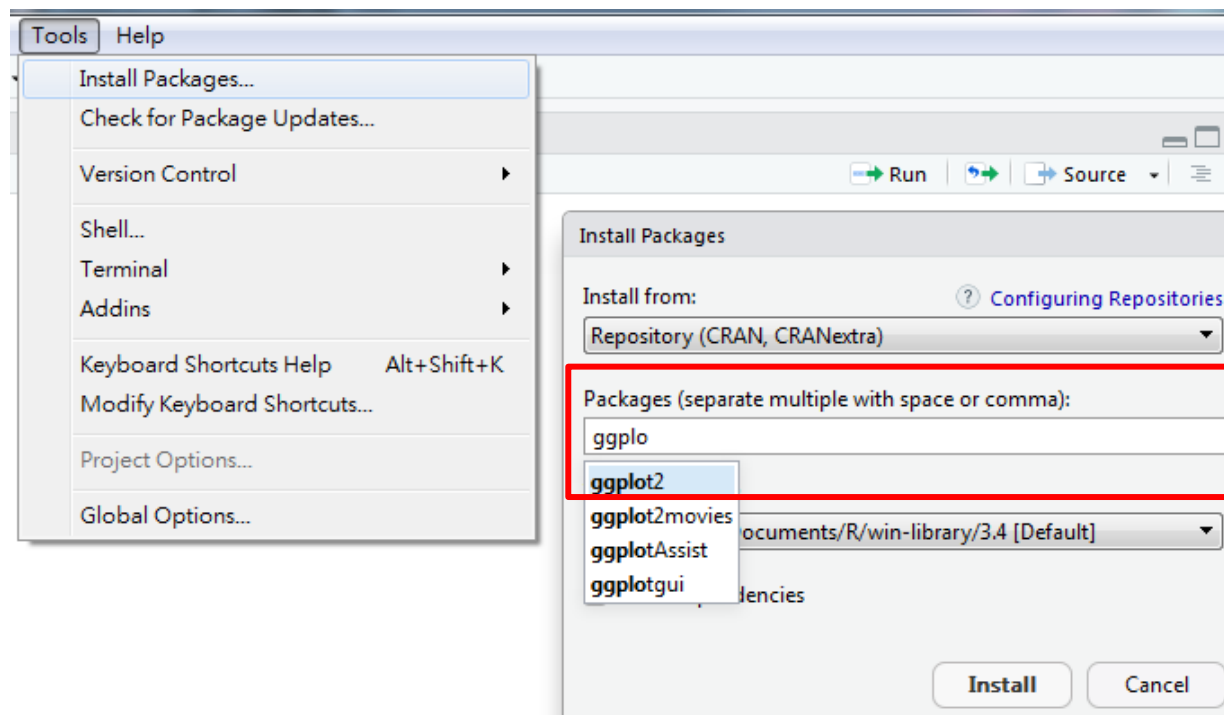# Chapter 8

BASIC DATA ANALYSIS (3)

回顧：在 RStudio 安裝套件

1. Tool -> Install Packages -> ggplot2

回顧：在 RStudio 安裝套件

2. 在 Console 會看到 相關的套件會一併安裝

```
> install.packages("ggplot2")
Installing package into 'C:/Users/user/Documents/R/win-library/3.4'
(as 'lib' is unspecified)
also installing the dependencies 'colorspace', 'assertthat', 'utf8', 'Rcpp', 'RColorBrewer', 'dichromat', 'munsell', 'labeling', 'R6
', 'viridisLite', 'cli', 'crayon', 'pillar', 'rlang', 'digest', 'gtable', 'plyr', 'reshape2', 'scales', 'tibble', 'lazyeval'

trying URL 'https://cran.rstudio.com/bin/windows/contrib/3.4/colorspace_1.3-2.zip'
Content type 'application/zip' length 447069 bytes (436 KB)
downloaded 436 KB

package 'colorspace' successfully unpacked and MD5 sums checked
package 'assertthat' successfully unpacked and MD5 sums checked
package 'utf8' successfully unpacked and MD5 sums checked
package 'Rcpp' successfully unpacked and MD5 sums checked
package 'RColorBrewer' successfully unpacked and MD5 sums checked
package 'dichromat' successfully unpacked and MD5 sums checked
package 'munsell' successfully unpacked and MD5 sums checked
package 'labeling' successfully unpacked and MD5 sums checked
package 'R6' successfully unpacked and MD5 sums checked
package 'viridisLite' successfully unpacked and MD5 sums checked
package 'cli' successfully unpacked and MD5 sums checked
package 'crayon' successfully unpacked and MD5 sums checked
package 'pillar' successfully unpacked and MD5 sums checked
package 'rlang' successfully unpacked and MD5 sums checked
package 'digest' successfully unpacked and MD5 sums checked
package 'gtable' successfully unpacked and MD5 sums checked
package 'plyr' successfully unpacked and MD5 sums checked
package 'reshape2' successfully unpacked and MD5 sums checked
package 'scales' successfully unpacked and MD5 sums checked
package 'tibble' successfully unpacked and MD5 sums checked
package 'lazyeval' successfully unpacked and MD5 sums checked
package 'ggplot2' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
        C:\Users\user\AppData\Local\Temp\RtmpC4QnI1\downloaded_packages
```
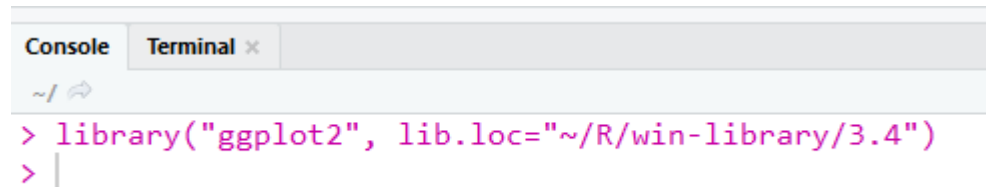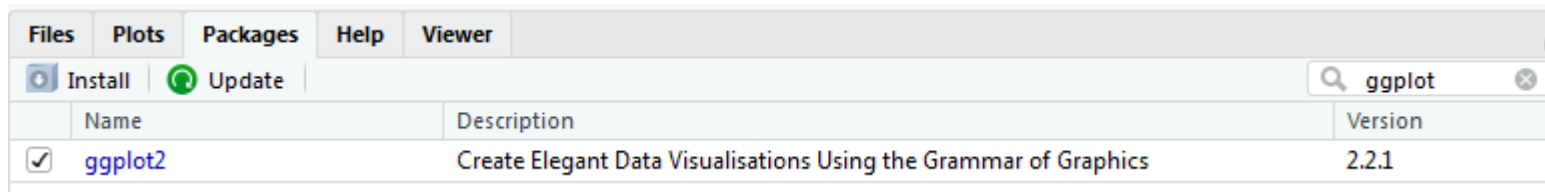
回顧：在 RStudio 安裝套件

3. 載入 ggplot2

ggplot2

ggplot2
Grammar of graphic ggplot(…) + geom_" 繪圖方式 "

```
1  library(ggplot2)
2
3  car_data <- cars
4  str(car_data)
5  head(car_data)
6
7  ggplot(car_data, aes(x=speed, y= dist)) + geom_point()
```
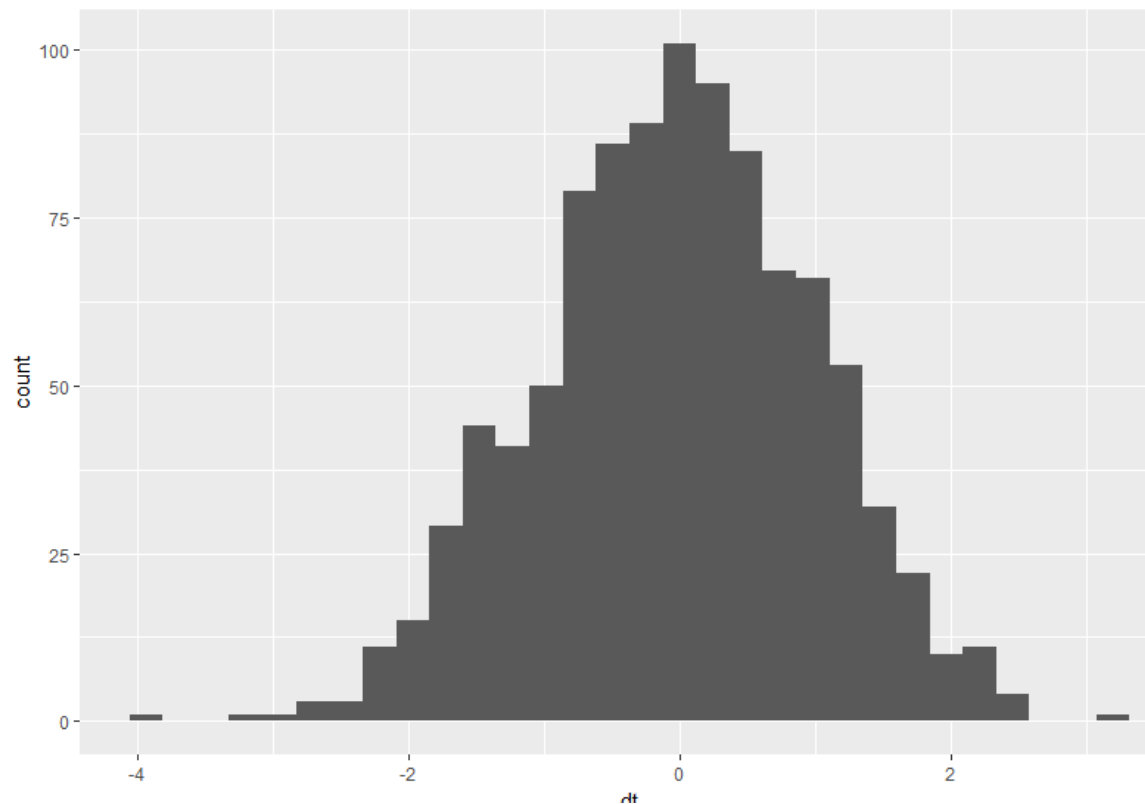
```
> car_data <- cars
> str(car_data)
'data.frame':   50 obs. of  2 variables:
 $ speed: num  4 4 7 7 8 9 10 10 10 11 ...
 $ dist : num  2 10 4 22 16 10 18 26 34 17 ...
> head(car_data)
  speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
>
> ggplot(car_data, aes(x=speed, y= dist)) + geom_point()
>
```
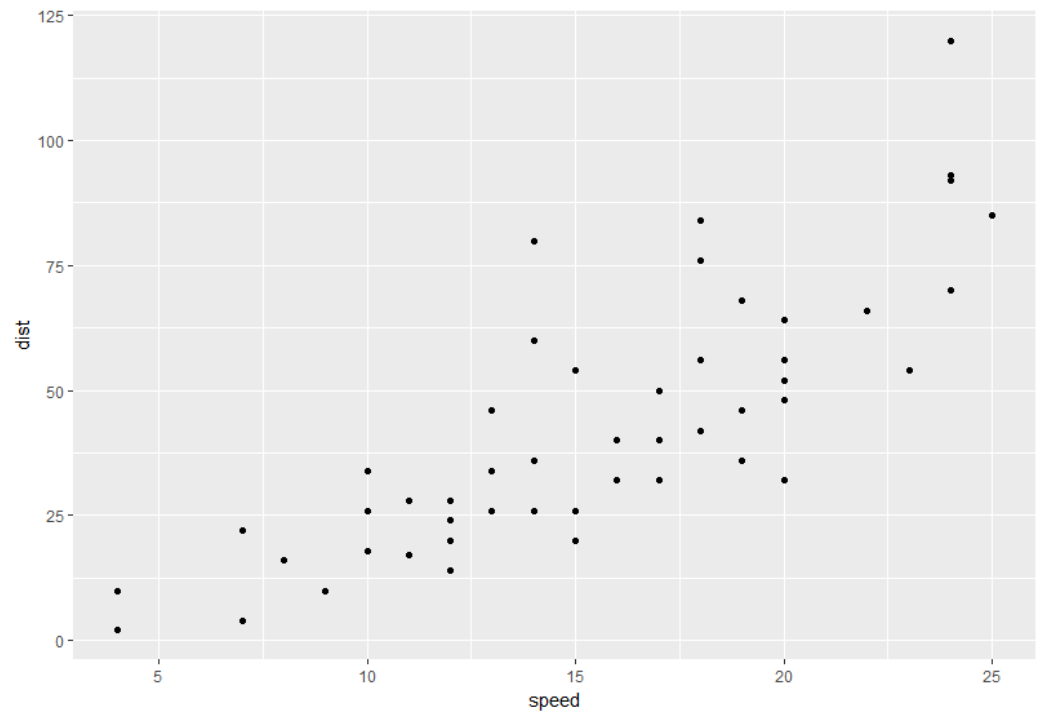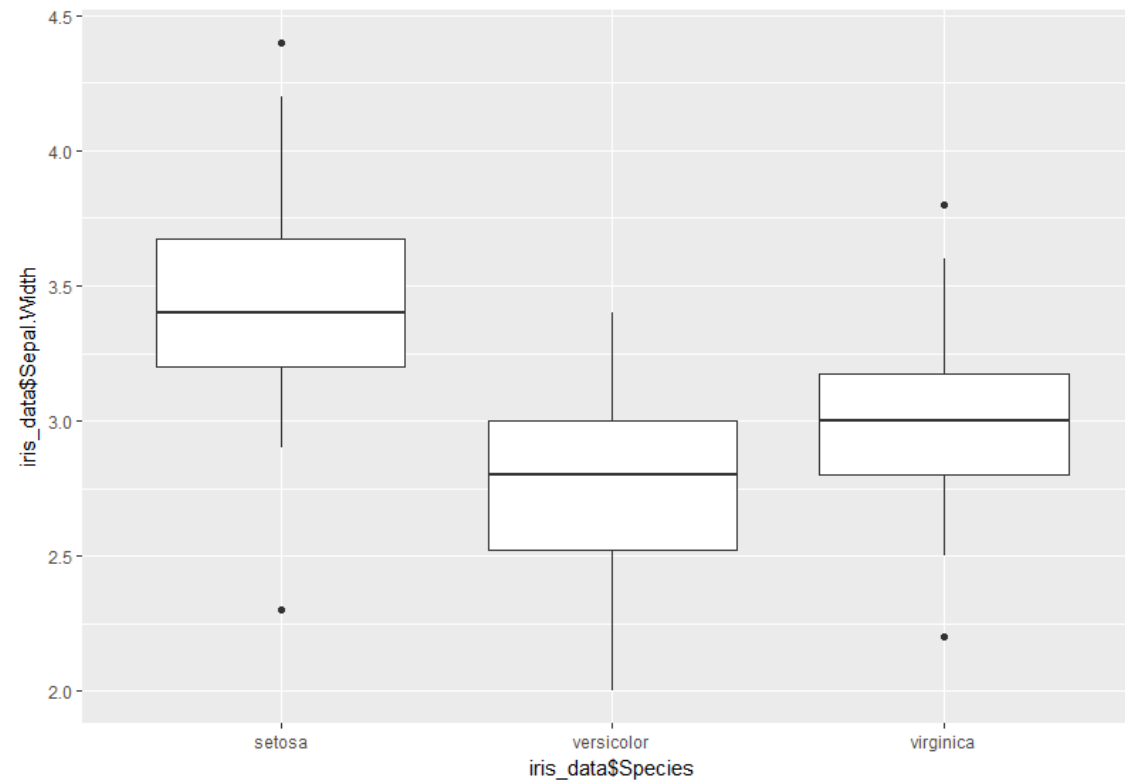
geom_histogram 直方圖

bindwitdh　用來設定分箱數
bins()

```
1  libaray(ggplot2)
2
3  dt <- rnorm(1000)
4
5  hist <- data.frame(dt)
6
7  ggplot(hist,aes(x=dt)) + geom_histogram(bindwidth =0.2)
```
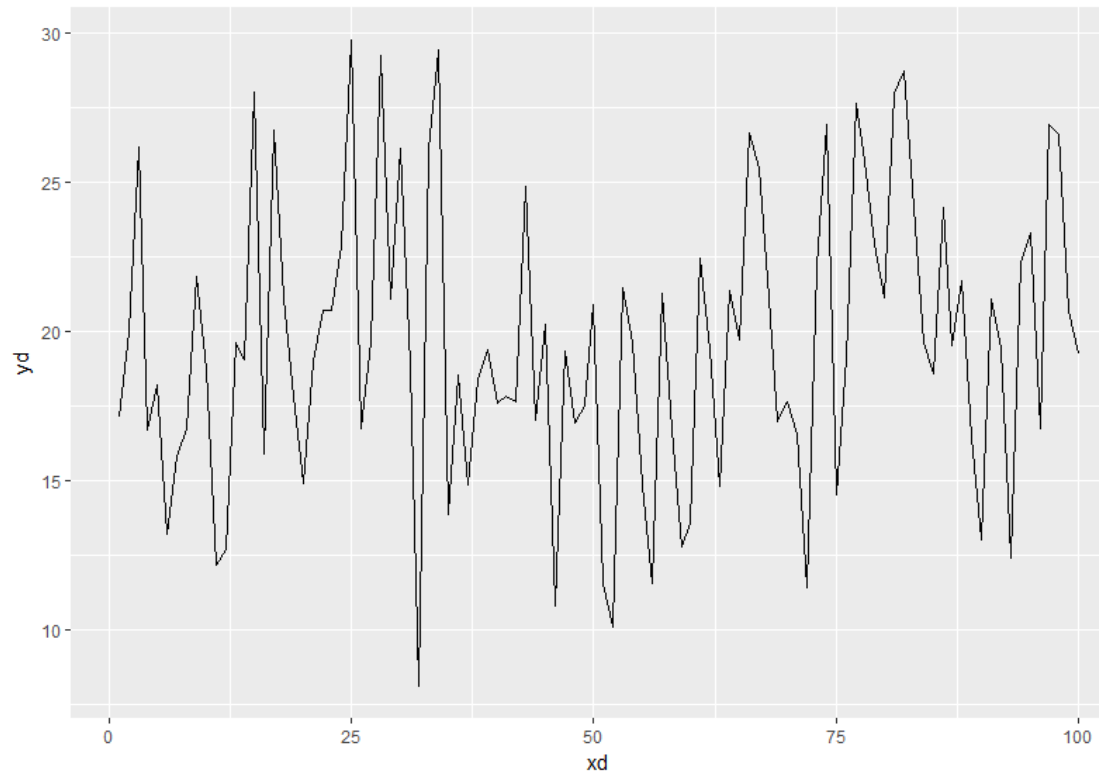
geom_boxplot() 盒鬚圖

```
1  library(ggplot2)
2
3  iris_data <- iris
4
5  ggplot(iris, aes(x= iris_data$Species,
6                   y= iris_data$Sepal.Width))+geom_boxplot()
```
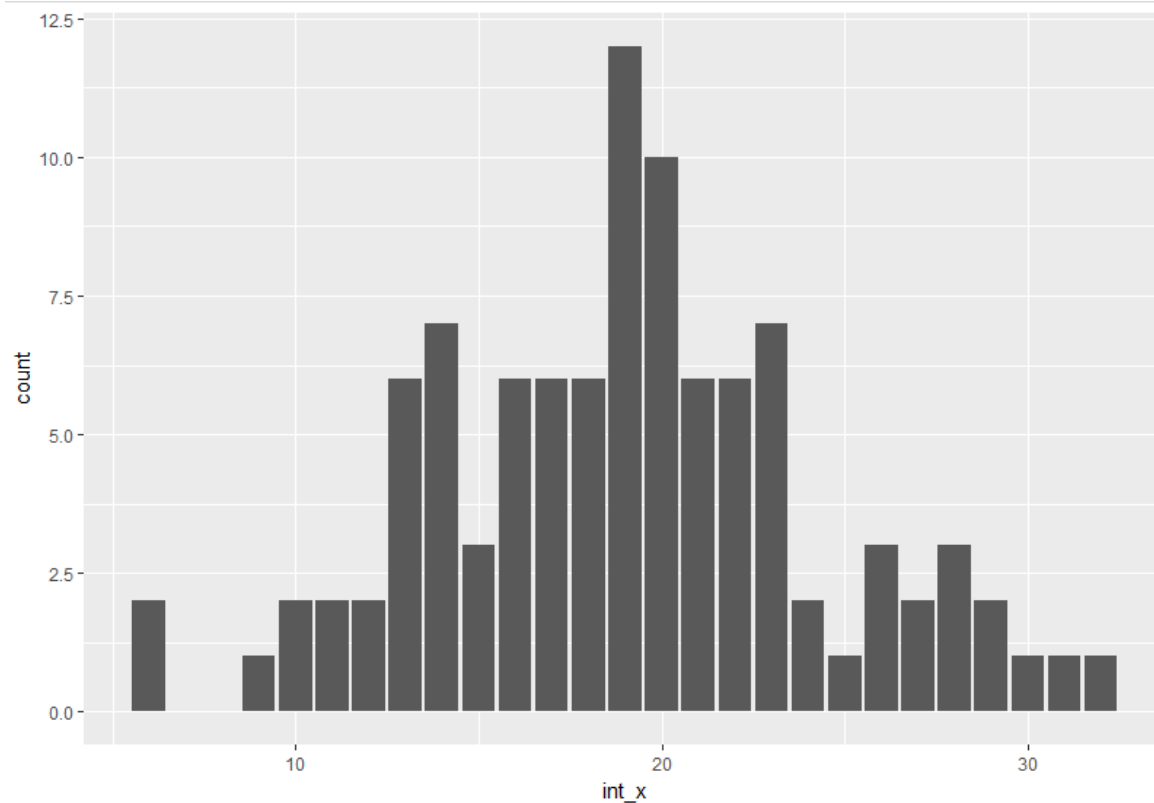
geom_line() 線圖

```
1  xd <-seq(1:100)
2  yd <- rnorm(100,20,5)
3
4  df <- data.frame(x=xd, y=yd)
5  ggplot(df, aes(x=xd, y=yd))+geom_line()
```
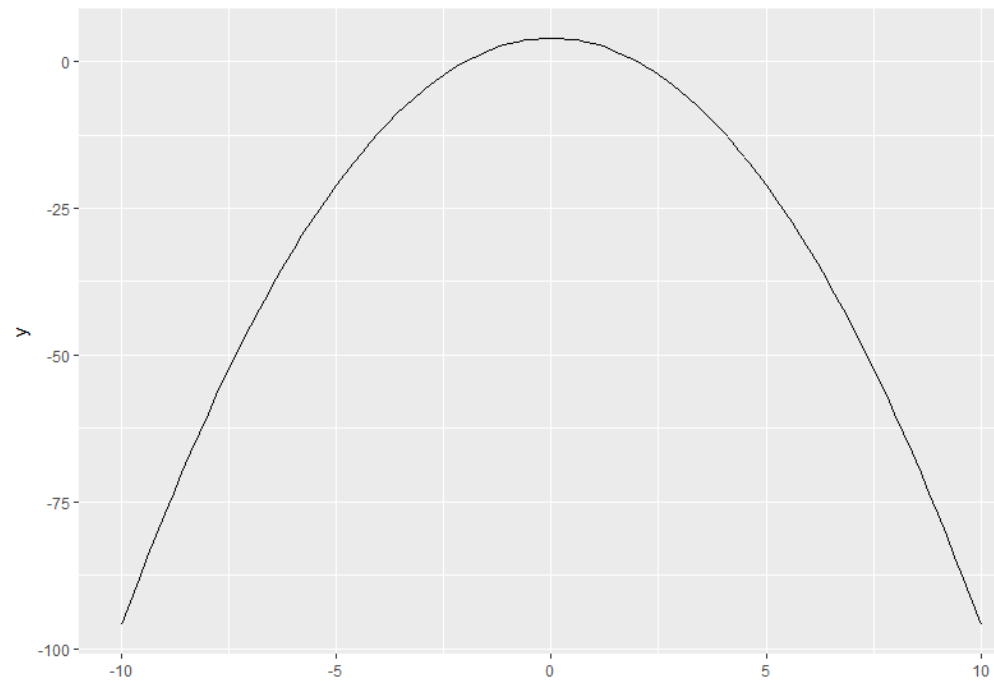
geom_bar() 長條圖

```
1  x <- rnorm(100,20,5)
2  int_x <-as.integer(x)
3  tb_x <- data.frame(int_x = int_x)
4
5  ggplot(tb_x, aes(x=int_x)) +geom_bar()
```

stat_function() 曲線圖

```r
1  library(ggplot2)
2
3  para_curve <- function(x)
4  {
5      return( -(x)^2+4)
6  }
7
8  rng <- data.frame(x= c(-10,10))
9
10 ggplot(rng,
11    aes(x=x))+stat_function(fun=para_curve, geom="line")
```

自訂圖形元素

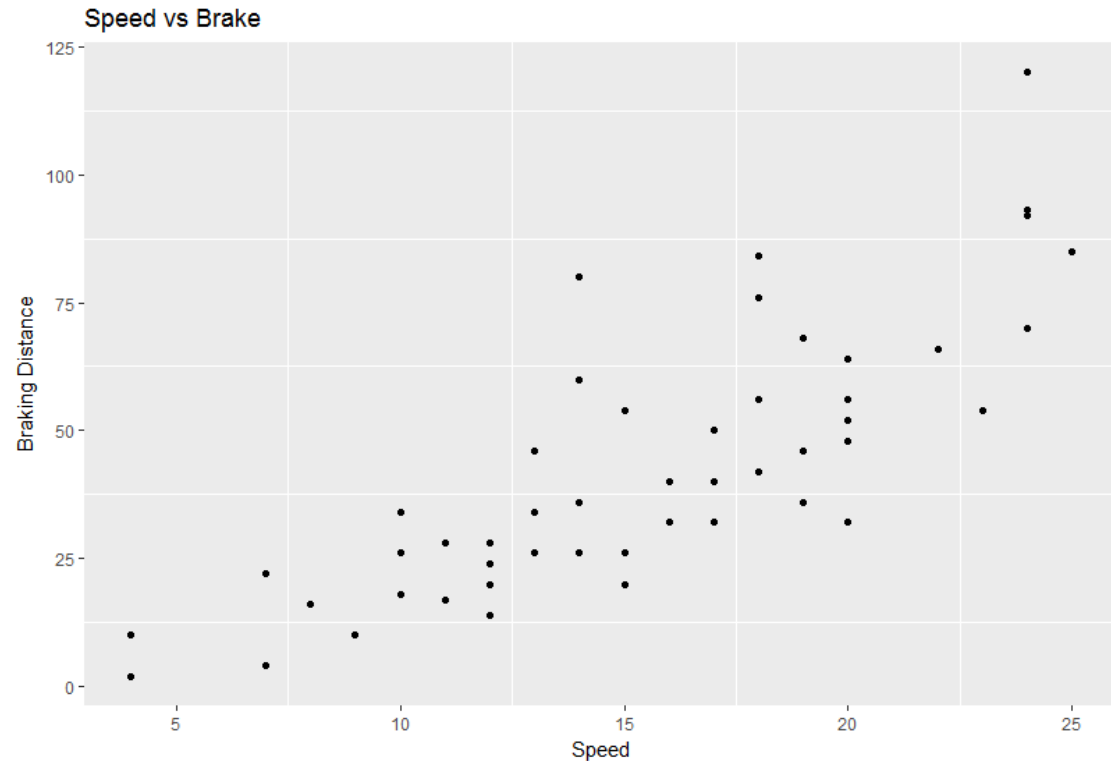| 參數	說明
ggtitle	標題
xlab	X 軸標題
ylab	Y 軸標題

```
1  library(ggplot2)
2  car_data <- cars
3  str(car_data)
4  head(car_data)
5
6  ggplot(car_data,
7          aes(x=speed, y=dist)) + geom_point()+
8    ggtitle("Speed vs Brake")  +
9    xlab("Speed") + ylab("Braking Distance")
```
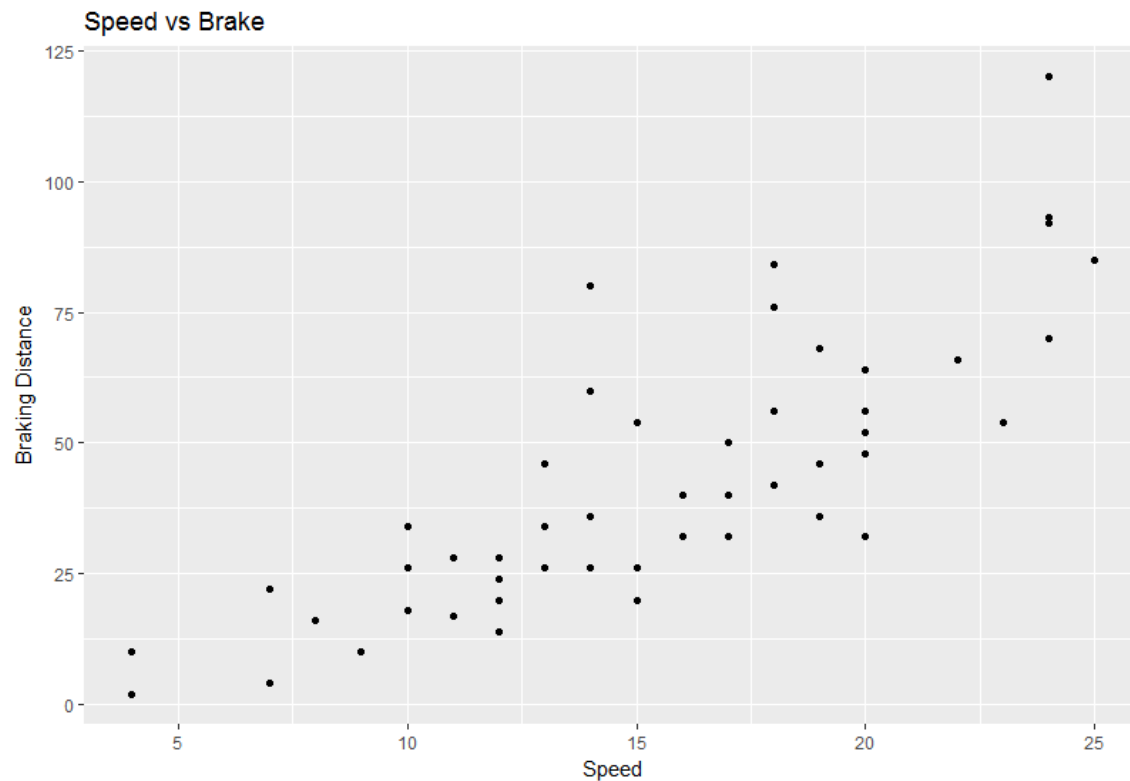
自訂圖形元素

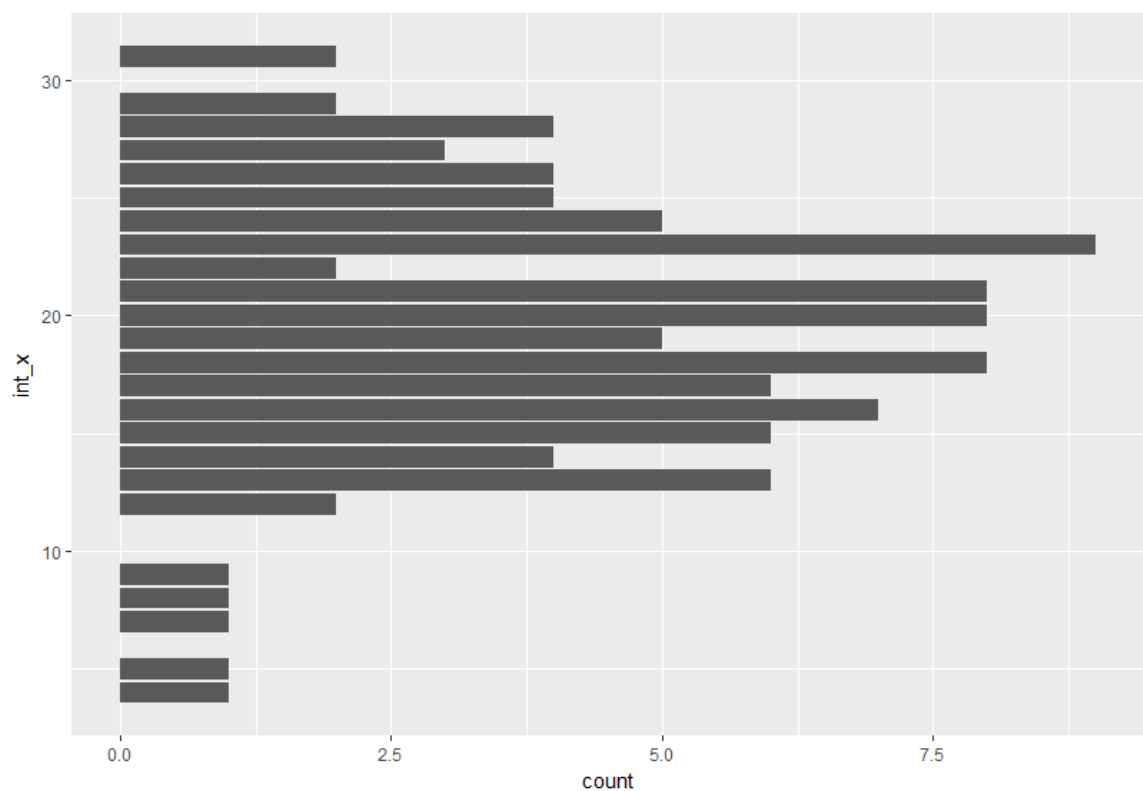| 參數	說明
Panel.grid.major	關掉主要格線
Panel.grid.minor	關掉次要格線
Panel.grid.major.x	關掉 x 軸主要格線
Panel.grid.minor.x	關掉 x 軸次要格線

```
1  library(ggplot2)
2  car_data <- cars
3  str(car_data)
4  head(car_data)
5
6  ggplot(car_data,
7        aes(x=speed, y=dist)) + geom_point()+
8    ggtitle("Speed vs Brake")  +
9    xlab("Speed") + ylab("Braking Distance") +
10   theme(panel.grid.major.x = element_blank(),
11         panel.grid.major.y = element_blank())
```


Speed vs Brake

自訂圖形元素

參數	說明
coord_flip()	水平

```
1  x <- rnorm(100,20,5)
2  int_x <-as.integer(x)
3  tb_x <- data.frame(int_x = int_x)
4
5  ggplot(tb_x, aes(x=int_x)) +
6  geom_bar()+ coord_flip()
```
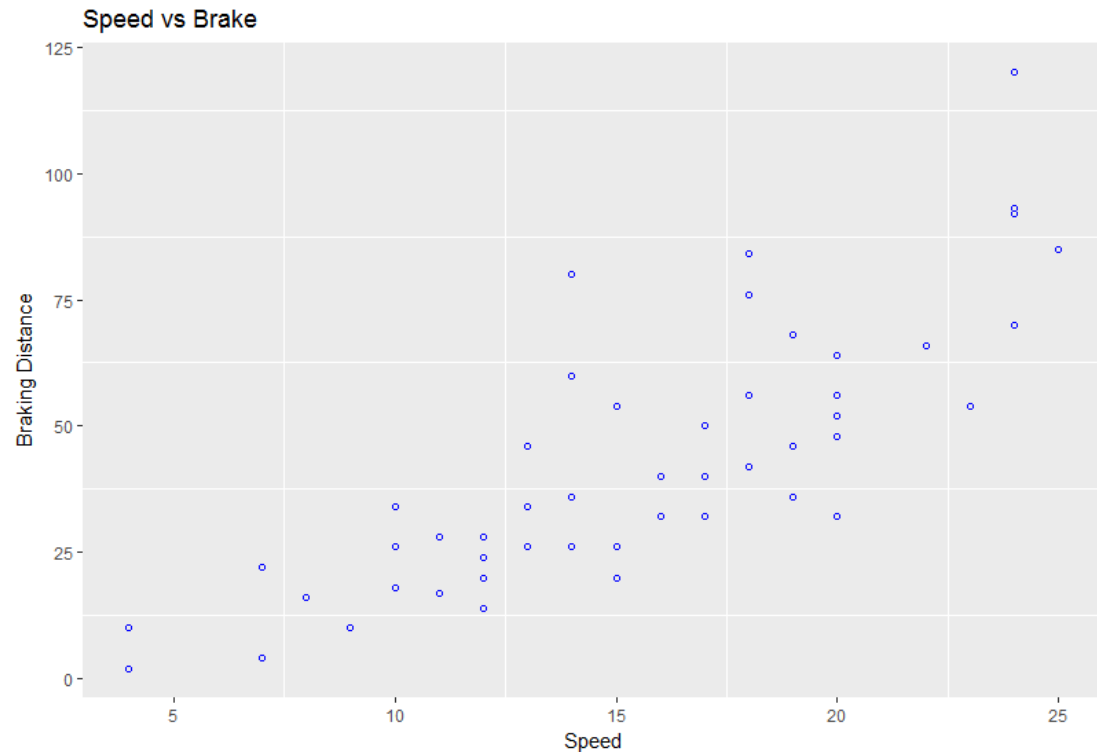
自訂圖形元素

| geom_point() 參數設定	
參數	說明
shape()	水平
colour	顏色

```
1  library(ggplot2)
2  car_data <- cars
3  str(car_data)
4  head(car_data)
5
6  ggplot(car_data,
7          aes(x=speed, y=dist)) +
8    geom_point(shape=1, colour="blue")+
9    ggtitle("Speed vs Brake")   +
10   xlab("Speed") + ylab("Braking Distance") +
11   theme(panel.grid.major.x = element_blank(),
12         panel.grid.major.y = element_blank())
```
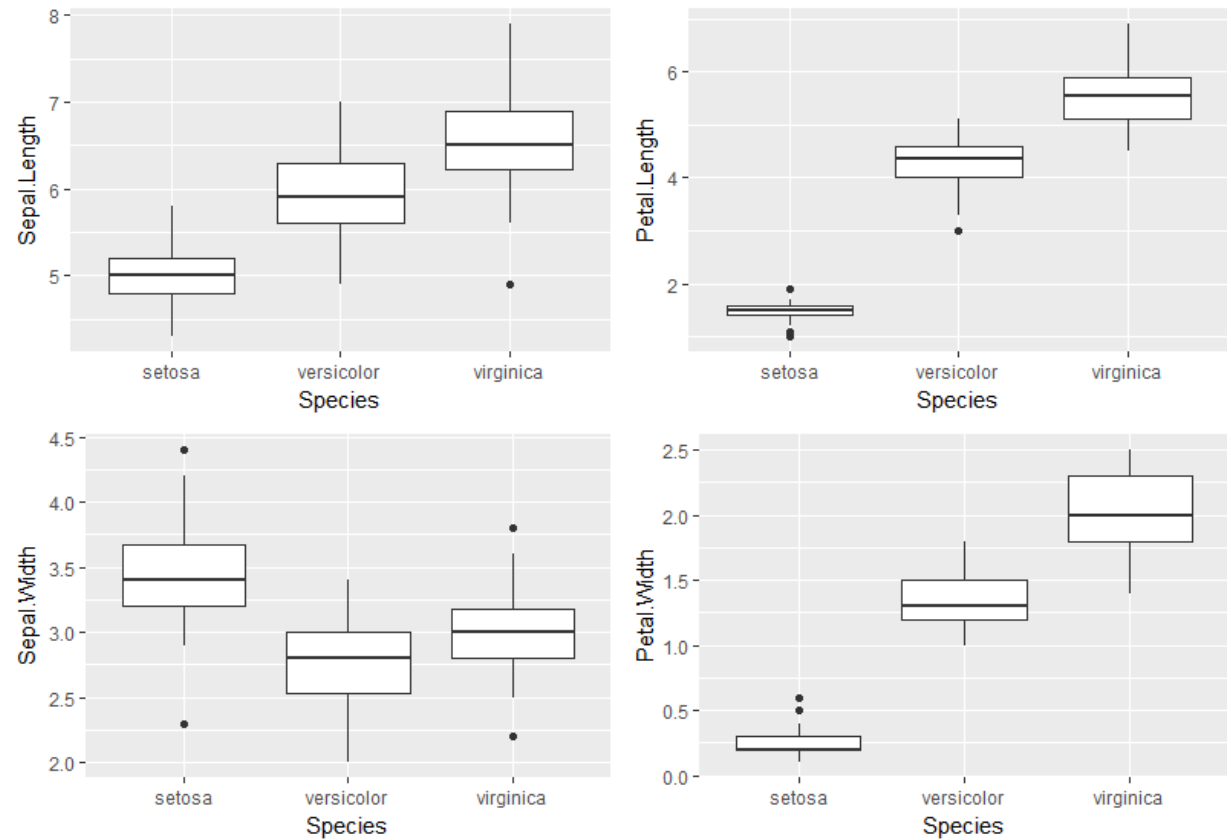

Speed vs Brake

繪製多個圖形

參數	說明
grid.arrange	切割區域 (m * n)

```r
1  library(ggplot2)
2  library(gridExtra)
3  x1  <- seq(1:100)
4  y1  <- rnorm(100,20,5)
5  df1 <- data.frame(x1,y1)
6  x2 <- seq(1:100)
7  y2 <- rnorm(100,20,5)
8  df2 <- data.frame(x2,y2)
9
10 f1 <- ggplot(df1, aes(x=x1,y=y1))+
11   geom_point()
12 f2 <- ggplot(df2, aes(x=x2,y=y2))+
13   geom_point()
14
15 grid.arrange(f1,f2, nrow =1, ncol=2)
```
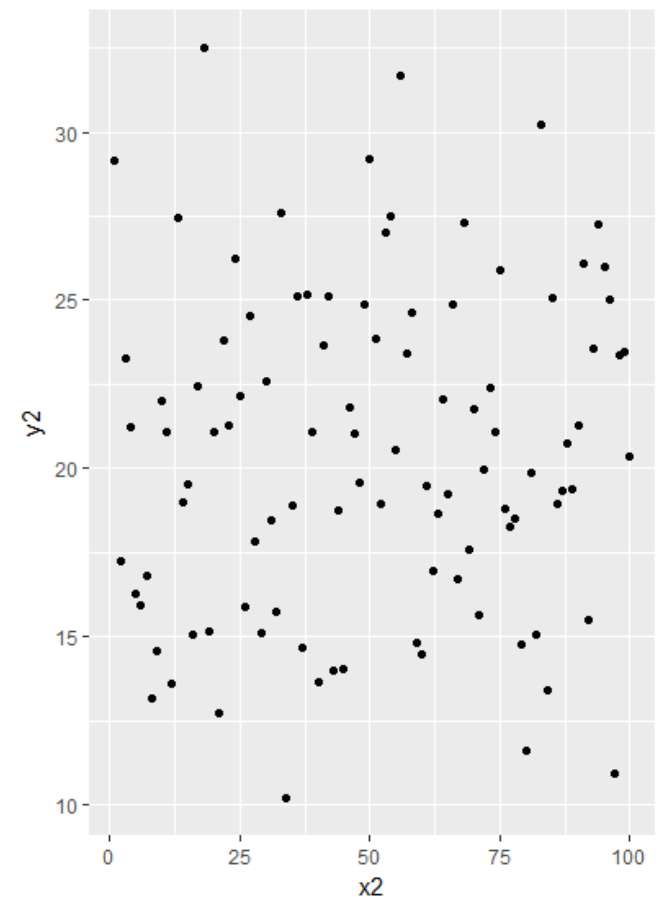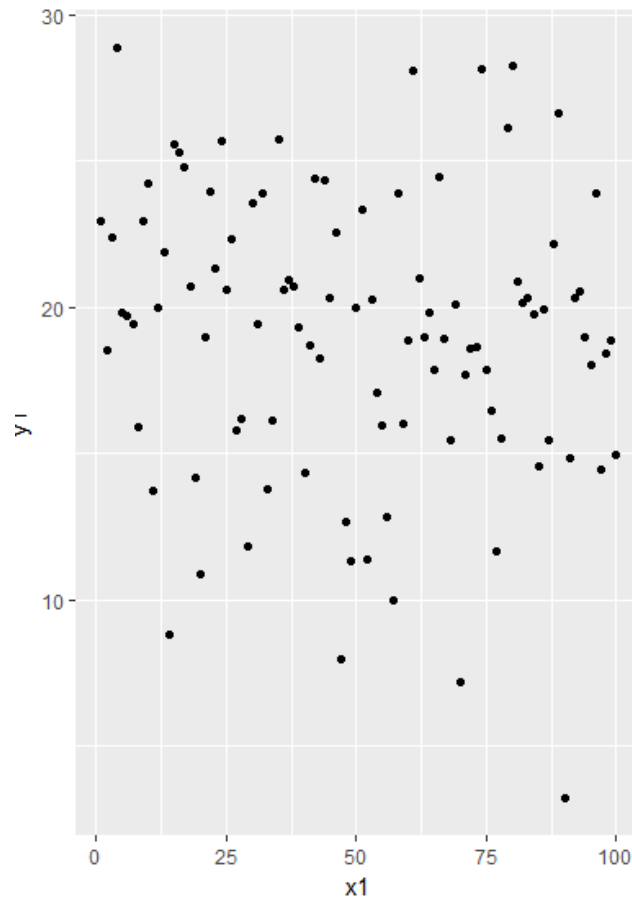
隨堂練習 1

1. 讀取 IRIS data
2. 畫出如以下圖形

Any Questions !?