03 Visual Perception

36-721 Statistical Graphics and Visualization

Jerzy Wieczorek

9/8/15

Last time

- Legible graphs: image format and quality
- Comprehensible graphs: labels, titles, and annotations
- Core charts in base R
- HW 1

Today

- Quantitative comparisons: basic perceptual tasks, distance
- Grouping and search: preattentive processing, gestalt, separable dimensions, alignment
- Cognition: derived variables, ranking
- Consistency: across small multiples, in design, with semantic associations
- R: choosing color, point symbol, line type
- `text`, `matplot`, `RColorBrewer`
- `mfrow`, `layout`, `mtext`

Today

Follow along:

- Editable code in `03_Perception_code.R`
- Code with output examples in `03_Perception_code.html`

Quantitative comparisons

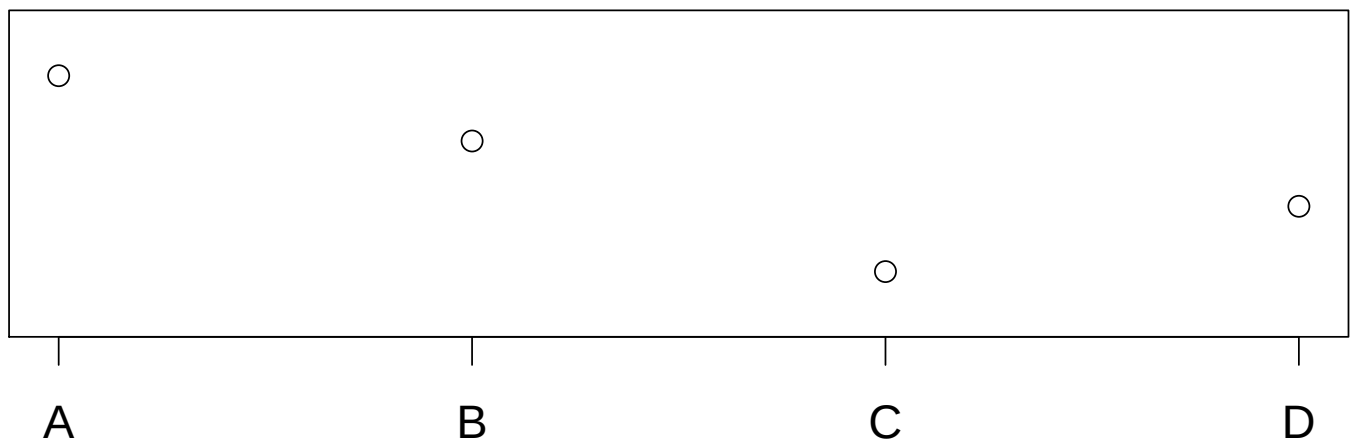
- Basic perceptual tasks
- Distance

Quantitative comparisons

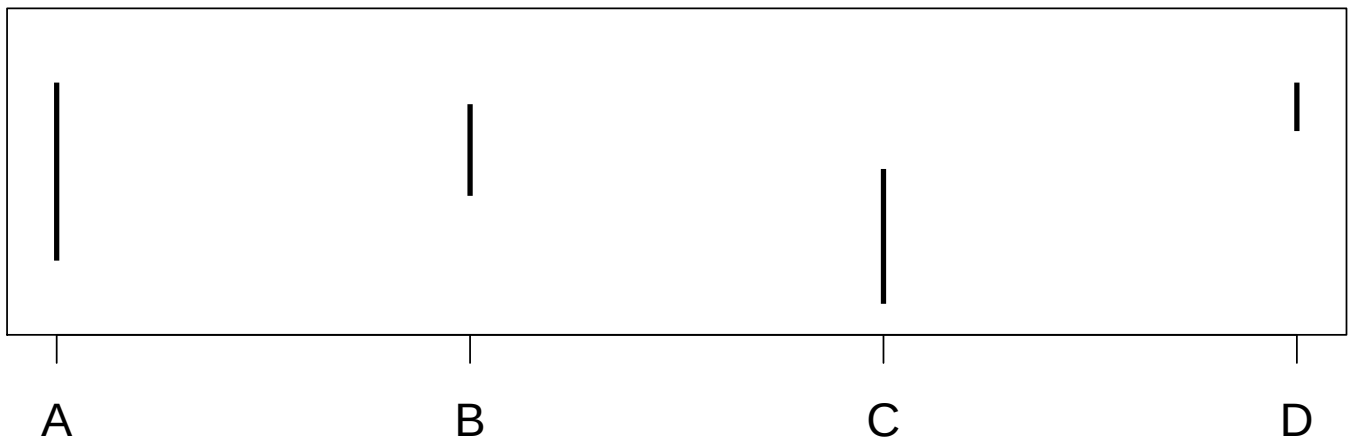
Experiment on next few slides:

	A	B	C	D
Positions	1	?	?	?
Lengths	1	?	?	?
Angles	1	?	?	?
Areas	1	?	?	?

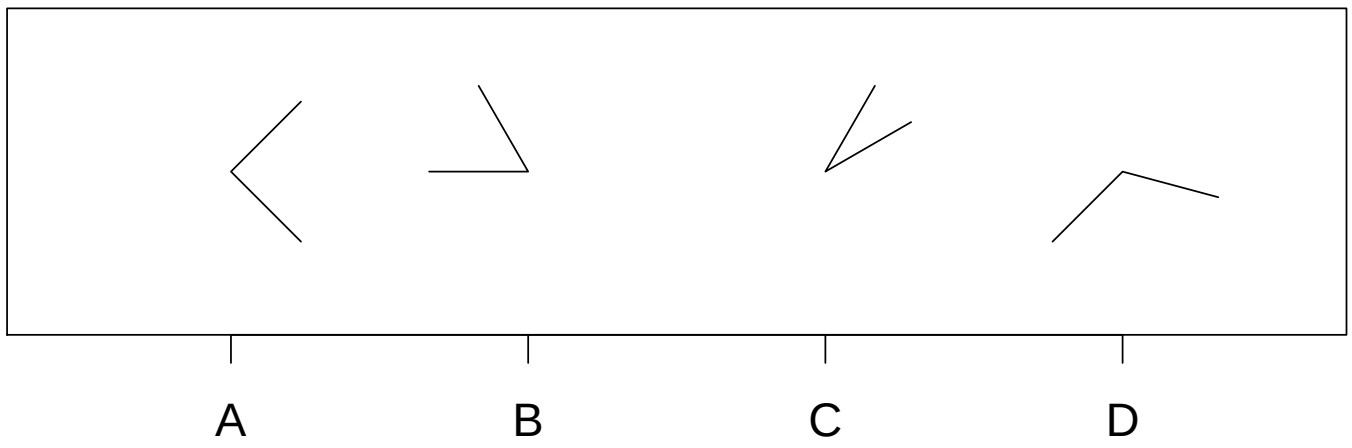
Quantitative perceptual tasks: position, aligned



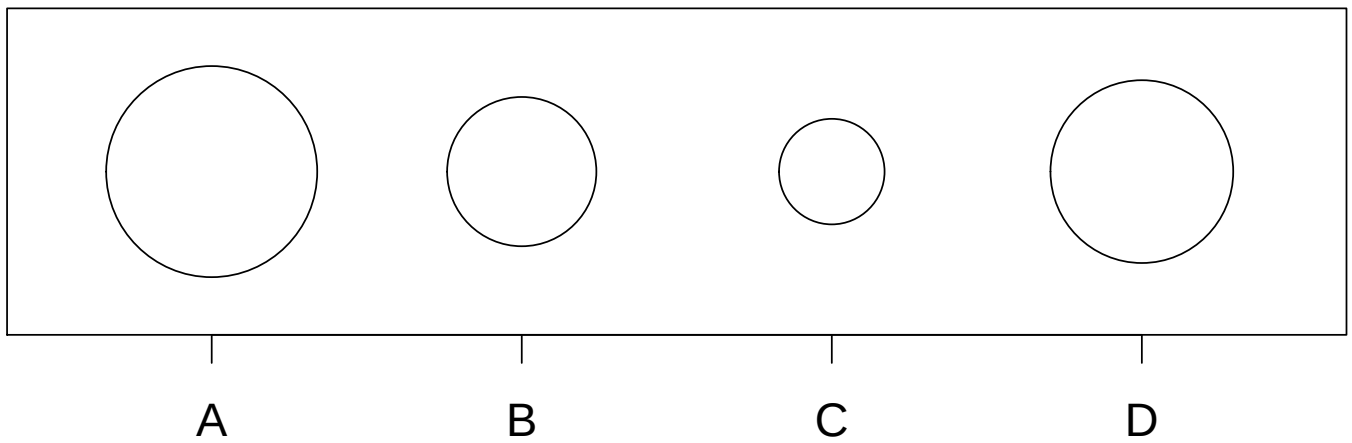
Quantitative perceptual tasks: length



Quantitative perceptual tasks: angle



Quantitative perceptual tasks: area



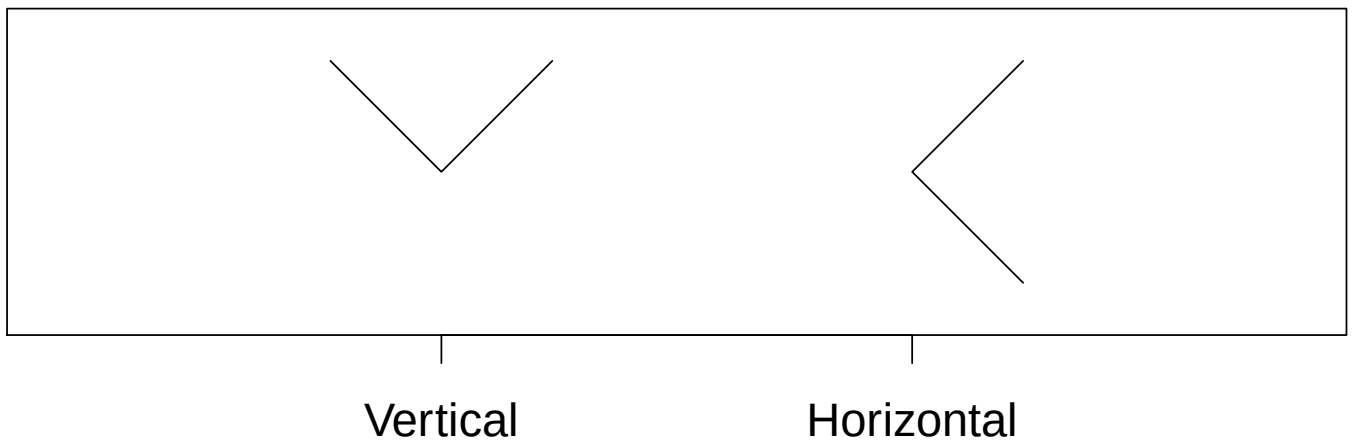
Quantitative perceptual tasks: answers

	A	B	C	D
Positions	1	3/4	1/4	2/4
Lengths	1	2/4	3/4	1/4
Angles	1	2/3	1/3	4/3
Areas	1	2/4	1/4	3/4

Cleveland and McGill (1984)

Cleveland, *The Elements of Graphing Data*

Quantitative perceptual tasks: effect of angle orientation

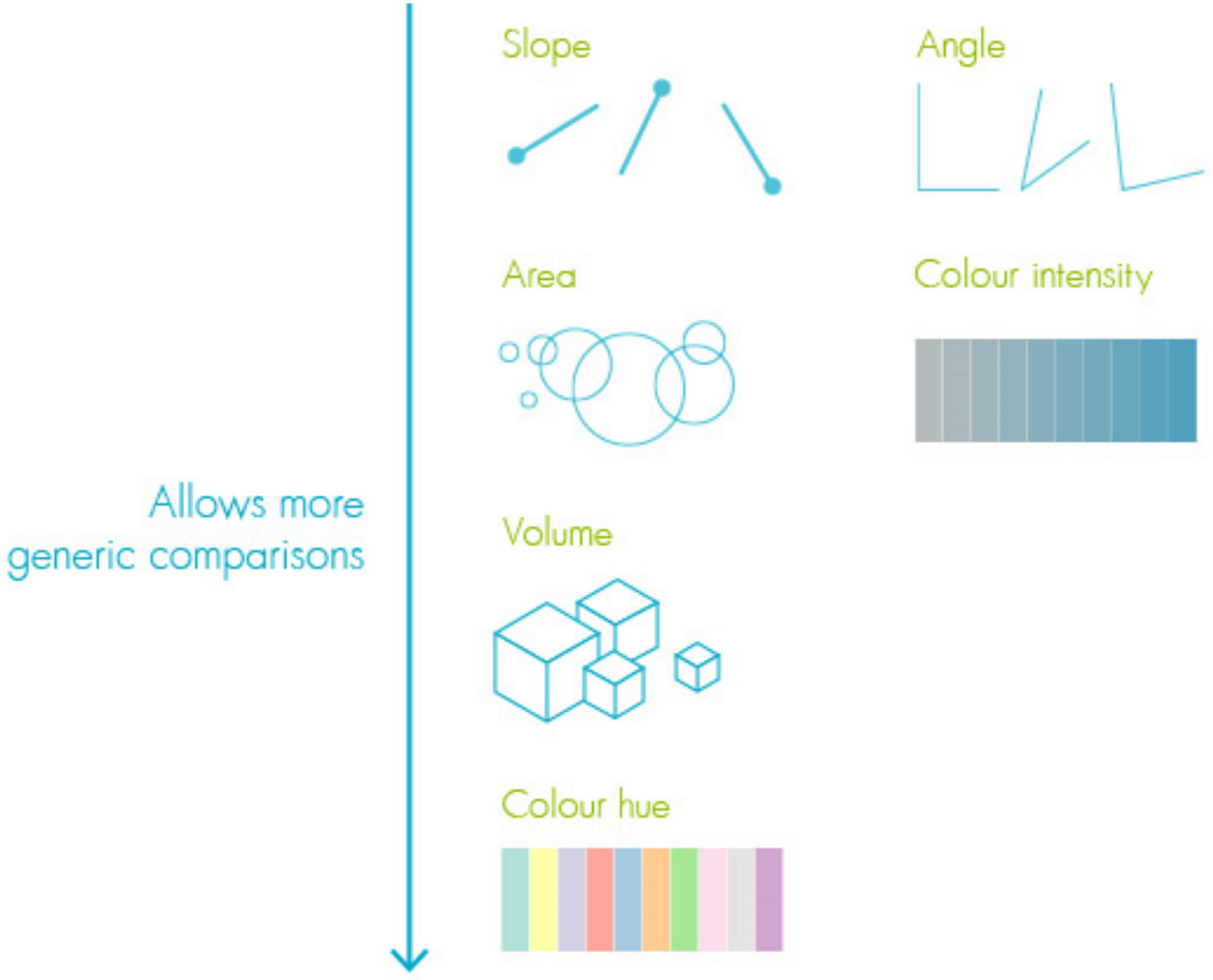


Same angle looks wider when bisector is horizontal.

Ordering of perceptual tasks

Cleveland and McGill's ordering

Ordering of perceptual tasks



Distance

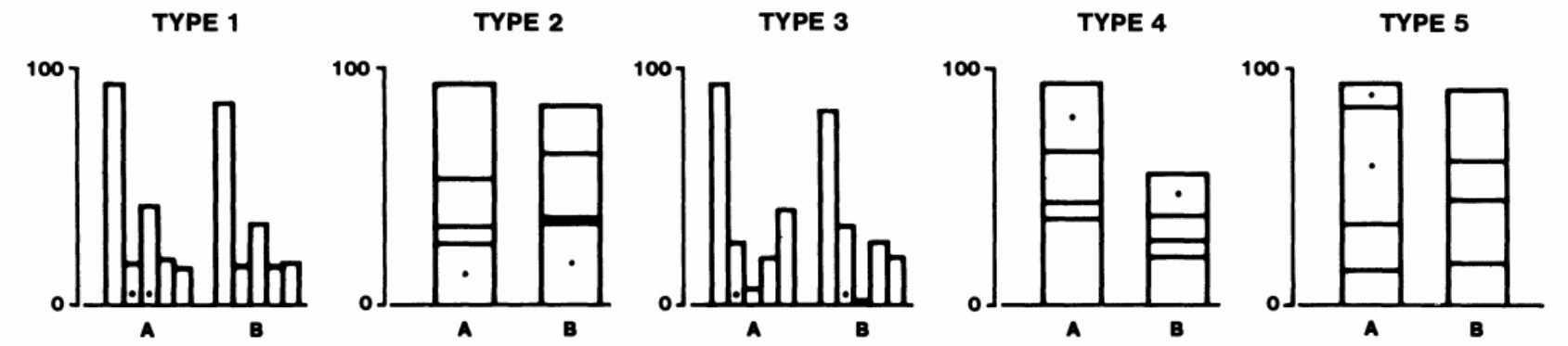


Figure 4. Graphs from position–length experiment.

Cleveland and McGill (1984)

Quantitative perceptual tasks

Lessons:

- Best to show quantitative variables with position or length
- Bars encode length, so start bars at 0; to zoom in, use dotplots (position) instead
- Avoid stacked bars (not aligned); use dots or lines (aligned baselines) instead
- Avoid pies, area, and volume entirely
- Choose and order hues sensibly; use Color Brewer
- Place things-to-be-compared near each other

Grouping and Search

- Preattentive processing
- Gestalt
- Separable dimensions
- Alignment

Preattentive processing: example task

Find and count the 6s

0 5 0 8 2 4 9 3 2 0 6 9 0 0 3 0 4 6 2 7

9 0 1 1 7 9 9 7 9 3 4 6 4 4 9 7 4 8 0 7

3 7 6 5 2 7 5 9 5 5 9 2 7 3 1 0 0 3 6 8

4 4 5 5 4 6 7 2 7 3 2 4 3 8 5 0 3 6 2 7

4 7 4 1 5 5 1 8 1 3 7 9 9 1 1 2 2 1 5 2

Preattentive processing: example task

Find and count the 6s now

0	5	0	8	2	4	9	3	2	0	6	9	0	0	3	0	4	6	2	7
9	0	1	1	7	9	9	7	9	3	4	6	4	4	9	7	4	8	0	7
3	7	6	5	2	7	5	9	5	5	9	2	7	3	1	0	0	3	6	8
4	4	5	5	4	6	7	2	7	3	2	4	3	8	5	0	3	6	2	7
4	7	4	1	5	5	1	8	1	3	7	9	9	1	1	2	2	1	5	2

Preattentive processing

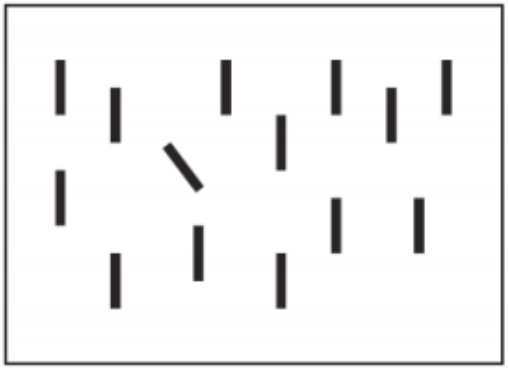
We automatically process and notice certain features, while others require conscious thought to find

We process faster when there are few categories to distinguish

Preattentive processing: features

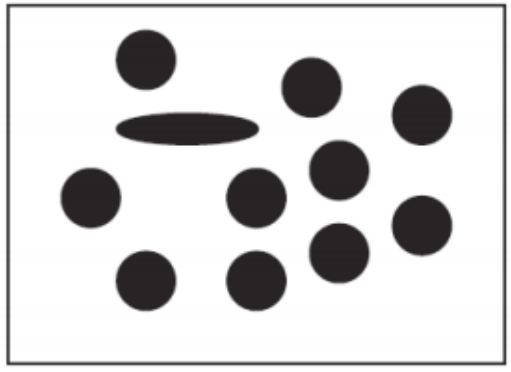
Colin Ware, *Information Visualization*

Orientation

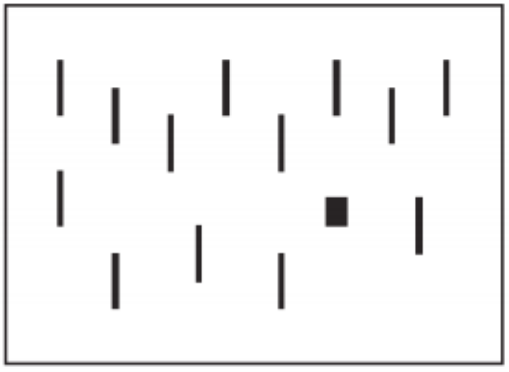


Curved/straight

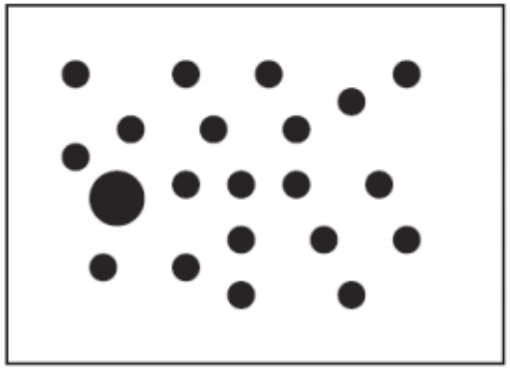
Shape



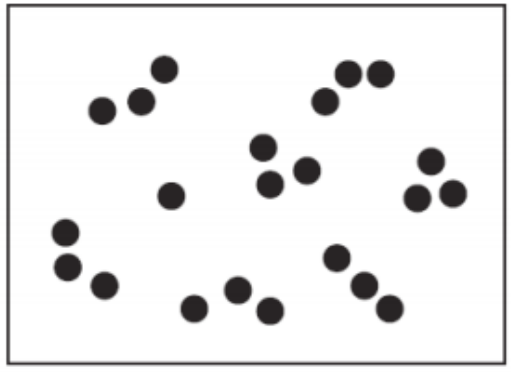
Shape



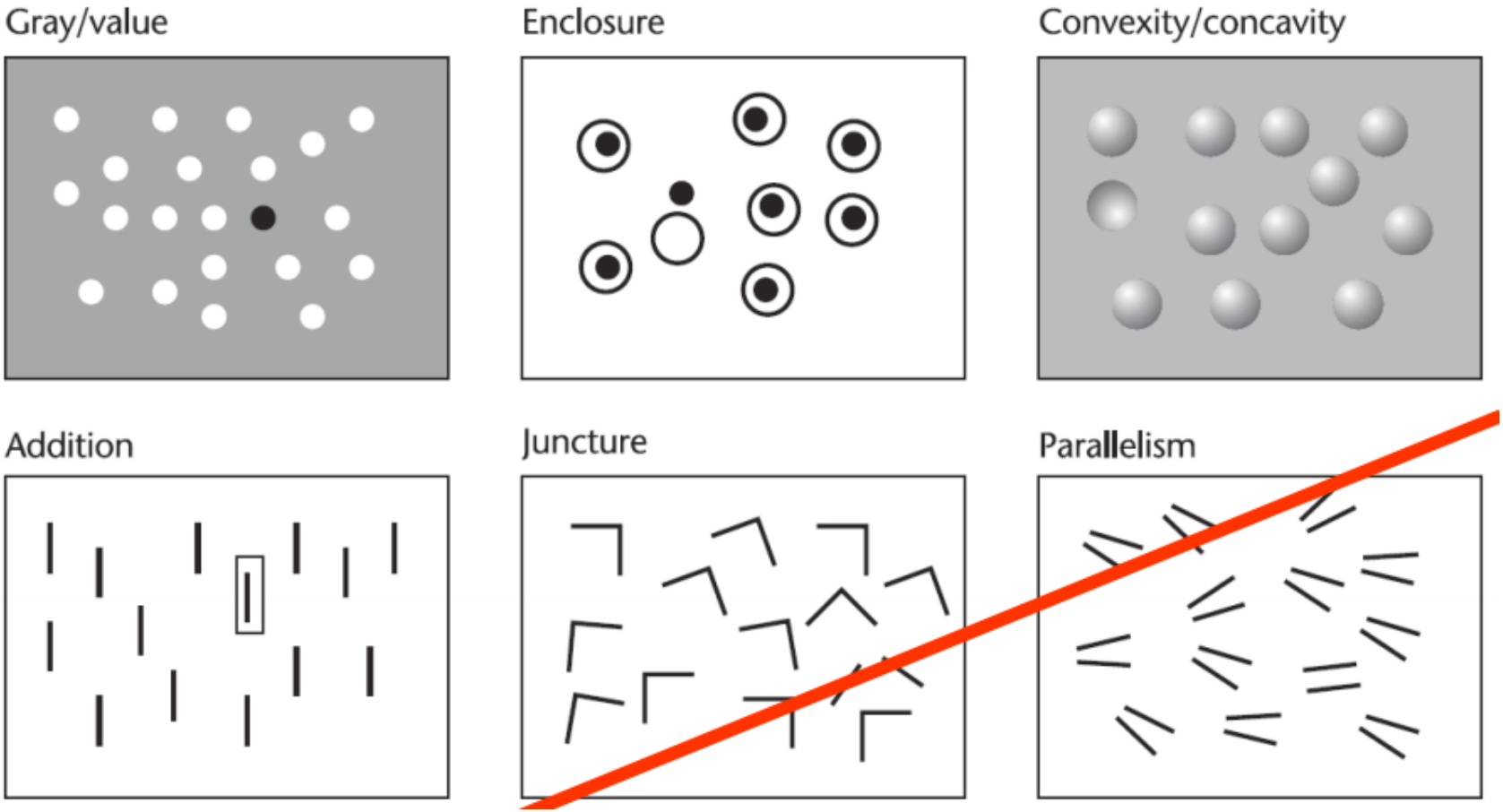
Size



Number



Preattentive processing: features



Preattentive processing

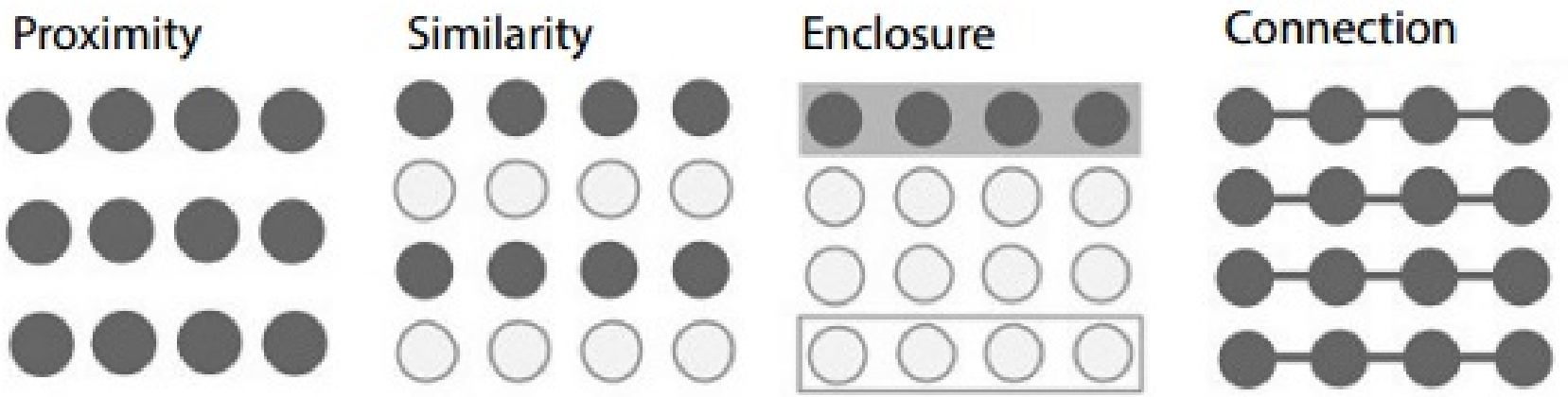
Lessons

- Distinguish categorical groups by features like hue & shape
- Hue also lets you use direct labels instead of a legend
- Don't try to show too many groups on one plot; use small multiples to show more sub-groups
- If highlighting one group, use a preattentive attribute

Gestalt

Gestalt = "pattern" in German

We automatically structure data into patterns / groups using certain features



Gestalt

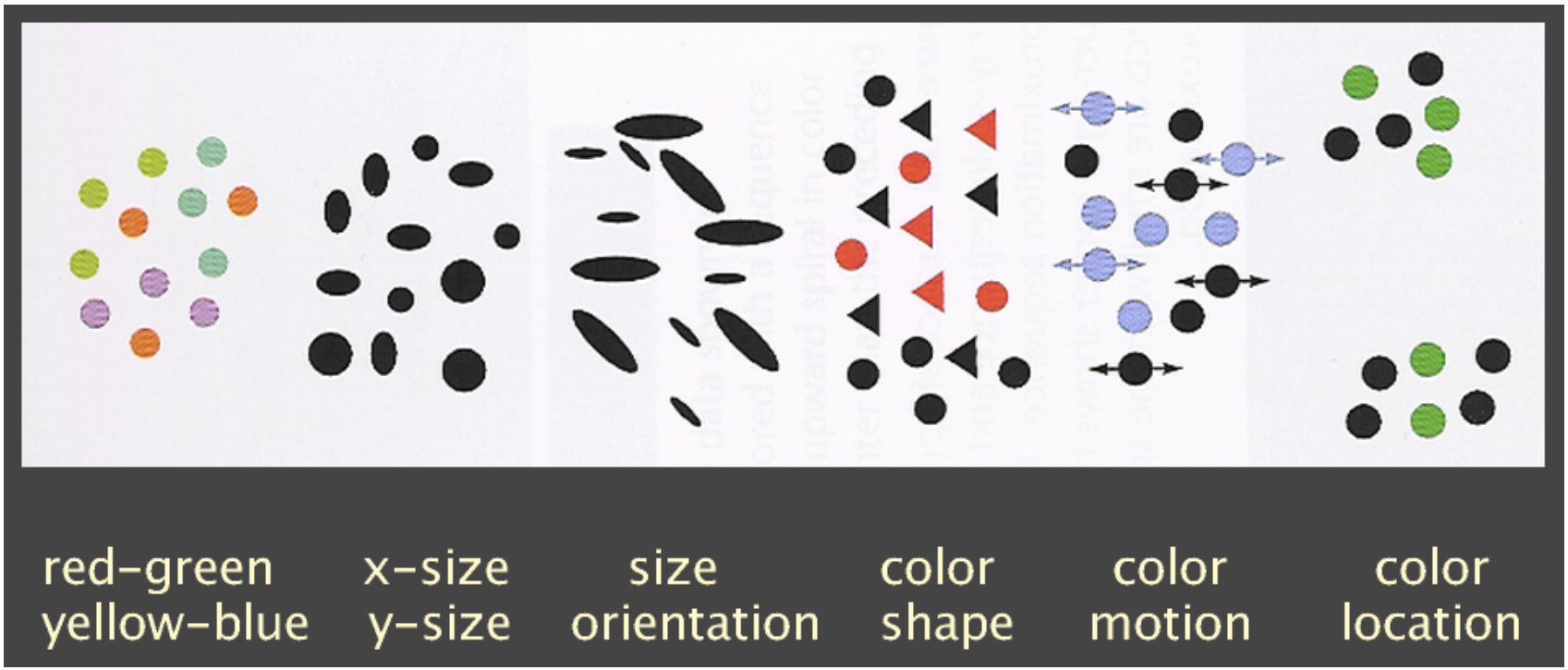
Lessons

- Distinguish categorical groups by similarity, proximity, or enclosure
- Use proximity to structure your layout (arrange small multiples)
- Use connection to show groups on line chart, parallel coordinates chart, network graph, etc.
- To highlight one group, use gestalt principles such as enclosure or similarity

Separable dimensions

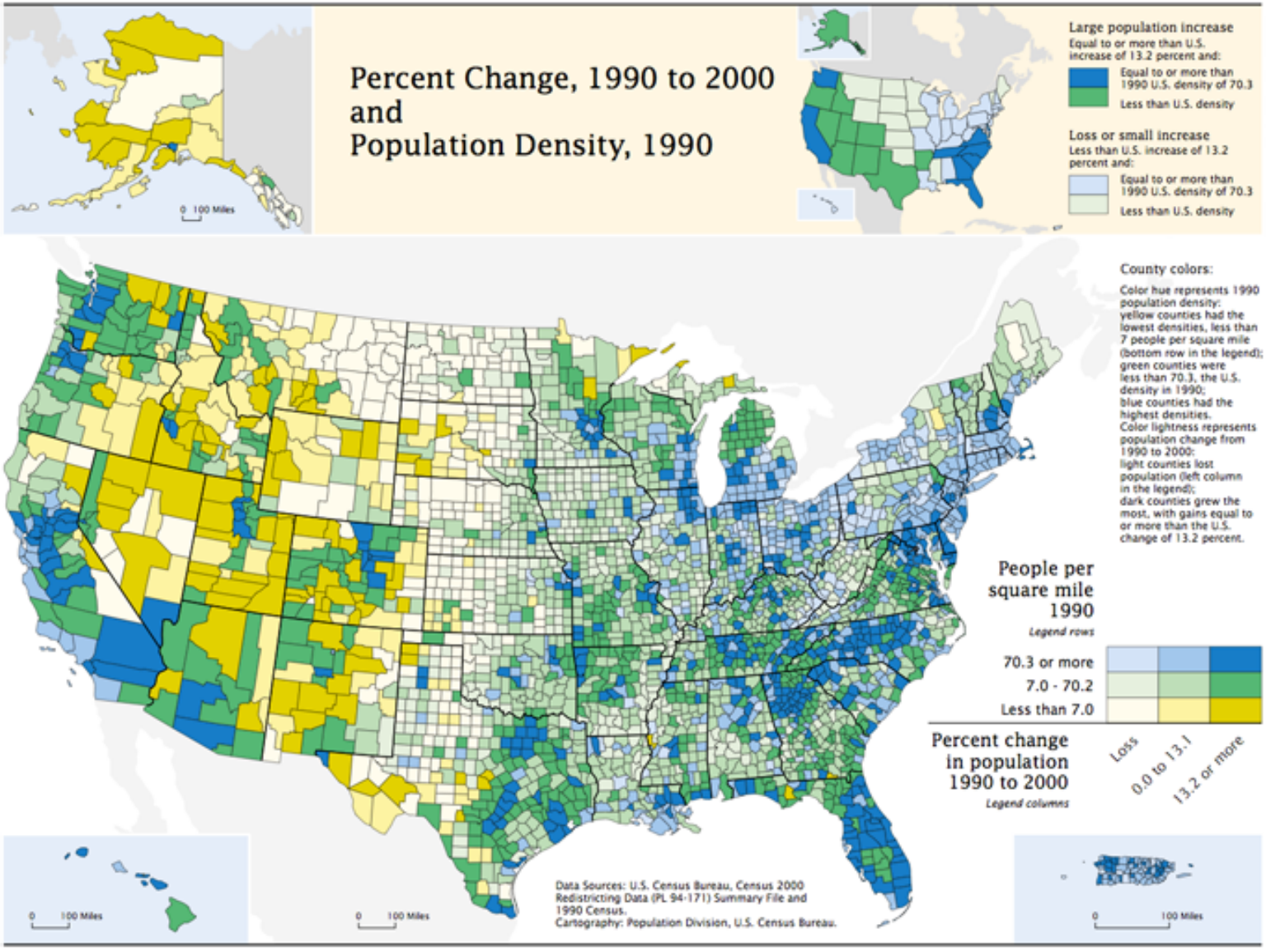
Some examples from Colin Ware, *Information Visualization*

<- More integral . . . More separable ->



Integral dimensions example

US Census Bureau map using hue and saturation



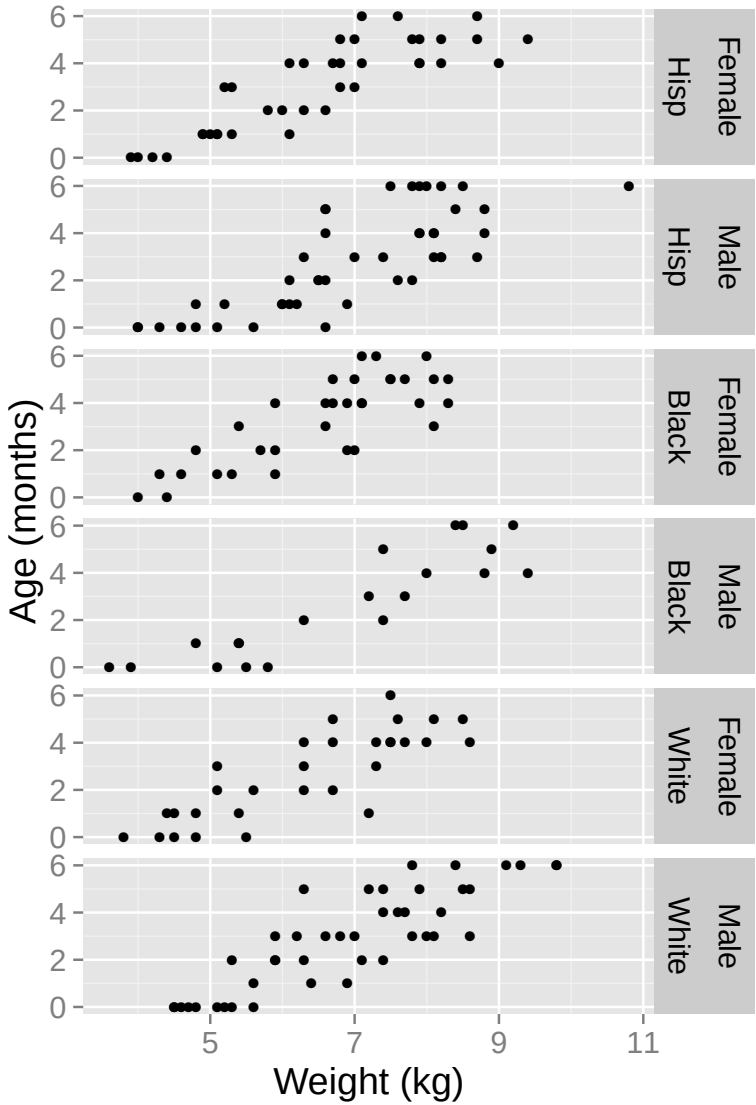
Separable dimensions

Lessons

- Use color and another variable (shape, size, orientation, motion)
- Use small multiples rather than different plotting symbols
- Avoid mixing 2 aspects of color, or 2 aspects of size
- Don't combine too many grouping variables at once

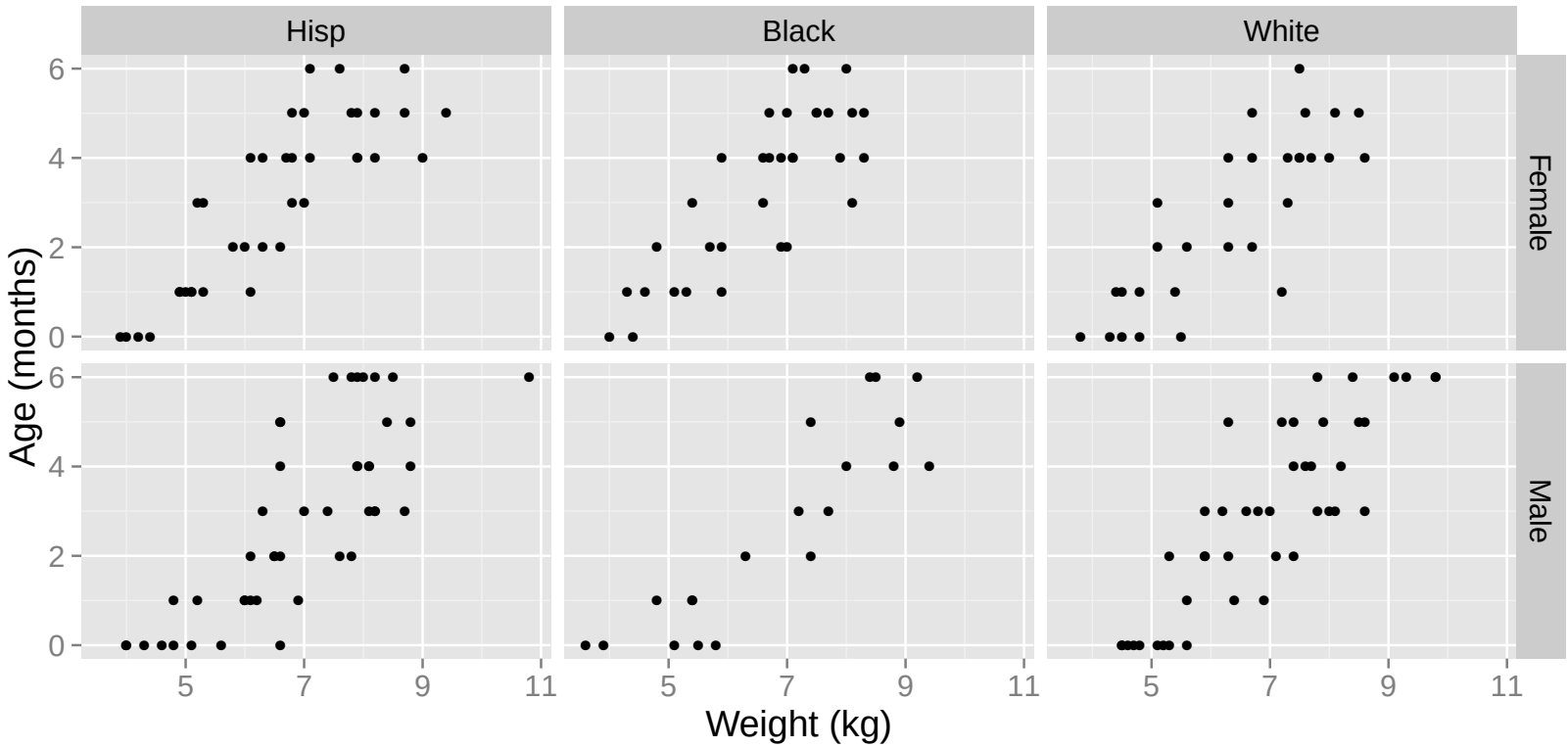
Alignment

Among male newborns, compare by race



Alignment

Among male newborns, compare by race:
easier search now, though harder comparison



Alignment

Lessons

- Decide on visual task, and helpfully align elements to be compared
- During EDA, try several arrangements

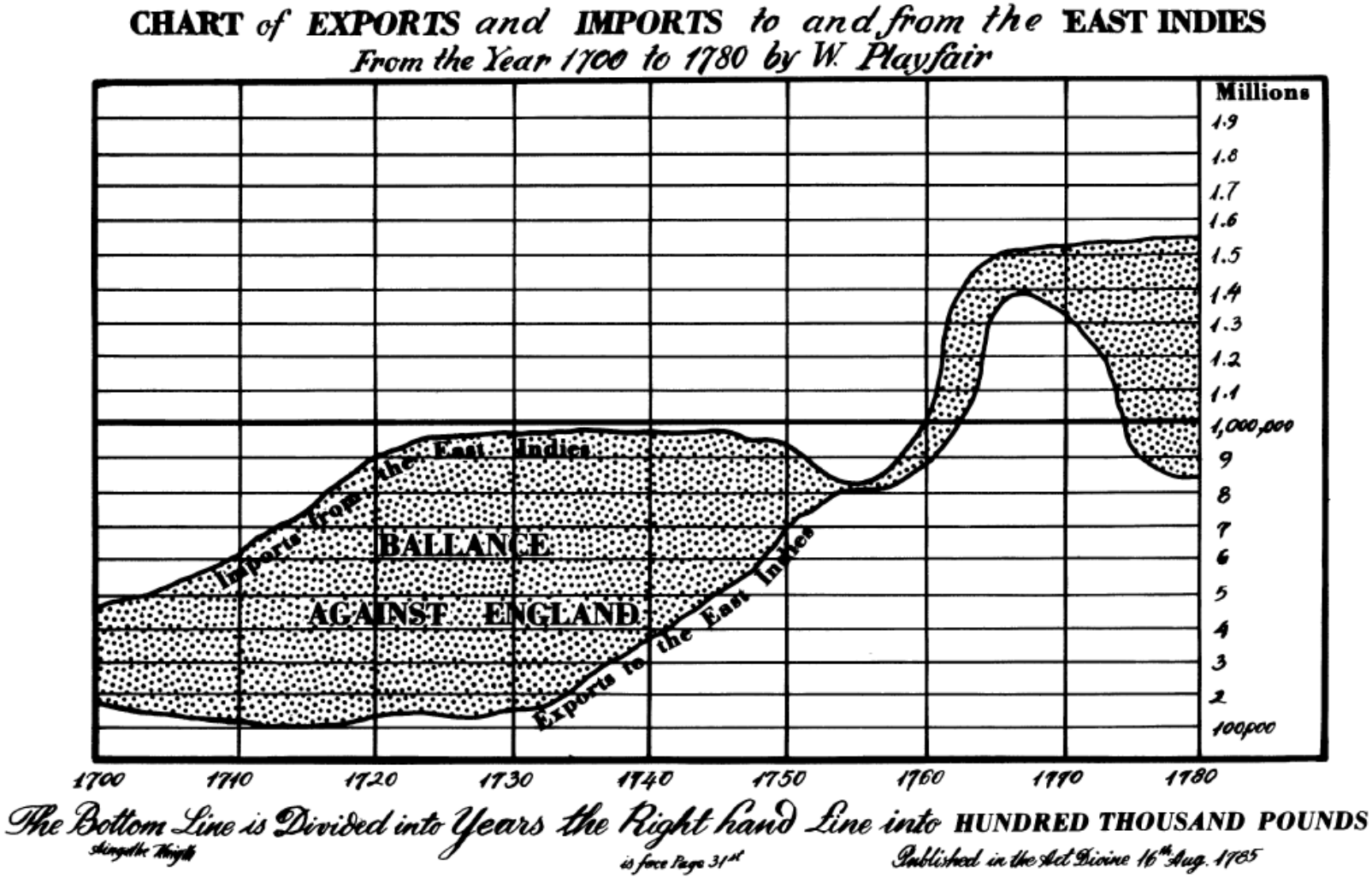
Cognition

- Derived variables
- Ranking

Derived variables

William Playfair, one of the earliest line charts

What does the difference look like?



Derived variables

Differences shown directly, by Cleveland and McGill

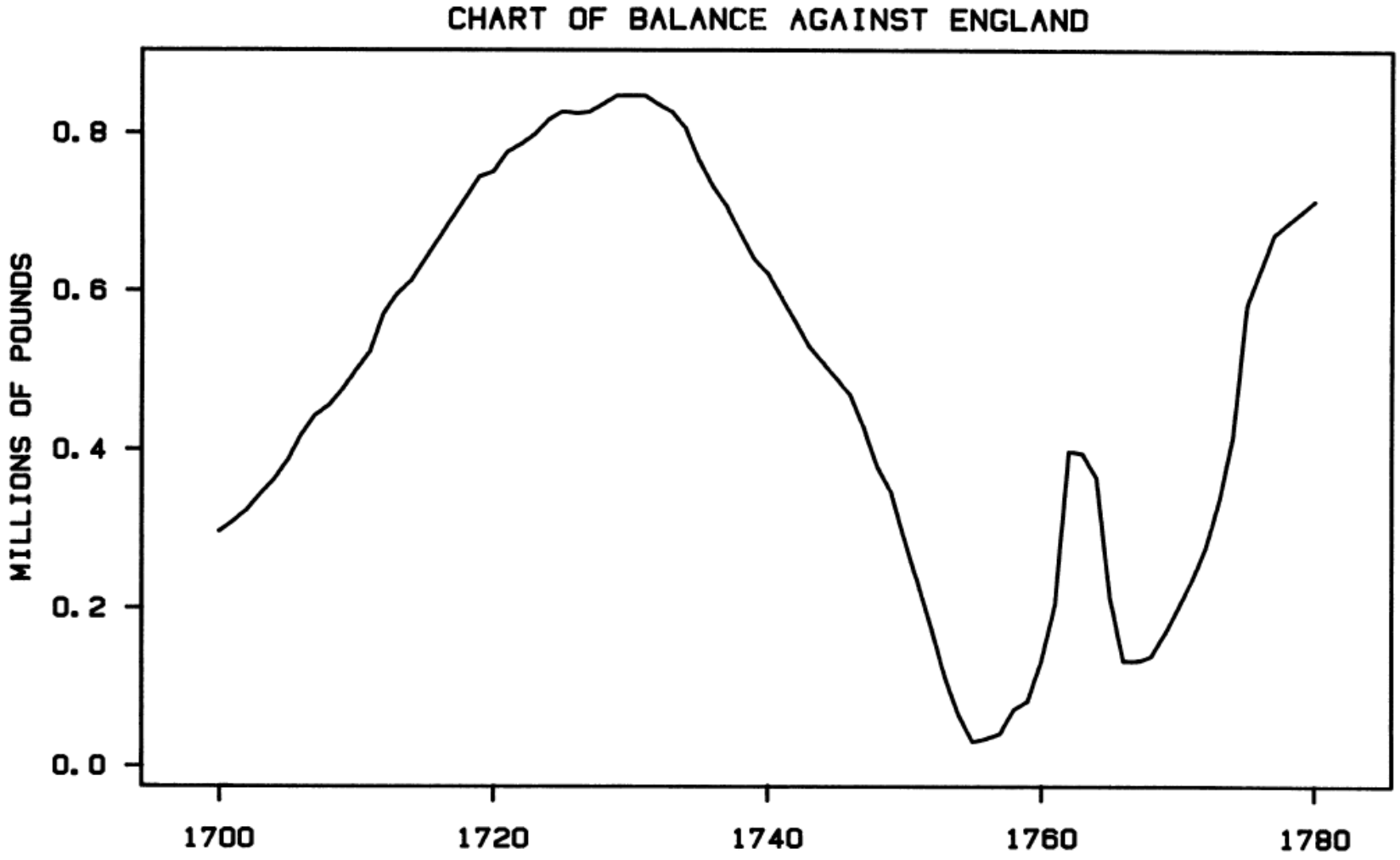
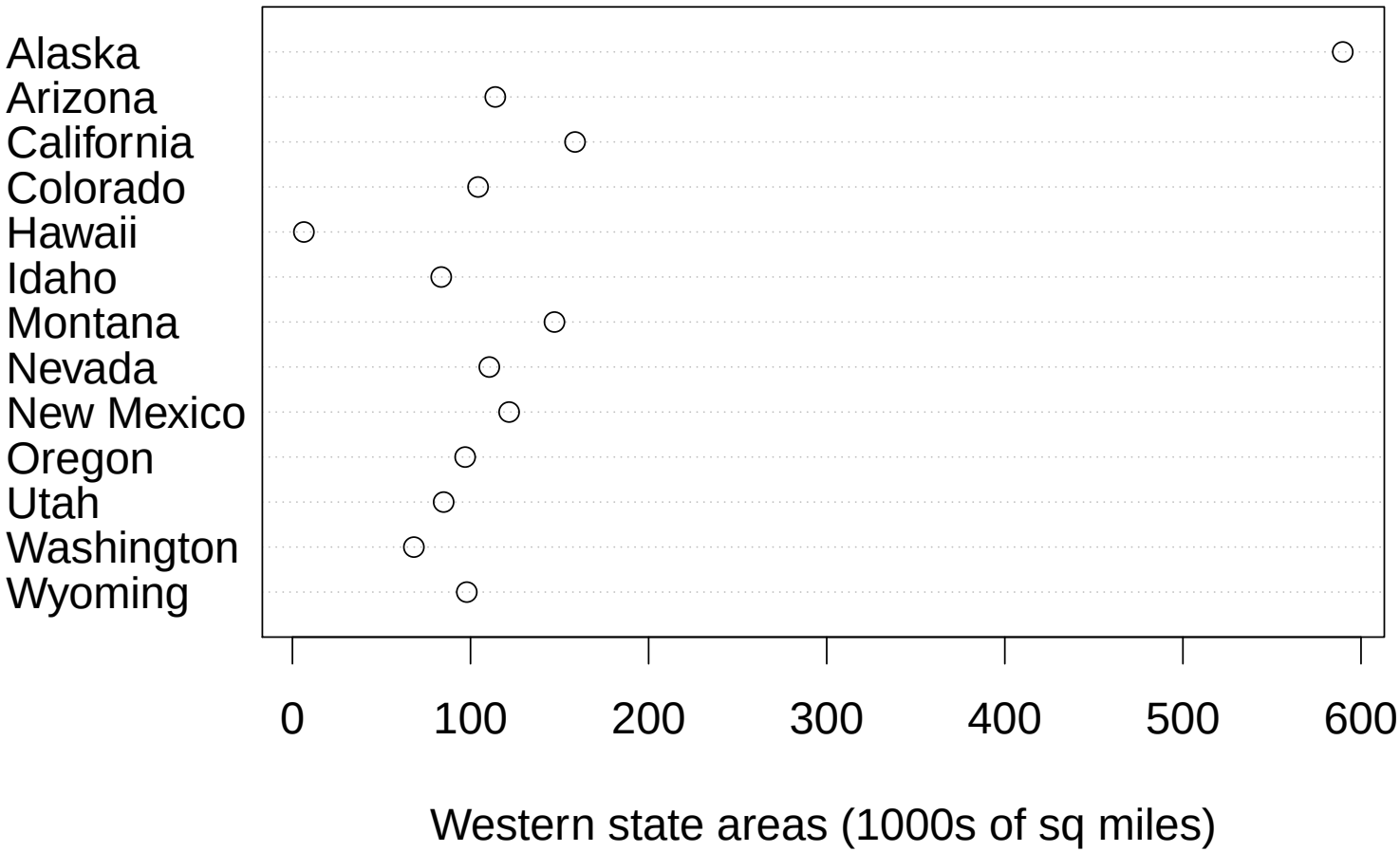
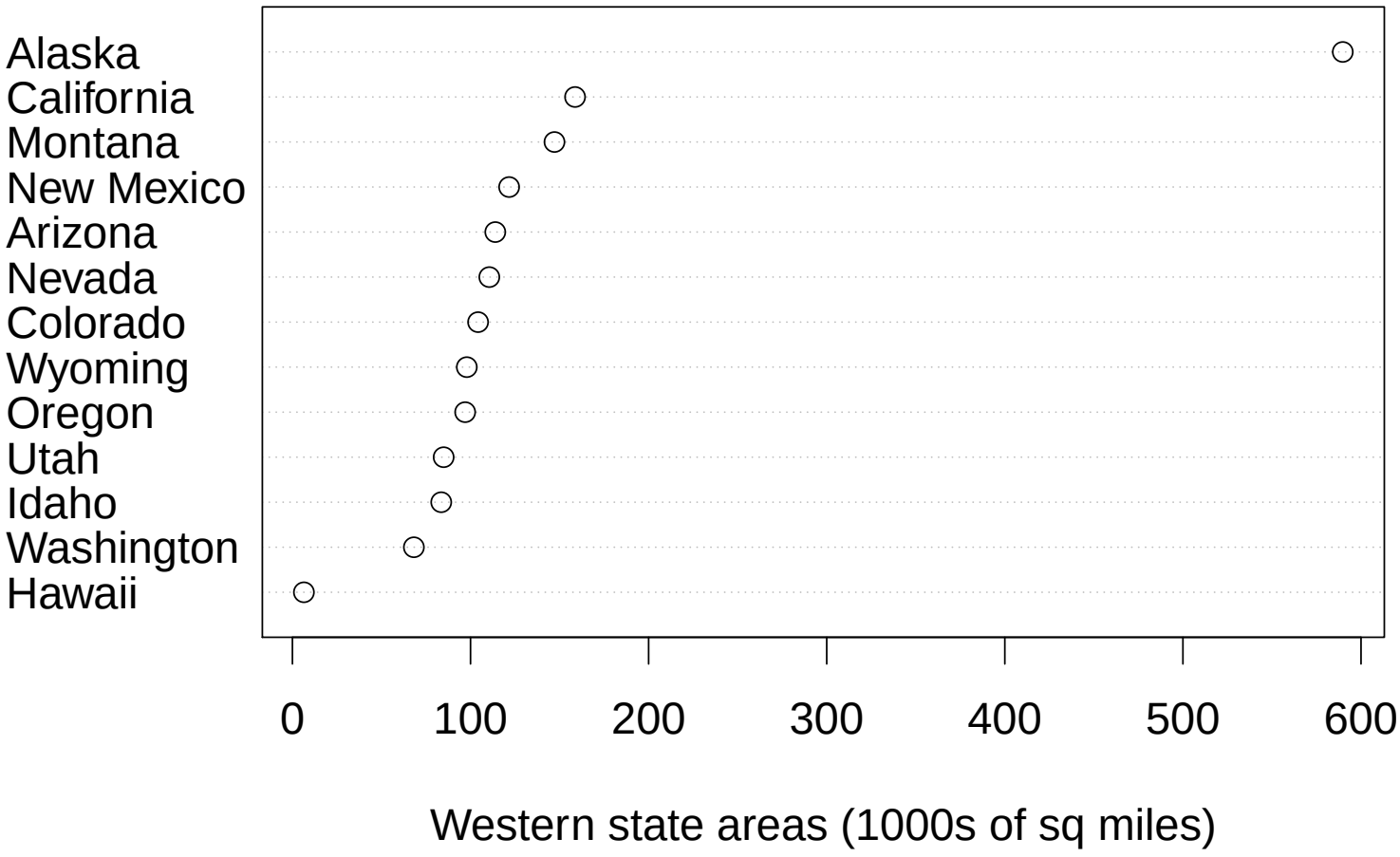


Figure 28. Playfair data.

Ranking: alphabetical



Ranking: informative



Derived variables and Ranking

Lessons

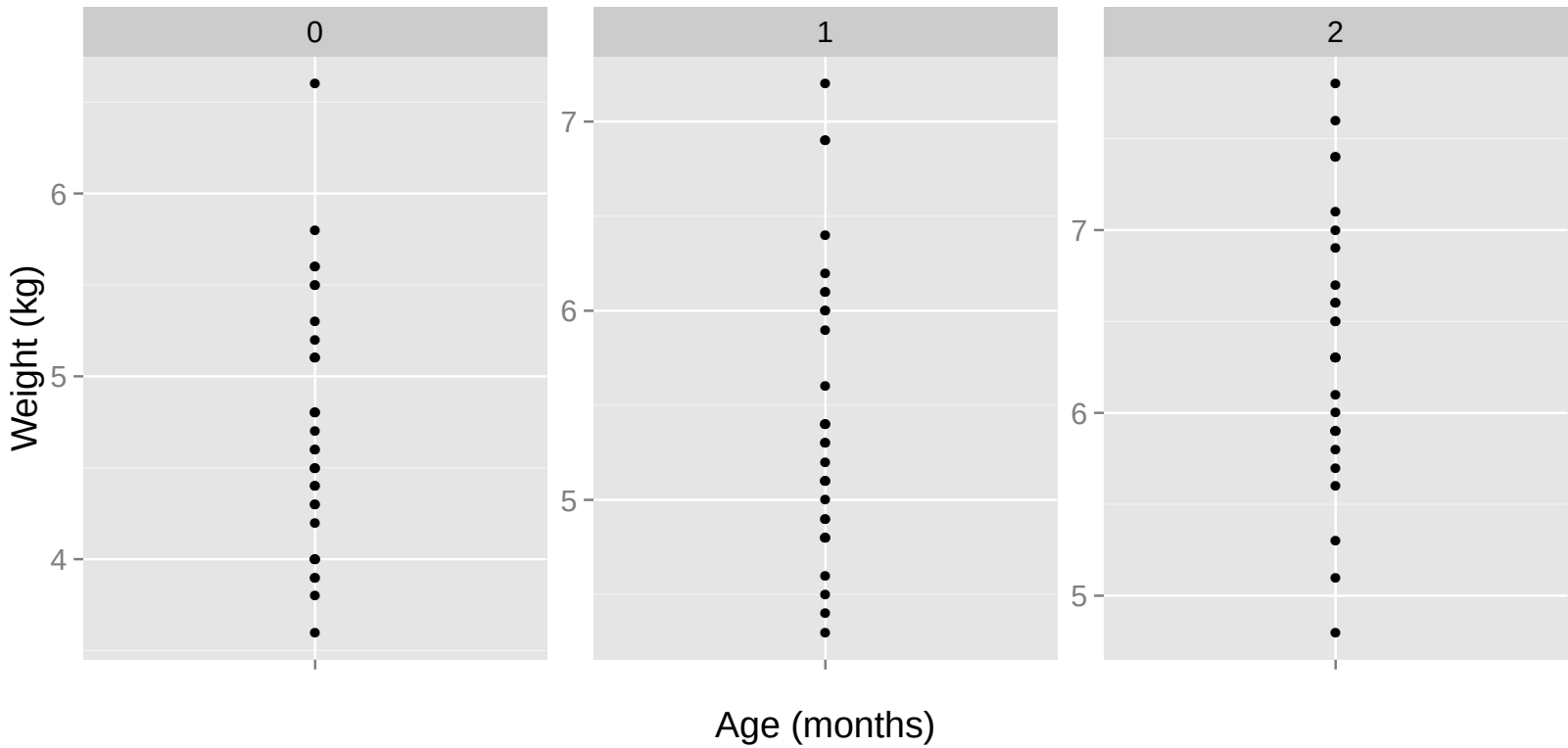
- If differences or ratios are interesting, compute and plot them directly
- Order your dots/bars meaningfully: ranked by a variable, not alphabetical

Consistency

- Across small multiples
- In design
- With semantic associations

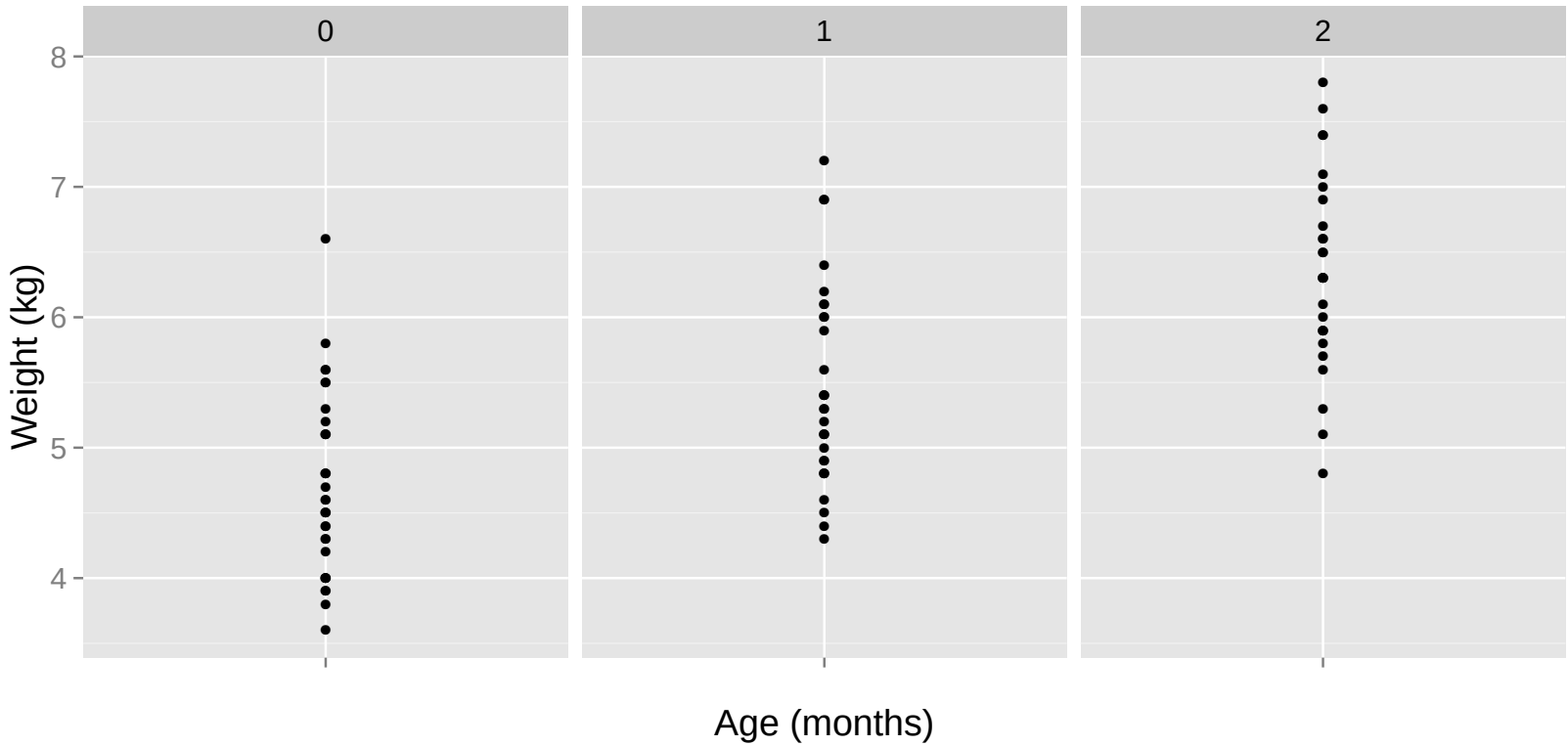
Consistency

Which age group weighs the least?



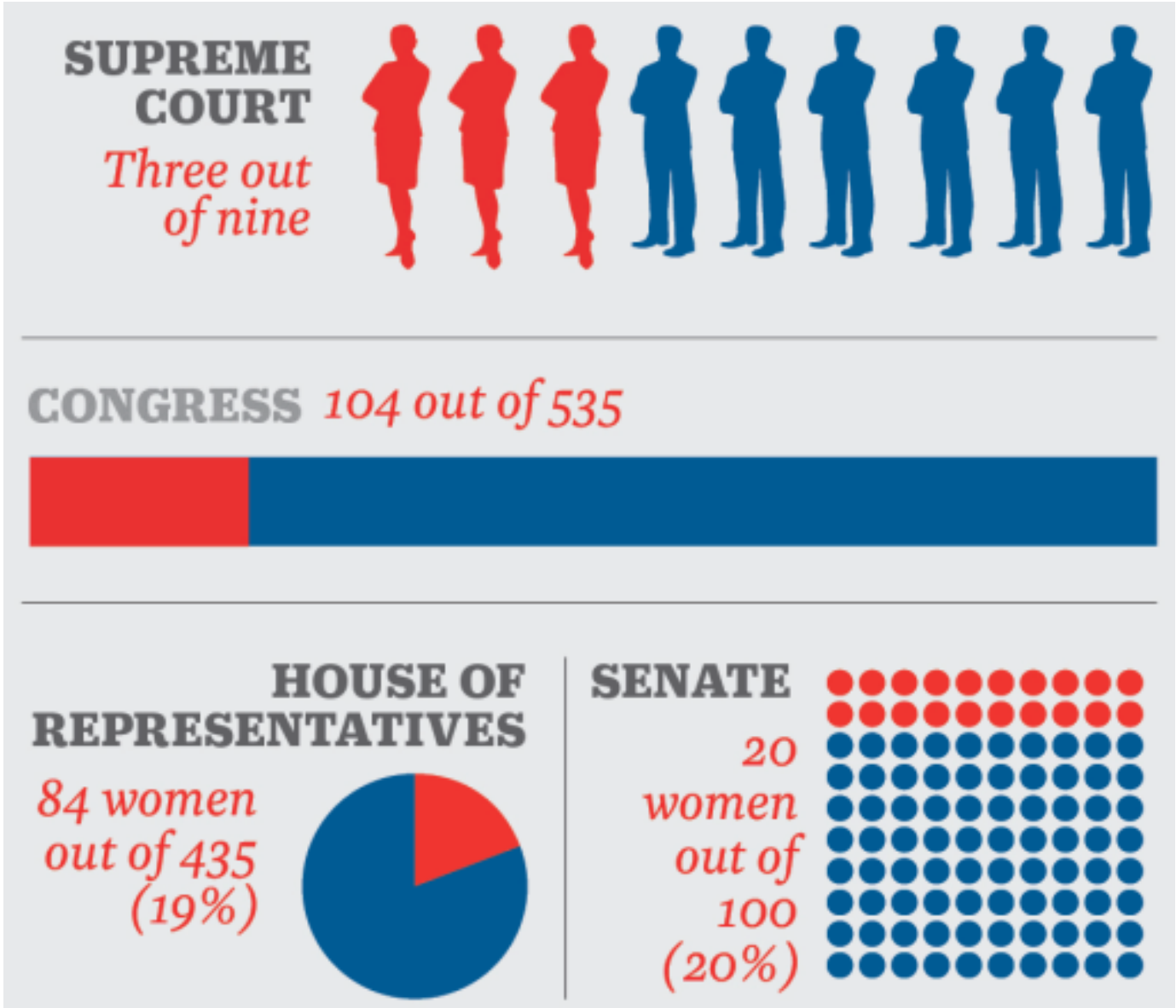
Consistency

Give all small multiples the same structure, usually **including axis limits,** to make comparisons easier and reduce cognitive load



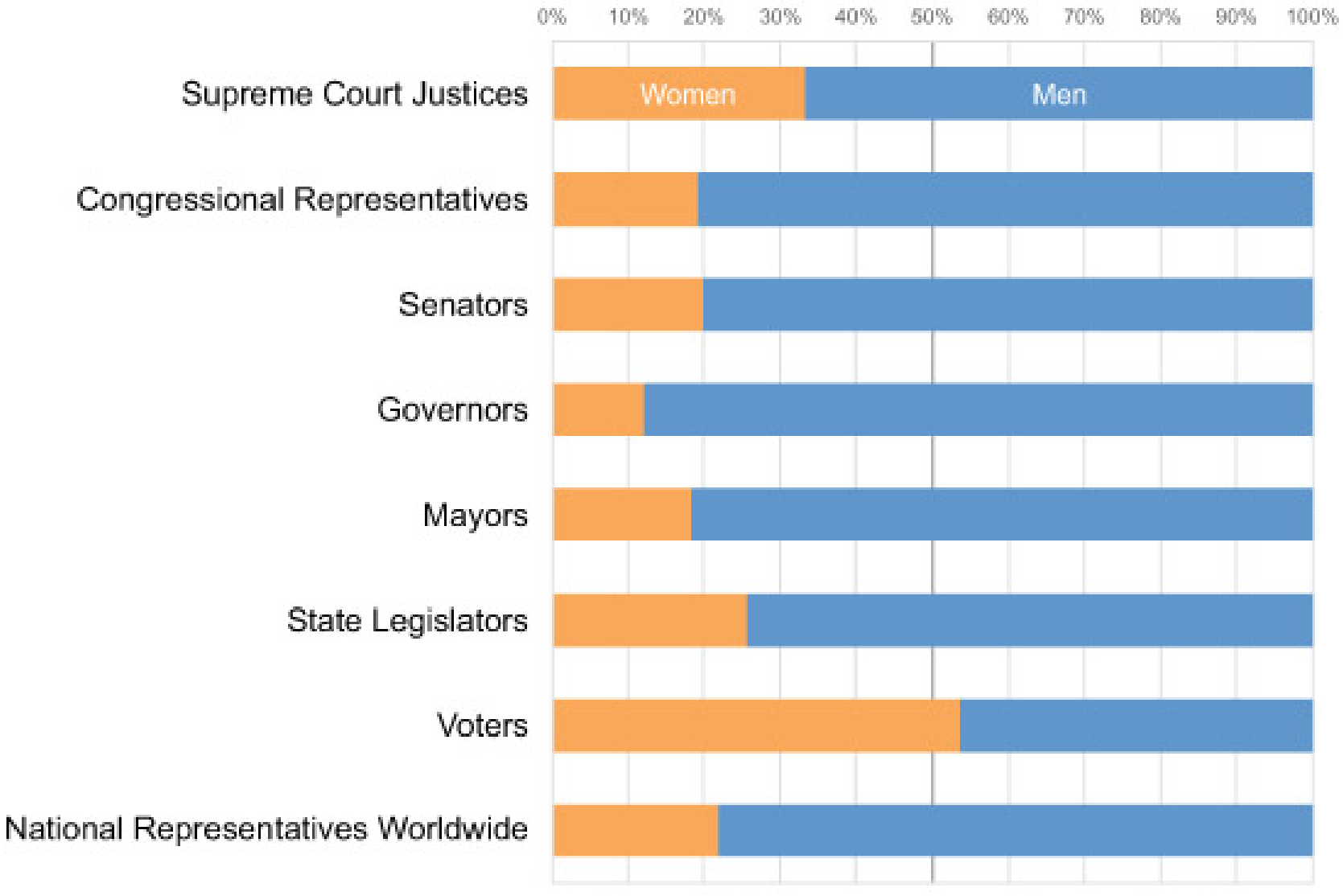
Consistency

Ensure design changes are meaningful (tied to data changes)



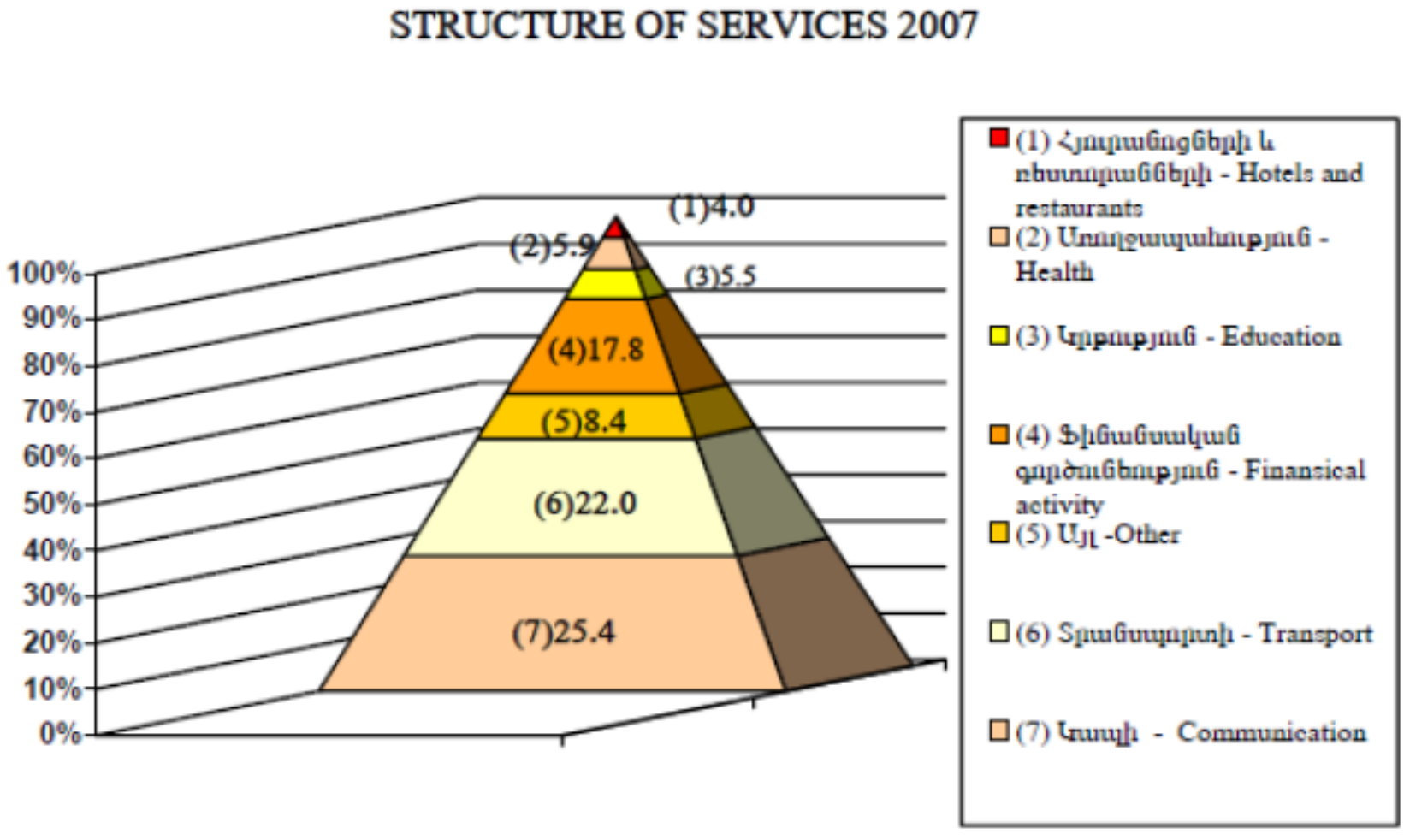
Consistency

More consistent redesign, Stephen Few



Consistency

Avoid meaningless visual variables like shadow or 3D



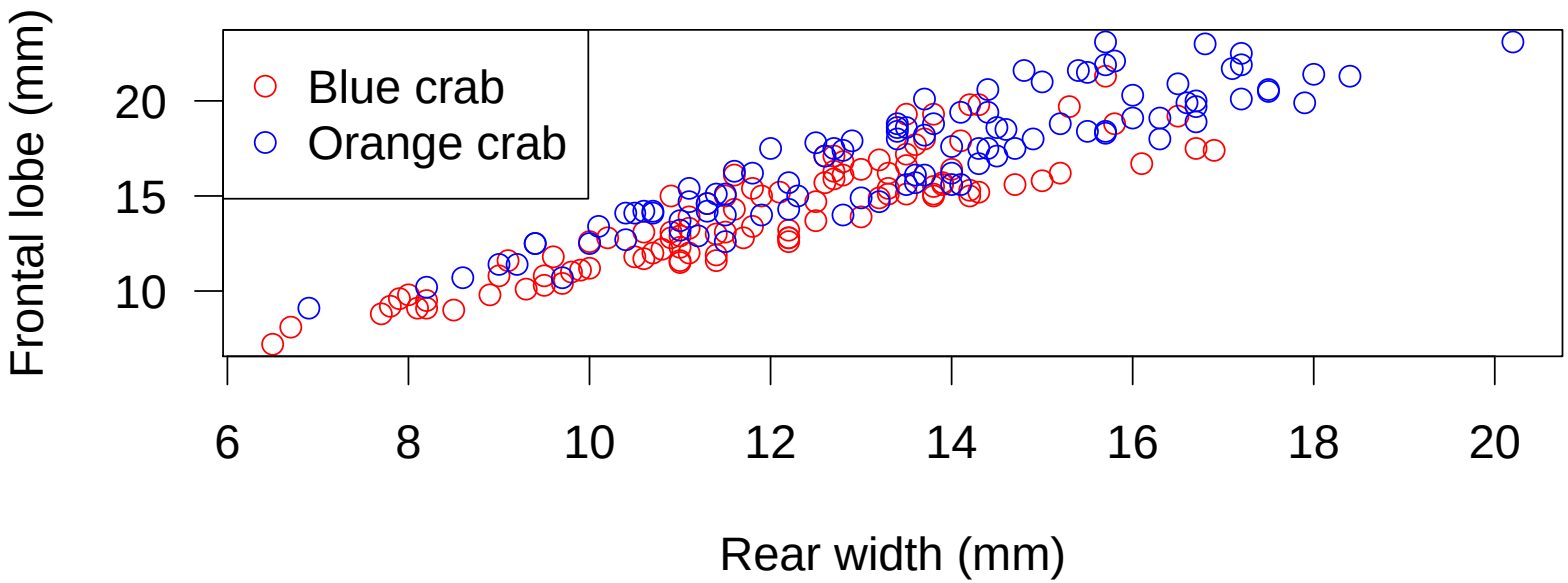
Consistency

Lessons

- Use consistent mappings (colors and shapes, axis limits) across graphs
- Don't reuse same mappings for a different data variable
- Avoid meaningless variety in design
- Avoid shadow, 3D, and other variables not mapped to data

Semantic associations

Orange vs blue crab species: I've seen this in a talk
(`crabs dataset` in `MASS` package)



Semantic associations

Lessons

- Use meaningful mappings: orange vs blue crab species = orange and blue symbols
- Use conventional mappings: blue = cold, red = hot
- "More = more": deeper saturation or larger size = higher value of variable

R

- Choosing color, point symbol, line type
- `text`, `matplot`, `RColorBrewer`
- `mfrow`, `layout`, `mtext`

For next time

- We'll cover the Grammar of Graphics framework, and how it is the basis for `ggplot2` and Tableau
- HW 2 due Saturday at 5pm, through Blackboard