資料倉儲導論-第2章 免費R語言,RStudio簡介,下載與安裝

大數據分析

- R/Python/Julia/SQL程式設計與應用 (R/Python/Julia/SQL Programming and Application)
- 資料視覺化 (Data Visualization)
- 機器學習 (Machine Learning)
- 統計品管 (Statistical Quality Control)
- 最佳化 (Optimization)

李明昌博士

alan9956@gmail.com

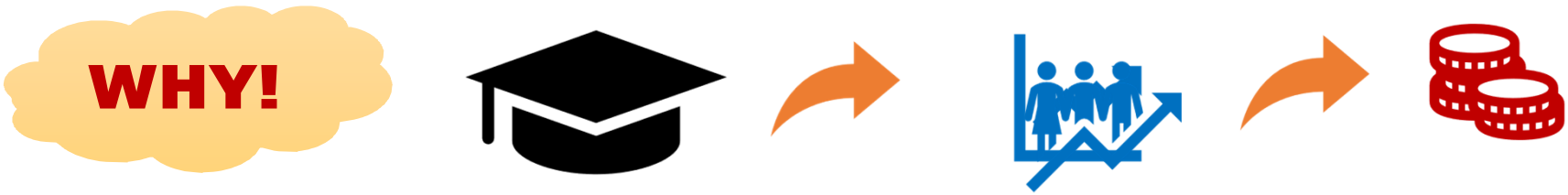
http://rwepa.blogspot.com/

大綱

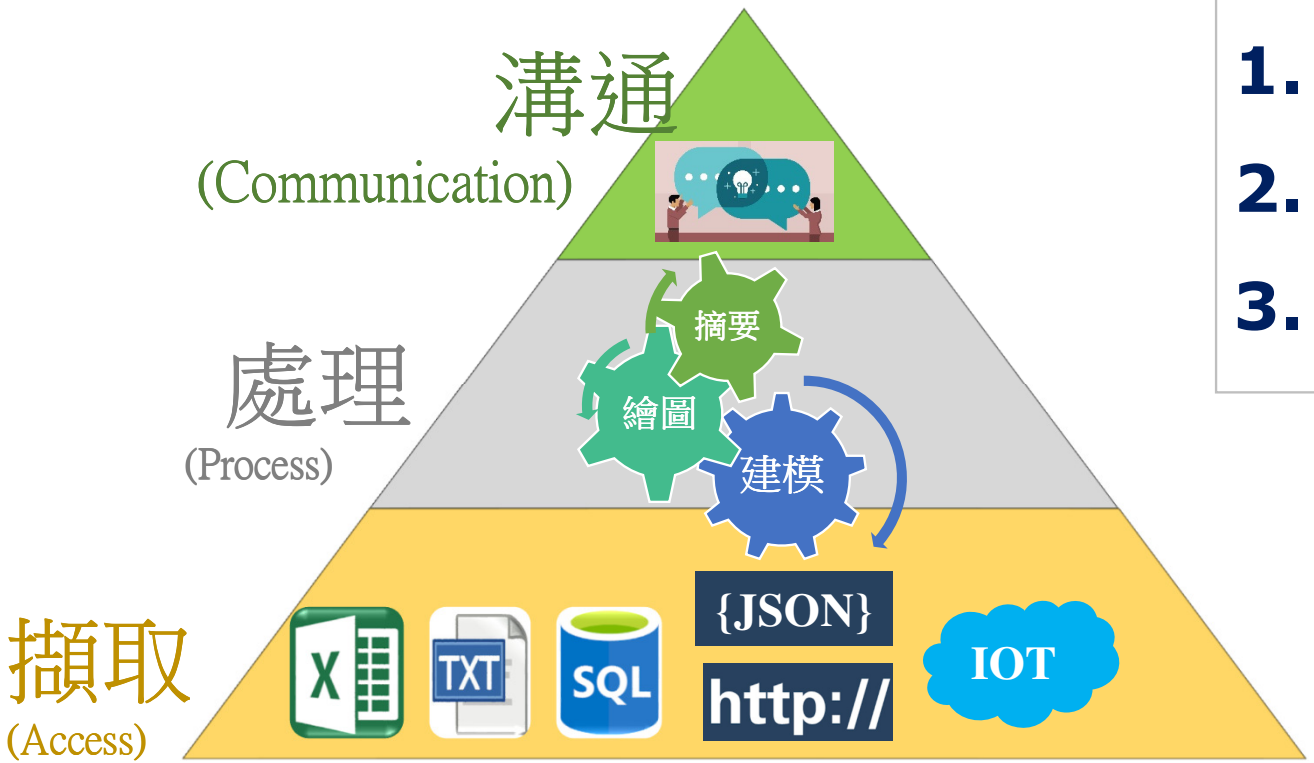
2.1 資料科學的心法

資料科學的心法

1. 熟悉教材內容
2. 掌握 APC方法
3. 掌握 摘要, 繪圖, 建模
4. 遇到問題時, 想辦法尋找答案
5. 將教材的資料集改為工作資料集
6. 參考網路應用文章 (進階) & 學術論文



★★★資料分析架構→APC方法



1.

2.

3.

R 入門資料分析與視覺化應用(7小時28分鐘)

- https://mastertalks.tw/products/r?ref=MCLEE

課程提供教學範例的原始程式檔案與資料集

- 主題

1. R, RStudio簡介與套件使用
2. 認識資料物件
3. 資料處理與分析
4. 資料視覺化應用

- 特色

1. 資料分析的**關鍵八步**
2. 提供必備**ggplot2**套件的應用知識與使用情境
3. 提供日期時間**zoo, xts套件**的整合應用操作
4. 提供**人力資源**資料與**銷售資料**，強化**實務資料**操作能力

R 商業預測應用(8小時53分鐘)

- https://mastertalks.tw/products/r-2?ref=MCLEE

課程提供教學範例的原始程式檔案與資料集

- 主題

1. R，RStudio工具操作
2. 非監督式學習業預測
3. 監督式學習商業預測
4. 財金資料預測應用

- 特色

1. 採用**最有效率**方式學習大數據R語言，並應用於**職場資料分析**與**商業預測應用**
2. 提供**多元線性迴歸**的必備知識
3. 提供**財金資料商業預測應用**的基礎與進階必學技能
4. 提供學員人力資源資料與**台指期tick資料**預測演練

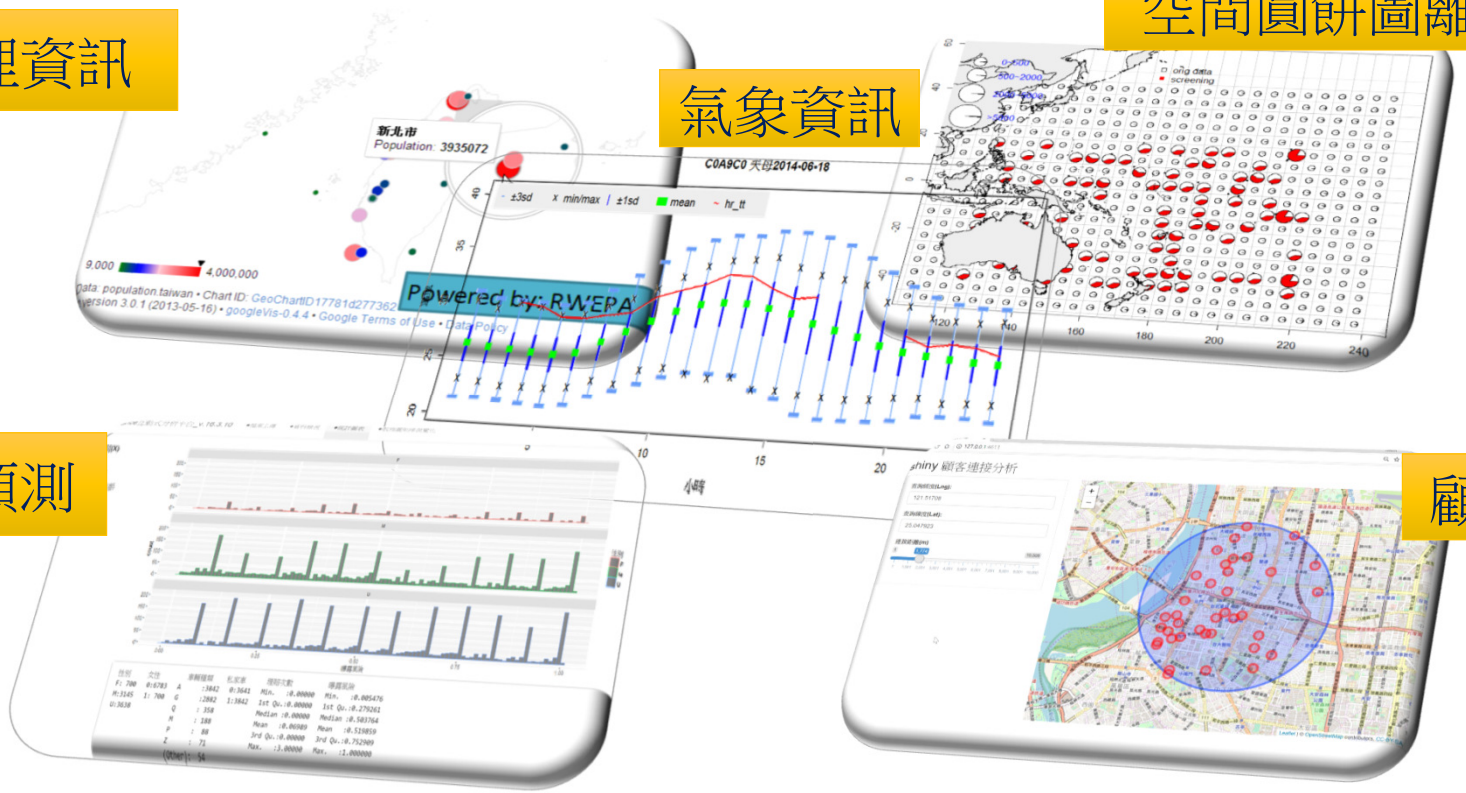
資料分析/視覺化應用

R + shiny → 互動式網頁

地理資訊

空間圓餅圖離群值分析

氣象資訊



保險預測

顧客連結資訊

中央氣象局 1,600萬筆資料

網頁呈現

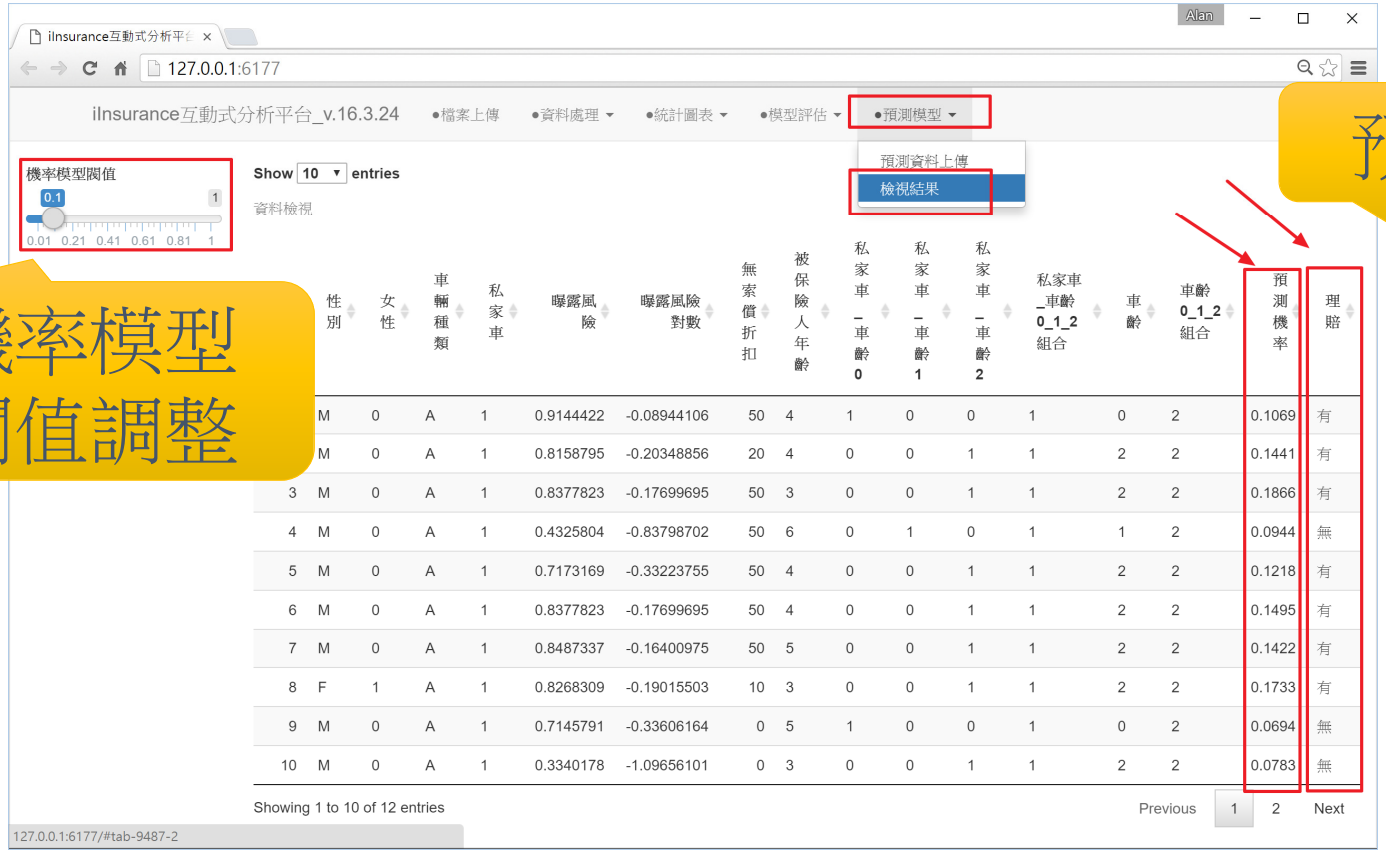
異常偵測

動態繪圖

客製化選單

R統計運算

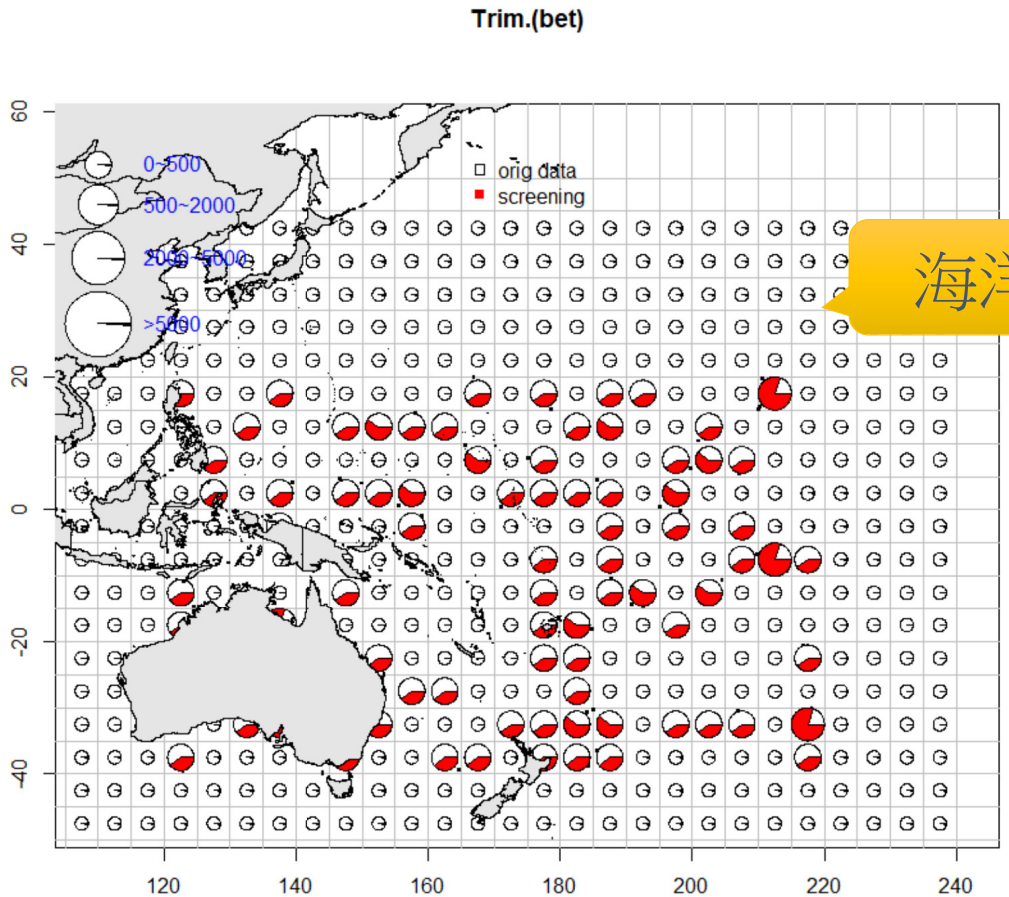
保險預測模型



機率模型閥值調整

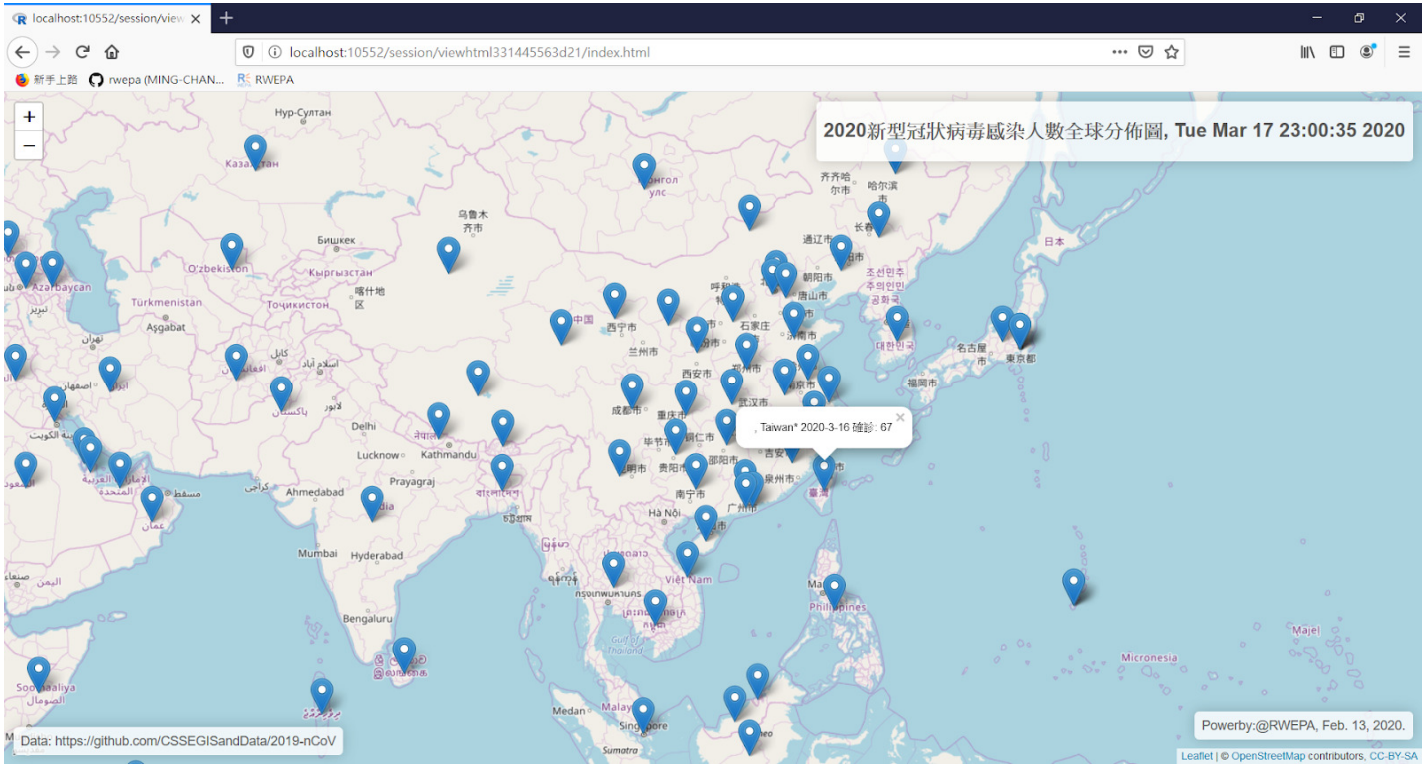
預測結果

空間圓餅圖離群值分析



海洋生態視覺化

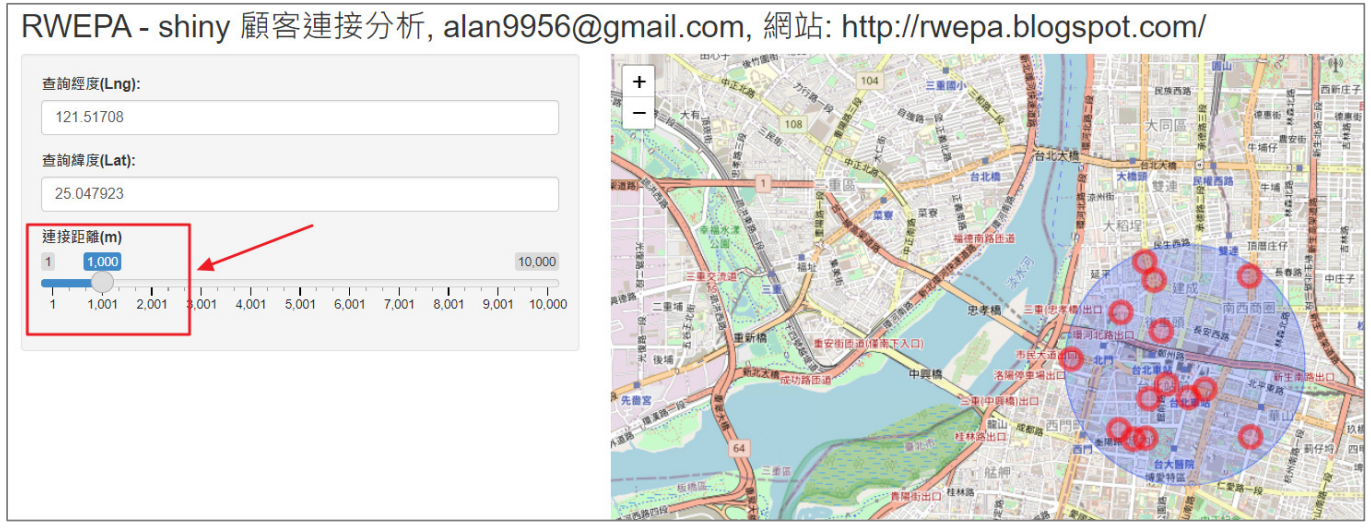
2020新型冠狀病毒視覺化



http://rwepa.blogspot.com/2020/02/2019nCoV.html

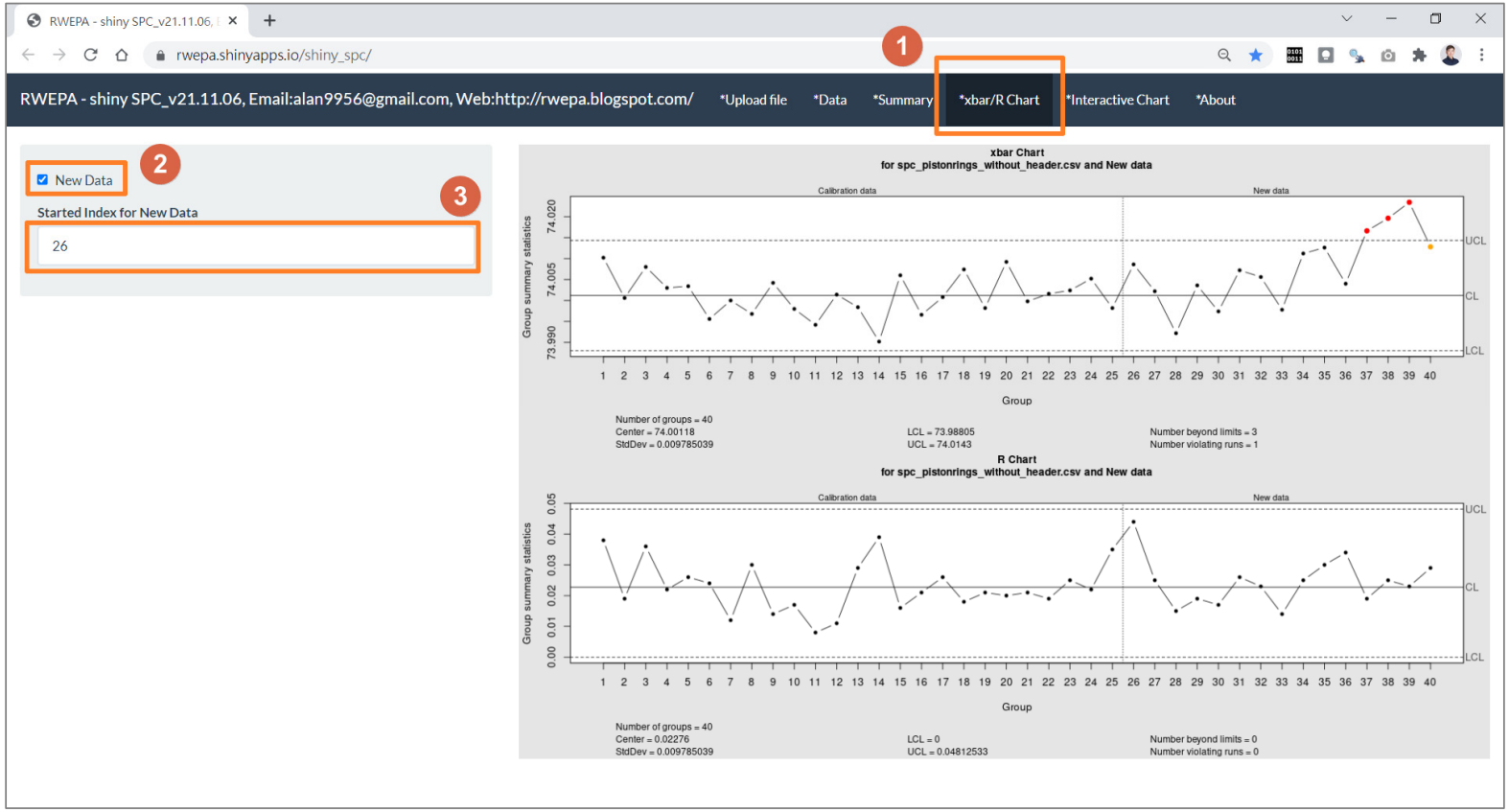
shiny 顧客連接分析

- https://rwepa.shinyapps.io/shinyCustomerConnect/



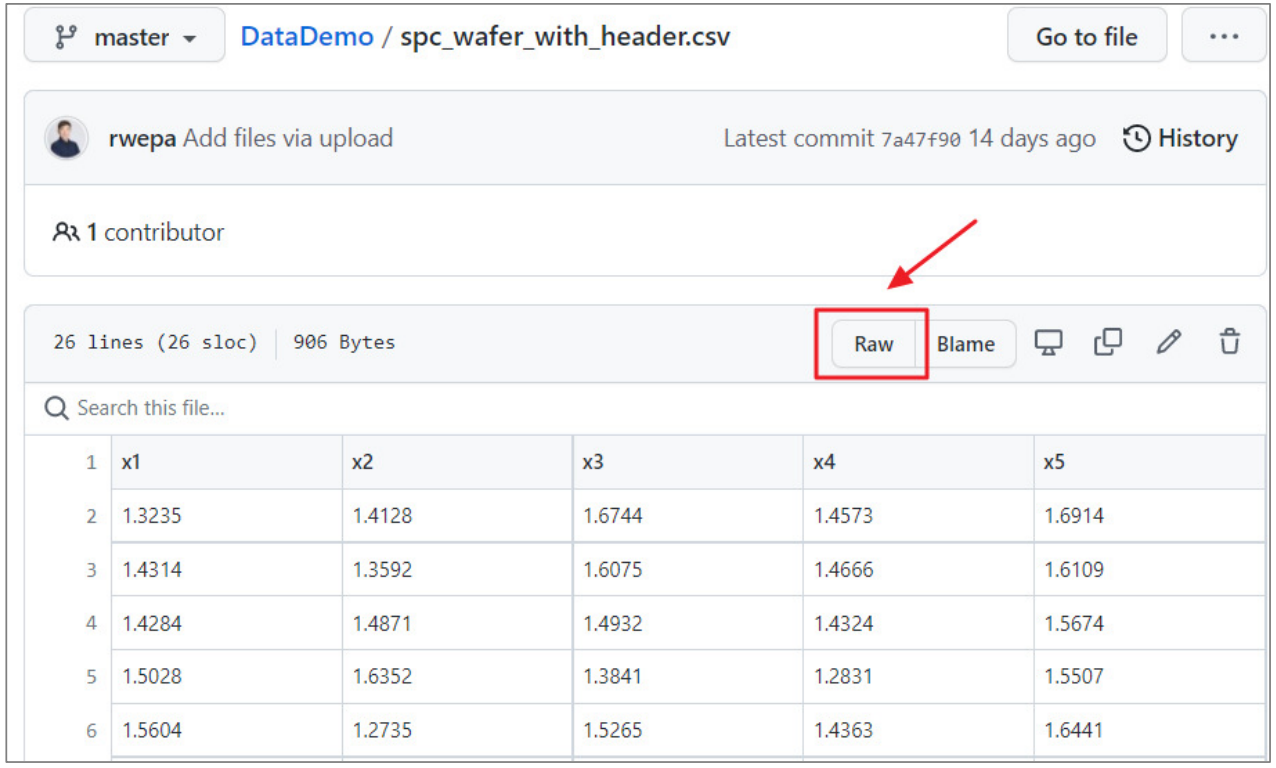
品質管制圖(quality control chart)應用

- http://rwepa.blogspot.com/2021/10/r-shiny-quality-control-chart.html



品質管制圖應用 (續)

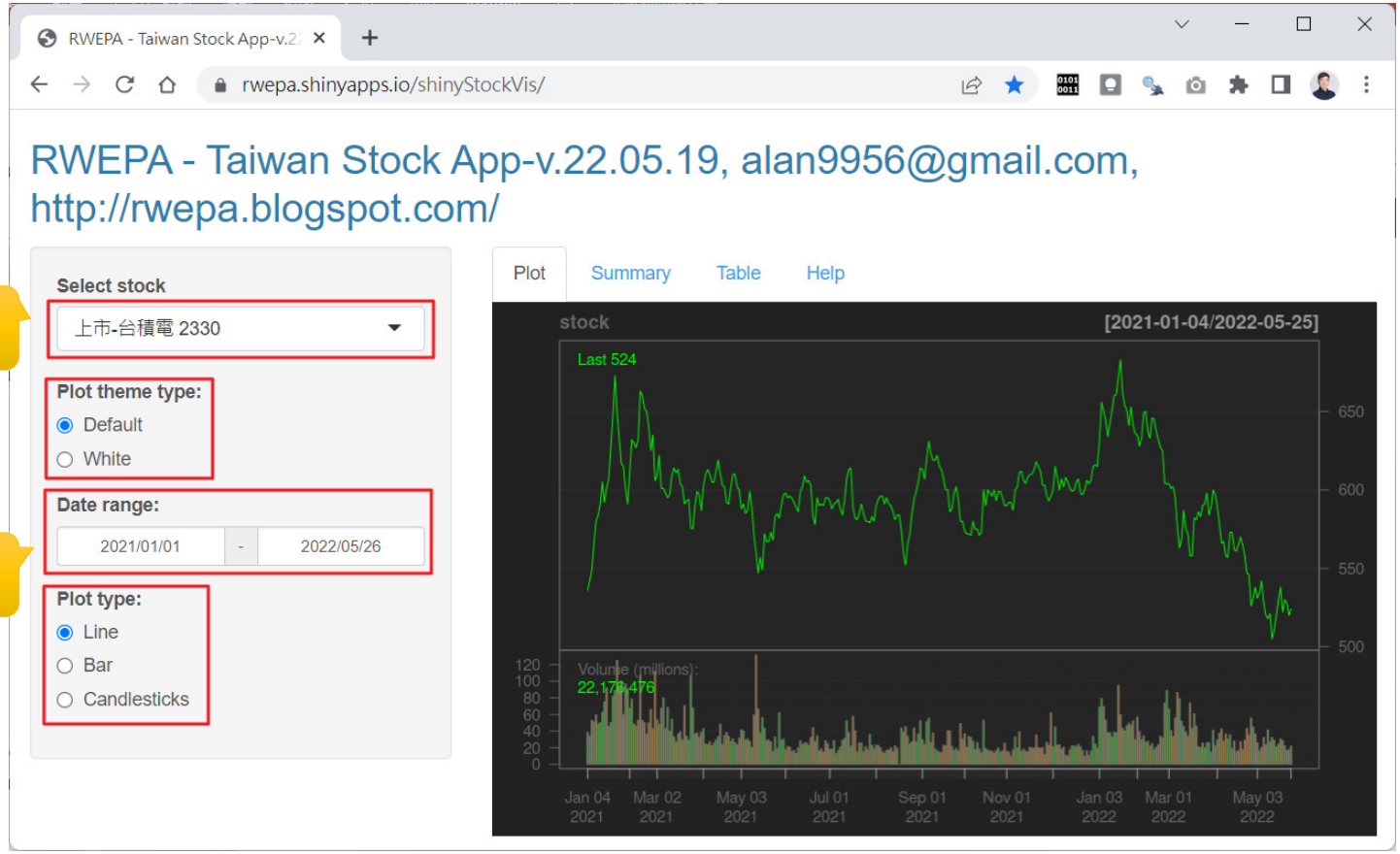
- https://github.com/rwepa/DataDemo/blob/master/spc_wafer_with_header.csv
- https://github.com/rwepa/DataDemo/blob/master/spc_pistonrings_without_header.csv



	x1	x2	x3	x4	x5
2	1.3235	1.4128	1.6744	1.4573	1.6914
3	1.4314	1.3592	1.6075	1.4666	1.6109
4	1.4284	1.4871	1.4932	1.4324	1.5674
5	1.5028	1.6352	1.3841	1.2831	1.5507
6	1.5604	1.2735	1.5265	1.4363	1.6441

Taiwan Stock App

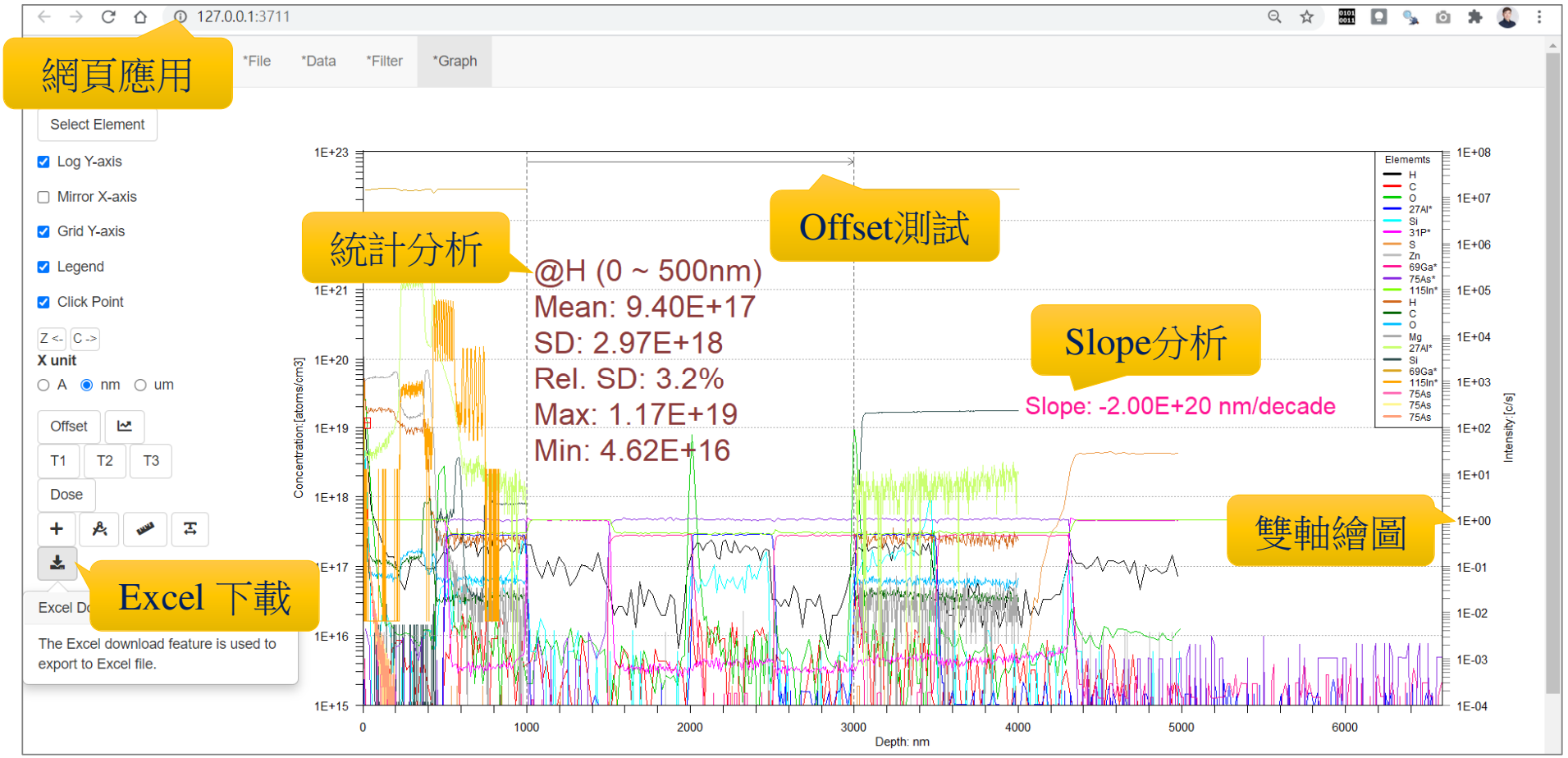
- https://rwepa.shinyapps.io/shinyStockVis/



下拉式選單

日期範圍

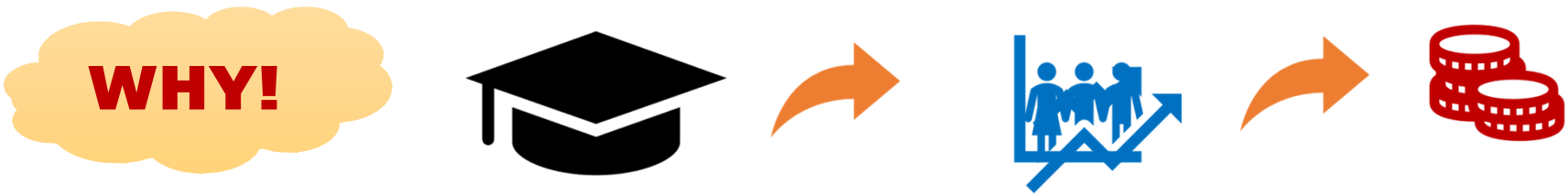
離子資料分析與視覺化應用



學習目標

如何學習 R?

- 熟悉教材內容
- 將教材內容的資料集改為工作資料集
- 遇到問題時, 想辦法尋找答案
- 掌握 APC方法
- 掌握 摘要, 繪圖, 建模
- 參考網路應用文章 (進階) & 學術論文

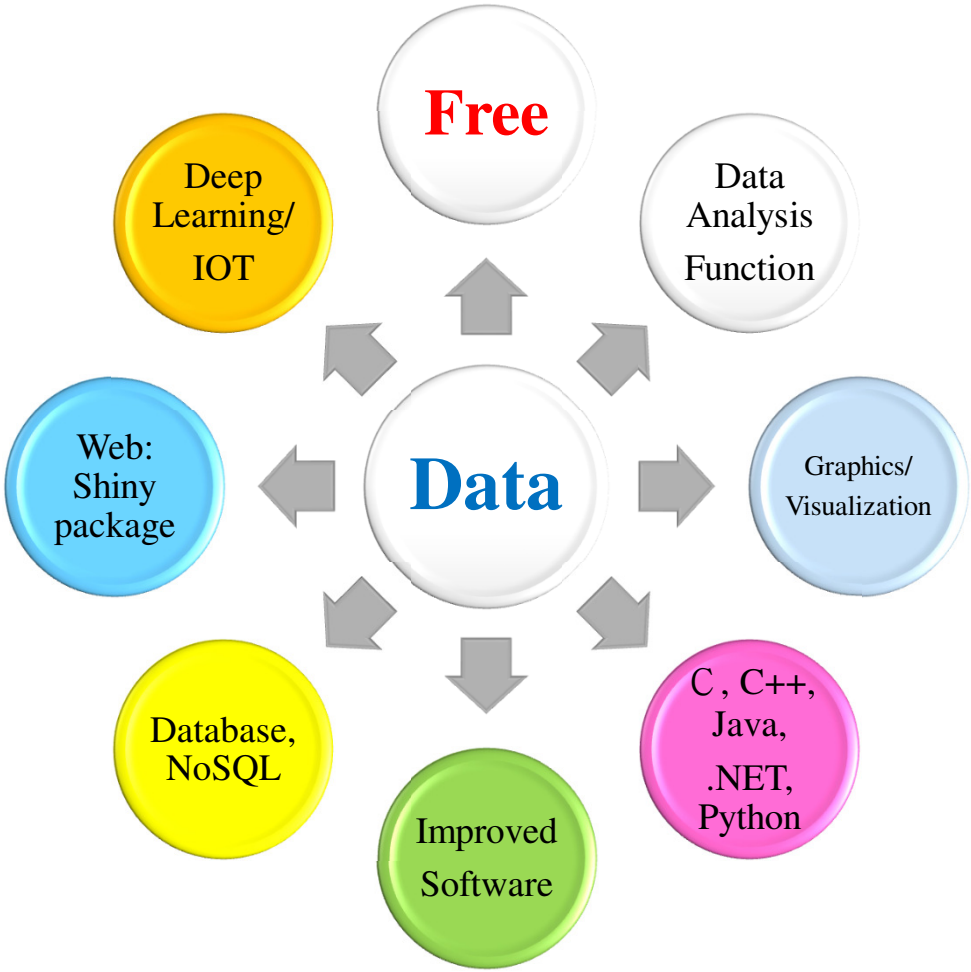


2.2 R語言簡介

認識R

- 1976 - 貝爾實驗室 John Chambers, Rick Becker, and Allan Wilks 研發S語言。
- 1993 - Ross Ihaka and Robert Gentlemen, University of Auckland, New Zealand 研發R 語言。
  - R 是一種基於 S 語言所發展出具備統計分析、繪圖與資料視覺化的程式語言。
- 1997年－R的核心開發團隊 (R development core team) 成立，專責R原始碼的修改與編寫。
  - 2000年2月 － R 1.0.0
  - 2013年3月 － R 2.15.3
  - 2022年6月 － R 4.2.1

R-八大功能


Free
Data Analysis Function
Graphics/ Visualization
C , C++, Java, .NET, Python
Improved Software
Database, NoSQL
Web: Shiny package
Deep Learning/ IOT
Data

2.3 R語言下載與安裝

R官方網頁

[Home]

Download

CRAN

R Project

About R
Logo
Contributors
What's New?
Reporting

The R Project for Statistical Computing

Getting Started

R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To **download R**, please choose your preferred CRAN mirror.

If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.

下載

統計計算

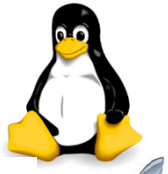
繪圖

R-下載

- 官網: http://www.r-project.org/
- 選取左側 Download \ CRAN
- 選取Taiwan CRAN

Taiwan
https://cran.csie.ntu.edu.tw/

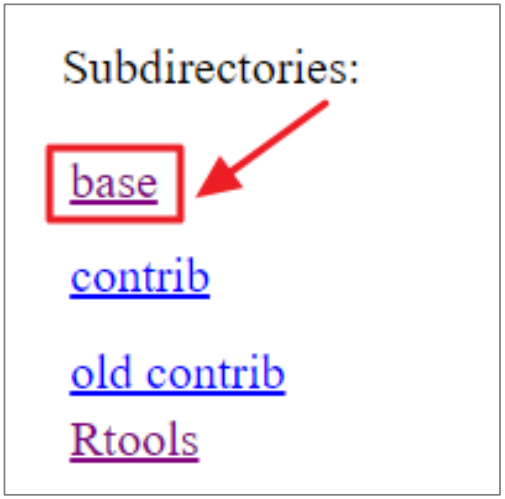
- 選取 Download R for Windows



- Download R for Linux
- Download R for (Mac) OS X
- Download R for Windows

R-下載 (續)

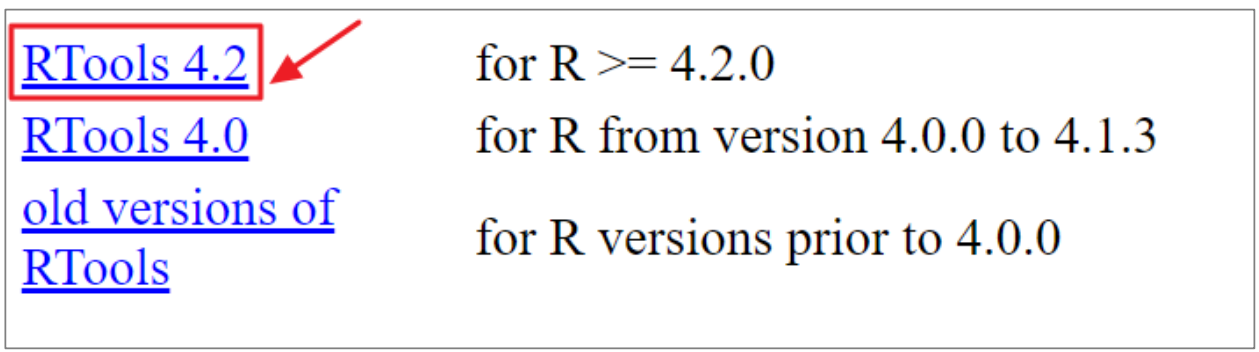
- 選取 base → 下載 [R-4.2.1-win.exe]

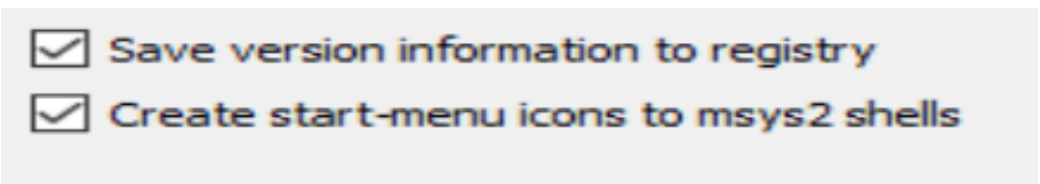


- R安裝路徑: 保留原路徑,不要修改
- 安裝參考說明, 2006
  - https://github.com/rwepa/DataDemo/blob/master/windows_intall_R.pdf

Rtools 下載與安裝

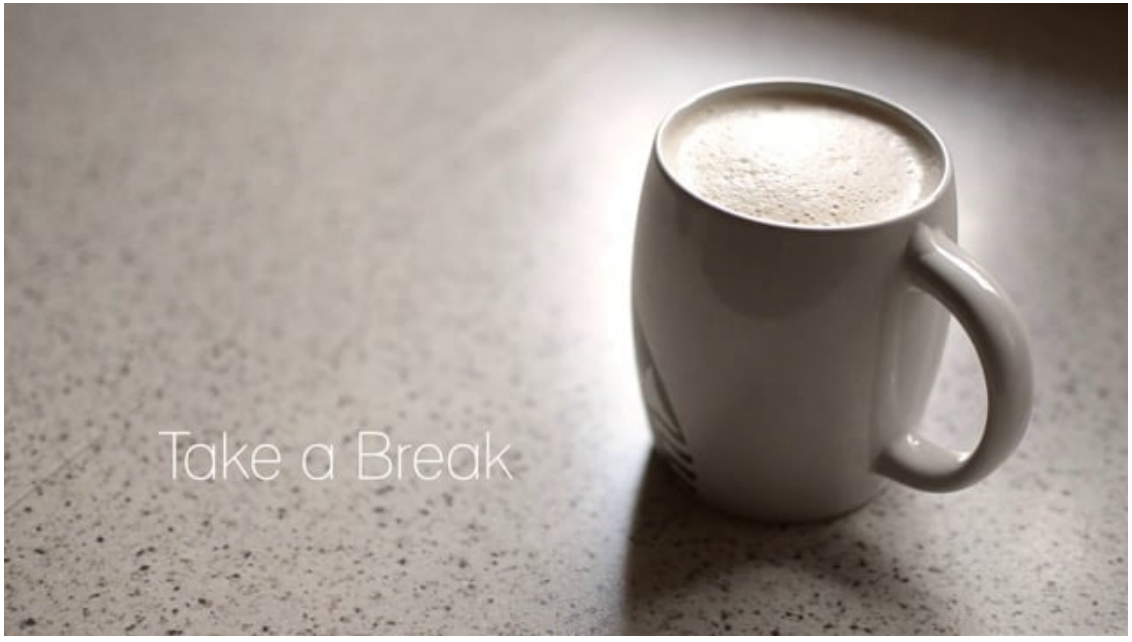
- Rtools for Windows: 一定要保留原路徑 C:\rtools42
- https://cloud.r-project.org/bin/windows/Rtools/rtools42/files/rtools42-5253-5107-signed.exe





2個都打勾

- 安裝R
- 安裝 Rtools



R Manuals (使用手冊)

The R Manuals

edited by the R Development Core Team.

The following manuals for R were created on Debian Linux and may differ from the manuals for Mac or Windows on platform-specific pages, but most parts version of the manuals for each platform are part of the respective R installations. The manuals change with R, hence we provide versions for the most recent version for the patched release version (R-patched) and finally a version for the forthcoming R version that is still in development (R-devel).

Here they can be downloaded as PDF files, EPUB files, or directly browsed as HTML:

Manual	R-release	R-patched
An Introduction to R is based on the former "Notes on R", gives an introduction to the language and how to use R for doing statistical analysis and graphics.	HTML | PDF | EPUB	HTML | PDF | EPUB
R Data Import/Export describes the import and export facilities available either in R itself or via packages which are available from CRAN.	HTML | PDF | EPUB	HTML | PDF | EPUB
R Installation and Administration	HTML | PDF | EPUB	HTML | PDF | EPUB
Writing R Extensions covers how to create your own packages, write R help files, and the foreign language (C, C++, Fortran, ...) interfaces.	HTML | PDF | EPUB	HTML | PDF | EPUB
A draft of **The R language definition** documents the language *per se*. That is, the objects that it works on, and the details of the expression evaluati process, which are useful to know when programming R functions.		
R Internals: a guide to the internal structures of R and coding standar the core team working on R itself.		
The R Reference Index: contains all help files of the R standard and recommended packages in printable form. (9MB, approx. 3500 pages)		

contributed documentation
(貢獻文件, 免費啦)

Translations of manuals into other languages than English are available from the contributed documentation section (only a few translations are available).

R Manuals (續)

Contributed Documentation

English --- Other Languages

Manuals, tutorials, etc. provided by users of R. The R core team does not take any responsibility for contents, but we appreciate the effort very much and encourage everybody to contribute to this list! To submit, follow the submission instructions on the CRAN main page. All material below is available directly from CRAN, you may also want to look at the list of other R documentation available on the Internet.

Note: Please use the directory listing to sort by name, size or date (e.g., to see which documents have been updated latetly).

English Documents

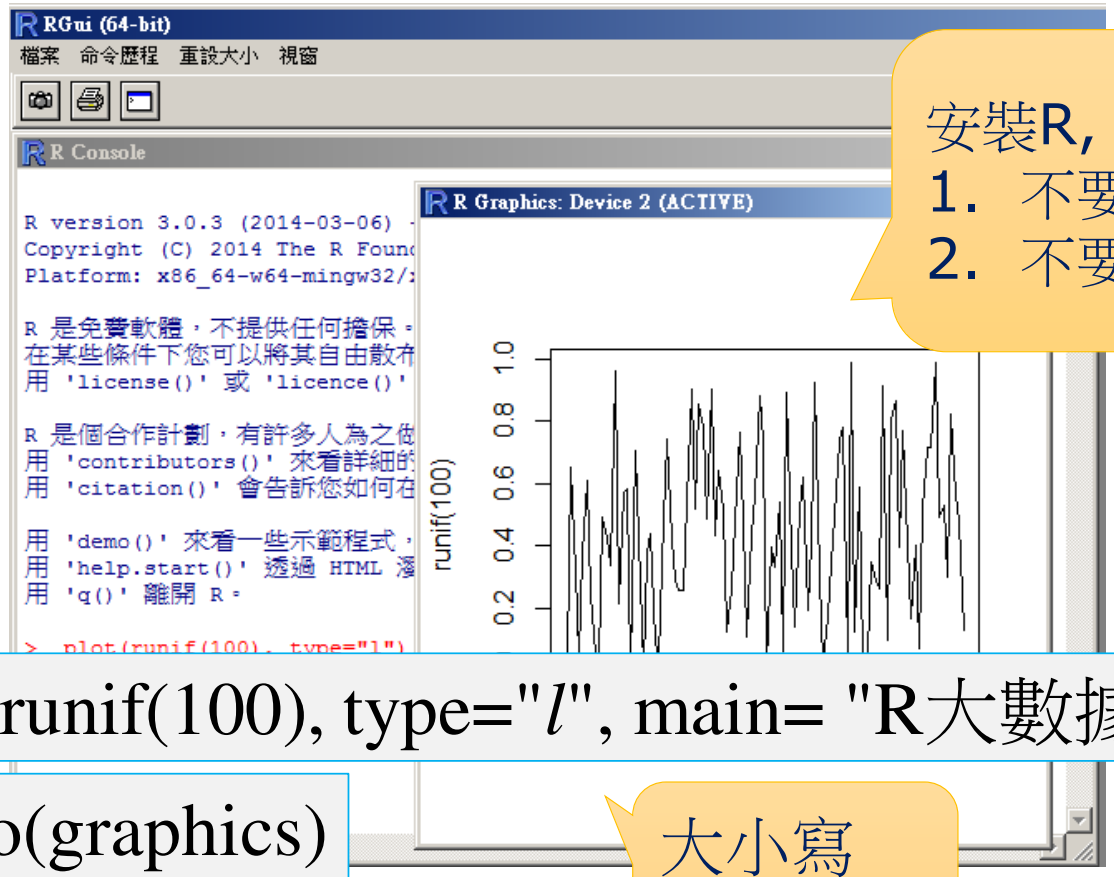
Documents with more than 100 pages:

好書!

- **"Visual Statistics. Use R!"** by Alexey Shipunov (PDF, 2016-06-06, 301 page
  are accessible from Alexey Shipunov's English R page .
- **"Using R for Data Analysis and Graphics - Introduction, Examples and Commentary"** by John Maindonald (PDF, data sets and scripts are available at JM's homepage).
- **"Practical Regression and Anova using R"** by Julian Faraway (PDF, data sets and scripts are available at the book homepage).

實作練習

R 執行畫面



安裝R, 登入名稱:
1. 不要使用空格
2. 不要使用中文字型

plot(runif(100), type="*l*", main= "R大數據分析")

demo(graphics)

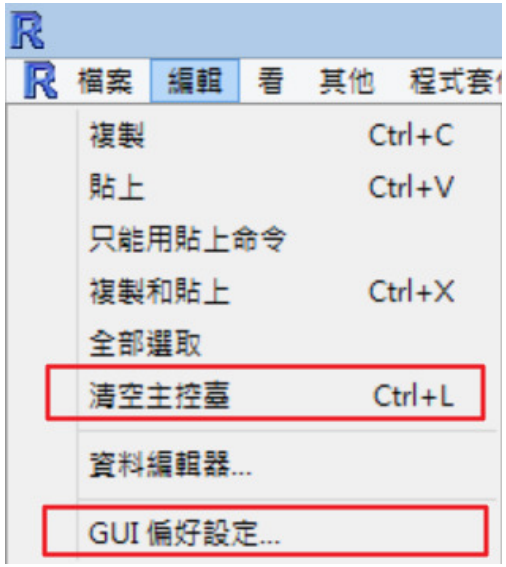
demo(persp)

大小寫
須一致

R 功能表

檔案

編輯



現行目錄 getwd()

儲存控制台-文字檔

輔助

實作練習

新增R檔案練習

- 步驟1: 選取程式碼
- 步驟2: 按 

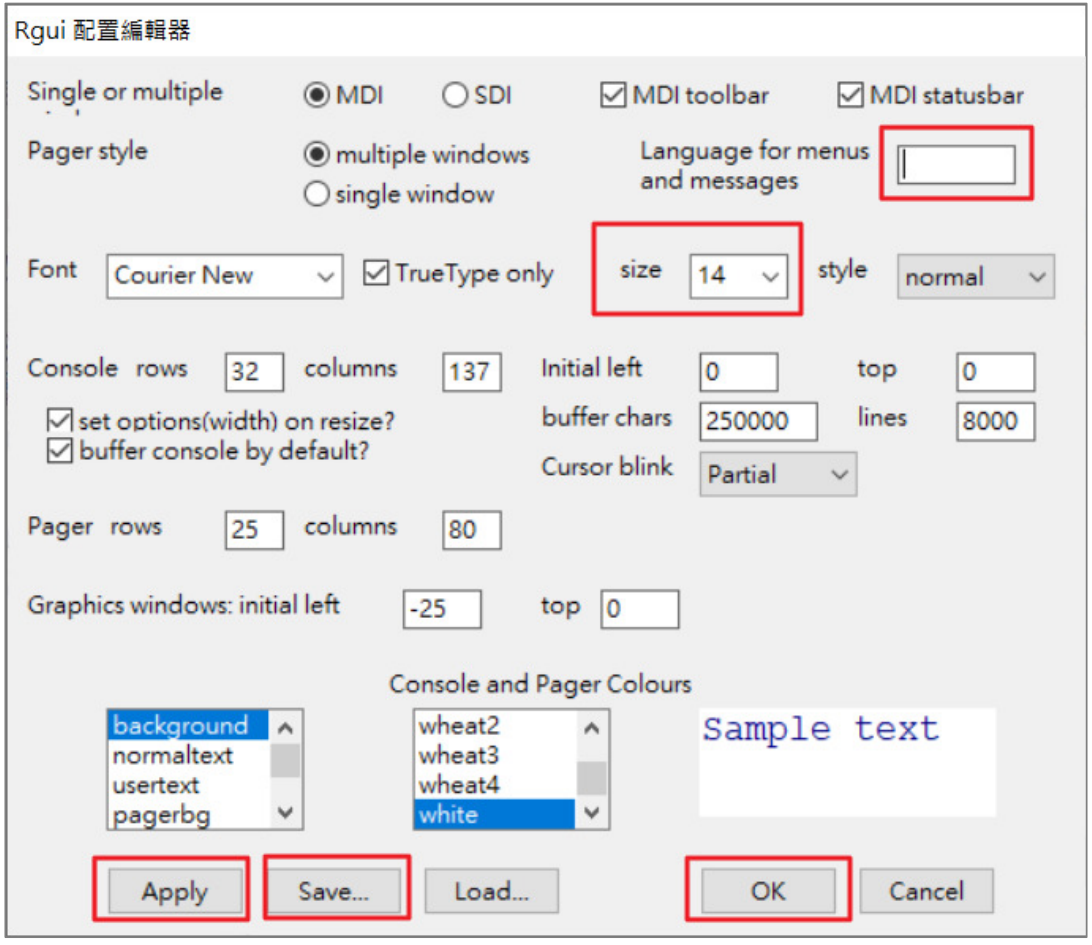
或 Ctrl +R
- 步驟3: 按 File \ Save


```
# 2022.09.04
plot(runif(10), type="b", main="R大數據分析")

x <- rnorm(10)

x
```

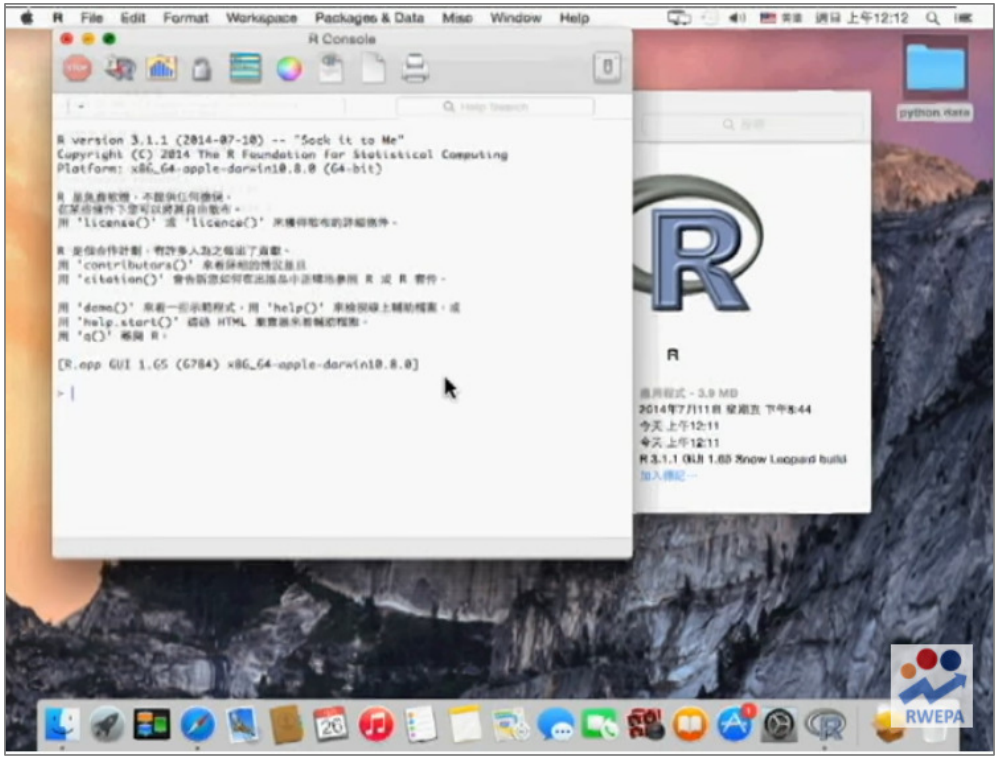
編輯 \ GUI 偏好設定



- Language: en 英文
- size: 字型大小

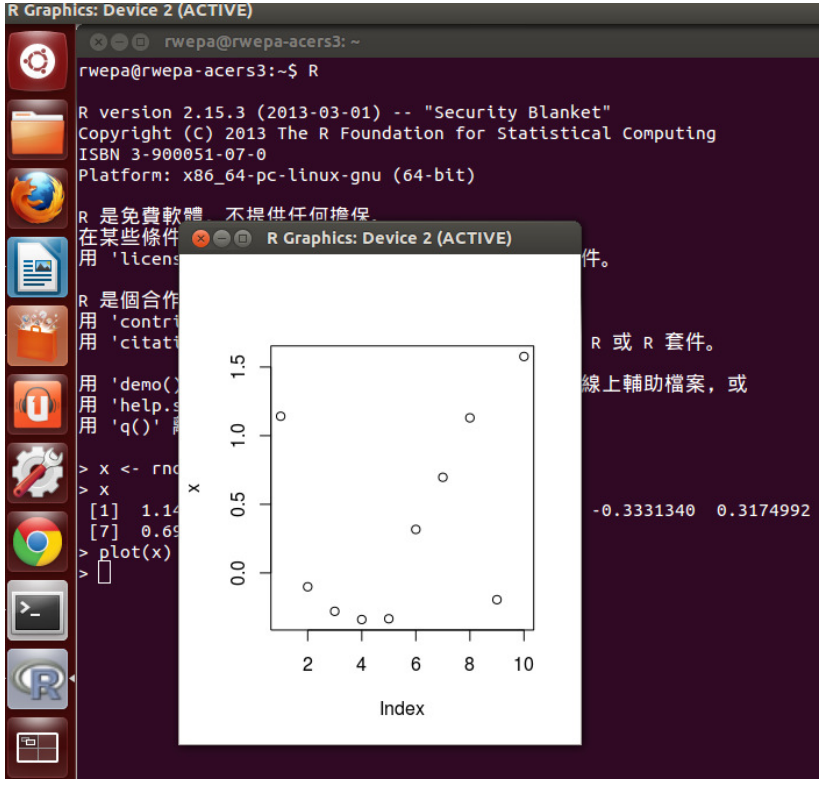
R for Mac

- https://youtu.be/72MYRBNo5Bk



R for Ubuntu

- http://rwepa.blogspot.com/2013/05/ubuntu-r.html



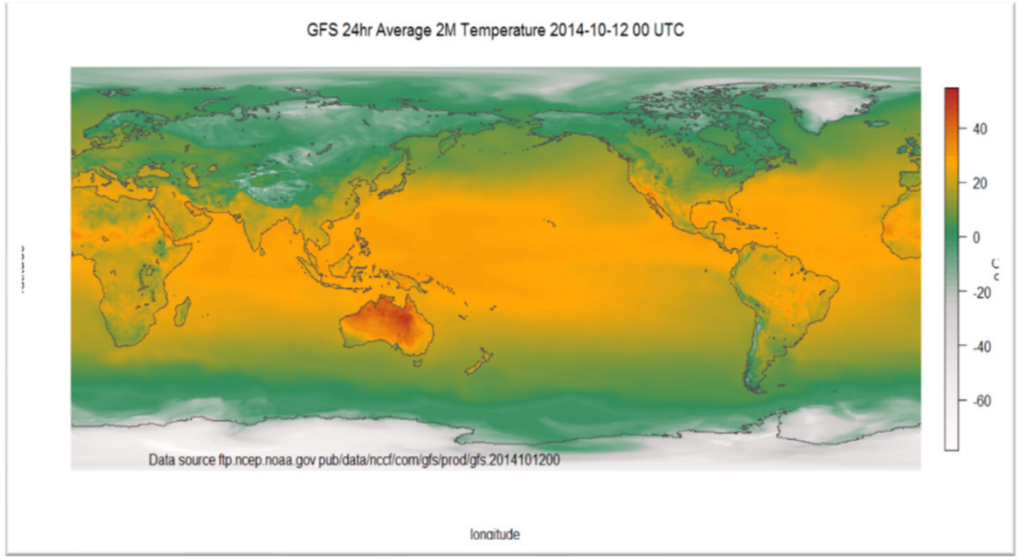
2.4 RStudio簡介

- 更名為 posit

https://posit.co/

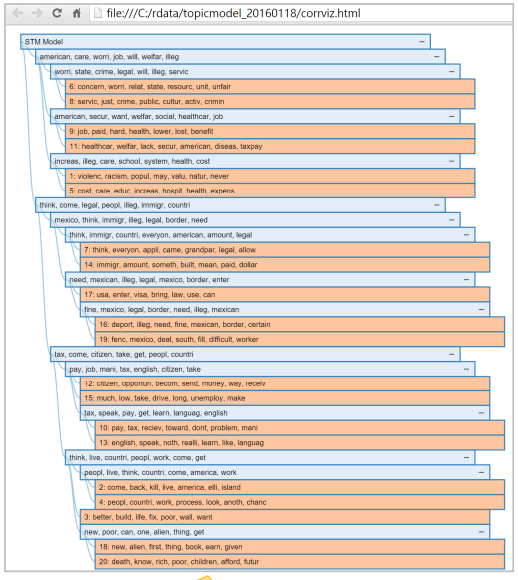
整合式開發環境 - RStudio

- http://www.rstudio.com/



視覺化應用

(全球2M氣溫圖)



主題模型

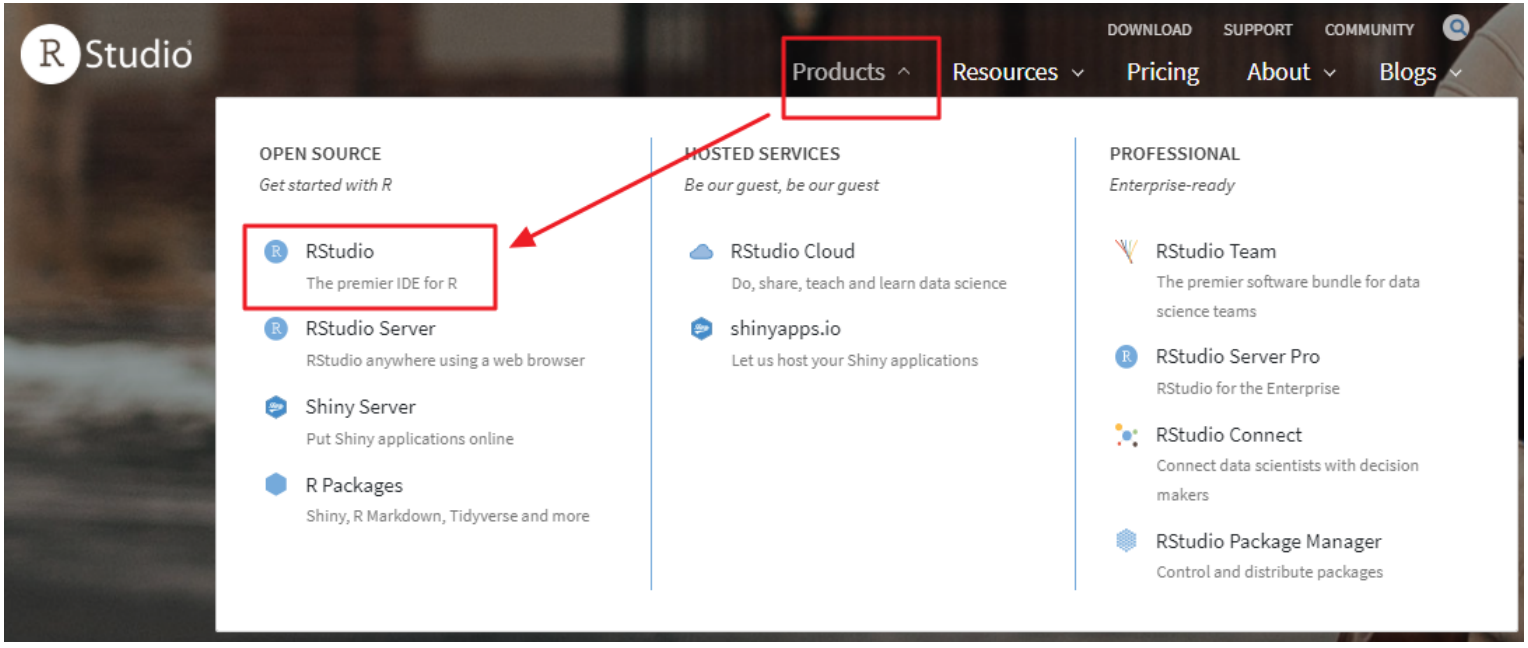
RStudio – 特性

- 支援智慧輸入 (按Tab)
- 高亮度顯示程式碼
- 整合R, Python 程式, 控制台, 變數清單, 繪圖視窗
- 整合資料庫匯入 SQL, Spark
- 整合R套件: shiny, rmarkdown, Quarto, …
- 安裝注意:
  - 先安裝R, 再安裝 RStudio
  - 安裝 RStudio時, 請先關閉R

2.5 RStudio下載與安裝

RStudio 下載

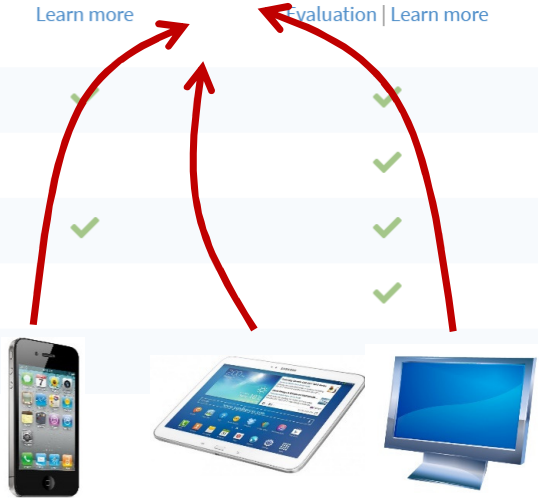
- http://www.rstudio.com/



RStudio 下載 (續)

	RStudio Desktop	RStudio Desktop	RStudio Server	RStudio Server Pro
	Open Source License	Commercial License	Open Source License	Commercial License
	Free	$995	Free	$4,975
	單機版	ear	伺服器版本	/year Named Users)
	DOWNLOAD	BUY	DOWNLOAD	BUY
	Learn more	Learn more	Learn more	valuation | Learn more
Integrated Tools for R	✓	✓	✓	✓
Priority Support		✓		✓
Access via Web Browser			✓	✓
Enterprise Security				✓
Project Sharing				

免費版



RStudio 下載 (續)

RStudio Desktop 2022.07.1+554 - Release Notes

1. Install R. RStudio requires R 3.3.0+.

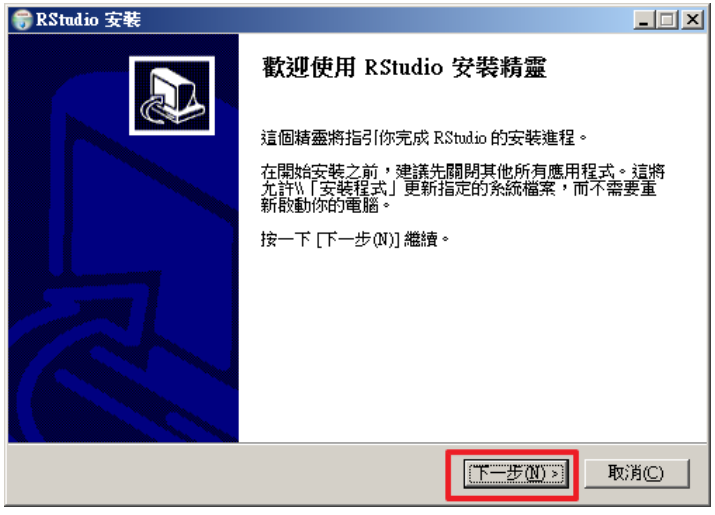
2. Download RStudio Desktop. Recommended for your system:

Requires Windows 10/11 (64-bit)

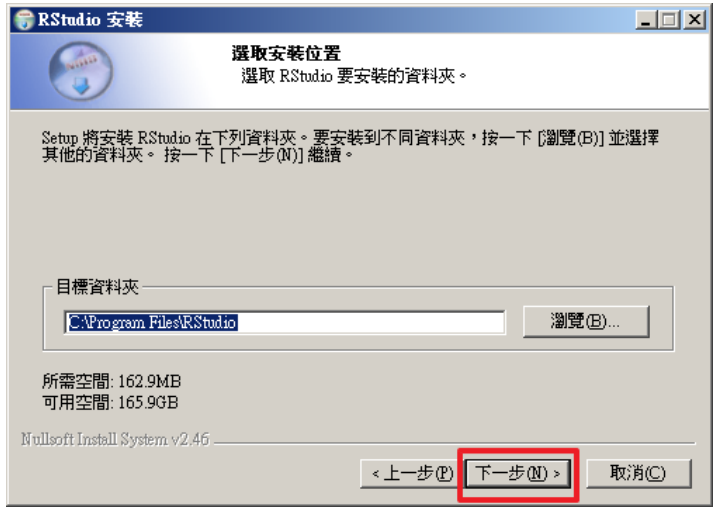
RStudio-2022.07.1-554.exe
(181 MB)

RStudio 安裝

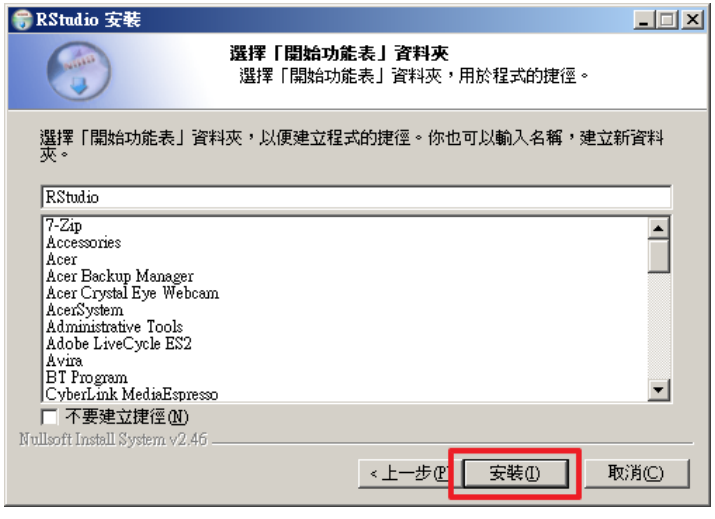
1



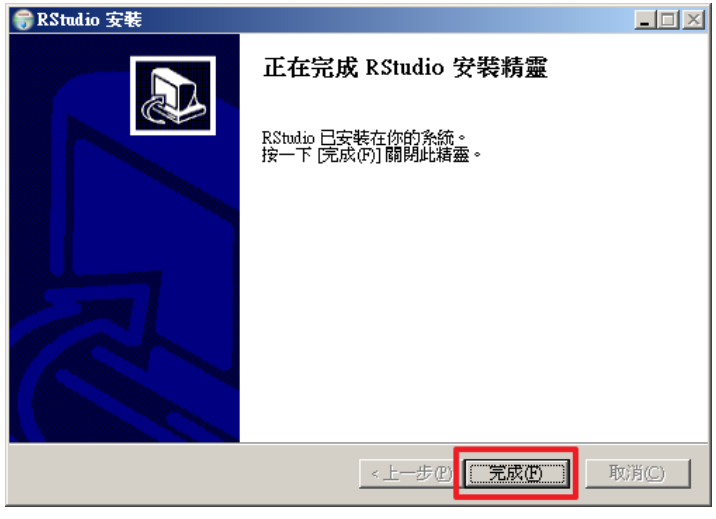
2



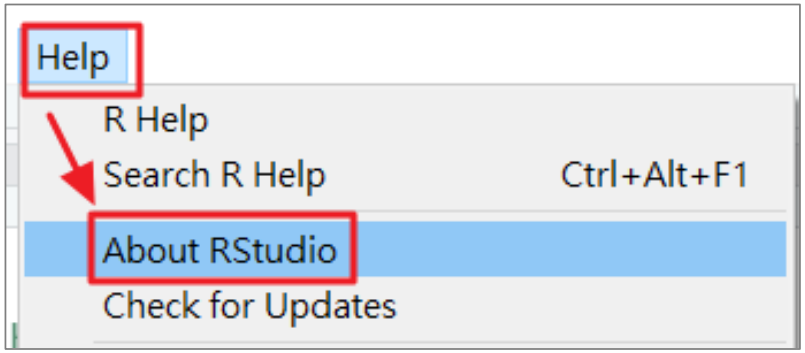
3



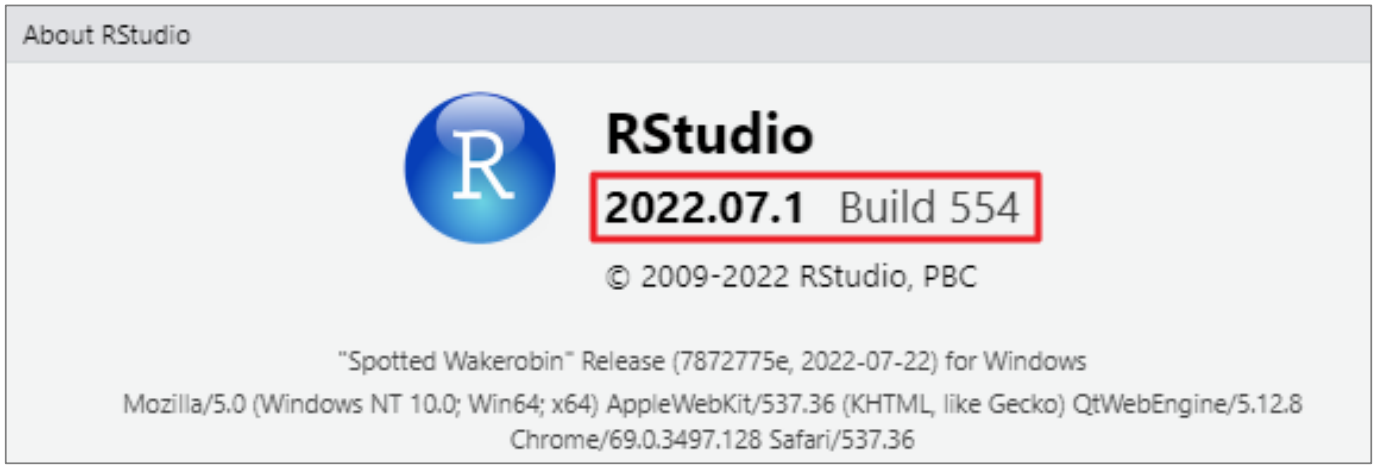
4



RStudio 版本訊息

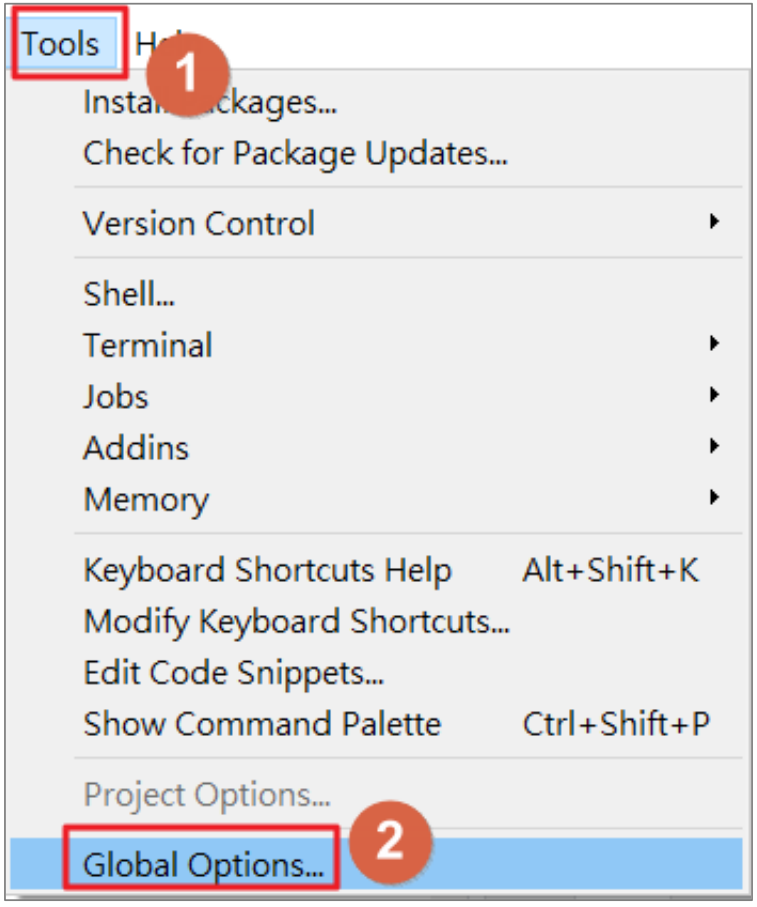


Help \ About RStudio

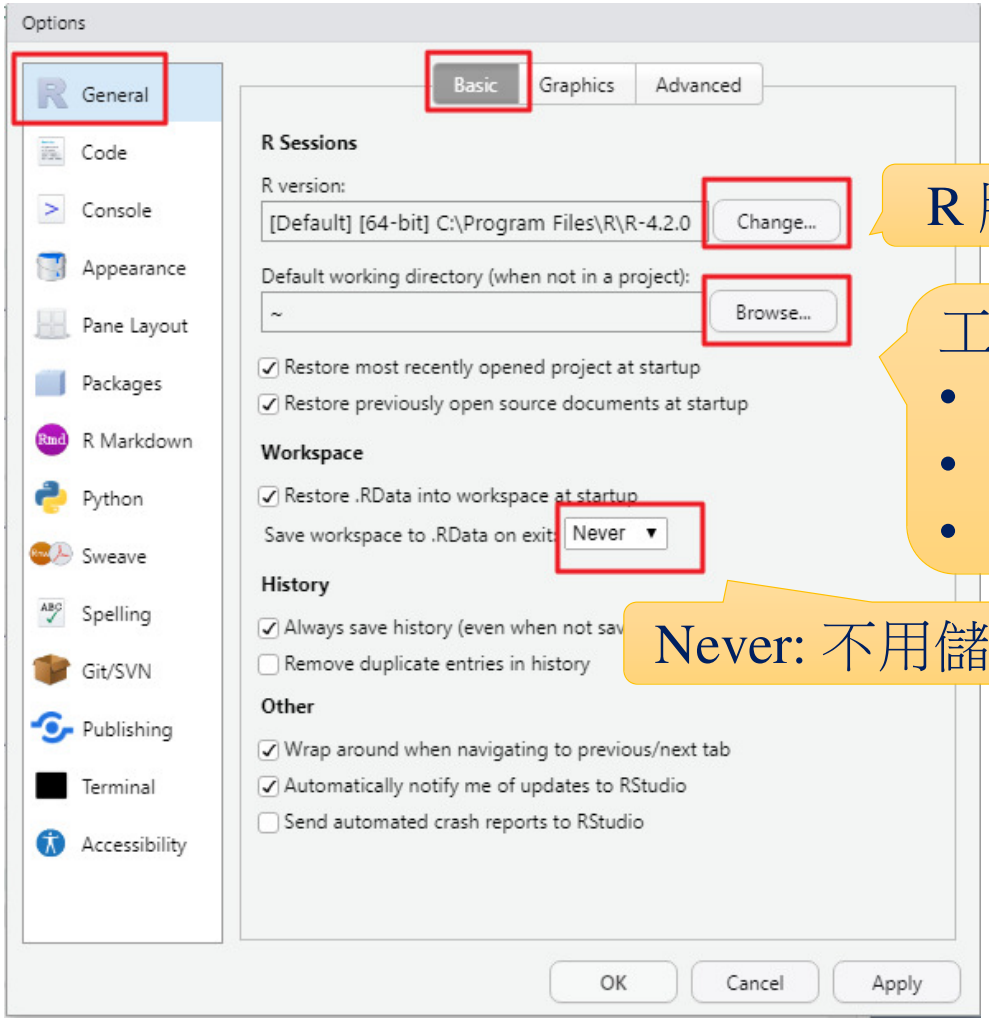


RStudio 選項設定

- Tools \ Global Options



General \ Basic



R 版本

工作目錄 ~

- Windows: 我的文件
- Mac: 使用者名稱
- Ubuntu: 使用者名稱

Never: 不用儲存RData

Windows 10.

```
> getwd()
[1] "C:/Users/asus/OneDrive/文件"
```

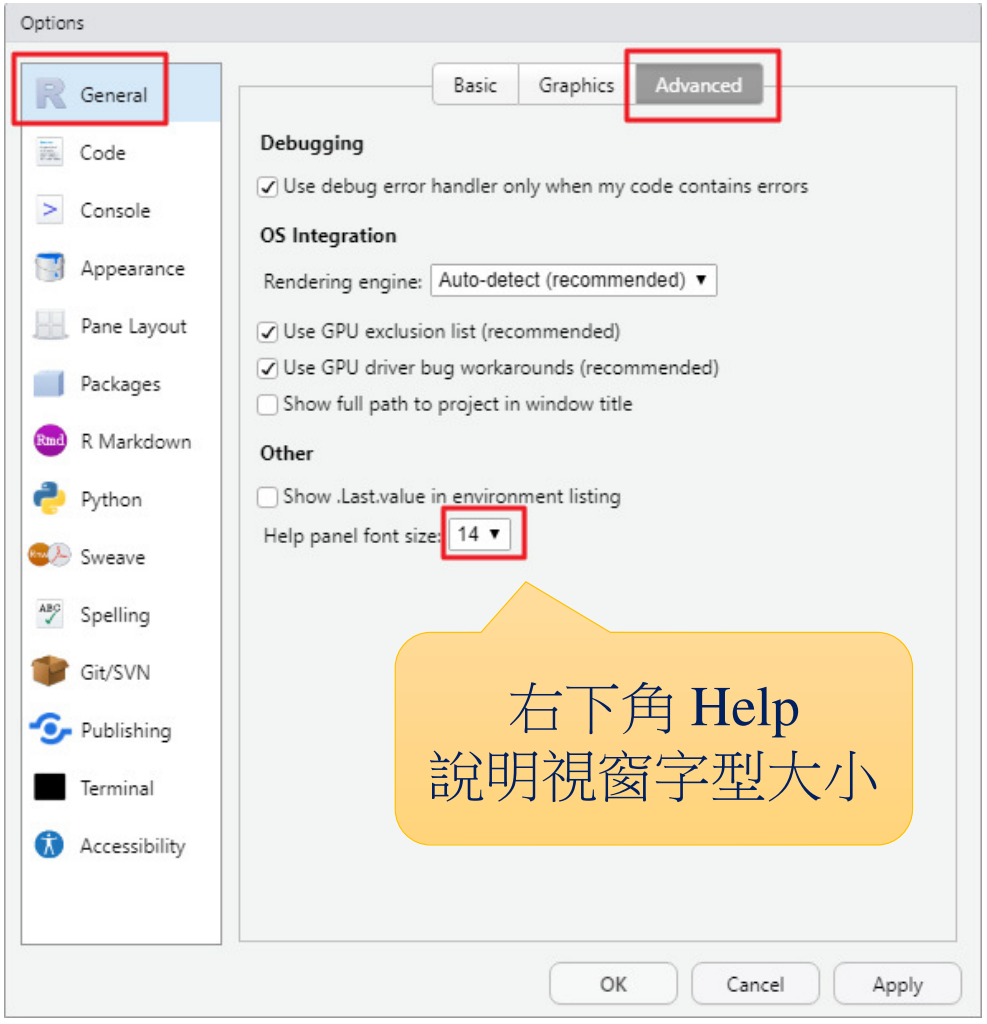
macOS Catalina

```
> getwd()
[1] "/Users/rwepa"
```

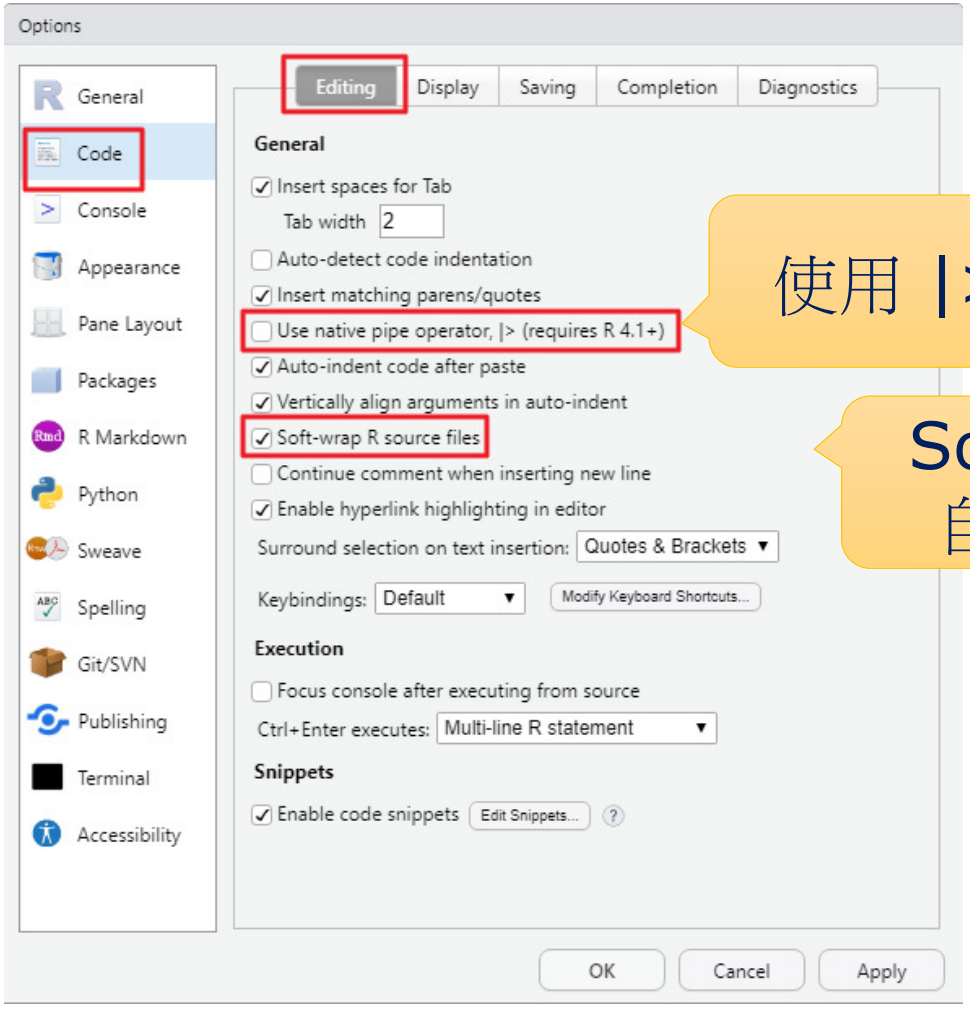
Ububtu 20.04

```
> getwd()
[1] "/home/rwepa"
```

General \ Advanced



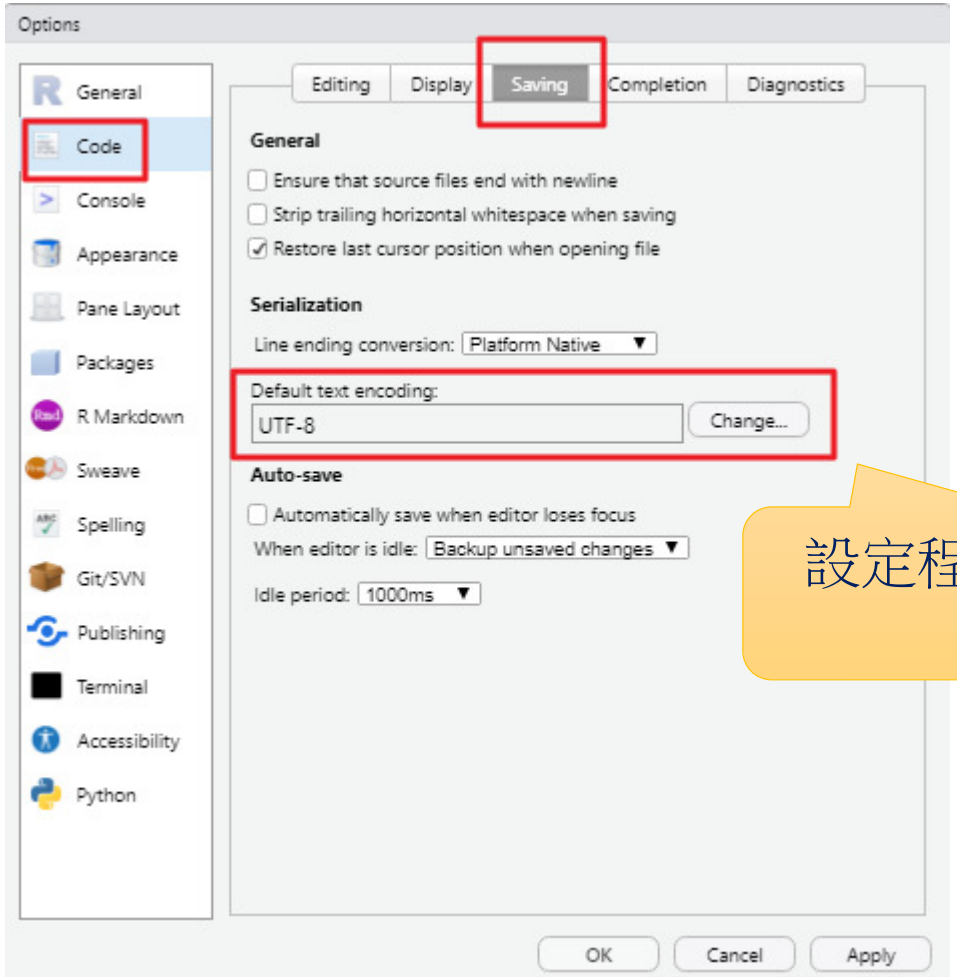
Code \ Editing



使用 |> 管線操作

Soft-wrap
自動換列

Code \ Saving

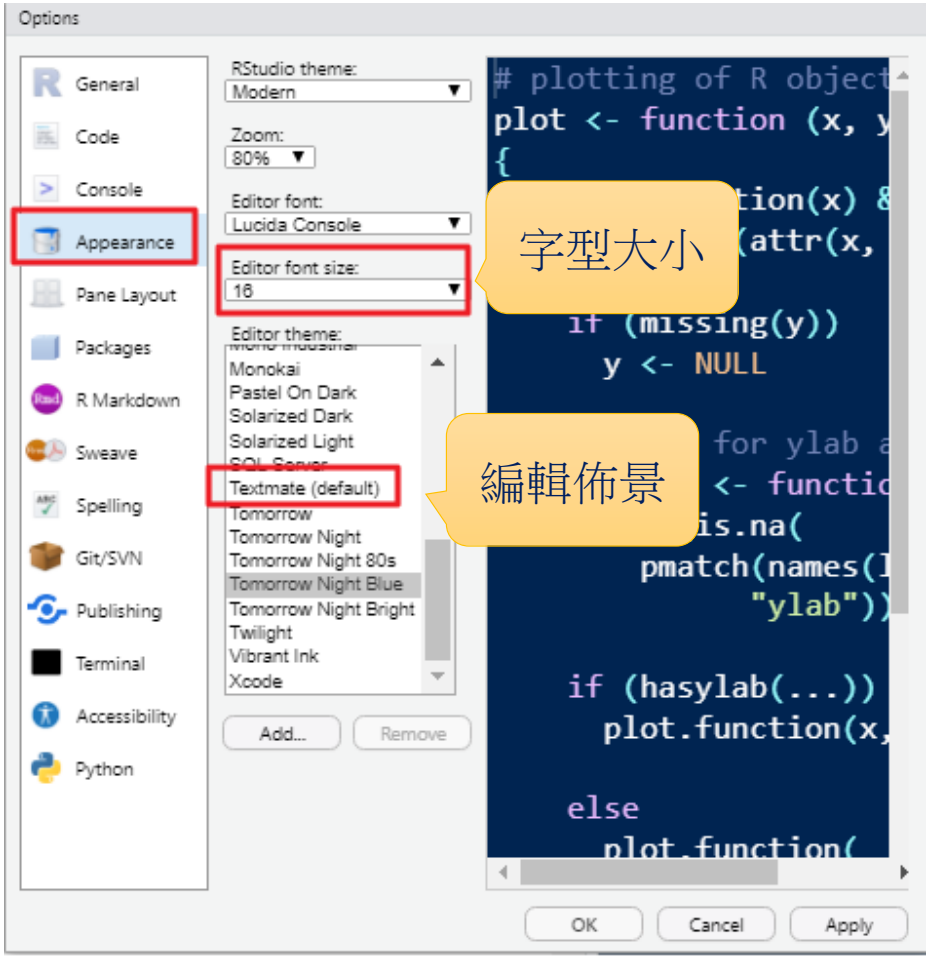


設定程式編碼 預設值
UTF-8

RStudio-選項設定(續)

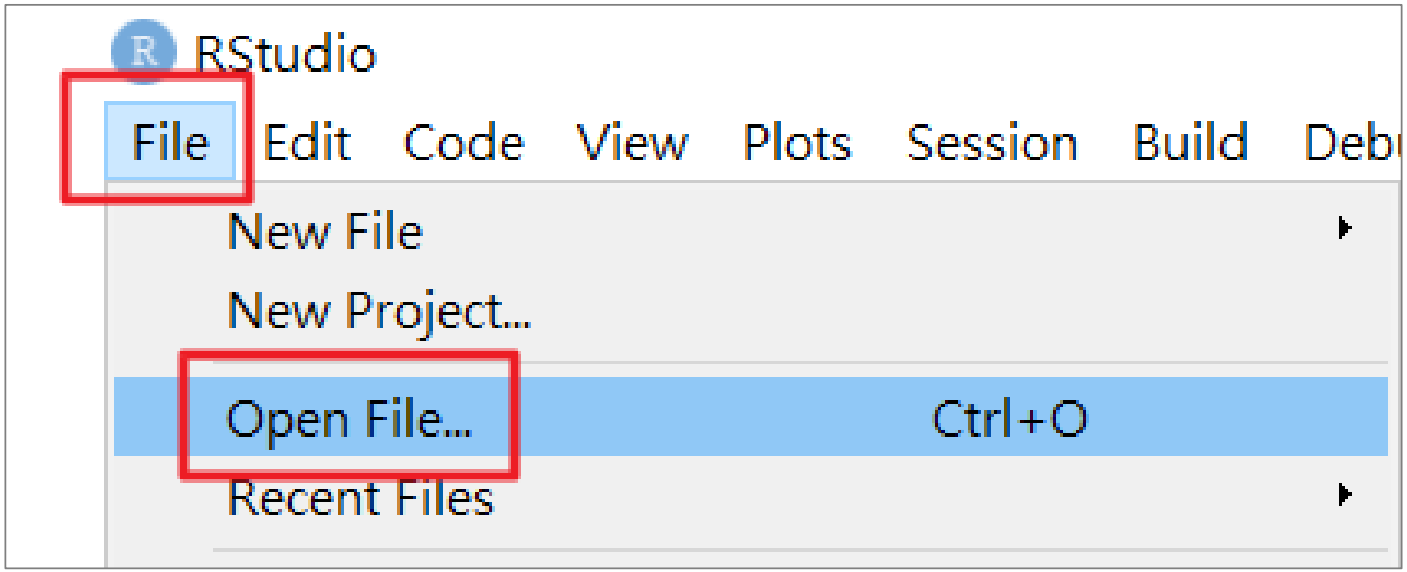
- Appearance \
  Editor theme
- 預設值:
  TextMate

設定完成,須重新啟動RStudio



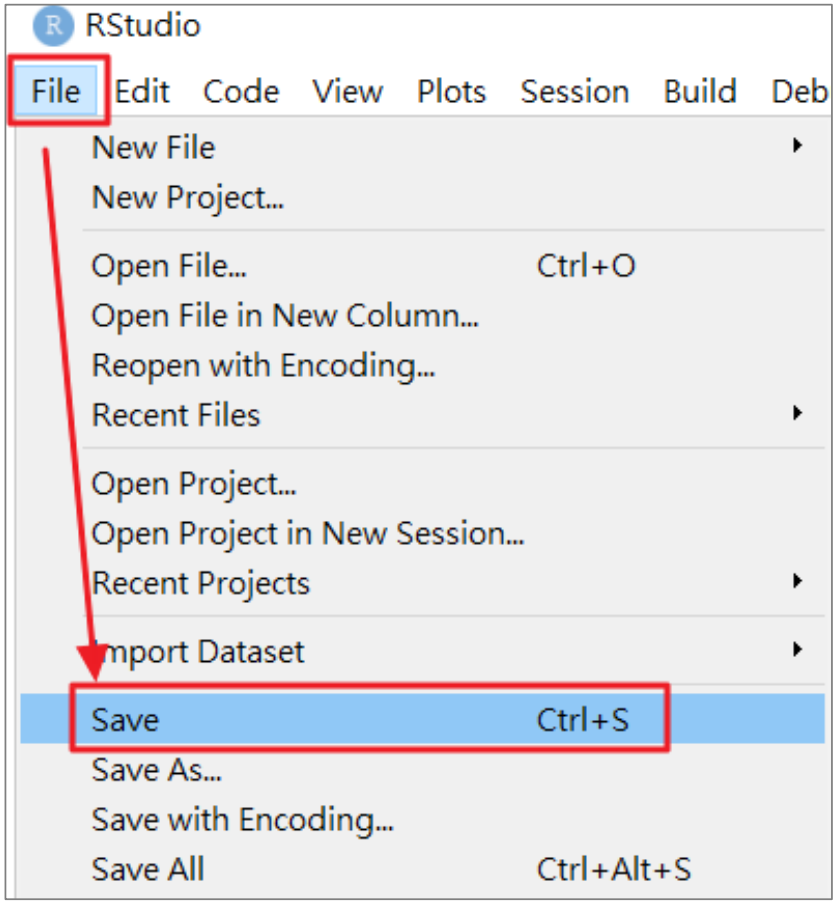
開啟檔案

- File \ Open File

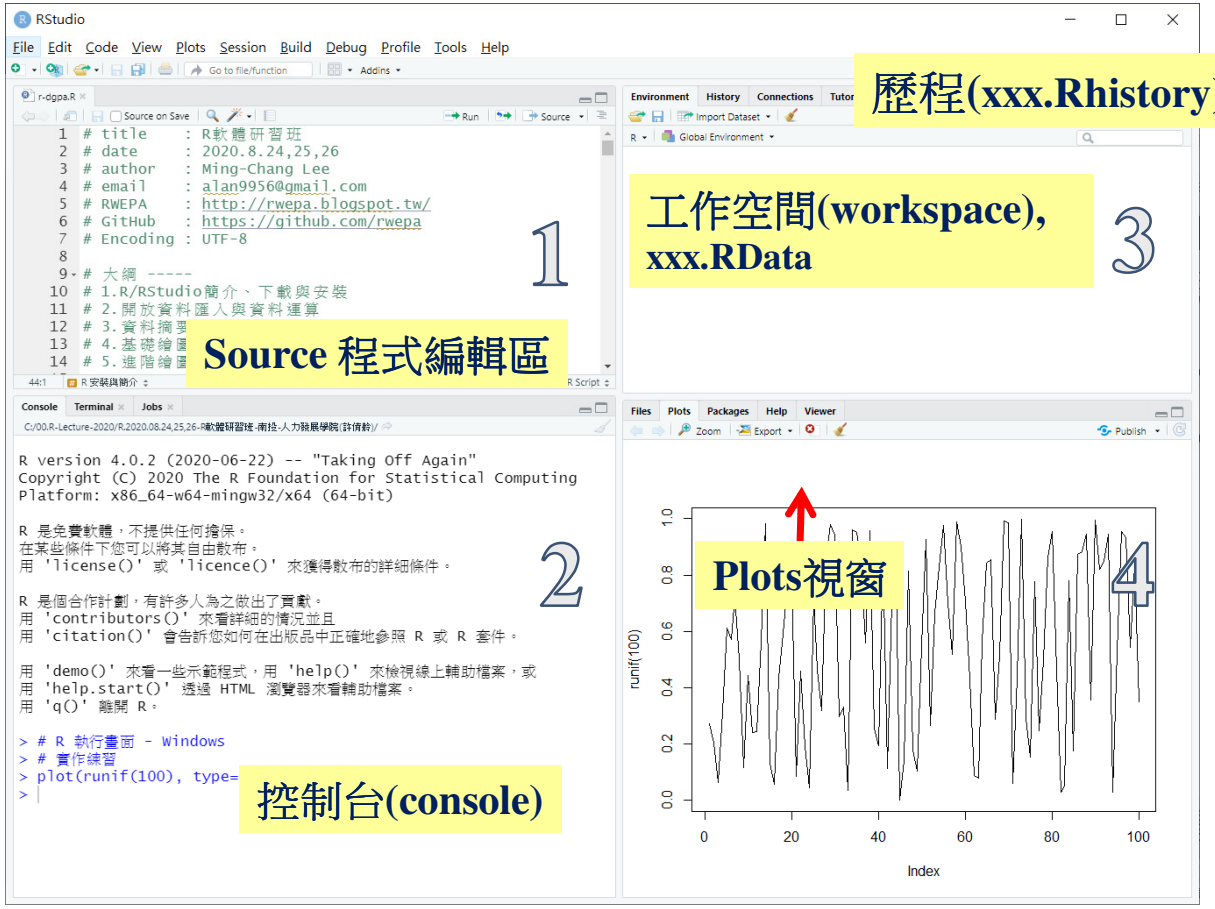


儲存檔案

- File \ Save (CTRL + S)



R/RStudio環境的基礎觀念



Ctrl + Shift + F10: 重新啟動R

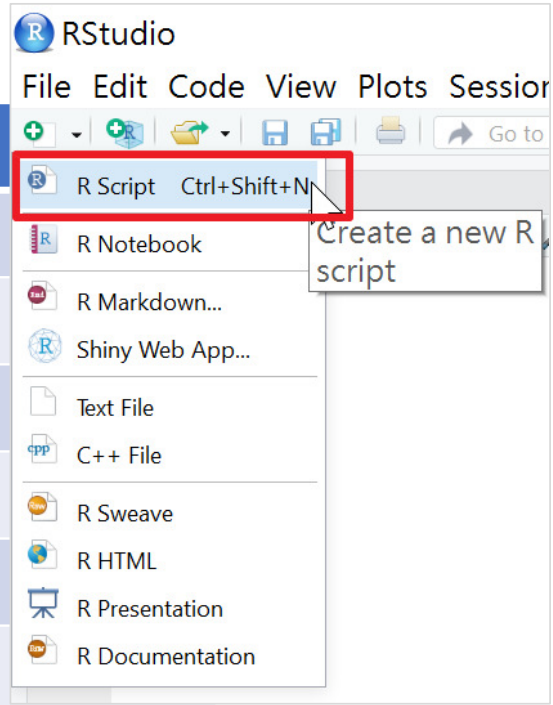
R + Editor

- R – 原生環境
- RStudio – IDE 整合介面
- Eclipse
  - StatET 4.x.0:
    An Eclipse based IDE (integrated development environment) plug-in for R.
  - https://projects.eclipse.org/projects/science.statet
- 如何安裝 Visual Studio R 工具
  -

R Tools for Visual Studio

RStudio 快速鍵

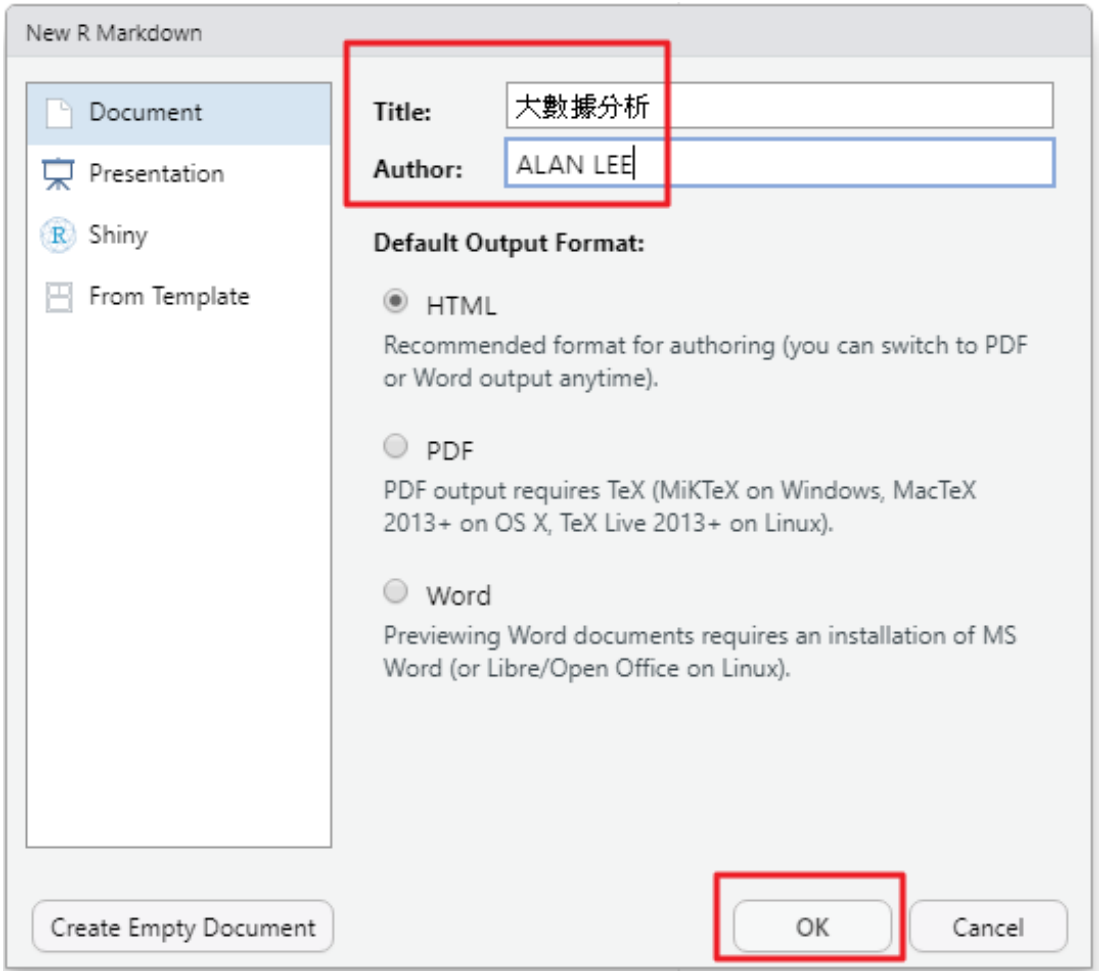
快速鍵	功能
Ctrl + Shift + N	建立新的R程式
Ctrl + S	儲存檔案
Ctrl + Shift + R	建立章節 (-----)
Alt + -	指派符號
Ctrl + Shift + C	註解
Ctrl + Enter	執行程式
Ctrl + Shift + F10	重新啟動R
Alt + Shift + K	快速鍵總表 (Esc 退出)



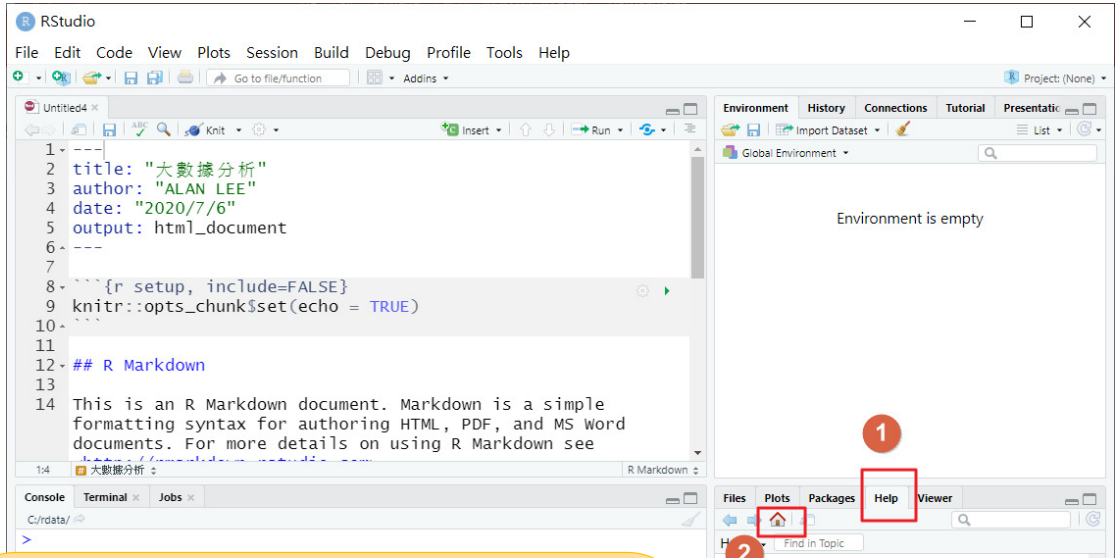
R Markdown

(R標記語言)

RStudio - Markdown



RStudio - Markdown (續)



1. Help
2. 
3. RStudio Cheat Sheets

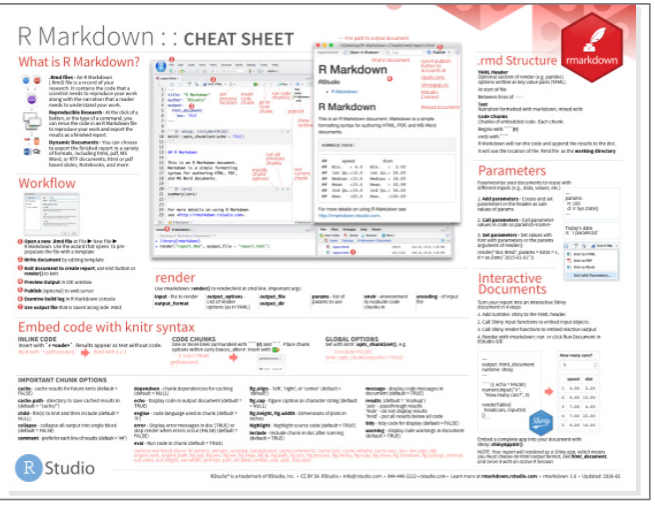
R Markdown Cheatsheet 線上說明

- https://www.rstudio.com/resources/cheatsheets/

R Markdown Cheatsheet

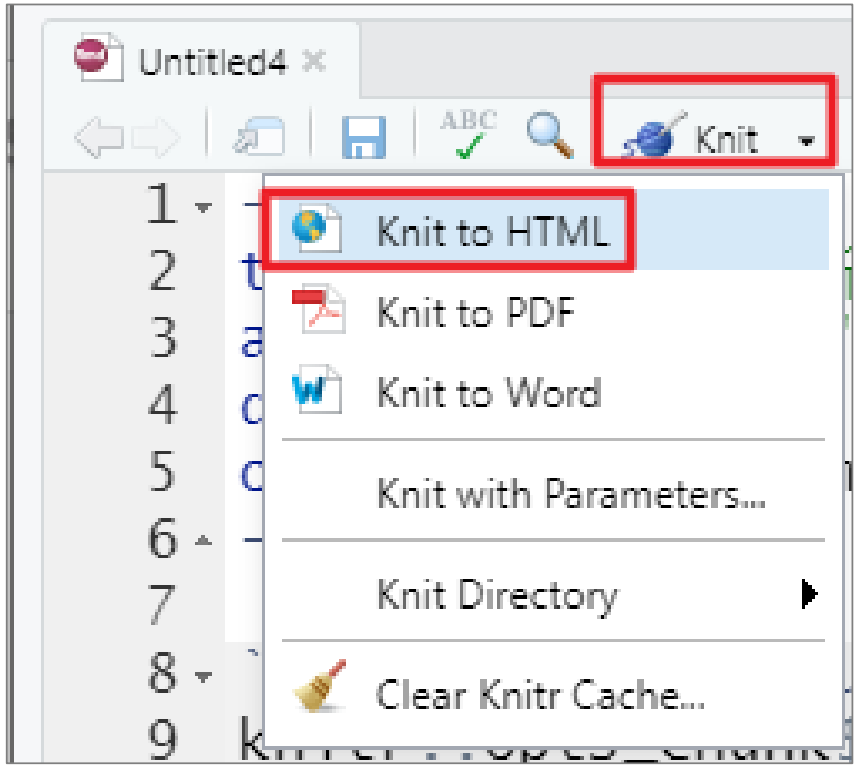
R Markdown is an authoring format that makes it easy to write reusable reports with R. You combine your R code with narration written in markdown (an easy-to-write plain text format) and then export the results as an html, pdf, or Word file. You can even use R Markdown to build interactive documents and slideshows. Updated February 16. (Old Version.

DOWNLOAD

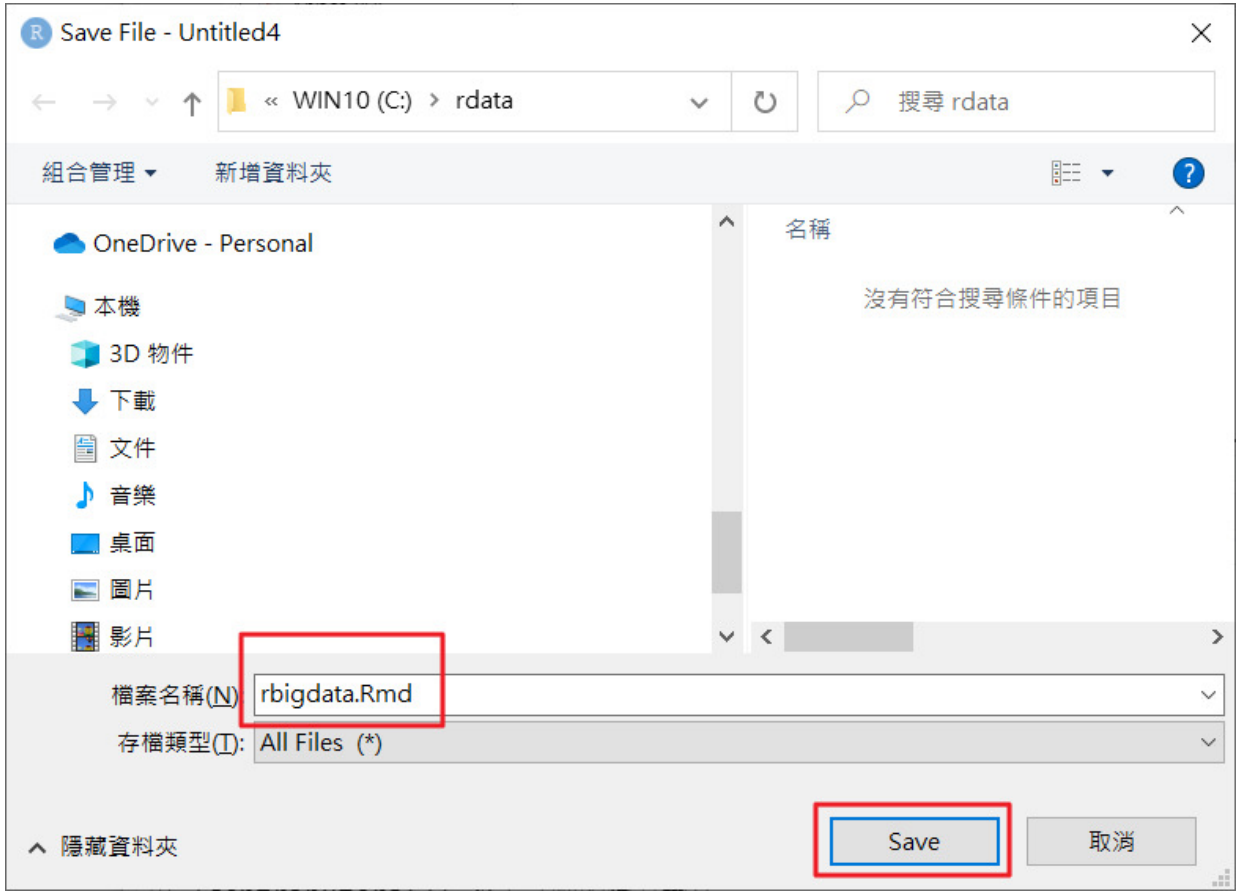


RStudio - Markdown (續)

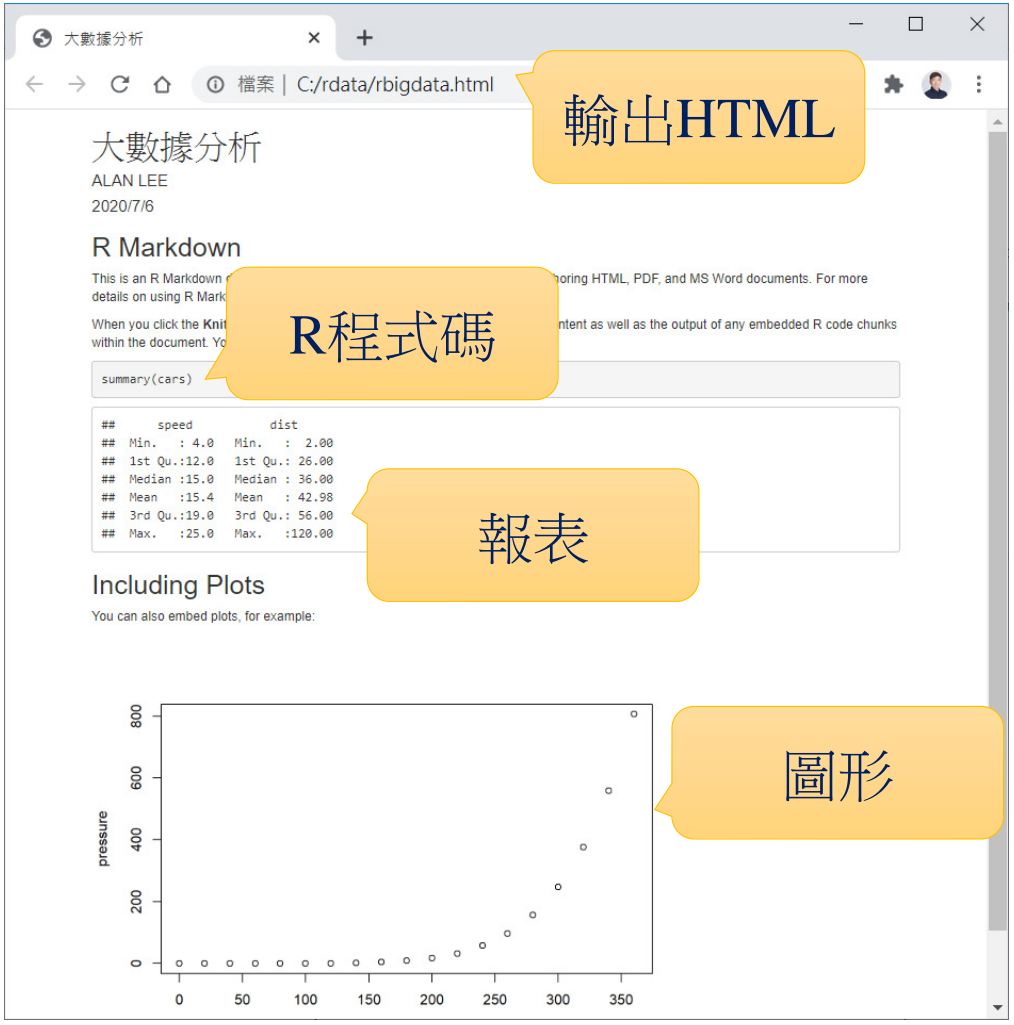
- Knit HTML
- Knit PDF
- Knit Word



RStudio - Markdown (續)

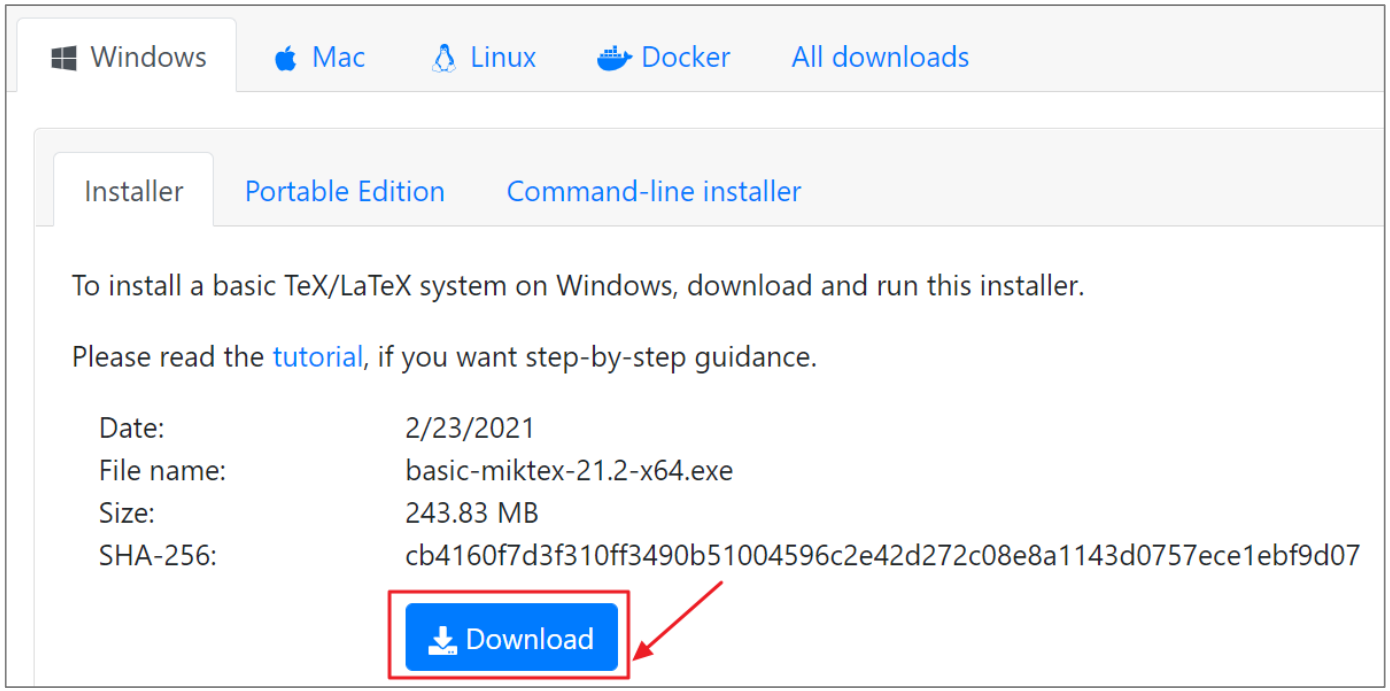


RStudio - Markdown : HTML

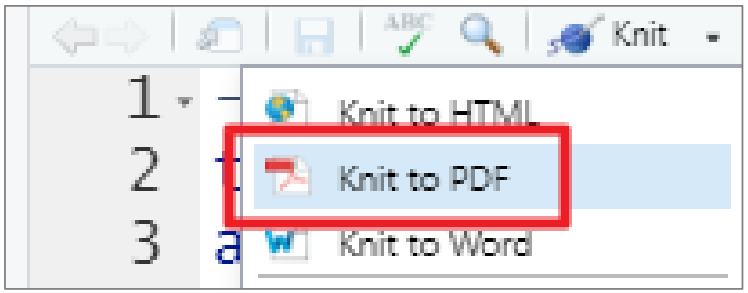


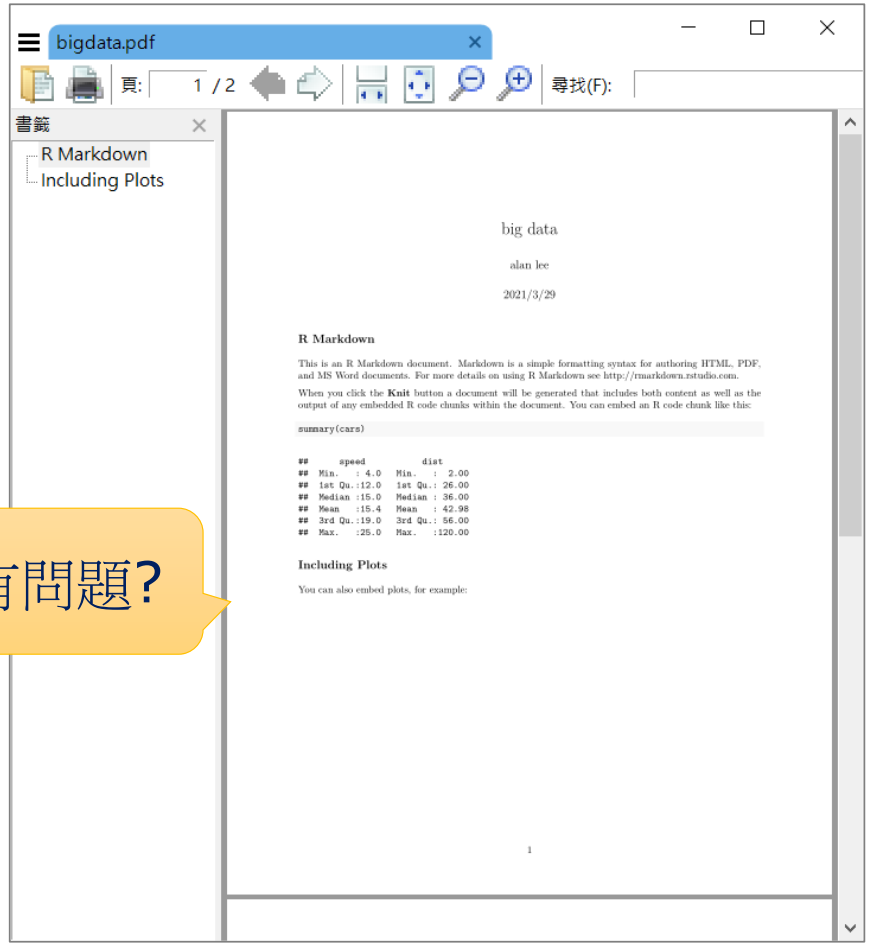
RStudio - Markdown : PDF

- 下載 Miktex: https://miktex.org/download
- basic-miktex-21.2-x64.exe (243.83MB)



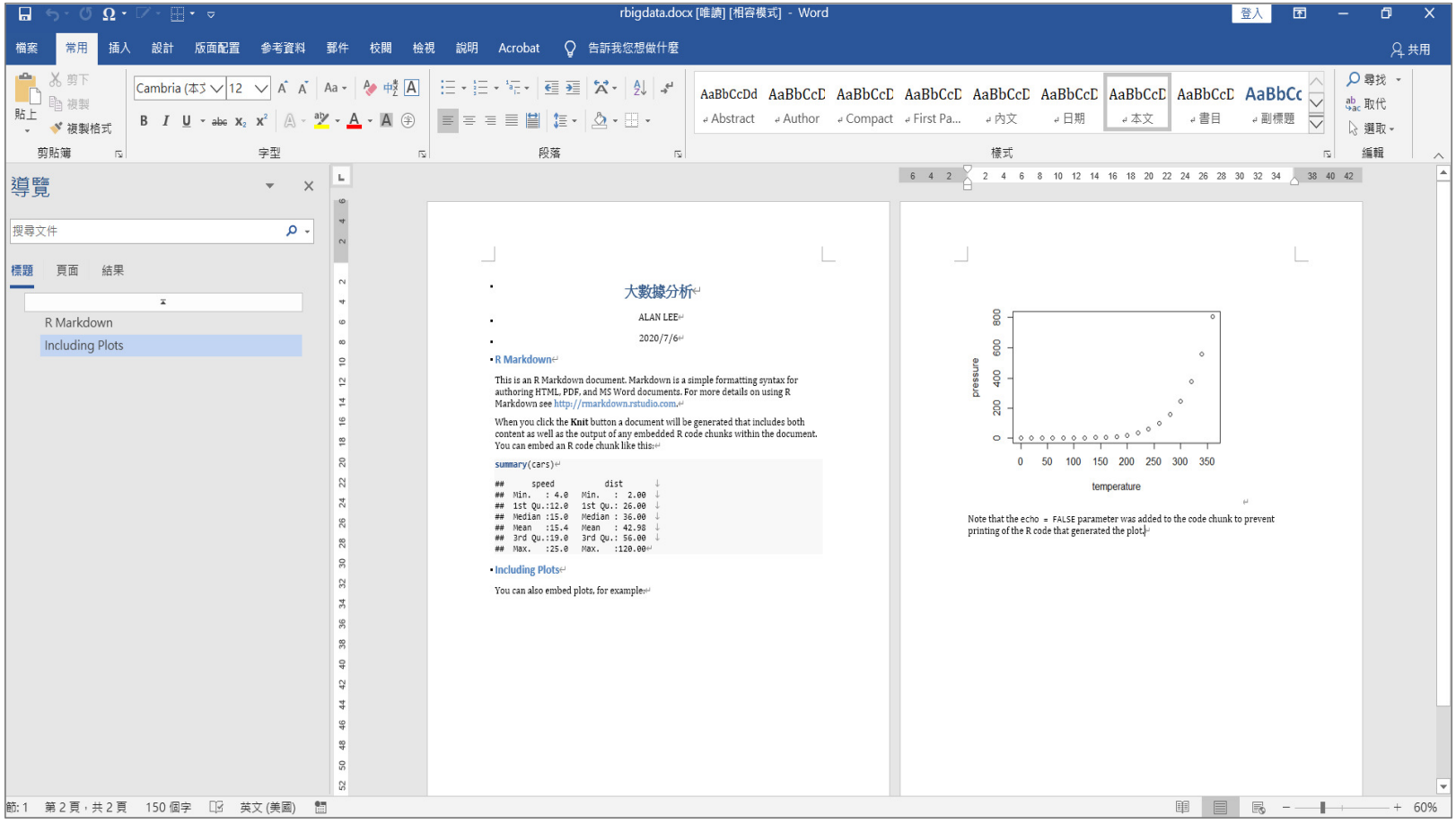
Knit to PDF





中文可能有問題?

RStudio - Markdown 轉換為 Word



參考資料

- RWEPA
  - http://rwepa.blogspot.com/
- R 入門資料分析與視覺化應用
  - https://mastertalks.tw/products/r?ref=MCLEE
- R 商業預測與應用
  - https://mastertalks.tw/products/r-2?ref=MCLEE

謝謝您的聆聽

Q & A

李明昌

EMAIL: *alan9956@gmail.com*

WEB: http://rwepa.blogspot.com/