

介紹R與R基本操作

2018/9/20

為什麼要用R

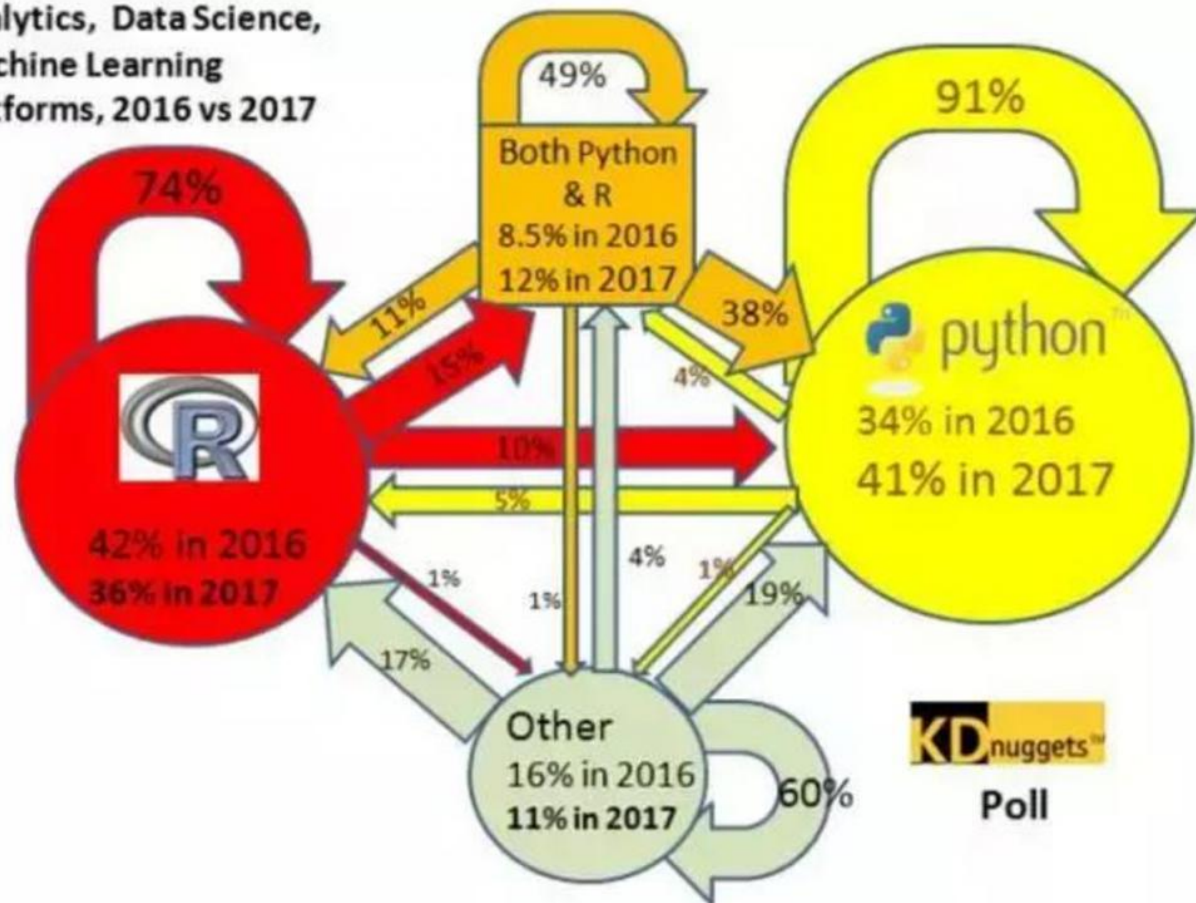
- 免費、開源
- 專門為統計和資料分析開發的語言，各種功能和函數琳琅滿目
- 安裝程式只有50Mb左右，因為體積輕便，運行起來系統負擔也小
- 各種OS的相容性好
- 因為用的人越來越多，又是開源，有很多“套件程式”為其錦上添花
- 有R GUI和RStudio兩種風格選擇
- 漂亮又靈活的圖

想挖出 **BIG DATA** 的秘密，這 **10** 個程式語言你不能不懂

- 若要列出所有程式語言，你能忘記其他的沒關係，但最不能忘的就是 **R**。從 **1997** 年悄悄地出現，最大的優勢就是它免費，為昂貴的統計軟體像是 **Matlab** 或 **SAS** 的另一種選擇。
- 但是在過去幾年來，它的身價大翻轉，變成了資料科學界眼中的寶。不只是木訥的統計學家熟知它，包括 **Wall Street** 交易員、生物學家，以及矽谷開發者，他們都相當熟悉 **R**。多元化的公司像是 **Google**、**Facebook**、美國銀行以及 **New York Times** 通通都使用 **R**，它的商業效用持續提高。
- R** 的好處在於它簡單易上手，透過 **R**，你可以從複雜的資料集中篩選你要的資料，從複雜的模型函數中操作資料，建立井然有序的圖表來呈現數字，這些都只需要幾行程式碼就可以了，打個比方，它就像是好動版本的 **Excel**。
- R** 最棒的資產就是活躍的動態系統，**R** 社群持續地增加新的軟件包，還有以內建豐富的功能集為特點。目前估計已有超過 **200** 萬人使用 **R**，最近的調查顯示，**R** 在資料科學界裡，到目前為止最受歡迎的語言，佔了回覆者的 **61%**（緊迫在後的是 **39%** 的 **Python**）。
- 它也吸引了 **Wall Street** 的注目。傳統而言，證券分析師在 **Excel** 檔從白天看到晚上，但現在 **R** 在財務建模的使用率逐漸增加，特別是視覺化工具，美國銀行的副總裁 **Niall O'Conno** 說，「**R** 讓我們俗氣的表格變得突出」。

R or Python ?

Analytics, Data Science,
Machine Learning
Platforms, 2016 vs 2017



ML 相關套件

Approach	Algorithm	R package	Python library
Regression	OLS	stats	scikit-learn
	Lasso	lars	scikit-learn
	SGD	sgd	scikit-learn
	SVR	e1071	scikit-learn
Classification	LDA	MASS	scikit-learn
	Random Forest	randomForest	scikit-learn
	KNN	class	scikit-learn
	SVM	e1071	scikit-learn
Clustering	K-means	stats	scikit-learn
	Hierarchical Clustering	stats	scikit-learn
	DBSCAN	dbscan	scikit-learn
	Spectral Clustering	kernlab	scikit-learn
Dimensionality reduction	PCA	prcomp	scikit-learn
	KPCA	kernlab	scikit-learn
	Isomap	vegan	scikit-learn
	LLE	lle	scikit-learn

安裝R



[Home]

Download

CRAN

R Project

About R

Logo

Contributors

What's New?

Reporting Bugs

Development Site

Conferences

Search

R Foundation

Foundation

Board

Members

Donors

Donate

Help With R

Getting Help

Documentation

The R Project for Statistical Computing

Getting Started

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

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

News

R version 3.3.3 (Another Canoe) has been released on Monday 2017-03-06.

- **useR! 2017** (July 4 - 7 in Brussels) has opened registration and more at <http://user2017.brussels/>

- Tomas Kalibera has joined the R core team. <https://fourdots.com/mirror/CRAN/>

- The R Foundation welcomes five new ordinary members: Tomas Kalibera, and Balasubramanian Narasimhan. <http://cran.stat.nus.edu.sg/>

- **The R Journal Volume 8/1** is available. <http://r.journal.r-project.org/>

- The **useR! 2017** conference will take place in Brussels. <http://cran.mirror.ac.za/>

- **R version 3.2.5 (Very, Very Secure Dishes)** has been released. <https://ftp.cixug.es/CRAN/>

- **Notice XQuartz users (Mac OS X)** A security issue has been discovered in the mechanism used by XQuartz. Avoid updating over insecure channels. <https://cran.rediris.es/>

- The **R Logo** is available for download in high-resolution. <http://cran.rediris.es/>

- **useR! 2016**, has taken place at Stanford University, CA. <https://ftp.acc.umu.se/mirror/CRAN/>

- <http://ftp.acc.umu.se/mirror/CRAN/>

Switzerland

<https://stat.ethz.ch/CRAN/>

<http://stat.ethz.ch/CRAN/>

Taiwan

<https://ftp.yzu.edu.tw/CRAN/>

<http://ftp.yzu.edu.tw/CRAN/>

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

Thailand

<http://mirrors.psu.ac.th/pub/cran/>

Turkey

<https://cran.pau.edu.tr/>

<http://cran.pau.edu.tr/>

<https://cran.ncc.metu.edu.tr/>

<http://cran.ncc.metu.edu.tr/>

Four Dots

National University of Singapore, Singapore

University of Cape Town

TENET, Johannesburg

Oficina de software libre (CIXUG)

Oficina de software libre (CIXUG)

Spanish National Research Network, Madrid

Spanish National Research Network, Madrid

Academic Computer Club, Umeå University

Academic Computer Club, Umeå University

ETH Zürich

ETH Zürich

Department of Computer Science and Engineering, Yuan Ze University

Department of Computer Science and Engineering, Yuan Ze University

National Taiwan University, Taipei

Prince of Songkla University, Hatyai

Pamukkale University, Denizli

Pamukkale University, Denizli

Middle East Technical University Northern Cyprus Campus, Mersin

Middle East Technical University Northern Cyprus Campus, Mersin

The Comprehensive R Archive Network

Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows and Mac** users most likely want one of these versions of R:

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

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use them. If you do not know what this means, you probably do not want to do it!

- The latest release (Monday 2017-03-06. Another Canoe) [R-3.3.3.tar.gz](#), read [what's new](#) in the latest version.

R for Windows

Subdirectories:

[base](#)

Binaries for base distribution (managed by Duncan Murdoch). This is what you want to [install R for the first time](#).

[contrib](#)

Binaries of contributed CRAN packages (for R >= 2.11.x; managed by Uwe Ligges). There is also information on [third party software](#) available for CRAN Windows services and corresponding environment and make variables.

[old contrib](#)

R-3.3.3 for Windows (32/64 bit)

[Rtools](#)

[Download R 3.3.3 for Windows](#) (71 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

Please do not sut

or to

d to Windows binaries.

You may also wa

If you want to double-check that the package you have downloaded matches the package distributed by CRAN, you can compare the [md5sum](#) of the .exe to the [fingerprint](#) on the master server. You will need a version of md5sum for windows: both [graphical](#) and [command line versions](#) are available.

Note: CRAN doe

Frequently asked questions

- [Does R run under my version of Windows?](#)
- [How do I update packages in my previous version of R?](#)
- [Should I run 32-bit or 64-bit R?](#)

Please see the [R FAQ](#) for general information about R and the [R Windows FAQ](#) for Windows-specific information.

Other builds

- Patches to this release are incorporated in the [r-patched snapshot build](#).
- A build of the development version (which will eventually become the next major release of R) is available in the [r-devel snapshot build](#).
- [Previous releases](#)

Note to webmasters: A stable link which will redirect to the current Windows binary release is [<CRAN MIRROR>/bin/windows/base/release.htm](#).

RStudio



[rstudio::conf](#)

[Products](#)

[Resources](#)

[Pricing](#)

[About Us](#)

[Blogs](#)



RStudio

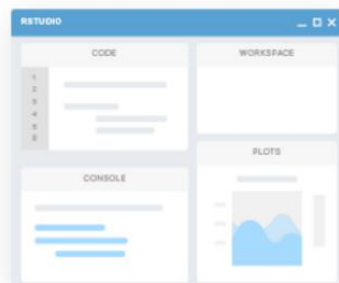
Open source and enterprise-ready
professional software for R

[Download RStudio](#)

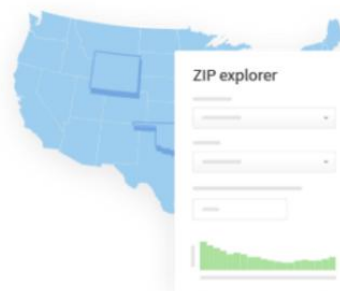
[Discover Shiny](#)

[shinyapps.io Login](#)

[Discover RStudio Connect](#)



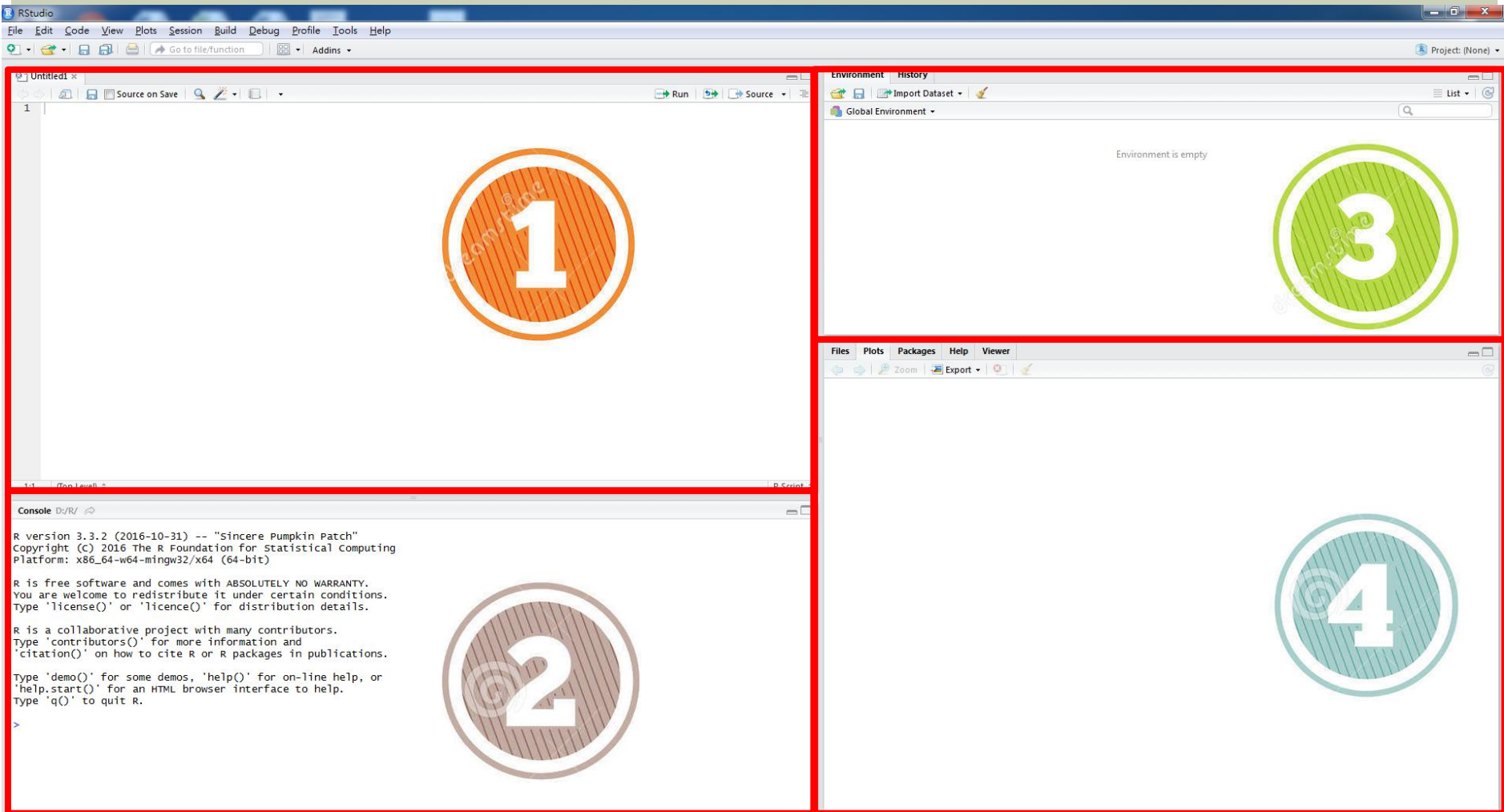
RStudio



Shiny



R Packages



你可以把R當作計算機

```
> 6+6
```

```
[1] 12
```

```
> 6-3
```

```
[1] 3
```

```
> 6*6
```

```
[1] 36
```

```
> 6/3
```

```
[1] 2
```

```
> 2^10
```

```
[1] 1024
```

```
> 9487%%10
```

```
[1] 7
```

```
> 9487%/10
```

```
[1] 948
```

變數定義與邏輯判斷

令x等於1，可以有三種寫法

```
> x = 1
```

```
> x <- 1
```

```
> 1 -> x
```

列印x可以有兩種方法

```
> print(x)
```

```
[1] 1
```

```
> x
```

```
[1] 1
```

邏輯判斷，R會回應TRUE或是FALSE

```
> x>0
```

```
[1] TRUE
```

```
> X>0
```

```
Error: object 'X' not found
```

向量

```
> x = c(1,2,3,4,5)
```

```
> x
```

```
[1] 1 2 3 4 5
```

```
> y = c(1,3,5,7,9)
```

```
> x+y
```

```
[1] 2 5 8 11 14
```

○ 常用的向量函數

seq 產生連續序列

```
> seq(1,10)
```

```
> seq(1,20,by=2)
```

```
> seq(1,20,length.out  
=2)
```

rep 產生重複序列

```
> rep(1,10)
```

rev 反序

```
> rev(1:10)
```

休息一下，換你動手做

- 產生一個yuntech向量，該向量為
8,7,6,5,7,6,5,4,6,5,4,3,5,4,3,2,4,3,2,1
- 產生一個ma314向量，該向量為
3,7,11,15,19,23,27,31,35,39,43

```
yuntech <- rev(c(seq(1,4),seq(2,5),seq(3,6),seq(4,7),seq(5,8)))  
ma314 <- seq(3,43,by=4)
```

向量合併與矩陣

```
> x = c(1,2,3,4,5)
> y = c(1,3,5,7,9)
# 行合併 (cbind)
> cbind(x,y)
```

```
      x y
[1,] 1 1
[2,] 2 3
[3,] 3 5
[4,] 4 7
[5,] 5 9
```

```
# 列合併 (rbind)
> rbind(x,y)
```

```
      [,1] [,2] [,3] [,4] [,5]
x      1   2   3   4   5
y      1   3   5   7   9
```

○ 產生矩陣

```
> d <- seq(1,25,1)
> matrix(d, nrow=5, ncol=5,
byrow=T)
```

```
>
matrix(c(x,y),nrow=2,ncol=5,byrow=
T)
```


資料篩選(1)

○ 一個向量中，如果只想要某些數值或元素

```
> x <- sample(1:100,20) #從1-100之中隨機抽取20個數字
```

```
> x
```

```
[1] 90 42 43 20 13 8 4 68 6 80 97 14 51 9 79 39 33 40 50 62
```

```
> x[x>20]
```

```
[1] 90 42 43 68 80 97 51 79 39 33 40 50 62
```

如果抽出來的數字只要奇數呢

```
> x[x%%2==1]
```

```
[1] 43 13 97 51 9 79 39 3
```

資料篩選(2) – 矩陣中的資料篩選

```
> d <- seq(1,25,1)
> y <- matrix(d, nrow=5, ncol=5, byrow=T)

> y[1,1]
> y[1,]
> y[,5]
> y[3,-1]
```