

Visualisation in R with ggplot2 and plyr

Hadley Wickham

Assistant Professor / Dobelman Family Junior Chair
Department of Statistics / Rice University

September 2011



1. Setup - <http://courses.had.co.nz>

2. Introductions

3. Course outline

HELLO

my name is

Hadley

Your turn

Throughout the course, there will be many opportunities for you to practice what you've learned and receive one-on-one help.

Please get things started by introducing yourself: what's your name and what sort of data do you work with.

Outline

Why? Data analysis

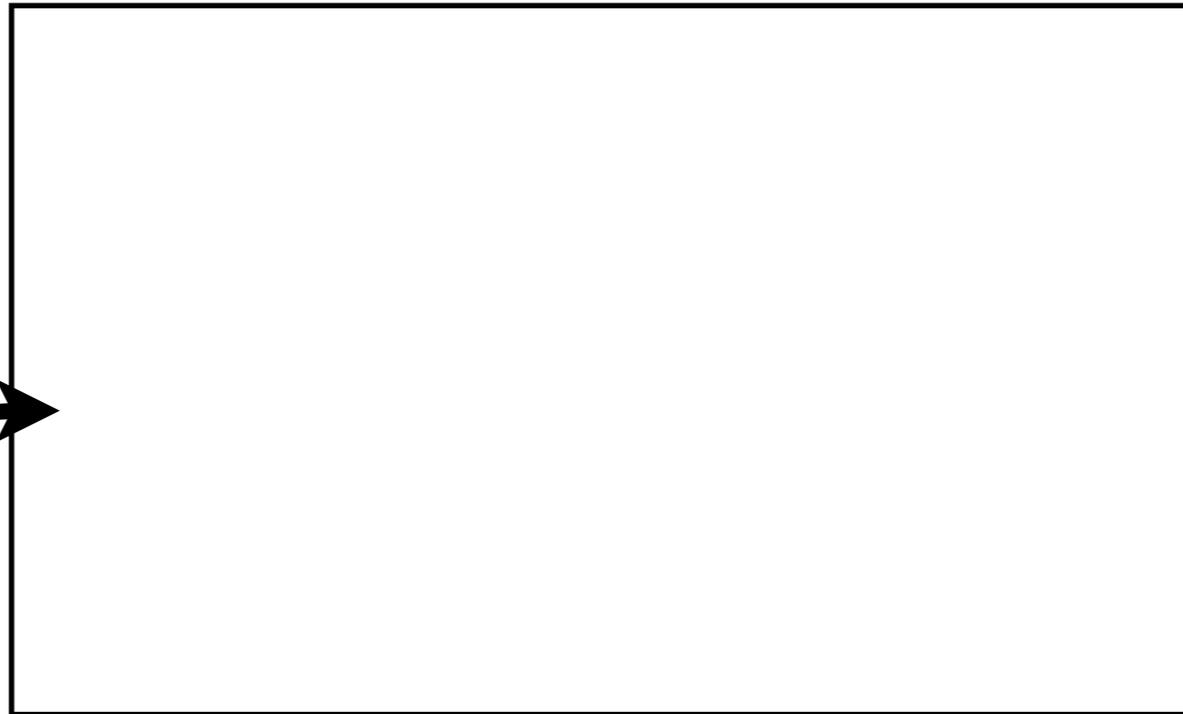
Why? Data analysis

Question

Why? Data analysis

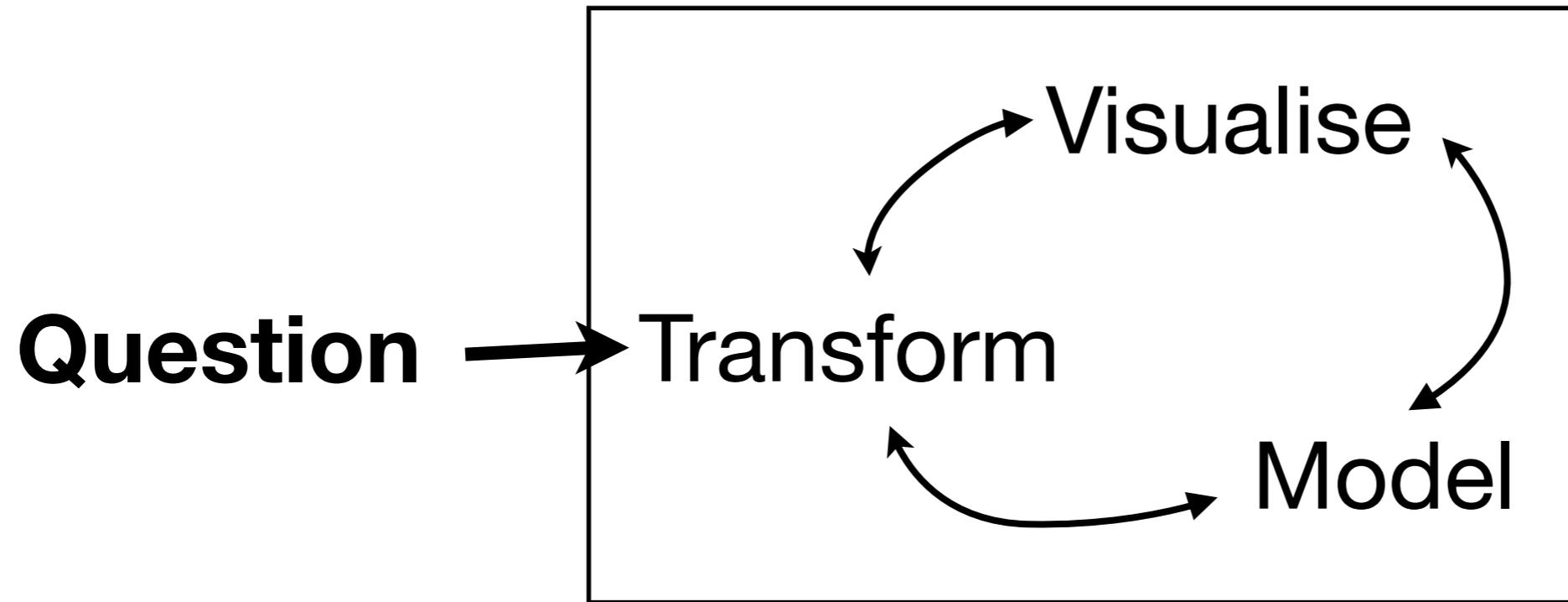
Understand

Question →

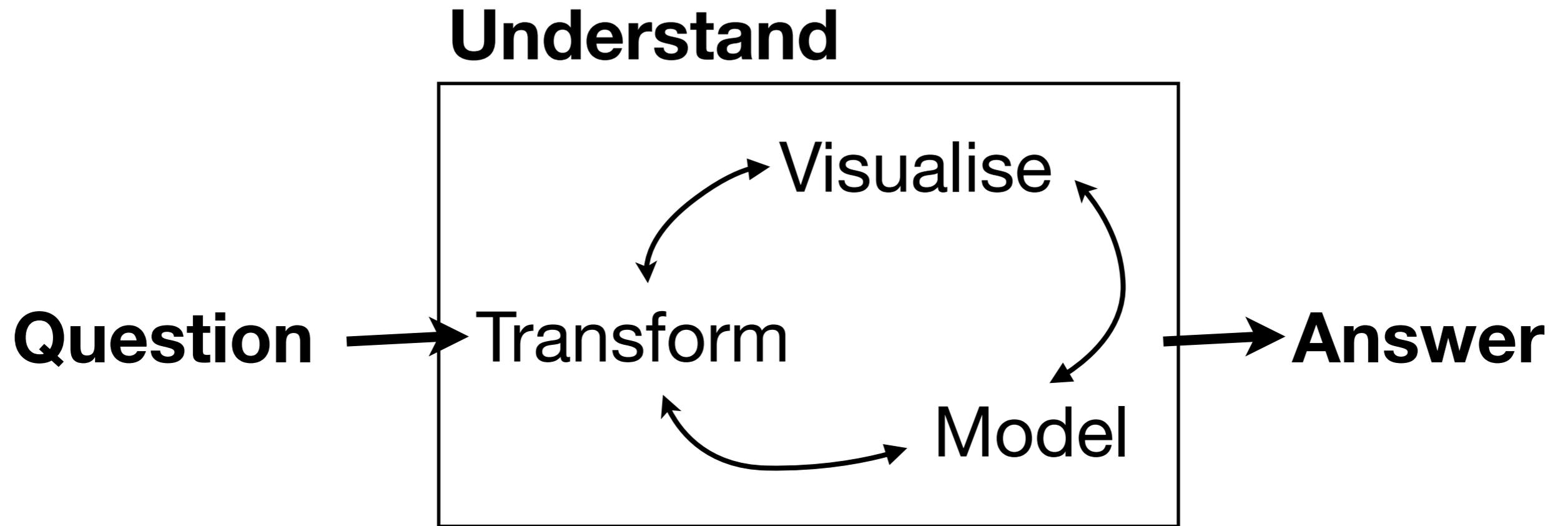


Why? Data analysis

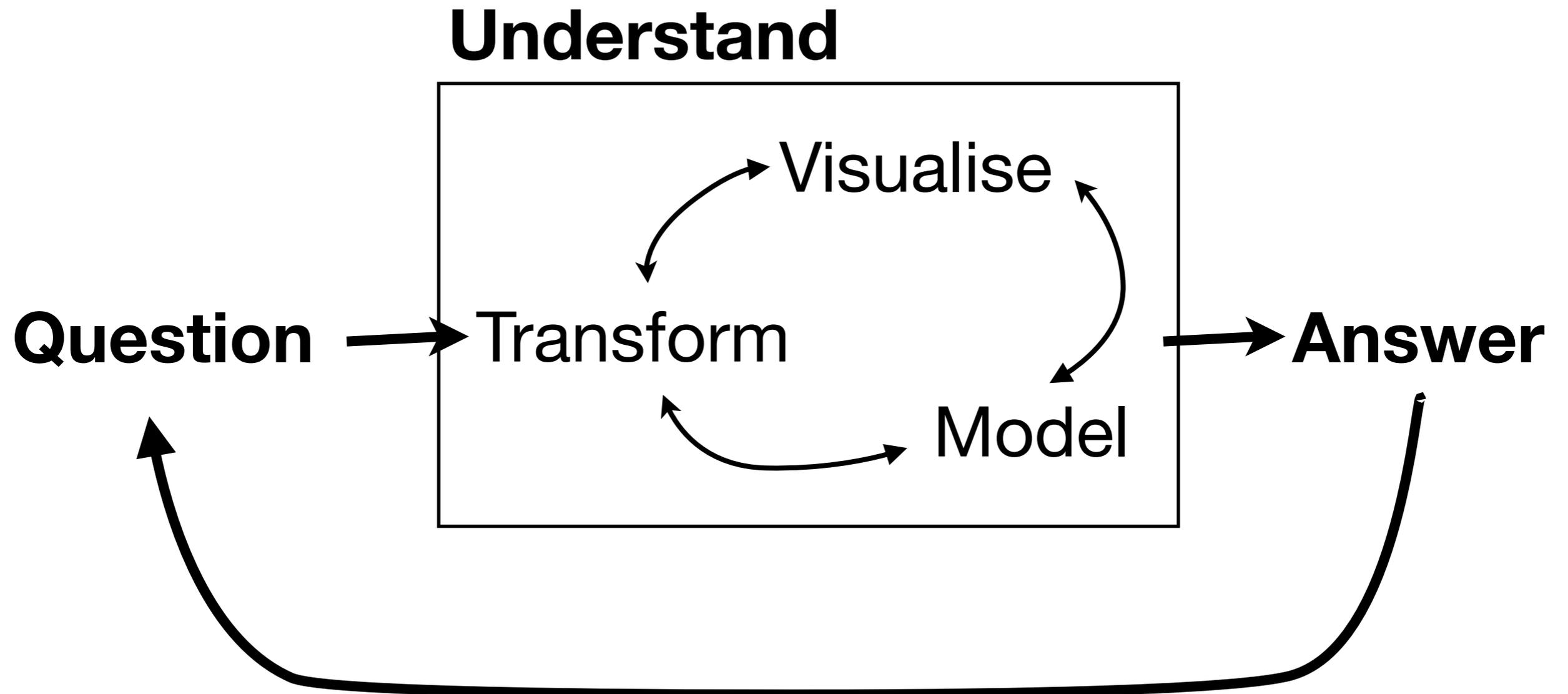
Understand

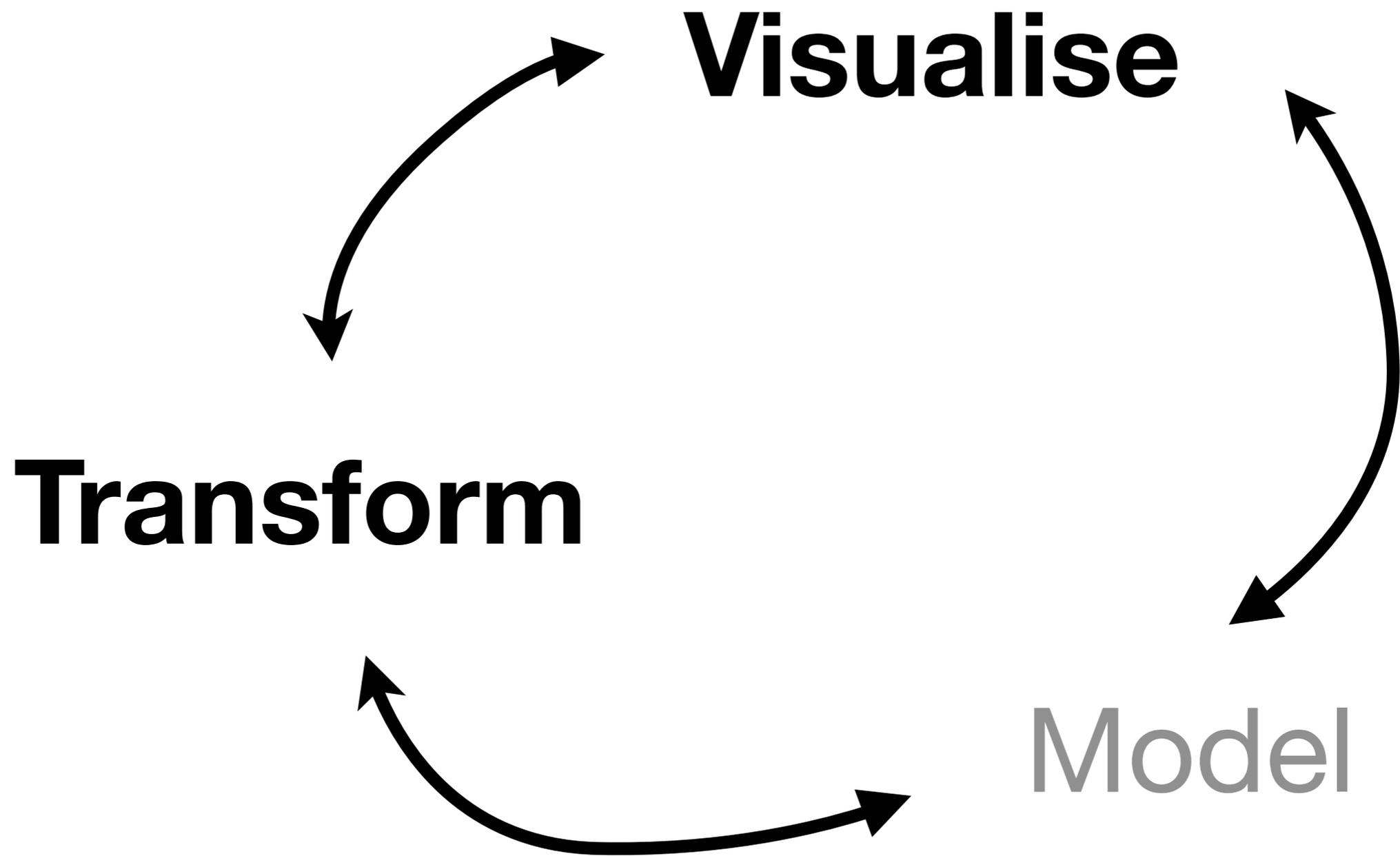


Why? Data analysis

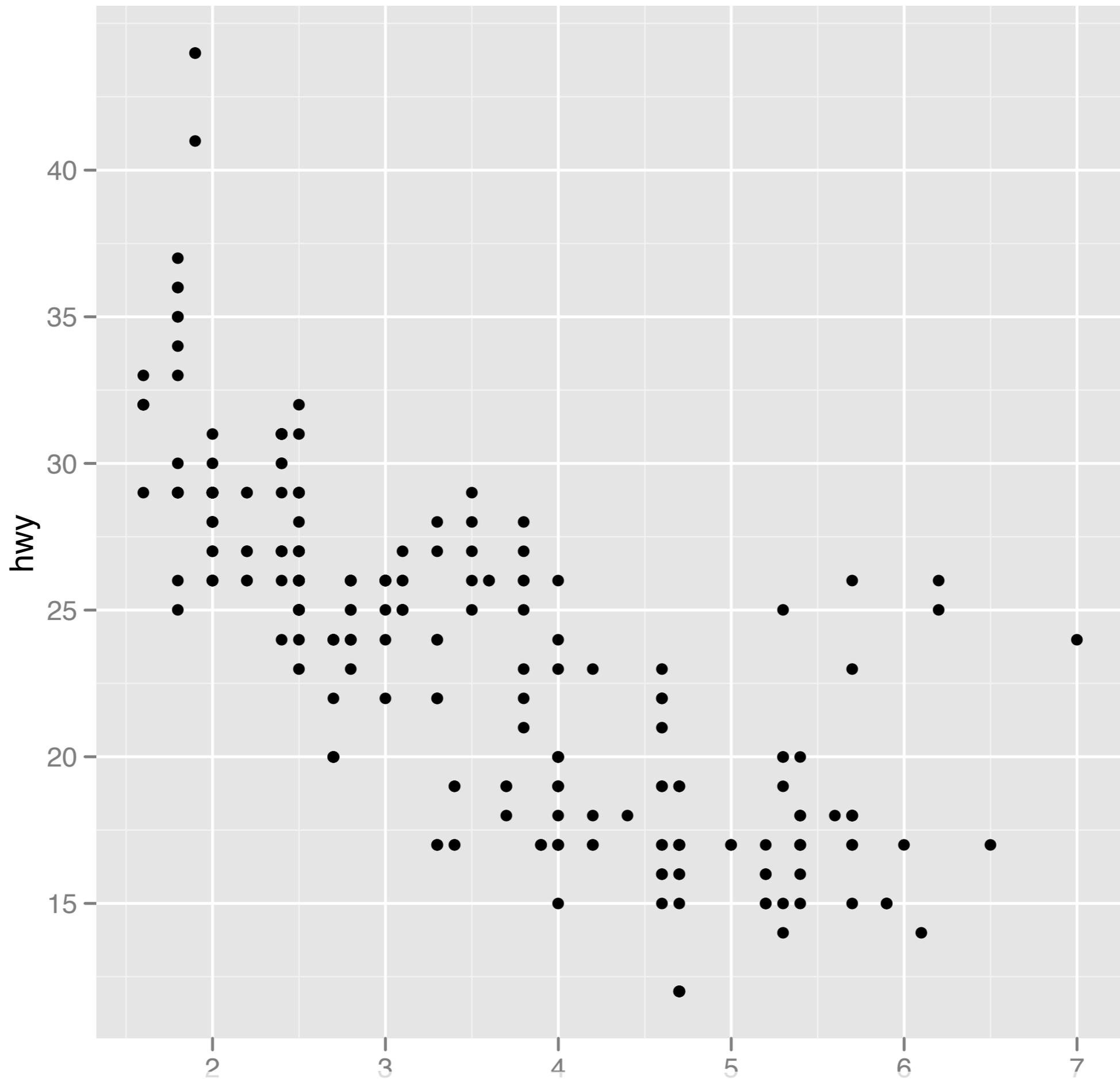


Why? Data analysis

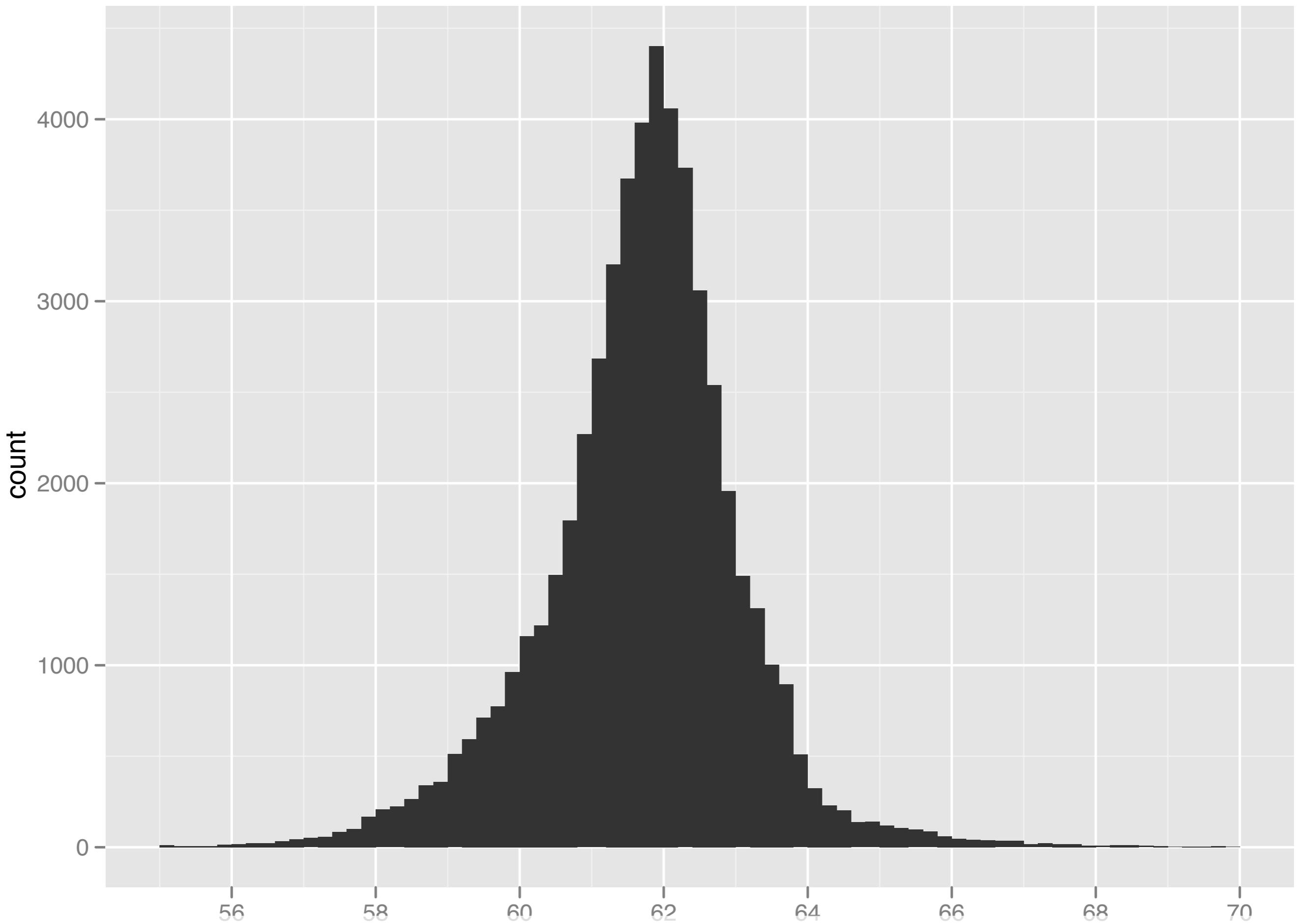




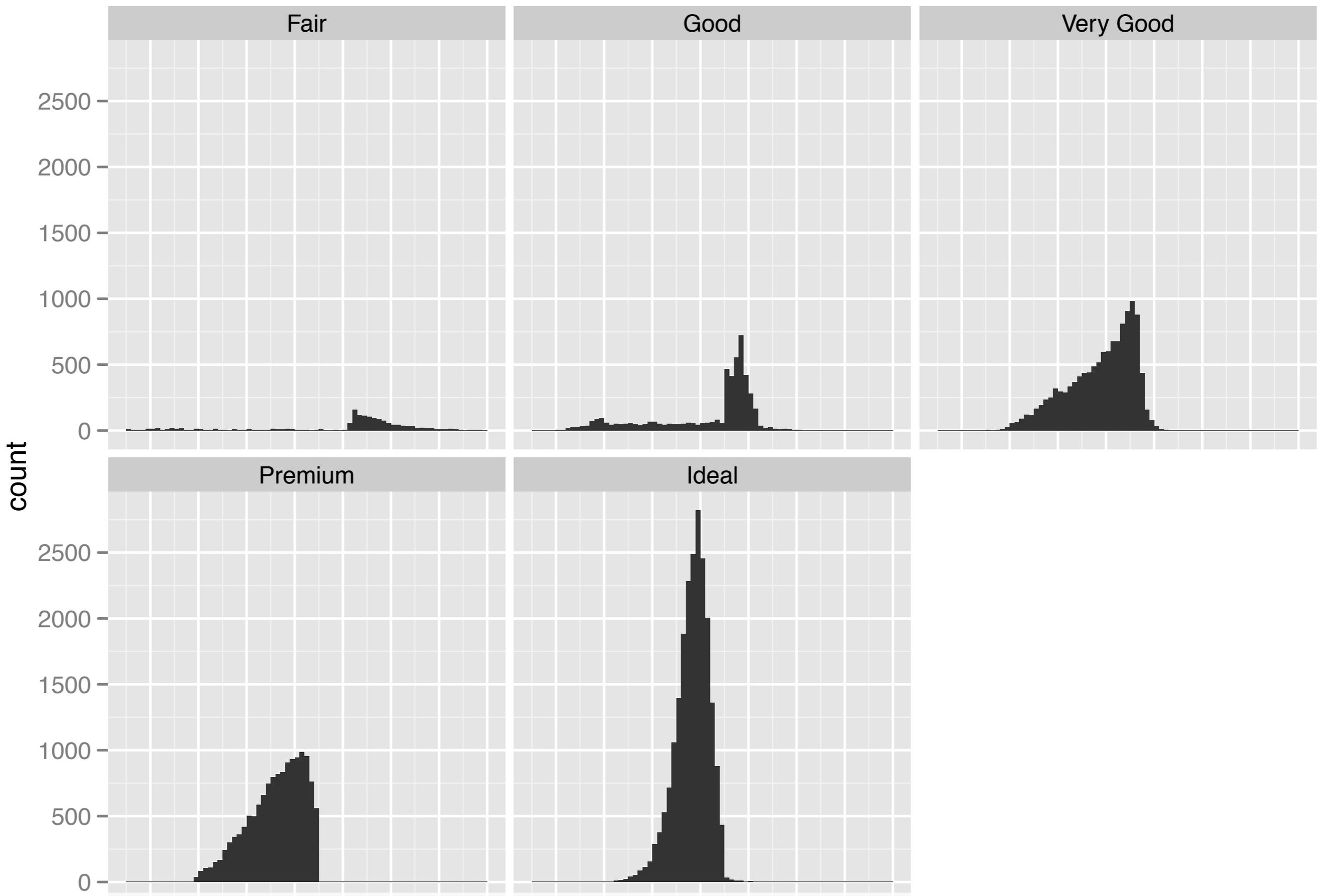
ggplot2 basics



```
qplot(displ, hwy, data = mpg)
```

```
qplot(depth, data = diamonds, binwidth = 0.2)
```



```

qplot(depth, data = diamonds, binwidth = 0.2) +
  xlim(55, 70) + facet_wrap(~cut)

```

Data manipulation

Function	Package
subset	base
summarise	plyr
mutate	plyr
arrange	plyr

Split

Apply

Combine

x	y
a	2
a	4
b	0
b	5
c	5
c	10

x	y
a	2
a	4

x	y
b	0
b	5

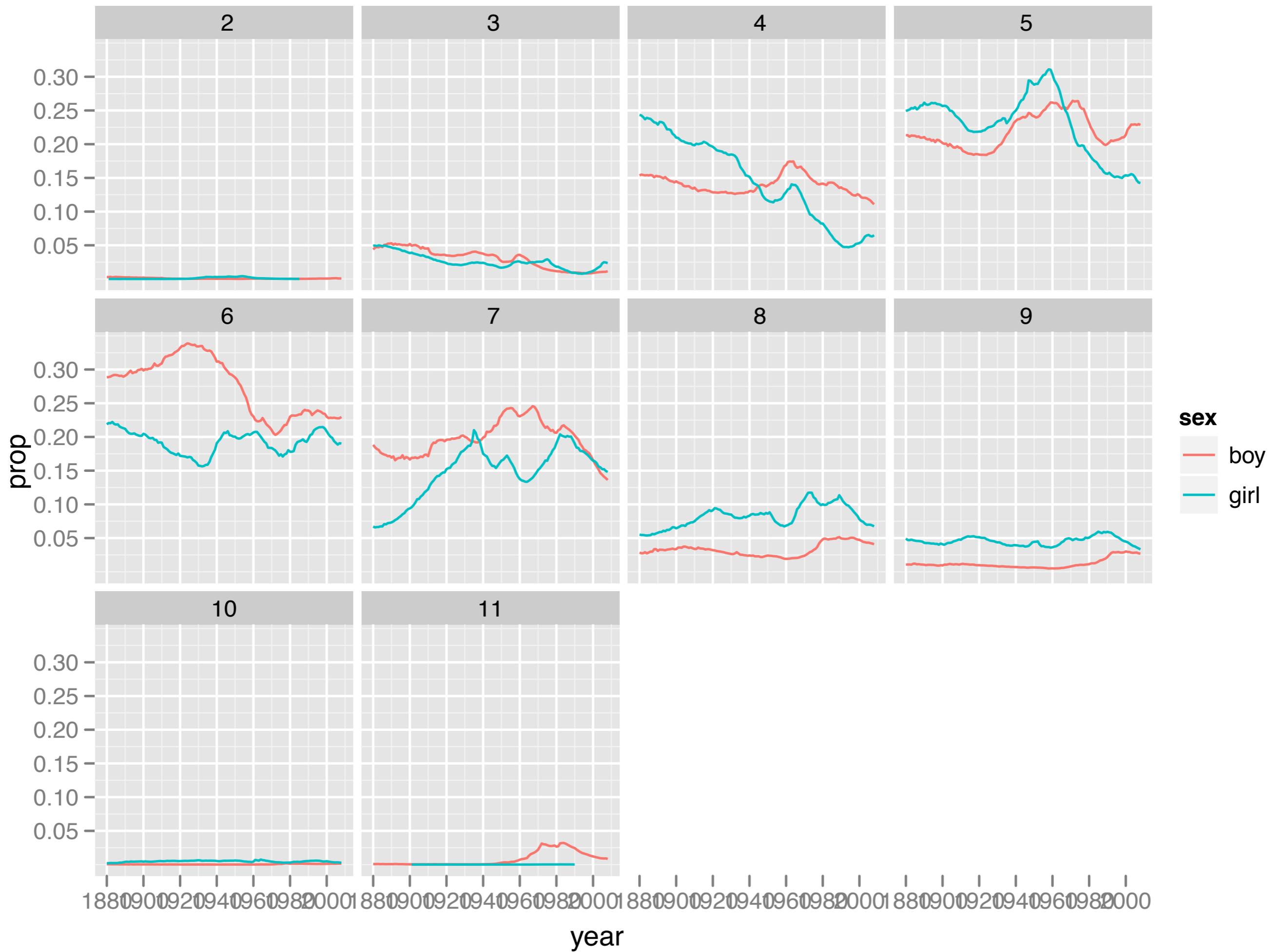
x	y
c	5
c	10

3

2.5

7.5

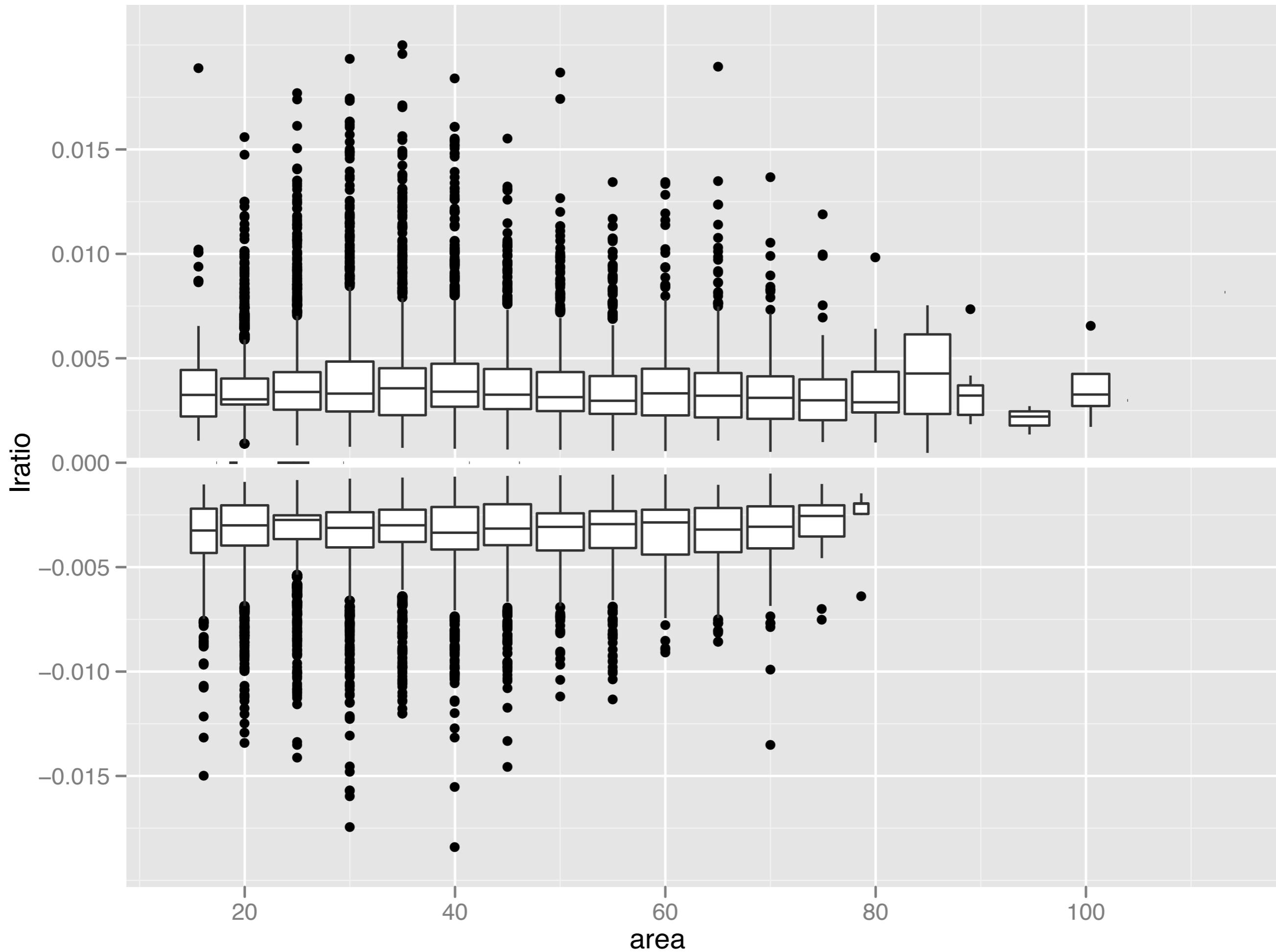
x	y
a	2
b	2.5
c	7.5



Graphics: critique and creation

What should I plot?
How can I plot it?

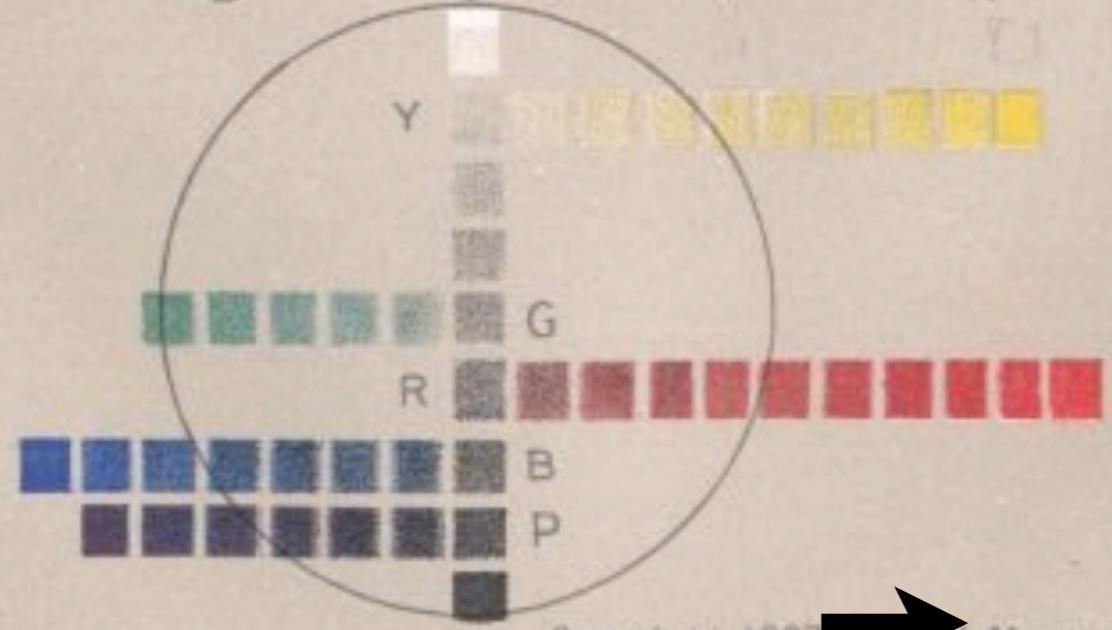
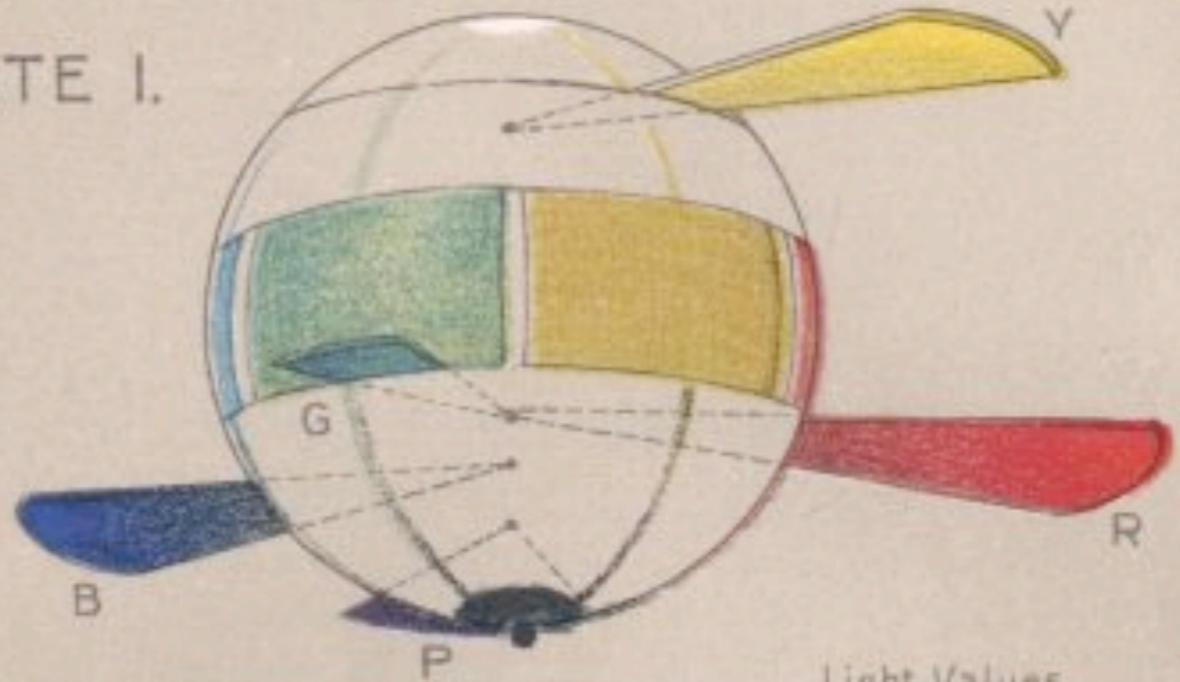




Polishing plots for presentation

1. **Scales:** used to override default perceptual mappings, and tune parameters of axes and legends.
2. **Coordinate systems:** override default Cartesian coordinate system
3. **Themes:** control presentation of non-data elements.
4. **Saving your work:** to include in reports, presentations, etc.

PLATE I.



Copyright 1907 by W. Munsell.



Learning a new
language is hard!