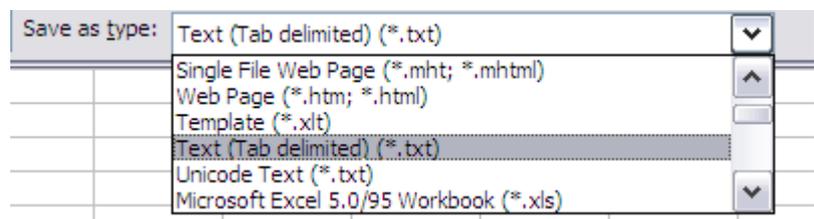# R HELP SHEET: Regression

CONTENTS
1. Creating a tab delimited data file using Excel
2. Conducting an Regression
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., *reserve*, *area, richness*), then save the file as a **Text (Table delimitated)(*.txt)** with an appropriate name (e.g., *RData_SpeciesRichness*) file using **Save as type.**

	A	B	C
1	reserve	area	richness
2	Mole	4840	22
3	Digya	3126	14
4	Bui	2074	17
5	Gbele	565	10
6	Kalakpa	325	12
7	Shai	58	5

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)

2. Conducting an Regression

> 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 Regression:

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.,species )
*******<-read.table(file.choose(),header=TRUE)
attach(*******)
names(*******)

#Conducting a regression
#(replace stars with appropriate text e.g., result, richness, area, result, result)
******<-lm(*******~****)
summary(*****)
confint(*****)
```

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.,speies)
> species<-read.table(file.choose(),header=TRUE)
> attach(species)
> names(species)
[1] "reserve"  "area"     "richness"
>
> #Conducting a regression
> #(replace stars with appropriate text e.g.,output, richness, area, output, output)
> output<-lm(richness~area)
> summary(output)
> confint(output)


Call:
lm(formula = richness ~ area)

Residuals:
    1       2       3       4       5       6
 0.4143  -2.8844  3.0011  0.1401  2.7983  -3.4693

Coefficients:
              Estimate    Std.Error    t value   Pr(>|t|)
(Intercept)   8.3102269   1.8215891    4.562     0.0103 *
area          0.0027429   0.0007245    3.786     0.0193 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.057 on 4 degrees of freedom
Multiple R-squared: 0.7818,    Adjusted R-squared: 0.7273
F-statistic: 14.33 on 1 and 4 DF,  p-value: 0.01934

                  2.5 %           97.5 %
(Intercept)   3.2526847891    13.367769102
area          0.0007312762    0.004754461
```

Orange annotations:
- **Intercept**
- **Effect Size (slope)**
- **Coefficient of determination (R^2)**
- **Statistic (F)**
- **P Value**
- **Degrees of Freedom**
- **95% CI of Effect size (slope)**

In summary the key information from the test is
$$y = 0.003\ x + 8.310;\ F_{1,4} = 14.332,\ P = 0.019;\ R^2 = 0.782$$