Undergrad Biostatistics - R Training - Week VI

Ege Ulgen

Data

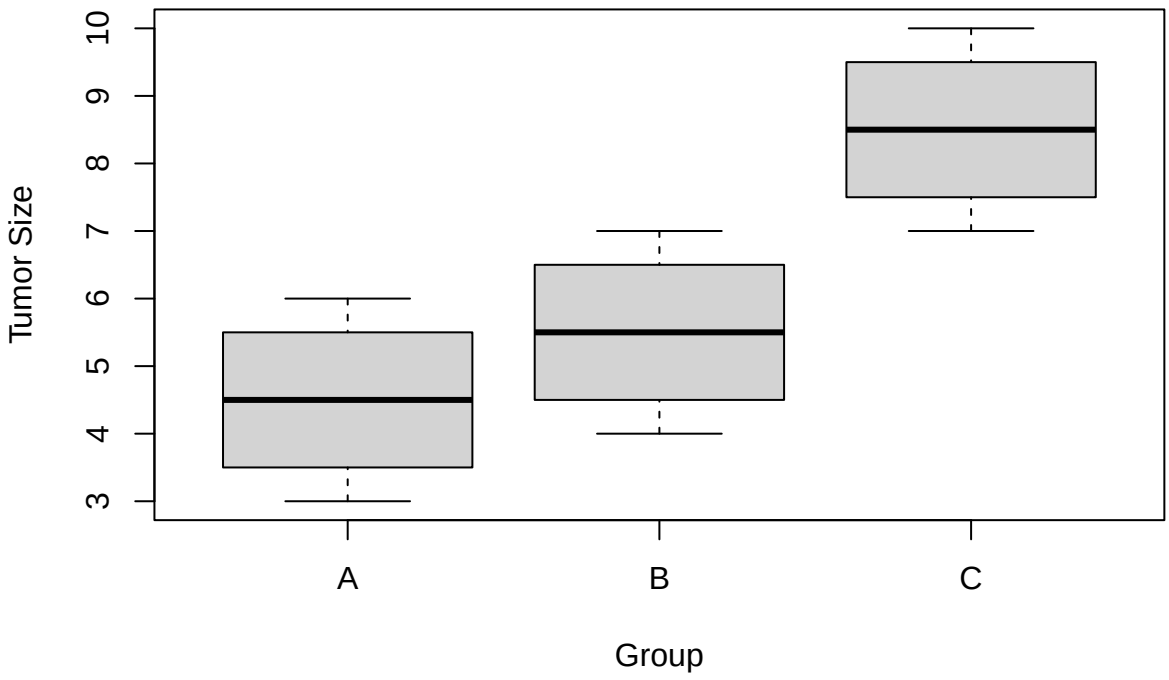
We'll create our own data for this exercise:

```
tumor_size_df <- data.frame(size = c(3, 4, 5, 6, 4, 5, 6, 7, 7, 8, 9, 10),
                            group = c("A","A","A","A","B","B","B","B","C","C","C","C"))
tumor_size_df
```

```
##    size group
## 1     3     A
## 2     4     A
## 3     5     A
## 4     6     A
## 5     4     B
## 6     5     B
## 7     6     B
## 8     7     B
## 9     7     C
## 10    8     C
## 11    9     C
## 12   10     C
```

Let's visualize the data:

```
boxplot(tumor_size_df$size~tumor_size_df$group, xlab = "Group", ylab = "Tumor Size")
```



ANOVA

```
res_aov <- aov(size ~ group, data = tumor_size_df)
summary(res_aov)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## group        2   34.7   17.33    10.4 0.0046 **
## Residuals    9   15.0    1.67
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Post-hoc Test

```
TukeyHSD(res_aov)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = size ~ group, data = tumor_size_df)
##
## $group
##     diff      lwr    upr   p adj
## B-A    1 -1.54874 3.5487 0.54025
## C-A    4  1.45126 6.5487 0.00451
```

```
## C-B    3  0.45126 5.5487 0.02317
```