Week 0 – Getting Ready

WORKSHEET 1 Installing R and RStudio

Aims and Objectives
This worksheet will walk you through the installation of R, R Studio and an R package. This worksheet will also enable you to test that your install is working.
Note that you should copy and paste the code and link **in the text format** as much as possible.

Step 1: Download and Install R

Link: https://cran.r-project.org/

Copy the link above and paste it onto a web browser.

For Windows user: Click on 'Download R for Windows' -> 'base' -> 'Download R {version} for Windows'. Then proceed to install it.

For Mac user: Click on 'Download R for macOS' -> 'R-{version}.pkg'. Then proceed to install it.

Step 2: Download and Install RStudio

Link: https://www.rstudio.com/products/rstudio/download/#download

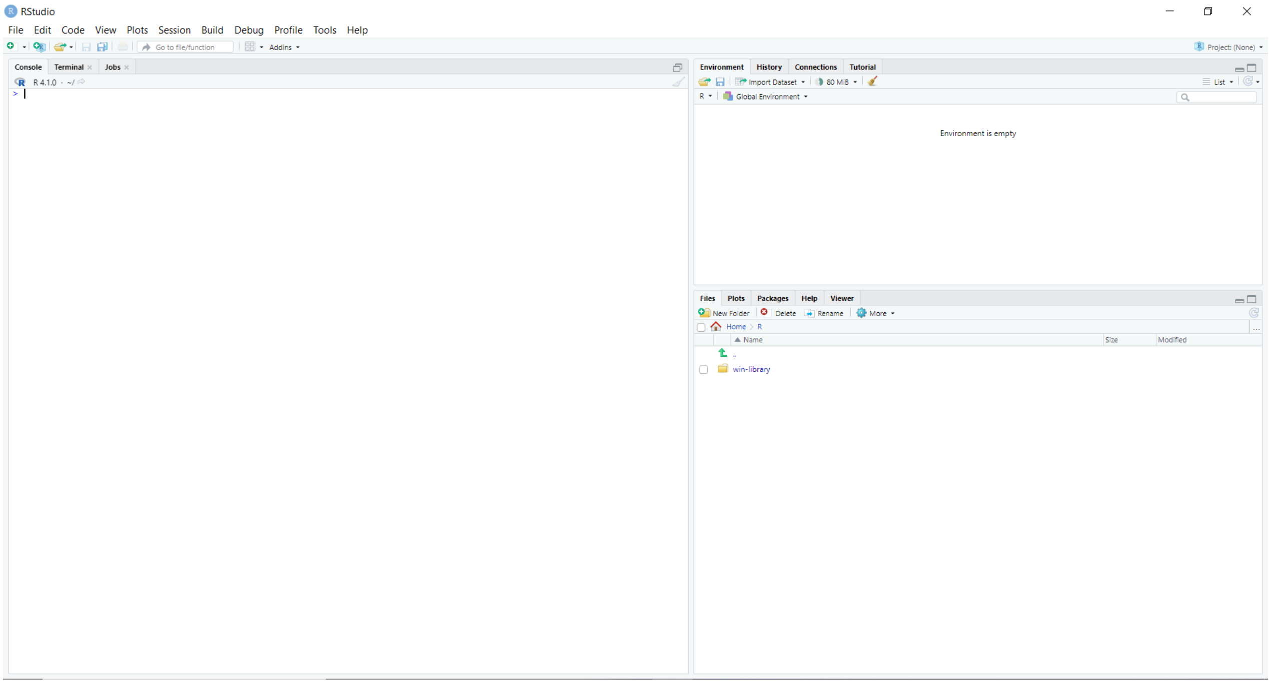
Copy the link above and paste it onto a web browser.

For Windows users: Click on 'DOWNLOAD RSTUDIO FOR WINDOWS. Then proceed to install it.

For MacOS users: Scroll down a little bit, and you will find a link for macOS. Click on 'RStudio-{version}.dmg'. Then proceed to install it.

Step 3: Starting up RStudio

When you open RStudio for the first time, you will get a window as shown in Figure 1.


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Console Terminal Jobs
R 4.1.0 · ~/
Environment History Connections Tutorial
Import Dataset
80 MiB
List
R
Global Environment
Files Plots Packages Help Viewer
New Folder Delete Rename More
Home R
Name Size Modified
..
win-library

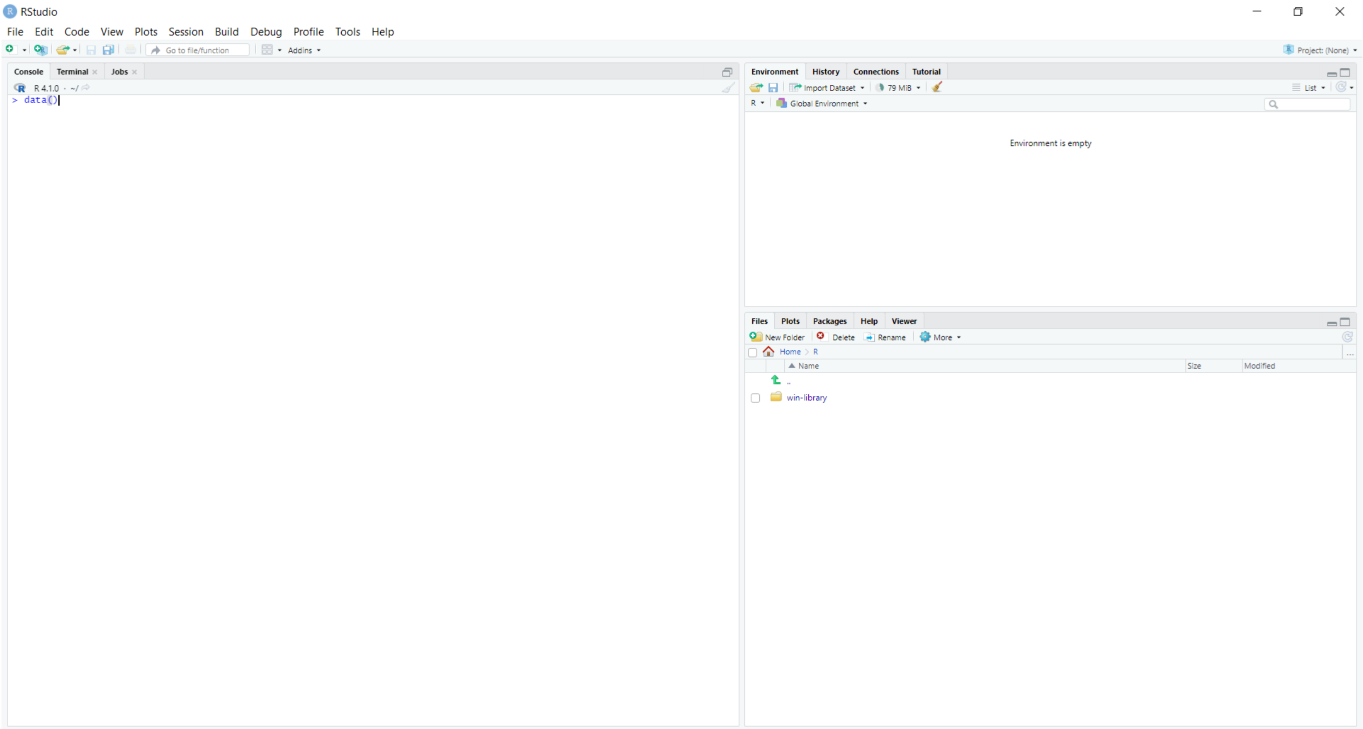
```
>
```

Environment is empty

Figure 1 R studio

The left panel is the console where you type in the commands, the top right panel (environment) shows the data and variables, and the bottom right panel is the working directory.

To test if R is working, we will load the built-in datasets. To view all of the built-in datasets, we need to type in `data()` into the console then press enter (see Figure 2) .

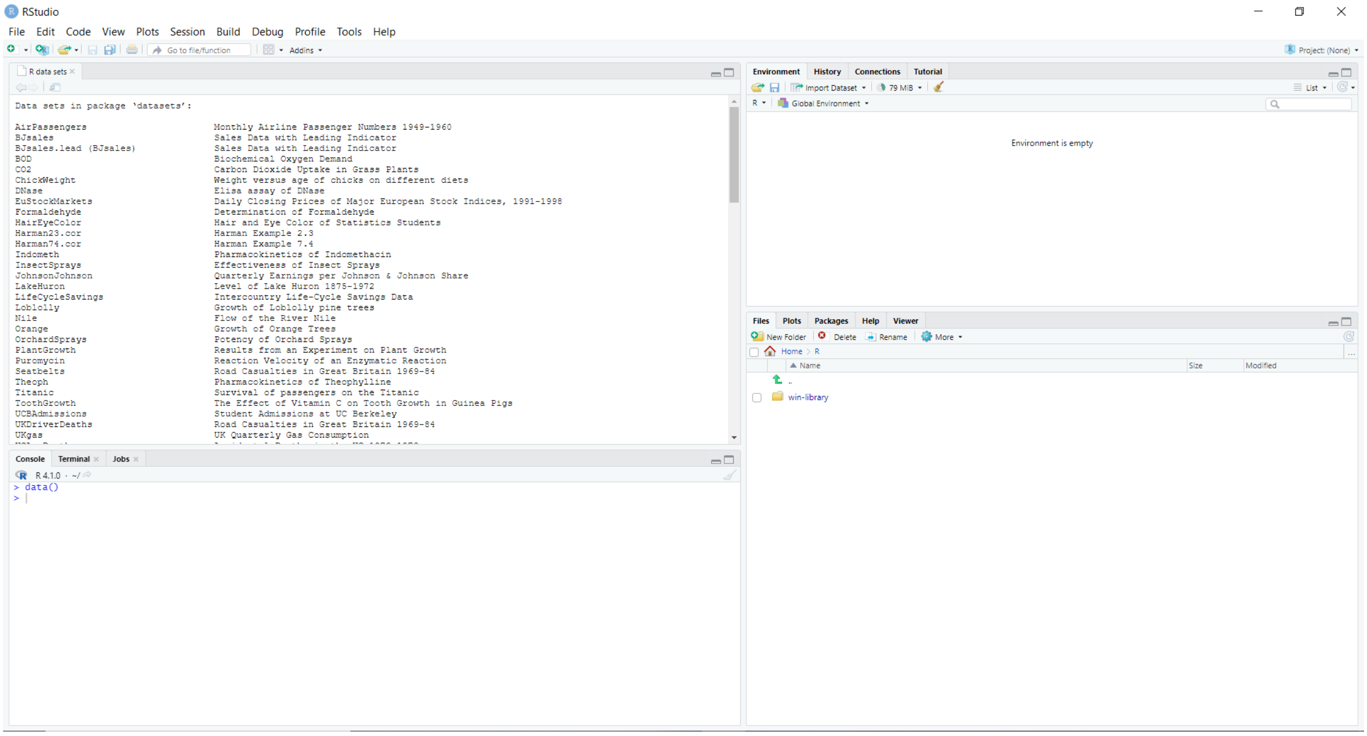

RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Project: (None)
Console Terminal Jobs
R 4.1.0 · ~/
Environment History Connections Tutorial
Import Dataset
79 MiB
List
R
Global Environment
Files Plots Packages Help Viewer
New Folder Delete Rename More
Home R
Name Size Modified
..
win-library

```
> data()
```

Environment is empty

Figure 2 Testing that R is working

This will open a new panel called ‘R data sets’ showing all the built-in datasets that we can use (See Figure 3) .

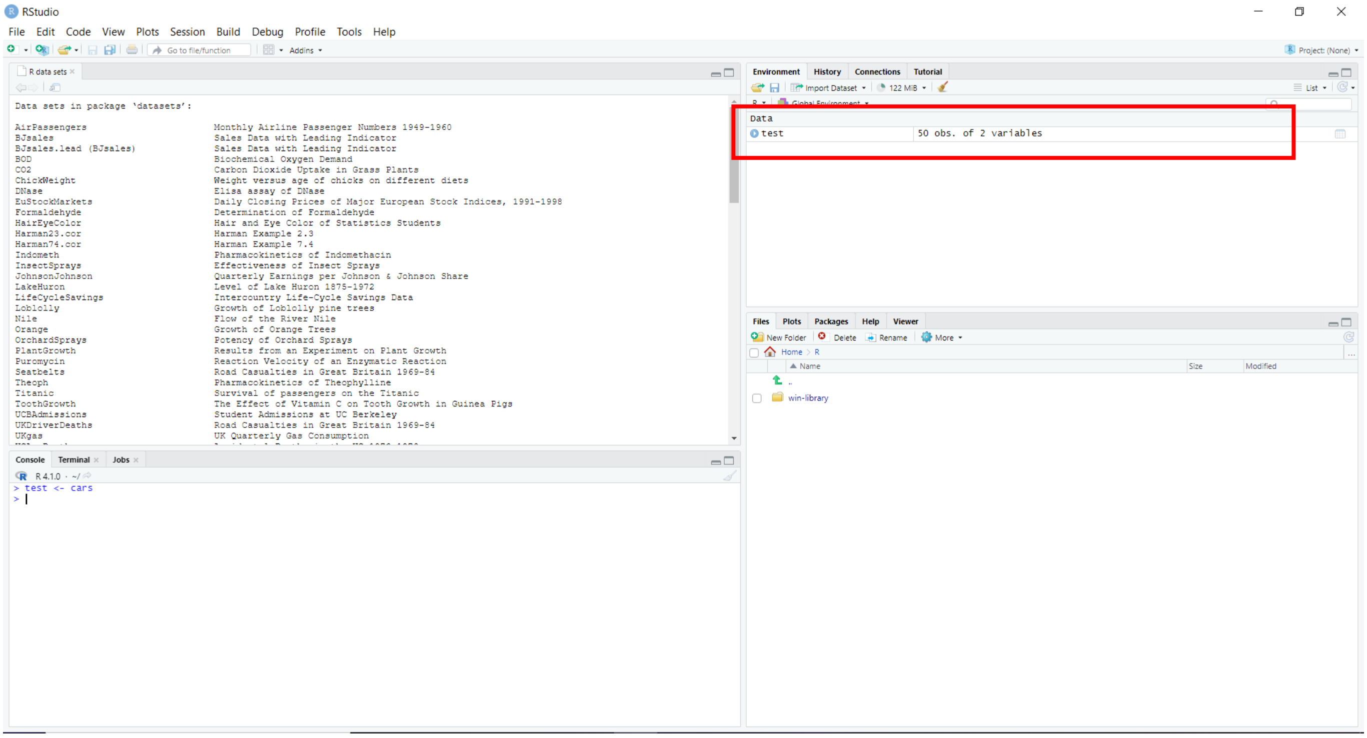


```
Data sets in package ‘datasets’:

AirPassengers                Monthly Airline Passenger Numbers 1949-1960
BJsales                      Sales Data with Leading Indicator
BJsales.lead (BJsales)       Sales Data with Leading Indicator
BOD                          Biochemical Oxygen Demand
CO2                          Carbon Dioxide Uptake in Grass Plants
ChickWeight                  Weight versus age of chicks on different diets
DNase                        Elisa assay of DNase
EuStockMarkets               Daily Closing Prices of Major European Stock Indices, 1991-1998
Formaldehyde                 Determination of Formaldehyde
HairEyeColor                 Hair and Eye Color of Statistics Students
Harman23.cor                 Harman Example 2.3
Harman74.cor                 Harman Example 7.4
Indometh                     Pharmacokinetics of Indomethacin
InsectSprays                 Effectiveness of Insect Sprays
JohnsonJohnson               Quarterly Earnings per Johnson & Johnson Share
LakeHuron                    Level of Lake Huron 1875-1972
LifeCycleSavings             Intercountry Life-Cycle Savings Data
Loblolly                     Growth of Loblolly pine trees
Nile                         Flow of the River Nile
Orange                       Growth of Orange Trees
OrchardSprays                Potency of Orchard Sprays
PlantGrowth                  Results from an Experiment on Plant Growth
Puromycin                    Reaction Velocity of an Enzymatic Reaction
Seatbelts                    Road Casualties in Great Britain 1969-84
Theoph                       Pharmacokinetics of Theophylline
Titanic                      Survival of passengers on the Titanic
ToothGrowth                  The Effect of Vitamin C on Tooth Growth in Guinea Pigs
UCBAdmissions                Student Admissions at UC Berkeley
UKDriverDeaths               Road Casualties in Great Britain 1969-84
UKgas                        UK Quarterly Gas Consumption
```

```
> data()
> 
```

Figure 3 The R data sets panel shows the built in data sets

Let’s take a closer look at the cars dataset. We will store cars in a variable called ‘test’ then explore the contents of this data set. Type in `test <- cars` then press enter. We will see ‘test’ in the environment panel (See Figure 4) . It is not necessary to store `cars` in as a new element `test` , but we want to make it easier to visualise what is inside that dataset for this demonstration.



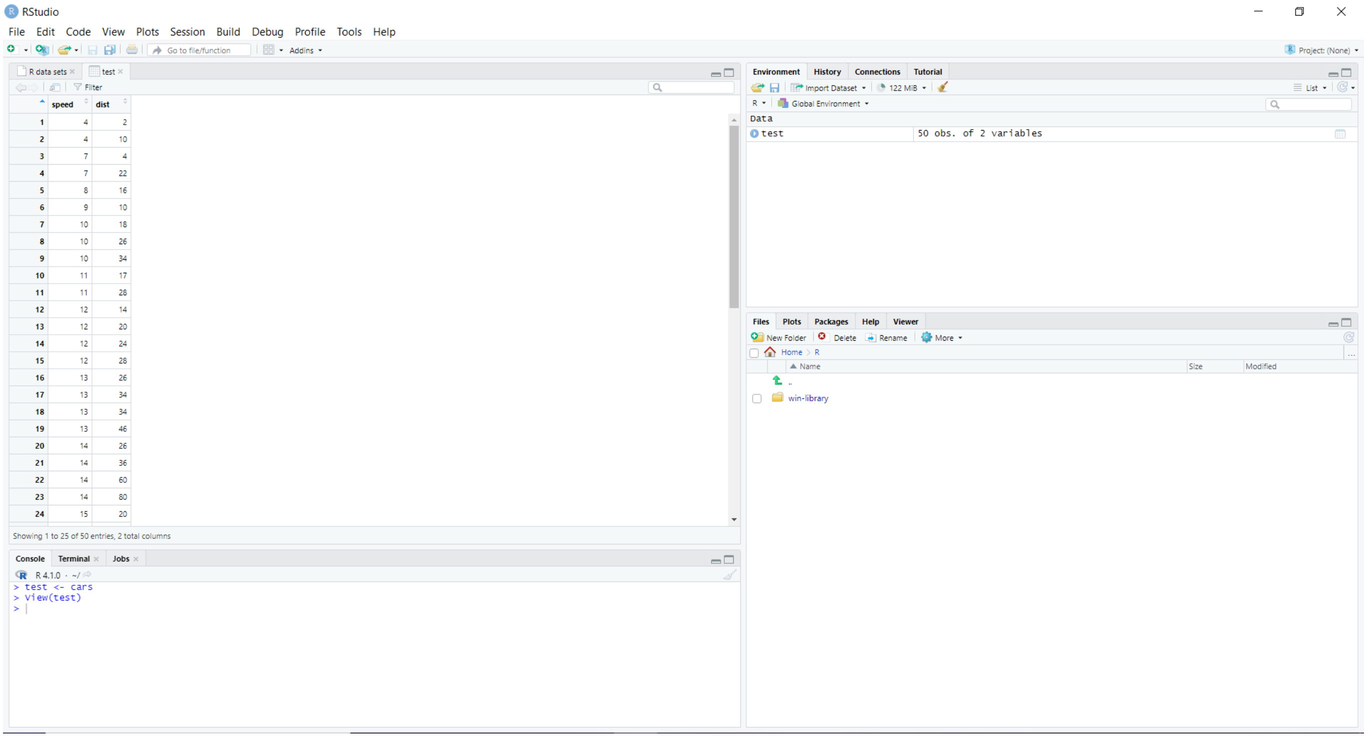
```
Data sets in package ‘datasets’:

AirPassengers                Monthly Airline Passenger Numbers 1949-1960
BJsales                      Sales Data with Leading Indicator
BJsales.lead (BJsales)       Sales Data with Leading Indicator
BOD                          Biochemical Oxygen Demand
CO2                          Carbon Dioxide Uptake in Grass Plants
ChickWeight                  Weight versus age of chicks on different diets
DNase                        Elisa assay of DNase
EuStockMarkets               Daily Closing Prices of Major European Stock Indices, 1991-1998
Formaldehyde                 Determination of Formaldehyde
HairEyeColor                 Hair and Eye Color of Statistics Students
Harman23.cor                 Harman Example 2.3
Harman74.cor                 Harman Example 7.4
Indometh                     Pharmacokinetics of Indomethacin
InsectSprays                 Effectiveness of Insect Sprays
JohnsonJohnson               Quarterly Earnings per Johnson & Johnson Share
LakeHuron                    Level of Lake Huron 1875-1972
LifeCycleSavings             Intercountry Life-Cycle Savings Data
Loblolly                     Growth of Loblolly pine trees
Nile                         Flow of the River Nile
Orange                       Growth of Orange Trees
OrchardSprays                Potency of Orchard Sprays
PlantGrowth                  Results from an Experiment on Plant Growth
Puromycin                    Reaction Velocity of an Enzymatic Reaction
Seatbelts                    Road Casualties in Great Britain 1969-84
Theoph                       Pharmacokinetics of Theophylline
Titanic                      Survival of passengers on the Titanic
ToothGrowth                  The Effect of Vitamin C on Tooth Growth in Guinea Pigs
UCBAdmissions                Student Admissions at UC Berkeley
UKDriverDeaths               Road Casualties in Great Britain 1969-84
UKgas                        UK Quarterly Gas Consumption
```

```
Data
test                    50 obs. of 2 variables
```

```
> test <- cars
> 
```

Figure 4 Looking at the contents of the 'cars" dataset

We can double click on ‘test’ in the environment panel or type `View(test)` onto the console to view what the cars dataset look like.



	speed	dist
1	4	2
2	4	10
3	7	4
4	7	22
5	8	16
6	9	10
7	10	18
8	10	26
9	10	34
10	11	17
11	11	28
12	12	14
13	12	20
14	12	24
15	12	28
16	13	26
17	13	34
18	13	34
19	13	46
20	14	26
21	14	36
22	14	60
23	14	80
24	15	20

Showing 1 to 25 of 50 entries, 2 total columns

```
> test <- cars
> View(test)
> |
```

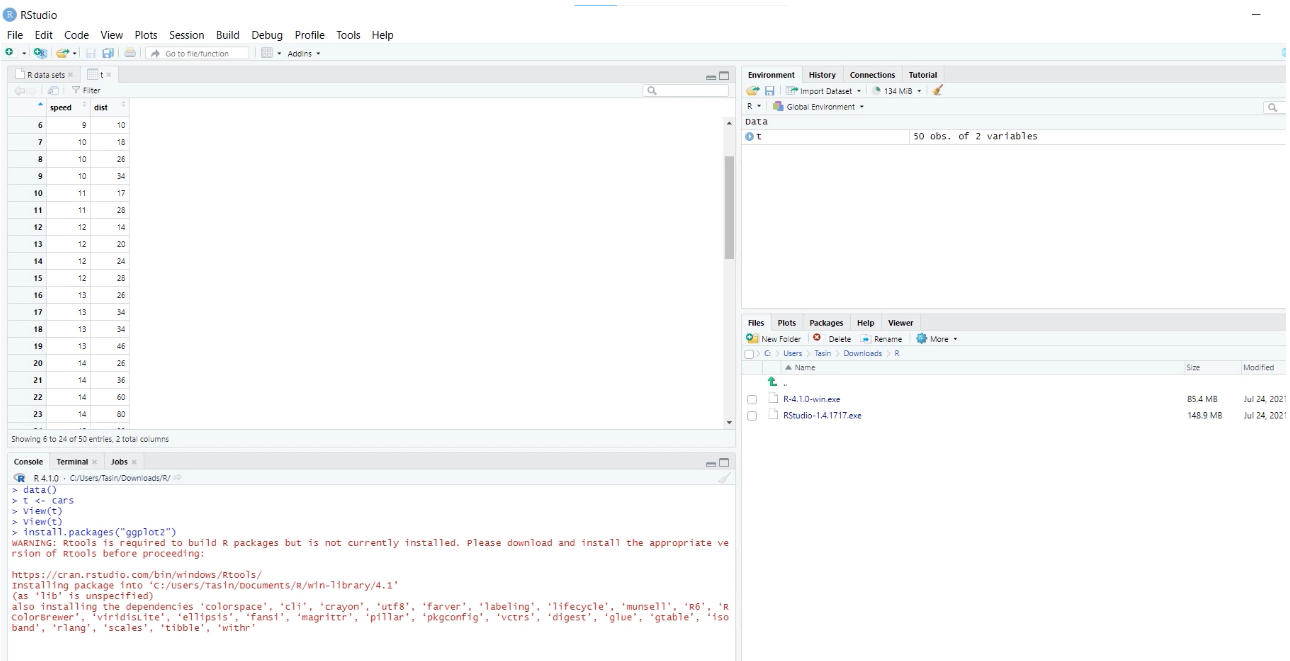
```
Data
test                50 obs. of 2 variables
```

Figure 5 Viewing the contents of the 'test' data

We should see some rows and columns; speed represents how fast the car travelled before hitting the brakes. Dist is the distance travelled by a vehicle after the brakes have been applied (See Figure 5).

Step 4: Installing ggplot2

In order to make full use of the R language you will occasionally be installing libraries of code referred to as packages. To install packages in R, you need to use the `install.packages()` command. ggplot2 can be installed by typing in `install.packages('ggplot2')` then hitting enter (See Figure 6) .


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
R data sets
t
Filter
Environment History Connections Tutorial
Import Dataset
134 MiB
Global Environment
Files Plots Packages Help Viewer
New Folder Delete Rename More
C: > Users > Tasin > Downloads > R
Console Terminal Jobs

	speed	dist
6	9	10
7	10	18
8	10	26
9	10	34
10	11	17
11	11	28
12	12	14
13	12	20
14	12	24
15	12	28
16	13	26
17	13	34
18	13	34
19	13	46
20	14	26
21	14	36
22	14	60
23	14	80

Showing 6 to 24 of 50 entries, 2 total columns

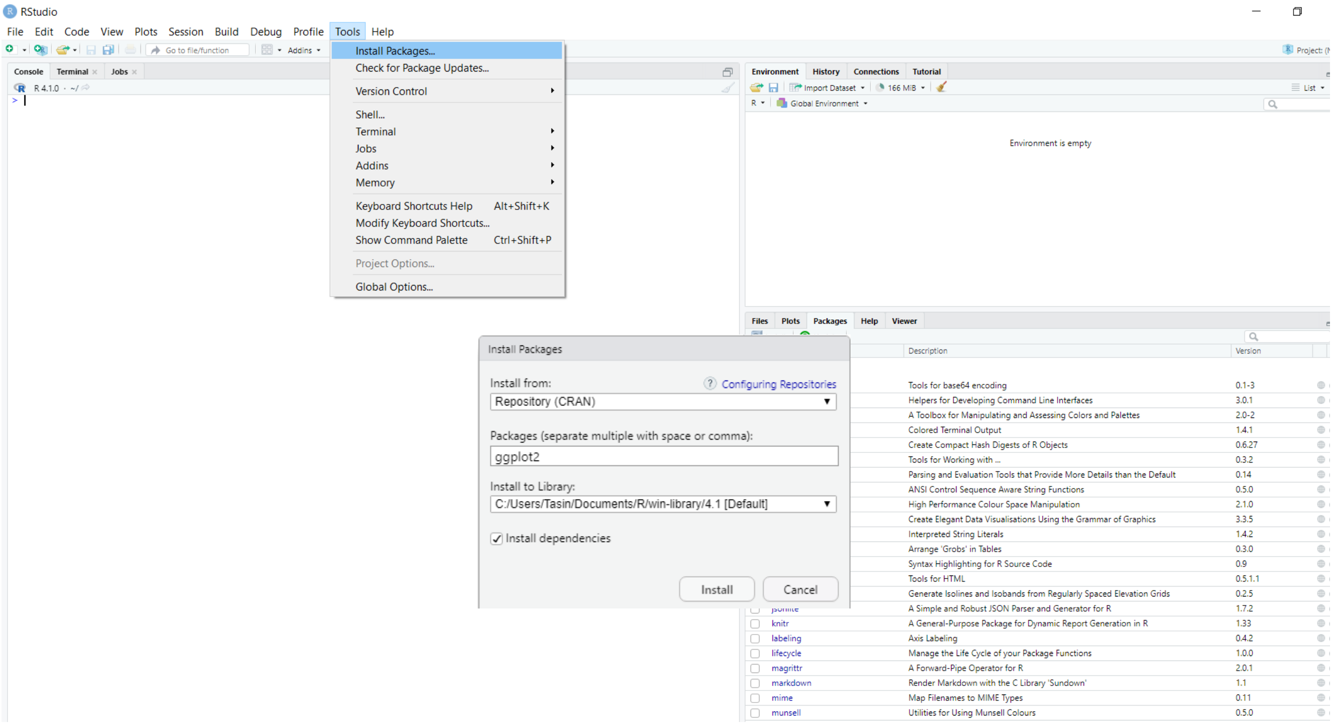
```
Data
t                    50 obs. of 2 variables
```

Name	Size	Modified
..		
R-4.1.0-win.exe	85.4 MB	Jul 24, 2021
RStudio-1.4.1717.exe	148.9 MB	Jul 24, 2021

```
R 4.1.0 · C:/Users/Tasin/Downloads/R/
> data()
> t <- cars
> View(t)
> View(t)
> install.packages("ggplot2")
WARNING: Rtools is required to build R packages but is not currently installed. Please download and install the appropriate ve
rsion of Rtools before proceeding:

https://cran.rstudio.com/bin/windows/Rtools/
Installing package into 'C:/Users/Tasin/Documents/R/win-library/4.1'
(as 'lib' is unspecified)
also installing the dependencies 'colorspace', 'cli', 'crayon', 'utf8', 'farver', 'labeling', 'lifecycle', 'munsell', 'R6', 'R
ColorBrewer', 'viridisLite', 'ellipsis', 'fansi', 'magrittr', 'pillar', 'pkgconfig', 'vctrs', 'digest', 'glue', 'gtable', 'iso
band', 'rlang', 'scales', 'tibble', 'withr'
```

Figure 6 Installing the 'ggplot2' package - see the lower left hand side of this screen

Alternatively you can use R-Studio menus : click on Tools -> Install Packages..., a window should pop up. In the packages field, type in `ggplot2` (See Figure 7) .


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function
Addins
Console Terminal Jobs
R 4.1.0 · ~/
Install Packages...
Check for Package Updates...
Version Control
Shell...
Terminal
Jobs
Addins
Memory
Keyboard Shortcuts Help Alt+Shift+K
Modify Keyboard Shortcuts...
Show Command Palette Ctrl+Shift+P
Project Options...
Global Options...
Environment History Connections Tutorial
Import Dataset
166 MiB
List
Global Environment
Environment is empty
Project: (
Files Plots Packages Help Viewer
Install
Cancel

Install Packages

Install from: Repository (CRAN) (?) Configuring Repositories

Packages (separate multiple with space or comma): ggplot2

Install to Library: C:/Users/Tasin/Documents/R/win-library/4.1 [Default]

☑ Install dependencies

		Description	Version
□		Tools for base64 encoding	0.1-3
□		Helpers for Developing Command Line Interfaces	3.0.1
□		A Toolbox for Manipulating and Assessing Colors and Palettes	2.0-2
□		Colored Terminal Output	1.4.1
□		Create Compact Hash Digests of R Objects	0.6.27
□		Tools for Working with ...	0.3.2
□		Parsing and Evaluation Tools that Provide More Details than the Default	0.14
□		ANSI Control Sequence Aware String Functions	0.5.0
□		High Performance Colour Space Manipulation	2.1.0
□		Create Elegant Data Visualisations Using the Grammar of Graphics	3.3.5
□		Interpreted String Literals	1.4.2
□		Arrange 'Grobs' in Tables	0.3.0
□		Syntax Highlighting for R Source Code	0.9
□		Tools for HTML	0.5.1.1
□		Generate Isolines and Isobands from Regularly Spaced Elevation Grids	0.2.5
□	jsonlite	A Simple and Robust JSON Parser and Generator for R	1.7.2
□	knitr	A General-Purpose Package for Dynamic Report Generation in R	1.33
□	labeling	Axis Labeling	0.4.2
□	lifecycle	Manage the Life Cycle of your Package Functions	1.0.0
□	magrittr	A Forward-Pipe Operator for R	2.0.1
□	markdown	Render Markdown with the C Library 'Sundown'	1.1
□	mime	Map Filenames to MIME Types	0.11
□	munsell	Utilities for Using Munsell Colours	0.5.0

Figure 7 Installiing packages using the R Studio Menus

Packages are collections of R functions, data, and compiled code in a well-defined format, created to add specific functionality. ggplot2 is a powerful graphics language for creating elegant and complex plots. It provides a more programmatic interface for specifying what variables to plot, how they are displayed, and general visual properties.

There are hundreds of such packages, if you want to find out and read about the most popular one see https://www.rstudio.com/products/rpackages/ and if you want an exhaustive list take a look at this https://cran.r-project.org/web/packages/available_packages_by_name.html .