

Tableau與R語言實務應用

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/>



個人簡介 <http://rwepa.blogspot.com/>

- 姓名：李明昌 (ALAN LEE)
- 現職：中華R軟體學會 常務理事
臺灣資料科學與商業應用協會 常務理事
- 學歷：中原大學 工業與系統工程所 博士
- 經歷：
 - 育達科技大學 資訊管理系(所) 專任助理教授
 - 佛光大學 兼任教師
 - 國立台北商業大學 兼任教師
 - 東吳大學 兼任教師
 - 崇友實業 行銷企劃專員
 - 國航船務代理股份有限公司 海運市場運籌管理員
- 大專院校、資策會、工業技術研究院、國家發展委員會、中央氣象局、公平交易委員會、各縣市政府與日本名古屋產業大學等公民營單位演講達300餘場，2900小時以上。
- 連絡資訊：alan9956@gmail.com



- iPAS 巨量資料分析師 證照推廣
- iPAS 營運智慧分析師 證照推廣



大綱

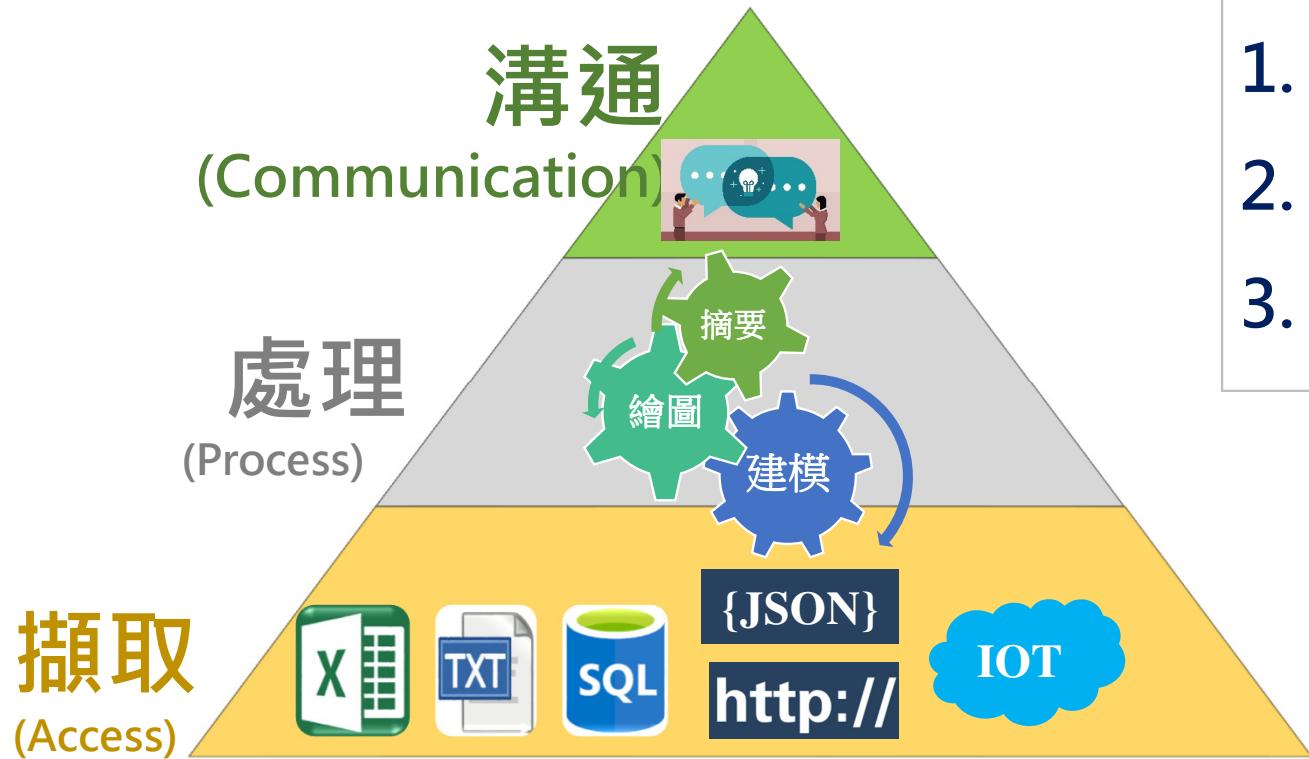
1. 資料分析的心法
2. R, RStudio簡介
3. Tableau與R整合應用

1. 資料分析的心法





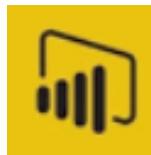
★★★資料分析架構→APC方法



- 1.
- 2.
- 3.



資料視覺化工具

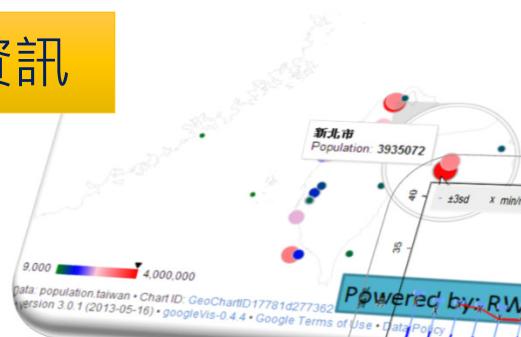
- R - <http://rwepa.blogspot.com/> 【免費】 
- Python - <http://rwepa.blogspot.com/2020/02/pythonprogramminglee.html> 【免費】 
- PowerBI- <https://powerbi.microsoft.com/zh-tw/> 【免費/付費】 
- Tableau - <https://www.tableau.com/> 【14天版免費/付費】 
- Excel 【付費】 



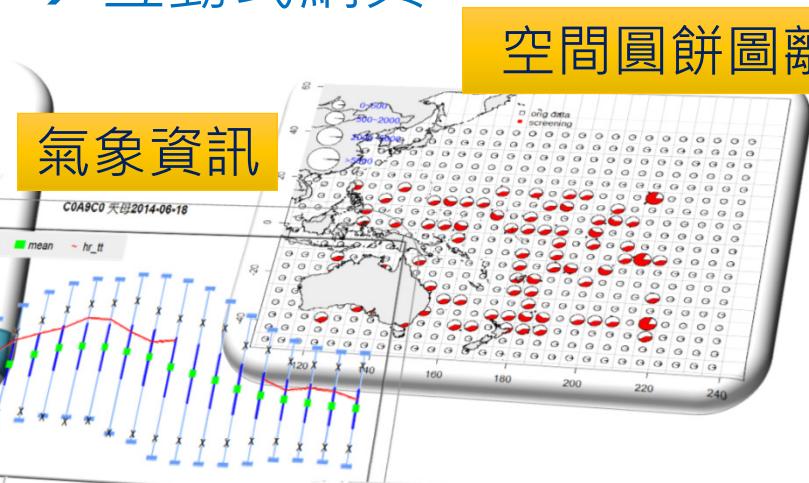
資料分析與視覺化應用

R + shiny → 互動式網頁

地理資訊

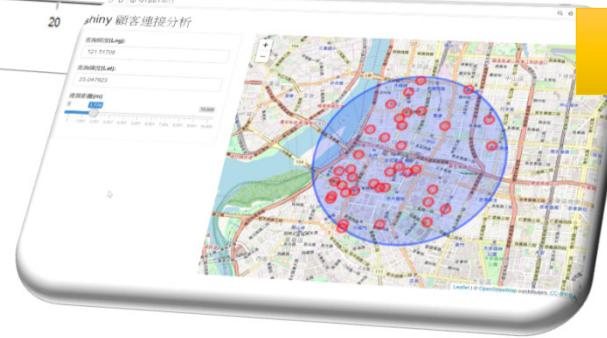


氣象資訊



空間圓餅圖離群值分析

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



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

The screenshot shows a web-based application interface for insurance data analysis. At the top, there are tabs for '檔案上傳', '資料處理', '統計圖表', '模型評估', and '預測模型'. The '預測模型' tab is selected. A red box highlights the '機率模型閥值' (Probability Model Threshold) slider, which is set to 0.1. Another red box highlights the '檢視結果' (View Results) button. A yellow callout points to the '預測結果' table on the right.

The table displays 10 entries of predicted results. The columns include: 性別 (Gender), 女性 (Female), 車輛種類 (Vehicle Type), 私家車 (Private Car), 曝露風險 (Exposure Risk), 曝露風險對數 (Exposure Risk Log), 無索償折扣 (No Claim Discount), 被保險人年齡 (Insured Person Age), 私家車一車齡 0 (Private Car Age 0), 私家車一車齡 1 (Private Car Age 1), 私家車一車齡 2 (Private Car Age 2), 私家車車齡 0_1_2 組合 (Private Car Age 0_1_2 Combination), 車齡 0_1_2 組合 (Age 0_1_2 Combination), 預測機率 (Prediction Probability), and 理賠 (Claim). The last two columns are highlighted with red boxes. Red arrows point from the '檢視結果' button to the table and from the table to the '預測機率' column.

性別	女性	車輛種類	私家車	曝露風險	曝露風險對數	無索償折扣	被保險人年齡	私家車一車齡 0	私家車一車齡 1	私家車一車齡 2	私家車車齡 0_1_2 組合	車齡 0_1_2 組合	預測機率	理賠		
M	0	A	1	0.9144422	-0.08944106	50	4	1	0	0	1	0	2	0.1069	有	
M	0	A	1	0.8158795	-0.20348856	20	4	0	0	1	1	2	2	0.1441	有	
3	M	0	A	1	0.8377823	-0.17699695	50	3	0	0	1	1	2	2	0.1866	有
4	M	0	A	1	0.4325804	-0.83798702	50	6	0	1	0	1	1	2	0.0944	無
5	M	0	A	1	0.7173169	-0.33223755	50	4	0	0	1	1	2	2	0.1218	有
6	M	0	A	1	0.8377823	-0.17699695	50	4	0	0	1	1	2	2	0.1495	有
7	M	0	A	1	0.8487337	-0.16400975	50	5	0	0	1	1	2	2	0.1422	有
8	F	1	A	1	0.8268309	-0.19015503	10	3	0	0	1	1	2	2	0.1733	有
9	M	0	A	1	0.7145791	-0.33606164	0	5	1	0	0	1	0	2	0.0694	無
10	M	0	A	1	0.3340178	-1.09656101	0	3	0	0	1	1	2	2	0.0783	無

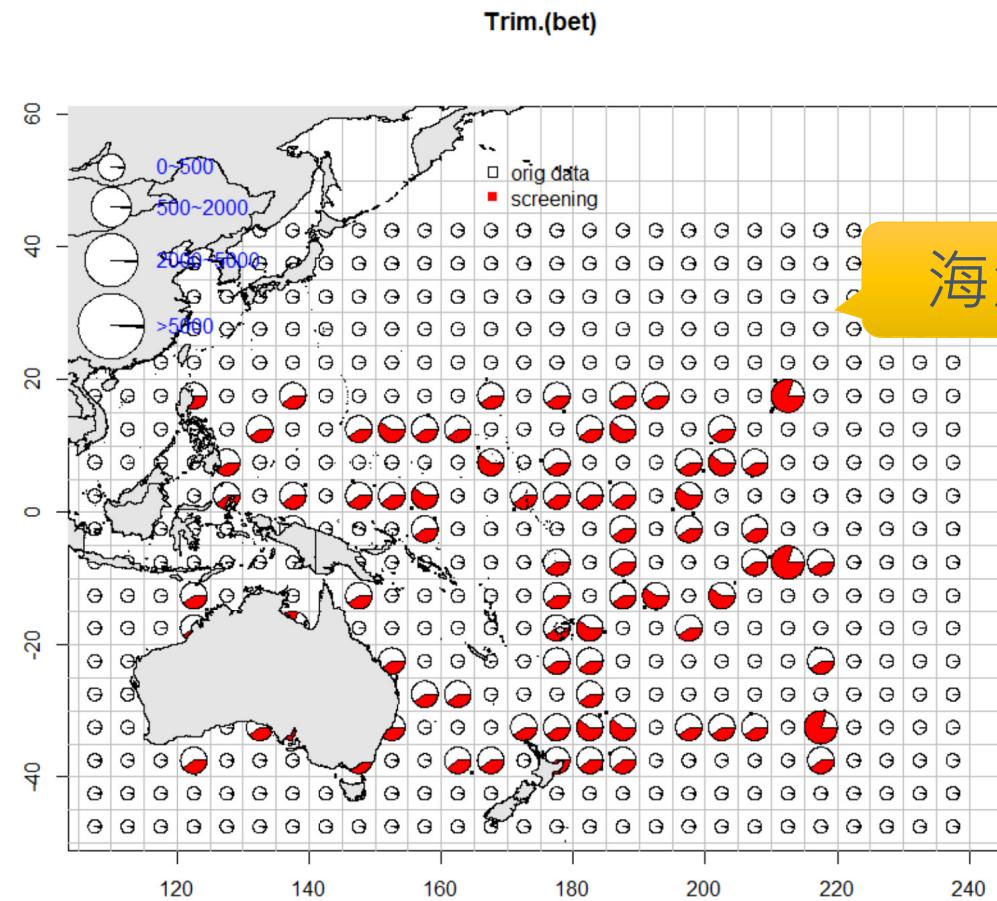
Showing 1 to 10 of 12 entries

Previous 1 2 Next

127.0.0.1:6177/#tab-9487-2

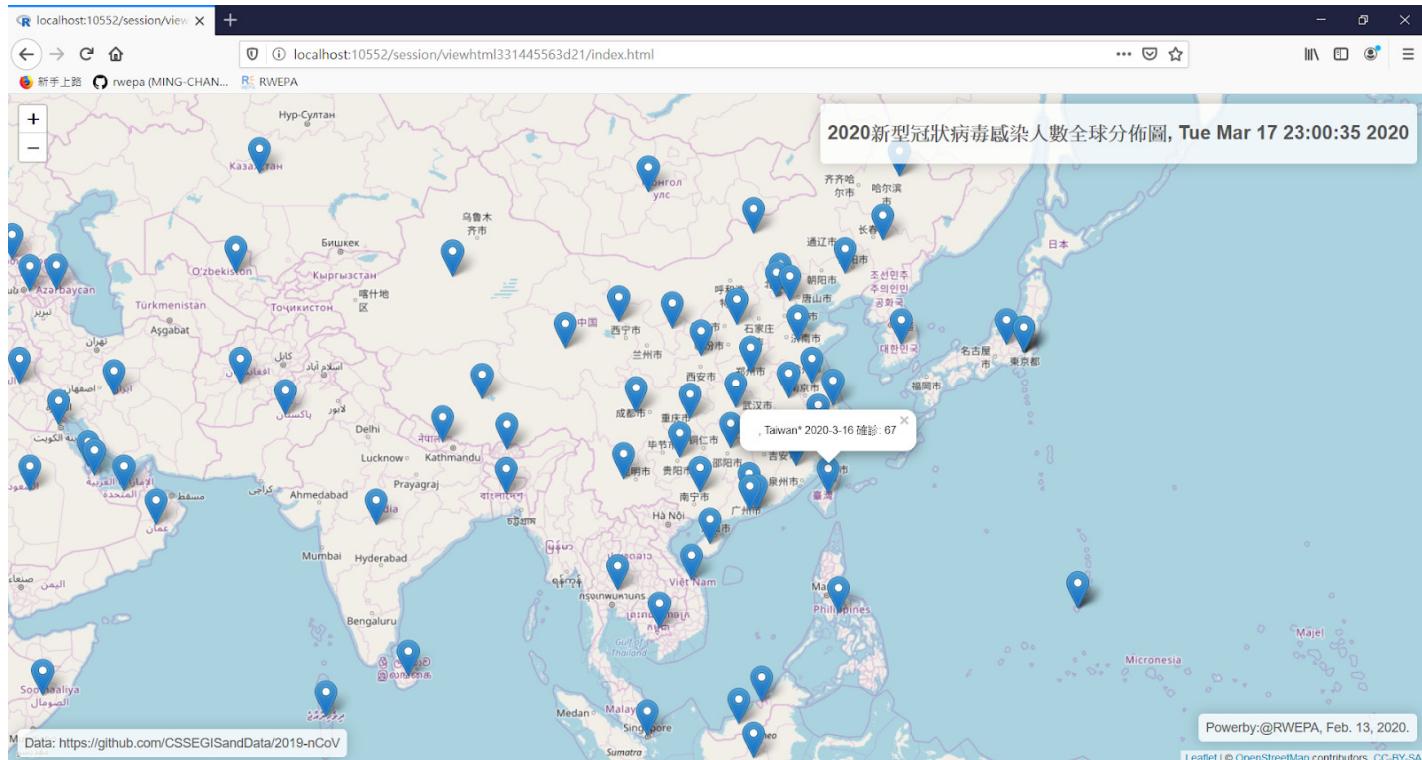


空間圓餅圖離群值分析





2020新型冠狀病毒視覺化

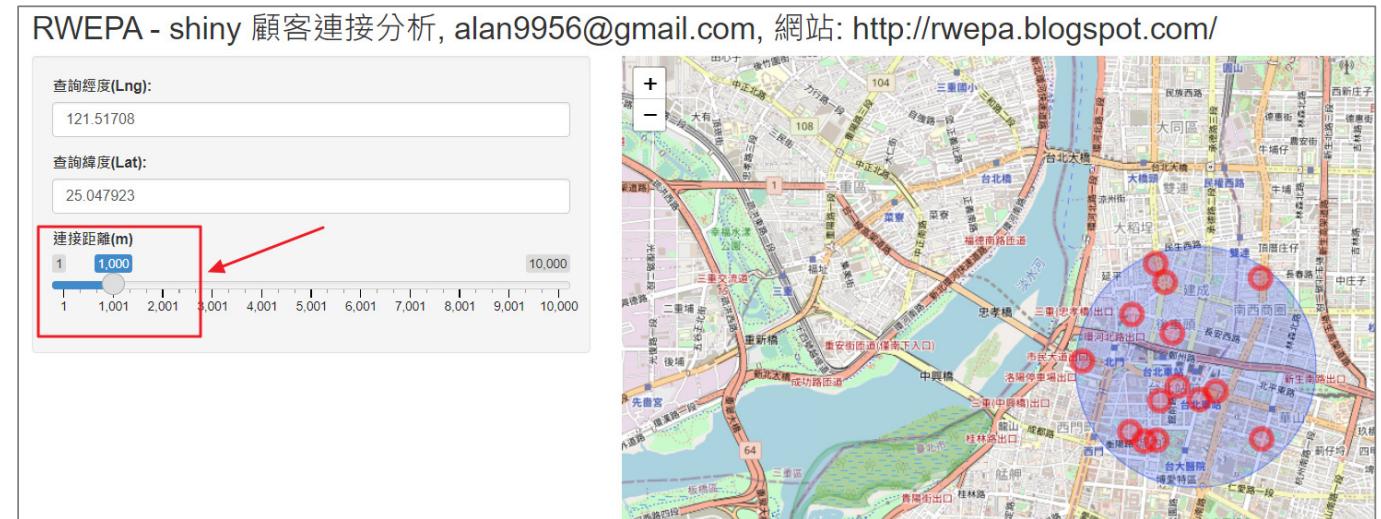


<http://rwepa.blogspot.com/2020/02/2019nCoV.html>



shiny 顧客連接分析

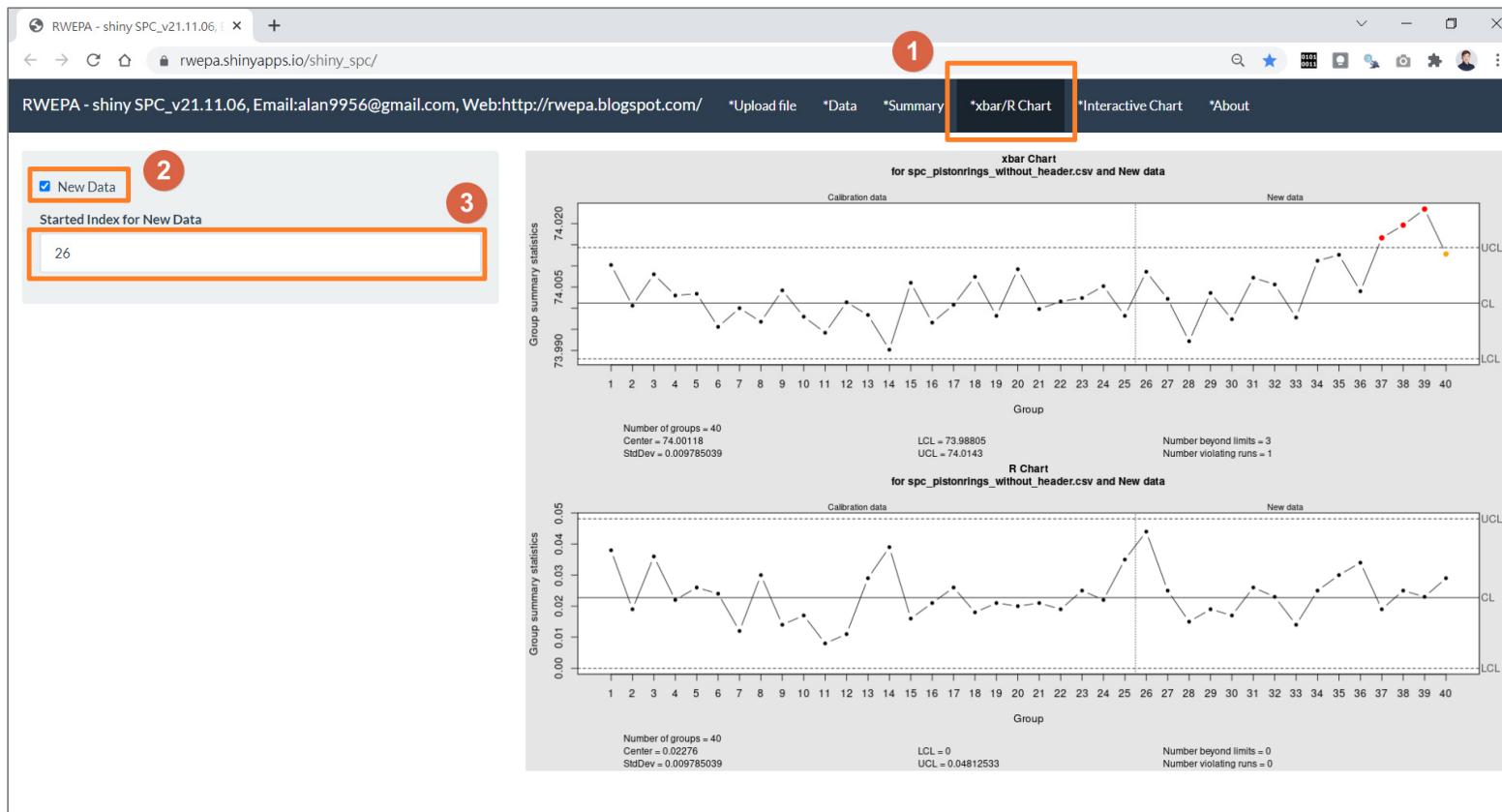
- <https://rwepa.shinyapps.io/shinyCustomerConnect/>





品質管制圖(quality control chart)應用

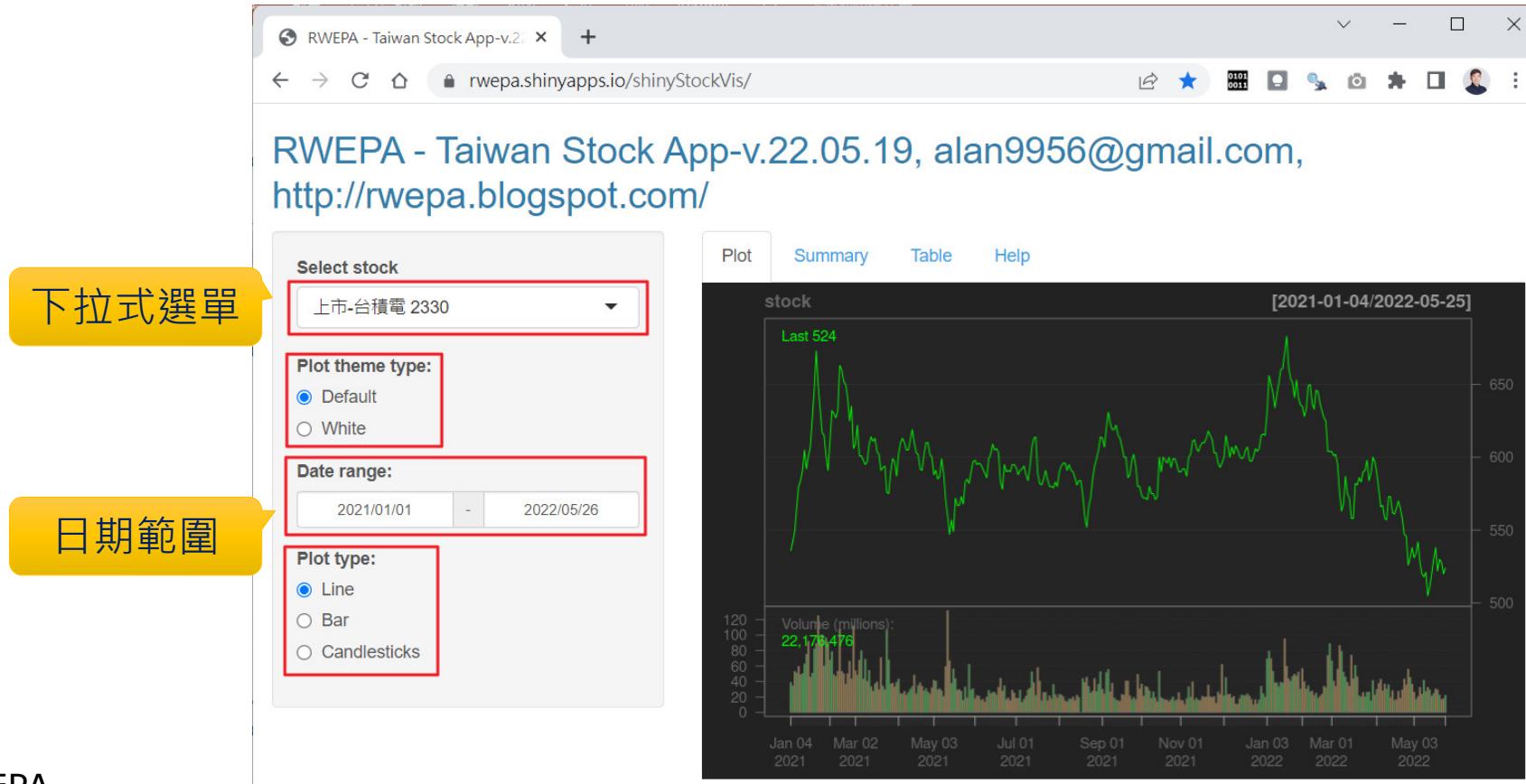
- <http://rwepa.blogspot.com/2021/10/r-shiny-quality-control-chart.html>





Taiwan Stock App

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





離子資料分析與視覺化應用





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. 非監督式學習商業預測
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

2. R, RStudio簡介





R 簡介



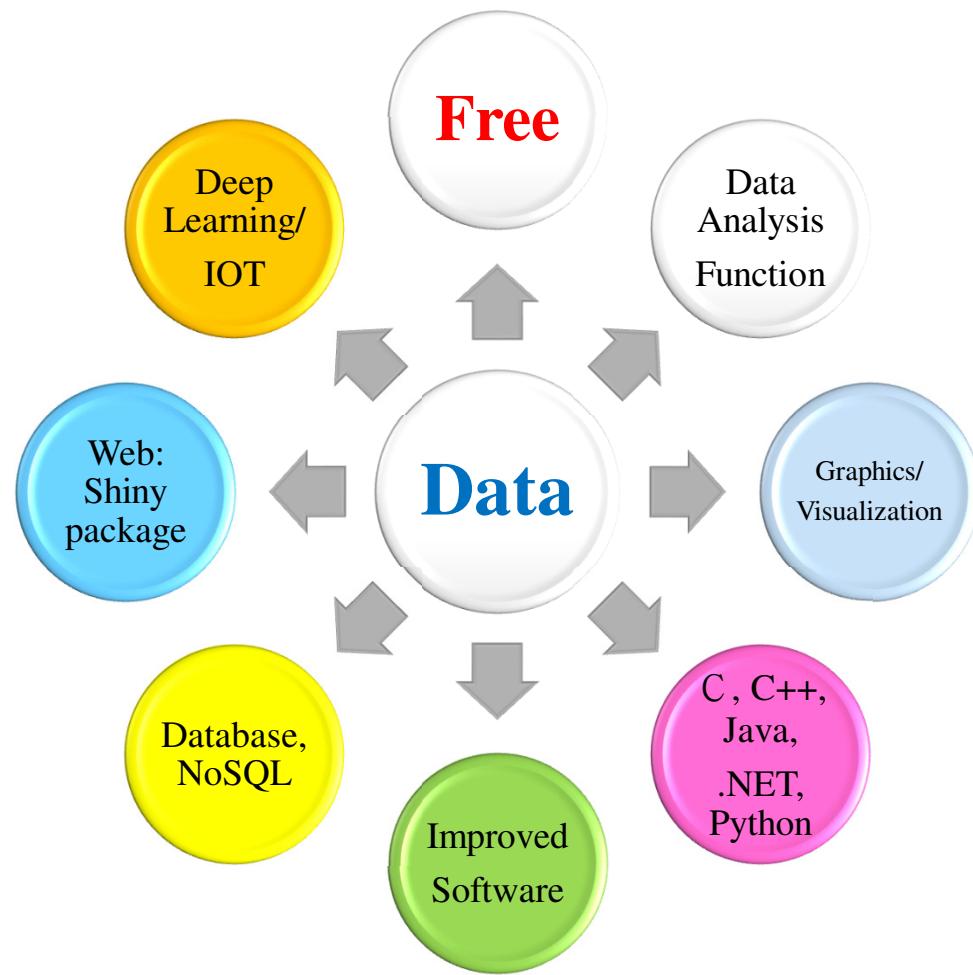
認識 R

- 1976 - 貝爾實驗室 John Chambers, Rick Becker, and Allan Wilks 研發 S 語言。
- 1993 - Ross Ihaka and Robert Gentleman, University of Auckland, New Zealand 將 S 語言研發為 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。
 - R 提供現代資料科學免費實務應用工具。
- 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10月 - R 4.2.2





R-八大功能





R 下載



R官方網頁



[Home]

Download

CRAN

下載

R Project

About R

Logo

Contributors

What's New?

Reporting

The R Project for Statistical Computing

Getting Started

統計計算與繪圖

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.

..



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

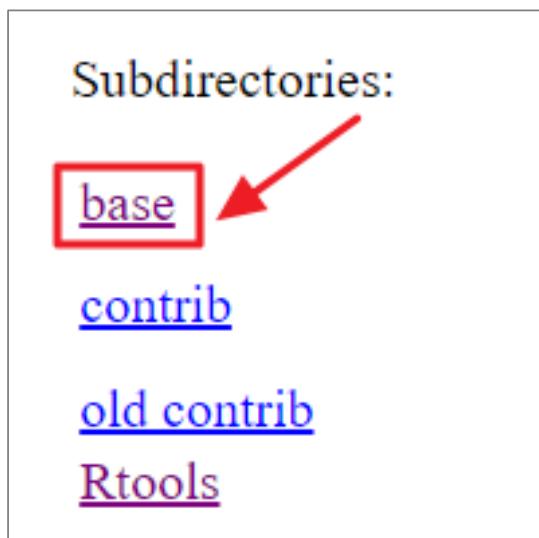
- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)





R-下載 (續)

- 選取 base → 下載 [R-4.2.2-win.exe]



[Download R-4.2.2 for Windows \(76 megabytes, 64 bit\)](#)
[README on the Windows binary distribution](#)
[New features in this version](#)

- 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/rtools42-5355-5357.exe>

Subdirectories:

- base
- contrib
- old contrib
- Rtools**

1

2

- RTools 4.2**
- RTools 4.0
- old versions of RTools

for R $\geq 4.2.0$
for R from version 4.0.0 to 4.1.3
for R versions prior to 4.0.0

Rtools42 may be installed from the [Rtools42 installer](#). It is recommended to use the defaults, including the default installation location of c:\rtools42.

3

Save version information to registry

Create start-menu icons to msys2 shells

2個都打勾

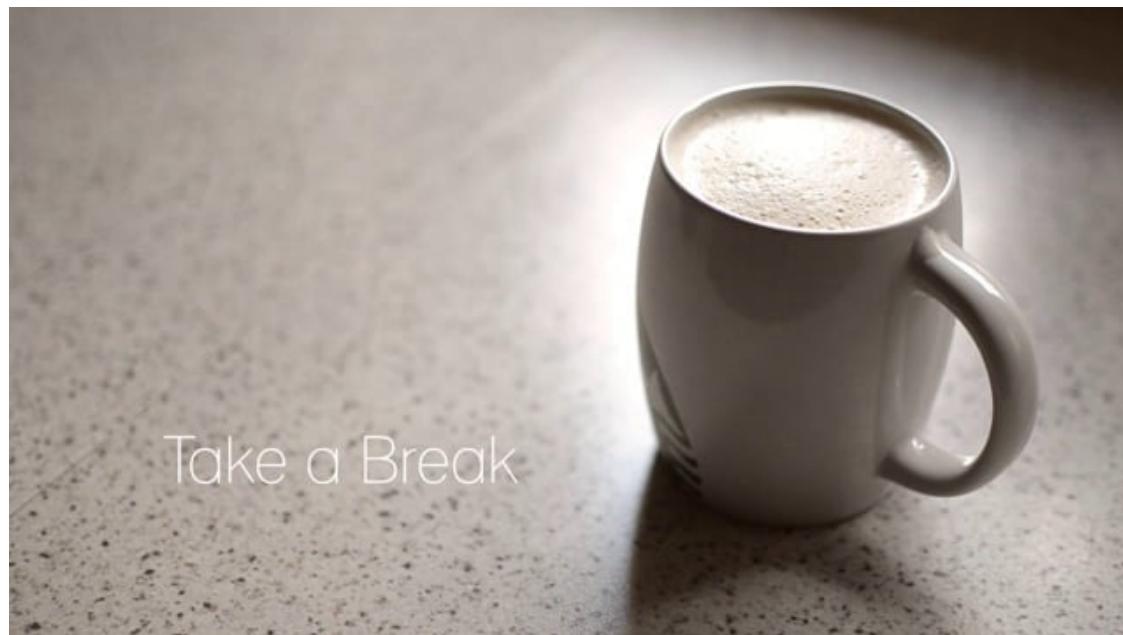


✓ 安裝R

✓ 安裝 Rtools

安裝R, Windows 登入名稱:

1. 不要使用空格
2. 不要使用中文字型



OneDrive問題: <http://rwepa.blogspot.com/2022/06/r-normalizepath-error-onedrive.html>



實作練習

R 執行畫面

R Gui (64-bit)

R Console

R Graphics: Device 2 (ACTIVE)

R version 4.1.1 (2021-08-10) -- "Kick Things"
Copyright (C) 2021 The R Foundation for Statistical Co\$
Platform: x86_64-w64-mingw32/x64 (64-bit)

R 是自由軟體，不提供任何擔保。
在某些條件下歡迎您將其散佈。
用 'license()' 或 'licence()' 來獲得散佈的詳細條件。

R 是個協作計劃，有許多人為之做出了貢獻。
用 'contributors()' 來看詳細的情況以及
用 'citation()' 會告訴您如何在出版品中正確地參照 R 或 \$

用 'demo()' 來看一些示範程式，用 'help()' 來檢視線上輔\$
用 'help.start()' 透過 HTML 瀏覽器來看輔助檔案。
用 'q()' 離開 R。

```
> 1+2  
[1] 3  
> plot(runif(100), type="l", main= "R大數據分析")  
>
```

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

runif(100)

Index

大小寫須一致

@RWEPA

R version 4.1.1 (2021-08-10)



線上參考文件



R Manuals (<https://cran.csie.ntu.edu.tw/>)



[CRAN](#)
[Mirrors](#)
[What's new?](#)
[Task Views](#)
[Search](#)

[About R](#)
[R Homepage](#)
[The R Journal](#)

[Software](#)
[R Sources](#)
[R Binaries](#)
[Packages](#)
[Other](#)

[Documentation](#)
[Manuals](#)
[FAQs](#)
[Contributed](#)

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 <i>per se</i> . That is, the objects that it works on, and the details of the expression evaluation process, which are useful to know when programming R functions.	HTML PDF EPUB	HTML PDF EPUB
R Internals : a guide to the internal structures of R and coding standards for the core team working on R itself.	HTML PDF EPUB	HTML PDF EPUB
The R Reference Index : contains all help files of the R standard and recommended packages in printable form. (9MB, approx. 3500 pages)	HTML PDF EPUB	HTML PDF EPUB

contributed documentation
(貢獻文件, 免費)

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 lately).

English Documents

Documents with more than 100 pages:

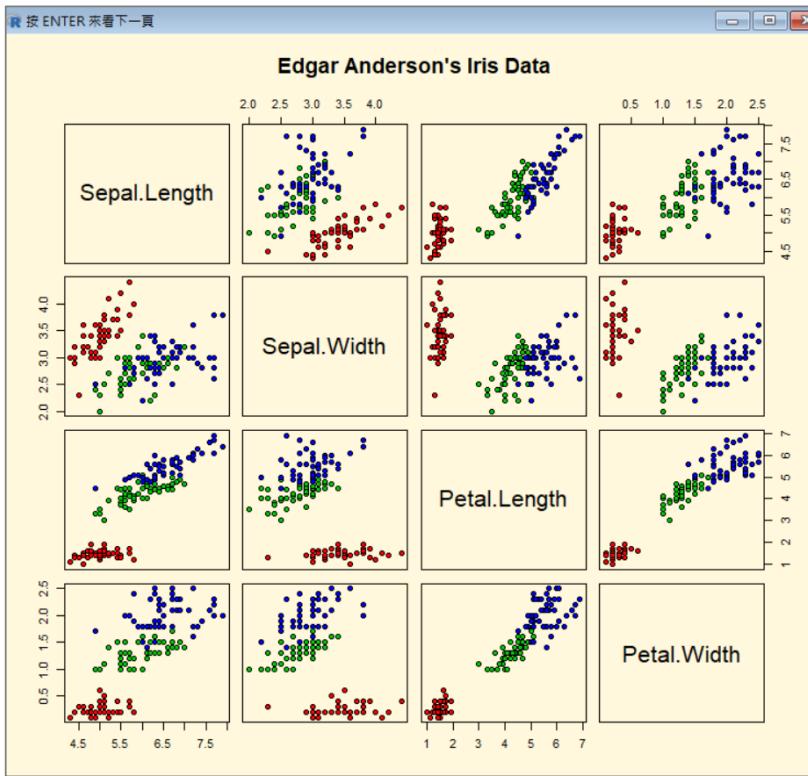
- “**Visual Statistics. Use R!**” by Alexey Shipunov ([PDF](#), 2016-06-06, 301 pages) are accessible from [Alexey Shipunov's English R page](#).
- “**Using R for Data Analysis and Graphics - Introduction, Examples and Commentary**” by John Maindonald ([PDF](#), data sets and scripts are available at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

好書!

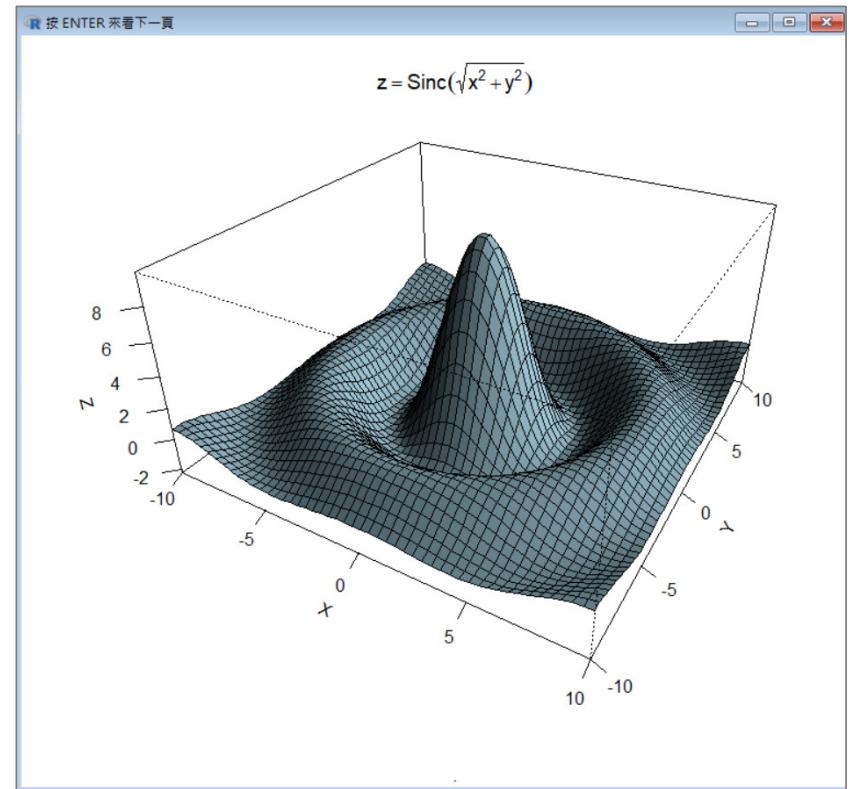


R 視覺化

- `demo(graphics)`



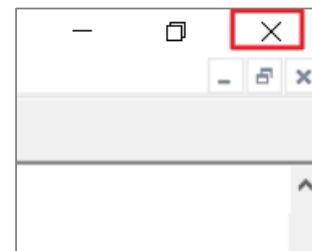
- `demo(persp)`



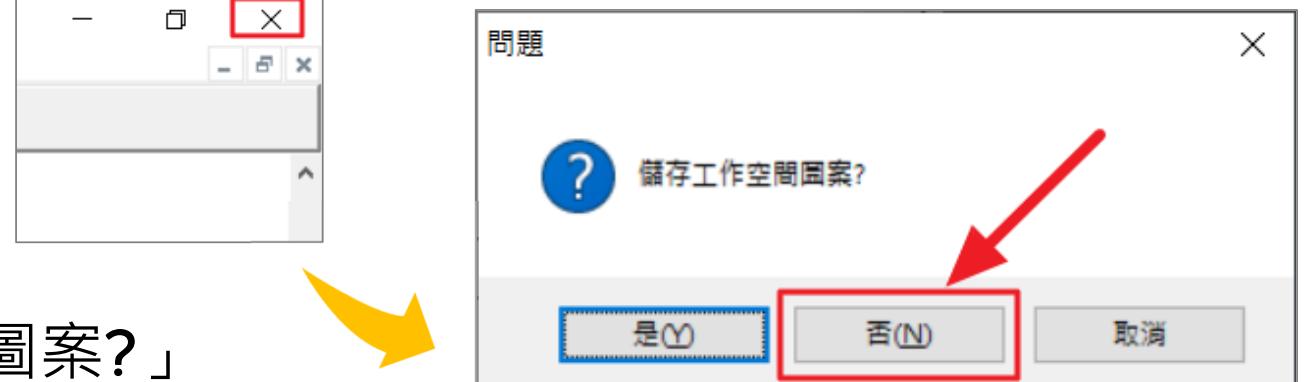


關閉 R

- 按右上角 X 關閉視窗



- 顯示「儲存工作空間圖案？」



- 是(Y): 在「我的文件」會儲存R的資料檔, 檔名: 「.RData」, 下次啟動R會自動載入

[原來儲存的工作空間已還原]
>

- 否(N): 不儲存R的資料檔, 直接關閉R, 一般情形, 選取此選項
- 取消: 返回R視窗

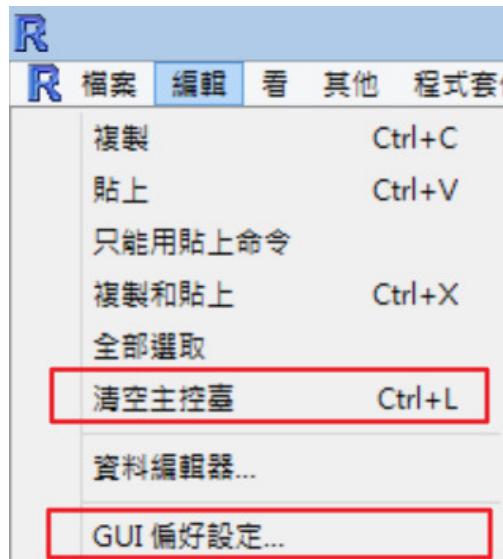


R 功能表

檔案



編輯



- 現行目錄 `getwd()`
- 變更目錄 `setwd()`

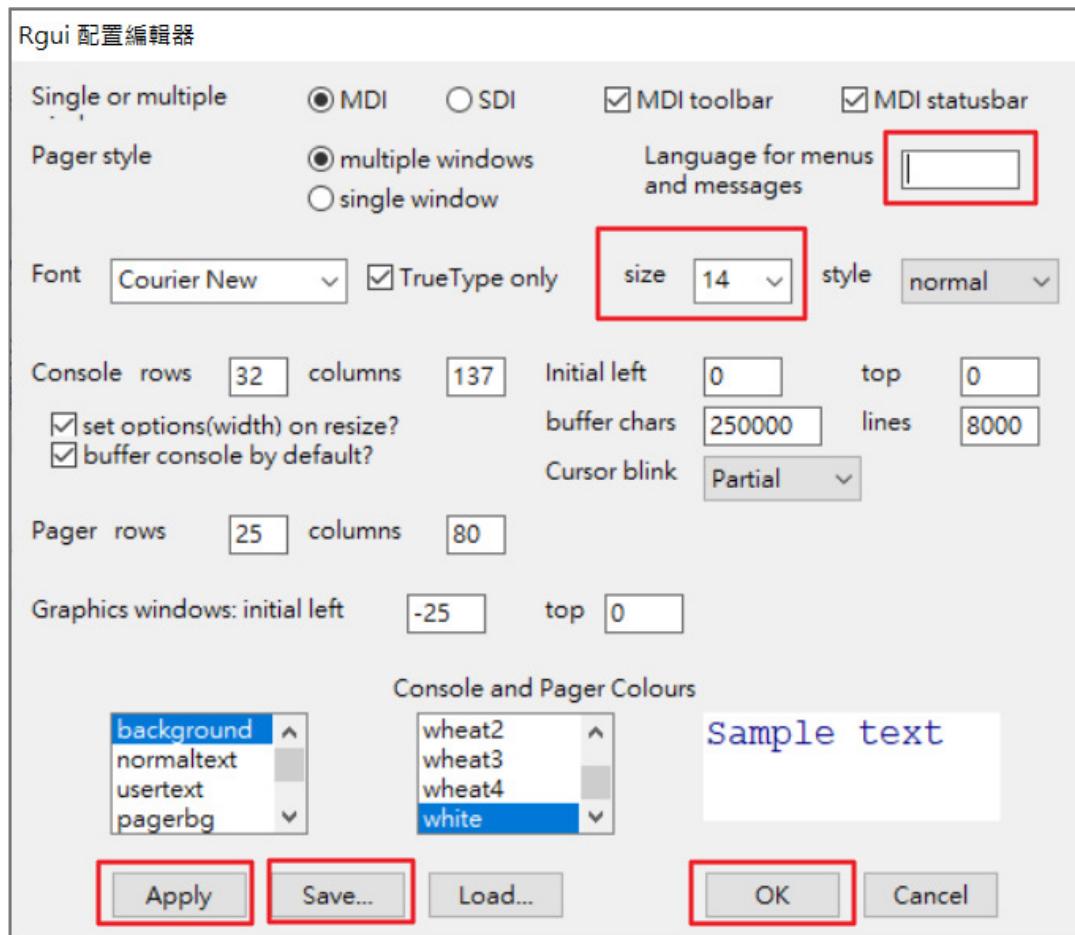
儲存控制台-文字檔

輔助





編輯 \ GUI 偏好設定



- Language: **en** (英文)
- size: 字型大小

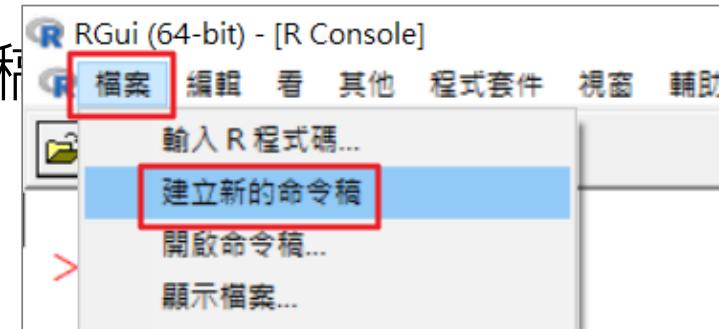


新增R檔案

- 檔案 \ 建立新的命令稿

實作練習

- 輸入以下內容



```
# 主題:第1個R程式
# 作者:李明昌
set.seed(168)
x <- runif(10)
x
plot(x, type = "b", main = "R大數據分析")
pairs(iris[-5], pch=16, col=iris$Species, main="iris資料集-散佈圖矩陣")
```

- 選取程式碼 \ 按  Run line or selection 或 Ctrl + R
- 按 File \ Save \ myfirst.R



新增R檔案練習-執行成果

程式碼

```
# 主題:第1個R程式
# 作者:李明昌
set.seed(168)
x <- runif(10)
x

plot(x, type = "b", main = "R大數據分析")
pairs(iris[-5], pch=16, col=iris$Species, main="iris資料集-散佈圖矩陣")
```

用 'citation()' 會告訴您如何在出版品中正確地參照 R 或 R 套件。

用 'demo()' 來看一些示範程式，用 'help()' 來檢視線上輔助檔案，或用 'help.start()' 透過 HTML 瀏覽器來看輔助檔案。

用 'q()' 離開 R。

控制台

```
> # 主題:第1個R程式
> # 作者:李明昌
> set.seed(168)
> x <- runif(10)
> x
[1] 0.3002227 0.9029981 0.6511521 0.9714271 0.8097960 0.5718119 0.1
> plot(x, type = "b", main = "R大數據分析")
> |
```

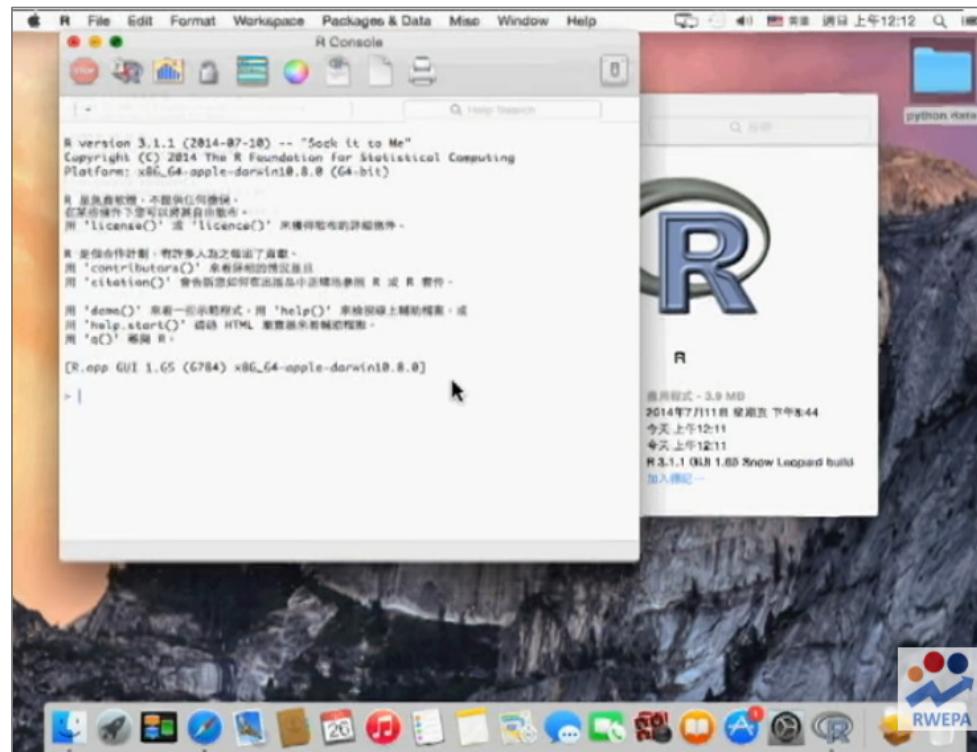
R version 4.1.1 (2021-08-10)

繪圖區



R for Mac

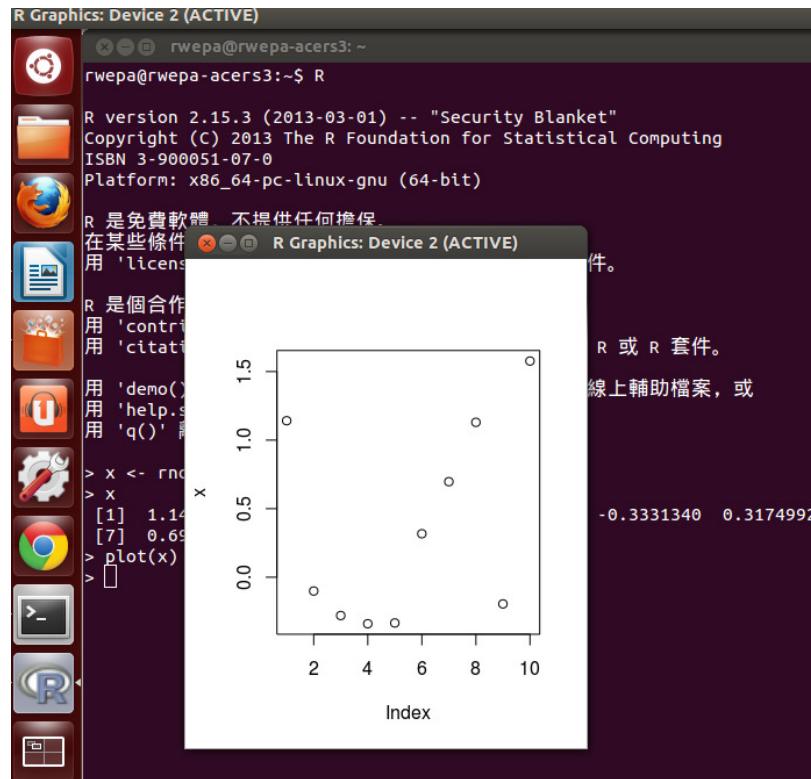
- <https://youtu.be/72MYRBNo5Bk>





R for Ubuntu

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



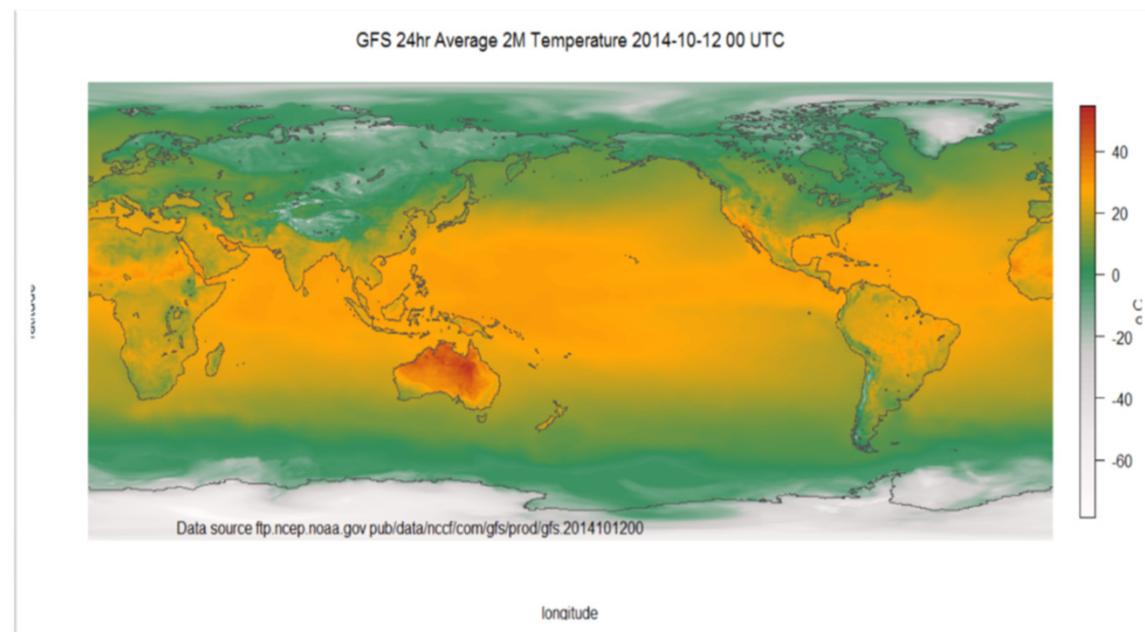


RStudio 簡介



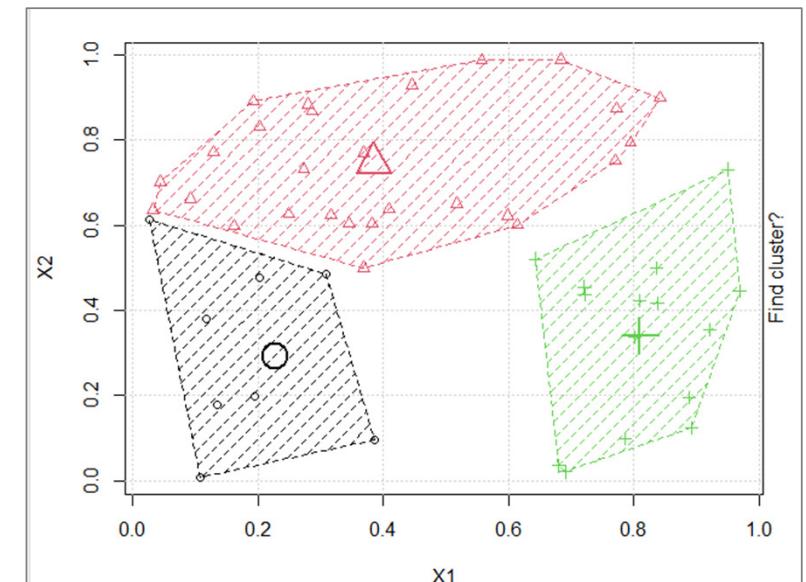
整合式開發環境 – RStudio IDE

- <https://posit.co/>



視覺化應用

(全球2M氣溫圖)



集群法



RStudio – 特性

- 支援智慧輸入 (按Tab)
- 高亮度顯示程式碼
- 整合R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
- 支援 RStudio外掛程式 (Addins)
- 安裝注意:
 - 先安裝R, 再安裝 RStudio
 - 安裝 RStudio時, 請先關閉R



RStudio 下載



RStudio 下載

- <https://posit.co/>

RStudio is now **Posit**,
our mission continues

At Posit, our goal is to make data science more open, intuitive, accessible, and collaborative. We provide tools that make it easy for individuals, teams, and enterprises to leverage powerful analytics and gain insights they need to make a lasting impact.





PRODUCTS \ RStudio IDE

The screenshot shows the posit website's header and a section titled 'PRODUCTS'. The 'PRODUCTS' menu item is highlighted with a red box and a red arrow points from it down to the 'RStudio IDE' link in the 'Open Source' category. Other visible links include 'Enterprise' and 'Cloud' under the 'Open Source' category, and 'Posit Team' and 'Posit Cloud' below.

posit

PRODUCTS ^

SOLUTIONS ▾

LEARN & SUPPORT ▾

EXPLORE MORE ▾

PRODUCTS

Explore our open source, cloud, and enterprise products

Open Source > Enterprise > Cloud >

RStudio IDE

Posit Team

Posit Cloud



DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Open Source Edition

The Premier IDE for R

Pricing

Free

[DOWNLOAD RSTUDIO DESKTOP](#)



[DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS](#)

Size: 190.49MB | [SHA-256: B38BF925](#) | Version: 2022.07.2+576 |

Released: 2022-09-21



RStudio 下載 (續)

RStudio Desktop	RStudio Desktop Pro	RStudio Server	RStudio Workbench
Open Source License	Commercial License	Open Source License	Commercial License
Free	\$995	Free	\$4,975
單機版 /year		伺服器版本 /year (5 Named Users)	
DOWNLOAD	BUY	DOWNLOAD	BUY
Learn more	Learn more	Learn more	Evaluation Learn more
Integrated Tools for R	✓	✓	✓
Priority Support	✓		✓
Access via Web Browser			✓

A red arrow points from the "Free" label under RStudio Desktop to the "DOWNLOAD" button. A yellow callout bubble with the text "免費版" (Free Version) points to the same "DOWNLOAD" button. Another red arrow points from the "Free" label under RStudio Server to the "伺服器版本" (Server Version) label. A third red arrow points from the "Evaluation | Learn more" link under RStudio Workbench to the "Evaluation | Learn more" link under RStudio Server.

The diagram illustrates the download options for RStudio. It shows four main products: RStudio Desktop (Single Machine Version), RStudio Desktop Pro, RStudio Server, and RStudio Workbench. RStudio Desktop and RStudio Server are offered as free versions, while RStudio Desktop Pro and RStudio Workbench are commercial licenses. The Single Machine Version and Server Version are both available for download. A yellow callout bubble labeled "免費版" (Free Version) points to the "DOWNLOAD" button for the Single Machine Version. Red arrows also point from the "Free" labels to the "伺服器版本" (Server Version) label and the "Evaluation | Learn more" link for RStudio Server.



RStudio 安裝

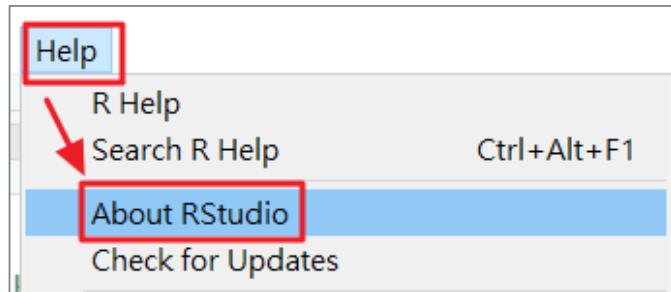


RStudio 安裝

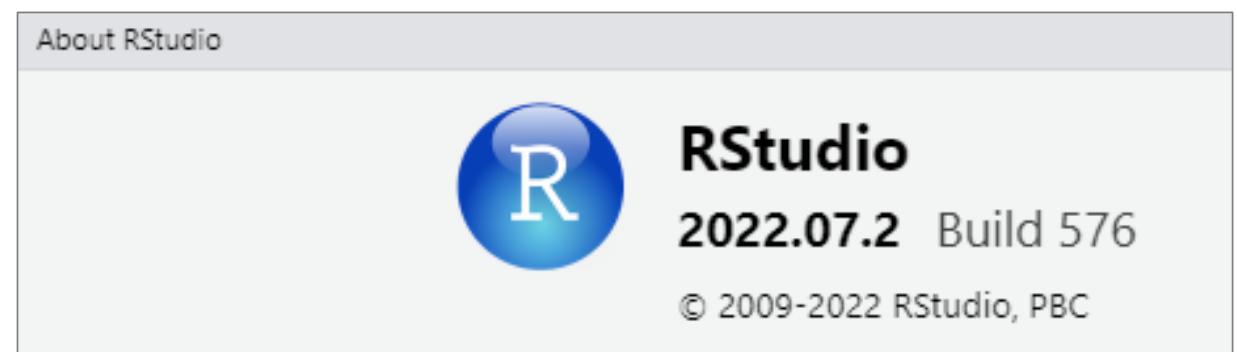




RStudio 版本訊息



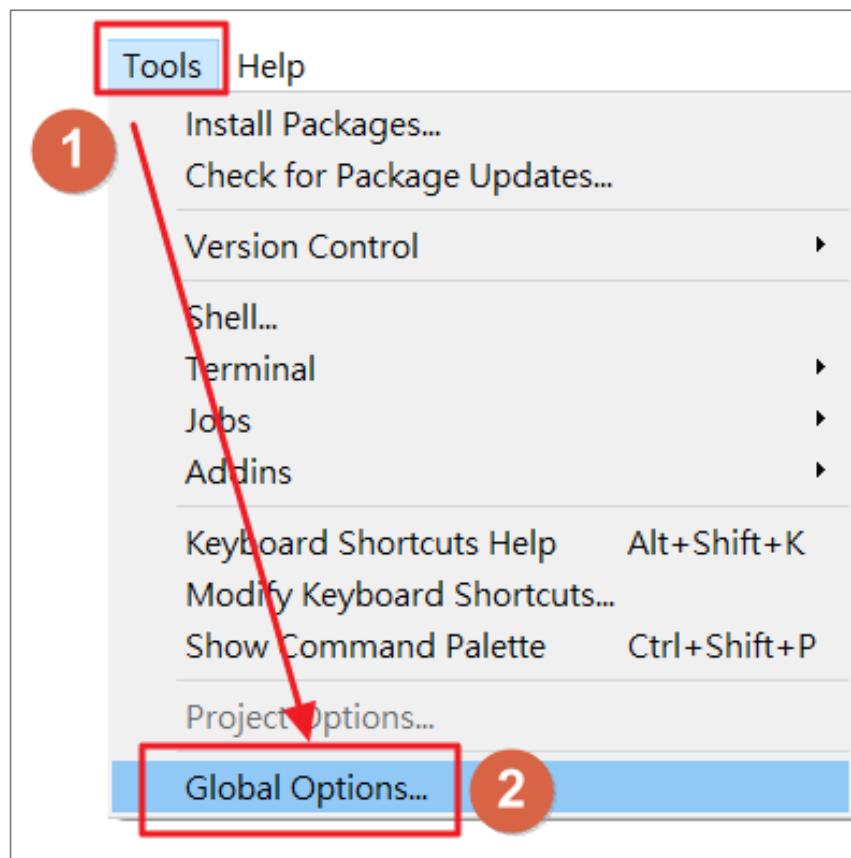
Help \ About RStudio





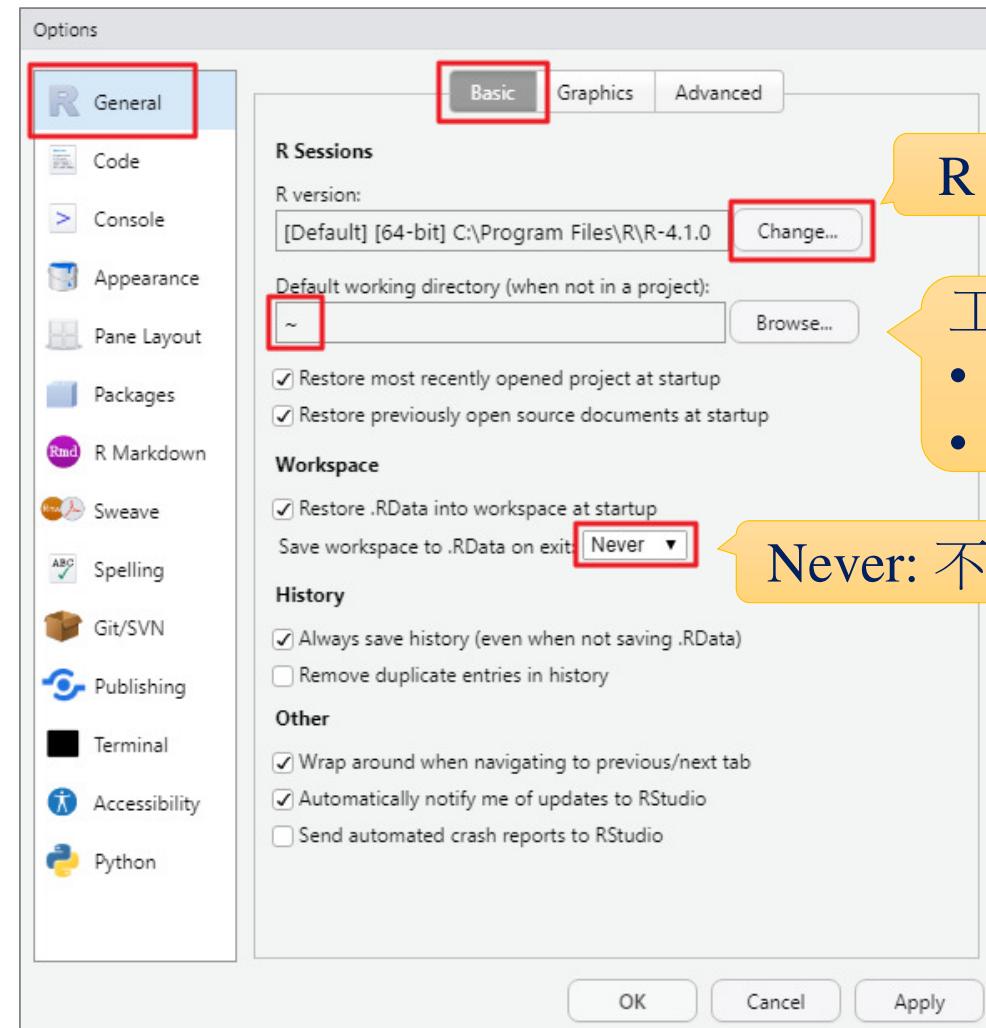
RStudio-選項設定

- Tools \ Global Options





General \ Basic



R 版本

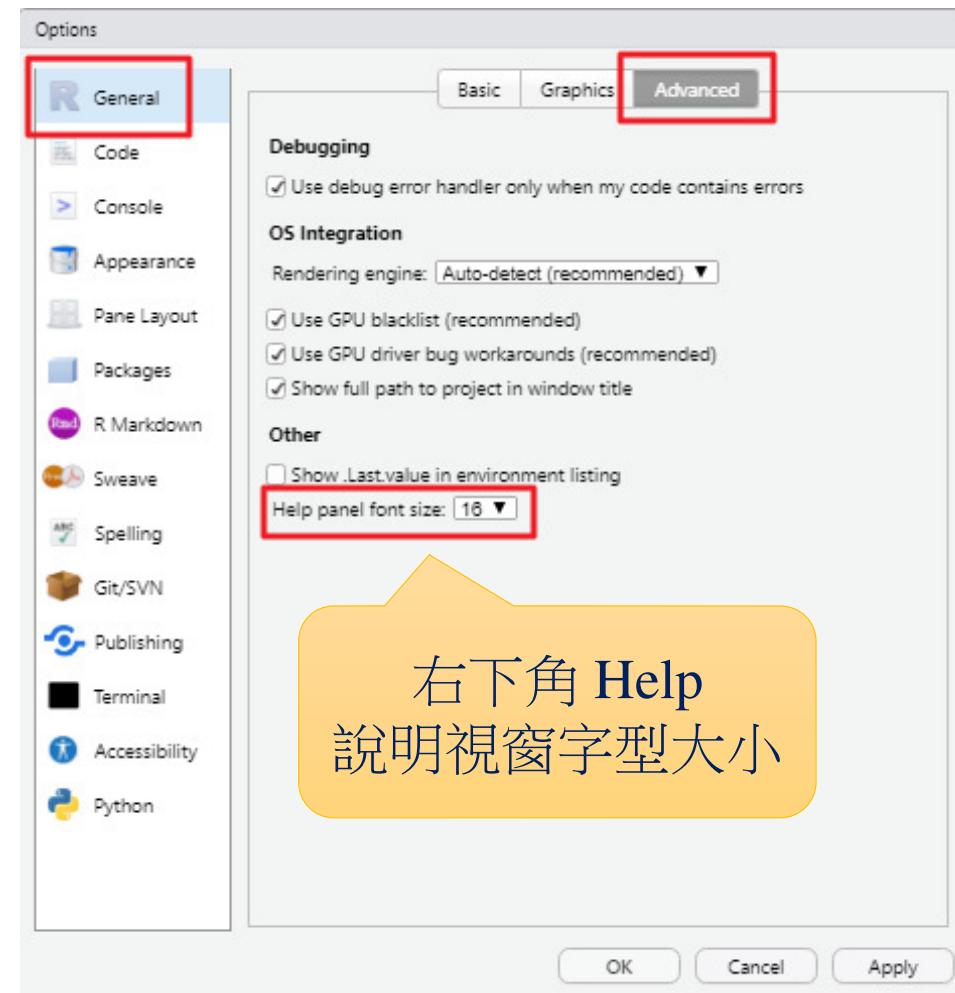
工作目錄:

- 預設值 我的文件
- 修改為 C:/rdata

Never: 不用儲存RData



General \ Advanced





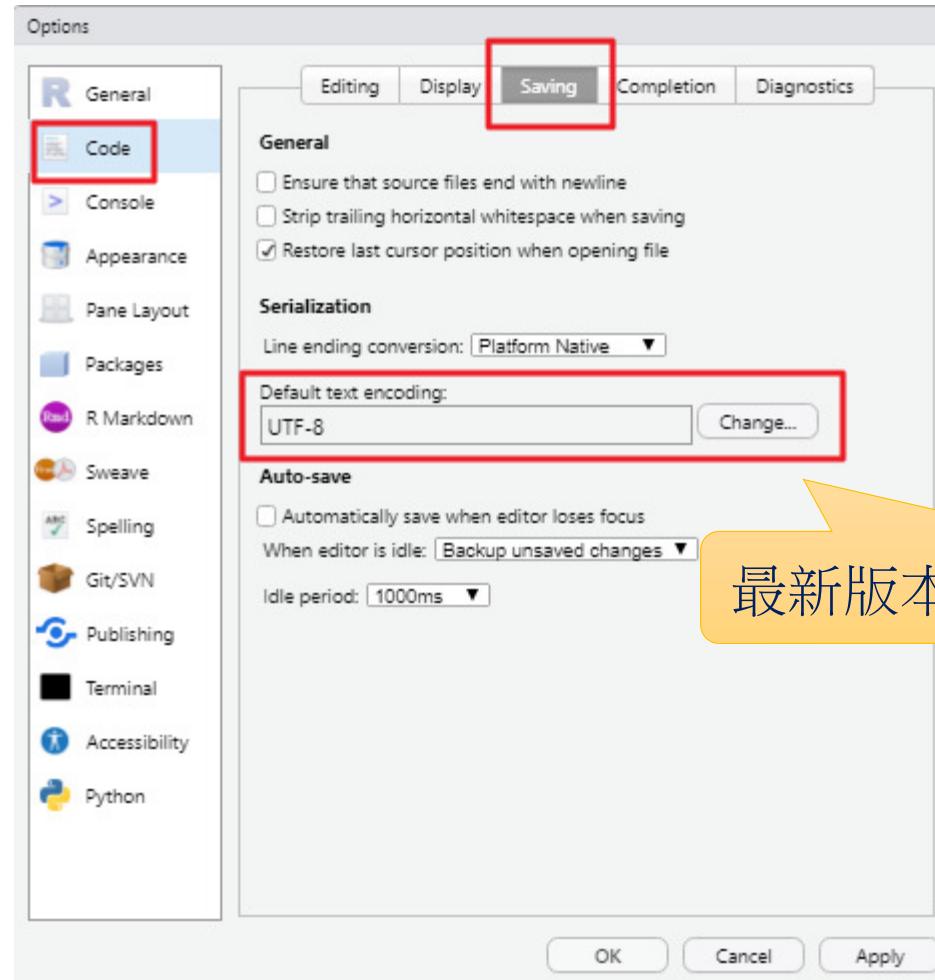
Code \ Editing

The screenshot shows the RStudio Options dialog with the 'Code' tab selected in the left sidebar. The 'Editing' tab is active at the top. In the 'General' section, the 'Soft-wrap R source files' checkbox is highlighted with a red box. A yellow callout bubble on the right contains the text 'Soft-wrap 自動換列'. Other visible options include 'Insert spaces for tab' (Tab width: 2), 'Auto-detect code indentation', 'Insert matching parens/quotes', 'Auto-indent code after paste', and 'Vertically align arguments in auto-indent'. The 'Execution' and 'Snippets' sections are also partially visible.

Soft-wrap
自動換列



Code \ Saving

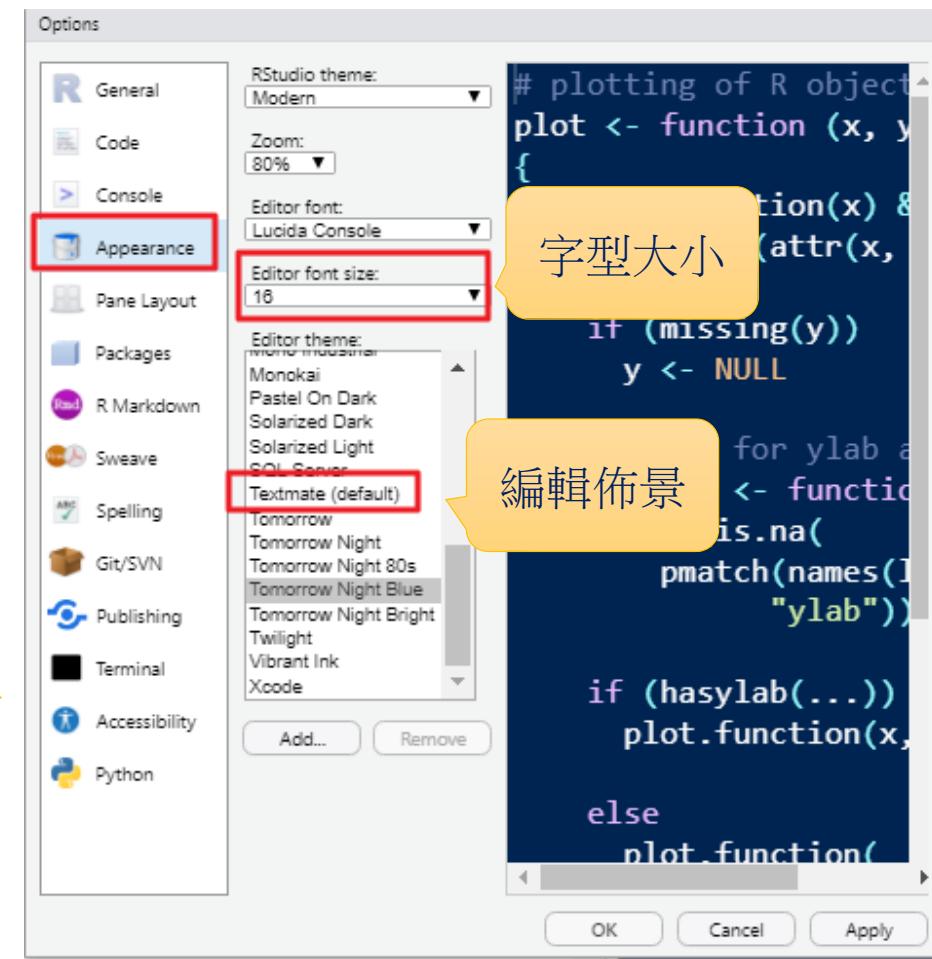




RStudio-選項設定(續)

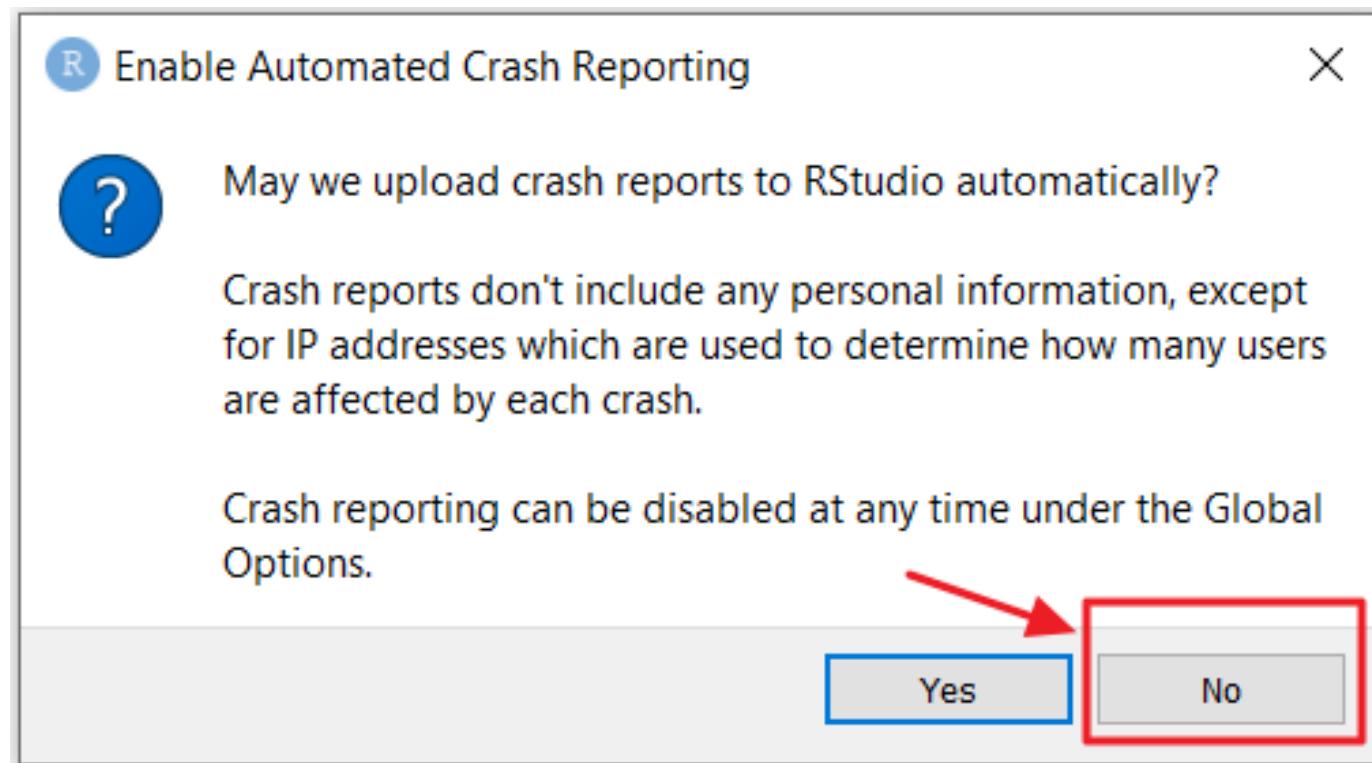
- Appearance \ Editor theme
- 預設值:
TextMate

設定完成,須重
新啟動RStudio





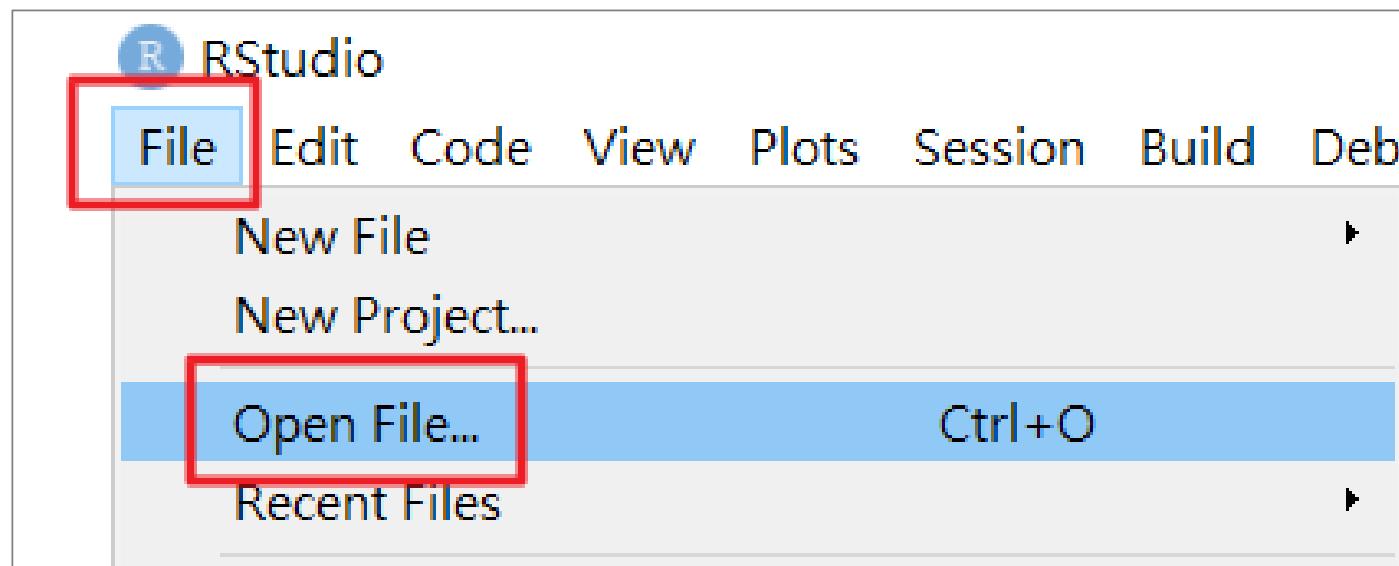
是否啟動自動當機報告 → No





開啟檔案

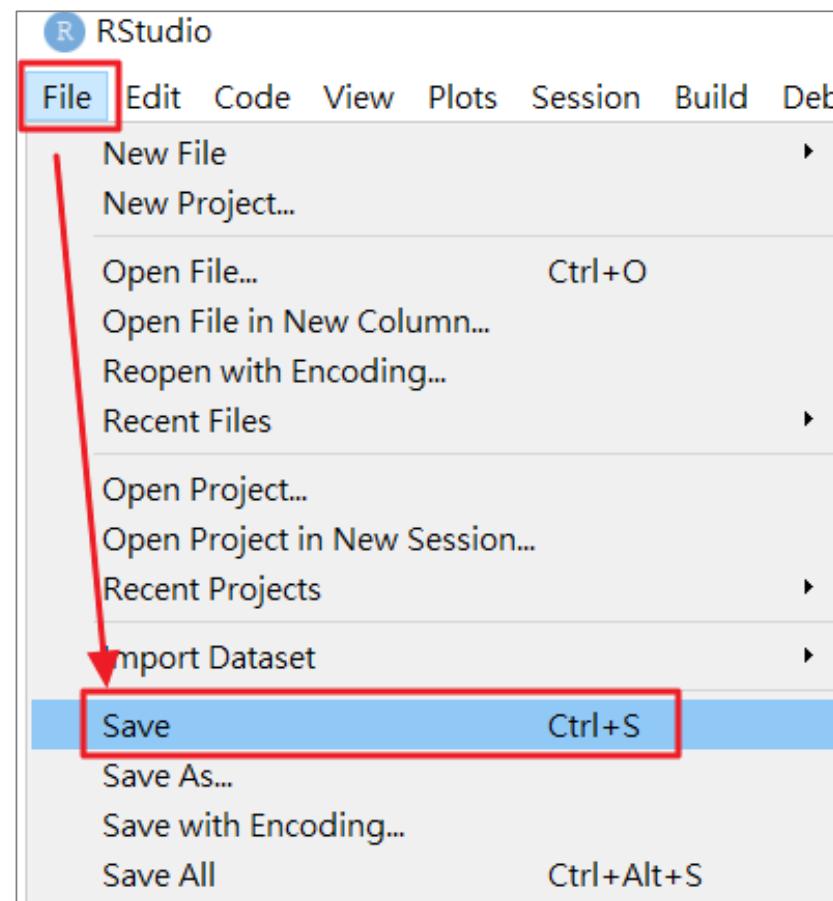
- File \ Open File





儲存檔案

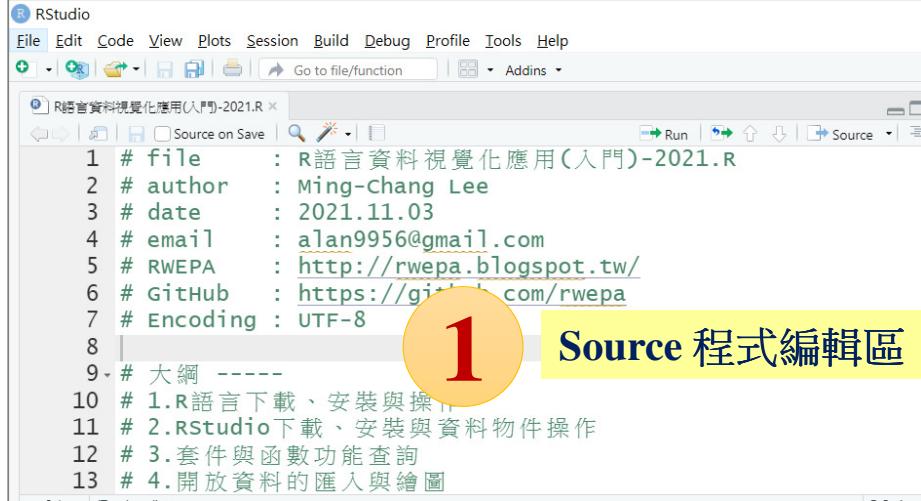
- File \ Save
(CTRL + S)





R/RStudio環境的基礎觀念

1 Source 程式編輯區

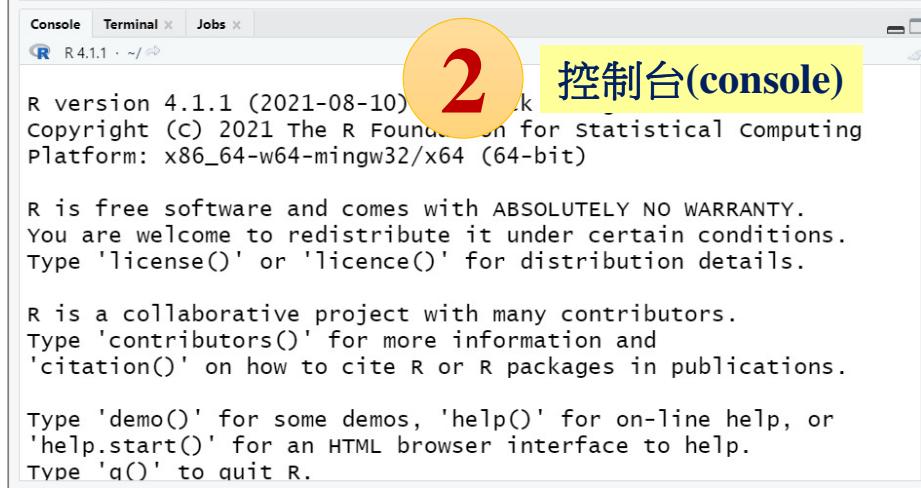


```

1 # file      : R語言資料視覺化應用(入門)-2021.R
2 # author   : Ming-chang Lee
3 # date     : 2021.11.03
4 # email    : alan9956@gmail.com
5 # RWEPA    : http://rwepa.blogspot.tw/
6 # GitHub   : https://github.com/rwepa
7 # Encoding : UTF-8
8
9 # 大綱 -----
10 # 1.R語言下載、安裝與操作
11 # 2.RStudio下載、安裝與資料物件操作
12 # 3.套件與函數功能查詢
13 # 4.開放資料的匯入與繪圖

```

2 控制台(console)



R version 4.1.1 (2021-08-10) -- "Click"
Copyright (c) 2021 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

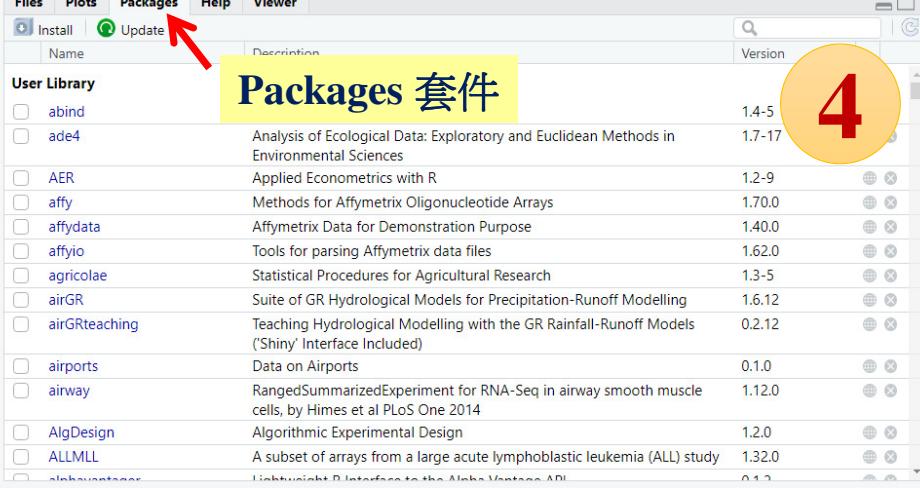
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

3 Environment (工作空間, xxx.Rdata)



Environment is empty

4 Packages 套件

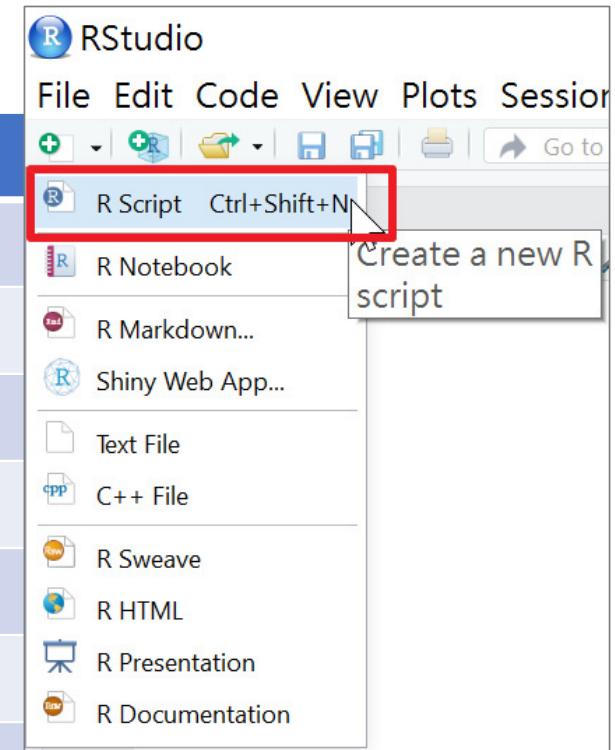


Name	Description	Version
abind	Analysis of Ecological Data: Exploratory and Euclidean Methods in Environmental Sciences	1.4-5
ade4	Applied Econometrics with R	1.7-17
AER	Methods for Affymetrix Oligonucleotide Arrays	1.2-9
affy	Affymetrix Data for Demonstration Purpose	1.70.0
affydata	Tools for parsing Affymetrix data files	1.40.0
affyio	Statistical Procedures for Agricultural Research	1.62.0
agricolae	Suite of GR Hydrological Models for Precipitation-Runoff Modelling	1.3-5
airGR	Teaching Hydrological Modelling with the GR Rainfall-Runoff Models ('Shiny' Interface Included)	1.6.12
airGTeaching	Data on Airports	0.2.12
airports	RangedSummarizedExperiment for RNA-Seq in airway smooth muscle cells, by Himes et al PLoS One 2014	0.1.0
airway	Algorithmic Experimental Design	1.12.0
AlgDesign	A subset of arrays from a large acute lymphoblastic leukemia (ALL) study	1.32.0
ALLMLL	Ljungkvist's R Interface to the Alpha Ventage API	0.1.2
alphavantagger		



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)

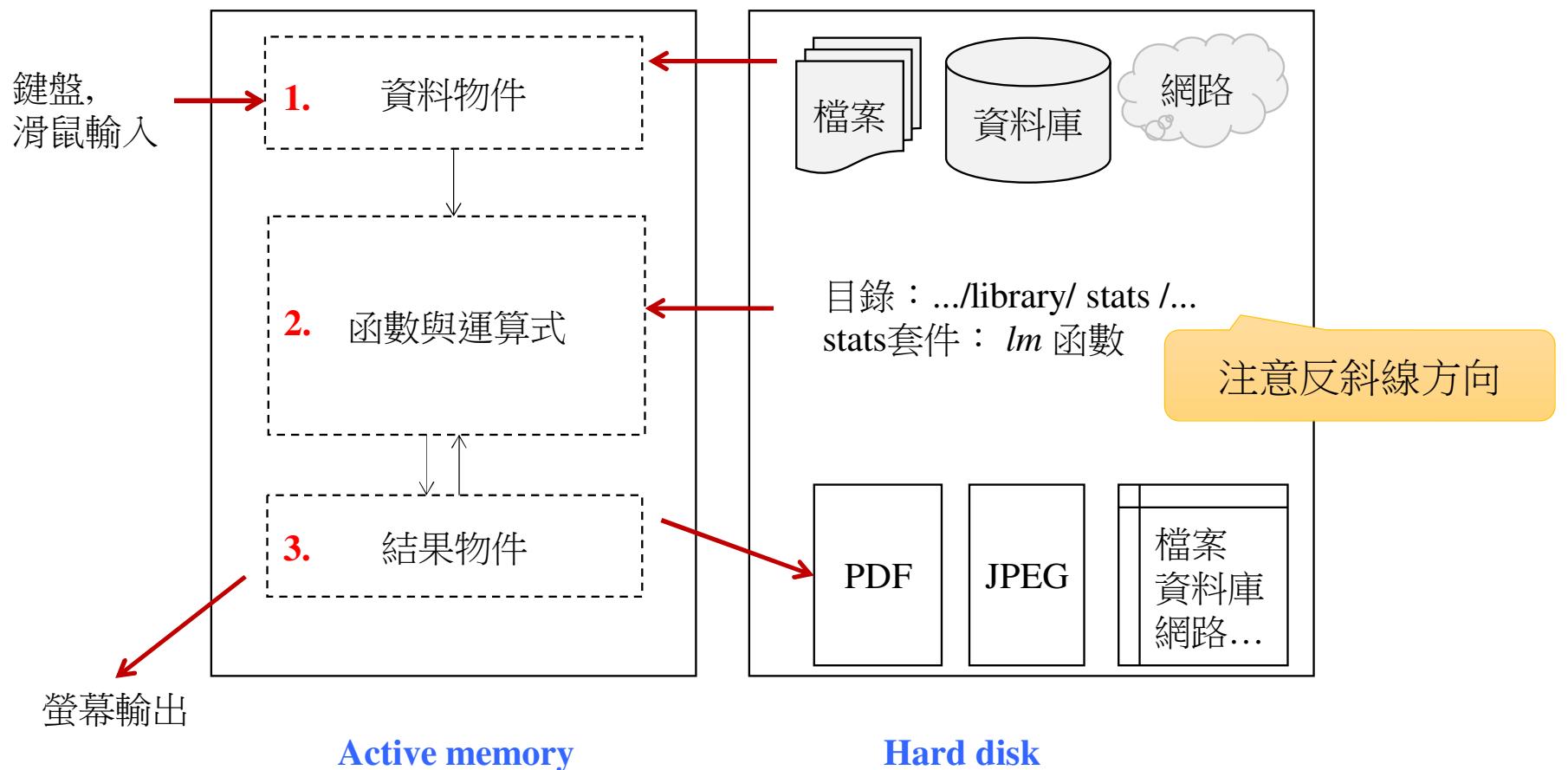




R使用環境



R運作方式





基本觀念

- 控制台(console)
- 歷程
 - xxx.Rhistory
- 套件(package)
- 工作空間(workspace)
 - xxx.RData
- 物件(object)

```
R Gui (64-bit) - [R Console]
R 檔案 編輯 看 其他 程式套件 視窗 輔助
[Icons]

R version 3.0.1 (2013-05-16) -- "Good Sport"
Copyright (C) 2013 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R 是免費軟體，不提供任何擔保。
在某些條件下您可以將其自由散布。
用 'license()' 或 'licence()' 來獲得散布的詳細條件。

R 是個合作計劃，有許多人為之做出了貢獻。
用 'contributors()' 來看詳細的情況並且
用 'citation()' 會告訴您如何在出版品中正確地參照 R 或 R 套件。

用 'demo()' 來看一些示範程式，用 'help()' 來檢視線上輔助檔案，或
用 'help.start()' 透過 HTML 瀏覽器來看輔助檔案。
用 'q()' 離開 R。

Welcome at Tue Aug 13 03:01:11 2013
> |
```



控制台的特定符號

- 命令提示字元(大於) > (等待使用者輸入資料)
- 指令未完提示字元(加號) + (表示尚未輸入完成)
- 註解提示字元(井字號) # (不會編譯註解)
- 結果行列顯示編號

```
> iris$Sepal.Length
[1] 5.1 4.9 4.7 4.6 5.0 5.4 4.6 5.0 4.4 4.9 5.4 4.8 4.8 4.3 5.8 5.7
[17] 5.4 5.1 5.7 5.1 5.4 5.1 4.6 5.1 4.8 5.0 5.0 5.2 5.2 4.7 4.8 5.4
[33] 5.2 5.5 4.9 5.0 5.5 4.9 4.4 5.1 5.0 4.5 4.4 5.0 5.1 4.8 5.1 4.6
[49] 5.3 5.0 7.0 6.4 6.9 5.5 6.5 5.7 6.3 4.9 6.6 5.2 5.0 5.9 6.0 6.1
[65] 5.6 6.7 5.6 5.8 6.2 5.6 5.9 6.1 6.3 6.1 6.4 6.6 6.8 6.7 6.0 5.7
[81] 5.5 5.5 5.8 6.0 5.4 6.0 6.7 6.3 5.6 5.5 5.5 6.1 5.8 5.0 5.6 5.7
[97] 5.7 6.2 5.1 5.7 6.3 5.8 7.1 6.3 6.5 7.6 4.9 7.3 6.7 7.2 6.5 6.4
[113] 6.8 5.7 5.8 6.4 6.5 7.7 7.7 6.0 6.9 5.6 7.7 6.3 6.7 7.2 6.2 6.1
[129] 6.4 7.2 7.4 7.9 6.4 6.3 6.1 7.7 6.3 6.4 6.0 6.9 6.7 6.9 5.8 6.8
[145] 6.7 6.7 6.3 6.5 6.2 5.9
```



物件命名原則

- R的大小寫有差異: *a* 與 *A* 是不同的物件
- R 也保留一些物件與指令人稱, 如 c, C, T, F 等為保留字 ("reserved words"), 命名時避免重覆, 以免引起人類困擾
- 物件名稱起始字元以英文字母(a~z, A~Z)為主
- 物件名稱起始不可以使用數字
- 物件名稱第二個字元以英文字母或數字(0-9)為主
- 物件名稱中間不可有空格
- 物件名稱中間可以使用下底線 _
- 物件名稱中間不建議使用「.」, 例: 將 Customer.Sales 改為 CustomerSales



Google's R Style Guide

- <https://google.github.io/styleguide/Rguide.html>
- 函數使用 BigCamelCase

```
# Good
DoNothing <- function() {
  return(invisible(NULL))
}
```

- 使用 **x <- 1**, 不要使用 **x = 1**
- **=** 用於函數之參數設定 **plot(..., type = "b")**



資料型別



資料型別

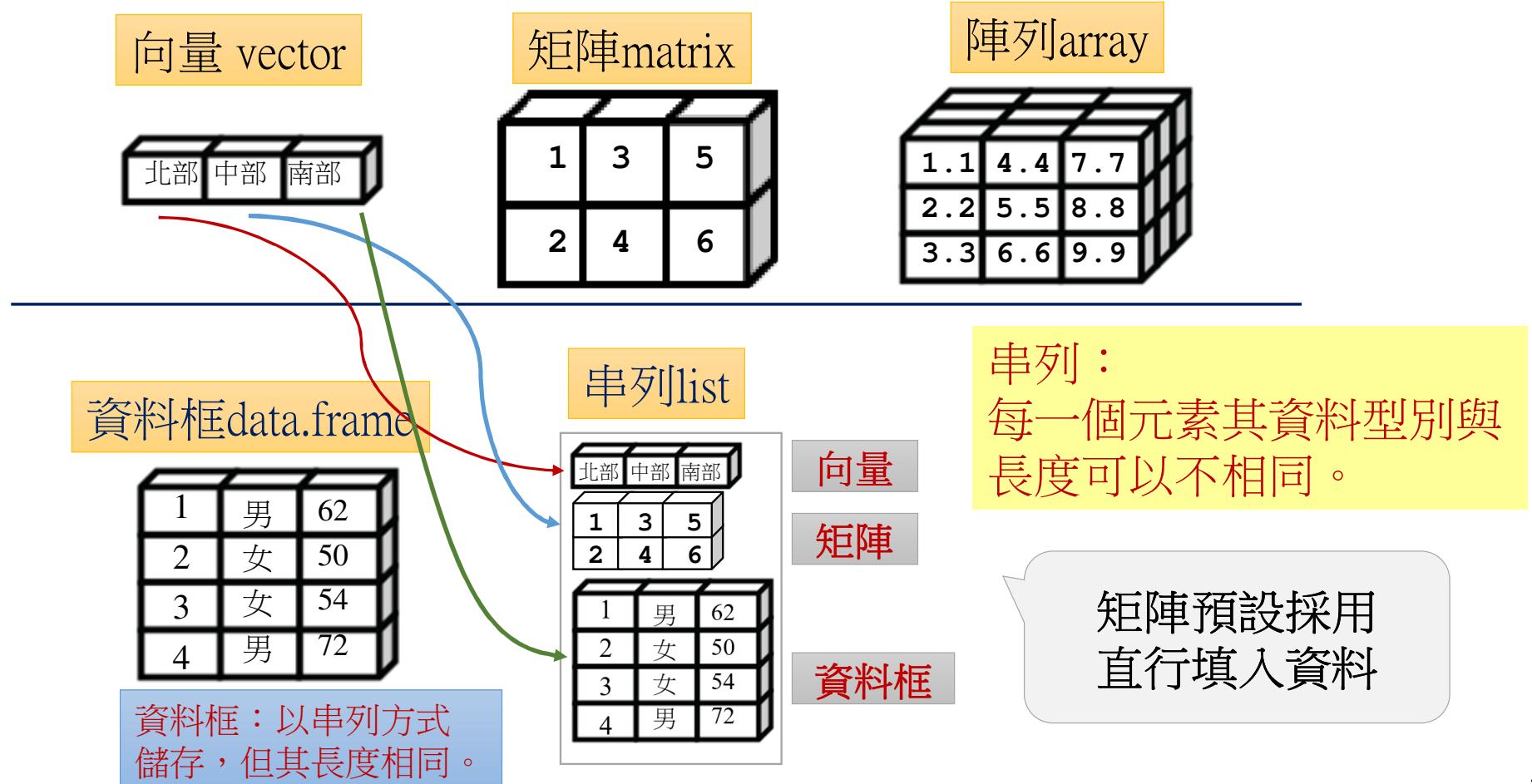
- 整數: 不可以有小數點之數值
- 數值: 可以包括小數點之數值
- 字串: 須使用 '台北市' 或 "台北市" 符號表示
- 邏輯值: 包括 TRUE, FALSE (全部大寫)
- NA: Not Available 不可使用, 表示遺漏值 Missing value



資料物件



資料物件





資料物件重要觀念

- 向量是最基本的物件
 - 向量包括整數向量、數值向量、字元向量與邏輯值向量
 - 所有資料的資料型別須相同, 向量不可以混合使用
 - 因子是一種較特別的向量, 儲存類別型變數, 例: {男,女}, {小, 中, 大}
- 矩陣與陣列
 - 矩陣是二維陣列, 陣列允許大於或等於2個維度
 - 所有資料的資料型別須相同
- 資料框類似 Excel 工作表, 同一直行資料表示向量
- 串列可以包含向量、矩陣、陣列與資料框等資料
- 統計與機器學習演算法輸入值須為資料框 (`data.frame`)
- 統計與機器學習演算法回傳值以串列(`list`)為主



建立向量函數 c (concatenate)

- 將一群數字、字串、邏輯值結合成向量
- 可以將多個向量結合成一個向量
- 向量物件具有屬性長度 `length` 與型式 `mode`
- 向量會將所有元素將強制(`coercion`)轉換成單一相同型態
- 因子 (`factor`) 是一種特別的向量，用於將資料依離散型變數做成分群 (`group`)



套件 package



套件

- 使用套件兩部曲 - 先安裝, 再載入套件
 - `install.packages("套件名稱")` # 安裝套件(一生一次)
 - `library(套件名稱)` # 載入套件(每次使用)
- 範例: 新增與載入 e1071套件(machine learning)

```
> install.packages("e1071")
trying URL 'http://cran.cs.pu.edu.tw/bin/windows/contrib/3.0/e1071_1.6-1.zip'
Content type 'application/zip' length 514468 bytes (502 Kb)
opened URL
downloaded 502 Kb

package 'e1071' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\Administrator\AppData\Local\Temp\RtmpoHS0Ak\downloaded_packages
> library(e1071)
Loading required package: class
```

example(svm, package="e1071")



已載入的套件 search()

```
> # 已載入套件
> search()
[1] ".GlobalEnv"
[3] "tools:rstudio"
[5] "package:graphics" package:graphics
[7] "package:utils"
[9] "package:methods"
[11] "package:base"
>
```

"package:e1071" package:e1071
"package:stats"
"package:grDevices"
"package:datasets" package:datasets
"Autoloads"





R套件 - 42類別

(2022.11.22)

- <https://cran.csie.ntu.edu.tw/web/packages/index.html>

Contributed Packages

Available Packages

Currently, the CRAN package repository features 18870 available packages.

[Table of available packages, sorted by date of publication](#)

[Table of available packages, sorted by name](#)

[CRAN Task Views](#) aim to provide some guidance which packages on CRAN are relevant for tasks related to a certain topic. They provide tools to automatically install all packages from each view. Currently, 42 views are available.



42類別 - 中文說明

2013年10月8日 星期二

Task Views - R套件

更新日期: 2022.11.22 - 42個套件類別

CRAN Task View:

<https://cran.csie.ntu.edu.tw/web/views/>

CRAN (Taiwan):

<https://cran.csie.ntu.edu.tw/>

選取 CRAN 網站左側 [Packages]，套件區分成以下類別，中文說明如下：

Agriculture, Agricultural Science, 農業科學

Bayesian, Bayesian Inference, 貝氏統計

CausalInference, Causal Inference, 因果推論

RWEPA網站 → task



R - 程式套件



- 安裝程式套件 `install.packages()`
- 載入程式套件 `library()`



RStudio 套件管理

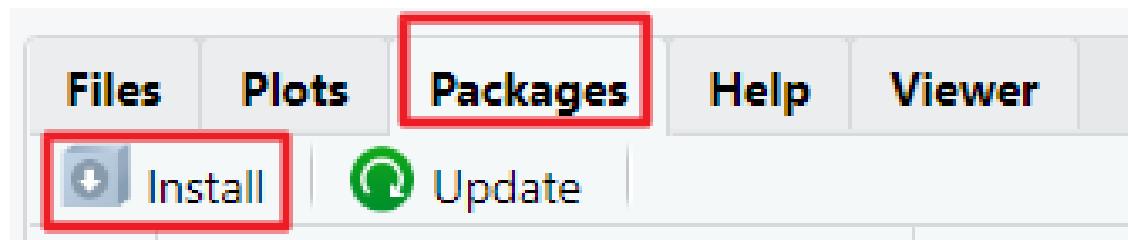
打勾表示已經
載入之套件

	Name	Description	Version	Install	Update
<input type="checkbox"/>	lavaan	Latent Variable Analysis	0.6-7		
<input type="checkbox"/>	lavaan.survey	Complex Survey Structural Equation Modeling (SEM)	1.1.3.1		
<input type="checkbox"/>	lazyeval	Lazy (Non-Standard) Evaluation	0.2.2		
<input type="checkbox"/>	lbfgs	Limited-memory BFGS Optimization	1.2.1		
<input type="checkbox"/>	lbfgsb3c	Limited Memory BFGS Minimizer with Bounds on Parameters with optim() 'C' Interface	2020-3.2		
<input type="checkbox"/>	la	Liouville Copulas	1.0.4		
<input type="checkbox"/>	latentDirichletAllocation	Collapsed Gibbs Sampling Methods for Topic Models	1.4.2		
<input type="checkbox"/>	lavaan	Interactive Visualization of Topic Models	0.3.2		
<input type="checkbox"/>	leafem	'leaflet' Extensions for 'mapview'	0.1.6		
<input checked="" type="checkbox"/>	leaflet	Create Interactive Web Maps with the JavaScript 'Leaflet' Library	2.0.4.1		
<input type="checkbox"/>	leaflet.providers	Leaflet Providers	1.9.0		
<input type="checkbox"/>	leafpop	Include Tables, Images and Graphs in Leaflet Pop-Ups	0.1.0		

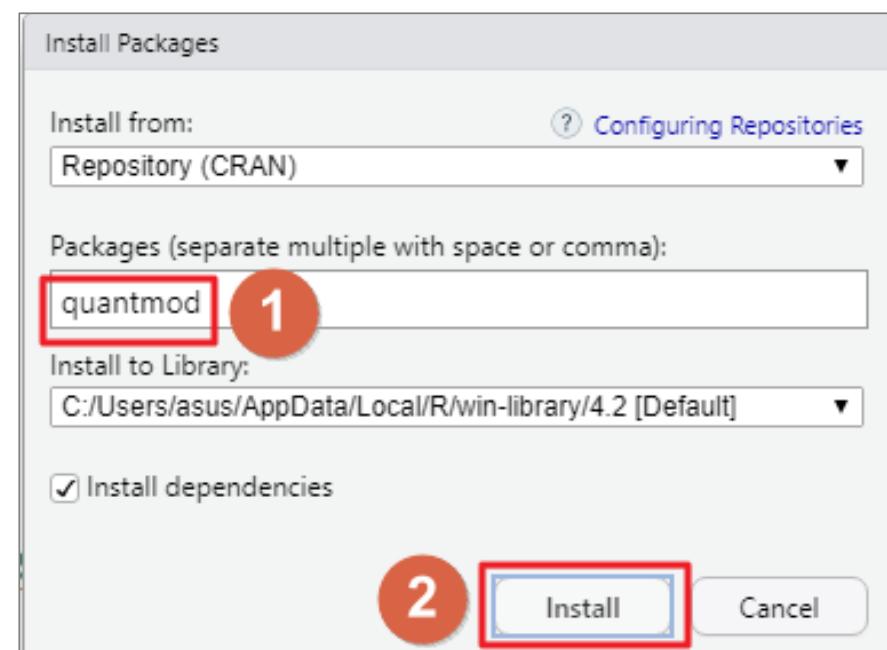


套件安裝

- Packages \ Install



- 輸入 qcc \ 按 Install





Google → r cran quantmod

- <https://cran.r-project.org/web/packages/quantmod/index.html>

quantmod: Quantitative Financial Modelling Framework

Specify, build, trade, and analyse quantitative financial trading strategies.

Version: 0.4.20
Depends: R (≥ 3.2.0), [xts](#) (≥ 0.9-0), [zoo](#), [TTR](#) (≥ 0.2), methods
Imports: [curl](#)
Suggests: [DBI](#), [RMySQL](#), [RSQLite](#), [timeSeries](#), [xml2](#), [downloader](#), [jsonlite](#) (≥ 1.1)
Published: 2022-04-29
Author: Jeffrey A. Ryan [aut, cph], Joshua M. Ulrich [cre, aut], Ethan B. Smith [ctb], [ctb]
Maintainer: Joshua M. Ulrich <josh.m.ulrich at gmail.com>
BugReports: <https://github.com/joshualrich/quantmod/issues>
License: [GPL-3](#)
URL: <http://www.quantmod.com> <https://github.com/joshualrich/quantmod>
NeedsCompilation: no
Materials: [NEWS](#)
In views: [Finance](#)
CRAN checks: [quantmod results](#)

Documentation:

Reference manual: [quantmod.pdf](#)



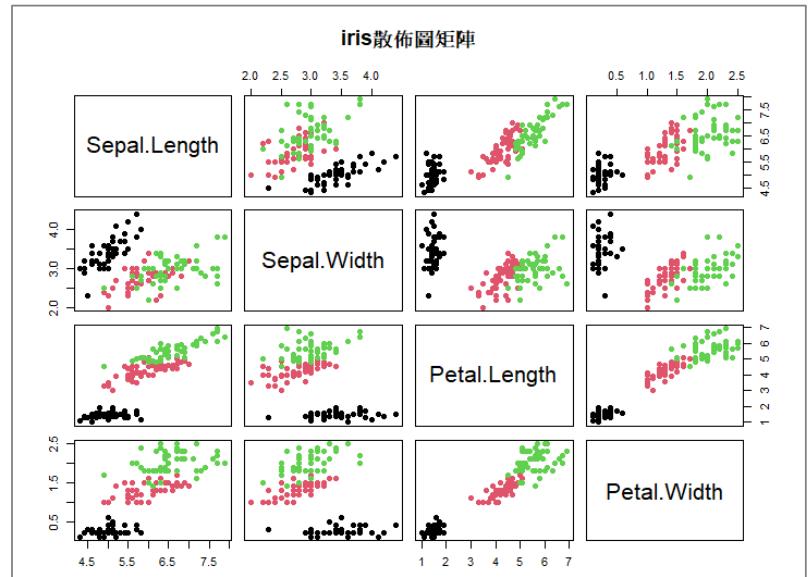
認識 iris 資料集



安德森鳶尾花卉資料集

https://en.wikipedia.org/wiki/Iris_flower_data_set

- iris # iris資料集
- str(iris) # 資料結構
- summary(iris) # 資料摘要
- iris\$Sepal.Length # 花萼長度
- iris\$Sepal.Width # 花萼寬度
- iris\$Petal.Length # 花瓣長度
- iris\$Petal.Width # 花瓣寬度
- iris\$Species # 屬種
- iris[1:3, 1:5] # 第1~3列, 第1~5行
- **boxplot**(iris\$Sepal.Width) # 盒鬚圖
- **plot**(iris\$Petal.Length, iris\$Petal.Width, pch=16, col=iris\$Species) # 散佈圖
- **pairs**(iris[-5], pch=16, col=iris\$Species, main="iris散佈圖矩陣")



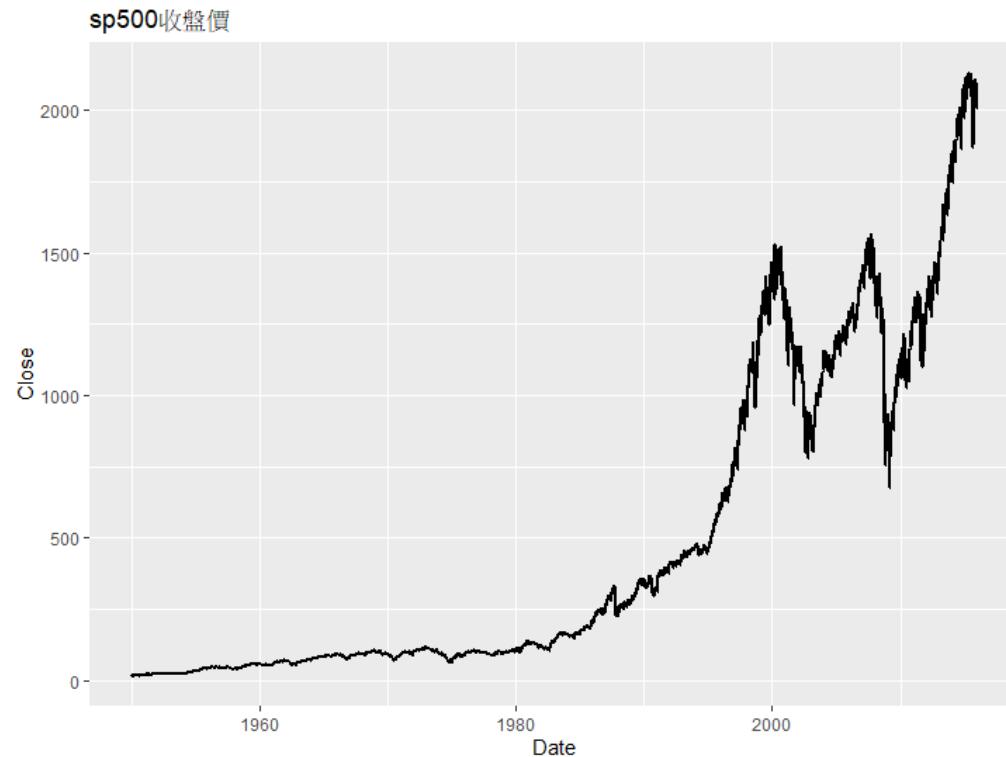


SP500 資料處理與視覺化



ggplot2, zoo 套件

- 資料集 - <https://github.com/rwepa/DataDemo/blob/master/sp500.RData>
- R 程式 - <https://github.com/rwepa/DataDemo/blob/master/sp500.R>



R demo

3. Tableau與R整合應用



Tableau 下載-方法 【★★★推薦方法】

- 選取下載 Tableau Desktop 網頁

<https://www.tableau.com/zh-tw/support/releases>

The screenshot shows the '版本資訊' (Version Information) section of the Tableau releases page. It lists five download links:

- Desktop** (highlighted with a red box and labeled 1)
- Prep**
- Server**
- Online** (highlighted with a red box and labeled 2)
- Bridge**

Each link has a dropdown icon and the version number below it: 2022.3, 2022.2, 2022.1, 2022.1, and 2022.1 respectively.

A yellow callout bubble with a black border and white text provides recommendations:

- 考量14天試用版的使用
- 安裝新版本後, 可能無法安裝舊版本
- 先選取【2022.1版本】, 到期後再選取【2022.2版本】



Tableau -英文版

- 2022-1-0 14天試用版
- 到期安裝 2022-1-1 等版本

The screenshot shows the Tableau application window with the following sections highlighted:

- Connect** (Section 1): Includes options for "Search for Data", "Tableau Server", "To a File" (Microsoft Excel, Text file, JSON file, Microsoft Access, PDF file, Spatial file, Statistical file, More...), "To a Server" (Microsoft SQL Server, MySQL, Oracle, Amazon Redshift, More...), and "Saved Data Sources".
- Open** (Section 2): Includes "Open a Workbook" and a list of recent files:
 1. 資料連結
 2. 最近檔案
 3. 教學探索
 4. 內建範例
- Discover** (Section 3): Includes "Training" (Getting Started, Connecting to Data, Visual Analytics, Understanding Tableau, More training videos...), "Resources" (Get Tableau Prep, Tableau Blueprint Assessment, Tableau Community Forums, Tableau Accelerators, Blog - Read latest post), and "More Samples".
- Sample Workbooks** (Section 4): Shows three sample workbooks: "Superstore" (a bar chart), "Regional" (a map of the United States), and "World Indicators" (a stacked bar chart).



Tableau - 中文版

The screenshot illustrates the Tableau application interface with several key areas highlighted:

- 1. 连線 (Connections):** Located in the left sidebar under the "連線" section. It includes options like "搜尋資料" (Search Data), "Tableau Server", "到檔案" (To File) with Microsoft Excel, JSON 檔案, Microsoft Access, PDF 檔案, 空間檔案, 統計檔案, and "更多..." (More...). A yellow circle highlights the "Tableau Server" option.
- 2. 開啟 (Open):** This area contains a "開啟工作簿" (Open Workbook) dialog box titled "歡迎使用" (Welcome). It states: "您的 14 天試用版已啟動。可透過以下這些很好的方法瞭解 Tableau。" (Your 14-day trial has started. You can learn about Tableau through the following great methods.) A yellow circle highlights the "開啟工作簿" button.
- 3. 探索 (Explore):** Located in the right sidebar under the "探索" section. It includes "訓練" (Training), "快速入門" (Quick Start), "連線至資料" (Connect to Data), "視覺分析" (Visual Analysis), "認識 Tableau" (Get to know Tableau), and "更多訓練影片..." (More training videos...). A yellow circle highlights the "訓練" link.
- 4. 資料連結 (Data Connection):** A large yellow callout box lists the following steps:
 1. 資料連結
 2. 最近檔案
 3. 教學探索
 4. 內建範例
- 範例工作簿 (Sample Workbooks):** Located at the bottom left, showing thumbnails for "超級市場" (Supermarket) and "世界指標" (World Metrics). A red box highlights this section.
- 更多範例 (More Examples):** Located at the bottom right, with a red box highlighting the "立即升級至 2022.2" (Upgrade to 2022.2) button.



工作簿 (Workbook)

- Tableau 使用**工作簿(Workbook)**結構，工作簿包含工作表、儀表板與故事。
 - **工作表 (Worksheet)**是一個具有資料欄位與視覺化的資料檢視。
 - 例：銷售量統計圖、利潤中心業績統計圖
 - **儀表板 (Dashboard)**是許多資料檢視的工作表，可以同步比較各種資料與圖表。
 - 例：銷售量統計圖、工廠不良率統計圖與利潤中心業績統計圖整合在儀表板。
 - 工作表和儀表板中的資料是相連的。當您修改工作表時，包含該工作表的任何儀表板也會變更，反之亦然。
 - **本文或故事 (Story)**是一系列共同目的以傳達資訊的工作表。
 - 例：建立故事以講述資料，示範決策與結果的關係或一個具有吸引力的案例。



Tableau二大欄位

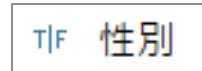
- Tableau二大欄位:
 - 維度 Dimensions: 字串, 布林 
 - 量值 Measures: 數值, 例: 實數, 整數, 使用總和(SUM)方式呈現
- Tableau支援四大資料型態, Create Calculated Field
 - REAL 實數
 - INTeger 整數
 - BOOlean 布林 (TRUE/FALSE)
 - STRING data 字串



Tableau + R

- 參考: <http://rwepa.blogspot.com/2019/03/tableau-r.html>
- 1. 下載 iris.csv
- 2. 安裝 Rserve 套件
- 3. 啟動 Rserve 套件
- 4. 開啟 Tableau \ Tableau 與R連結設定
- 5. 建立計算欄位 (Create Calculated Field)
- 6. 建立視覺化集群結果
- 7. 將計算欄位拖拽至 [Color 色彩]



Rserve 套件

Rserve: Binary R server

Rserve acts as a socket server (TCP/IP or local sockets) which allows binary requests to be sent to R. Every connection has a separate workspace and working directory. Client-side implementations are available for popular languages such as C/C++ and Java, allowing any application to use facilities of R without the need of linking to R code. Rserve supports remote connection, user authentication and file transfer. A simple R client is included in this package as well.

Version: 1.8-10
Depends: R (\geq 1.5.0)
Suggests: [RSclient](#)
Published: 2021-11-25
Author: Simon Urbanek
Maintainer: Simon Urbanek <Simon.Urbanek at r-project.org>
License: [GPL-2](#) | file [LICENSE](#)
URL: <http://www.rforge.net/Rserve/>





Rserve 套件 (續)

- 啟動 R 或 RStudio
- 安裝並載入 Rserve 套件
- 執行 Rserve() 函數

```
install.packages("Rserve")
library(Rserve)
Rserve()
```

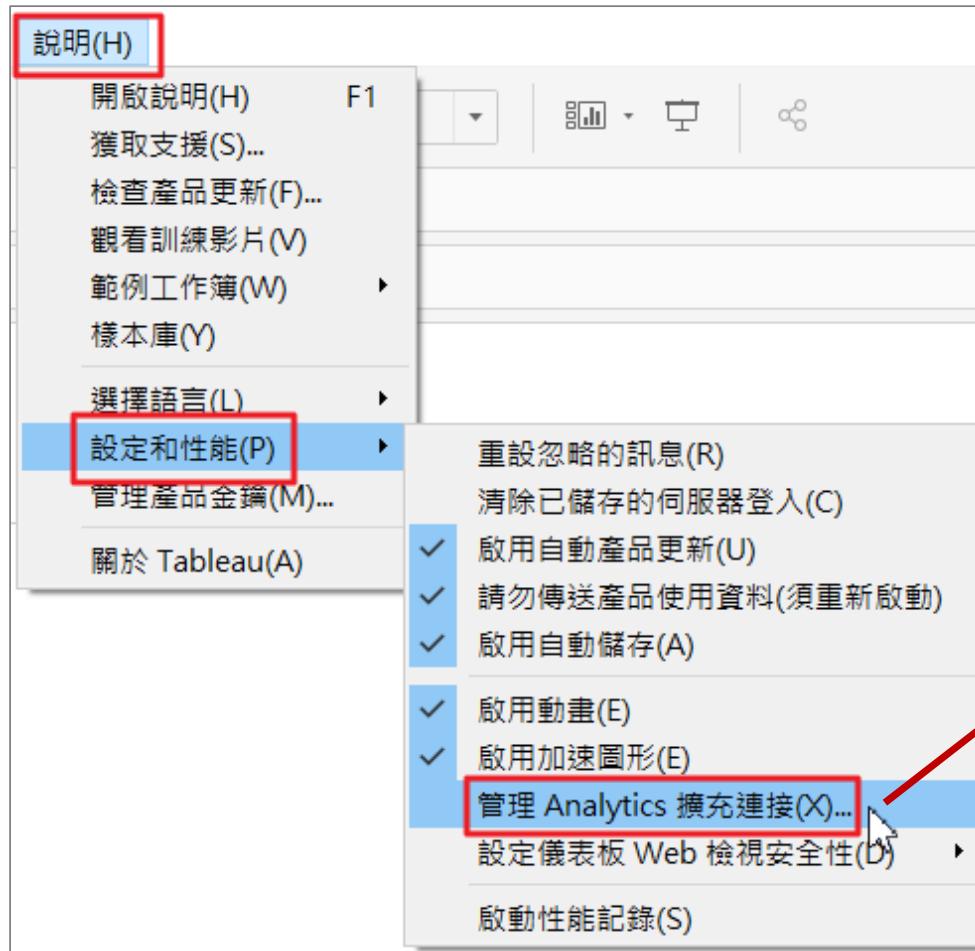
```
> install.packages("Rserve")
Installing package into 'C:/Users/asus/AppData/Local/R/win-library/4.2'
(as 'lib' is unspecified)
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.2/Rserve_1.8-10.zip'
Content type 'application/zip' length 4980938 bytes (4.8 MB)
downloaded 4.8 MB

package 'Rserve' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
      C:\Users\asus\AppData\Local\Temp\RtmpEH2ihK\downloaded_packages
> library(Rserve)
>
> Rserve()
Starting Rserve...
"C:\Users\asus\AppData\Local\R\WIN-LI~1\4.2\Rserve\libs\x64\Rserve.exe"
>
```



Tableau 與 R 連結設定





管理 Analytics 擴充功能連線

管理 Analytics 擴充功能連線

新增 RServer 連線

1 需要 SSL
localhost

2 使用使用者名稱和密碼登入
使用者名稱
密碼

3





Tableau + R 完成圖

The screenshot shows a Tableau dashboard titled "Tableau_with_R". The data source is "iris". The visualization is a scatter plot with "Petal.Length" on the x-axis and "Petal.Width" on the y-axis. The data points are colored by cluster: 1 (blue), 2 (orange), and 3 (green). A legend on the right side shows the color mapping for the clusters.

In the bottom-left corner, there is a "Cluster" dialog box with the following content:

```
沿著表(横向)計算結果。
SCRIPT_INTC('set.seed(168);
result <- kmeans(data.frame(.arg1,.arg2,.arg3,.arg4), 3);
result$cluster;');
SUM([Sepal.Length]),SUM([Sepal.Width]),SUM([Petal.Length]),SUM([Petal.Width]))
```

The "Marks" card in the center shows the following settings:

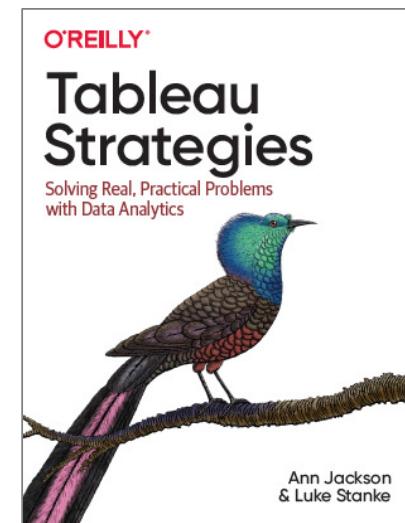
- 标记 (Marks): 圆 (Circle)
- 色彩 (Color): 使用 "Cluster" 作为颜色
- 大小 (Size): 不显示
- 标签 (Label): 不显示
- 详细资料 (Details): 不显示
- 工具提示 (Tool Tip): 不显示
- 集群 (Clusters): 高亮显示

The "Rows" shelf contains "Petal.Length" and the "Columns" shelf contains "Petal.Width".



參考資料

- RWEPA: <http://rwepa.blogspot.com/>
- Tableau 與 R 整: <http://rwepa.blogspot.com/2019/03/tableau-r.html>
- Youtube-Tableau: <https://www.youtube.com/user/tableausoftware>
- 大數據時代必學的超吸睛視覺化工具與技術：Excel+Tableau成功晉升資料分析師, 彭其捷, 博碩出版.
- Ann Jackson and Luke Stanke, Tableau Strategies: Solving Real, Practical Problems with Data Analytics, 2021.
<https://www.amazon.com/Tableau-Strategies-Practical-Problems-Analytics/dp/149208008X>





謝謝您的聆聽

Q & A



李明昌

alan9956@gmail.com

<http://rwepa.blogspot.tw/>