

R for bioinformatics, data summarisation and statistics

HUST Bioinformatics course series

Wei-Hua Chen (CC BY-NC 4.0)

25 三月, 2020

section 1: TOC

test2

ls -l

```

## total 78360
## -rw-r--r--@ 1 wchen staff      205  3 25 17:45 R for bioinformatics 2019.Rproj
## -rw-r--r--@ 1 wchen staff      943  7 22  2019 Untitled.Rmd
## -rw-r--r--@ 1 wchen staff      145  7 18  2019 _output.yaml
## -rw-r--r--@ 1 wchen staff    4549  7 18  2019 beamer_default.tex
## drwxr-xr-x  3 wchen staff       96  8  8  2019 cheatsheets
## drwxr-xr-x 10 wchen staff      320  9 24  2019 data
## drwxr-xr-x 14 wchen staff      448  9 24  2019 images
## -rw-r--r--@ 1 wchen staff      166  8  3  2019 mystyle.sty
## drwxr-xr-x  3 wchen staff       96  7 19  2019 rscripts
## -rw-r--r--@ 1 wchen staff     2509  2 25 10:49 talk00.Rmd
## -rw-r--r--  1 wchen staff    95484  2 25 10:46 talk00.log
## -rw-r--r--@ 1 wchen staff 1706791  2 25 10:46 talk00.pdf
## -rw-r--r--@ 1 wchen staff     5656  9 27 13:16 talk00_github.Rmd
## -rw-r--r--  1 wchen staff    104404  9 27 13:16 talk00_github.log
## -rw-r--r--@ 1 wchen staff   5183829  9 27 13:16 talk00_github.pdf
## -rw-r--r--@ 1 wchen staff     13227 11 28 15:52 talk00_试讲.Rmd
## -rw-r--r--  1 wchen staff     113384 11 28 15:53 talk00_试讲.log
## -rw-r--r--@ 1 wchen staff   8355725 11 28 15:53 talk00_试讲.pdf
## -rw-r--r--@ 1 wchen staff     17126 11 28 15:22 talk01.Rmd
## -rw-r--r--  1 wchen staff     114820 11 28 10:57 talk01.log
## -rw-r--r--@ 1 wchen staff   7844853 11 28 10:57 talk01.pdf
## -rw-r--r--@ 1 wchen staff     13064  2 25 15:13 talk02.Rmd
## -rw-r--r--  1 wchen staff     118626  2 25 15:13 talk02.log
## -rw-r--r--@ 1 wchen staff   701362  2 25 15:13 talk02.pdf

```

技能树

前置技能

是否需要？

编程？	不需要	plus
Linux？	不需要	plus

装入必需的 packages ...

```
library(tidyverse);
```

```
## -- Attaching packages -----
```

```
## v ggplot2 3.2.1      v purrr  0.3.2
## v tibble  2.1.3      v dplyr  0.8.3
## v tidyr   1.0.0      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0
```

```
## -- Conflicts -----
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

ggplot2 的四个基本组成

1 图层 (layers)

- `geom_< 图层名 >`

2 scale: 控制数据至美学属性的 mapping

- `scale_< 属性 mapping 方式 >`, e.g. `scale_color_identity()`

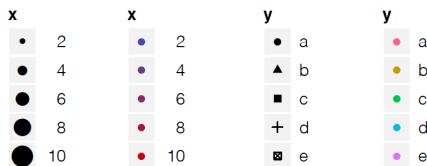


Figure 1: 数据的 4 种 scale 方法

ggplot2 的 scale

- `scale_color_...`
- `scale_shape_...`
- `scale_size_...`
- `scale_fill_...`

与坐标系统联动的函数

- `scale_x_log()`
- `scale_y_log()`

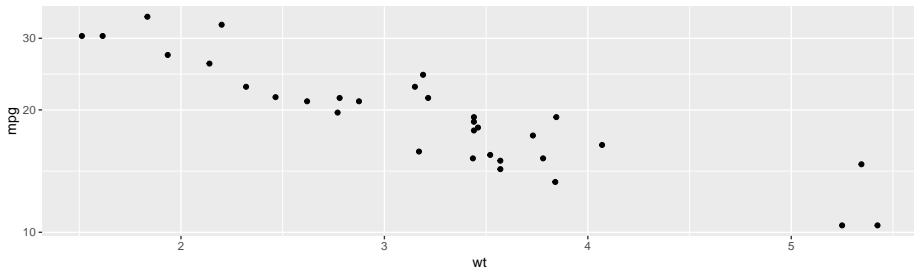
更多内容可以见 《ggplot2: elegant graphics for data analysis》一书的第 6 章。

ggplot2 要素 3: 坐标系统

- 正常
- log-transform

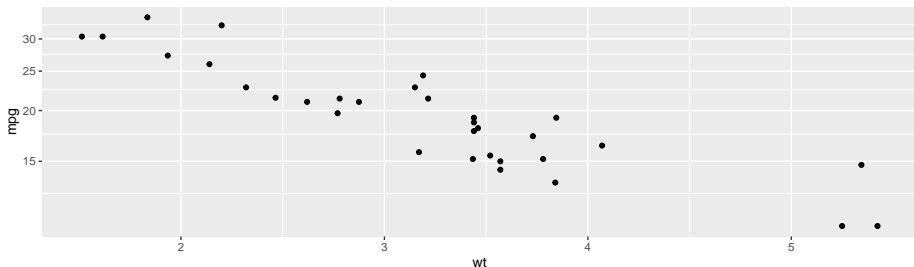
示例:

```
ggplot(mtcars, aes( wt , mpg)) + geom_point() +  
  scale_y_log10()
```



ggplot2 要素 3: 坐标系统, cont.

```
ggplot(mtcars, aes( wt , mpg)) + geom_point() +  
  coord_trans( y = "log10" );
```



`coord_trans()` 的其它参数:

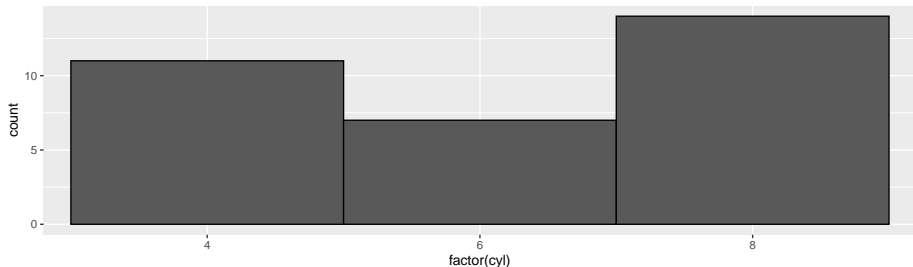
- `limx`, `limy`: 限制 `xy` 的显示范围

ggplot2 要素 3: 坐标系统, cont.

其它函数

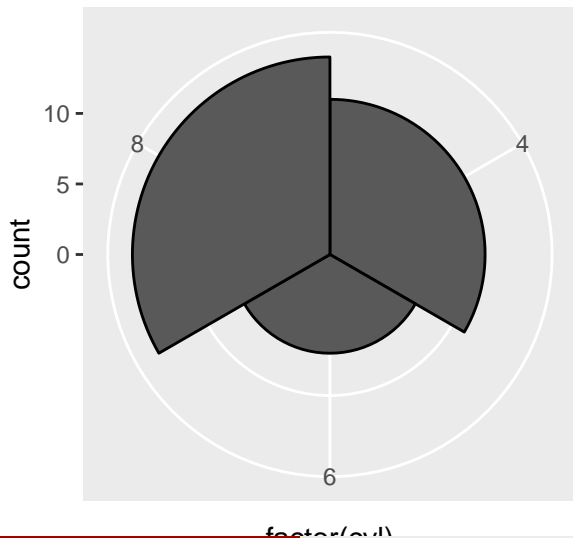
- `coord_flip()` : x, y 轴互换; 竖 bar 变横 bar;
- `coord_polar()` :

```
plot1 <- ggplot(mtcars, aes(x = factor(cyl))) +  
  geom_bar(width = 1, colour = "black");  
plot1;
```



ggplot2 要素 3: 坐标系统, cont.

```
plot1 + coord_polar();
```



ggplot2 要素 4: faceting ...

```
qplot(displ, hwy, data=mpg, facets = . ~ year) + geom_smooth();
```

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

