R Tutorial 1.3 筆記 (基本知識與指令)

首字母大寫為word 自動，Ｒ內為小寫

* ＝、<-：Assign value
* 物件名稱規則：
  + 大小寫不可混用
  + 可使用標點符號
  + 數字不可為名稱開頭
* Print()：顯示某物件
* Ls()：查看現有物件
* Rm()：移除物件
* “ ”：引號中內容視為charater資料型態
* ^：次方
* Sqrt()：開根號
* Log()：log（log2()也可）
* Exp()：指數（anti log）
* Abs()：絕對值
* 如果指令沒打完，Ｒ會在下一行顯示＋
* 方向鍵上可以呼叫上一個指令
* ＃….：用於註記或提醒自己，Ｒ會呼略＃後的內容
* C()：向量

R Tutorial 1.4 筆記 （基本知識與指令）

* ：：create連續整數
* Seq(from=數列的第一個數字, to=數列的最後一個數字, by=間隔)：create數列
* Rep(數字或其他指令, times=次數)：重複數字的數列（ex: 1,1,1,1、character也可以）
* X[1]：提取物件Ｘ中的第一個分量

X[-1]：提取物件Ｘ中除了第一個分量以外的所有分量

X[1:3]：提取物件Ｘ中的第一個到第三個分量

X[c(1,5)]：提取物件Ｘ中的第一個和第五個分量

X[X<6]：提取物件Ｘ中小於6的分量

* Matrix(c(1,2,3,4,5,6), nrow=3, byrow=TRUE)：將123456照row的方式排列成3row的矩陣（byrow=FALSE則按column的順序排列）

1. 2
2. 4

5 6

* 物件名[r, c]：提取該矩陣第r row第c column的值（if r或c=空白，則提取所有row/column）

R Tutorial 1.5 筆記 （將excel導入Ｒ）

* 若要將excel data導入Ｒ，可將excel data儲存為**csv.**或**txt.**
* Read.csv(file.choose(), header=T)：導入儲存的**csv檔**
  + file.choose()：會呼叫出視窗讓我們可以透過視窗找檔案而不需要用key的把檔案位置告訴Ｒ
  + header=T：Ｔ＝TRUE，此指令告訴Ｒ這個data的第一行是不是變數名稱（FALSE：則代表此data從第一行開始就是data）
* read.table(file.choose()或file=檔案路徑, header=T, sep=”,”) ：導入儲存的**csv檔**或**txt檔**
  + sep”,”：告訴Ｒ此data以什麼做區隔，若改為”\t”則可導入txt檔
* read.delim(file.choose(), header=T)：導入儲存的**txt檔**

R Tutorial 1.5b 筆記 (將excel導入Ｒstudio)

* 可以直接透過Ｒstudio功能導入excel檔案
* 可以選擇要導入的工作表、導入的範圍、將特定內容設定為N\A、設定該column的資料型態等
* 右下角會呈現此次import的程式碼，可以直接複製貼上用於下一次的導入

R Tutorial 1.6 筆記 (將data導出Ｒ)

* Write.table(data物件名稱, file=導出data檔案名稱, sep=”,”)
  + 若要清除row name，可在指令中加入row.names=F
  + 若要將導出檔案儲存在別的資料夾，file=要寫明該資料夾的位置與檔案名
* Write.csv()：不需要打sep＝

R Tutorial 1.7 筆記 (將data導入、編輯)

* Dim()：查看data有幾row,column
* Head()：顯示前6row data
* Tail()：顯示後6row data
* Names()：查看data.frame的header

R Tutorial 1.8 筆記 (處理data)

* Data.frame名稱＄column名稱：extract出該行的data(ex.影片中的LungCapData$Age)
* Mean()：平均數
* Attach()：將data attach至Ｒ的記憶，之後直接數入column nameＲ就會認得是指此data的那一行的資料，而不需要重複的使用＄。缺點則是需佔用Ｒ的記憶空間，且容易被overwritten
* Detach()：將data自Ｒ的記憶中抹除
* Levels()：顯示factor（一種資料型態）中的level

(ex: data中其中一行名稱為gender,且為factor>>>顯示“female”,”male”)

* Summary()：基本的數據指標比如平均值
* 許多data coding時會使用0跟1代表yes與no,此時需使用as.factor將其資料型態轉換，使用summary或其他指令時才會出現我們想要的結果

R Tutorial 1.9 筆記 (創造子集)

此段落舉例的指令是在已經執行attach(LungCapData)的情況下

* ==：單個=是assign value的意思，雙等號則是真正等於的意思，等於的factor須加雙引號

（ex. If age ==”female，注意，此用法只有在age的資料型態為factor時邏輯才成立）

* 從data中創造子集的方式：

影片Ex: MaleUnder15 <- LungCapData[Gender==”male”& Age<15,]

* + []內的意思是：只要性別為male且年齡小於15的row，column全都要

R Tutorial 1.10 筆記 (logic)

此段落舉例的指令是在已經執行attach(LungCapData)的情況下

* 創造一個邏輯變數（logical variable）
  + Temp <- Age>15

Temp的值：TRUE, FALSE, FALSE…..

* + Temp <- as.numeric( Age>15)

Temp的值：1, 0, 0…..（以數字0與1代表Ｔ或Ｆ）

* + FemSmoke <- Gender==”female”& Smoke==”yes”
* Cbind(data,data)：將兩個data或向量以column的方向結合
  + New <- cbind( LungCapData, FemSmoke)
* Rbind(data,data)：將兩個data或向量以row的方向結合
* Ls()：目前的所有variable

R Tutorial 1.11 筆記 (wd)

* Setwd(wd路徑)：設定wd,路徑可以用～省略（ex: ~/desktop/R課程）
* Save.image(檔名或路徑)：儲存workspace（若是至目前wd則檔名即可）
* Load(檔名或路徑或file.choose())：開啟儲存之檔案（若是自目前wd則檔名即可）
* 上述操作皆可使用menu完成

檔名會為Rdata

R Tutorial 1.12 筆記 (writing scripts)

* Cmd+enter=run

R Tutorial 1.13 筆記 (packages)

* Install.packages()
* Library(套件名)：才可使用該套件內指令（結束Ｒsession時library會消失，因此每次都要執行一次此指令才可使用套件內指令）

R Tutorial 1.14 筆記 (客製化Ｒ介面)

* 影片option位置為舊版，新版稱為global option

R Tutorial 1.15 筆記 (apply functions)

* Function會較forloop更有效率、需要的程式碼也較少（較少機會出現error）
* Apply( X=資料名稱, MARGIN=用row(1)還是column(2)計算, FUN=要引用的function, na.rm=T/F 是否移除NA)

可使用的function比如

* + Mean 平均
  + Max 最大值
  + Quantile 前x比例的平均, probs=c(想計算的比例)
  + Plot 畫圖, type=”1（折線圖）”, main=”圖表標題”, xlab=”x軸標題）, ylab=”y軸標題“
  + Sum
  + summary
* colMeans(資料名稱, na.rm=T/F)
* rowSums(資料名稱, na.rm=T/F)
* plot() 畫圖
* point() 在圖上加有顏色的點
* 也可以寫自己的function

R Tutorial 1.16 筆記 (tApply functions)

* apply function to a **subsets** of variable or factor
* tapply( X=想計算的data名稱或子集名稱, INDEX= 分類的依據變數（比如是否吸菸）, FUN=function, 其他條件, simplify=T/F(R是否要簡化結果))
  + INDEX也可使用多個分類依據（eg. list( Smoke, Gender)）
* By()作用與tapply類似，只是回傳的結果為向量（vector format）