你想呈現什麼? by R
What would you like to show? by R

範例檔案

比較
Comparison

單變量
One Variable

- 長條圖 Single Bar Chart — Rcode
- 群組長條圖 — Rcode

單項目 / 多項目

雙變量
Two Variable

- 不等寬柱狀圖 Variable Width Column — Rcode

靜止 Static

三或多變量
Three or more Variable

- Table or Table with Embedded Charts — Rcode

隨時間變化
Changing Over Time

- 雷達圖(Radar chart) Circular Area Chart — Rcode

具循環週期性 Cyclical Data

- 折線圖 Line Chart — Rcode

非循環週期性 Non-Cyclical Data

多時間 Many Periods

少時間 Few Periods

- 柱狀圖 Column Chart — Rcode
- 折線圖 Line Chart — Rcode

單一或少量項目 Single or Few Items

多項目 Many Items

關聯
Relationship

雙變量
Two Variables

- 散佈圖 Scatter Chart — Rcode

三變量
Three Variables

- 氣泡圖 Bubble Chart — Rcode

組成
Composition

隨時間變化
Changing Over Time

少週期 Few Periods

- 只有相對差異 Only Relative Differences Matter
  - 百分比堆疊長條圖 Stacked 100% Column Chart — Rcode
- 相對與絕對差異 Relative and Absolute Differences Matter
  - 堆疊長條圖 Stacked Column Chart — Rcode

多週期 Many Periods

- 只有相對差異 Only Relative Differences Matter
  - 百分比堆疊圖面積圖 Stacked 100% Stacked Area Chart — Rcode
- 相對與絕對差異 Relative and Absolute Differences Matter
  - 堆疊面積圖 Stacked Area Chart — Rcode

靜止
Static

- 組成中的組成 Components of Components
  - 帶有子類別 — Rcode
- 累積或累減總和 Accumulation or Subtraction to Total
  - 瀑布圖 Waterfall Chart — Rcode

總和的簡單分類 Simple Share of Total

- 圓餅圖 Pie Chart — Rcode

分佈
Distribution

單變量
One Variable

- 少數據 Few Item
  - 直方圖 Column Histogram — Rcode
- 多數據 Many Item
  - 密度圖(Density Plot) Line Histogram — Rcode

雙變量
Two Variables

- 散佈圖 Scatter Chart — Rcode

三變量
Three Variables

- 曲面圖 3D Area Char — Rcode

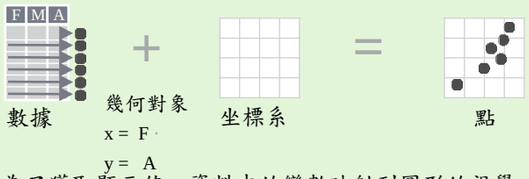
包含地理資訊

- Rcode
- Rcode

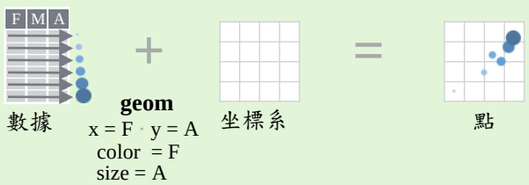
ggplot2數據可視化::速查表

基礎

ggplot2 基於圖形語法，使用相同的元件（數據集、坐標系和表示資料點的幾何物件）來構建圖片。



為了獲取顯示值，資料中的變數映射到圖形的視覺屬 性，如大小、顏色以及x和y位置。



完成以下範本來構建圖形

```
ggplot (data = <DATA> ) +                          必要
<geom_function>(mapping = aes( <mappings> ),
 stat = <stat> , position = <position> ) +          非必要，
<Coordinate_function> +                            有預設值
<Facet_function> +
<Scale_function> +
<Theme_function>
```

ggplot(data = mpg, **aes(**x = cty, y = hwy**))** 通過添加圖層來完成圖形，每層添加一個**geom**函數。

last_plot() 返回上一個圖片。

ggsave("plot.png", width = 5, height = 5**)** 將最後一個圖片保存至工作目錄中名為"plot.png"的5'x 5'文件。 檔案類型與文件副檔名相匹配。

參數 通用參數

color and **fill** - string ("red", "#RRGGBB") #框線或填充顏色

linetype - integer or string (0 = "blank", 1 = "solid", 2 = "dashed", 3 = "dotted", 4 = "dotdash", 5 = "longdash", 6 = "twodash") #線條樣式

lineend - string ("round", "butt", or "square") #線端點樣式

linejoin - string ("round", "mitre", or "bevel") #線段點樣式

size - integer (line width in mm)

shape - integer/shape name or a single character ("a") #資料點樣式

幾何對象

使用geom函數表達資料，使用geom的美學屬性表示變量。 每個函數繪製一個圖層。

基本圖像

```
a <- ggplot(economics, aes(date, unemploy))
b <- ggplot(seals, aes(x = long, y = lat))
```

a + geom_blank() and **a + expand_limits()**

b + geom_curve(aes(yend = lat + 1, xend = long + 1), curvature = 1**)** - x, xend, y, yend, alpha, angle, color, curvature, linetype, size

a + geom_path(lineend = "butt", linejoin = "round", linemitre = 1**)** x, y, alpha, color, group, linetype, size

a + geom_polygon(aes(alpha = 50)**)** - x, y, alpha, color, fill, group, subgroup, linetype, size

b + geom_rect(aes(xmin = long, ymin = lat, xmax = long + 1, ymax = lat + 1)**)** - xmax, xmin, ymax, ymin, alpha, color, fill, linetype, size

a + geom_ribbon(aes(ymin = unemploy - 900, ymax = unemploy + 900)**)** - x, ymax, ymin, alpha, color, fill, group, linetype, size

線段

常用參數: x, y, alpha, color, linetype, size

b + geom_abline(aes(intercept = 0, slope = 1)**)**
b + geom_hline(aes(yintercept = lat)**)**
b + geom_vline(aes(xintercept = long)**)**

b + geom_segment(aes(yend = lat + 1, xend = long + 1)**)**
b + geom_spoke(aes(angle = 1:1155, radius = 1)**)**

單變數 連續

```
c <- ggplot(mpg, aes(hwy)); c2 <- ggplot(mpg)
```

c + geom_area(stat = "bin"**)** x, y, alpha, color, fill, linetype, size

c + geom_density(kernel = "gaussian"**)** x, y, alpha, color, fill, group, linetype, size, weight

c + geom_dotplot() x, y, alpha, color, fill

c + geom_freqpoly() x, y, alpha, color, group, linetype, size

c + geom_histogram(binwidth = 5**)** x, y, alpha, color, fill, linetype, size, weight

c2 + geom_qq(aes(sample = hwy)**)** x, y, alpha, color, fill, linetype, size, weight

離散變數

```
d<- ggplot(mpg, aes(fl))
```

d + geom_bar() x, alpha, color, fill, linetype, size, weight

雙變數

兩個都連續

```
e <- ggplot(mpg, aes(cty, hwy))
```

e + geom_label(aes(label = cty), nudge_x = 1, nudge_y = 1**)** - x, y, label, alpha, angle, color, family, fontface, hjust, lineheight, size, vjust

e + geom_point() x, y, alpha, color, fill, shape, size, stroke

e + geom_quantile() x, y, alpha, color, group, linetype, size, weight

e + geom_rug(sides = "bl"**)** x, y, alpha, color, linetype, size

e + geom_smooth(method = lm**)** x, y, alpha, color, fill, group, linetype, size, weight

e + geom_text(aes(label = cty), nudge_x = 1, nudge_y = 1**)** - x, y, label, alpha, angle, color, family, fontface, hjust, lineheight, size, vjust

一個離散，一個連續

```
f <- ggplot(mpg, aes(class, hwy))
```

f + geom_col() x, y, alpha, color, fill, group, linetype, size

f + geom_boxplot() x, y, lower, middle, upper, ymax, ymin, alpha, color, fill, group, linetype, shape, size, weight

f + geom_dotplot(binaxis = "y", stackdir = "center"**)** x, y, alpha, color, fill, group

f + geom_violin(scale = "area"**)** x, y, alpha, color, fill, group, linetype, size, weight

兩個都是離散變數

```
g <- ggplot(diamonds, aes(cut, color))
```

g + geom_count() x, y, alpha, color, fill, shape, size, stroke

e + geom_jitter(height = 2, width = 2**)** x, y, alpha, color, fill, shape, size

連續二元分佈

```
h <- ggplot(diamonds, aes(carat, price))
```

h + geom_bin2d(binwidth = c(0.25, 500)**)** x, y, alpha, color, fill, linetype, size, weight

h + geom_density_2d() x, y, alpha, color, group, linetype, size

h + geom_hex() x, y, alpha, color, fill, size

連續函數

```
i <- ggplot(economics, aes(date, unemploy))
```

i + geom_area() x, y, alpha, color, fill, linetype, size

i + geom_line() x, y, alpha, color, group, linetype, size

i + geom_step(direction = "hv"**)** x, y, alpha, color, group, linetype, size

視覺化誤差

```
df <- data.frame(grp = c("A", "B"), fit = 4:5, se = 1:2)
j <- ggplot(df, aes(grp, fit, ymin = fit - se, ymax = fit + se))
```

j + geom_crossbar(fatten = 2**)** - x, y, ymax, ymin, alpha, color, fill, group, linetype, size

j + geom_errorbar() - x, ymax, ymin, alpha, color, group, linetype, size, width
Also **geom_errorbarh()**.

j + geom_linerange() x, ymin, ymax, alpha, color, group, linetype, size

j + geom_pointrange() - x, y, ymin, ymax, alpha, color, fill, group, linetype, shape, size

地圖

```
data <- data.frame(murder = USArrests$Murder,
                   state = tolower(rownames(USArrests)))
map <- map_data("state")
 k <-ggplot(data, aes(fill = murder))
```

k + geom_map(aes(map_id = state), map = map**)**
+ expand_limits(x = map$long, y = map$lat**)**
map_id, alpha, color, fill, linetype, size

三變數

```
seals$z <- with(seals, sqrt(delta_long^2 + delta_lat^2)); l <- ggplot(seals, aes(long, lat))
```

l + geom_contour(aes(z = z)**)** x, y, z, alpha, color, group, linetype, size, weight

l + geom_contour_filled(aes(fill = z)**)** x, y, alpha, color, fill, group, linetype, size, subgroup

l + geom_raster(aes(fill = z), hjust = 0.5, vjust = 0.5, interpolate = FALSE**)** x, y, alpha, fill

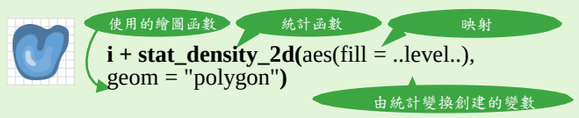
l + geom_tile(aes(fill = z)**)** x, y, alpha, color, fill, linetype, size, width

統計變換 另一種構建圖層的方法

統計變換構建新變數來繪圖（例如，count，prop）。

通過更改geom函數的默認統計信息，**geom_bar(stat="count")** 或者使用統計變化功能來繪圖**stat_count(geom="bar")**，其調用默認圖片來創建一個圖層（相當於geom函數）。

使用 **..name..** 語法將統計變化映射到坐標。



c + stat_bin(binwidth = 1, boundary = 10**)**
x, y | ..count.., ..ncount.., ..density.., ..ndensity..

c + stat_count(width = 1**) x, y** | ..count.., ..prop..

c + stat_density(adjust = 1, kernel = "gaussian"**)**
x, y | ..count.., ..density.., ..scaled..

e + stat_bin_2d(bins = 30, drop = T**)**
x, y, fill | ..count.., ..density..

e + stat_bin_hex(bins = 30**) x, y, fill** | ..count.., ..density..

e + stat_density_2d(contour = TRUE, n = 100**)**
x, y, color, size | ..level..

e + stat_ellipse(level = 0.95, segments = 51, type = "t"**)**

l + stat_contour(aes(z = z)**) x, y, z, order** | ..level..

l + stat_summary_hex(aes(z = z), bins = 30, fun = max**) x, y, z, fill** | ..value..

l + stat_summary_2d(aes(z = z), bins = 30, fun = mean**)**
x, y, z, fill | ..value..

f + stat_boxplot(coef = 1.5**)**
x, y | ..lower.., ..middle.., ..upper.., ..width.. , ..ymin.., ..ymax..

f + stat_ydensity(kernel = "gaussian", scale = "area"**) x, y** | ..density.., ..scaled.., ..count.., ..n.., ..violinwidth.., ..width..

e + stat_ecdf(n = 40**) x, y** | ..x.., ..y..

e + stat_quantile(quantiles = c(0.1, 0.9), formula = y ~ log(x), method = "rq"**) x, y** | ..quantile..

e + stat_smooth(method = "lm", formula = y ~ x, se = T, level = 0.95**) x, y** | ..se.., ..x.., ..y.., ..ymin.., ..ymax..

ggplot() + xlim(-5, 5**) + stat_function(**fun = dnorm, n = 20, geom = "point"**) x** | ..x.., ..y..

ggplot() + stat_qq(aes(sample = 1:100)**)**
x, y, sample | ..sample.., ..theoretical..

e + stat_sum() x, y, size | ..n.., ..prop..

e + stat_summary(fun.data = "mean_cl_boot"**)**

h + stat_summary_bin(fun = "mean", geom = "bar"**)**

e + stat_identity()

e + stat_unique()

尺規 使用**scales**包覆蓋預設值

將映射資料縮放到較為美觀的比例。

添加新的尺規來改變映射。

n <- d + geom_bar(aes(fill = fl))

scale_ 開頭 | 調整參數 | 預設參數 | 尺規特定屬性的參數

n + scale_fill_manual(
values = c("skyblue", "royalblue", "blue", "navy"), **limits** = c("d", "e", "p", "r"), breaks =c("d", "e", "p", "r"), **name** = "fuel", labels = c("D", "E", "P", "R")**)**

映射中包含值的範圍 | 圖例/軸中使用的標題 | 圖例/軸中使用的標籤 | 圖例/軸的分隔

尺規的一般用法

使用大多數參數

scale_*_continuous() - 將資料的連續取值映射為圖形屬性的取值
scale_*_discrete() - 將資料的離散取值映射為圖形屬性的取值
scale_*_binned() - 將資料的連續取值映射為離散的統計項
scale_*_identity() - 使用資料的值作為圖形屬性的取值
scale_*_manual(values = c()**)** - 將資料的離散取值作為手工指定的圖形屬性的取值
scale_*_date(date_labels = "%m/%d"),
date_breaks = "2 weeks"**)** - 將資料值視為日期
scale_*_datetime() - 將數據x視為時間

參數和scale_x_date()一樣。有關標簽格式請參閱?striptime。

調整X和Y的比例

調整x和y的標尺(使用x為例)
scale_x_log10() - 以log10比例繪製
scale_x_reverse() - 反轉x軸方向
scale_x_sqrt() - 以平方根繪製x

顏色和填充比例（離散）

n + scale_fill_brewer(palette = "Blues"**)**
選擇調色板: RColorBrewer::display.brewer.all()

n + scale_fill_grey(start = 0.2, end = 0.8, na.value = "red"**)**

顏色和填充比例（連續）

o <- c + geom_dotplot(aes(fill = ..x..))
o + scale_fill_distiller(palette = "Blues"**)**

o + scale_fill_gradient(low="red", high="yellow"**)**

o + scale_fill_gradient2(low = "red", high = "blue", mid = "white", midpoint = 25**)**

o + scale_fill_gradientn(colors = topo.colors(6)**)**
也見: rainbow(), heat.colors(), terrain.colors(), cm.colors(), RColorBrewer::brewer.pal()

形狀和尺寸比例

p <- e + geom_point(aes(shape = fl, size = cyl))

p + scale_shape() + scale_size()
p + scale_shape_manual(values = c(3:7)**)**
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

p + scale_radius(range = c(1,6)**)**
p + scale_size_area(max_size = 6**)**

坐標系

r <- d + geom_bar()

r + coord_cartesian(xlim = c(0, 5)**)** - xlim, ylim
默認笛卡爾坐標系

r + coord_fixed(ratio = 1/2**)**
ratio, xlim, ylim - x和y單位之間固定長寬比的笛卡爾座標

ggplot(mpg, aes(y = fl)**) + geom_bar()**
通過切換x和y參數映射翻轉笛卡爾座標

r + coord_polar(theta = "x", direction=1**)**
theta, start, direction – 極座標

r + coord_trans(y = "sqrt"**)** - x, y, xlim, ylim
轉換後的笛卡爾座標。將xtrans和ytrans設置為視窗函數的名稱。

π + coord_quickmap()
π + coord_map(projection = "ortho", orientation = c(41, -74, 0)**)** - projection, xlim, ylim

從mapproj包中映射投影(mercator (default), azequalarea, lagrange, etc.)

位置調整

位置調整決定了如何安排原本會占據相同空間的圖例

s <- ggplot(mpg, aes(fl, fill = drv))

s + geom_bar(position = "dodge")
並排排列元素

s + geom_bar(position = "fill")
堆疊元素並標準化高度

e + geom_point(position = "jitter")
將隨機抖動添加到每個元素的X和Y位置以避免重疊

e + geom_label(position = "nudge")
標籤稍微遠離資料點

s + geom_bar(position = "stack")
堆疊元素

每個位置調整都可以重新編寫為具有手動寬度和高度參數的函數:
s + geom_bar(position = position_dodge(width = 1))

主題

r + theme_bw()
網格白色背景

r + theme_gray()
灰色背景（默認主題）

r + theme_dark()
黑色背景

r + theme_classic()

r + theme_light()

r + theme_linedraw()

r + theme_minimal()
簡單主題

r + theme_void()
清空主題

r + theme() 自訂主題的各個方面，例如軸、圖例、面板和構面屬性。
r + ggtitle("Title") + theme(plot.title.postion = "plot")
r + theme(panel.background = element_rect(fill = "blue"))

分面

根據一個或多個離散變數劃分子圖。

t <- ggplot(mpg, aes(cty, hwy)) + geom_point()

t + facet_grid(cols = vars(fl))
基於fl的列分面

t + facet_grid(rows = vars(year))
基於year的行分面

t + facet_grid(rows = vars(year), cols = vars(fl))
列和行的分面圖

t + facet_wrap(vars(fl))
包裹成矩形佈局的分面圖

設置scales限制分面坐標軸
t + facet_grid(rows = vars(drv), cols = vars(fl), scales = "free")
x和y軸適應各自的分面
"free_x" - 限制調整x軸
"free_y" - 限制調整y

軸設置labeller屬性調整分面的標籤
t + facet_grid(cols = vars(fl), labeller = label_both)
t + facet_grid(rows = vars(fl),

fl: c	fl: d	fl: e	fl: p	fl: r

labeller = label_bquote(alpha ^ .(fl)**))**

α^c	α^d	α^e	α^p	α^r

標籤與圖例

使用**labs()**標記圖中的元素。
t + labs(x = "New x axis label", **y** = "New y axis label",
title ="Add a title above the plot",
subtitle = "Add a subtitle below title",
caption = "Add a caption below plot",
alt = "Add alt text to the plot",
<AES> = "New <AES> legend title"**)**

t + annotate(geom = "text", x = 8, y = 9, label = "A"**)**
Places a geom with manually selected aesthetics.使用手動選擇的參數調整幾何物件的放置位元置。

p + guides(x = guide_axis(n.dodge = 2)**)** 使用guide_axis(n.dodge 或 angle)避免擁擠或重疊的標籤。

n + guides(fill = "none"**)**設置圖例類型**: colorbar, legend, or none (no legend)**

n + theme(legend.position = "bottom"**)**
放置圖例："bottom", "top", "left", or "right"

n + scale_fill_discrete(name = "Title", labels = c("A", "B", "C", "D", "E")**)**
使用scale函數設置圖例標籤

縮放

沒有裁剪（推薦）

t + coord_cartesian(xlim = c(0, 100), ylim = c(10, 20)**)**

裁剪（刪除看不見的資料點）

t + xlim(0, 100**) + ylim(**10, 20**)**

t + scale_x_continuous(limits = c(0, 100)**) + scale_y_continuous(**limits = c(0, 100)**)**