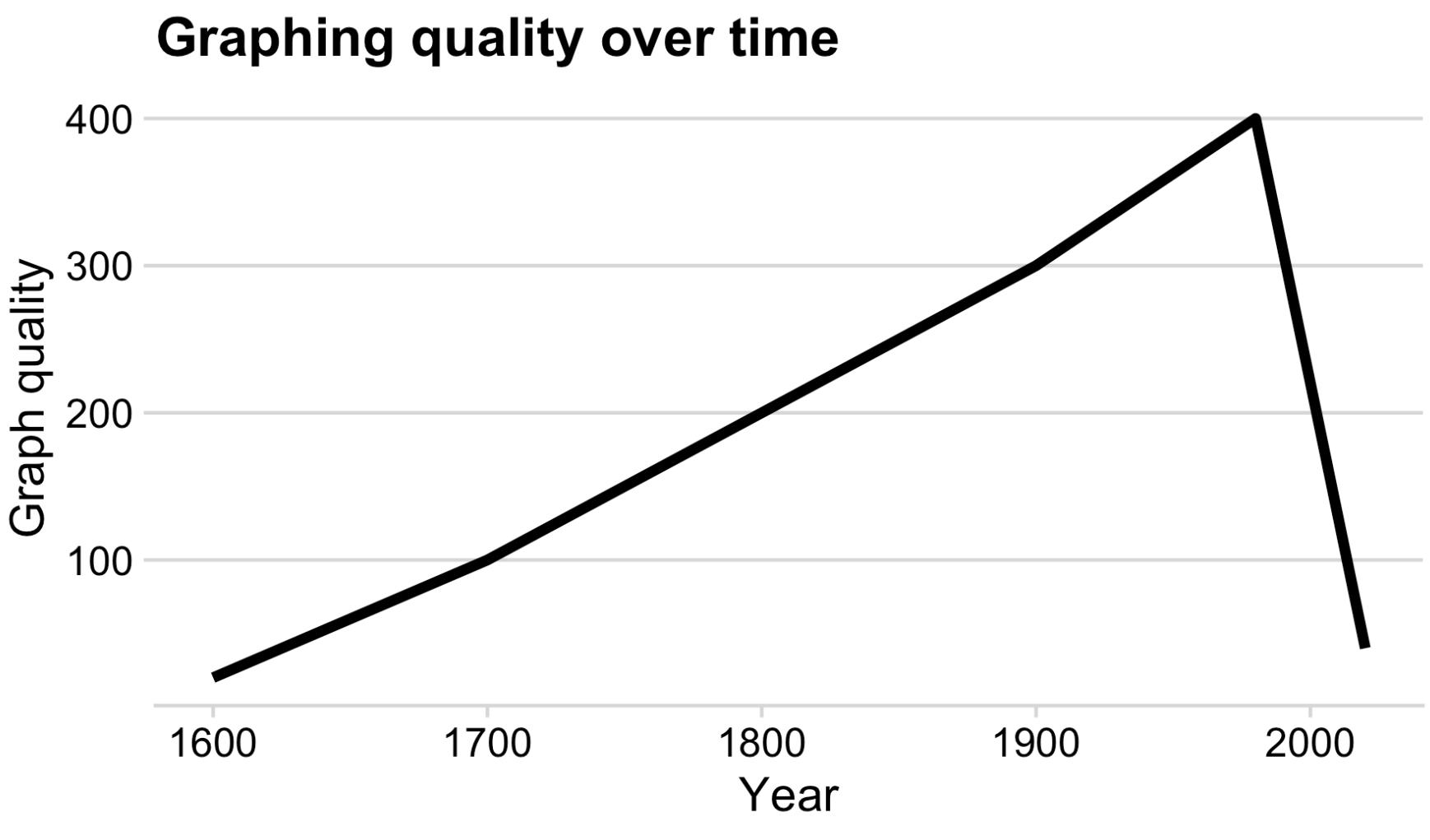


# Visualizing information

EMSE 4197 | John Paul Helveston | January 29, 2020



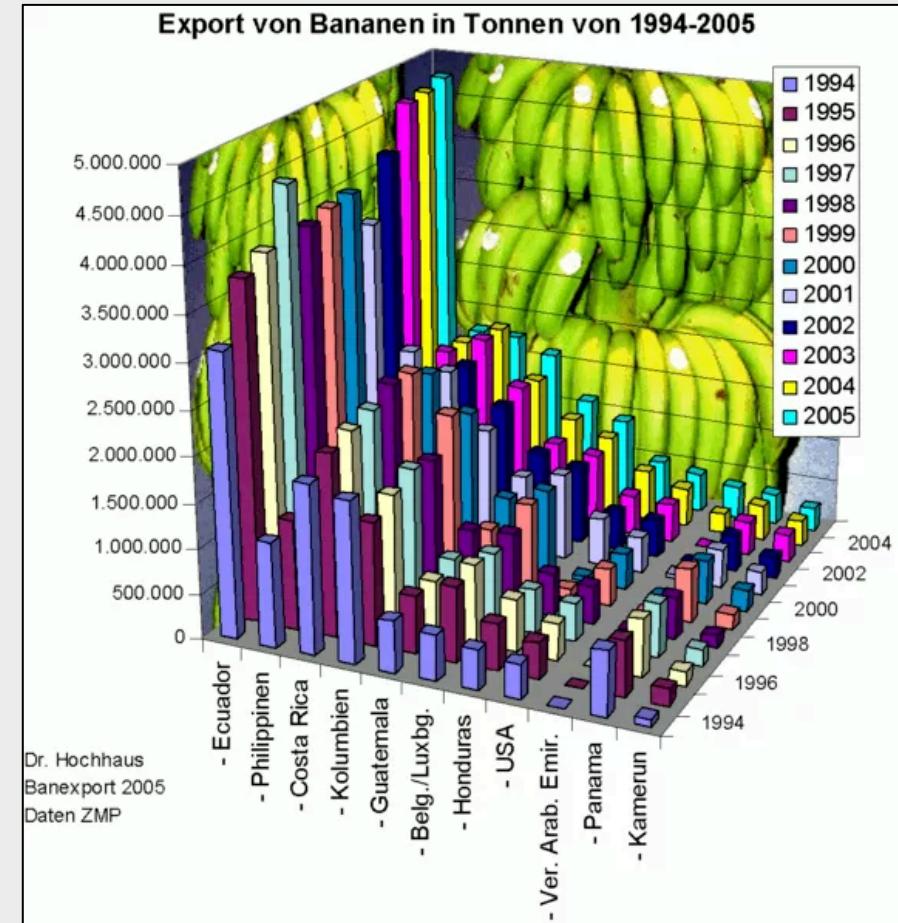
*“Having word processing software  
doesn’t make us great writers.”*

- Stephen Few

# We don't write paragraphs like this:

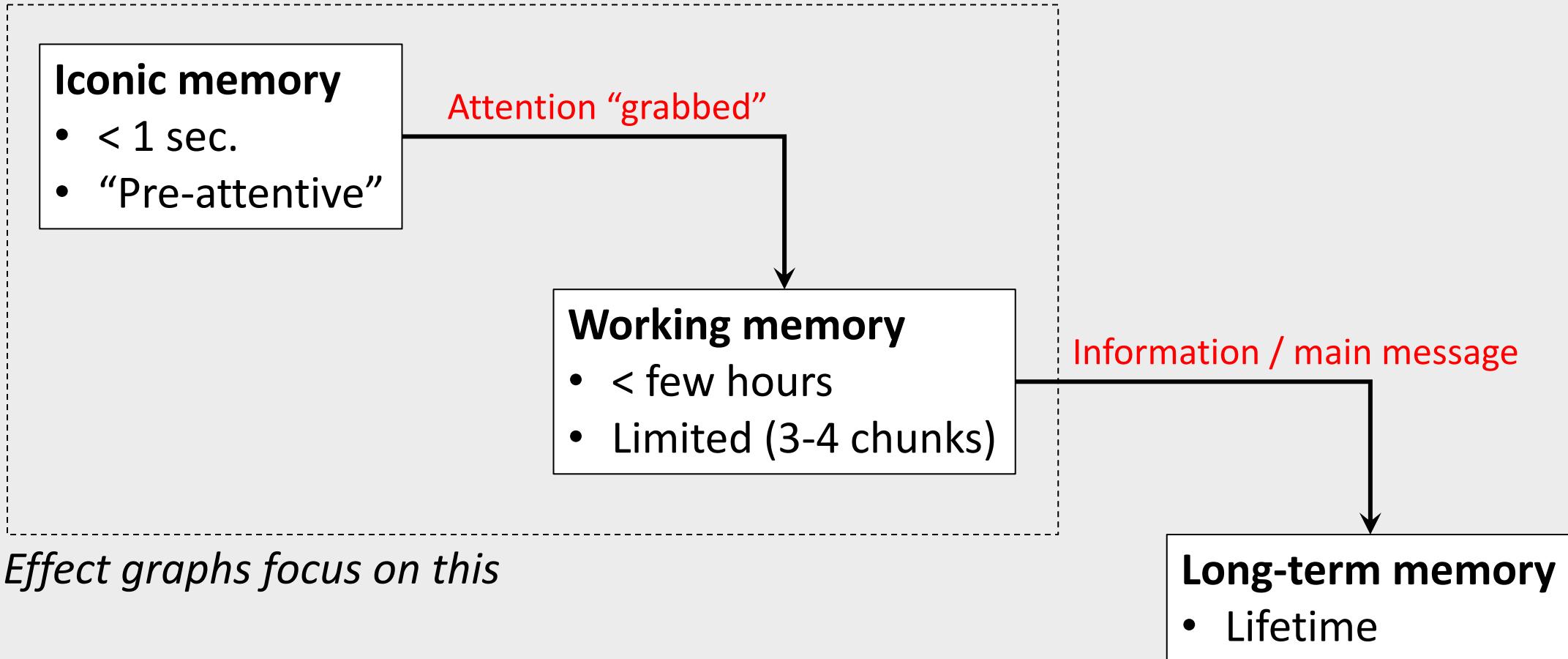
People **sometimes do** this [use poor graphic choices] because they've seen **similar charts in newspapers** or on the web and they're naively following a **bad example**. People who know better **sometimes do** this **because** they care more about **the visual impact** than the **clarity of communication**. *If we wanted* to tell the **truth in** a way people can easily understand, this **is not** an **effective approach**.

# So don't make graphs like this:



Good visualizations optimize for  
the human visual-memory system

# A (very) simplified model of visual-memory system



# Two objectives of effective graphs:

1. Grab & direct attention (iconic memory)
2. Reduce processing demands (working memory)

# The power of pre-attentive processing

Count all the “5”s:

821134907856412043612  
304589640981709812734  
123450986124790812734  
029860192837401489363  
123479827961203459816  
234009816256908127634  
123459087162342015237  
123894789237498230192

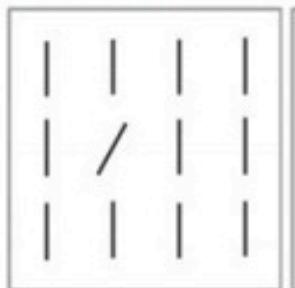
# The power of pre-attentive processing

Count all the “5”s:

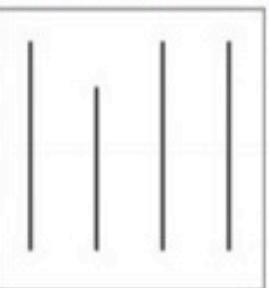
8211349078**5**6412043612  
304**5**89640981709812734  
1234**5**0986124790812734  
029860192837401489363  
1234798279612034**5**9816  
2340098162**5**6908127634  
1234**5**908716234201**5**237  
123894789237498230192

## Form

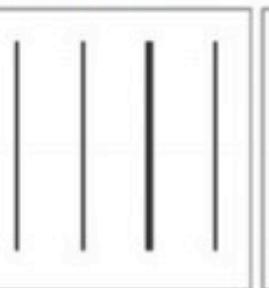
Orientation



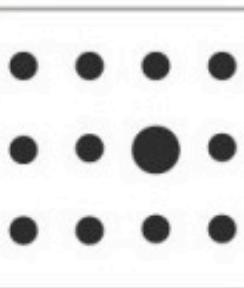
Line Length



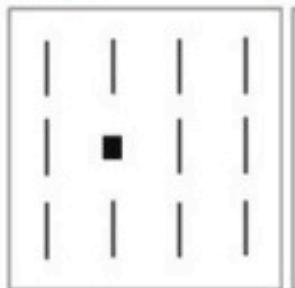
Line Width



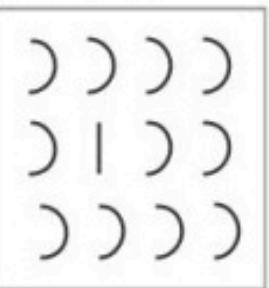
Size



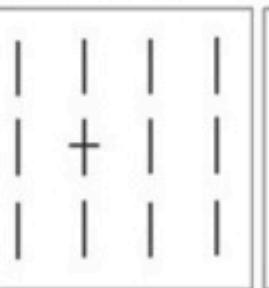
Shape



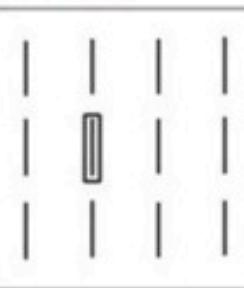
Curvature



Added Marks

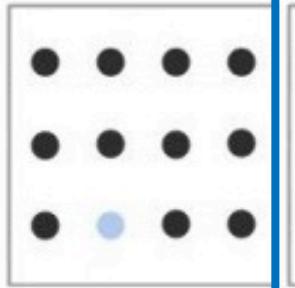


Enclosure

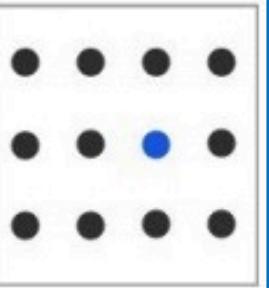


## Color

Intensity

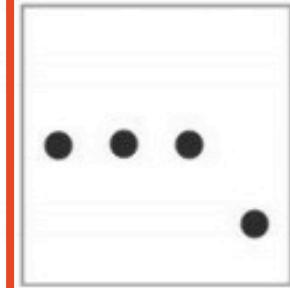


Hue



## Spatial Position

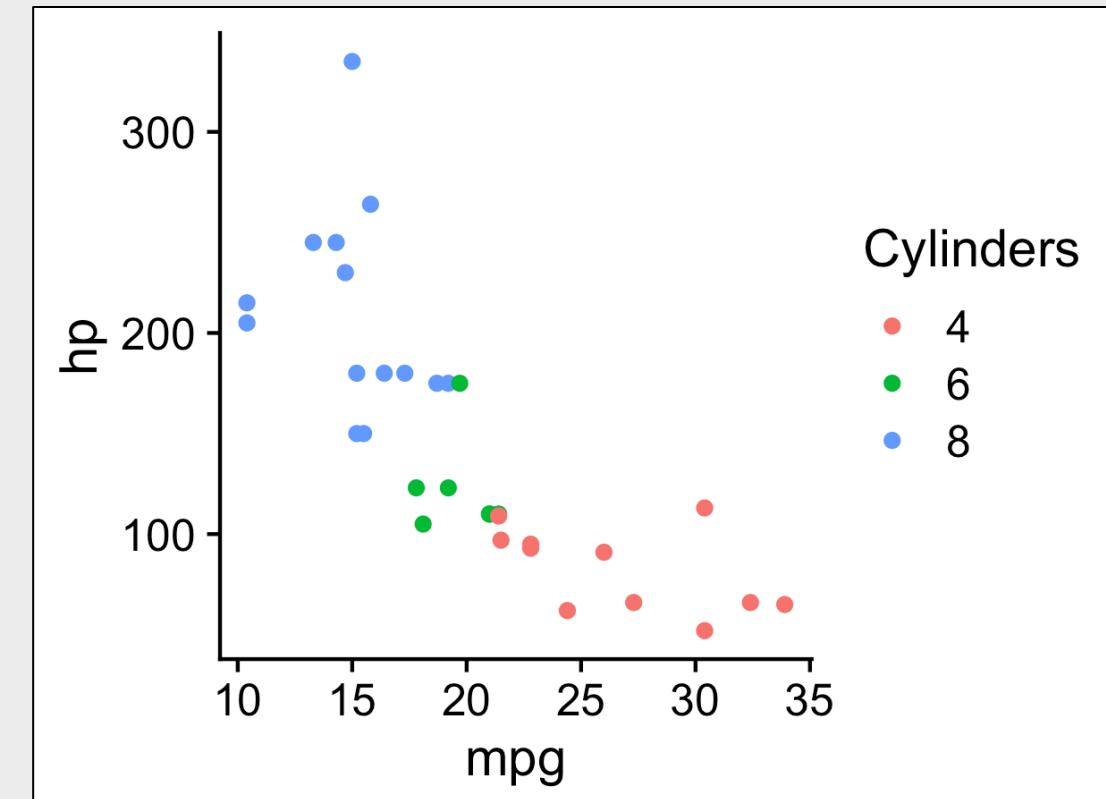
2-D Position



# Pre-attentive attributes

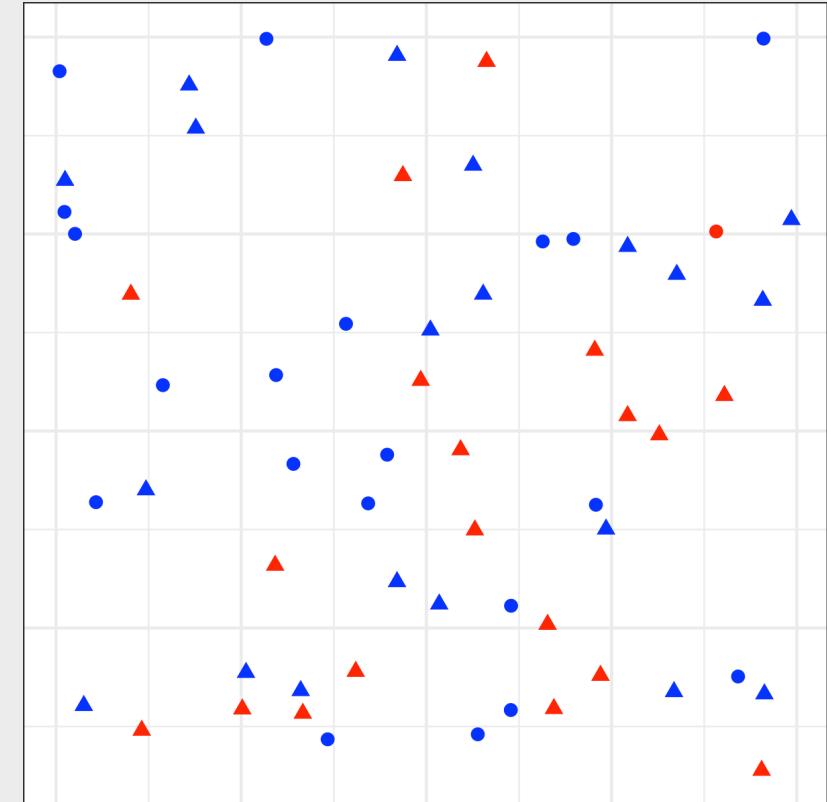
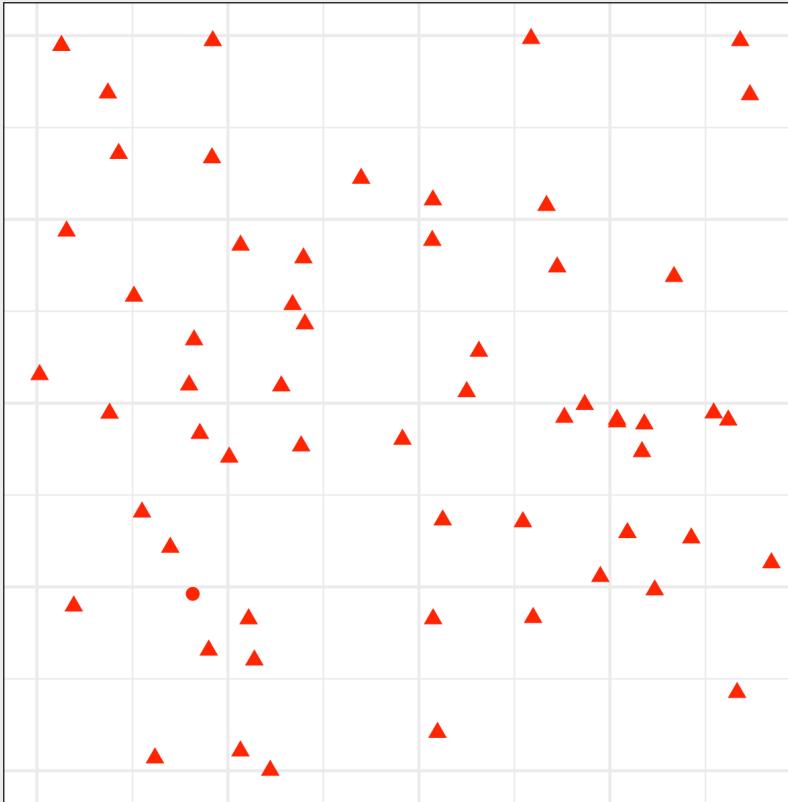
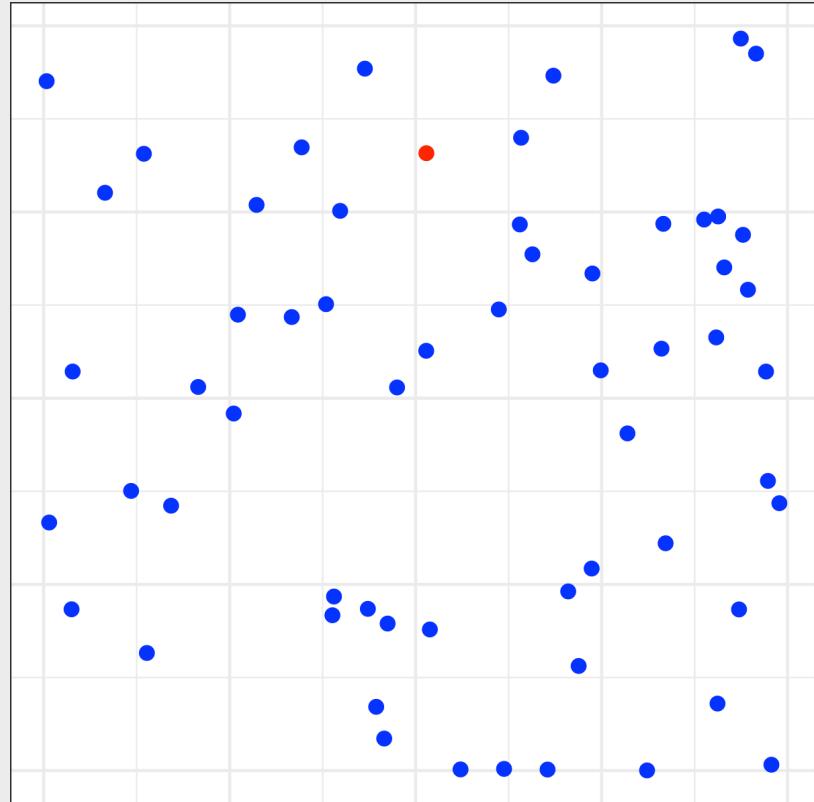
Quantitative data

Categorical data

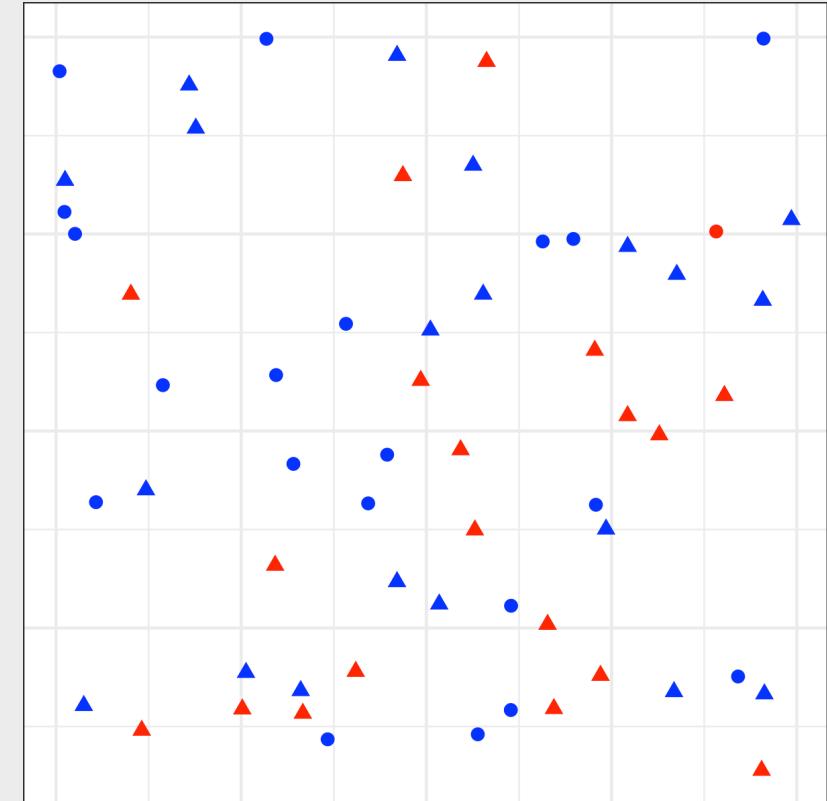
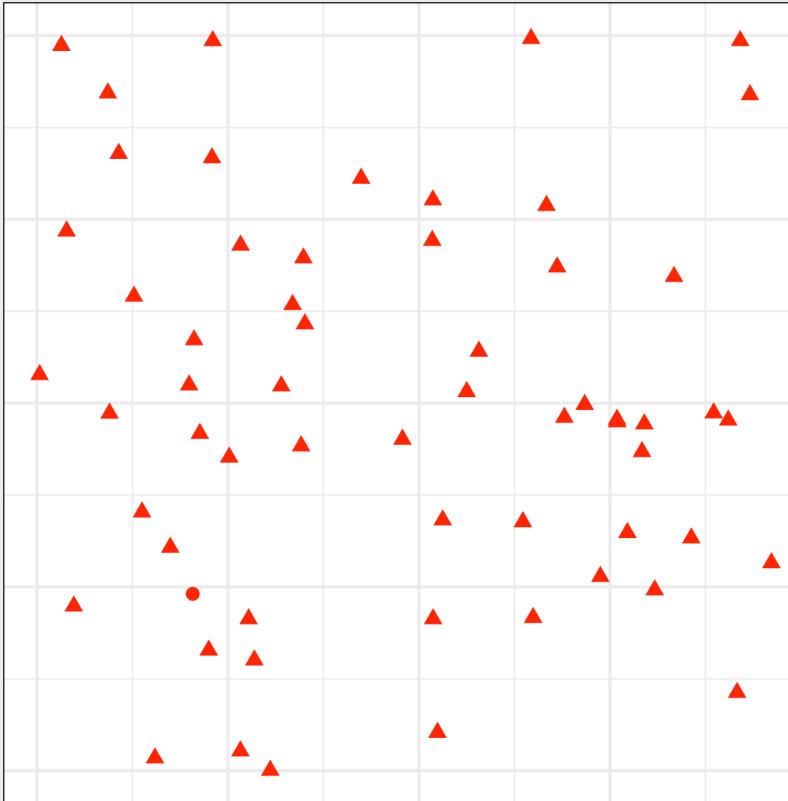
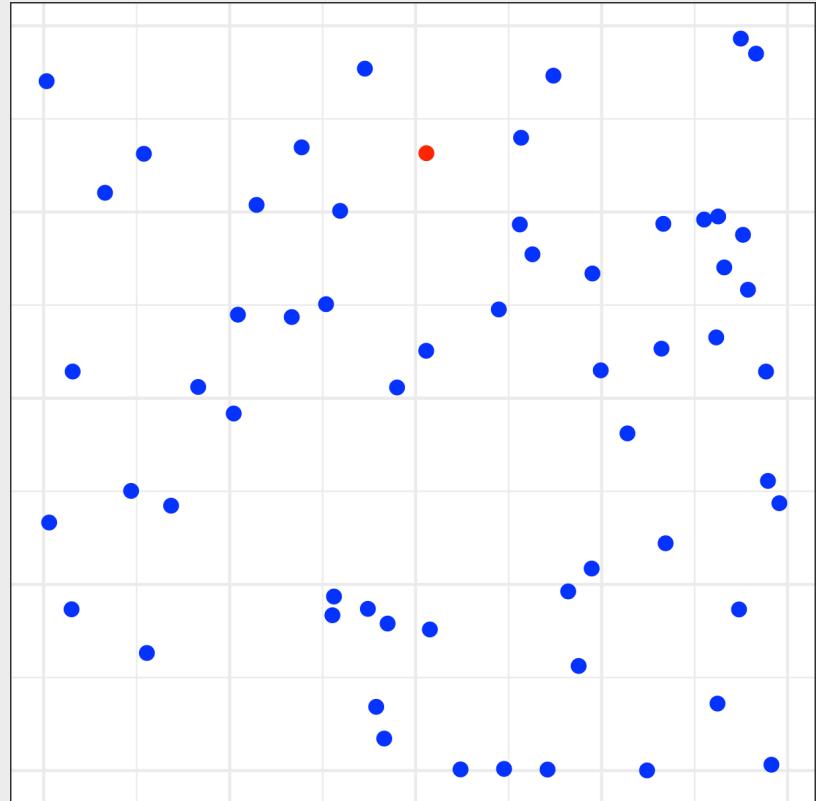


Not all pre-attentive  
attributes are equal

# Is there a red dot?

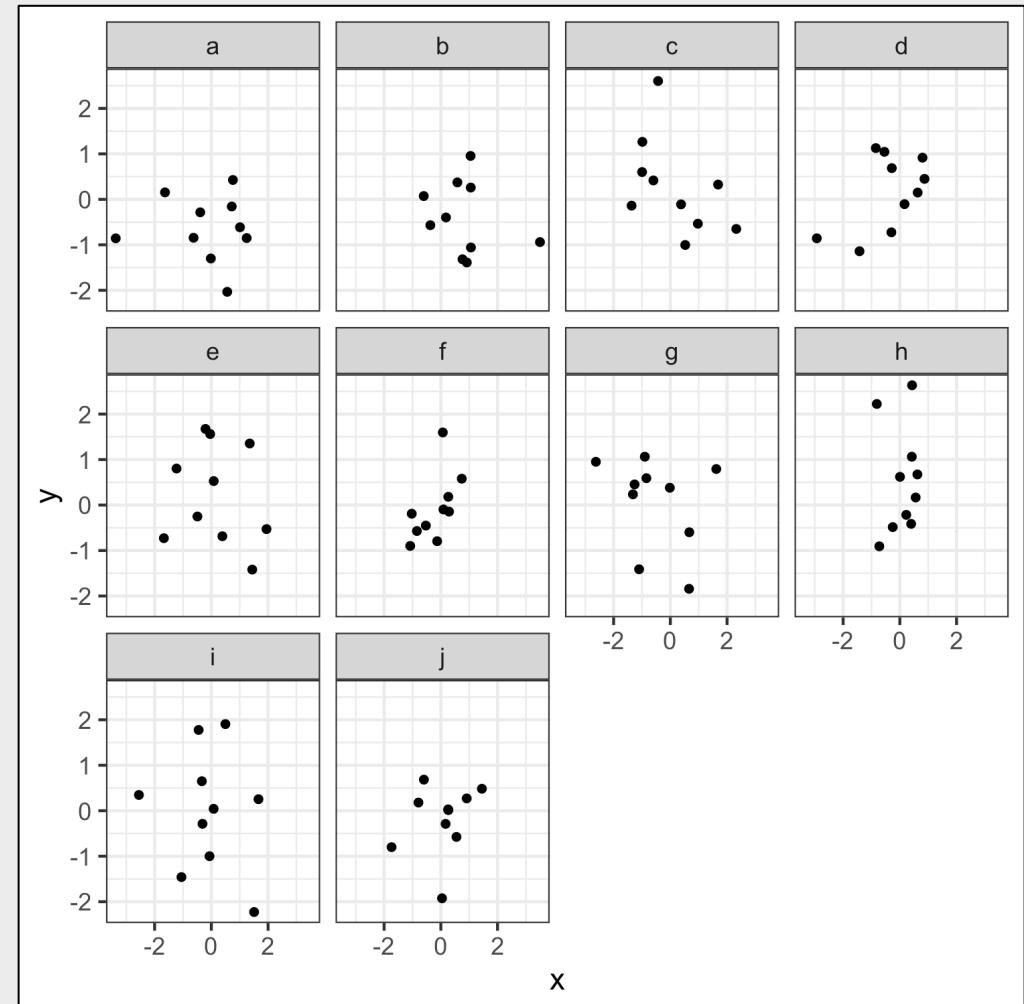
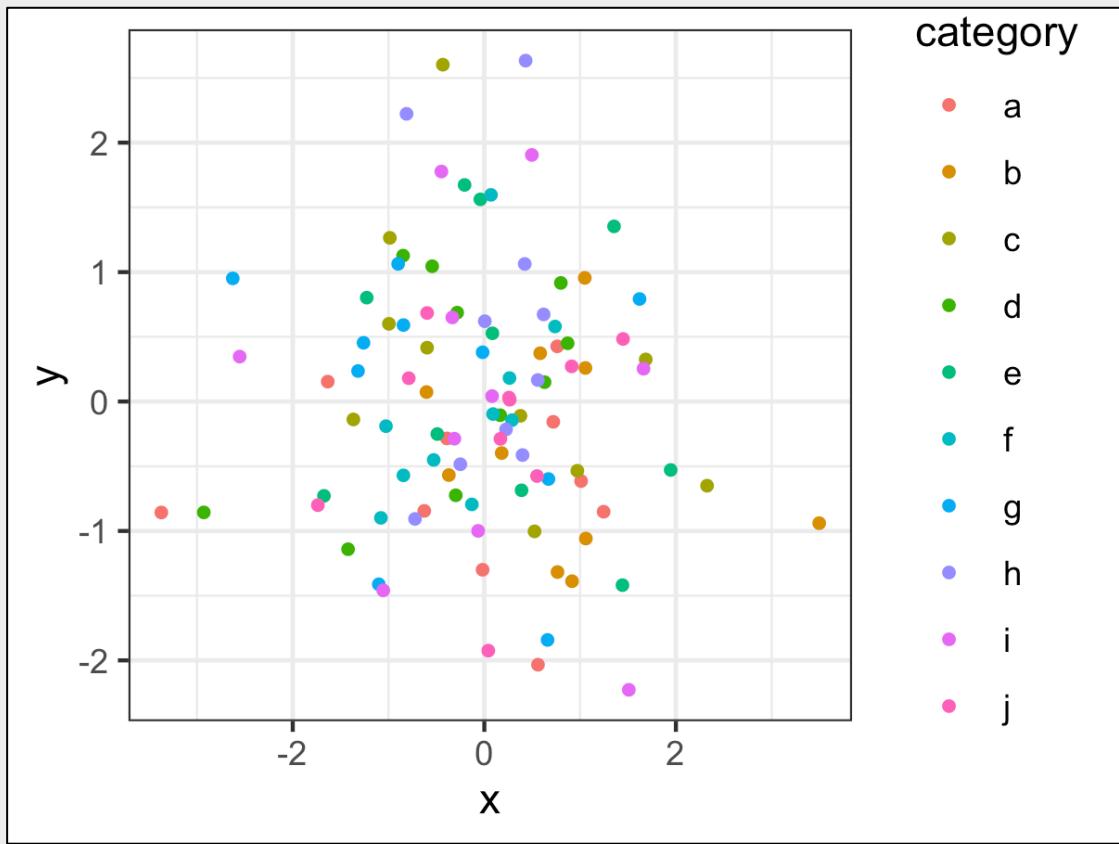


# For categorical data



Hue > Shape

Less is more

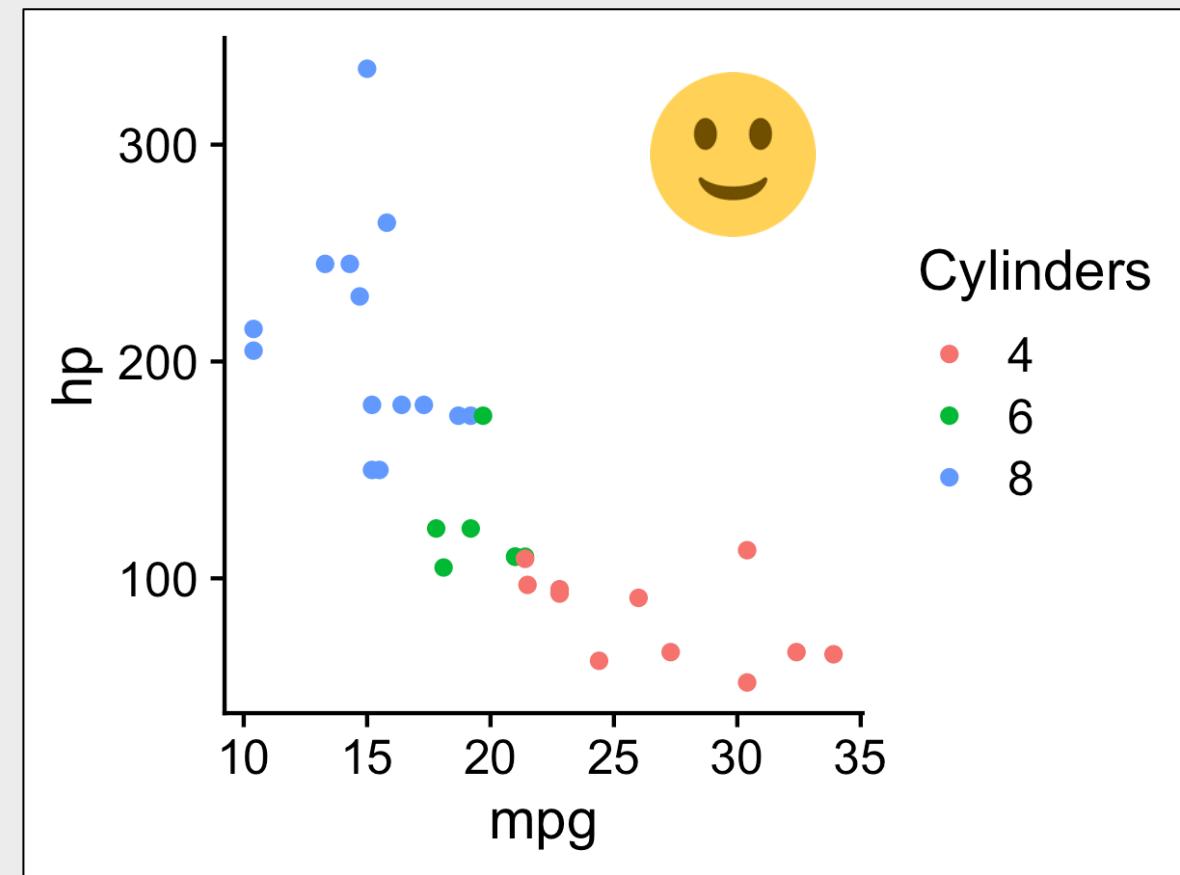


# Graph components

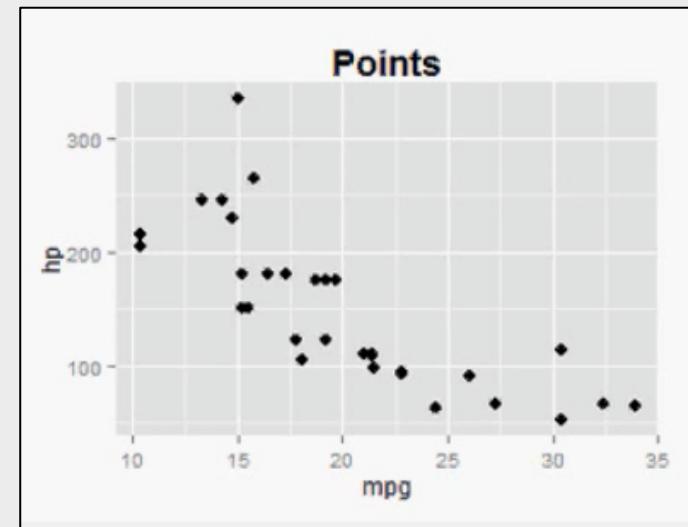
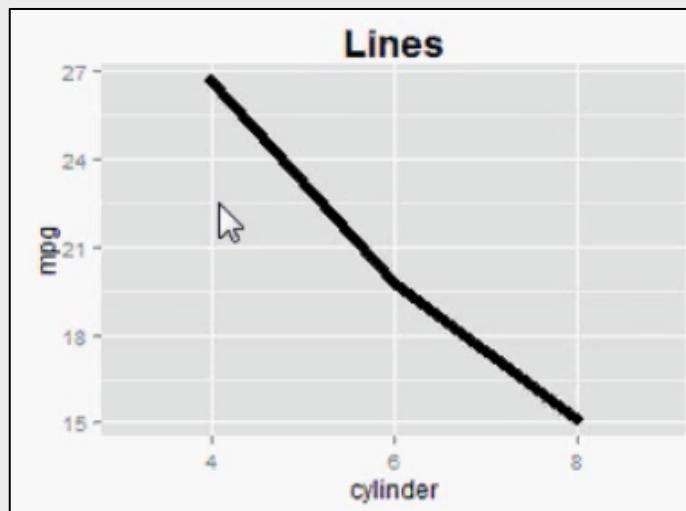
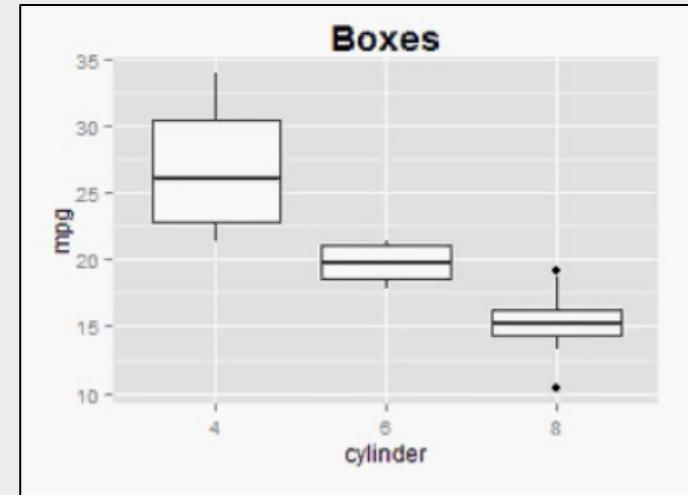
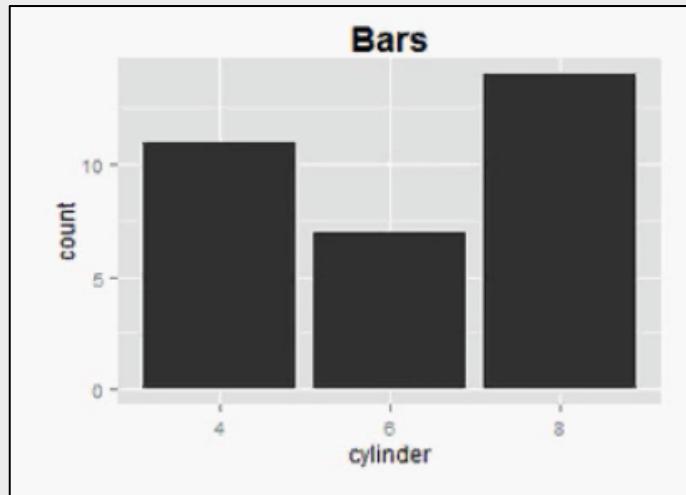
# Graph components

## 1. Geoms:

*points, lines, boxes, bars, etc.*

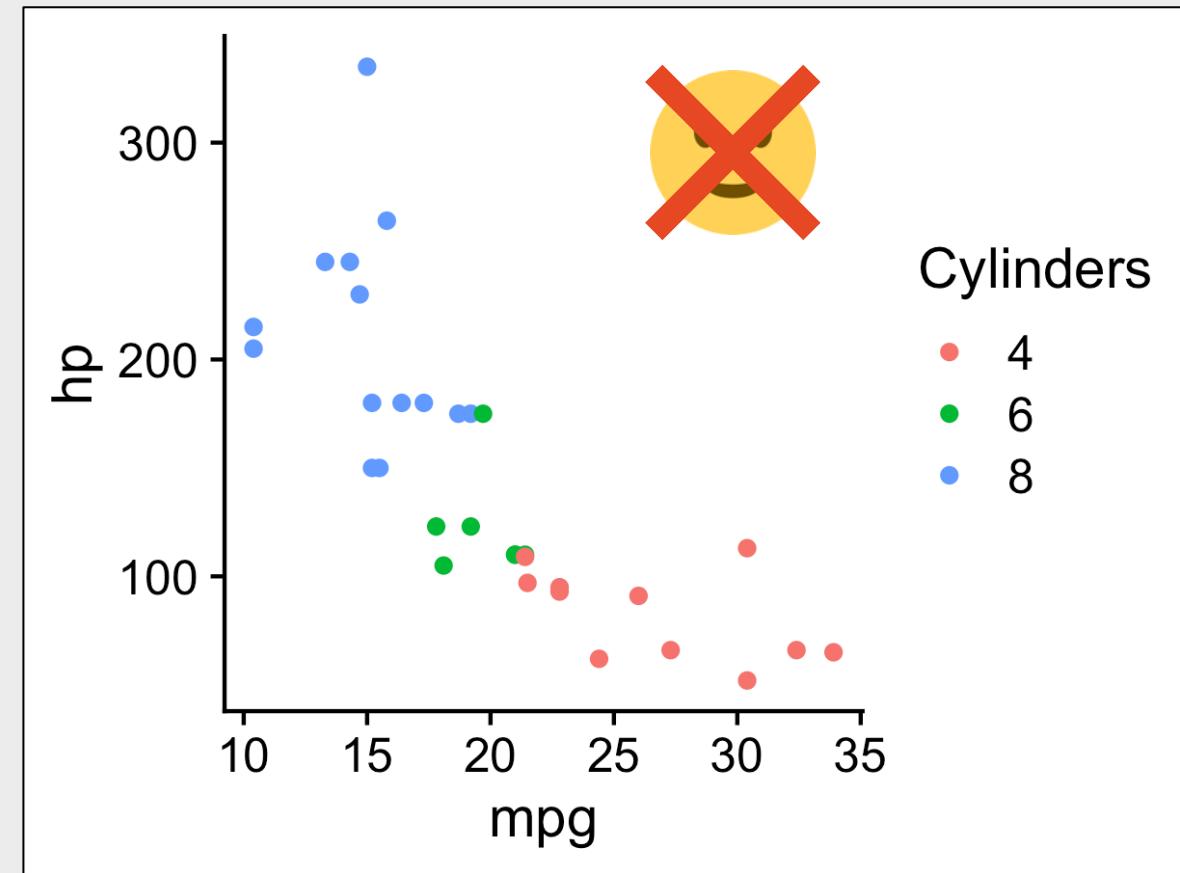


# Geoms: geometric shapes used to represent data



# Graph components

1. Geoms:  
*points, lines, boxes, bars, etc.*
  2. Pre-attentive attributes:  
*position, color, shape, curvature, etc.*
  3. Non-data ink:  
*scales, grid lines, legend, labels, etc.*
- No “chart junk” (Tufte, 2001)



*“Erase non-data ink.”*  
- Edward Tufte

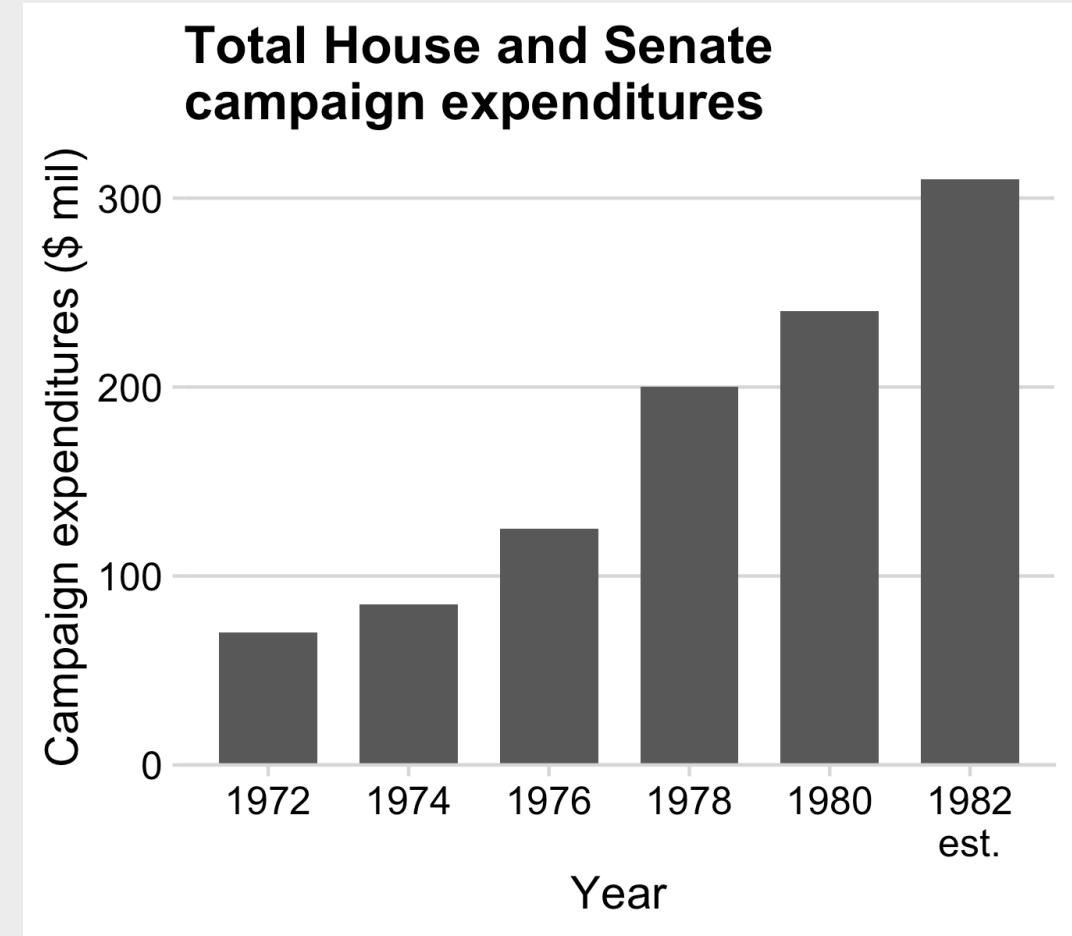


Figure 1.6: 'Monstrous Costs' by Nigel Holmes, in Healy, 2018

# Lessons from Psychology research

## Graphical Perception and Graphical Methods for Analyzing Scientific Data

William S. Cleveland and Robert McGill

Graphs provide powerful tools both for analyzing scientific data and for communicating quantitative information. The computer graphics revolution, which began in the 1960's and has intensified during the past several years, stimulated the invention of graphical meth-

mation from graphs; theory and experimental data are then used to order the tasks on the basis of accuracy. The ordering has an important application: data should be encoded so that the visual decoding involves tasks as high in the ordering as possible, that is, tasks per-

**Summary.** Graphical perception is the visual decoding of the quantitative and qualitative information encoded on graphs. Recent investigations have uncovered basic principles of human graphical perception that have important implications for the display of data. The computer graphics revolution has stimulated the invention of many graphical methods for analyzing and presenting scientific data, such as box plots, two-tiered error bars, scatterplot smoothing, dot charts, and graphing on a log base 2 scale.

ods: types of graphs and types of quantitative information to be shown on graphs (1-4). One purpose of this article is to describe and illustrate several of these new methods.

What has been missing, until recently, in this period of rapid graphical invention and deployment is the study of graphs and the human visual system. When a graph is constructed, quantitative and categorical information is encoded, chiefly through position, shape, size, symbols, and color. When a person looks at a graph, the information is visually decoded by the person's visual sys-

tem with greater accuracy. This is illustrated by several examples in which some much-used graphical forms are presented, set aside, and replaced by new methods.

### Elementary Tasks for the Graphical Perception of Quantitative Information

The first step is to identify elementary graphical-perception tasks that are used to visually extract quantitative information from a graph. (By "quantitative information" we mean numerical values

al field that comes without apparent mental effort. We also perform cognitive tasks such as reading scale information, but much of the power of graphs—and what distinguishes them from tables—comes from the ability of our preattentive visual system to detect geometric patterns and assess magnitudes. We have examined preattentive processes rather than cognition.

We have studied the elementary graphical-perception tasks theoretically, borrowing ideas from the more general field of visual perception (7, 8), and experimentally by having subjects judge graphical elements (1, 5). The next two sections illustrate the methodology with a few examples.

### Study of Graphical Perception: Theory

Figure 2 provides an illustration of theoretical reasoning that borrows some ideas from the field of computational vision (8). Suppose that the goal is to judge the ratio,  $r$ , of the slope of line segment BC to the slope of line segment AB in each of the three panels. Our visual system tells us that  $r$  is greater than 1 in each panel, which is correct. Our visual system also tells us that  $r$  is closer to 1 in the two rectangular panels than in the square panel; that is, the slope of BC appears closer to the slope of AB in the two rectangular panels than in the square panel. This, however, is incorrect;  $r$  is the same in all three panels.

The reason for the distortion in judging Fig. 2 is that our visual system is geared to judging angle rather than slope. In their work on computational theories of vision in artificial intelligence, Marr (8) and Stevens (9) have investigated how people judge the slant and tilt (10) of the surfaces of three-dimensional objects. They argue that we judge slant and tilt as

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science, New Series*, 229(4716), 828-833.

## Graphical Perception and Graphical Methods for Analyzing Scientific Data

William S. Cleveland and Robert McGill

Graphs provide powerful tools both for analyzing scientific data and for communicating quantitative information. The computer graphics revolution, which began in the 1960's and has intensified during the past several years, stimulated the invention of graphical meth-

mation from graphs; theory and experimental data are then used to order the tasks on the basis of accuracy. The ordering has an important application: data should be encoded so that the visual decoding involves tasks as high in the ordering as possible, that is, tasks per-

**Summary.** Graphical perception is the visual decoding of the quantitative and qualitative information encoded on graphs. Recent investigations have uncovered basic principles of human graphical perception that have important implications for the display of data. The computer graphics revolution has stimulated the invention of many graphical methods for analyzing and presenting scientific data, such as box plots, two-tiered error bars, scatterplot smoothing, dot charts, and graphing on a log base 2 scale.

ods: types of graphs and types of quantitative information to be shown on graphs (1-4). One purpose of this article is to describe and illustrate several of these new methods.

What has been missing, until recently, in this period of rapid graphical invention and deployment is the study of graphs and the human visual system. When a graph is constructed, quantitative and categorical information is encoded, chiefly through position, shape, size, symbols, and color. When a person looks at a graph, the information is visually decoded by the person's visual sys-

tem with greater accuracy. This is illustrated by several examples in which some much-used graphical forms are presented, set aside, and replaced by new methods.

### Elementary Tasks for the Graphical Perception of Quantitative Information

The first step is to identify elementary graphical-perception tasks that are used to visually extract quantitative information from a graph. (By "quantitative information" we mean numerical values

al field that comes without apparent mental effort. We also perform cognitive tasks such as reading scale information, but much of the power of graphs—and what distinguishes them from tables—comes from the ability of our preattentive visual system to detect geometric patterns and assess magnitudes. We have examined preattentive processes rather than cognition.

We have studied the elementary graphical-perception tasks theoretically, borrowing ideas from the more general field of visual perception (7, 8), and experimentally by having subjects judge graphical elements (1, 5). The next two sections illustrate the methodology with a few examples.

### Study of Graphical Perception: Theory

Figure 2 provides an illustration of theoretical reasoning that borrows some ideas from the field of computational vision (8). Suppose that the goal is to judge the ratio,  $r$ , of the slope of line segment BC to the slope of line segment AB in each of the three panels. Our visual system tells us that  $r$  is greater than 1 in each panel, which is correct. Our visual system also tells us that  $r$  is closer to 1 in the two rectangular panels than in the square panel; that is, the slope of BC appears closer to the slope of AB in the two rectangular panels than in the square panel. This, however, is incorrect;  $r$  is the same in all three panels.

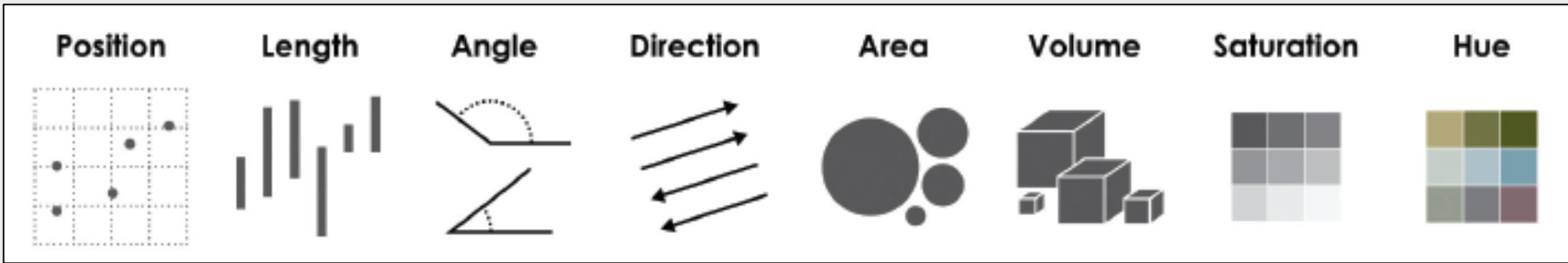
The reason for the distortion in judging Fig. 2 is that our visual system is geared to judging angle rather than slope. In their work on computational theories of vision in artificial intelligence, Marr (8) and Stevens (9) have investigated how people judge the slant and tilt (10) of the surfaces of three-dimensional objects. They argue that we judge slant and tilt as

Table 1. Ordering elementary tasks by accuracy, according to theoretical arguments and experimental results. Graphs should exploit tasks as high in the ordering as possible. The tasks are ordered from most accurate to least.

Rank	Aspect judged
1	Position along a common scale
2	Position on identical but nonaligned scales
3	Length
4	Angle
5	Slope (with $\theta$ not too close to $0, \pi/2$ , or $\pi$ radians)
6	Area
7	Volume
	Density
	Color saturation
	Color hue

Cleveland, W. S., & McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science, New Series*, 229(4716), 828-833.

# Pattern recognition hierarchy:



← More accurate

Less accurate →

Full disclosure:

Much of the next 50-ish slides are  
from John Rauser's [talk on YouTube](#)

# Cleveland's three visual operations of pattern perception:

- **Estimation** —→ • Discrimination  $X \neq Y$
- Assembly • Ranking  $X > Y$
- Detection • Ratioing  $X / Y$

*“At the heart of quantitative reasoning is a single question: compared to what?”*

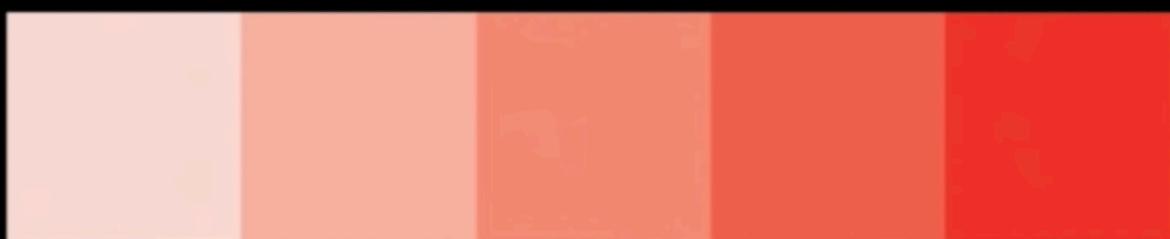
- Edward Tufte

1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

luminance



saturation

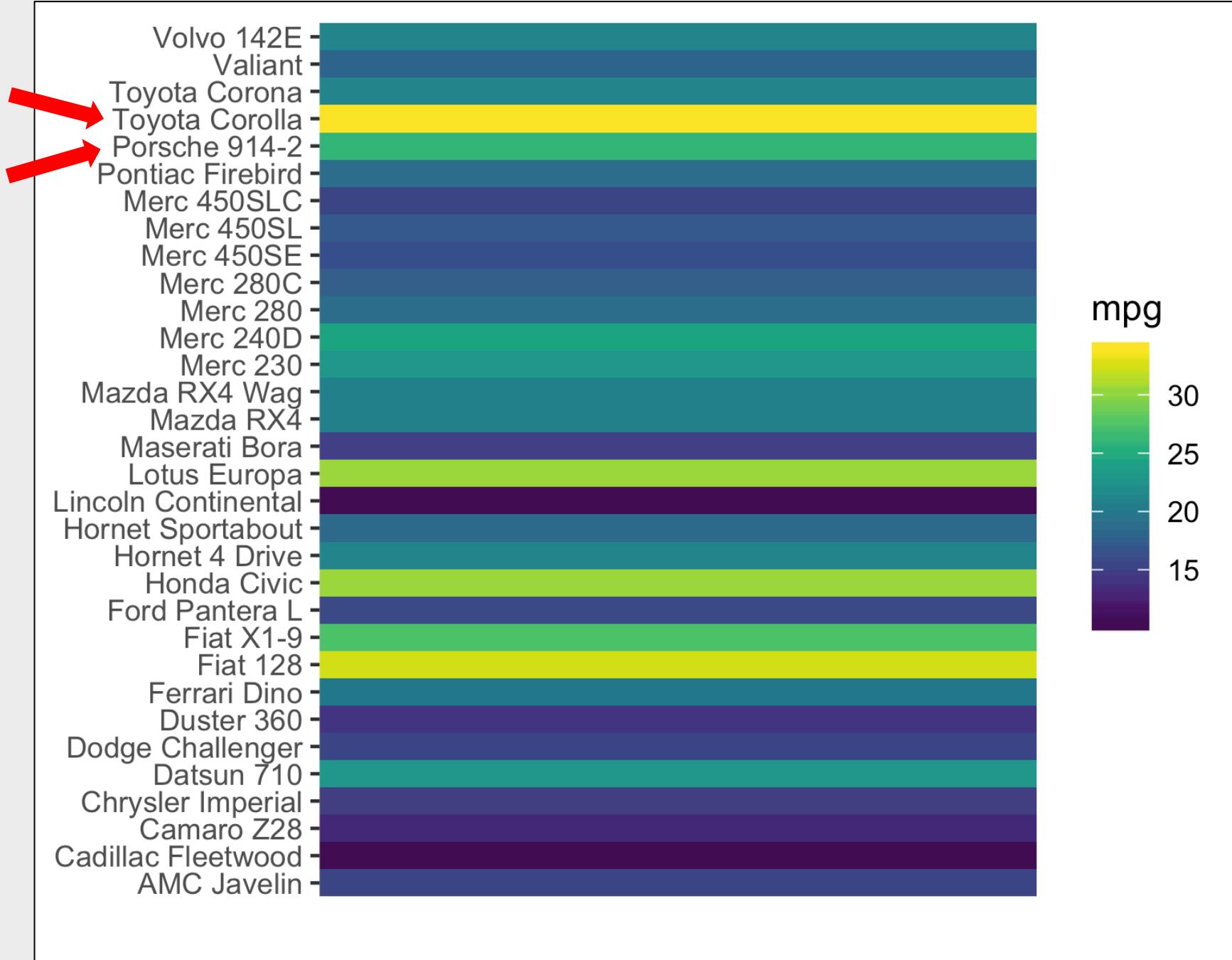


hue



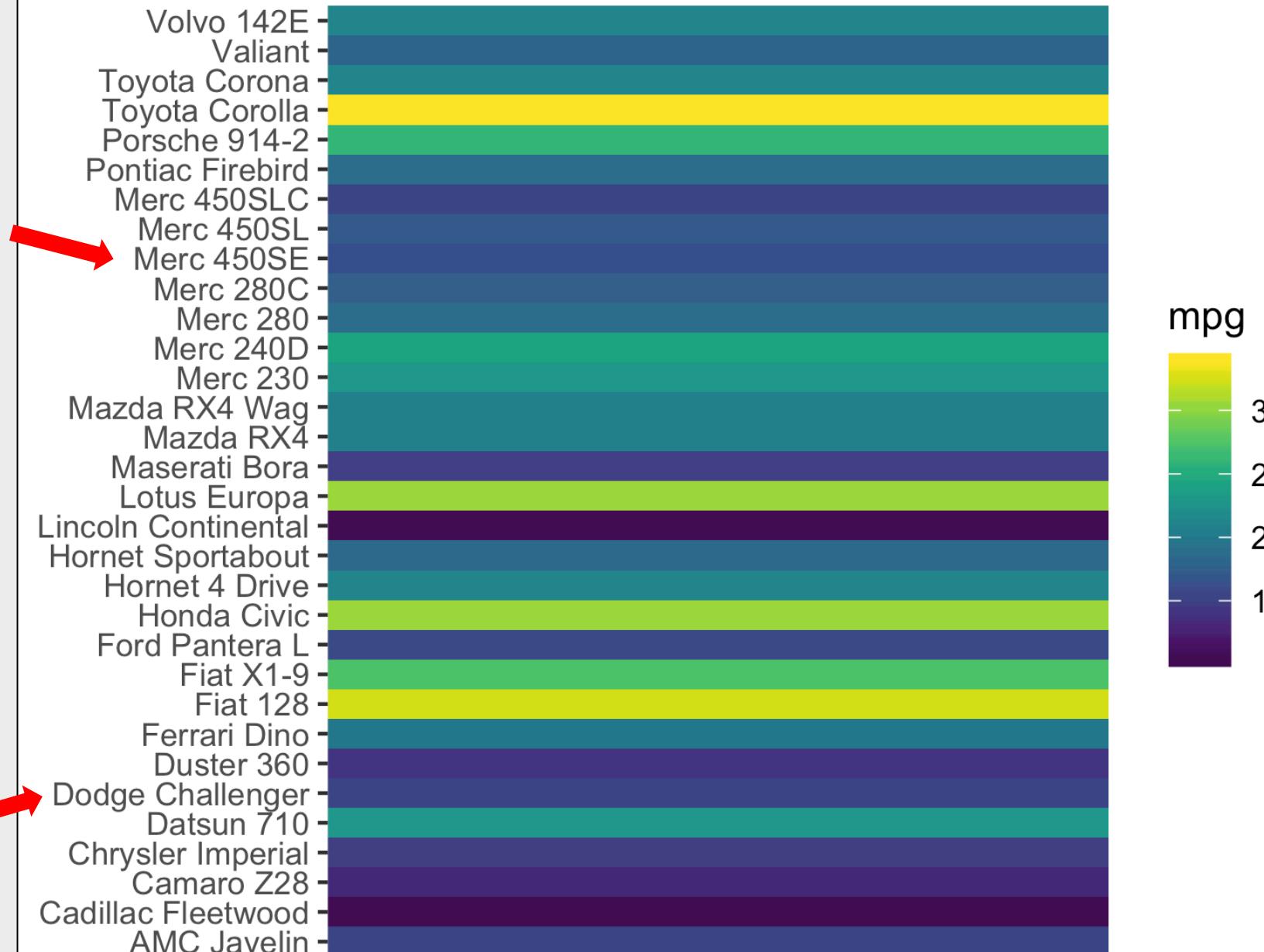
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Discrimination?



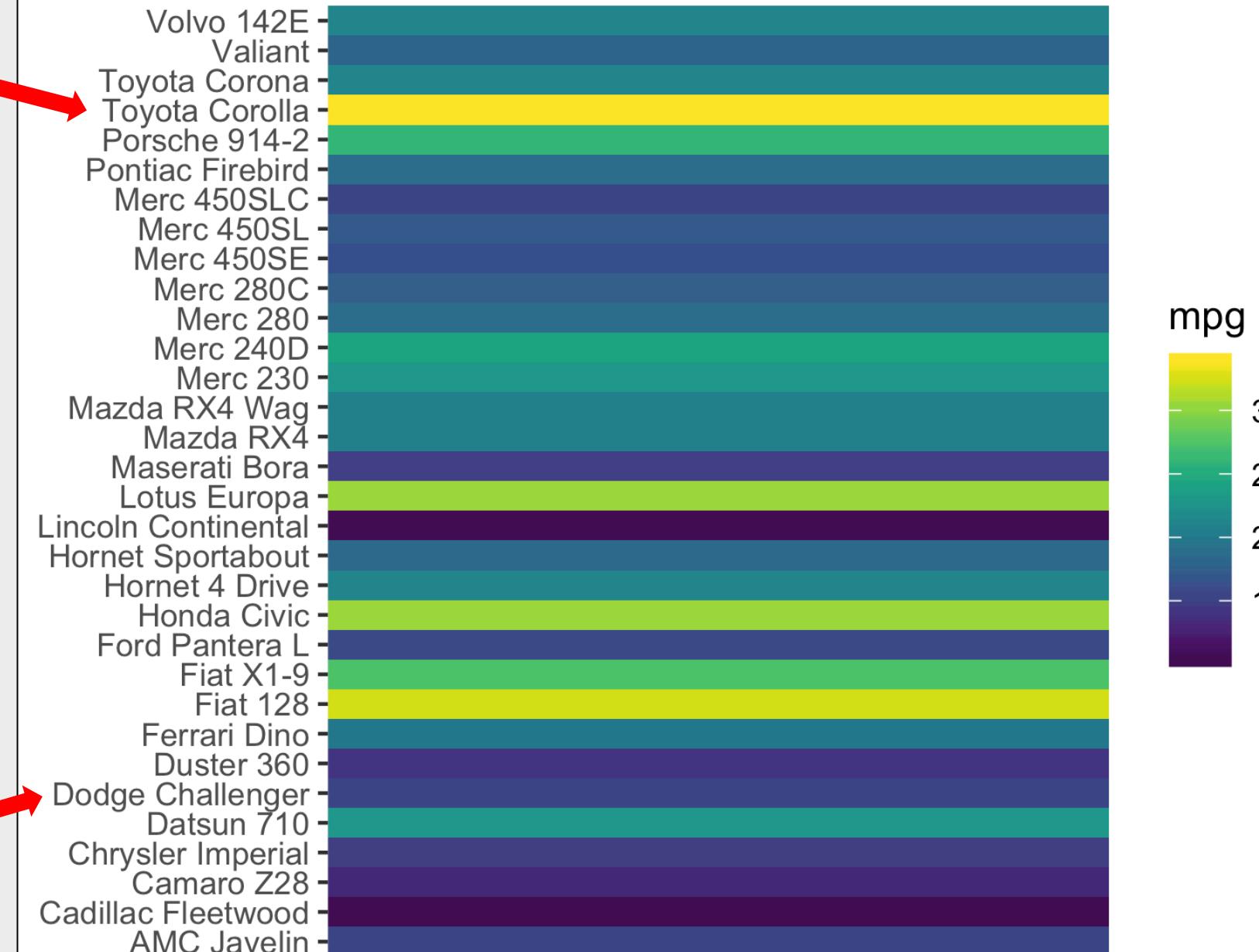
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Discrimination?



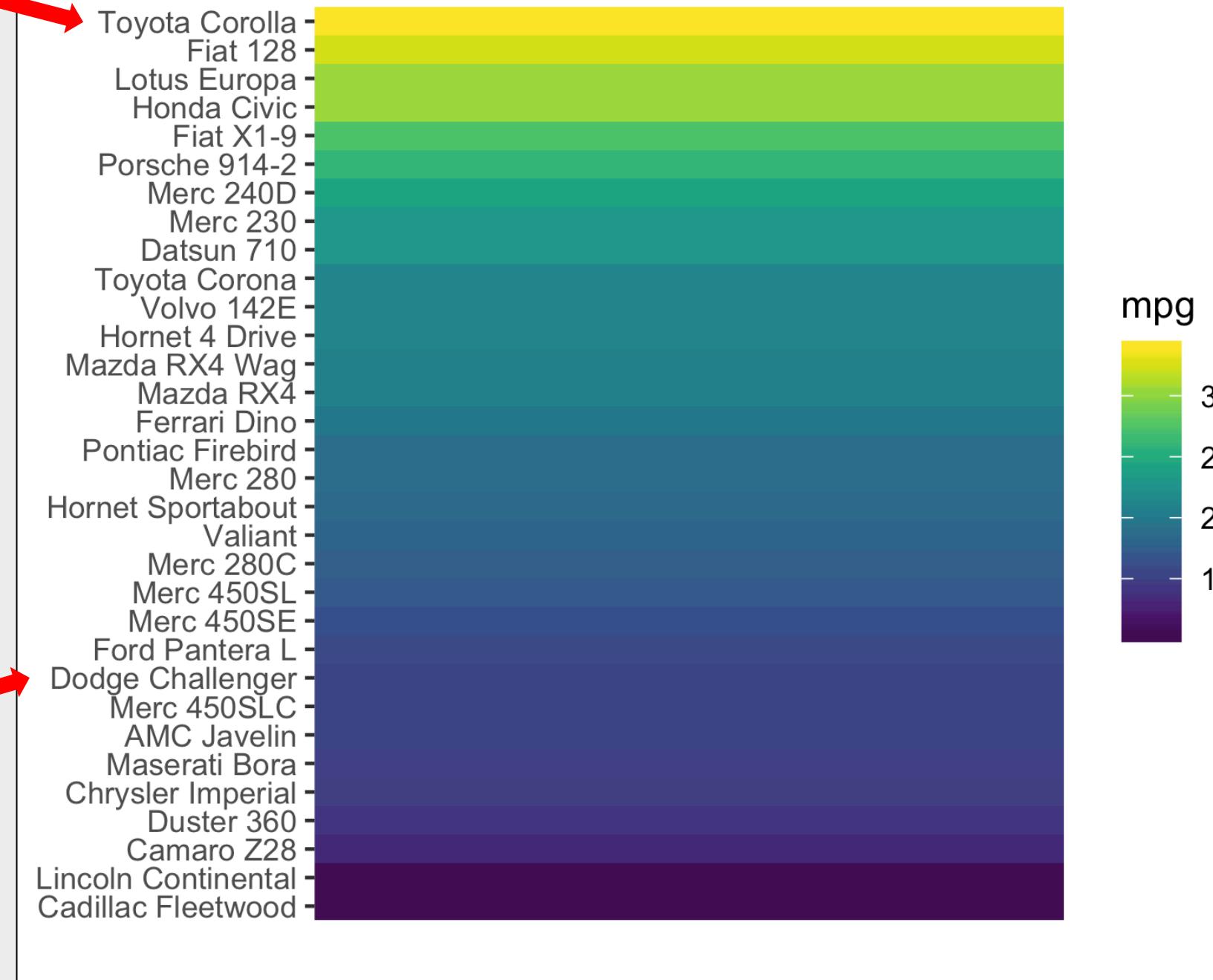
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ranking?



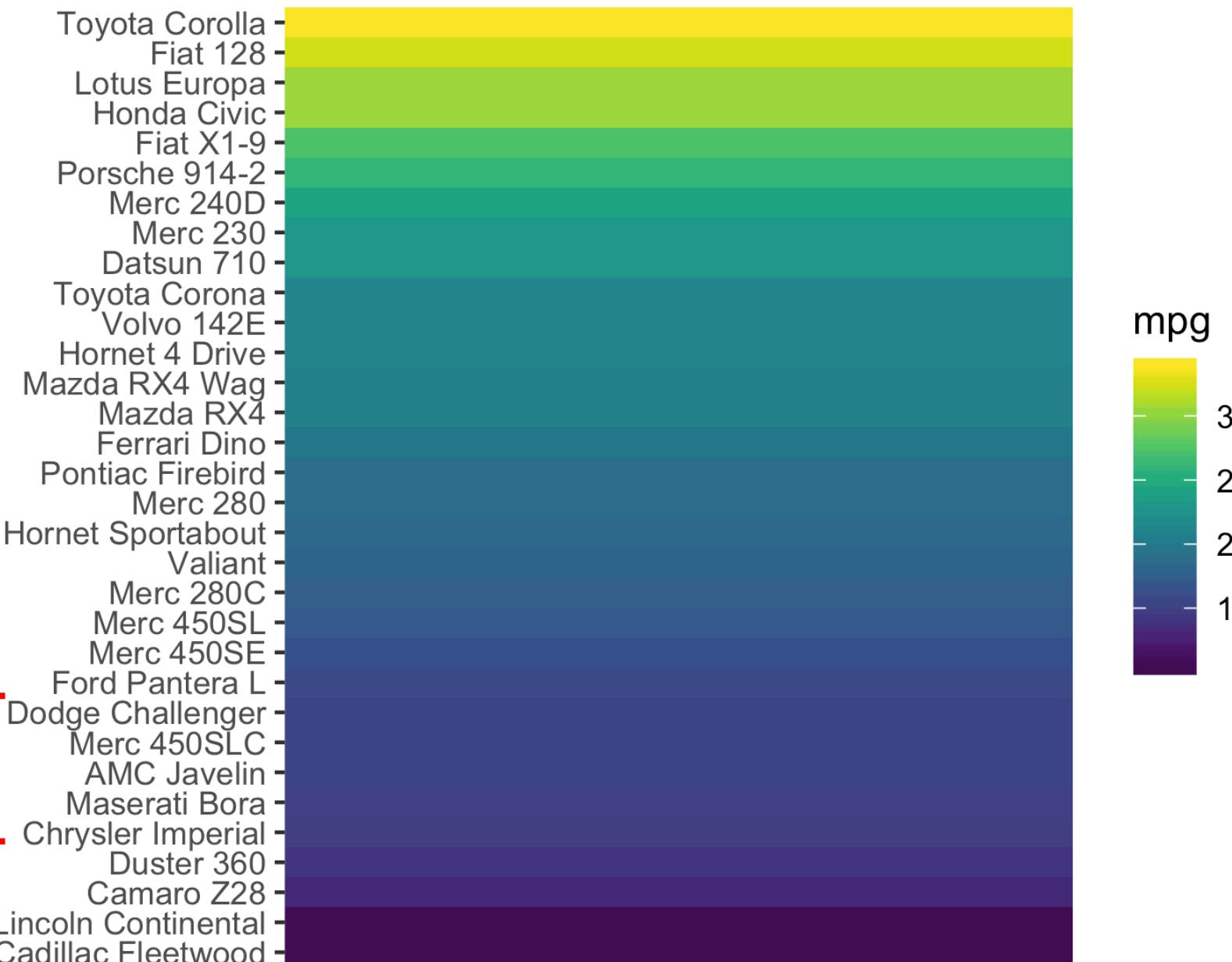
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ranking?

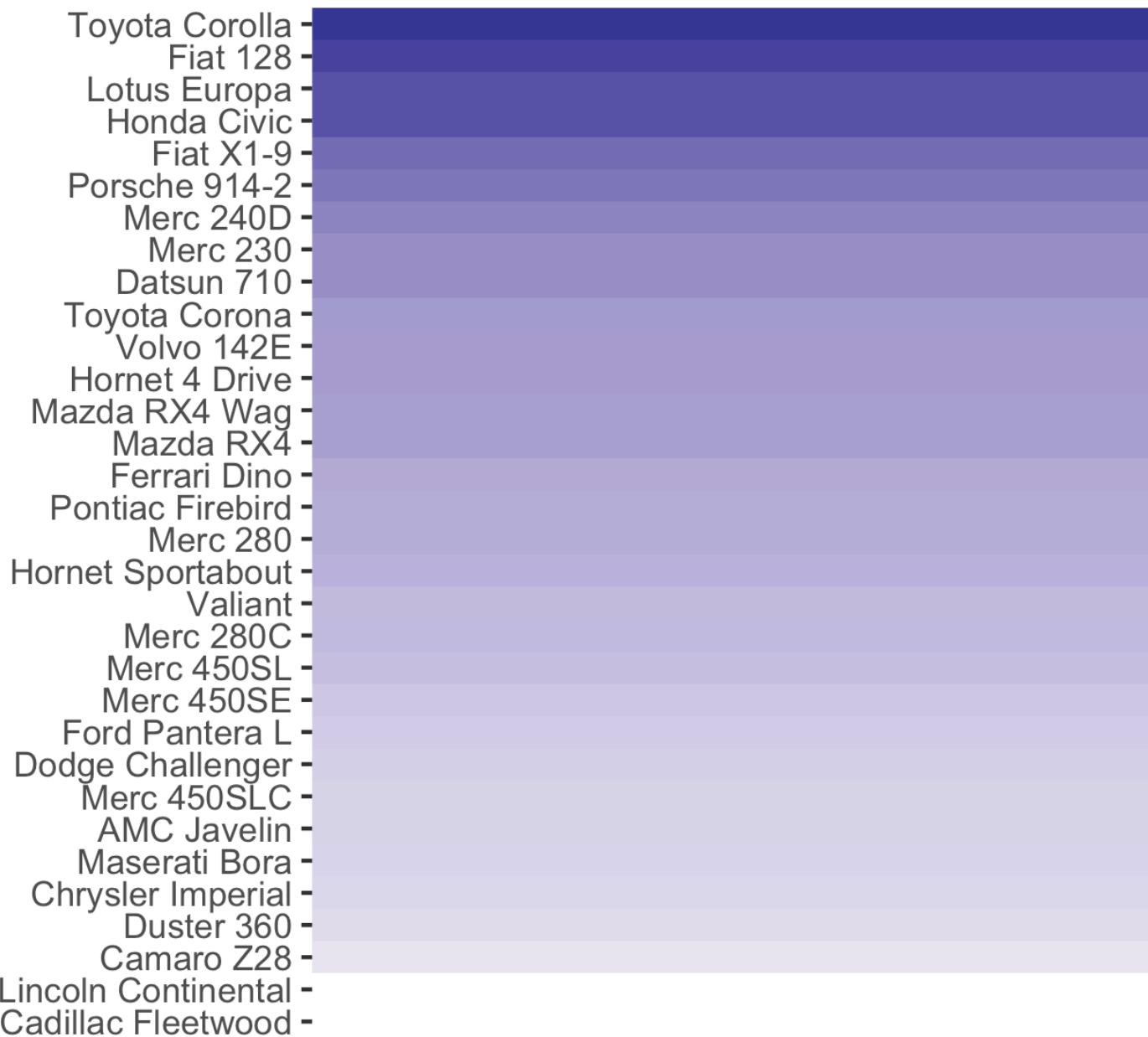


1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

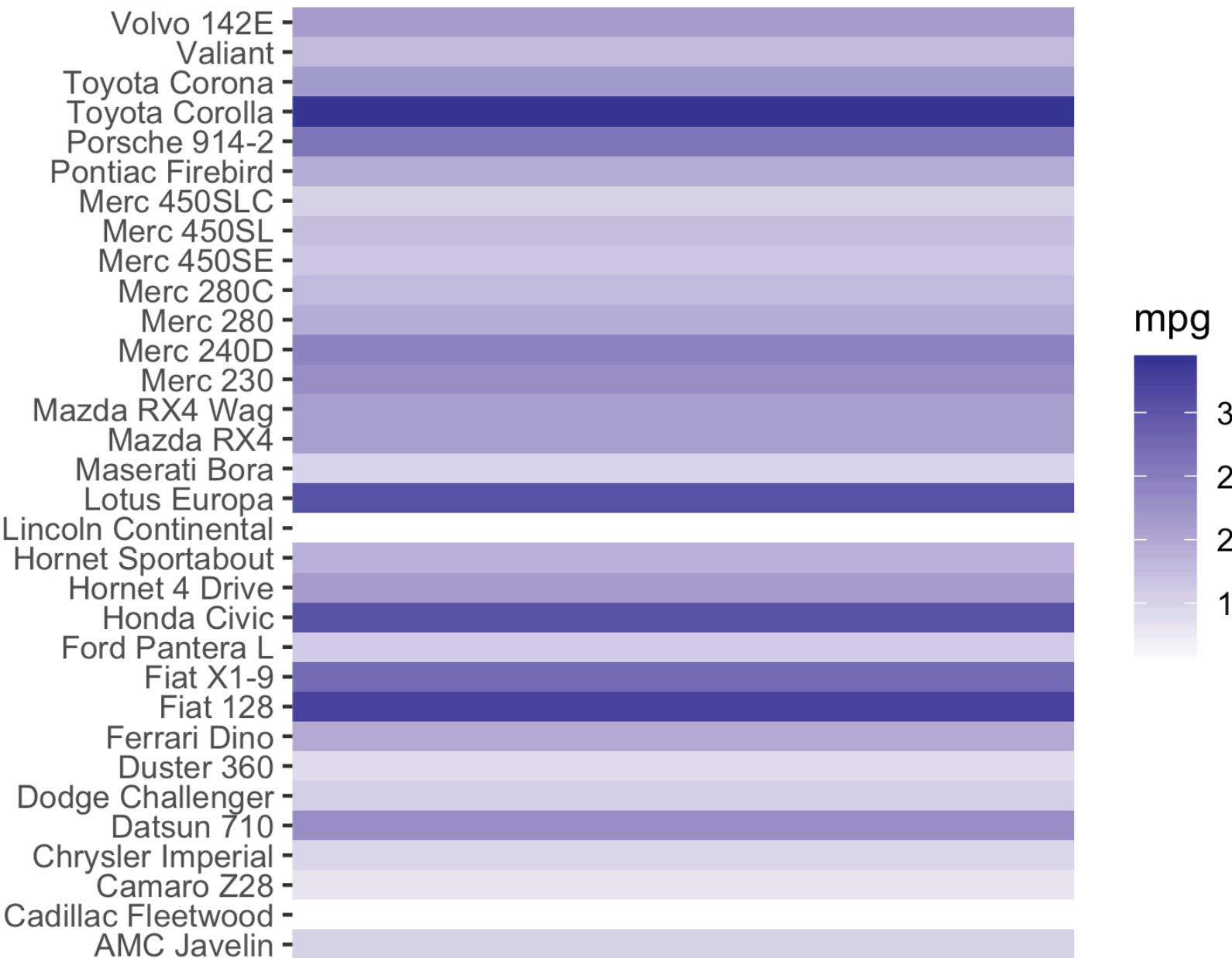
Discrimination?



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

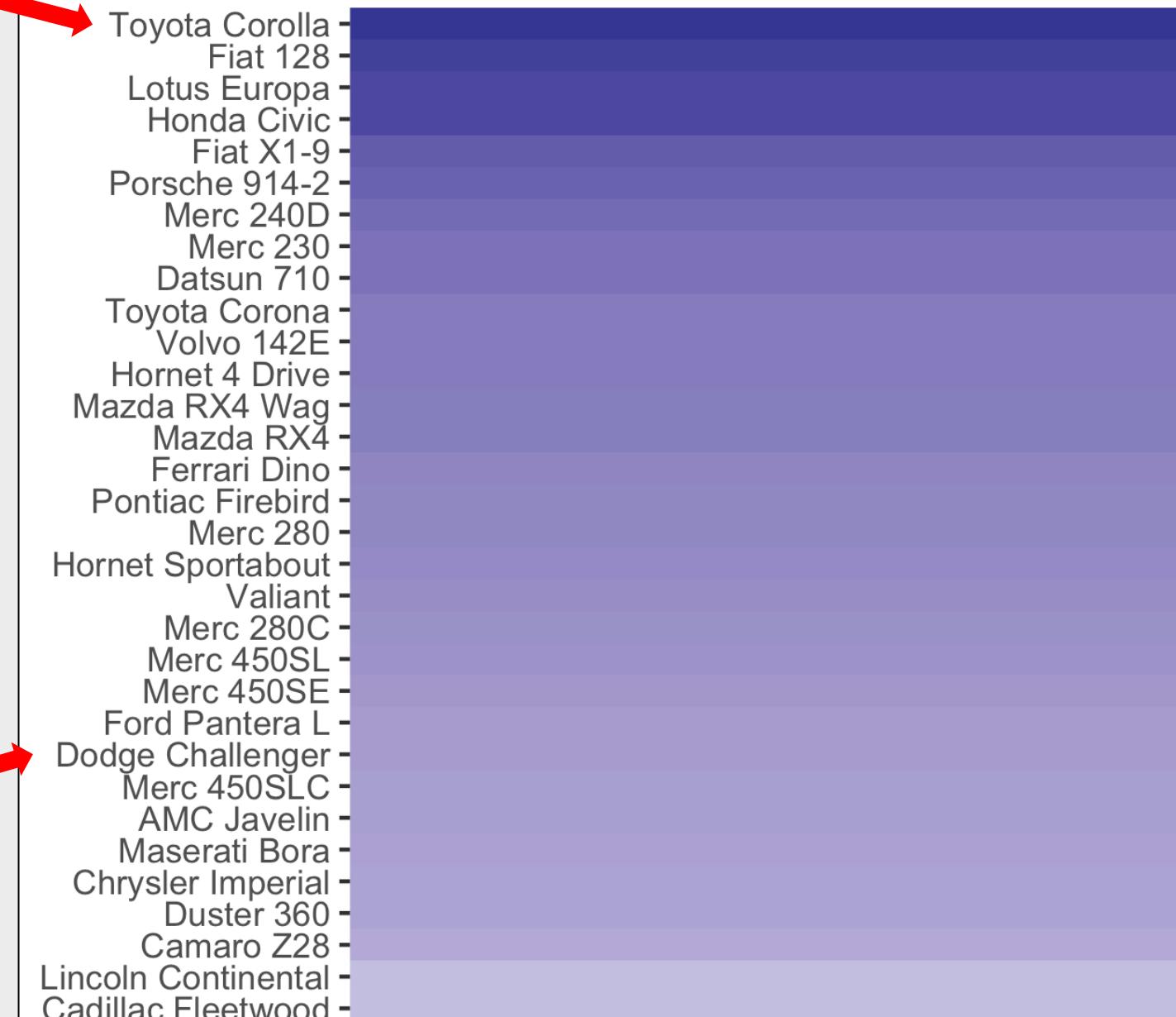


1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ratioing?



mpg

30

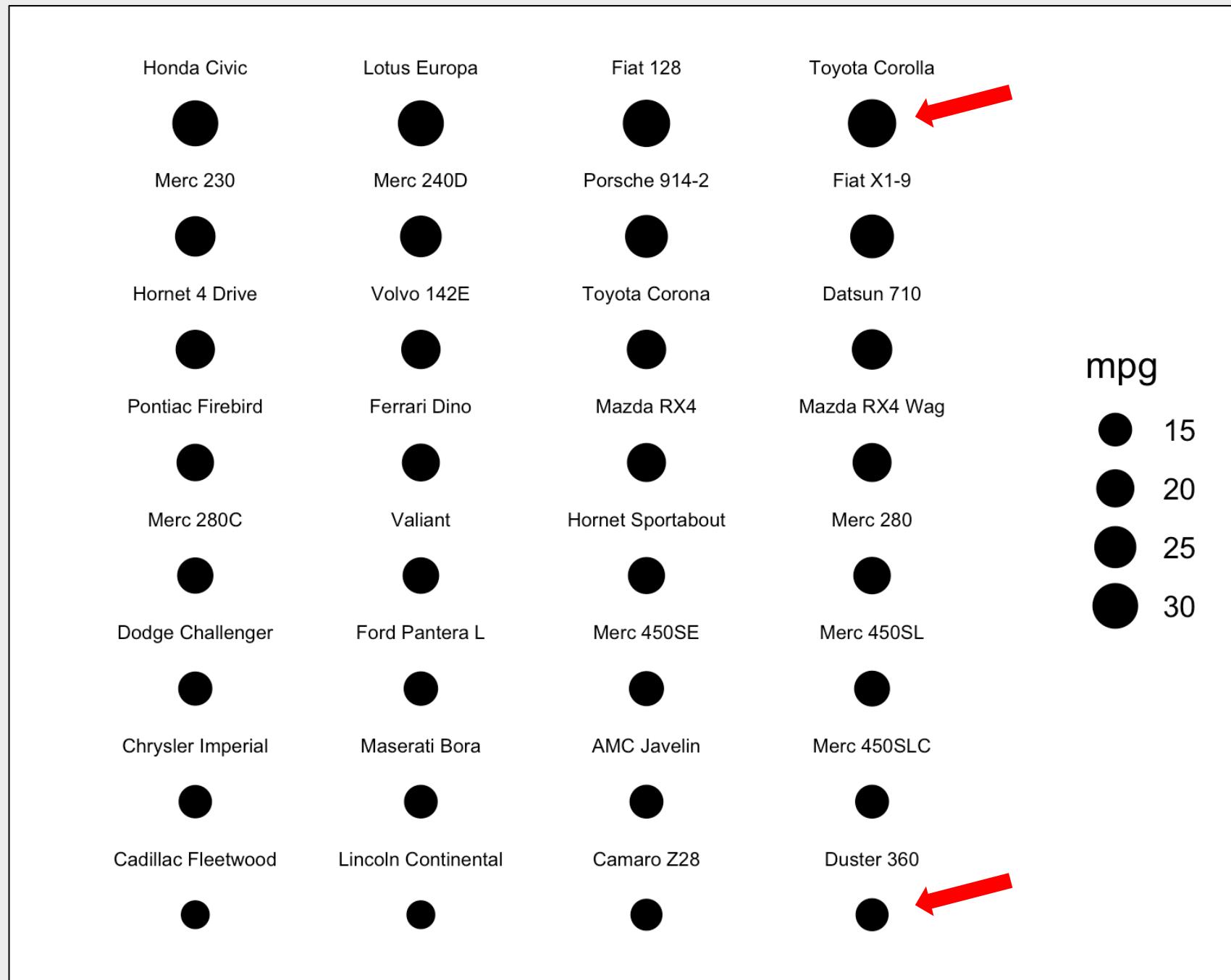
20

10

0

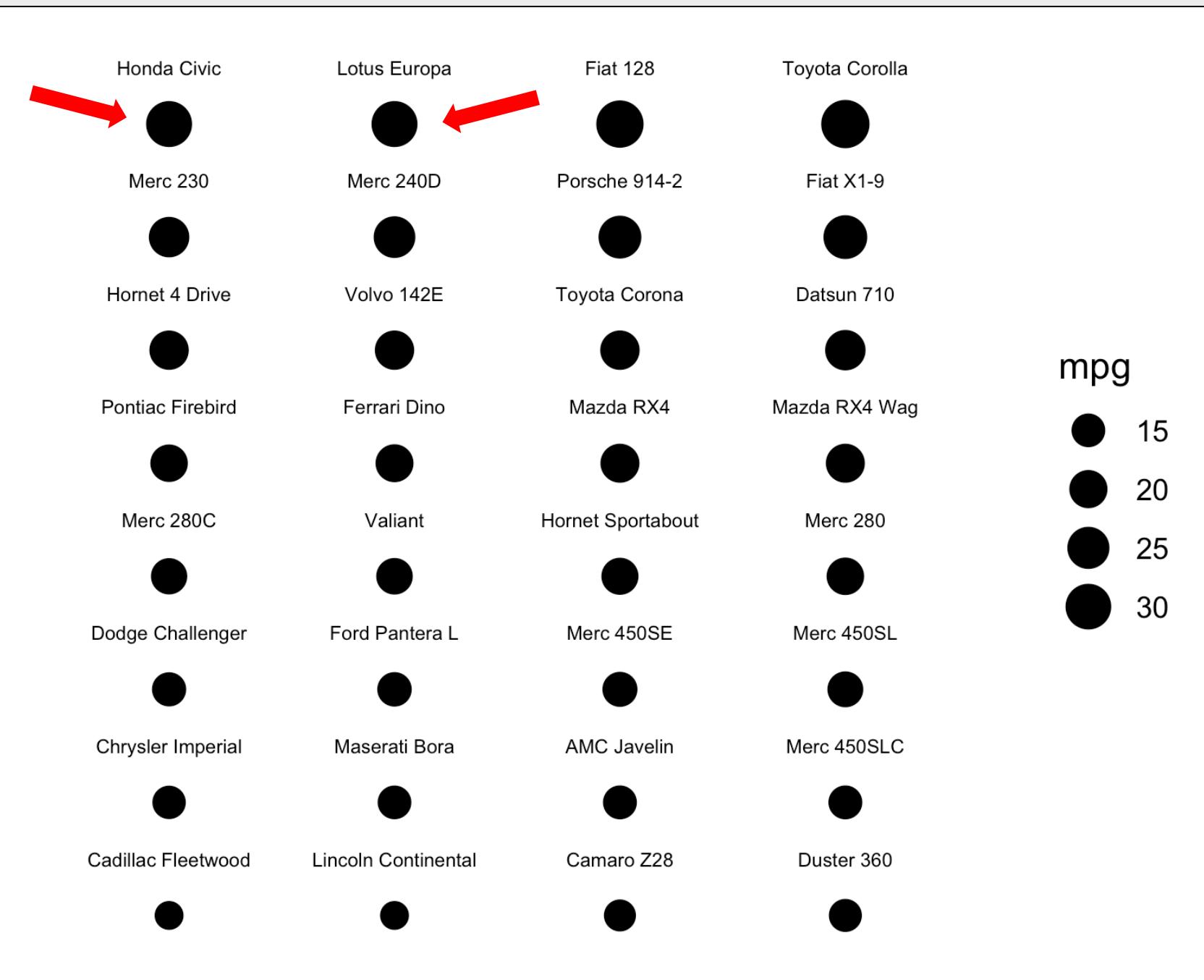
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ratioing?

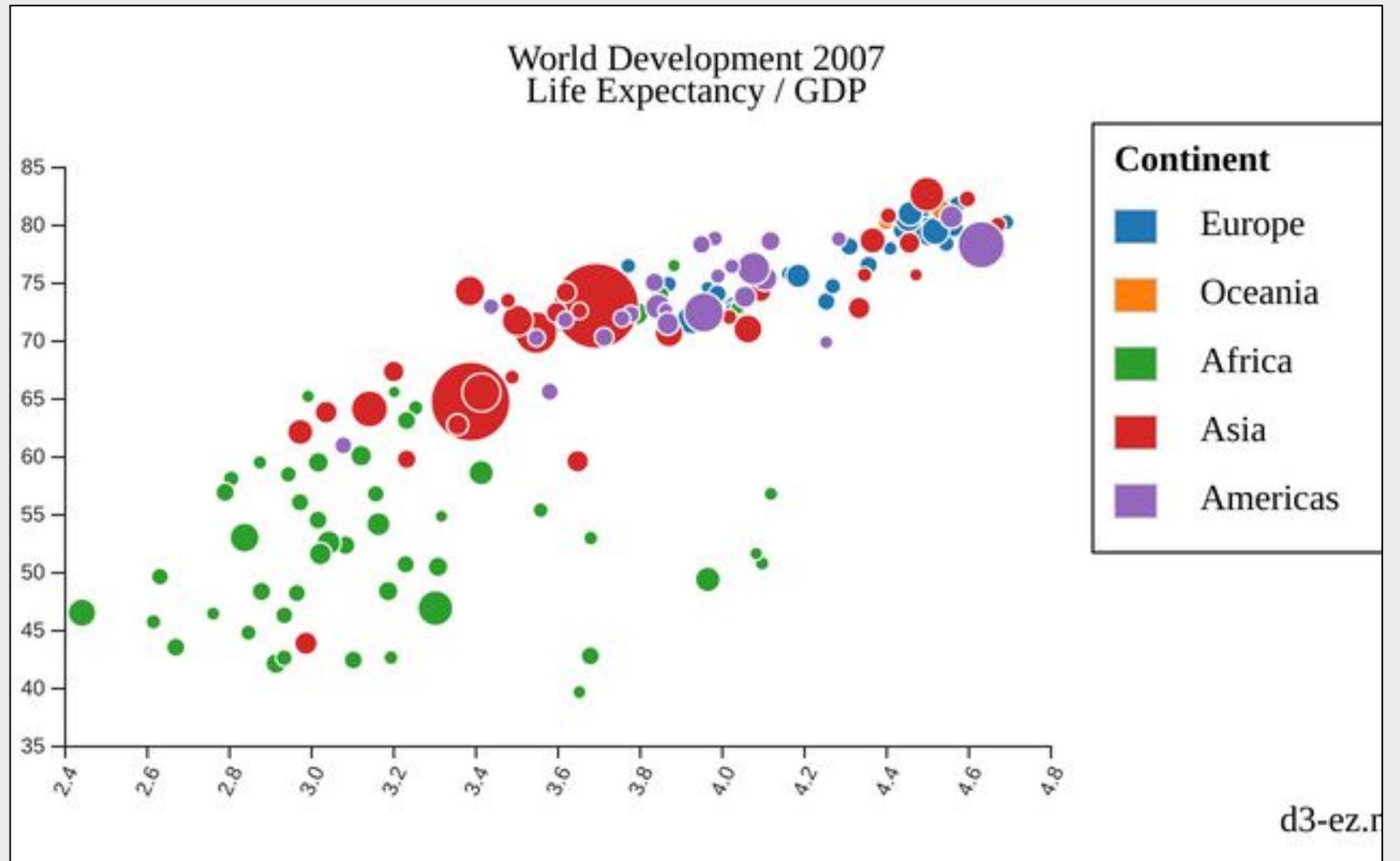


1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Discrimination?  
Ranking?



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ratioing?

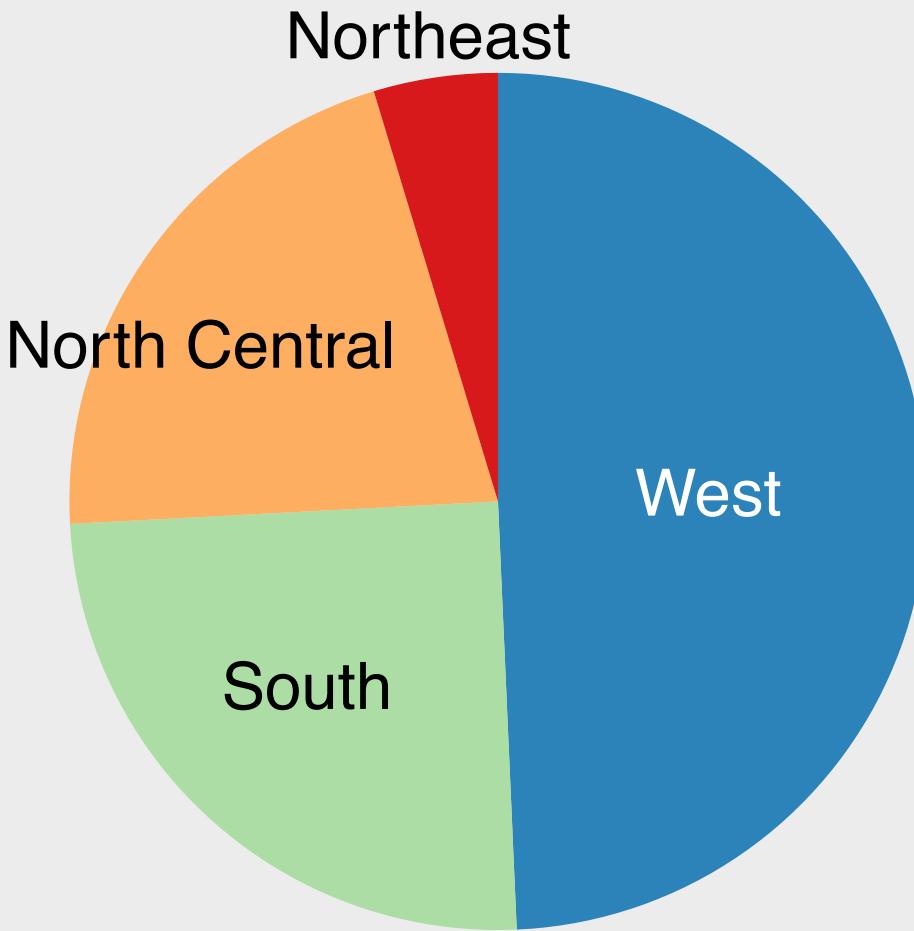


1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Discrimination?  
Ranking?



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue



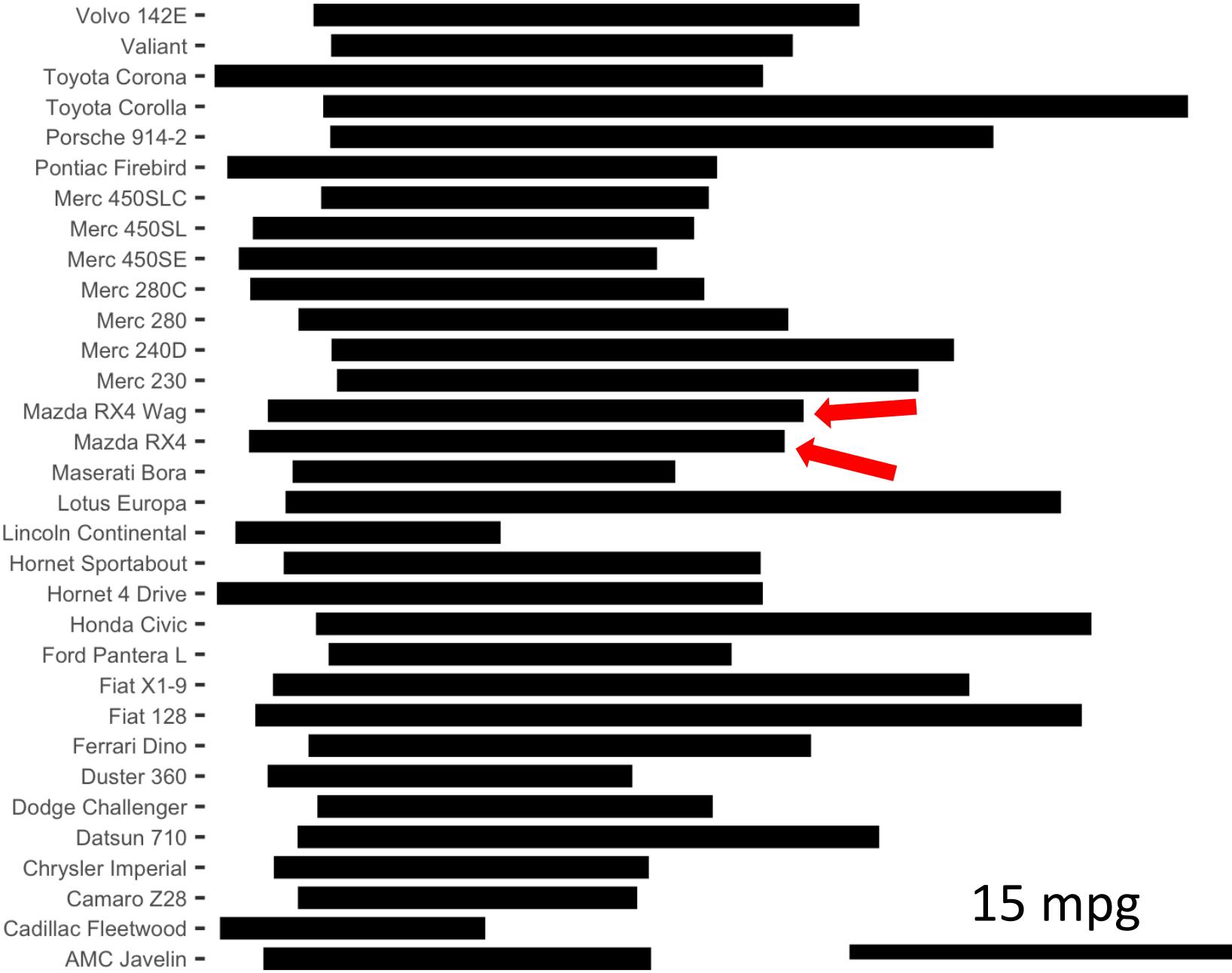
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ratioing?



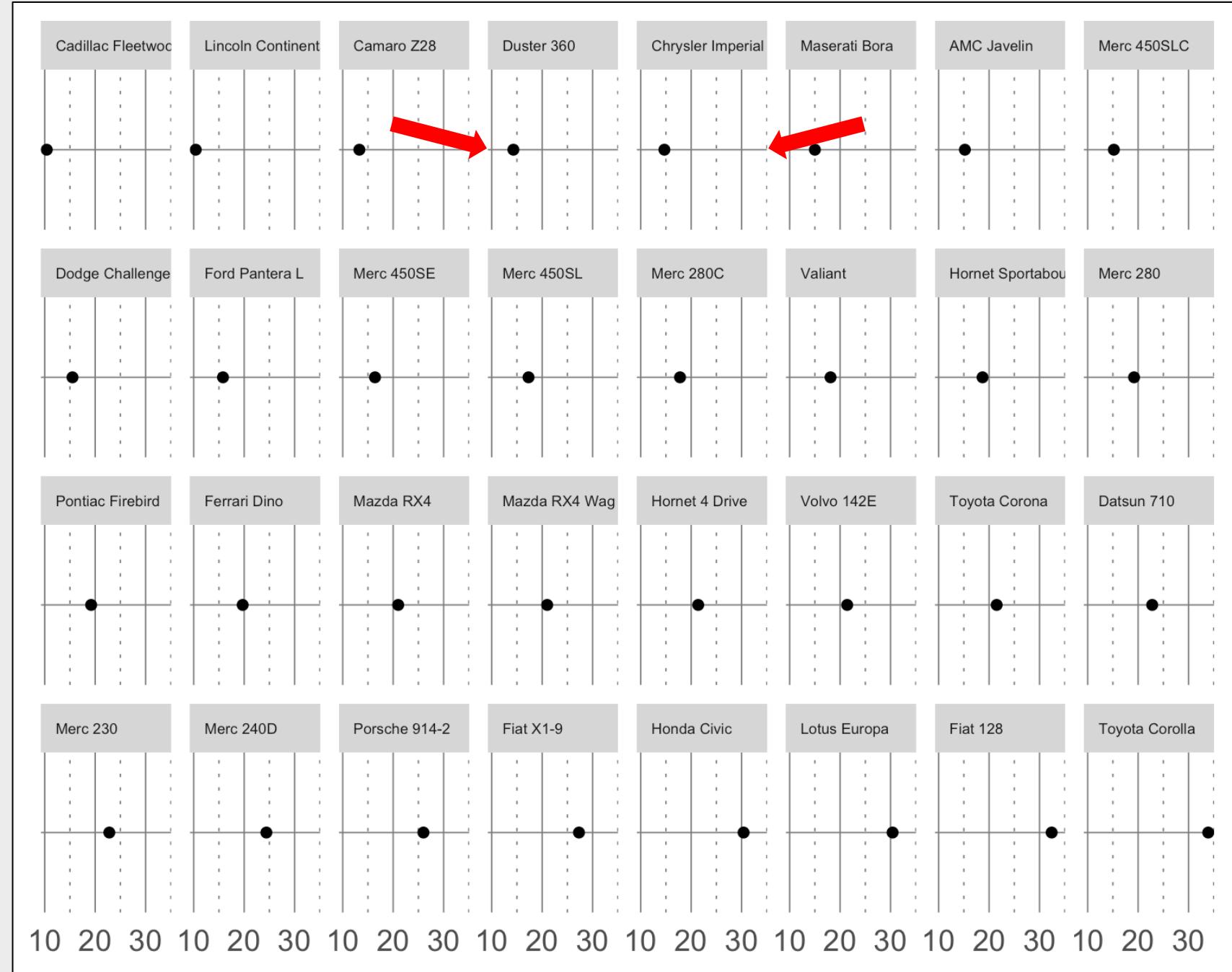
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Discrimination?  
Ranking?



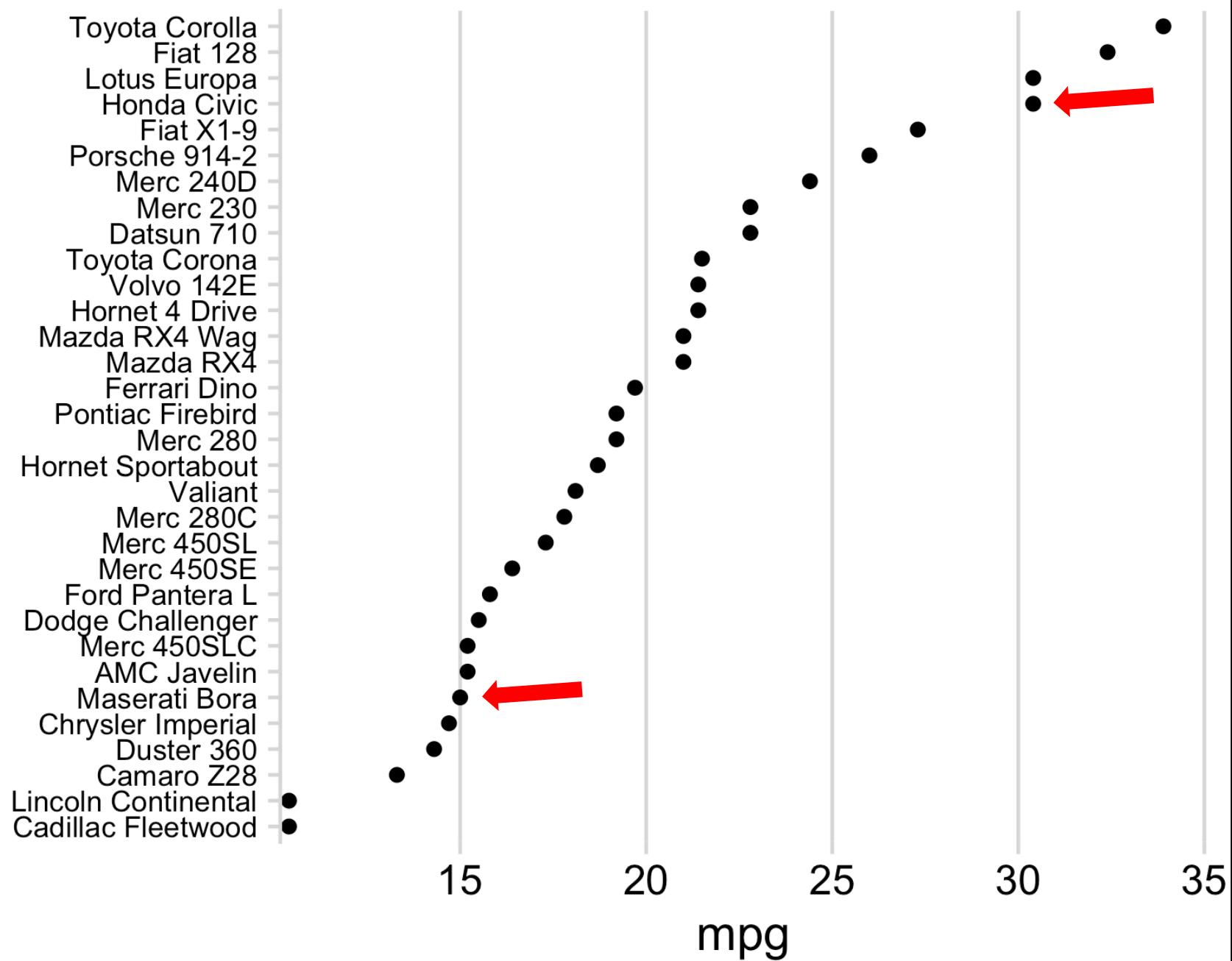
1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Discrimination?  
Ranking?



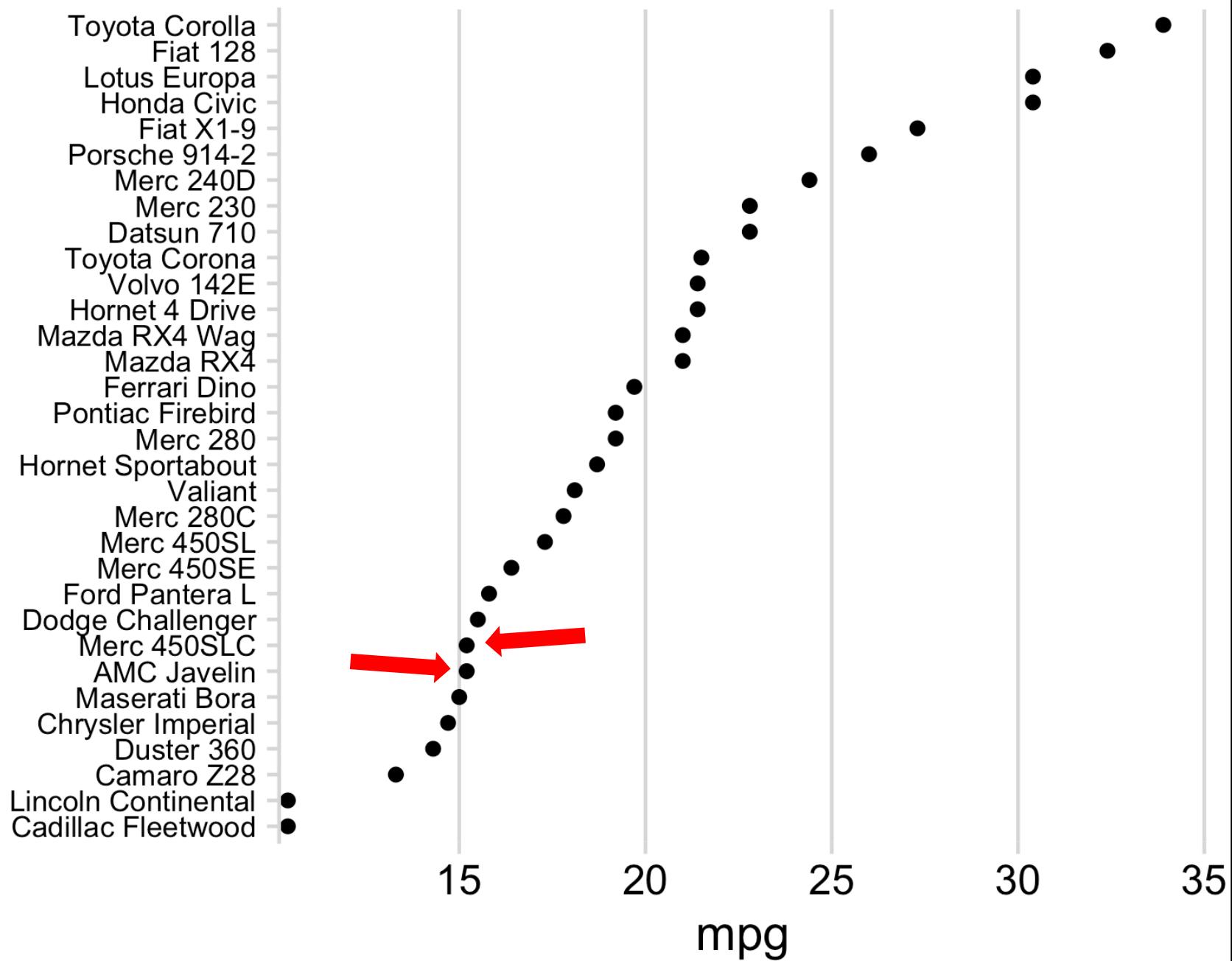
1. Position on a common scale
  2. Position on non-aligned scales
  3. Length
  4. Angle
  5. Area
  6. Color saturation
  7. Color hue

# Ratioing?



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

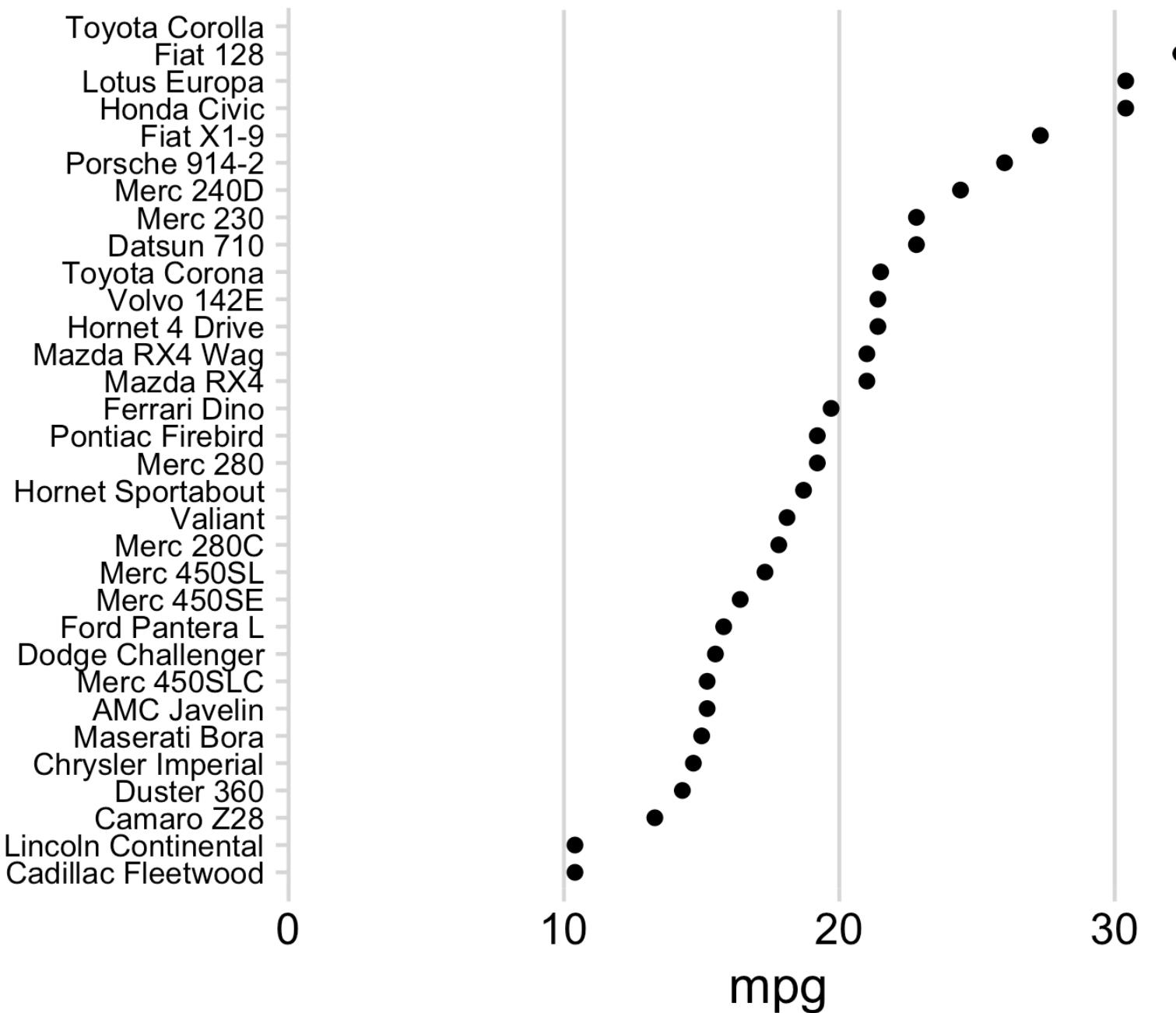
Discrimination?  
Ranking?



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

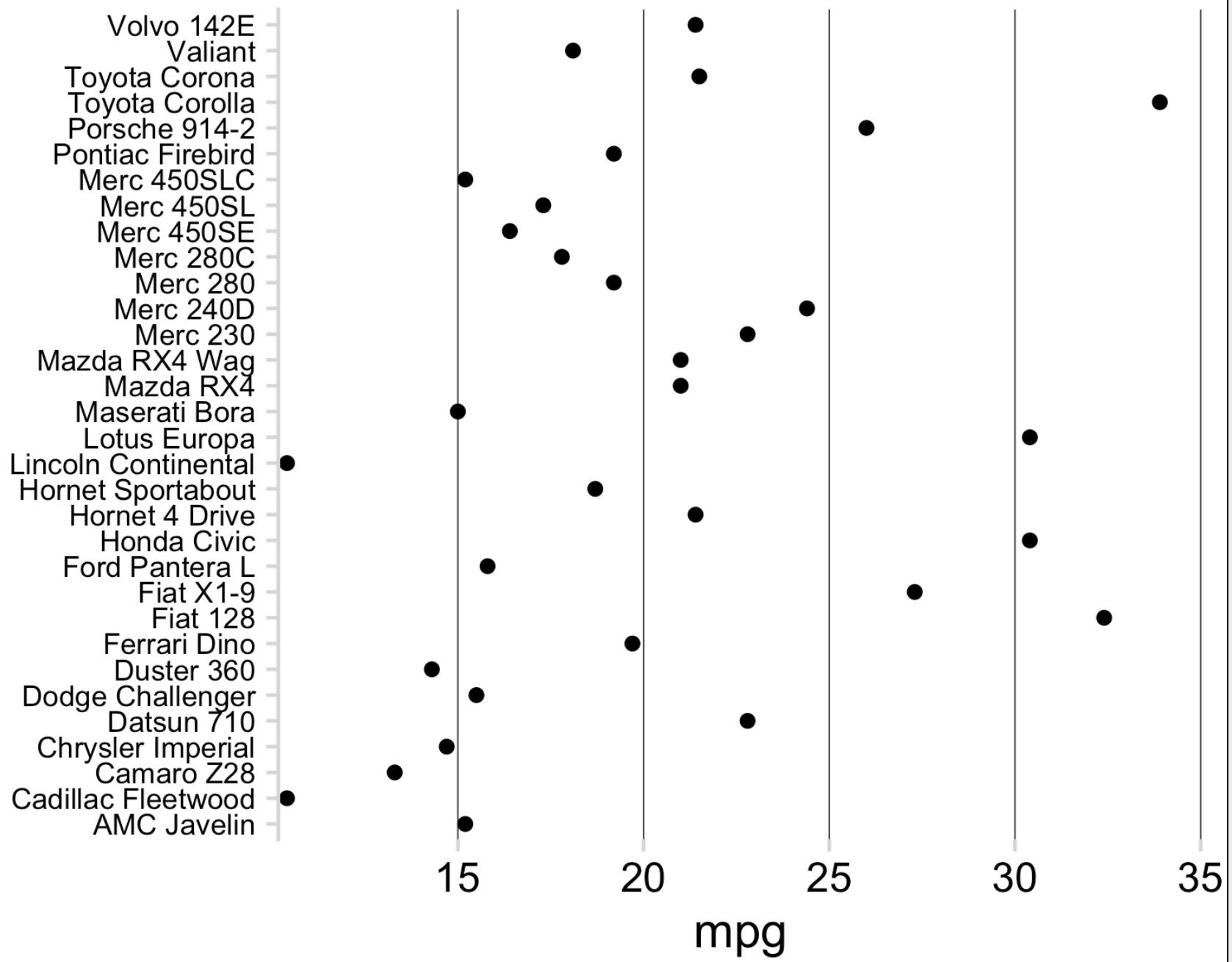
No need to scale to 0:

- Lowers resolution
- Isn't needed for accurate ratioing



1. Position on a common scale
2. Position on non-aligned scales
3. Length
4. Angle
5. Area
6. Color saturation
7. Color hue

Ordering still crucial



# Cleveland's three visual operations of pattern perception:

- Estimation
- **Assembly** → The grouping of graphical elements
- Detection

# Gestalt Psychology

The whole has a reality that is entirely separate from the parts



**WWF**

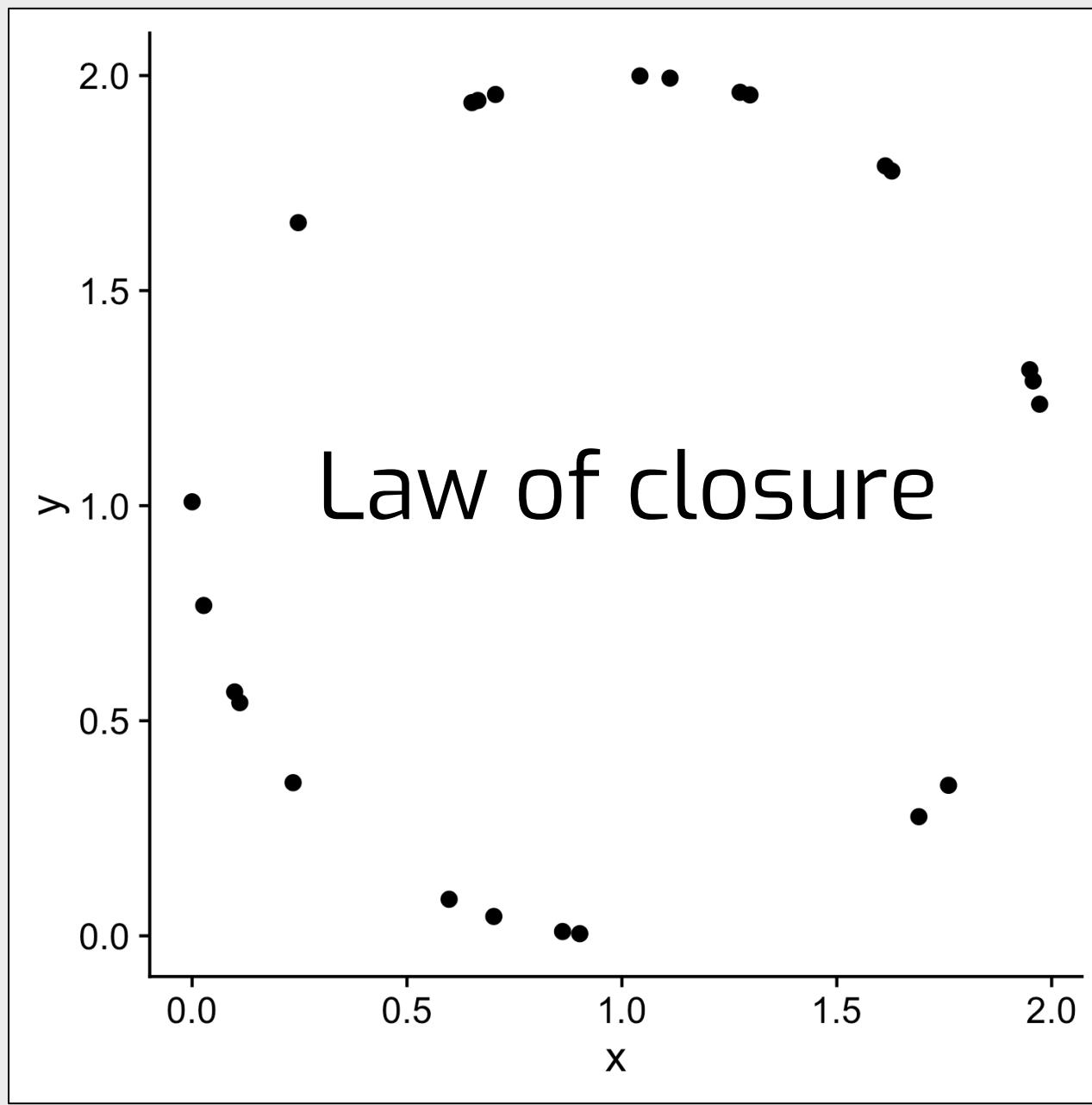
# Reification





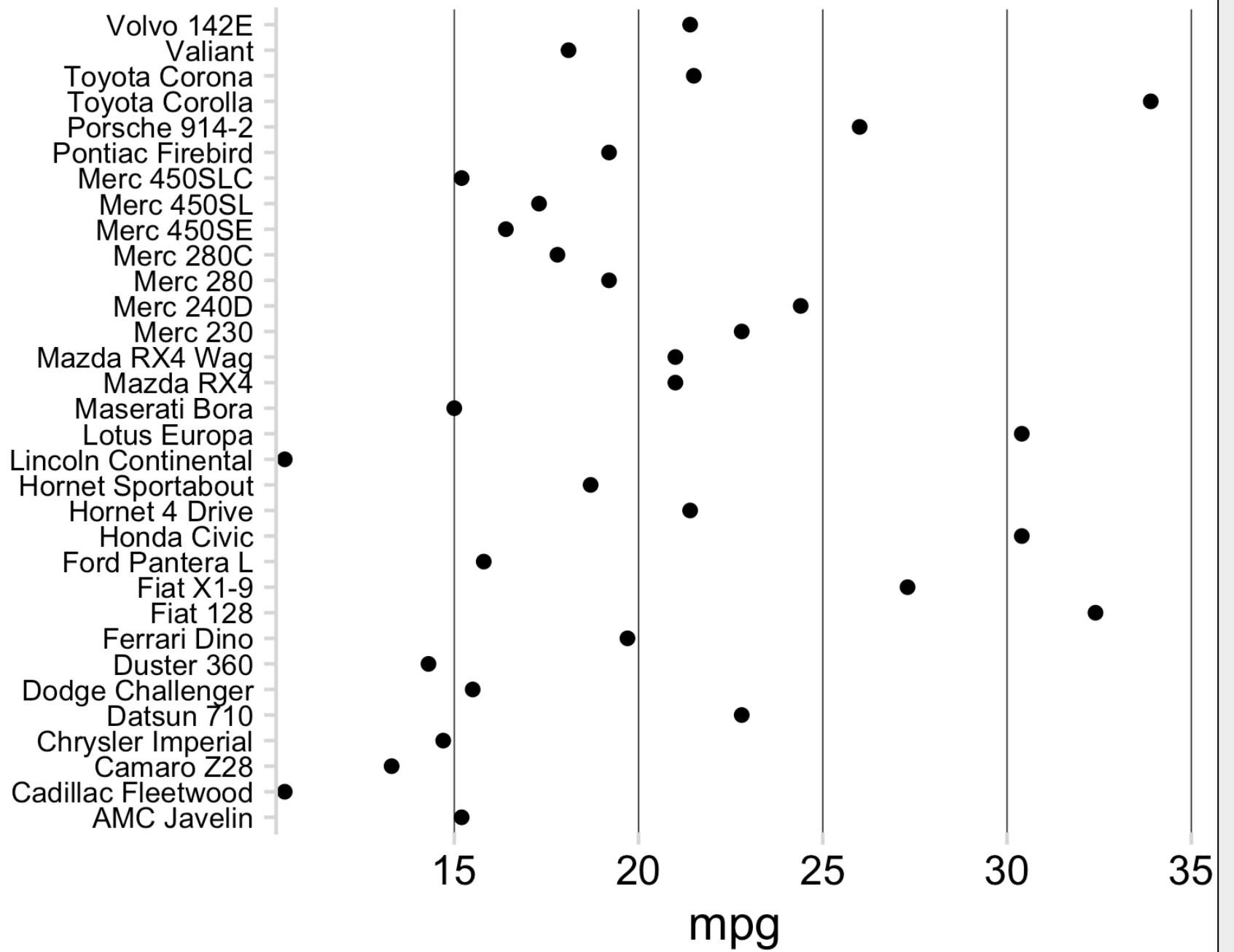
Emergence

X	y
1.972	1.236
1.112	1.994
0	1.009
0.665	1.942
0.235	0.356
0.247	1.658
1.275	1.961
0.702	0.045
1.76	0.35
1.691	0.277
1.628	1.778
1.957	1.29
0.111	0.542
0.902	0.005
0.598	0.085
1.613	1.79
1.298	1.955
0.651	1.937
1.949	1.316
0.099	0.567
0.862	0.01
0.027	0.768
0.706	1.956
1.042	1.999



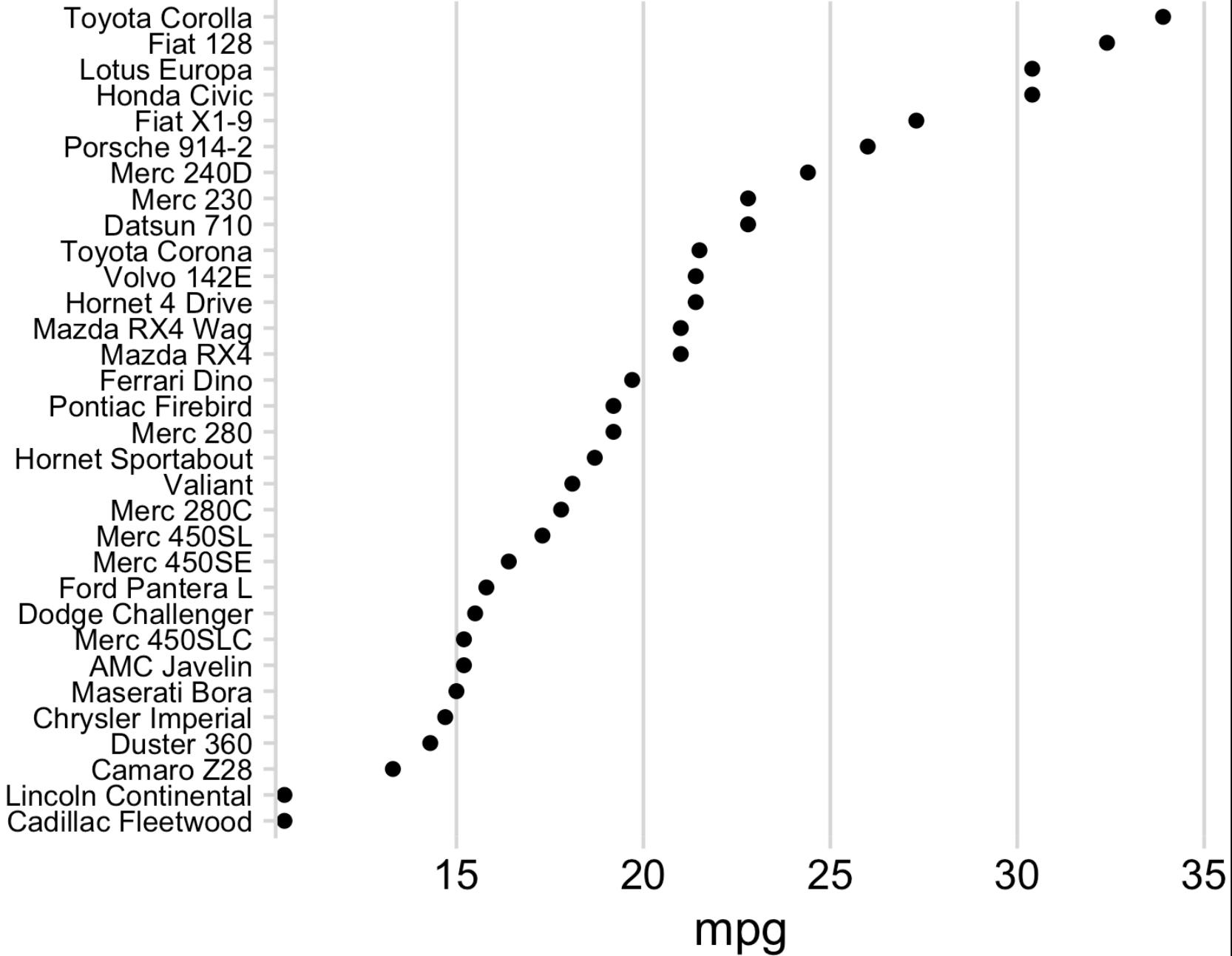
# Prägnanz

We strongly prefer to interpret stimuli as regular, simple, and orderly



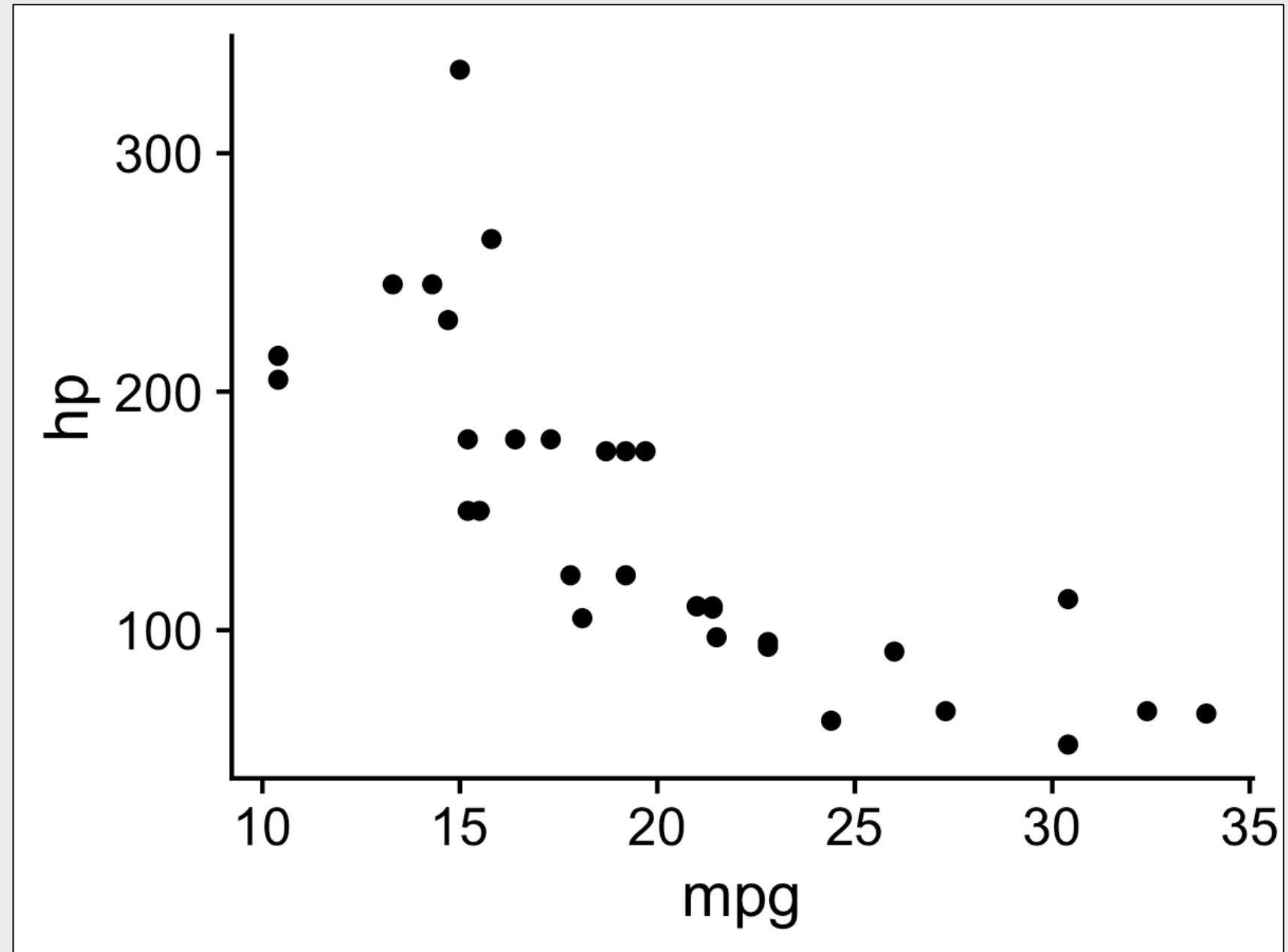
# Law of continuity

We will group together objects that follow an established direction



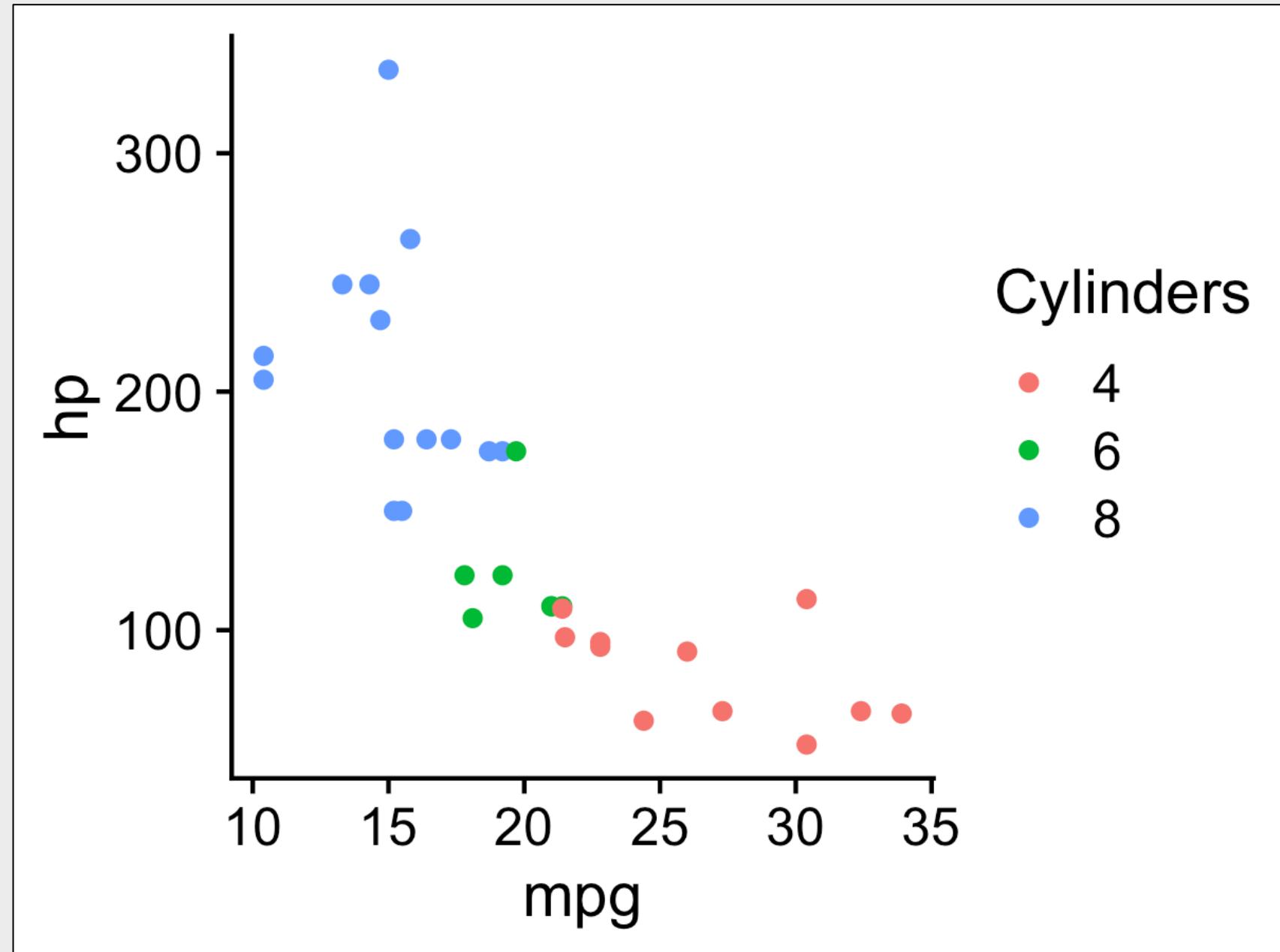
# Law of continuity

We will group together objects that follow an established direction



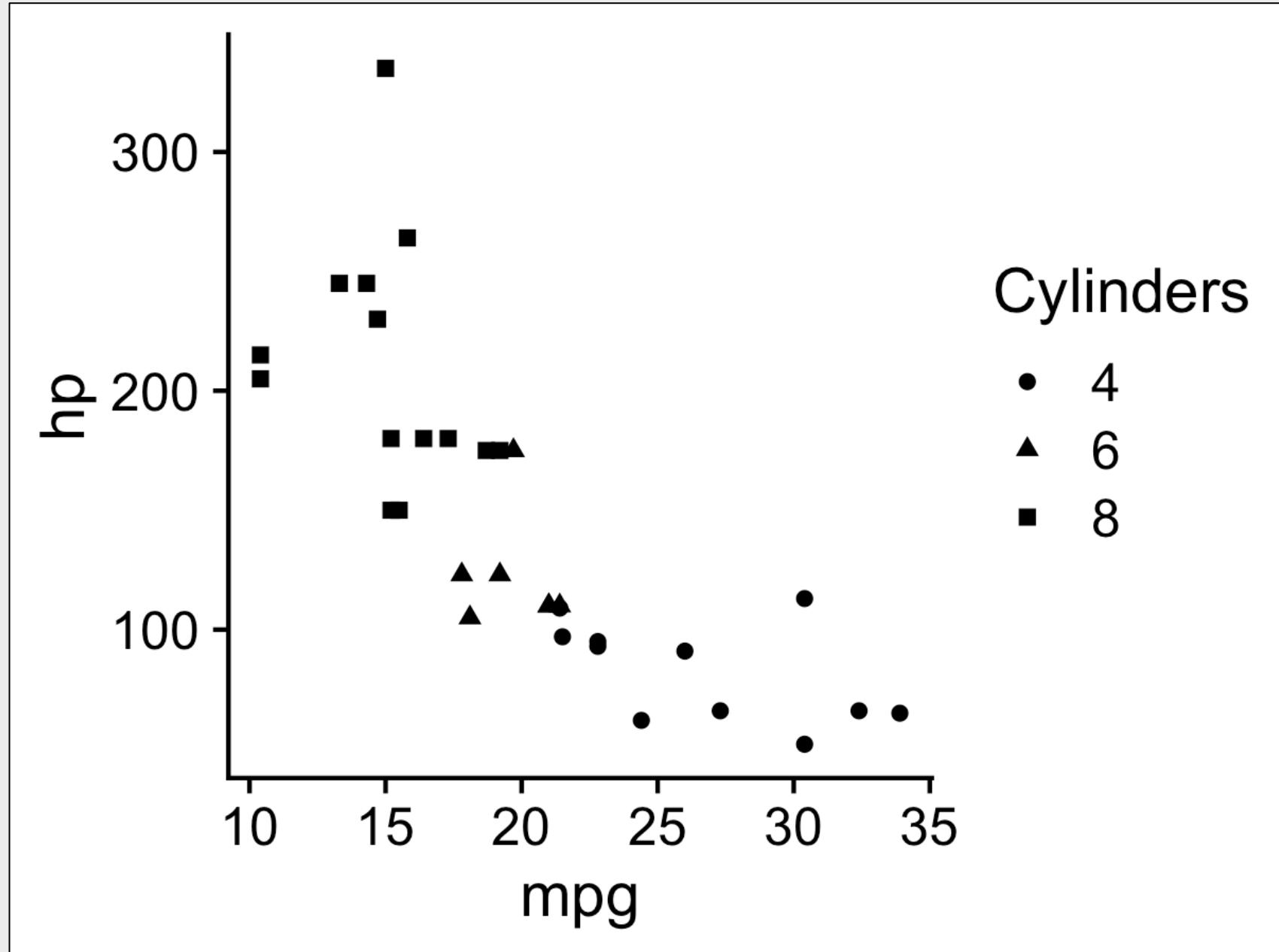
# Law of similarity

We tend to see elements that are physically *similar* as part of the same object



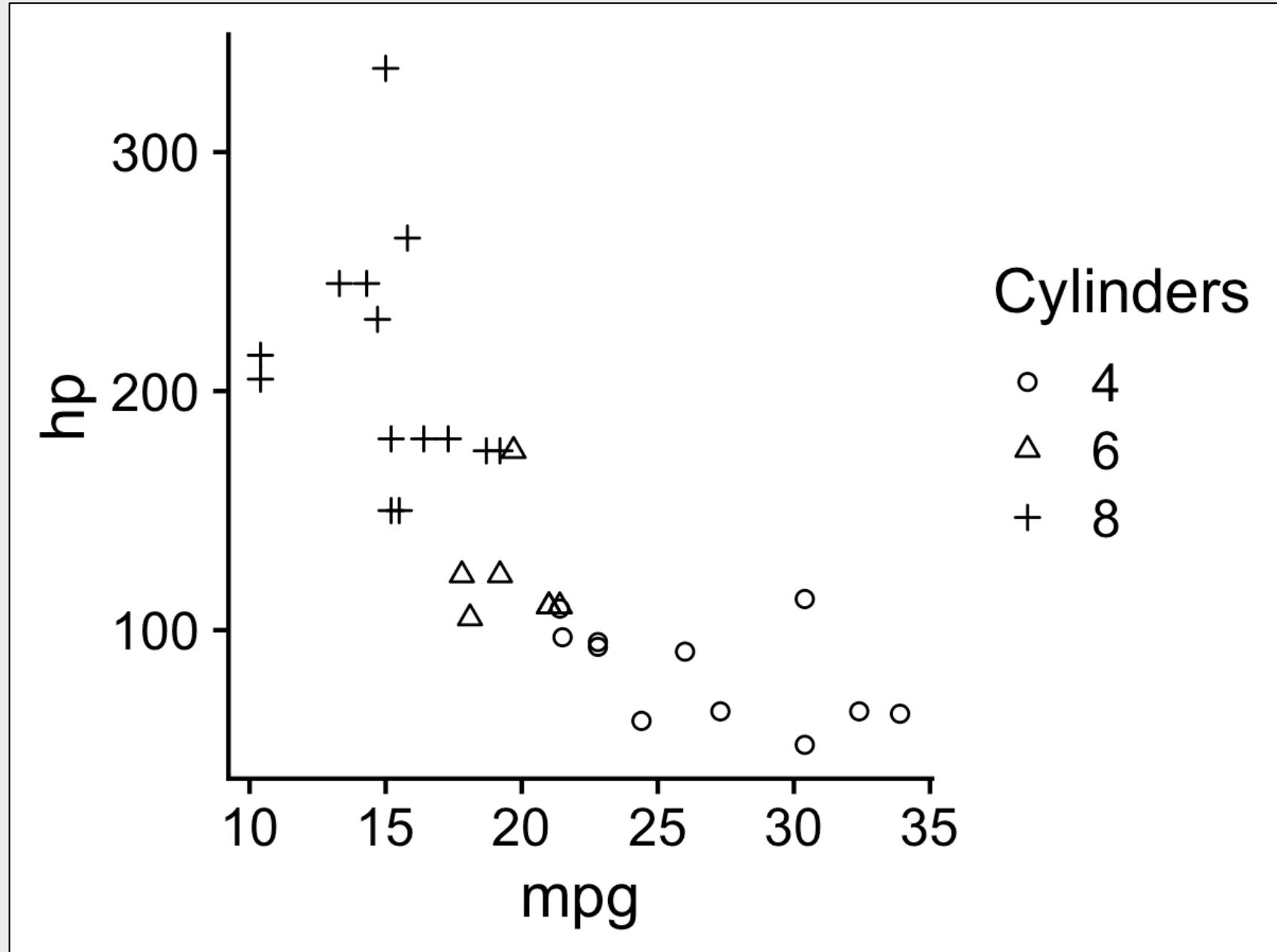
# Law of similarity

We tend to see elements that are physically *similar* as part of the same object



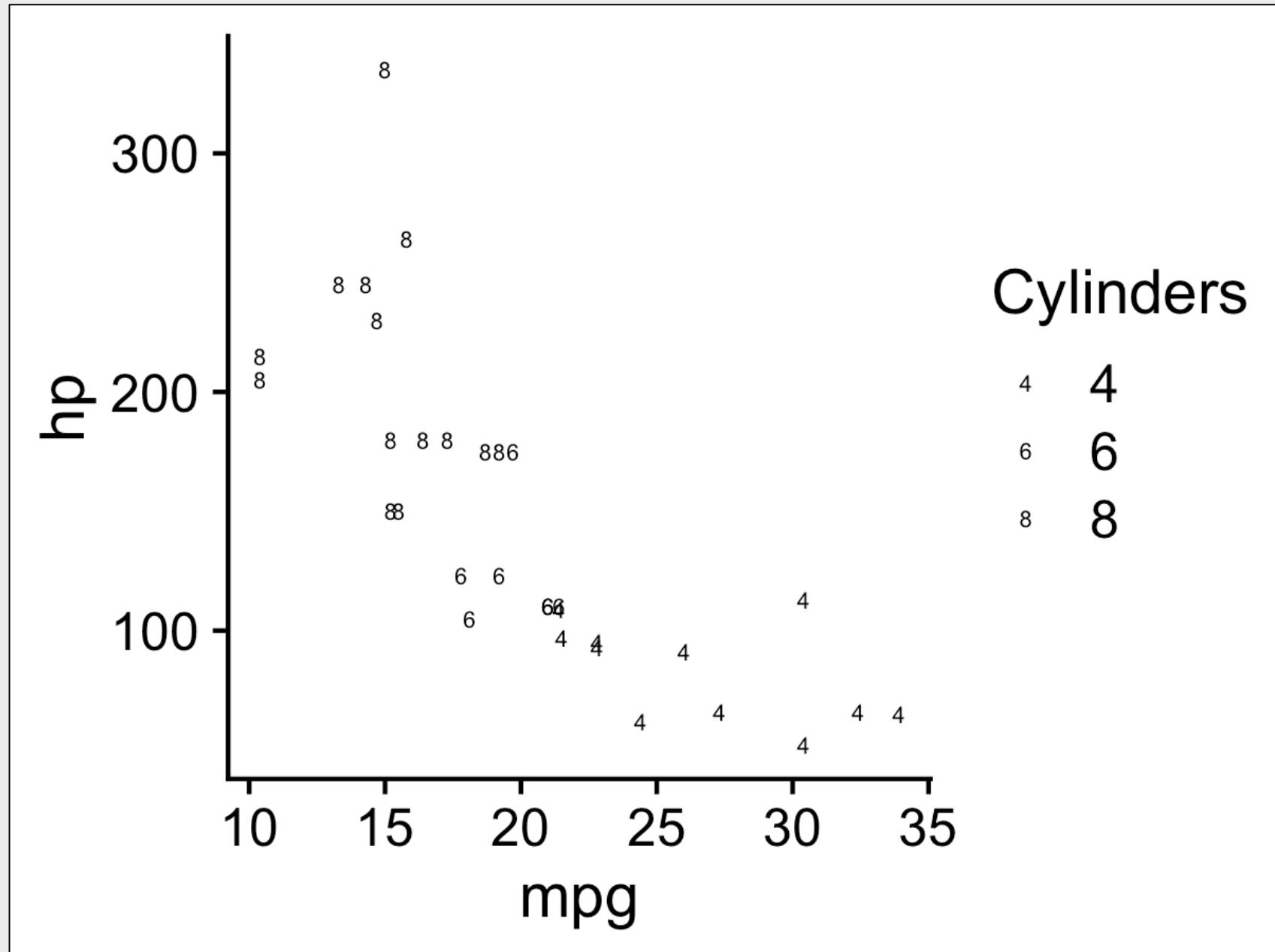
# Law of similarity

We tend to see elements that are physically *similar* as part of the same object



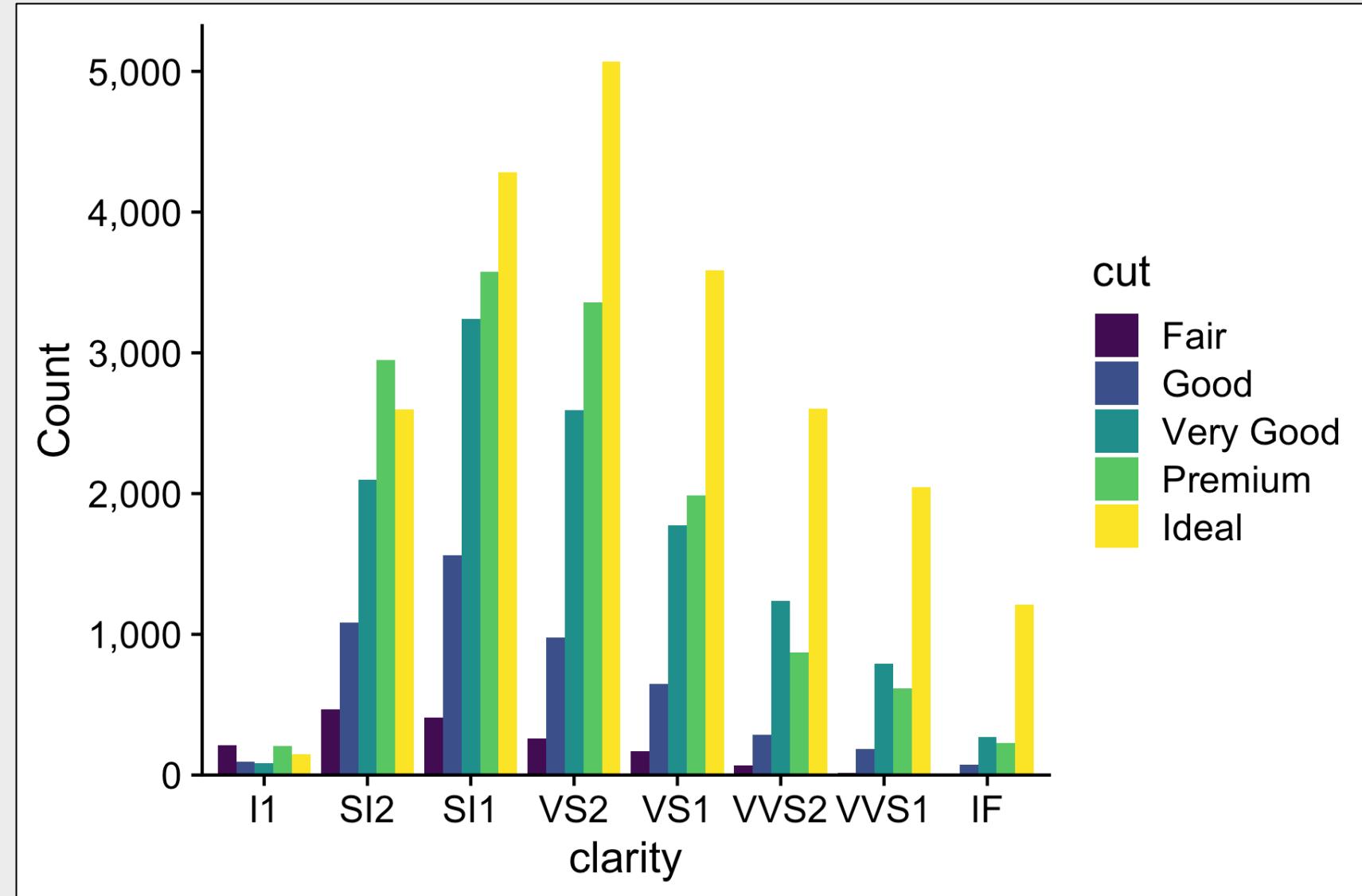
# Law of similarity

We tend to see elements that are physically *similar* as part of the same object



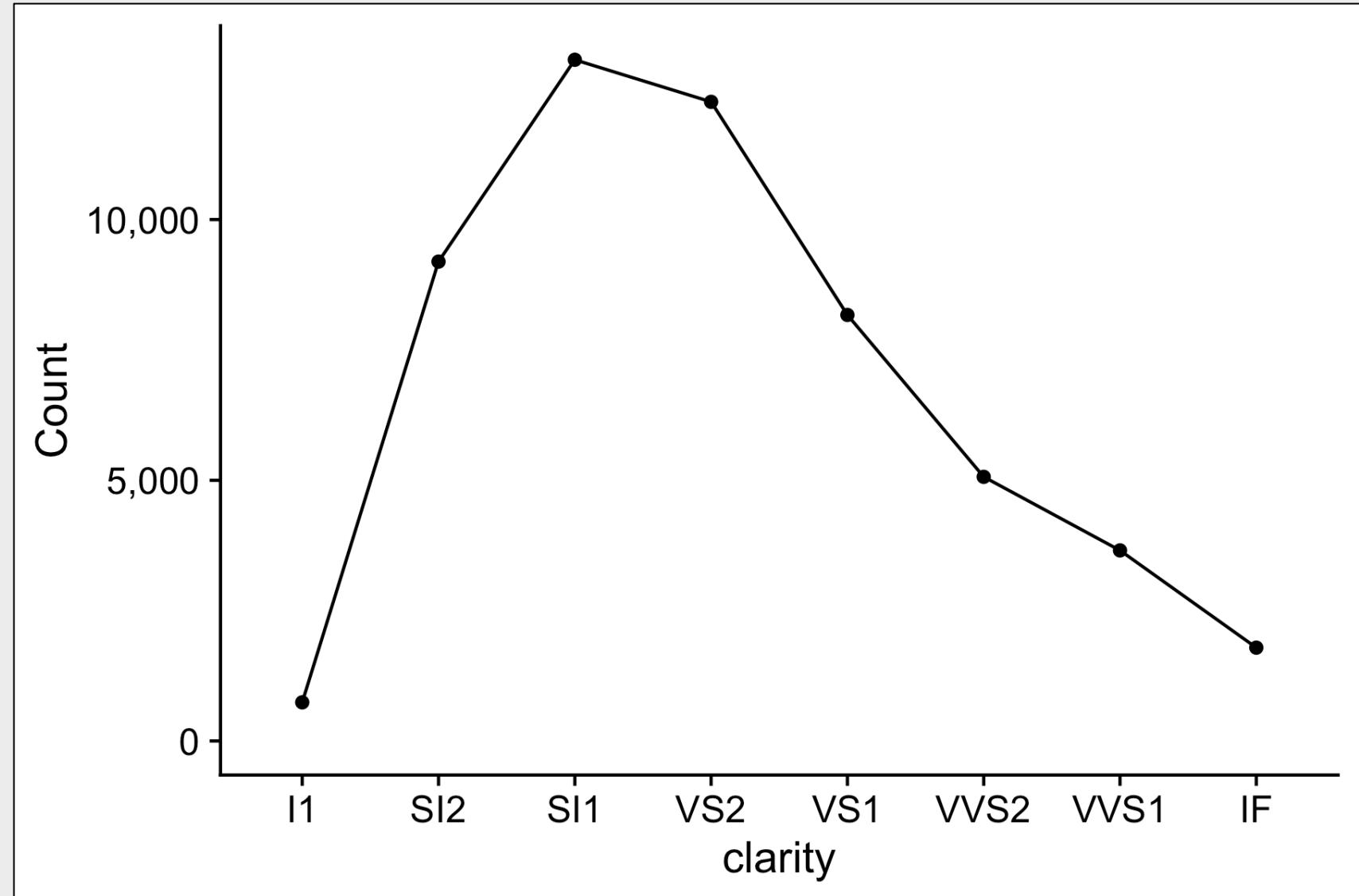
# Law of proximity

We tend to see elements that are physically *near* each other as part of the same object



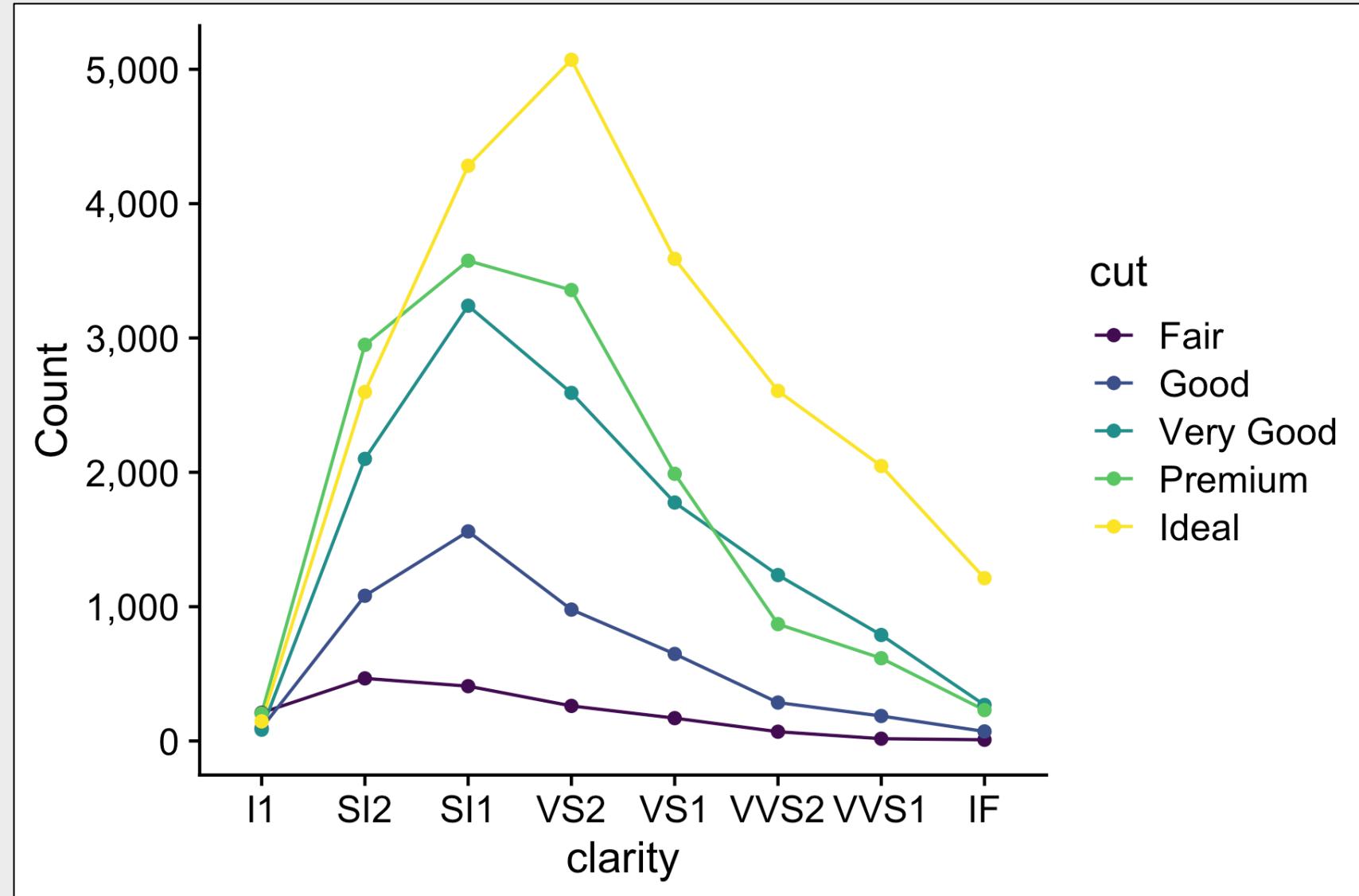
# Law of proximity

We tend to see elements that are physically *near* each other as part of the same object



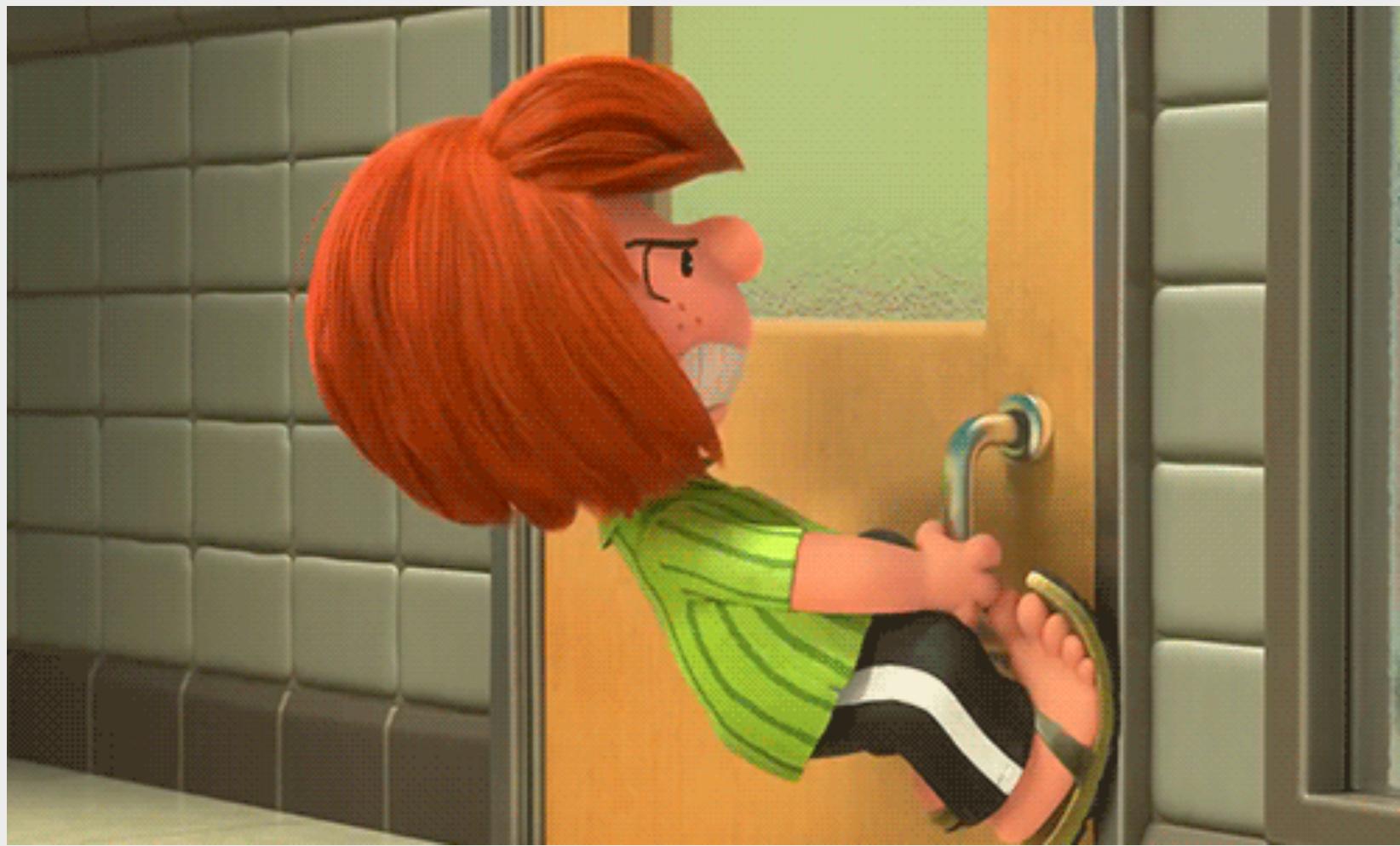
# Law of proximity

We tend to see elements that are physically *near* each other as part of the same object



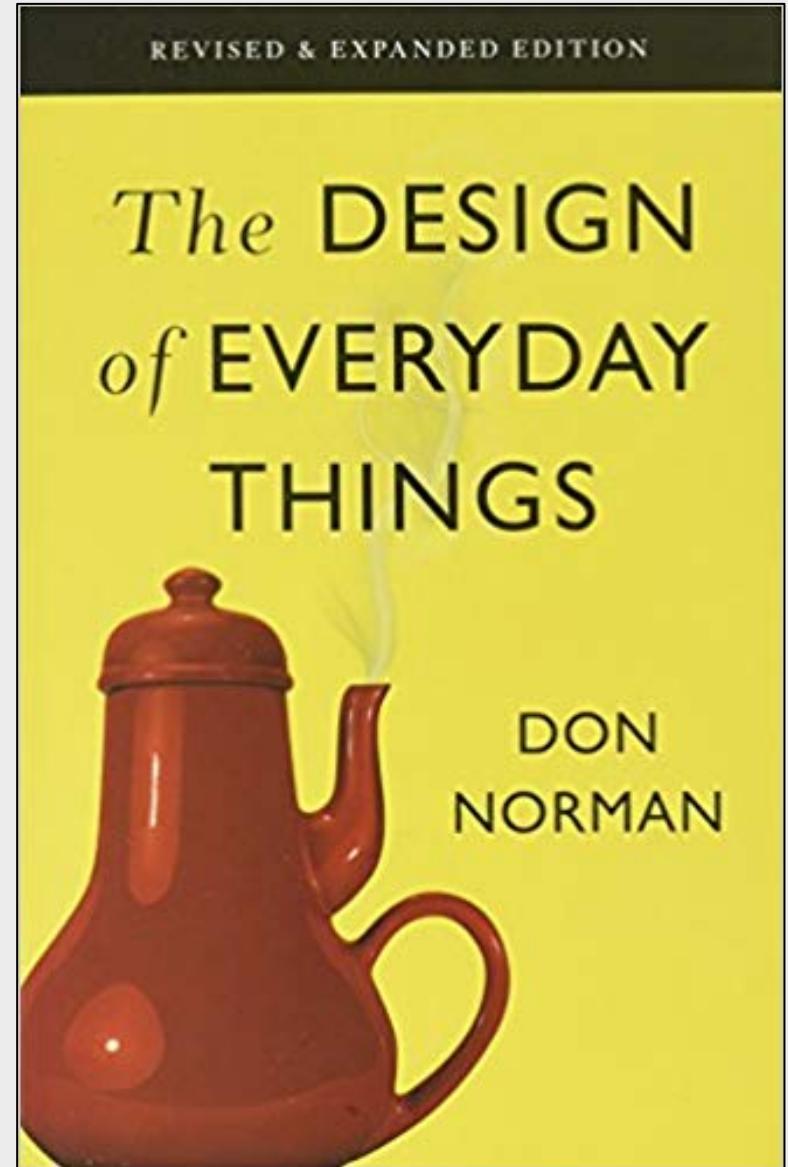
# Cleveland's three visual operations of pattern perception:

- Estimation
- Assembly
- Detection



## Norman door (n.):

1. A door where the design tells you to do the opposite of what you're actually supposed to do.
2. A door that gives the wrong signal and needs a sign to correct it.





Norman door

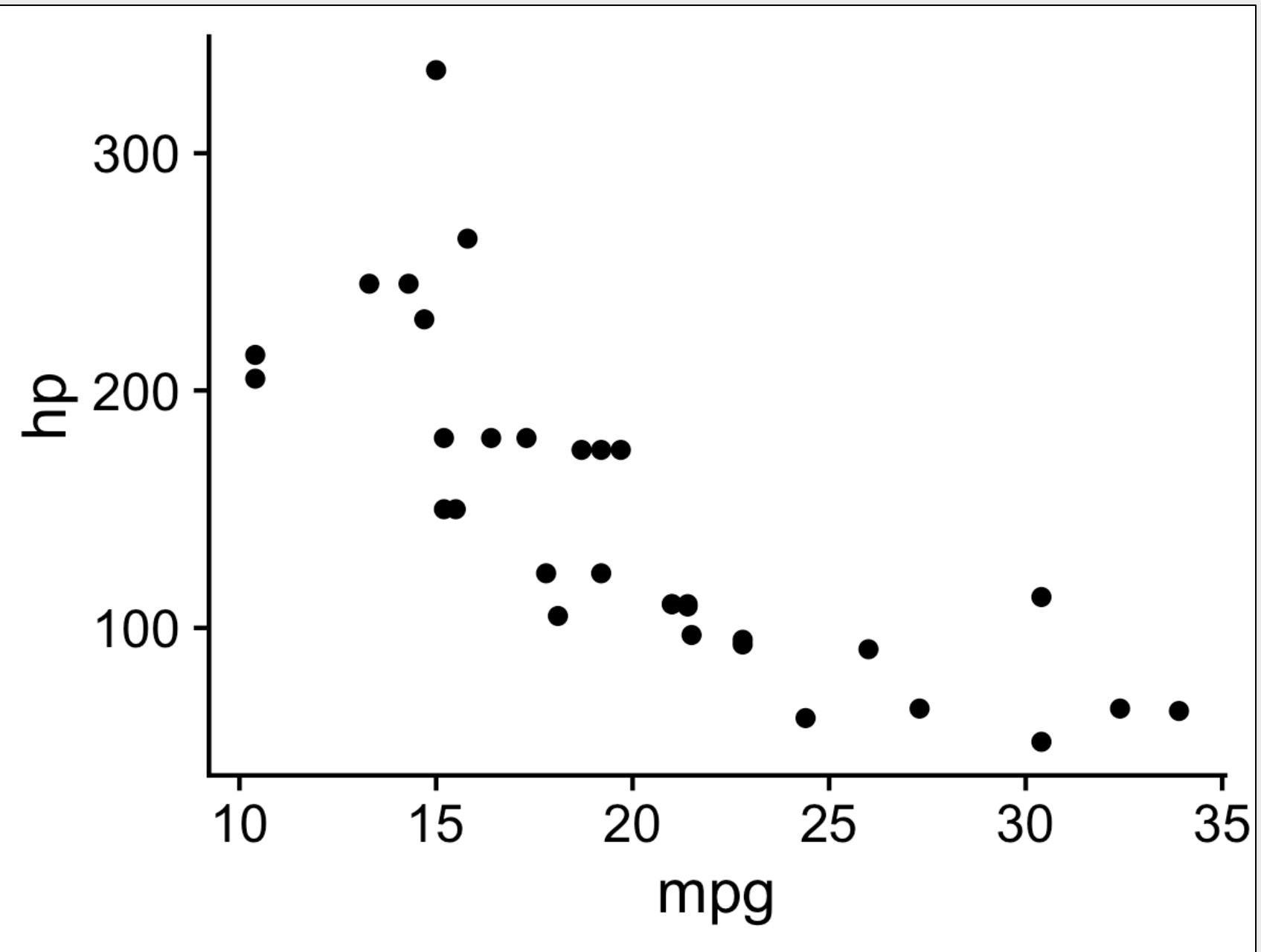


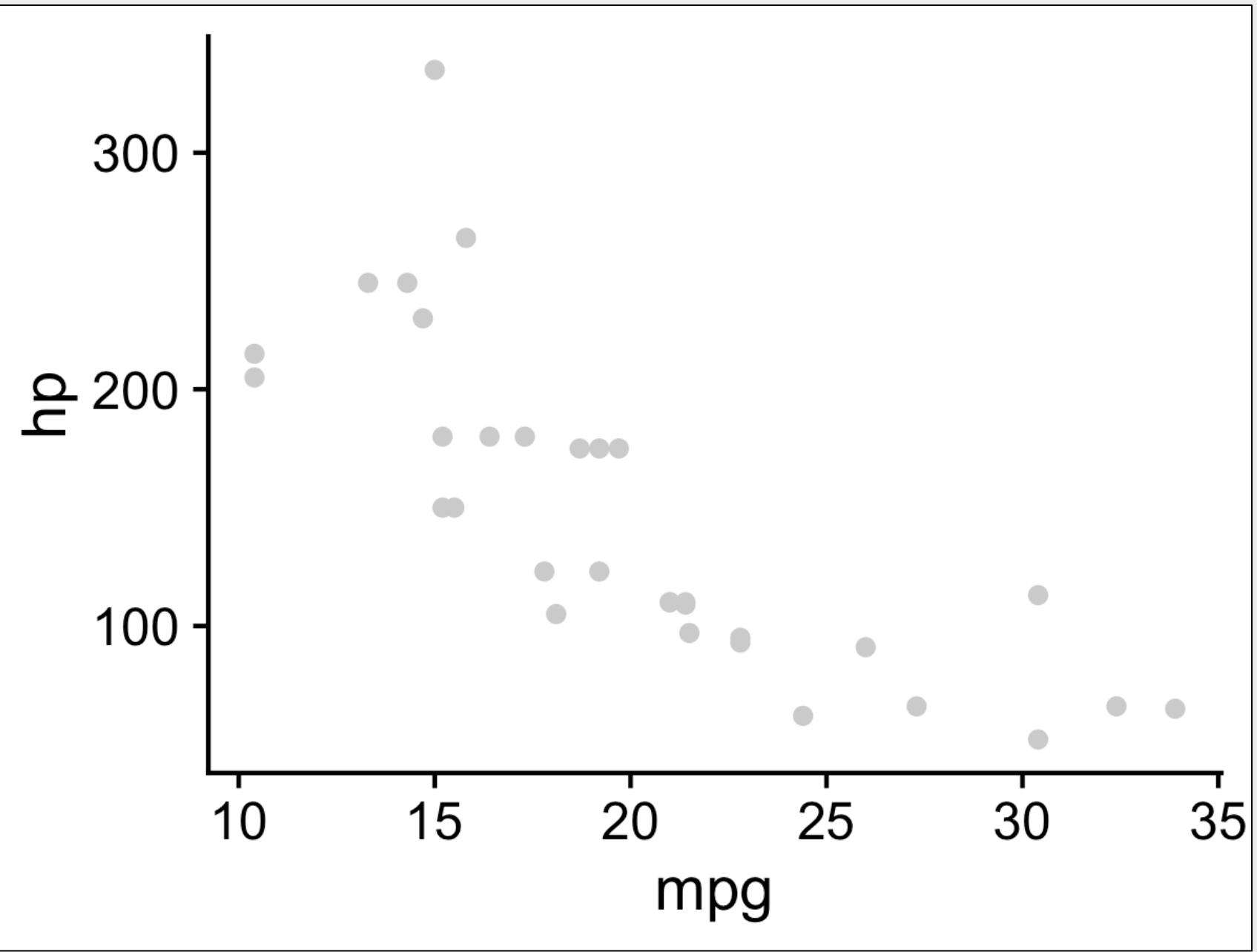
Non-Norman door

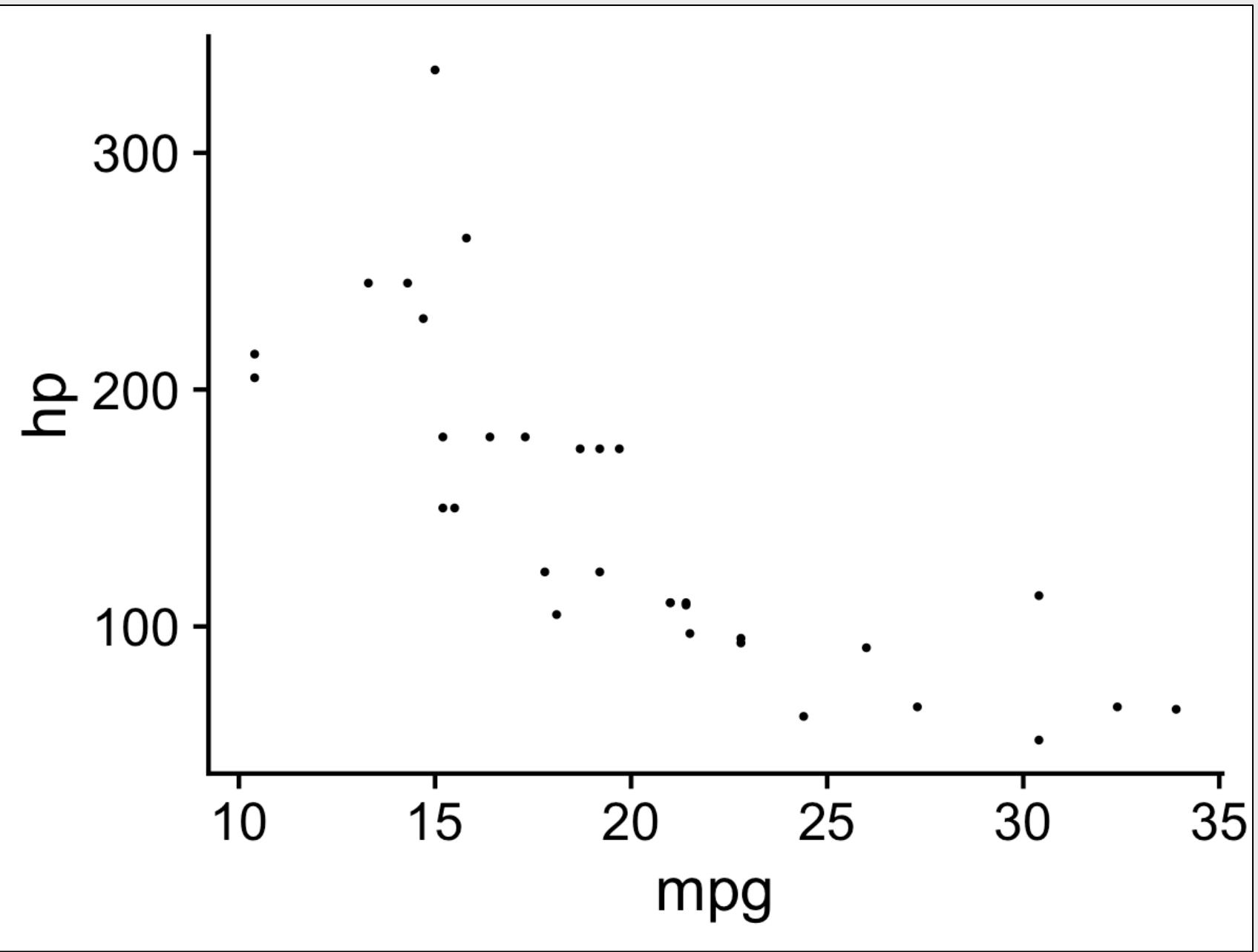


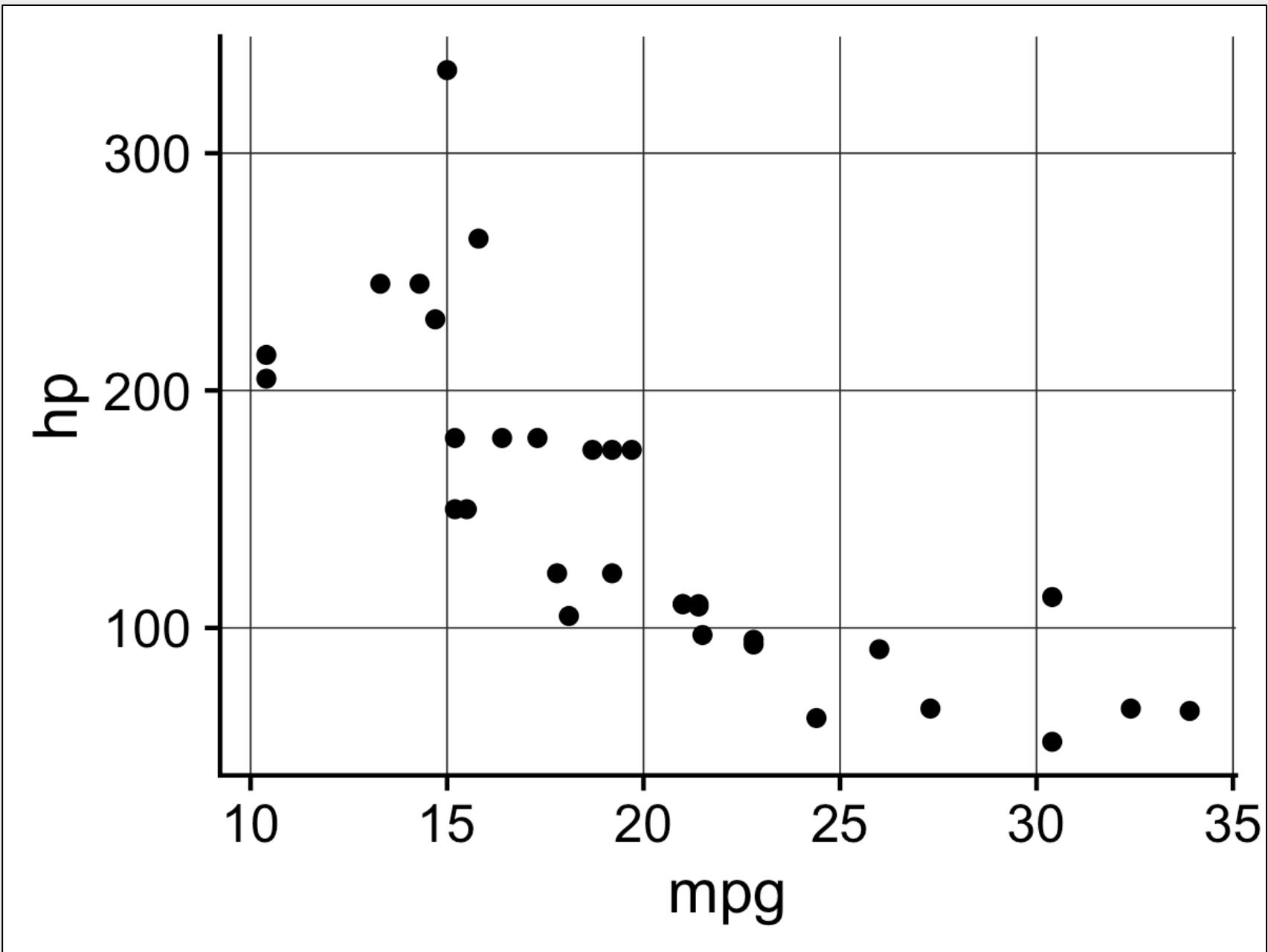
# Cleveland's three visual operations of pattern perception:

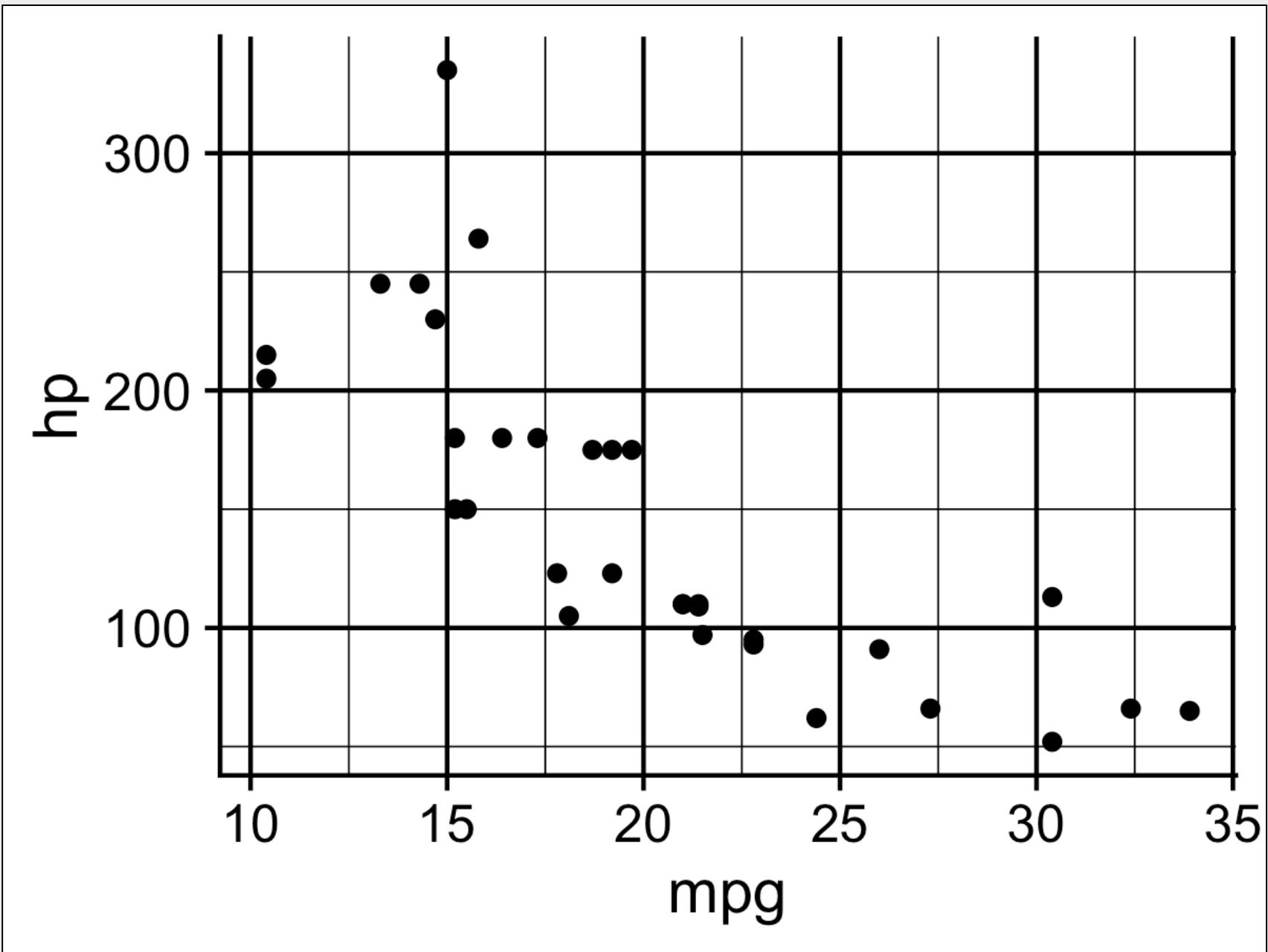
- Estimation
- Assembly
- **Detection** → Recognizing that a geometric object encodes a physical value

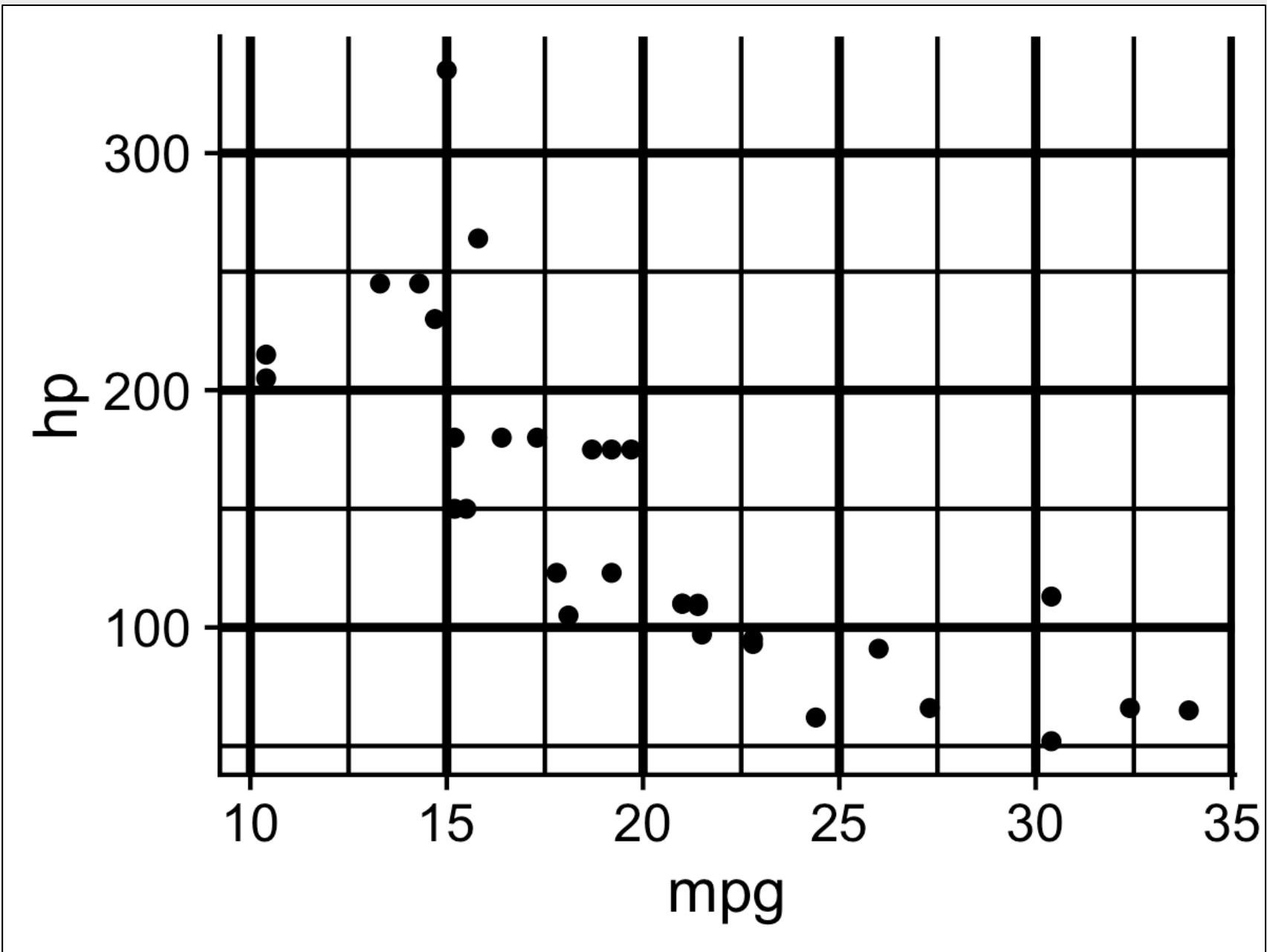




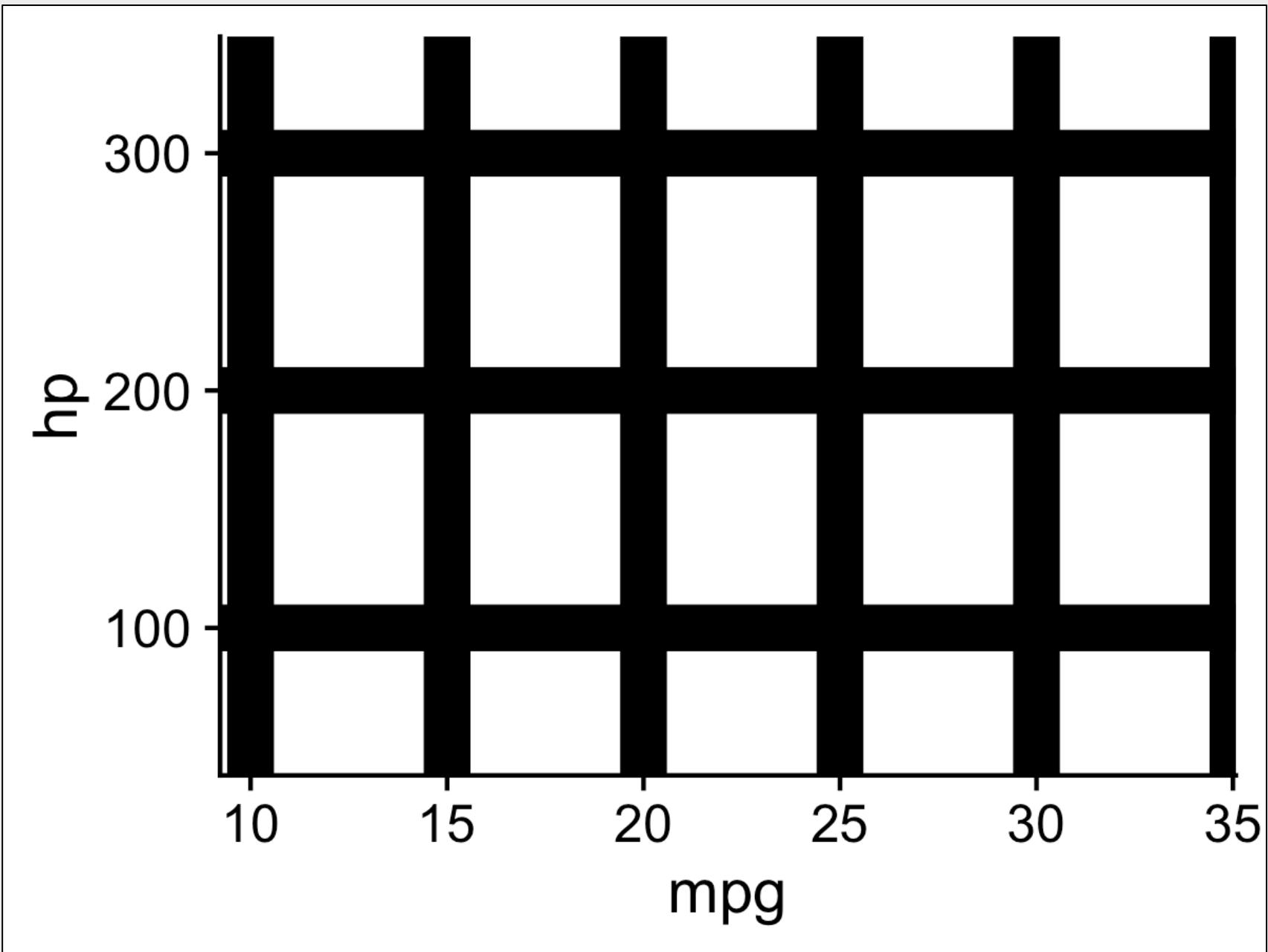






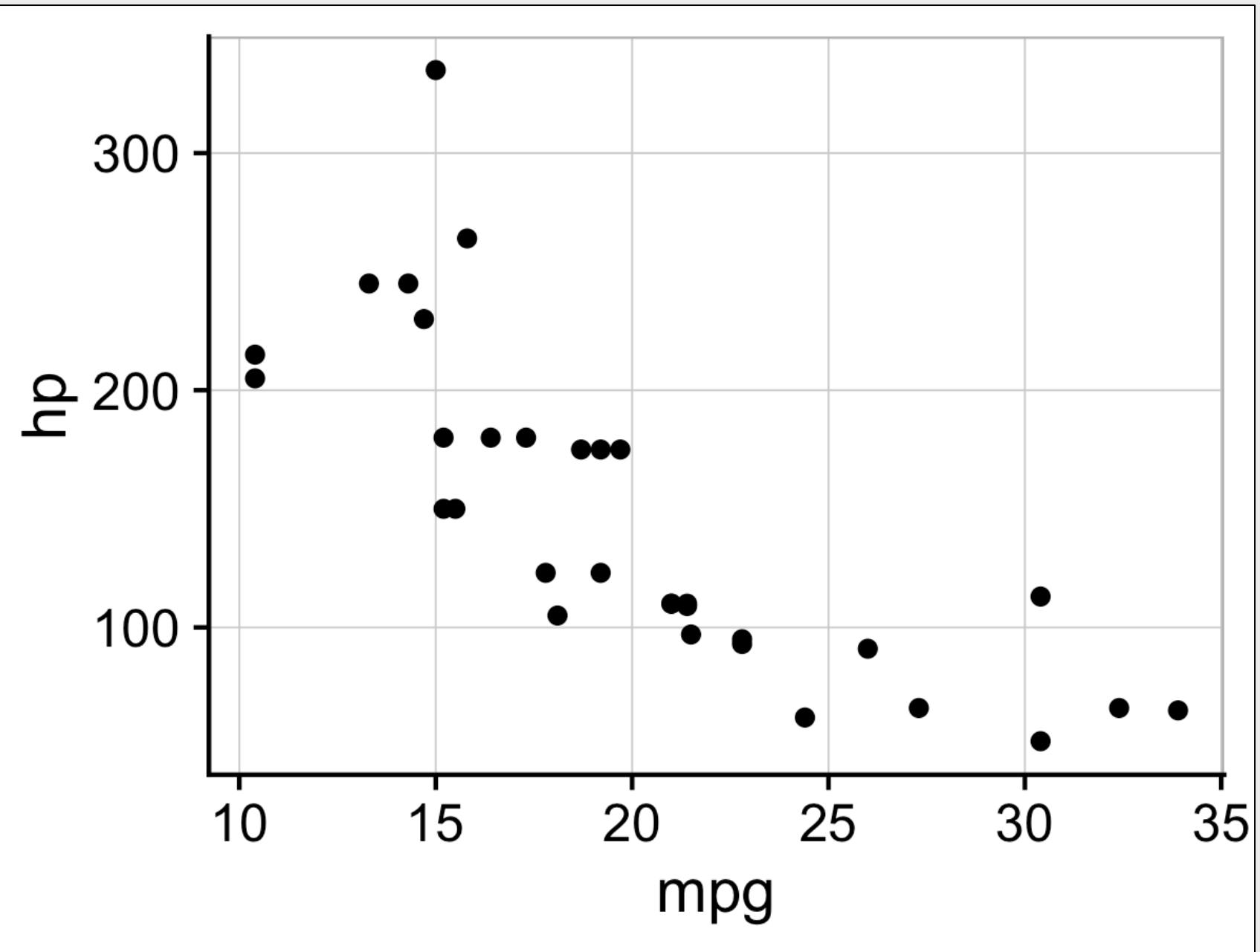


## The Hermann Grid illusion



*“Above all else, show the data”*

- Edward Tufte



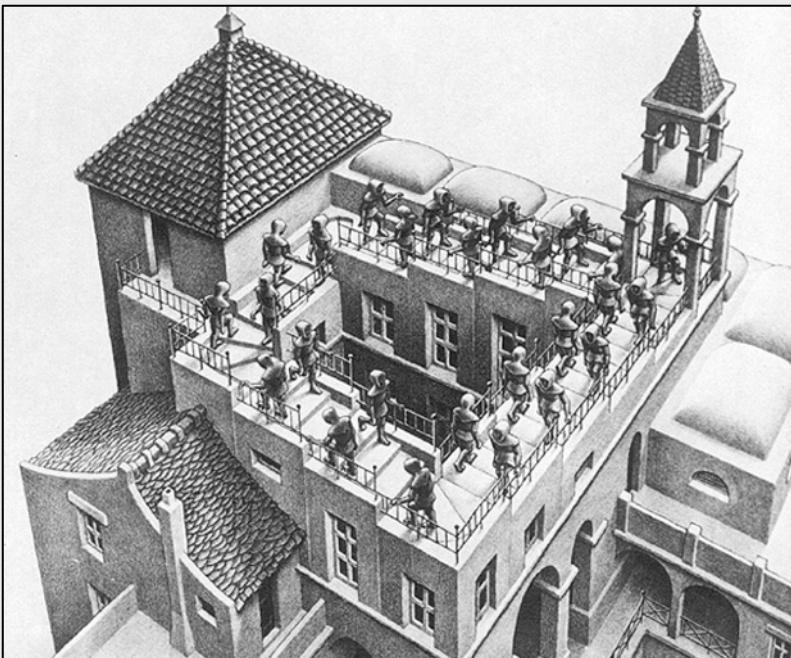
# 10 lessons from research on visual perception

1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

# Don't: 3D plots

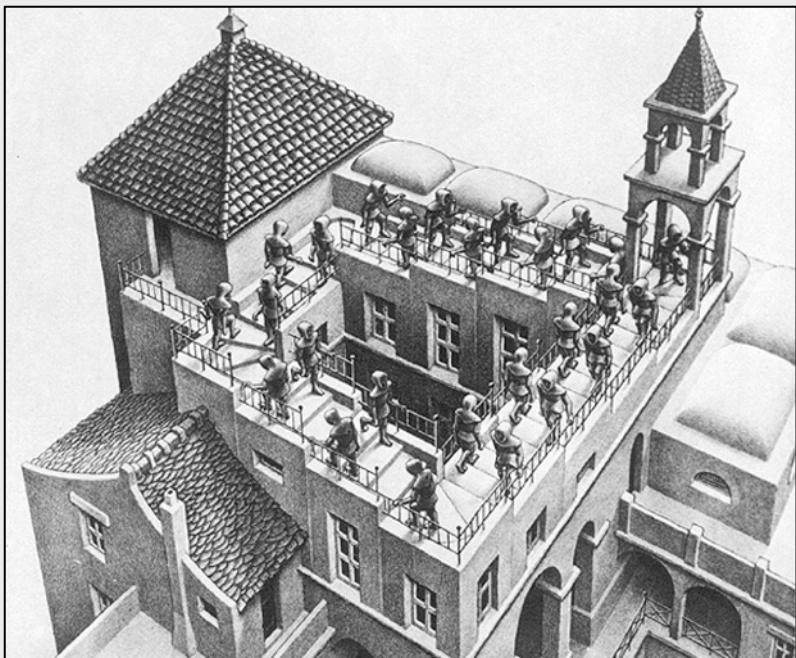
1) Humans are not that good at  
distinguishing 3D space



M.C. Escher (1898-1972)  
Upstairs and Downstairs  
(1960; lithography)

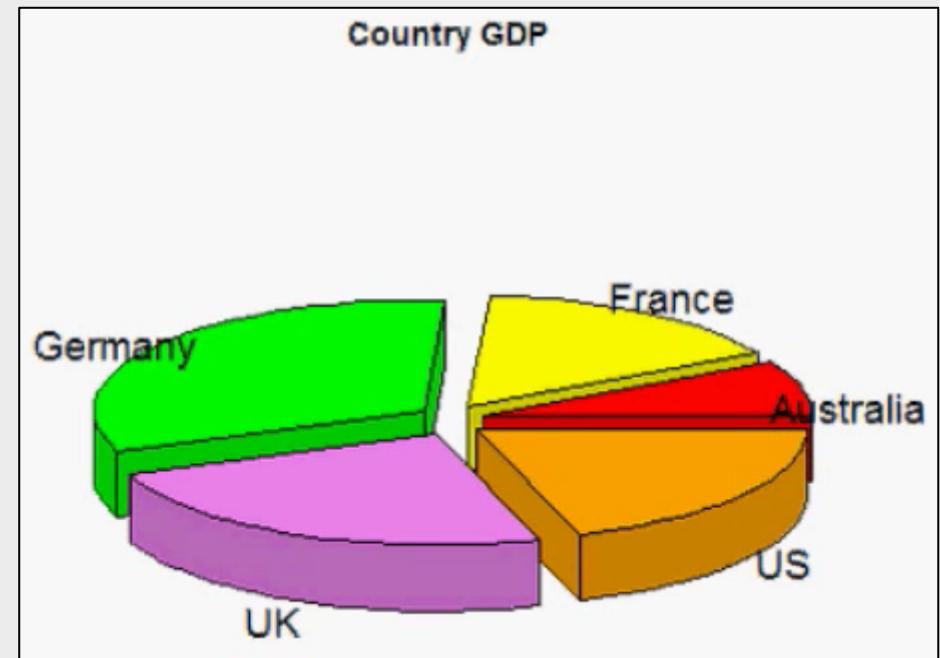
# Don't: 3D plots

1) Humans are not that good at  
distinguishing 3D space



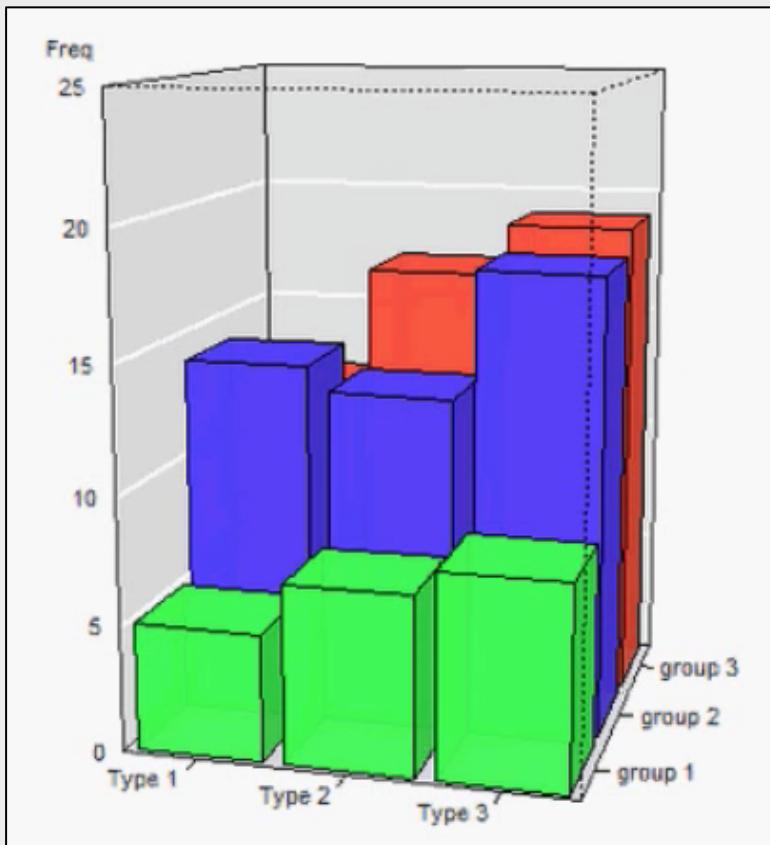
M.C. Escher (1898-1972)  
Upstairs and Downstairs  
(1960; lithography)

Ink proportions != true proportions

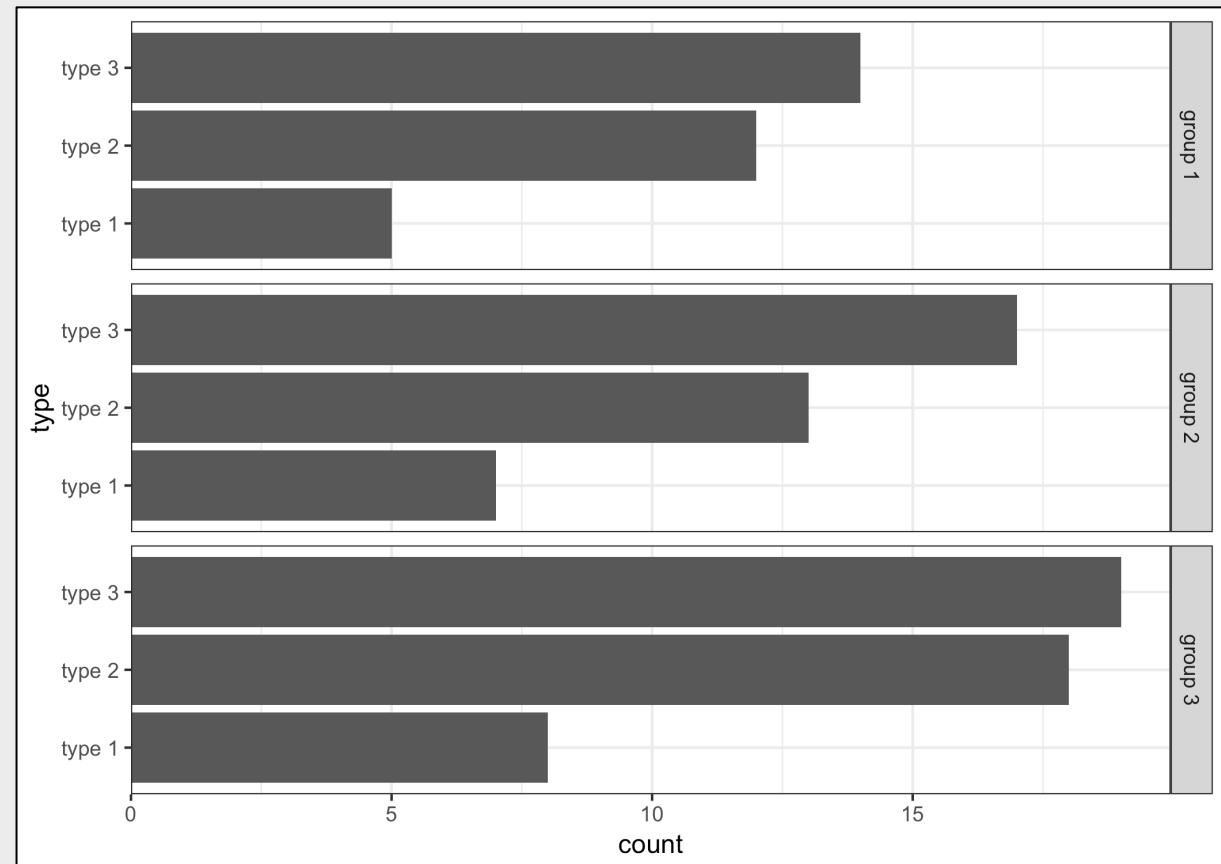
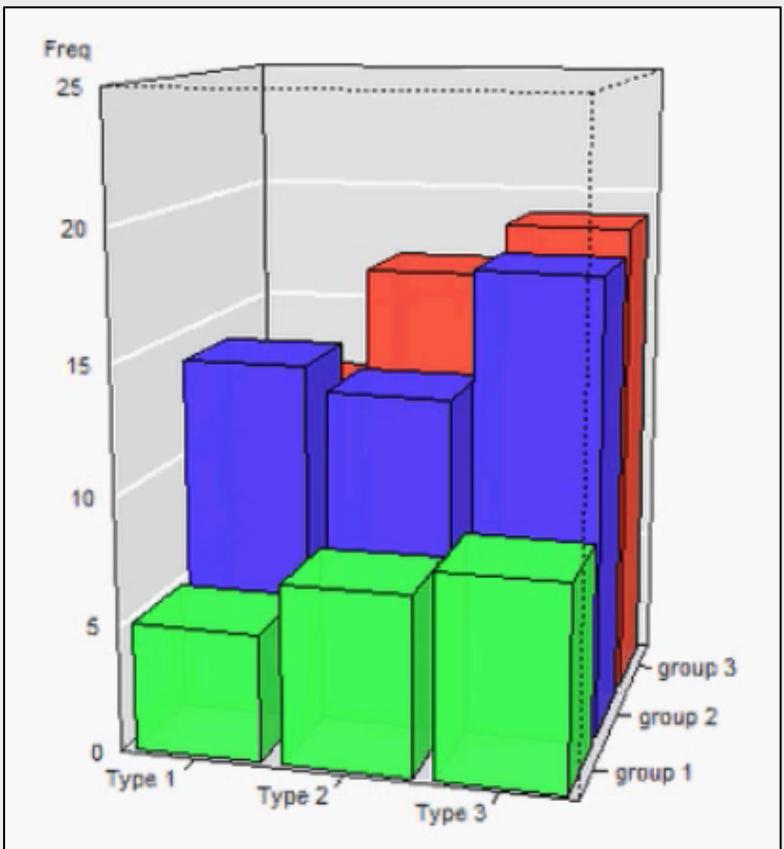


# Don't: 3D plots

2) Occlusion: geoms are obscured

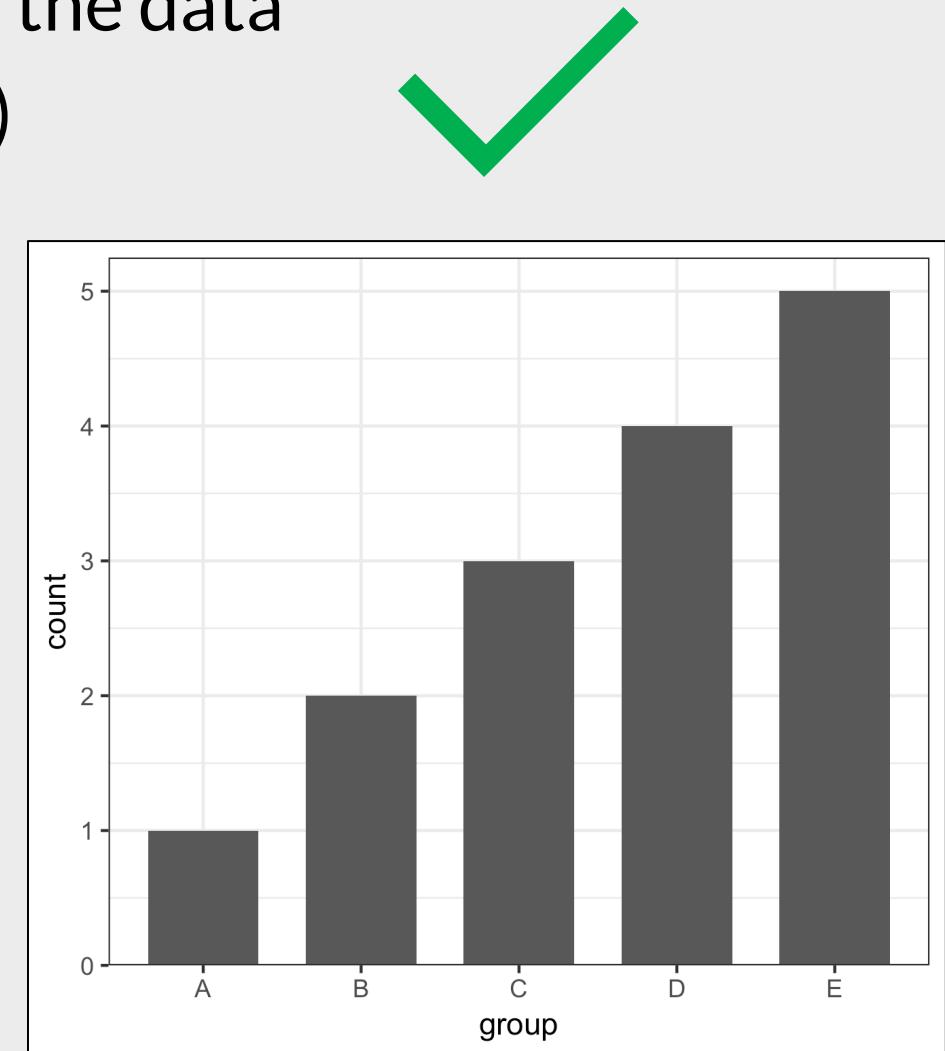
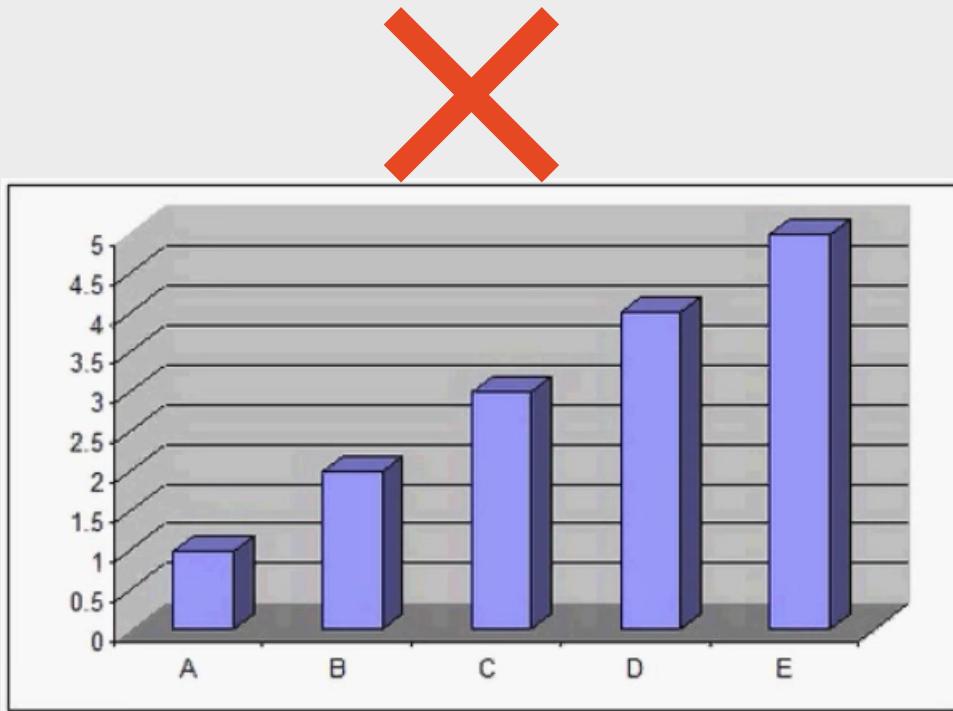


# A better alternative to 3D: faceting



# Don't: 3D plots

3) The third dimension distracts from the data  
(this is what Tufte calls “chart junk”)

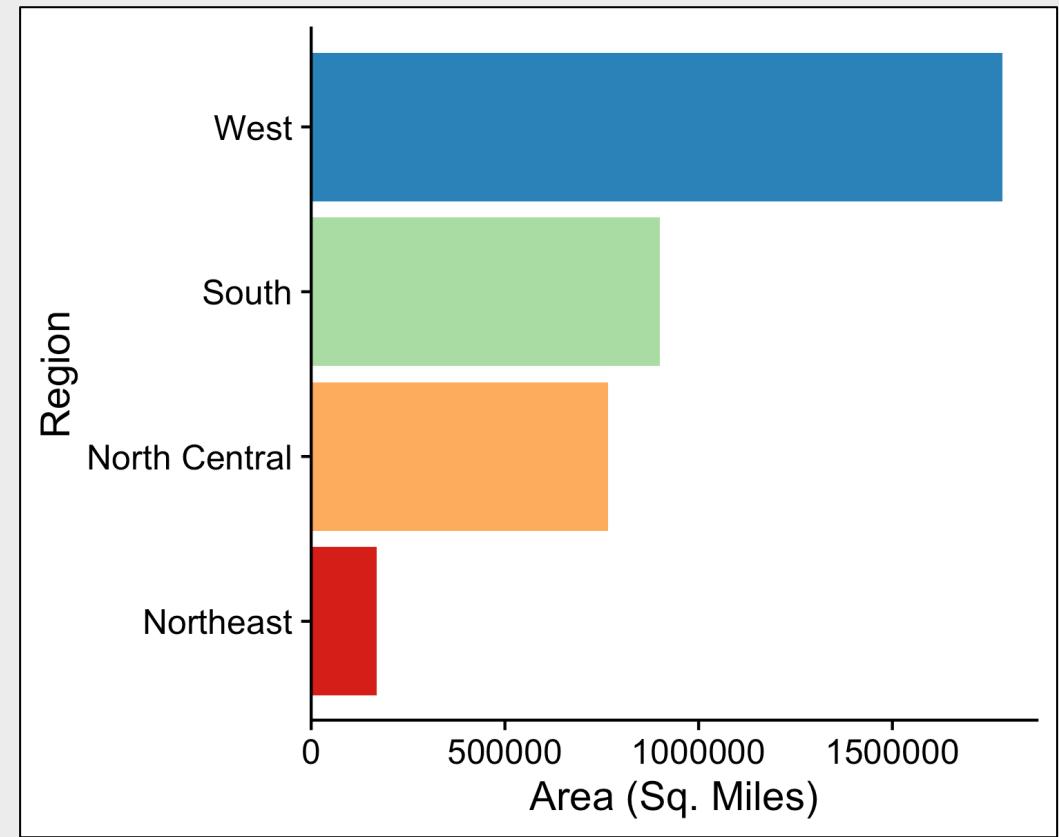
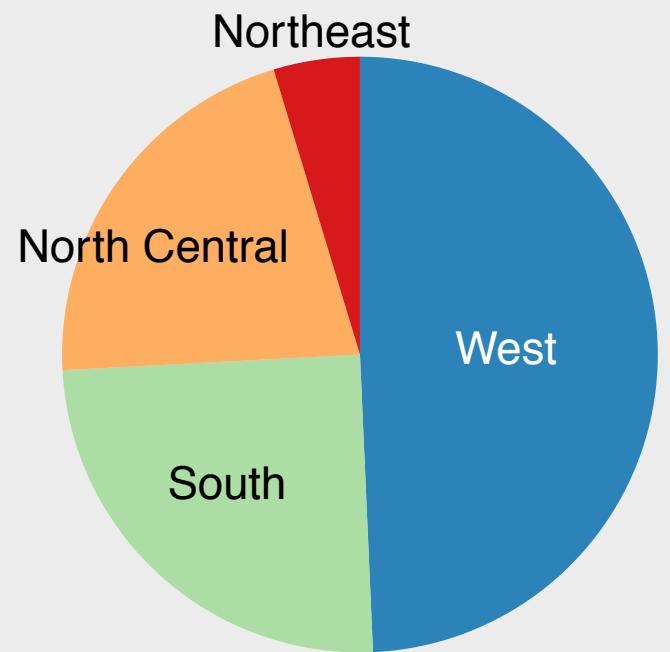


1. Don't make 3D plots
2. **Don't use pie charts for proportions**
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

*“Pie charts are the information visualization equivalent of a roofing hammer to the frontal lobe”*

- Coda Hale





**Sky**



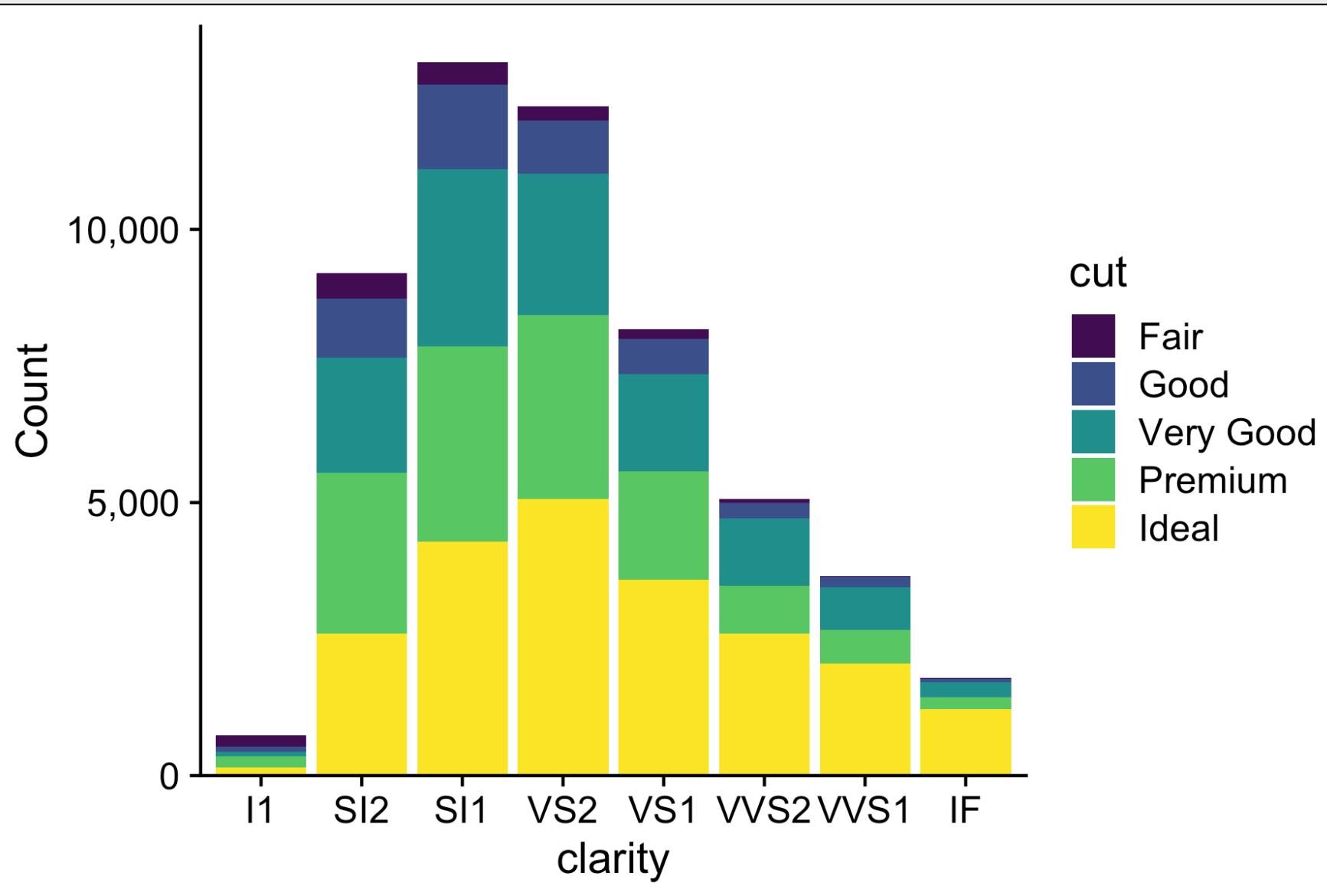
**Sunny side of pyramid**

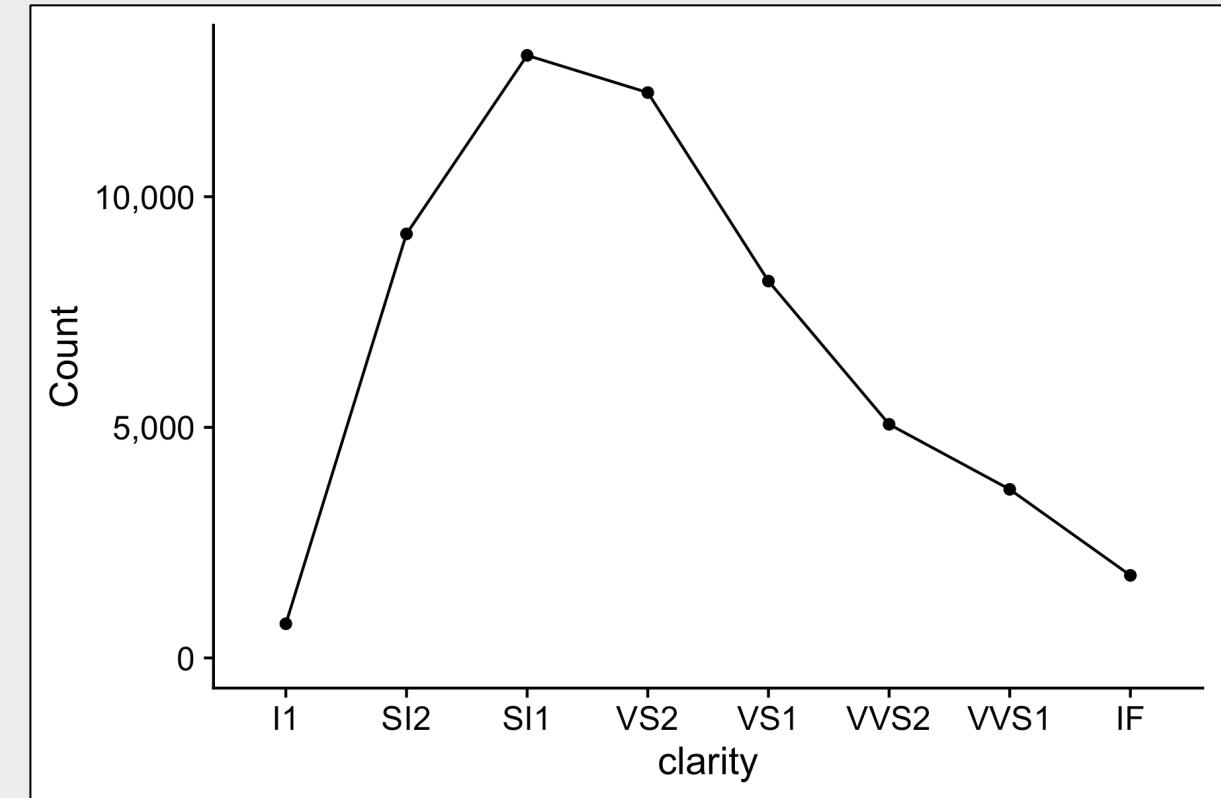
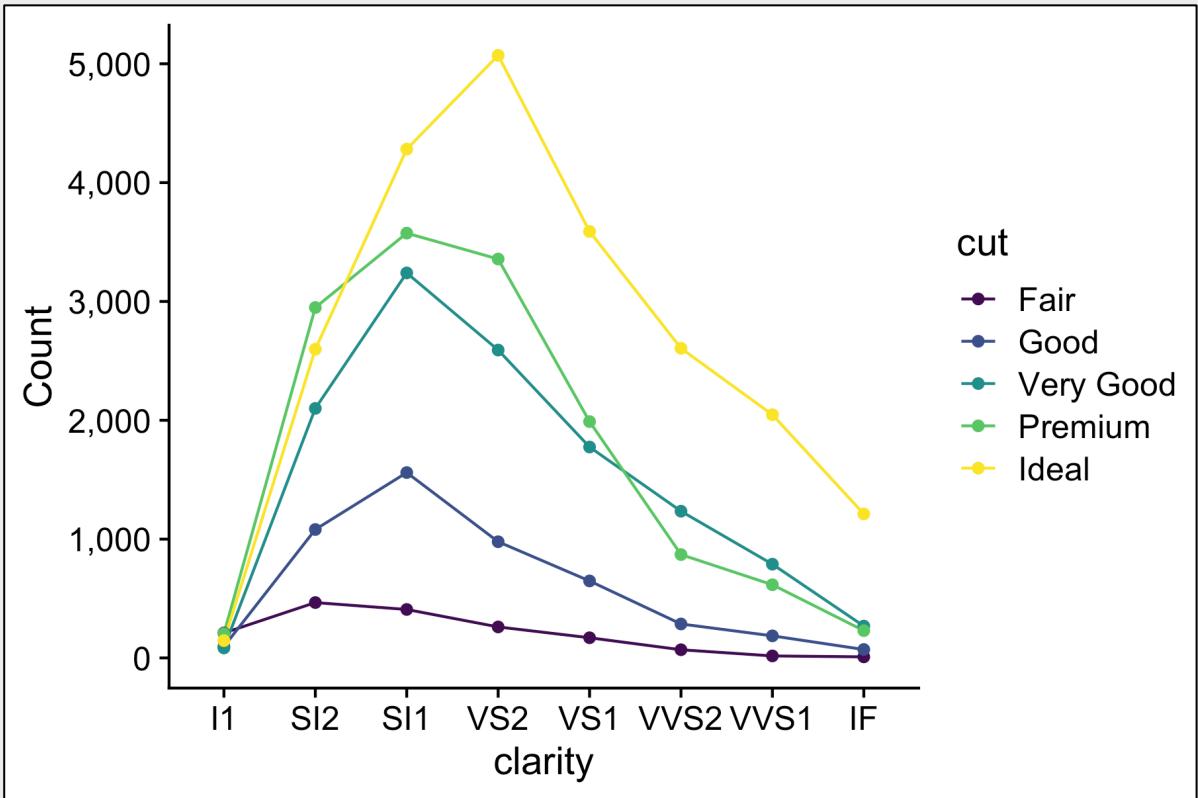


**Shady side of pyramid**

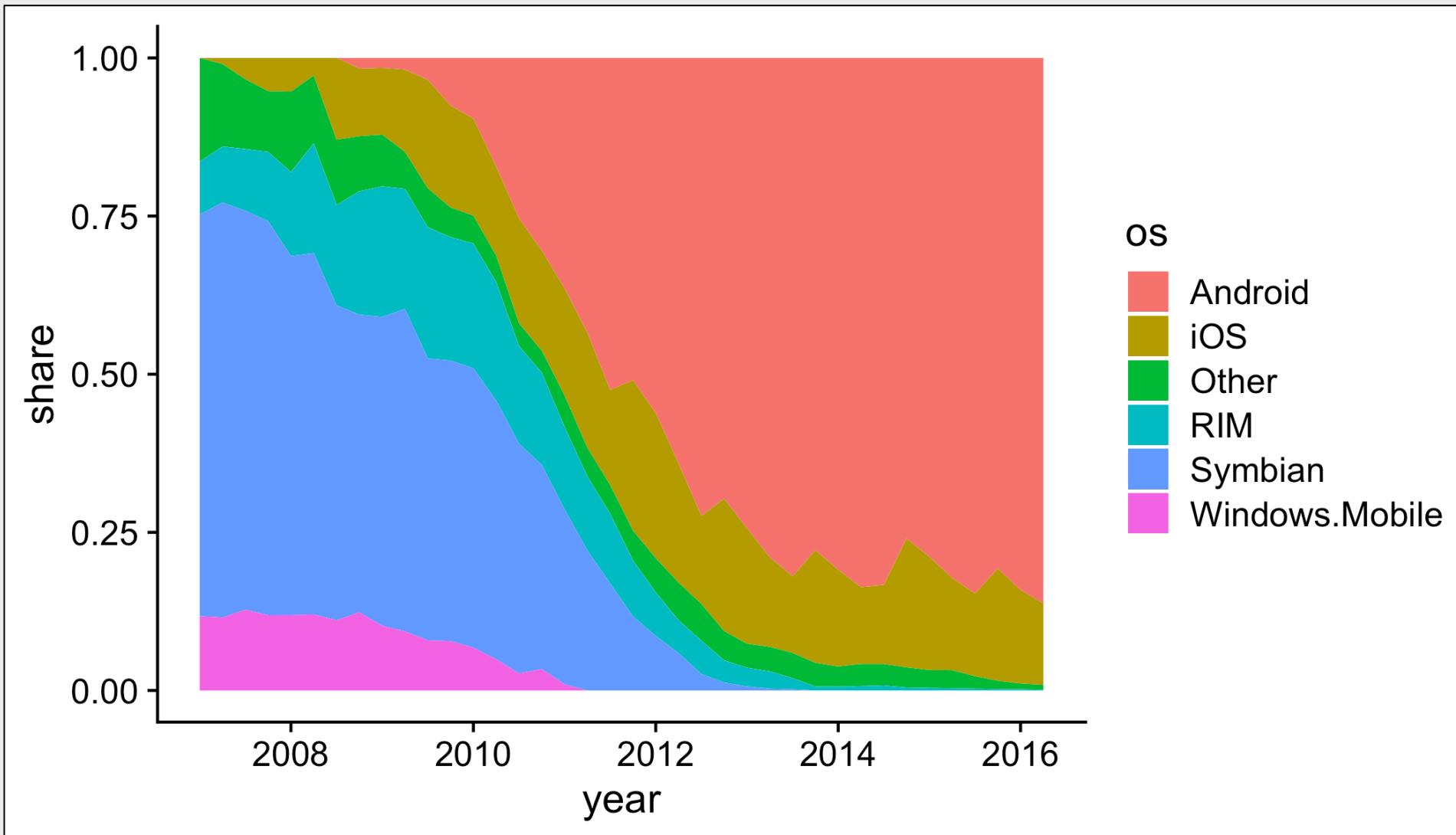
1. Don't make 3D plots
2. Don't use pie charts for proportions
3. **Don't stack bars**
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time





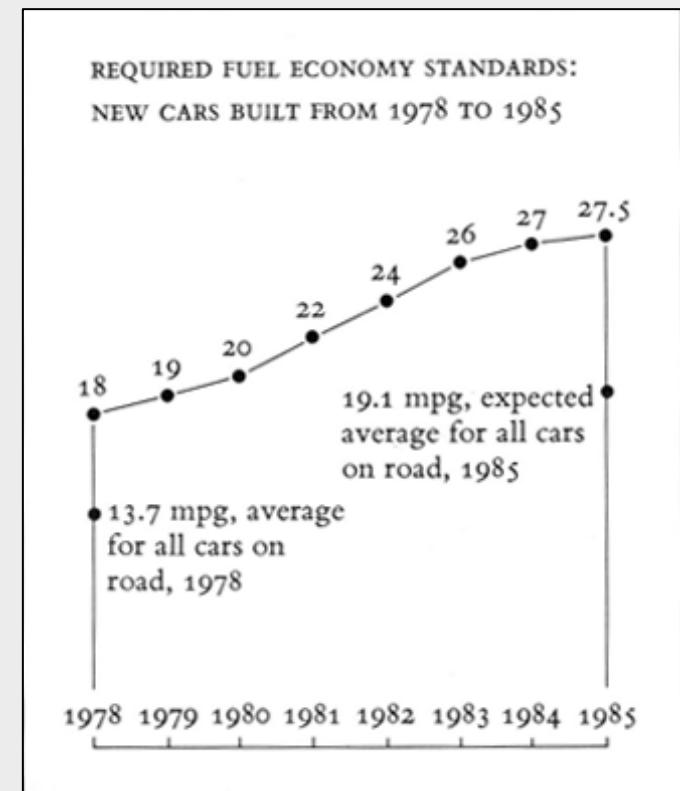
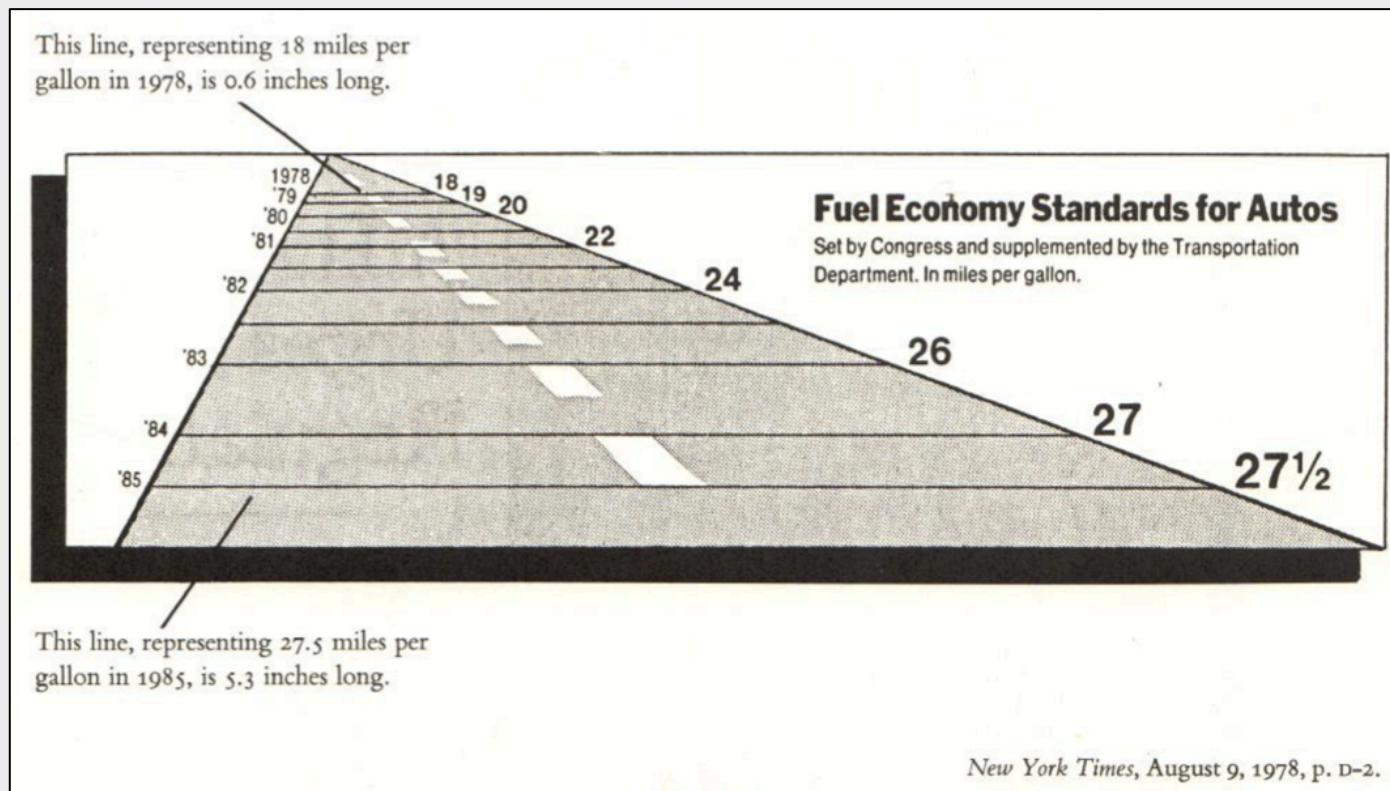
# Exception



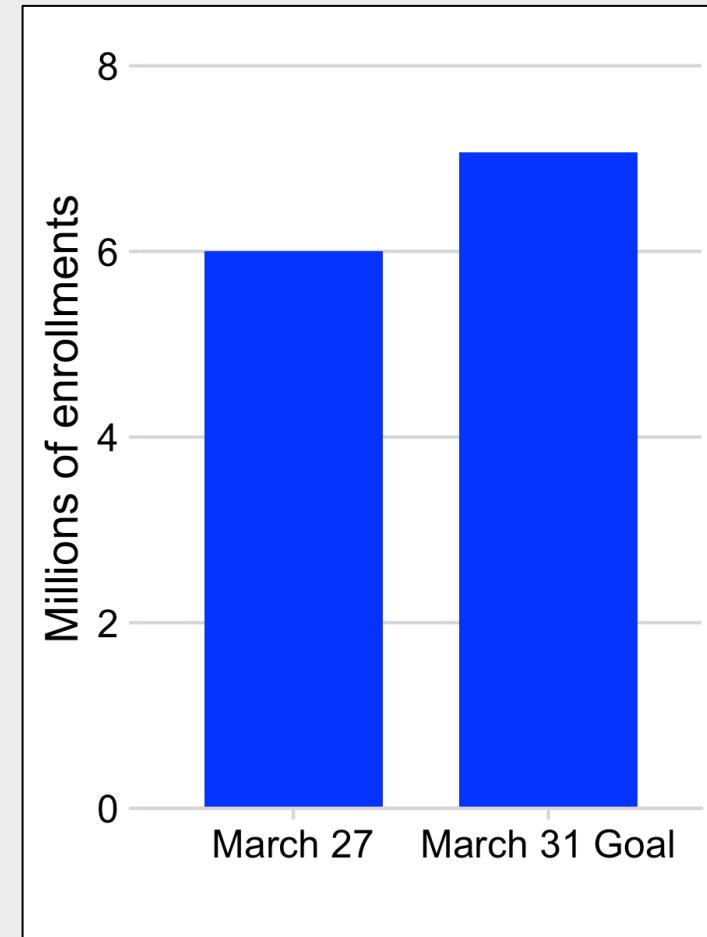
1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. **Don't lie**
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

$$\text{“Lie Factor”} = \frac{\text{Size of effect shown in graphic}}{\text{Size of effect in data}} = \frac{\frac{5.3 - 0.6}{0.6}}{\frac{27.5 - 18}{18}} = \frac{7.83}{0.53} = 14.8$$

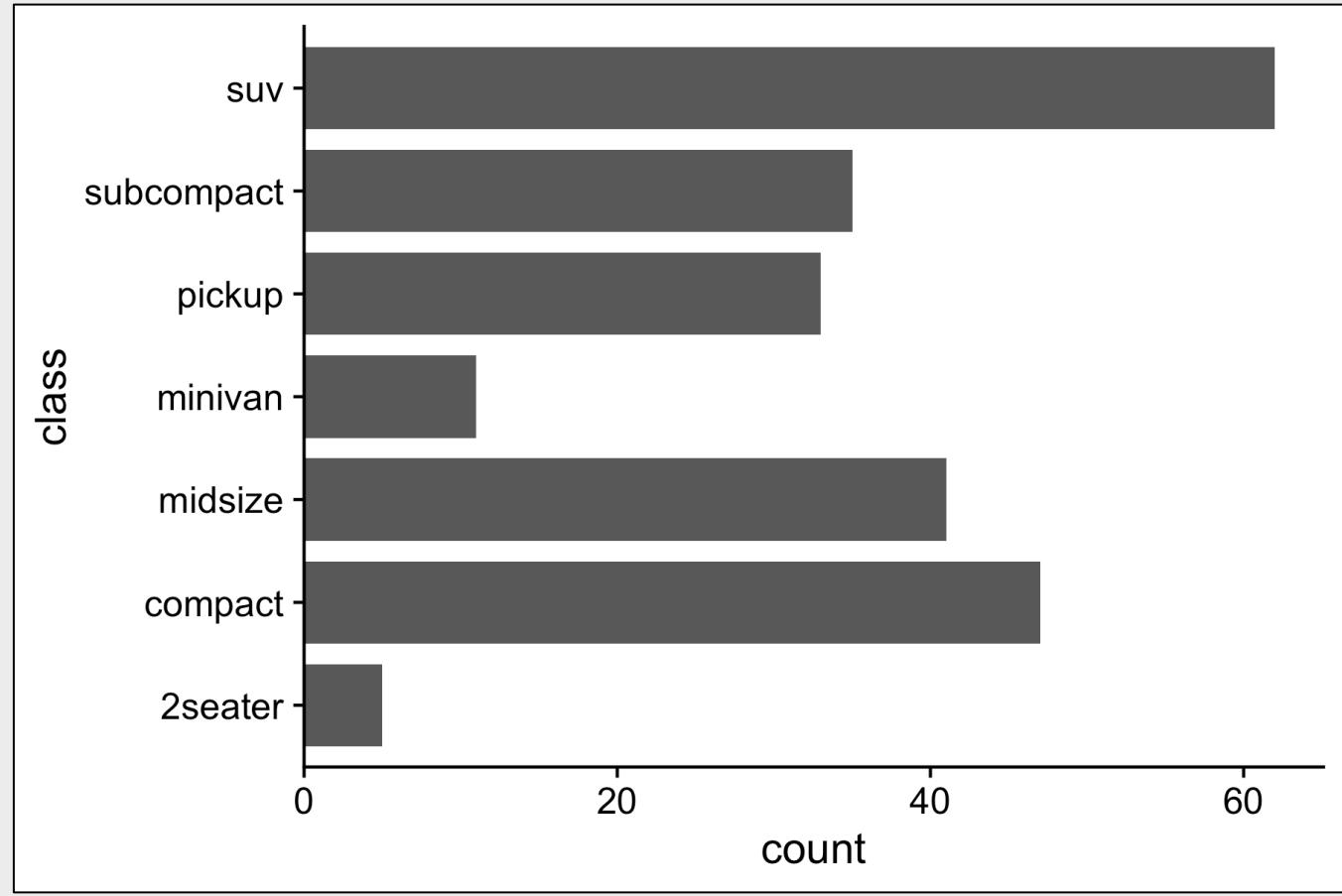
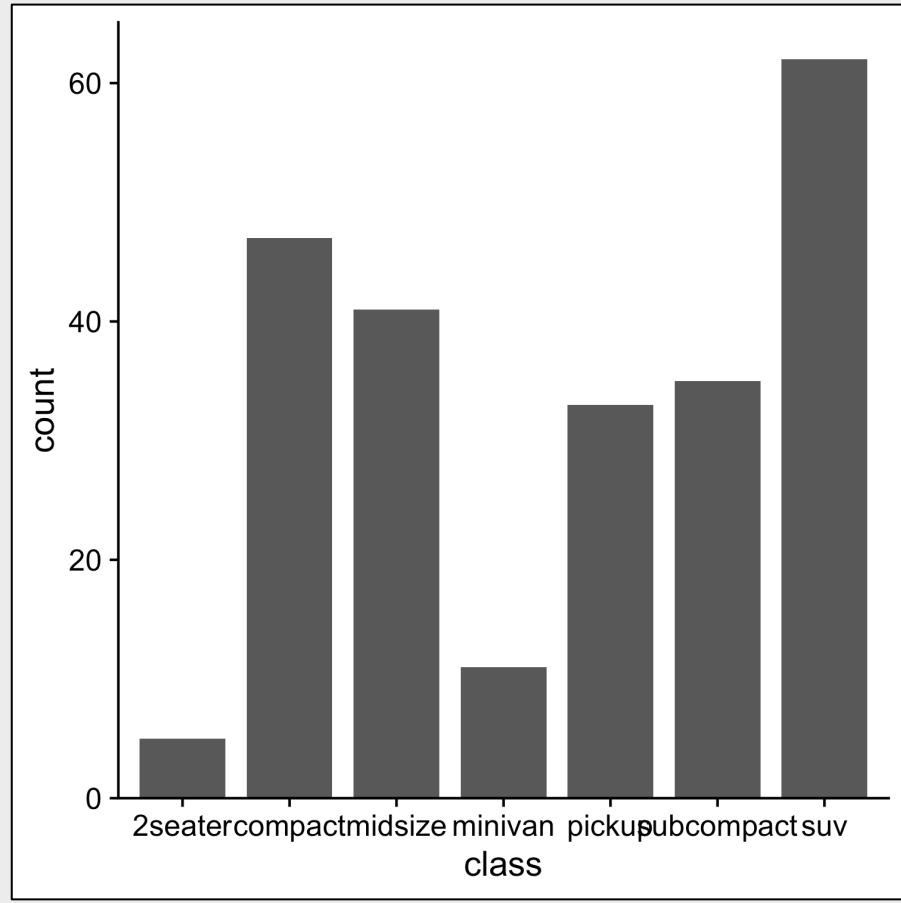


# Do: Start bar charts at 0



1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. **Do rotate and sort categorical axes**
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

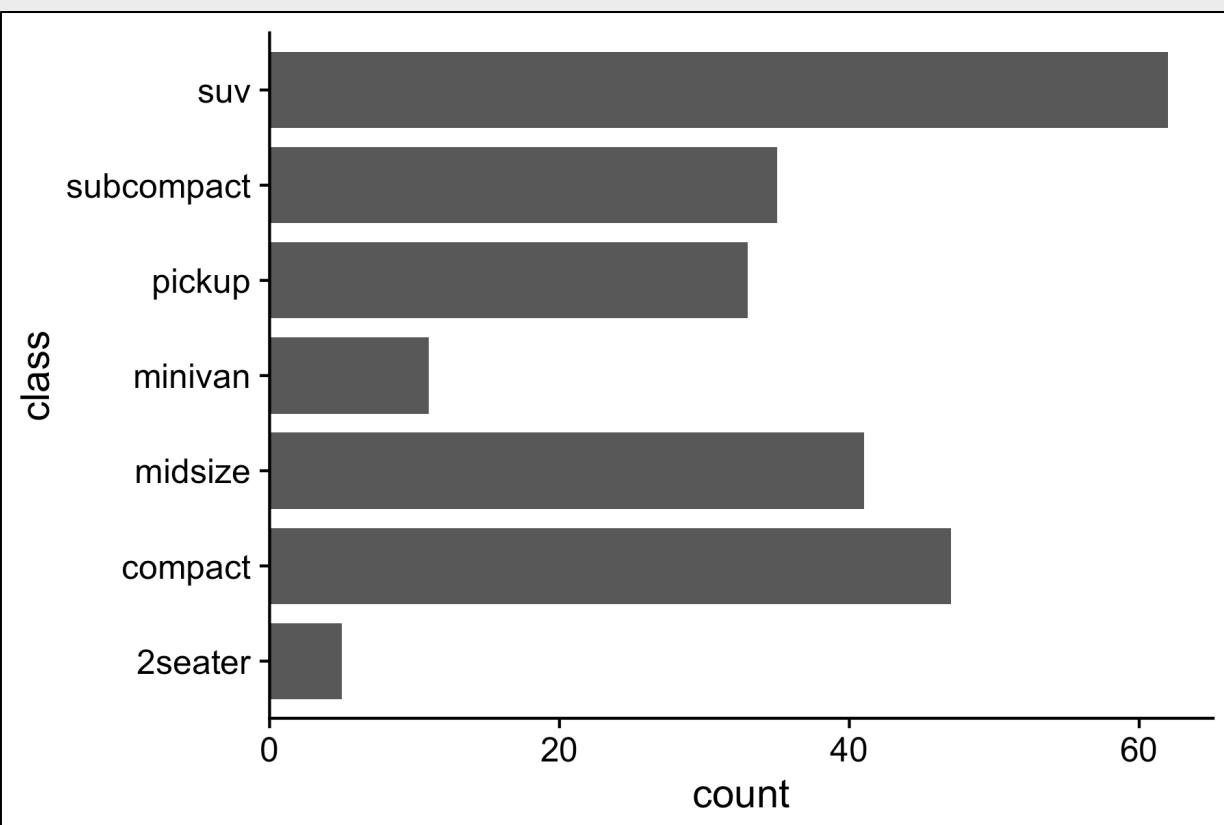
\* ...most of the time



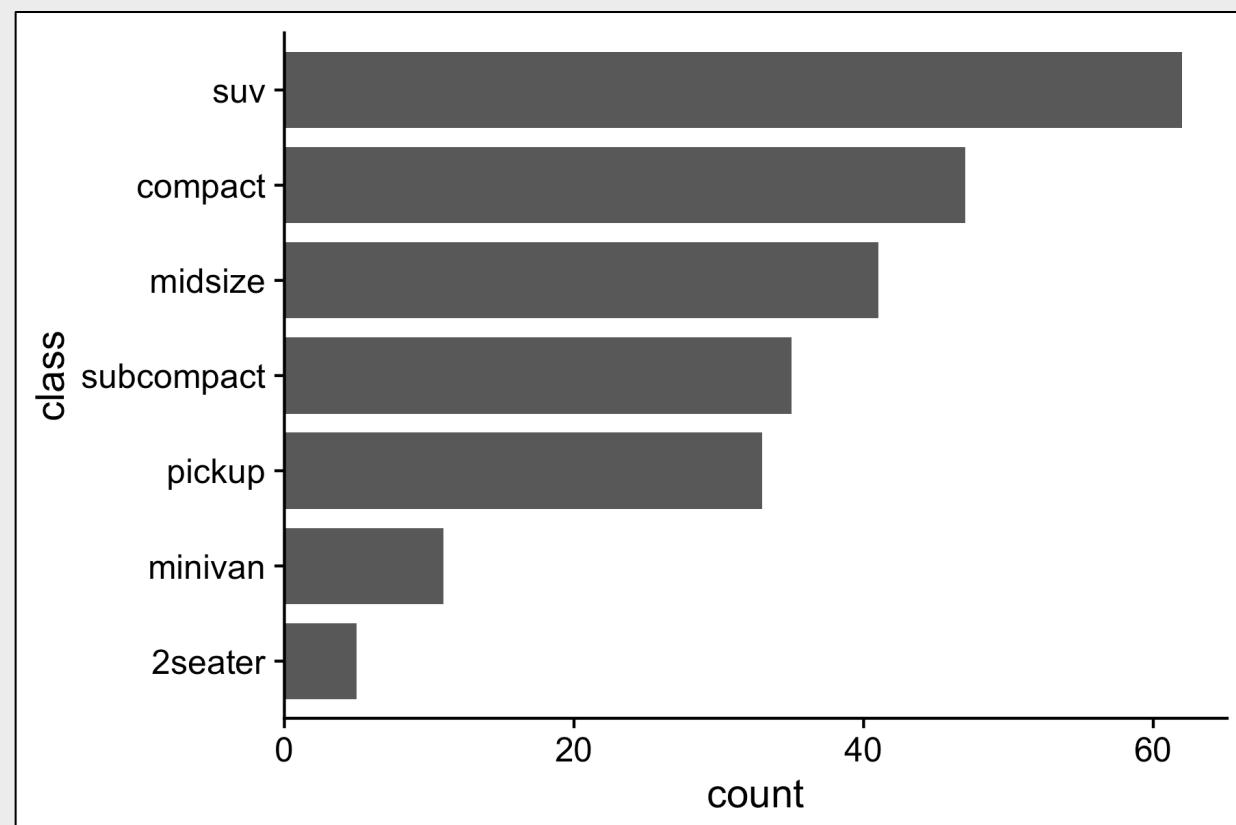
# Default ordering is almost always wrong



Ordered by alphabet



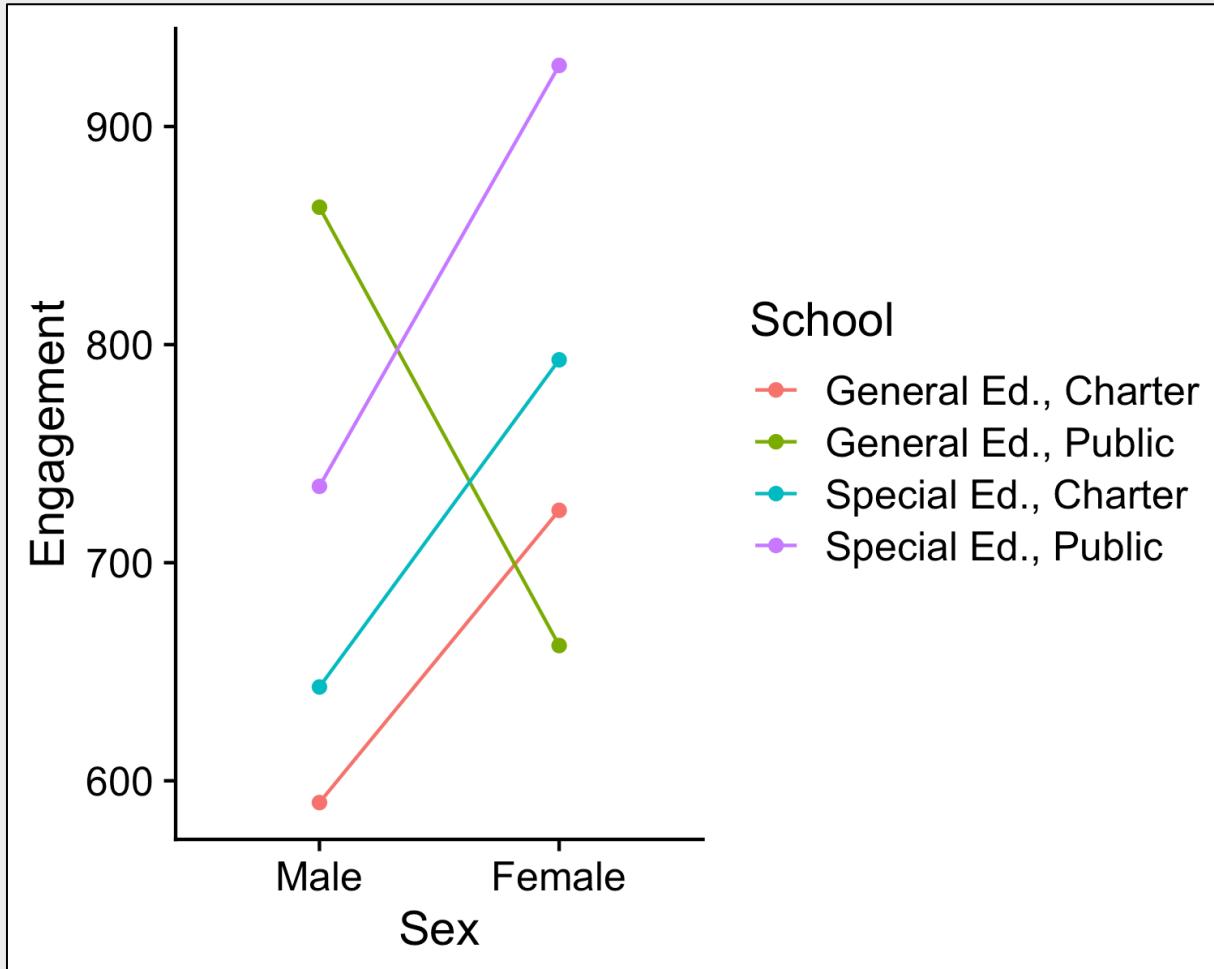
Ordered by count



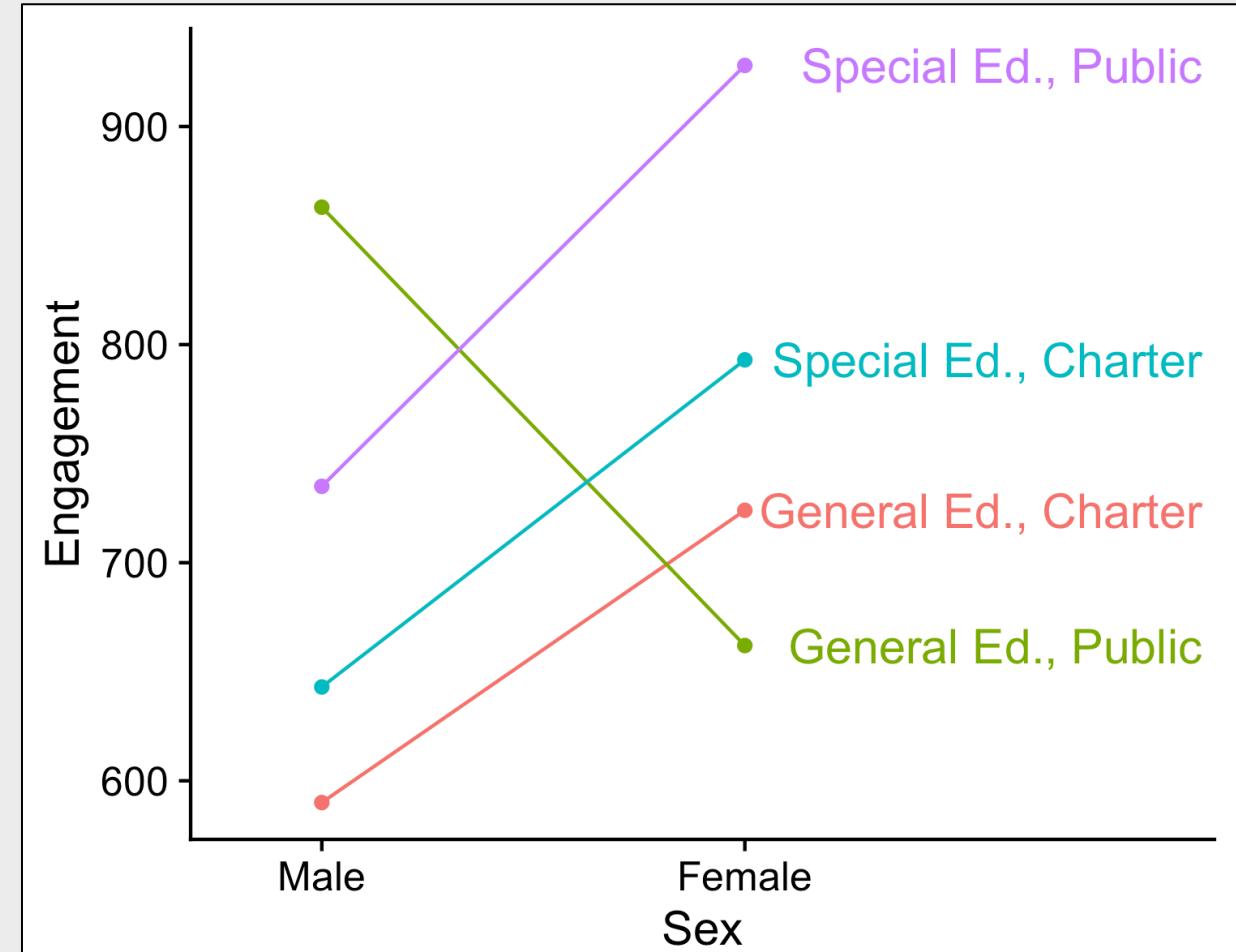
1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. **Do eliminate legends & directly label geoms**
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

**X**



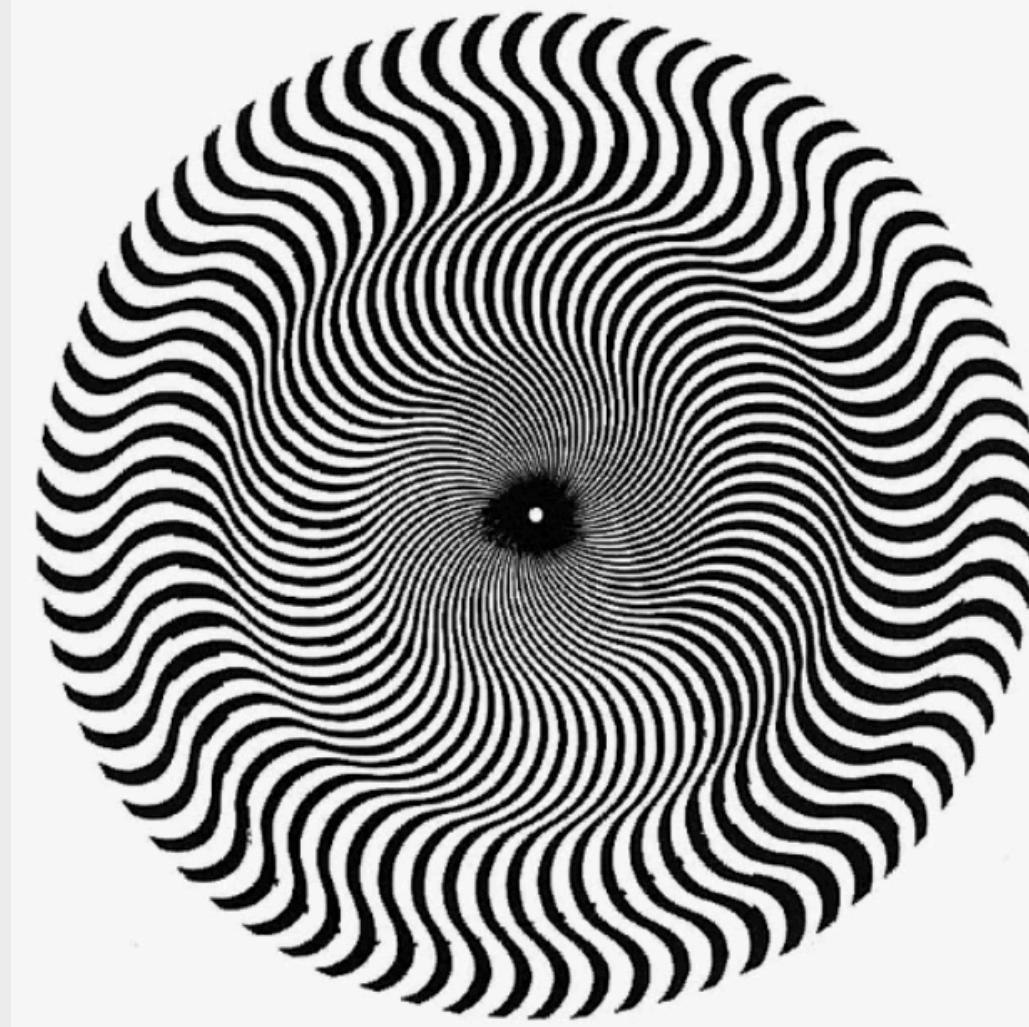
**✓**

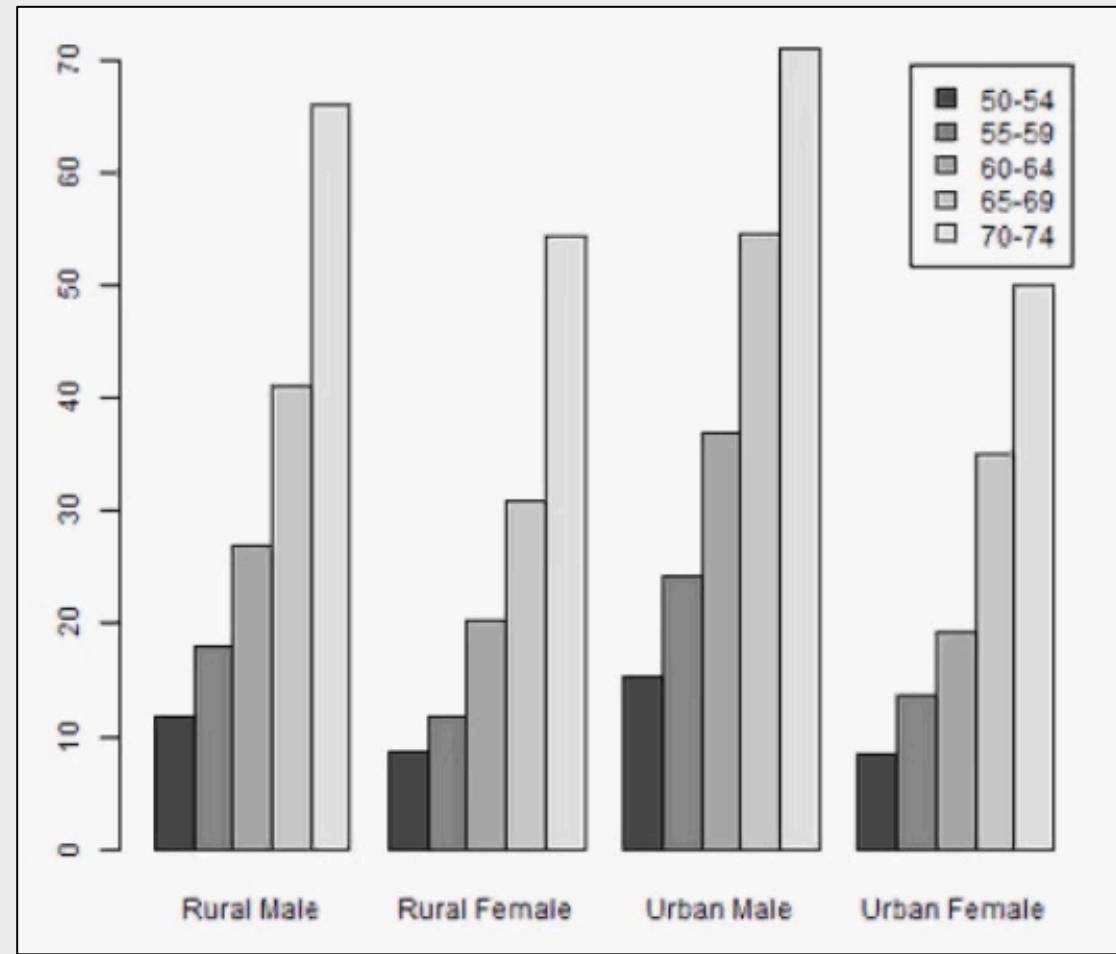
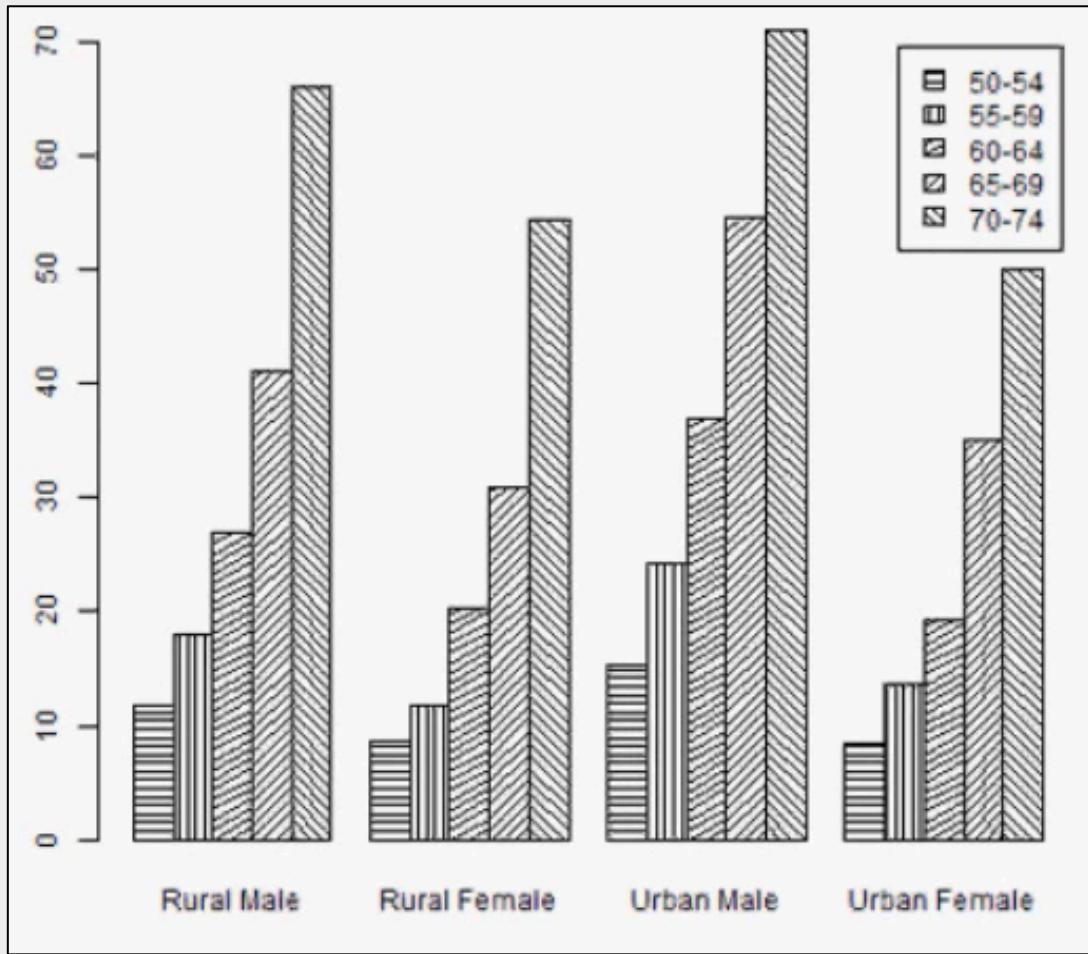


1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. **Don't use pattern fills**
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

# Don't use pattern fills

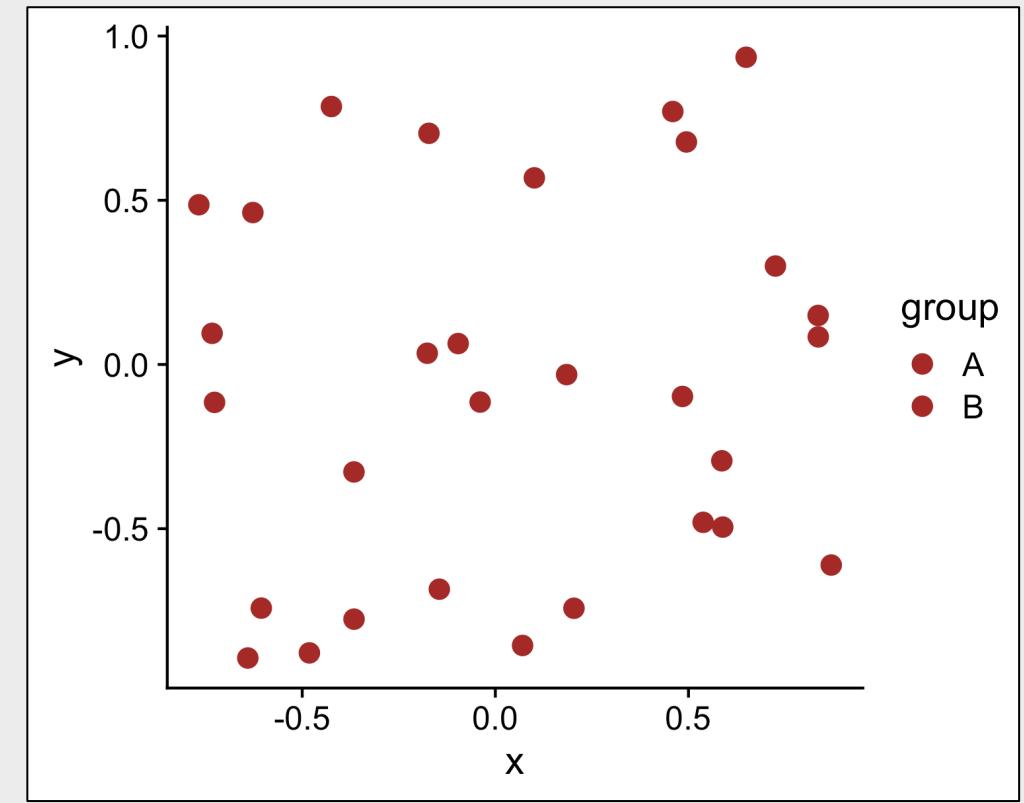
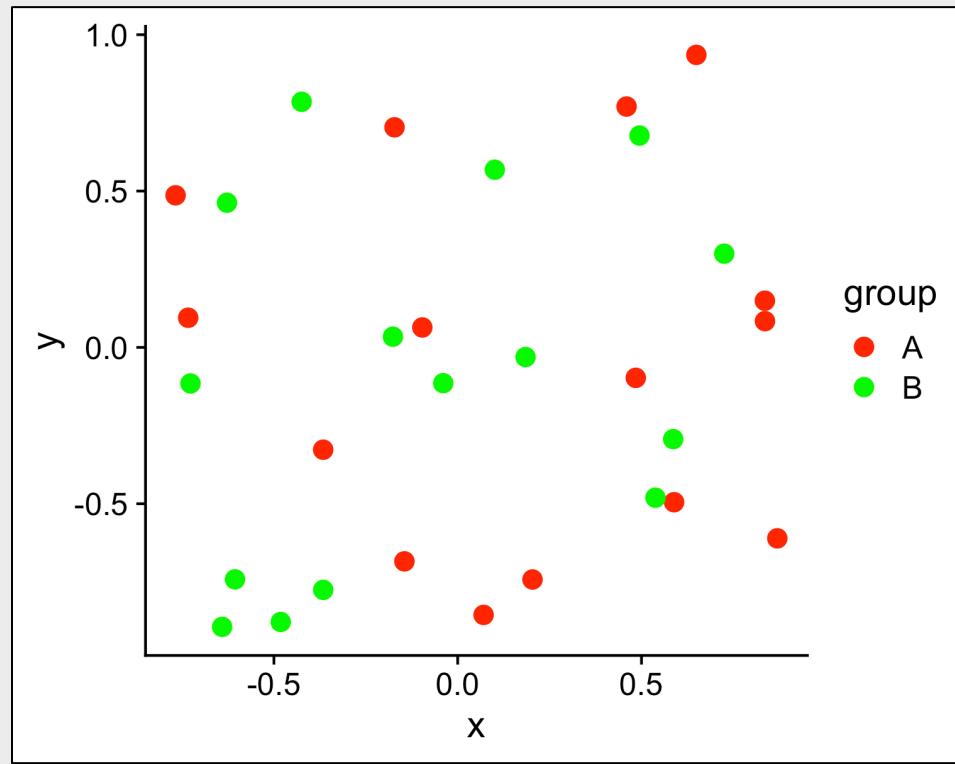




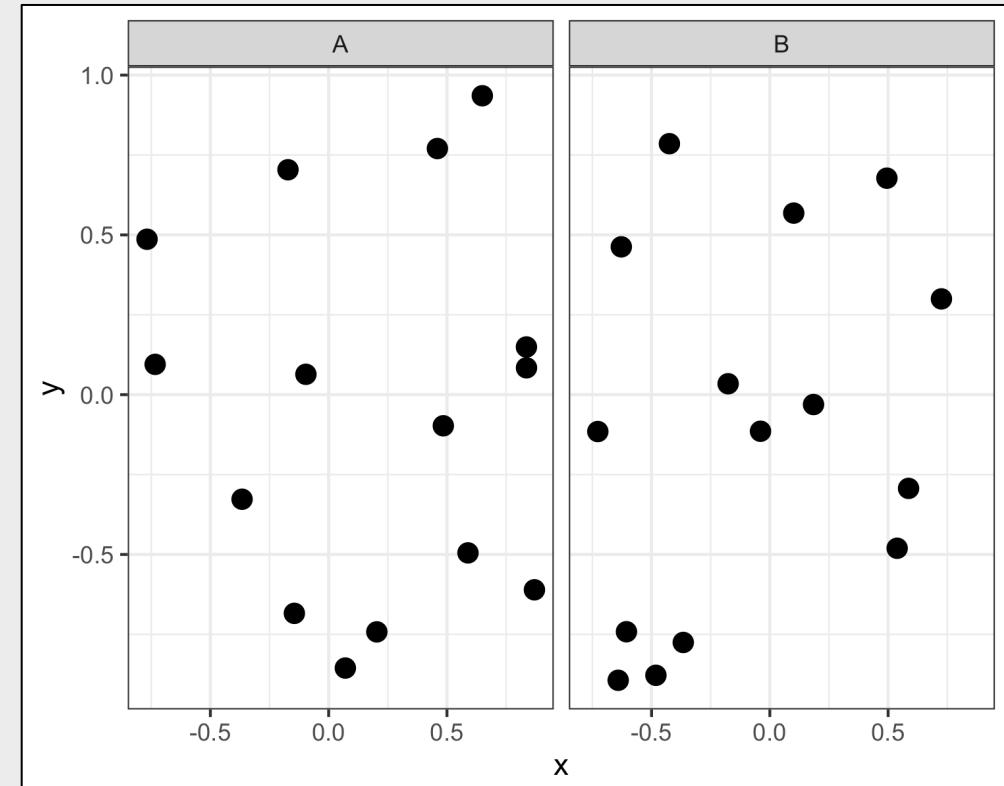
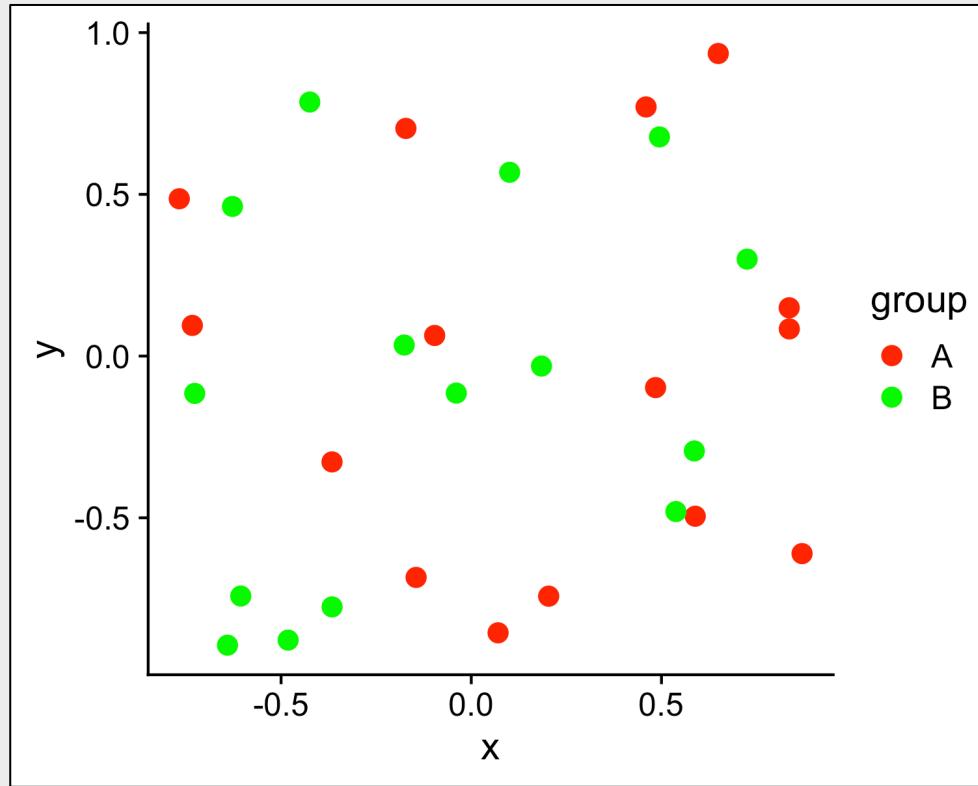
1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. **Don't use red & green together**
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

10% of males and 1% of females are color blind



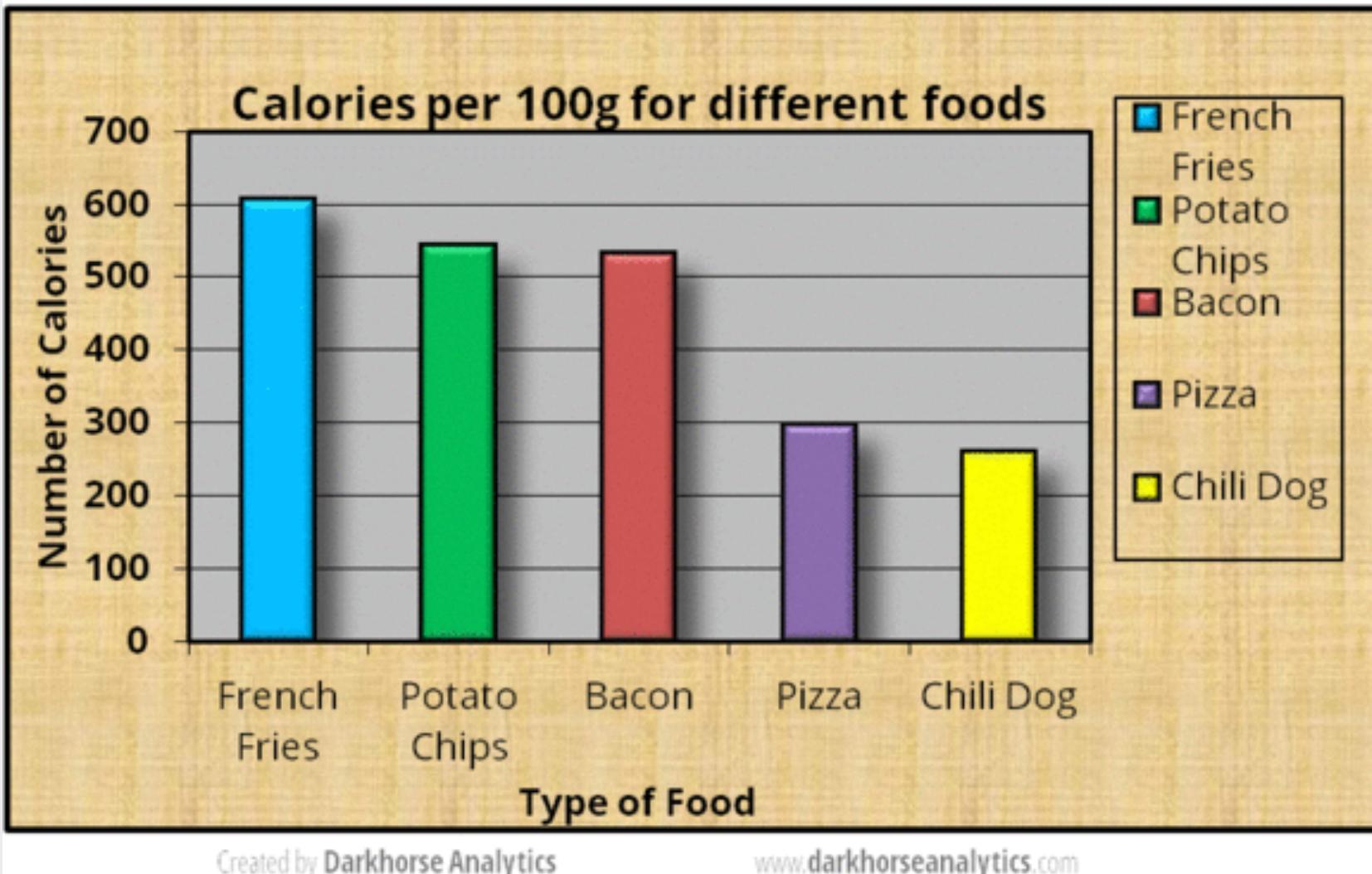
# Facets can be used to avoid color altogether



1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. **Do remove chart chunk**
10. Do consider tables for small data sets

\* ...most of the time

# Before



# Remove to improve (the **data-ink** ratio)

Created by Darkhorse Analytics

[www.darkhorseanalytics.com](http://www.darkhorseanalytics.com)

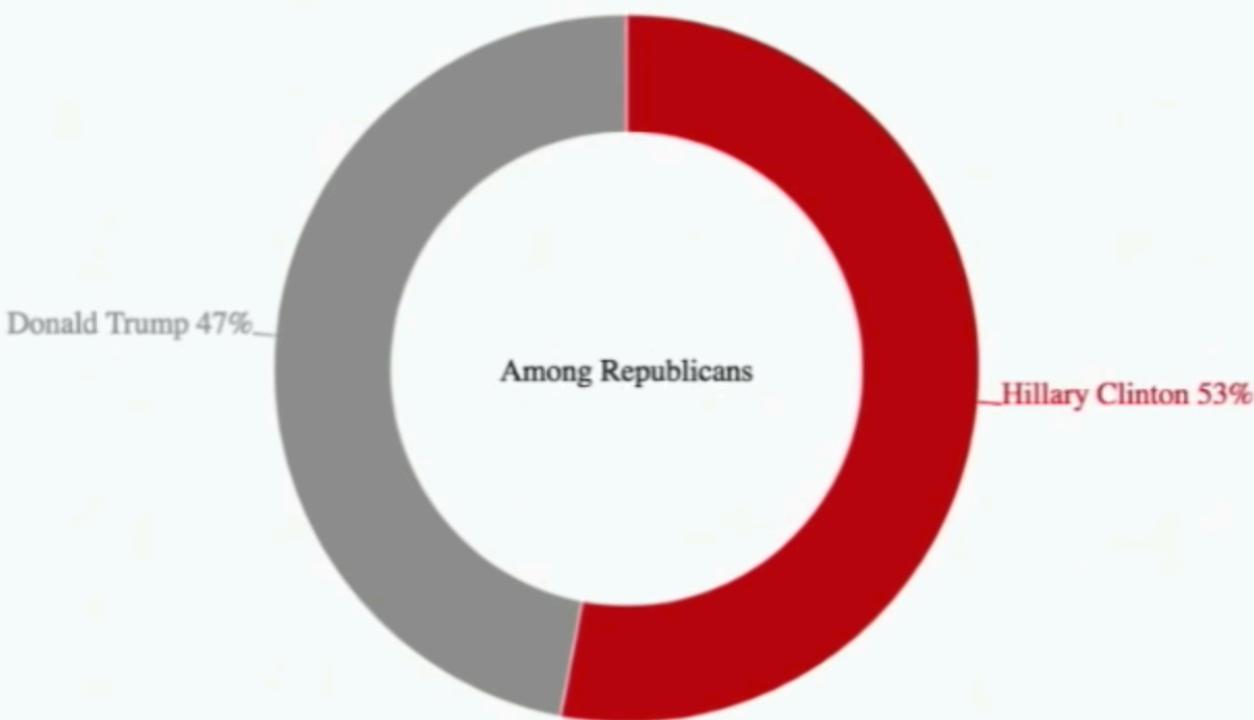
1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\* ...most of the time

## Who do you think did a better job in tonight's debate?

Among Republicans

Among Democrats



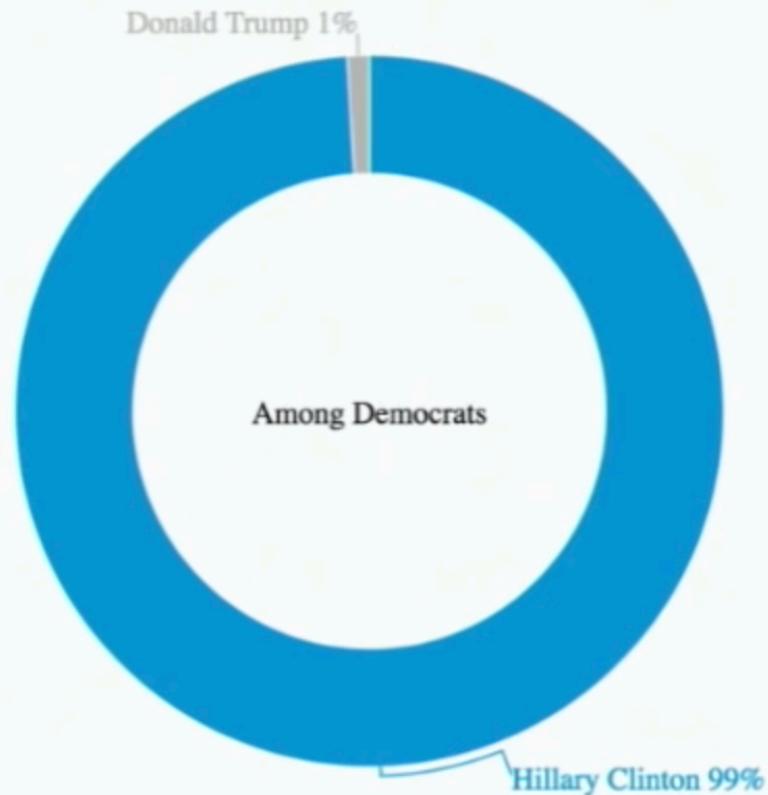
Share

POLITICO

## Who do you think did a better job in tonight's debate?

Among Republicans

Among Democrats



Share

POLITICO

Who do you think did a better job in tonight's debate?

---

	Clinton	Trump
Among democrats	99%	1%
Among republicans	53%	47%

---

Two objectives of effective graphs:

1. Grab & direct attention (iconic memory)
2. Reduce processing demands (working memory)

Graph components:

1. Geoms:
  - points, lines, boxes, bars, etc.
2. Pre-attentive attributes:
  - position, color, shape, curvature, etc.
3. Non-data ink:
  - scales, grid lines, legend, labels, etc.
4. No chart junk!

Pattern recognition hierarchy:

- Position on a common scale
- Position on non-aligned scales
- Length
- Angle
- Area
- Color saturation
- Color hue

Cleveland's three visual operations of pattern perception:

1. Estimation:
  - Discrimination  $X \neq Y$
  - Ranking  $X > Y$
  - Ratioing  $X / Y$
2. Assembly:
  - The grouping of graphical elements
  - Prägnanz: We strongly prefer to interpret stimuli as regular, simple, and orderly
3. Detection:
  - Recognizing that a geometric object encodes a physical value
  - Above all else, show the data

10 lessons\* from research on visual perception:

1. Don't make 3D plots
2. Don't use pie charts for proportions
3. Don't stack bars
4. Don't lie
5. Do rotate and sort categorical axes
6. Do eliminate legends & directly label geoms
7. Don't use pattern fills
8. Don't use red & green together
9. Do remove chart chunk
10. Do consider tables for small data sets

\*most of the time

# Your turn

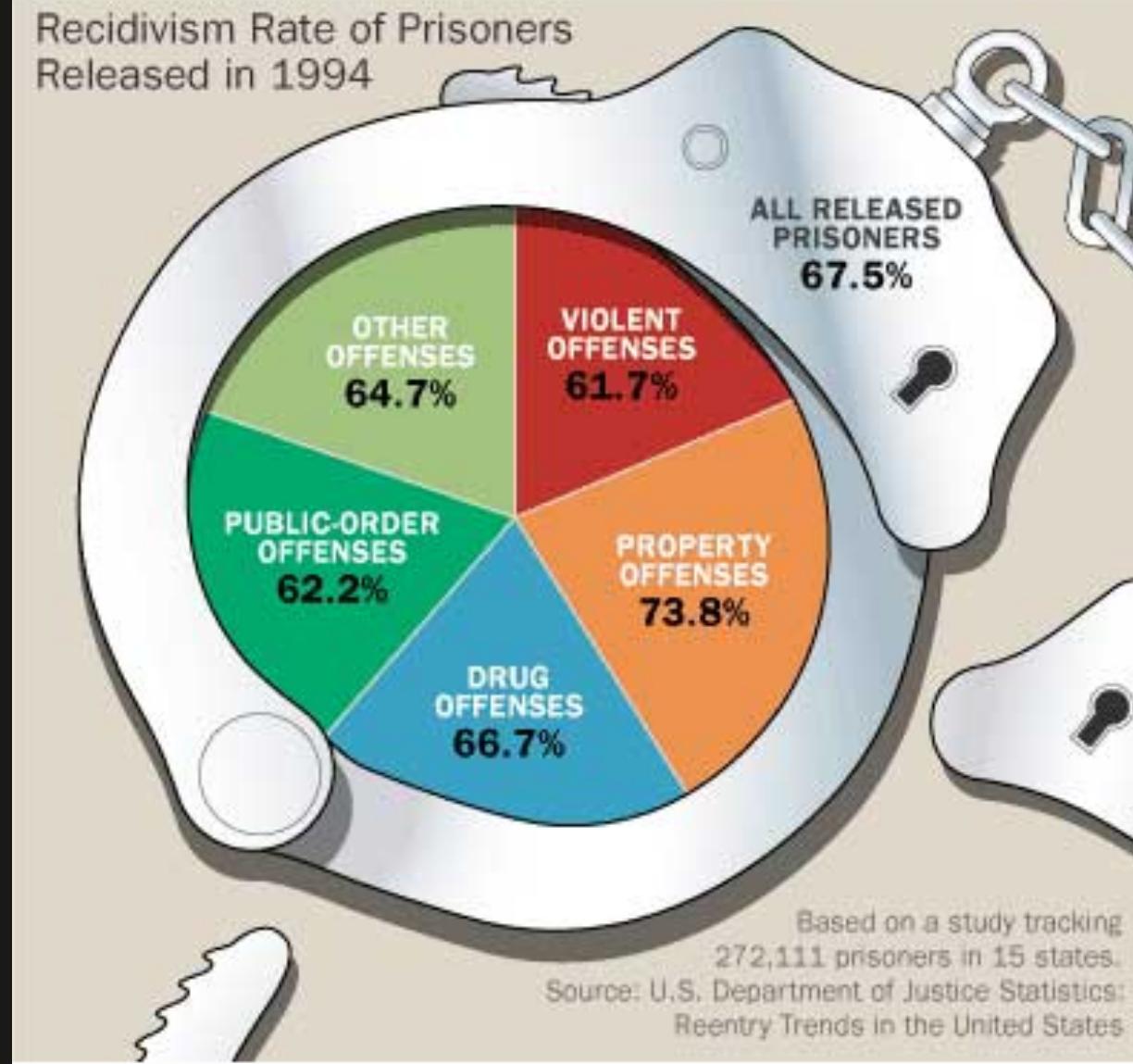
1) Identify the following:

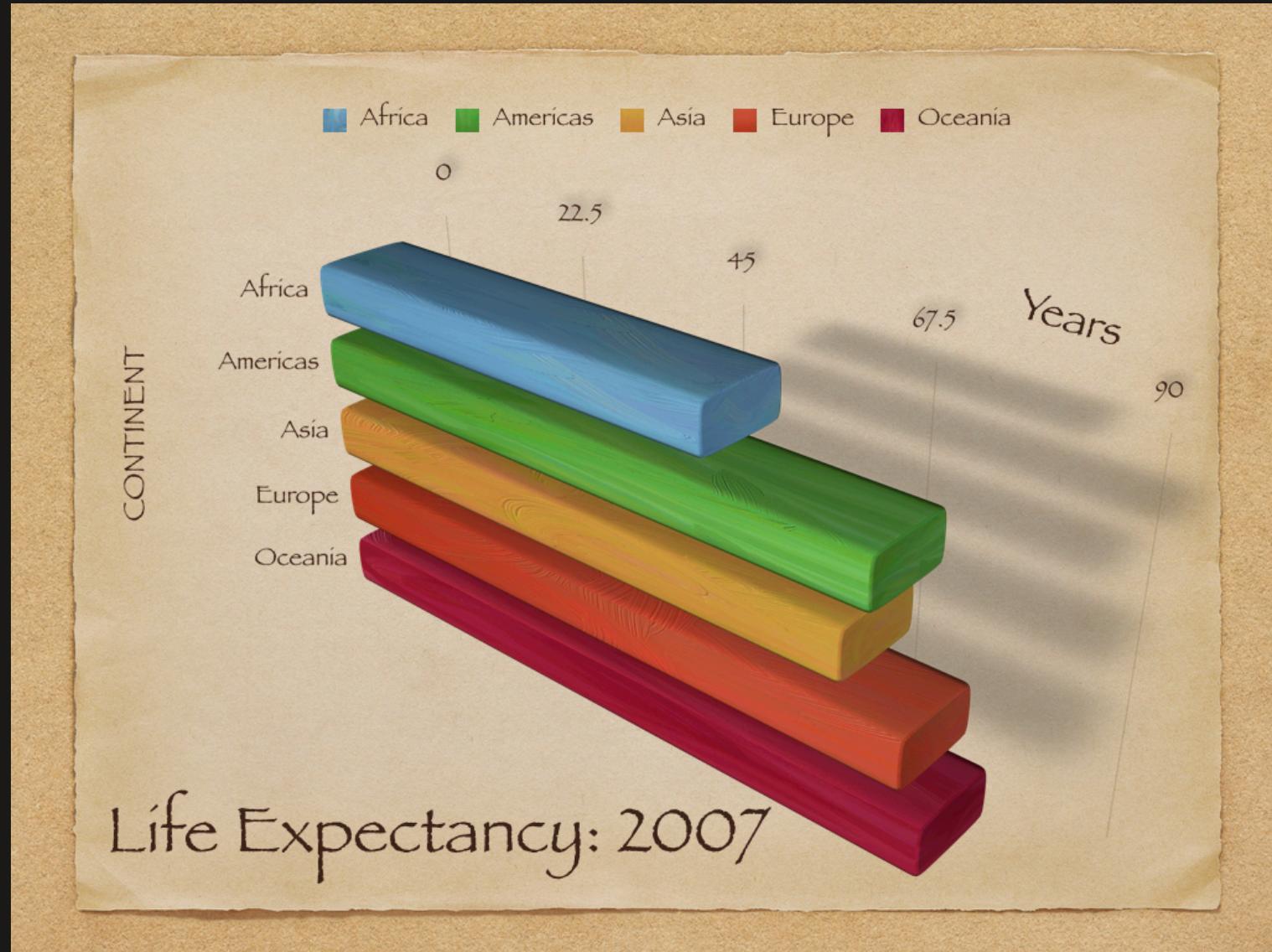
- The geoms
- The pre-attentive attributes
- Where the graphic falls on Cleveland's pattern recognition hierarchy

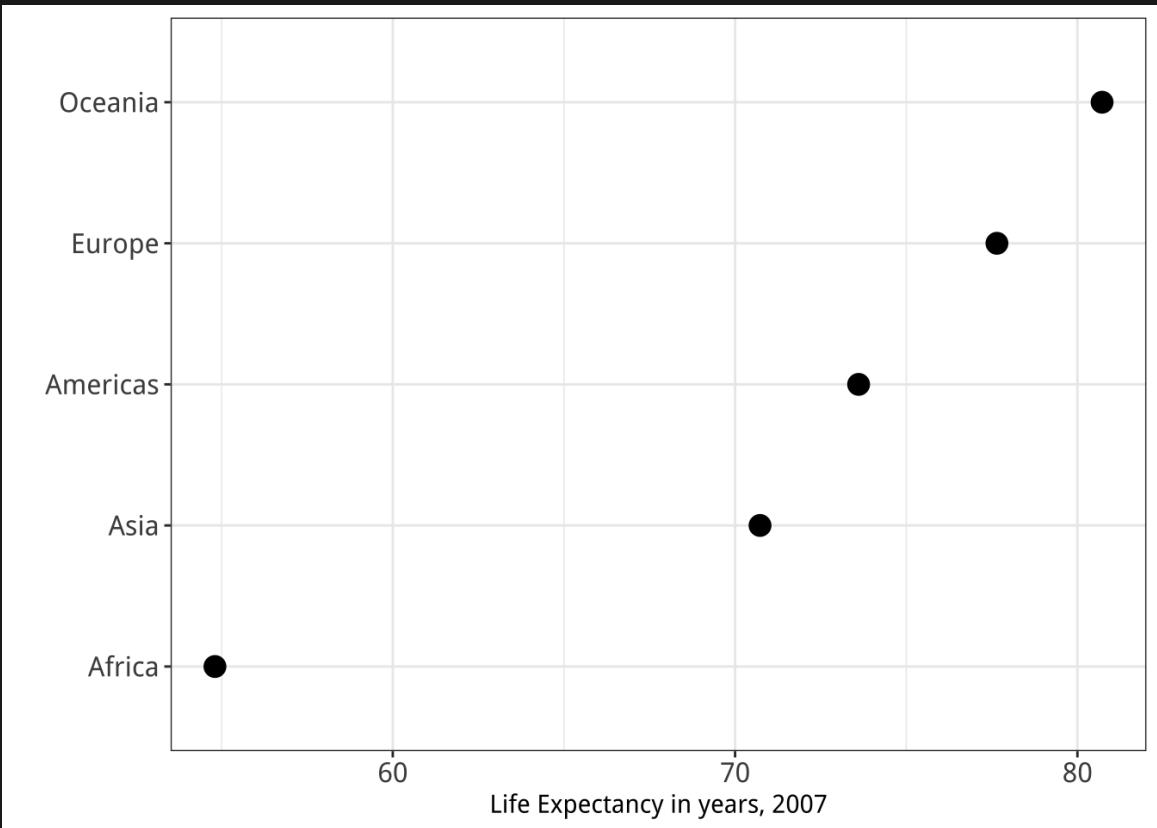
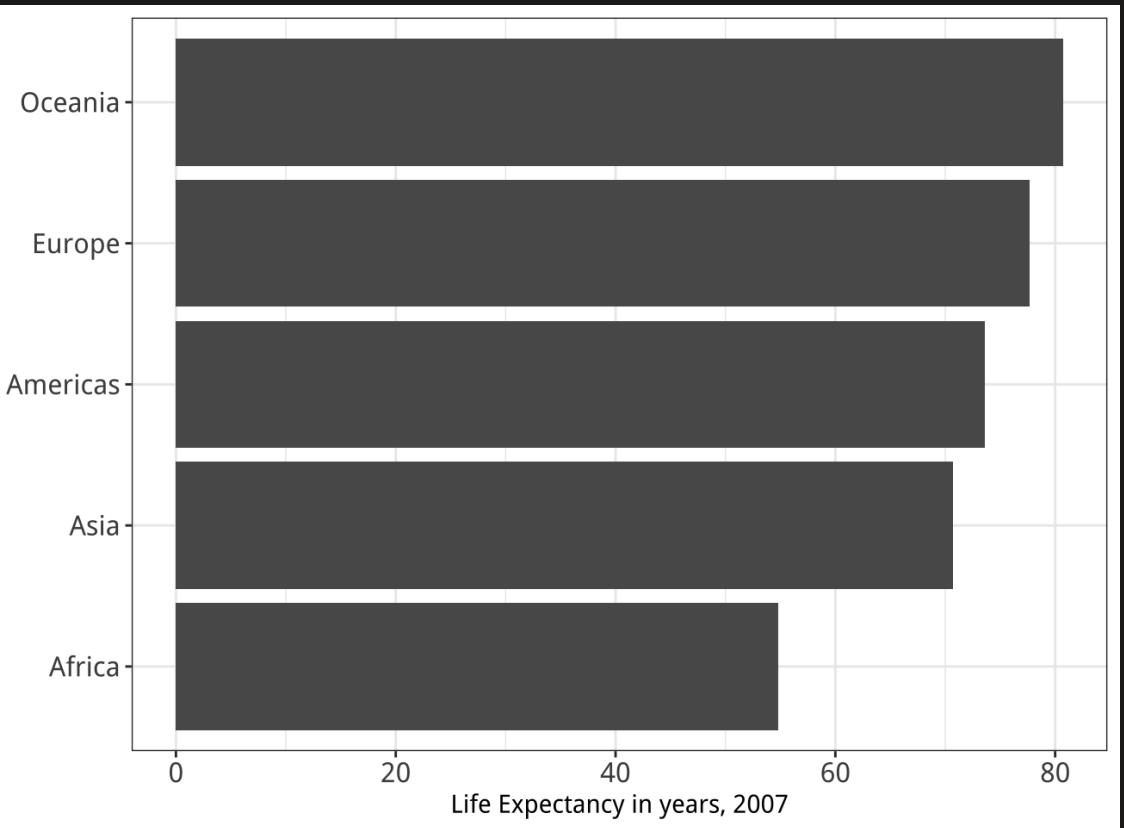
2) Identify the design rules that are broken (if any)

3) Suggest at least two improvements

Recidivism Rate of Prisoners  
Released in 1994

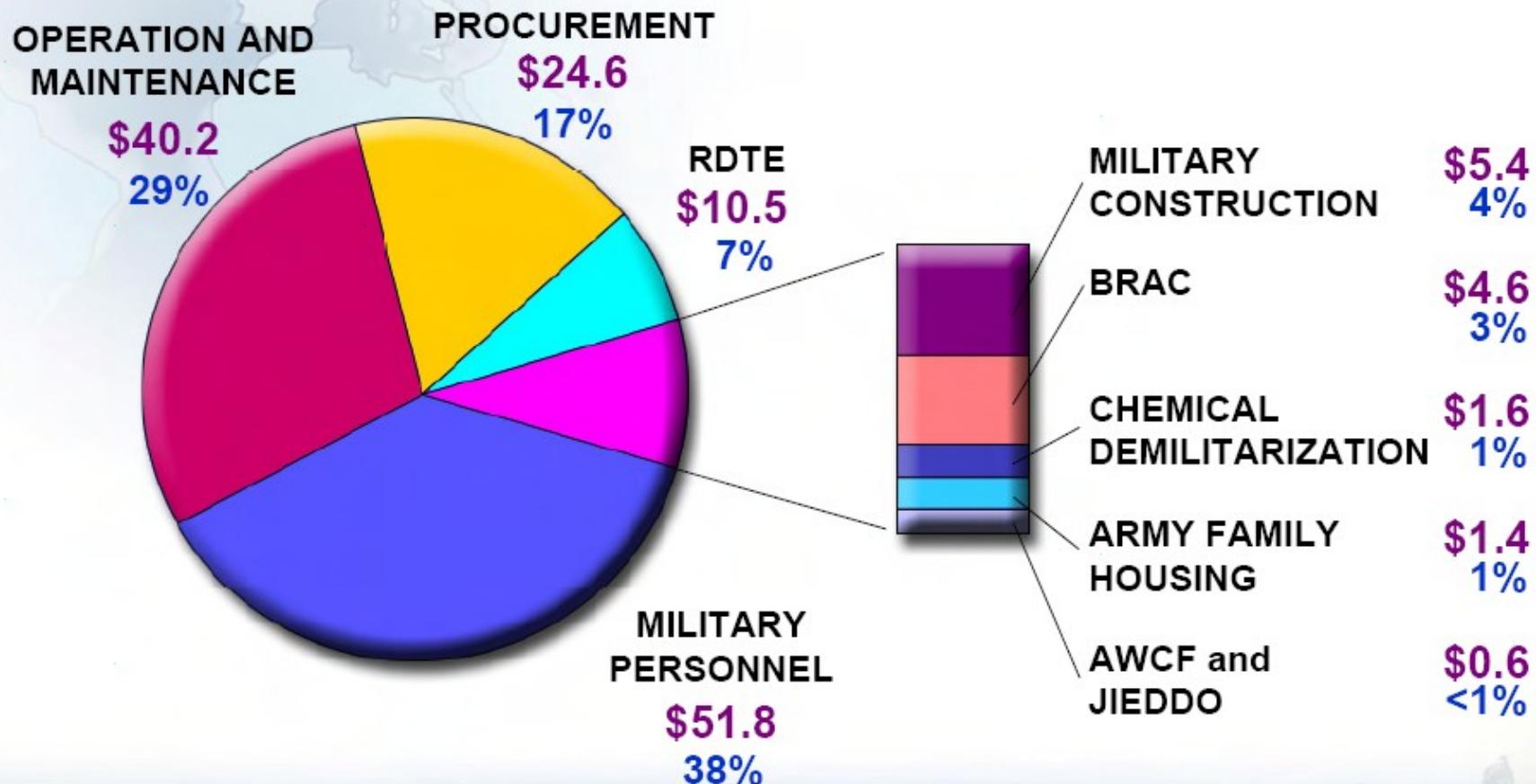






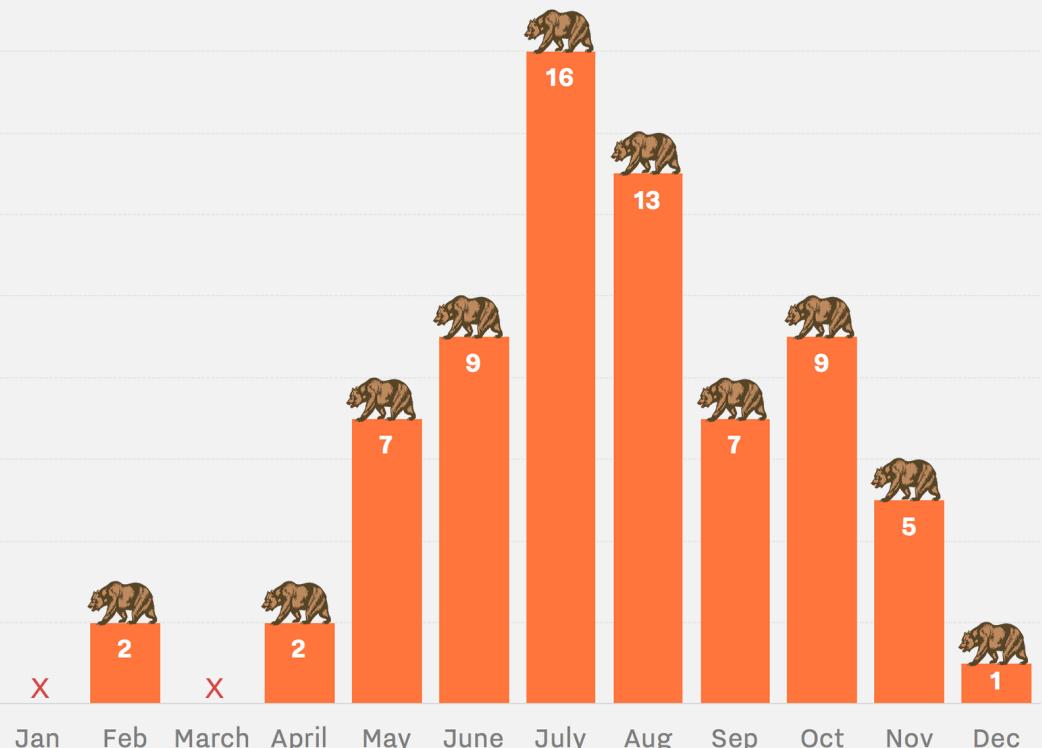


# FY09 Obligation Authority



## Most fatal bear attacks occur in July and August

Total fatal bear attacks (grizzly, black, and polar), 1900 to present



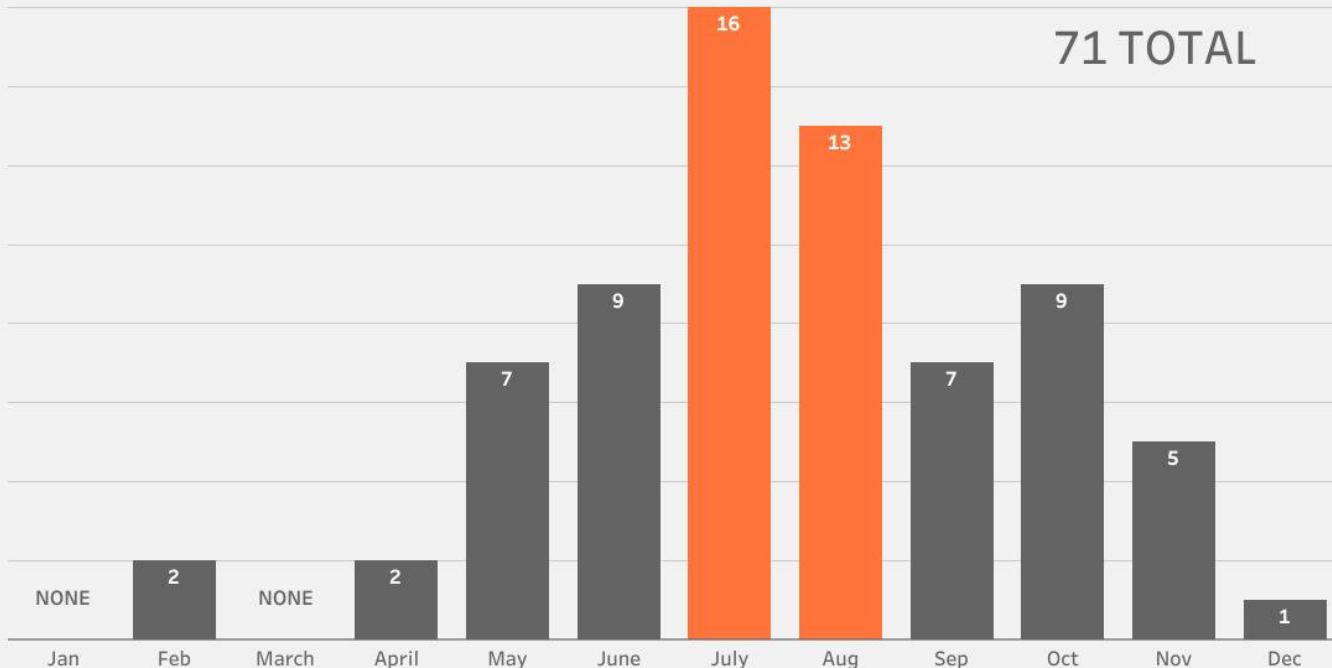
Source: News archives, Wikipedia

Vox

## BEAR ATTACKS IN U.S. PARKS & WILDERNESS AREAS

Most fatal bear attacks occur in July and August

Total fatal bear attacks by grizzly, black and polar bears from 1900 to present



Source: News archives, Wikipedia (as of 10/2016)

Created by Jeffrey A. Shaffer | MakeoverMonday 2019WK21

# ONE HUNDRED AND THIRTY EIGHT

Fatal Wild Bear Attacks in North America - 1900 to 2018

Only 17% occurred during winter and spring

*Be careful in the **summer** & **autumn***

20

Spring

65

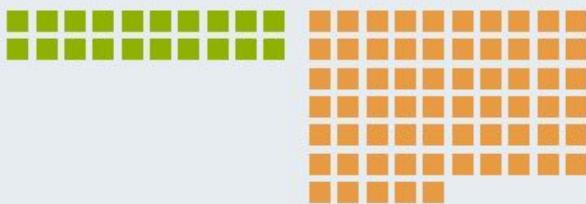
Summer

49

Autumn

4

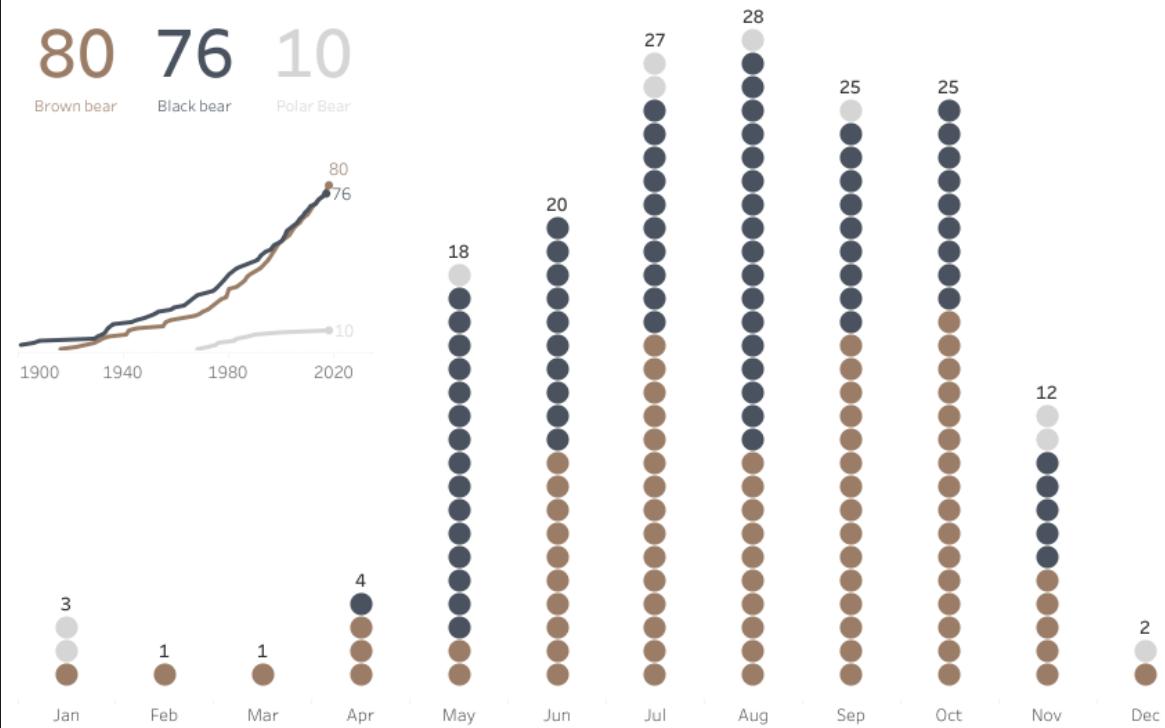
Winter



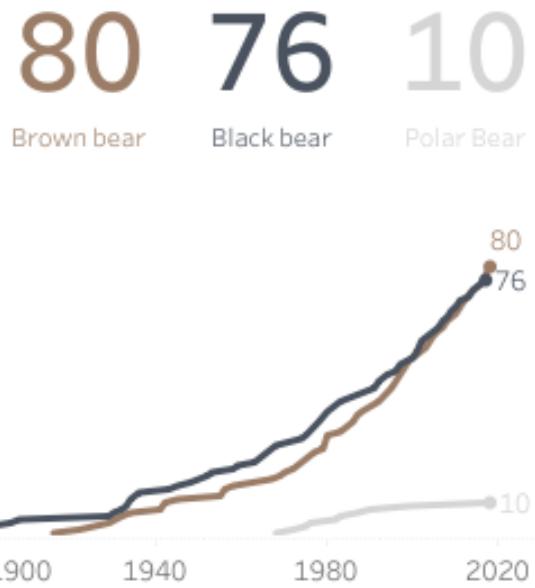
Data Source: en.wikipedia.org/wiki/List\_of\_fatal\_bear\_attacks\_in\_North\_America

Data Preparation: Ali Sanne | Design: Daniel Caroli | Shield Design by Icon Works

## BEAR ATTACKS IN NORTH AMERICA: 1900-2018

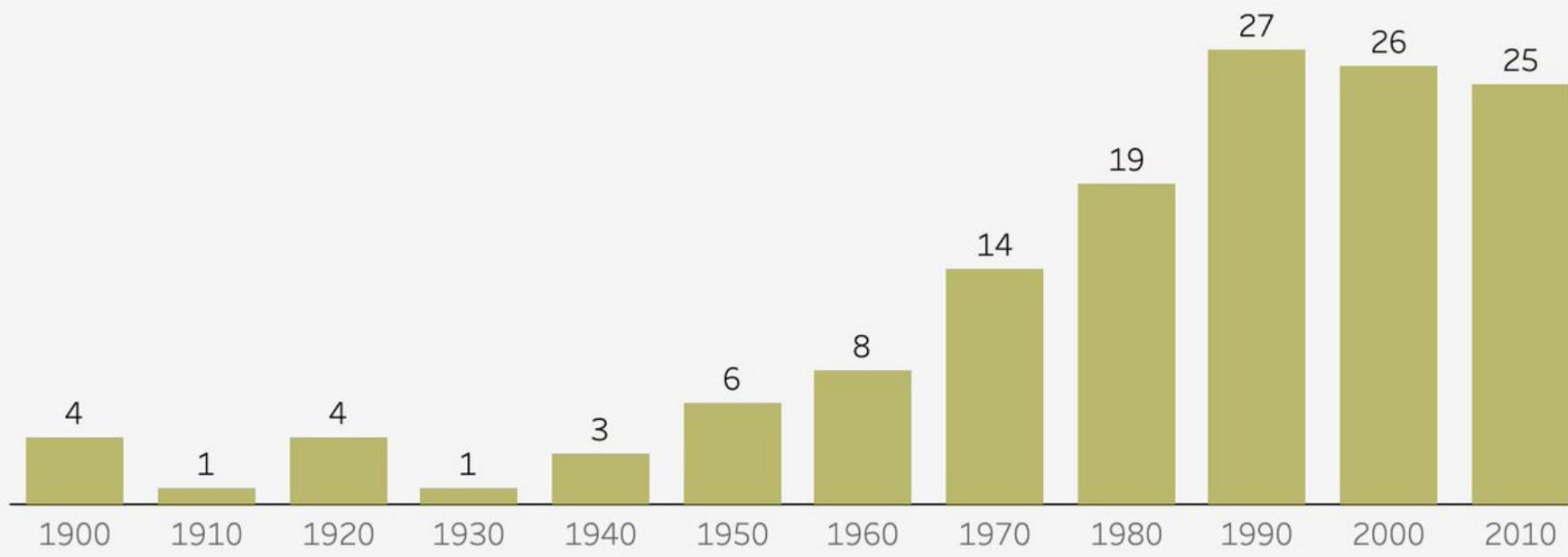


## BEAR ATTACKS



## FATAL WILD BEAR ATTACKS IN NORTH AMERICA BY DECADE

DATA COVERS THE PERIOD 1900-2018



SOURCE: Ali Sanne via Wikipedia | DESIGN: @CharlieHTableau

# Extra Slides



Hillary Clinton

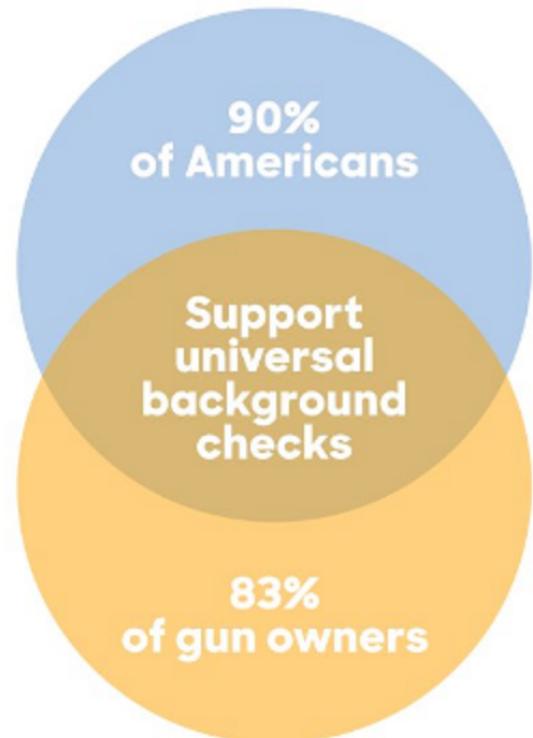
Follow

Dear Congress,

Let's get this done.

Thanks,

The vast majority of Americans



RETWEETS  
2,308 LIKES  
5,333



**People who know  
how to make  
Venn Diagrams**

**Hillary's graphic  
design staff**



Hillary Clinton   
@HillaryClinton

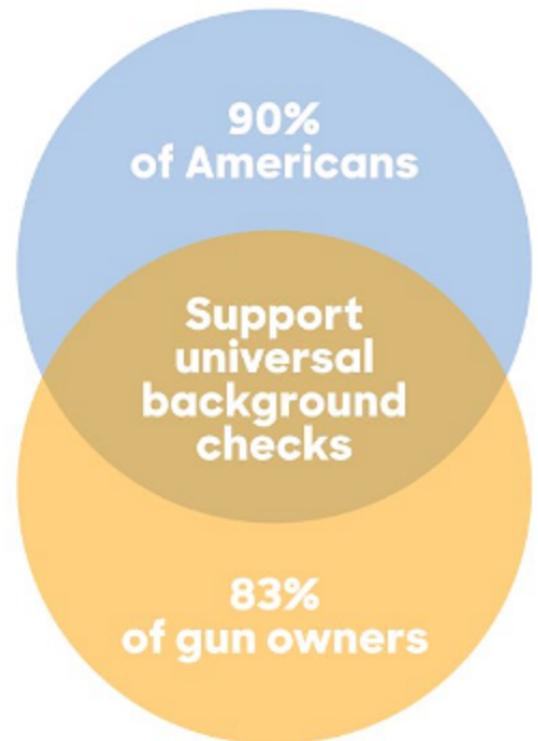
Follow

Dear Congress,

Let's get this done.

Thanks,

The vast majority of Americans



RETWEETS 2,308 LIKES 5,333



Hillary Clinton   
@HillaryClinton

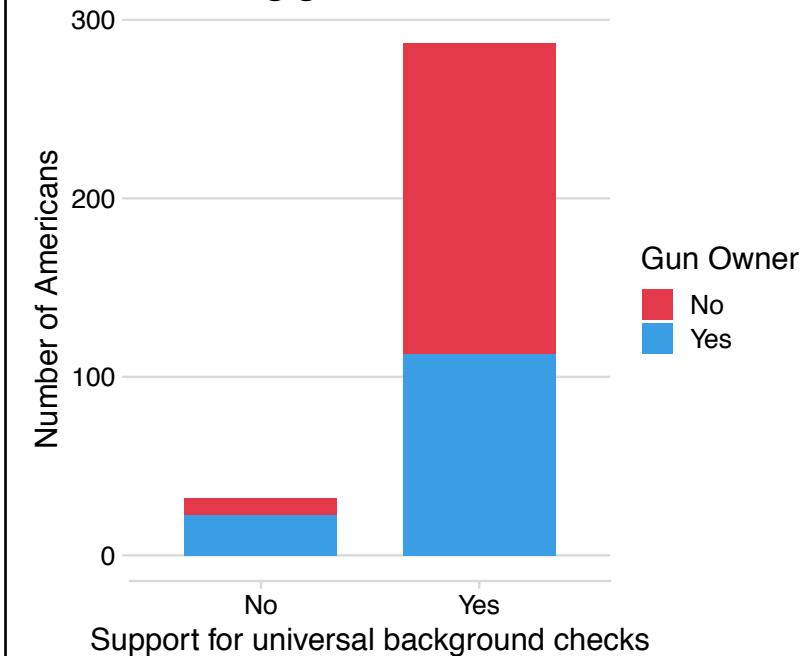
Follow

Dear Congress,

Let's get this done.

Thanks,

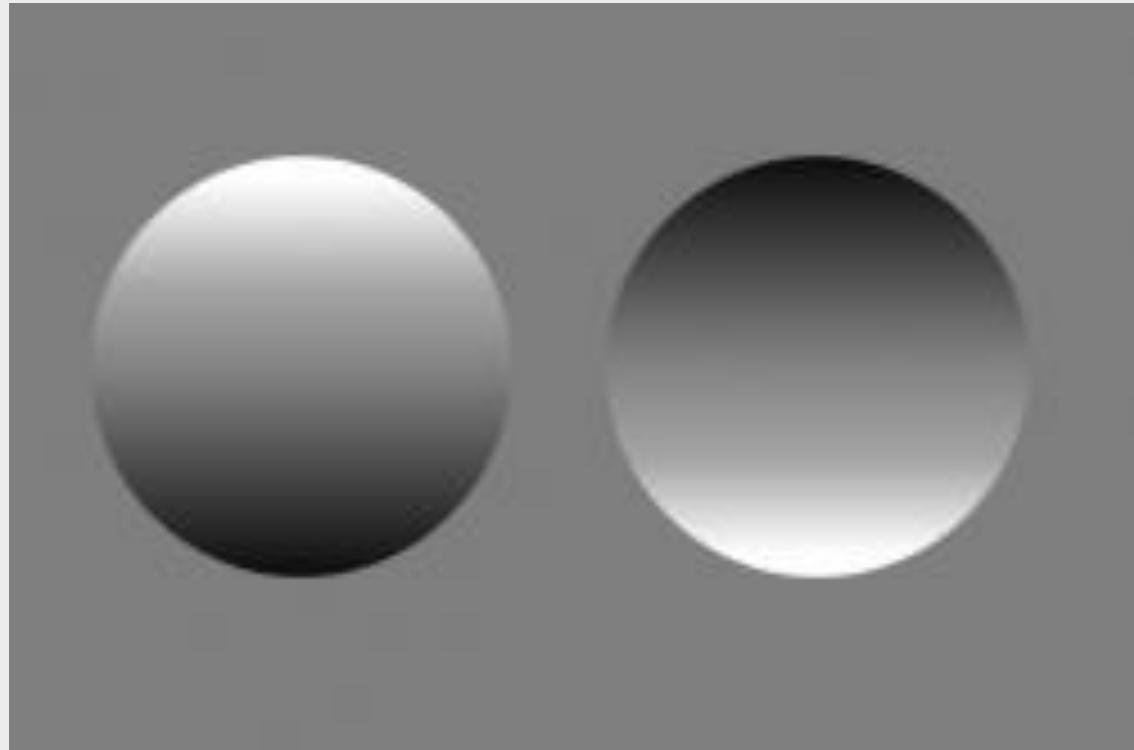
**The vast majority of Americans support universal background checks, including gun owners**

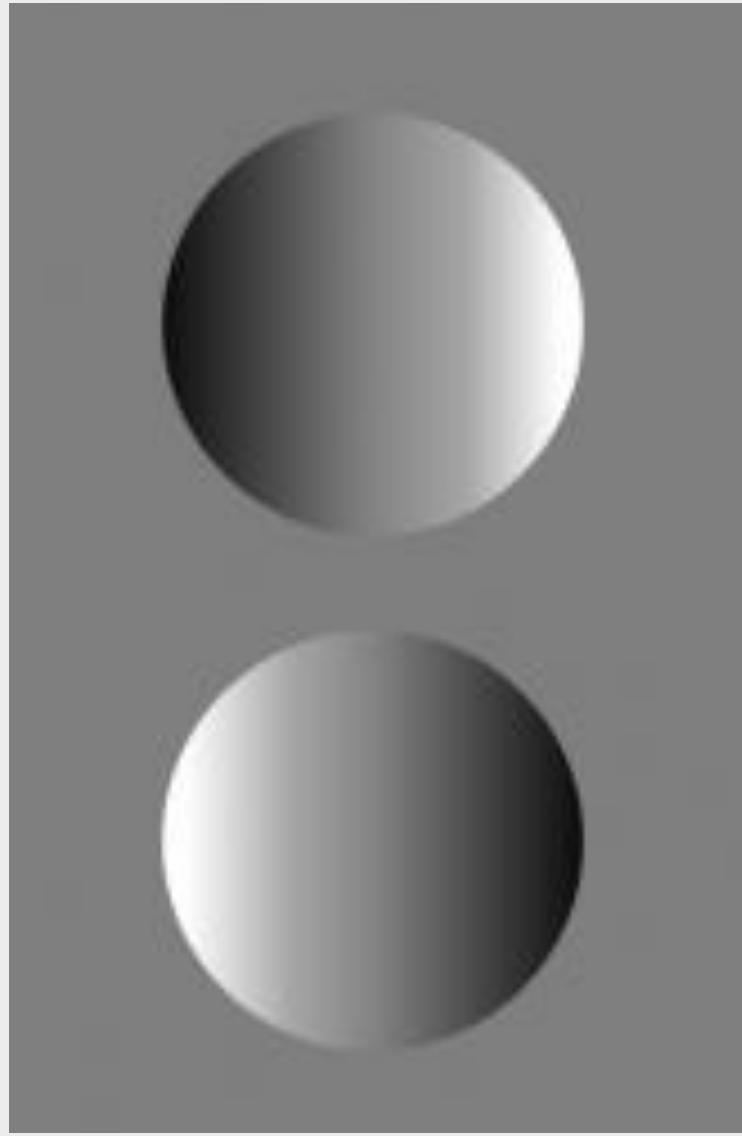


RETWEETS 2,308 LIKES 5,333



2014 2015 2016





# Do: Consider using *graphs* instead of *tables*

**Table 5**

Regression Coefficient for weighted U.S. and China models in the WTP space

Attribute	Coef.	Model 1: MNL		Model 2: MXL	
		U.S.	China	U.S.	China
Price	$\mu$	0.052 (0.002)***	0.033 (0.002)***	0.066 (0.003)***	0.039 (0.002)***
Powertrain type (base = CV)					
HEV	$\mu$	-1.176 (1.611)	4.882 (1.917)	-0.418 (1.585)	4.962 (1.992)
	$\sigma$	-	-	0.188 (4.664)	19.403 (7.723)
PHEV10	$\mu$	0.027 (1.782)	-1.291 (2.069)	0.822 (1.796)	-1.748 (2.098)
	$\sigma$	-	-	2.197 (5.428)	6.055 (8.872)
PHEV20	$\mu$	1.695 (1.751)	-1.242 (2.031)	3.207 (1.734)	-2.245 (2.074)
	$\sigma$	-	-	8.664 (5.719)	4.041 (6.792)
PHEV40	$\mu$	2.650 (1.774)	0.930 (2.023)	3.304 (1.741)	0.380 (2.039)
	$\sigma$	-	-	7.141 (5.466)	9.108 (6.179)
BEV75	$\mu$	-20.137 (1.978)***	-6.032 (2.088)***	-18.453 (1.934)***	-7.627 (2.365)***
	$\sigma$	-	-	4.175 (6.232)	29.843 (7.417)***
BEV100	$\mu$	-19.496 (1.984)***	-8.151 (2.144)***	-18.947 (1.965)***	-10.377 (2.286)***
	$\sigma$	-	-	1.898 (5.368)	8.600 (7.340)
BEV150	$\mu$	-13.691 (1.959)***	1.305 (2.050)	-12.727 (1.959)***	0.616 (2.075)
	$\sigma$	-	-	10.486 (6.061)	6.973 (6.406)
Brand (base = German)					
American	$\mu$	8.188 (1.289)***	-10.574 (1.560)***	7.432 (1.268)***	-7.612 (1.687)***
	$\sigma$	-	-	0.665 (3.439)	19.299 (5.866)***
Japanese	$\mu$	0.934 (1.289)	-18.098 (1.689)***	-0.577 (1.289)	-15.169 (1.790)***
	$\sigma$	-	-	11.765 (3.508)***	23.666 (5.941)***
Chinese	$\mu$	-19.008 (1.550)***	-9.674 (1.509)***	-19.848 (1.666)***	-6.049 (1.691)***
	$\sigma$	-	-	8.078 (4.173)	34.541 (6.544)***
S. Korean	$\mu$	-9.510 (1.398)***	-19.361 (1.725)***	-10.412 (1.378)***	-17.774 (2.124)***
	$\sigma$	-	-	12.335 (3.850)***	54.771 (6.171)***
Cost and performance					
PHEV fast-charge	$\mu$	3.944 (1.330)***	7.615 (1.565)***	3.331 (1.335)	7.567 (1.653)***
	$\sigma$	-	-	8.882 (4.396)	20.119 (5.449)***
BEV fast-charge	$\mu$	3.343 (1.478)	6.662 (1.599)***	0.030 (1.821)	6.428 (1.668)***
	$\sigma$	-	-	26.237 (3.871)***	11.567 (5.360)
Operating cost	$\mu$	-1.598 (0.106)***	-3.214 (0.242)***	-1.626 (0.104)***	-3.467 (0.267)***
	$\sigma$	-	-	0.076 (0.247)	3.275 (0.968)***
Acceleration time	$\mu$	-1.172 (0.255)***	-4.651 (0.299)***	-1.269 (0.293)***	-4.878 (0.319)***
	$\sigma$	-	-	5.766 (0.880)***	3.359 (0.949)***
LL:		-3425.6	-6788.8	-3373.1	-6720.9
Null model LL:		-4360.5	-7487.3	-4360.6	-7487.3
AIC:		6883.3	13609.6	6808.3	13503.8
McFadden $R^2$ :		0.21	0.09	0.23	0.10
Adj. McFadden $R^2$ :		0.21	0.09	0.22	0.10
Num. of Obs:		5760	6720	5760	6720

Standard errors of estimates are presented in parenthesis. Coefficient units are USD \$1000. \* <0.05. \*\* <0.01. \*\*\* <0.001.

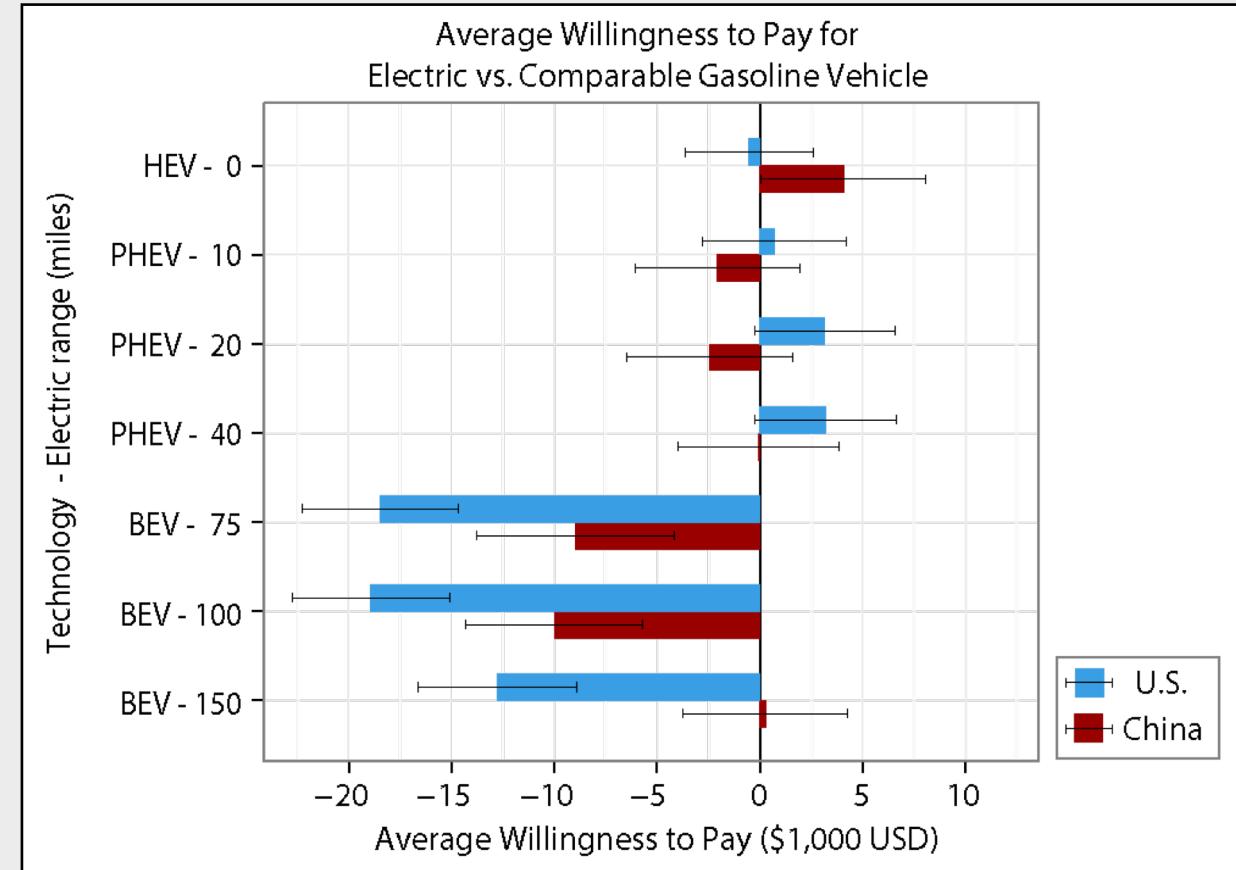
# Do: Consider using graphs instead of tables

**Table 5**

Regression Coefficient for weighted U.S. and China models in the WTP space

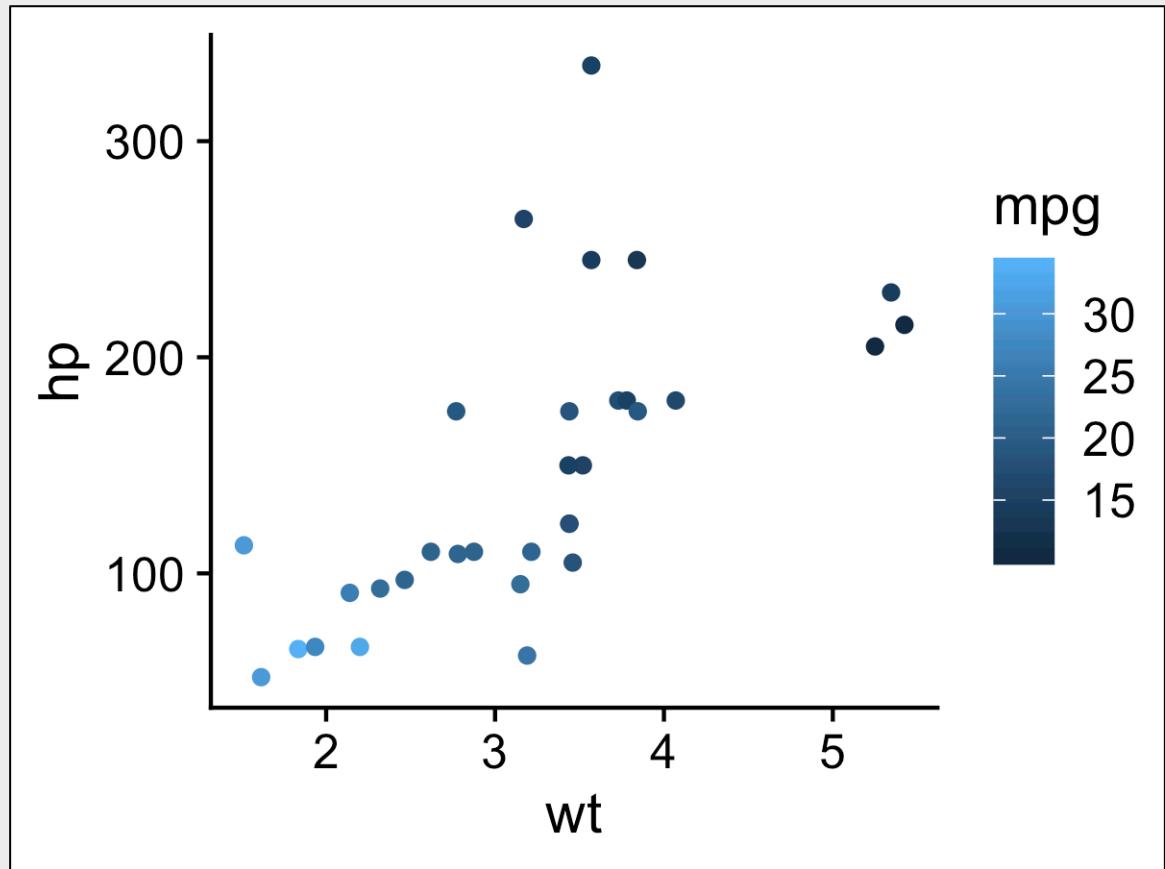
Attribute	Coef.	Model 1: MNL		Model 2: MXL	
		U.S.	China	U.S.	China
Price	$\mu$	0.052 (0.002)***	0.033 (0.002)***	0.066 (0.003)***	0.039 (0.002)***
Powertrain type (base = CV)	$\mu$	-1.176 (1.611)	4.882 (1.917)	-0.418 (1.585)	4.962 (1.992)
HEV	$\mu$	-	-	0.188 (4.664)	19.403 (7.723)
PHEV10	$\mu$	0.027 (1.782)	-1.291 (2.069)	0.822 (1.796)	-1.748 (2.098)
PHEV20	$\mu$	1.695 (1.751)	-1.242 (2.031)	2.197 (5.428)	6.055 (8.872)
PHEV40	$\mu$	2.650 (1.774)	0.930 (2.023)	3.207 (1.734)	-2.245 (2.074)
PHEV40	$\sigma$	-	-	8.664 (5.719)	4.041 (6.792)
BEV75	$\mu$	-20.137 (1.978)***	-6.032 (2.088)***	-18.453 (1.934)***	-7.627 (2.365)***
BEV100	$\mu$	-19.496 (1.984)***	-8.151 (2.144)***	-18.947 (1.965)***	-10.377 (2.286)***
BEV150	$\mu$	-13.691 (1.959)***	1.305 (2.050)	-12.727 (1.959)***	8.600 (7.340)
BEV150	$\sigma$	-	-	10.486 (6.061)	0.616 (2.075)
Brand (base = German)					
American	$\mu$	8.188 (1.289)***	-10.574 (1.560)***	7.432 (1.268)***	-7.612 (1.687)***
American	$\sigma$	-	-	0.665 (3.439)	19.299 (5.866)***
Japanese	$\mu$	0.934 (1.289)	-18.098 (1.689)***	-0.577 (1.289)	-15.169 (1.790)***
Chinese	$\mu$	-19.008 (1.550)***	-9.674 (1.509)***	11.765 (3.508)***	23.666 (5.941)***
Chinese	$\sigma$	-	-	-19.848 (1.666)***	-6.049 (1.691)***
S. Korean	$\mu$	-9.510 (1.398)***	-19.361 (1.725)***	8.078 (4.173)	34.541 (6.544)***
S. Korean	$\sigma$	-	-	-10.412 (1.378)***	-17.774 (2.124)***
Cost and performance					
PHEV fast-charge	$\mu$	3.944 (1.330)***	7.615 (1.565)***	3.331 (1.335)	7.567 (1.653)***
PHEV fast-charge	$\sigma$	-	-	8.882 (4.396)	20.119 (5.449)***
BEV fast-charge	$\mu$	3.343 (1.478)	6.662 (1.599)***	0.030 (1.821)	6.428 (1.668)***
BEV fast-charge	$\sigma$	-	-	26.237 (3.871)***	11.567 (5.360)
Operating cost	$\mu$	-1.598 (0.106)***	-3.214 (0.242)***	-1.626 (0.104)***	-3.467 (0.267)***
Operating cost	$\sigma$	-	-	0.076 (0.247)	3.275 (0.968)***
Acceleration time	$\mu$	-1.172 (0.255)***	-4.651 (0.299)***	-1.269 (0.293)***	-4.878 (0.319)***
Acceleration time	$\sigma$	-	-	5.766 (0.880)***	3.359 (0.949)***
LL:		-3425.6	-6788.8	-3373.1	-6720.9
Null model LL:		-4360.5	-7487.3	-4360.6	-7487.3
AIC:		6883.3	13609.6	6808.3	13503.8
McFadden $R^2$ :		0.21	0.09	0.23	0.10
Adj. McFadden $R^2$ :		0.21	0.09	0.22	0.10
Num. of Obs:		5760	6720	5760	6720

Standard errors of estimates are presented in parenthesis. Coefficient units are USD \$1000. \* <0.05. \*\* <0.01. \*\*\* <0.001.



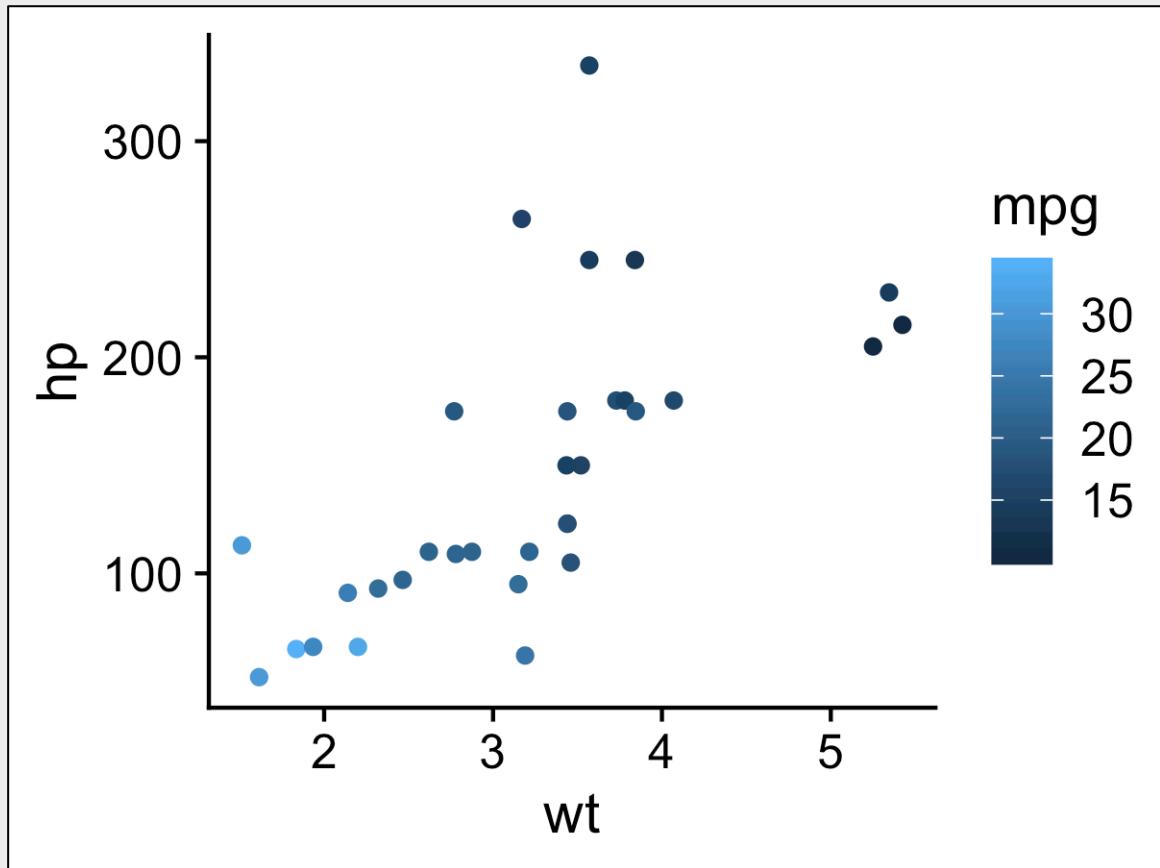
## Gradient:

# Change in continuous variable



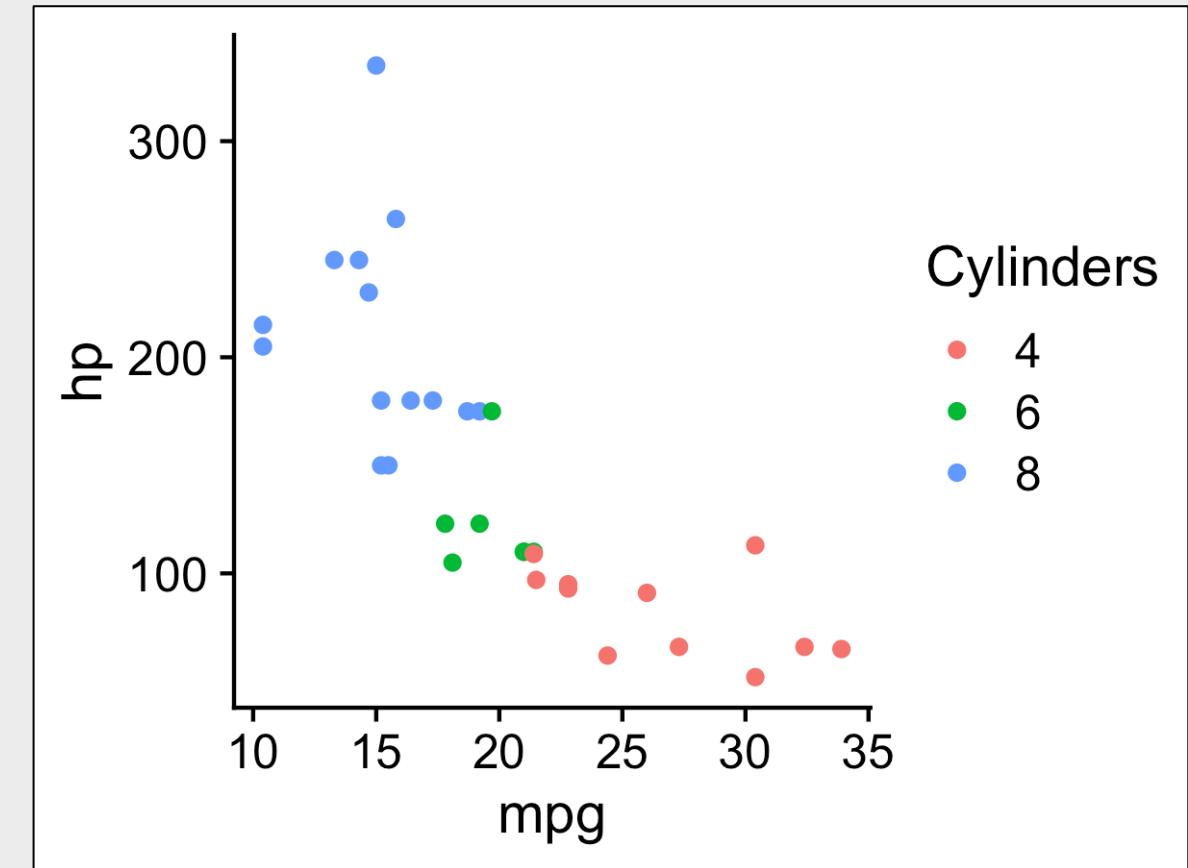
**Gradient:**

Change in continuous variable



**Discrete:**

Identify (a few) categorical variables



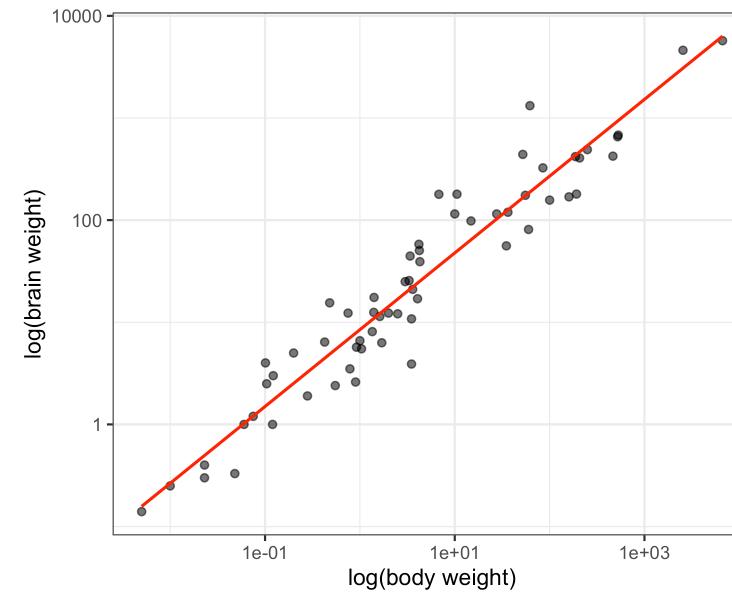
## Tables:

Data organized into rows and columns

Mammal	Body (kg)	Brain (g)
Arctic fox	3.385	44.5
Owl monkey	0.48	15.5
Mountain beaver	1.35	8.1
Cow	465	423
Grey wolf	36.33	119.5
Goat	27.66	115

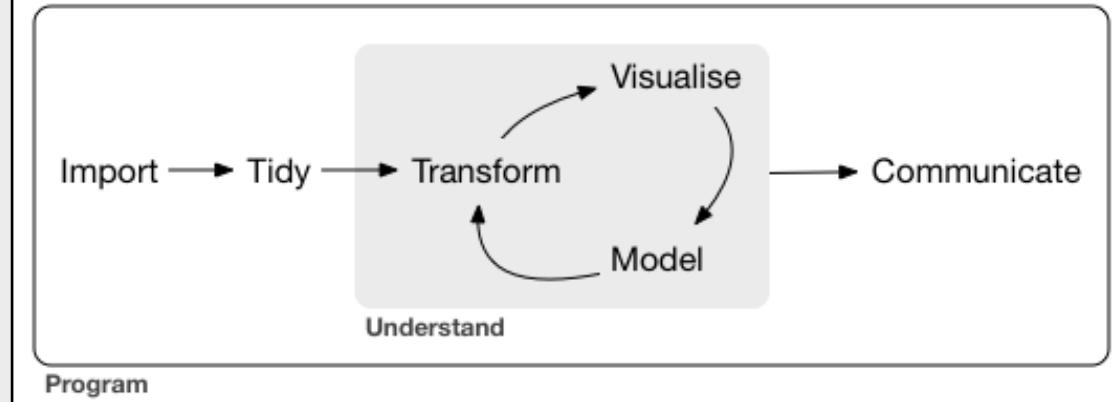
## Graphs:

Visual representation of information in tables



## Diagrams:

Visual representation of a process



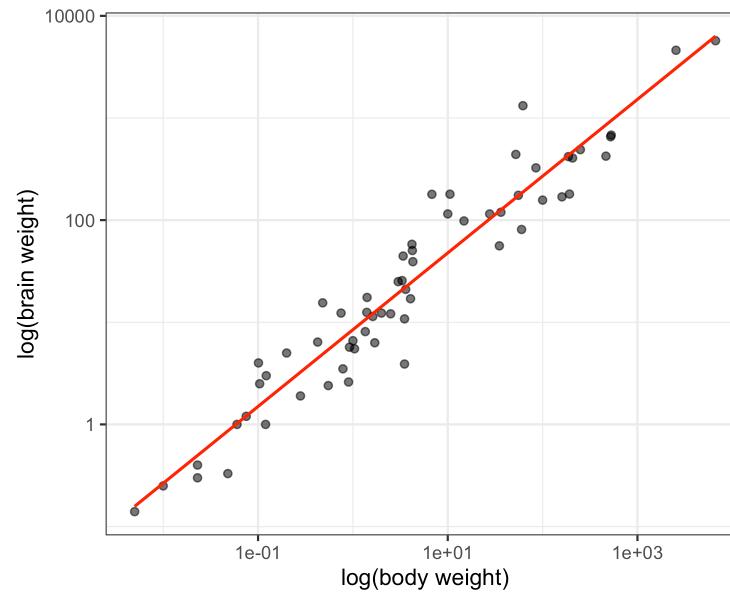
## Tables:

Data organized into rows and columns

Mammal	Body (kg)	Brain (g)
Arctic fox	3.385	44.5
Owl monkey	0.48	15.5
Mountain beaver	1.35	8.1
Cow	465	423
Grey wolf	36.33	119.5
Goat	27.66	115

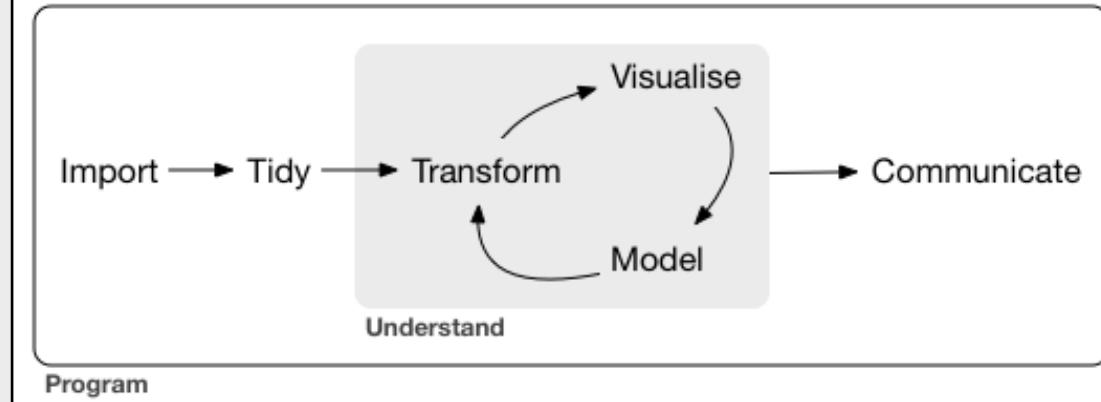
## Graphs:

Visual representation of information in tables



## Diagrams:

Visual representation of process



# “Visualizing data helps us think<sup>1</sup>”

A		B		C		D		
x	y	x	y	x	y	x	y	
10	8.04	10	9.14	10	7.46	8	6.58	
8	6.95	8	8.14	8	6.77	8	5.76	
13	7.58	13	8.74	13	12.74	8	7.71	
9	8.81	9	8.77	9	7.11	8	8.84	
11	8.33	11	9.26	11	7.81	8	8.47	
14	9.96	14	8.1	14	8.84	8	7.04	
6	7.24	6	6.13	6	6.08	8	5.25	
4	4.26	4	3.1	4	5.39	19	12.5	
12	10.84	12	9.13	12	8.15	8	5.56	
7	4.82	7	7.26	7	6.42	8	7.91	
5	5.68	5	4.74	5	5.73	8	6.89	
Sum:	99	82.51	99	82.51	99	82.5	99	82.51
Mean:	9	7.5	9	7.5	9	7.5	9	7.5
St. Dev:	3.3	2	3.3	2	3.3	2	3.3	2

# “Visualizing data helps us think<sup>1</sup>”

