

中華民國品質學會第57屆年會暨2021國際品質管理研討會 論壇：開源AI品質工具 - 統計品管與實務應用

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，2600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. Python 實作DoE常用的分析方法
2. Shiny 套件簡介
3. 互動式統計品管網頁實作
4. Q & A

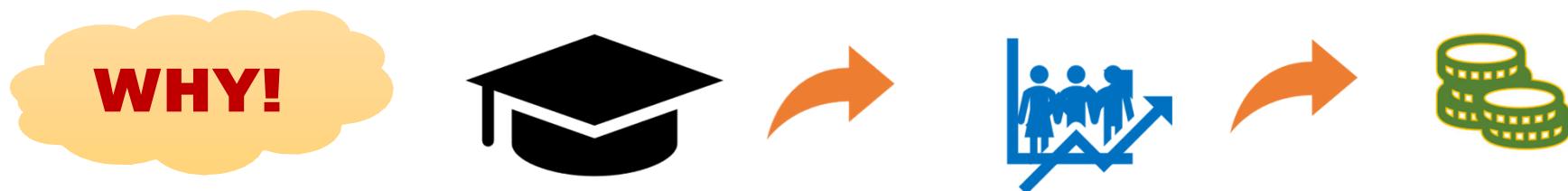
資料分析的心法

★★★資料分析架構→APC方法



如何學習 Python/R?

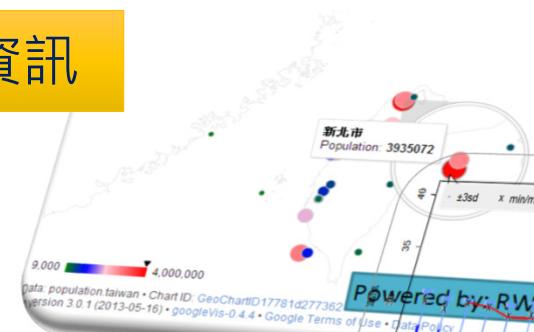
- 熟悉教材內容
- 將教材內容的資料集改為工作資料集(企業, 學術)
- 遇到問題時, 想辦法尋找答案
- 掌握 APC方法
- 掌握 ①摘要 ②繪圖 ③建模
- 參考網路應用文章 (進階) & 學術論文



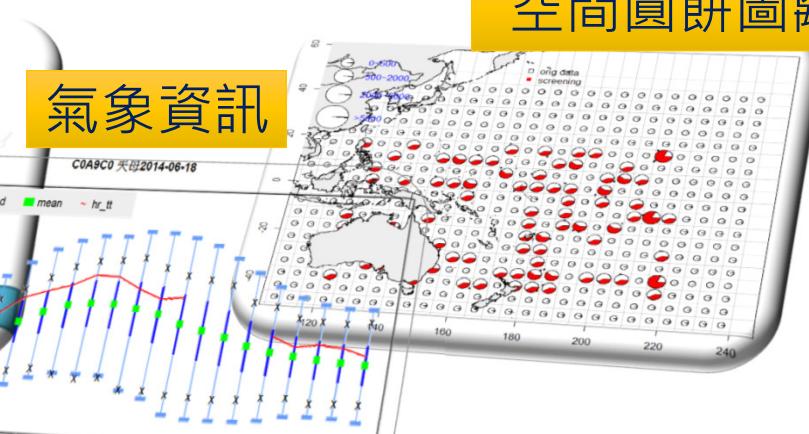
資料分析/視覺化應用

R + shiny → 互動式網頁

地理資訊



氣象資訊

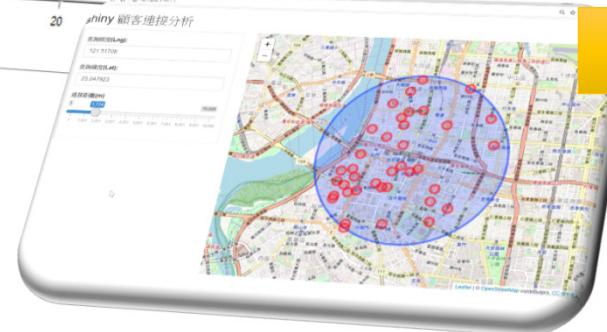


空間圓餅圖離群值分析

保險預測



顧客連結資訊



中央氣象局 1,600萬筆資料

網頁呈現



保險預測模型

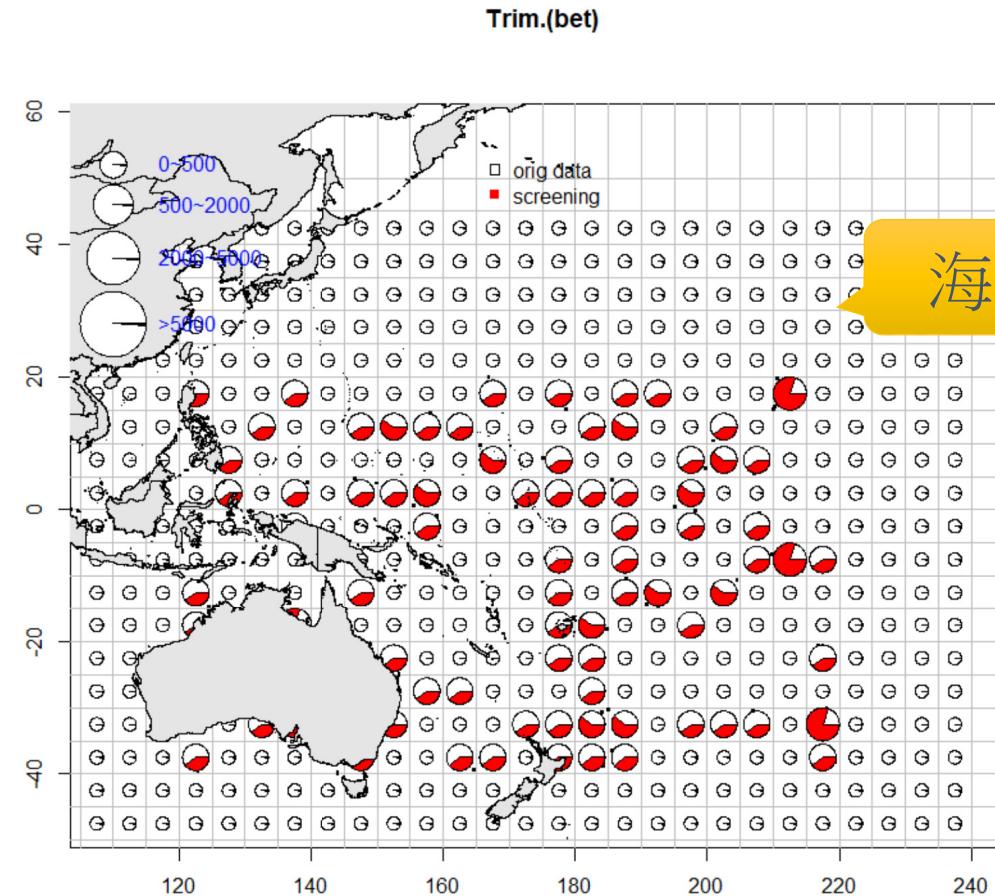
機率模型閾值調整

預測結果

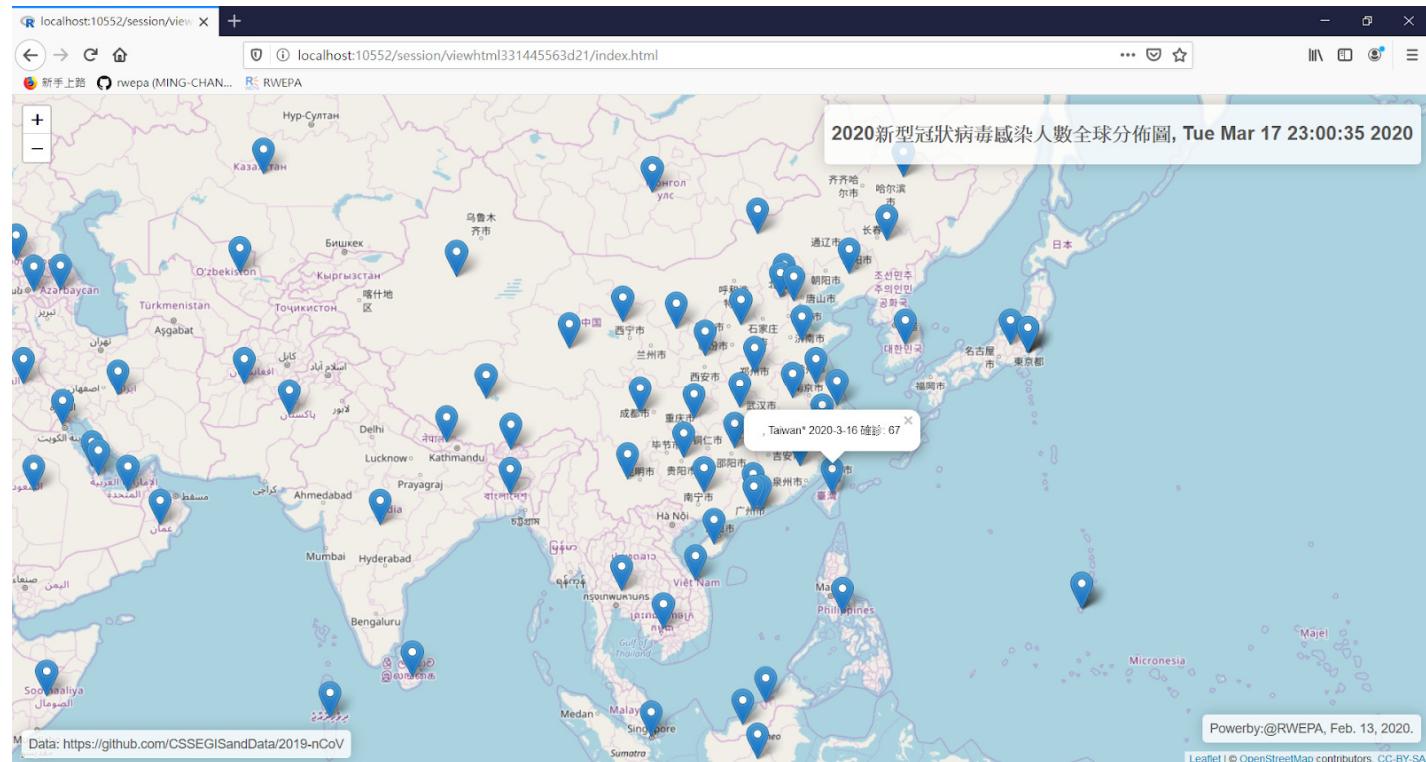
The screenshot shows the iInsurance interactive analysis platform version v.16.3.24. The interface includes a navigation bar with options like '檔案上傳', '資料處理', '統計圖表', '模型評估', and '預測模型'. A red box highlights the '預測模型' dropdown. Below it, a '機率模型閾值' (Probability Model Threshold) slider is set to 0.1, with a red box around its current position. Another red box highlights the '檢視結果' (View Results) button in a modal window. A large yellow callout points to the '預測機率' (Prediction Probability) column in the main table, which lists values such as 0.1069, 0.1441, etc. A red box also highlights the '理賠' (Claim) column. The table has various columns including 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 - 車齡 0 (Private Car - Age 0), 私家車 - 車齡 1 (Private Car - Age 1), 私家車 - 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 bottom of the table shows 'Showing 1 to 10 of 12 entries' and page navigation buttons.

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 - 車齡 0	私家車 - 車齡 1	私家車 - 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

空間圓餅圖離群值分析



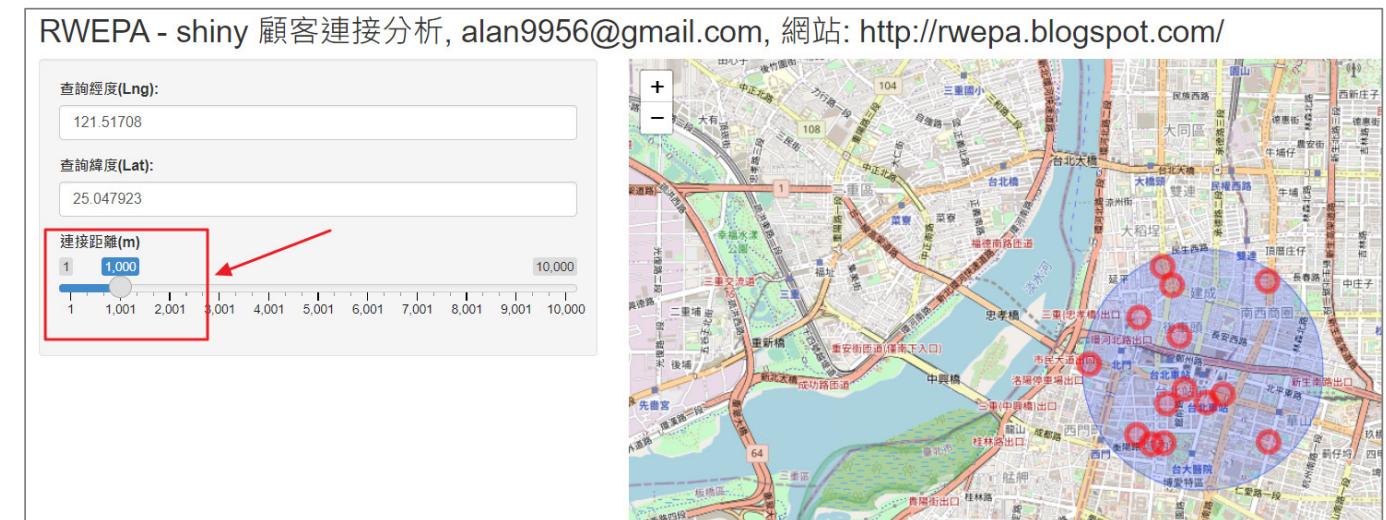
2020新型冠狀病毒視覺化



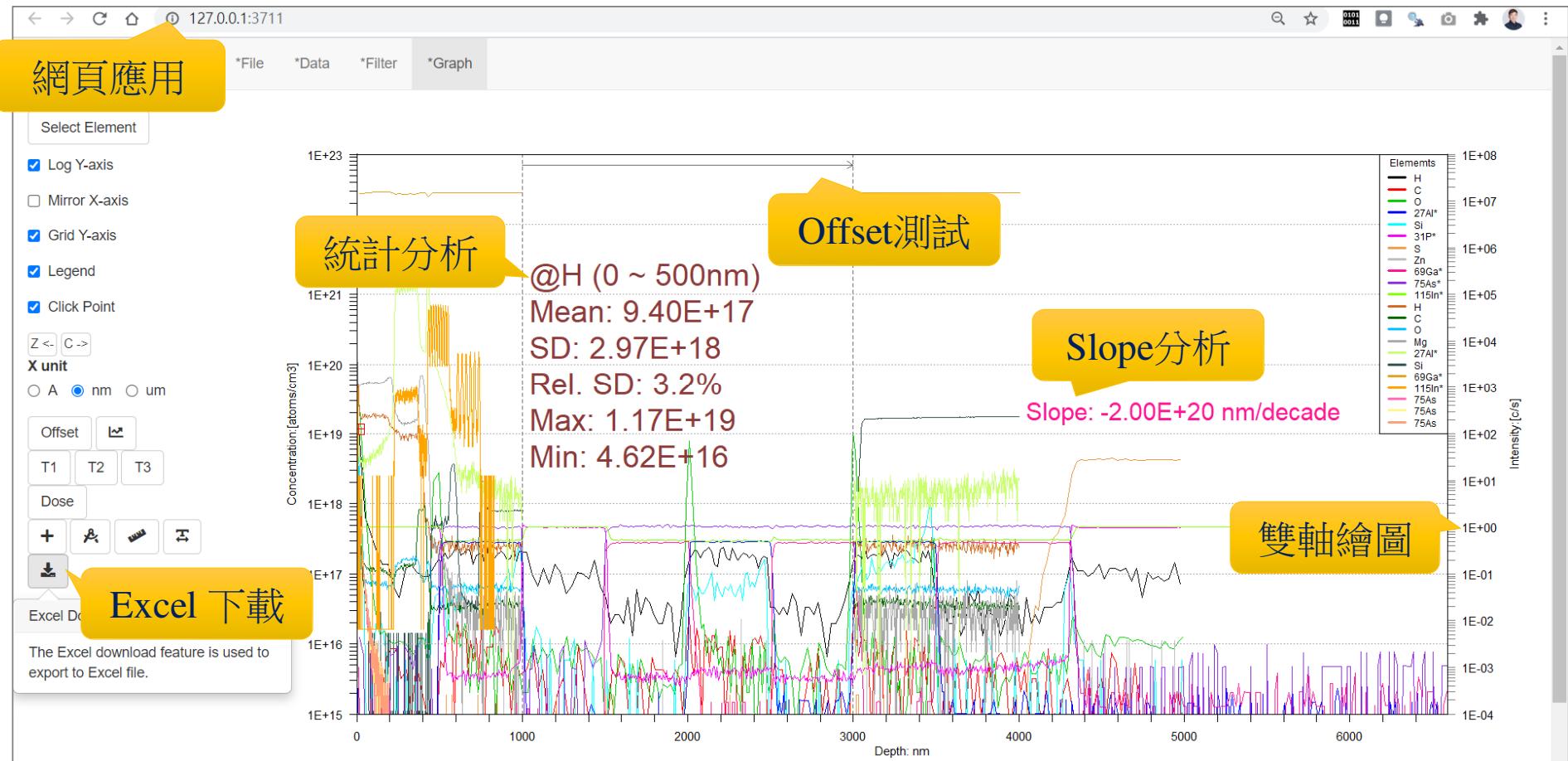
<http://rwepa.blogspot.com/2020/02/2019nCoV.html>

RWEPA - shiny 顧客連接分析

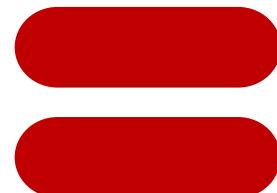
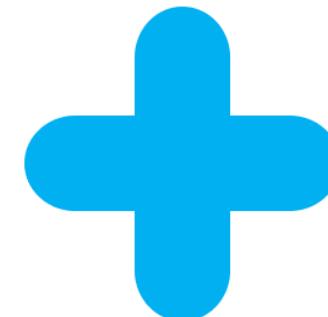
- <https://rwepa.shinyapps.io/shinyCustomerConnect/>



離子資料分析與視覺化應用



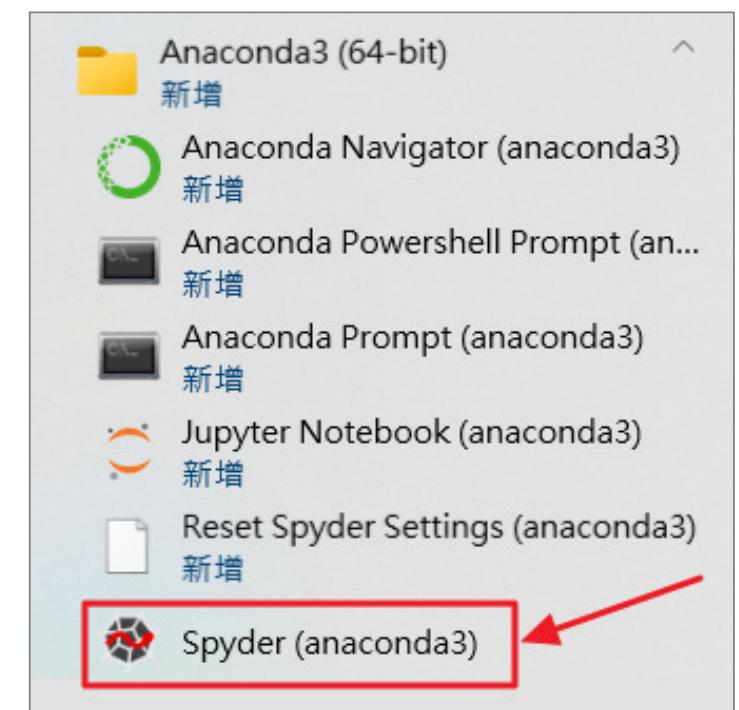
學習目標



1. Python 實作DoE常用的分析方法

Anaconda and Spyder

- Anaconda
 - <https://www.anaconda.com/products/individual>
 - 包括 Spyder, Jupyter Notebook 等工具
- Spyder
 - Python 跨平台整合開發工具
 - DoE (Design Of Experiment 實驗設計)
 - 機率 → 統計 → 實驗設計
 - 實驗設計
 - 一因子實驗-變異數分析
 - 2水準因子實驗



實驗設計

實驗設計

- 實驗設計(Design Of Experiments, DOE)，又稱為試驗設計，屬於科學探究的一種規劃設計，其目的是獲得特定的過程或是系統的相關訊息。
- 實驗中盡可能排除實驗變數以外的影響和干擾，以提高實驗效果的精準度。
- 實驗目標：
 - 製程良率的提昇
 - 找出最大或最小反應值的產品設計參數
 - 比較在不同可控制變數條件下所獲得的反應值
 - 建立數學模型以預測未來的反應值

實驗設計發展



- 1920-1930
- 英國, Sir Ronald A. Fisher
- 農業實驗設計,三原則
- 1951
- 英國, George E. P. Box and K. B. Wilson
- 反應曲面(Response Surface Methodology)
- 中心組合設計 (central composite design)
- 1970
- 日本, 田口玄一
- 穩健參數設計
- 2021
- 機器學習
- 深度學習
- 科學計算

實驗設計的用語

- 實驗 (Experiment) 是一種行動，是實驗者改變至少一個正在研究的變數，然後觀察此變數的影響。一般被動收集資料不是實驗。
- 實驗單位 (Experimental unit)：接受研究的對象，如特定商場(單位)、餐廳或特定製程、產品、消費者。
- 因子 (Factor, Treatment factor, Independent variable)：能夠控制或調整的變數。
- 水準 (Level)：一個因子的各種不同程度。例如：溫度水準為 {30、40、50°C}。
- 處理 (Treatment)：因子中一個或以上水準的組合。溫度30 °C 、溫度40 °C 處理。
- 反應變數 (Response variable)：實驗欲研究的目標變數。

實驗設計的用語 (續)

- **背景變數(Background variable, Lurking variable 潛在變數)**是實驗者未知或無法控制的變數，它可能對實驗結果產生影響。在穩健規劃的實驗設計中，這些潛在變數的影響應該平衡，以避免改變研究結論。
- **效果 (Effect)**是由因子或自變數的改變而引起的反應變化，一般使用反應變數加以估算。
 - 在進行實驗設計後，可以通過觀察到的反應數據計算來估算效果。該估計值稱為計算效果 (Calculated effect)。
 - 在進行實驗之前，研究人員可能知道效果應該有多大才能具有實際重要性。這稱為實務效果(Practical effect)。
- **交絡 (Confounding)**是指實驗過程中，操控的獨立變數與被忽略的其他獨立變項之間，存在一些相互干擾的現象，而影響到獨立變數與依變數之間因果關係的認定。
 - 在這種情況下，無法確定哪個因素導致響應或因變量出現任何觀察到的變化。
 - 例：金屬試驗在50°C油水與100 °C鹽水實驗中，溫度與溶液效果產生交絡現象。

單因子實驗-變異數分析

- One way Analysis of Variance
- Single factor Analysis of Variance

參考: Douglas C. Montgomery, Design and Analysis of Experiments, 10th Edition, Wiley, 2020.

ANOVA資料結構

處理	觀測值			總和	平均值
1	y_{11}	y_{12}	\dots	y_{1n}	$y_{1\cdot}$
2	y_{21}	y_{22}	\dots	y_{2n}	$\bar{y}_{2\cdot}$
\dots				\dots	
a	y_{a1}	y_{a2}	\dots	y_{an}	$\bar{y}_{a\cdot}$
					$\bar{y}_{\cdot\cdot}$

- 平均值模型(Means model)

- $y_{ij} = \mu_i + \varepsilon_{ij}, i = 1, 2, \dots, a; j = 1, 2, \dots, n$

- 效果模型(Effects model)

- 考慮 $\mu_i = \mu + \tau_i, i = 1, 2, \dots, a$

- $y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \mu$ 總平均, τ_i 第*i*個處理效果

- 處理有 a 個(或因子有 a 個水準)
- y_{ij} : 第 i 個處理，第 j 個觀測值
- μ_i : 第 i 處理的平均數, 即 $\bar{y}_{i\cdot}$
- ε_{ij} : 隨機誤差, $E(\varepsilon_{ij}) = 0$, 即 $E(y_{ij}) = \mu_i$

假設檢定

- 考慮隨機誤差 ε_{ij} 獨立且服從常態隨機變數.
- ε_{ij} 平均數為 0, 變異數為 σ^2
- y_{ij} 為彼此獨立 $y_{ij} \sim N(\mu + \tau_i, \sigma^2)$

資料模型

- $y_{i\cdot} = \sum_{j=1}^n y_{ij}, i = 1, 2, \dots, a$
- $\bar{y}_{i\cdot} = \frac{y_{i\cdot}}{n}$
- $y_{..} = \sum_{i=1}^a \sum_{j=1}^n y_{ij}$
- $\bar{y}_{..} = \frac{y_{..}}{N}, N = an$ 總觀測值個數
- 檢定 a 個處理的平均值是否相等
 - $H_0: \mu_1 = \mu_2 = \dots = \mu_a$
 - $H_1: \mu_i \neq \mu_j$ 至少有一對 (i, j)

	處理	觀測值			總和	平均值
1	y_{11}	y_{12}	\dots	y_{1n}	$y_{1\cdot}$	$\bar{y}_{1\cdot}$
2	y_{21}	y_{22}	\dots	y_{2n}	$y_{2\cdot}$	$\bar{y}_{2\cdot}$
\dots					\dots	
a	y_{a1}	y_{a2}	\dots	y_{an}	$y_{a\cdot}$	$\bar{y}_{a\cdot}$
					$y_{..}$	$\bar{y}_{..}$

$$\begin{aligned}
 \because \mu &= \frac{\sum_{i=1}^a \mu_i}{a} = \frac{\sum_{i=1}^a (\mu + \tau_i)}{a} \\
 &= \frac{a\mu + \sum_{i=1}^a \tau_i}{a} = \mu + \frac{\sum_{i=1}^a \tau_i}{a} \\
 \therefore \sum_{i=1}^a \tau_i &= 0, \text{ 其中 } \tau_i = \mu_i - \mu
 \end{aligned}$$



- $H_0: \tau_1 = \tau_2 = \dots = \tau_a = 0$
- $H_1: \text{至少有一個 } i \text{ 使得 } \tau_i \neq 0$

總平方和分解法

- 總平方和 (Total Sum of Squares)

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$\begin{aligned} SS_T &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^n [(\bar{y}_{i..} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i..})]^2 \\ &= n \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{..})^2 + 2 \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_{i..} - \bar{y}_{..}) (y_{ij} - \bar{y}_{i..}) + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i..})^2 \end{aligned}$$

$$= \boxed{n \sum_{i=1}^a (\bar{y}_{i..} - \bar{y}_{..})^2} + \boxed{\sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i..})^2} \quad \sum_{j=1}^n (y_{ij} - \bar{y}_{i..}) = y_{i..} - n\bar{y}_{i..} = y_{i..} - n \frac{y_{i..}}{n} = 0$$

= 平均與總平均差的平方和 + 處理內觀測值與處理平均的差平方和
 = 處理平均間差異 + 隨機誤差

- $SS_T = SS_{Treatment} + SS_E =$ 處理平方和(處理平均間) + 誤差平方和(處理內)

變異數分析表

變異來源	平方和	自由度	均方	F_0
處理間	$SS_{Treatment}$	$a - 1$	$MS_{Treatment} = \frac{SS_{Treatment}}{a - 1}$	$F_0 = \frac{MS_{Treatment}}{MS_E}$
誤差	SS_E	$N - a$	$MS_E = \frac{SS_E}{N - a}$	
總和	SS_T	$N - 1$		

註:

1. 總樣本數 $N = a \times n$
2. 處理間平方和 $SS_{Treatment} = n \sum_{i=1}^a (\bar{y}_{i\cdot} - \bar{y}_{..})^2$
3. 誤差平方和 $SS_E = SS_T - SS_{Treatment}$
4. 總平方和 $SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$
5. 誤差項自由度 $(N - 1) - (a - 1) = N - a$
6. 如果 $F_0 \geq F_\alpha(a - 1, N - a)$, 則表示 **p值 < α**, 即拒絕 H_0 , 接受不同處理平均數有不同.

統計模組

- `scipy.stats`



- Scientific computing tools for Python (科學計算)
- 模組包含：統計、最佳化、線性代數、積分、插值、特殊函數、快速傅立葉變換、訊號處理和圖像處理、常微分方程式。

- `statsmodels`

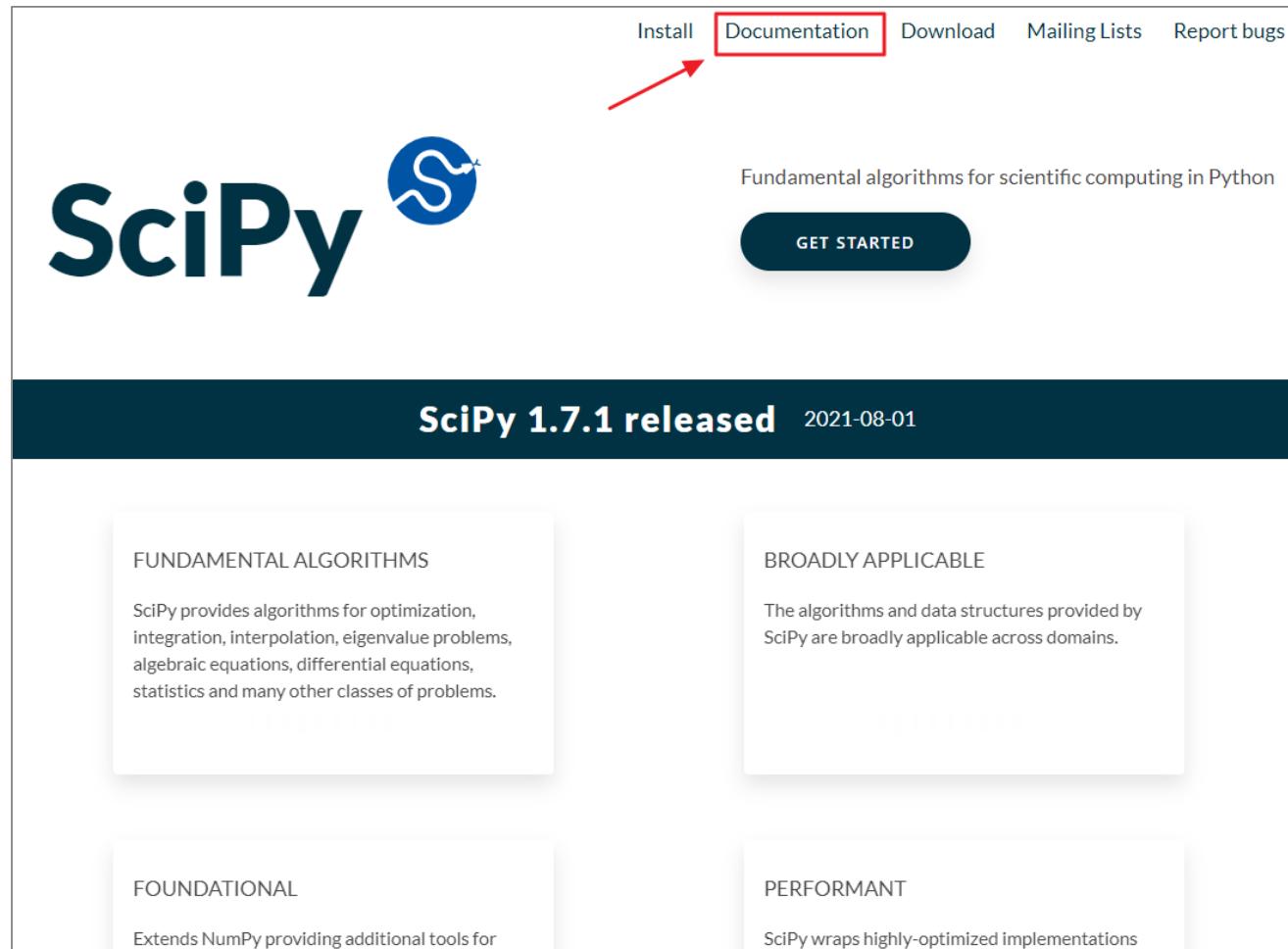


statsmodels v0.13.0

- 提供敘述性統計、統計檢定、統計模型，繪圖功能。
- Statsmodels 加強 SciPy的統計功能。

scipy.stats 模組

<https://scipy.org/>



The screenshot shows the official website for SciPy. At the top, there is a navigation bar with links: Install, Documentation, Download, Mailing Lists, and Report bugs. The 'Documentation' link is highlighted with a red box and a red arrow pointing to it from the left. Below the navigation bar, the SciPy logo is displayed, followed by the text 'Fundamental algorithms for scientific computing in Python' and a 'GET STARTED' button. A dark blue banner across the middle of the page announces 'SciPy 1.7.1 released' on '2021-08-01'. The main content area is divided into four sections: 'FUNDAMENTAL ALGORITHMS' (describing optimization, integration, interpolation, eigenvalue problems, etc.), 'BROADLY APPLICABLE' (describing the broad applicability across domains), 'FOUNDATIONAL' (extending NumPy with additional tools), and 'PERFORMANT' (wrapping highly-optimized implementations). Each section has a brief description and a small icon.

Install Documentation Download Mailing Lists Report bugs

SciPy S

Fundamental algorithms for scientific computing in Python

GET STARTED

SciPy 1.7.1 released 2021-08-01

FUNDAMENTAL ALGORITHMS

SciPy provides algorithms for optimization, integration, interpolation, eigenvalue problems, algebraic equations, differential equations, statistics and many other classes of problems.

BROADLY APPLICABLE

The algorithms and data structures provided by SciPy are broadly applicable across domains.

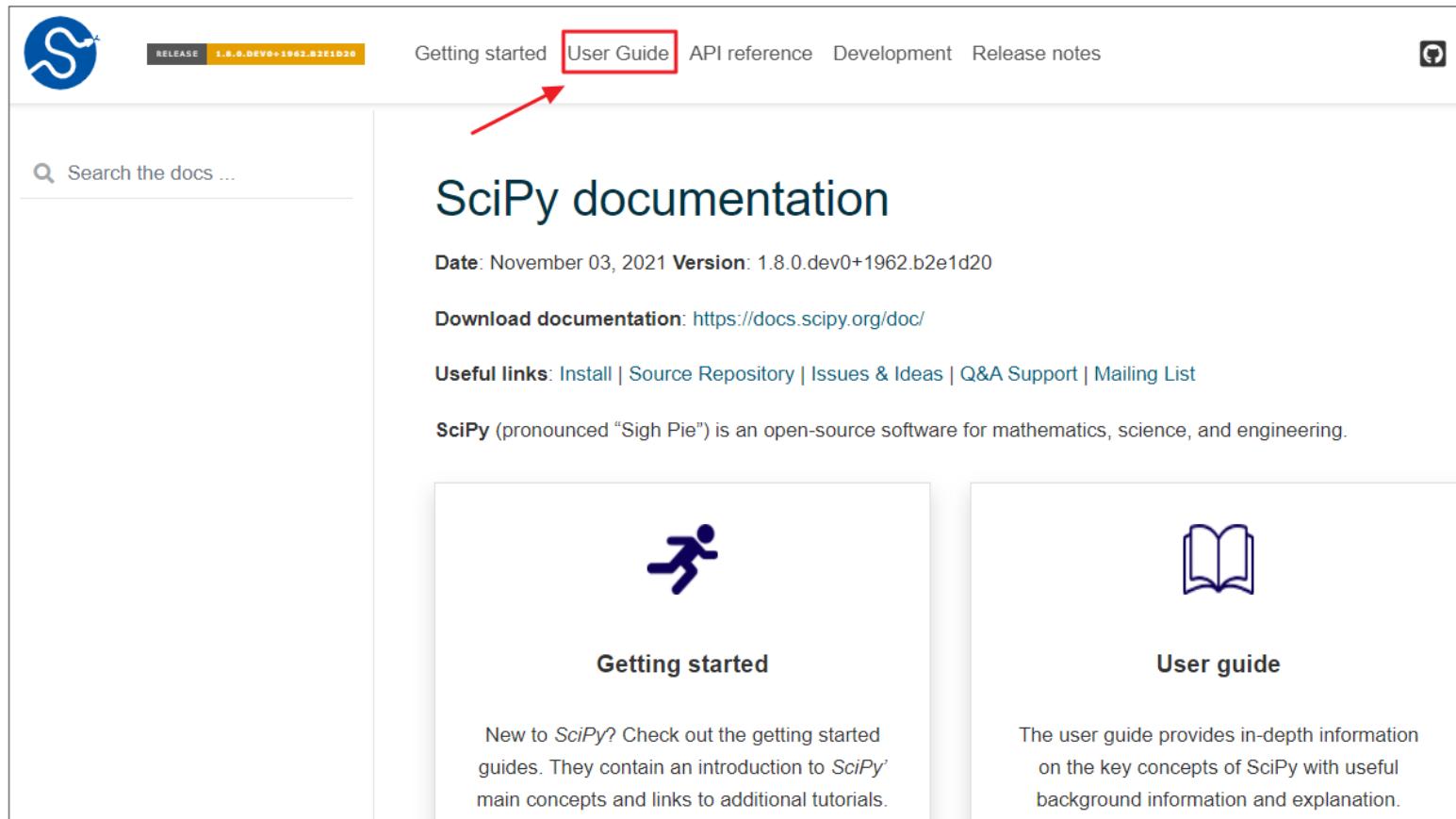
FOUNDATIONAL

Extends NumPy providing additional tools for

PERFORMANT

SciPy wraps highly-optimized implementations

User Guide



The screenshot shows the SciPy documentation homepage. At the top, there is a navigation bar with links: Getting started, **User Guide**, API reference, Development, and Release notes. The "User Guide" link is highlighted with a red box and a red arrow points to it from below. To the right of the navigation bar is a small icon of a person with a gear. On the left side of the page is a search bar with the placeholder "Search the docs ...". The main content area features the title "SciPy documentation" in large blue text. Below the title, it says "Date: November 03, 2021 Version: 1.8.0.dev0+1962.b2e1d20". It also includes a "Download documentation" link (<https://docs.scipy.org/doc/>) and a "Useful links" section with links to Install, Source Repository, Issues & Ideas, Q&A Support, and Mailing List. A brief description follows: "SciPy (pronounced ‘Sigh Pie’) is an open-source software for mathematics, science, and engineering." Below this, there are two main sections: "Getting started" and "User guide". Each section has an icon (a running figure for "Getting started" and an open book for "User guide") and a brief description.

Getting started **User Guide** API reference Development Release notes

SciPy documentation

Date: November 03, 2021 Version: 1.8.0.dev0+1962.b2e1d20

Download documentation: <https://docs.scipy.org/doc/>

Useful links: [Install](#) | [Source Repository](#) | [Issues & Ideas](#) | [Q&A Support](#) | [Mailing List](#)

SciPy (pronounced “Sigh Pie”) is an open-source software for mathematics, science, and engineering.

 Getting started

New to SciPy? Check out the getting started guides. They contain an introduction to SciPy's main concepts and links to additional tutorials.

 User guide

The user guide provides in-depth information on the key concepts of SciPy with useful background information and explanation.

SciPy User Guide

- Introduction
- Special functions (scipy.special)
- Integration (scipy.integrate) 積分
- Optimization (scipy.optimize) 最佳化
- Interpolation (scipy.interpolate) 內差法
- Fourier Transforms (scipy.fft)
- Signal Processing (scipy.signal)
- Linear Algebra (scipy.linalg)
- Sparse eigenvalue problems with ARPACK
- Compressed Sparse Graph Routines (scipy.sparse.csgraph)
- Spatial data structures and algorithms (scipy.spatial)
- **Statistics (scipy.stats)** 統計
- Multidimensional image processing (scipy.ndimage)
- File IO (scipy.io)

scipy.stats.f_oneway

scipy.stats.f_oneway

```
scipy.stats.f_oneway(*args, axis=0)
```

[source]

Perform one-way ANOVA.

The one-way ANOVA tests the null hypothesis that two or more groups have the same population mean. The test is applied to samples from two or more groups, possibly with differing sizes.

Parameters: `sample1, sample2, ... : array_like`

The sample measurements for each group. There must be at least two arguments. If the arrays are multidimensional, then all the dimensions of the array must be the same except for `axis`.

axis : int, optional

Axis of the input arrays along which the test is applied. Default is 0.

Returns: `statistic : float`

The computed F statistic of the test.

pvalue : float

The associated p-value from the F distribution.

statsmodels

<https://www.statsmodels.org/stable/>

statsmodels v0.13.0

Search

statsmodels
6.8k Stars · 2.4k Forks

Versions

statistical models, hypothesis tests, and data exploration

statsmodels v0.13.0

Installing statsmodels

Getting started

User Guide

Examples

API Reference

About statsmodels

Developer Page

Release Notes

 statsmodels

statsmodels is a Python module that provides classes and functions for the estimation of many different statistical models, as well as for conducting statistical tests, and statistical data exploration. An extensive list of result statistics are available for each estimator. The results are tested against existing statistical packages to ensure that they are correct. The package is released under the open source Modified BSD (3-clause) license. The online documentation is hosted at statsmodels.org.

Contents

Introduction

Citation

Index

Show Source

Introduction

statsmodels supports specifying models using R-style formulas and pandas DataFrames. Here is a simple example using ordinary least squares:

statsmodels.formula.api.ols

statsmodels.formula.api.ols

```
statsmodels.formula.api.ols(formula, data, subset=None, drop_cols=None, *args, **kwargs)
```

Create a Model from a formula and dataframe.

Parameters

formula : str or generic Formula object

The formula specifying the model.

data : array_like

The data for the model. See Notes.

subset : array_like

An array-like object of booleans, integers, or index values that indicate the subset of df to use in the model. Assumes df is a pandas.DataFrame.

<https://www.statsmodels.org/dev/generated/statsmodels.formula.api.ols.html>

單因子實驗-範例

- 考慮蝕刻製程，研究 RF功率(X)與蝕刻率(Y)的關係
 - RF功率有4個水準，160,180,200,220W (完全隨機試驗)
 - 每個RF功率測試5片晶圓
 - 反應變數是蝕刻率
 - 實驗設計計劃
 - 水準 $a = 4$
 - 重複 $n = 5$
 - 只考慮RF功率，因此是一因子實驗
 - 全部實驗次數20
- | RF Power
(W) | Observed Etch Rate (Å/min) | | | | |
|-----------------|----------------------------|-----|-----|-----|-----|
| | 1 | 2 | 3 | 4 | 5 |
| 160 | 575 | 542 | 530 | 539 | 570 |
| 180 | 565 | 593 | 590 | 579 | 610 |
| 200 | 600 | 651 | 610 | 637 | 629 |
| 220 | 725 | 700 | 715 | 685 | 710 |
- 埃格斯特朗 (Ångström, 簡稱埃，符號Å)
 - Å是一個長度計量單位。
 - $1 \text{ Å} = 10^{-10} \text{ 米} = 0.1 \text{ 奈米}.$

scipy.stats 實作

```
In [1]: from scipy.stats import f_oneway  
  
In [2]: w160 = [575, 542, 530, 539, 570]  
  
In [3]: w180 = [565, 593, 590, 579, 610]  
  
In [4]: w200 = [600, 651, 610, 637, 629]  
  
In [5]: w220 = [725, 700, 715, 685, 710]  
  
In [6]: rf = f_oneway(w160, w180, w200, w220)  
  
In [7]: rf  
Out[7]: F_onewayResult(statistic=66.79707321945864, pvalue=2.882865908493268e-09)  
  
In [8]: type(rf)  
Out[8]: scipy.stats.stats.F_onewayResult  
  
In [9]: rf.statistic      # 66.79707321945864  
Out[9]: 66.79707321945864  
  
In [10]: rf.pvalue        # 2.882865908493268e-09  
Out[10]: 2.882865908493268e-09  
  
In [11]: rf.pvalue < 0.05 # True, Do not accept H0, 不同RF功率(X)的蝕刻率(Y)不全相同  
Out[11]: True
```

statsmodels.formula.api.ols 實作

```
In [1]: import statsmodels.api as sm

In [2]: from statsmodels.formula.api import ols

In [3]: import pandas as pd

In [4]: df = pd.read_excel(io = 'C:/rdata/rf.xlsx', sheet_name = 'rf')

In [5]: df.head(n=3)
Out[5]:
    rf  etchrate
0  w160      575
1  w160      542
2  w160      530

In [6]: mod = ols('etchrate ~ rf', data = df).fit()

In [7]: aov_table = sm.stats.anova_lm(mod, typ=2)

In [8]: print(aov_table)
          sum_sq      df           F      PR(>F)
rf        66870.55   3.0  66.797073  2.882866e-09
Residual  5339.20  16.0         NaN        NaN
```

結果與 scipy 相同

2.Shiny 套件簡介

認識 R

- 1976 - 貝爾實驗室 John Chambers, Rick Becker, and Allan Wilks 研發 S 語。
- 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。
 - 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
 - 2021年11月 - R 4.1.2
 - 下載 <http://www.r-project.org/>

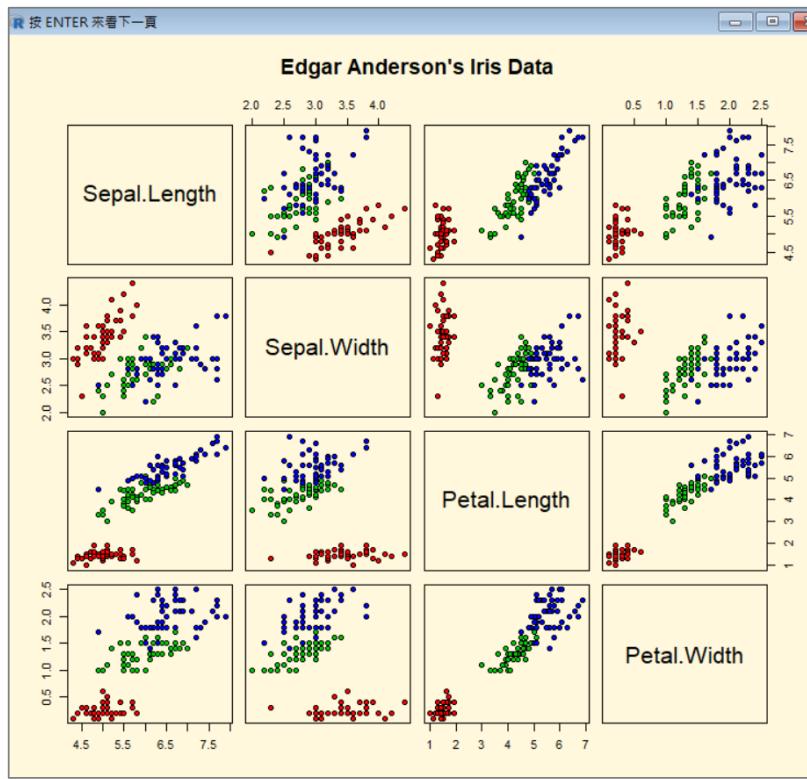


R-八大功能

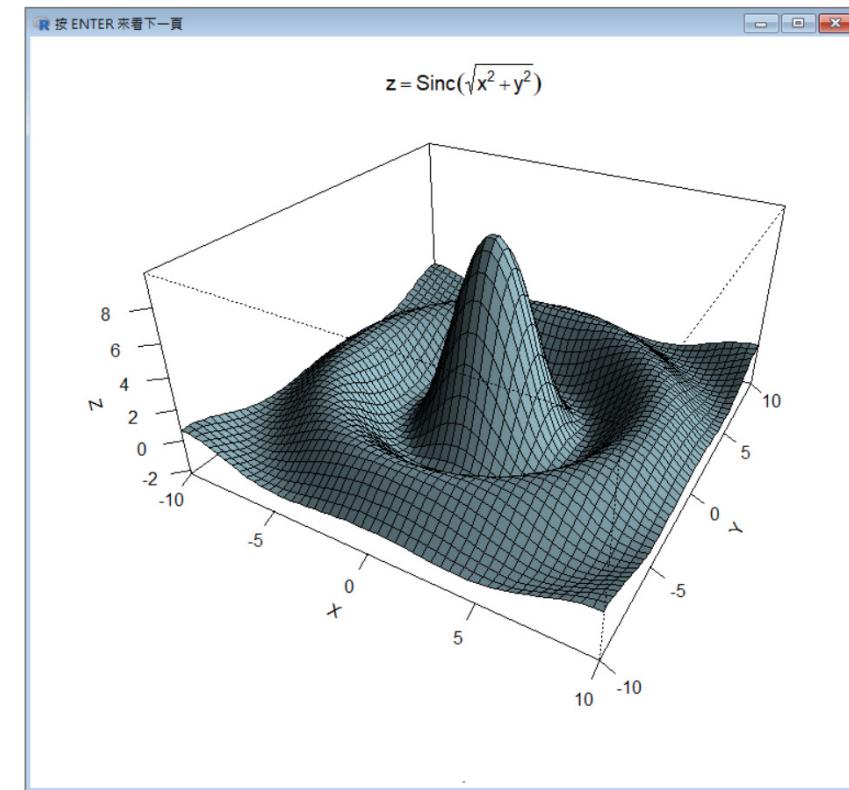


R 視覺化

- `demo(graphics)`

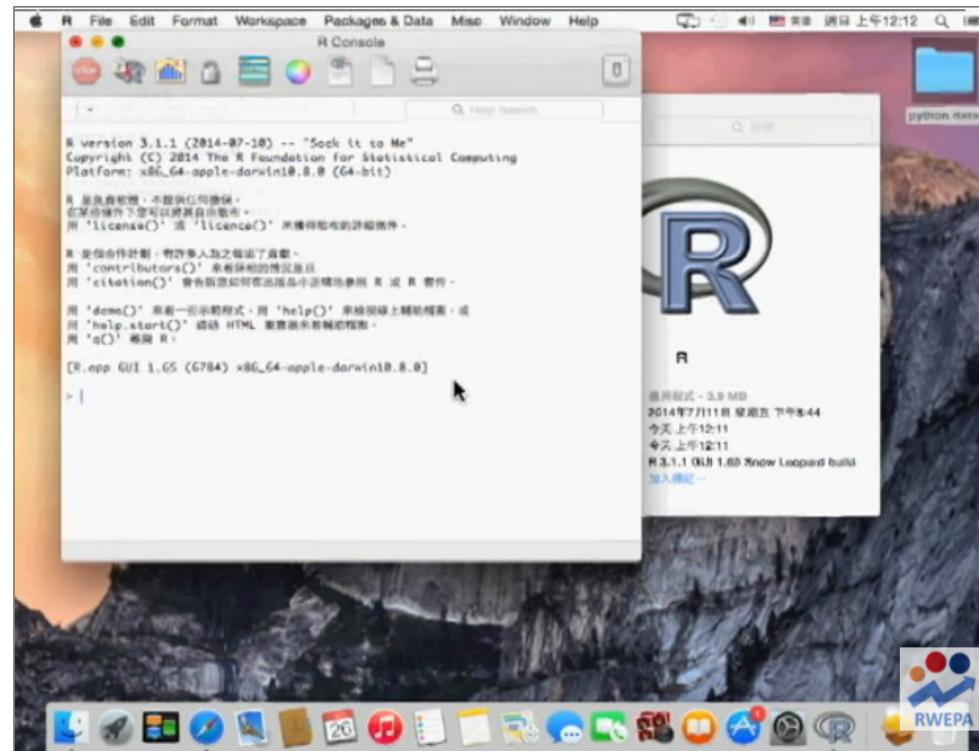


- `demo(persp)`



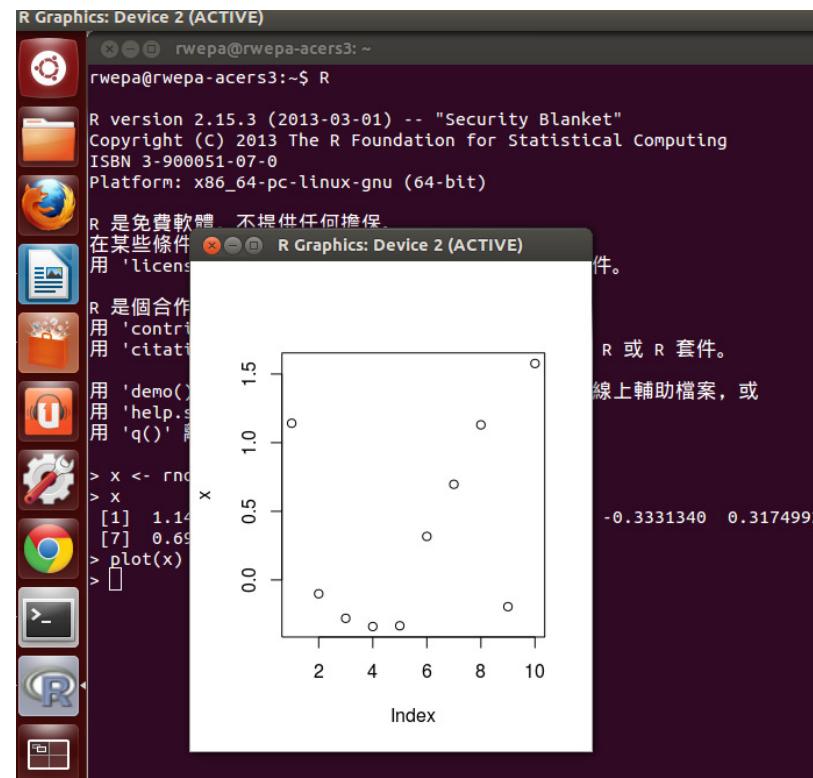
R for Mac

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



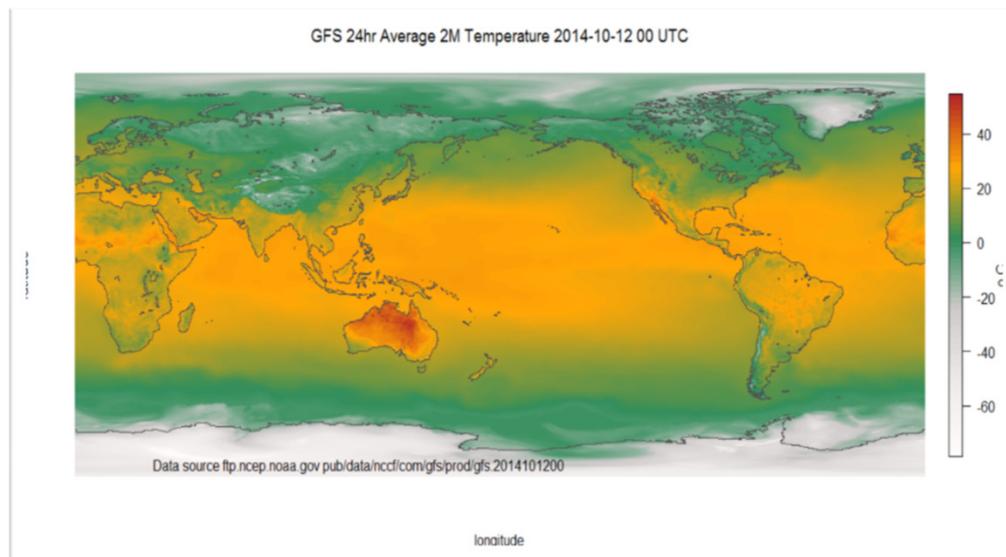
R for Ubuntu

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



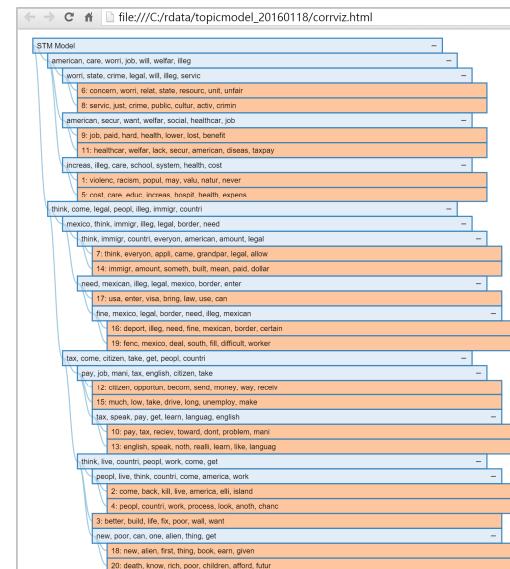
整合式開發環境 - RStudio

- <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
 - Tools\Global Options \ Code \ Saving \ Default text encoding: Change: UTF-8

R and RStudio 下載 - <http://rwepa.blogspot.com/>

Service : alan9956@gmail.com

大數據分析,資料視覺化,R,PYTHON,shiny程式
設計,統計品管,最佳化,企業服務,業師協同教學.

RWEPA 搜尋此網誌 (例: task)

- 關於作者
- GitHub DataDemo
- R與實驗設計應用影片(6)
- ★★★R入門資料分析與視覺化(付費,字幕)
- ★★★R商業預測與應用(付費,字幕)
- iPAS-R-tutorial(繪圖中文字型solved)
- iPAS-Python-tutorial
- R教學-基礎篇/程式碼(免費)
- Python程式設計PDF(免費)
- ★R 4.1.2-Wndows下載 1
- ★RStudio Desktop-2021.09.0+351下載 2

- R-4.1.2-win.exe
- RStudio-2021.09.0+351.exe

基礎互動式shiny

- <https://shiny.rstudio.com/>
- <https://www.rstudio.com/resources/webinars/>

自動安裝套件

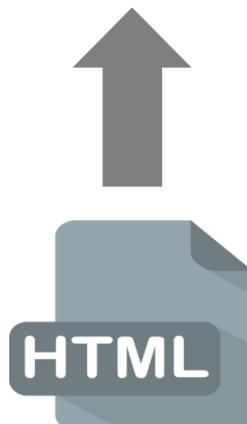
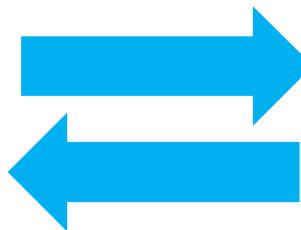
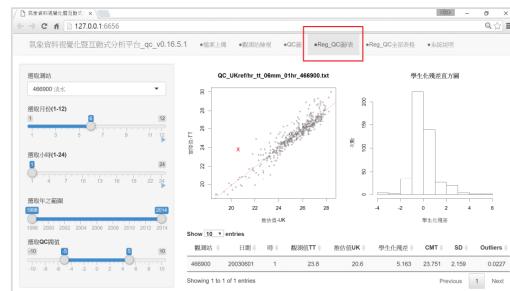
```
# 自動安裝套件 -----
usedpackages <- c("shiny", "ggplot2")

verify.packages <- function(need.packages) {
  for (x in need.packages) {
    if (!x %in% installed.packages()[,"Package"])
      install.packages(x)
  }
}
verify.packages(usedpackages)
```

shiny 簡介

- 隨著資訊科技技術的進步，如何提供WEB化的應用服務。
- shiny 套件提供更方便，互動式與動態更新等應用
 - 輸入 - 文字方塊、下拉式選單、按鈕。
 - 處理 - 執行 R 運算、建立模型。
 - 輸出 - 網頁中呈現文字摘要、表格與圖表等。

shiny (server) 架構



ui.R



server.R

CRAN shiny 套件

shiny: Web Application Framework for R

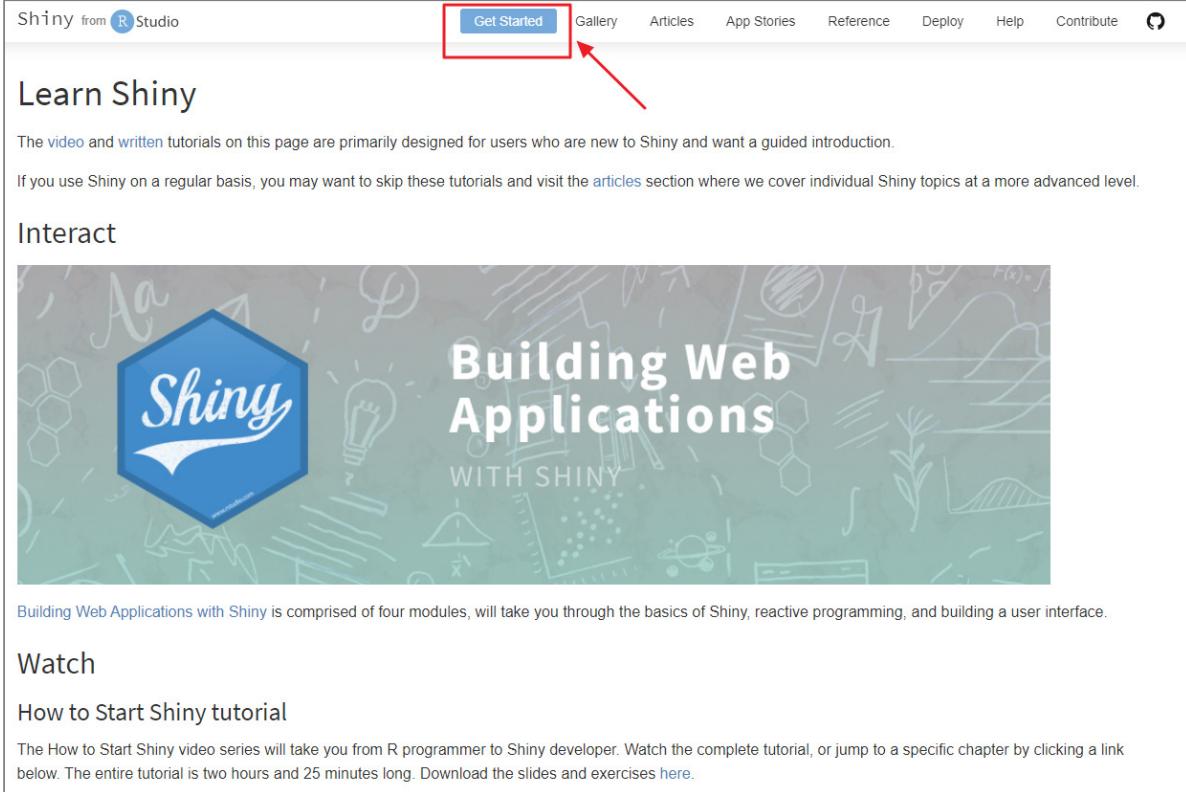
互動式網頁應用

Makes it incredibly easy to build interactive web applications with R. Automatic "reactive" binding between inputs and outputs and extensive prebuilt widgets make it possible to build beautiful, responsive, and powerful applications with minimal effort.

Version:	1.7.1
Depends:	R (\geq 3.0.2), methods
Imports:	utils, grDevices, httpuv (\geq 1.5.2), mime (\geq 0.3), jsonlite (\geq 0.9.16), xtable , fontawesome (\geq 0.2.1), htmltools (\geq 0.5.2), R6 (\geq 2.0), sourcetools , later (\geq 1.0.0), promises (\geq 1.1.0), tools, crayon , rlang (\geq 0.4.10), fastmap (\geq 1.1.0), withr , commonmark (\geq 1.7), glue (\geq 1.3.2), bslib (\geq 0.3.0), cachem , ellipsis , lifecycle (\geq 0.2.0)
Suggests:	datasets, Cairo (\geq 1.5-5), testthat (\geq 3.0.0), knitr (\geq 1.6), markdown , rmarkdown , ggplot2 , reactlog (\geq 1.0.0), magrittr , shinytest (\geq 1.4.0.9003), yaml , future , dygraphs , ragg , showtext , sass
Published:	2021-10-02
Author:	Winston Chang  [aut, cre], Joe Cheng [aut], JJ Allaire [aut], Carson Sievert  [aut], Barret Schloerke  [aut], Yihui Xie [aut], Jeff Allen [aut], Jonathan McPherson [aut], Alan Dipert [aut], Barbara Borges [aut], RStudio [cph], jQuery Foundation [cph] (jQuery library and jQuery UI library),

Shiny 參考資料-Get Started

- <http://shiny.rstudio.com/tutorial/>



The screenshot shows the 'Get Started' section of the Shiny website. The top navigation bar includes links for 'Get Started' (which is highlighted with a red box and a red arrow pointing to it), 'Gallery', 'Articles', 'App Stories', 'Reference', 'Deploy', 'Help', and 'Contribute'. Below the navigation, there's a heading 'Learn Shiny' followed by a paragraph about video and written tutorials. There's also a note for regular users about skipping the 'articles' section. The main visual is a large image of the 'Building Web Applications WITH SHINY' book cover, which features the Shiny logo and various scientific and technical illustrations. Below the image, a caption describes the book's four modules covering basics, reactive programming, and user interfaces. At the bottom, there's a 'Watch' section for the 'How to Start Shiny tutorial'.

Learn Shiny

The video and written tutorials on this page are primarily designed for users who are new to Shiny and want a guided introduction.

If you use Shiny on a regular basis, you may want to skip these tutorials and visit the [articles](#) section where we cover individual Shiny topics at a more advanced level.

Interact



**Building Web Applications
WITH SHINY**

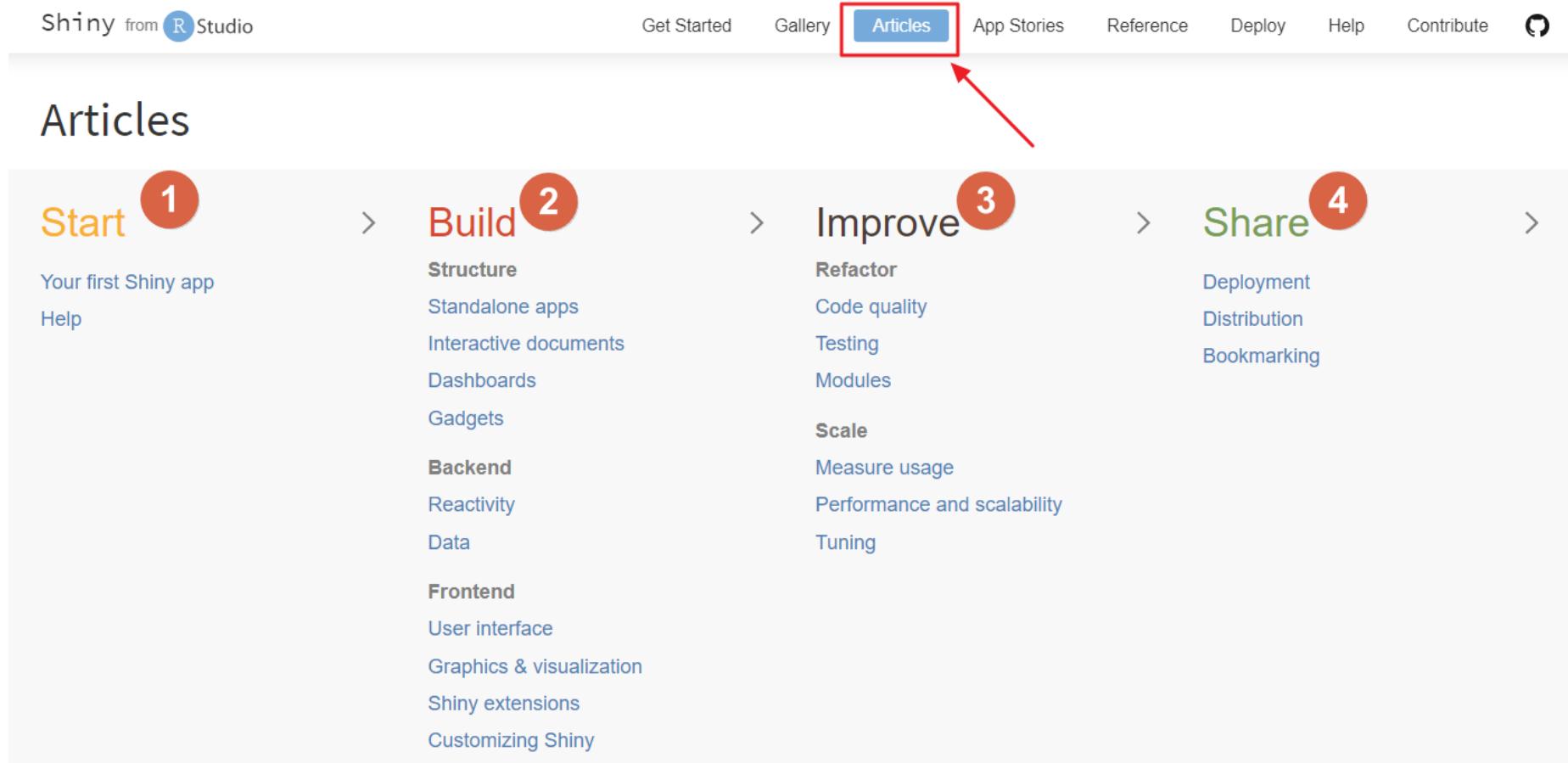
Building Web Applications with Shiny is comprised of four modules, will take you through the basics of Shiny, reactive programming, and building a user interface.

Watch

[How to Start Shiny tutorial](#)

The How to Start Shiny video series will take you from R programmer to Shiny developer. Watch the complete tutorial, or jump to a specific chapter by clicking a link below. The entire tutorial is two hours and 25 minutes long. Download the slides and exercises [here](#).

Shiny 參考資料-Articles



Shiny from R Studio

Get Started Gallery **Articles** App Stories Reference Deploy Help Contribute 

Articles

Start 1 > Build 2 > Improve 3 > Share 4 >

Start Your first Shiny app Help	Build Structure Standalone apps Interactive documents Dashboards Gadgets Backend Reactivity Data	Improve Refactor Code quality Testing Modules Scale Measure usage Performance and scalability Tuning	Share Deployment Distribution Bookmarking
--	--	--	---

Frontend
User interface
Graphics & visualization
Shiny extensions
Customizing Shiny

shiny 內建範例資料

- 範例(examples)目錄

`R\R-4.x.x\library\shiny\examples`

```
> # 顯示內建11個範例
> dir(paste0(.libPaths(), "/shiny/examples"))
[1] "01_hello"        "02_text"         "03_reactivity"  "04_mpg"
[5] "05_sliders"      "06_tabssets"     "07_widgets"     "08_html"
[9] "09_upload"       "10_download"    "11_timer"
>
```

內建11個shiny範例

shiny 範例 - 01_hello

```
> # shiny example - 01_hello  
> library(shiny)  
> runExample("01_hello")
```

- 左側: user-interface

Hello Shiny!

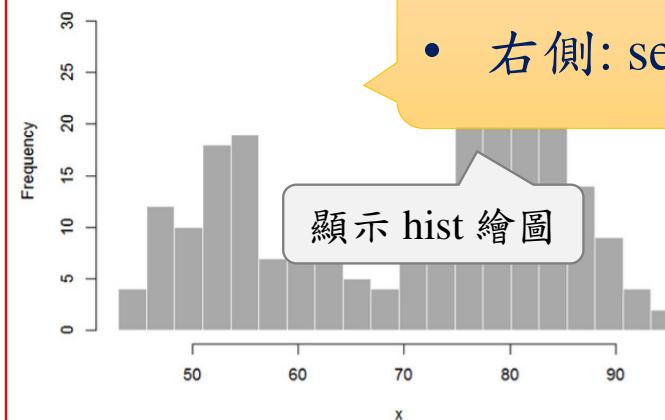
Number of bins:



slider: 可調整樣本大小

- runExample("範例名稱")
- 區分ui.R, server.R
- 亦可合併為 app.R

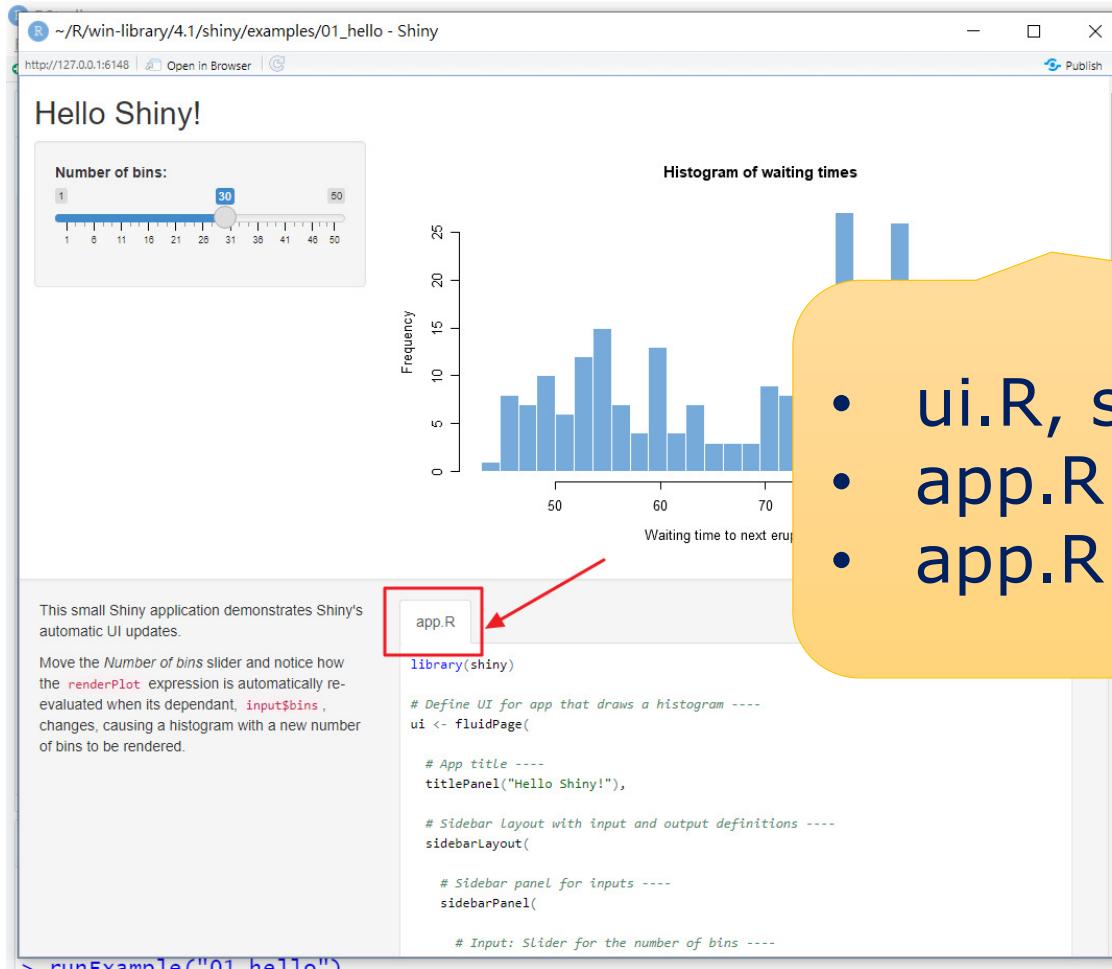
Histogram of x



顯示 hist 繪圖

- 右側: server-interface

方法1-app.R



The screenshot shows a Shiny application window titled "Hello Shiny!". On the left, there is a sidebar with a slider titled "Number of bins" set to 30. The main area displays a histogram titled "Histogram of waiting times" with the x-axis labeled "Waiting time to next eruption" and the y-axis labeled "Frequency". The histogram has several peaks, with the highest peak around 50-55 units. Below the histogram, a red arrow points to the word "app.R" in the code editor, which is highlighted with a red box.

```
library(shiny)

# Define UI for app that draws a histogram ----
ui <- fluidPage(
  # App title ----
  titlePanel("Hello Shiny!"),

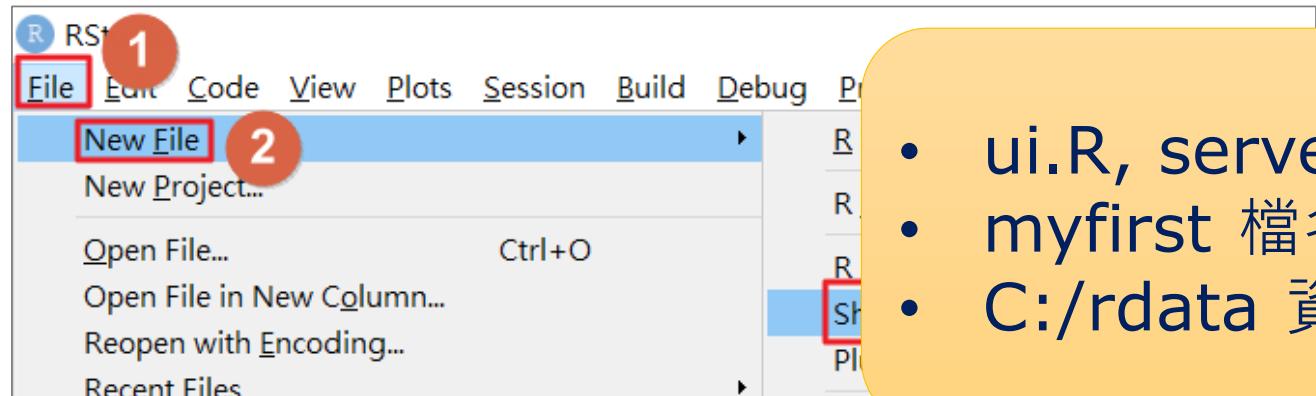
  # Sidebar layout with input and output definitions ----
  sidebarLayout(
    # Sidebar panel for inputs ----
    sidebarPanel(
      # Input: Slider for the number of bins ----
      sliderInput("bins", "Number of bins:", min = 1, max = 50, value = 30)
    ),
    # Main panel for displaying outputs ----
    mainPanel(
      # Output: Histogram ----
      renderPlot(histogram)
    )
  )
)
```

> runExample("01_hello")

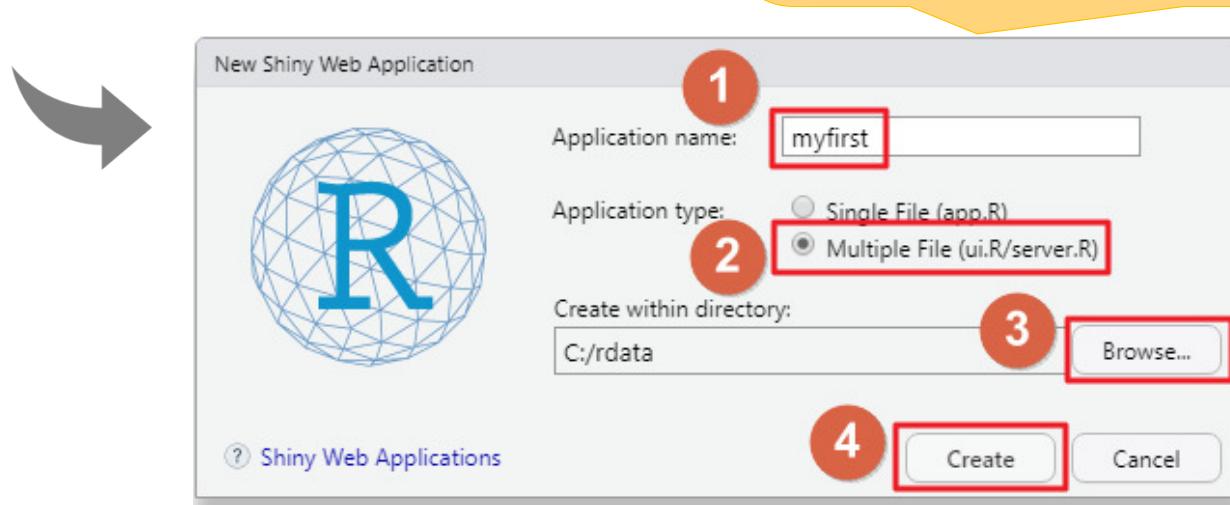
Listening on http://127.0.0.1:6148

- ui.R, server.R 檔名不可更換
- app.R 檔名可以更換
- app.R 適用於小程序式,方便測試

方法2-Multiple File(ui.R/server.R)



- ui.R, server.R 檔名不可更換
- myfirst 檔名可以更換
- C:/rdata 資料夾名稱可更換



ui.R

```

10 library(shiny)
11
12 # Define UI for application that draws a histogram
13 shinyUI(fluidPage(
14
15   # Application title
16   titlePanel("old Faithful Geyser Data"),
17
18   # Sidebar with a slider input for number of bins
19   sidebarLayout(
20     sidebarPanel(
21       sliderInput("bins",
22         "Number of bins:",
23         min = 1,
24         max = 50,
25         value = 30),
26
27
28     # Show a plot of the generated distribution
29     mainPanel(
30       plotOutput("distPlot")
31     )
32   )
33 ))

```

fluidPage()

1

titlePanel: 標題

2

sidebarInput: 滑桿(輸入功能)

輸入參數 bins
value 預設值

輸出採用 distPlot 參數

server.R

```
10 library(shiny)
11
12 # Define server logic required to draw a histogram
13 shinyServer(function(input, output) {
14
15   output$distPlot <- renderPlot({ • distPlot 繪圖參數
16
17     # generate bins based on input$bins from ui.R
18     x      <- faithful[, 2]
19     bins <- seq(min(x), max(x), length.out = input$bins + 1)
20
21     # draw the histogram with the specified number of bins
22     hist(x, breaks = bins, col = 'darkgray', border = 'white')
23
24   })
25
26 })
```

- distPlot 繪圖參數

renderPlot

- 輸入參數: **x** 內建資料集 faithful
- 輸入參數: 使用者輸入 **bins**

shiny 範例 - 02_text

Shiny Text

Choose a dataset:

rock

Number of observations to view:

10

area	peri	shape	perm
Min. : 1016	Min. : 308.6	Min. : 0.09033	Min. : 6.30
1st Qu.: 5305	1st Qu.: 1414.9	1st Qu.: 0.16226	1st Qu.: 76.45
Median : 7487	Median : 2536.2	Median : 0.19886	Median : 130.50
Mean : 7188	Mean : 2682.2	Mean : 0.21811	Mean : 415.45
3rd Qu.: 8870	3rd Qu.: 3989.5	3rd Qu.: 0.26267	3rd Qu.: 777.50
Max. : 12212	Max. : 4864.2	Max. : 0.46413	Max. : 1300.00

area	peri	shape	perm
4990	2791.90	0.09	6.30
7002	3892.60	0.15	6.30
7558	3930.66	0.18	6.30
7352	3869.32	0.12	6.30
7943	3948.54	0.12	17.10
7979	4010.15	0.17	17.10
9333	4345.75	0.19	17.10
8209	4344.75	0.16	17.10
8393	3682.04	0.20	119.00
6425	3098.65	0.16	119.00

shiny 範例 - 03_reactivity

Reactivity

The diagram illustrates the reactive relationship in a Shiny application. A red arrow points from the 'Caption' input field on the left to the 'Data Summary-RWEPA網站2021' output table on the right, indicating that the output is dependent on the input.

Caption:
Data Summary-RWEPA網站2021

Choose a dataset:
rock

Number of observations to view:
5

Data Summary-RWEPA網站2021

area	peri	shape	perm
Min. : 1016	Min. : 308.6	Min. : 0.09033	Min. : 6.30
1st Qu.: 5305	1st Qu.: 1414.9	1st Qu.: 0.16226	1st Qu.: 76.45
Median : 7487	Median : 2536.2	Median : 0.19886	Median : 130.50
Mean : 7188	Mean : 2682.2	Mean : 0.21811	Mean : 415.45
3rd Qu.: 8870	3rd Qu.: 3989.5	3rd Qu.: 0.26267	3rd Qu.: 777.50
Max. : 12212	Max. : 4864.2	Max. : 0.46413	Max. : 1300.00

area	peri	shape	perm
4990	2791.90	0.09	6.30
7002	3892.60	0.15	6.30
7558	3930.66	0.18	6.30
7352	3869.32	0.12	6.30
7943	3948.54	0.12	17.10

輸入控制項

輸入控制項

函數	功能
actionButton	每次按鈕+1
checkboxGroupInput	核取方塊-群組
checkboxInput	核取方塊 (用於YES/NO)
dateInput	選取單一日期
dateRangeInput	選取日期範圍 (開始-結束)
fileInput	上傳檔案
helpText	輔助說明
numericInput	數值輸入
radioButtons	選取按鈕 (單選題)
selectInput	下拉式選單
sliderInput	左右滑桿
submitButton	送出按鈕
textInput	文字方塊

selectInput 下拉式選單

selectInput {shiny}

R Documentation

Create a select list input control

Description

Create a select list that can be used to choose a single or multiple items from a list of values.

Usage

① ② ③

```
selectInput(inputId, label, choices, selected = NULL, multiple = FALSE,  
           selectize = TRUE, width = NULL, size = NULL)
```

```
selectizeInput(inputId, ..., options = NULL, width = NULL)
```

Arguments

inputId

The input slot that will be used to access the value.

label

Display label for the control, or `NULL` for no label.

choices

List of values to select from. If elements of the list are named then that name rather than the value is displayed to the user.

輸入資料須為list

selectInput 下拉式選單(續)

```
1 ## selectInput example
2 library(shiny)
3 if (interactive()) {
4
5   ui <- fluidPage(
6     selectInput("variable", "Variable:",
7                 c("Cylinders" = "cy1",
8                  "Transmission" = "am",
9                  "Gears" = "gear"))),
10    tableOutput("data")
11  )
12
13  server <- function(input, output) {
14    output$data <- renderTable({
15      mtcars[, c("mpg", input$variable), drop = FALSE]
16    }, rownames = TRUE)
17  }
18
19 shinyApp(ui, server)
20 }
```



萬一資料
筆數很多!@#

輸出控制項-文字/表報



render函數對照表

使用時機	ui.R (依字母排序)	server.R
table	dataTableOutput	renderDataTable
raw HTML	htmlOutput	renderUI
image	imageOutput	renderImage
plot	plotOutput	renderPlot
table	tableOutput	renderTable
text	textOutput	renderText
raw HTML	uiOutput	renderUI
text	verbatimTextOutput	renderPrint

版面配置

- <http://shiny.rstudio.com/articles/layout-guide.html>

1. 側邊佈局 Sidebar Layout

Hello Shiny!

Number of bins:

1

30

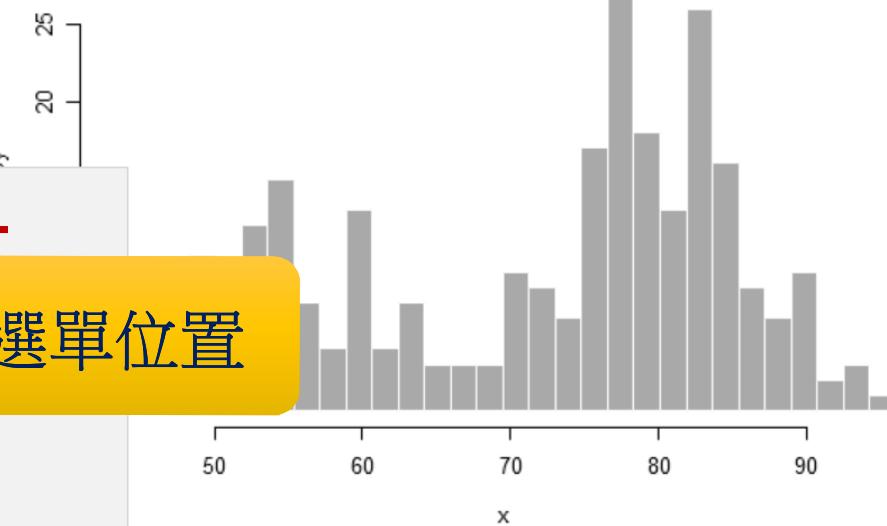
50

```
sidebarLayout(position = "right",
```

```
  sidebarPanel(  
    # Inputs  
  mainPanel(  
    # Outputs  
)
```

調整選單位置

Histogram of x

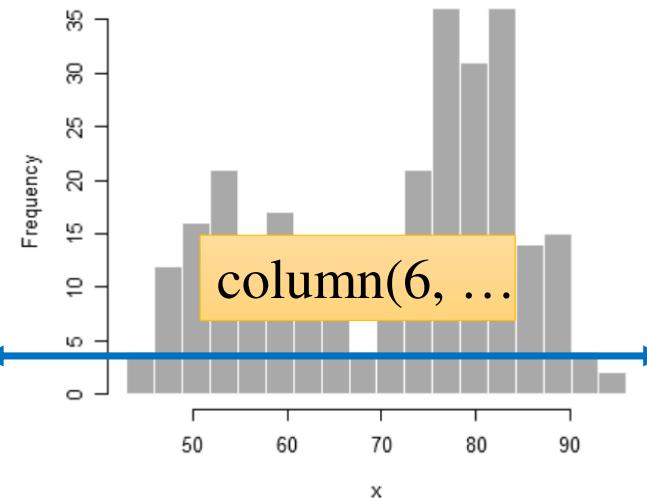


2. 網格佈局 Grid Layout

Hello Shiny - Grid Layout!

Number of bins:

Histogram of x



column(6, ...)

column(6, ...)

Grid Layout – ui.R

```
library(shiny)

# Define UI for application that draws a histogram
shinyUI(fluidPage(
  titlePanel("Hello Shiny - Grid Layout!"),
  fluidRow(
    column(6,
      wellPanel(
        sliderInput("bins", "Number of bins:",
                   min = 1, max = 50, value = 30)
      )
    ),
    column(6,
      plotOutput("distPlot")
    )
  )))

```

- $6+6=12$
- try: $4+8=12$

3.分頁佈局 Tabsets Layout

Tabsets

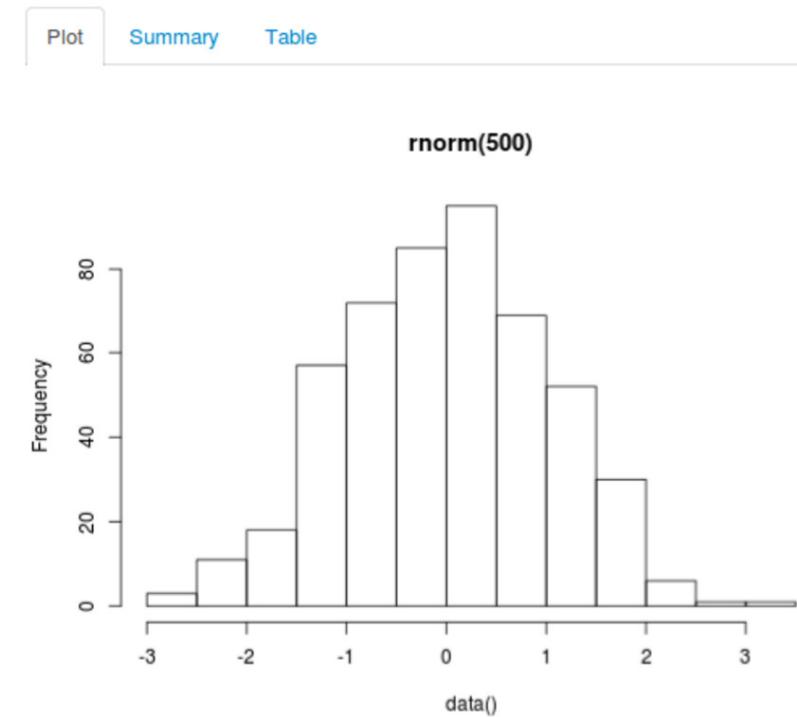
1 2 3

Distribution type:

- Normal
- Uniform
- Log-normal
- Exponential

Number of observations:

1 500 1,000



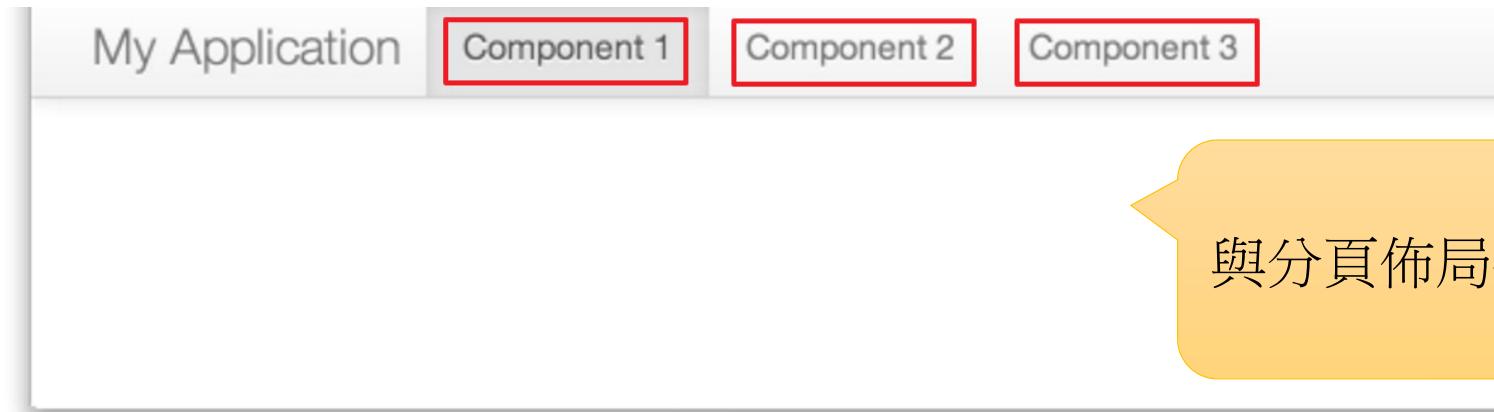
Tabssets Layout – ui.R

```
shinyUI(fluidPage(  
  titlePanel("Tabssets"),  
  sidebarLayout(  
    sidebarPanel(  
      # Inputs excluded for brevity  
    ),  
    mainPanel(  
      tabsetPanel(  
        tabPanel("Plot", plotOutput("plot")),  
        tabPanel("Summary", verbatimTextOutput("summary")),  
        tabPanel("Table", tableOutput("table"))  
      )  
    )  
  ))
```

tabsetPanel()

3個分頁 tabPanel

4. 瀏覽選單佈局 Navbar Pages

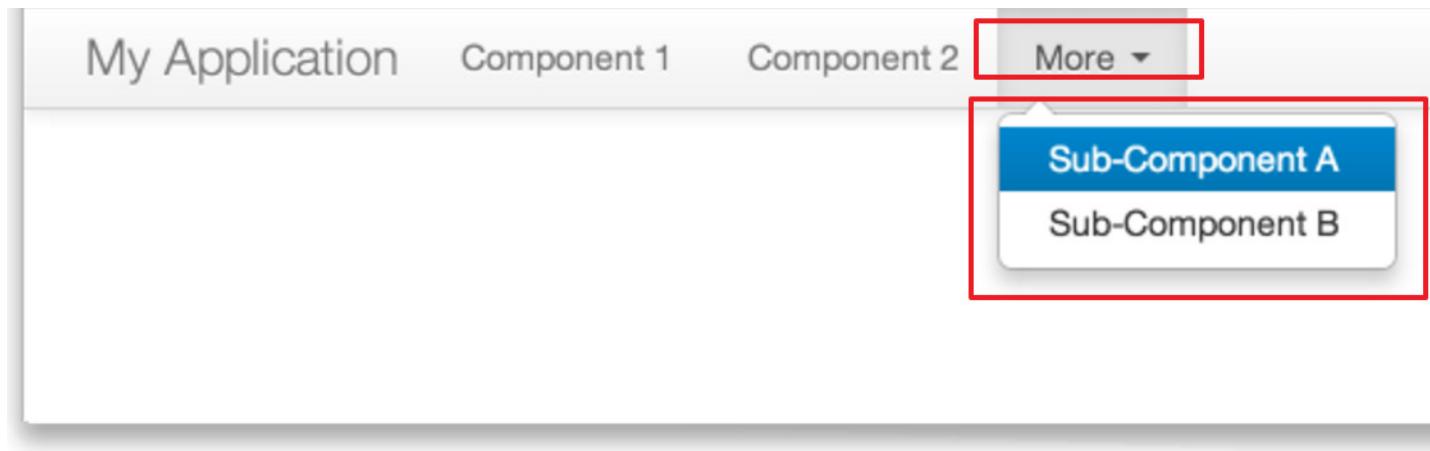


My Application Component 1 Component 2 Component 3

與分頁佈局有何不同?

```
shinyUI(navbarPage("My Application",
  tabPanel("Component 1"),
  tabPanel("Component 2"),
  tabPanel("Component 3")
))
```

5. 瀏覽下拉式選單佈局



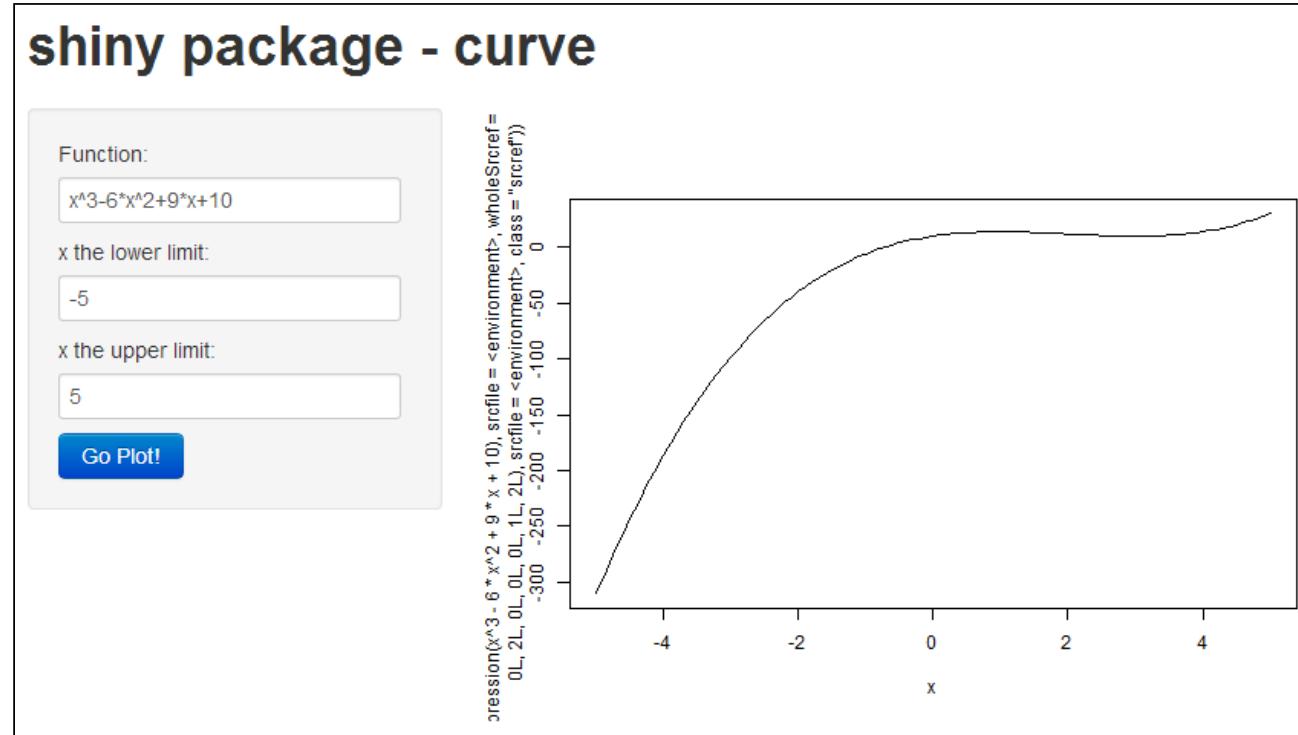
The screenshot shows a top navigation bar with tabs labeled "My Application", "Component 1", "Component 2", and "More". A red box highlights the "More" tab, which has a dropdown menu. The dropdown menu contains two items: "Sub-Component A" and "Sub-Component B", also highlighted with a red box.

```
shinyUI(navbarPage("My Application",
  tabPanel("Component 1"),
  tabPanel("Component 2"),
  navbarMenu("More",
    tabPanel("Sub-Component A"),
    tabPanel("Sub-Component B"))
))
```

navbarMenu()

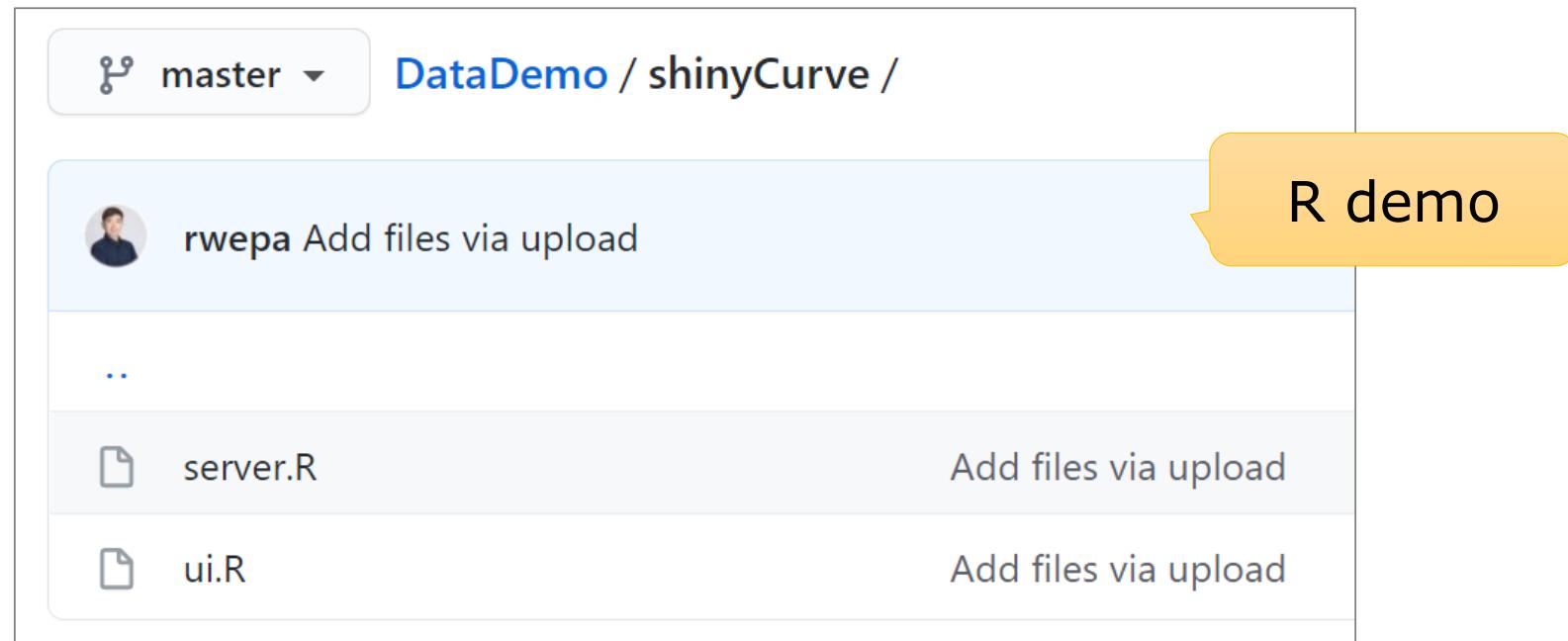
shiny 範例 – shinyCurve - 展示

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



shiny 範例 – shinyCurve - 程式碼

- <https://github.com/rwepa/DataDemo/tree/master/shinyCurve>
- 將 server.R 與 ui.R 複製到工作目錄的單一資料夾中, 例如
「C:\rdata\shinyCurve」



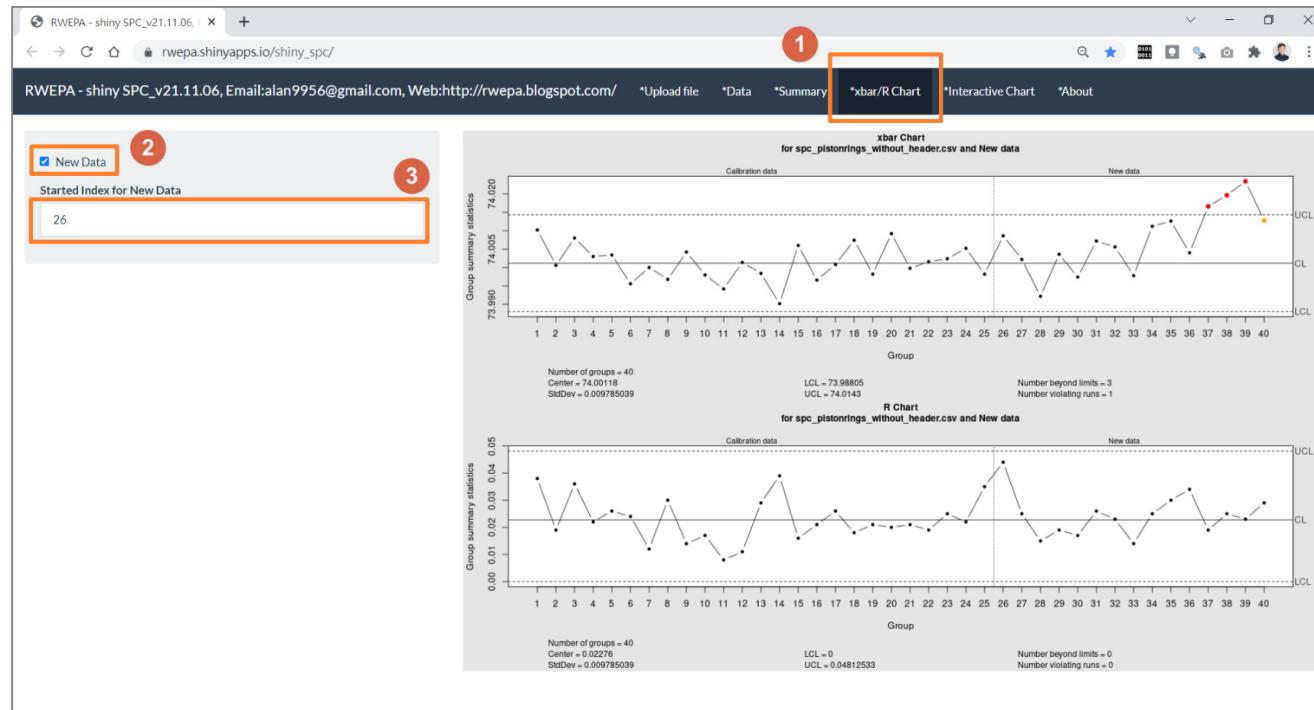
3.互動式統計品管網頁實作

使用套件

Packages	Feature
shiny	Web Application Framework for R
shinythemes	Themes for Shiny
DT	DataTables JavaScript library
qcc	Quality Control Charts
plotly	Create Interactive Web Graphics via 'plotly.js'

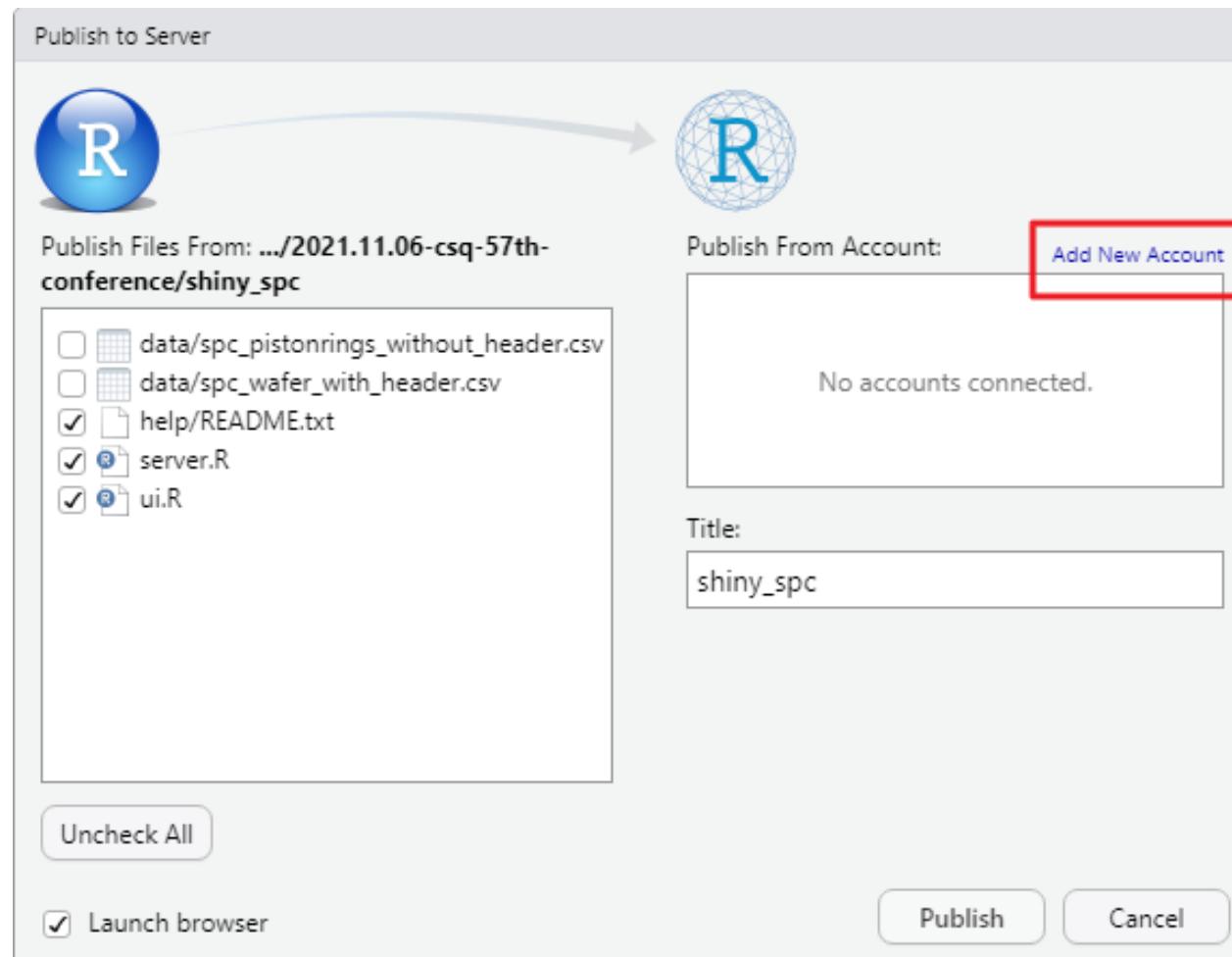
R shiny - 品質管制圖(quality control chart)應用

- <http://rwepa.blogspot.com/2021/10/r-shiny-quality-control-chart.html>
- R 程式碼下載 https://github.com/rwepa/shiny_spc
- 線上示範 https://rwepa.shinyapps.io/shiny_spc/



R demo

發佈至 rwepa.shinyapp.io\shiny_spc



Connect Account

Connect Account

Connect Publishing Account



To publish content, you first need to connect RStudio to an account on the service you want to publish to.

Once you've authorized this computer to publish content to an account, you can publish any time without re-entering your credentials.

Next

Cancel

ShinyApps.io

Connect Account

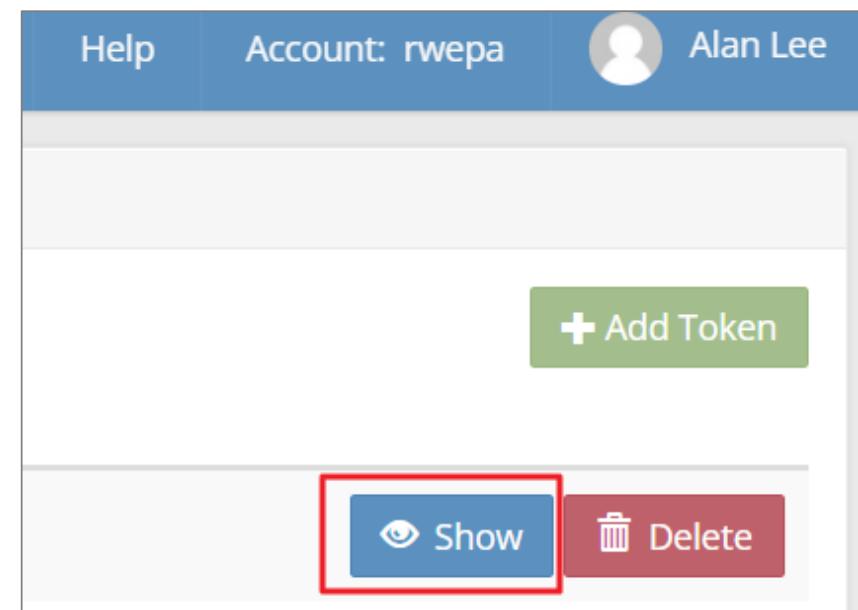
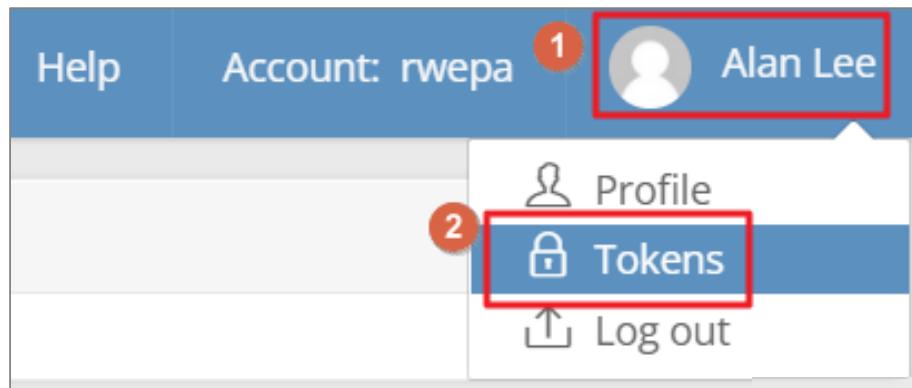
Back Connect Account

**ShinyApps.io**
A cloud service run by RStudio. Publish Shiny applications and interactive documents to the Internet.

**RStudio Connect**
RStudio Connect is a server product from RStudio for secure sharing of applications, reports, plots, and APIs.

Cancel

複製 Tokens



按 CTRL + C

貼上 Tokens

Connect Account

Back Connect ShinyApps.io Account



Go to [your account on ShinyApps](#) and log in.
Click your name, then choose **Tokens** from your account menu.
Click **Show** on the token you want to use, then **Show Secret** and **Copy to Clipboard**. Paste the result here:

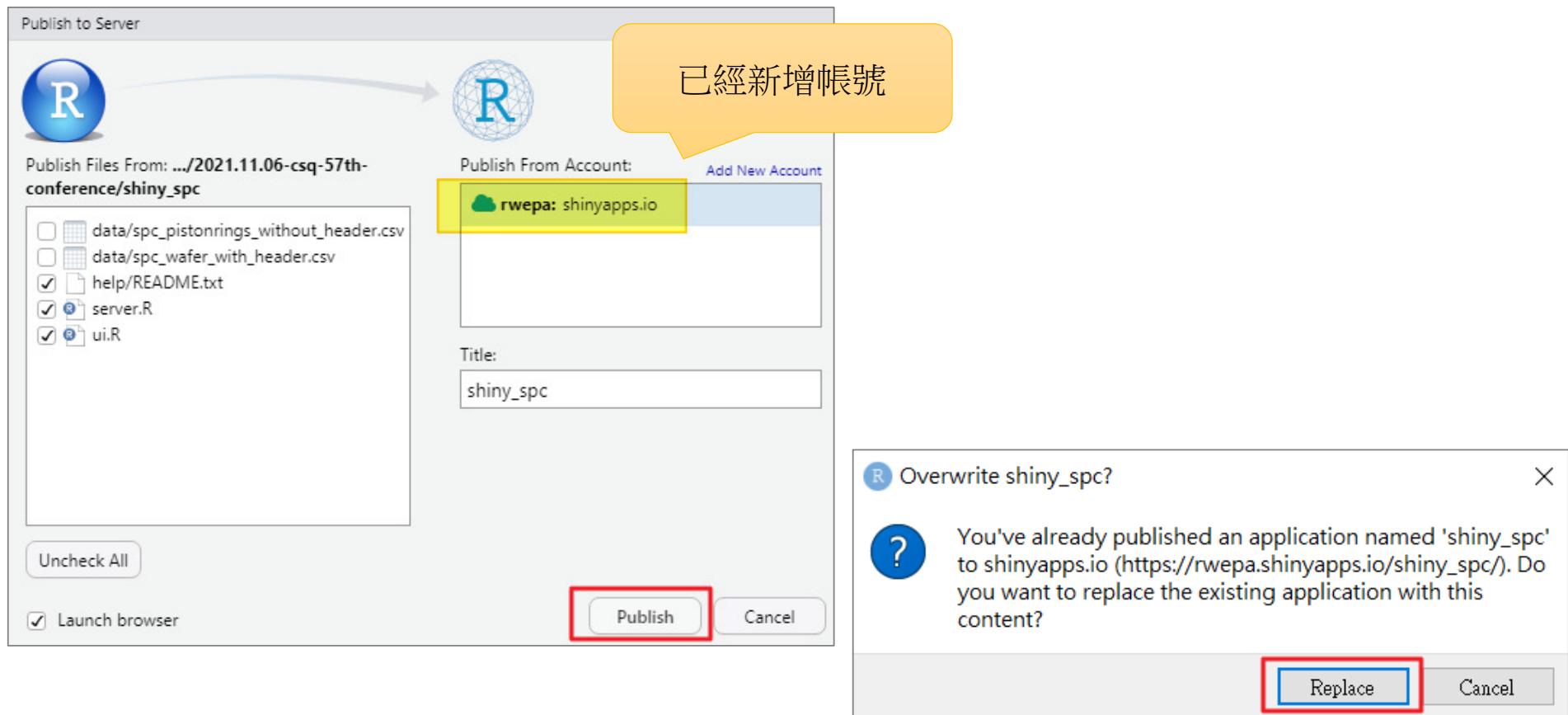


按 CTRL + V

Need a ShinyApps.io account? [Get started here.](#)

按 Publish

Publish to Server



已經新增帳號

Publish Files From: .../2021.11.06-csq-57th-conference/shiny_spc

- data/spc_pistonrings_without_header.csv
- data/spc_wafer_with_header.csv
- help/README.txt
- server.R
- ui.R

Uncheck All

Launch browser

Publish Cancel

Publish From Account: Add New Account

rwepa: shinyapps.io

Title: shiny_spc

Overwrite shiny_spc?

You've already published an application named 'shiny_spc' to shinyapps.io (https://rwepa.shinyapps.io/shiny_spc/). Do you want to replace the existing application with this content?

Replace Cancel

Python 模組

模組	功能	
Numpy	Large, multi-dimensional arrays and matrices	
Scipy	Optimization, linear algebra, integration, FFT, signal	
Pandas	DataFrame object for data manipulation	
Matplotlib	Static, animated, and interactive visualizations	
Statsmodels	Statistical models	
Scikit-learn	Machine learning library	
Tensorflow	Deep learning	
Biopython	Biological computation	
Scanpy	Single-cell analysis	
Django, Flask	Web	
Plotly, dash, bokeh	Interactive visualization	  

R 套件

模組	功能	
dplyr	A grammar of data manipulation	
data.table	Extension of data.frame	
ggplot2	Create Elegant Data Visualizations Using the Grammar of Graphics	
shiny	Web application framework	
caret	Classification and Regression Training	
mlr3	Provides R6 objects for efficient, object-oriented programming on the building blocks of machine learning	
Tensorflow	R Interface to 'TensorFlow'	
Bioconductor	Tools for the analysis genomic data	
SingleCellExperiment	Orchestrating Single-Cell Analysis with Bioconductor	

參考資料

- RWEPA
 - <http://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)
- Python 程式設計-李明昌 <免費電子書>
 - <http://rwepa.blogspot.com/2020/02/pythonprogramminglee.html>
- 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

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