R HELP SHEET: Anova

CONTENTS
1. Creating a tab delimited data file using Excel
2. Conducting an Anova
3. Identifying the key elements of the output

1. Creating a tab delimited data file using Excel
Open Excel and type data from your dependent variable into one column and data from your independent variable into another column with appropriate headings at the top of each column (e.g., *site, nitrogen*), then save the file as a **Text (Table delimitated)(*.txt)** with an appropriate name (e.g., *RData_Reedbeds*) file using **Save as type.**

	A	B
1	site	nitrogen
2	1	2.92
3	1	2.88
4	1	3.25
5	1	2.64
6	1	3.28
7	2	3.06
8	2	2.6
9	2	2.55
10	2	2.42
11	2	2.35
12	3	3.41
13	3	3.23
14	3	3.93
15	3	3.74
16	3	3.18

Save as type: | Text (Tab delimited) (*.txt)

Single File Web Page (*.mht; *.mhtml)
Web Page (*.htm; *.html)
Template (*.xlt)
Text (Tab delimited) (*.txt)
Unicode Text (*.txt)
Microsoft Excel 5.0/95 Workbook (*.xls)
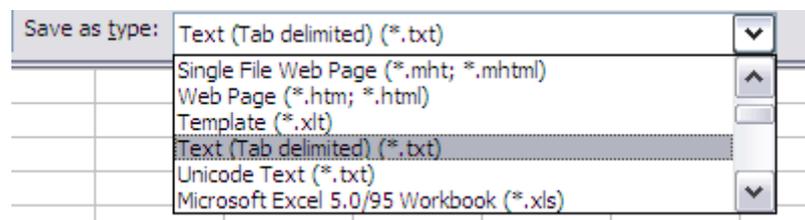
(Note this example left has *site* coded as numbers 1, 2 and 3 but it could also be written in words e.g., *Site1, Site2, Site3*).

2. Conducting an Anova

The text in green after the hash (**#**) sign is just **notes** to help you remember what's in the output: it does not get R to actually "do" anything. The text in blue is **R code** with stars representing words that are specific to the example: you need to replace this with text specific to your data as shown in the output in section 3.

To get R to conduct an Anova:
    Open an **R-Editor** window by selecting **File** then **New script**.
    Type in (or copy and paste) the notes and code below.
    Replace the stars with appropriate text as indicated in notes.
    Highlight everything and press **Ctrl R**.

```
#Importing data from tab delimited file
#(replace stars with e.g.,reeds )
*****<-read.table(file.choose(),header=TRUE)
attach(*****)
names(*****)


#Conducting an Anova
#(replace stars with appropriate text e.g., anova, nitrogen, site, anova, anova)
******<-aov(***~as.factor(******))
summary(*****)
TukeyHSD(*****)
```

3. Identifying the key elements of the output

Following the instructions above will produce the following output in the **R Console** window: the **key elements** are annotated in orange.

```
> #Importing data from tab delimited file
> #(replace stars with an appropriate object name e.g.,reeds)
> reeds<-read.table(file.choose(),header=TRUE)
> attach(reeds)
> names(reeds)
[1] "site"    "nitrogen"
>
> #Conducting an Anova
> #(replace stars with appropriate text e.g.,anova,nitrogen,site)
> anova<-aov(nitrogen~as.factor(site))
> summary(anova)
               Df    Sum Sq    Mean Sq    F value    Pr(>F)
as.factor(site) 2    2.043     1.0217     11.95      0.00139 **
Residuals      12    1.026     0.0855
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> TukeyHSD(anova)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = nitrogen ~ as.factor(site))

$`as.factor(site)`
        diff      lwr           upr          p adj
2-1    -0.398    -0.89135525    0.09535525   0.1207456
3-1     0.504     0.01064475    0.99735525   0.0451782
3-2     0.902     0.40864475    1.39535525   0.0010226
```

Annotations: **Degrees of Freedom**, **Statistic**, **P Value**, **Effect sizes**, **95% CI of Effect sizes**

In summary the key information from the test is

$F_{2,12}=11.95, P=0.001$

And the unstandardized effect size (estimated difference between the populations) for example:

difference (site 2 – site 1) = -0.398, 95% CI [-0.891, -0.095]