

# MATH 4753 Laboratory 1

## Introduction to R

---

In this lab you will learn the basics of R. This program is free and you are encouraged to obtain a copy for your Mac, PC or Linux machine. Install it and then download and install R studio (this is a nice front end to R and is also free).

### *Objectives*

In this lab you will learn how to:

1. Read data into R.
2. Create data in R.
3. Manipulate data frames.
4. Create vectors.
5. Create matrices.
6. Create tables.
7. Create barplots.
8. Create pie charts.
9. Create boxplots.
10. Create scatter plots.
11. Be familiar with R markdown and knitting documents

### *Tasks*

There are a number of front-ends for R

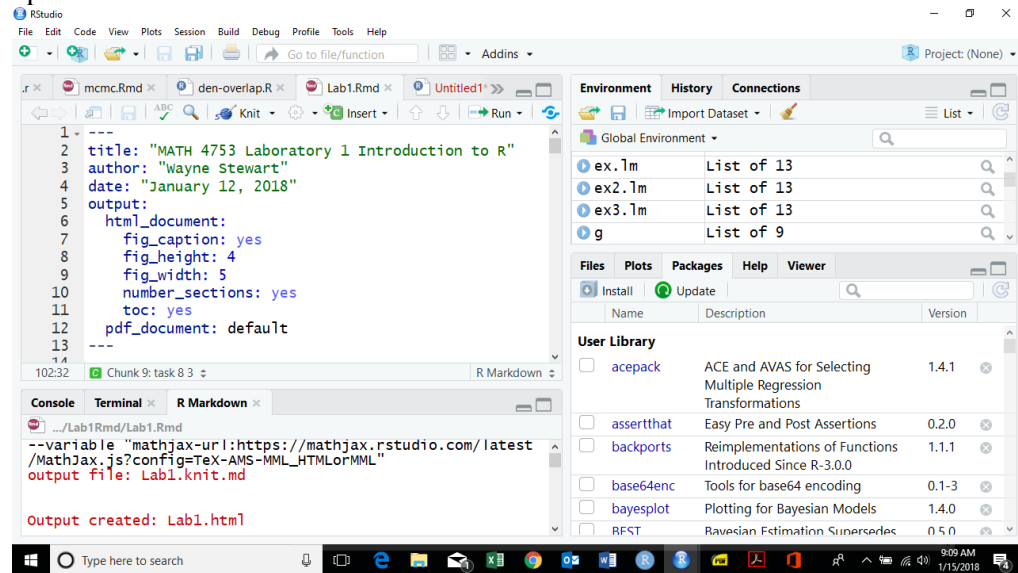
We will however use Rstudio for the course but you may trial other front ends as you have time.

All answers to questions should be written up in R markdown and then knitted into an html. Save and place the .Rmd, and .html files in the dropbox when completed.

**Only the html file will be graded** (the rmd document will be viewed occasionally to help with coding problems)

- **Task 1**
  - Download from CANVAS the zipped data files, “Dataxls”
  - Unzip the contents into a directory on your desktop (call it LAB1)
  - Download the file “lab1.r”
  - Place this file with the others in LAB1.
  - Start Rstudio
  - Open “lab1.r” from within Rstudio.
  - Go to the “session” menu within Rstudio and “set working directory” to where the source files are located.
  - Obtain the working directory by issuing the command `getwd()` :
- **Task 2**
  - Find the file “DDT.xls” inside LAB1
  - Open it in Excel
  - Save As type CSV(comma delimited) “\*.csv”

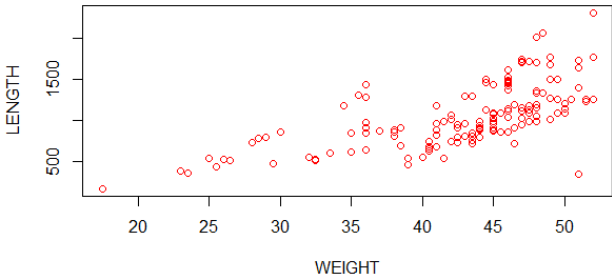
- Use `read.table()`, `read.csv()` or the menu on Rstudio to read the data into R, this function will already be available within the script `lab1.r` which you have opened in Rstudio.



- Obtain the first six lines of the data using “`head ( )`”:
- Task 3
  - We read the data into an object called “`ddt`”. This object contains all of the data. What are the qualitative variables in “`ddt`”?
  - What are the quantitative variables in “`ddt`”?
  - How many SPECIES are in the `ddt` data set?
  - Subset the `ddt` data set so that we see only the observations with SPECIES, LMBASS and WEIGHT bigger than 800gms
  - Subset the `ddt` data set so that we see only those observations where the RIVER is SCM and DDT bigger than 4.0
- CLICKER QUESTIONS (these are to be answered below since most will not have a clicker yet)

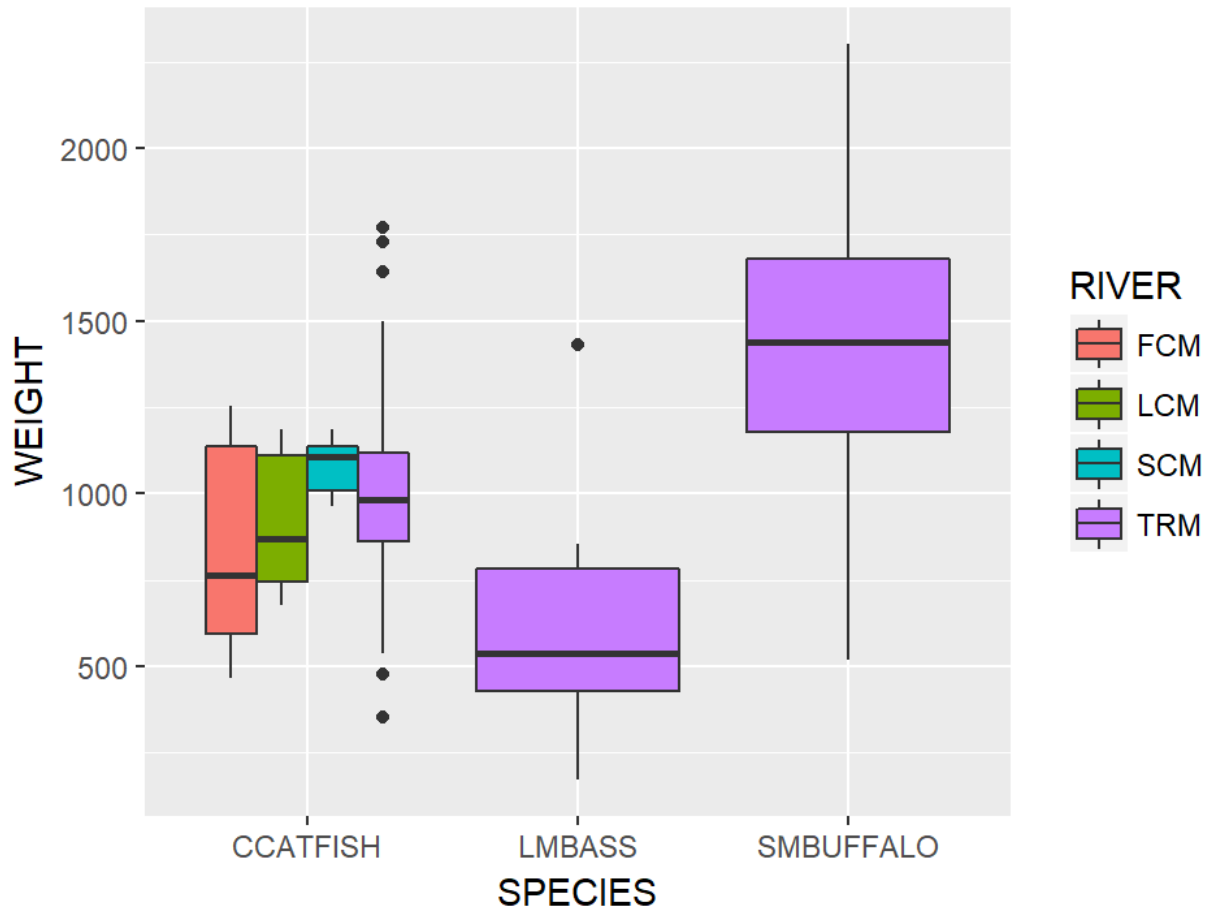


Q.1 Find the mean length of fish in the <code>ddt</code> data frame!	Ans.
Q.2 What is the standard deviation of the weight of fish in the <code>ddt</code> data frame?	Ans.
Q3. Is this the plot of LENGTH vs WEIGHT from the <code>ddt</code> data frame?	Ans. 1=Yes, 0=No

	
<p>Q4. If <math>v=1:20</math> what is the last value of <math>v/20</math>?</p>	<p>Ans.</p>

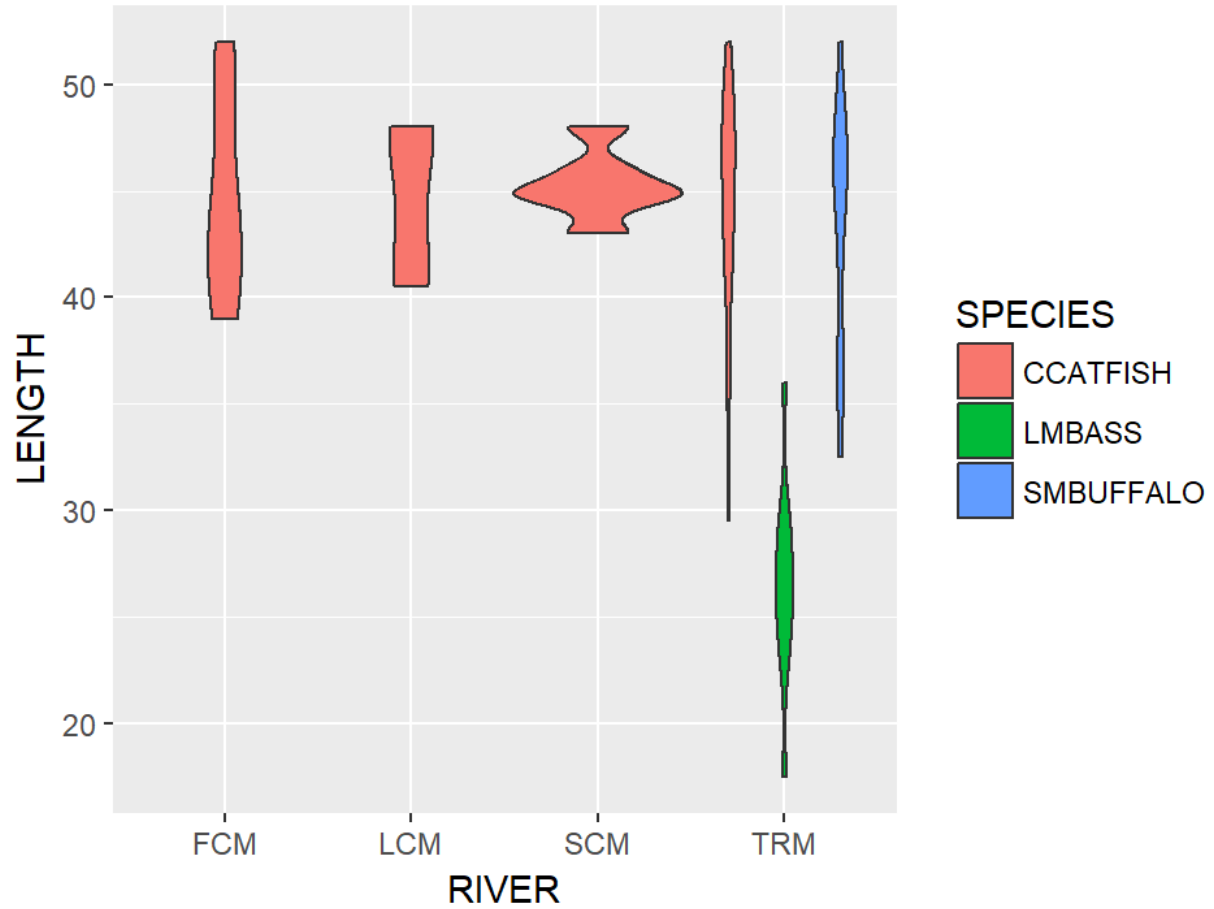
- Task 4
  - Make a table of the rivers.
  - Make a color coded barplot of the rivers.
  - Make a table of the rivers crossed with the species of fish.
  - Make a barplot of the rivers crossed with the species of fish.
- Task 5
  - Make a pie chart of Fish SPECIES
  - Make a pie chart of RIVERS
- Task 6
  - Make a boxplot of the variable DDT
  - Make a boxplot of the variable WEIGHT
  - Make a boxplot of the variable LENGTH
- Task 7
  - Make a coplot of LENGTH Vs WEIGHT given RIVER
  - Make a coplot of DDT Vs WEIGHT given SPECIES
- Task 8 Make sure all code is placed in with your answers!
  - Use `ggplot` and make the following boxplots (be sure to put your name on the title using `ggtitle()`)

Put your name here



- Use ggplot and make the following violin plots (name on the title)

Put your name here



- Use ggplot and make the following scatter plots (name on the title)

Put your name here

