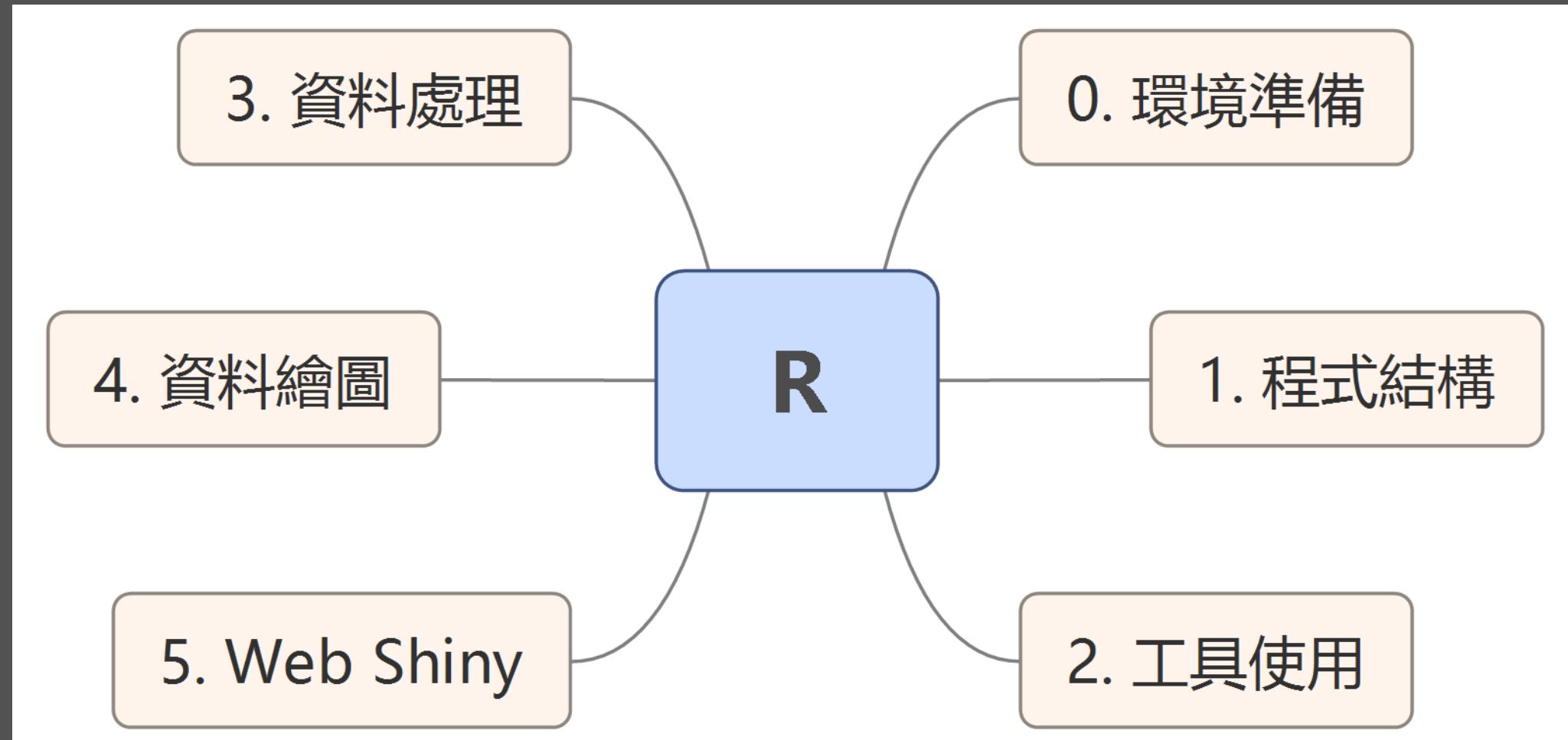


R 工具使用

目的

在此說明 R 基本的匯入匯出方案。



R 讀取與載入資料

1. CSV
2. EXCEL
3. 資料庫讀取
4. 統計軟體資料
5. R 二進位檔
6. R 內建資料集
7. 網路爬資料

R 讀取與載入資料

1. CSV

- (1) read.table
- (2) read.csv

```
theUrl = "https://raw.githubusercontent.com/kancheng/easyR/master/C02/im_2010.csv"
udf = read.table(file = theUrl, header = TRUE, sep = ",")  
head(udf)
dim(udf)
```

```
> theUrl = "https://raw.githubusercontent.com/kancheng/easyR/master/C02/im_2010.csv"
> udf = read.table(file = theUrl, header = TRUE, sep = ",")  
> head(udf)
   sid loan pjm ines cala calb sta ec ac mana sadm mis pom itdc cppg
1 SD990595    0  94   91   60   68   71  73  62   89   87   77   73   73   30
2 SD990596    0  89   74   60    0   60  60  73   84   89   74   67   77   81
3 SD990597    0  85   87   61   60   70  42  65   85   88   85   76   64   36
4 SD990598    0  90   89   25   30   74  46  75   86   90   80   82   81   61
5 SD990599    0  79   79   69   60   65  78  61   84   90   71   74   61   35
6 SD990600    1  66   69   50   65   60  56  64   86   90   69   72   81   78
  dtst oopg pcpg adcpq nwkpm itdcn inkpg xmwd idbs dbms
1   30   60   60    35    79    67    71    70    60    76
2   23   17   60    44    63    69    71    60    48    64
3   27   60   60    40    78    72    75    71    60    92
4   60   64   60    60    81    79    92    83    66    97
5   36   39   60    31    66    67    80    69    60    76
6   37   60   60    60    77    69    85    71    64    62
> dim(udf)
[1] 143  25
> |
```

R 讀取與載入資料

```
udf2 = read.table (file = theUrl, sep = ",")  
head(udf2)
```

```
> udf2 = read.table (file = theUrl, sep = ",")  
> head(udf2)  
    V1     V2     V3     V4     V5     V6     V7     V8     V9     V10    V11    V12    V13    V14    V15    V16    V17  
1      sid loan pjm ines cala calb sta ec ac mana sadm mis pom itdc cppg dtst oopg  
2 SD990595     0   94   91   60   68   71   73   62   89   87   77   73   73   30   30   60  
3 SD990596     0   89   74   60     0   60   60   73   84   89   74   67   77   81   23   17  
4 SD990597     0   85   87   61   60   70   42   65   85   88   85   76   64   36   27   60  
5 SD990598     0   90   89   25   30   74   46   75   86   90   80   82   81   61   60   64  
6 SD990599     0   79   79   69   60   65   78   61   84   90   71   74   61   35   36   39  
    V18     V19     V20     V21     V22     V23     V24     V25  
1 pcpg  adpcpg nwkpm itdcn inkpg xmwd idbs dbms  
2 60     35     79     67     71     70     60     76  
3 60     44     63     69     71     60     48     64  
4 60     40     78     72     75     71     60     92  
5 60     60     81     79     92     83     66     97  
6 60     31     66     67     80     69     60     76  
> |
```

R 讀取與載入資料

當中的 file、header、sep 為
read.table 函數中的引數

file ->為匯進來的名稱

header ->是否將資料的第一橫列設
為直行

sep -> 資料的分隔符號

sep 通常用的分隔

\t -> tab 鍵

； -> 分號

當中有個引數 stringsAsFactors (預設為 TRUE) , 以防 character 直行轉為 factor, 尤其在面對很多重複的字元(character)時可以利用, 同樣的這個引數在 data.frame 也能用到。

Description

Reads a file in table format and creates a data frame from it, with cases corresponding to lines and variables to fields in the file.

Usage

```
read.table(file, header = FALSE, sep = "", quote = "\""",
           dec = ".", numerals = c("allow.loss", "warn.loss", "no.loss"),
           row.names, col.names, as.is = !stringsAsFactors,
           na.strings = "NA", colClasses = NA, nrows = -1,
           skip = 0, check.names = TRUE, fill = !blank.lines.skip,
           strip.white = FALSE, blank.lines.skip = TRUE,
           comment.char = "#",
           allowEscapes = FALSE, flush = FALSE,
           stringsAsFactors = default.stringsAsFactors(),
           fileEncoding = "", encoding = "unknown", text, skipNul = FALSE)

read.csv(file, header = TRUE, sep = ",", quote = "\""",
         dec = ".", fill = TRUE, comment.char = "", ...)

read.csv2(file, header = TRUE, sep = ";", quote = "\""",
          dec = ",", fill = TRUE, comment.char = "", ...)

read.delim(file, header = TRUE, sep = "\t", quote = "\""",
           dec = ".", fill = TRUE, comment.char = "", ...)

read.delim2(file, header = TRUE, sep = "\t", quote = "\""",
            dec = ",", fill = TRUE, comment.char = "", ...)
```

R 讀取與載入資料

```
x = 10:1  
y = -4:5  
q = c("Hockey", "Football", "Baseball",  
      "Curling", "Rugby", "Lacrosse",  
      "Basketball", "Tennis", "Cricket",  
      "Soccer")
```

```
theDF = data.frame( First = x,  
                    Second = y, Sport = q,  
                    stringsAsFactors = FALSE)
```

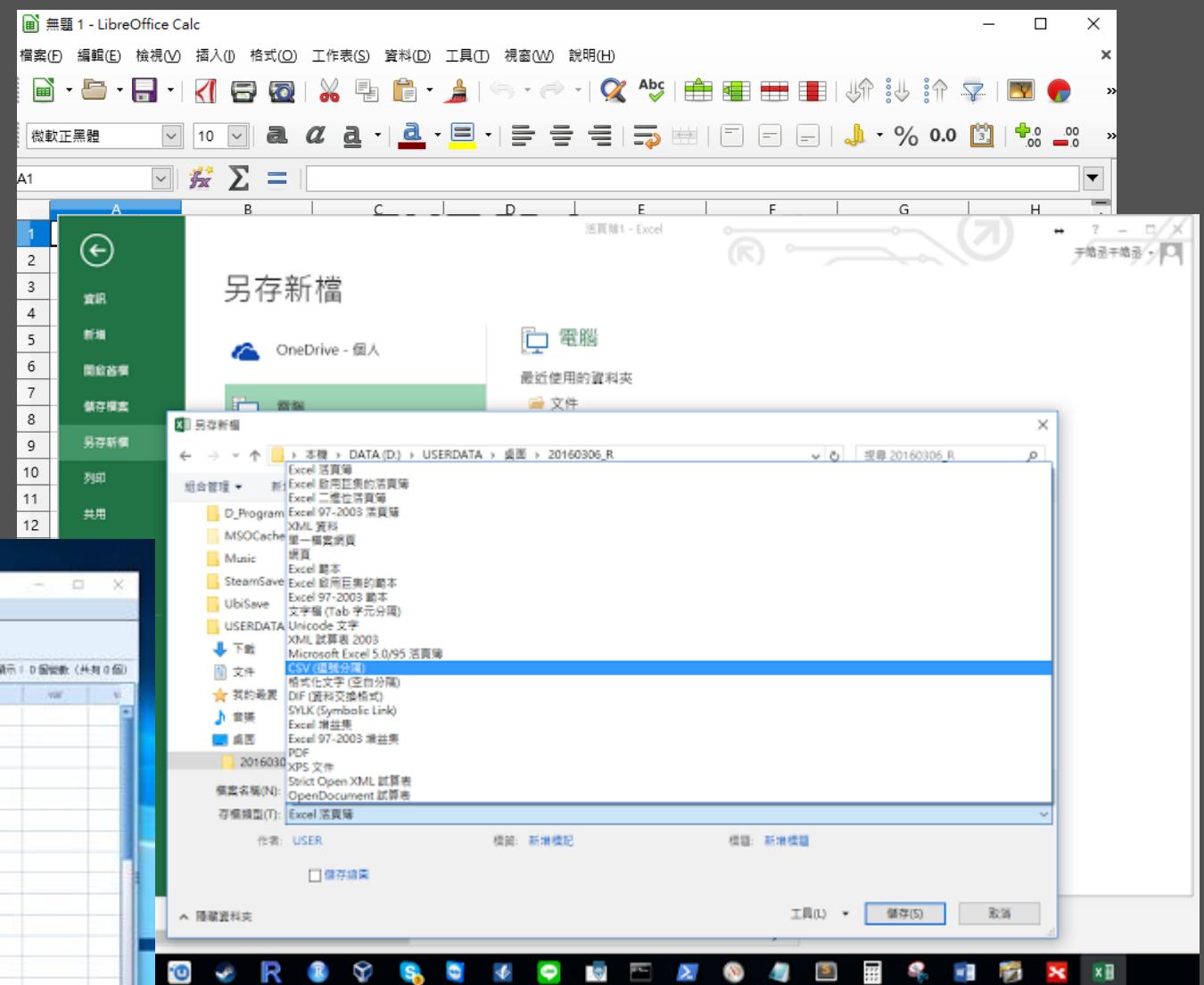
```
theDF
```

```
theDF$Sport
```

```
> x = 10:1  
> y = -4:5  
> q = c("Hockey", "Football", "Baseball",  
+ "Curling", "Rugby", "Lacrosse",  
+ "Basketball", "Tennis", "Cricket",  
+ "Soccer")  
> theDF = data.frame( First = x,  
+ Second = y, Sport = q,  
+ stringsAsFactors = FALSE)  
> theDF  
   First Second     Sport  
1     10     -4    Hockey  
2      9     -3   Football  
3      8     -2   Baseball  
4      7     -1    Curling  
5      6      0     Rugby  
6      5      1   Lacrosse  
7      4      2 Basketball  
8      3      3     Tennis  
9      2      4    Cricket  
10     1      5   Soccer  
> theDF$Sport  
[1] "Hockey"      "Football"     "Baseball"     "Curling"      "Rugby"  
[6] "Lacrosse"    "Basketball"   "Tennis"       "Cricket"      "Soccer"  
> |
```

R 讀取與載入資料

其實 Excel 等試算表軟體或者統計軟體直接轉存成 CSV 即可，當然過程要注意該作業系統編碼的問題。



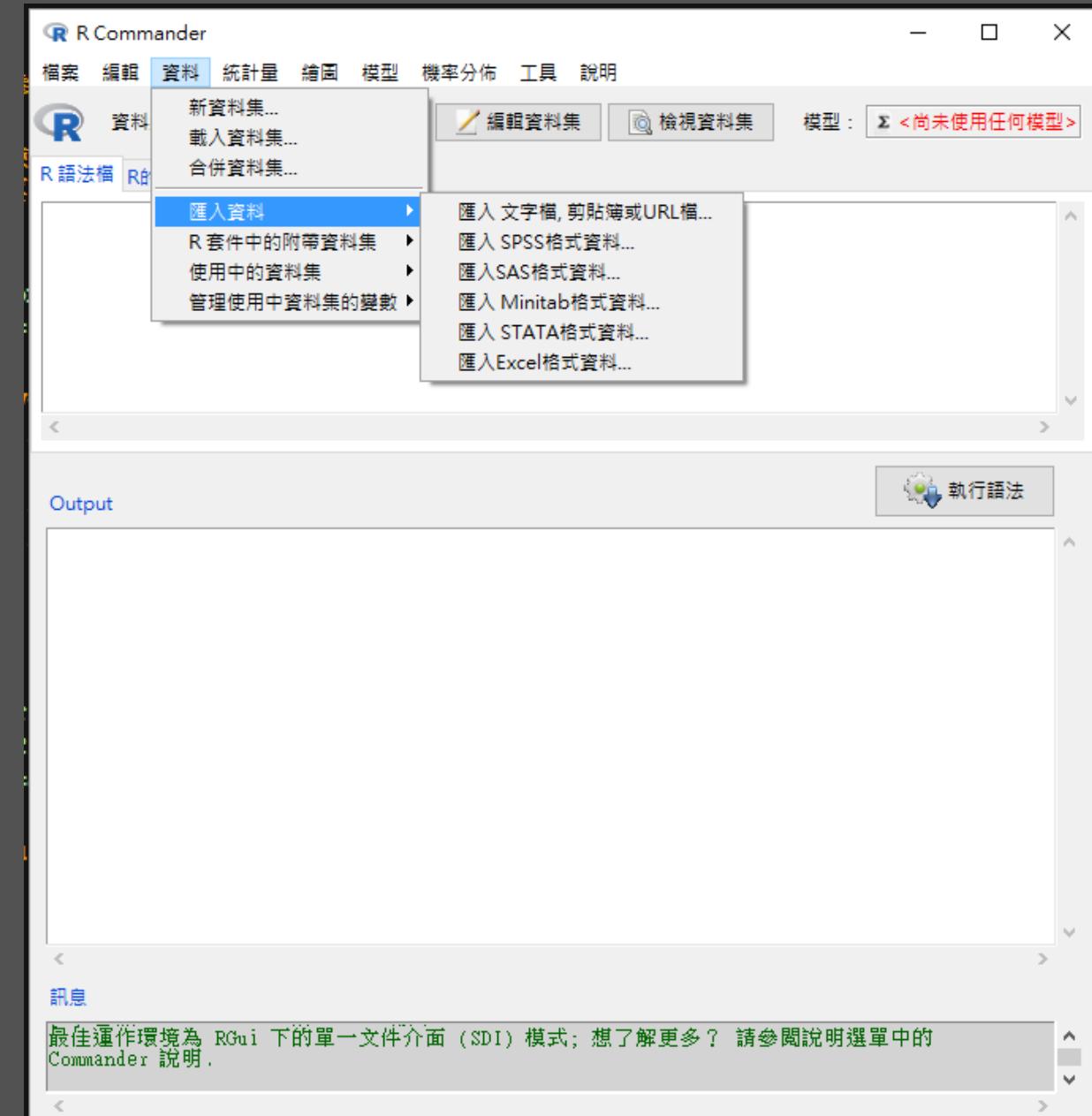
R Commander

這邊就不解釋了，除了轉存CSV，你可以直接用套件進行動作。

Packages 推薦 -> gdata、XLConnect、xlsReadWrite

RODBC Packages 函數 -> odbcConnectExcel2007

```
# install.packages("Rcmdr")
library("Rcmdr")
```



R Excel

若無安裝好 JAVA , xlsx 這個套件會無法使用。

1. xls

使用 R Commander

2. xlsx

使用 xlsx 套件

```
> install.packages("xlsx")
--- Please select a CRAN mirror for use in this session ---
also installing the dependencies 'rJava', 'xlsxjars'

嘗試 URL 'https://cran.rsm.ac.jp/bin/windows/contrib/3.3/rJava_0.9-8.zip'
Content type 'application/zip' length 713501 bytes (696 KB)
downloaded 696 KB

嘗試 URL 'https://cran.rsm.ac.jp/bin/windows/contrib/3.3/xlsxjars_0.6.1.zip'
Content type 'application/zip' length 9485180 bytes (9.0 MB)
downloaded 9.0 MB

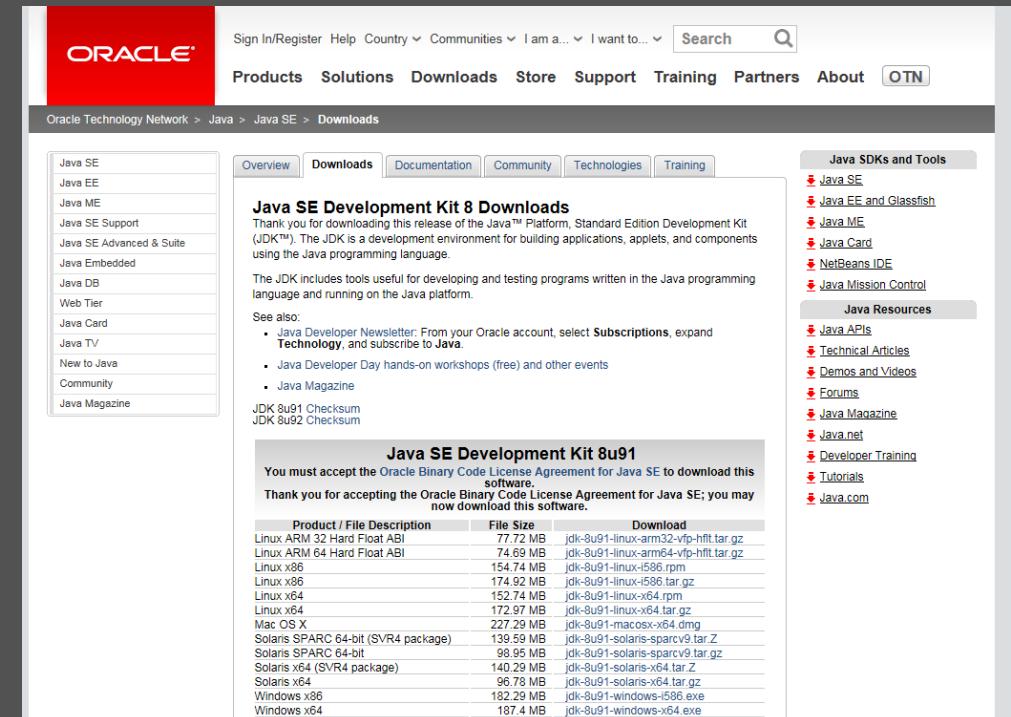
嘗試 URL 'https://cran.rsm.ac.jp/bin/windows/contrib/3.3/xlsx_0.5.7.zip'
Content type 'application/zip' length 400993 bytes (391 KB)
downloaded 391 KB

package 'rJava' successfully unpacked and MD5 sums checked
package 'xlsxjars' successfully unpacked and MD5 sums checked
package 'xlsx' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:/Users/F502/AppData/Local/Temp/Rtmp6x5yl8/downloaded_packages
> |
```

```
> require("xlsx")
Loading required package: xlsx
Loading required package: rJava
Error : .onLoad failed in loadNamespace() for 'rJava', details:
  call: fun(libname, pkgname)
  error: JAVA_HOME cannot be determined from the Registry
Failed with error: 'package 'rJava' could not be loaded'
> |
```

直接 google 搜尋 jdk，找到 Oracle 在其下載頁面中，選擇我同意並挑選相應的版本安裝即可。



R Excel

用 LibreOffice 自行建立的 Excel (.xlsx) 格式的資料

	A	B	C	D	E	F
1	sid	name	cal	st	ec	
2	1 kan		60	80	98	
3	2 hao		80	75	70	
4	3 cheng		90	90	81	
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						

```
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。


> setwd("C:/Rtest")
> require("xlsx")
Loading required package: xlsx
Loading required package: rJava
Loading required package: xlsxjars
> kandata = read.xlsx("C:/Rtest/rexce.xlsx",1)
> kandata
   sid name cal st ec
1   1  kan  60 80 98
2   2  hao  80 75 70
3   3 cheng  90 90 81
> |
```

```
install.packages("xlsx")
require("xlsx")
data = read.xlsx("C:/Path/EFile.xlsx", 1)
data
```

R ODBC MS SQL Server

MS SQL Server 必須安裝好，然後 Windows 版本中要有 ODBC。

```
# install.packages("RODBC")
```

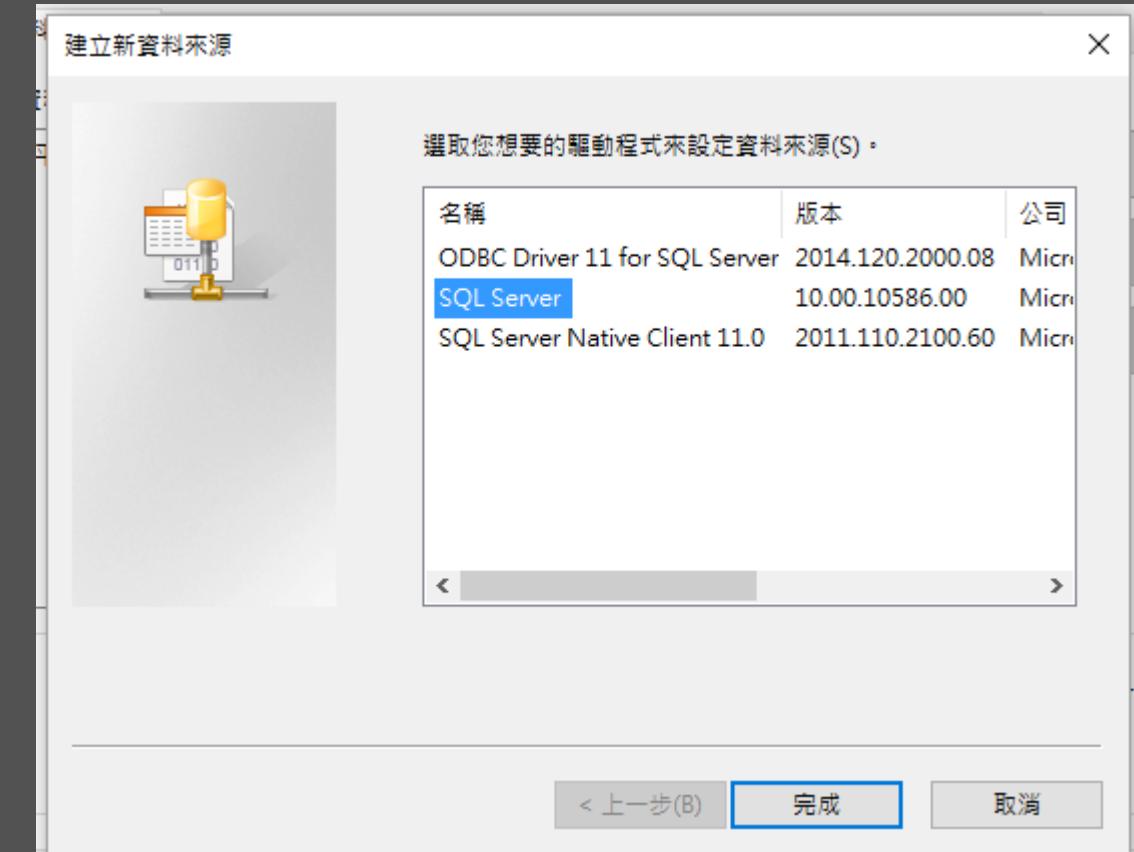
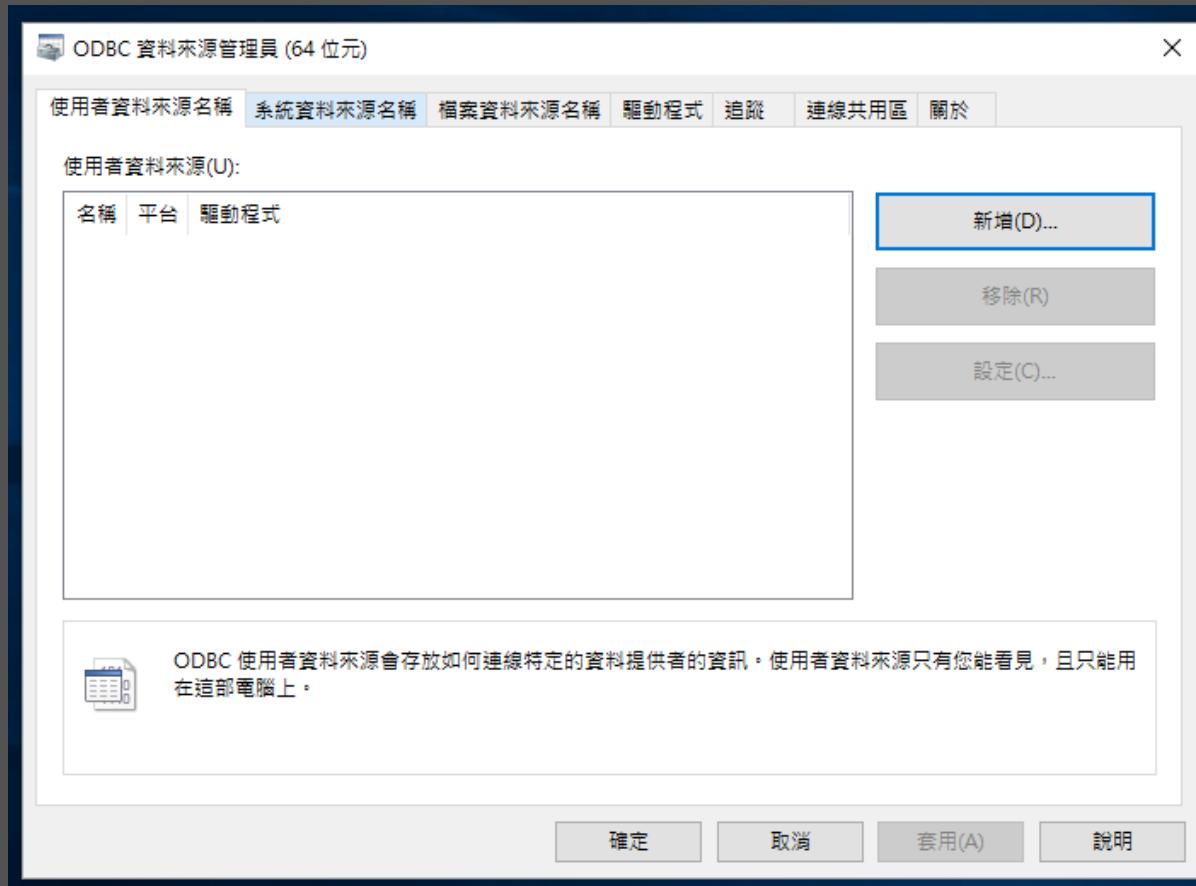


```
> install.packages("RODBC")
Warning in install.packages("RODBC") :
  'lib' = "C:/Program Files/R/R-3.2.4revised/library'" is not writable
--- Please select a CRAN mirror for use in this session ---
嘗試 URL 'https://cloud.r-project.org/bin/windows/contrib/3.2/RODBC_1.3-12.zip'
Content type 'application/zip' length 829435 bytes (809 KB)
downloaded 809 KB

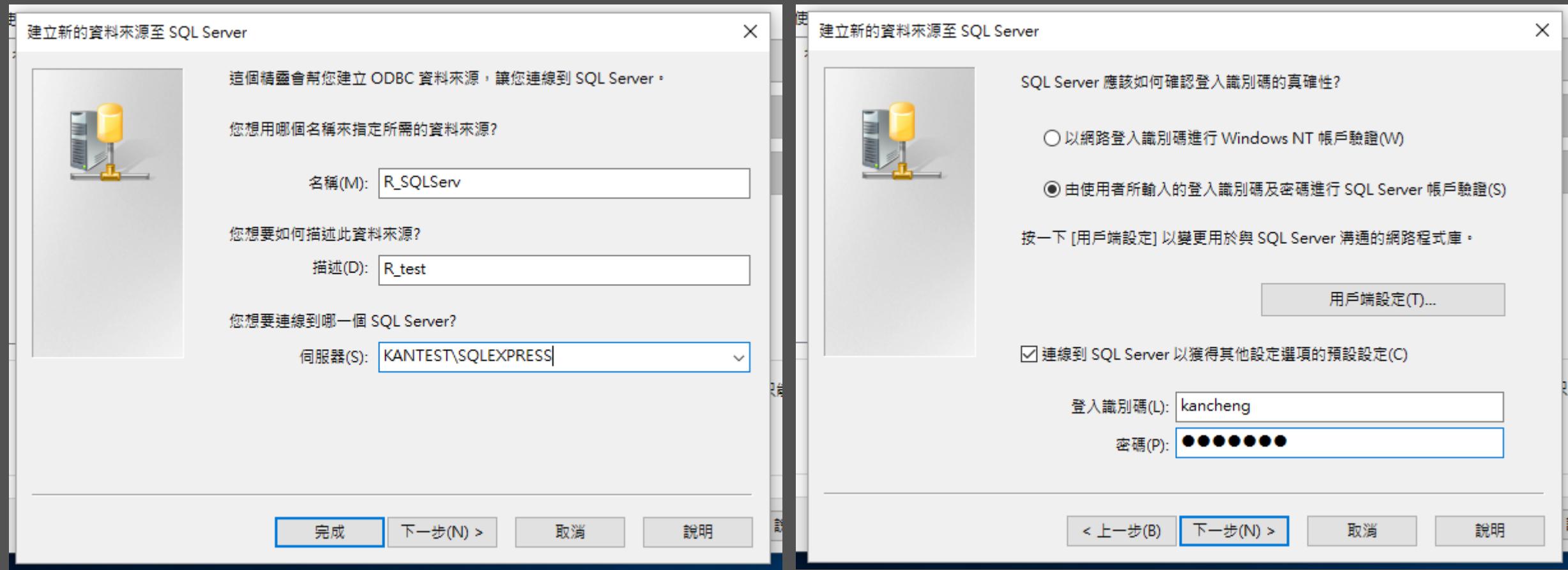
package 'RODBC' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
      C:\Users\root\AppData\Local\Temp\RtmpEXZYdb\downloaded_packages
> |
```

R ODBC MS SQL Server

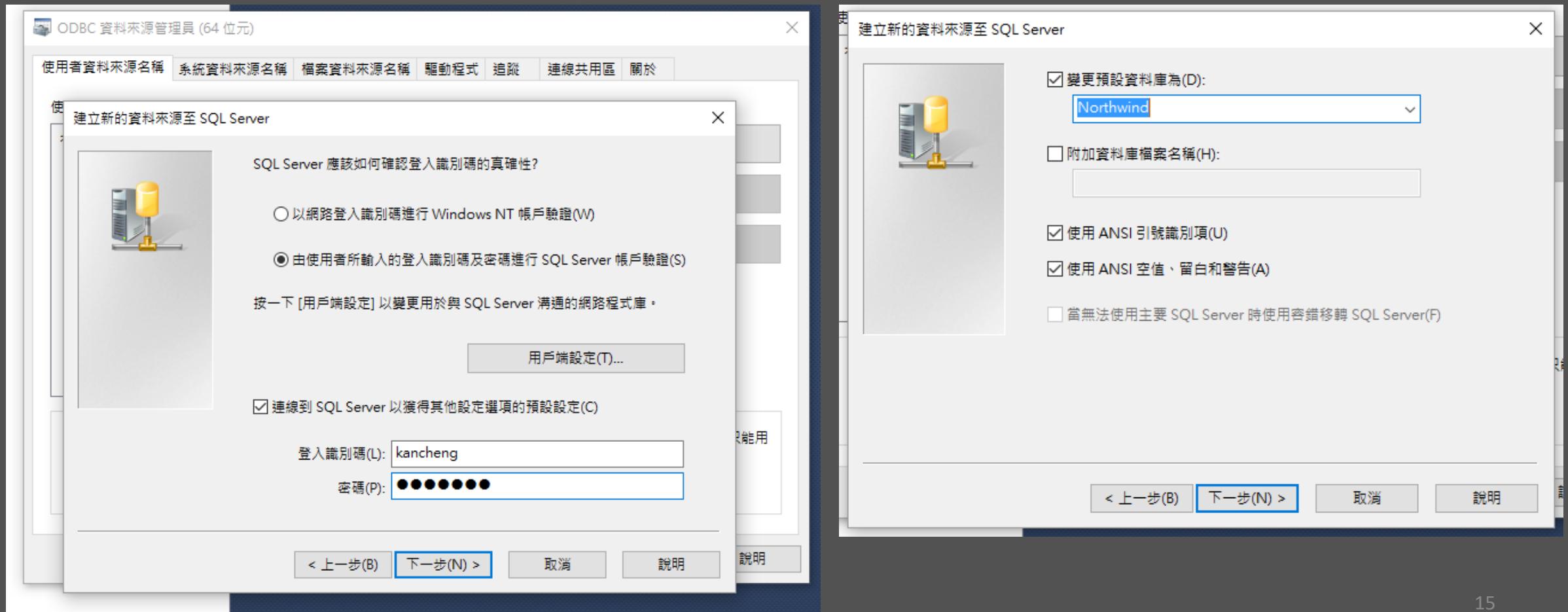


R ODBC MS SQL Server



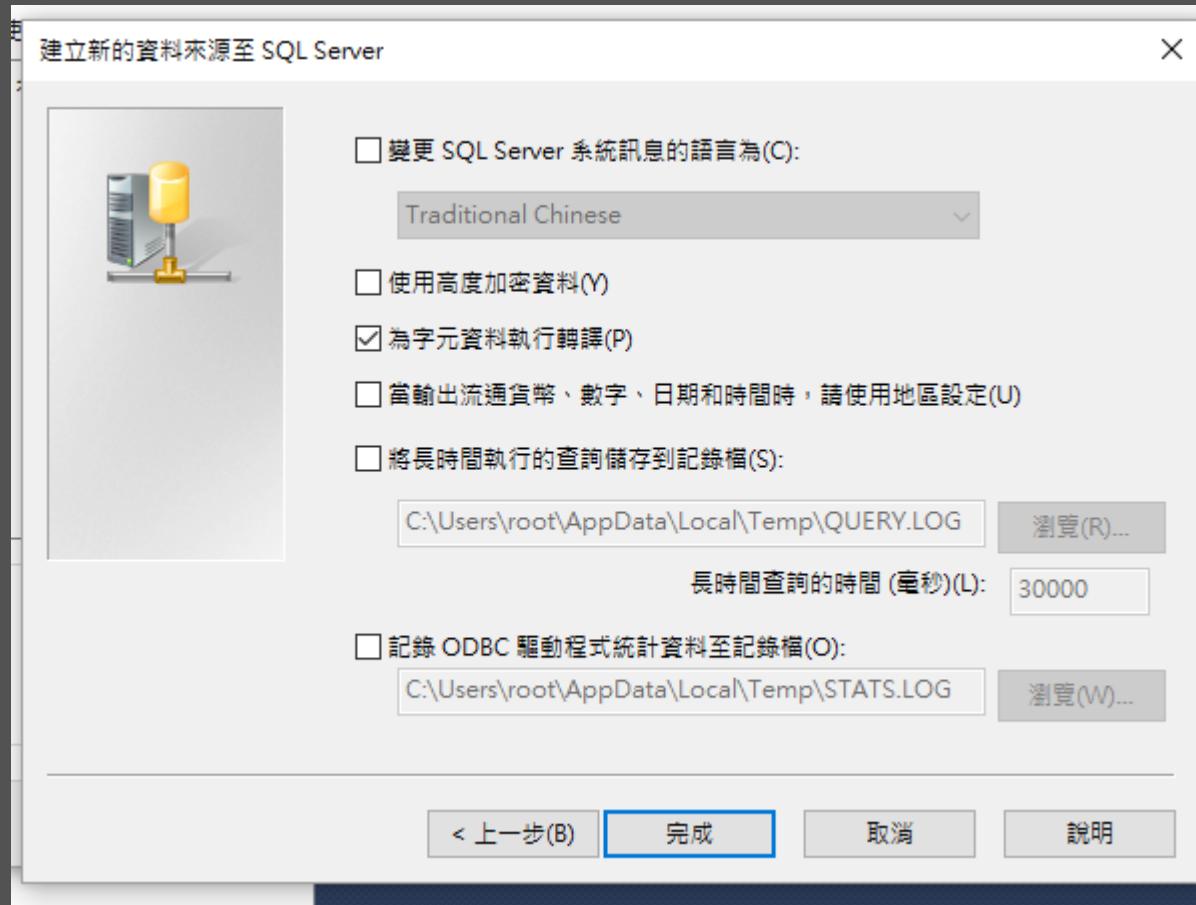
R ODBC MS SQL Server

檢查 MSSQL、MSSQL內的DB與帳號、TCP/IP 這些設定有沒有問題!!!
設定自己之前在 MS 下載的 範例資料庫 Northwind



R ODBC MS SQL Server

完成 !!!



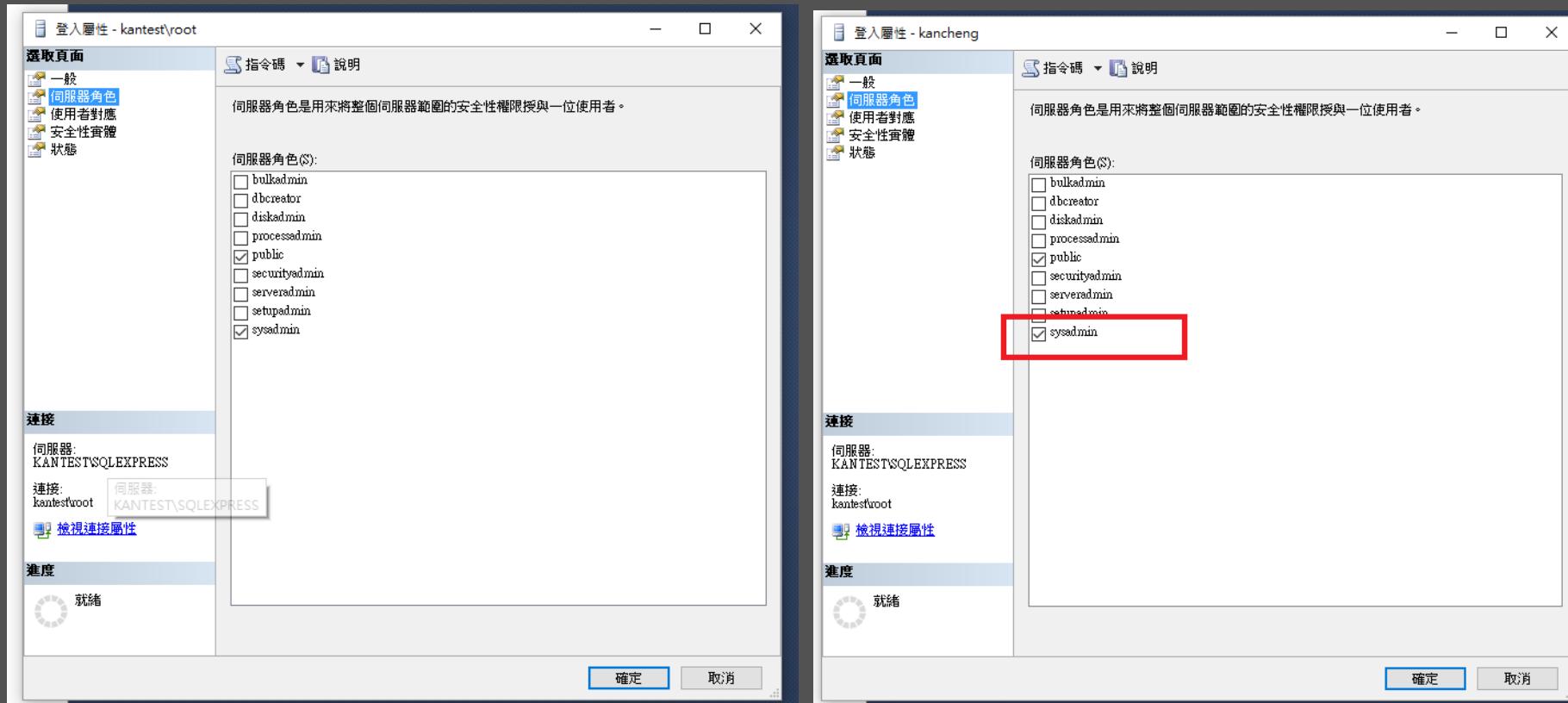
R ODBC MS SQL Server

MS SQL 帳號會因為權限不足的問題，我將權限設定用成跟MSSQL 主機模式下的 Windows系統帳號權限一樣 (也就是 sysadmin)

主機帳號權限

kantest -> 主機名稱

root -> 主機系統權限帳號名稱
觀察 sysadmin 權限!!!



R ODBC MS SQL Server

使用 R 利用 RODBC 對 MS SQL 的 Northwind 資料庫進行查詢
這是這回所要利用的表 -> dbo.Orders

The screenshot shows the Microsoft SQL Server Management Studio interface. On the left, the Object Explorer displays the Northwind database structure, including tables like Customers, Orders, and Products. In the center, a query window titled 'SQLQuery3.sql' is open, containing the following SQL code:

```
USE [Northwind]
GO

SELECT [OrderID]
      ,[CustomerID]
      ,[EmployeeID]
      ,[OrderDate]
      ,[RequiredDate]
      ,[ShippedDate]
      ,[ShipVia]
      ,[Freight]
      ,[ShipName]
      ,[ShipAddress]
      ,[ShipCity]
      ,[ShipRegion]
      ,[ShipPostalCode]
      ,[ShipCountry]
   FROM [dbo].[Orders]
```

Below the code, the 'Results' tab shows the query results:

OrderID	CustomerID	EmployeeID	OrderDate	RequiredDate	ShippedDate	ShipVia	F
1	10248	VINET	5	1996-07-04 00:00:00.000	1996-08-01 00:00:00.000	1996-07-16 00:00:00.000	3
2	10249	TOMSP	6	1996-07-05 00:00:00.000	1996-08-16 00:00:00.000	1996-07-10 00:00:00.000	1
3	10250	HANAR	4	1996-07-08 00:00:00.000	1996-08-05 00:00:00.000	1996-07-12 00:00:00.000	2
4	10251	VICTE	3	1996-07-08 00:00:00.000	1996-08-05 00:00:00.000	1996-07-15 00:00:00.000	1
5	10252	SURPD	4	1996-07-09 00:00:00.000	1996-08-06 00:00:00.000	1996-07-11 00:00:00.000	2
6	10253	HANAR	3	1996-07-10 00:00:00.000	1996-07-24 00:00:00.000	1996-07-16 00:00:00.000	2
7	10254	CHOPS	5	1996-07-11 00:00:00.000	1996-08-08 00:00:00.000	1996-07-23 00:00:00.000	2
8	10255	RICSU	9	1996-07-12 00:00:00.000	1996-08-09 00:00:00.000	1996-07-15 00:00:00.000	3
9	10256	WELLI	3	1996-07-15 00:00:00.000	1996-08-12 00:00:00.000	1996-07-17 00:00:00.000	2
10	10257	HILAA	4	1996-07-16 00:00:00.000	1996-08-13 00:00:00.000	1996-07-22 00:00:00.000	3
11	10258	ERNSH	1	1996-07-17 00:00:00.000	1996-08-14 00:00:00.000	1996-07-23 00:00:00.000	1
12	10259	CENTC	4	1996-07-18 00:00:00.000	1996-08-15 00:00:00.000	1996-07-25 00:00:00.000	3

The screenshot shows the R console window with the following text:

```
R version 3.2.4 Revised (2016-03-16 r70336) -- "Very Secure Dishes"
Copyright (C) 2016 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考它們。 

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

Below the text, there is a screenshot of the 'SQL Server 登入' (SQL Server Login) dialog box. It shows the following fields:

資料來源:	R_SQLServ
<input type="checkbox"/> 使用信任連線(U)	
登入識別碼(L):	kancheng
密碼(P):	*****

Buttons on the right include '確定' (OK), '取消' (Cancel), '說明(H)' (Help), and '選項(O) >>' (Options).

R ODBC MS SQL Server

```
require(RODBC)
sdb=odbcConnect("R_SQLServ")
```

```
> require(RODBC)
Loading required package: RODBC
> sdb = odbcConnect("R_SQLServ")
> ordtable = sqlQuery(sdb, "SELECT * FROM dbo.Orders", stringsAsFactors = FALSE)
> class(ordtable)
[1] "data.frame"
> names(ordtable)
[1] "OrderID"      "CustomerID"    "EmployeeID"    "OrderDate"     "RequiredDate"   "ShippedDate"   "ShipVia"       "Freight"
[9] "ShipName"      "ShipAddress"    "ShipCity"      "ShipRegion"    "ShipPostalCode" "ShipCountry"
> head(ordtable,n=3)
  OrderID CustomerID EmployeeID OrderDate RequiredDate ShippedDate ShipVia Freight
1  10248      VINET          5 1996-07-04 1996-08-01 1996-07-16      3  32.38 Vins et alcools Chevalier 59 rue de l'Abbaye
2  10249      TOMSP          6 1996-07-05 1996-08-16 1996-07-10      1  11.61      Toms Spezialitaten Luisenstr. 48
3  10250      HANAR          4 1996-07-08 1996-08-05 1996-07-12      2  65.83      Hanari Carnes  Rua do Paco, 67
  ShipCity ShipRegion ShipPostalCode ShipCountry
1      Reims      <NA>        51100     France
2    Munster      <NA>        44087    Germany
3 Rio de Janeiro      RJ        05454-876     Brazil
> |
```

R ODBC MS SQL Server

基本上可以使用 `sqlQuery()` 連線進入 MSSQL 並且使用 SQL, `head()` 可以檢視 table 的內容。

```
> table(ordtable$ShipCountry)

Argentina      Austria      Belgium      Brazil      Canada      Denmark      Finland      France      Germany      Ireland      Italy
       16          40          19          83          30          18          22          77          122         19          28
Mexico        Norway      Poland      Portugal     Spain      Sweden      Switzerland      UK          USA      Venezuela
       28           6           7          13          23          37          18          56          122         46

> mean(ordtable$Freight)
[1] 78.2442
> odbcClose(sdb)
> |
```

```
ordtable = sqlQuery(sdb, "SELECT * FROM dbo.Orders", stringsAsFactors = FALSE)
class(ordtable)
names(ordtable)
head(ordtable,n=3)
table(ordtable$ShipCountry)
mean(ordtable$Freight)
odbcClose(sdb)
```

R ODBC MS SQL Server

使用 R 利用 RODBC 對 MS SQL 的 Northwind 資料庫進行新增
這是這回所要利用的表 -> dbo.Shippers

The screenshot shows the Microsoft SQL Server Management Studio interface. On the left is the Object Explorer pane, which displays the database structure of the Northwind database, including tables like Shippers, Customers, and Orders. In the center is the SQL Query window titled 'SQLQuery4.sql - KANTEST\SQLEXPRESS.Northwind (kantest\root (51)) - Microsoft SQL Server Management Studio'. It contains the following T-SQL code:

```
USE [Northwind]
GO

SELECT [ShipperID]
      ,[CompanyName]
      ,[Phone]
   FROM [dbo].[Shippers]
GO
```

Below the query window is the Results pane, which displays the output of the query. The results show three rows of data from the Shippers table:

ShipperID	CompanyName	Phone
1	Speedy Express	(503) 555-9831
2	United Package	(503) 555-3199
3	Federal Shipping	(503) 555-9931

At the bottom of the screen, the taskbar shows various open applications, including a browser and file explorer windows.

R ODBC MS SQL Server

跟前面查詢一樣，利用 ODBC 與 RODBC 輸入帳號登入 MSSQL

```
> require(RODBC)
Loading required package: RODBC
> sdb = odbcConnect("R_SQLServ")
> ordtable = sqlQuery(sdb, "SELECT * FROM dbo.Shippers", stringsAsFactors = FALSE)
> ordtable
  ShipperID      CompanyName      Phone
1          1 Speedy Express (503) 555-9831
2          2 United Package (503) 555-3199
3          3 Federal Shipping (503) 555-9931
> dim(ordtable)[1]
[1] 3
> names(ordtable)
[1] "ShipperID"    "CompanyName"   "Phone"
> |
```

在此使用 R 新增 SQL 對其 table 增加一筆資料

```
> insert.sql = paste("INSERT INTO dbo.Shippers", "(CompanyName, Phone)", "VALUES", "(", "'WEPA'", "'7654321'", ")")
> sqlQuery(sdb, insert.sql)
character(0)
> |
```

R ODBC MS SQL Server

成功

```
> ordtabletest = sqlQuery(sdb, "SELECT * FROM dbo.Shippers", stringsAsFactors = FALSE)
> ordtabletest
  ShipperID    CompanyName        Phone
1          1  Speedy Express (503) 555-9831
2          2 United Package (503) 555-3199
3          3 Federal Shipping (503) 555-9931
4          4           WEPA        7654321
>
```

odbcClose() -> 無 !!!

```
> ordtabletest = sqlQuery(sdb, "SELECT * FROM dbo.Shippers", stringsAsFactors = FALSE)
> ordtabletest
  ShipperID    CompanyName        Phone
1          1  Speedy Express (503) 555-9831
2          2 United Package (503) 555-3199
3          3 Federal Shipping (503) 555-9931
4          4           WEPA        7654321
>
>
> odbcClose(sdb)
> |
```

```
require(RODBC)
sdb = odbcConnect("R_SQLServ")
ordtable = sqlQuery(sdb,
  "SELECT * FROM dbo.Shippers",
  stringsAsFactors = FALSE)
ordtable
dim(ordtable)[1]
names(ordtable)
insert.sql = paste("INSERT INTO
dbo.Shippers",
  "(CompanyName, Phone) ",
  "VALUES", "(", "'WEPA'", "'7654321'",
  ")")
sqlQuery(sdb, insert.sql)
ordtabletest = sqlQuery(sdb,
  "SELECT * FROM dbo.Shippers",
  stringsAsFactors = FALSE)
ordtabletest
odbcClose(sdb)
```

R MariaDB 與 RMYSQL套件設定 by Windows

Rtool

<https://cran.r-project.org/bin/windows/Rtools/>

Rtools Downloads

Some of the tools are incompatible with obsolete versions of R. We maintain one actively updated version of the tools, and other "frozen" snapshots of them. We recommend the latest release of R.

The current version of this file is recorded here: [VERSION.txt](#).

Download	R compatibility	Frozen?
Rtools34.exe	R 3.3.x and later	No
Rtools33.exe	R 3.2.x to 3.3.x	Yes
Rtools32.exe	R 3.1.x to 3.2.x	Yes
Rtools31.exe	R 3.0.x to 3.1.x	Yes

R MariaDB 與 RMYSQL套件設定 by Windows

Change History

Changes since R 3.3.0

The version number has been updated to 3.4.x.y.

Support for the gcc 4.6.3 toolchain has been dropped.

We have updated Tcl/Tk to version 8.6.4. If you are using Rtools3

TCL_VERSION = 85

to

TCL_VERSION = 86

in R_HOME/src/gnuwin32/fixed/etc/Makeconf.

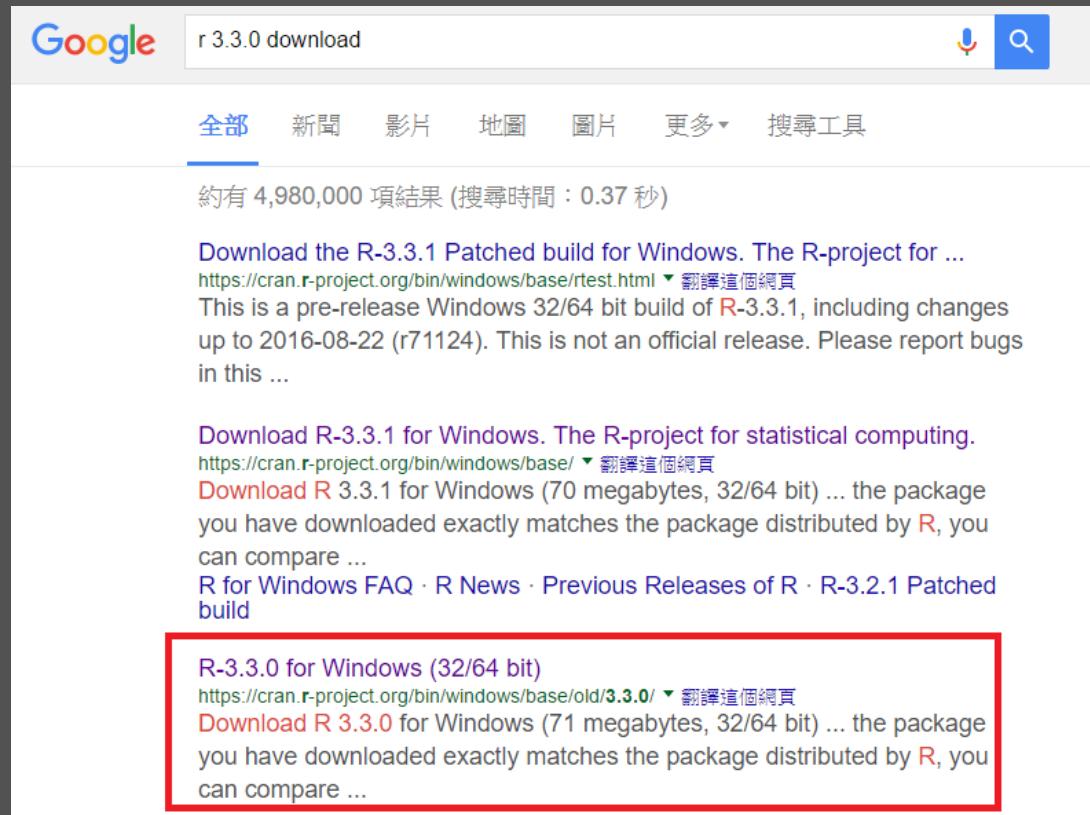
Changes since R 3.2.2

Both the gcc 4.6.3 toolchain and a toolchain based on gcc 4.9.3 and compilers are selected.

Changes since R 3.1.3

Rtool

請根據 Rtool 來找 R 的版本



R MariaDB 與 RMYSQL 套件設定 by Windows

MariaDB - <https://mariadb.org/>

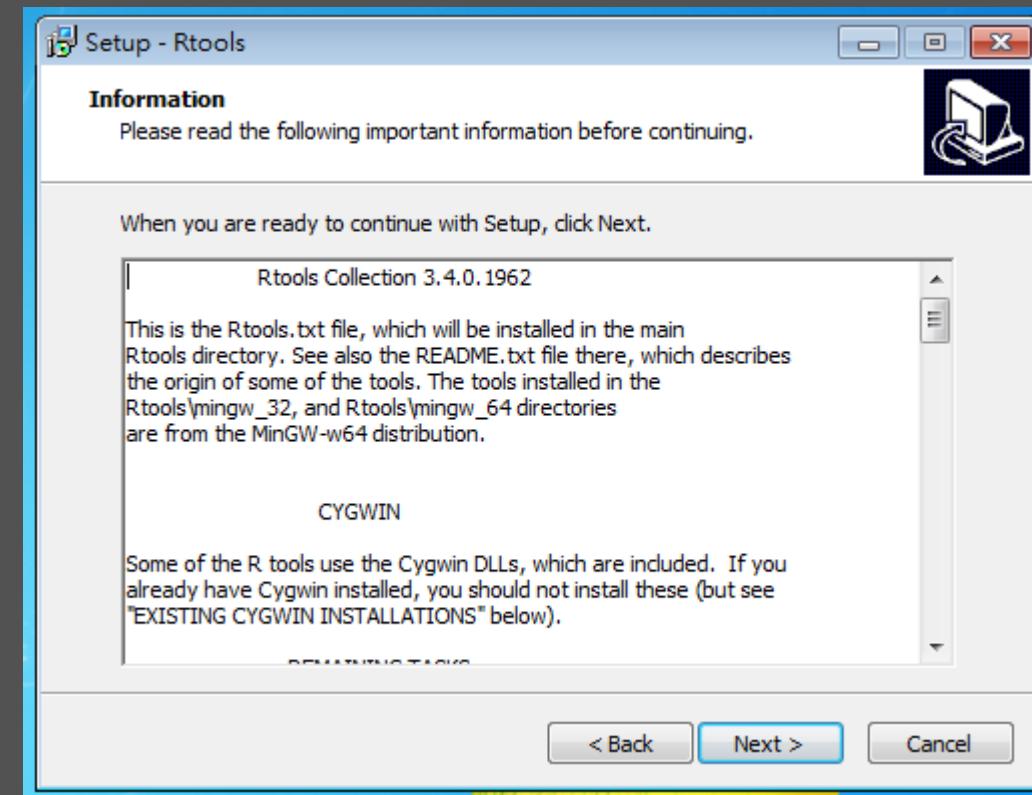
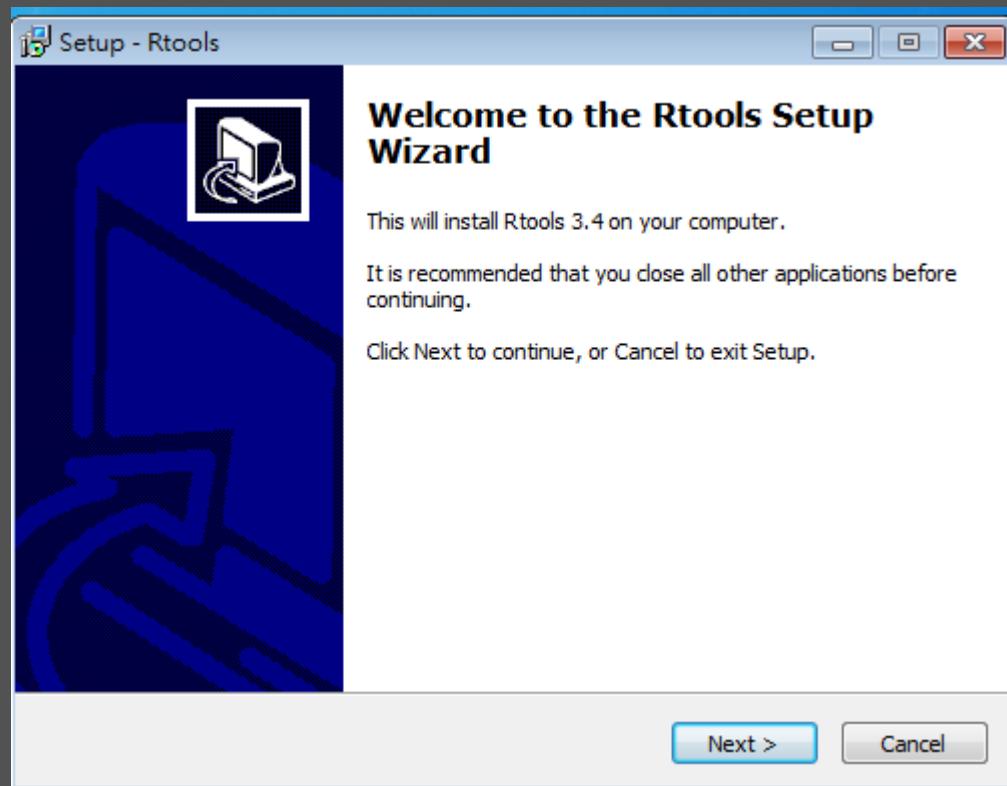
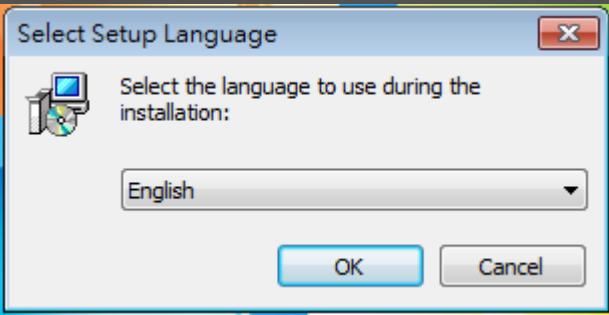
進入下載後選擇 MariaDB 5.5 Series 的版本



R MariaDB 與 RMYSQL 套件設定 by Windows

Rtool 安裝

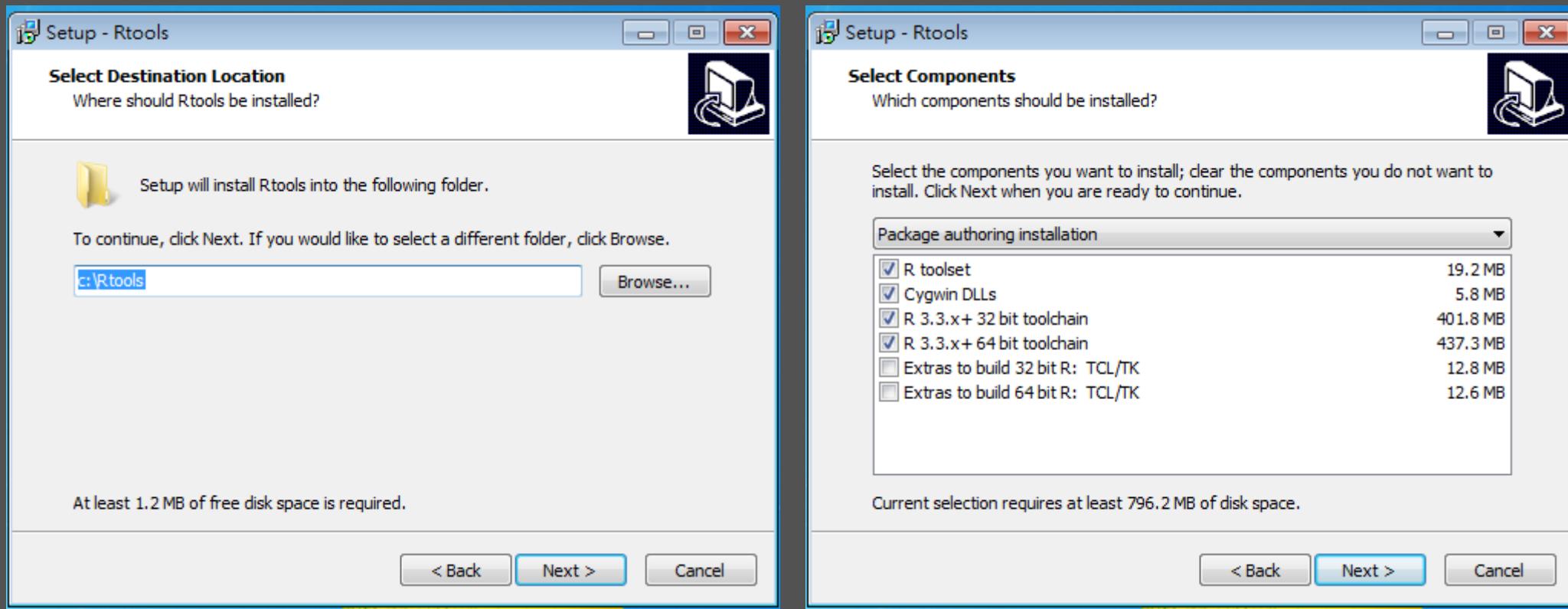
下一步或用預設即可 ...



R MariaDB 與 RMYSQL 套件設定 by Windows

Rtool 安裝

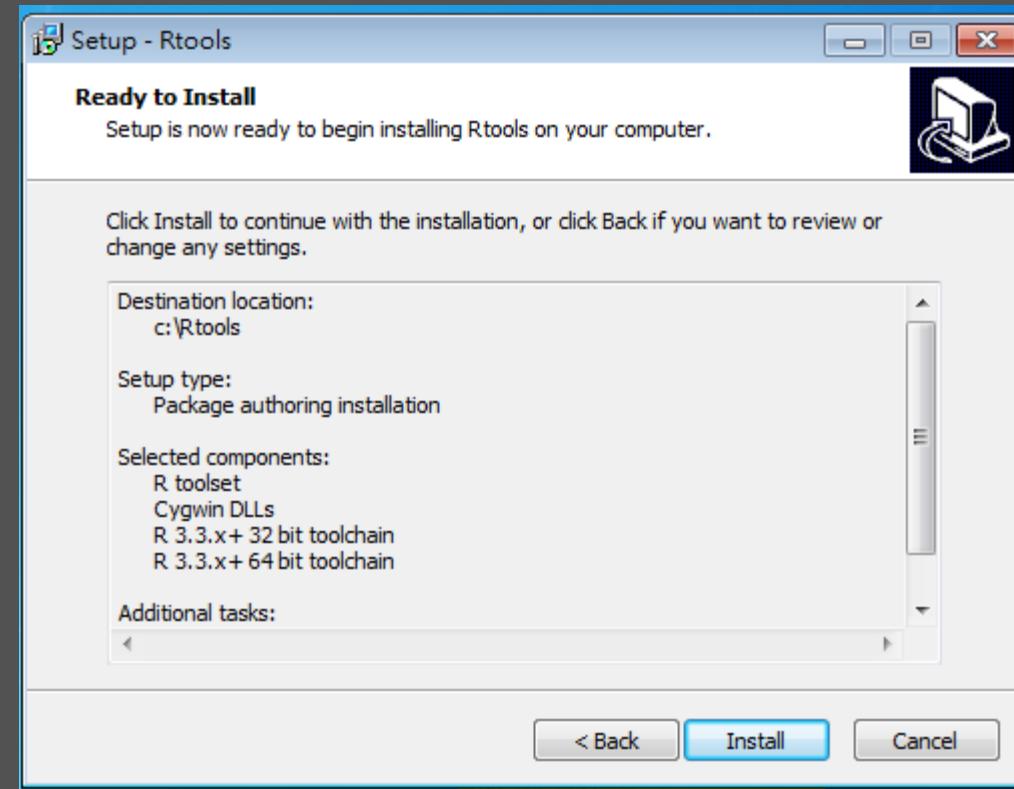
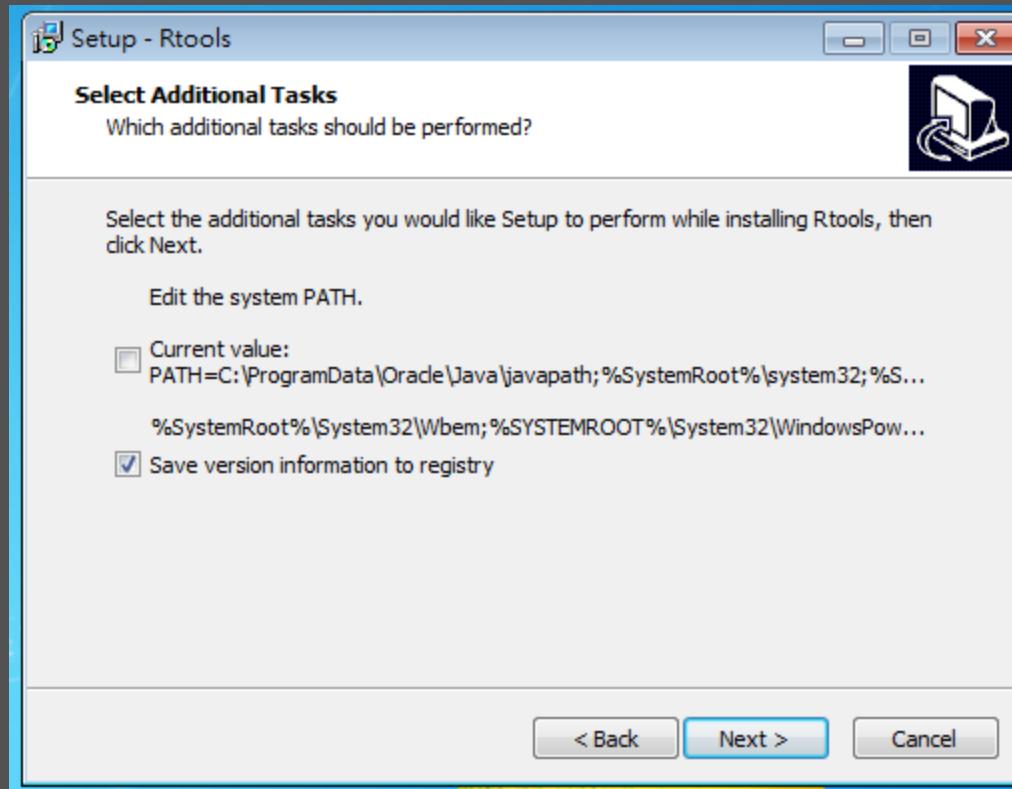
下一步或用預設即可 ...



R MariaDB 與 RMYSQL 套件設定 by Windows

Rtool 安裝

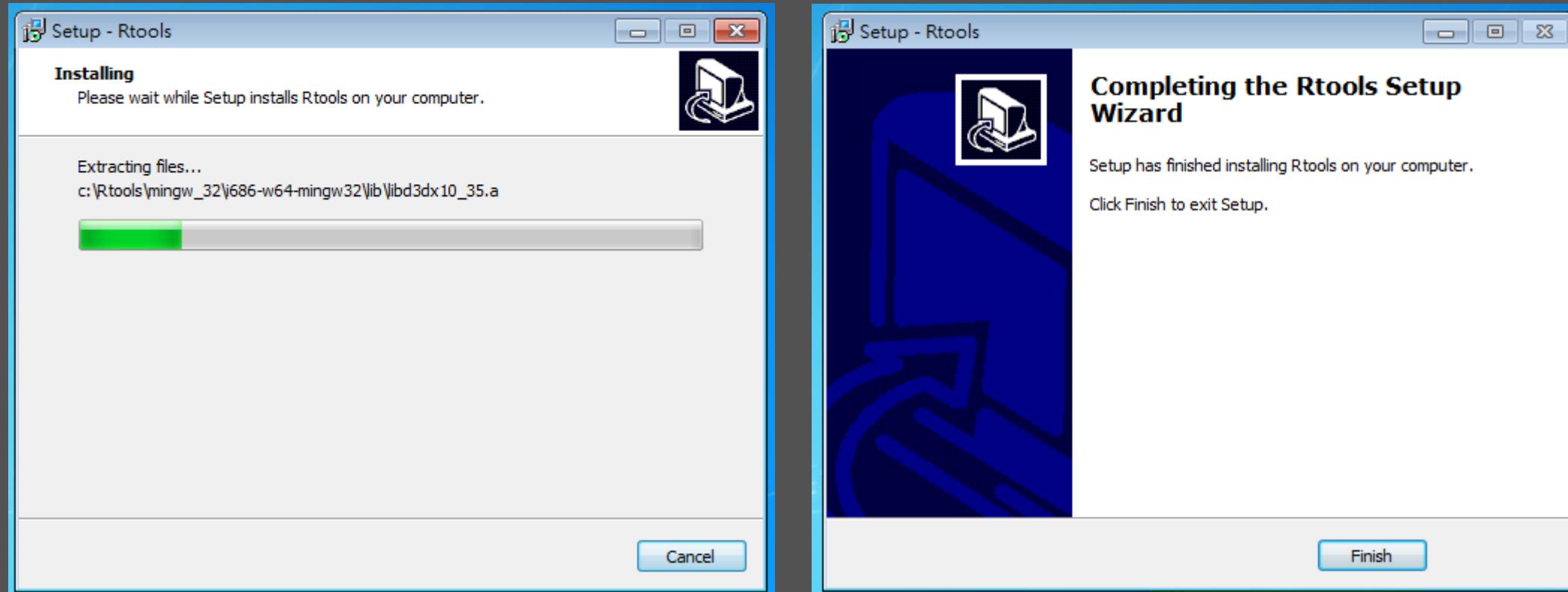
下一步或用預設即可 ...



R MariaDB 與 RMYSQL 套件設定 by Windows

Rtool 安裝

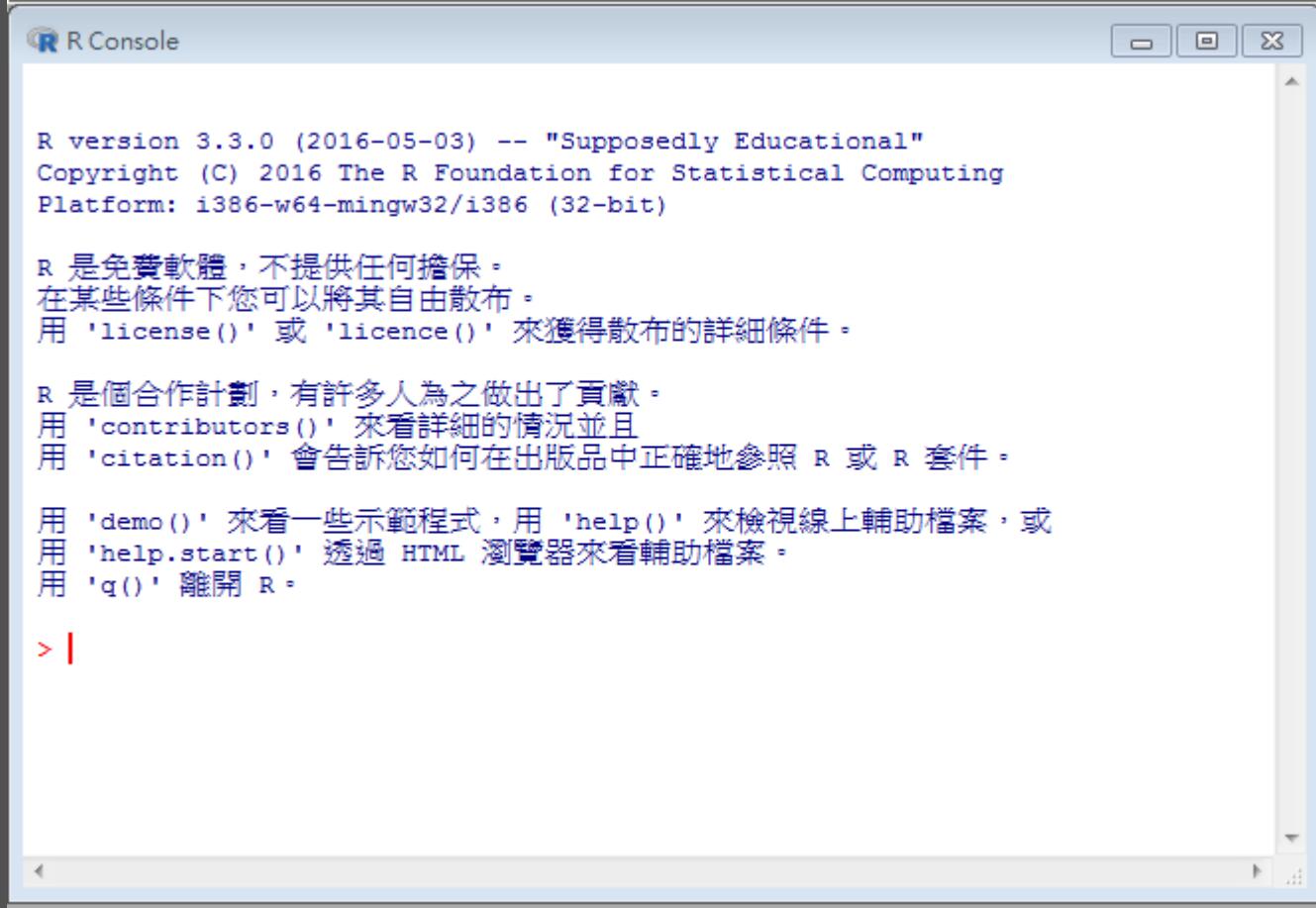
下一步或用預設即可 ...



R MariaDB 與 RMYSQL 套件設定 by Windows

R

安裝細節不難也是下一步即可，這裡要用 32bit 的部分去玩



The screenshot shows the R Console window with the following text displayed:

```
R version 3.3.0 (2016-05-03) -- "Supposedly Educational"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: i386-w64-mingw32/i386 (32-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。

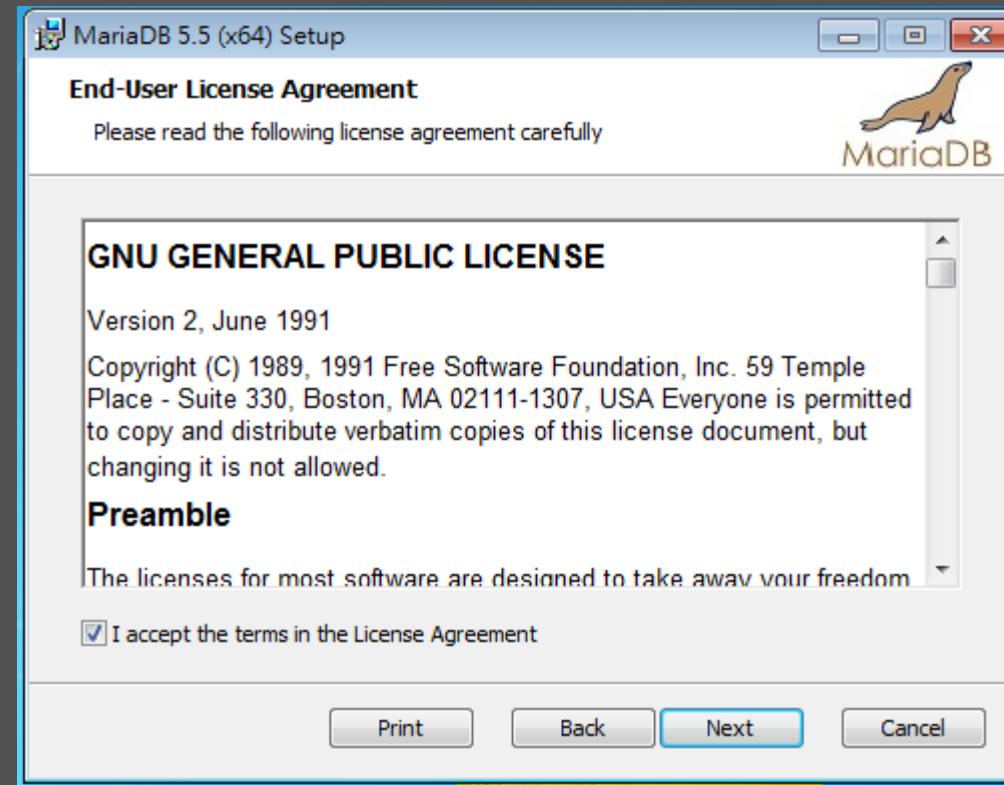
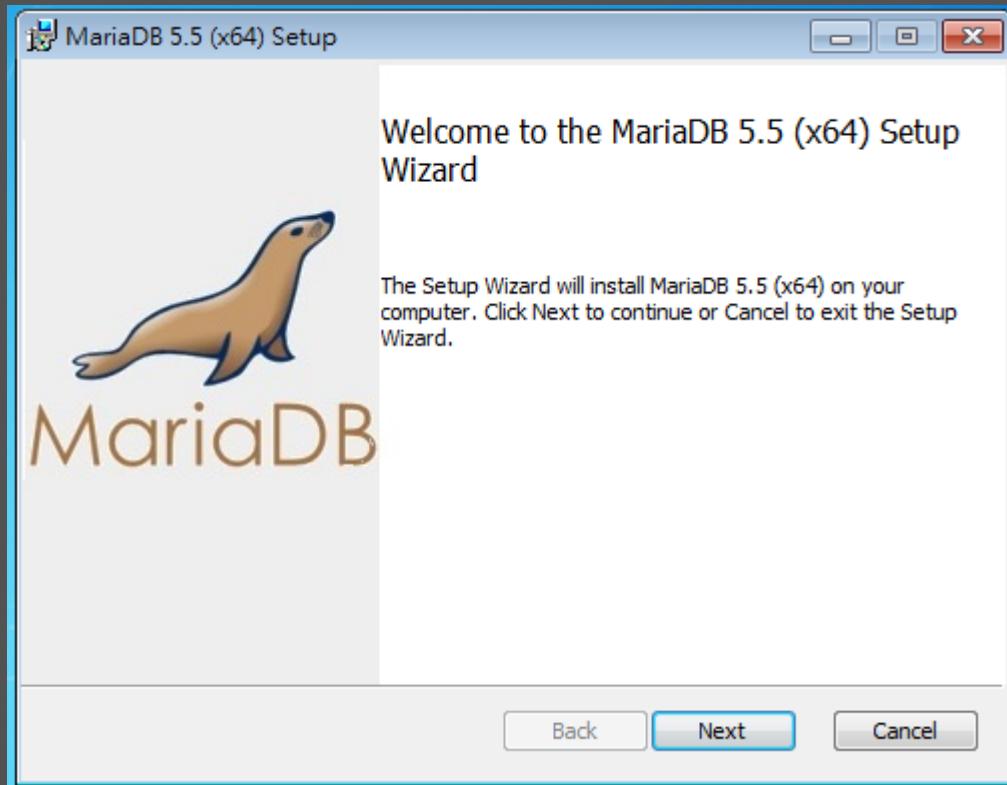
用 'demo()' 來看一些示範程式，用 'help()' 來檢視線上輔助檔案，或
用 'help.start()' 透過 HTML 瀏覽器來看輔助檔案。
用 'q()' 離開 R。


> |
```

R MariaDB 與 RMYSQL 套件設定 by Windows

MariaDB

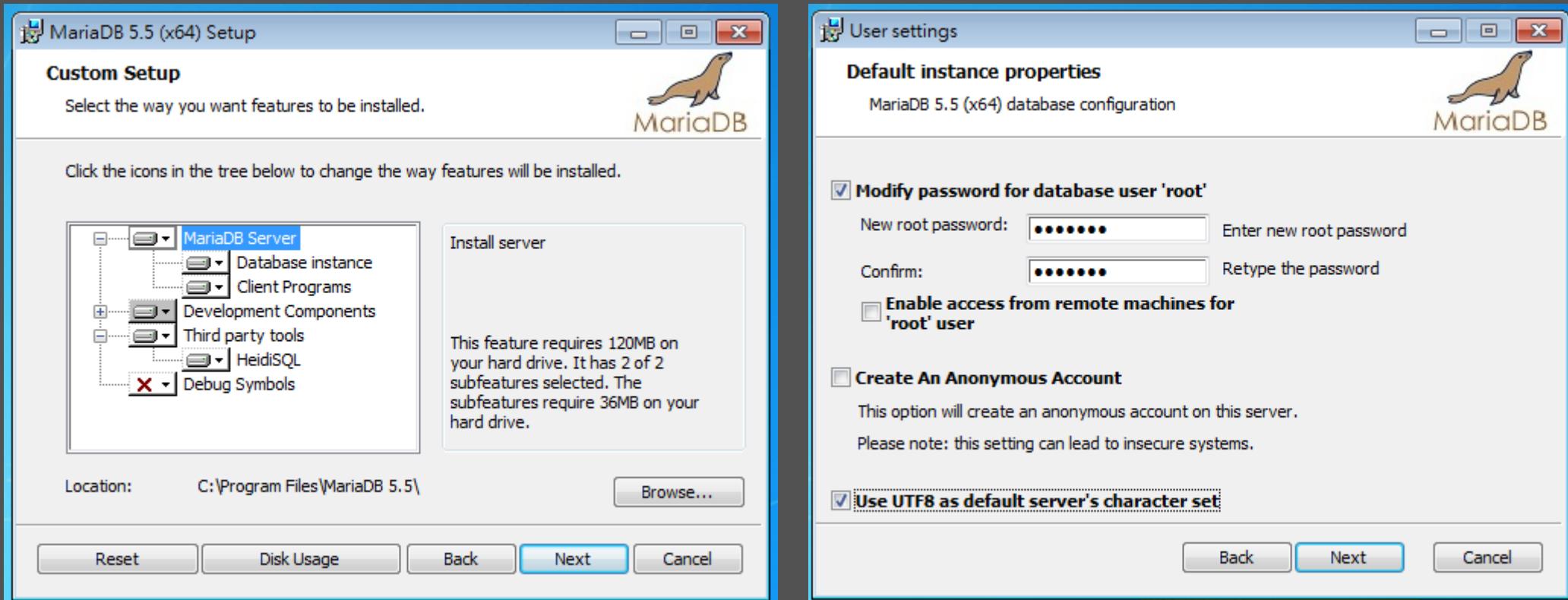
資料庫帳密設定與文字編碼的地方要注意，因為後面我們要用。



R MariaDB 與 RMYSQL 套件設定 by Windows

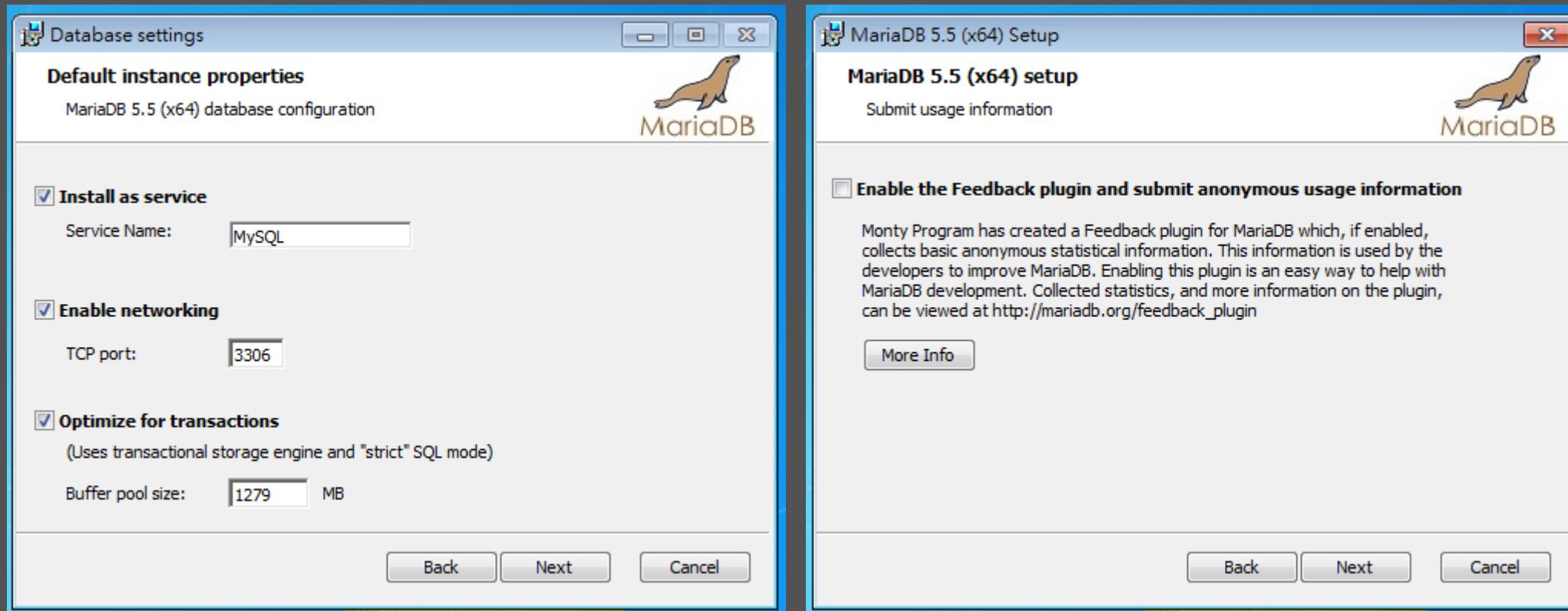
MariaDB

這裡要注意設定資料庫的密碼，我在這裡設定UTF8編碼。



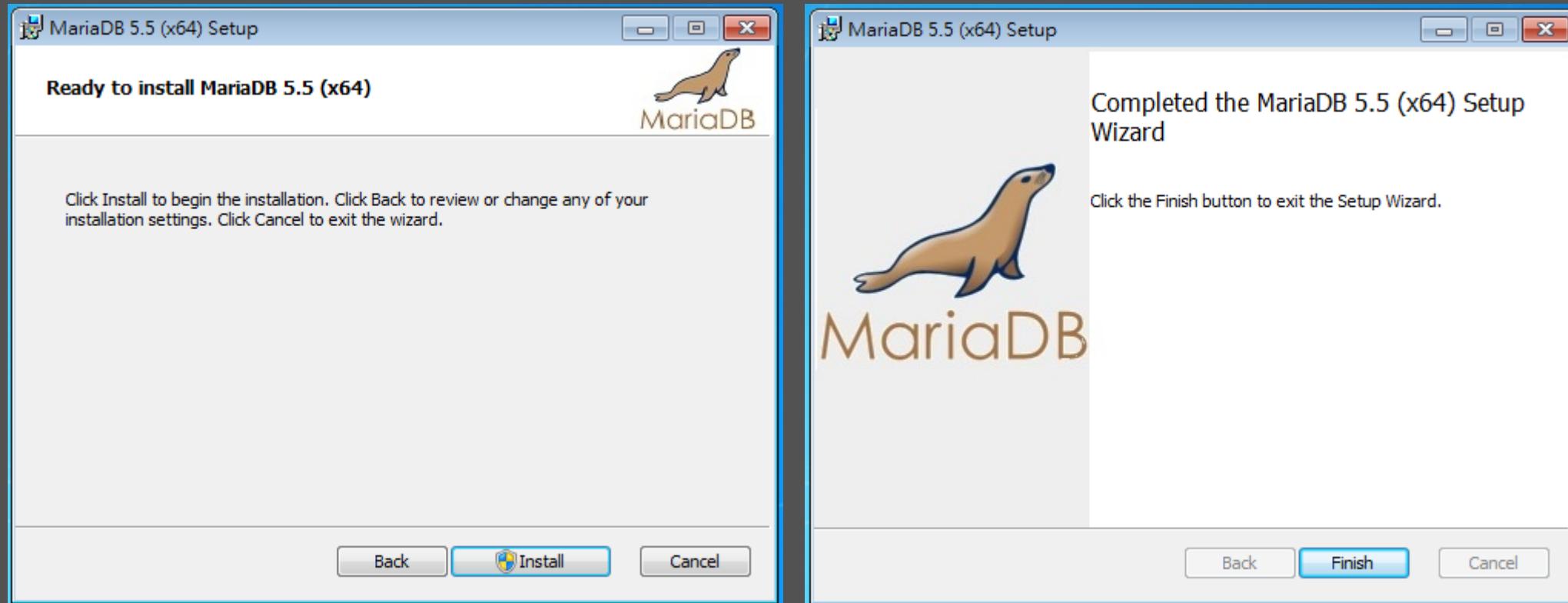
R MariaDB 與 RMYSQL 套件設定 by Windows

MariaDB



R MariaDB 與 RMYSQL 套件設定 by Windows

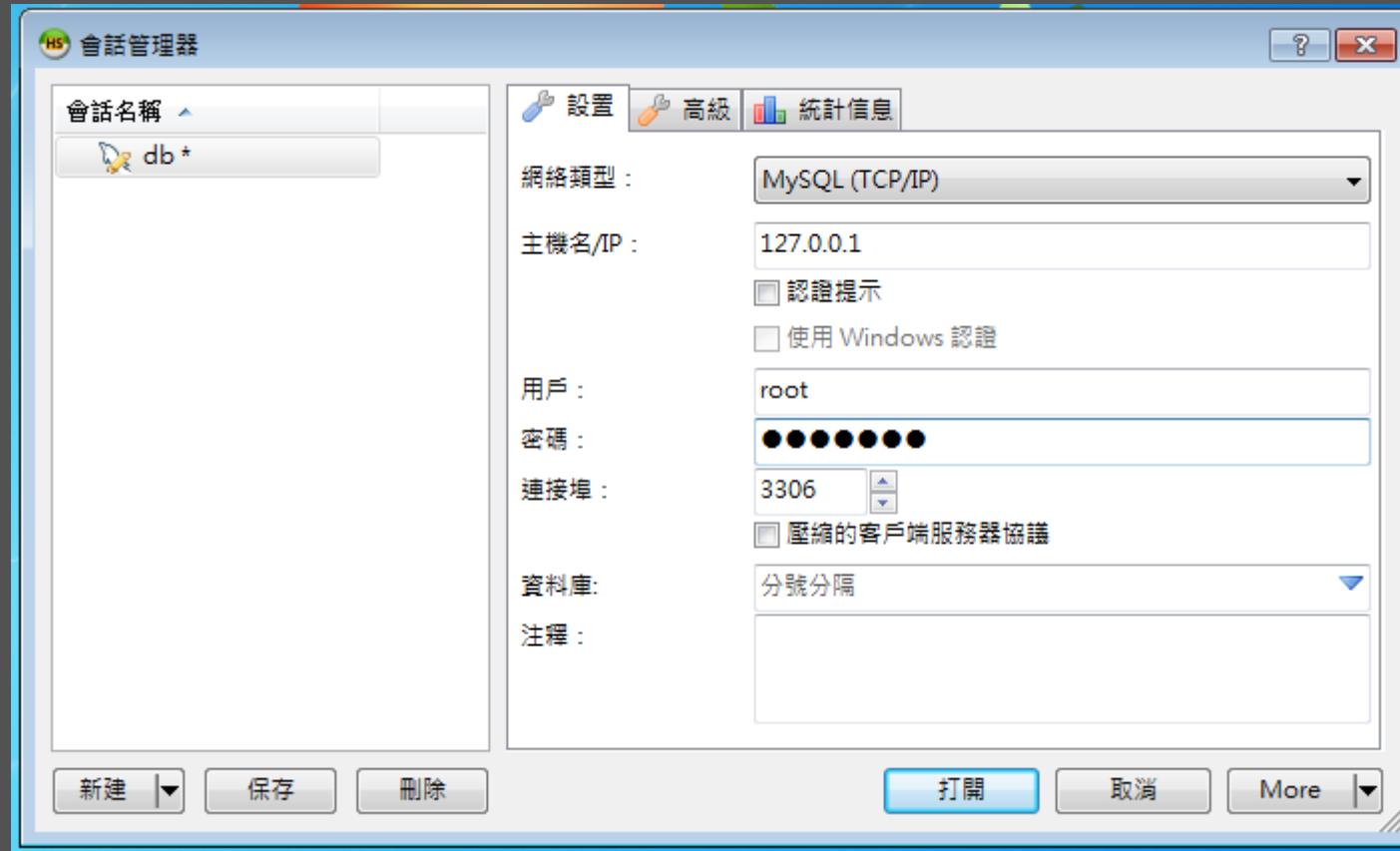
MariaDB



R MariaDB 與 RMYSQL 套件設定 by Windows

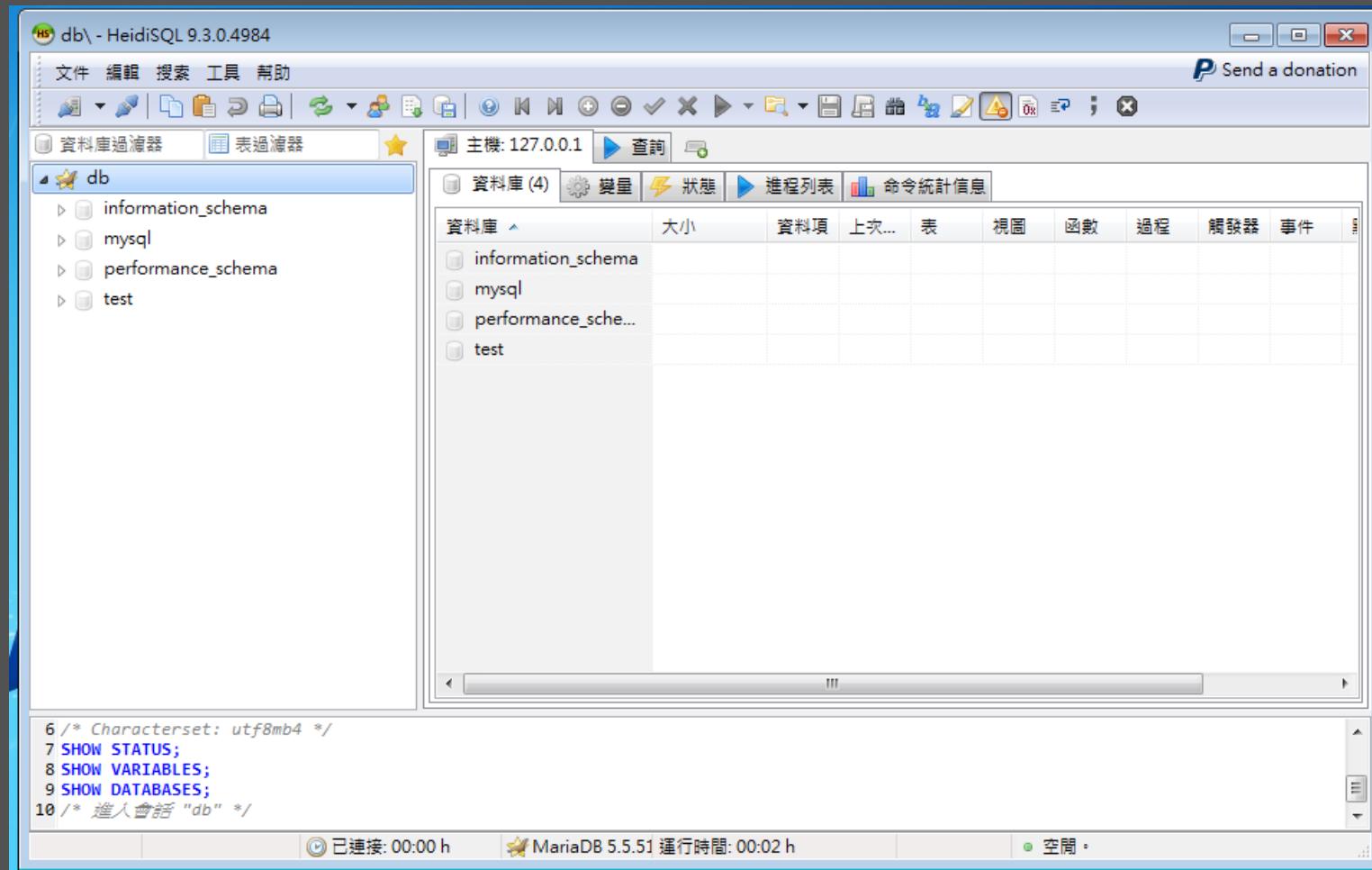
MariaDB

在這裡我們用安裝 MariaDB 中就有的介面，HeidiSQL 新增一個會話名稱輸入剛剛設定的密碼即可登入。



R MariaDB 與 RMYSQL 套件設定 by Windows

MariaDB
資料庫 OK!!!



R MariaDB 與 RMYSQL套件設定 by Windows

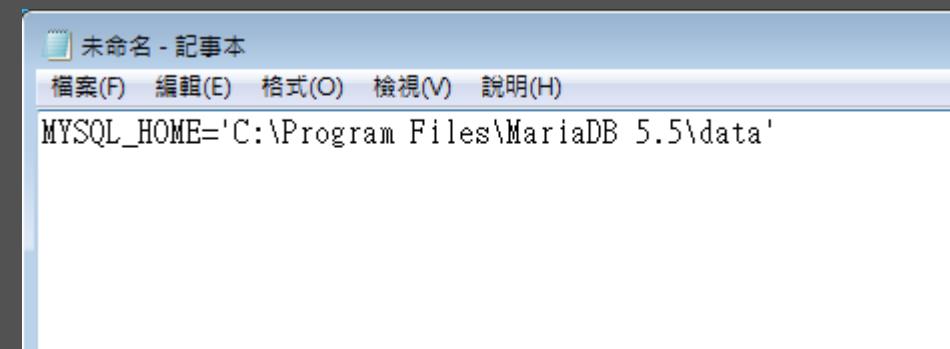
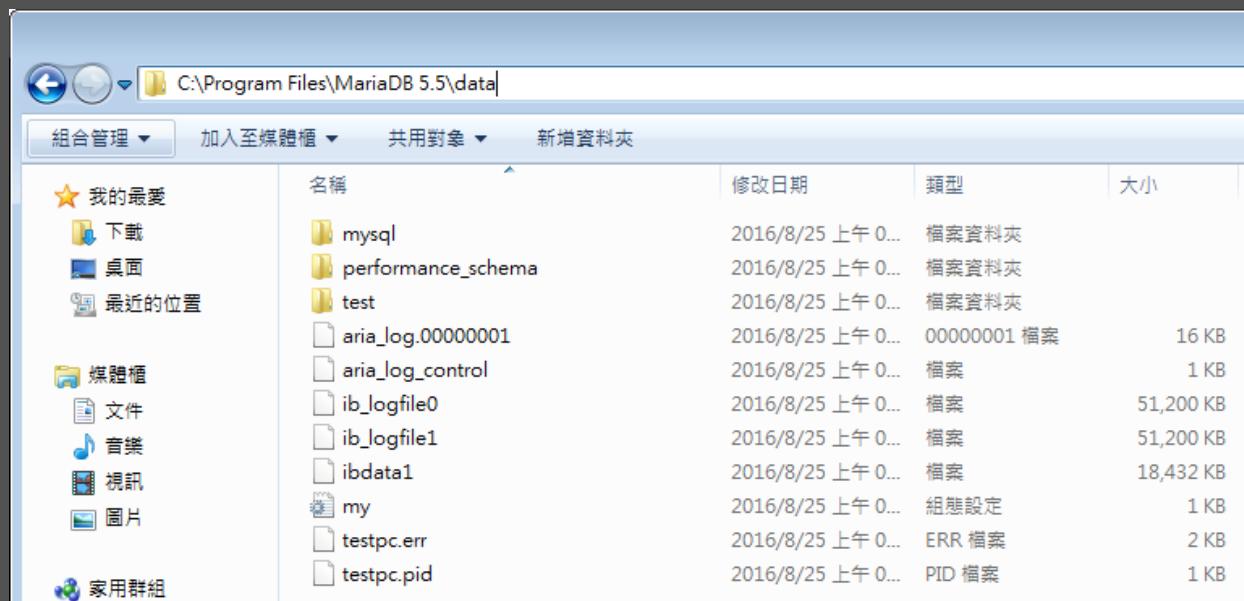
Renviron.site

接下來要設定在 R 3.3.0 下面的 /etc 目錄新增一個名為 Renviron.site 的檔案，我們可以使用任何編輯器去做（為了方便就直接用記事本了）

Renviron.site 內容為

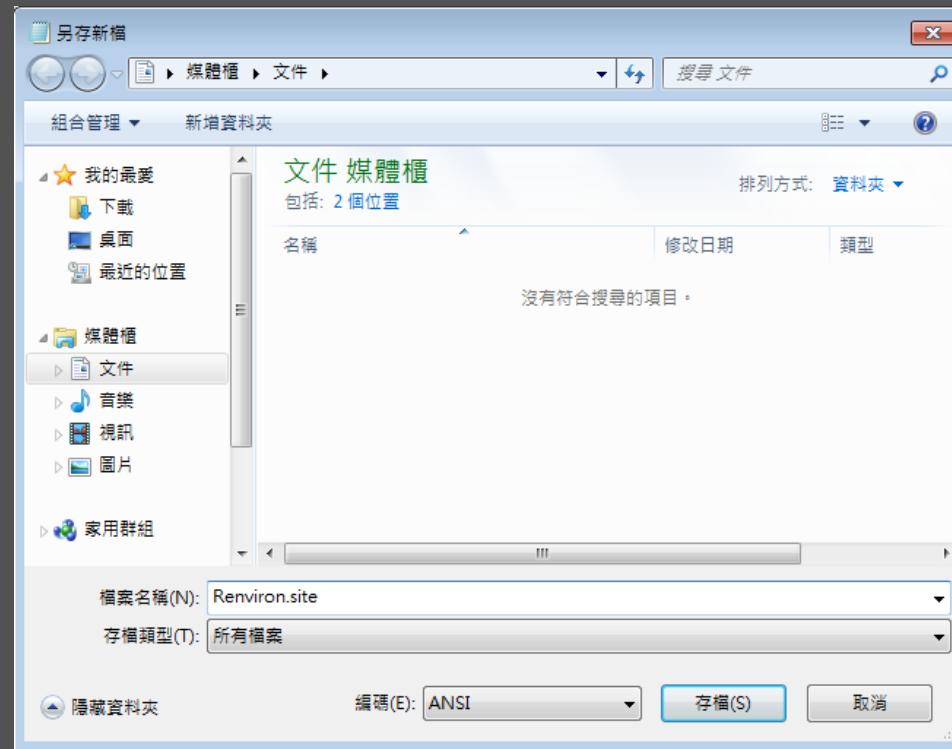
MYSQL_HOME='MariaDB路徑'

其實就是在 C 槽 Program Files 安裝的 MariaDB 下面的 /data 目錄，完成後就放到其版本下的 /etc 目錄....



R MariaDB 與 RMYSQL 套件設定 by Windows

Renvironment.site



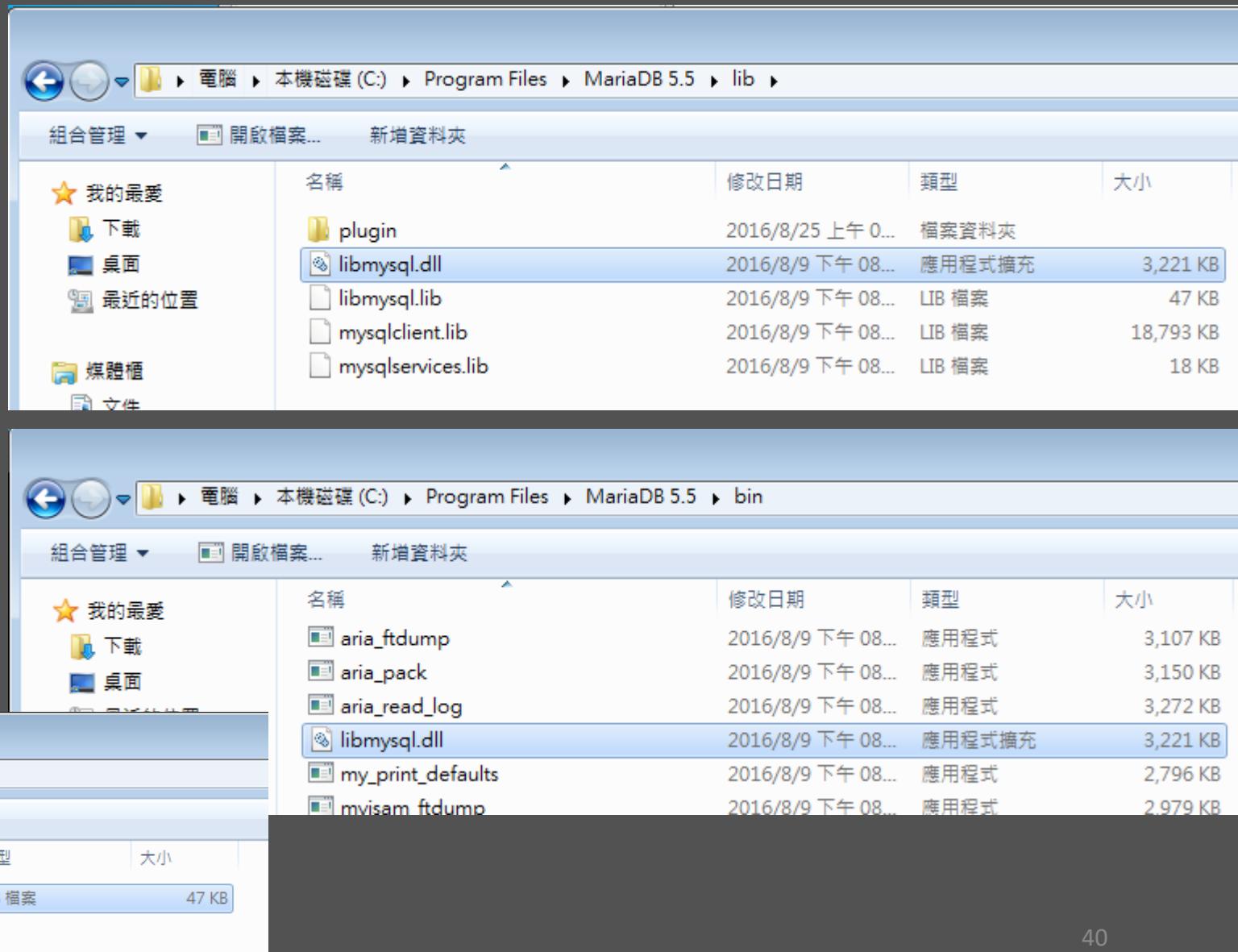
名稱	修改日期	類型	大小
i386	2016/8/25 上午 0...	檔案資料夾	
x64	2016/8/25 上午 0...	檔案資料夾	
curl-ca-bundle	2015/11/28 上午 ...	安全性憑證	251 KB
Rcmd_environ	2013/6/20 上午 1...	檔案	2 KB
Rconsole	2010/3/17 下午 0...	檔案	2 KB
Rdevga	2011/10/3 上午 1...	檔案	1 KB
Renvironment	2016/8/25 上午 0...	SITE 檔案	1 KB
repositories	2016/3/17 上午 0...	檔案	2 KB
rgb	2010/3/17 下午 0...	文字文件	30 KB
Rprofile	2013/5/20 上午 1...	SITE 檔案	1 KB

R MariaDB 與 RMYSQL 套件設定 by Windows

MariaDB 的設定

1. 將 MariaDB 下 /lib 目錄中的 libmysql.dll 複製至 /bin 目錄之中
2. 將 MariaDB 下 /lib 目錄中的 libmysql.lib 複製至 /lib/opt 目錄之下

...由於沒有 /opt 目錄所以直接建一個



R MariaDB 與 RMySQL 套件設定 by Windows

完成

想要知道前面 `Renvironment.site` 有無成功可用 `Sys.getenv('MYSQL_HOME')`

`Sys.getenv('MYSQL_HOME')`

```
> Sys.getenv('MYSQL_HOME')
[1] "C:\\Program Files\\MariaDB 5.5\\data"
> |
```

安裝 RMySQL 套件

`install.packages("RMySQL")`

```
> install.packages("RMySQL")
Installing package into 'C:/Users/test/Documents/R/win-library/3.3'
(as 'lib' is unspecified)
嘗試 URL 'https://cran.usthb.dz/bin/windows/contrib/3.3/RMySQL_0.10.9.zip'
Content type 'application/zip' length 1914362 bytes (1.8 MB)
downloaded 1.8 MB

package 'RMySQL' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\test\AppData\Local\Temp\RtmpgFzybL\downloaded_packages
> |
```

R MariaDB 與 RMYSQL套件設定 by Windows

DB

```
CREATE DATABASE school;
SHOW DATABASES;
USE school;
CREATE TABLE stdscore
(stdscoreid int auto_increment primary key,
stdno char(7),
name char(50),
birthdate date,
courseno char(6),
courname char(50),
teacher char(50),
score int);

SHOW TABLES;

INSERT INTO stdscore (stdno, name, birthdate, courseno, courname, teacher, score)
VALUES('9643001', 'Linda', '1982/1/15', 'MS1038', 'Statistics', 'Henry Chen', 88);
INSERT INTO stdscore (stdno, name, birthdate, courseno, courname, teacher, score)
VALUES('9643001', 'Linda', '1982/1/15', 'MS1014', 'Data Structure', 'James Lee', 82);
INSERT INTO stdscore (stdno, name, birthdate, courseno, courname, teacher, score)
VALUES('9643018', 'Tracy', '1982/11/18', 'MS1014', 'Data Structure', 'James Lee', 77);

INSERT INTO stdscore (stdno, name, birthdate, courseno, courname, teacher, score)
VALUES('9643025', 'Tom', '1982/9/17', 'MS1038', 'Statistics', 'Henry Chen', 65);
SELECT * FROM stdscore;
```

The screenshot shows the MySQL Workbench interface. The top part is a query editor window titled '查詢*' (Query) with the connection set to '主機: 127.0.0.1'. It contains the SQL code provided in the text block above. The bottom part shows the 'SCHEMATA (1×5)' tab, which lists the databases: information_schema, mysql, performance_schema, school, and test. The 'TABLE NAMES (1×1)' tab is also visible.

R MariaDB 與 RMySQL 套件設定 by Windows

R

```
library(RMySQL)
con = dbConnect(MySQL(), user="root", password="hitachi", dbname="school",
host="localhost")
dbListTables(con)
dbListFields(con, "stdscore")
data.all = dbReadTable(con, "stdscore")
class(data.all)
data.all
data.select = dbGetQuery(con, "select * from stdscore where courseno='MS1038'")
data.select
summary(MySQL(), verbose = TRUE)
summary(con, verbose = TRUE)
summary(data.all, verbose = TRUE)
dbListConnections(MySQL())
dbDisconnect(con)
```

```
> library(RMySQL)
Loading required package: DBI
Warning message:
package 'RMySQL' was built under R version 3.3.1
> con = dbConnect(MySQL(), user="root", password="hitachi", dbname="school", host="localhost")
> dbListTables(con)
[1] "stdscore"
> dbListFields(con, "stdscore")
[1] "stdscoreid" "stdno"      "name"       "birthdate"   "courseno"   "courname"   "teacher"    "score"
> data.all = dbReadTable(con, "stdscore")
> class(data.all)
[1] "data.frame"
> data.all
  stdscoreid stdno name birthdate courseno      courname teacher score
1          1 9643001 Linda 1982-01-15 MS1038 Statistics Henry Chen    88
2          2 9643001 Linda 1982-01-15 MS1014 Data Structure James Lee     82
3          3 9643018 Tracy 1982-11-18 MS1014 Data Structure James Lee     77
> data.select = dbGetQuery(con, "select * from stdscore where courseno='MS1038'")
> data.select
  stdscoreid stdno name birthdate courseno      courname teacher score
1          1 9643001 Linda 1982-01-15 MS1038 Statistics Henry Chen    88
```

R MariaDB 與 RMYSQL套件設定 by Ubuntu Linux

更新與安裝(在此之前要先裝好 R、Ubuntu Linux (LAMP))

```
~$ sudo apt-get update
```

```
~$ sudo apt-get install r-cran-rmysql
```

```
kancheng@ubndjsrserv:~$ sudo apt-get update
下載:1 http://cran.csie.ntu.edu.tw/bin/linux/ubuntu xenial/ InRelease [3,590 B]
略過:2 http://dl.google.com/linux/chrome/deb stable InRelease
已有:3 http://tw.archive.ubuntu.com/ubuntu xenial InRelease
已有:4 http://tw.archive.ubuntu.com/ubuntu xenial-updates InRelease
已有:5 http://tw.archive.ubuntu.com/ubuntu xenial-backports InRelease
已有:6 http://dl.google.com/linux/chrome/deb stable Release
略過:1 http://cran.csie.ntu.edu.tw/bin/linux/ubuntu xenial/ InRelease
下載:8 http://security.ubuntu.com/ubuntu xenial-security InRelease [94.5 kB]
```

```
kancheng@ubndjsrserv:~$ sudo apt-get install r-cran-rmysql -y
正在讀取套件清單... 完成
正在重建相依關係
正在讀取狀態資料... 完成
以下套件為自動安裝，並且已經無用：
 libcdaudio0 libdirectfb-1.2-9 libenca0 libmpcdec6 libslv2-9
Use 'sudo apt autoremove' to remove them.
The following additional packages will be installed:
 libmariadb2 r-cran-dbi
下列【新】套件將會被安裝：
 libmariadb2 r-cran-dbi r-cran-rmysql
升級 0 個，新安裝 3 個，移除 0 個，有 0 個未被升級。
需要下載 429 kB 的套件檔。
此操作完成之後，會多佔用 888 kB 的磁碟空間。
```

R MariaDB 與 RMySQL 套件設定 by Ubuntu Linux

安裝 RMySQL 套件

```
install.packages("RMySQL")
```

根據上述的訊息....

安裝 libmariadb-client-lgpl-dev 套件，
並找出
/tmp/RtmpFBqpwR/downloaded_packages 下的 RMySQL 與 DB 目錄，並對
R 做設定。

```
/tmp/Rtmp*****/downloaded_packages
```

```
> install.packages("RMySQL")
Installing package into '/home/kancheng/R/x86_64-pc-linux-gnu-library/3.3'
(as 'lib' is unspecified)
嘗試 URL 'https://cran.ms.unimelb.edu.au/src/contrib/RMySQL_0.10.9.tar.gz'
Content type 'application/x-gzip' length 54012 bytes (52 KB)
=====
downloaded 52 KB

* installing *source* package 'RMySQL' ...
** package 'RMySQL' successfully unpacked and MD5 sums checked
Using PKG_CFLAGS=
Using PKG_LIBS=-lmysqlclient
----- ANTICONF ERROR -----
Configuration failed because libmysqlclient was not found. Try installing:
* deb: libmariadb-client-lgpl-dev (Debian, Ubuntu 16.04)
      libmariadbclient-dev (Ubuntu 14.04)
* rpm: mariadb-devel | mysql-devel (Fedora, CentOS, RHEL)
* csw: mysql56_dev (Solaris)
* brew: mariadb-connector-c (OSX)
If libmysqlclient is already installed, check that 'pkg-config' is in your
PATH and PKG_CONFIG_PATH contains a libmysqlclient.pc file. If pkg-config
is unavailable you can set INCLUDE_DIR and LIB_DIR manually via:
R CMD INSTALL --configure-vars='INCLUDE_DIR=... LIB_DIR=...'
-----
ERROR: configuration failed for package 'RMySQL'
* removing '/home/kancheng/R/x86_64-pc-linux-gnu-library/3.3/RMySQL'

The downloaded source packages are in
  '/tmp/RtmpFBqpwR/downloaded_packages'
Warning message:
In install.packages("RMySQL") :
  installation of package 'RMySQL' had non-zero exit status
> █
```

R MariaDB 與 RMySQL 套件設定 by Ubuntu Linux

安裝 RMySQL 套件

```
~$ sudo apt-get install libmariadb-client-lgpl-dev -y
```

```
~$ ls /tmp/RtmpFBqpwR/downloaded_packages
```

```
~$ whereis mysql
```

```
~$ R CMD INSTALL --configure-args='--with-mysql-dir=/usr/lib/mysql'  
/tmp/RtmpFBqpwR/downloaded_packages/RMySQL_0.10.9.tar.gz
```

```
kancheng@ubndjsrserv:~$ sudo apt-get install libmariadb-client-lgpl-dev -y  
正在讀取套件清單... 完成  
正在重建相依關係  
正在讀取狀態資料... 完成  
以下套件為自動安裝，並且已經無用：  
    libcdaudio1 libdirectfb-1.2-9 libenca0 libmpcdec6 libslv2-9  
Use 'sudo apt autoremove' to remove them.  
下列【新】套件將會被安裝：  
    libmariadb-client-lgpl-dev  
升級 0 個，新安裝 1 個，移除 0 個，有 0 個未被升級。
```

```
kancheng@ubndjsrserv:~$ ls /tmp/RtmpFBqpwR/downloaded_packages  
RMySQL_0.10.9.tar.gz  
kancheng@ubndjsrserv:~$ whereis mysql  
mysql: /usr/bin/mysql /usr/lib/mysql /etc/mysql /usr/share/mysql /usr/share/man/man1/mysql.1.gz  
kancheng@ubndjsrserv:~$  
kancheng@ubndjsrserv:~$ R CMD INSTALL --configure-args='--with-mysql-dir=/usr/lib/mysql' /tmp/RtmpFBqp  
wR/downloaded_packages/RMySQL_0.10.9.tar.gz  
* installing to library '/home/kancheng/R/x86_64-pc-linux-gnu-library/3.3'  
* installing *source* package 'RMySQL' ...  
** package 'RMySQL' successfully unpacked and MD5 sums checked  
Found mysql_config cflags and libs!  
Using PKG_CFLAGS=-I/usr/include/mariadb -g -O2 -fstack-protector-strong -Wformat -Werror=format-securi  
ty -D_FORTIFY_SOURCE=2 -Wunused -Wno-uninitialized  
Using PKG_LIBS=-L/usr/lib/x86_64-linux-gnu -lmariadb  
** libs
```

R MariaDB 與 RMYSQL套件設定 by Ubuntu Linux

進入資料庫，建立之後的測試資料(SQL)

```
~$ mysql -u root -p
```

```
> create database rmysql;  
> grant all on rmysql.* to rmysql@'%' identified by 'rmysql';  
> grant all on rmysql.* to rmysql@localhost identified by 'rmysql';  
> use rmysql
```

```
> CREATE TABLE t_user(  
id INT PRIMARY KEY AUTO_INCREMENT,  
user varchar(12) NOT NULL UNIQUE  
)ENGINE=INNODB DEFAULT CHARSET=utf8;
```

```
> INSERT INTO t_user(user) values('A1'),('AB'),('fens.me');
```

```
> SELECT * FROM t_user;
```

```
kancheng@ubndjsr:~$ mysql -u root -p  
Enter password:  
Welcome to the MariaDB monitor.  Commands end with ; or \g.  
Your MariaDB connection id is 7  
Server version: 10.0.25-MariaDB-0ubuntu0.16.04.1 Ubuntu 16.04  
  
Copyright (c) 2000, 2016, Oracle, MariaDB Corporation Ab and others.  
  
Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.  
  
MariaDB [(none)]>  
MariaDB [(none)]> create database rmysql;  
Query OK, 1 row affected (0.00 sec)  
  
MariaDB [(none)]> grant all on rmysql.* to rmysql@'%' identified by 'rmysql';  
Query OK, 0 rows affected (0.00 sec)  
  
MariaDB [(none)]> grant all on rmysql.* to rmysql@localhost identified by 'rmysql';  
Query OK, 0 rows affected (0.00 sec)  
  
MariaDB [(none)]> use rmysql  
Database changed  
  
MariaDB [rmysql]> CREATE TABLE t_user(  
    -> id INT PRIMARY KEY AUTO_INCREMENT,  
    -> user varchar(12) NOT NULL UNIQUE  
    -> )ENGINE=INNODB DEFAULT CHARSET=utf8;  
Query OK, 0 rows affected (0.04 sec)  
  
MariaDB [rmysql]>  
MariaDB [rmysql]> INSERT INTO t_user(user) values('A1'),('AB'),('fens.me');  
Query OK, 3 rows affected (0.01 sec)  
Records: 3  Duplicates: 0  Warnings: 0  
  
MariaDB [rmysql]>  
MariaDB [rmysql]> SELECT * FROM t_user;  
+----+----+  
| id | user |  
+----+----+  
| 1  | A1  |  
| 2  | AB  |  
| 3  | fens.me |  
+----+----+  
3 rows in set (0.00 sec)  
  
MariaDB [rmysql]> █
```

R MariaDB 與 RMySQL 套件設定 by Ubuntu Linux

進到之前裝的 phpmyAdmin 可以看到資料庫已經建立起來了....

The screenshot shows a dual-monitor setup. On the left monitor, a Chrome browser window displays the phpMyAdmin interface for the 'rmysql' database. The 'Structure' tab is selected, showing a single table named 't_user' with three columns: id, user, and password. The table has 3 InnoDB rows with data: (1, 'A1'), (2, 'AB'), and (3, 'fens.me'). On the right monitor, an R terminal window is open, displaying the R license and copyright information, followed by a session where the RMySQL package is loaded and a query is run against the 't_user' table.

```
library("RMySQL")
conn = dbConnect(MySQL(),
                 dbname = "rmysql",
                 username = "root",
                 password = "hitachi")
users = dbGetQuery(conn,
                  "SELECT * FROM t_user")
dbDisconnect(conn)
users
```

```
R version 3.3.1 (2016-06-21) -- "Bug in Your Hair"
Copyright (C) 2016 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (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。

> library(RMySQL)
Loading required package: DBI
> conn = dbConnect(MySQL( ), dbname = "rmysql", username="root", password="hitachi")
>
> users = dbGetQuery(conn, "SELECT * FROM t_user")
> dbDisconnect(conn)
[1] TRUE
> users
   id    user
1  1     A1
2  2     AB
3  3 fens.me
>
```

R 統計軟體 -R Commander & foreign

1. SPSS

```
library(foreign)
# SPSS
spss = read.spss("C:/.../spss_ex.sav",
                 use.value.labels=TRUE,
                 max.value.labels=Inf,
                 to.data.frame=TRUE)
colnames(spss) = tolower(colnames(spss))
```

2. SAS

```
# SAS
.Datasets = read.xport("C:/.../sas_ex.xpt")
sas = .Datasets
remove(.Datasets)
```

3. STATA

```
# STATA
stata = read.dta("C:/.../stata_ex.dta",
                  convert.dates=TRUE,
                  convert.factors=TRUE,
                  missing.type=TRUE,
                  convert.underscore=TRUE,
                  warn.missing.labels=TRUE)
```

4. EXCEL

```
# Excel
excel <- readXL("C:/.../excel_ex.xls",
                  rownames=FALSE,
                  header=TRUE, na="",
                  sheet="Tomato_First",
                  stringsAsFactors=TRUE)
```

RData

預設會存在使用者家目錄的文件目錄

save() 函數 - 儲存

save(tomato, file = "tomato.rdata") -> save(物件名稱, file = "檔名.rdata")

儲存多個物件

save(n, r, w, file = "multiple.rdata") -> save(物件名稱1, 物件名稱2, 物件名稱3, file = "檔名.rdata")

rm() 函數 - 移除

rm(tomato) -> rm(物件名稱)

head() 函數 - 檢視

head(tomato) -> head(物件名稱)

load() 函數 - 載入

load("tomato.rdata") -> load("檔名.rdata")

R 內建資料集

1. ggplot2 套件的 diamonds 資料 (鑽石)

```
require(ggplot2)  
data(diamonds)  
head(diamonds)
```

```
> require(ggplot2)  
Loading required package: ggplot2  
> data(diamonds)  
> head(diamonds)  
   carat       cut color clarity depth table price     x     y     z  
1  0.23      Ideal    E     SI2   61.5     55   326 3.95 3.98 2.43  
2  0.21      Premium  E     SI1   59.8     61   326 3.89 3.84 2.31  
3  0.23       Good   E     VS1   56.9     65   327 4.05 4.07 2.31  
4  0.29      Premium I     VS2   62.4     58   334 4.20 4.23 2.63  
5  0.31       Good   J     SI2   63.3     58   335 4.34 4.35 2.75  
6  0.24  Very Good J     VVS2   62.8     57   336 3.94 3.96 2.48  
> |
```

2. IRIS 翠尾花 生物資料集

```
data(iris)  
head(iris)
```

```
> data(iris)  
> head(iris)  
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species  
1          5.1         3.5          1.4         0.2  setosa  
2          4.9         3.0          1.4         0.2  setosa  
3          4.7         3.2          1.3         0.2  setosa  
4          4.6         3.1          1.5         0.2  setosa  
5          5.0         3.6          1.4         0.2  setosa  
6          5.4         3.9          1.7         0.4  setosa  
> |
```

R - readHTMLTable

```
require(XML)
theURL = "http://XXXX"
Data = readHTMLTable(theURL, which = 1, header = FALSE, stringsAsFactors = FALSE)
```

XML v3.98-1.9 Other versions ▾
by ORPHANED
21,129 ⚡ 96th
Monthly downloads > Percentile
https://www.rdocumentation.org/packages/XML Copy

Tools for Parsing and Generating XML Within R and S-Plus
Many approaches for both reading and creating XML (and HTML) documents (including DTDs), both local and accessible via HTTP or FTP. Also offers access to an 'XPath' interpreter.

which -> 可選要讀取的列表

header -> 表頭的設定

stringsAsFactors -> 防範 factor 和 character 直行上的問題

readHTMLTable

Read Data From One Or More HTML Tables

This function and its methods provide somewhat robust methods for extracting data from HTML tables in an HTML document. One can read all the tables in a document given by filename or (`http:` or `ftp:`) URL, or having already parsed the document via `htmlParse`. Alternatively, one can specify an individual `<table>` node in the document.

The methods attempt to do some heuristic computations to determine the header labels for the columns, the name of the table, etc.

Keywords `data, IO`

From XML v3.98-1.9
by ORPHANED 96th Percentile

R - 汇入多個 CSV 檔案

```
setwd("C:/....")
getwd()
csvpath = "C:/..../"
csvfilesn = list.files( path = csvpath, pattern="*.csv")
tmprt = function(rtcsv){
  read.csv( rtcsv, stringsAsFactors = FALSE)
}
data = lapply(paste(csvpath,csvfilesn, sep = ""), tmprt)
csvfilesn
data = data[[1]]
```

R - 利用自訂函數來匯入多個 CSV 檔案

```
a = c(10:1)
b = c(11:20)
c = c(21:30)
test = data.frame( a, b, c)
test
colnames(test) = c("kan", "hao", "cheng")
test
write.table( test, file = "helloWorld.CSV", sep = ",")
getwd()
write.table( test, file= "HW.CSV",
             quote = FALSE, sep = ",",
             row.names = FALSE)
```

	A	B	C	D	E
1	kan	hao	cheng		
2	10	11	21		
3	9	12	22		
4	8	13	23		
5	7	14	24		
6	6	15	25		
7	5	16	26		
8	4	17	27		
9	3	18	28		
10	2	19	29		
11	1	20	30		
12					

	kan	hao	cheng
1	10	11	21
2	9	12	22
3	8	13	23
4	7	14	24
5	6	15	25
6	5	16	26
7	4	17	27
8	3	18	28
9	2	19	29
10	1	20	30

	A	B	C	D
1	kan	hao	cheng	
2	10	11	21	
3	9	12	22	
4	8	13	23	
5	7	14	24	
6	6	15	25	
7	5	16	26	
8	4	17	27	
9	3	18	28	
10	2	19	29	
11	1	20	30	
12				

```
> a = c(10:1)
> b = c(11:20)
> c = c(21:30)
> test = data.frame( a, b, c)
> test
   a   b   c
1 10 11 21
2  9 12 22
3  8 13 23
4  7 14 24
5  6 15 25
6  5 16 26
7  4 17 27
8  3 18 28
9  2 19 29
10 1 20 30
> colnames(test) = c("kan", "hao", "cheng")
> test
  kan hao cheng
1 10 11 21
2  9 12 22
3  8 13 23
4  7 14 24
5  6 15 25
6  5 16 26
7  4 17 27
8  3 18 28
9  2 19 29
10 1 20 30
> write.table( test, file = "helloWorld.CSV", sep = ",")
> getwd()
[1] "D:/USERDATA/Documents"
> write.table( test, file= "HW.CSV", quote = FALSE, sep = ",", row.names = FALSE)
> |
```

R 多個套件安裝與檢查

```
# set a CRAN mirror
#Automatic redirection to servers worldwide, currently sponsored by
Rstudio
local({r = getOption("repos")}
r["CRAN"] = "https://cloud.r-project.org/"
options(repos=r)}

# install package name
pkgs = c("devtools", "xml2", "plyr", "dplyr", "data.table", "text2vec",
"jiebaR", "fmsb", "e1071", "rpart", "randomForest", "GGally", "lubridate",
"stringr", "ggplot2", "scales", "reshape2", "data.table", "coefplot", "Rcmdr",
"broom", "xlsx", "shiny")

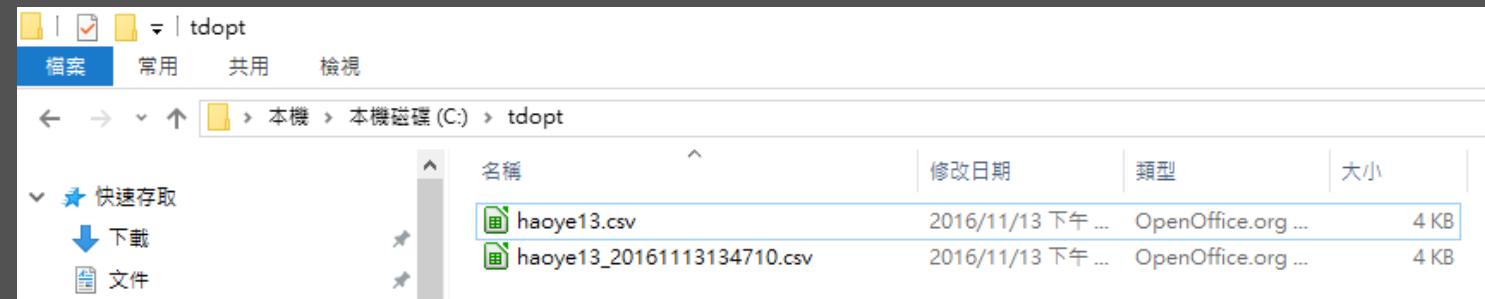
pkgs = pkgs[!(pkgs %in% installed.packages()[,"Package"] )]

#install
if(length(pkgs)) install.packages(pkgs)
```

R 使用自訂函數來處理匯出檔案覆蓋的問題

```
# 1
wrta = function(xdo , ycsv){
write.table( xdo, file= ycsv, quote = FALSE, sep = "", row.names = FALSE)
}
setwd("C:/[工作目錄路徑]")
getwd()

# 2
wrta = function(xdo , ycsv, swd = getwd()){
setold = getwd()
setwd(swd)
dirseh = dir()
if( length(which(ycsv == dirseh)) > 0){
  sdc = as.character(Sys.time())
  sdc1t = gsub( ":", " ", sdc)
  sdc2t = gsub( "-", " ", sdc1t)
  sdc3t = paste(strsplit( sdc2t ,split = " ", fixed = T)[[1]],collapse="")
  s2pt = strsplit( ycsv ,split=".")fixed=T)
  ycsvhd = paste0( s2pt[[1]][1], "_", sdc3t, ".", s2pt[[1]][2])
  write.table( xdo, file= ycsvhd, quote = FALSE, sep = "", row.names = FALSE)
}else{
  write.table( xdo, file= ycsv, quote = FALSE, sep = "", row.names = FALSE)
}
setwd(setold)
}
```



```
sdop = "C:/[工作目錄路徑] "
head(iris, 3)
getwd()
wrta(iris, "haoye13.csv")
wrta(iris, "haoye13.csv")
wrta(iris, "haoye13.csv", sdop)
wrta(iris, "kan13.csv", sdop)
```

R 資料輸入 - scan

1.what - 可以指定輸入的內容，如 logical、integer、numeric、complex、character、raw 與 list。

2. nmax - 為設定輸入資料的最大次數。

3. file - 所要輸入的檔案名稱

`scan {base}`

Read Data Values

Description

Read data into a vector or list from the console or file.

Usage

```
scan(file = "", what = double(), nmax = -1, n = -1, sep = "",  
      quote = if(identical(sep, "\n")) "" else "'\"'", dec = ".,",  
      skip = 0, nlines = 0, na.strings = "NA",  
      flush = FALSE, fill = FALSE, strip.white = FALSE,  
      quiet = FALSE, blank.lines.skip = TRUE, multi.line = TRUE,  
      comment.char = "", allowEscapes = FALSE,  
      fileEncoding = "", encoding = "unknown", text, skipNul = FALSE)
```

R 資料輸入 - scan

```
sc1 = scan()
```

```
0 9
```

```
2 9
```

```
sc1
```

```
sc2 = scan(nmax = 1)
```

```
0 9 2 9
```

```
sc2
```

```
sc3 = scan(what = "character", quiet = TRUE)
```

```
Hi KanCheng
```

```
class(sc3)
```

```
NROW(sc3)
```

```
sc3
```

```
> sc1 = scan()
1: 0 9
3: 2 9
5:
Read 4 items
> sc1
[1] 0 9 2 9
>
> sc2 = scan(nmax = 1)
1: 0 9 2 9
Read 1 item
>
```

```
sc4 = scan(what = "character", quiet = TRUE)
```

```
"Hi KanCheng"
```

```
class(sc4)
```

```
NROW(sc4)
```

```
sc4
```

```
> sc4 = scan(what = "character", quiet = TRUE)
1: "Hi KanCheng"
2:
> class(sc4)
[1] "character"
> NROW(sc4)
[1] 1
> sc4
[1] "Hi KanCheng"
> |
```

```
sc5 = scan( what = list(name = "character",
+ age = "numeric",
+ pass = "logical"))
```

```
john 24 true
```

```
kan 23 false
```

```
jun 20 true
```

```
class(sc5)
```

```
summary(sc5)
```

```
sc5
```

```
> sc5 = scan( what = list(name = "character",
+ age = "numeric",
+ pass = "logical"))
1: john 24 true
2: kan 23 false
3: jun 20 true
4:
Read 3 records
> class(sc5)
[1] "list"
> summary(sc5)
  Length Class  Mode
name  3     -none- character
age   3     -none- character
pass  3     -none- character
> sc5
$name
[1] "john" "kan"  "jun"

$age
[1] "24"  "23"  "20"

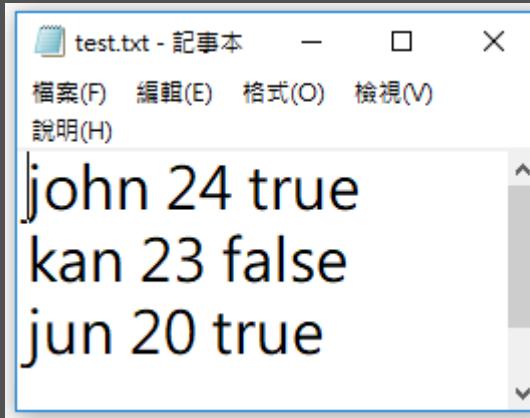
$pass
[1] "true" "false" "true"
> |
```

```
> sc3 = scan(what = "character", quiet = TRUE)
1: Hi KanCheng
3:
> class(sc3)
[1] "character"
> NROW(sc3)
[1] 2
> sc3
[1] "Hi"      "KanCheng"
>
```

R 資料輸入 - scan

```
getwd()  
setwd("C:/rws/scan")  
getwd()  
dir()
```

```
sc.data = scan(file = "test.txt",  
               what = list(name = "character",  
                           age = "numeric",  
                           pass = "logical"))  
sc.data  
sc.mat = as.data.frame(sc.data)  
class(sc.mat)  
sc.mat
```



```
> getwd()  
[1] "C:/rws"  
> setwd("C:/rws/scan")  
> getwd()  
[1] "C:/rws/scan"  
> dir()  
[1] "test.txt"  
>  
> sc.data = scan(file = "test.txt",  
+ what = list(name = "character",  
+ age = "numeric",  
+ pass = "logical"))  
Read 3 records  
> sc.data  
$name  
[1] "john" "kan" "jun"  
  
$age  
[1] "24" "23" "20"  
  
$pass  
[1] "true" "false" "true"  
  
> sc.mat = as.data.frame(sc.data)  
> class(sc.mat)  
[1] "data.frame"  
> sc.mat  
   name age  pass  
1 john 24  true  
2 kan 23 false  
3 jun 20  true  
> |
```

R 汇出執行結果紀錄 - sink

輸出範例資料

```
head(iris)
summary(iris[1])
summary(iris[2])
summary(lm(iris$Sepal.Length ~ iris$Sepal.Width))
```

```
> summary(lm(iris$Sepal.Length ~ iris$Sepal.Width))

Call:
lm(formula = iris$Sepal.Length ~ iris$Sepal.Width)

Residuals:
    Min      1Q  Median      3Q     Max 
-1.5561 -0.6333 -0.1120  0.5579  2.2226 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  6.5262    0.4789   13.63 <2e-16 ***
iris$Sepal.Width -0.2234    0.1551   -1.44    0.152  
---
Signif. codes:  0 '****' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

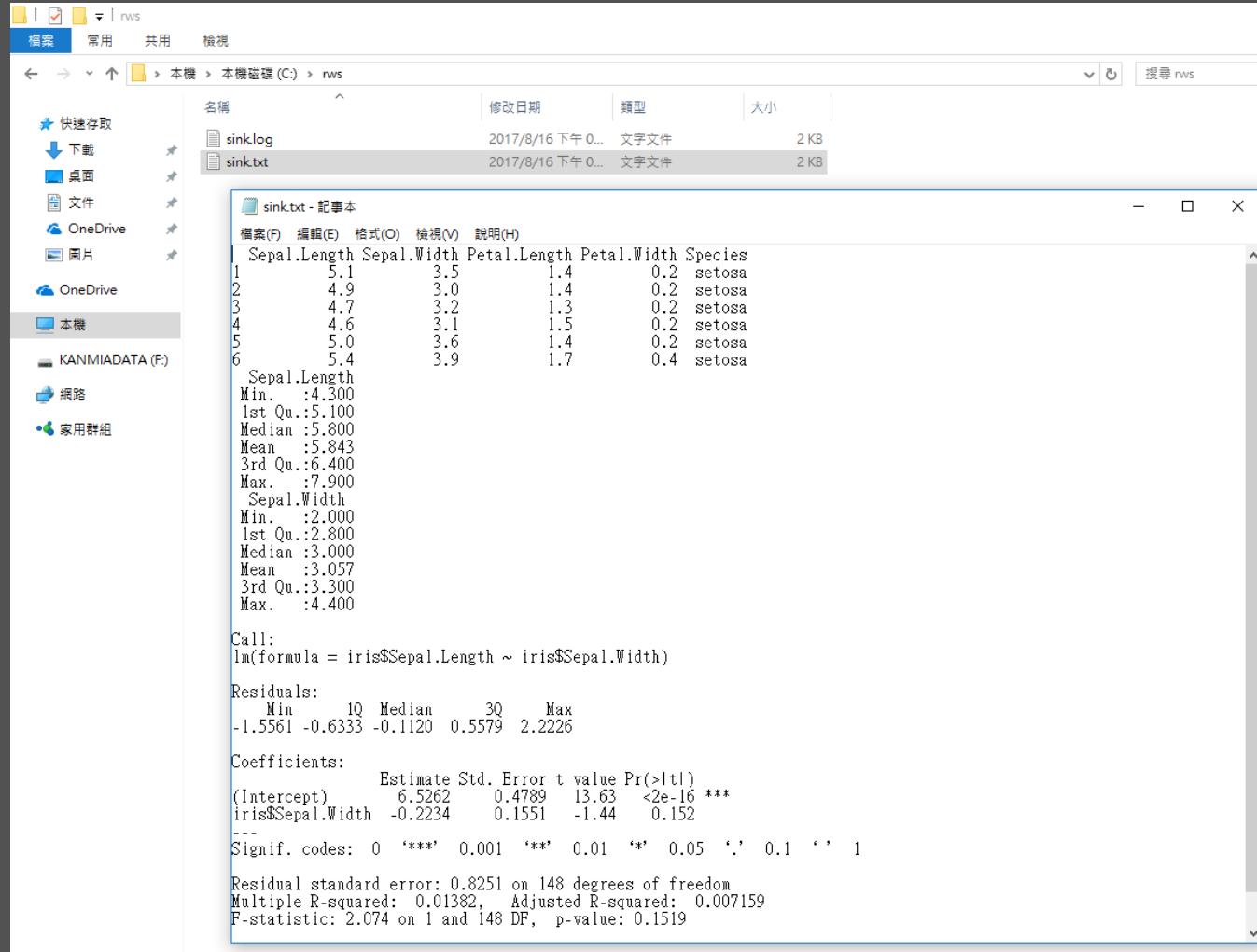
Residual standard error: 0.8251 on 148 degrees of freedom
Multiple R-squared:  0.01382, Adjusted R-squared:  0.007159 
F-statistic: 2.074 on 1 and 148 DF,  p-value: 0.1519

> |
```

```
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1       3.5        1.4       0.2  setosa
2          4.9       3.0        1.4       0.2  setosa
3          4.7       3.2        1.3       0.2  setosa
4          4.6       3.1        1.5       0.2  setosa
5          5.0       3.6        1.4       0.2  setosa
6          5.4       3.9        1.7       0.4  setosa
> summary(iris[1])
  Sepal.Length
  Min. :4.300
  1st Qu.:5.100
  Median :5.800
  Mean   :5.843
  3rd Qu.:6.400
  Max.   :7.900
> summary(iris[2])
  Sepal.Width
  Min. :2.000
  1st Qu.:2.800
  Median :3.000
  Mean   :3.057
  3rd Qu.:3.300
  Max.   :4.400
```

R 汇出執行結果紀錄 - sink

在該函數第一行用 sink() 輸入字串路徑與所要的檔名與副檔名，結束後用 sink() 汇出。



在此有嘗試過 ".log" 和 ".txt" 的版本。

```
sink("c:/rws/sink.log")
head(iris)
summary(iris[1])
summary(iris[2])
summary(lm(iris$Sepal.Length ~ iris$Sepal.Width))
sink()
```

```
sink("c:/rws/sink.txt")
head(iris)
summary(iris[1])
summary(iris[2])
summary(lm(iris$Sepal.Length ~ iris$Sepal.Width))
sink()
```

R 計算執行時間

1. mat 為自己寫的自訂函數，塞一個連續數值向量就會加 1。

2. data 為 1 - 100 的連續數列資料。

可以使用 system.time({ }) 來進行測試。

```
> data = 1:100
> mat = function(obj){
+   out = NULL
+   for( addnb in 1:length(obj)){
+     out[addnb] = obj[addnb] + 1
+   }
+   out
+ }
>
> system.time({
+   mat(data)
+ })
    user  system elapsed
      0        0       0
> |
```

```
data = 1:100
mat = function(obj){
  out = NULL
  for( addnb in 1:length(obj)){
    out[addnb] = obj[addnb] + 1
  }
  out
}
```

```
system.time({
  mat(data)
})
```

R 記憶體資訊

當中的單位標準為 MB，而使用 32 位元版本最多只會有 4095。

```
> # 當前使用的記憶體  
> # current memory in use  
> memory.size(max = FALSE)  
[1] 182.88  
>  
> # 從OS獲得的最大記憶體  
> # maximum memory obtained from the OS  
> memory.size(max = TRUE)  
[1] 186.69  
>  
> # 當前記憶體限制  
> # current memory limit  
> memory.size(max = NA)  
[1] 65499  
>  
> # 當前記憶體限制  
> # current memory limit  
> memory.limit(size = NA)  
[1] 65499  
>  
> #增加記憶體限制  
> # increase memory limit  
> memory.limit(size = size)  
Error in memory.limit(size = size) : object 'size' not found  
> |
```

```
# 當前使用的記憶體  
# current memory in use  
memory.size(max = FALSE)
```

```
# 從OS獲得的最大記憶體  
# maximum memory obtained from the OS  
memory.size(max = TRUE)
```

```
# 當前記憶體限制  
# current memory limit  
memory.size(max = NA)
```

```
# 當前記憶體限制  
# current memory limit  
memory.limit(size = NA)
```

```
#增加記憶體限制  
# increase memory limit  
memory.limit(size = size)
```

R 記憶體資訊

R 文件

```
memory.size {utils}
```

Report on Memory Allocation

Description

`memory.size` reports the current or maximum memory allocation of the `malloc` function used in this version of R.

`memory.limit` reports or increases the limit in force on the total allocation.

Usage

```
memory.size(max = FALSE)
```

```
memory.limit(size = NA)
```

Arguments

`max` logical. If TRUE the maximum amount of memory obtained from the OS is reported, if FALSE the amount currently in use, if NA the memory limit.

`size` numeric. If NA report the memory limit, otherwise request a new limit, in Mb. Only values of up to 4095 are allowed on 32-bit R builds, but see 'Details'.

R 記憶體資訊

R 物件所佔的記憶體資訊

在此建立一個包含了 1 - 1000000 連續資料，並叫出
iris 資料集作為測試。

x = 1:1000000
NROW(x)
head(x)
tail(x)
object.size(x)
head(iris)
object.size(iris)

```
> x = 1:1000000
> NROW(x)
[1] 1000000
> head(x)
[1] 1 2 3 4 5 6
> tail(x)
[1] 999995 999996 999997 999998 999999 1000000
> object.size(x)
4000040 bytes
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5          1.4         0.2  setosa
2          4.9         3.0          1.4         0.2  setosa
3          4.7         3.2          1.3         0.2  setosa
4          4.6         3.1          1.5         0.2  setosa
5          5.0         3.6          1.4         0.2  setosa
6          5.4         3.9          1.7         0.4  setosa
> object.size(iris)
7088 bytes
> |
```

R 查詢系統與版本資訊

```
# R 版本資訊  
R.Version()
```

```
version
```

```
# R 系統環境資訊  
Sys.getenv("R_ARCH")
```

```
# 系統地區資訊  
Sys.getlocale()
```

```
# 系統資訊  
Sys.info()
```

```
# 顯示 R 版本、系統與套件等資訊  
sessionInfo()
```

```
# 系統時間包含日月時分秒  
Sys.time()
```

```
# 估算 R 物件所占記憶體大小  
x = 1:1000000  
object.size(x)
```

```
# 系統年月日  
Sys.Date()
```

```
# 系統時區  
Sys.timezone()
```

```
> # R 系統環境資訊  
> Sys.getenv("R_ARCH")  
[1] "/x64"  
>  
> Sys.getenv()  
ALLUSERSPROFILE  
APPDATA  
CLASSPATH  
CommonProgramFiles  
CommonProgramFiles(x86)  
CommonProgramW6432  
COMPUTERNAME  
ComSpec  
DISPLAY  
FP_NO_HOST_CHECK  
FPS_BROWSER_APP_PROFILE_STRING  
FPS_BROWSER_USER_PROFILE_STRING  
GFORTRAN_STDERR_UNIT  
GFORTRAN_STDOUT_UNIT  
GIT_LFS_PATH
```

```
> # R 版本資訊  
> R.Version()  
$platform  
[1] "x86_64-w64-mingw32"  
  
$arch  
[1] "x86_64" > version  
  
$os  
[1] "mingw32" platform x86_64-w64-mingw32  
  
$system  
[1] "x86_64, mingw32" arch x86_64  
os mingw32  
system x86_64, mingw32  
status  
major 3  
minor 4.1  
year 2017  
month 06  
day 30  
svn_rev 72865  
language R  
version.string R version 3.4.1 (2017-06-30)  
nickname Single Candle  
>
```

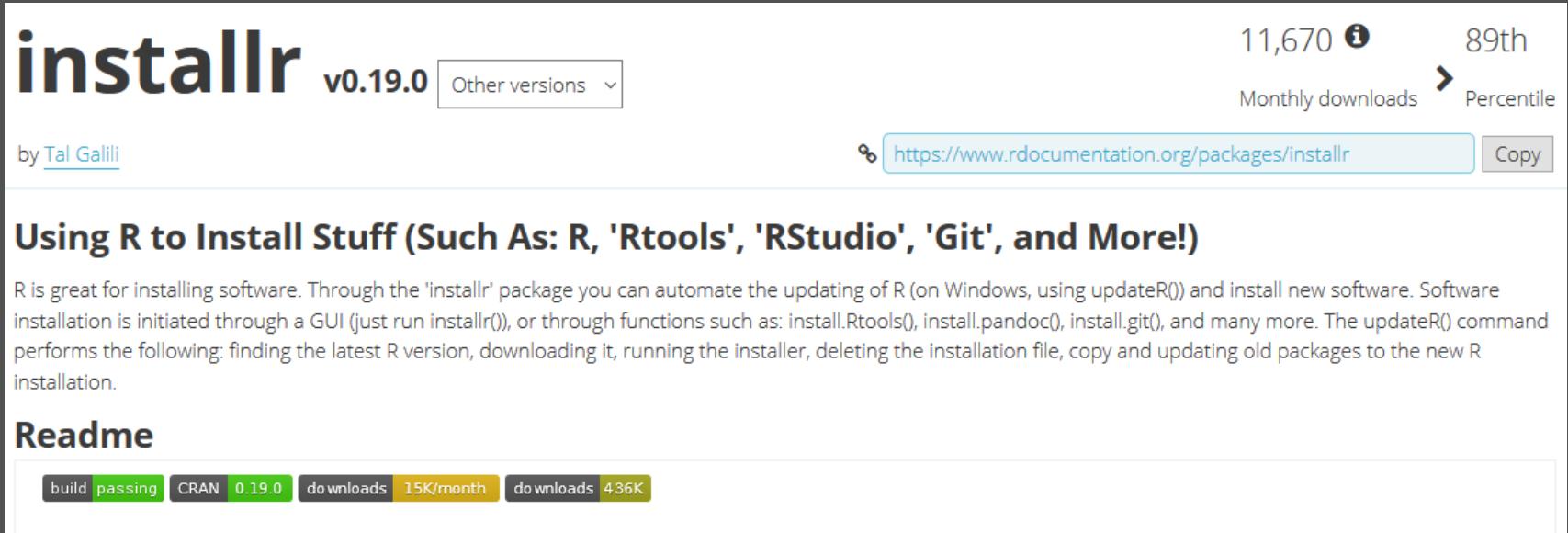
```
C:\ProgramData  
C:\Users\USER\AppData\Roaming  
. ;C:\Program Files (x86)\QuickTime\QTSystem\QTJava.zip  
C:\Program Files\Common Files  
C:\Program Files (x86)\Common Files  
C:\Program Files\Common Files  
HC-B15MC17  
C:\WINDOWS\system32\cmd.exe  
:0  
NO  
Internet Explorer  
Default  
-1  
-1  
C:\Program Files\Git LFS
```

Packages - installr

```
# install.packages("installr")
```

```
library("installr")
```

```
installr()
```



The screenshot shows the CRAN package page for 'installr' version 0.19.0. The page title is 'installr v0.19.0'. It features a summary section with a 'Using R to Install Stuff (Such As: R, 'Rtools', 'RStudio', 'Git', and More!)' heading and a detailed description of the package's purpose. Below this is a 'Readme' section. At the bottom, there are build status badges for 'build passing' and 'CRAN 0.19.0', along with download statistics: 'downloads 15K/month' and 'downloads 436K'. The top right corner displays the package's popularity metrics: 11,670 monthly downloads (89th percentile).

installr v0.19.0

by Tal Galili

11,670 1 89th

Monthly downloads > Percentile

https://www.rdocumentation.org/packages/installr

Copy

Using R to Install Stuff (Such As: R, 'Rtools', 'RStudio', 'Git', and More!)

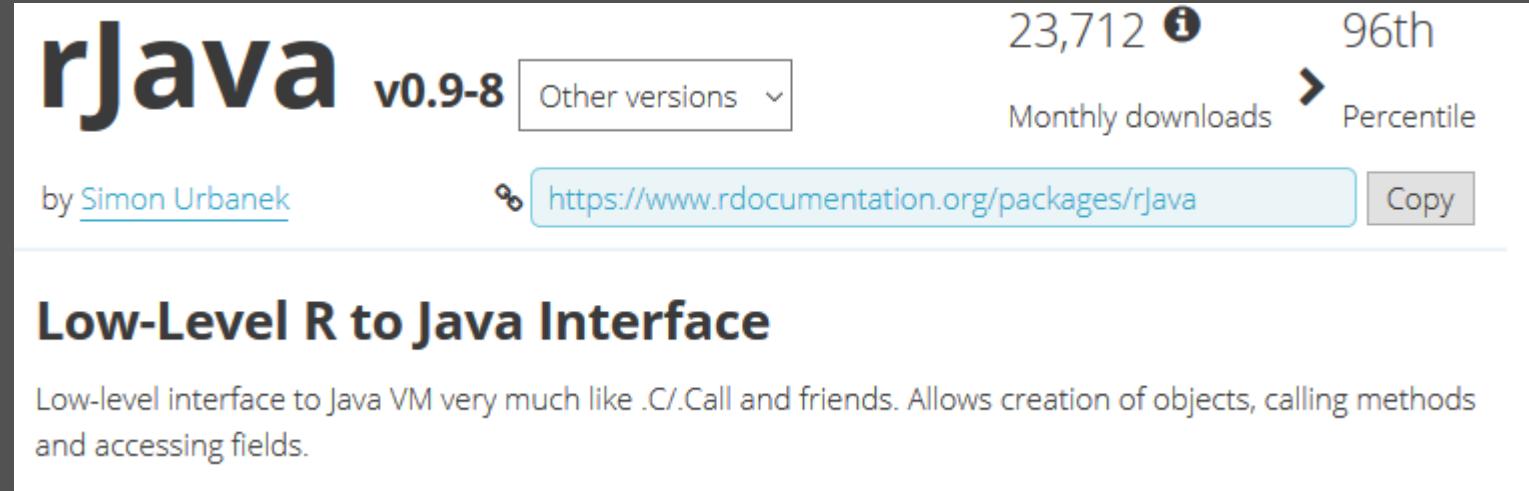
R is great for installing software. Through the 'installr' package you can automate the updating of R (on Windows, using updateR()) and install new software. Software installation is initiated through a GUI (just run installr()), or through functions such as: install.Rtools(), install.pandoc(), install.git(), and many more. The updateR() command performs the following: finding the latest R version, downloading it, running the installer, deleting the installation file, copy and updating old packages to the new R installation.

Readme

build passing CRAN 0.19.0 downloads 15K/month downloads 436K

Packages - rJava

```
# install.packages("rJava")  
  
library("rJava")
```



The screenshot shows the CRAN package page for rJava. At the top, the package name "rJava" is displayed in large, bold, dark blue letters, followed by the version "v0.9-8". To the right of the version are two buttons: "Other versions" and a link to "Monthly downloads" which shows 23,712 at the 96th percentile. Below this, the author "Simon Urbanek" is mentioned. A prominent button provides a direct link to the documentation: "https://www.rdocumentation.org/packages/rJava". A "Copy" button is also present. The main content area features a section titled "Low-Level R to Java Interface" with a brief description: "Low-level interface to Java VM very much like .C/.Call and friends. Allows creation of objects, calling methods and accessing fields."

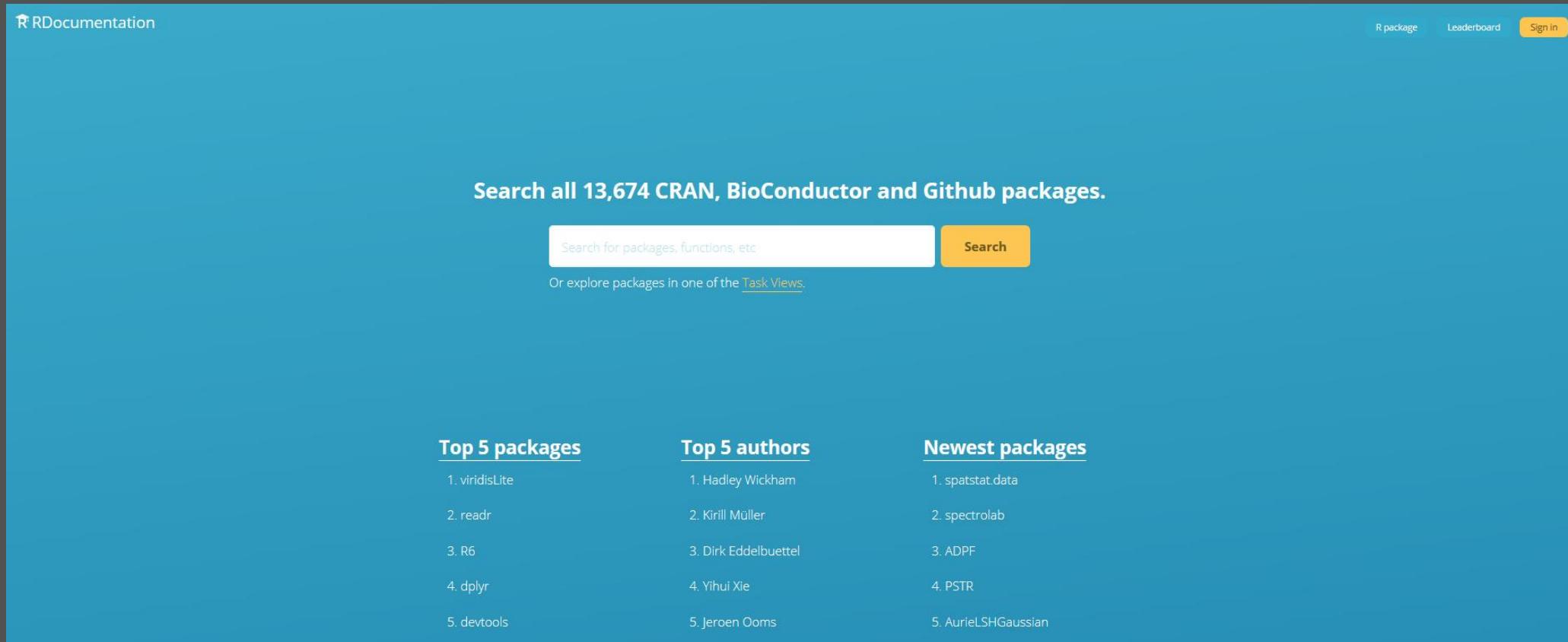
Packages - ROracle

```
# install.packages("ROracle")  
library("ROracle ")
```

The screenshot shows the RDocumentation.org page for the ROracle package. At the top, the package name "ROracle" is displayed in large, bold letters, followed by "v1.3-1". To the right, there are statistics: "604" with an info icon, "Other versions" dropdown, "Monthly downloads" (with a right arrow), and a "Copy" button for the URL. Below this, the author is listed as "by Rajendra S Pingte" and the URL "https://www.rdocumentation.org/packages/ROracle" is provided with a copy icon. The main title of the package is "OCI Based Oracle Database Interface for R". A descriptive text below the title states: "Oracle Database interface (DBI) driver for R. This is a DBI-compliant Oracle driver based on the OCI."

RDocumentation

<https://www.rdocumentation.org/>



The screenshot shows the RDocumentation homepage with a teal header and footer and a white central content area. The header includes the R logo and the text "RDocumentation". The footer has links for "R package", "Leaderboard", and "Sign in". The main content features a search bar with placeholder text "Search for packages, functions, etc." and a yellow "Search" button. Below the search bar is a link to "Task Views". The page displays three sections: "Top 5 packages", "Top 5 authors", and "Newest packages", each with a numbered list of items.

Top 5 packages

1. viridisLite
2. readr
3. R6
4. dplyr
5. devtools

Top 5 authors

1. Hadley Wickham
2. Kirill Müller
3. Dirk Eddelbuettel
4. Yihui Xie
5. Jeroen Ooms

Newest packages

1. spatstat.data
2. spectrolab
3. ADPF
4. PSTR
5. AurielLSH Gaussian

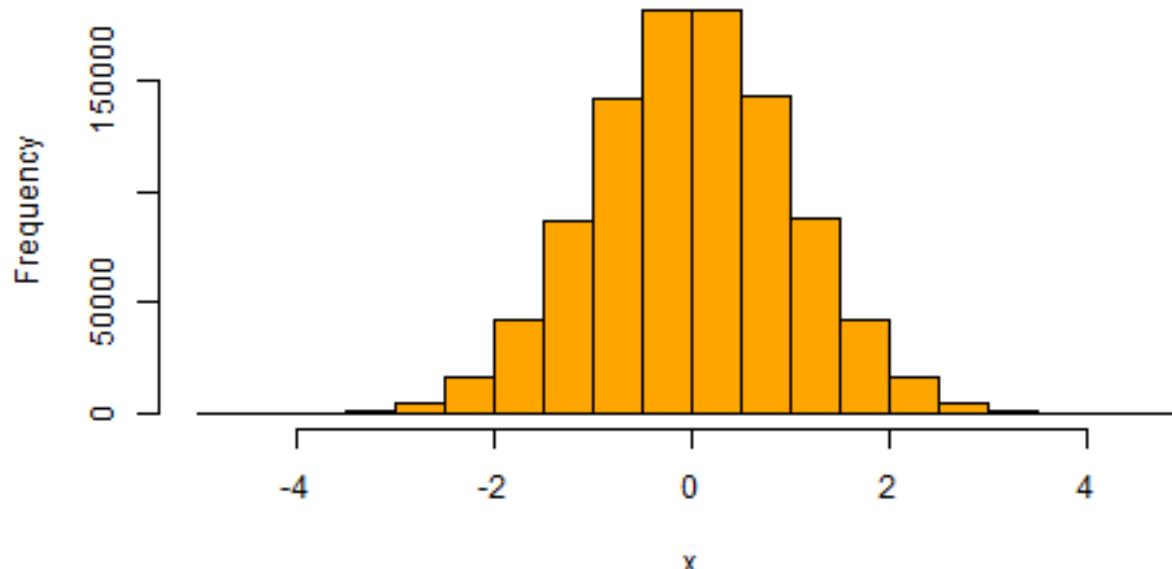
R & PHP

本機 > 本機磁碟 (C:) > devServer >			
	名稱	修改日期	類型
上	下	上	下
	output	2017/8/14 上午 0...	檔案資料夾
	R	2017/8/14 上午 0...	PHP 來源檔案 1 KB
	script	2017/8/14 上午 0...	R 來源檔案 1 KB

① 127.0.0.1/R.php?n=950000

input n value : 提交

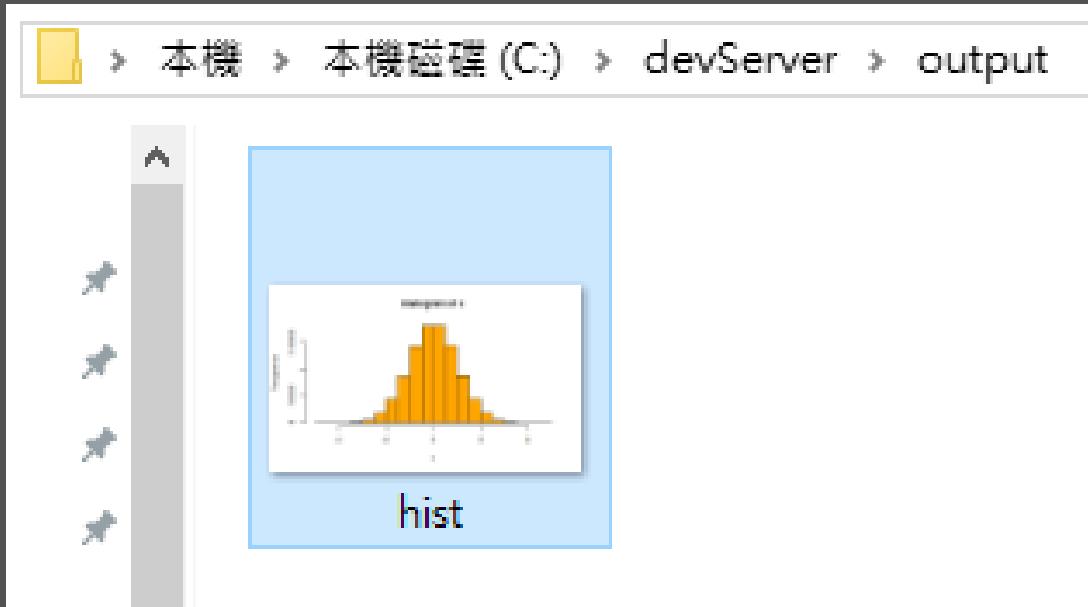
Histogram of x



設定好 Apache 與 php 後，在此建立名為 R.php 與 script.R 的兩個檔案與一個名為 output 目錄，用於存放 script.R 輸出的結果。

R & PHP

可以觀察到 PHP 將命令傳給 R 時，會將所輸出的繪圖暫存於 output。



R & Julia

RCall 是連上 R，而
RDatasets 則是連上 R
原本的資料集。
要注意的地方，在 Julia
中使用 R 時，要按 [\\$]
鍵切換至 R 的模式。

而要回到原本的樣子則
是直接用 [Backspace]
鍵即可。
至於在 R 的模式當中要
用 Julia 的物件則是要
在名稱前面加上 "\$"。

```
julia> Pkg.add("RCall")
INFO: Cloning cache of CategoricalArrays from https://github.com/JuliaData/CategoricalArrays.jl.git
INFO: Cloning cache of NullableArrays from https://github.com/JuliaStats/NullableArrays.jl.git
INFO: Cloning cache of RCall from https://github.com/JuliaInterop/RCall.jl.git
INFO: Cloning cache of WinRes from https://github.com/simonbyrne/WinRes.jl.git
INFO: Installing CategoricalArrays v0.1.0
INFO: Installing NullableArrays v0.0.10
INFO: Installing RCall v0.6.4
INFO: Installing WinRes v0.2.0
INFO: Package database updated

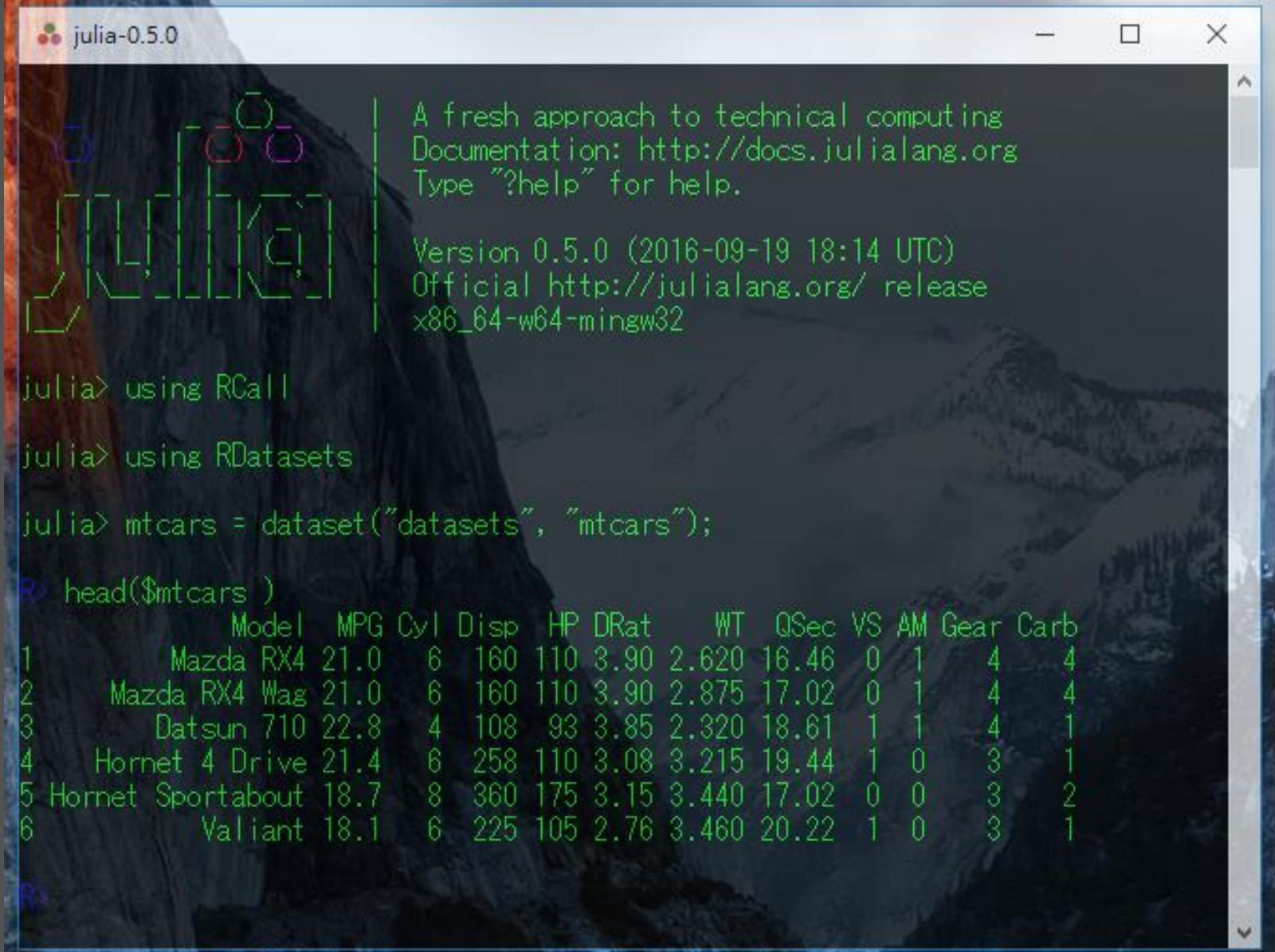
julia> using RCall
INFO: Precompiling module RCall...
INFO: Using R installation at D:\RVR-3.3.2
```

```
julia> Pkg.add("RDatasets")
INFO: Cloning cache of RData from https://github.com/JuliaStats/RData.jl.git
INFO: Cloning cache of RDatasets from https://github.com/johnmyleswhite/RDatasets.jl.git
INFO: Installing RData v0.0.4
INFO: Installing RDatasets v0.2.0
INFO: Package database updated

julia> using RDatasets

julia>
```

R & Julia



A fresh approach to technical computing
Documentation: <http://docs.julialang.org>
Type "?help" for help.

Version 0.5.0 (2016-09-19 18:14 UTC)
Official <http://julialang.org/> release
x86_64-w64-mingw32

```
julia> using RCall
julia> using RDatasets
julia> mtcars = dataset("datasets", "mtcars");
julia> head($mtcars )
      Model   MPG   Cyl Disp HP DRat   WT QSec VS AM Gear Carb
1 Mazda RX4 21.0     6 160 110 3.90 2.620 16.46  0  1    4    4
2 Mazda RX4 Wag 21.0     6 160 110 3.90 2.875 17.02  0  1    4    4
3 Datsun 710 22.8     4 108  93 3.85 2.320 18.61  1  1    4    1
4 Hornet 4 Drive 21.4     6 258 110 3.08 3.215 19.44  1  0    3    1
5 Hornet Sportabout 18.7     8 360 175 3.15 3.440 17.02  0  0    3    2
6 Valiant 18.1     6 225 105 2.76 3.460 20.22  1  0    3    1
```