

```
# IRDR04 Research Tools Module
# Week 8 Tutorial
# 27/11/18 9-11am @ Birkbeck, Malet Street 414/415 Public Cluster
# Tutor: Dr Alison Fairbrass, UCL Centre for Biodiversity & Environment Research, alison.fairbrass@gmail.com
```

```
# # # Walk-through # # #
```

```
# Load compensation.csv datafile from Moodle
Growth <- read.csv("Week8//compensation.csv")
```

```
# An alternative way to load data
Growth <- read.csv(file.choose())
```

```
# Look at the file
Growth
# and its structure
str(Growth)
```

```
# now let's plot it using R base plot() function; we're interested in knowing whether
# Root biomass (Root) has an influence or correlates with Fruit production (Fruit)
```

```
plot(Growth$Root, Growth$Fruit)
```

```
# A better and more intuitive way to do it uses the data = argument
```

```
plot(Fruit ~ Root, data = Growth)
```

```
# so we can see that the relationship looks linear! But also that it has two components.
# Let's tidy it up
```

```
# Add some axis labels and you can vary their size using the list(cex combination of arguments)
plot(Fruit ~ Root, data = Growth,
     xlab = list("Root Biomass", cex = 1.5), #add half to the size of the default font
     ylab = list("Fruit Production", cex = 0.5)) # halves the default font
```

```
# Now let's consider the points on the graph
# ?par gives you help on the graphics parameters
# ?points specifically on the points.
```

```
plot(Fruit ~ Root, data = Growth,
     xlab = list("Root Biomass", cex = 1), #reset size to default of 1
     ylab = list("Fruit Production", cex = 1), #reset size to default of 1
     cex = 2, pch = 21) # the pch argument controls the type 21 = filled circle, bg = colour.
```

```
# You can see that there are two clusters of points here relating to the treatment - so let's explore that.
# It would be instructive to give them different colours to illustrate the patterns.
```

```
plot(Fruit ~ Root, data = Growth,
     xlab = list("Root Biomass", cex = 1),
     ylab = list("Fruit Production", cex = 1),
     cex = 2, pch = 21, col=Grazing)
```

```
# Now lets add a legend
legend(4.2,120,unique(Growth$Grazing),col=1:length(Growth$Grazing), pch=21)
```

```
# First linear regression output
# use the Nelson.csv data
# This is an experiment on 9 batches of flour beetles assessing their weight loss measured in mg
# at different humidity levels ranging from 0-93% humidity
# The experiment lasted 6 days
```

```
# Load data
Flour <- read.csv(file.choose())
```

```
# Look at it
str(Flour)
```

```
# Load the car package which contains the scatterplot function
library(car)
# Draw some pictures to assess linearity and normality of the data
scatterplot(WEIGHTLOSS ~ HUMIDITY, data = Flour,
            xlab = "% Humidity",
            ylab = "Beetle weight loss (mg)",
            pch = 21, col = "blue", bg = "blue")
```

```
# Run a linear model using the lm() function from base R
Flour.lm <- lm(WEIGHTLOSS ~ HUMIDITY, data = Flour) # don't mix up your response and explanatory variables!
```

```
# Look at the output
summary(Flour.lm) # refer to lecture slides to see what the numbers mean i.e. for interpretation
```

```
# Now check assumption
# 1. Homoskedasticity
```

```
plot(fitted(Flour.lm), residuals(Flour.lm))
```

```
# 2. Normality of residuals
hist(residuals(Flour.lm))
qqnorm(residuals(Flour.lm))
qqline(residuals(Flour.lm))
```

```
# recreate the plot with a regression line
# Draw some pictures to assess linearity - you've plotted this so you can see it is!
plot(WEIGHTLOSS ~ HUMIDITY, data = Flour,
     xlab = "% Humidity",
     ylab = "Beetle weight loss (mg)",
     pch = 21, col = "blue", bg = "blue")
abline(Flour.lm) # add the line using the abline() function
```