STAT 222 Spring 2021 HW6

Matthew Zhao

```
## Warning: package 'knitr' was built under R version 3.6.2
```

```
wtlift = read.table("http://www.stat.uchicago.edu/~yibi/s222/weight.lifting.txt", h=T)
```

Q1 — 7 points

```
library(mosaic)
mean(rate ~ 1, data=wtlift)
##       1
## 31.1364
mean(rate ~ A, data=wtlift)
##       1       2
## 32.2727 30.0000
mean(rate ~ B, data=wtlift)
##       1       2       3
## 26.5909 32.2273 34.5909
mean(rate ~ B+A, data=wtlift)
##     1.1     2.1     3.1     1.2     2.2     3.2
## 27.9091 33.0000 35.9091 25.2727 31.4545 33.2727
```

Using the code above, we estimate $\mu = 31.1364$.

$\alpha_i = \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdots}$
$\alpha_1 = 32.2727 - 31.1364 = 1.1363$, $\alpha_2 = 30.0000 - 31.1364 = -1.1364$

Remark (not required):

```
wtlift$A = as.factor(wtlift$A)
wtlift$B = as.factor(wtlift$B)
contrasts(wtlift$A) = contr.sum(2)
contrasts(wtlift$B) = contr.sum(3)
lm1 = lm(rate ~ A * B, data=wtlift)
lm1$coef
## (Intercept)          A1          B1          B2       A1:B1       A1:B2
##   31.136364    1.136364   -4.545455    1.090909    0.181818   -0.363636
```

Q2 — 8 points

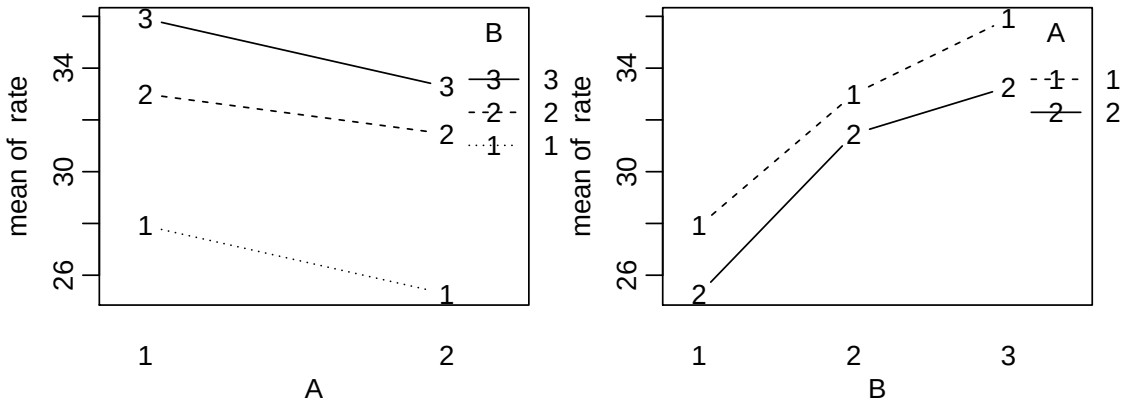
```
SST = var(wtlift$rate)*(66-1)
```

You may verify your calculation in R as follows:

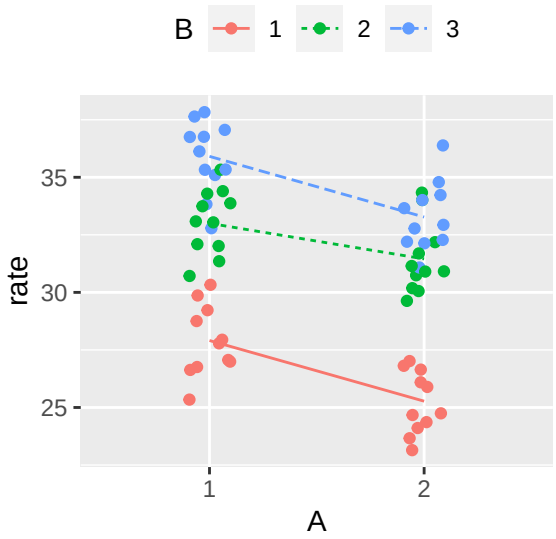
```
lm1 = lm(rate ~ A * B, data=wtlift)
options(scipen=6, digits=8)  # increasing the number of digits in the output
anova(lm1)
## Analysis of Variance Table
##
## Response: rate
##           Df  Sum Sq Mean Sq  F value         Pr(>F)
## A          1  85.227  85.227  39.0625 0.000000046796
## B          2 743.273 371.636 170.3333     < 2.22e-16
## A:B        2   4.364   2.182   1.0000        0.37393
## Residuals 60 130.909   2.182
options(scipen=6, digits=5)  # restoring the digit setting back to 5
```

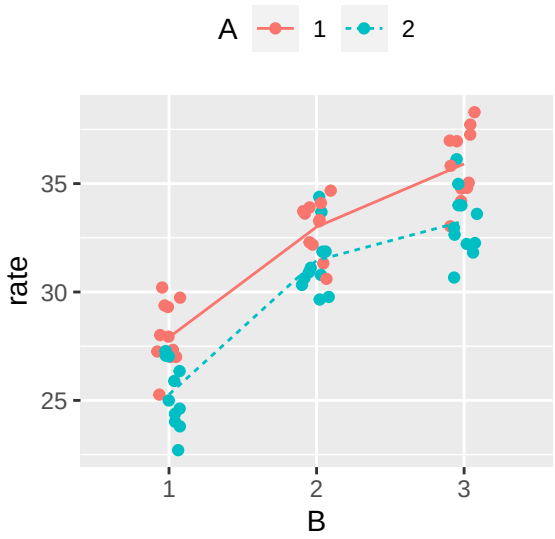
Q3 — 3 points

```
par(mai=c(.6,.6,.1,.3),mgp=c(2,.6,0))
with(wtlift, interaction.plot(A, B, rate, type="b"))
with(wtlift, interaction.plot(B, A, rate, type="b"))
```




```
wtlift$A = as.factor(wtlift$A)
wtlift$B = as.factor(wtlift$B)
library(ggplot2)
ggplot(wtlift, aes(x=A,y=rate, color=B))+
  geom_point(position = position_jitter(width = .1))+
  stat_summary(fun="mean",geom="line",aes(group=B,linetype=B)) +
  theme(legend.position="top")
ggplot(wtlift, aes(x=B,y=rate,color=A))+
  geom_point(position = position_jitter(width = .1))+
  stat_summary(fun="mean",geom="line",aes(group=A,linetype=A)) +
  theme(legend.position="top")
```





Q4 — 4 points

Q5 — 3 points