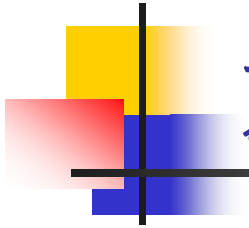
Using R – R Commander

MING-CHANG LEE

Department of Information Management
Yu Da College of Business

December 6, 2006

Email: alan9956@ydu.edu.tw
Web : http://web.ydu.edu.tw/~alan9956/



大綱

1. R Commander package 簡介
2. 下載/啓動 R Commander
3. 資料輸入(1. 直接輸入 2. 匯入資料)
4. 資料摘要
5. 圖形
6. 機率分配
7. 迴歸分析

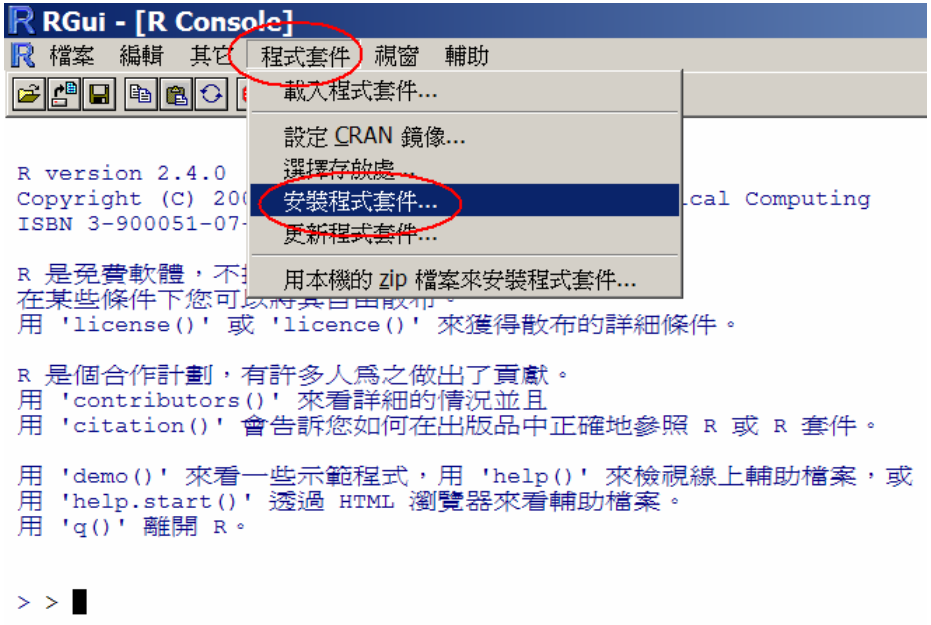
1. R Commander package簡介

- R Commander 套件是由 Fox(2006) 以 R 爲基礎所發展出來的統計分析套件 (package) 。
- 該套件包括功能列、工具列、編輯視窗、輸出視窗與訊息視窗等圖形使用介面 (GUI)，使用 R Commander 的圖形介面可方便資料統計分析。

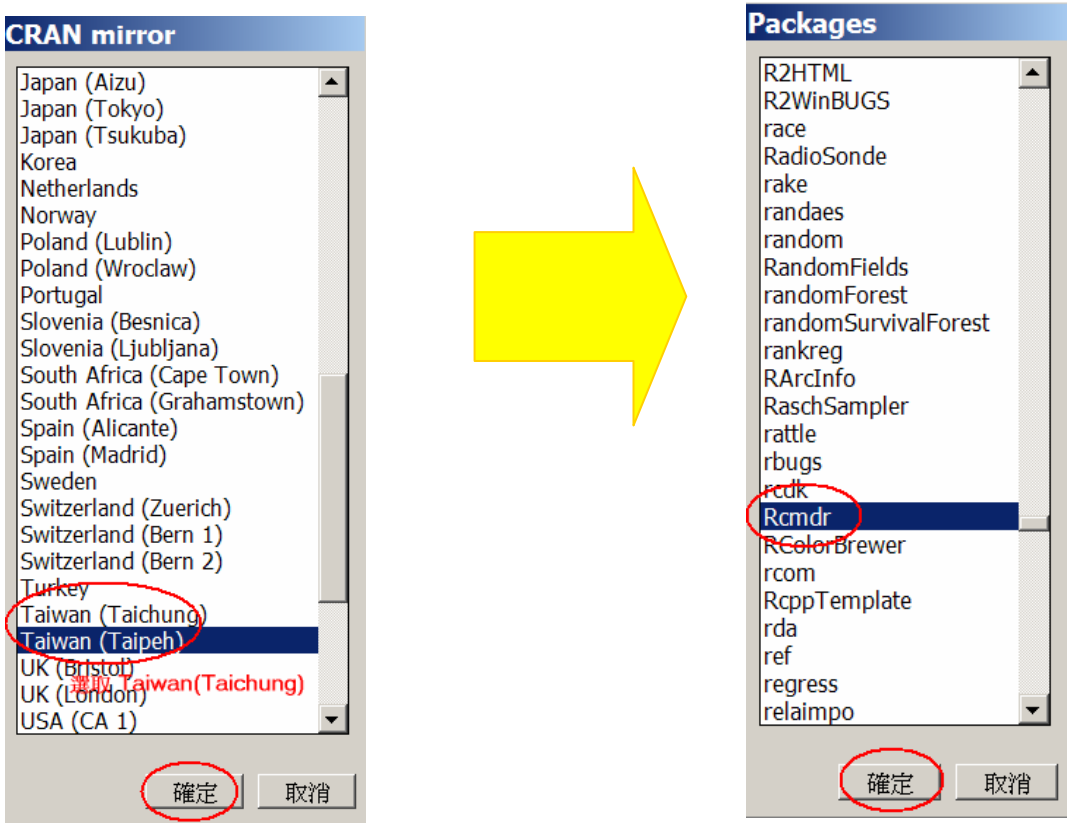
Reference:
Fox, J. (2006) The R Commander: A Basic-Statistics GUI for R,
Available from: http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/

2. 下載/啓動 R Commander

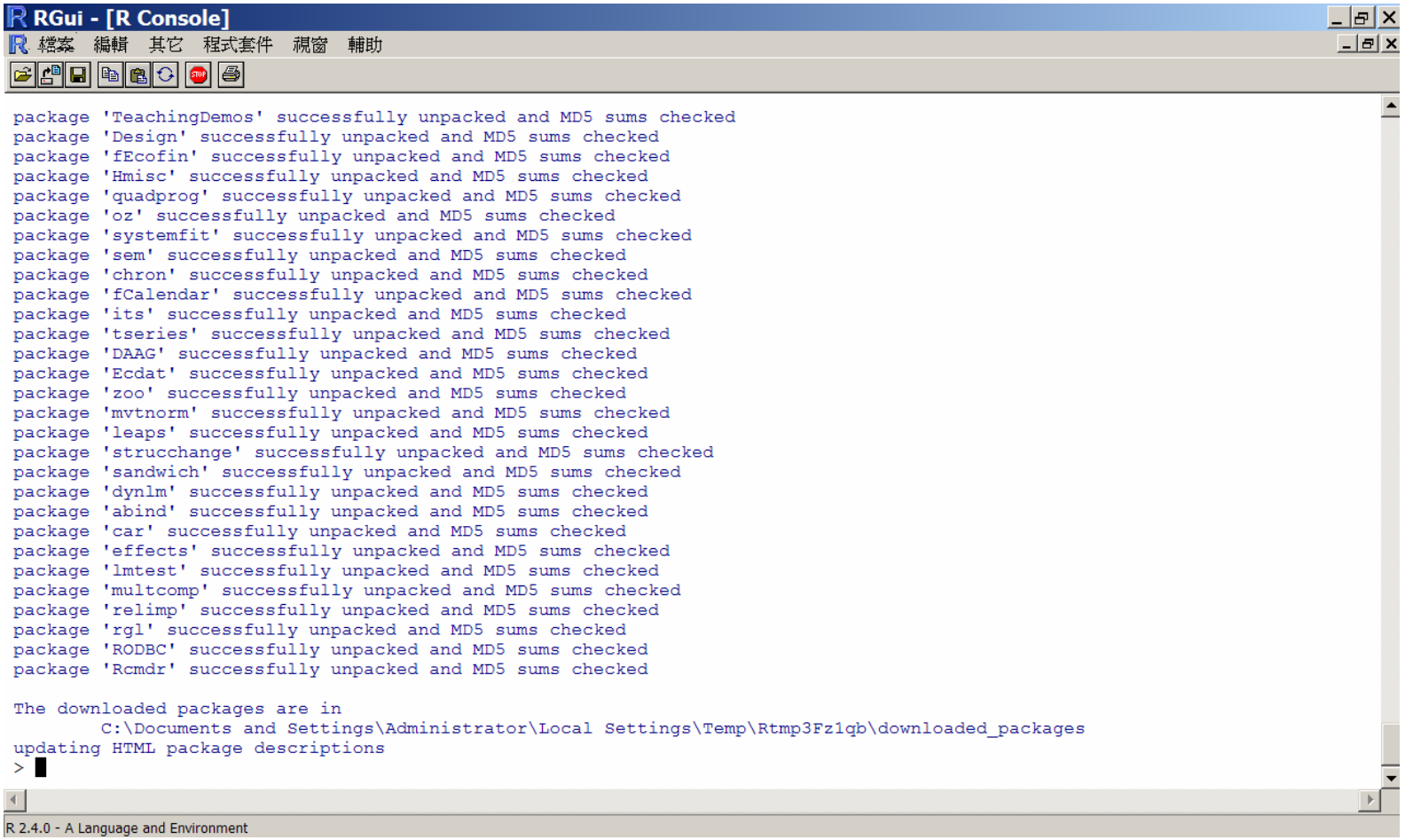
- 下載 R Commander 套件
  - 確定已連線網路
  - 啓動 R → 程式套件\安裝程式套件...\



- 選取 Taiwan (Taichung) → 確定 → 選取 Rcmdr → 確定



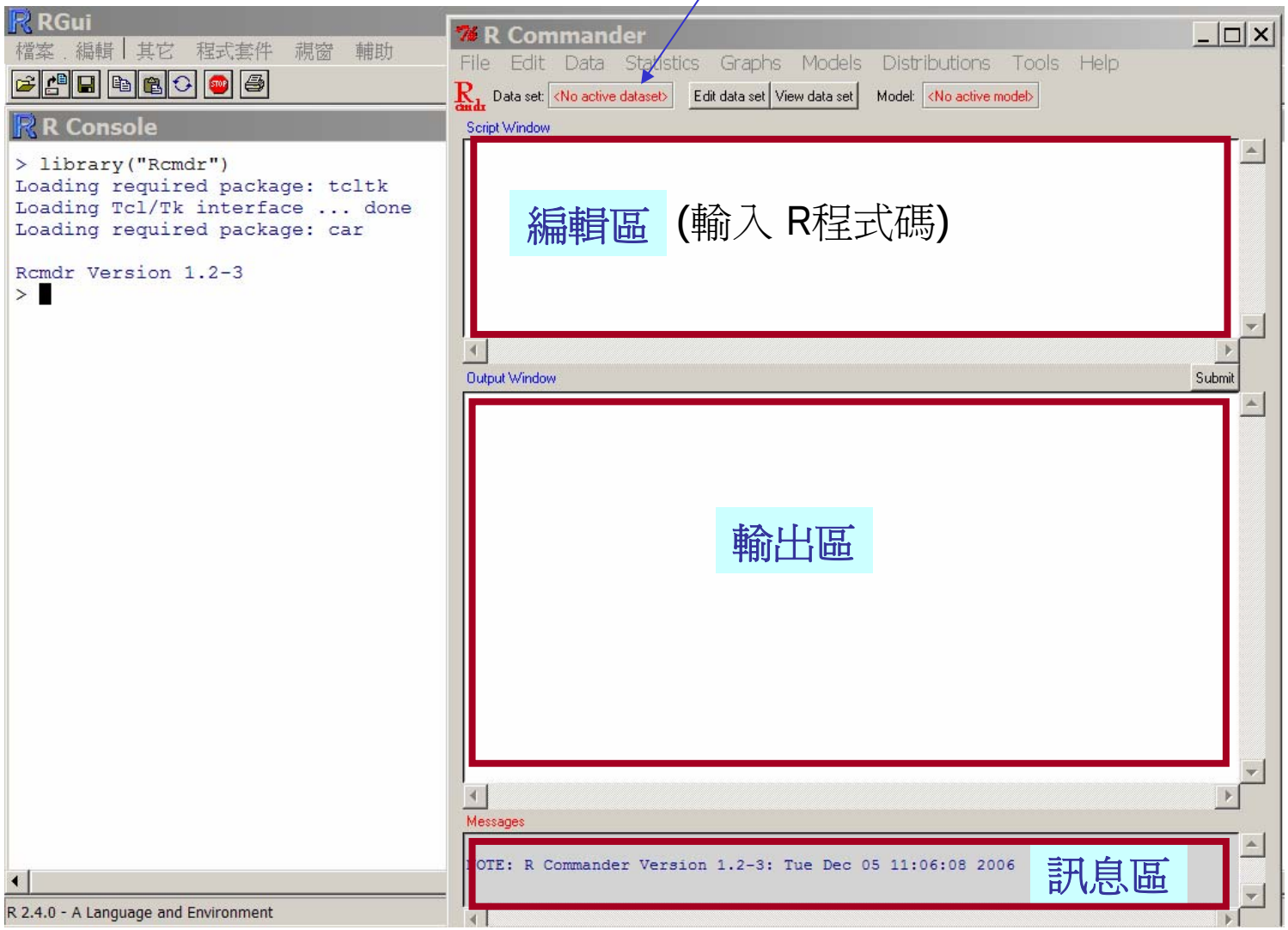
完成後畫面


RGui - [R Console]
檔案 編輯 其它 程式套件 視窗 輔助
package 'TeachingDemos' successfully unpacked and MD5 sums checked
package 'Design' successfully unpacked and MD5 sums checked
package 'fEcofin' successfully unpacked and MD5 sums checked
package 'Hmisc' successfully unpacked and MD5 sums checked
package 'quadprog' successfully unpacked and MD5 sums checked
package 'oz' successfully unpacked and MD5 sums checked
package 'systemfit' successfully unpacked and MD5 sums checked
package 'sem' successfully unpacked and MD5 sums checked
package 'chron' successfully unpacked and MD5 sums checked
package 'fCalendar' successfully unpacked and MD5 sums checked
package 'its' successfully unpacked and MD5 sums checked
package 'tseries' successfully unpacked and MD5 sums checked
package 'DAAG' successfully unpacked and MD5 sums checked
package 'Ecdat' successfully unpacked and MD5 sums checked
package 'zoo' successfully unpacked and MD5 sums checked
package 'mvtnorm' successfully unpacked and MD5 sums checked
package 'leaps' successfully unpacked and MD5 sums checked
package 'strucchange' successfully unpacked and MD5 sums checked
package 'sandwich' successfully unpacked and MD5 sums checked
package 'dynlm' successfully unpacked and MD5 sums checked
package 'abind' successfully unpacked and MD5 sums checked
package 'car' successfully unpacked and MD5 sums checked
package 'effects' successfully unpacked and MD5 sums checked
package 'lmtest' successfully unpacked and MD5 sums checked
package 'multcomp' successfully unpacked and MD5 sums checked
package 'relimp' successfully unpacked and MD5 sums checked
package 'rgl' successfully unpacked and MD5 sums checked
package 'RODBC' successfully unpacked and MD5 sums checked
package 'Rcmdr' successfully unpacked and MD5 sums checked
The downloaded packages are in
C:\Documents and Settings\Administrator\Local Settings\Temp\Rtmp3Fz1qb\downloaded_packages
updating HTML package descriptions
>
R 2.4.0 - A Language and Environment

啓動 R Commander

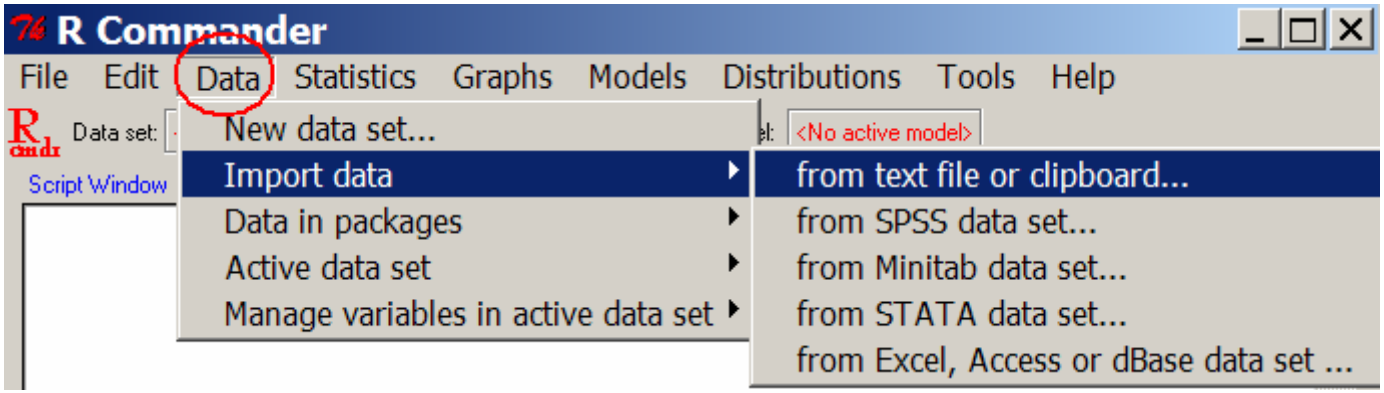
> library(“Rcmdr”)

按此鈕可切換不同資料集

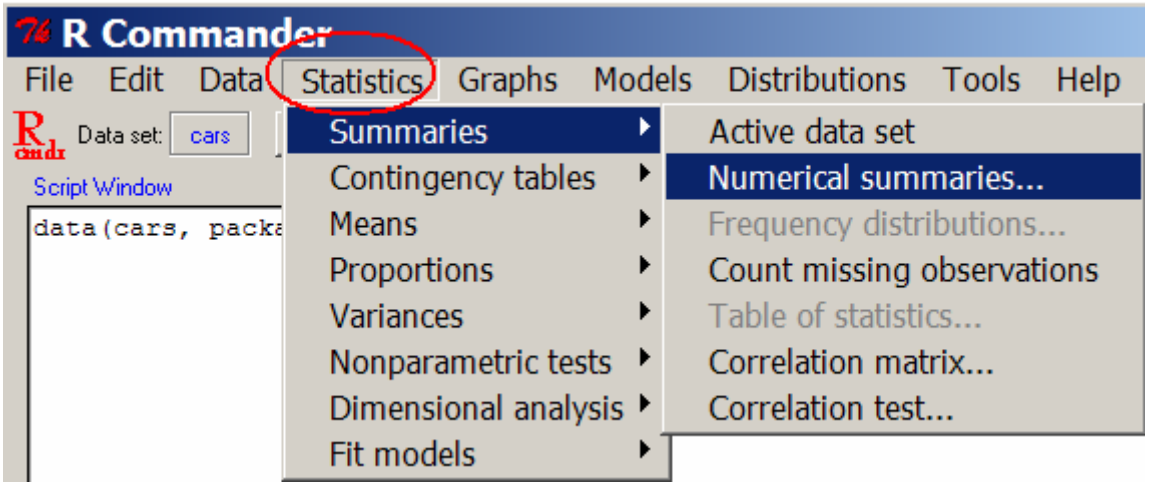


瀏覽功能(1/3)

Data:輸入/編輯資料

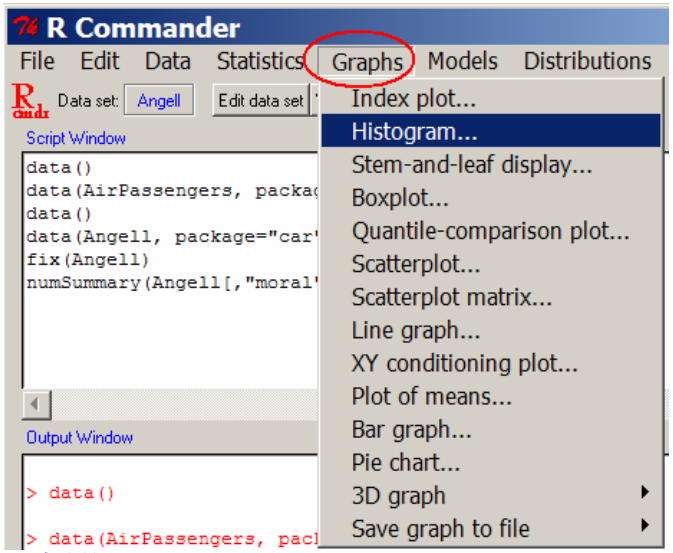


Statistics: 統計檢定分析

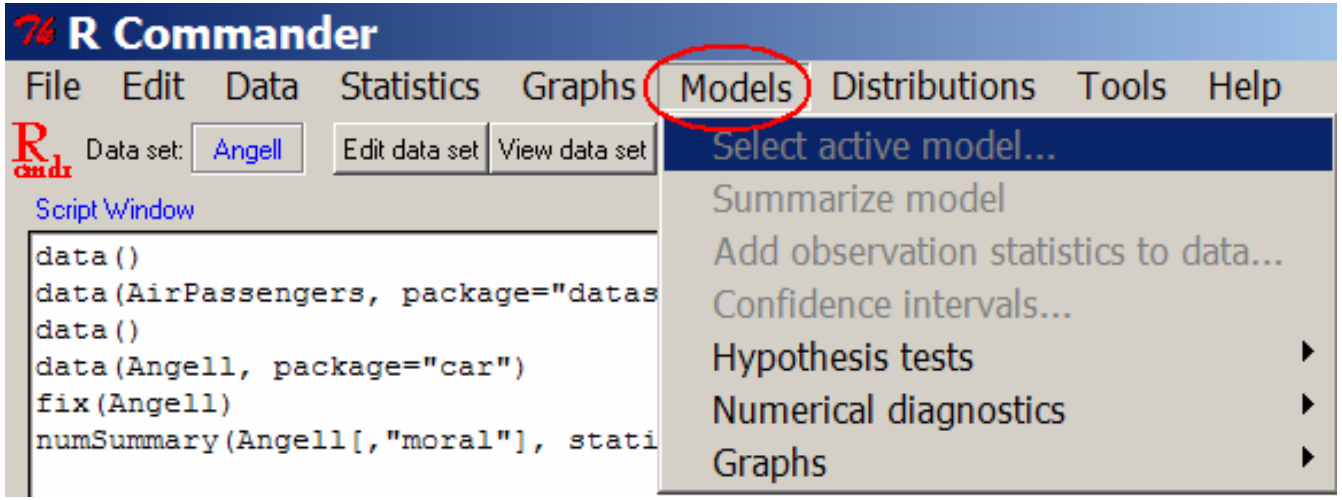


瀏覽功能(2/3)

Graphs: 畫圖

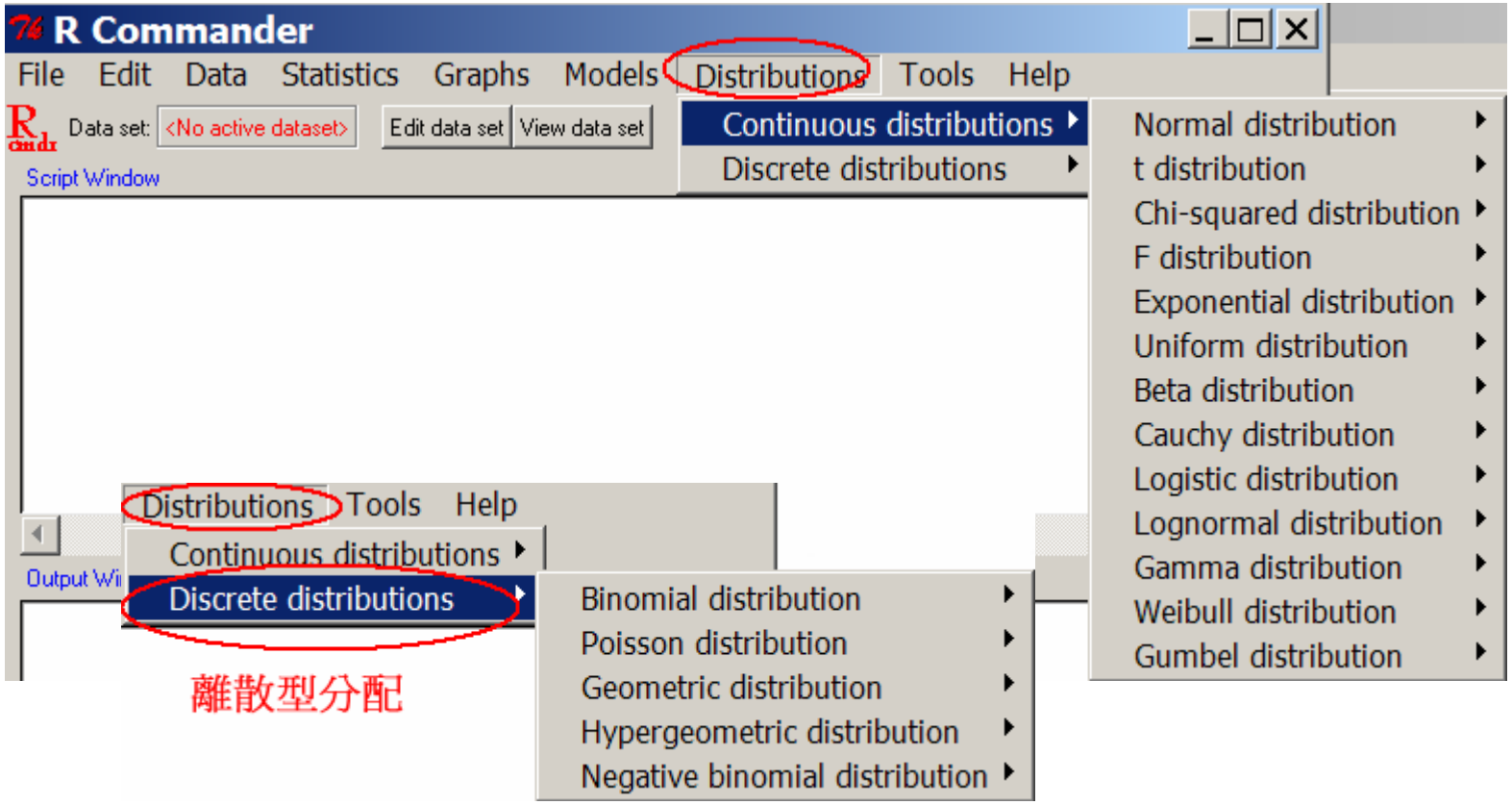


Models: 建立統計模型



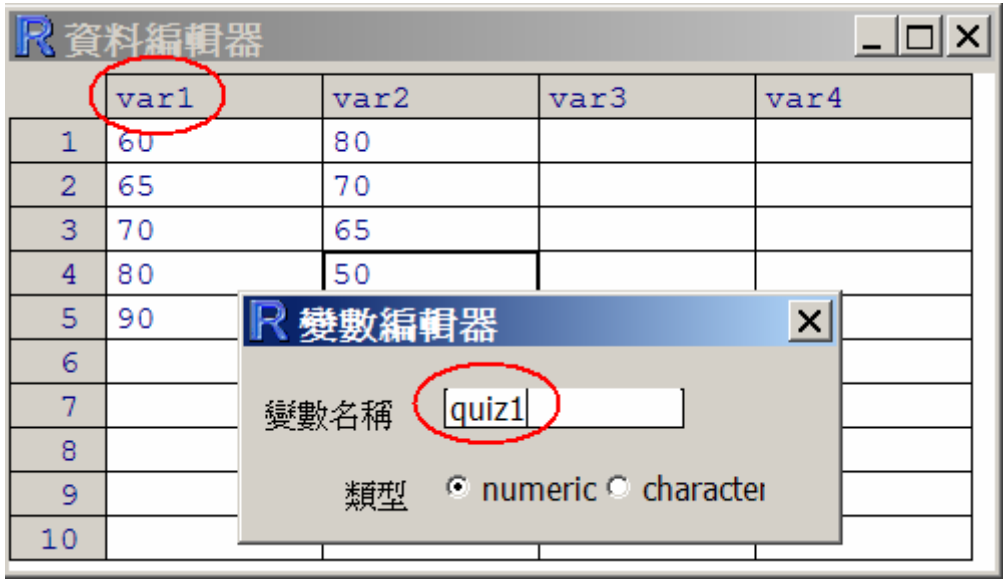
瀏覽功能(3/3)

Distributions: 產生連續型/離散型機率分配 樣本



3.資料輸入(方法1. 直接輸入)

- Data –> New data set…
- 按 var1 可直接更改變數名稱,變數名稱不可有空白
- 適用於少量資料



資料輸入(方法2. 匯入檔案)

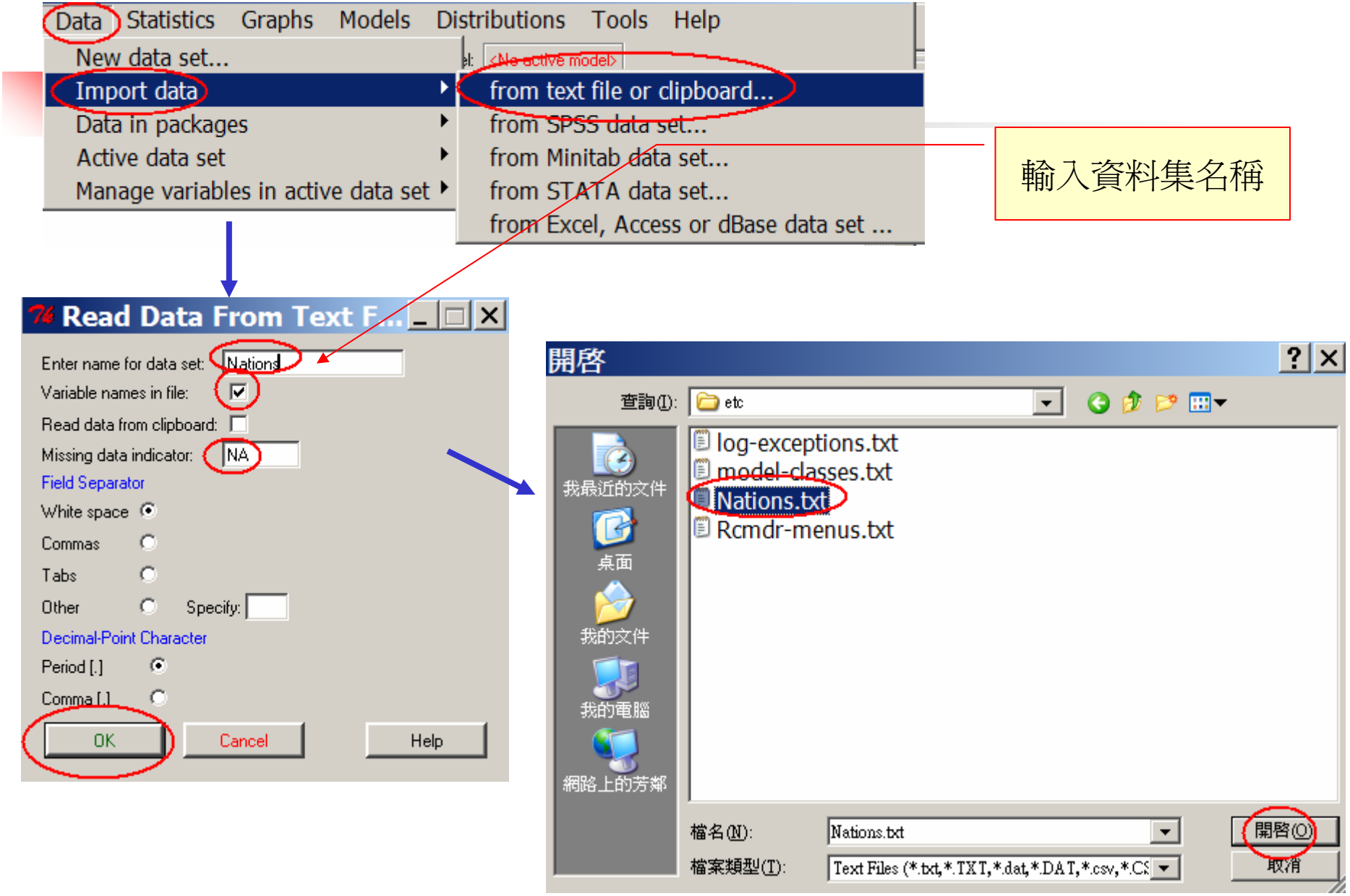
- 讀入文字檔:

C:\Program Files\R\R-2.4.0\library\Rcmdr\etc\Nations.txt
文字檔 第1列爲 變數名稱,有5個變數,含遺漏值 NA

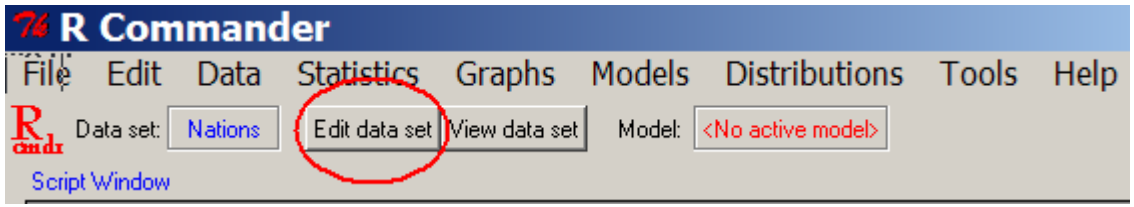
1. 數值-TFR (the total fertility rate, expressed as number of children per woman)
2. 數值- contraception (the rate of contraceptive use among married women, in percent)
3. 數值- infant.mortality (the infant-mortality rate per 1000 live births)
4. 數值- GDP (gross domestic product per capita, in U.S. dollars)
5. 字串- region.

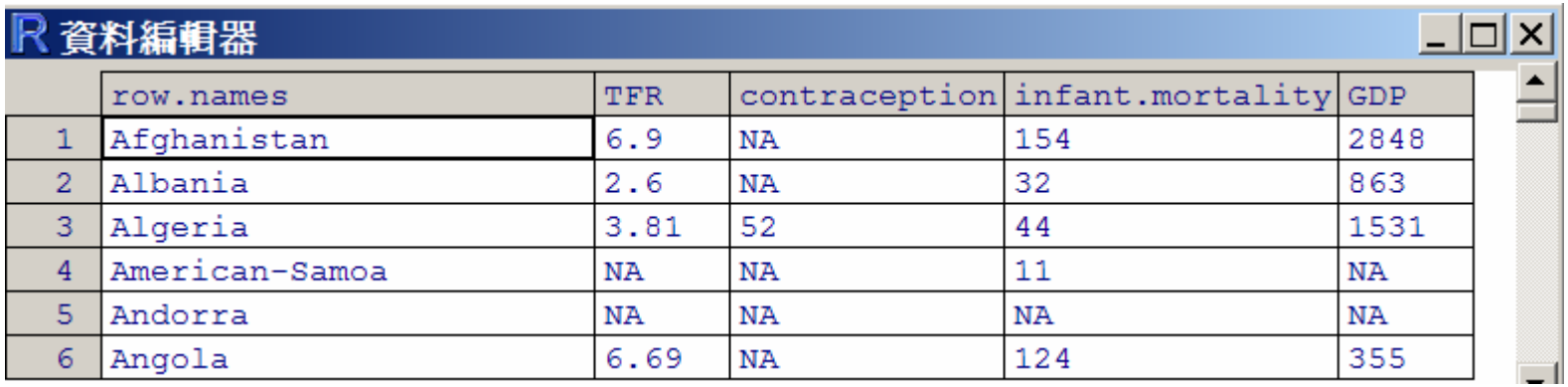
列名稱:國家名稱

	TFR	contraception	infant.mortality	GDP	region
Afghanistan	6.90	NA	154	2848	Asia
Albania	2.60	NA	32	863	Europe
Algeria	3.81	52	44	1531	Africa
American-Samoa	NA	NA	11	NA	Oceania
Andorra	NA	NA	NA	NA	Europe
Angola	6.69	NA	124	355	Africa


Data
Statistics
Graphs
Models
Distributions
Tools
Help
New data set...
Import data
Data in packages
Active data set
Manage variables in active data set
from text file or clipboard...
from SPSS data set...
from Minitab data set...
from STATA data set...
from Excel, Access or dBase data set ...
輸入資料集名稱
Read Data From Text F...
Enter name for data set:
Nations
Variable names in file:
Read data from clipboard:
Missing data indicator:
NA
Field Separator
White space
Commas
Tabs
Other
Specify:
Decimal-Point Character
Period [.]
Comma [,]
OK
Cancel
Help
開啓
查詢(I):
etc
log-exceptions.txt
model-classes.txt
Nations.txt
Rcmdr-menus.txt
我最近的文件
桌面
我的文件
我的電腦
網路上的芳鄰
檔名(N):
Nations.txt
檔案類型(T):
Text Files (*.txt,*.TXT,*.dat,*.DAT,*.csv,*.CS
開啓(O)
取消

- 按 **Edit data set** 可編輯資料

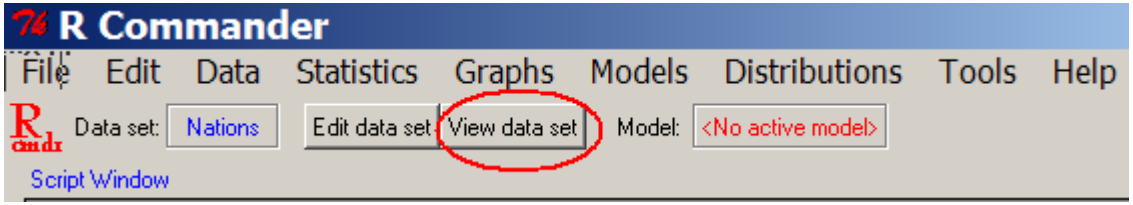


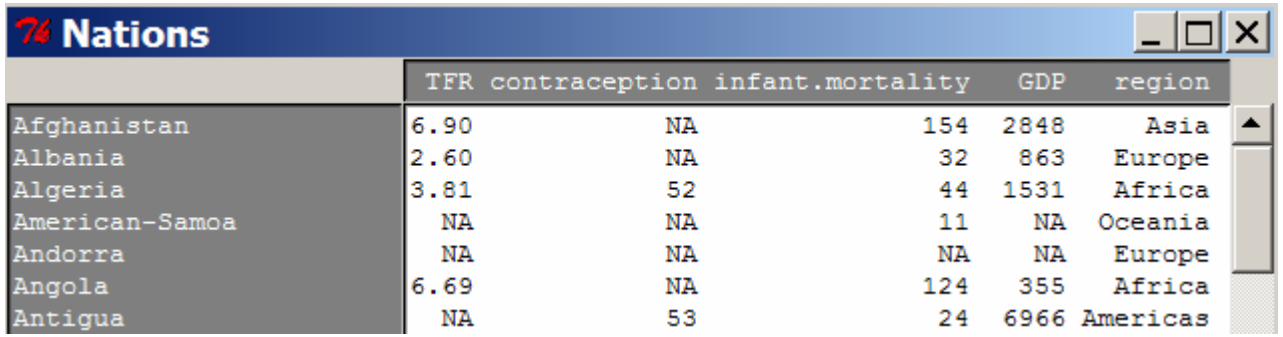


	row.names	TFR	contraception	infant.mortality	GDP
1	Afghanistan	6.9	NA	154	2848
2	Albania	2.6	NA	32	863
3	Algeria	3.81	52	44	1531
4	American-Samoa	NA	NA	11	NA
5	Andorra	NA	NA	NA	NA
6	Angola	6.69	NA	124	355

- 修改資料後，按 X 即可關閉編輯視窗

• 按 **View data set** 可檢視資料





	TFR	contraception	infant.mortality	GDP	region
Afghanistan	6.90	NA	154	2848	Asia
Albania	2.60	NA	32	863	Europe
Algeria	3.81	52	44	1531	Africa
American-Samoa	NA	NA	11	NA	Oceania
Andorra	NA	NA	NA	NA	Europe
Angola	6.69	NA	124	355	Africa
Antigua	NA	53	24	6966	Americas

•資料是 “data frames” 型式

•在Rcmdr編輯區中可顯示匯入資料指令 **read.table( )**：

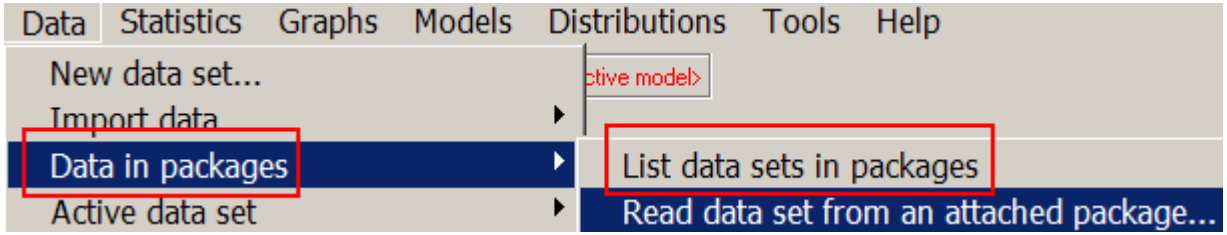
```
Nations <- read.table("C:/Program Files/R/R-2.4.0/library/Rcmdr/etc/Nations.txt", header=TRUE, sep="", na.strings="NA", dec=".", strip.white=TRUE)
```

顯示 package 的資料檔

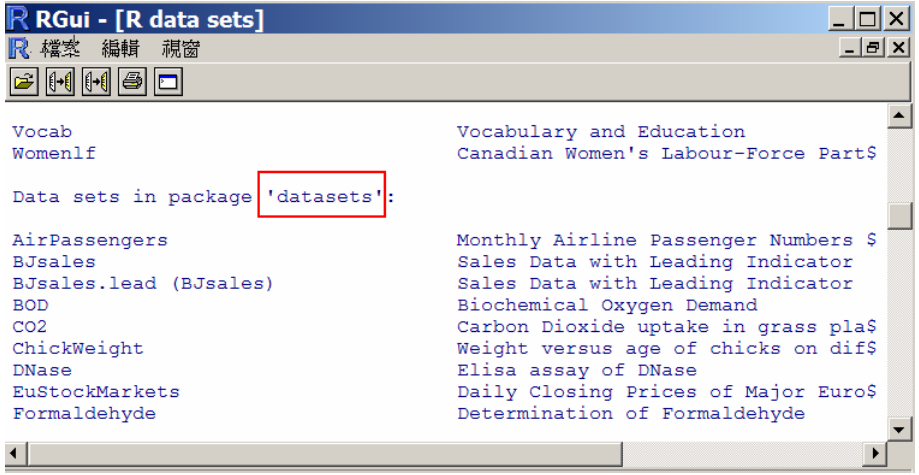
- R 的packages 中含有許多資料集可供測試.

  Data \ Data in packages \List data sets in packages

  顯示所有可用資料集,其中 datasets package 有許多資料可供測試。



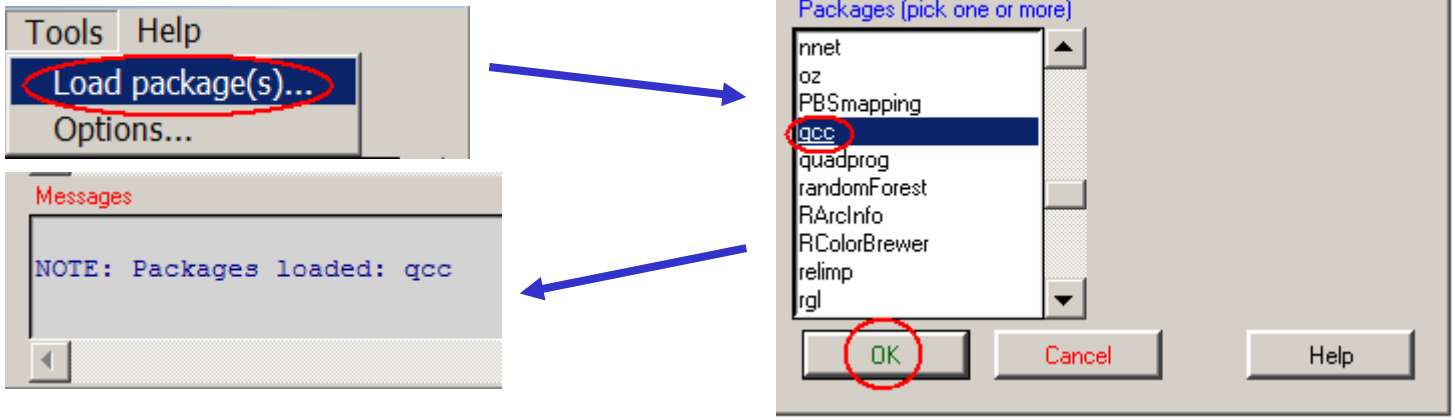

```
> data()
```



Data sets in package 'car':
…
Data sets in package 'datasets':
…

先載入package再匯入資料檔

- 如果須使用的資料沒有在預設的資料集中,則可先載入package。
- Tools \ Load package(s)… \ 選取 qcc \ 按 ok, 訊息區會有載入package 內容。



- Data \
  Data in packages \
  List data sets in packages,
  結果參考右圖

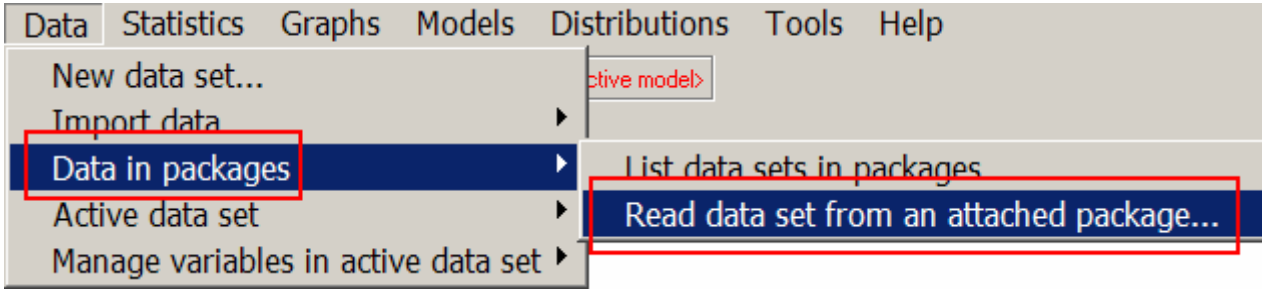
```
Data sets in package 'car':
…
Data sets in package 'datasets':
…
Data sets in package 'qcc':
…
```

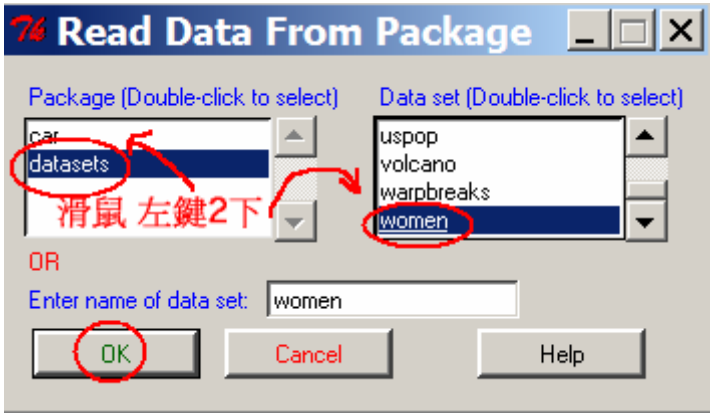
匯入 package 資料檔

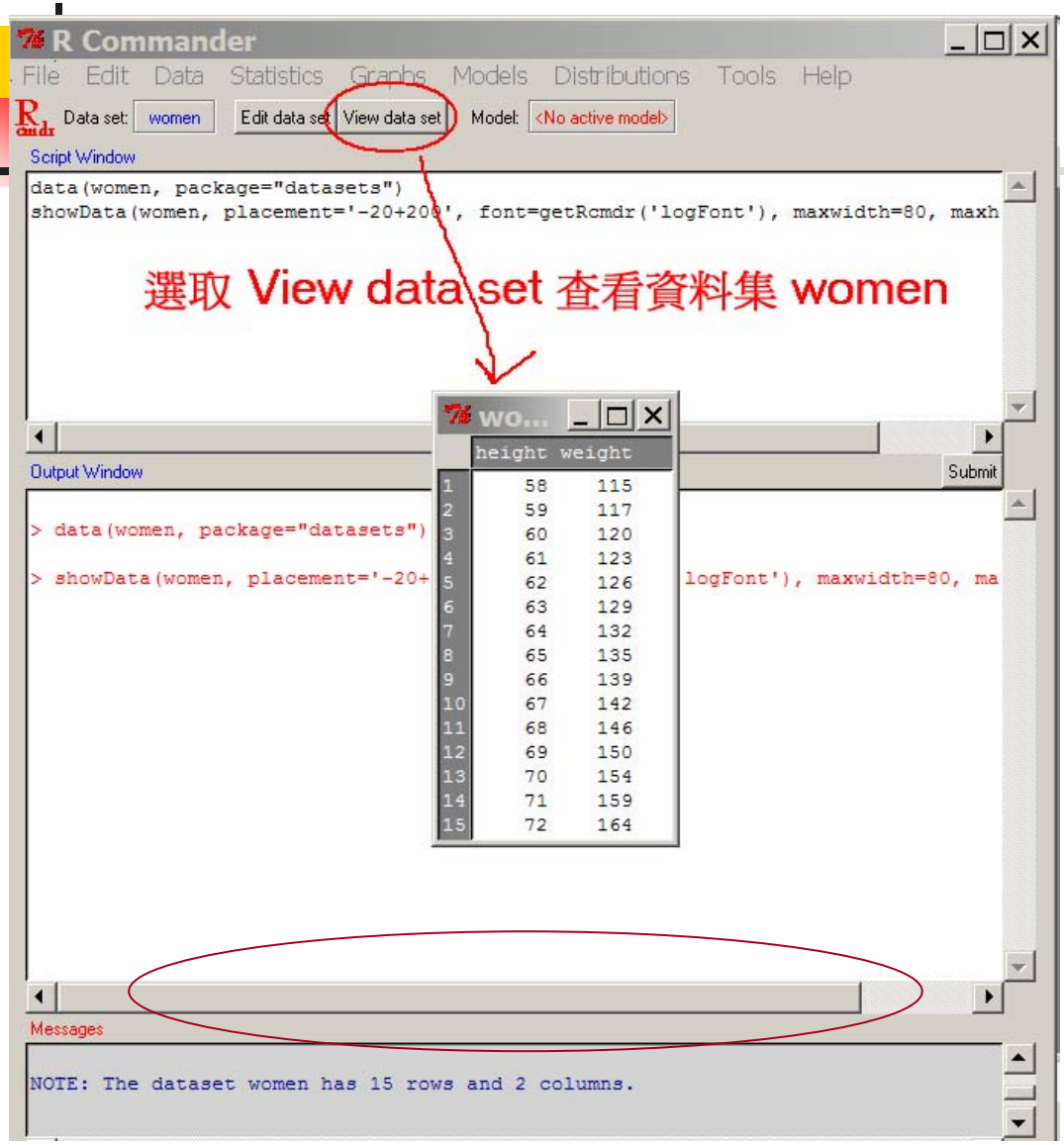
- 匯入R package 的資料集

Read data set from an attached package…

```
data(iris3, package="datasets")
```

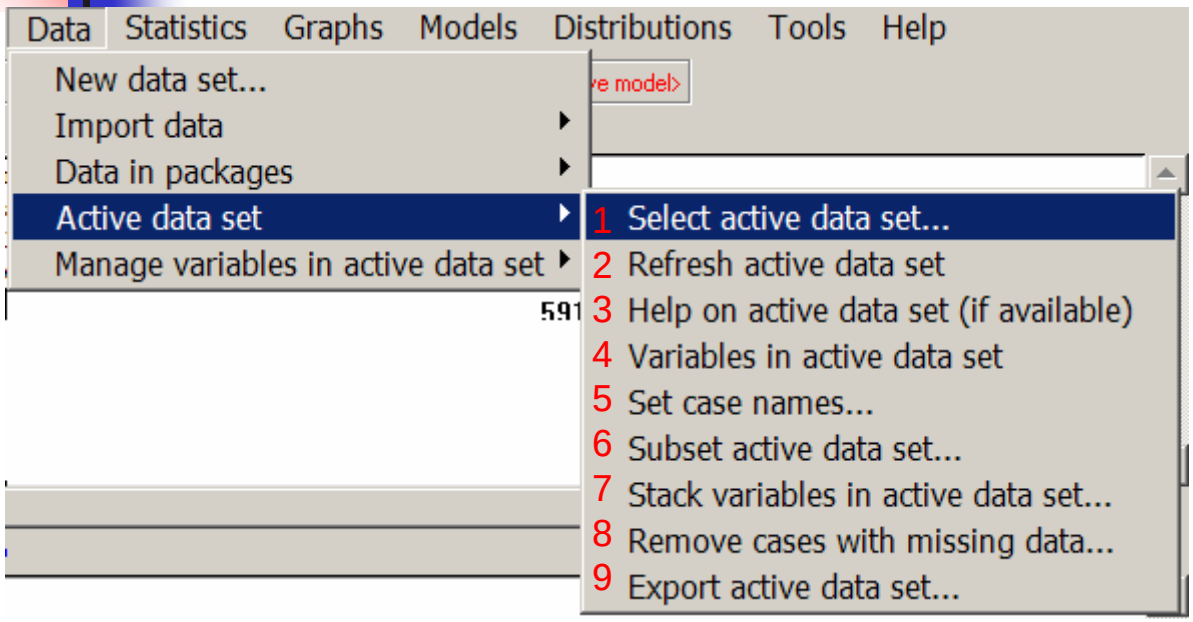






TRY ! 匯入 datasets package的 trees 資料集

Active data set 功能表

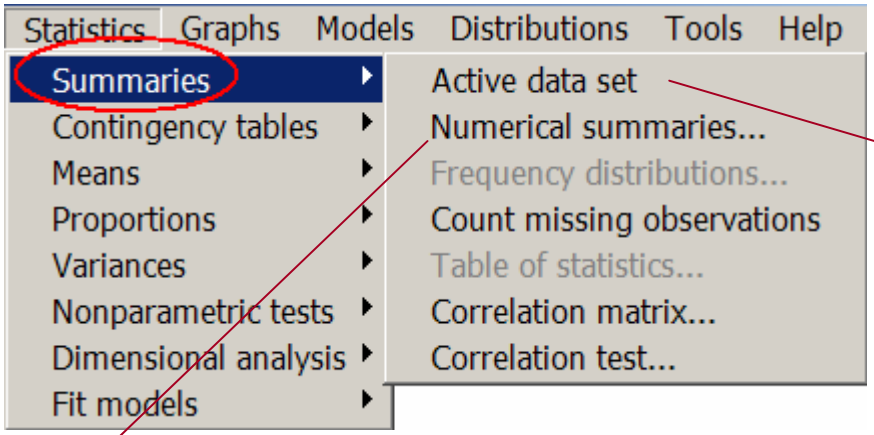


1. 如果有使用二個以上資料集,選擇目前須分析的資料集
2. 更新資料集
3. 查詢資料集的說明檔
4. 顯示資料集的變數名稱
5. 選取有列名稱的變數
6. 選取現有資料集的子集合
7. 串連資料集(向量)
8. 刪除含遺漏值的資料
9. 輸出資料

```
help("trees")
names(trees)
```

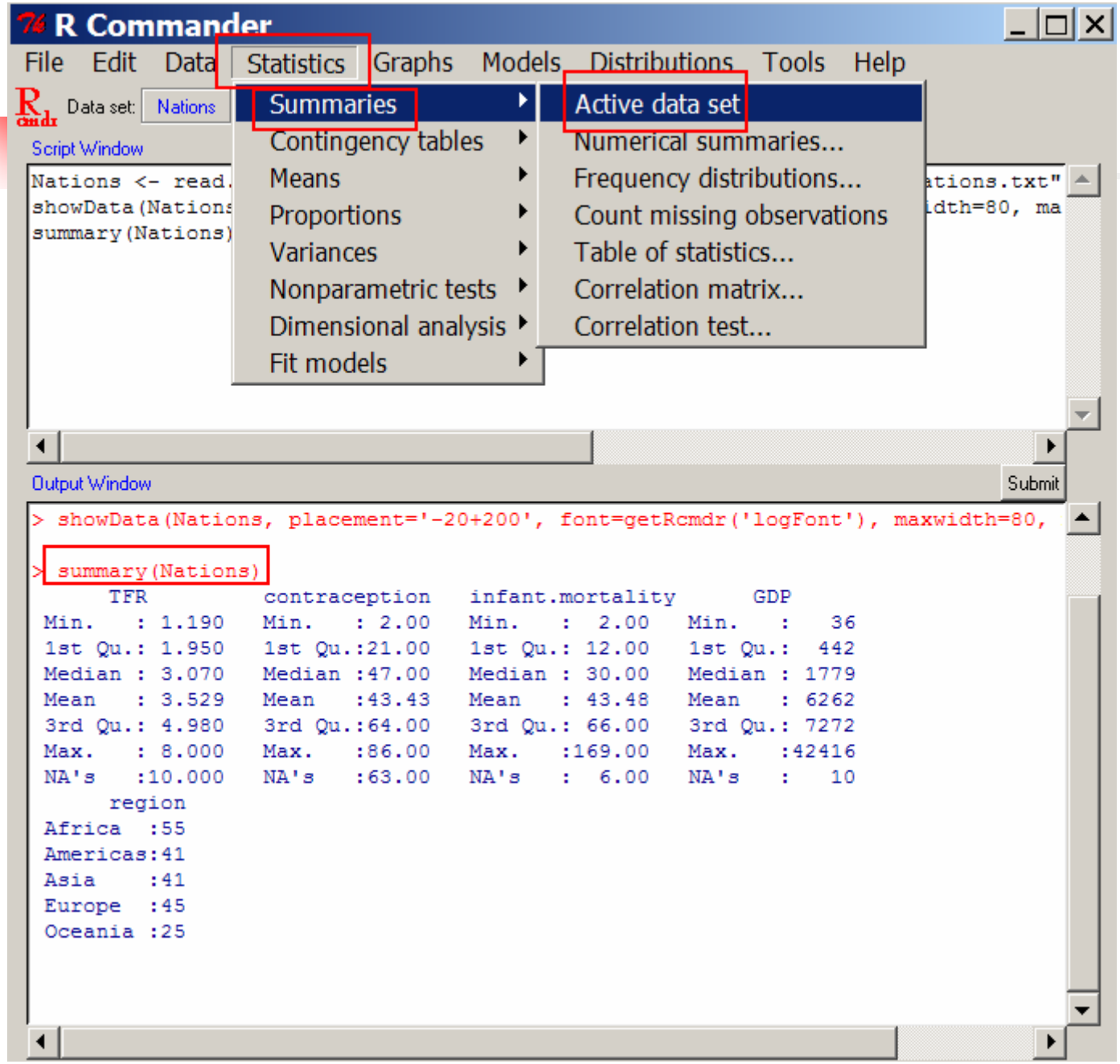
4. 資料摘要

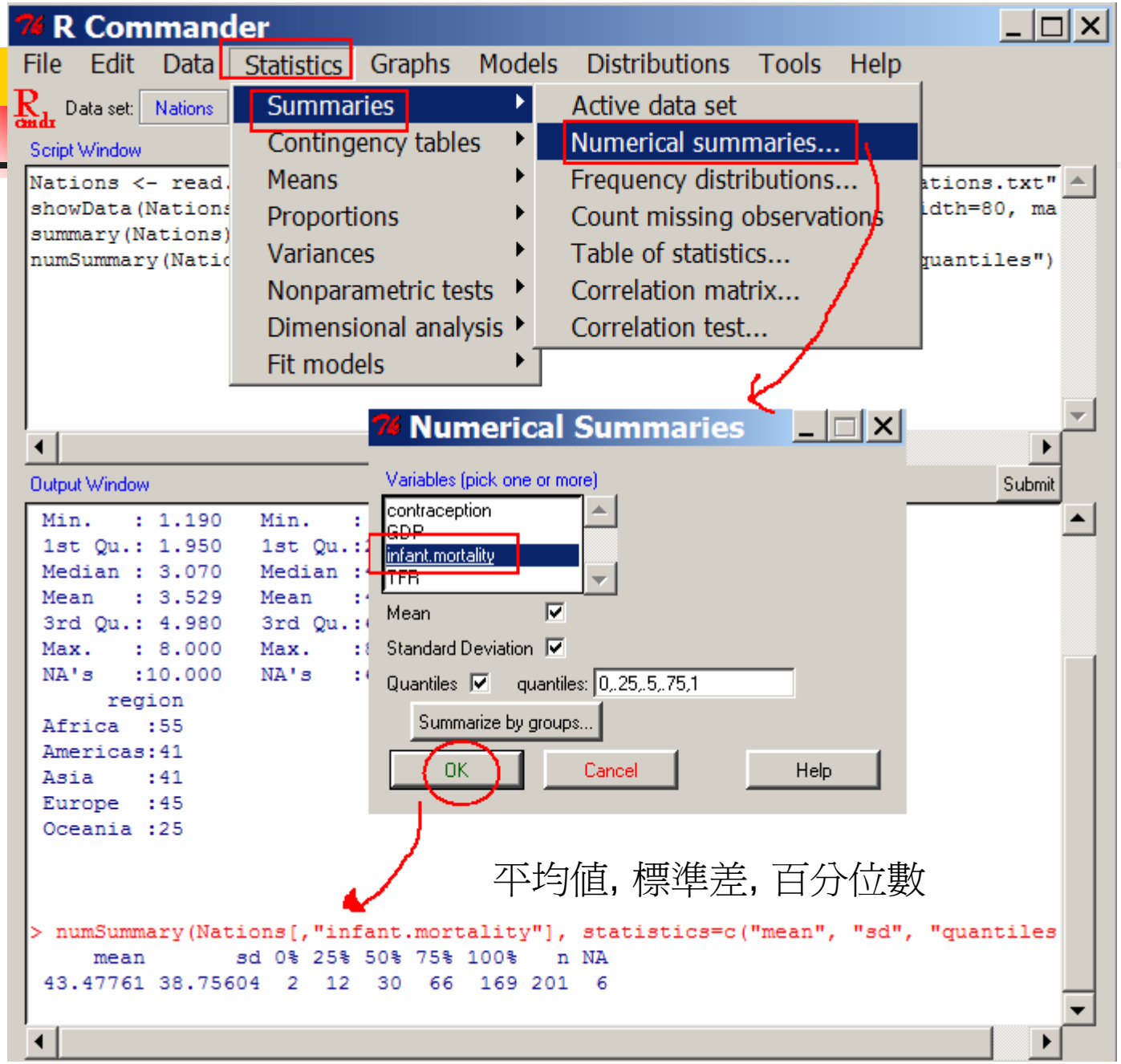
Statistics 功能表

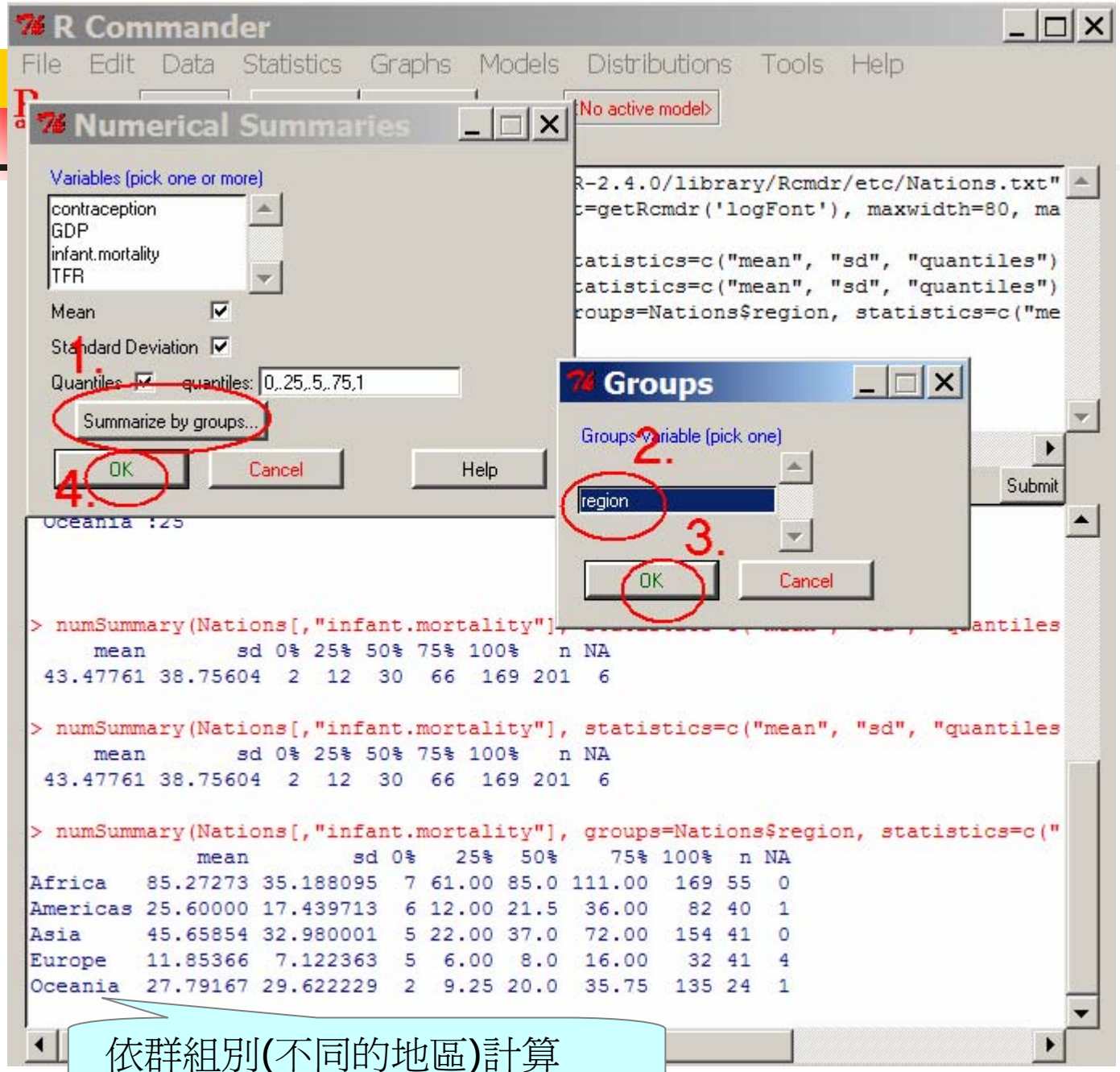


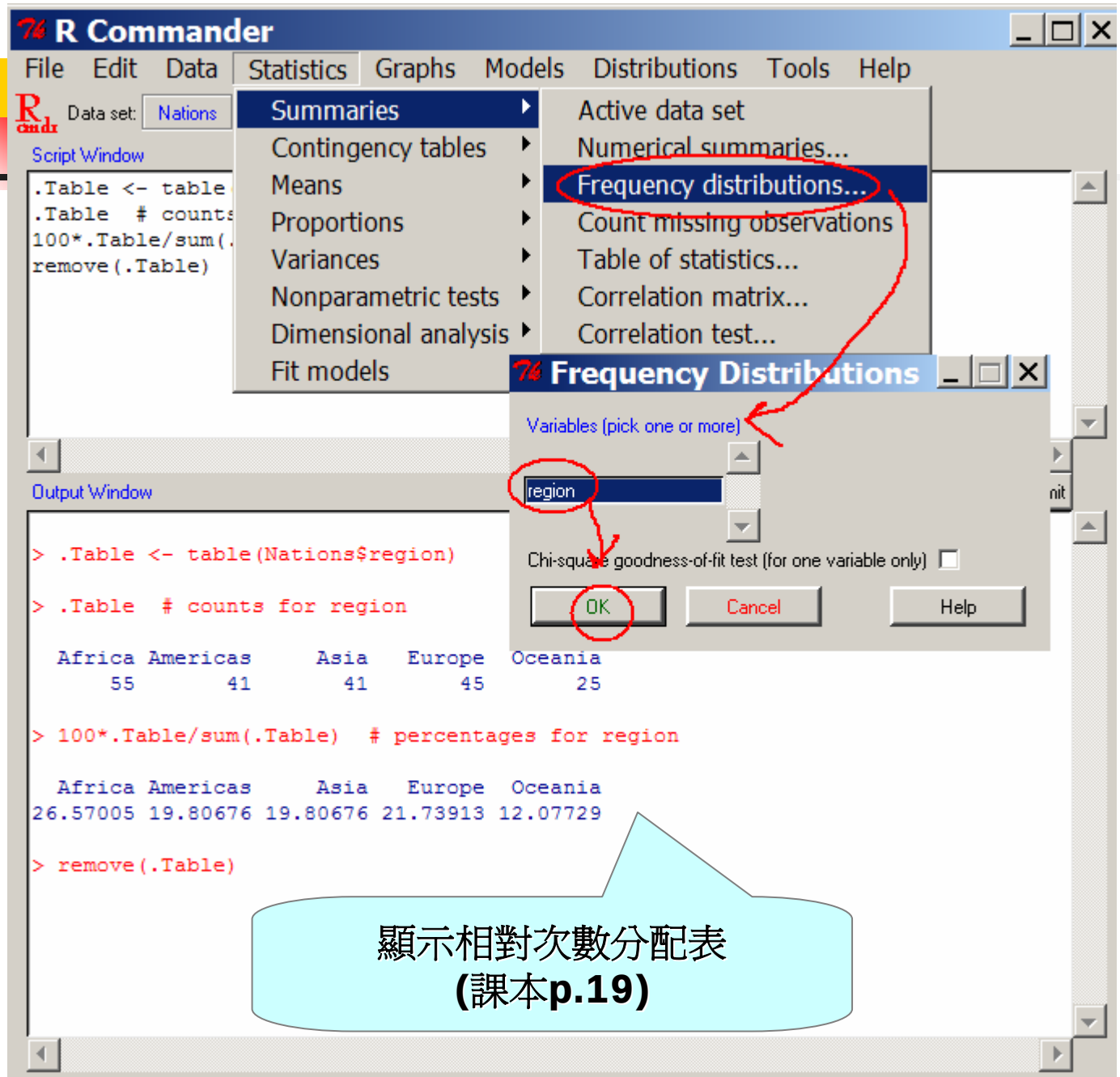

```
> summary(trees)
     Girth           Height       Volume     
 Min.   : 8.30   Min.   :63   Min.   :10.20  
 1st Qu.:11.05   1st Qu.:72   1st Qu.:19.40  
 Median :12.90   Median :76   Median :24.20  
 Mean   :13.25   Mean   :76   Mean   :30.17  
 3rd Qu.:15.25   3rd Qu.:80   3rd Qu.:37.30  
 Max.   :20.60   Max.   :87   Max.   :77.00  
```

```
> numSummary(trees[,"Girth"], statistics=c("mean", "sd", "quantiles"))
     mean       sd  0%   25%  50%   75% 100%  n
 13.24839 3.138139 8.3 11.05 12.9 15.25 20.6 31
```


R Commander
File Edit Data Statistics Graphs Models Distributions Tools Help
Data set: Nations
Summaries
Contingency tables
Means
Proportions
Variances
Nonparametric tests
Dimensional analysis
Fit models
Active data set
Numerical summaries...
Frequency distributions...
Count missing observations
Table of statistics...
Correlation matrix...
Correlation test...
Script Window
Nations <- read.
showData(Nations
summary(Nations)
ations.txt"
idth=80, ma
Output Window
Submit
> showData(Nations, placement='-20+200', font=getRcmdr('logFont'), maxwidth=80,
> summary(Nations)
TFR contraception infant.mortality GDP
Min. : 1.190 Min. : 2.00 Min. : 2.00 Min. : 36
1st Qu.: 1.950 1st Qu.:21.00 1st Qu.: 12.00 1st Qu.: 442
Median : 3.070 Median :47.00 Median : 30.00 Median : 1779
Mean : 3.529 Mean :43.43 Mean : 43.48 Mean : 6262
3rd Qu.: 4.980 3rd Qu.:64.00 3rd Qu.: 66.00 3rd Qu.: 7272
Max. : 8.000 Max. :86.00 Max. :169.00 Max. :42416
NA's :10.000 NA's :63.00 NA's : 6.00 NA's : 10
region
Africa :55
Americas:41
Asia :41
Europe :45
Oceania :25


R Commander
File Edit Data Statistics Graphs Models Distributions Tools Help
Data set: Nations
Summaries
Contingency tables
Means
Proportions
Variances
Nonparametric tests
Dimensional analysis
Fit models
Active data set
Numerical summaries...
Frequency distributions...
Count missing observations
Table of statistics...
Correlation matrix...
Correlation test...
Script Window
Nations <- read.
showData(Nations
summary(Nations)
numSummary(Natic
ations.txt"
idth=80, ma
quantiles")
Numerical Summaries
Variables (pick one or more)
contraception
GDP
infant.mortality
TFR
Mean
Standard Deviation
Quantiles
quantiles: 0,.25,.5,.75,1
Summarize by groups...
OK
Cancel
Help
Submit
Output Window
Min. : 1.190
1st Qu.: 1.950
Median : 3.070
Mean : 3.529
3rd Qu.: 4.980
Max. : 8.000
NA's :10.000
region
Africa :55
Americas:41
Asia :41
Europe :45
Oceania :25
平均值, 標準差, 百分位數
> numSummary(Nations[,"infant.mortality"], statistics=c("mean", "sd", "quantiles
mean sd 0% 25% 50% 75% 100% n NA
43.47761 38.75604 2 12 30 66 169 201 6


R Commander
File Edit Data Statistics Graphs Models Distributions Tools Help
<No active model>
Numerical Summaries
Variables (pick one or more)
contraception
GDP
infant.mortality
TFR
Mean
Standard Deviation
Quantiles
quantiles: 0,.25,.5,.75,1
1.
Summarize by groups...
OK
Cancel
Help
4.
Groups
Groups variable (pick one)
2.
region
3.
OK
Cancel
Submit
R-2.4.0/library/Rcmdr/etc/Nations.txt"
t=getRcmdr('logFont'), maxwidth=80, ma
tatistics=c("mean", "sd", "quantiles")
tatistics=c("mean", "sd", "quantiles")
roups=Nations$region, statistics=c("me
Oceania :25
> numSummary(Nations[,"infant.mortality"]
mean sd 0% 25% 50% 75% 100% n NA
43.47761 38.75604 2 12 30 66 169 201 6
> numSummary(Nations[,"infant.mortality"], statistics=c("mean", "sd", "quantiles
mean sd 0% 25% 50% 75% 100% n NA
43.47761 38.75604 2 12 30 66 169 201 6
> numSummary(Nations[,"infant.mortality"], groups=Nations$region, statistics=c("
mean sd 0% 25% 50% 75% 100% n NA
Africa 85.27273 35.188095 7 61.00 85.0 111.00 169 55 0
Americas 25.60000 17.439713 6 12.00 21.5 36.00 82 40 1
Asia 45.65854 32.980001 5 22.00 37.0 72.00 154 41 0
Europe 11.85366 7.122363 5 6.00 8.0 16.00 32 41 4
Oceania 27.79167 29.622229 2 9.25 20.0 35.75 135 24 1
依群組別(不同的地區)計算


R Commander
File Edit Data Statistics Graphs Models Distributions Tools Help
Data set: Nations
Summaries
Contingency tables
Means
Proportions
Variances
Nonparametric tests
Dimensional analysis
Fit models
Active data set
Numerical summaries...
Frequency distributions...
Count missing observations
Table of statistics...
Correlation matrix...
Correlation test...
Script Window
.Table <- table
.Table # counts
100*.Table/sum(.
remove(.Table)
Frequency Distributions
Variables (pick one or more)
region
Chi-square goodness-of-fit test (for one variable only)
OK
Cancel
Help
Output Window
> .Table <- table(Nations$region)
> .Table # counts for region
Africa Americas Asia Europe Oceania
55 41 41 45 25
> 100*.Table/sum(.Table) # percentages for region
Africa Americas Asia Europe Oceania
26.57005 19.80676 19.80676 21.73913 12.07729
> remove(.Table)
顯示相對次數分配表
(課本p.19)

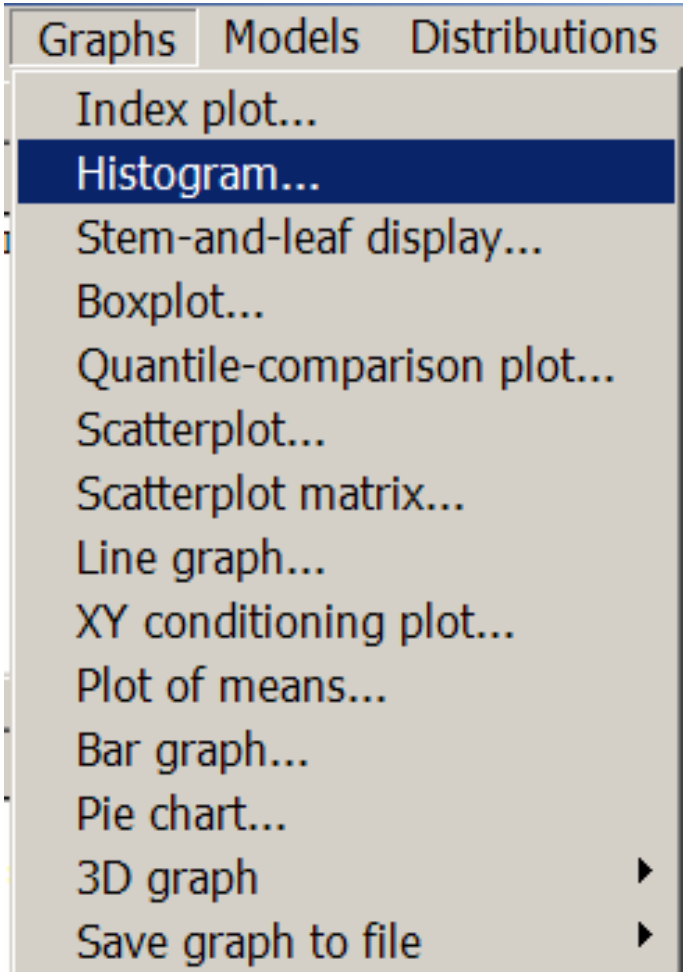
設定工作環境

- options(digits=3)
  設定 整數+小數點 顯示位數爲3位,預設值爲7.

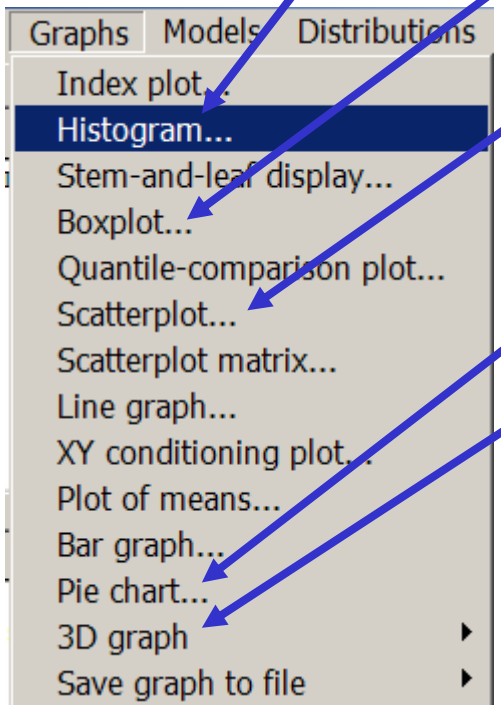
```
> x <-sqrt(2)
> x
[1] 1.414214
> options(digits=3)
> y <- sqrt(2)
> y
[1] 1.41
> 
```

- ? options
  顯示options輔助說明

5. 圖形

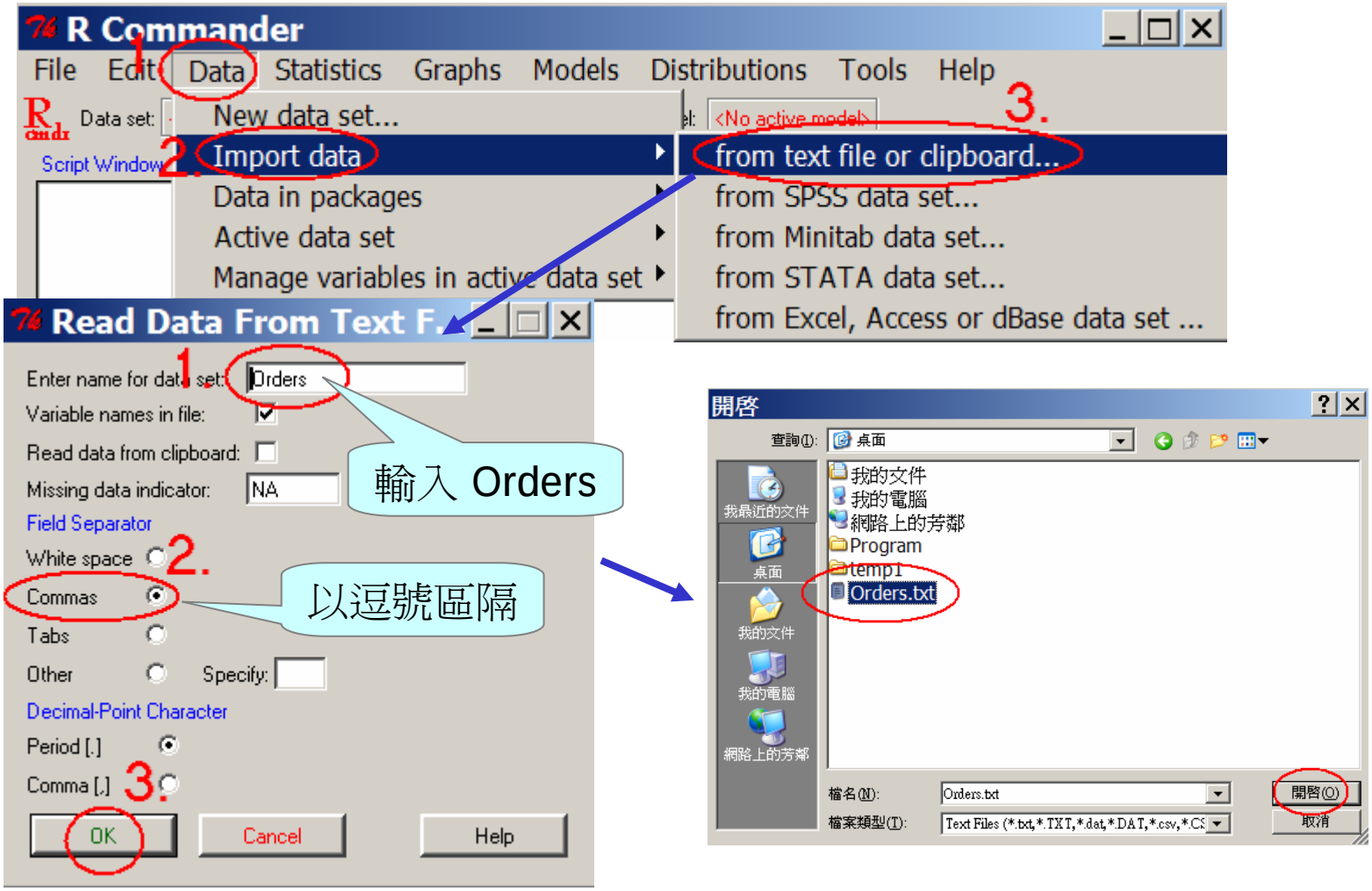

Graphs
Models
Distributions
Index plot...
Histogram...
Stem-and-leaf display...
Boxplot...
Quantile-comparison plot...
Scatterplot...
Scatterplot matrix...
Line graph...
XY conditioning plot...
Plot of means...
Bar graph...
Pie chart...
3D graph
Save graph to file

- 圖形包括 直方圖, 盒形圖, 散佈圖, 圓形圖, 3D繪圖等.



Ex1: 繪圖直方圖

- Orders 資料集



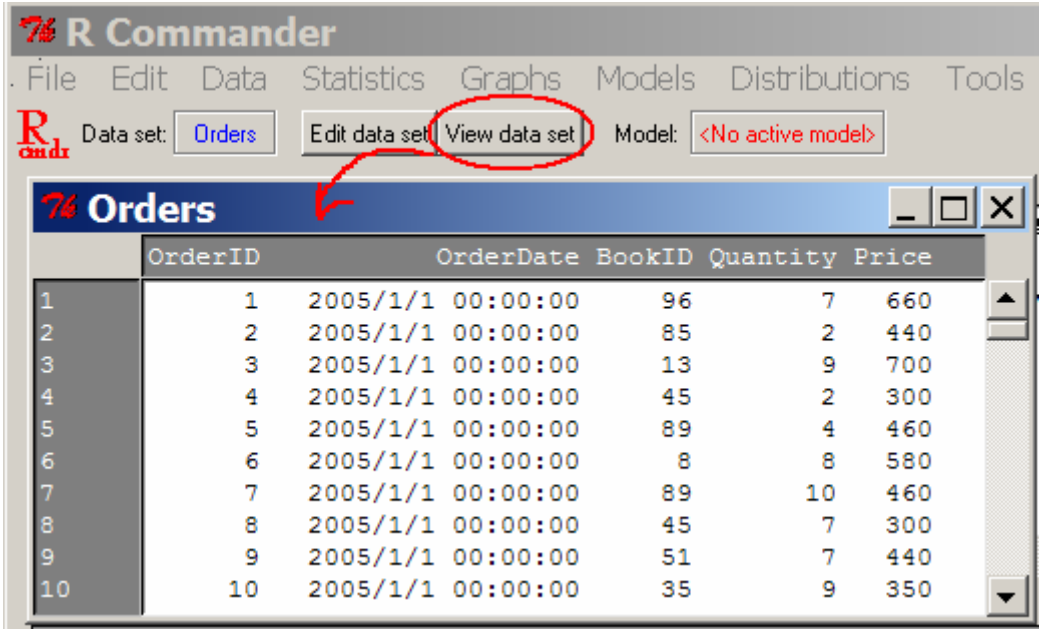
Ex1: 繪圖直方圖 (cont.)

- 完成後訊息視窗顯示資料集 Orders 包括100,000筆資料, 5個欄位.

Messages

```
NOTE: The dataset Orders has 100000 rows and 5 columns.
```

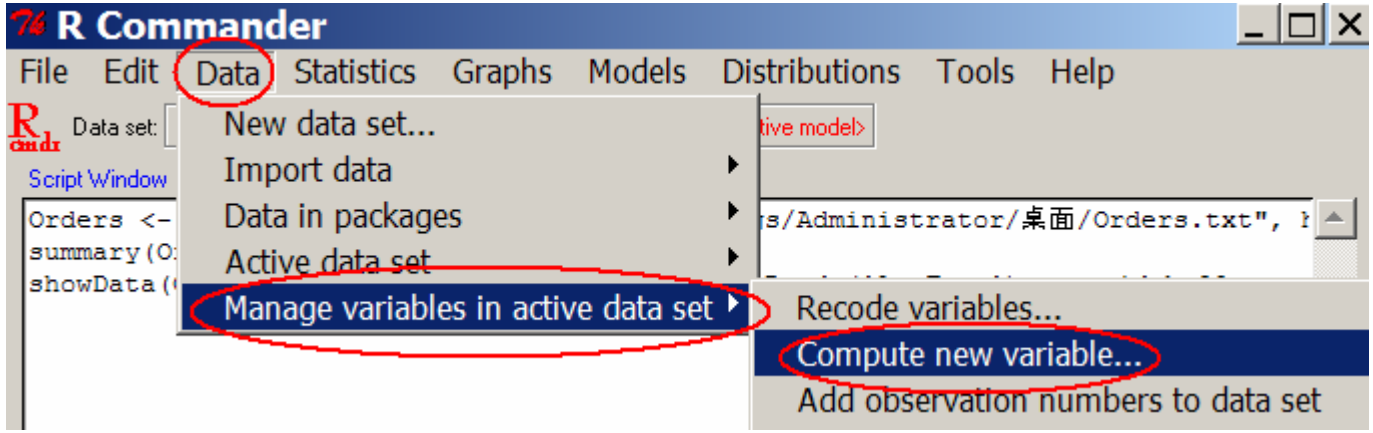
- 檢視資料集 View data set

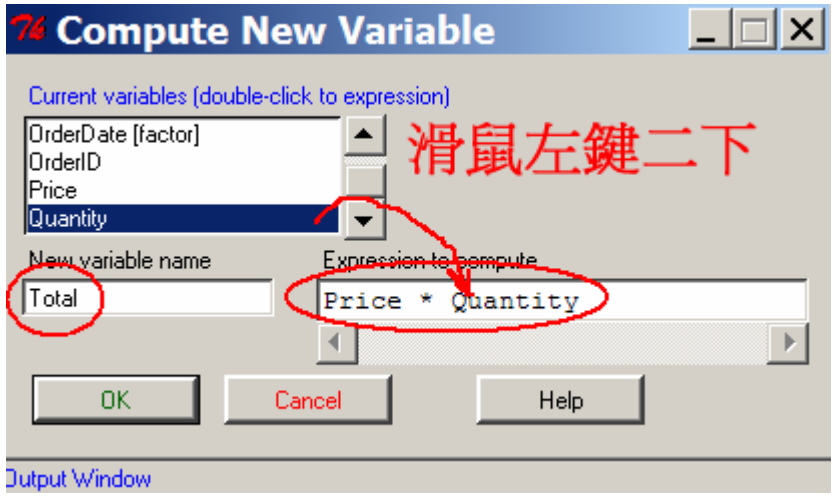


	OrderID	OrderDate	BookID	Quantity	Price
1	1	2005/1/1 00:00:00	96	7	660
2	2	2005/1/1 00:00:00	85	2	440
3	3	2005/1/1 00:00:00	13	9	700
4	4	2005/1/1 00:00:00	45	2	300
5	5	2005/1/1 00:00:00	89	4	460
6	6	2005/1/1 00:00:00	8	8	580
7	7	2005/1/1 00:00:00	89	10	460
8	8	2005/1/1 00:00:00	45	7	300
9	9	2005/1/1 00:00:00	51	7	440
10	10	2005/1/1 00:00:00	35	9	350

Ex1: 繪圖直方圖 (cont.)

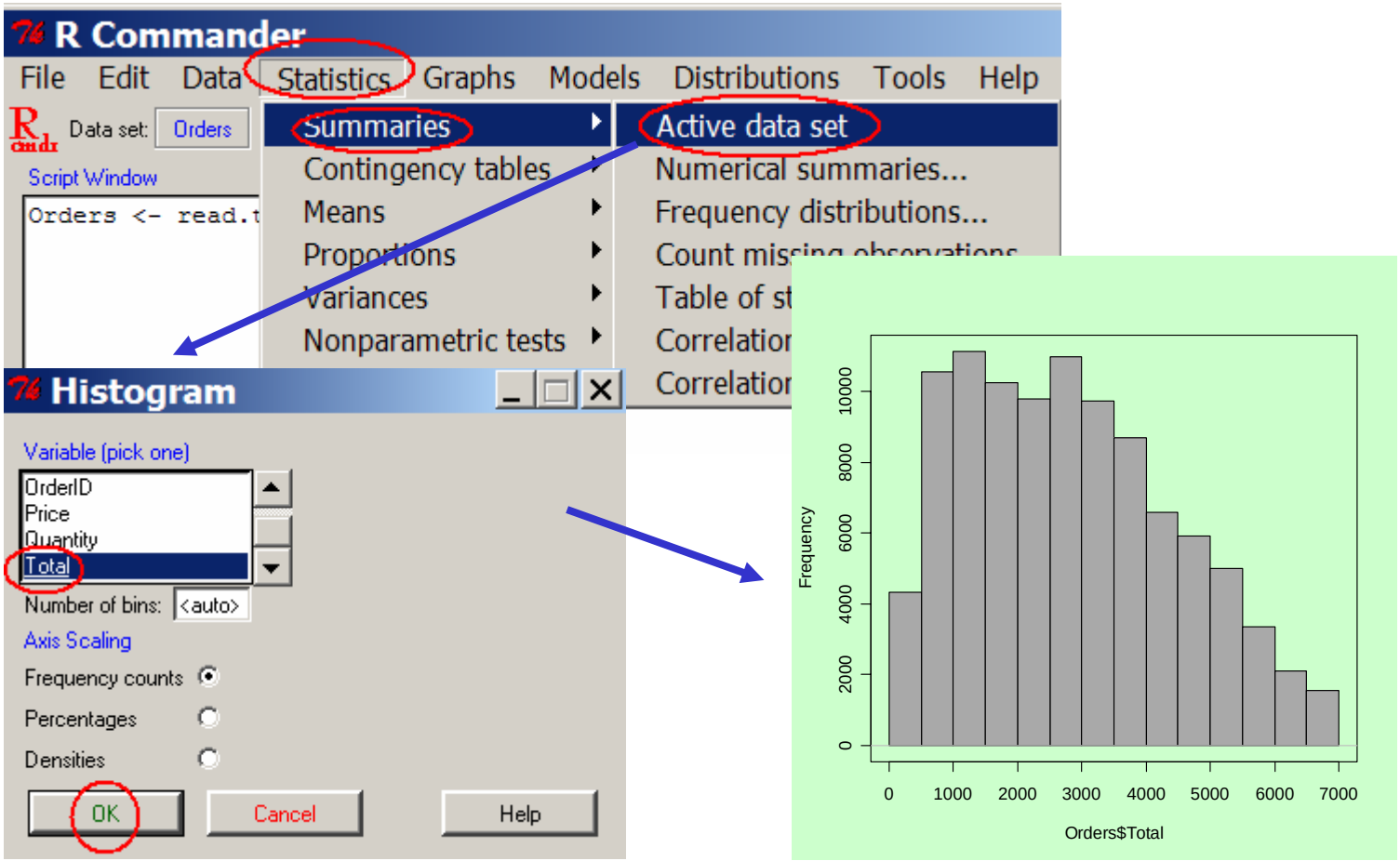
- 加入新變數 Total.





Ex1: 繪圖直方圖 (cont.)

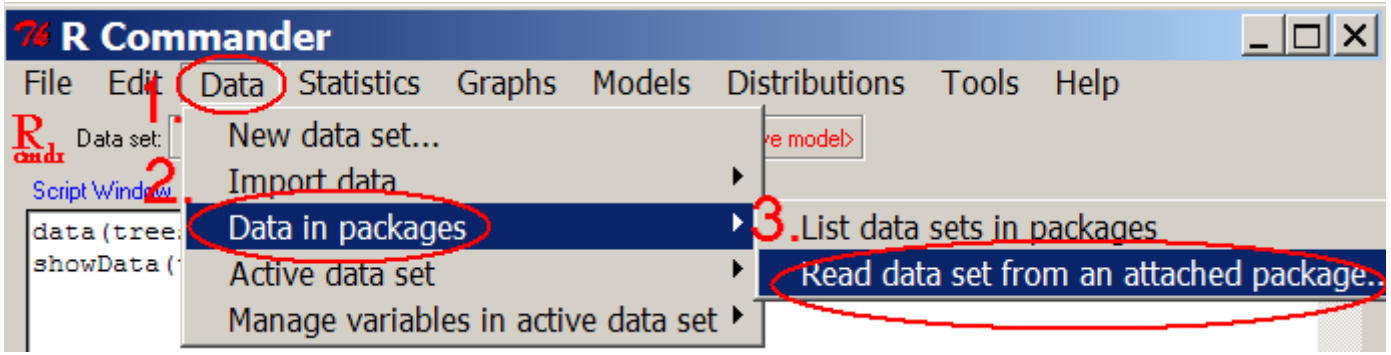
- 繪製直方圖 Histogram



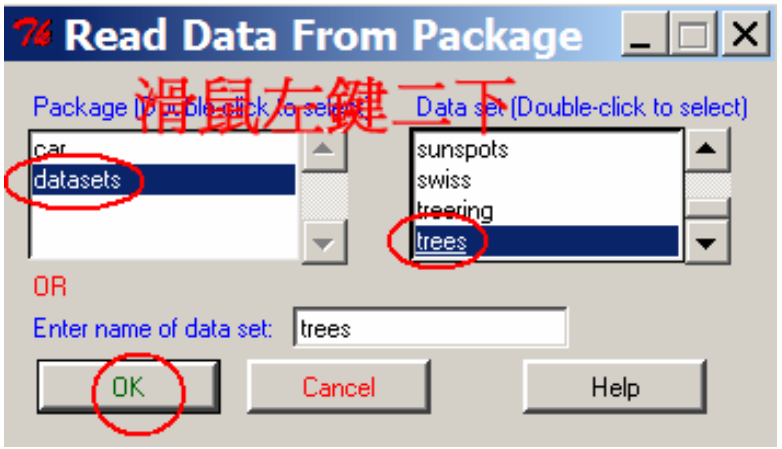
Ex2: 繪圖3D圖

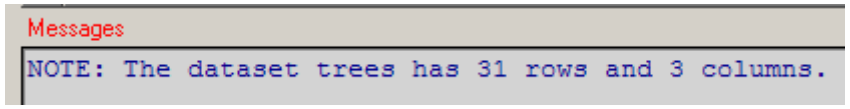
- 匯入 trees資料集

(1)

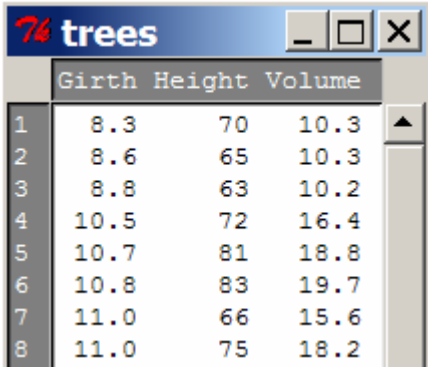


(2)





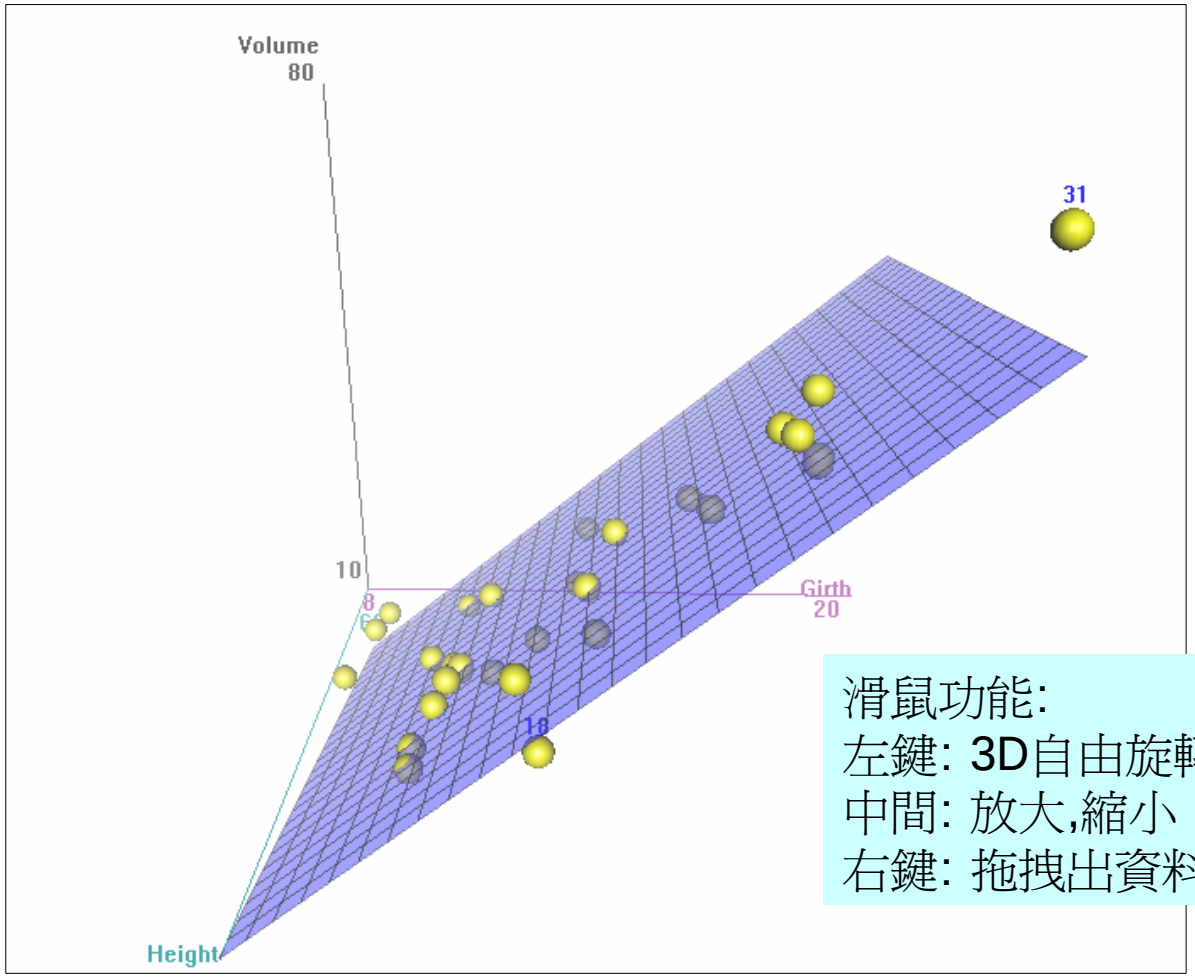
(3)



trees

	Girth	Height	Volume
1	8.3	70	10.3
2	8.6	65	10.3
3	8.8	63	10.2
4	10.5	72	16.4
5	10.7	81	18.8
6	10.8	83	19.7
7	11.0	66	15.6
8	11.0	75	18.2

Ex2: 繪圖3D圖(cont.)



滑鼠功能:
左鍵: 3D自由旋轉
中間: 放大,縮小
右鍵: 拖拽出資料點範圍,會顯示資料集編號

6. 機率分配

分配	R 分配名稱	參數
Beta	**beta**	shapel, shape2
Binomial	**binom**	size, prob
Cauchy	**cauchy**	location, scale
Chi-squared	**chisq**	df, ncp
Exponential	**exp**	rate
F	**f**	dfl, df2, ncp
Gamma	**gamma**	shape, rate
Geometric	**geom**	prob
Hypergeometric	**hyper**	m, n, k
Log-normal	**lnorm**	meanlog, sdlog
Loistic	**logis**	location, scale
Negative binomial	**nbinom**	size, prob
Normal	**norm**	mean, sd
Poisson	**pois**	lambda
Student's t	**t**	df, ncp
Uniform	**unif**	min, max
Weibull	**weibull**	shape, scale
Wilcoxon	**wilcox**	m, n

機率分配- d, p, q, r

- The standard distributions:
  - d:機率密度函數(Probability Density Functions, pdf or p.d.f.)
  - p:累積分配函數 (Cumulative distribution function, CDF) $F(x) = P(X \leq x)$
  - q:百分比函數 (Quantile function)
  - r:隨機產生分配的資料

EXCEL 函數

	A	B
1	0.058440944	=NORMDIST(1.96,0,1,FALSE)
2	0.975002105	=NORMDIST(1.96,0,1,TRUE)
3	1.959963985	=NORMINV(0.975,0,1)

```
> dnorm(1.96,0,1)
[1] 0.05844094
> pnorm(1.96,0,1)
[1] 0.9750021
> qnorm(0.975,0,1)
[1] 1.959964
> rnorm(5,0,1)
[1] -0.3986337  0.9078262  0.7505547 -0.5570301 -0.7823251
```

Normal distribution (Mean, sd)

Function	Usage
Density	dnorm(x, mean=0, sd=1, log = FALSE)
distribution function	pnorm(q, mean=0, sd=1, lower.tail = TRUE, log.p = FALSE)
quantile function	qnorm(p, mean=0, sd=1, lower.tail = TRUE, log.p = FALSE)
random generation	rnorm(n, mean=0, sd=1)

x, q: vector of quantiles.
p: vector of probabilities.
n: number of observations.
mean: vector of means.
sd: vector of standard deviations.
log, log.p logical: if TRUE, probabilities p are given as log(p).
lower.tail logical: if TRUE (default), probabilities are P[X <= x],
otherwise, P[X > x].

範例: dnorm(), pnorm(), qnorm(), rnorm()

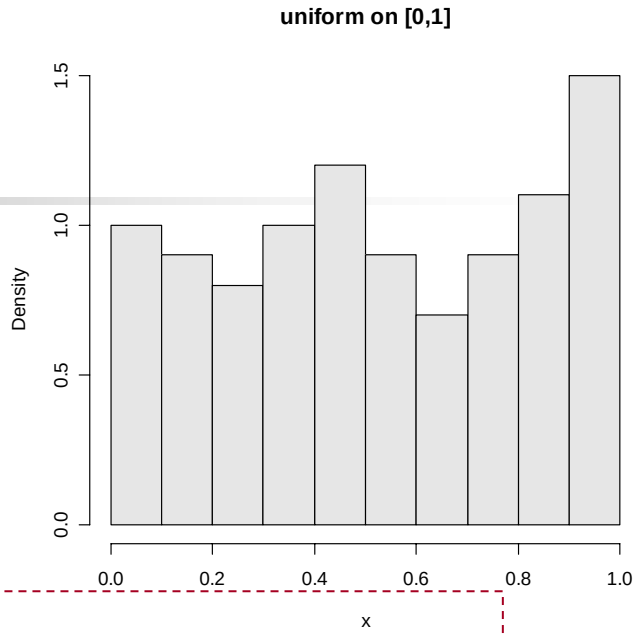
```
> dnorm(1.645)
[1] 0.1031108

> pnorm(1.645)
[1] 0.950015
> pnorm(1.96)
[1] 0.9750021
> pnorm(2)
[1] 0.9772499

> qnorm(0.95,0,1)
[1] 1.644854

> rnorm(5,0,1)
[1] -0.0714020  0.3427821 -1.1009895  0.6383514  0.2380390
```

隨機產生資料




```
> runif(1,0,2) # time at light
[1] 1.088542 # also runif(1,min=0,max=2)

> runif(5,0,2) # time at 5 lights
[1] 0.8577781 1.4196343 1.5049239 1.9123655 0.6559980

> runif(5) # 5 random numbers in [0,1]
[1] 0.1717392 0.5073215 0.7584391 0.1981516 0.8141901

> x=runif(100) # get the random numbers
> hist(x,probability=TRUE,col=gray(.9),main="uniform on [0,1]")
```

範例:二項分配

- 已知某產品之不良率爲0.1,隨機抽取10個產品檢查,至多有3個產品爲不良品的機率爲何?

解答:

方法1: 直接計算

$$P(X \le 3) = \sum_{i=0}^{3} f(x) = \sum_{x=0}^{3} C_x^{10} (0.1)^x (0.9)^{10-x}$$, 查表可得0.9872

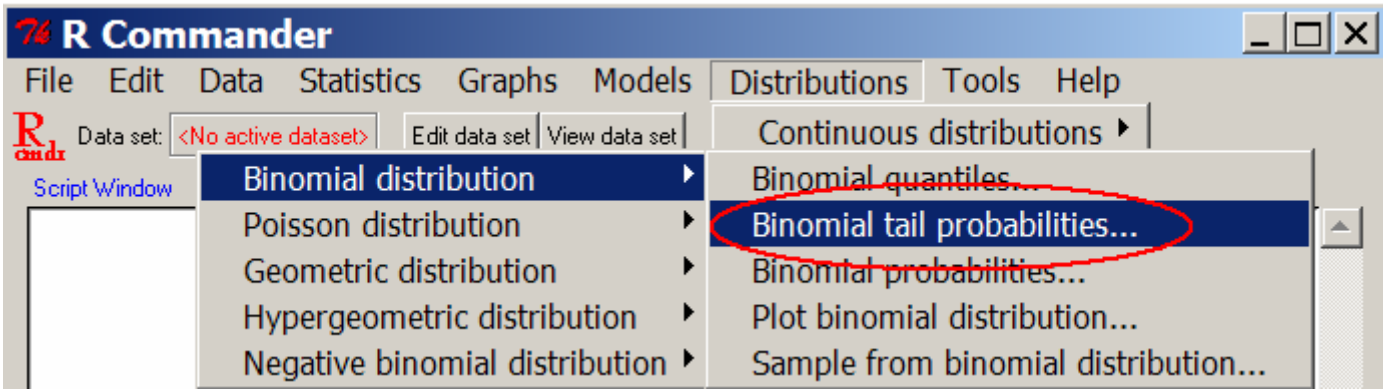
方法2: R

```
> pbinom(3, 10, 0.1)
[1] 0.9872048
```

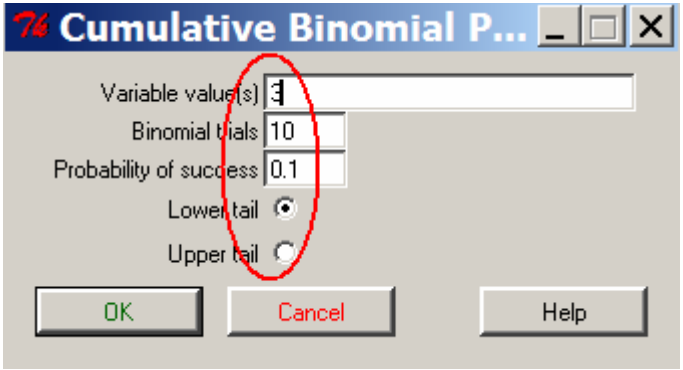
範例:二項分配(cont.)

方法3: R Commander

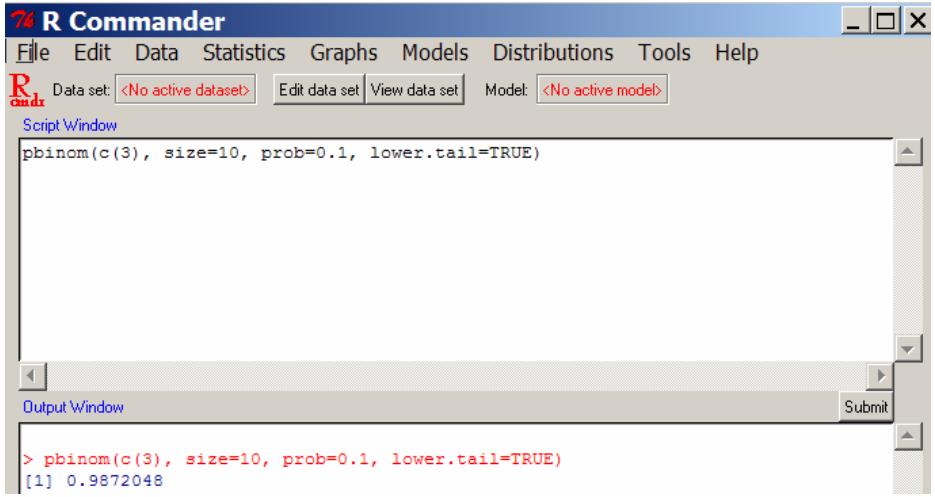
(1).



(2).



(3).



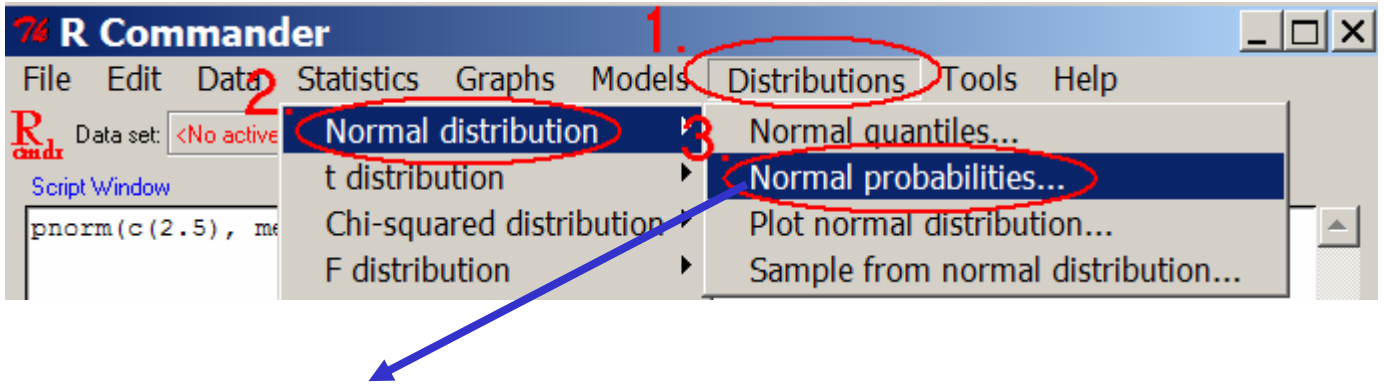
範例:常態分配

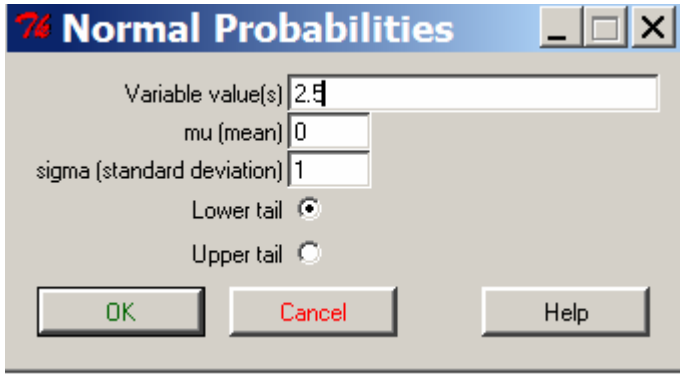
例題 7.3

若 $R.V.\ Z \sim N(0,1)$，求(1)$P(Z \le 2.5) = ?$ ，$P(Z \le 2.41) = ?$

(2) $P(-2 \le Z \le 3) = ?$

(3)求常數 a ，使得$P(Z \le a) = 0.95$。






```
> pnorm(c(2.5), mean=0, sd=1, lower.tail=TRUE)
[1] 0.9937903
```

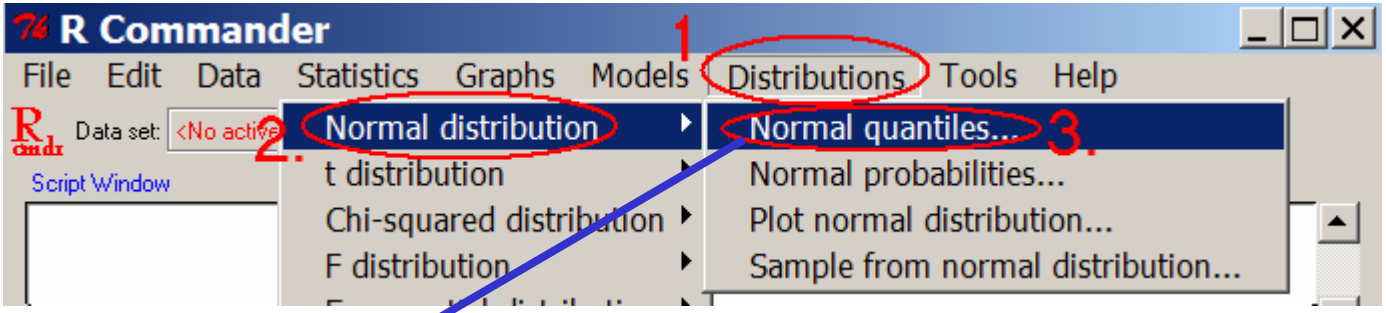
TRY !

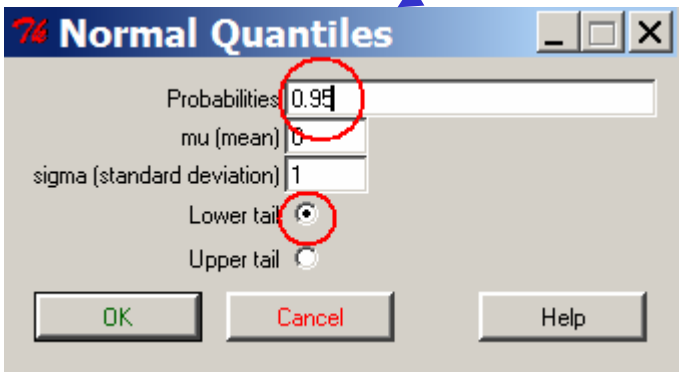
(1) P(Z≦2.41)=?

(2) P(-2 ≦ Z ≦ 3)=?

範例:常態分配(cont.)

(3)






```
> qnorm(c(0.95), mean=0, sd=1, lower.tail=TRUE)
[1] 1.644854
```

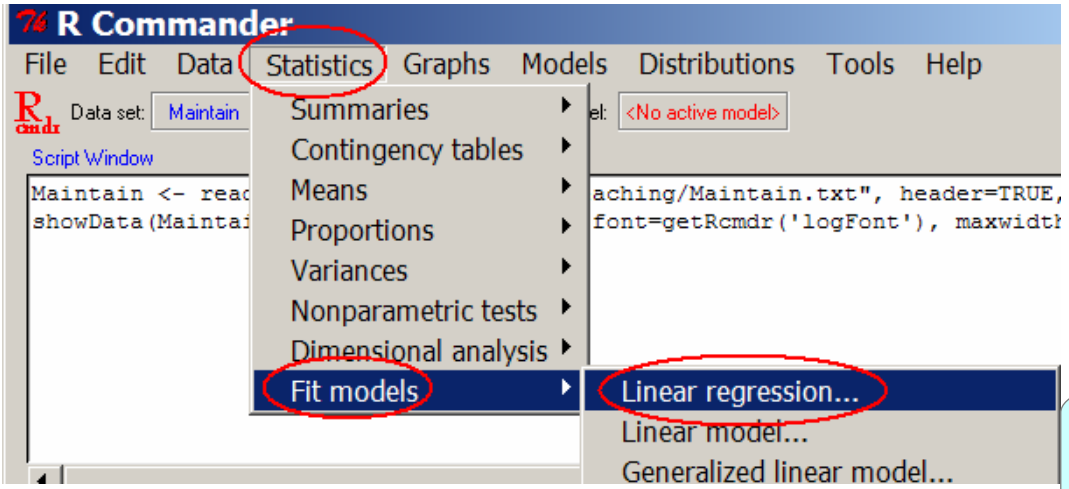
7. 迴歸分析 Regression Analysis

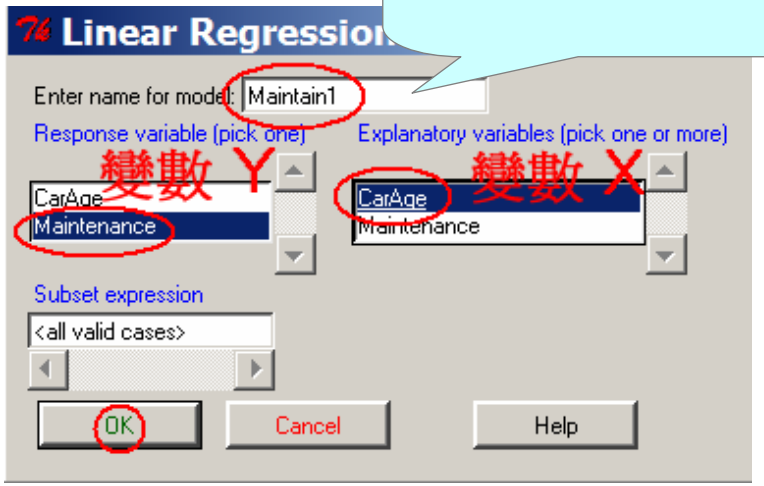
- 考慮某品牌汽車之車齡與保養費用資料如下:

CarAge	1	2	2	3	3	4	4	5
Maintenance	6500	12000	13000	15000	20000	20000	25000	30000

Q:利用最小平方法找出 $\hat{y} = \hat{\alpha} + \hat{\beta}x, x$:車齡 , y:保養費用

- Statistics\Fit models\Linear regression





輸入模式名稱

輸出畫面

```
> Maintain1 <- lm(Maintenance~CarAge, data=Maintain)
> summary(Maintain1)
```

Output Window

```
Call:
lm(formula = Maintenance ~ CarAge, data = Maintain)

Residuals:
    Min      1Q  Median      3Q     Max
-3270.8  -750.0   437.5  1291.7  2312.5

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)    937.5     2034.6   0.461 0.661184
CarAge        5583.3      627.9   8.892 0.000113 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2175 on 6 degrees of freedom
Multiple R-Squared: 0.9295,     Adjusted R-squared: 0.9177
F-statistic: 79.07 on 1 and 6 DF,  p-value: 0.0001127
```

$$\hat{y} = 937.5 + 5583.3x,$$

$$\hat{\alpha} = 937.5$$

$$\hat{\beta} = 5583.3$$

$$R^2 = 0.9295$$

- R^2:判定係數, coefficient of determination

Linear model

```
> names(Maintain1)
 [1] "coefficients"  "residuals"     "effects"       "rank"
 [5] "fitted.values" "assign"        "qr"            "df.residual"
 [9] "xlevels"       "call"          "terms"         "model"

> Maintain1$coefficient
(Intercept)      CarAge
    937.500    5583.333

> Maintain1$fitted.values
        1         2         3         4         5         6         7         8
6520.833 12104.167 12104.167 17687.500 17687.500 23270.833 23270.833 28854.167
```

- names: 物件之內容
- Maintain1$coefficient: Maintain1物件中的 coefficient 值
- Maintain1$fitted.values: 迴歸模型的預測值

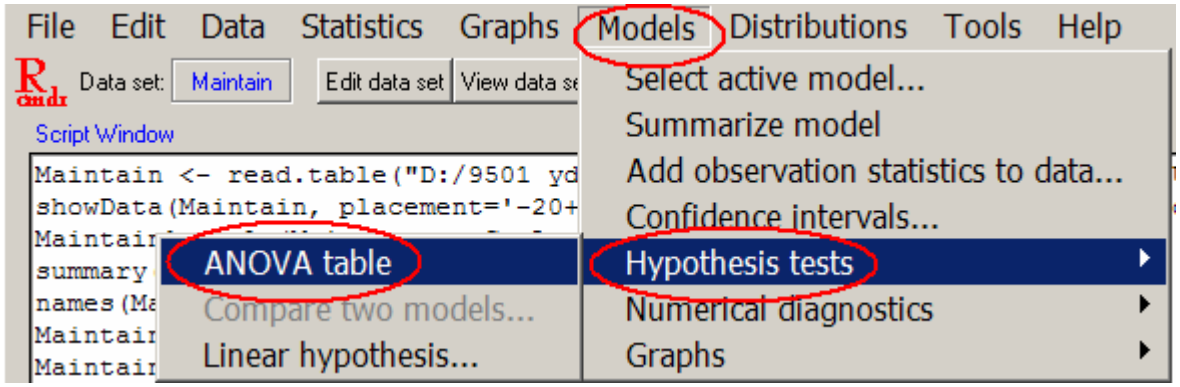
Q:此迴歸模型是吾具有解釋能力

- 利用ANOVA

表 13.3　檢定迴歸模型之變異數分析表

變異來源	平方和	自由度	均方	f 值
迴歸模型	SSR	1	$MSR=\frac{SSR}{1}$	$f_0=\frac{SSR}{S^2}$
隨機誤差	SSE	$n-2$	$MSE=S^2=\frac{SSE}{n-2}$	
總和	SST	$n-1$		

(f_0 -- 大)



Q:此迴歸模型是吾具有解釋能力(cont.)

H_0:此模型不具解釋能力(β =0)
H_1:此模型具解釋能力

```
> qf(c(0.05), df1=1, df2=6, lower.tail=FALSE)
[1] 5.987378
```

```
> Anova(Maintain1)
Anova Table (Type II tests)

Response: Maintenance
             Sum Sq Df F value    Pr(>F)
CarAge    374083333  1  79.072 0.0001127 ***
Residuals  28385417  6
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

註: 查表 $F_{0.05}(1,6)=5.9874$

因 79.072 > 5.987378 , 所以拒絕 H_0,即此迴歸模型具有解釋能力
由p- value (0.0001127 很小)可直接觀察 reject H_0.

THANKS

Q & A