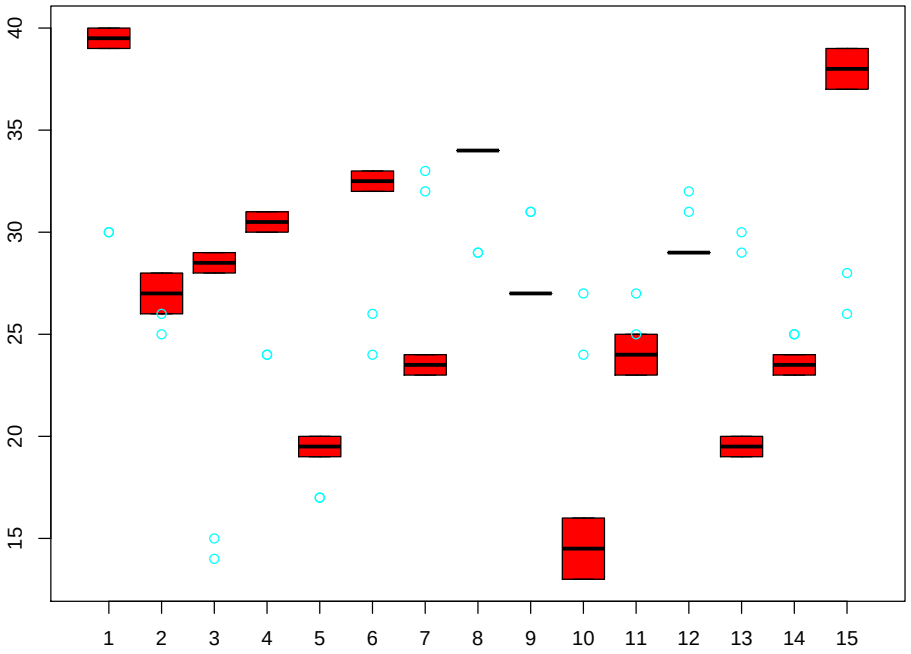
Solution to Series 5

1. **a)** Plot the data. Read in the data:

```
> paint <- read.table(file="http://stat.ethz.ch/Teaching/Datasets/paint.txt",header=TRUE)
> paint$SAMPLE <- as.factor(paint$SAMPLE)
> paint$BATCH <- as.factor(paint$BATCH)
> plot(paint$BATCH[paint$SAMPLE==1],paint$MOISTURE[paint$SAMPLE==1],col=2)
> points(paint$BATCH[paint$SAMPLE==2],paint$MOISTURE[paint$SAMPLE==2],col=5)
```



b) Write down the analysis of variance model. Why are all factors random?

$$Y_{ijk} = \mu + batch_i + sample_{j(i)} + \epsilon_{ijk}$$

$batch_i$ random effect of BATCH i, $batch_i \sim \mathcal{N}(0, \sigma^2_{batch})$,
$sample_{j(i)}$ random effect of SAMPLE j, $sample_{j(i)} \sim \mathcal{N}(0, \sigma^2_{sample})$,
ϵ_{ijk} measurement error, $\epsilon_{ijk} \sim \mathcal{N}(0, \sigma^2_e)$.
All factors are random because BATCHs are chosen at random from the population of batches and SAMPLEs are chosen from the population of all samples.

c) Construct the skeleton analysis of variance table with degrees of freedom and expected values.

Score	df	$E(MS)$
batch	14	$\sigma^2_e + 2\sigma^2_{sample} + 4\sigma^2_{batch}$
batch:sample	15	$\sigma^2_e + 2\sigma^2_{sample}$
Residuals	30	σ^2_e
Total	59	

d) Calculate estimates for σ^2_{BATCH} and σ^2_{SAMPLE}. Compute the F test statistics, first by hand and then use R .

```
> mod1 <- aov(MOISTURE~BATCH/SAMPLE,data=paint)
> summary(mod1)
```

```
             Df Sum Sq Mean Sq F value Pr(>F)
BATCH        14 1210.9   86.50   94.36 <2e-16 ***
BATCH:SAMPLE 15  869.7   57.98   63.26 <2e-16 ***
Residuals    30   27.5    0.92
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

$$\hat{\sigma}_e^2 = 0.917$$

$$\hat{\sigma}_{sample}^2 = (57.983 - 0.917)/2 = 28.533$$

$$\hat{\sigma}_{batch}^2 = (86.495 - 57.983)/4 = 7.128$$

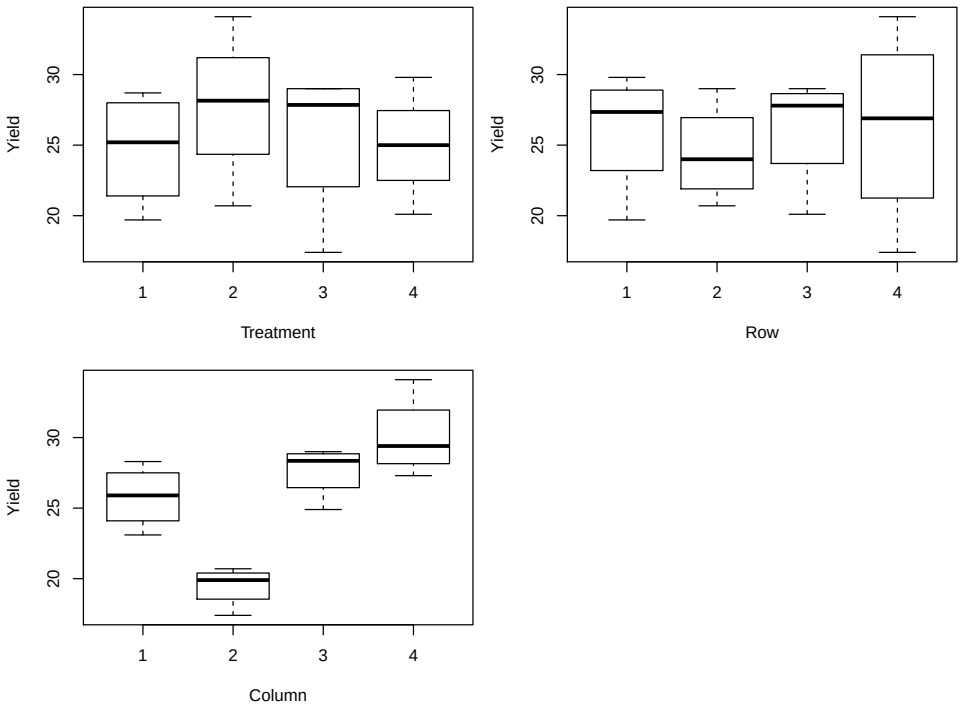
$$F_{\sigma_{batch}^2=0} = \frac{\hat{\sigma}_e^2 + 2\hat{\sigma}_{sample}^2 + 4\hat{\sigma}_{batch}^2}{\hat{\sigma}_e^2 + 2\hat{\sigma}_{sample}^2} = \frac{0.917 + 2 \cdot 28.533 + 4 \cdot 7.128}{0.917 + 2 \cdot 28.533} = 1.4917$$

$$F_{\sigma_{sample}^2=0} = \frac{\hat{\sigma}_e^2 + 2\hat{\sigma}_{sample}^2}{\hat{\sigma}_e^2} = 63.255$$

```
> mod2 <- aov(MOISTURE~BATCH+Error(SAMPLE%in% BATCH),data=paint)
> summary(mod2)

Error: SAMPLE:BATCH
          Df Sum Sq Mean Sq F value Pr(>F)
BATCH     14 1210.9   86.50   1.492  0.226
Residuals 15  869.7   57.98

Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 30   27.5  0.9167
```

2. **a)** Plot the data.

```
> peanut <- read.table(file="http://stat.ethz.ch/Teaching/Datasets/Peanut.txt",header=TRUE)
> peanut$Row <- as.factor(peanut$Row)
> peanut$Column <- as.factor(peanut$Column)
> peanut$Treatment <- as.factor(peanut$Treatment)

> par(mfrow=c(2,2))
> plot(Yield ~ Treatment+Row+Column, data=peanut)
```



b) Carry out the analysis of variance and report your findings.

```
> modP1 <- aov(Yield ~ Treatment+Row+Column, data=peanut)
> summary(modP1)
```

```
            Df Sum Sq Mean Sq F value  Pr(>F)
Treatment    3  23.42    7.81   1.953 0.22255
Row          3   9.43    3.14   0.786 0.54394
Column       3 245.91   81.97  20.507 0.00148 **
Residuals    6  23.98    4.00
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

We see that the factor Treatment is not significant.

c) Treatment 1 is the control treatment. Even though the anova F is not significant, do any of the other treatments have a significantly higher yield, if a one-sided test is used?

```
> TukeyHSD(modP1,"Treatment", conf.level=0.90)
```

```
  Tukey multiple comparisons of means
    90% family-wise confidence level

Fit: aov(formula = Yield ~ Treatment + Row + Column, data = peanut)

$Treatment
      diff        lwr      upr     p adj
2-1  3.075 -0.9887413 7.138741 0.2320297
3-1  0.825 -3.2387413 4.888741 0.9334796
4-1  0.275 -3.7887413 4.338741 0.9971031
3-2 -2.250 -6.3137413 1.813741 0.4479925
4-2 -2.800 -6.8637413 1.263741 0.2914481
4-3 -0.550 -4.6137413 3.513741 0.9782233
```

If we look at the first three rows of the Treatment differences table we see that all of the confidence intervals of 2-1, 3-1 and 4-1 contain 0. So we can say that non of the one-sided tests are significant. That is, non of the other treatments have a significantly higher yield compared to the control treatment number 1.

The breeder decided to replicate the experiment. He used three blocks at different locations in the field, but in each location the north-south and east-west gradients were identified. The data are in the file `Peanut2.txt`.

d) Which factors are nested, which are crossed? The factors Column and Row are nested in the factor Rep which represents the three different positions on the field.

e) Give an analysis of variance table and test whether any pairwise differences are significant.

```
> peanut2 <- read.table(file="http://stat.ethz.ch/Teaching/Datasets/Peanut2.txt",header=TRUE)
> peanut2$Row <- as.factor(peanut2$Row)
> peanut2$Column <- as.factor(peanut2$Column)
> peanut2$Treatment <- as.factor(peanut2$Treatment)
> peanut2$Rep <- as.factor(peanut2$Rep)
> modP2 <- aov(Yield ~ Treatment+Rep/(Row+Column), data=peanut2)
> summary(modP2)
            Df Sum Sq Mean Sq F value   Pr(>F)
Treatment    3   95.7   31.92   3.962   0.0199 *
Rep          2   15.8    7.92   0.983   0.3888
Rep:Row      9   66.1    7.35   0.912   0.5311
Rep:Column   9  530.8   58.97   7.321 4.55e-05 ***
Residuals   24  193.3    8.06
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> TukeyHSD(modP2,"Treatment", conf.level=0.90)
  Tukey multiple comparisons of means
    90% family-wise confidence level

Fit: aov(formula = Yield ~ Treatment + Rep/(Row + Column), data = peanut2)

$Treatment
           diff        lwr         upr     p adj
2-1  3.36666667  0.5619194  6.17141398 0.0363778
3-1  0.50000000 -2.3047473  3.30474732 0.9724701
4-1 -0.06666667 -2.8714140  2.73808065 0.9999296
3-2 -2.86666667 -5.6714140 -0.06191935 0.0899752
4-2 -3.43333333 -6.2380806 -0.62858602 0.0320424
4-3 -0.56666667 -3.3714140  2.23808065 0.9607866
```

We see that Treatment 2 is significantly better then the control.

3. To do an anova fit you can reshape your data as follows

```
> tennis.max=reshape(tennis[,c(1:5,9,13)],varying=c("max1","max2","max3"),idvar="id",
   timevar="period",v.names="max",direction="long")
> tennis.max$Treatment[tennis.max$order==1 & tennis.max$period==1]="Motrin"
> tennis.max$Treatment[tennis.max$period==2]="Washout"
> tennis.max$Treatment[tennis.max$order==1 & tennis.max$period==3]="Placebo"
> tennis.max$Treatment[tennis.max$order==2 & tennis.max$period==3]="Motrin"
> tennis.max$Treatment[tennis.max$order==2 & tennis.max$period==1]="Placebo"
> tennis.max$Treatment=factor(tennis.max$Treatment)
> tennis.max$id=factor(tennis$id)
> tennis.max$sex=factor(tennis$sex)
> levels(tennis.max$sex)=c("male","female")
> tennis.max$period=factor(tennis.max$period)
```

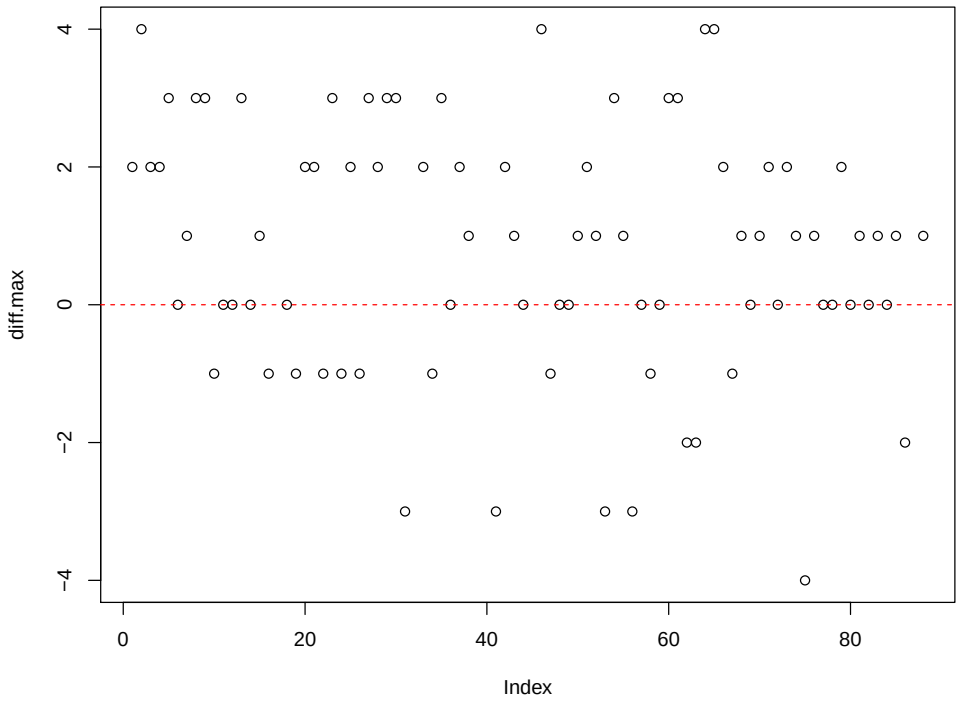
Is the treatment with Motrin efficient?

To find this out: First do a pairwise t.test and wilcox.test of the four different ways to compare mentioned before. Then do an anova fit for only the maximum activity comparison taking into account id, period and treatment.

Is the washout period long enough or could there still be a carry-over effect?

Read in the data:

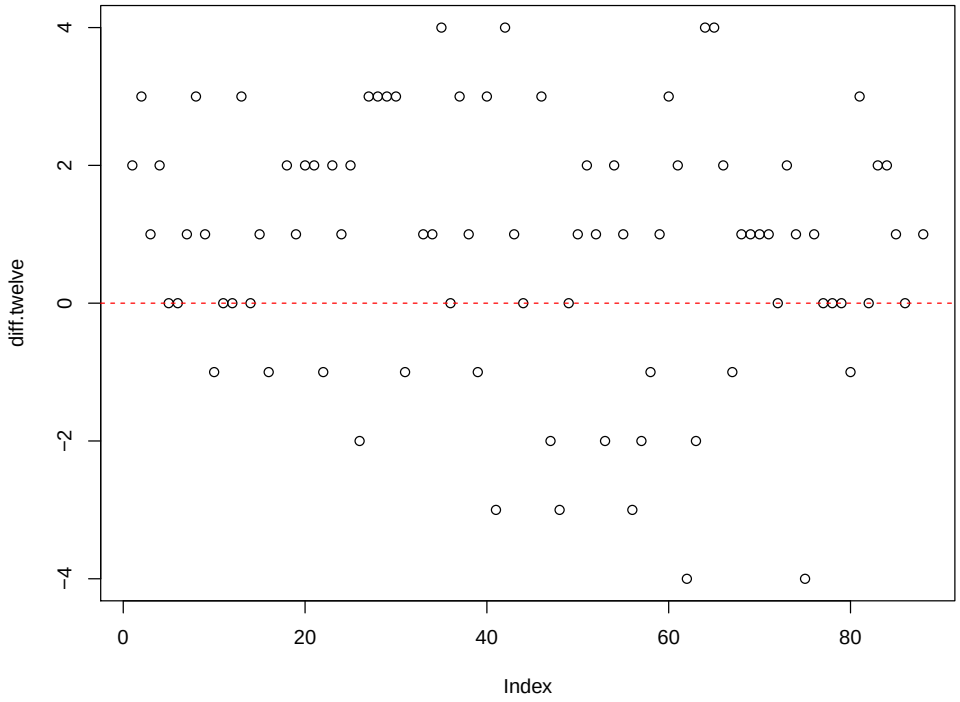
```
> tennis <- read.table(file="http://stat.ethz.ch/Teaching/Datasets/TENNIS.dat")
> names(tennis)=c("id","age","sex","order","max1","twelve1","ave1","overall1",
 "max2","twelve2","ave2","overall2","max3","twelve3","ave3","overall3")
> tennis$sex[tennis$sex==9] <- 1##just fill in something for the missing data to not complicate th
> head(tennis)
> for (i in 3:16)
   tennis[,i][tennis[,i]==9 | tennis[,i]==0]=NA##identify the invalid values

> diff.max=tennis$max1-tennis$max3
> diff.max[tennis$order==2]=-diff.max[tennis$order==2]
> plot(diff.max)
> abline(h=0,col=2,lty=2)
```




```
> summary(diff.max)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-4.0000  0.0000  1.0000  0.7927  2.0000  4.0000       6

> t.test(diff.max)

        One Sample t-test

data:  diff.max
t = 3.9354, df = 81, p-value = 0.0001746
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.3919132 1.1934527
sample estimates:
mean of x
0.7926829

> wilcox.test(diff.max)

        Wilcoxon signed rank test with continuity correction
```

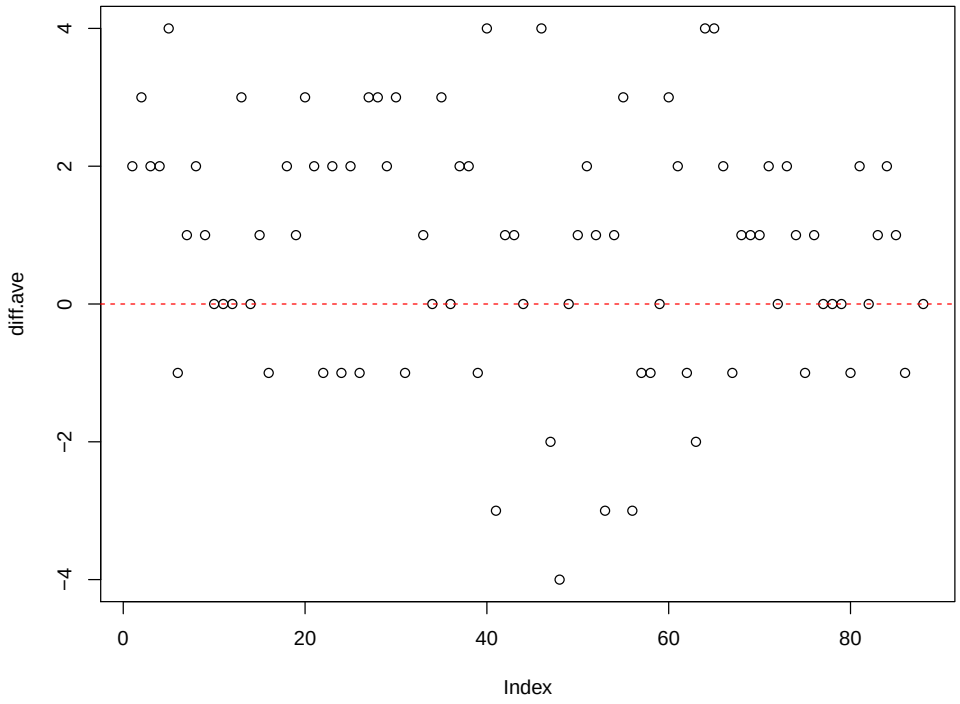
```
data:  diff.max
V = 1578.5, p-value = 0.000271
alternative hypothesis: true location is not equal to 0
```

```
> diff.twelve=tennis$twelve1-tennis$twelve3
> diff.twelve[tennis$order==2]=-diff.twelve[tennis$order==2]
> plot(diff.twelve)
> abline(h=0,col=2,lty=2)
```

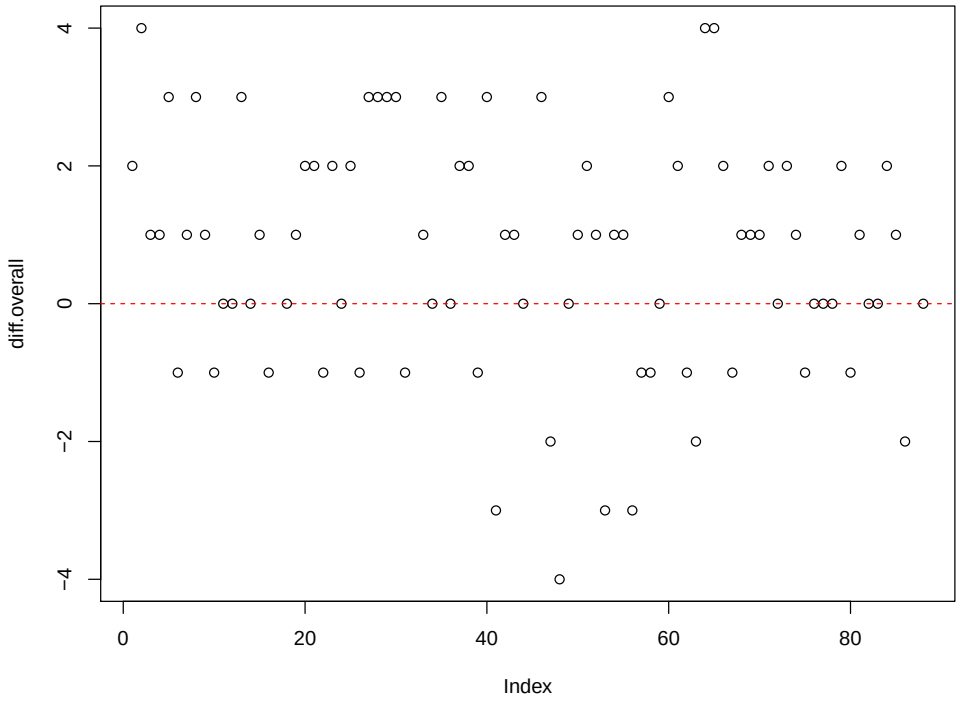



```
> summary(diff.twelve)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-4.0000  0.0000  1.0000  0.7976  2.0000  4.0000       4

> t.test(diff.twelve)

        One Sample t-test

data:  diff.twelve
t = 3.9551, df = 83, p-value = 0.0001604
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.3965036 1.1987345
sample estimates:
mean of x
 0.797619

> wilcox.test(diff.twelve)

        Wilcoxon signed rank test with continuity correction

data:  diff.twelve
V = 1855, p-value = 0.0002831
alternative hypothesis: true location is not equal to 0
```

```
> diff.ave=tennis$ave1-tennis$ave3
> diff.ave[tennis$order==2]=-diff.ave[tennis$order==2]
> plot(diff.ave)
> abline(h=0,col=2,lty=2)
```




```
> summary(diff.ave)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-4.0000  0.0000  1.0000  0.8214  2.0000  4.0000       4

> t.test(diff.ave)

        One Sample t-test

data:  diff.ave
t = 4.2674, df = 83, p-value = 5.218e-05
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.4385796 1.2042776
sample estimates:
mean of x
0.8214286

> wilcox.test(diff.ave)

        Wilcoxon signed rank test with continuity correction

data:  diff.ave
V = 1869, p-value = 6.026e-05
alternative hypothesis: true location is not equal to 0

> diff.overall=tennis$overall1-tennis$overall3
> diff.overall[tennis$order==2]=-diff.overall[tennis$order==2]
> plot(diff.overall)
> abline(h=0,col=2,lty=2)
```




```
> summary(diff.overall)

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
-4.0000  0.0000  1.0000  0.7143  2.0000  4.0000       4

> t.test(diff.overall)

        One Sample t-test

data:  diff.overall
t = 3.7938, df = 83, p-value = 0.0002808
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 0.339812 1.088759
sample estimates:
mean of x
0.7142857

> wilcox.test(diff.overall)

        Wilcoxon signed rank test with continuity correction

data:  diff.overall
V = 1705.5, p-value = 0.0003242
alternative hypothesis: true location is not equal to 0
```

We see that the degree of pain while on Motrin significantly better than degree of pain on placebo for all four pain measurements.

Anova approach for max, ave, twelve and overall Example for max:

```
> tennis.max=reshape(tennis[,c(1:5,9,13)],varying=c("max1","max2","max3"),idvar="id",
   timevar="period",v.names="max",direction="long")
> tennis.max$Treatment[tennis.max$order==1 & tennis.max$period==1]="Motrin"
```

```
> tennis.max$Treatment[tennis.max$period==2]="Washout"
> tennis.max$Treatment[tennis.max$order==1 & tennis.max$period==3]="Placebo"
> tennis.max$Treatment[tennis.max$order==2 & tennis.max$period==3]="Motrin"
> tennis.max$Treatment[tennis.max$order==2 & tennis.max$period==1]="Placebo"
> tennis.max$Treatment=factor(tennis.max$Treatment)
> tennis.max$id=factor(tennis$id)
> ##tennis.max$sex[tennis.max$sex==9]=NA##not used anyway
> tennis.max$sex=factor(tennis$sex)
> levels(tennis.max$sex)=c("male","female")
> tennis.max$period=factor(tennis.max$period)
> mod1=aov(max~id+period+Treatment,data=tennis.max[tennis.max$period!=2,])
> summary(mod1)

            Df Sum Sq Mean Sq F value   Pr(>F)
id          86 181.32   2.108   1.576   0.0203 *
period       1  29.03  29.030  21.703 1.25e-05 ***
Treatment    1  24.46  24.460  18.286 5.22e-05 ***
Residuals   80 107.01   1.338
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
7 observations deleted due to missingness

> model.tables(mod1,type="means")

Tables of means
Grand mean

3.449704

 id
    701 702 704 705 706 707 708 712 713 718 720 721 723 725
      4   2   5   4   3 1.5   5   3 1.5   4   5 1.5 2.5   4
rep   2   2   1   2   2 2.0   2   2 2.0   2   2 2.0 2.0   2
    727 729 732 733 734 735 736 740 741 742 744 745 746 747
      2   3   2 3.5   5   2 3.5 3.5 3.5 2.5   2   3 3.5   5
rep   2   2   2 2.0   2   1 2.0 2.0 2.0 2.0   2   2 2.0   2
    749 752 756 759 761 762 763 766 767 768 769 771 772 774
    2.5 4.5   5 2.5   4   4 4.5 3.5 4.5   3 4.5 2.5 1.5   3
rep 2.0 2.0   2 2.0   2   2 2.0 2.0 2.0   2 2.0 2.0 2.0   2
    775 776 777 779 780 781 782 783 784 785 786 789 792 793
    3.5 2.5 3.5   4 3.5 3.5   4   5   4 4.5   2   5 3.5   5
rep 2.0 2.0 2.0   2 2.0 2.0   1   2   2 2.0   1   1 2.0   2
    794 796 798 799 804 806 808 809 810 811 812 813 815 817
    2.5   4   2 4.5   2   2 5.5   3 3.5 3.5 4.5 4.5 3.5   3
rep 2.0   2   2 2.0   2   2 2.0   2 2.0 2.0 2.0 2.0 2.0   2
    821 823 824 825 826 827 831 832 836 839 840 842 843 844
    3.5   5 3.5 2.5   5   4   4   4   4 3.5 3.5   2 3.5   4
rep 2.0   2 2.0 2.0   2   2   2   2   2 2.0 2.0   2 2.0   2
    845 846 848
      2   4 1.5
rep   2   2 2.0

 period
        1     3
     3.053  3.87
rep 87.000 82.00

 Treatment
    Motrin Placebo
```

```
     3.822   3.073
rep 85.000  84.000

>
```

Is there any carry-over effect?
If there is a lasting effect of Motrin, the treatment difference in group A is smaller than in group B, in other words the average response in group A is larger than in group B.

```
> mod2=aov(max~order,data=tennis.max[tennis.max$period!=2,])
> summary(mod2)

             Df Sum Sq Mean Sq F value Pr(>F)
order         1    0.8  0.7624   0.373  0.542
Residuals   167  341.1  2.0423
7 observations deleted due to missingness
```

This not the case.
Another possibility: Are there any differences in the washout responses?

```
> mod3=aov(max~order,data=tennis.max[tennis.max$period==2,])
> summary(mod3)

            Df Sum Sq Mean Sq F value Pr(>F)
order        1   1.78   1.775   0.973  0.327
Residuals   83 151.45   1.825
3 observations deleted due to missingness
```

No.