

Unit 03: IRIS - 鳶尾花 資料集

連 豐 力

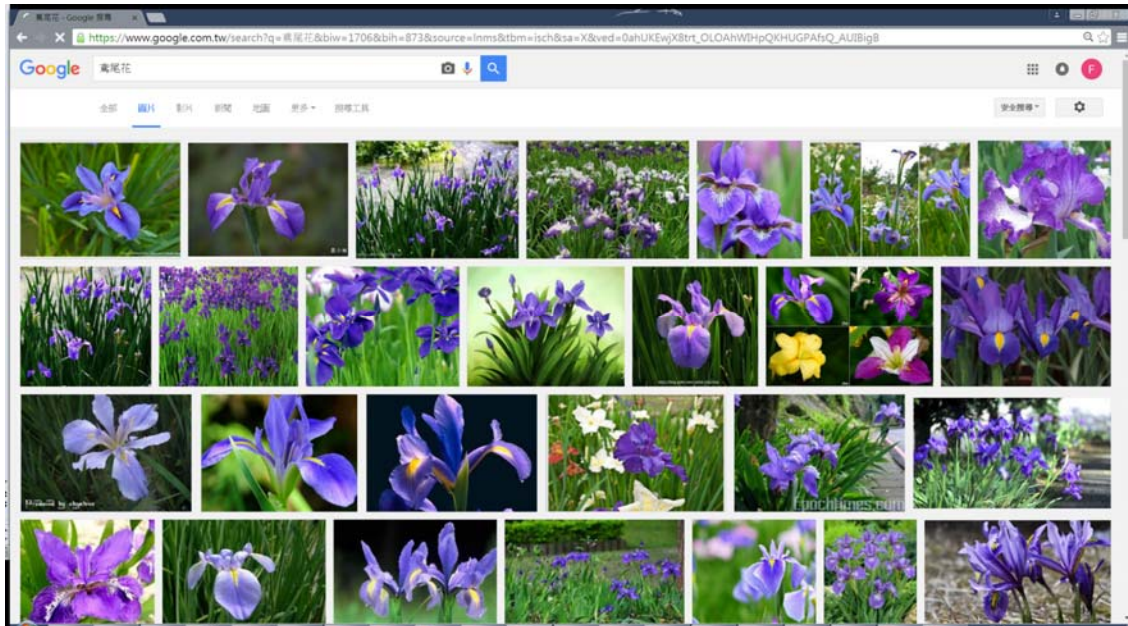
臺大電機系

Sep 2016 - Jan 2017

大綱

計算機程式設計 - 2016F
Chap 03: 鳶尾花 IRIS
Feng-Li Lian @ NTU-EE

- 以 鳶尾花 - IRIS 數據為例
- 數據所在位置與數據的內容
- 初步分析數據
- 繪製圖形 -
 - 圖形位置安排
 - 一維圖：
 - 直方圖，盒鬚圖，莖葉圖，長條記錄圖，圓餅圖，機率分布圖，經驗累積分布圖，常態機率圖
 - 多維圖：
 - 散點圖，散點直方核密度，多重分布，三維散點圖



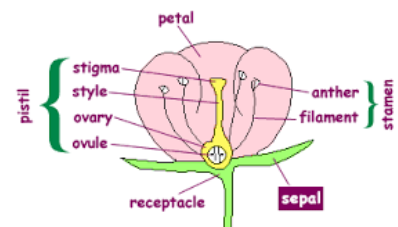
- 3

鳶尾花 - IRIS

- 鳶尾花 (iris) 資料集
 - 非常著名的生物資訊資料集之一
 - 取自美國加州大學歐文分校的機械學習資料庫

- 資料的筆數為150筆，

- 共有五個欄位：
 1. 花萼長度 (Sepal Length) : 計算單位是公分。
 2. 花萼寬度 (Sepal Width) : 計算單位是公分。
 3. 花瓣長度 (Petal Length) : 計算單位是公分。
 4. 花瓣寬度 (Petal Width) : 計算單位是公分。
 5. 類別 (Class) : 可分為 **Setosa** , **Versicolor** 和 **Virginica** 三個品種。



- 4

```
Console - / ?
> iris
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1 4.9 3.0 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
7 4.6 3.4 1.4 0.3 setosa
8 5.0 3.4 1.5 0.2 setosa
9 4.4 2.9 1.4 0.2 setosa
10 4.9 3.1 1.5 0.1 setosa
11 5.4 3.7 1.5 0.2 setosa
12 4.8 3.4 1.6 0.2 setosa
13 4.8 3.0 1.4 0.1 setosa
14 4.3 3.0 1.1 0.1 setosa
15 5.8 4.0 1.2 0.2 setosa
16 5.7 4.4 1.5 0.4 setosa
17 5.4 3.9 1.3 0.4 setosa
18 5.1 3.5 1.4 0.3 setosa
19 5.7 3.8 1.7 0.3 setosa
20 5.1 3.8 1.5 0.3 setosa
21 5.4 3.4 1.7 0.2 setosa
22 5.1 3.7 1.5 0.4 setosa
23 4.6 3.6 1.0 0.2 setosa
24 5.1 3.3 1.7 0.5 setosa
25 4.8 3.4 1.9 0.2 setosa
26 5.0 3.0 1.6 0.2 setosa
27 5.0 3.4 1.6 0.4 setosa
28 5.2 3.5 1.5 0.2 setosa
29 5.2 3.4 1.4 0.2 setosa
30 4.7 3.2 1.6 0.2 setosa
31 4.8 3.1 1.6 0.2 setosa
32 5.4 3.4 1.5 0.4 setosa
33 5.2 4.1 1.5 0.1 setosa
34 5.5 4.2 1.4 0.2 setosa
35 4.9 3.1 1.5 0.2 setosa
36 5.0 3.2 1.2 0.2 setosa
37 5.5 3.5 1.3 0.2 setosa
38 4.9 3.6 1.4 0.1 setosa
39 4.4 3.0 1.3 0.2 setosa
40 5.1 3.4 1.5 0.2 setosa
41 5.0 3.5 1.3 0.3 setosa
42 4.5 2.3 1.3 0.3 setosa
43 4.4 3.2 1.3 0.2 setosa
44 5.0 3.5 1.6 0.6 setosa
45 5.1 3.8 1.9 0.4 setosa
46 4.8 3.0 1.4 0.3 setosa
47 5.1 3.8 1.6 0.2 setosa
48 4.6 3.2 1.4 0.2 setosa
49 5.3 3.7 1.5 0.2 setosa
50 5.0 3.3 1.4 0.2 setosa
51 7.0 3.2 4.7 1.4 versicolor
52 6.4 3.2 4.5 1.5 versicolor
53 6.9 3.1 4.9 1.5 versicolor
54 5.5 2.3 4.0 1.3 versicolor
55 6.5 2.8 4.6 1.5 versicolor
56 5.7 2.8 4.5 1.3 versicolor
57 6.3 3.3 4.7 1.6 versicolor
58 4.9 2.4 4.3 1.0 versicolor
59 6.6 2.9 4.6 1.3 versicolor
60 5.2 2.7 3.9 1.4 versicolor
61 5.0 2.0 3.5 1.0 versicolor
```

```
91 5.5 2.6 4.4 1.2 versicolor
92 6.1 3.0 4.6 1.4 versicolor
93 5.8 2.6 4.0 1.2 versicolor
94 5.0 2.3 3.3 1.0 versicolor
95 5.6 2.7 4.2 1.3 versicolor
96 5.7 3.0 4.2 1.2 versicolor
97 5.7 2.9 4.2 1.3 versicolor
98 6.2 2.9 4.3 1.3 versicolor
99 5.1 2.5 3.0 1.1 versicolor
100 5.7 2.8 4.1 1.3 versicolor
101 6.3 3.3 6.0 2.5 virginica
102 5.8 2.7 5.1 1.9 virginica
103 7.1 3.0 5.9 2.1 virginica
104 6.3 2.9 5.6 1.8 virginica
105 6.5 3.0 5.8 2.2 virginica
106 7.6 3.0 6.6 2.1 virginica
107 4.9 2.5 4.5 1.7 virginica
108 7.3 2.9 6.3 1.8 virginica
109 6.7 2.5 5.8 1.8 virginica
110 7.2 3.6 6.1 2.5 virginica
111 6.5 3.2 5.1 2.0 virginica
112 6.4 2.7 5.3 1.9 virginica
113 6.8 3.0 5.5 2.1 virginica
114 5.7 2.5 5.0 2.0 virginica
115 5.8 2.8 5.1 2.4 virginica
116 6.4 3.2 5.3 2.3 virginica
117 6.5 3.0 5.5 1.8 virginica
118 7.7 3.8 6.7 2.2 virginica
119 7.7 2.6 6.9 2.3 virginica
120 6.0 2.2 5.0 1.5 virginica
121 6.9 3.2 5.7 2.3 virginica
122 5.6 2.8 4.9 2.0 virginica
123 7.7 2.8 6.7 2.0 virginica
124 6.3 2.7 4.9 1.8 virginica
125 6.7 3.3 5.7 2.1 virginica
126 7.2 3.2 6.0 1.8 virginica
127 6.2 2.8 4.8 1.8 virginica
128 6.1 3.0 4.9 1.8 virginica
129 6.4 2.8 5.6 2.1 virginica
130 7.2 3.0 5.8 1.6 virginica
131 7.4 2.8 6.1 1.9 virginica
132 7.9 3.8 6.4 2.0 virginica
133 6.4 2.8 5.6 2.2 virginica
134 6.3 2.8 5.1 1.5 virginica
135 6.1 2.6 5.6 1.4 virginica
136 7.7 3.0 6.1 2.3 virginica
137 6.3 3.4 5.6 2.4 virginica
138 6.4 3.1 5.5 1.8 virginica
139 6.0 3.0 4.8 1.8 virginica
140 6.9 3.1 5.4 2.1 virginica
141 6.7 3.1 5.6 2.4 virginica
142 6.9 3.1 5.1 2.3 virginica
143 5.8 2.7 5.1 1.9 virginica
144 6.8 3.2 5.9 2.3 virginica
145 6.7 3.3 5.7 2.5 virginica
146 6.7 3.0 5.2 2.3 virginica
147 6.3 2.5 5.0 1.9 virginica
148 6.5 3.0 5.2 2.0 virginica
149 6.2 3.4 5.4 2.3 virginica
150 5.9 3.0 5.1 1.8 virginica
```

- 5

數據所在位置與數據的內容

- `iris[i, j]`
 - `iris[1, 1]`
 - `iris[1, 2]`
 - `iris[1, 3]`
 - `iris[1, 4]`
 - `iris[1, 5]`
 - `iris[2, 1]`
 - `iris[2, 2]`
 - `iris[1,]`
 - `iris[2,]`
 - `iris[3,]`
 - `iris[, 1]`
 - `iris[, 2]`
 - `iris[, 3]`

- 6

- `mydata <- iris`
- `mydata`
- `mydata[i, j]`
 - `mydata[1, 1]`
 - `mydata[3,]`

初步分析數據

- `mydata <- iris`
- `mydata`
- `str(mydata)`

```
> str( mydata )
'data.frame': 150 obs. of 5 variables:
 $ Sepal.Length: num  5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num  3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
```

- `summary(mydata)`

```
> summary( mydata )
 Sepal.Length      Sepal.Width      Petal.Length      Petal.Width      Species
Min.   :4.300   Min.   :2.000   Min.   :1.000   Min.   :0.100   setosa   :50
1st Qu.:5.100   1st Qu.:2.800   1st Qu.:1.600   1st Qu.:0.300   versicolor:50
Median :5.800   Median :3.000   Median :4.350   Median :1.300   virginica :50
Mean   :5.843   Mean   :3.057   Mean   :3.758   Mean   :1.199
3rd Qu.:6.400   3rd Qu.:3.300   3rd Qu.:5.100   3rd Qu.:1.800
Max.   :7.900   Max.   :4.400   Max.   :6.900   Max.   :2.500
```

- 9

- `mydata <- iris`
- `mydata`
- `head(mydata, n = 5)`

```
> head( mydata, n = 5 )
 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
```

- `tail(mydata, n = 5)`

```
> tail( mydata, n = 5 )
 Sepal.Length Sepal.Width Petal.Length Petal.Width Species
146          6.7           3.0           5.2           2.3 virginica
147          6.3           2.5           5.0           1.9 virginica
148          6.5           3.0           5.2           2.0 virginica
149          6.2           3.4           5.4           2.3 virginica
150          5.9           3.0           5.1           1.8 virginica
```

- 10

- `mydata <- iris`
- `mydata`

- `mydata[mydata$Species == "setosa", 1]`
- `mydata[mydata$Species == "versicolor", 1]`
- `mydata[mydata$Species == "virginica", 1]`

- `mydata[mydata$Species == "setosa", 1:2]`
- `mydata[mydata$Species == "setosa", 1:3]`
- `mydata[mydata$Species == "setosa", 2:4]`

- `subset(mydata, Species == "setosa", select = Sepal.Length)`
- `subset(mydata, Species == "setosa", select = c(Sepal.Length, Sepal.Width))`

- `mydata <- iris`
- `mydata`

- `max(mydata[, 1])`
- `min(mydata[, 1])`
- `c(max(mydata[, 1]), min(mydata[, 1]))`
- `MaxMin <- c(max(mydata[, 1]), min(mydata[, 1]))`
- `range(mydata[, 1])`
- `mean (mydata[, 1])`
- `sd(mydata[, 1])`
- `median(mydata[, 1])`

- `mystat <- c(max(mydata[, 1]), mean(mydata[, 1]), median(mydata[, 1]), min(mydata[, 1]), sd(mydata[, 1]))`
- `summary(mydata[, 1])`

- `mydata <- iris`
- `mydata`

- `mydata_sort <- sort(mydata[, 1])`
- `mydata_sort[150*0.5]`
- `mydata_sort[150*0.25]`
- `mydata_sort[150*0.75]`

- `summary(mydata[, 1])`

繪製圖形

- `plot(data)`
 - # 個別繪製圖形
 - `mydata[, 1]`
 - `plot(mydata[, 1])`
 - `mydata[, 2]`
 - `plot(mydata[, 2])`
 - `mydata[, 3]`
 - `plot(mydata[, 3])`
 - `mydata[, 4]`
 - `plot(mydata[, 4])`
 - `mydata[, 5]`
 - `plot(mydata[, 5])`

- `hist(data)`
 - # 個別繪製圖形
 - `mydata[, 1]`
 - `hist(mydata[, 1])`
 - `mydata[, 2]`
 - `hist(mydata[, 2])`
 - `mydata[, 3]`
 - `hist(mydata[, 3])`
 - `mydata[, 4]`
 - `hist(mydata[, 4])`
 - `mydata[, 5]`
 - `hist(mydata[, 5])`

圖形位置安排

17

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
 Chap 03: 鳶尾花 IRIS
 Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- `M` 是圖形分佈的矩陣，
- `widths`、`heights` 各是設定 `M` 矩陣長、寬的比例，其基準點是左上角

`matrix(c(1, 2, 3, 4) , 2, 2, byrow = T)`

1	2
3	4

`widths = c(1,1), heights = c(1,1))`

`matrix(c(1, 2, 3, 4, 5, 6) , 3, 2, byrow = T)`

1	2
3	4
5	6

`widths = c(1,3), heights = c(1,2))`

1	2
3	4

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
 - `M`是圖形分佈的矩陣，
 - `widths`、`heights`各是設定`M`矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `plot(mydata[, 3])`
 - `plot(mydata[, 4])`

- 19

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
 - `M`是圖形分佈的矩陣，
 - `widths`、`heights`各是設定`M`矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `hist(mydata[, 1])`
 - `hist(mydata[, 2])`
 - `hist(mydata[, 3])`
 - `hist(mydata[, 4])`

- 20

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- M是圖形分佈的矩陣，
- `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
 - # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `plot(mydata[, 3])`
 - `plot(mydata[, 4])`

- 21

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- M是圖形分佈的矩陣，
- `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
 - # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`

- 22

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
 - M是圖形分佈的矩陣，
 - `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4, 5, 6, 7, 8) , 4, 2, byrow = T), widths = c(1,1), heights = c(1,1,1,1))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`
 - `plot(mydata[, 3])`
 - `hist(mydata[, 3])`
 - `plot(mydata[, 4])`
 - `hist(mydata[, 4])`

- 23

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
 - M是圖形分佈的矩陣，
 - `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4, 5, 6, 7, 8) , 4, 2, byrow = T), widths = c(1,2), heights = c(1,1,1,1))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`
 - `plot(mydata[, 3])`
 - `hist(mydata[, 3])`
 - `plot(mydata[, 4])`
 - `hist(mydata[, 4])`

- 24

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
 - M是圖形分佈的矩陣，
 - `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(2,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`

- 25

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
 - M是圖形分佈的矩陣，
 - `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,1), heights = c(2,1))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`

- 26

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- M是圖形分佈的矩陣，
- `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
 - # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,3), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`

- 27

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- M是圖形分佈的矩陣，
- `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
 - # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 4) , 2, 2, byrow = T), widths = c(1,3), heights = c(1,3))`
 - `plot(mydata[, 1])`
 - `hist(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `hist(mydata[, 2])`

- 28

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- M是圖形分佈的矩陣，
- `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
 - # 幾張圖繪製在一起
 - `layout(matrix(c(1, 2, 3, 0) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `plot(mydata[, 3])`

- 29

繪圖 - 圖形位置安排

計算機程式設計 - 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `layout(M, widths, heights)`
- M是圖形分佈的矩陣，
- `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
 - # 幾張圖繪製在一起
 - `layout(matrix(c(1, 0, 2, 3) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `plot(mydata[, 2])`
 - `plot(mydata[, 3])`

- 30

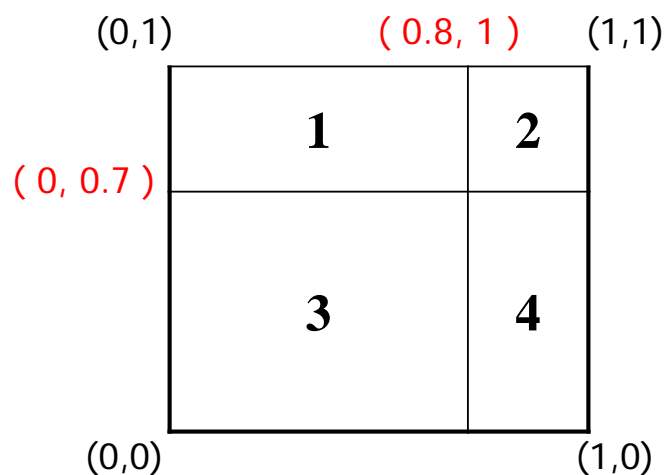
繪圖 - 圖形位置安排

- `layout(M, widths, heights)`
 - M是圖形分佈的矩陣，
 - `widths`、`heights`各是設定M矩陣長、寬的比例，其基準點是左上角
-
- # 幾張圖繪製在一起
 - `layout(matrix(c(1, 0, 0, 2) , 2, 2, byrow = T), widths = c(1,1), heights = c(1,1))`
 - `plot(mydata[, 1])`
 - `plot(mydata[, 2])`

- 31

繪圖 - 圖形位置安排

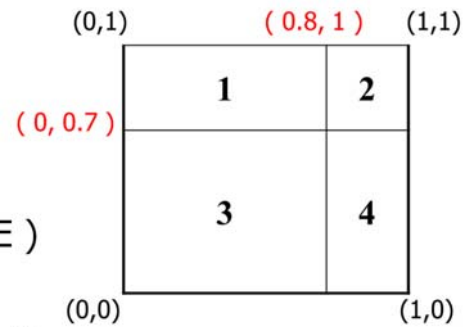
- `par(fig = c(x1, x2, y1, y2))`
- `par(fig = c(0, 0.8, 0.7, 1))`
圖1 的 左下角座標 (x1,y1) 是 (0, 0.7) ,
圖1 的 右上角座標 (x2,y2) 是 (0.8, 1)



- 32

繪圖 - 圖形位置安排

- `par(fig = c(x1, x2, y1, y2))`
- `par(fig = c(0, 0.8, 0.7, 1))`
圖1 的 左下角座標 (x1,y1) 是 (0, 0.7) ,
圖1 的 右上角座標 (x2,y2) 是 (0.8, 1)



- # 幾張圖繪製在一起
- `par(fig=c(0, 0.8, 0, 0.7), new=TRUE)`
- `plot(mydata[, 1])`
- `par(fig=c(0, 0.8, 0.7, 1), new=TRUE)`
- `plot(mydata[, 2])`
- `par(fig=c(0.8, 1, 0, 0.7), new=TRUE)`
- `plot(mydata[, 3])`
- `par(fig=c(0.8, 1, 0.7, 1), new=TRUE)`
- `plot(mydata[, 4])`

- 33

大綱

基本數據圖示法

- `plot(mydata)`
- `plot(mydata[, 1:4])`

- `boxplot(mydata[, 1])`
- `rug(mydata[, 1], side = 4)`

- `boxplot(mydata[, 1], horizontal = TRUE)`
- `rug(mydata[, 1], side = 1)`

- `par(mfrow = c(1, 2), mex = 0.8, mar = c(5,5,3,1) + 0.1)`

- `hist(mydata[, 1], freq = TRUE, breaks = "Sturges")`
- `rug(mydata[, 1], side = 1)`

- `hist(mydata[, 1], prob = TRUE, breaks = "Sturges", col = "lightblue", border = "magenta")`
- `rug(mydata[, 1], side = 1)`

- `hist(mydata[, 1], freq = TRUE, breaks = "Scott")`
- `hist(mydata[, 1], freq = TRUE, breaks = "freedman-Diaconis")`

- `hist(mydata[, 1], freq = TRUE, breaks = 10)`
- `hist(mydata[, 1], freq = TRUE, breaks = seq(from = 4, to = 8, by = 0.2))`

- `stem(mydata[, 1], scale = 1.0)`
- `stem(mydata[, 1], scale = 0.5)`
- `sum(mydata[, 1]) == 4.4)`
- `sum(mydata[, 1]) == 4.6)`

- `stripchart(mydata[, 1], method = "overplot", at = 0.7)`
- `text(6, 0.65, "overplot")`
- `stripchart(mydata[, 1], method = "stack", add = TRUE, at = 0.85)`
- `text(6, 0.8, "stack")`
- `stripchart(mydata[, 1], method = "jitter", add = TRUE, at = 1.2)`
- `text(6, 1.05, "jitter")`
- `title(main = "Strip Chart")`

繪圖 – pie chart 圓餅圖

計算機程式設計 – 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `x <- cut(mydata[, 1], breaks = 6)`
- `y <- table(x)`
-
- `pie(y)`

- `par(mfrow = c(2,2))`
- `pie(y)`
- `pie(y, clockwise = TRUE)`
- `pie(y, col = terrain.colors(6))`
- `pie(y, col = gray(seq(from = 0.4, to = 1.0, length = 6)))`

- 39

繪圖 – density 機率分布圖

計算機程式設計 – 2016F
Chap 03: 鸞尾花 IRIS
Feng-Li Lian @ NTU-EE

- `x <- mydata[, 1]`

- `par(mfrow = c(1,2))`

- `plot(density(x), col = "red", main = "Kernel Density Estimate")`
- `rug(x, side = 1)`

- `hist(x, prob = TRUE, breaks = "Sturges", main = "Histogram & KDE")`
- `lines(density(x), col = "red")`
- `rug(x, side = 1)`

- 40



繪圖 – ECDF 經驗累積分布圖

計算機程式設計 – 2016F
Chap 03: 鸚尾花 IRIS
Feng-Li Lian @ NTU-EE

- `x <- mydata[, 1]`
- `plot.ecdf(x)`

- 41



繪圖 – normal QQ 常態機率圖

計算機程式設計 – 2016F
Chap 03: 鸚尾花 IRIS
Feng-Li Lian @ NTU-EE

- `x <- mydata[, 1]`
- `qqnorm(x)`
- `qqline(x, col = "red", lwd = 2)`

- 42

- iris
- `x <- iris[, 1:4]`

- `plot(x)`
- `pairs(x)`
- `pairs(x, panel = panel.smooth)`

scatterplot

- 43

進階繪圖函數與功能

44

- iris
- x <- iris[, 1:4]

scatterplot

- `panel.hist <- function(x, ...) {`
- `usr <- par("usr"); on.exit(par(usr))`
- `par(usr = c(usr[1:2], 0, 1.5))`
- `h <- hist(x, plot = FALSE)`
- `breaks <- h$breaks; nB <- length(breaks)`
- `y <- h$counts; y <- y / max(y)`
- `rect(breaks[-nB], 0, breaks[-1], y, col = "cyan", ...)`
- `lines(density(x, na.rm = TRUE), col = "red")`
- `}`
- `pairs(x, panel = panel.smooth, pch = 1, bg = "lightcyan",`
- `diag.panel = panel.hist, font.labels = 2, cex.labels = 1.2)`

- 45

- iris

scatterplot 不同品種之 散點圖

- `pairs(iris[, 1:4], pch = c(1, 2, 4)[iris$Species], col = c("red", "green", "blue")[iris$Species])`

- 46

- 第一品種之中，
花萼長度, 花萼寬度, 花瓣長度, 花瓣寬度, 分布情形
- `setosa <- iris[iris$Species == "setosa", 1:4]`
- `boxplot(setosa, names = c("sep.len", "sep.wid", "pet.len", "pet.wid"),
main = "Iris setosa")`

- 三個品種，
花萼長度, 花萼寬度, 花瓣長度, 花瓣寬度, 分布情形
- `par(mfrow = c(1, 2), mex = 0.5, mar = c(5, 4, 4, 2) + 0.1)`
- `with(iris, boxplot(Sepal.Length ~ Species, main = "Sepal length"))`
- `with(iris, boxplot(Sepal.Length ~ Species, notch = TRUE, main =
"Sepal length"))`

- 三個品種，
花萼長度, 花萼寬度, 花瓣長度, 花瓣寬度，分布情形
- 依照不同種類，先分成三群
- `par(mfrow = c(1, 2), mex = 0.5, mar = c(5, 4, 4, 2) + 0.1)`
- `sx <- with(iris, split(Sepal.Length, Species))`
- `boxplot(sx, main = "Sepal length")`
- `boxplot(sx, notch = TRUE, main = "Sepal length")`

- 49

- 花萼長度 與 花萼寬度 之間的關係
- 依照不同種類，先分成三群
- `sx <- with(iris, split(Sepal.Length, Species))`
- `sy <- with(iris, split(Sepal.Width, Species))`
- `plot(0, xlim = range(sx), ylim = range(sy), type = "n", xlab = "x", ylab = "y")`
- `points(sx[[1]], sy[[1]], pch = 1, col = 1)`
- `points(sx[[2]], sy[[2]], pch = 2, col = 2)`
- `points(sx[[3]], sy[[3]], pch = 3, col = 3)`
- `for (i in 1:3) abline(lm(sy[[i]] ~ sx[[i]]), col = i)`
- `legend("topright", legend = c("setosa", "versicolor", "virginica"), lty = 1, pch = 1:3, col = 1:3)`

不同品種之
散點圖

- 50

- 花萼長度與花萼寬度之間的關係
 - 依照不同種類，先分成三群
- 不同品種之散點圖
- `x <- iris[[1]]`
 - `y <- iris[[2]]`
 - `species <- iris[[5]]`
 - `library(lattice)`
 - `xyplot(y ~ x, groups = species, type = c("g", "p", "r"), auto.key = TRUE)`

- 花萼長度與花萼寬度之間的關係
 - 依照不同種類，先分成三群
- 不同品種分開之散點圖
- `x <- iris[[1]]`
 - `y <- iris[[2]]`
 - `species <- iris[[5]]`
 - `library(lattice)`
 - `xyplot(y ~ x | species, type = c("g", "p", "r"), auto.key = TRUE)`

- (花萼長度, 花萼寬度, 花瓣長度)
- data(iris)
- x <- iris[, 1]
- y <- iris[, 2]
- z <- iris[, 3]
- library(lattice)
- `cloud(z ~ x * y, groups = iris$Species, pch = 1:3, col = 1:3,`
- `scales = list(arrows = FALSE),`
- `light.source = c(10, 0, 10))`

作業

HW02

On 9/28, 2016

- 主旨：數據處理與繪圖指令功能
- 繳交下面檔案，檔案名稱：HW02_學號
 - 歷史指定檔案： HW02_B01921001.Rhistory
 - 主要指定檔案： HW02_B01921001.R
 - 報告檔案： HW02_B01921001.pptx 或者 .pdf
- 請說明電腦使用方式，例如：
 - 筆電：Windows 10 + R + Rstudio
 - 筆電：Max OS X + R + Rstudio
- 繳交方式與期限：
 - E-mail 上面四個檔案到：ntucp105f@gmail.com
 - E-mail 主旨：HW02_B01921001 (就是，作業編號_您的學號)
 - 繳交期限：10/5 (Wed), 2016, 11pm 以前
 - HW02，每位繳交自己的作業，或者兩人繳交一份
 - 兩人完成作業者，請註明工作模式，例如：
 - A 進行指令測試與報告編輯，然後給 B 重複測試
 - A 進行指令測試，B 進行報告編輯
 - A 進行指令測試與報告編輯，B 僅看過報告，但沒有實際測試
 - A 進行指令測試與報告編輯，B 沒有參與指令測試，也沒有看過報告

- 55

HW02

On 9/28, 2016

- 主旨：數據處理與繪圖功能測試
- 請參考 U03 講義，以及指令歷史檔
- 自行挑選一組數據分析指令，以及一組圖形繪製指令
- 執行這些指令
- 把從頭到尾的執行過程，編輯於 .R 檔之中，並且依序執行這些指令
- 把執行的過程，或者是產生的數據/圖形等，整理到報告檔 (pptx or pdf)
- 報告檔中，請編輯：
 - 所使用的的指令
 - 產生的結果，數據 and/or 圖形
 - 解釋說明該指令的功能，產生的結果，該結果的意義，特點等

- 56