

Example R Markdown Script

Adapted from:

```
# Tidy data and efficient manipulation # # Coding Club tutorial #  
# January 18th 2017 # # Sandra Angers-Blondin (s.angers-blondin@ed.ac.uk) # # John Godlee # #  
24/Jan/2017 # #####
```

Use this example R script to practice compiling an R Markdown file.

Try to make a well commented, easy to follow record of what is going on so that others can easily follow.

Download the datasets for this example script from:

<https://github.com/ourcodingclub/CC3-DataManip>

Install and load the relevant packages

```
library(dplyr) # an excellent data manipulation package library(tidyr) # a package to format your data  
library(pander) #to create pretty tables
```

Set your working directory to the folder where you have downloaded the datasets

```
setwd()
```

Import data

```
elongation <- read.csv("EmpetrumElongation.csv", sep = ";") # stem elongation measurements on crow-  
berry germination <- read.csv("Germination.csv", sep = ";") # germination of seeds subjected to toxic  
solutions
```

Tidying the data

```
#Putting the data into long format using gather() elongation_long <- gather(elongation, Year, Length,  
c(X2007, X2008, X2009, X2010, X2011, X2012)) #gather() works like this: data, key, value, columns to  
gather. Here we want the lengths (value) to be gathered by year (key). Note that you are completely  
making up the names of the second and third arguments, unlike most functions in R. head(elongation_long)
```

Investigating the data

Create a boxplot of 'elongation_long' to visualise elongation for each year.

This set of boxplots can be added to your R Markdown document by putting the code in a code chunk

```
boxplot(Length ~ Year, data = elongation_long, xlab = "Year", ylab = "Elongation (cm)", main = "Annual growth of Empetrum hermaphroditum")
```

Use filter() to keep only the rows of germination' for speciesSR'

```
germinSR <- filter(germination, Species == 'SR')
```

Let's have a look at the distribution of germination across SR

This histogram can be added to your R Markdown document by simply putting the code in a code chunk

Try adding some plain text to your R markdown document to explain the histogram

```
hist(germinSR$Nb_seeds_germin, breaks = 8)
```

Use mutate() to create a new column of the germination percentage using the total number of seeds and the number of seeds that germinated

```
germin_percent <- mutate(germination, Percent = Nb_seeds_germin / Nb_seeds_tot * 100)
```

Use a pipe to get a table of summary statistics for each Seed type

```
germin_summ <- germin_percent %>% group_by(Species) %>% summarise("Mean germination per" = mean(Nb_seeds_germin), "Max germination per" = max(Nb_seeds_germin), "Min germination per" = min(Nb_seeds_germin))
```

Make a table of 'germin_summ' in your R markdown document using pander(), the instructions can be found in the tutorial