Q3 (a-c) - Data 2: beers tasting

(a) Find the rating mean for each country and type. Find also the cell mean for each treatment combination (county and type combination).

```
beers = read.table("/Users/doganakad/Desktop/uoft/second semester/sta303/Assignments/A1/beers.csv",sep=
str(beers)
```

```
## 'data.frame':    36 obs. of  4 variables:
##  $ name   : Factor w/ 36 levels "1554 Black","60minute",..: 2 28 15 32 3 26 20 25 13 16 ...
##  $ type   : Factor w/ 2 levels "IPA","Lager": 1 1 1 1 1 1 1 1 1 1 ...
##  $ country: Factor w/ 3 levels "Belgium","UK",..: 3 3 3 3 3 3 1 1 1 1 ...
##  $ rating : num  4.09 4.19 4.27 4.22 3.89 4.48 4.21 3.81 3.99 4.04 ...
```

```
beers$name = as.factor(beers$name) # put name into a factor variable
beers$type = as.factor(beers$type) # put type into a factor variable
beers$country = as.factor(beers$country) # put country into a factor variable
rating = beers$rating
name = beers$name
type = beers$type
country = beers$country

# Find the rating mean for each country
with(beers, tapply(rating, country, mean))
```

```
##  Belgium       UK      USA
## 3.654167 3.535833 3.775833
```

```
# Find the rating mean for each type of beers
with(beers, tapply(rating, type, mean))
```

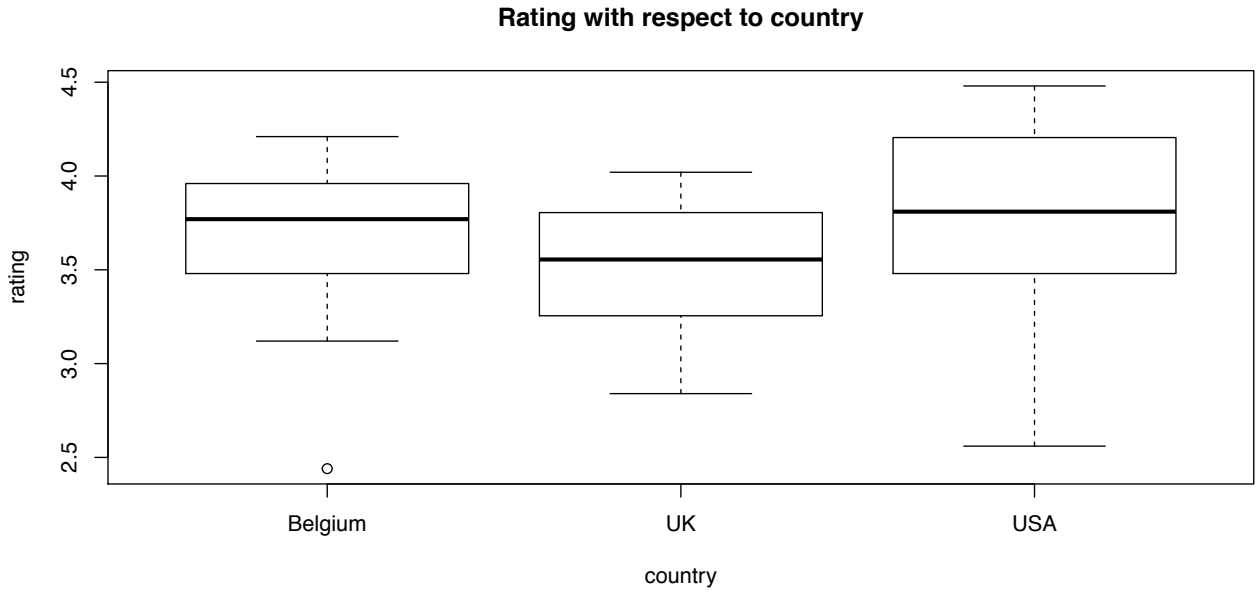
```
##      IPA    Lager
## 3.922778 3.387778
```

```
# Find the cell mean for each treatment combination
with(beers, tapply(rating, list(type, country), mean))
```

```
##        Belgium       UK      USA
## IPA   3.950000 3.628333 4.190000
## Lager 3.358333 3.443333 3.361667
```

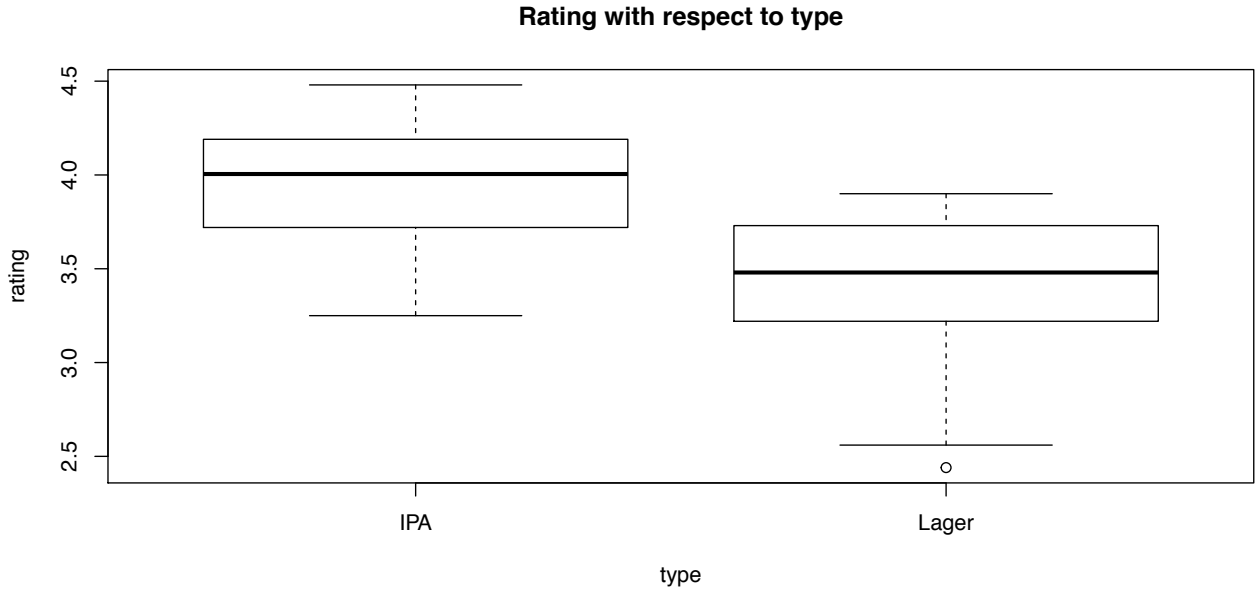
(b) Create box-plot of rating with respect to two factors, type and country. What can you say about the difference of rating mean for each factor ?

```
# boxplot of rating with respect to country
boxplot(rating~country,data=beers, main="Rating with respect to country", xlab="country", ylab="rating"
```



Comments with respect to country: Medians and interquartile ranges for each country is really close to each other.Therefore there is a similar variation in countries. Their means are also close to each other if we look at the result at Q3-a. Combining these factors, we see evidence of low between-group variance, so we have a little evidence that the main effect is significant.

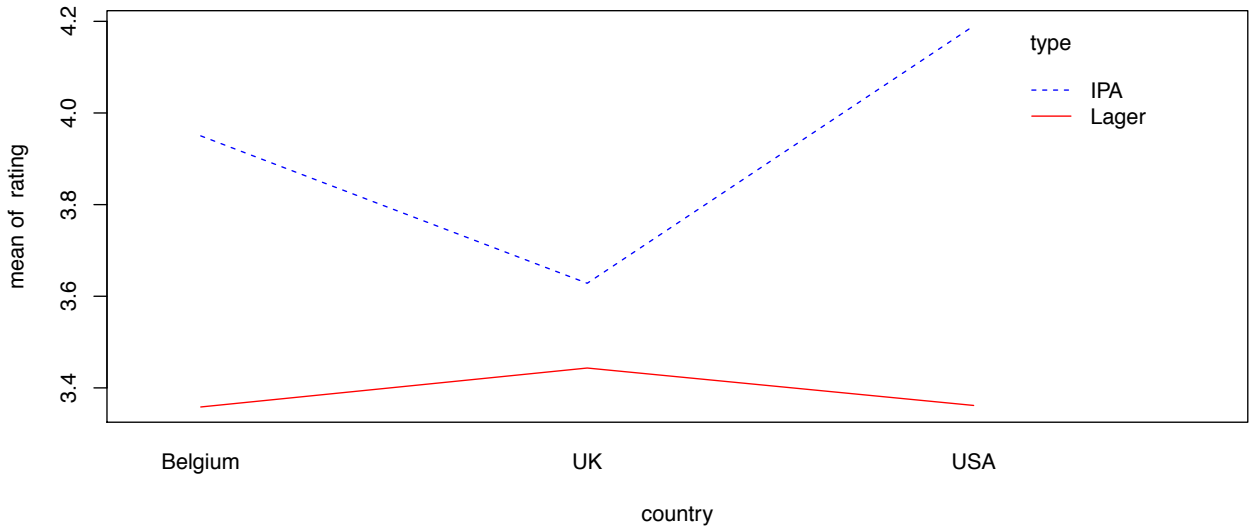
```
# boxplot of rating with respect to type
boxplot(rating~type,data=beers, main="Rating with respect to type", xlab="type", ylab="rating")
```



Comments with respect to type: IPA has higher rating than Lager. Their means are significantly different, not even close to each other if we look at the result from Q3-a. IPA's median is higher than Lager's upper quantile. Combining all of these we can say that effect of beer type on rating is signifiant.

(c) Create the interaction plot. What could you say about the main effect and interaction effect ?

```
# interaction plot
 with(beers ,interaction.plot(country, type, rating, col=c("blue","red"), legend = TRUE))
```



Comments: If we change the level of country, IPA has a bigger/stronger change than Lager which indicates that country has a bigger main effect on IPA. Looking at the plot, we can also say that since increasing the level of country from Belgium to UK causes the mean rating of IPA to drop substantially and increasing from UK to USA causes the mean to increase, there is a reinforcement effect. This is the exact opposite for Lager, so the effect of country is surpressed when the type is Lager. Also there is a significant horizontal distance between the lines for IPA and lager which concludes that beer type has a significant effect. Finally, looking at the lines for both types, since increasing country changes the line we can also say that country has a significant effect.