RMarkdown_Demo_1

Erik Bolch

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

Install and load the relevant packages

Changed tidyr to tidyverse

```r
library(dplyr) # an excellent data manipulation package
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyverse) # a package to format your data
```

```
## -- Attaching packages -----------------------------------------------------------------

## v ggplot2 3.2.1     v readr   1.3.1
## v tibble  2.1.3     v purrr   0.3.3
## v tidyr   1.0.0     v stringr 1.4.0
## v ggplot2 3.2.1     v forcats 0.4.0


## -- Conflicts --------------------------------------------------------------------------
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(pander) #to create pretty tables
```

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

Commented out because not recommended for Rmd files

```r
#setwd("./data/CC-3-DataManip-master")
```

Import data

Germination file didn't exist, changed to dragons which is included.

```r
elongation <- read_csv("./data/CC-3-DataManip-master/EmpetrumElongation.csv") # stem elongation measure
```

```
## Parsed with column specification:
## cols(
##   Zone = col_double(),
##   Indiv = col_double(),
##   X2007 = col_double(),
##   X2008 = col_double(),
##   X2009 = col_double(),
##   X2010 = col_double(),
##   X2011 = col_double(),
##   X2012 = col_double()
## )
```

```r
dragons <- read_csv("./data/CC-3-DataManip-master/dragons.csv") # dragons subjected to spices
```

```
## Parsed with column specification:
## cols(
##   dragon.ID = col_double(),
##   species = col_character(),
##   tabasco = col_double(),
##   jalapeno = col_double(),
##   wasabi = col_double(),
##   paprika = col_double()
## )
```

Tidying the data and Putting the data into long format using gather()

```r
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 b
head(elongation_long)
```

```
## # A tibble: 6 x 4
##    Zone Indiv Year  Length
##   <dbl> <dbl> <chr>  <dbl>
## 1     2   373 X2007    5.1
## 2     2   379 X2007    8.1
## 3     2   383 X2007    9.3
## 4     2   389 X2007   15
## 5     2   390 X2007    3.5
## 6     2   395 X2007    6.1
```

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")
```

## Annual growth of Empetrum hermaphroditum
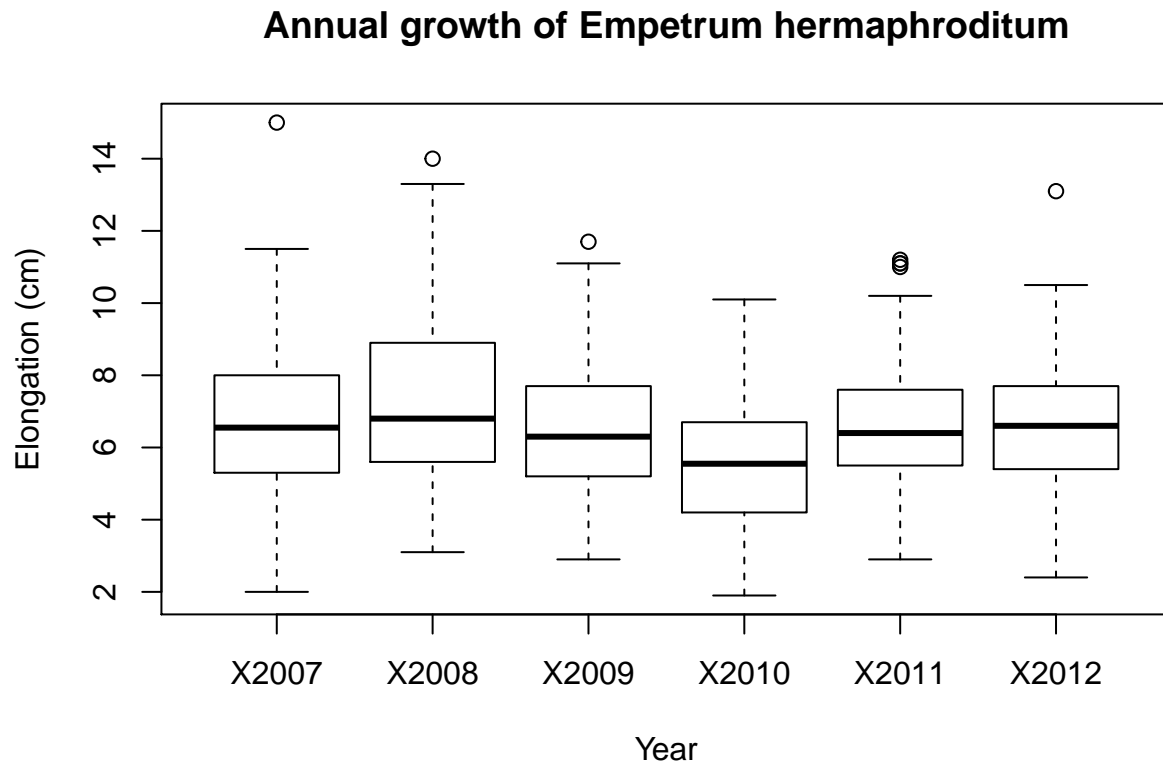
Use filter() to keep only the rows of 'dragon' for species 'hungarian_horntail'

```
dragon.hh <- filter(dragons, species == 'hungarian_horntail')
```

Let's have a look at the distribution of dragons species exposed to spices 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(dragon.hh$wasabi, breaks = 8)
```

Histogram of dragon.hh$wasabi
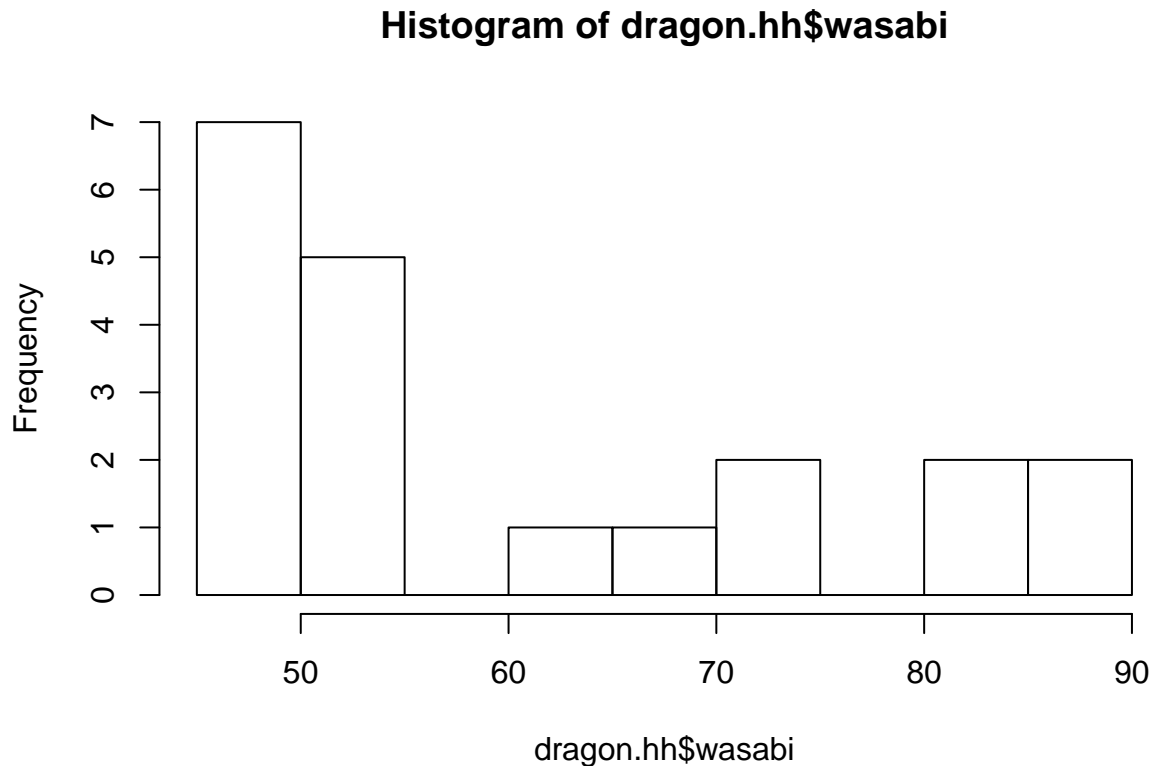
Use mutate() to create a new column of the dragon percentage using the total number total and the number of dragons exposed to jalapeno out of total spice exposure

```
dragon_j_percent <- mutate(dragons, Percent_Jalapeno = jalapeno / (tabasco+jalapeno+wasabi+paprika) * 1
head(dragon_j_percent)
```

```
## # A tibble: 6 x 7
##    dragon.ID species      tabasco jalapeno wasabi paprika Percent_Jalapeno
##        <dbl> <chr>          <dbl>    <dbl>  <dbl>   <dbl>            <dbl>
## 1          1 hungarian_hor~    124      100     45      25             34.0
## 2          4 hungarian_hor~    156      110     47      30             32.1
## 3          7 hungarian_hor~    147       90     55      25             28.4
## 4         10 hungarian_hor~    138      136     47      23             39.5
## 5         13 hungarian_hor~    154      124     53      23             35.0
## 6         16 hungarian_hor~    110       92     48      32             32.6
```

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

```
dragon_summ <- dragon_j_percent %>%
  group_by(species) %>%
  summarise("Mean jalapeno exposure per" = mean(jalapeno), "Max jalapeno exposure per" = max(jalapeno),
```

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

```
pander(dragon_summ)
```

Table 1: Table continues below

species	Mean jalapeno exposure per	Max jalapeno exposure per
hungarian_horntail	119.2	141
swedish_shortsnout	152.4	195
welsh_green	121.1	152

Min jalapeno exposure
87
113
98