# CMPU4091 Visualising Data

# Practical Class Week 1

# Getting started with R, RStudio and ggplot

**Objectives:**

* Install the necessary technology (R and RStudio)
* Familiarize yourself with R and RStudio
* Familiarize yourself with Data Exploration using R:
  + Display documentation about a dataset
  + Load a dataset and examine its structured
  + Display and analyse the data with functions like head(), nrow(), summary(), and str().
* Learn the basics of using ggplot to create basic plots:
  + Bar Charts
  + Pie Charts
  + Histograms
  + Customizing colours, labels and using themes.
  + Manipulating data to suit different plot types.
* Practice!

Contents

[Step 1: Set up R and R Studio on your own machine. 1](#_Toc188888521)

[Step 2: Install and load necessary packages 2](#_Toc188888522)

[Step 3: Working with a script 3](#_Toc188888523)

[Step 4: Open and run L1-datasaurus.r 4](#_Toc188888524)

[Step 5: Working with ggplot 4](#_Toc188888525)

[Step 6: Complete the exercises at the end of P1-ggplot-beginner.r 4](#_Toc188888526)

# Step 1: Set up R and R Studio on your own machine.

* Download and install R: <https://cran.r-project.org/>
* Download and install RStudio: <https://posit.co/products/open-source/rstudio/>
* Launch RStudio
  + A window with three panes should appear (see below)
  + The largest pane is a console window. This is where you’ll run your R code and see results.
  + In the other panes (some hidden) are a text editor, a graphics window, a debugger, a file manager, and much more. You’ll learn more about those as you go through the lab.

A screenshot of a computer

Description automatically generated

# Step 2: Install and load necessary packages

* There is a core set of functions available in what is called base R. But many of the functions we'll use in R are not built in. Instead, they are organized into packages that can be installed and then loaded as needed.
* To install a package you use the install.packages() command passing in the name of the package you want to install.
* We need to install for the moment tidyverse and datasauRus
  + At the console enter the commands

install.packages("tidyverse")

install.packages("datasauRus")

Note: you may get warnings about Rtools, packages being built with older versions of R – you can ignore these

* Once you have installed the packages each time you want to use them you need to load them:

library(tidyverse)

library(datasauRus)

* Now you have access to all the data and functions included with these two packages for the duration of your RStudio session.
* Note: to clear the console – use the sweeping brush icon at the top right of the console pane or use CTRL + L (Windows) or Command + L (Mac)

# Step 3: Working with a script

* Open the file **L1-Anscombe.r** 
  + **File -> Open or click on** 
* The script pane will open showing the R code:

A screenshot of a computer

Description automatically generated

* This script will load the datasets package (included with base R)
* It will load the Anscombe datasets
* It will set up a 2 x 2 grid in which 4 scatterplots will be displayed
* It will create 4 plots using each of the 4 Anscombe datasets and display them.
* To run it, click on the Source button  at the top right of the script pane.
  + The plots will appear in the Plots pane – bottom right.

A screenshot of a computer

Description automatically generated

* Alternatively, you can run it line by line, just position your cursor in the line of code and bit CTRL + Enter (Windows) or Command + Enter (Mac).
  + Or run a set of lines by selecting them and then CTRL + Enter (Windows) or Command + Enter (Mac).

# Step 4: Open and run L1-datasaurus.r

* + You may get warnings about datasaurus being built with an older version of R – ignore.
  + Just get used to the interface.

# Step 5: Working with ggplot

* + Open the R script **P1-ggplot-beginner.r** we talked about during the lecture.
  + Run the code
    1. It is recommended that you run it line by line to follow what is going on
    2. READ the comments – they explain clearly how each aspect of the code words which will help you with the exercises later.

# Step 6: Complete the exercises at the end of P1-ggplot-beginner.r

* + Note: You may use AI assistance, but if you do so, explain clearly in your code comments where you got your information and the prompts you supplied.

**Note: There is no need to demonstrate your code this week.**