/5 + 1/5) )
diff.allow <- diff.se * model.cc.90

ci <- c(diff - diff.allow, diff + diff.allow)
print("90% C.I. for difference between mean of blue and orange groups:"); print(ci)
```

```
## [1] "90% C.I. for difference between mean of blue and orange groups:"
```

```
## [1] -10.176748   7.376748
```

9.

Part 5 concluded that the group means were not significantly different from one another. *Part 7* agrees with this since the means of the other groups were included in each group's own confidence interval. *Part 8* is

consistent with both of these, as 0 lies within the confidence interval of the difference between the blue and orange groups, in other words their means must be presumed to be equal.