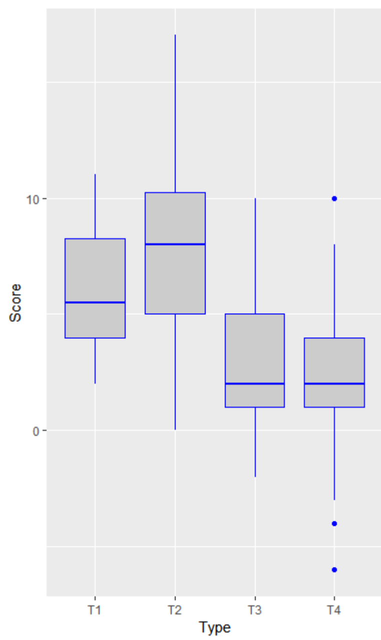
Ex 8.3.1

```
> hw831 <- read.csv("C:/Users/Teaching/ma4310/EXR_C08_S03_01.csv")
> hw831$type = factor(hw831$type,labels = c("T1", "T2", "T3","T4"))
> hw831$subject = factor(hw831$subject,levels=unique(hw831$subject))
> require(ggplot2)
> require(ggplot2)
> ggplot(hw831, aes(x = type, y = score)) +
+     geom_boxplot(fill = "grey80", colour = "blue") +
+     scale_x_discrete() + xlab("Type") +
+     ylab("Score")
```




```
> hw831.model = lm(score~subject+type, data = hw831)
> anova(hw831.model)

Two-way anova - additive model

Analysis of Variance Table

Response: score
          Df Sum Sq Mean Sq F value    Pr(>F)
subject   23 481.12  20.918  2.4923  0.001887 **
type       3 498.37 166.125 19.7930 2.306e-09 ***
Residuals 69 579.12   8.393
---

> qf(0.95,df1=3,df2=69)
[1] 2.737492
```

Conclusion: F = 19.79 > 2.737492. Also p < 0.001. Hence reject H0. Improvement scores differed significantly among the four training types.

Ex 8.3.2

```
> # NOTE: I MANUALLY MODIFIED DATA INPUT FILE TO FIX FORMAT
> hw832 <- read.csv("C:/Users/Teaching/ma4310/EXR_C08_S03_02b.csv")
> hw832$type = factor(hw832$type,levels=unique(hw832$type))
> hw832$subject = factor(hw832$subject,levels=unique(hw832$subject))
> hw832.model = lm(score~subject+type, data = hw832)
> anova(hw832.model)
```

```
Two-way anova - additive model

Analysis of Variance Table

Response: score
          Df  Sum Sq Mean Sq F value    Pr(>F)    
subject    9 1694868  188319 24.1161 1.121e-10 ***
type       3  153813   51271  6.5658  0.001775 ** 
Residuals 27  210839    7809
```

Conclusion: F = 6.57, p = 0.002. Reject H0. Tic frequencies differed significantly among the time periods

Ex 8.4.1

```
> hw841 <- read.csv("C:/Users/Teaching/ma4310/EXR_C08_S04_01.csv")
> hw841$time = factor(hw841$time,levels=unique(hw841$time))
> hw841$subj = factor(hw841$subj,levels=unique(hw841$subj))
> hw841.model = lm(outcome~subj+time, data = hw841)
> anova(hw841.model)
```

```
Two-way anova - additive model

Analysis of Variance Table

Response: outcome
          Df  Sum Sq Mean Sq F value    Pr(>F)    
subj       9  64.225   7.136  5.4313 0.0002846 ***
time       3 192.275  64.092 48.7801 4.939e-11 ***
Residuals 27  35.475   1.314
```

Conclusion: F = 48.78, p < 0.001. Reject H0. There was significant difference in the four time periods

Ex 8.4.2

```
> hw842 <- read.csv("C:/Users/Teaching/ma4310/EXR_C08_S04_02.csv")
> hw842$LOAD = factor(hw842$LOAD,levels=unique(hw842$LOAD))
> hw842$SPECIMEN = factor(hw842$SPECIMEN,levels=unique(hw842$SPECIMEN))
> hw842.model = lm(LAXITY~SPECIMEN+LOAD, data = hw842)
> anova(hw842.model)

Two-way anova - additive model

Analysis of Variance Table

Response: LAXITY
          Df Sum Sq Mean Sq F value    Pr(>F)
SPECIMEN   6 195951   32659 66.5885  8.96e-14 ***
LOAD       4  14892    3723  7.5909 0.0004221 ***
Residuals 24  11771     490
```

Conclusion: F = 7.59, p < 0.001. Reject H0. Graft laxity was significantly different among the five load levels

Ex 8.5.2

```
> hw852 <- read.csv("C:/Users/Teaching/ma4310/EXR_C08_S05_02.csv")
> hw852$PSYCH = factor(hw852$PSYCH,labels = c("T1", "T2", "T3","T4"))
> hw852$PHYS = factor(hw852$PHYS)
> hw852.model = lm(MONTHS~PHYS+PSYCH+PHYS:PSYCH, data = hw852)
> anova(hw852.model)

Two-way anova - additive model with interaction

Analysis of Variance Table

Response: MONTHS
           Df Sum Sq Mean Sq F value    Pr(>F)
PHYS        5 13.796  2.7592  7.5913 2.552e-05 ***
PSYCH       3 90.408 30.1359 82.9112 < 2.2e-16 ***
PHYS:PSYCH 15 32.001  2.1334  5.8694 1.203e-06 ***
Residuals  48 17.447  0.3635
```

Conclusion: Since the *p* values are significant for physical therapy treatment, psychological treatment and interaction, we reject the null hypothesis for each of those variables.

QUIZ #4 A study investigated the awareness of deficit profiles among stroke patients undergoing rehabilitation. The study included 35 patients with a stroke lesion in the **right hemisphere** and 19 patients with a lesion on the **left hemisphere**. In addition, lesion size was grouped as **size-2**: 1-3 cm; **size-3**: 3-5 cm"; and **size-4:** 5 cm or greater. The outcome variables is a measure of each patient's total unawareness of their own limitations, with scores ranging from 8 to 24 and higher scores indicating more unawareness.

(a) Explore the data and make sure each factor is identified correctly. Is the experiment design balanced or unbalanced? Justify your conclusion.
(b) Apply the correct Anova method to test for a difference in lesion size, hemisphere, and interaction, using significance level 0.05 for all tests. Discuss the conclusion of the test.
(c) Use the Tukey's HSD procedure to test for significant differences among individual pairs of means, if appropriate. Justify your conclusion.

```
> dataq4 <- read.csv("C:/Users/dlabate/Desktop/Teaching/ma4310/dataq4.csv")
> str(dataq4)
'data.frame':   54 obs. of  3 variables:
 $ Score: int  11 13 10 11 9 10 9 8 10 13 ...
 $ Side : chr  "L" "L" "L" "L" ...
 $ Size : int  2 2 2 2 2 2 2 2 2 3 ...
> dataq4$Side <- factor(dataq4$Side, levels = c("L","R"),labels = c("L","R"))
> dataq4$Size <- factor(dataq4$Size, levels = c(2,3,4),labels = c("2","3","4"))
> str(dataq4)
'data.frame':   54 obs. of  3 variables:
 $ Score: int  11 13 10 11 9 10 9 8 10 13 ...
 $ Side : Factor w/ 2 levels "L","R": 1 1 1 1 1 1 1 1 1 1 ...
 $ Size : Factor w/ 3 levels "2","3","4": 1 1 1 1 1 1 1 1 1 2 ...
113 obs. of  2 variables:
```

a) Analysis of variance

```
> table(dataq4$Side, dataq4$Size)

     2  3  4
  L  9  5  5
  R 14  9 12
```

The table shows that the experiment has unequal numbers of subjects in each group.
Hence, we need to run the Anova test with **Unbalanced Design.**

b) We apply the analysis of variance hypothesis testing

```
> res.aov <- aov(Score ~ Side*Size, data = dataq4)
> library(car)
> Anova(res.aov, type = "III")

Anova Table (Type III tests)

Response: Score
            Sum Sq Df  F value  Pr(>F)
(Intercept) 920.11  1 227.8446 < 2e-16 ***
Side          0.18  1   0.0452 0.83253
Size         30.10  2   3.7269 0.03129 *
Side:Size    10.54  2   1.3050 0.28062
Residuals   193.84 48
```

Conclusion: since p-value < 0.05 for the size factor, the **lesion size is a significant factor** in the awareness deficit profiles. On the other hand, **side and interaction are not statistically significant factors**.

NOTE: if we do not apply the Type III test, we obtain a different result

```
> summary(res.aov)
```

```
             Df Sum Sq Mean Sq F value Pr(>F)
Side          1   1.29   1.287   0.319 0.5750
Size          2  29.98  14.991   3.712 0.0317 *
Side:Size     2  10.54   5.270   1.305 0.2806
Residuals    48 193.84   4.038
```

c) <u>We apply the Tukey's HSD procedure</u>

```
> TukeyHSD(res.aov, which = "Size")

Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = Score ~ Side * Size, data = dataq4)

$Size
         diff        lwr      upr     p adj
3-2 0.5824733 -1.0649968 2.229943 0.6709014
4-2 1.7373036  0.1828198 3.291787 0.0252228
4-3 1.1548304 -0.5991994 2.908860 0.2587441
```

Conclusion: based on the values of p adj, only the difference between the sizes 2 and 4 is statistically significant.