

R, RStudio 簡介

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

<https://www.youtube.com/@alan9956>

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

alan9956@gmail.com

大綱

- 1.RWEPA簡介
- 2.資料分析架構
- 3.R簡介與安裝
- 4.RStudio簡介與安裝

1. RWEPA 簡介

RWEPA簡介: <http://rwepa.blogspot.com/>

- 姓名：李明昌 (ALAN LEE)
- 現職：中華R軟體學會 理事
臺灣資料科學與商業應用協會 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335餘場，3072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

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**

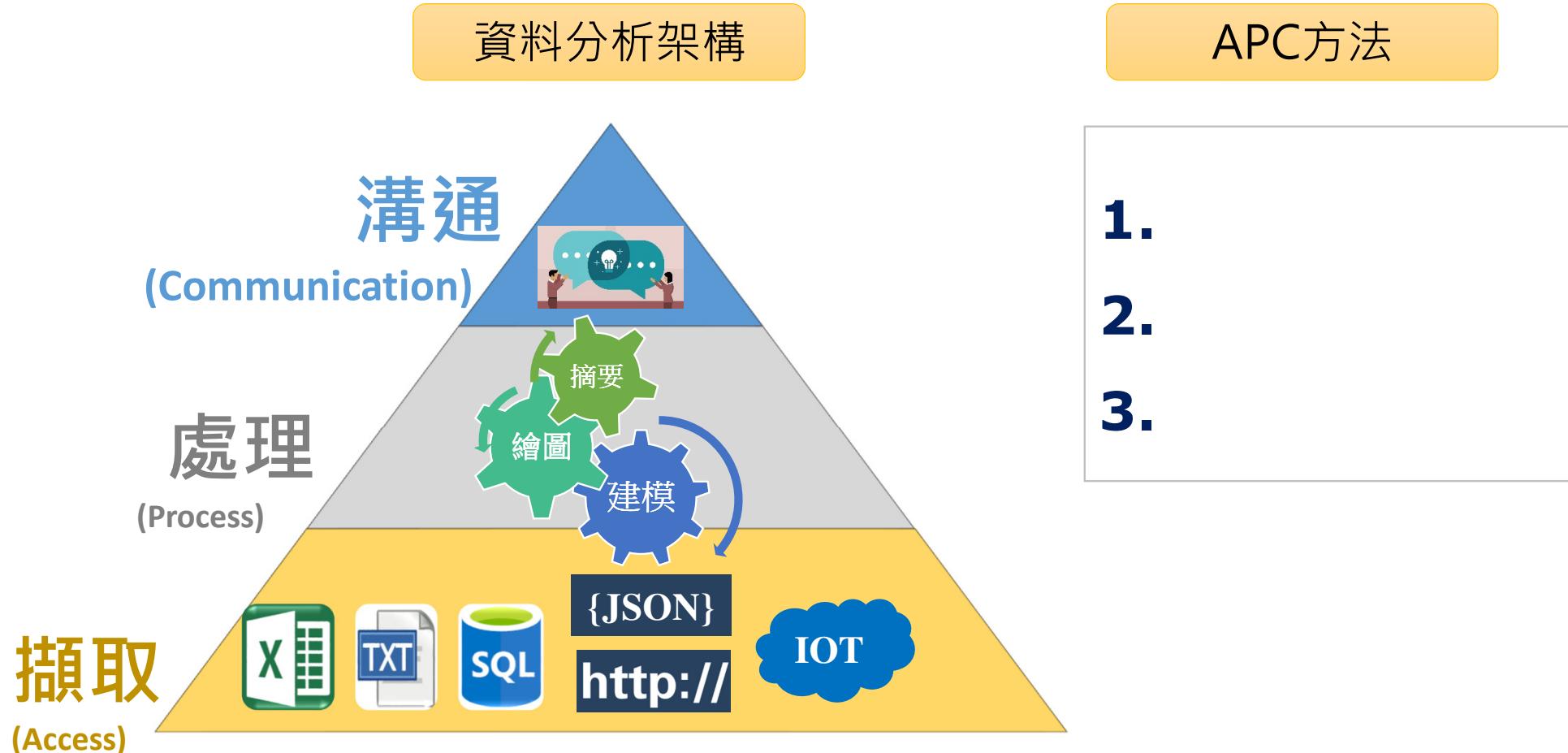
1. R · RStudio工具操作
2. 非監督式學習商業預測
3. 監督式學習商業預測
4. 財金資料預測應用

- **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. 資料分析架構

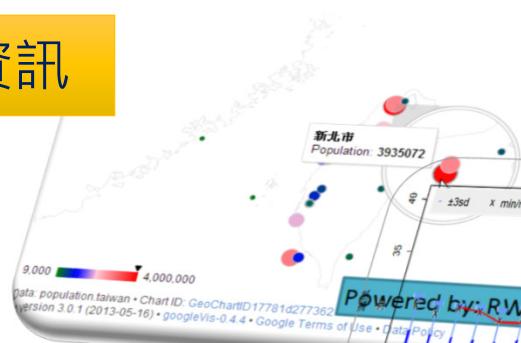
★★★資料分析架構暨APC方法



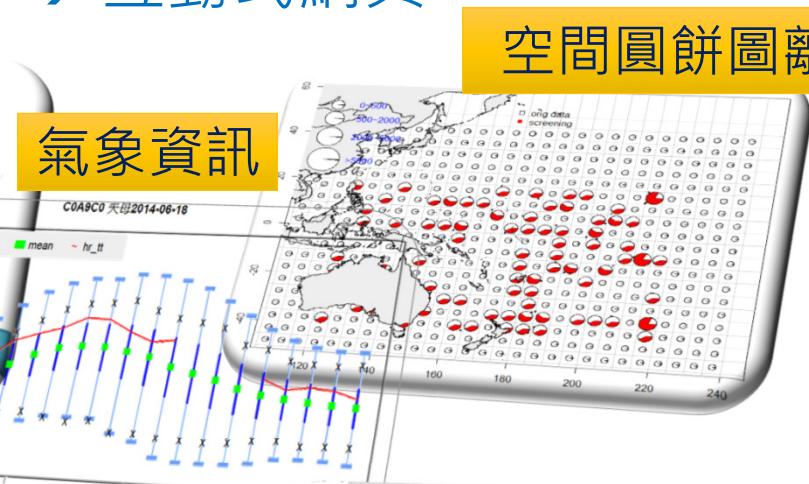
資料分析/視覺化應用

R + shiny → 互動式網頁

地理資訊

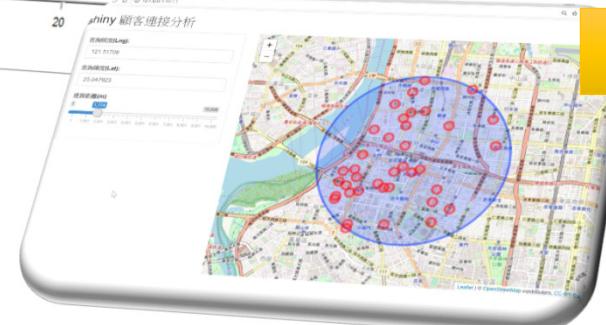


氣象資訊



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(14,328個檔案)

網頁呈現



客製化選單

R統計運算

14,328 個檔案, 18 個資料夾

動態繪圖

保險預測模型

The screenshot shows the iinsurance interactive analysis platform version v.16.3.24. The top navigation bar includes links for '檔案上傳', '資料處理', '統計圖表', '模型評估', and '預測模型'. A red box highlights the '機率模型閥值' (Probability Model Threshold) slider set at 0.1. Another red box highlights the '檢視結果' (View Results) button. A yellow callout bubble labeled '機率模型閥值調整' (Probability Model Threshold Adjustment) points to the slider. A yellow callout bubble labeled '預測結果' (Prediction Results) points to the results table.

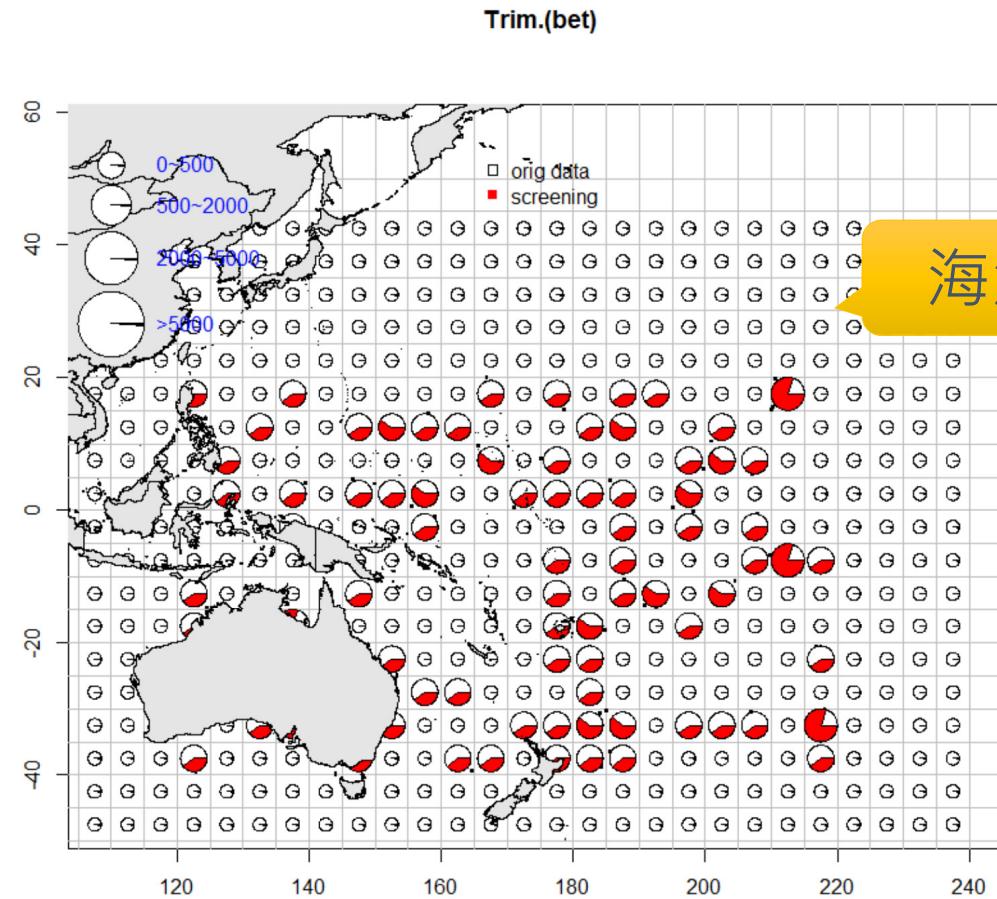
性別	女性	車輛種類	私家車	曝露風險		無索償折扣	被保險人年齡	私家車 一車齡 0	私家車 一車齡 1	私家車 一車齡 2	私家車 車齡 0_1_2 組合	車齡 0_1_2 組合	預測機率	理賠		
				曝露風險對數	對數											
M	0	A	1	0.9144422	-0.08944106	50	4	1	0	0	1	0	2	0.1069	有	
M	0	A	1	0.8158795	-0.20348856	20	4	0	0	1	1	2	2	0.1441	有	
3	M	0	A	1	0.8377823	-0.17699695	50	3	0	0	1	1	2	2	0.1866	有
4	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

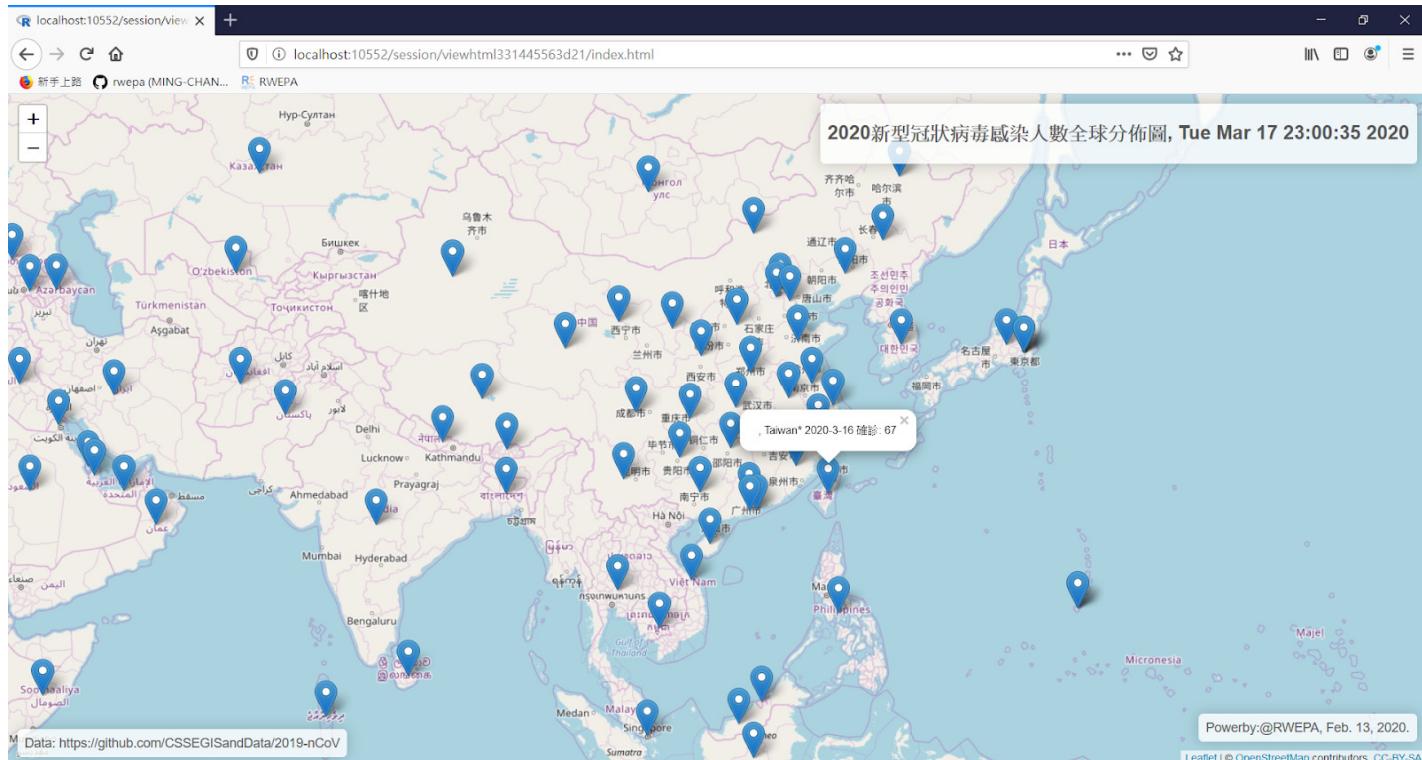
Previous 1 2 Next

127.0.0.1:6177/#tab-9487-2

空間圓餅圖離群值分析



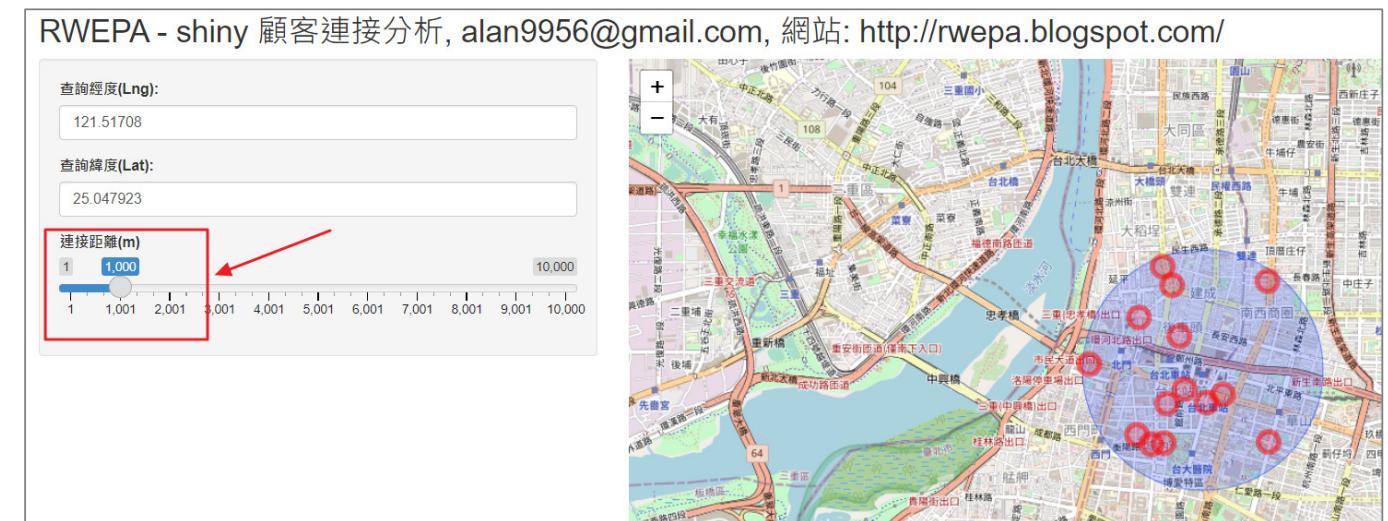
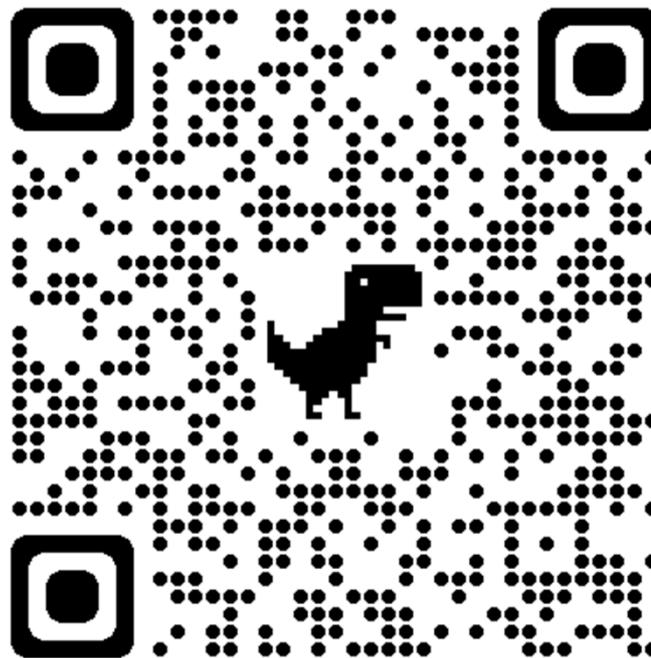
2020新型冠狀病毒視覺化



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

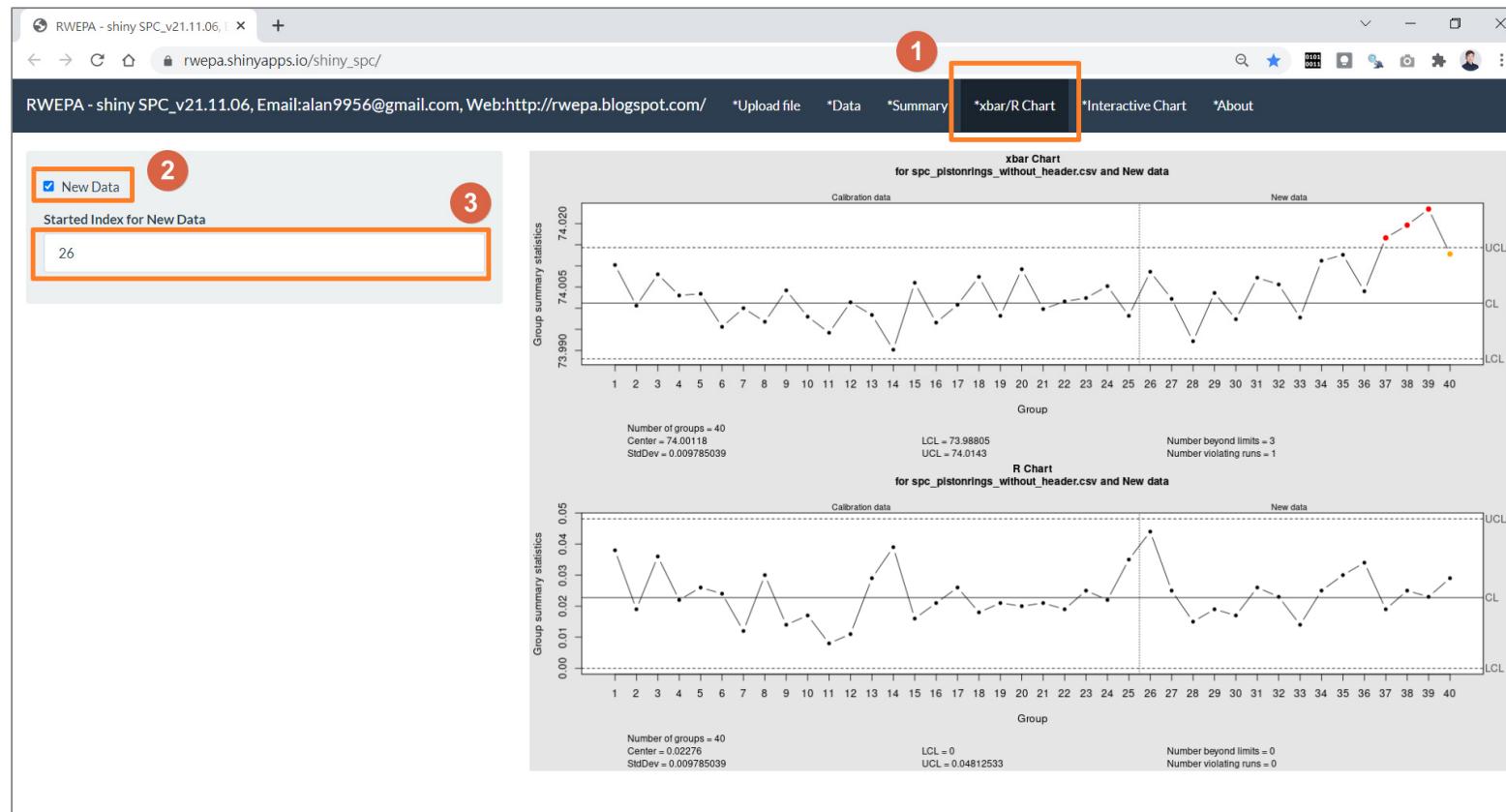
RWEPA - shiny 顧客連接分析

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



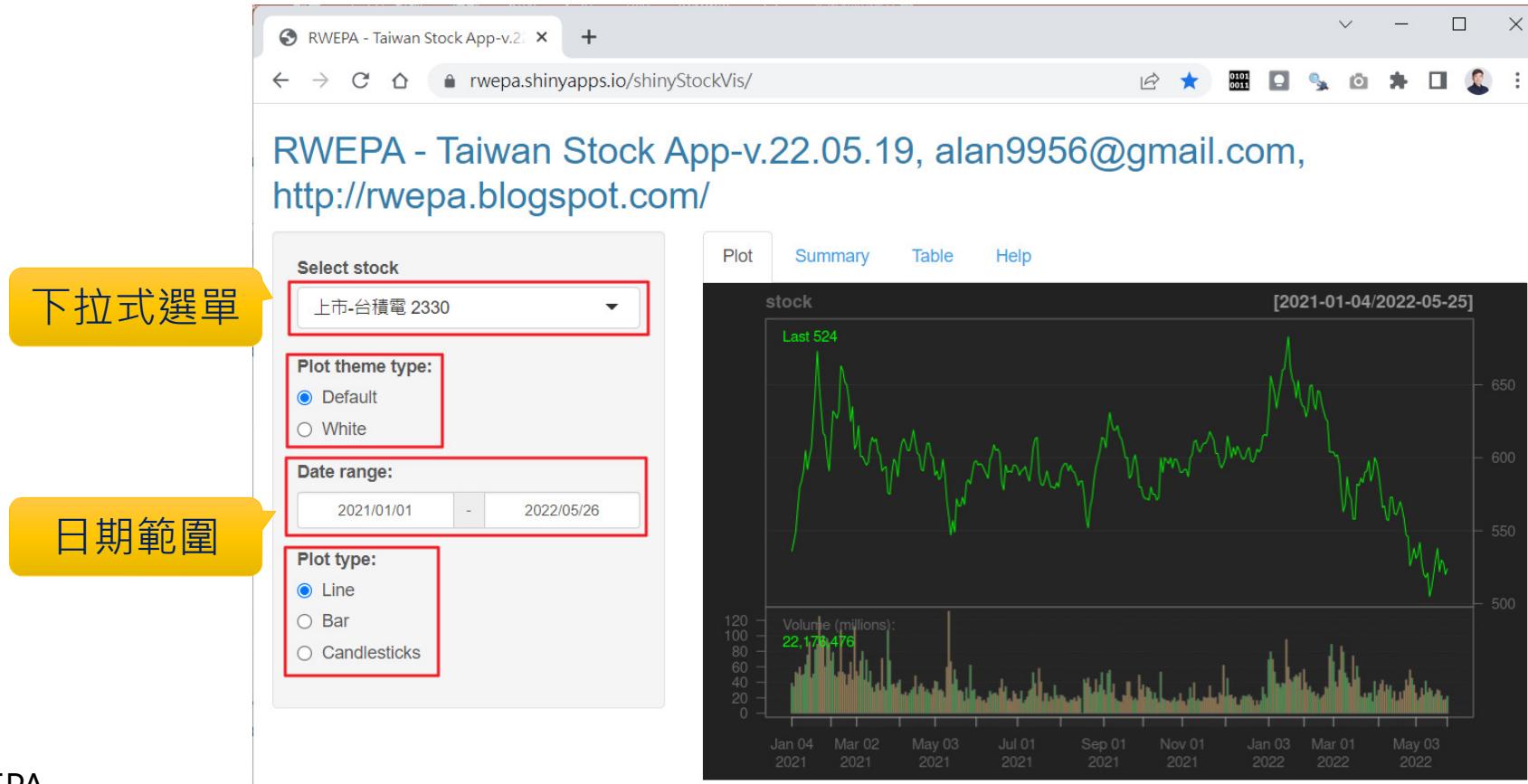
品質管制圖(quality control chart)應用

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

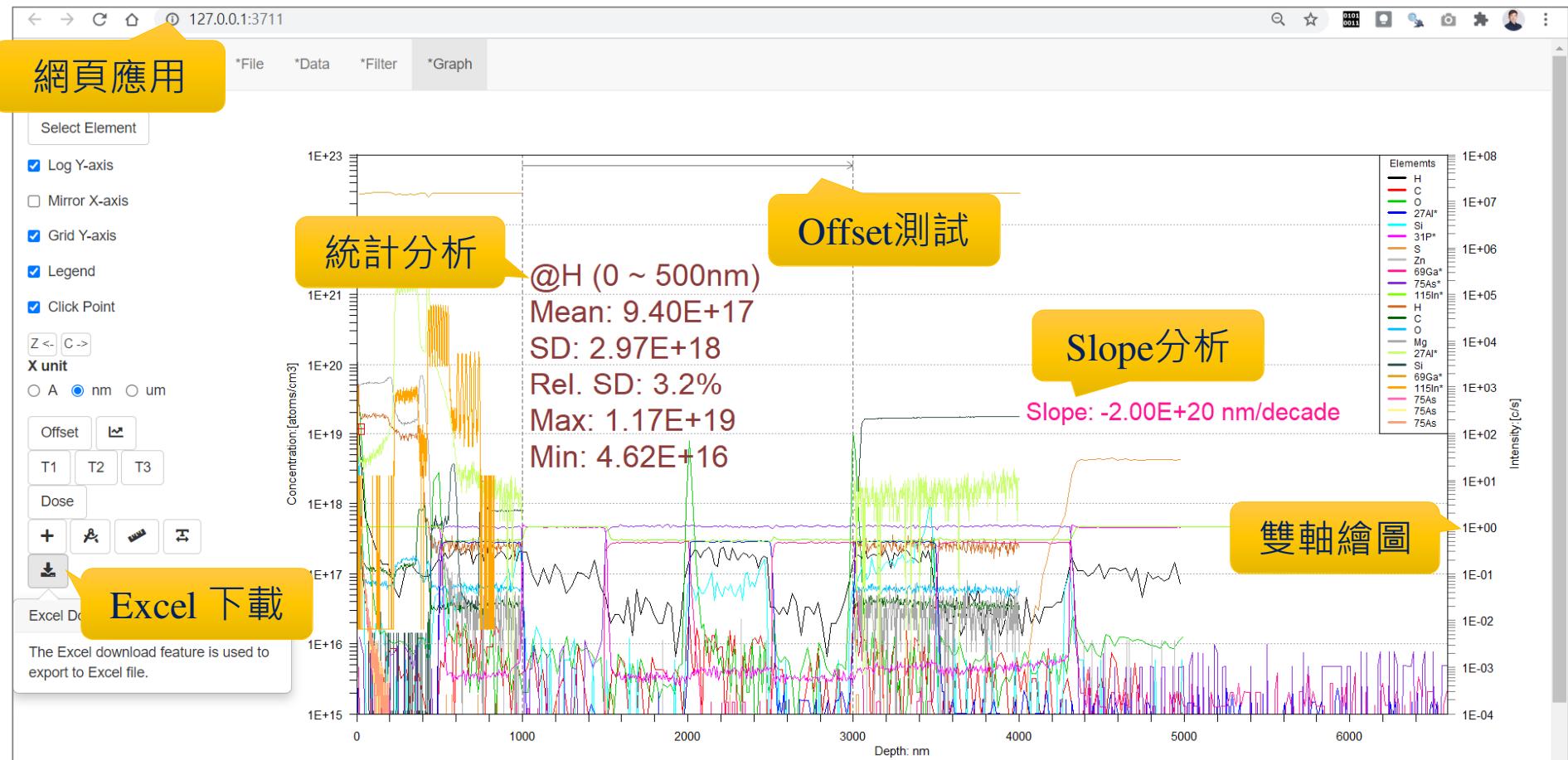


Taiwan Stock App

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

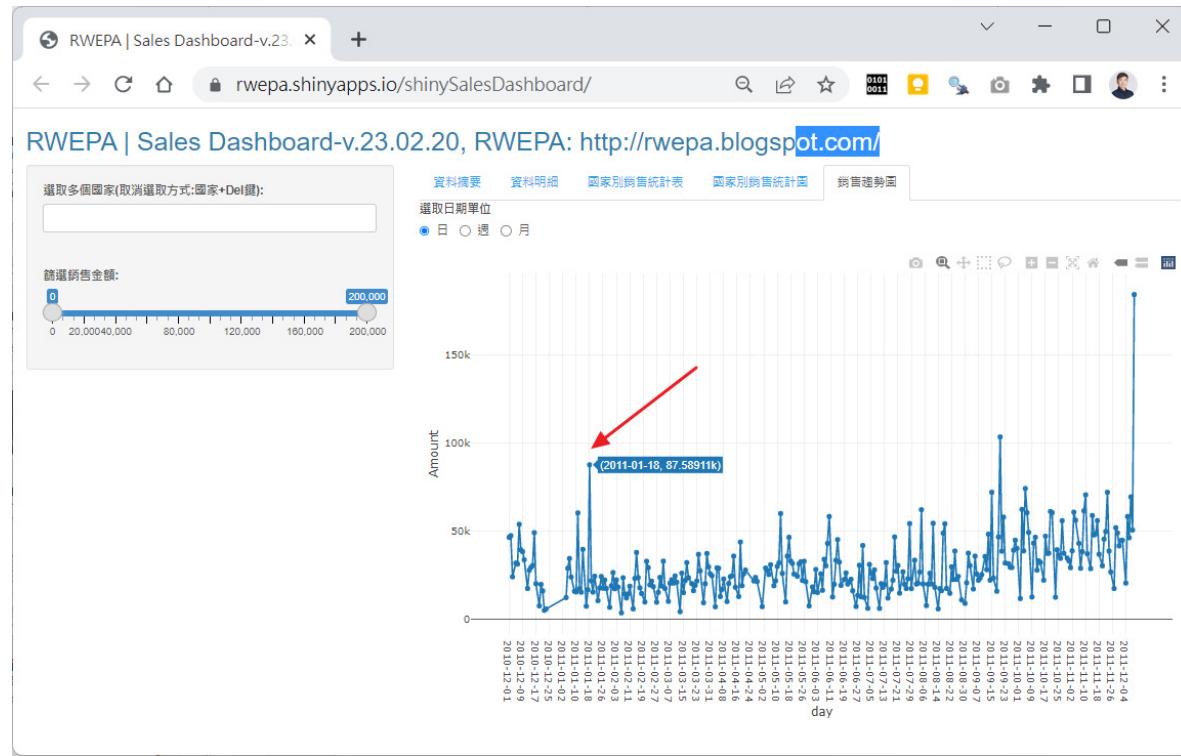


離子資料分析與視覺化應用



shiny銷售儀表板

- Shiny: <https://rwepa.shinyapps.io/shinySalesDashboard/>
- Ubuntu Shiny Server: <https://shiny.rwepa.net/shiny-sales/>



shiny銷售儀表板 (續)

- RWEPA | shiny企業實務應用 第4集-shiny銷售儀表板【中文字幕】
- YouTube:
<https://youtu.be/4GgZlf8heQk>
- Slide:
<https://rwepa.quarto.pub/r-shiny-04-sales-project/>
- Code:
https://github.com/rwepa/business_analytics/tree/main/r-shiny-04-sales-project



小明算命師

- Ubuntu Shiny Server: <https://shiny.rwepa.net/shiny-hr-teller/>



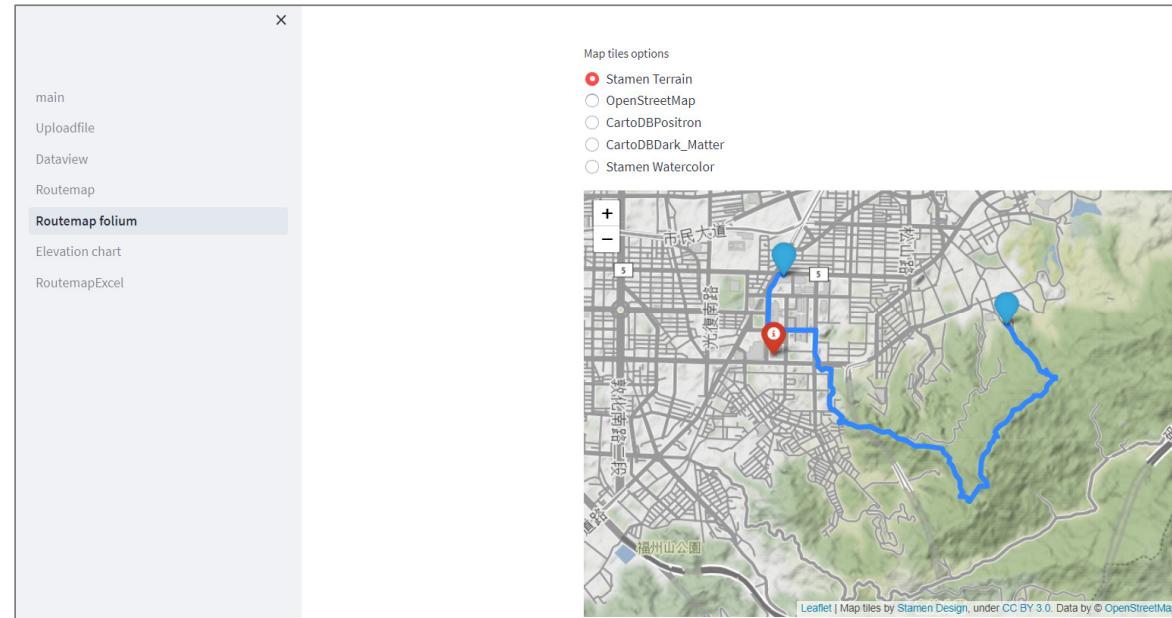
1 / 22

小明算命師 (續)

- RWEPA | shiny企業實務應用 第6集-小明算命師(下) - 第1季完結篇
- YouTube:
<https://youtu.be/rrD6KV3eV-w>
- Slide:
<https://rwepa.quarto.pub/r-shiny-06-hr-teller/>
- Code:
https://github.com/rwepa/business_analytics/tree/main/r-shiny-06-hr-teller

登山路線視覺化分析平台 (Python + Streamlit)

- YouTube: https://youtu.be/-_zghs2qrlg
- 系統展示: <https://rwepa-climb.streamlit.app/>



Power BI資料視覺化 - RFM分析

- 🌸 YouTube: <https://youtu.be/Lkr9HmzLTtg>
- 🌸 Link: <http://rwepa.blogspot.com/2023/07/rwepa-rfm-analysis-using-power-bi.html>

Customer Segmentation Using RFM Analysis, 2023



Calendar



最近消費 (recency) :
顧客上次消費時間愈近，用戶價值愈大。

消費頻率 (frequency) :
顧客在一段時間中，總購買次數，購買頻率愈高，用戶價值愈大。

消費金額 (monetary) :
顧客總消費金額，消費金額愈高，用戶價值愈大。

Author : Ming-Chang Lee
YouTube : <https://www.youtube.com/@alan9956>
RWEPA : <http://rwepa.blogspot.tw/>
GitHub : <https://github.com/rwepa>
Email : alan9956@gmail.com

RFM分析 x RFM視覺化分析 RECENTY FREQUENCY Monetary +

Tableau 資料視覺化

- Tableau Talks
 - <https://github.com/rwepa/Talks>
- Tableau資料分析與視覺化工具實作教師工作坊(初階)
 - https://github.com/rwepa/Talks/blob/main/tableau_tutorial_basic.pdf
- Tableau資料分析與視覺化工具實作教師工作坊(進階)
 - https://github.com/rwepa/Talks/blob/main/tableau_tutorial_advanced.pdf
- Tableau與R語言實務應用
 - https://github.com/rwepa/Talks/blob/main/tableau_r.pdf
- Tableau與資料庫實務應用
 - https://github.com/rwepa/Talks/blob/main/tableau_mysql.pdf
- Tableau智慧製造應用
 - <https://public.tableau.com/app/profile/ming.chang.lee/vizzes>

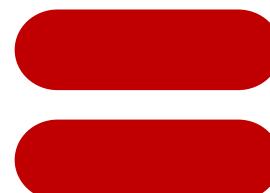
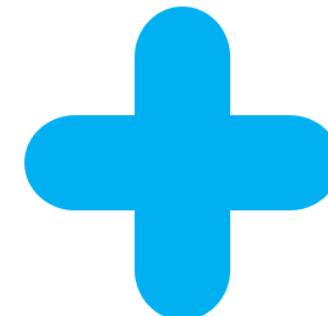
如何學習 R?

- 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

```
> 尋找答案 <- list(方法1 = c("同事", "同學", "朋友等"),
+                         方法2 = "Google",
+                         方法3 = "alan9956@gmail.com")
> print(尋找答案)
$方法1
[1] "同事"    "同學"    "朋友等"
$方法2
[1] "Google"
$方法3
[1] "alan9956@gmail.com"
```



學習目標



3.R簡介與安裝

認識 R

- 1976 - 貝爾實驗室 John Chambers, Rick Becker, and Allan Wilks 將 S 語言研發為 R。
- 1993 - Ross Ihaka and Robert Gentleman, 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
 - 2024年2月 - R 4.3.3



R-八大功能



R官方網頁



[Home]

Download

CRAN

下載

R Project

About R

Logo

Contributors

What's New?

Reporting Bugs

Conferences

Search

Get Involved: Mailing

Lists

Get Involved:

The R Project for Statistical Computing

Getting Started

統計計算 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.

News

- **R version 4.3.3 (Angel Food Cake) prerelease versions** will appear starting Monday 2024-02-19. Final release is scheduled for Thursday 2024-02-29.
- **Registration for useR! 2024** has opened with early bird deadline March 31 2024.
- **R version 4.3.2 (Eye Holes)** has been released on 2023-10-31.

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-4.3.1-win.exe]

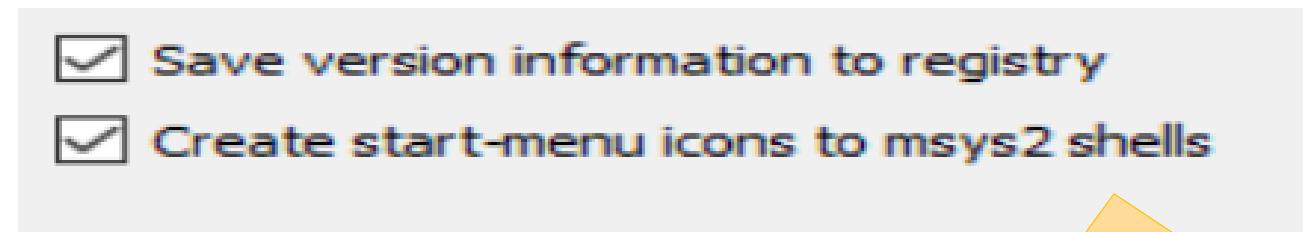


The diagram illustrates the process of downloading R for Windows. On the left, a list of subdirectories is shown: Subdirectories: base, contrib, old contrib, and Rtools. The base item is highlighted with a red rectangular box and has a red arrow pointing to it from the text above. A yellow curved arrow points from this list to a larger box on the right. Inside this box, there are three blue hyperlinks: Download R-4.3.1 for Windows (79 megabytes, 64 bit), README on the Windows binary distribution, and New features in this version.

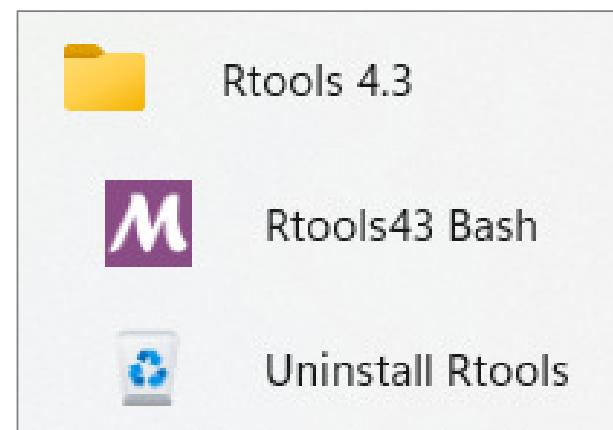
- R安裝路徑: 保留原路徑,不要修改
- 安裝參考說明, 2006
https://github.com/rwepa/DataDemo/blob/master/windows_intall_R.pdf

Rtools 下載與安裝

- Rtools for Windows: 一定要保留原路徑 C:\rtoolsXX



安裝完成：程式集 \ Rtools xx



✓ 安裝 R

✓ 安裝 Rtools



R Manuals

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.		
R Internals : a guide to the internal structures of R and coding standards 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 lately).

English Documents

Documents with more than 100 pages:

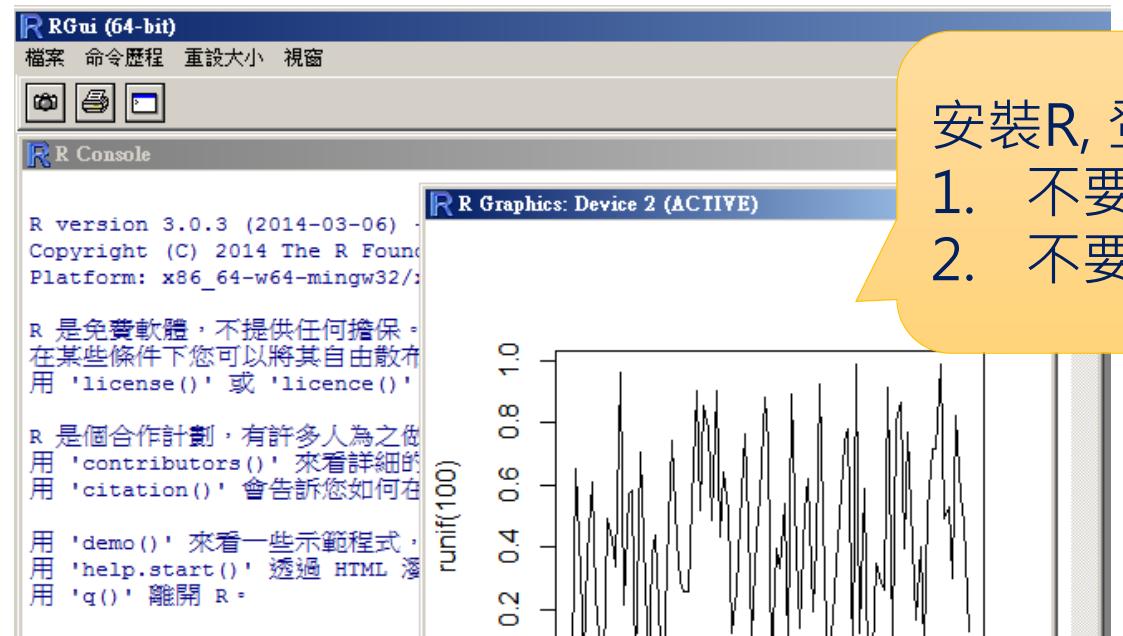
- “**Visual Statistics. Use R!**” by Alexey Shipunov ([PDF](#), 2016-06-06, 301 pages) are accessible from [Alexey Shipunov's English R page](#).
- “**Using R for Data Analysis and Graphics - Introduction, Examples and Commentary**” by John Maindonald ([PDF](#), data sets and scripts are available at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



實作練習

R 執行畫面



The screenshot shows the RGui (64-bit) interface. The title bar says "RGui (64-bit)". The menu bar includes "檔案", "命令歷程", "重設大小", and "視窗". The toolbar has icons for file operations. The "R Console" window displays the R startup message and a plot command. The "R Graphics: Device 2 (ACTIVE)" window shows a line plot of 100 random uniform numbers.

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

demo(graphics)

demo(persp)

大小寫須一致

參考文獻:
citation()

安裝R, 登入名稱:

1. 不要使用空格
2. 不要使用中文字型

R 功能表

檔案



編輯



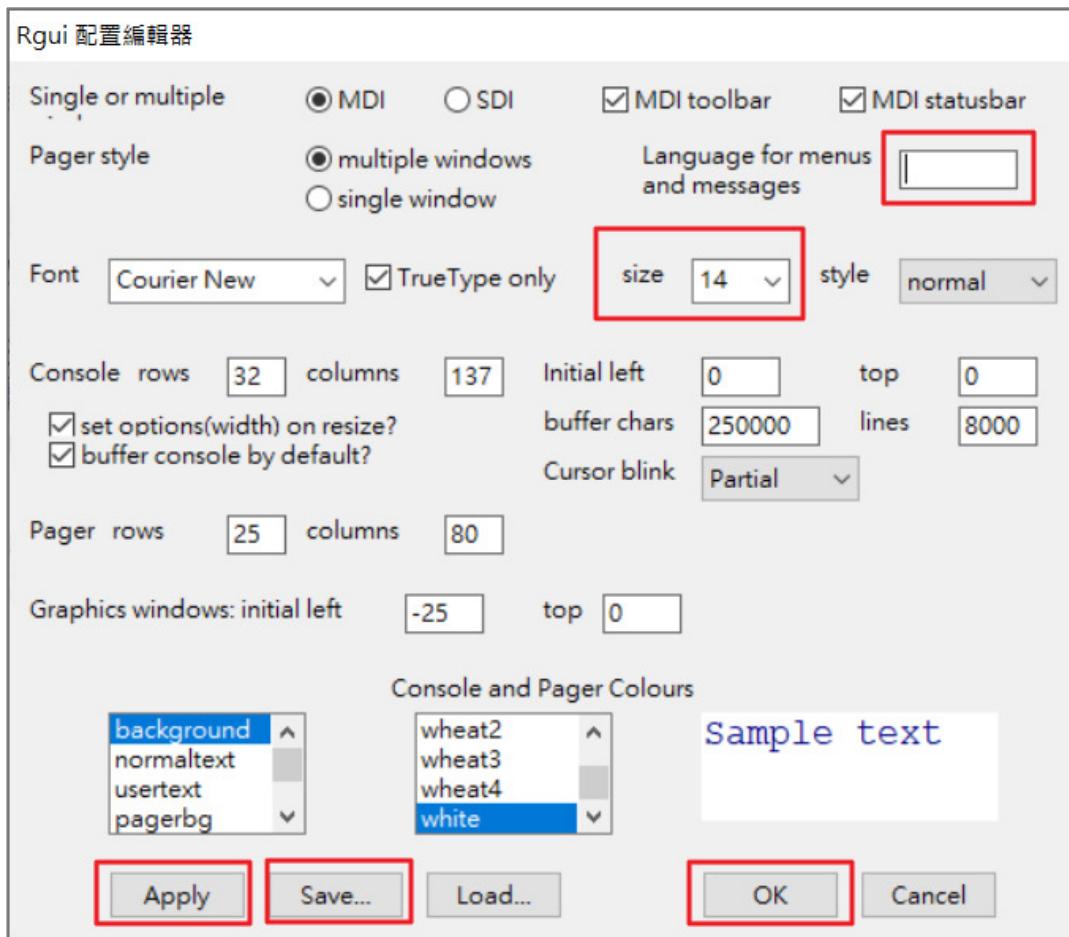
現行目錄 getwd()

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

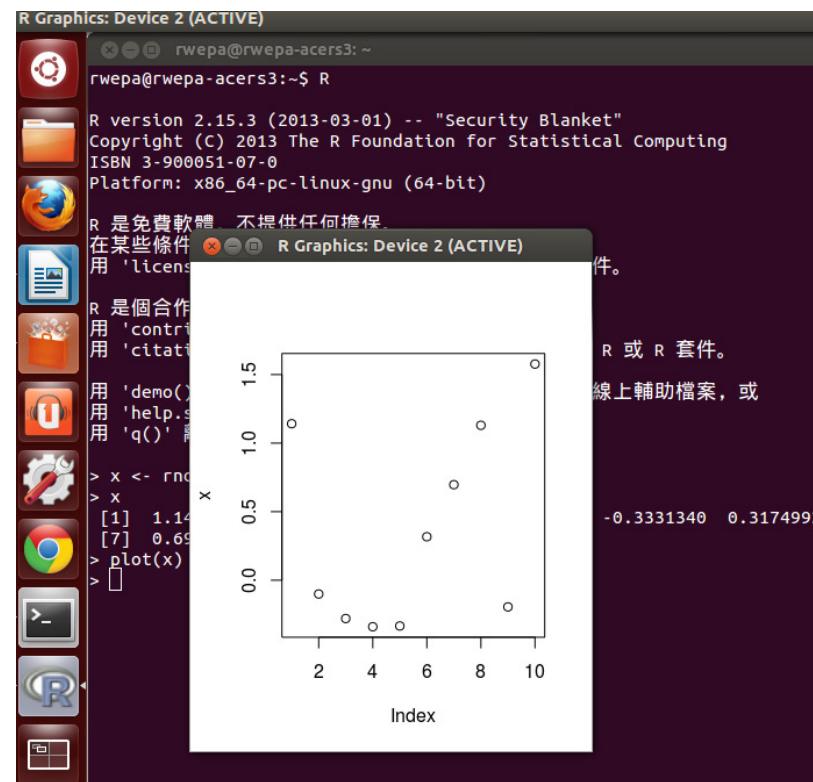
R for Mac

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



R for Ubuntu

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



新增R檔案練習

The screenshot shows the RGui interface with three main windows:

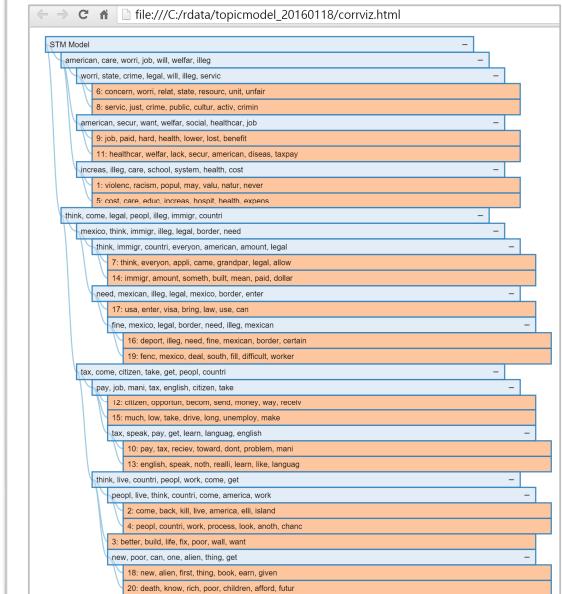
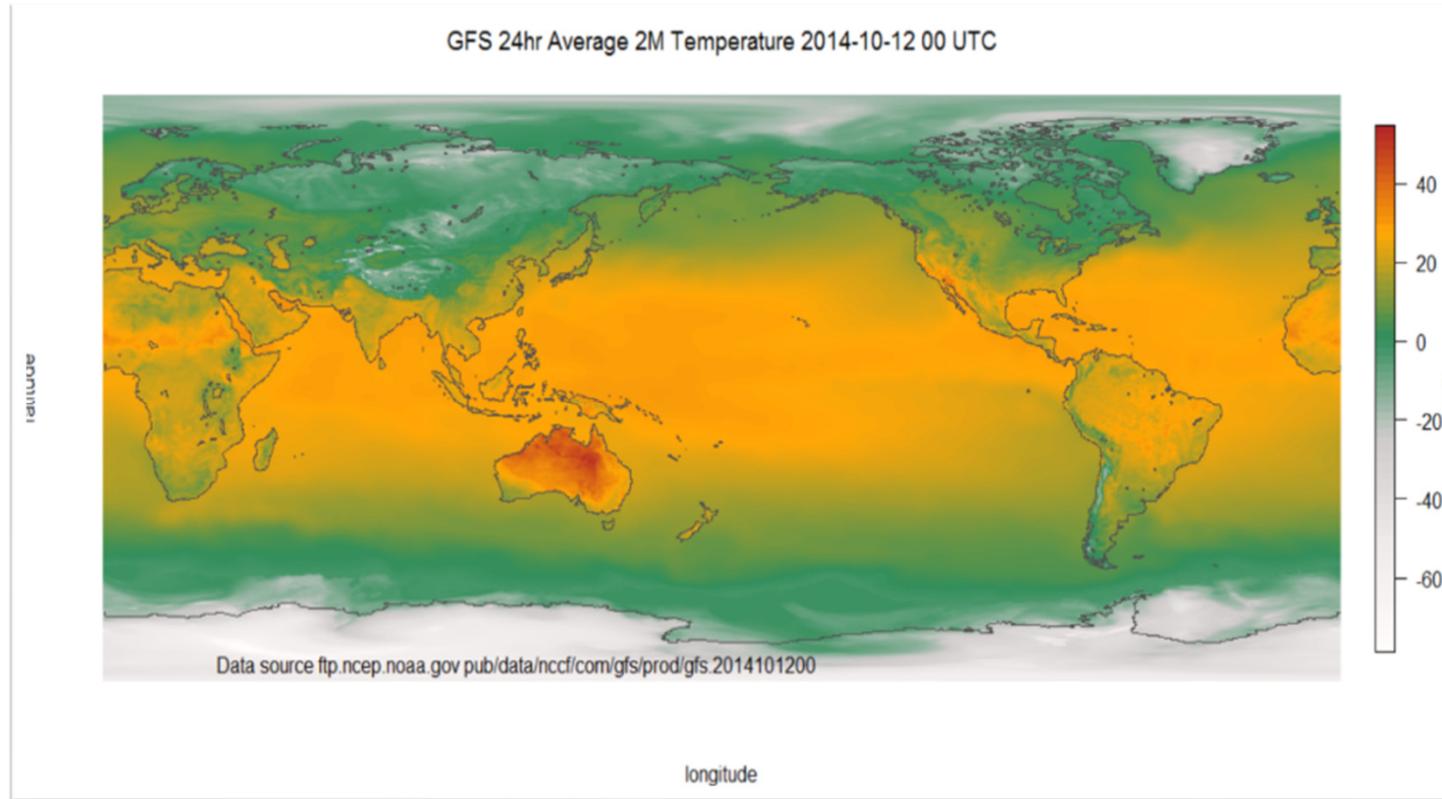
- R Editor (Left):** Displays R code for generating a scatter plot matrix. The code includes plotting uniform random numbers, creating a vector of normal random numbers, and using the `pairs` function to create a scatter plot matrix for the Iris dataset, titled "RWEPA-iris資料集散佈圖矩陣".
- R Console (Bottom Left):** Shows the R command line with partial code execution, including the creation of a scatter plot.
- R Graphics (Bottom Right):** Displays a scatter plot matrix for the Iris dataset, titled "RWEPA-iris資料集散佈圖矩陣". The matrix contains 15 plots showing the relationships between Sepal.Length, Sepal.Width, Petal.Length, and Petal.Width across three species (Setosa, Versicolor, Virginica).

A yellow callout box highlights the following steps:

- 步驟1: 選取程式碼
- 步驟2: 按 Run line or selection 或 Ctrl + R
- 步驟3: 按 File \ Save \ MyFirstProject.R

4.RStudio簡介與安裝

整合式開發環境 - RStudio



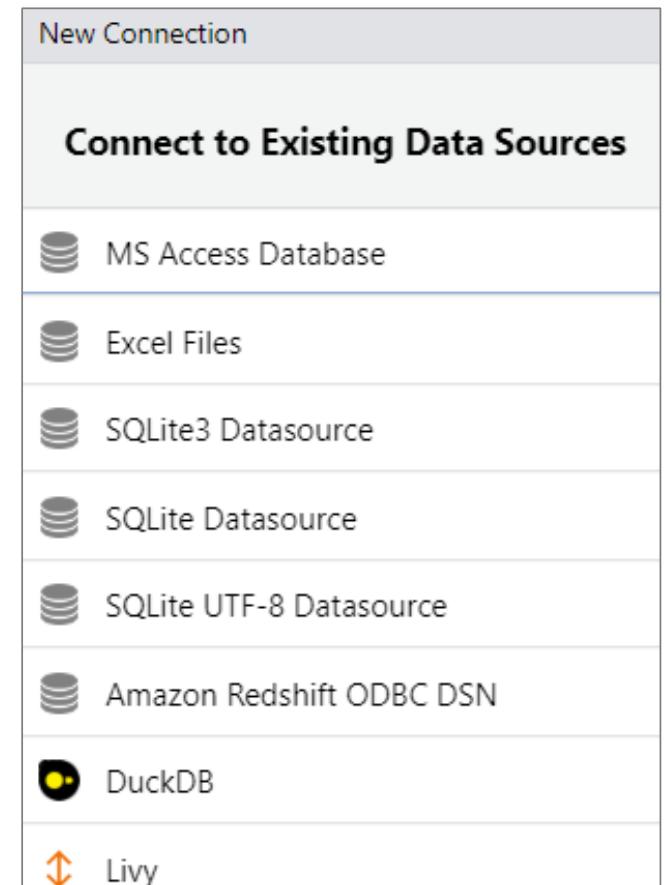
主題模型

視覺化應用

(全球2M氣溫圖)

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
- 整合連接資料庫: SQL, Spark
- 整合R套件: shiny, rmarkdown, Quarto
- 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

• <http://www.rstudio.com/>



<https://posit.co/>

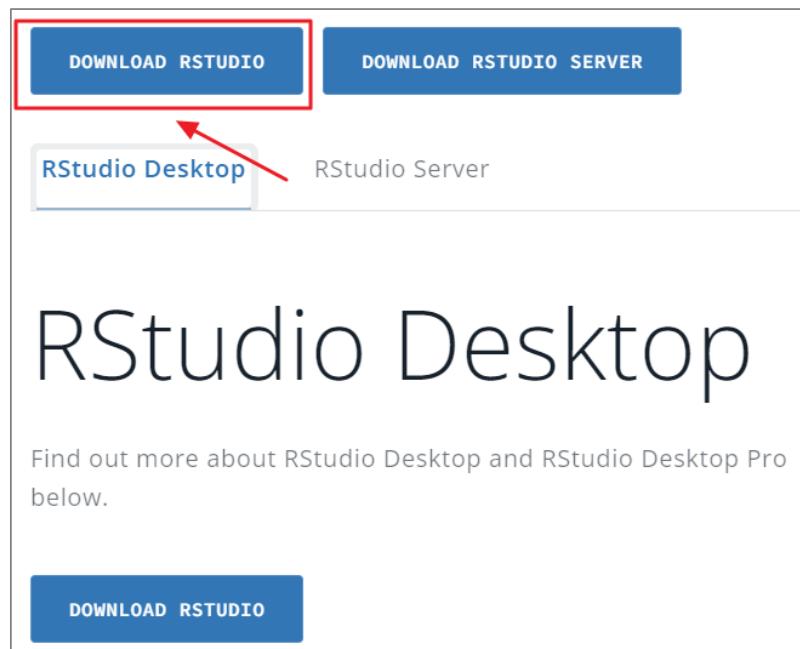
The screenshot shows the posit.co homepage. At the top, there's a navigation bar with links: PRODUCTS (highlighted with a red box), SOLUTIONS, LEARN & SUPPORT, EXPLORE MORE, PRICING, and a search icon. To the right of the search icon is a blue button labeled "DOWNLOAD RSTUDIO" with a red box around it. A red arrow points from the crossed-out URL above to this button. Below the navigation bar, the main content area features a large heading "Deployment made easy" in a large, dark font. Underneath, a paragraph explains the deployment capabilities: "Deploy all of your work, including Shiny, Streamlit, and Dash applications. Models. Quarto documents. Jupyter Notebooks. Reports. Dashboards. Even APIs. With customizable access controls and authentication options that make IT happy." To the right of the text, there's a photograph of a person working at a computer, with a monitor displaying a Shiny dashboard for "Penguin Penguins dataset".

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' row to the 'DOWNLOAD' button under the 'Desktop' section. A yellow callout bubble labeled '免費版' (Free Version) is positioned over the 'Desktop' section. Three red arrows point from the 'Server' and 'Workbench' sections towards the bottom icons representing a smartphone, tablet, and desktop monitor.

RSTUDIO-2023.09.0-463.EXE 下載

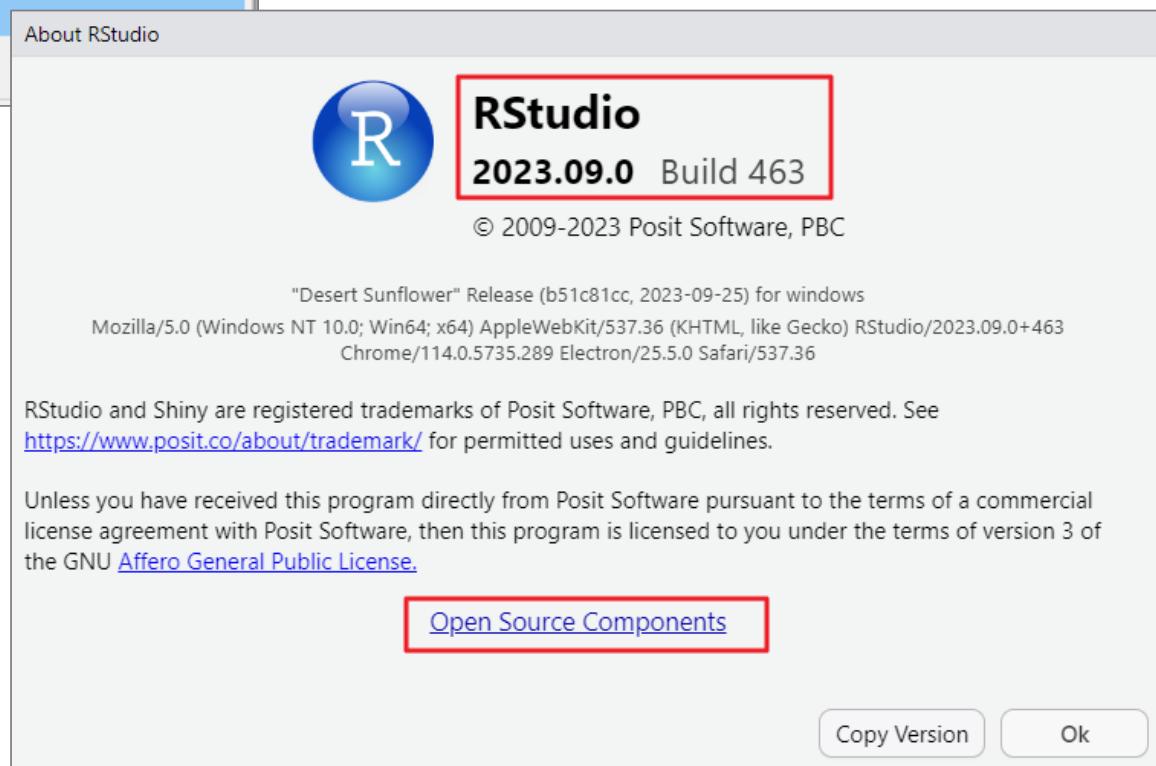
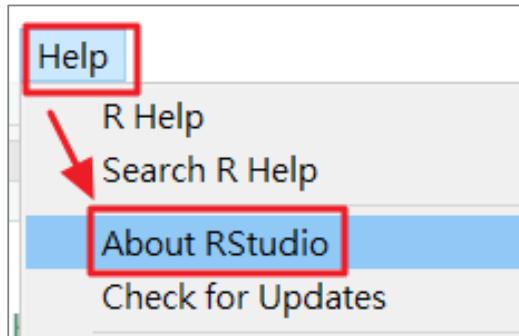


OS	Download
Windows 10/11	RSTUDIO-2023.09.0-463.EXE
macOS 11+	RSTUDIO-2023.09.0-463.DMG
Ubuntu 20/Debian 11	RSTUDIO-2023.09.0-463-AMD64.DEB
Ubuntu 22	RSTUDIO-2023.09.0-463-AMD64.DEB

RStudio 安裝

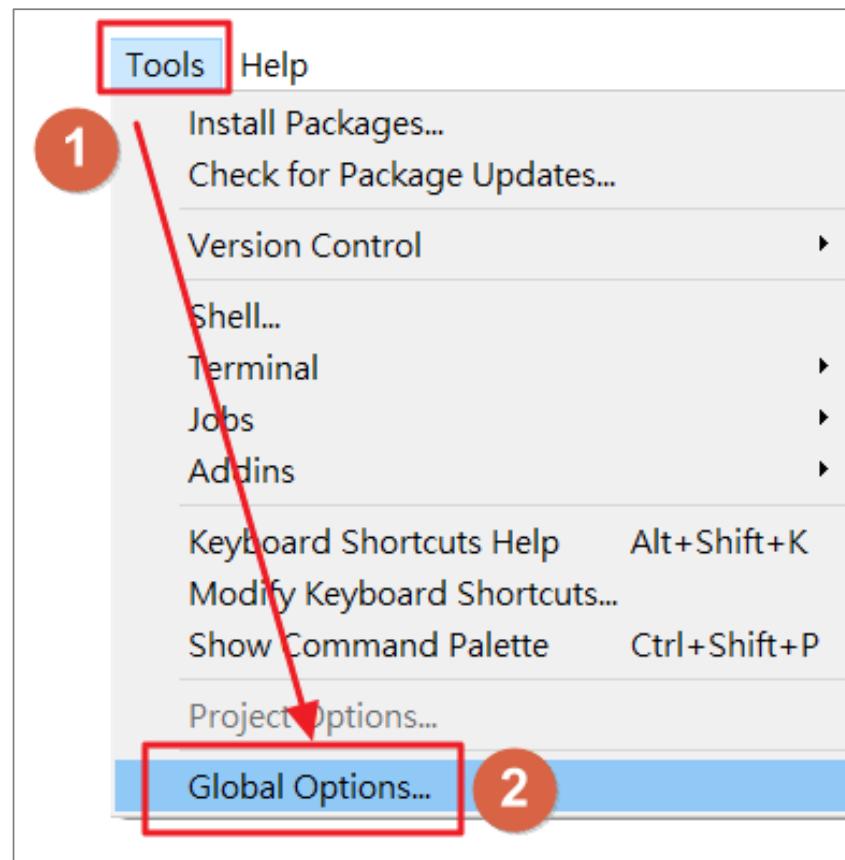


RStudio 版本訊息

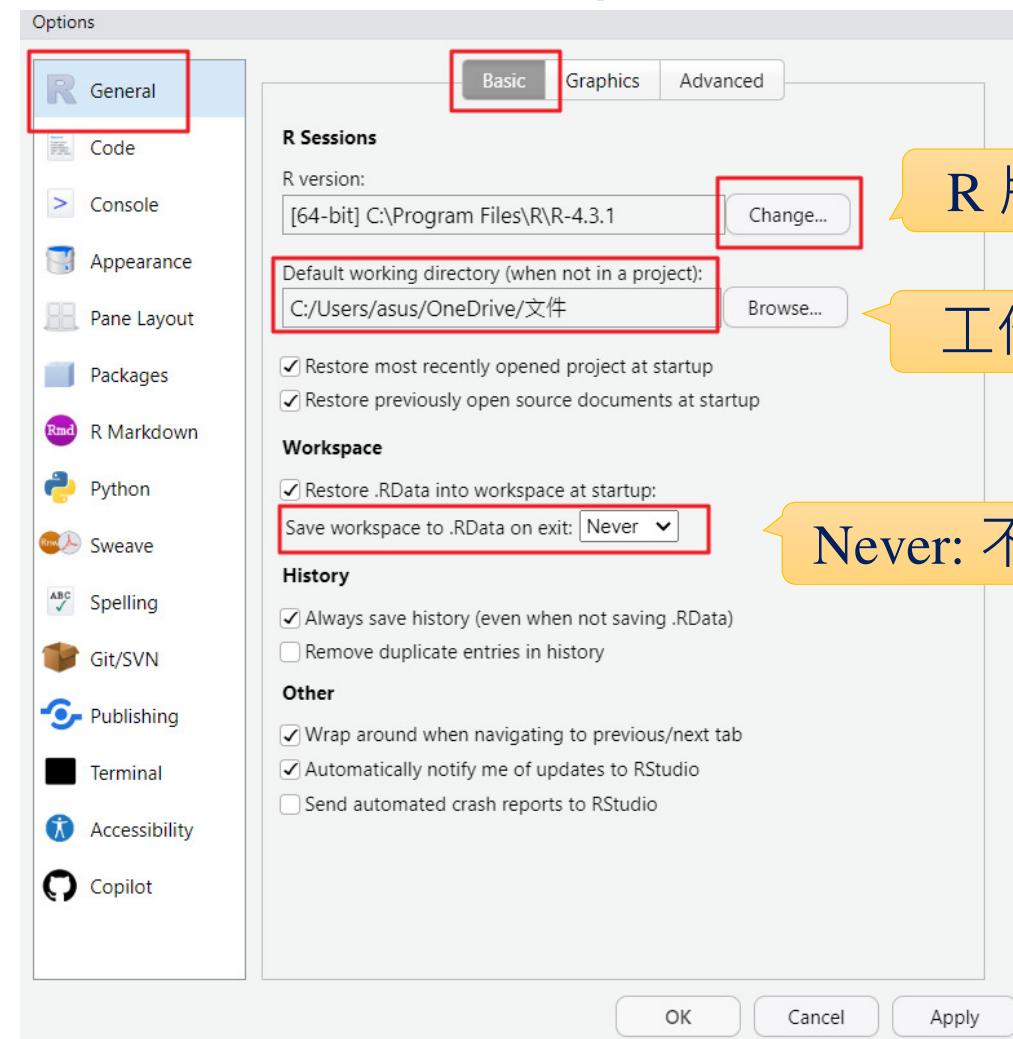


RStudio-選項設定

- Tools \ Global Options



General \ Basic

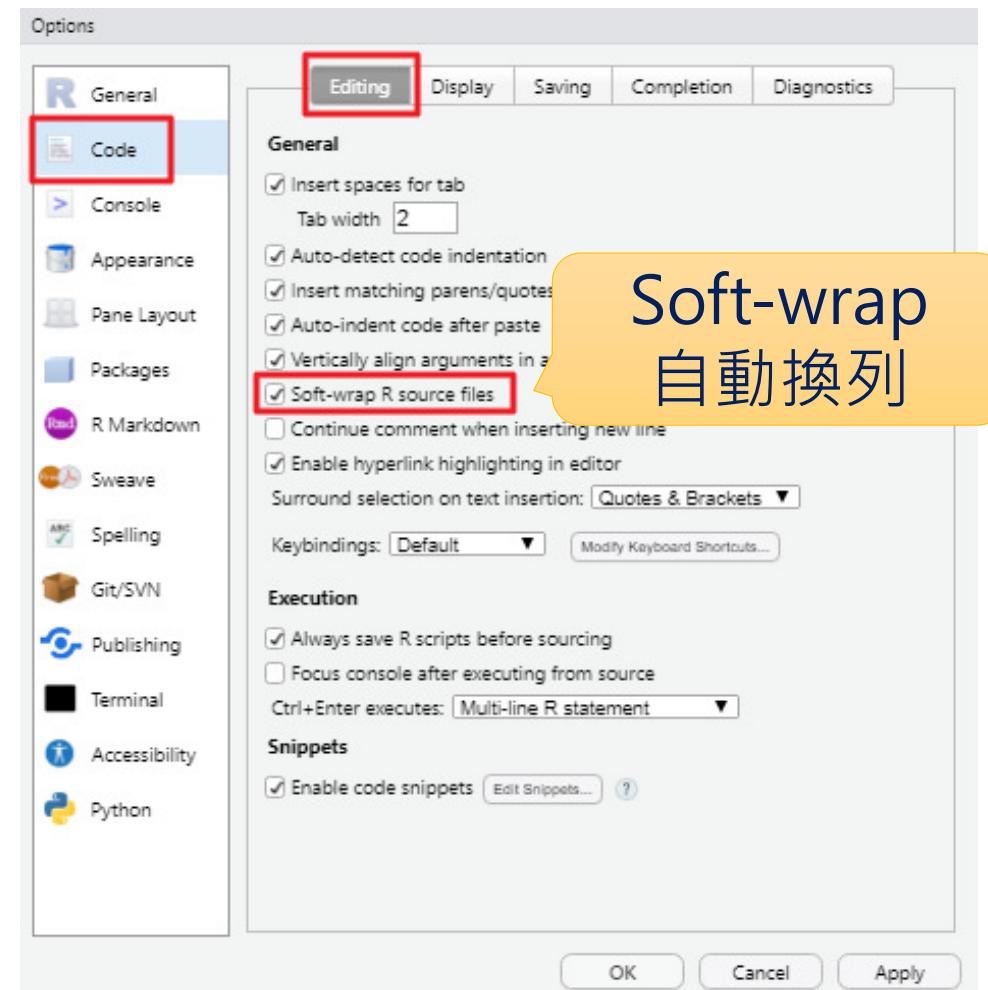


R 版本

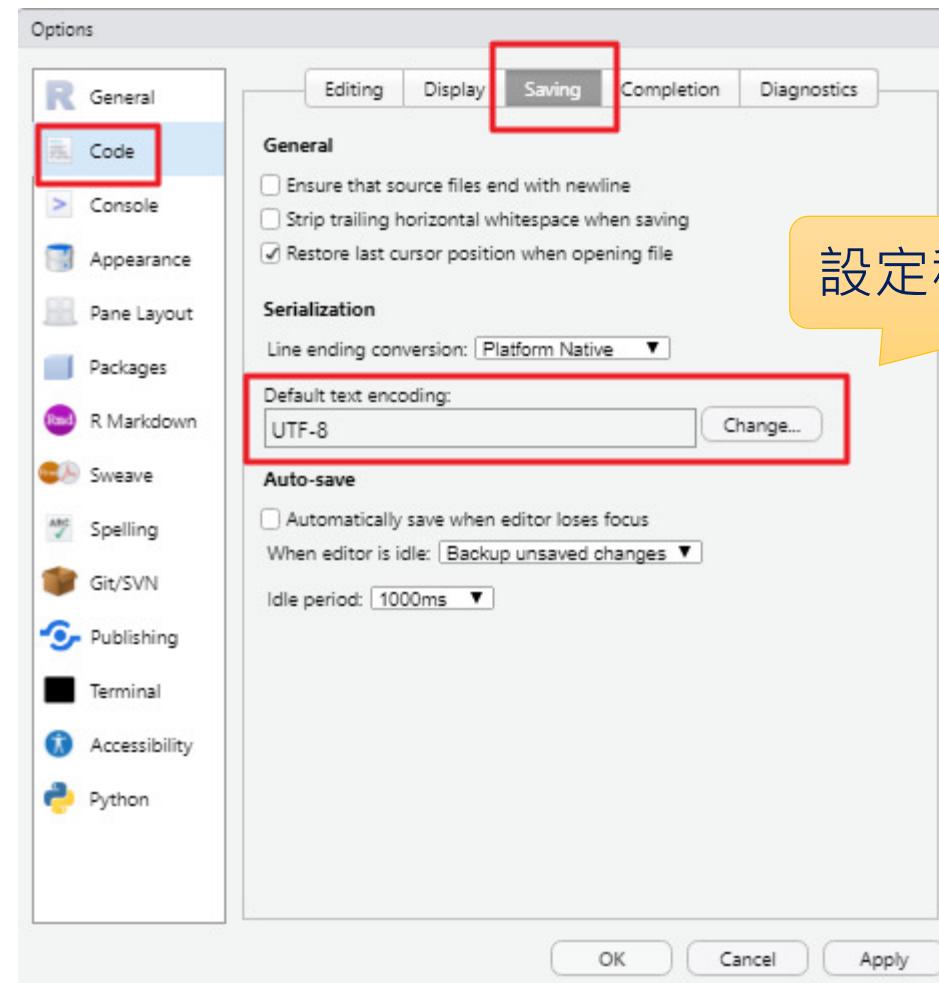
工作目錄 C:/rdata

Never: 不用儲存RData

Code \ Editing



Code \ Saving

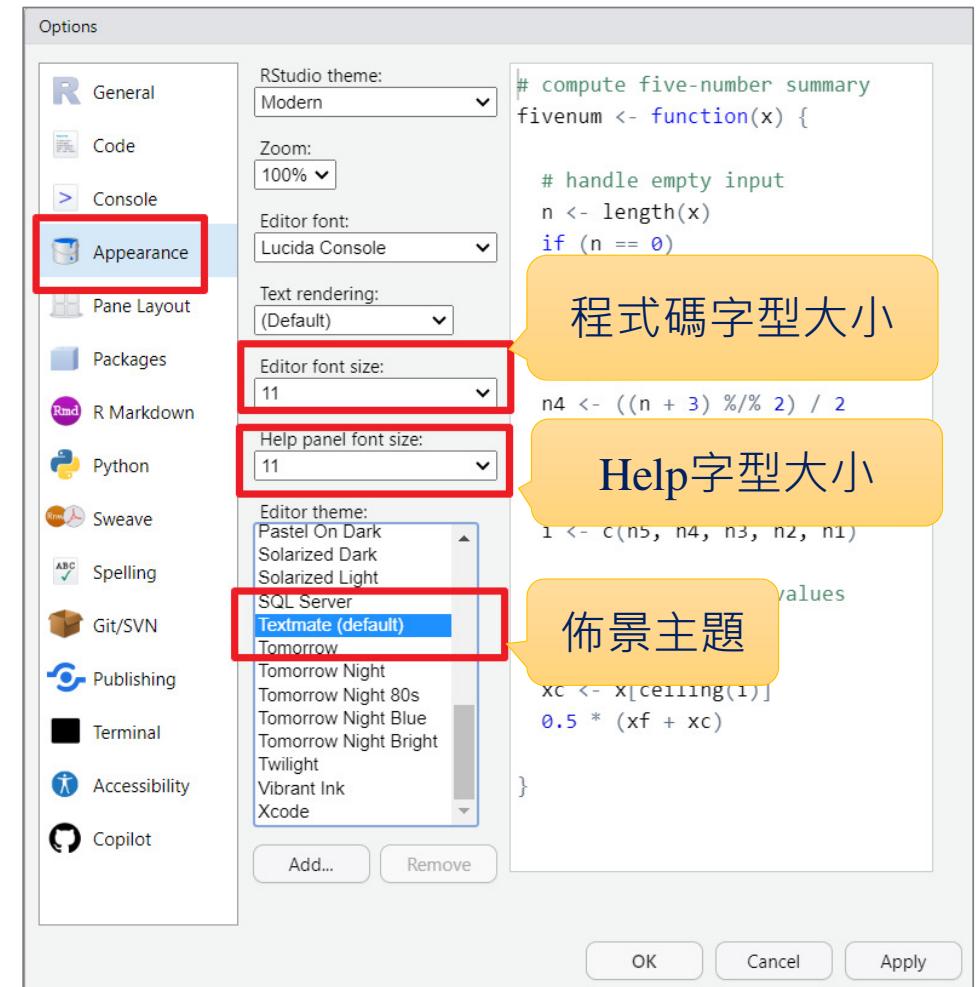


設定程式編碼 UTF-8

RStudio-選項設定(續)

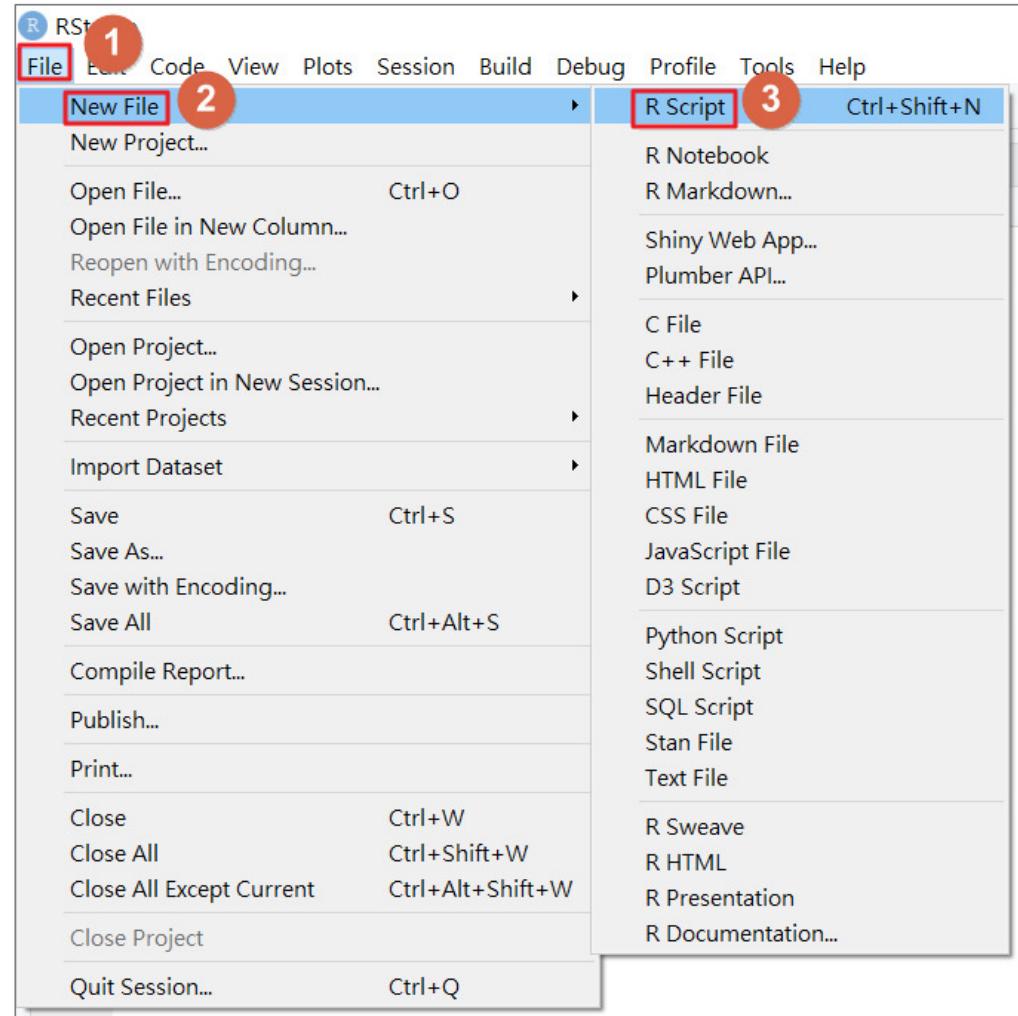
- Appearance \ Editor font size
- Appearance \ Help panel font size
- Editor theme \ TextMate

設定完成，
可能須重新啟動RStudio



新增檔案

- File \ New File \ R Script
(CTRL + SHIFT + N)

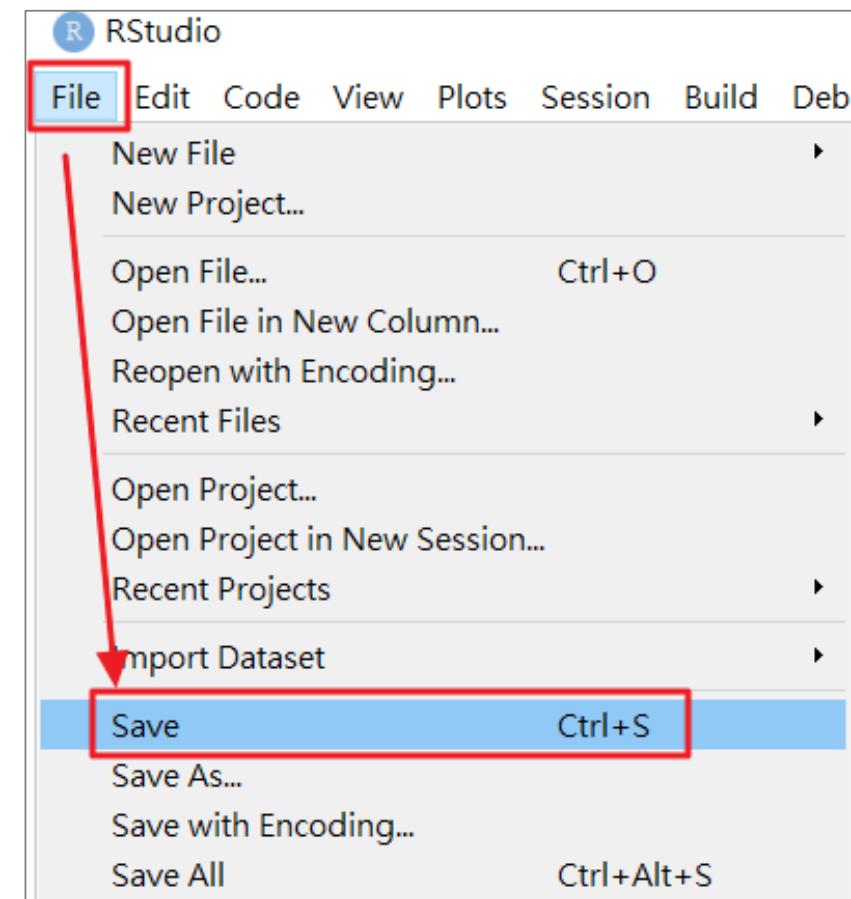


儲存檔案

- 輸入練習2之R程式碼

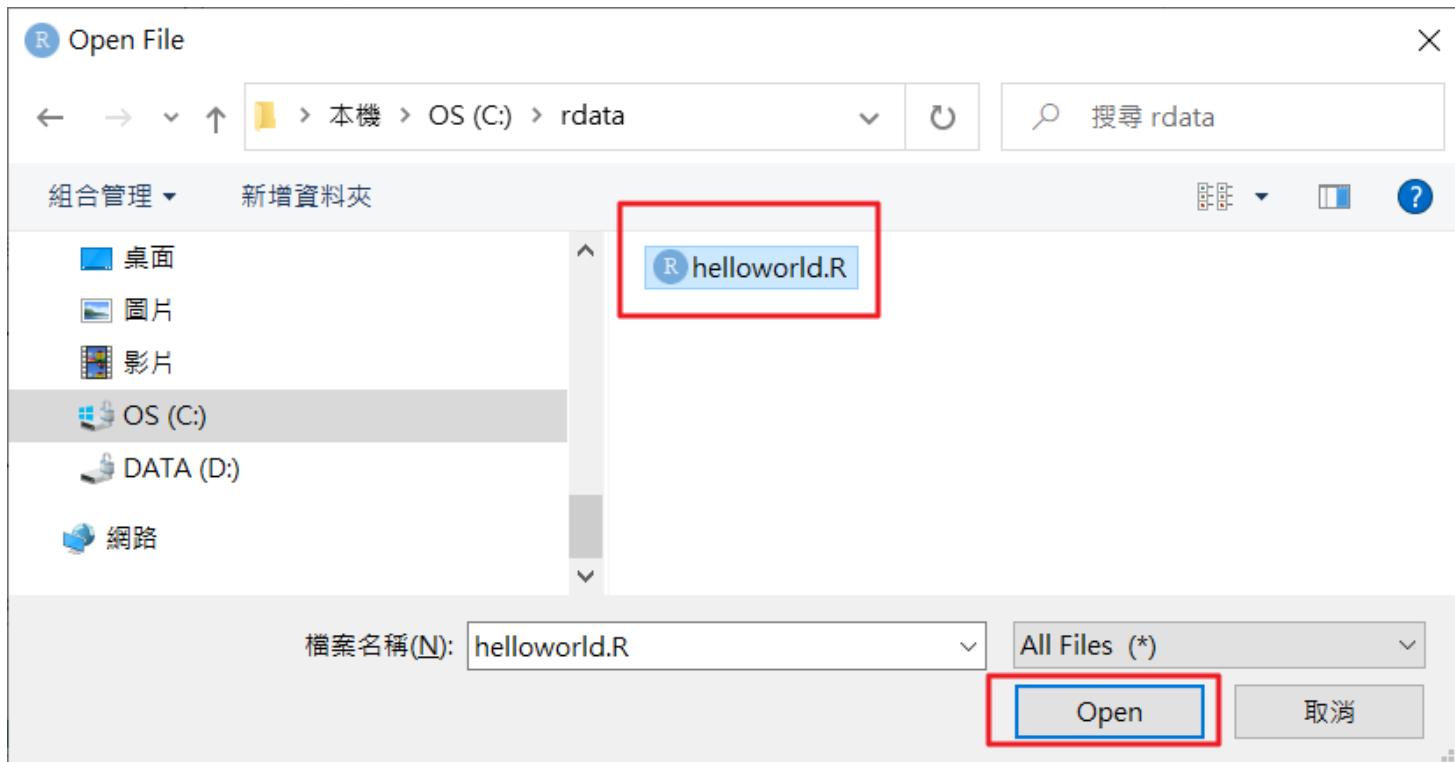
```
1 plot(runif(10), type="b", main= "R大數據分析")
2 x <- rnorm(10)
3 x
4 pairs(iris[-5],
5       pch=16,
6       col=iris$Species,
7       main="RWEPA-iris資料集散佈圖矩陣")
```

- File \ Save (CTRL + S) →
C:\rdata\helloworld.R

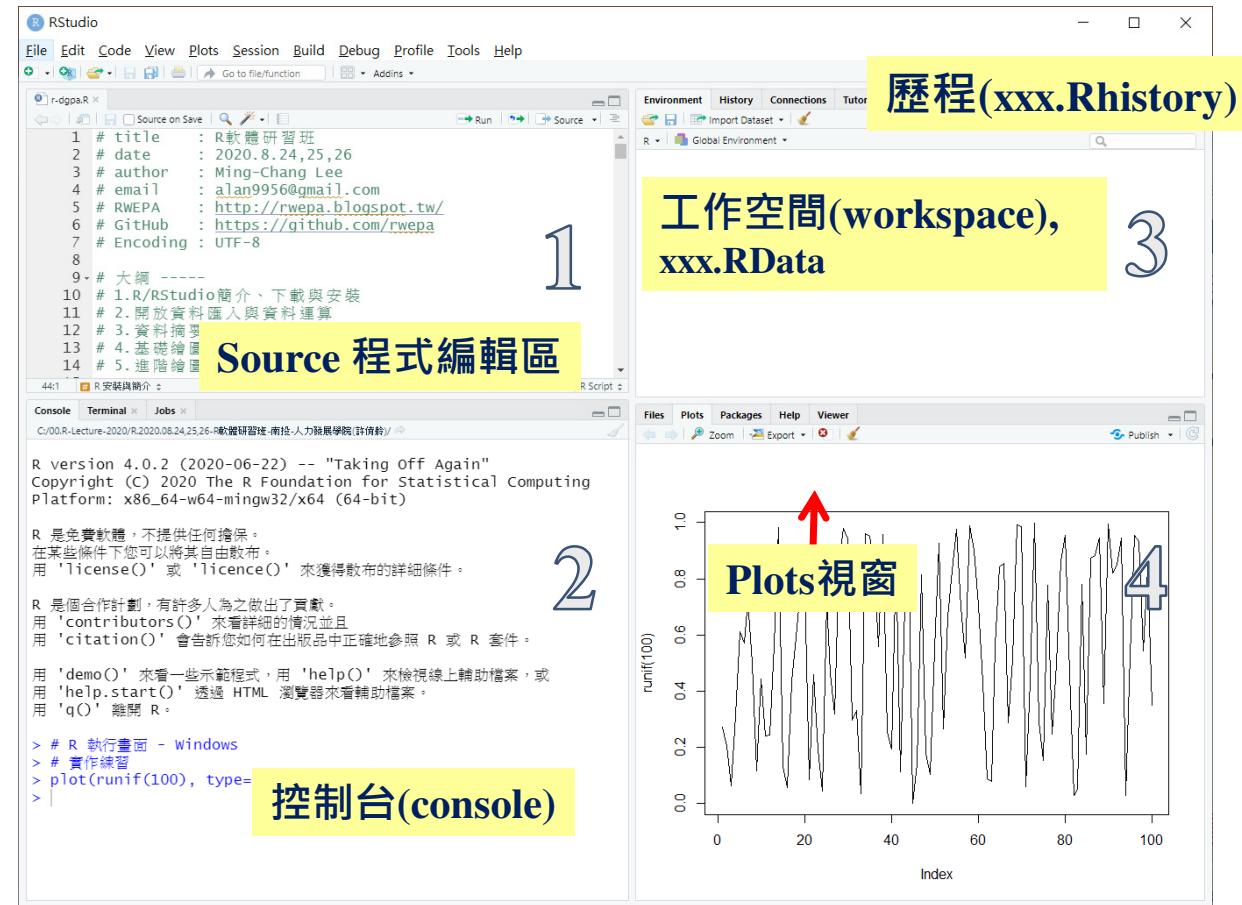


開啟檔案

- File \ Open File \ helloworld.R



R/RStudio環境的基礎觀念



Ctrl + Shift + F10: 重新啟動R

資料型別

資料型別

- 整數
- 數值
- 字串: 須使用 '台北市' 或 "台北市" 符號
- 邏輯值: 包括 TRUE, FALSE

字串範例

The screenshot shows the RStudio interface with the following details:

- Source Editor:** Displays R code for concatenating strings.

```
1 x1 <- c('台北市', '新北市')
2 x1
3 x2 <- c("台北市", "新北市")
4 x2
5
6 x1 == x2
7 |
```
- Environment View:** Shows the values of x1 and x2.

x1	chr [1:2] "台北市" "新北市"
x2	chr [1:2] "台北市" "新北市"
- Console:** Displays the R session output.

```
> x1 <- c('台北市', '新北市')
> x1
[1] "台北市" "新北市"
>
> x2 <- c("台北市", "新北市")
> x2
[1] "台北市" "新北市"
>
> x1 == x2
[1] TRUE TRUE
> |
```
- Panels:** Includes History, Connections, Tutorial, Plots, Packages, Help, Viewer, and Presentation tabs.

資料物件

資料物件

向量 vector

北部	中部	南部
----	----	----

矩陣 matrix

1	3	5
2	4	6

陣列 array

1.1	4.4	7.7
2.2	5.5	8.8
3.3	6.6	9.9

資料框 data.frame

串列 list

類似 Excel
工作表

1	男	62
2	女	50
3	女	54
4	男	72

向量
矩陣

資料框

串列：
每一個元素其資料型別
與長度可以不相同。

參考資料

- RWEPA
 - <https://rwepa.blogspot.com/>
- iPAS-R-tutorial(繪圖中文字型solved)
 - [https://github.com/rwepa/DataDemo/blob/master/aMarvelousR_Lee\(pp238\).R](https://github.com/rwepa/DataDemo/blob/master/aMarvelousR_Lee(pp238).R)
- R入門資料分析與視覺化應用教學(付費)
 - <https://mastertalks.tw/products/r?ref=MCLEE>
- R商業預測與應用(付費)
 - <https://mastertalks.tw/products/r-2?ref=MCLEE>

謝謝您的聆聽

Q & A



李明昌

alan9956@gmail.com

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