ch19output_2

```
############################ Example: Disk Drive Service#################################################
#page 868 Problem 19.16, the staff of a service center for electornic equipment includes
#three technicians who specilize in repairing three widely used makes of disk drives for
#destop computer.
#We want to study the effects of technician (factor A)
# and make of disk drive (factor B) on the service time y.

disk<- read.table(file="~/Desktop/jenn/teaching/stat445545/data/CH19PR16.txt",
                  col.names=c("y","tech","make","obs"))
nt<-nrow(disk)
nt
```

```
## [1] 45
```

```
disk
```

```
##     y tech make obs
## 1  62    1    1   1
## 2  48    1    1   2
## 3  63    1    1   3
## 4  57    1    1   4
## 5  69    1    1   5
## 6  57    1    2   1
## 7  45    1    2   2
## 8  39    1    2   3
## 9  54    1    2   4
## 10 44    1    2   5
## 11 59    1    3   1
## 12 53    1    3   2
## 13 67    1    3   3
## 14 66    1    3   4
## 15 47    1    3   5
## 16 51    2    1   1
## 17 57    2    1   2
## 18 45    2    1   3
## 19 50    2    1   4
## 20 39    2    1   5
## 21 61    2    2   1
## 22 58    2    2   2
## 23 70    2    2   3
## 24 66    2    2   4
## 25 51    2    2   5
## 26 55    2    3   1
## 27 58    2    3   2
## 28 50    2    3   3
## 29 69    2    3   4
## 30 49    2    3   5
## 31 59    3    1   1
## 32 65    3    1   2
## 33 55    3    1   3
## 34 52    3    1   4
## 35 70    3    1   5
```

```
## 36 58    3    2   1
## 37 63    3    2   2
## 38 70    3    2   3
## 39 53    3    2   4
## 40 60    3    2   5
## 41 47    3    3   1
## 42 56    3    3   2
## 43 51    3    3   3
## 44 44    3    3   4
## 45 50    3    3   5
```

```
disk$tech <- factor(disk$tech)
disk$make <- factor(disk$make)
attach(disk)

##summary statistics
library(plyr)
# Calculate the cell means for each (tech, make) combination
tech.mean  <- ddply(disk, .(tech), summarise, m = mean(y))
tech.mean
```

```
##   tech        m
## 1    1 55.33333
## 2    2 55.26667
## 3    3 56.86667
```

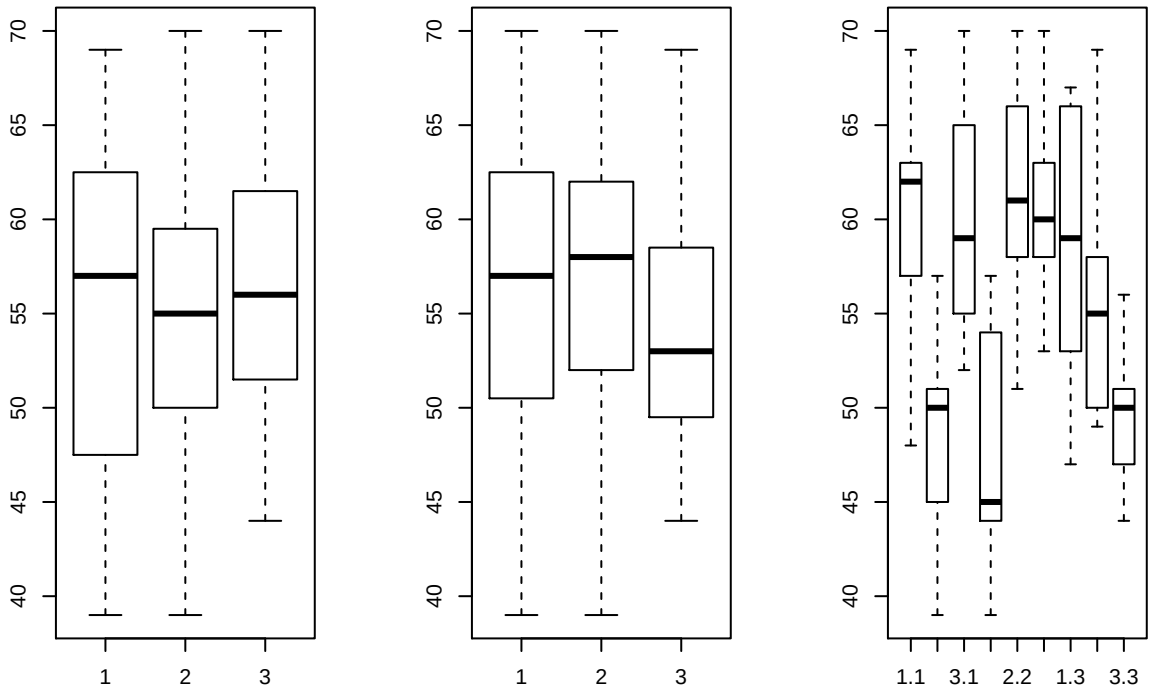
```
make.mean  <- ddply(disk, .(make), summarise, m = mean(y))
make.mean
```

```
##   make        m
## 1    1 56.13333
## 2    2 56.60000
## 3    3 54.73333
```

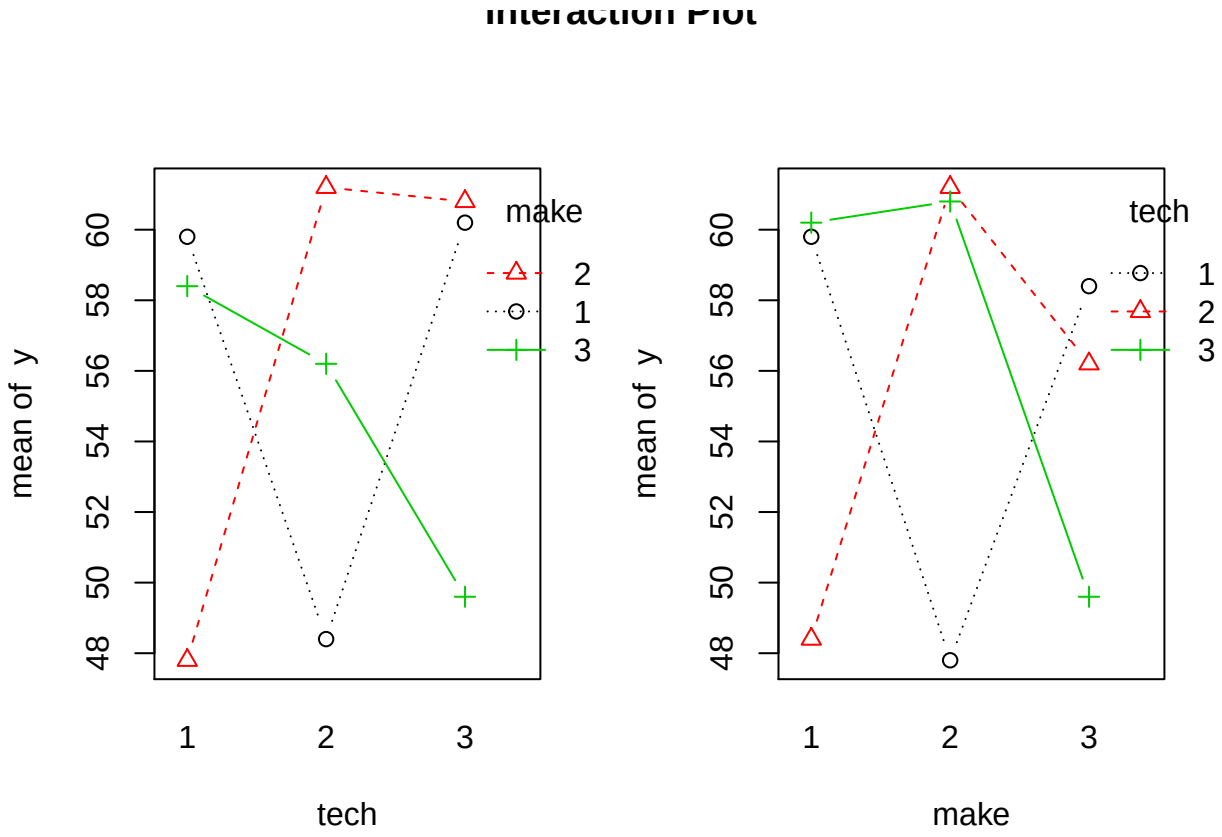
```
cellmean <- ddply(disk, .(tech,make), summarise, m = mean(y))
cellmean
```

```
##   tech make    m
## 1    1    1 59.8
## 2    1    2 47.8
## 3    1    3 58.4
## 4    2    1 48.4
## 5    2    2 61.2
## 6    2    3 56.2
## 7    3    1 60.2
## 8    3    2 60.8
## 9    3    3 49.6
```

```
##boxplots
par(mfrow=c(1,3))
boxplot(y ~ tech)
boxplot(y ~ make)
boxplot(y ~ tech + make)
```




```
##View interaction plots, also called profile plots
par(mfrow=c(1,2))
interaction.plot(tech,make,y,type='b',
                 col=1:3, pch=1:3)
#pch: a vector of plotting symbols or characters, with sensible default.
interaction.plot(make,tech,y,type='b',
                 col=1:3, pch=1:3)
title(main="Interaction Plot", outer=TRUE) #note that interaction is significant
```




```
##fit anova model
myfit.disk = aov(y~tech*make)
summary(myfit.disk)  #ANOVA table, notice that interaction effect is highly significant
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## tech         2   24.6   12.29   0.236 0.790779
## make         2   28.3   14.16   0.272 0.763283
## tech:make    4 1215.3  303.82   5.841 0.000994 ***
## Residuals   36 1872.4   52.01
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##Estimate \mu_{11} with a 99% CI, \bar y_{11.} +- t s
cellmean
```

```
##   tech make    m
## 1    1    1 59.8
## 2    1    2 47.8
## 3    1    3 58.4
## 4    2    1 48.4
## 5    2    2 61.2
## 6    2    3 56.2
## 7    3    1 60.2
## 8    3    2 60.8
## 9    3    3 49.6
```

```
s11<-sqrt(52.01/5)  #var = MSE/n_i
y11_CI<-c(cellmean$m[1]-qt(1-0.01/2, 36)*s11,cellmean$m[1]+qt(1-0.05/2, 36)*s11)
names(y11_CI)<-c("lower","upper")
```

```
y11_CI
```

```
##    lower    upper
## 51.02908 66.34104
```

```
##Tukey's multiple comparison for treatment combinations
diskTukey<-TukeyHSD(myfit.disk,"tech:make",ordered = TRUE, conf.level=.95)
diskTukey
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = y ~ tech * make)
##
## $`tech:make`
##         diff       lwr      upr     p adj
## 2:1-1:2  0.6 -14.43865 15.63865 1.0000000
## 3:3-1:2  1.8 -13.23865 16.83865 0.9999787
## 2:3-1:2  8.4  -6.63865 23.43865 0.6555252
## 1:3-1:2 10.6  -4.43865 25.63865 0.3552211
## 1:1-1:2 12.0  -3.03865 27.03865 0.2095773
## 3:1-1:2 12.4  -2.63865 27.43865 0.1771874
## 3:2-1:2 13.0  -2.03865 28.03865 0.1359755
## 2:2-1:2 13.4  -1.63865 28.43865 0.1130579
## 3:3-2:1  1.2 -13.83865 16.23865 0.9999991
## 2:3-2:1  7.8  -7.23865 22.83865 0.7366003
## 1:3-2:1 10.0  -5.03865 25.03865 0.4316582
## 1:1-2:1 11.4  -3.63865 26.43865 0.2658953
## 3:1-2:1 11.8  -3.23865 26.83865 0.2273106
## 3:2-2:1 12.4  -2.63865 27.43865 0.1771874
## 2:2-2:1 12.8  -2.23865 27.83865 0.1487668
## 2:3-3:3  6.6  -8.43865 21.63865 0.8715885
## 1:3-3:3  8.8  -6.23865 23.83865 0.5990906
## 1:1-3:3 10.2  -4.83865 25.23865 0.4054029
## 3:1-3:3 10.6  -4.43865 25.63865 0.3552211
## 3:2-3:3 11.2  -3.83865 26.23865 0.2867388
## 2:2-3:3 11.6  -3.43865 26.63865 0.2460831
## 1:3-2:3  2.2 -12.83865 17.23865 0.9999018
## 1:1-2:3  3.6 -11.43865 18.63865 0.9965056
## 3:1-2:3  4.0 -11.03865 19.03865 0.9928966
## 3:2-2:3  4.6 -10.43865 19.63865 0.9825836
## 2:2-2:3  5.0 -10.03865 20.03865 0.9710966
## 1:1-1:3  1.4 -13.63865 16.43865 0.9999970
## 3:1-1:3  1.8 -13.23865 16.83865 0.9999787
## 3:2-1:3  2.4 -12.63865 17.43865 0.9998112
## 2:2-1:3  2.8 -12.23865 17.83865 0.9994109
## 3:1-1:1  0.4 -14.63865 15.43865 1.0000000
## 3:2-1:1  1.0 -14.03865 16.03865 0.9999998
## 2:2-1:1  1.4 -13.63865 16.43865 0.9999970
## 3:2-3:1  0.6 -14.43865 15.63865 1.0000000
## 2:2-3:1  1.0 -14.03865 16.03865 0.9999998
## 2:2-3:2  0.4 -14.63865 15.43865 1.0000000
```

```r
plot(diskTukey, sub="Tukey Honest Significant Differences")

##Bonferroni comparison,
#the nature of the interaction effects is to be studied by making,
#for each technician, do all three pairwise comparisons among the disk drive makes in order
#to identity the make of disk drive for
#which the technician's mean service time is lowest.
#Family confidence coefficient for each set of three pairwise comparisons is 95%
cellmean
```

```
##   tech make    m
## 1    1    1 59.8
## 2    1    2 47.8
## 3    1    3 58.4
## 4    2    1 48.4
## 5    2    2 61.2
## 6    2    3 56.2
## 7    3    1 60.2
## 8    3    2 60.8
## 9    3    3 49.6
```

```r
D1<-cellmean$m[1]-cellmean$m[2]
D2<-cellmean$m[1]-cellmean$m[3]
D3<-cellmean$m[2]-cellmean$m[3]
D4<-cellmean$m[4]-cellmean$m[5]
D5<-cellmean$m[4]-cellmean$m[6]
D6<-cellmean$m[5]-cellmean$m[6]
D7<-cellmean$m[7]-cellmean$m[8]
D8<-cellmean$m[7]-cellmean$m[9]
D9<-cellmean$m[8]-cellmean$m[9]
D<-c(D1,D2,D3,D4,D5,D6,D7,D8,D9)
nameD<-c("A1B1-A1B2","A1B1-A1B3","A1B2-A1B3","A2B1-A2B2","A2B1-A2B3","A2B2-A2B3",
         "A3B1-A3B2","A3B1-A3B3","A3B2-A3B3")
#find s(D), check number of obsns in the cell
table(tech,make)  #5 each
```

```
##     make
## tech 1 2 3
##    1 5 5 5
##    2 5 5 5
##    3 5 5 5
```

```r
sD<-sqrt(2*52.01/5)  #var(D)=2*MSE/n_i
sD
```

```
## [1] 4.56114
```

```r
b<-qt(1-0.05/(2*3),36) #each set of three pairwise comparisons, g=3
lower<-rep(0,9)
upper<-rep(0,9)
for(i in 1:9){
  lower[i] =D[i]-b*sD
  upper[i]<-D[i]+b*sD
}
D_CI<-data.frame(nameD,D,lower,upper)
D_CI
```

```
##       nameD     D       lower      upper
## 1 A1B1-A1B2  12.0   0.5467929 23.4532071
## 2 A1B1-A1B3   1.4 -10.0532071 12.8532071
## 3 A1B2-A1B3 -10.6 -22.0532071  0.8532071
## 4 A2B1-A2B2 -12.8 -24.2532071 -1.3467929
## 5 A2B1-A2B3  -7.8 -19.2532071  3.6532071
## 6 A2B2-A2B3   5.0  -6.4532071 16.4532071
## 7 A3B1-A3B2  -0.6 -12.0532071 10.8532071
## 8 A3B1-A3B3  10.6  -0.8532071 22.0532071
## 9 A3B2-A3B3  11.2  -0.2532071 22.6532071
```

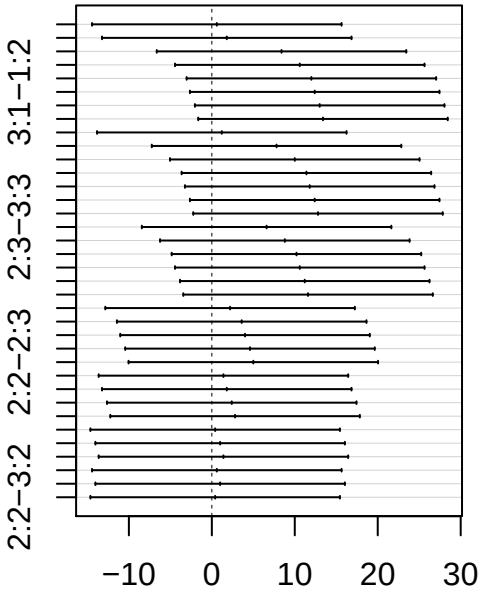
```
#With family confidence coefficient of 95%, technician 1 has the lowest service time when
#working with make 2, and is significantly shorter than working with make 1;
#With family confidence coefficient of 95%, technician 2 has the lowest service time when
#working with make 1, and is significantly shorter than working make 2 and 3;
#With family confidence coefficient of 95%, technician 3 has the lowest service time when
#working with make 3, but is not significantly shorter than working other makes.
#Technician interact with makes.

##Contrasts: suppose each technician services 10 machines of each make per week.
#How much time could be saved per week, on average, if technician 1 services only
##make 2, technician 2 services only make 1, and technician 3 services only make 3? use 99% CI
#L= 10*(\mu_{11}+\mu_{12} + \mu_{13}) + 10*(\mu_{21}+\mu_{22} + \mu_{23}) +
#10*(\mu_{31}+\mu_{32} + \mu_{33}) - 30*\mu_{12} - 30*\mu_{21}-30*\mu_{33}
coefL<-c(10,-20,10,-20,10,10,10,10,-20)
hatL<-sum(coefL*cellmean$m)
varL<-sum(coefL^2)*52.01/5 #var =(\sum c_i^2) MSE/n_i
sL<-sqrt(varL)
sL
```

```
## [1] 136.8342
```

```
yL_CI<-c(hatL-qt(1-0.01/2, 36)*sL,hatL+qt(1-0.01/2, 36)*sL)
names(yL_CI)<-c("lower","upper")
yL_CI
```

```
##     lower     upper
##  277.8815 1022.1185
```

95% family−wise confidence leve



Differences in mean levels of tech:make
Tukey Honest Significant Differences