

COS30045

DATA VISUALISATION

TOPIC 04: VISUAL VARIABLES

Commonwealth of Australia
Copyright Act 1968

Notice for paragraph 135ZXA (a) of the *Copyright Act 1968*

Warning

This material has been reproduced and communicated to you by or on behalf of Swinburne University of Technology under Part VB of the *Copyright Act 1968* (the *Act*).

The material in this communication may be subject to copyright under the *Act*. Any further reproduction or communication of this material by you may be the subject of copyright protection under the *Act*.

Do not remove this notice.

IMAGES USED IN THIS PRESENTATION: FROM MICROSOFT CLIPART (UNLESS OTHERWISE SPECIFIED)

Acknowledgment: This unit is partially based on the Visualisation unit developed by Alex Lex and Hanspeter Hfister (with permission).

Source of images used in this lecture available as links in the PDF version.



VISUAL VARIABLES: MARKS AND CHANNELS

What:

- ▶ marks and channels: graphic and visual elements for representing data

Why:

- ▶ different graphical and visual elements are better at representing different types of data, choose the right elements for your data

How:

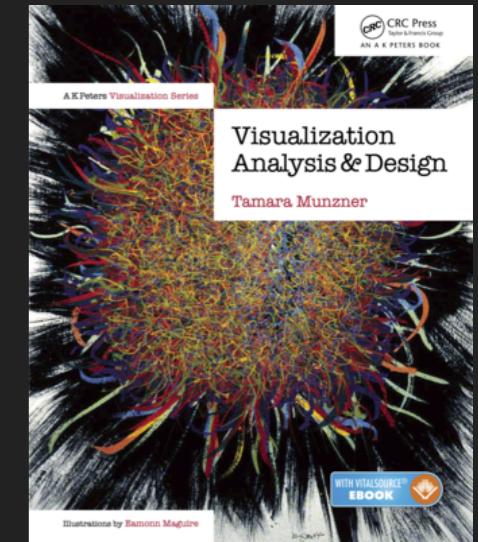
- ▶ understand the pros and cons of using different graphical elements for different types of data



VISUAL VARIABLES

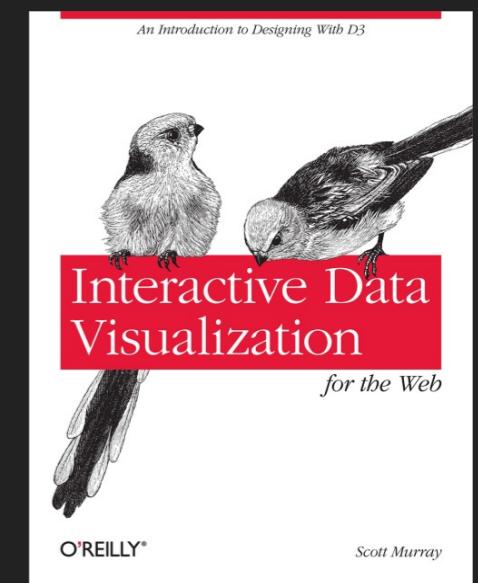
Lecture Readings

- ▶ VAD Ch 5



Tutorials

- ▶ IDV Ch 5 and 6



HOW MANY DIFFERENT WAYS CAN WE REPRESENT TWO NUMBERS?

For example,

2 and 4

VISUAL VARIABLES

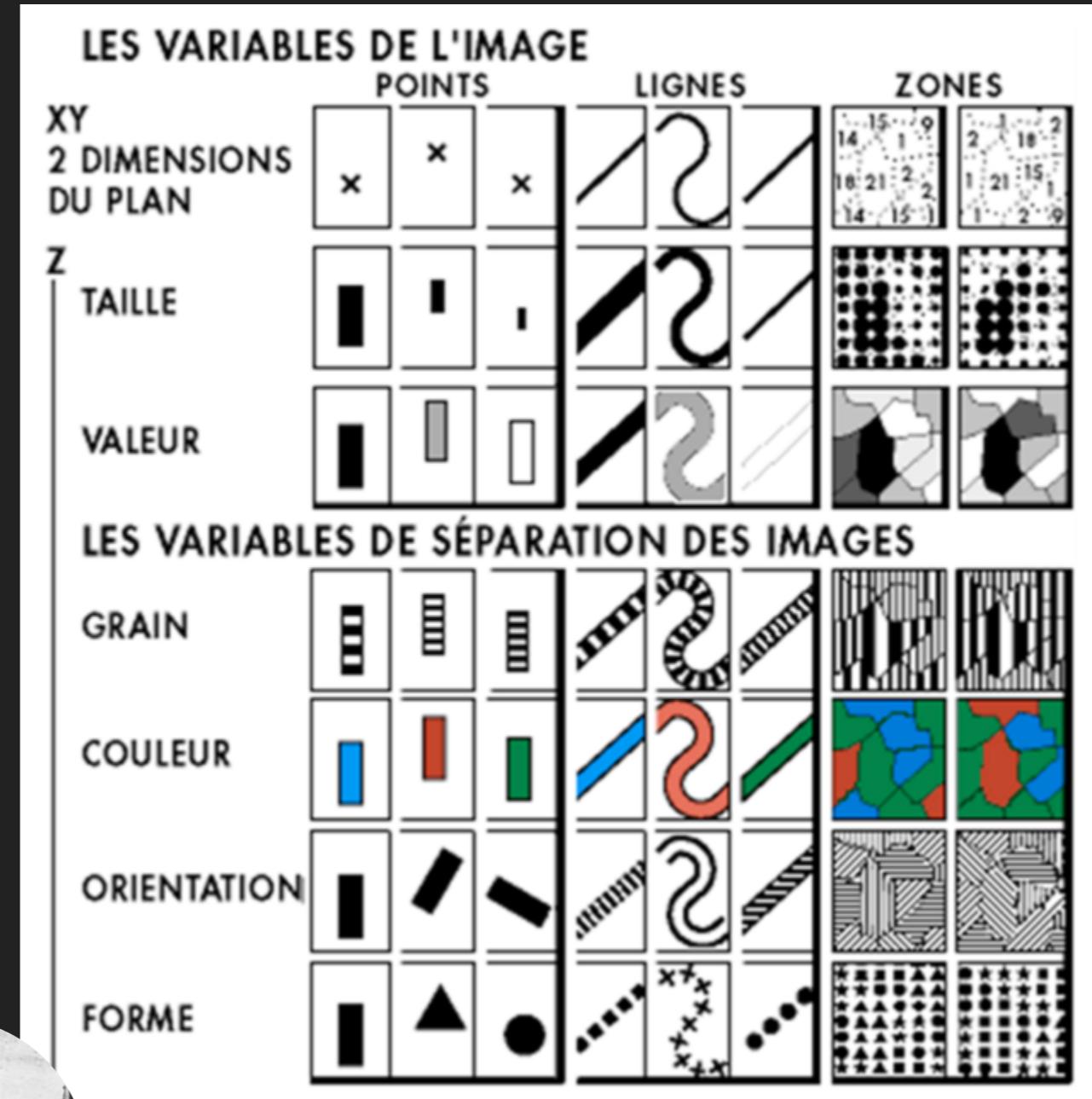
Jacques Bertin

- ▶ developed theoretical principles for visual encoding

[Graphics] is a strict and simple system of signs, which anyone can learn to use and which leads to better understanding.

Jacques Bertin

French cartographer (1918-2010)



Semiology of Graphics (1967)

MARKS AND CHANNELS

Mark

- ▶ graphic (geometric) elements that depict items or links

Channel

- ▶ visual element used to control the appearance of a mark



MARKS FOR ITEMS

→ Points



→ Lines



→ Areas



Figure 5.2. Marks are geometric primitives.

MARKS FOR LINKS (NETWORK DATA SETS)

Marks as Items/Nodes

→ Points



→ Lines



→ Areas



Marks as Links

→ Containment



hierarchal relationships
based on area

→ Connection



pair wise connection
between items

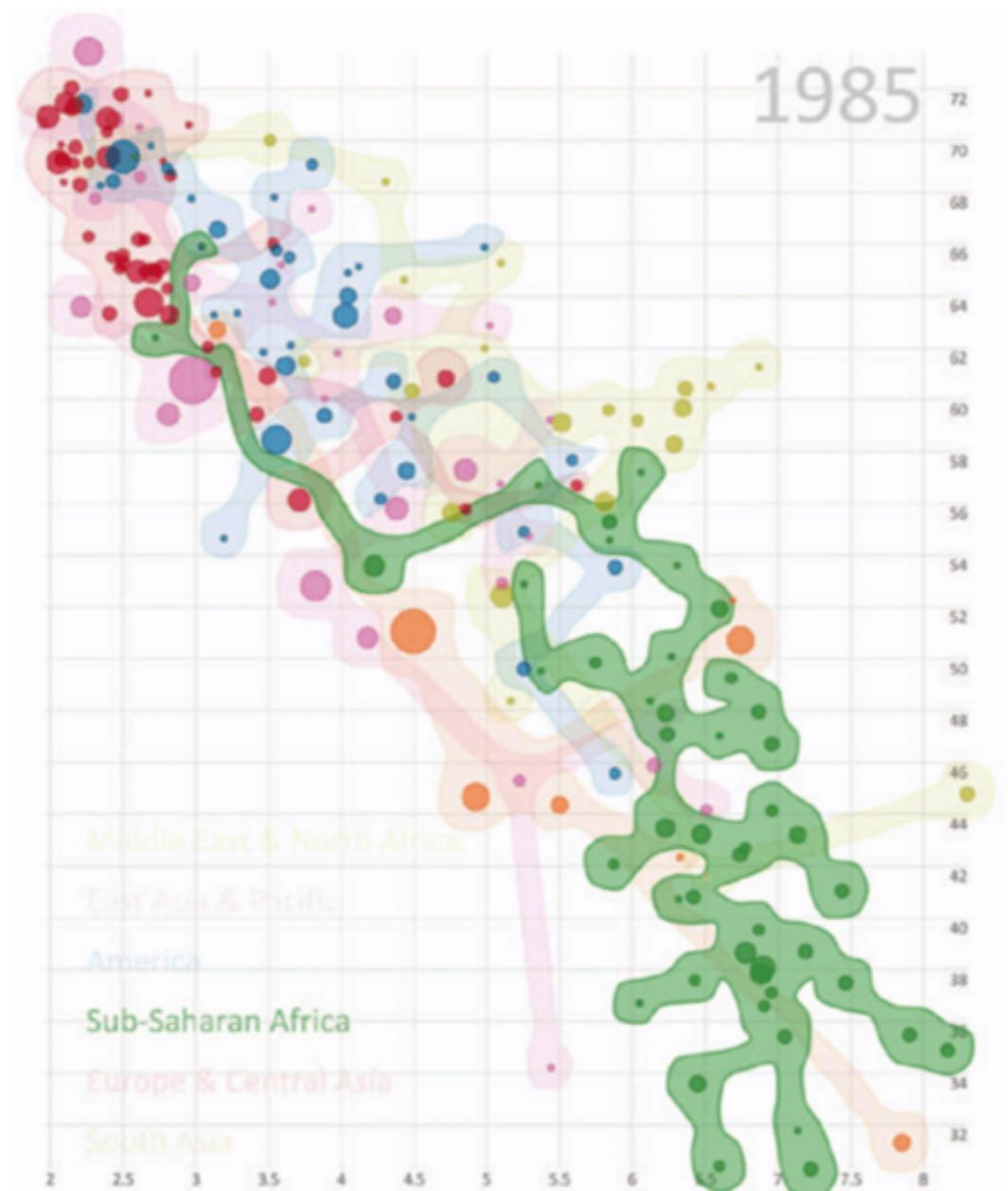
- ▶ Links can not be represented by points

LINK MARK: CONTAINMENT

Bubble Set: Collins, Penn, and Carpendale (2009)



Sets of geographically-defined items in lower Manhattan, showing hotels (orange), subway stations (brown), and medical clinics (purple). Medical clinics are noticeably absent on the West side, and there is a cluster of clinics and hotels near transit in the Northeast corner.



A scatterplot of fertility rate by life expectancy by country. Hovering on a set member causes all non-members and other sets to be made transparent, clarifying set membership. Here, enclosure eases discovery of the outliers in the upper left, as well as giving a general impression about the spatial distribution of the set.

VISUAL VARIABLES: CHANNELS

④ Position

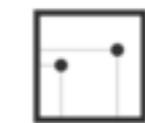
→ Horizontal



→ Vertical



→ Both



④ Color



④ Shape



④ Tilt

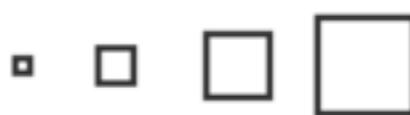


④ Size

→ Length



→ Area



→ Volume

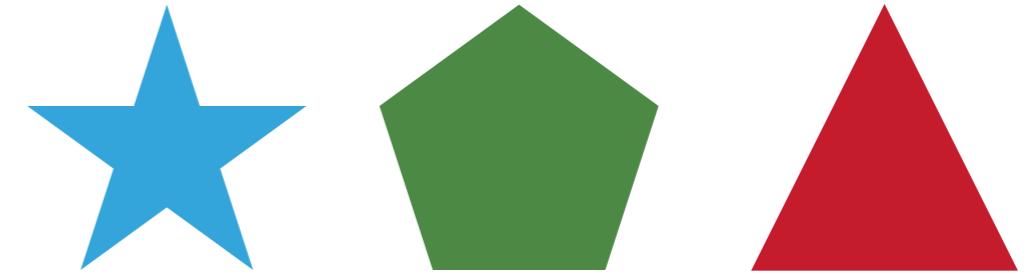


Figure 5.3. Visual channels control the appearance of marks.

CHANNEL TYPES

Identity channels

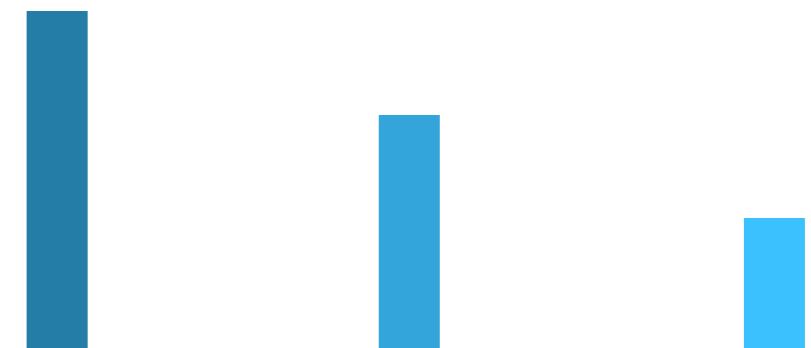
- ▶ information about **what or where**
- ▶ encode (qualitative) **categorial** attributes



What shape, what colour....
but not how much shape or colour.

Magnitude channels

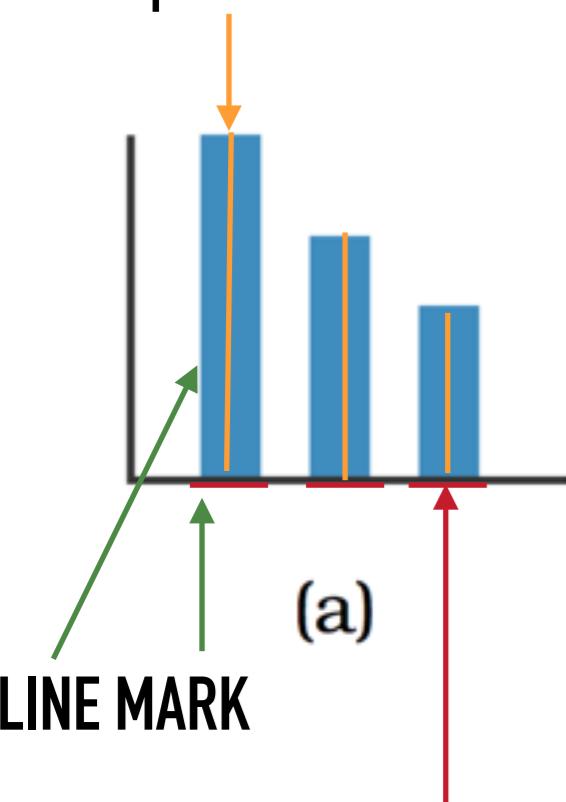
- ▶ information about **how much**
- ▶ encode **quantitative** attributes



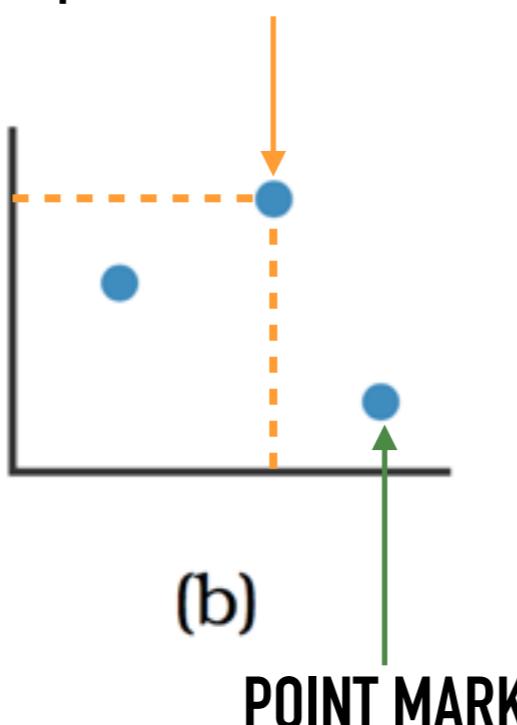
How much longer or darker is that line?

USING MARKS AND CHANNELS

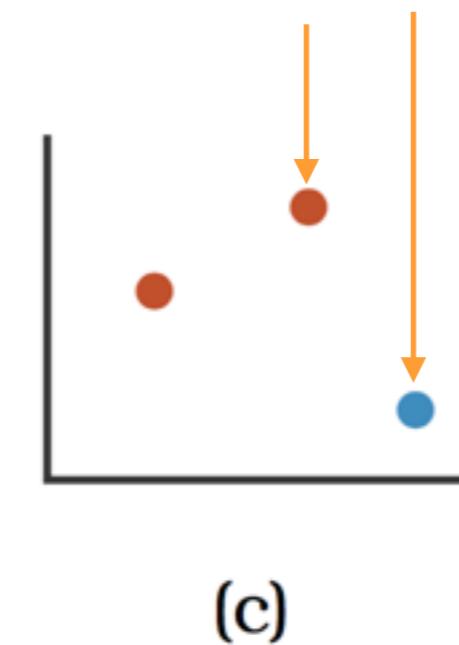
Quantitative vertical spatial position channel



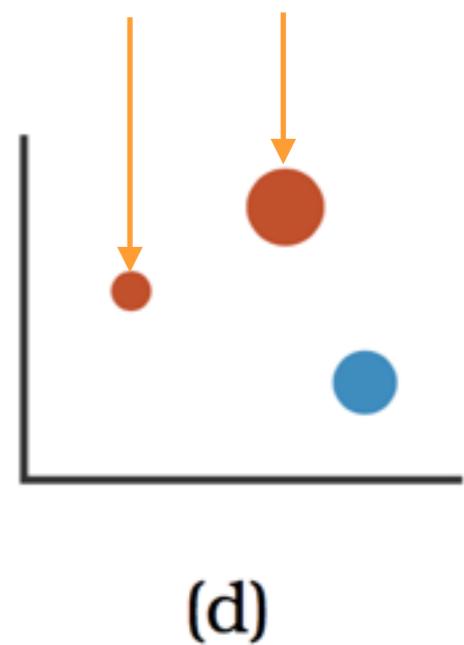
Quantitative vertical and horizontal spatial position channel



(Qualitative) categorial attribute - colour channel

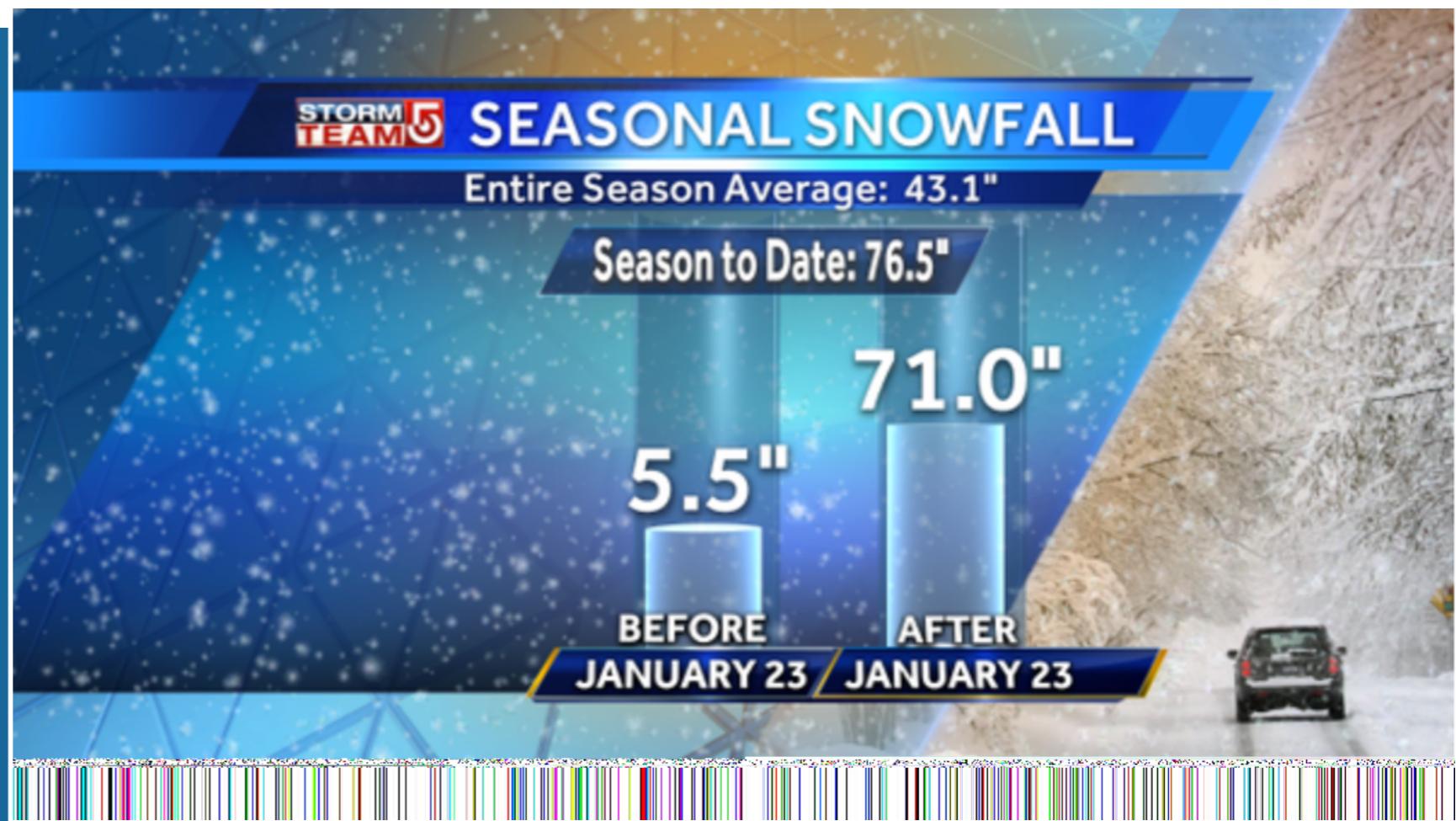
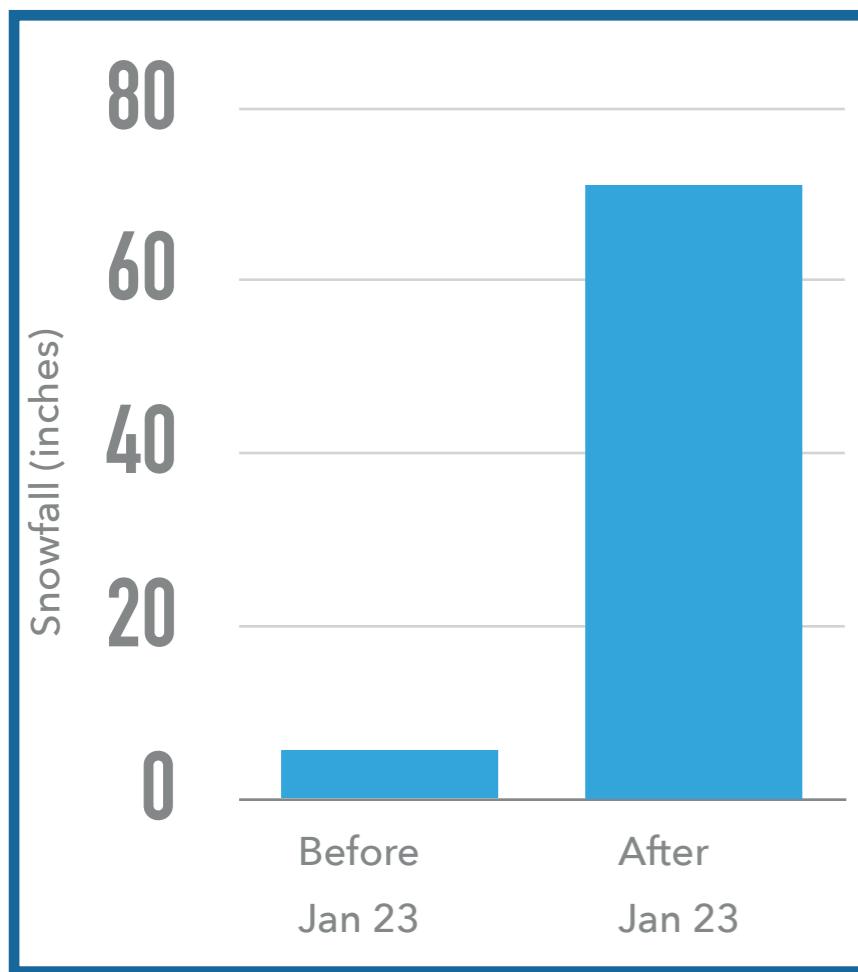


(Quantitative) attribute size channel



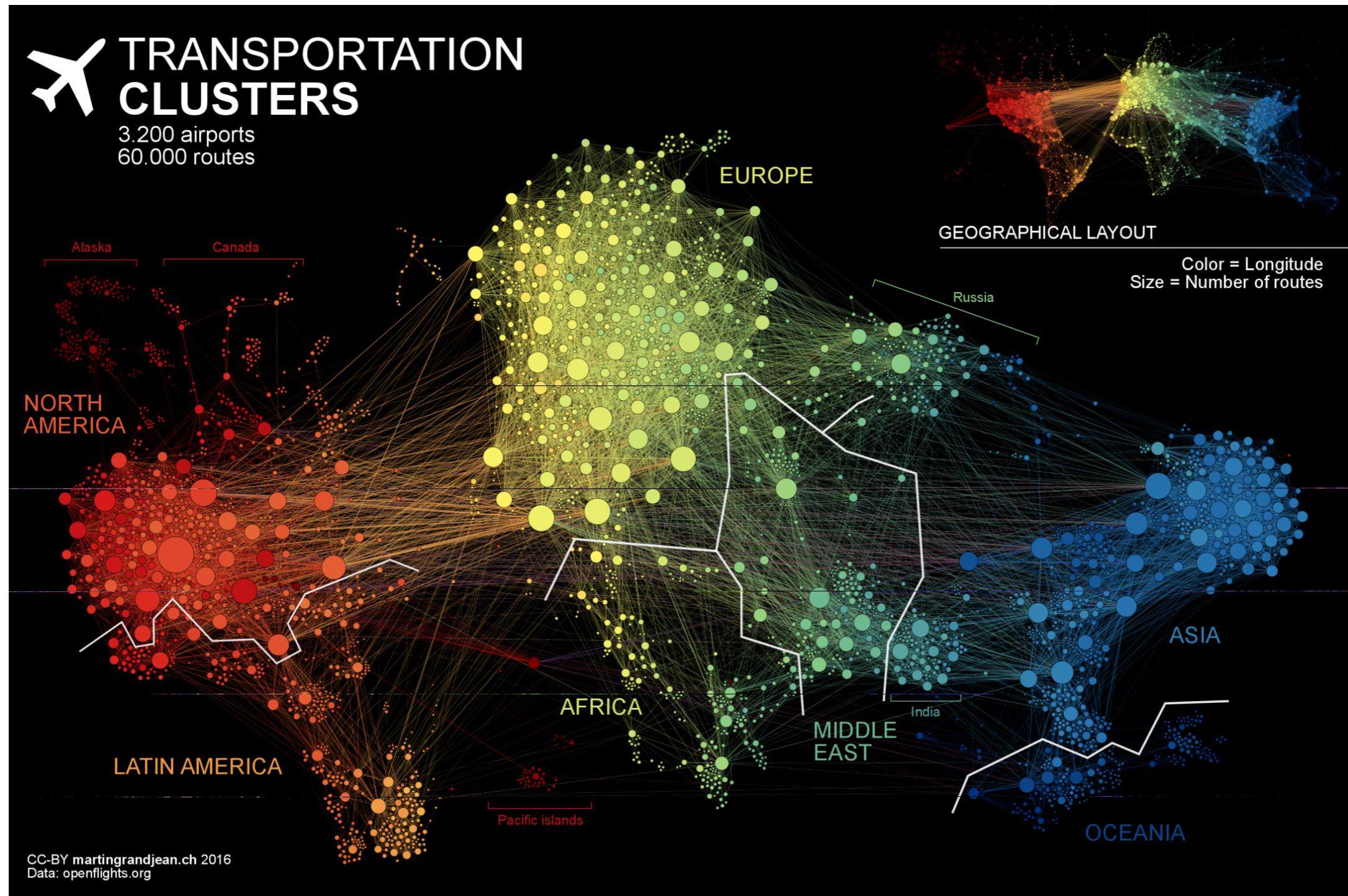
(Qualitative) categorial attribute - horizontal spatial position channel

USING MARKS AND CHANNELS: SCALE DISTORTION (LIE FACTOR)



VISUAL VARIABLES

USING MARKS AND CHANNELS IN A NETWORK



CHANNEL CHOICE GUIDELINES

When choosing channels consider

- ▶ Expressiveness
- ▶ ‘...visual encoding should express all of, and only, the information in the data set attributes.’
- ▶ if data is unordered (i.e., categorical, do not imply order by using magnitude channel)



What is being encoded in this visualisation?

Fruit and Veg are categorial attributes, but variation in saturation implies quantitative relationship between items categories that does not exist

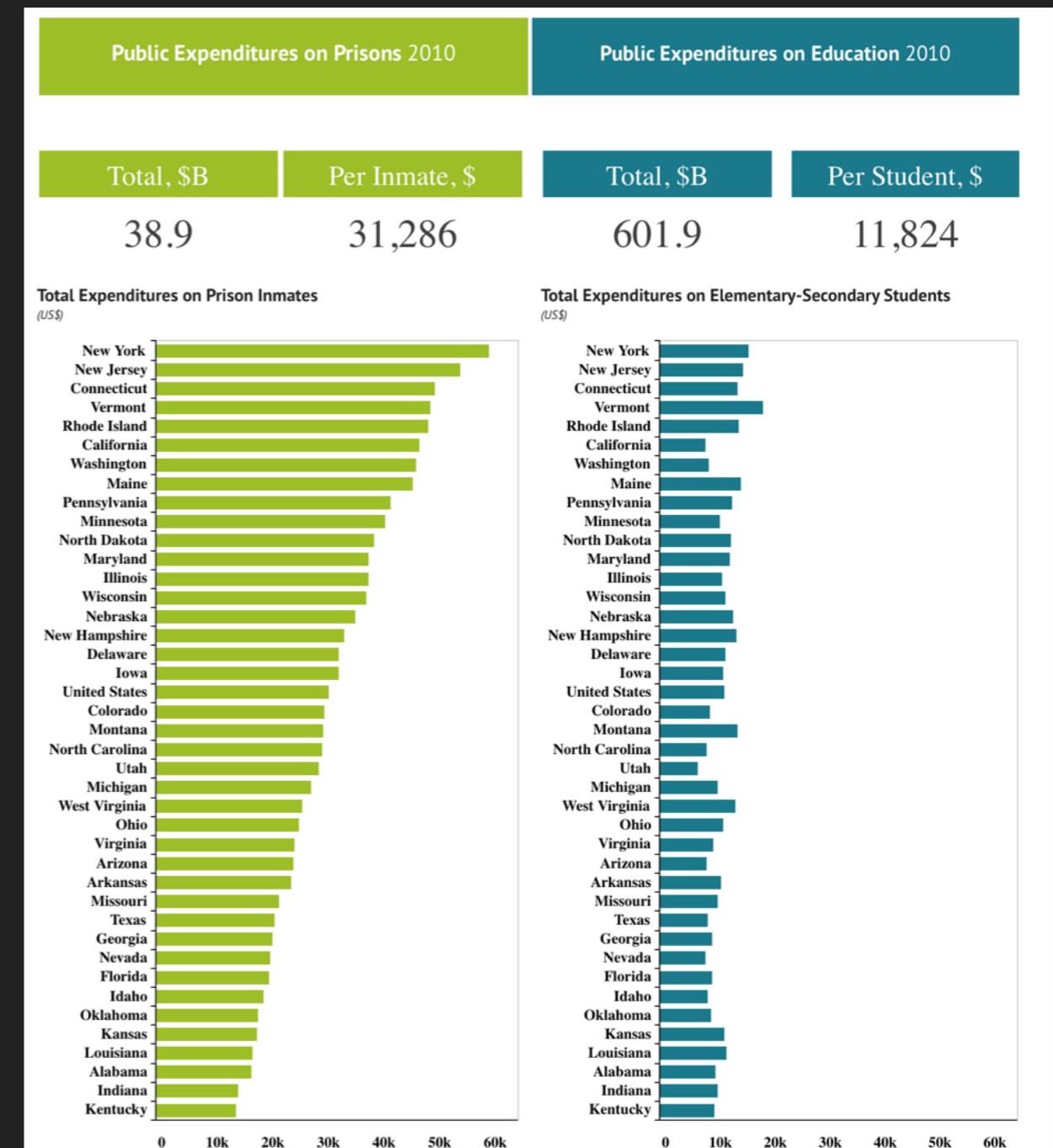
X

CHANNEL CHOICE GUIDELINES

When choosing channels consider

▶ Effectiveness

- ▶ ‘...the importance of the attribute should match the salience of the channel; that is its noticeability.’
- ▶ most important aspects of data should be use the most effective channels so as to be the most noticeable



CHANNEL EFFECTIVENESS

Channels: Expressiveness Types and Effectiveness Ranks

→ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



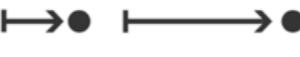
Tilt angle



Area (2D size)



Depth (3D position)



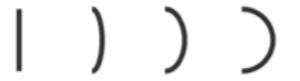
Color luminance



Color saturation



Curvature



Volume (3D size)



→ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



Shape



▲ Most
Effectiveness
▼ Least

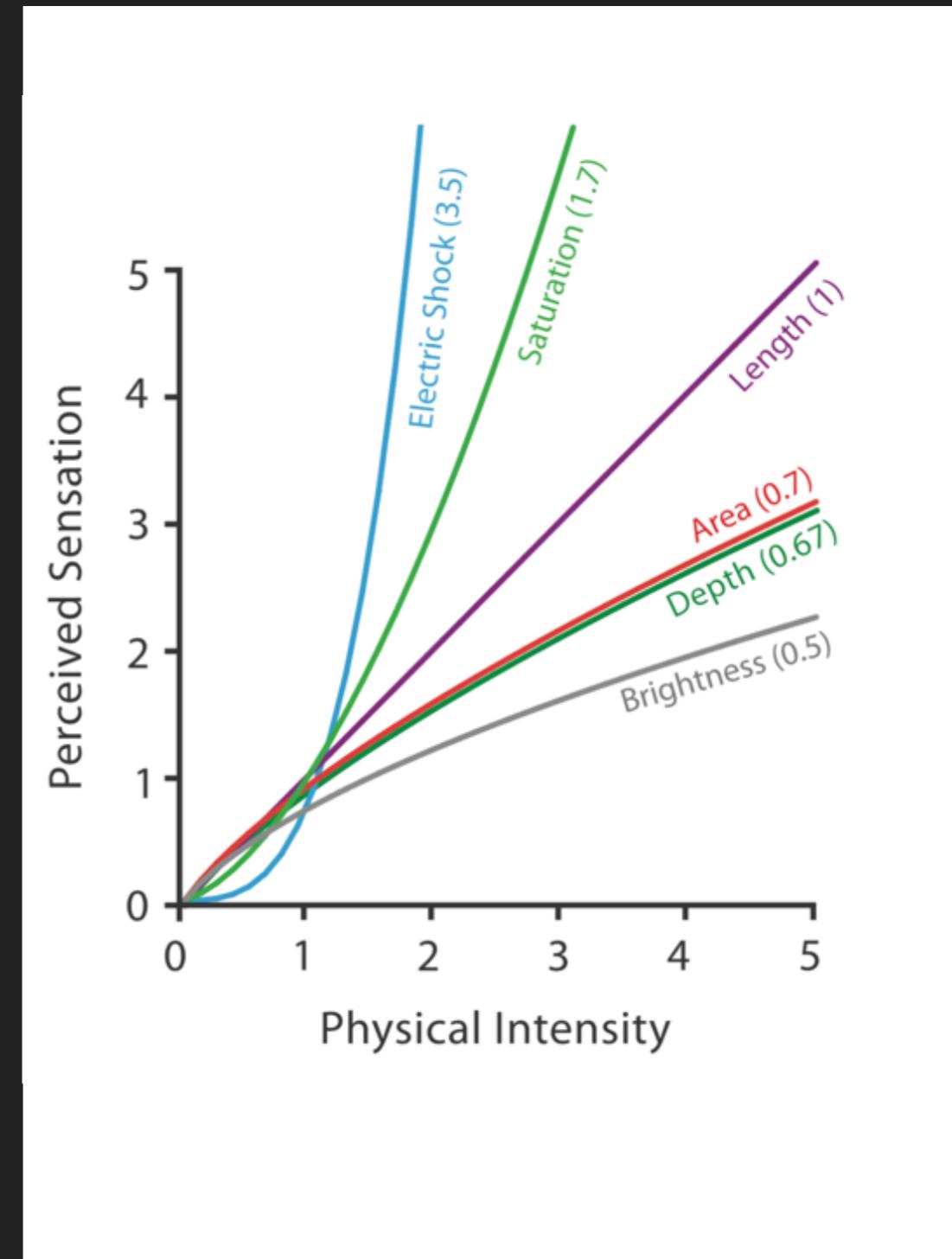
[Same]

- ▶ Attributes encoded by position most effective

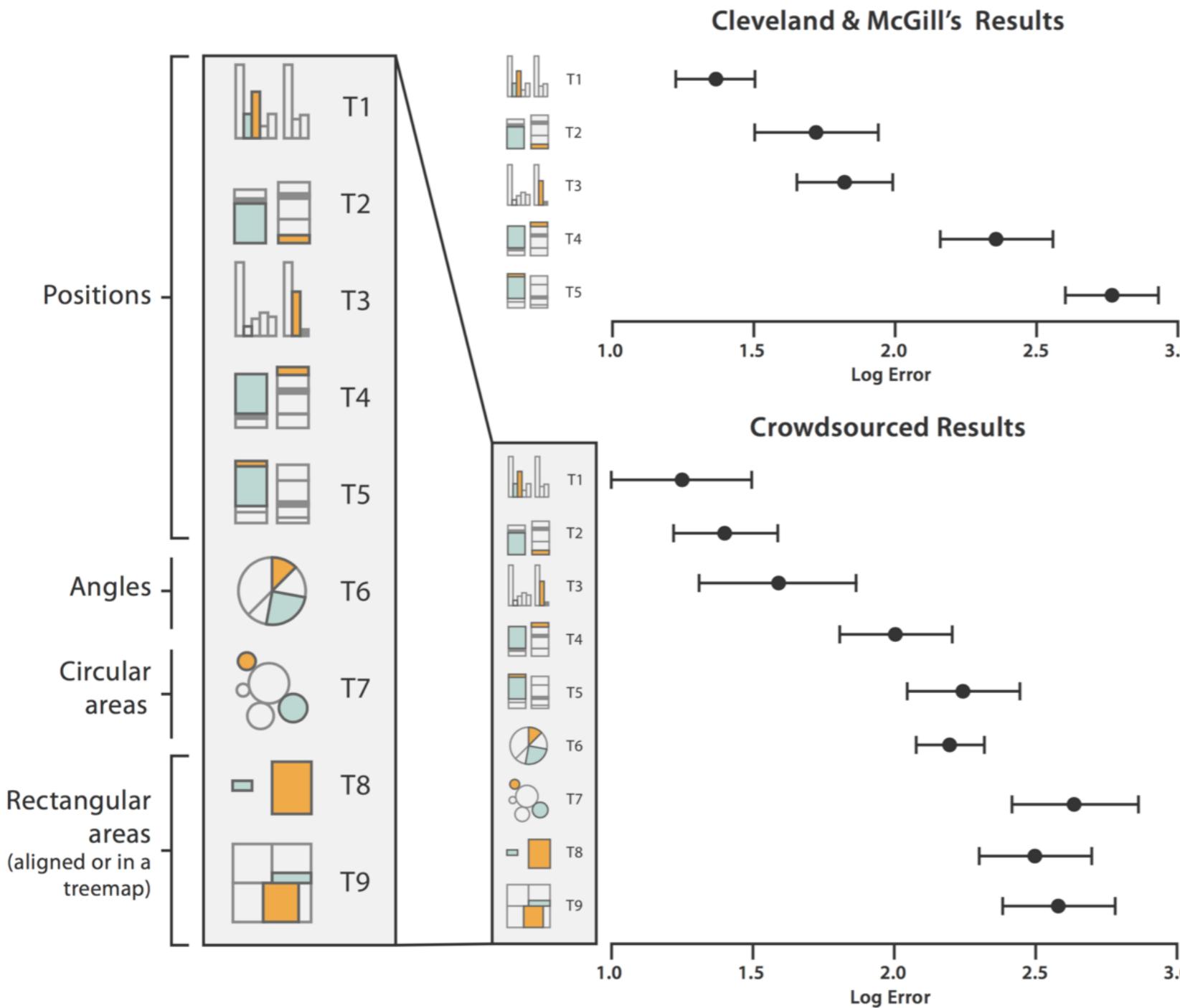
CHANNEL EFFECTIVENESS

Steven's Psychophysical power law

- ▶ Accuracy of perceptual judgement
 - ▶ small change in electric shock => perceived as large change (over estimate change)
 - ▶ change in physical stimuli => same change in perceived length (accurate estimation of change)
 - ▶ large change in brightness => perceived as small change (underestimate change)



CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION



- ▶ Note: Log scale (so circles and rectangular areas are much worse than position)

Figure 5.8. Error rates across visual channels, with recent crowdsourced results replicating and extending seminal work from Cleveland and McGill [Cleveland and McGill 84a]. After [Heer and Bostock 10, Figure 4].

CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION

- ▶ How much longer is B (Same, 2 x, 3x, 4x...)?



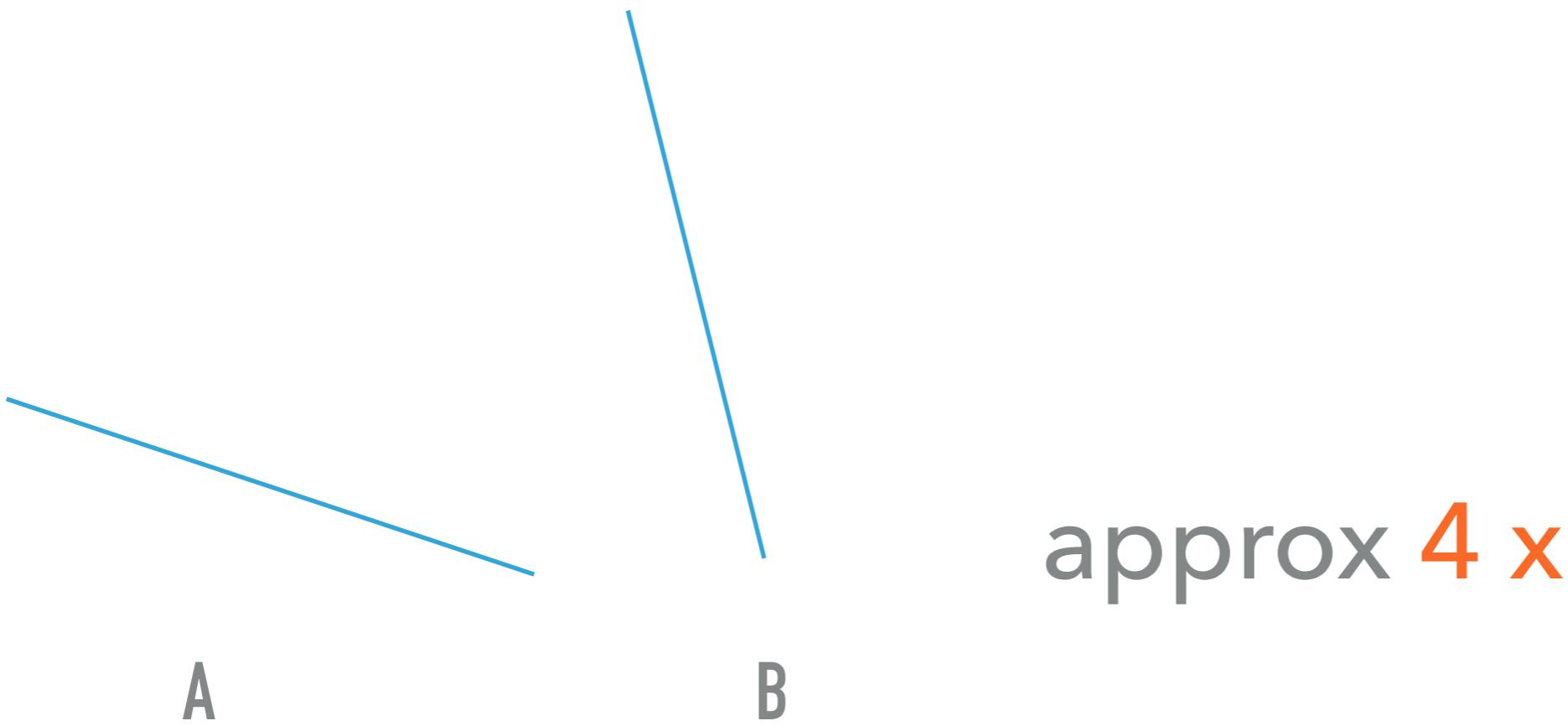
CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION

- ▶ How much longer is B?



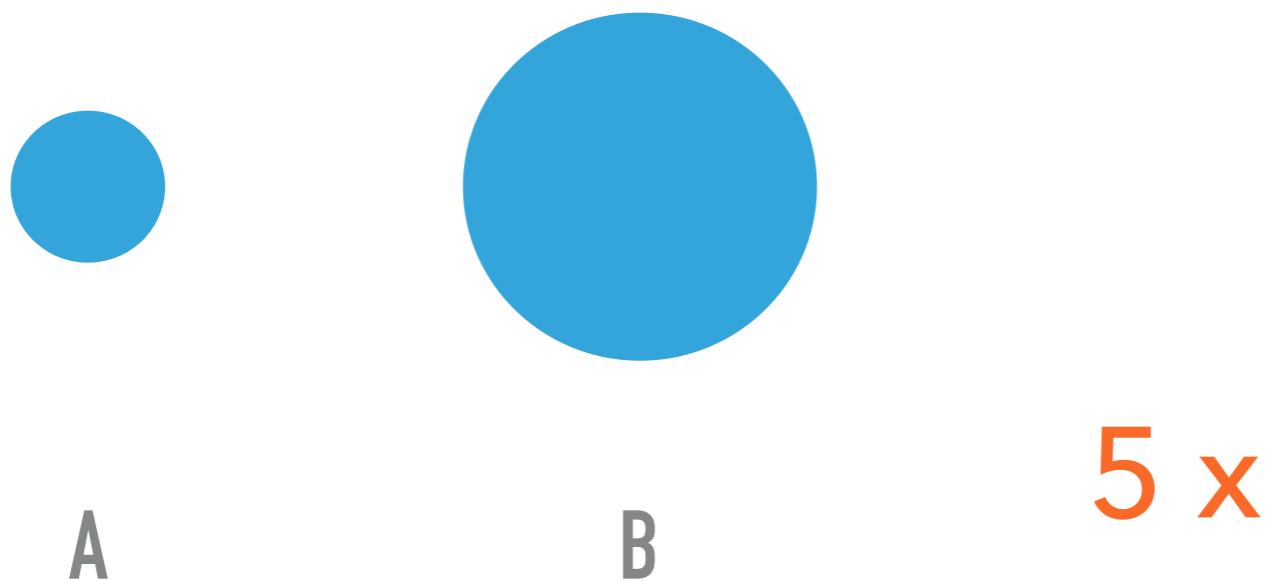
CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION

- ▶ How much steeper is B?

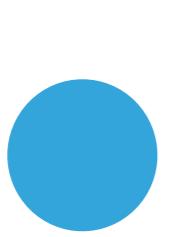


CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION

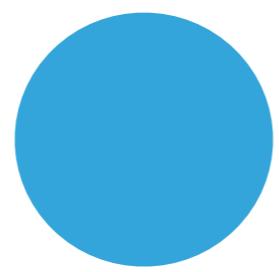
- ▶ How much larger is B?



CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION



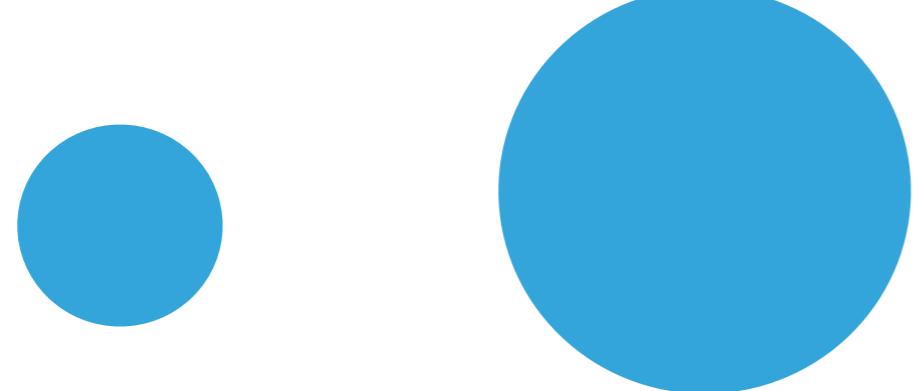
A



B

3 x area

CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION



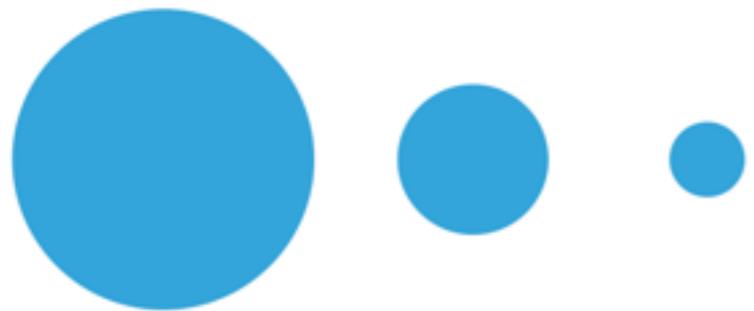
A

B

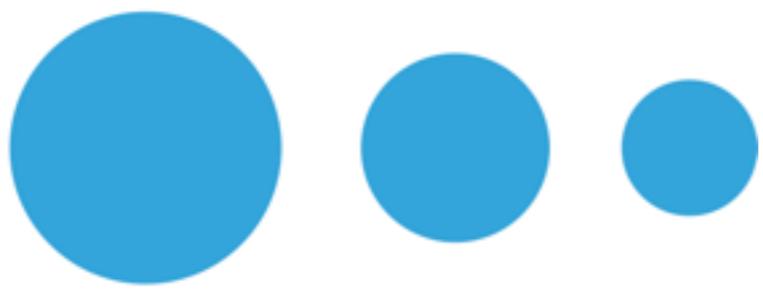
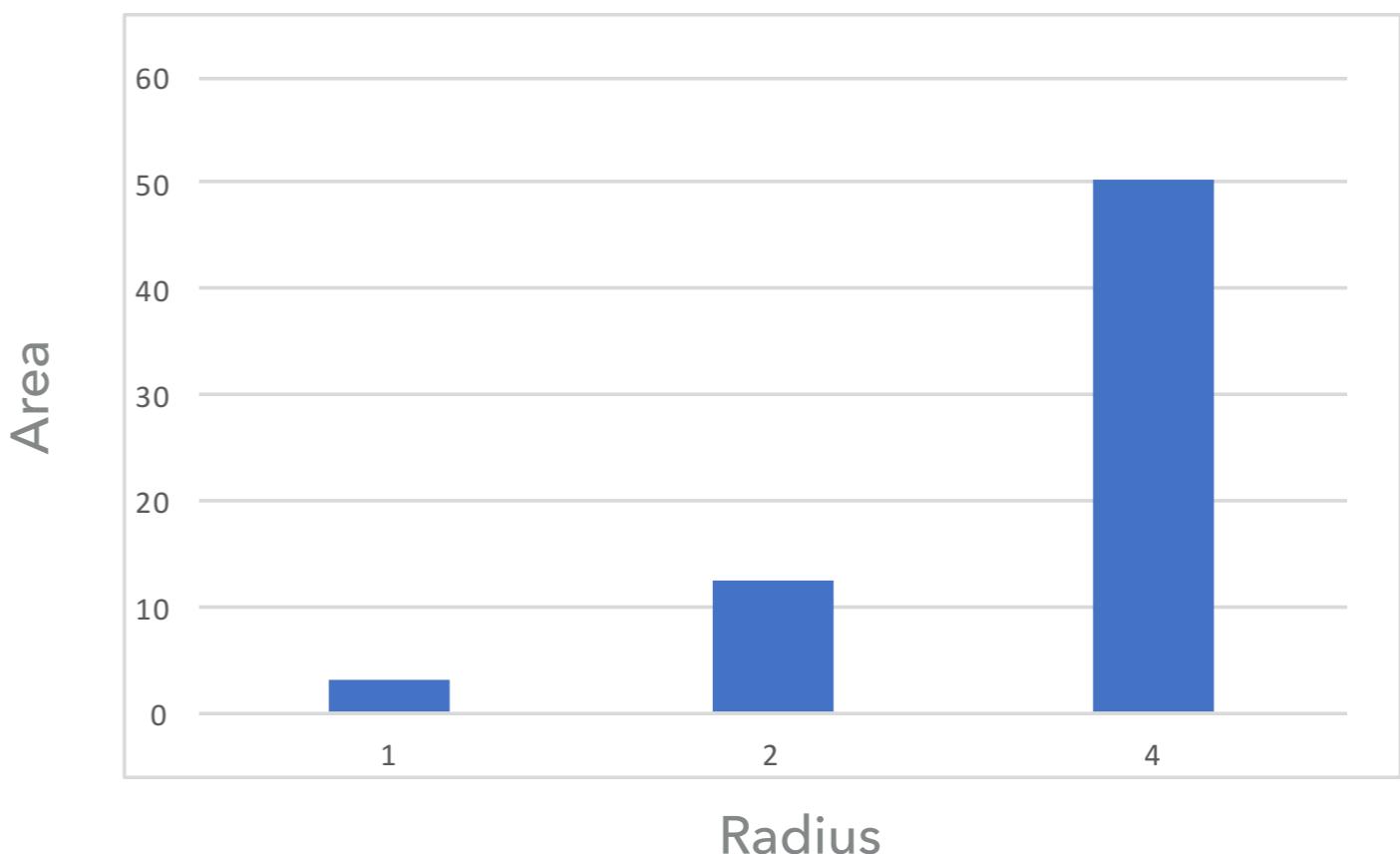
2 x diameter

4 x area ($\text{Area} = \pi r^2$)

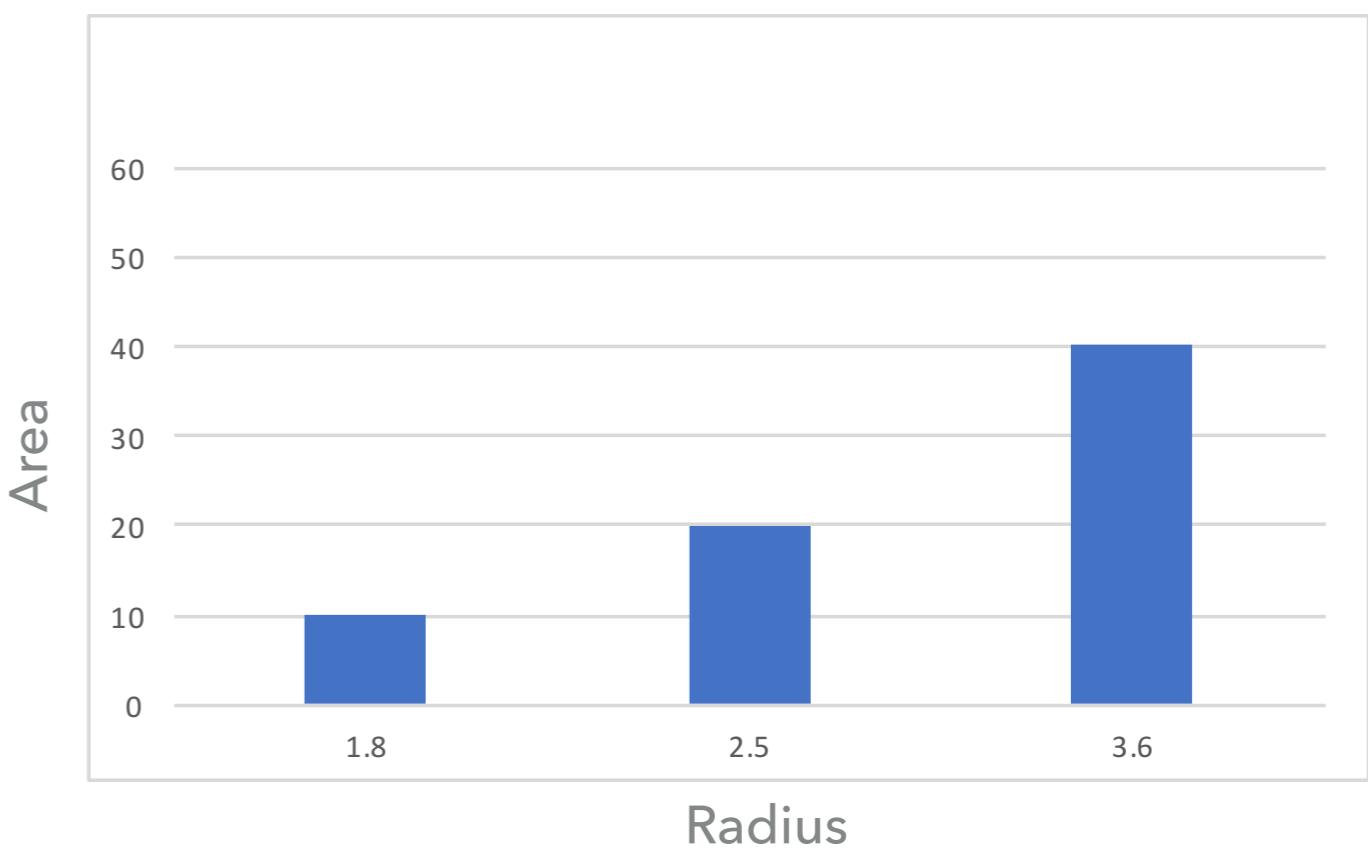
AREA vs RADIUS CODING



Double the radius



Double the area



CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION



2 x darker

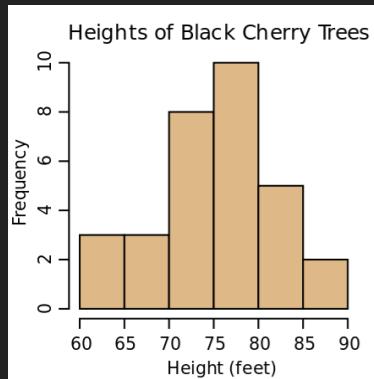
CHANNEL EFFECTIVENESS: ACCURACY OF PERCEPTION



3 x darker

CHANNEL DISCRIMINABILITY

- ▶ Can the user discriminate between items?
 - ▶ determined by the number of 'bins' available in a channel
 - ▶ line width (small number of bins)
 - ▶ line height (large number of bins)



Bin: a non-overlapping interval on a range of values

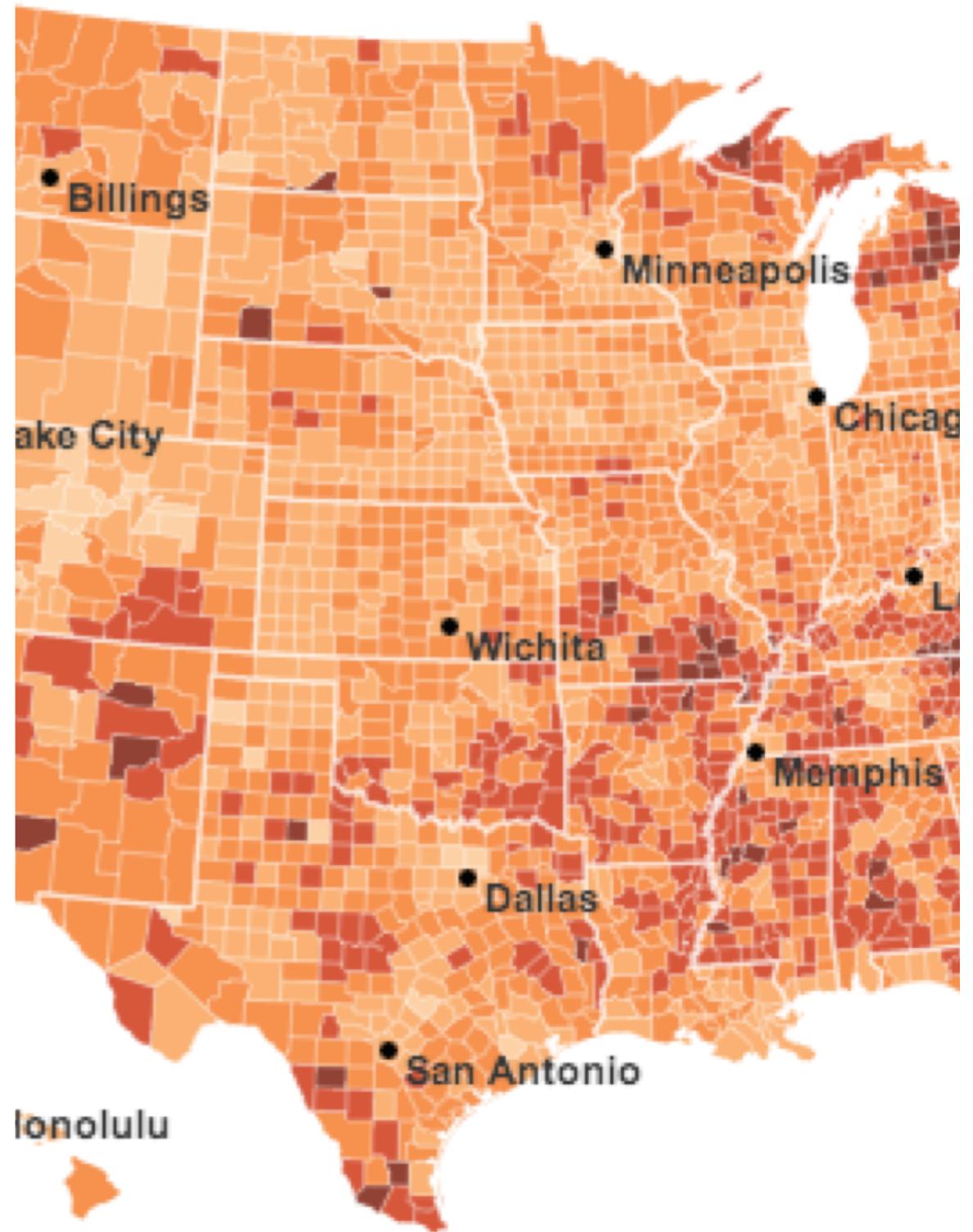
Bin size:
Route Total (millions of minutes):

- 0-100
- 101-250
- 251-500



MARKS AND CHANNELS ACTIVITY

- ▶ Learn how to decode charts using the language and rules of visual encoding



If the nature of the data suggests the shape of the graphic, follow that suggestion.

Edward Tufte

Statistician and data
visualisation guru





VISUAL VARIABLES: MARKS AND CHANNELS

What:

- ▶ marks and channels: graphic and visual elements for representing data

Why:

- ▶ different graphical and visual elements are better at representing different types of data, choose the right elements for your data

How:

- ▶ understand the pros and cons of using different graphical elements for different types of data