Visual Encoding Principles

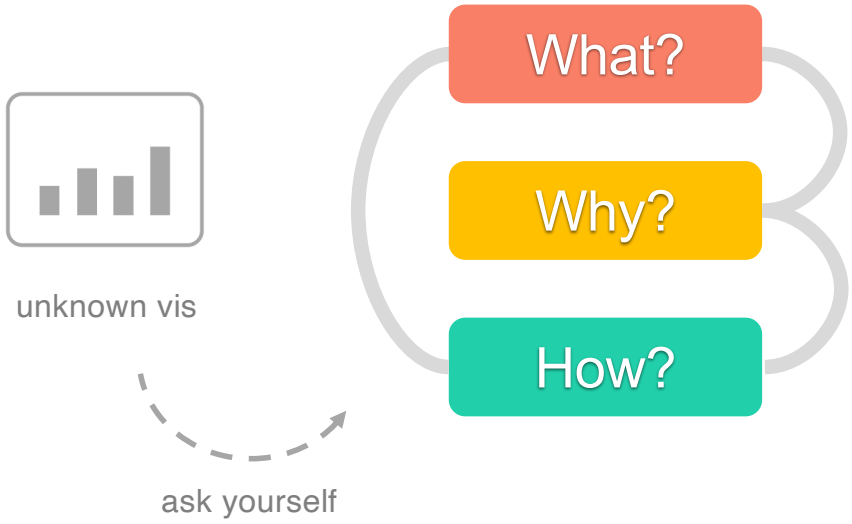
Torsten Möller

Readings

- Munzner, “Visualization Analysis and Design”:
  - Chapter 5 (Marks and Channels)
  - Chapter 10 (Map Color and Other Channels)
- Ware, “Information Visualization – Perception for Design”:
  - Chapter 4 (Color)
  - Chapter 5 (Visual Salience and Finding Information)
- Ward, Grinstein and Keim, “Interactive Data Visualization”:
  - Chapter 3 (Human Perception and Information Processing)
  - Chapter 4 (Visualization Foundations)
- Mackinlay, “Automating the design of graphical presentations of relational information”, ACM ToG, 5(2), 110-141, 1986, doi: 10.1145/22949.22950

Three-part analysis framework

to analyze any existing visualization



What data is shown?

Why is the task being performed?

How is the visualization constructed?

Munzner, Visualization Analysis and Design, 2014, p. 17 ff.

Three-part analysis framework

to analyze any existing visualization

What?

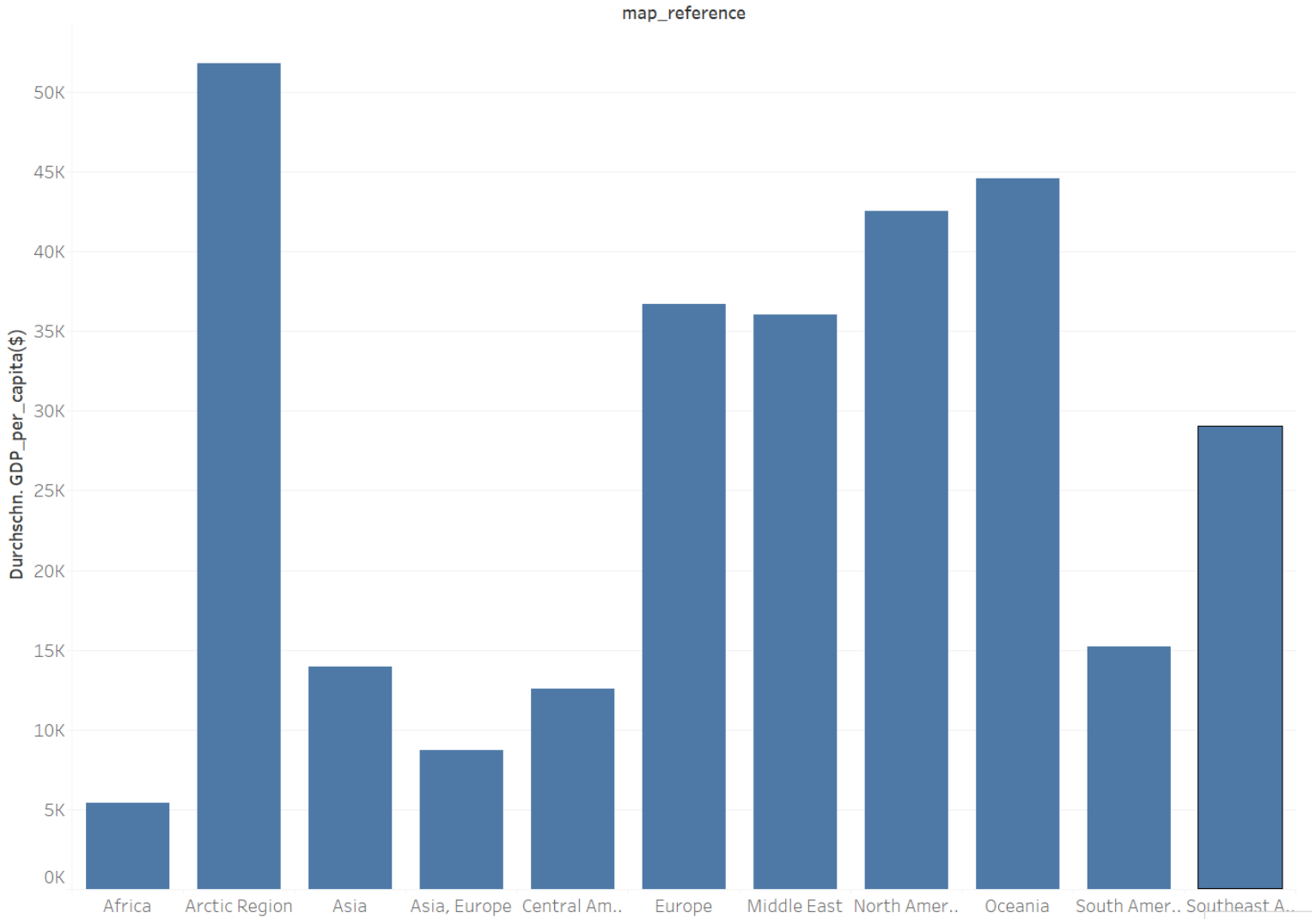
Why?

How?

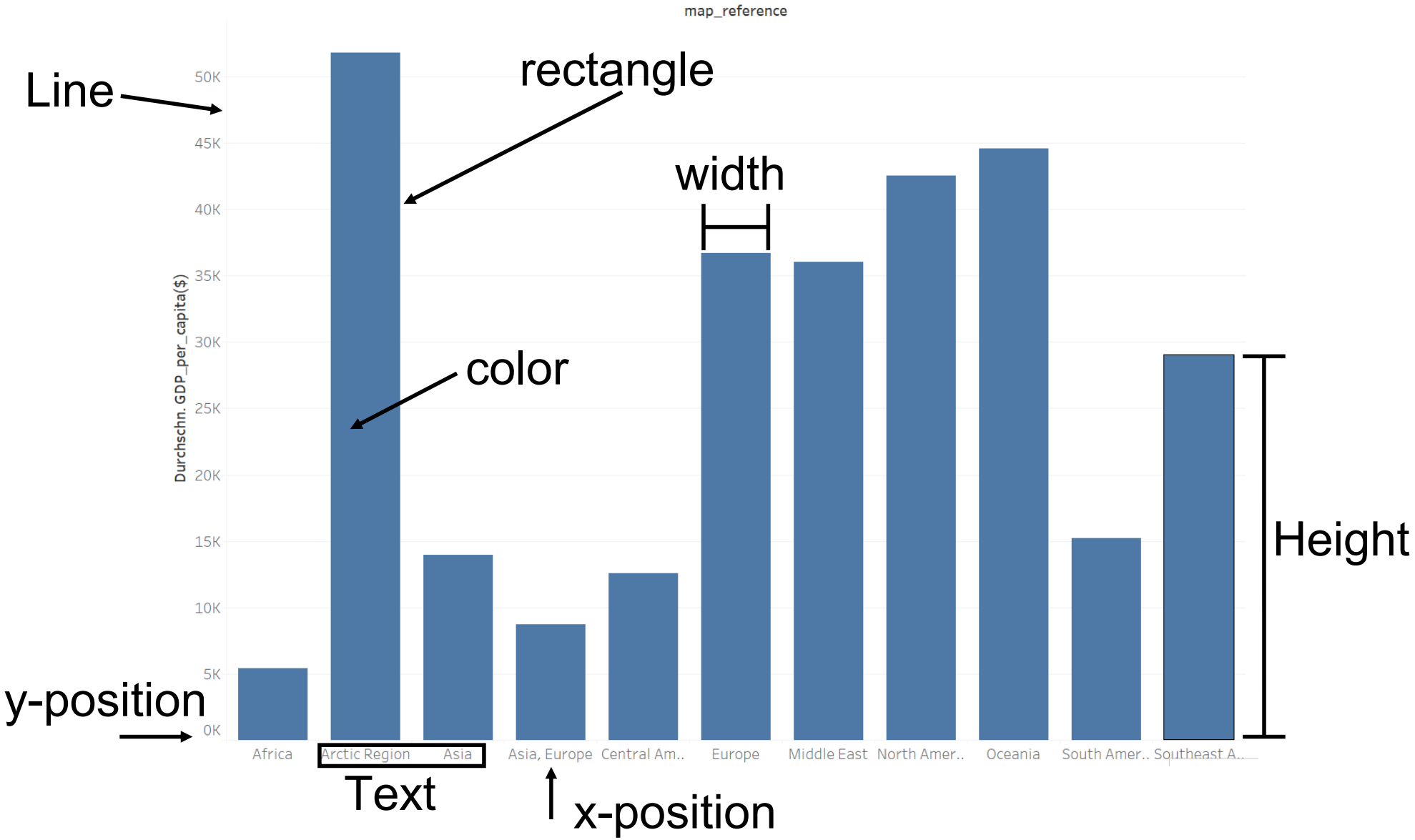
- What? – Data Abstraction
- Why? – Task Abstraction
- How? – Visual Encoding
  - Marks and channels
    Examples, Channel types, Mark types, Expressiveness + Effectiveness
  - Channel effectiveness
    Accuracy, Discriminability, Separability, Popout
  - Channel characteristics
    Spatial position, Relative vs. absolute, Color (Color theory basics, Perception issues & deficiencies, Categorical, Ordinal)

Marks and channels

Parts of a visualization



Parts of a visualization



Visual language is a sign system

- Image perceived as a set of signs
- Sender encodes information in signs
- Receiver decodes information from signs

- Jacques Bertin
  - French cartographer [1918-2010]
  - Semiology of Graphics [1967]
  - Theoretical principles for visual encodings

Bertin, Semiology of Graphics, 1983

According to Bertin ...

Marks

Points Lines Areas

Channels

Position

Size

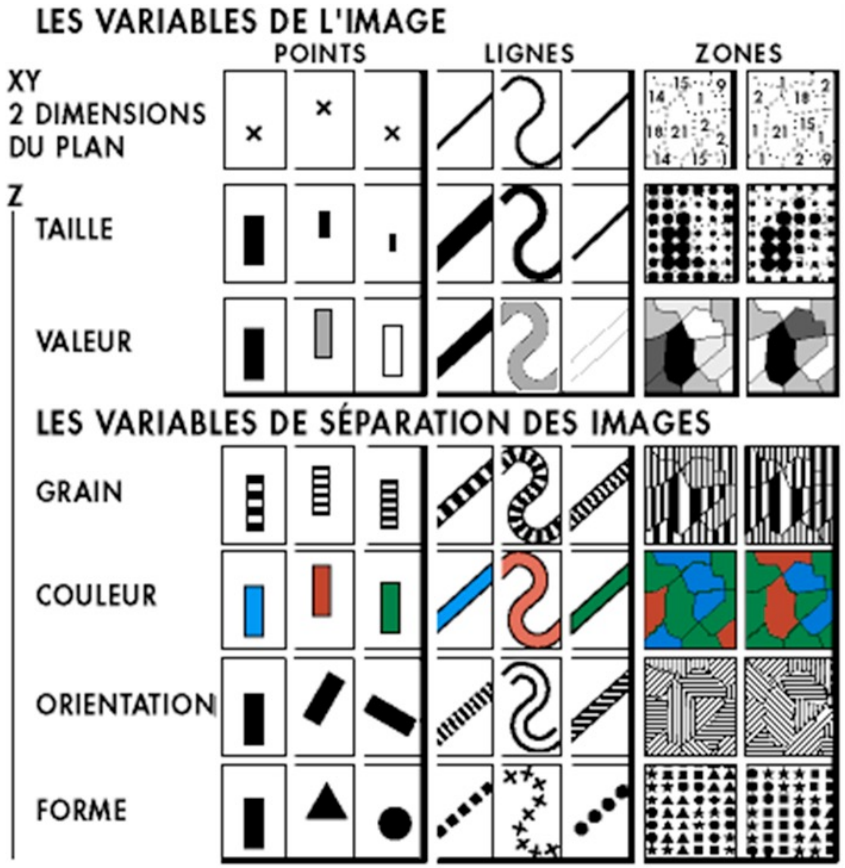
(Grey)Value

Texture

Color

Orientation

Shape

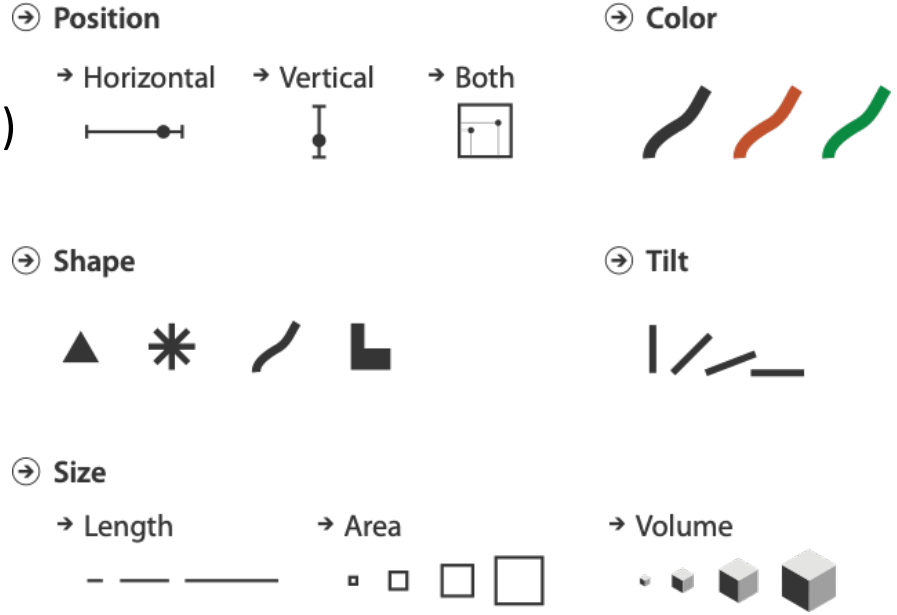


Bertin, Semiology of Graphics, 1983

Marks and Channels (Munzner)

- Mark: basic graphical element / geometric primitive:
  - point (0D)
  - line (1D)
  - area (2D)
  - volume (3D)

- Channel: control appearance (of a mark)
  - position
  - color (hue, saturation, luminance)
  - shape
  - tilt / angle
  - size (length, area, volume)
  - curvature
  - motion
  - etc.



Munzner, Visualization Analysis and Design, 2014, p. 96

Marks and Channels in Polaris

(predecessor to Tableau)

property	marks	ordinal/nominal mapping	quantitative mapping
shape	glyph	○ □ + △ S U	
size	rectangle, circle, glyph, text		
orientation	rectangle, line, text		
color	rectangle, circle, line, glyph, y-bar, x-bar, text, gantt bar		

Stolte and Hanrahan, Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases, 2000

Marks and Channels:

Tableau example

- Let us now use different visual encodings
- Dataset:
  - World Happiness dataset (description)
    composed from multiple sources:
    - The World Happiness Report: scores based on answers to a life evaluation question, where respondents rate the quality of their current lives on a scale from 0 to 10.
    - The CIA Factbook contains almanac style information about the countries of the world, such as population, GDP family income, internet access, cell phone subscriptions, etc.
    - The Human Development Index
    - The Inequality-adjusted Human Development Index
- Create new chart(s) using the following fields: Country, GDP per capita, Internet Access, Happiness Score, Generosity

Marks and Channels:

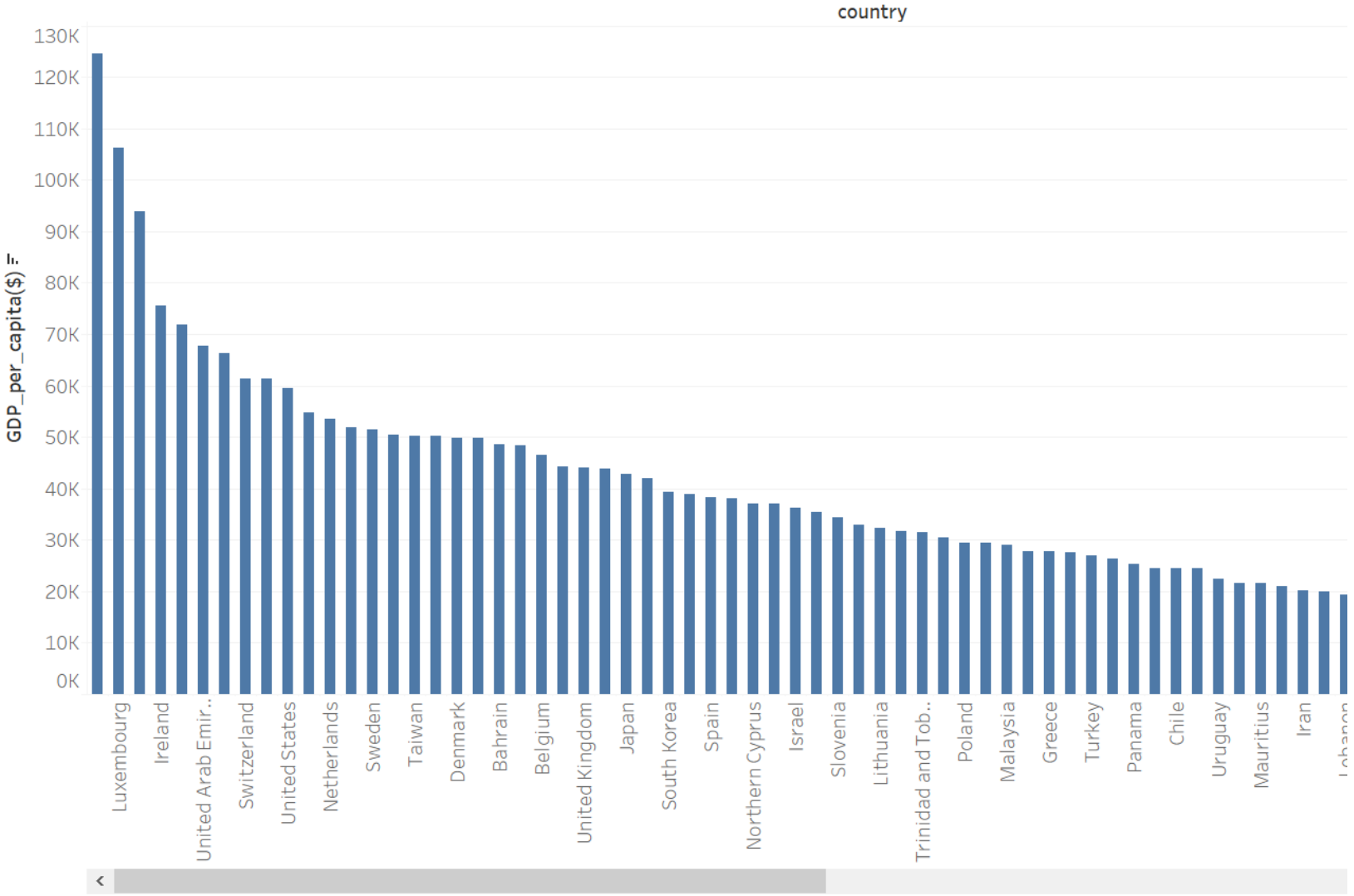
Tableau example

- Use a rectangle mark, length channel for GDP of each country
- Use point marks and position channels for **GDP** and **Internet Access.** Is there a correlation (visually) between the two?
- Which countries have low **Happiness** but a high **GDP**?
- Which countries have low **GDP**, but still high **Generosity**?

Marks and Channels:

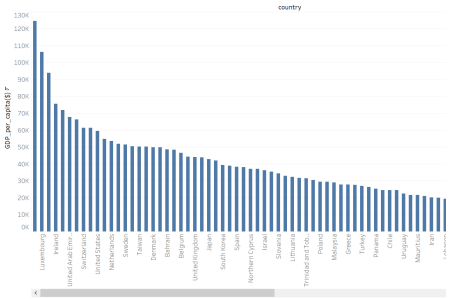
Tableau example

GDP per capita for all countries?

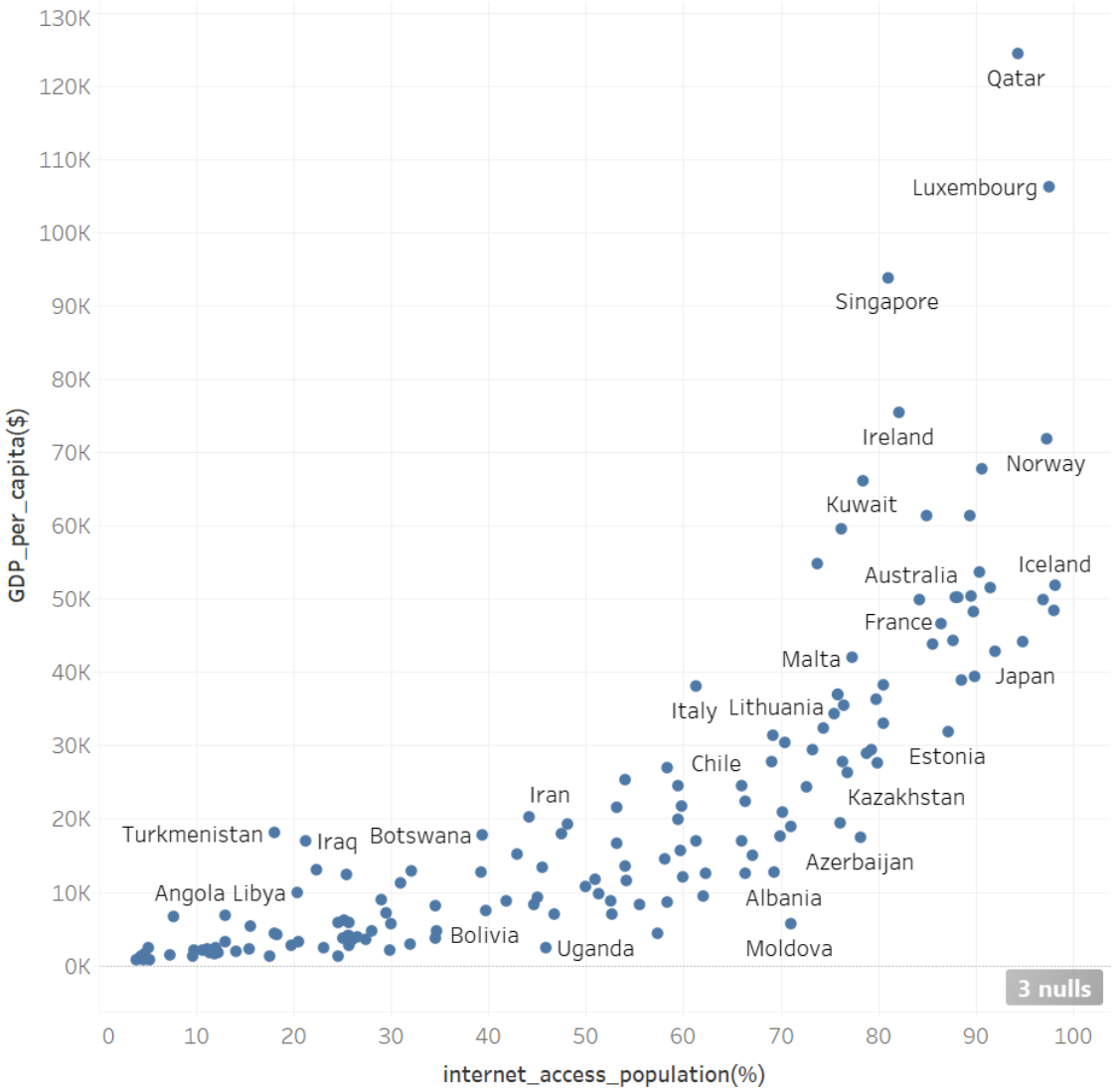


Marks and Channels:

Tableau example

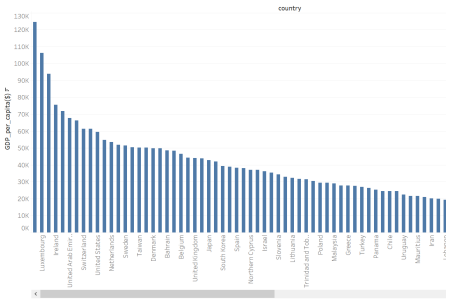


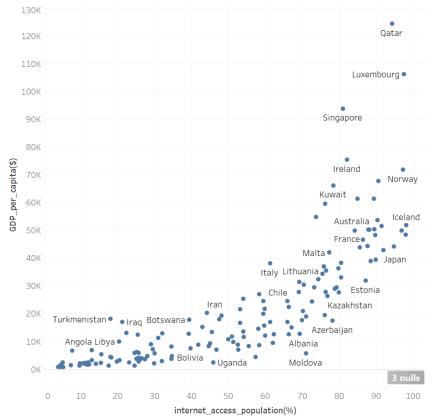
Internet access vs. GDP?



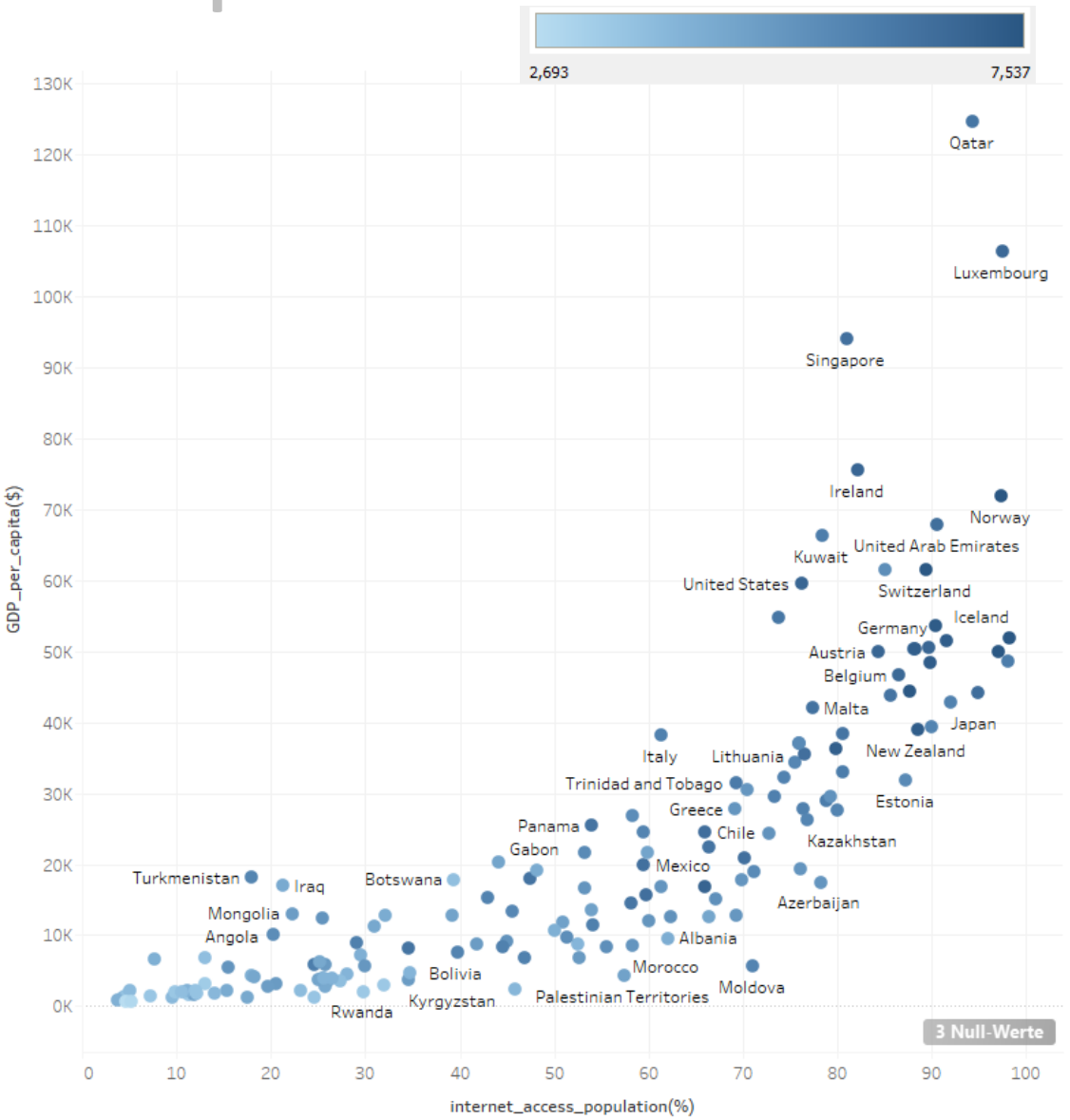
Marks and Channels:

Tableau example





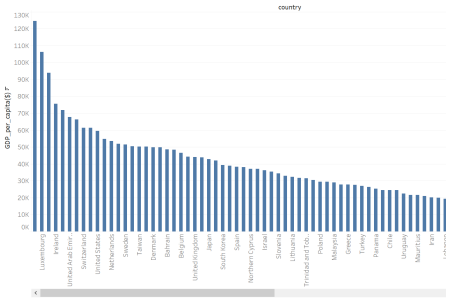
Low Happiness Score, high GDP?

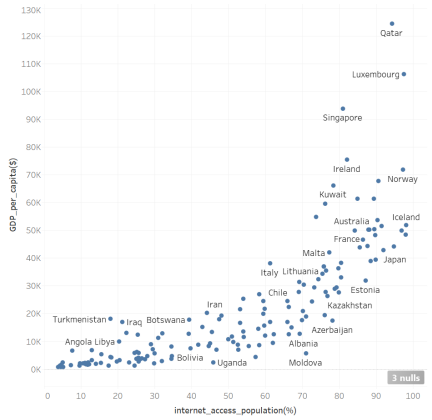


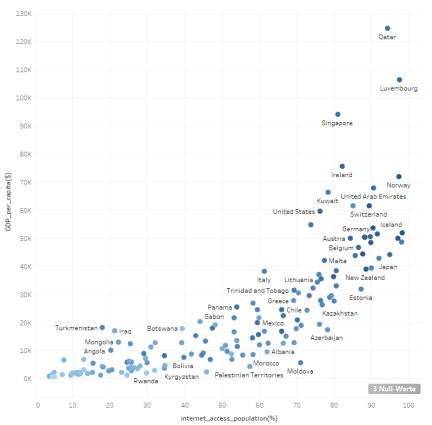
Marks and Channels:

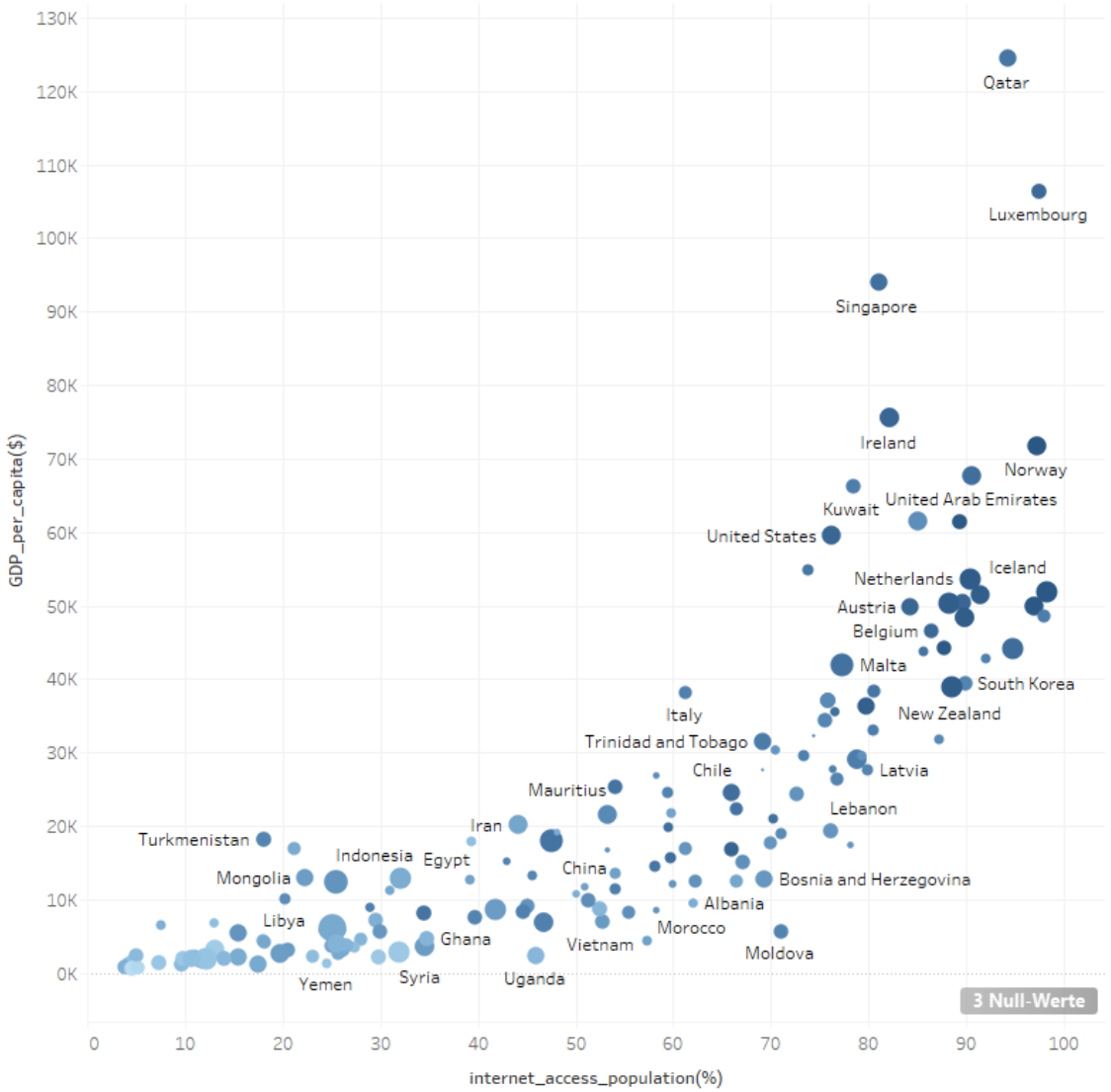
Tableau example

Low GDP, high Generosity?









Marks and Channels:

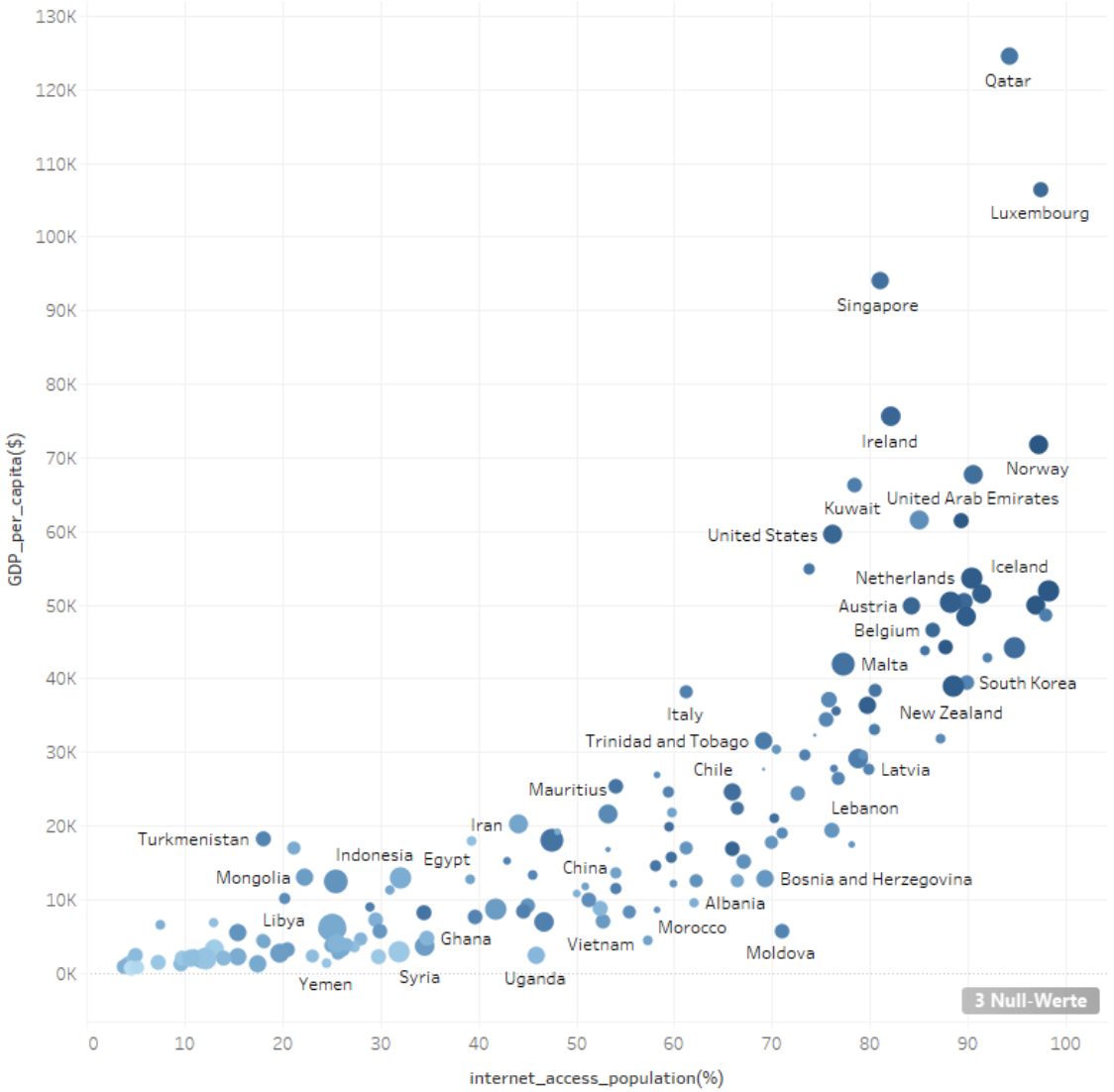
Tableau example

Marks:

- Circles - Countries

Channels:

- Rows (position along y axis): GDP per Capita ($)
- Columns (position along x axis): Internet access %
- Color (saturation/ luminance): Happiness score
- Size: Generosity



Channel types

- *What*: **categorical**

- *How Much*: **ordered** (ordinal, quantitative)

Channel types

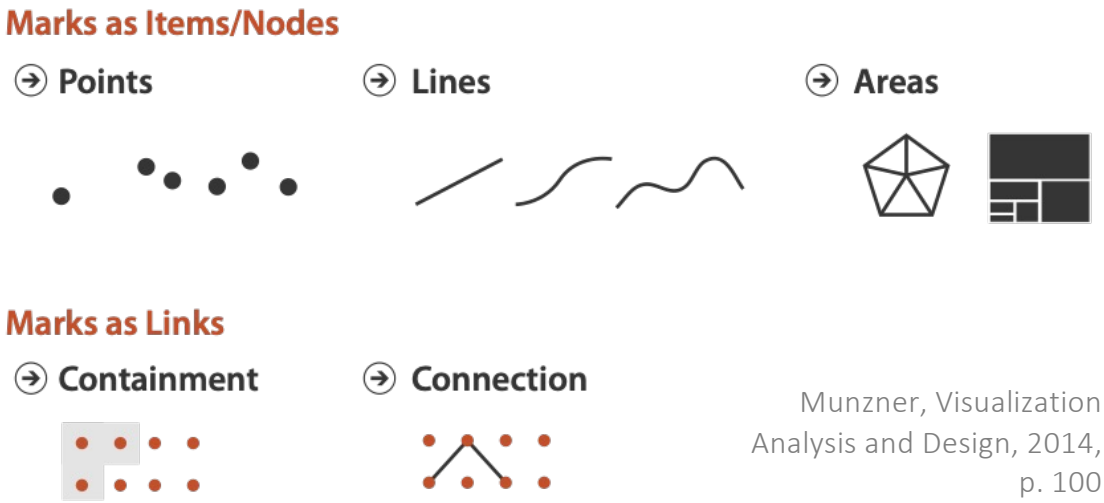
- *What*: **categorical**
  - shape
  - spatial region
  - color (hue)
- *How Much*: **ordered** (ordinal, quantitative)

Channel types

- *What*: **categorical**
  - shape
  - spatial region
  - color (hue)
- *How Much*: **ordered** (ordinal, quantitative)
  - length (1D)
  - area (2D)
  - volume (3D)
  - tilt
  - position
  - color (lightness)

Mark types

- Tables: item = point
- Network: node+link
- Link types:
  - **Connection**: relationship between two nodes
  - **Containment**: hierarchy

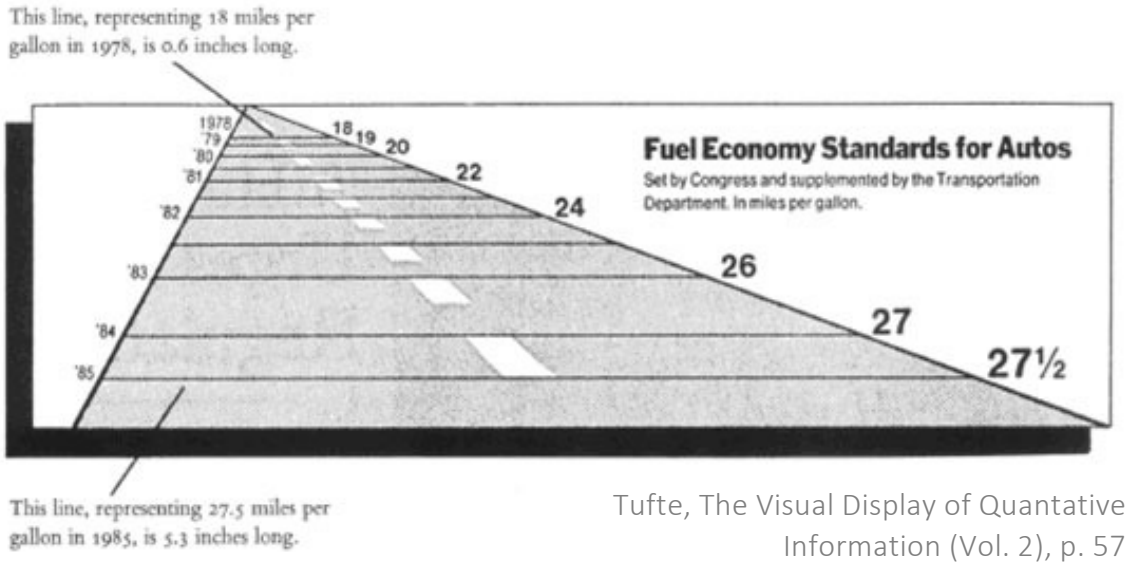


Munzner, Visualization Analysis and Design, 2014, p. 100

Expressiveness + Effectiveness

Expressiveness principle:

- Visual encoding should express **all of, and only**, the information in the dataset attributes
- Lie factor

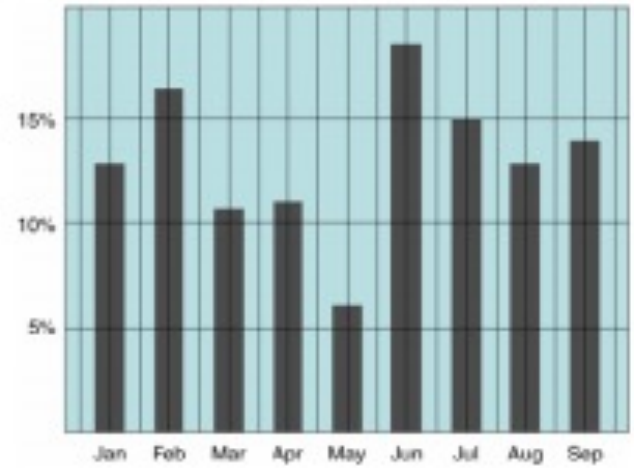


Tufte, The Visual Display of Quantative Information (Vol. 2), p. 57

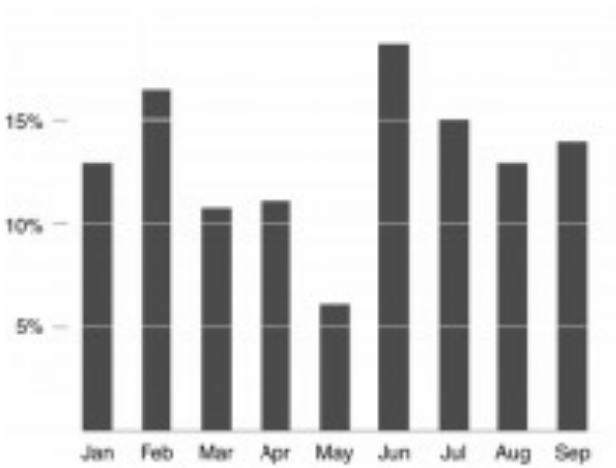
Expressiveness + Effectiveness

Effectiveness principle:

- Importance of the attribute should match the salience of the channel
- Data-ink ratio



Low Data/Ink



High Data/Ink

Channel effectiveness

Which channel is most accurate?*

- Color
- Shape
- Curvature
- Area
- Motion
- Length
- Angle
- Position
- Depth
- Volume

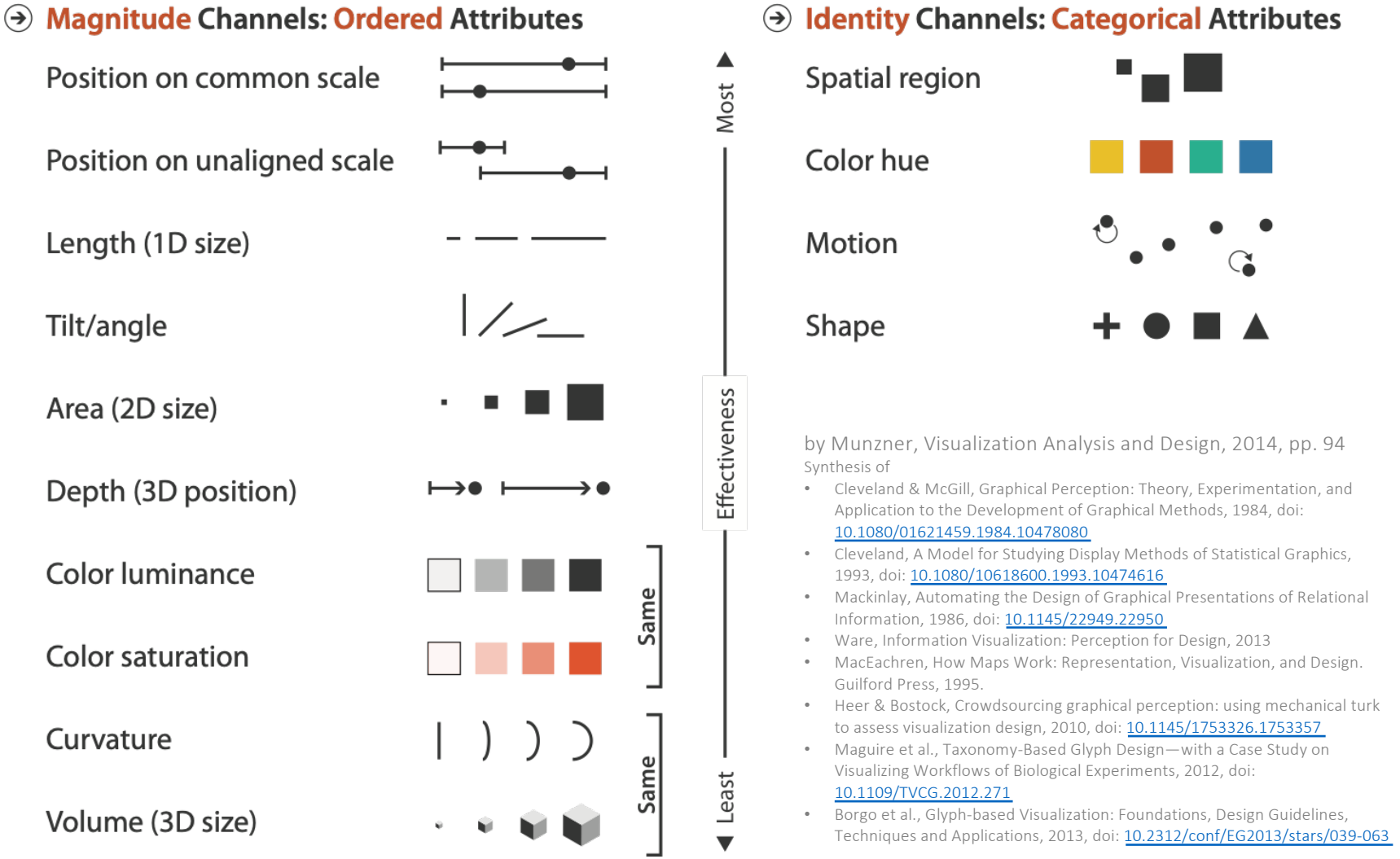
* for the task of comparing two values

Which channel is most accurate?*

* for the task of comparing two values

- Color
- Shape
- Curvature
- Area
- Motion
- **Length** ← Ordered attributes
- Angle
- **Position** ← Ordered attributes
- Depth
- Volume

Effectiveness ranking



by Munzner, Visualization Analysis and Design, 2014, pp. 94

Synthesis of

- Cleveland & McGill, Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods, 1984, doi: 10.1080/01621459.1984.10478080
- Cleveland, A Model for Studying Display Methods of Statistical Graphics, 1993, doi: 10.1080/10618600.1993.10474616
- Mackinlay, Automating the Design of Graphical Presentations of Relational Information, 1986, doi: 10.1145/22949.22950
- Ware, Information Visualization: Perception for Design, 2013
- MacEachren, How Maps Work: Representation, Visualization, and Design. Guilford Press, 1995.
- Heer & Bostock, Crowdsourcing graphical perception: using mechanical turk to assess visualization design, 2010, doi: 10.1145/1753326.1753357
- Maguire et al., Taxonomy-Based Glyph Design—with a Case Study on Visualizing Workflows of Biological Experiments, 2012, doi: 10.1109/TVCG.2012.271
- Borgo et al., Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications, 2013, doi: 10.2312/conf/EG2013/stars/039-063

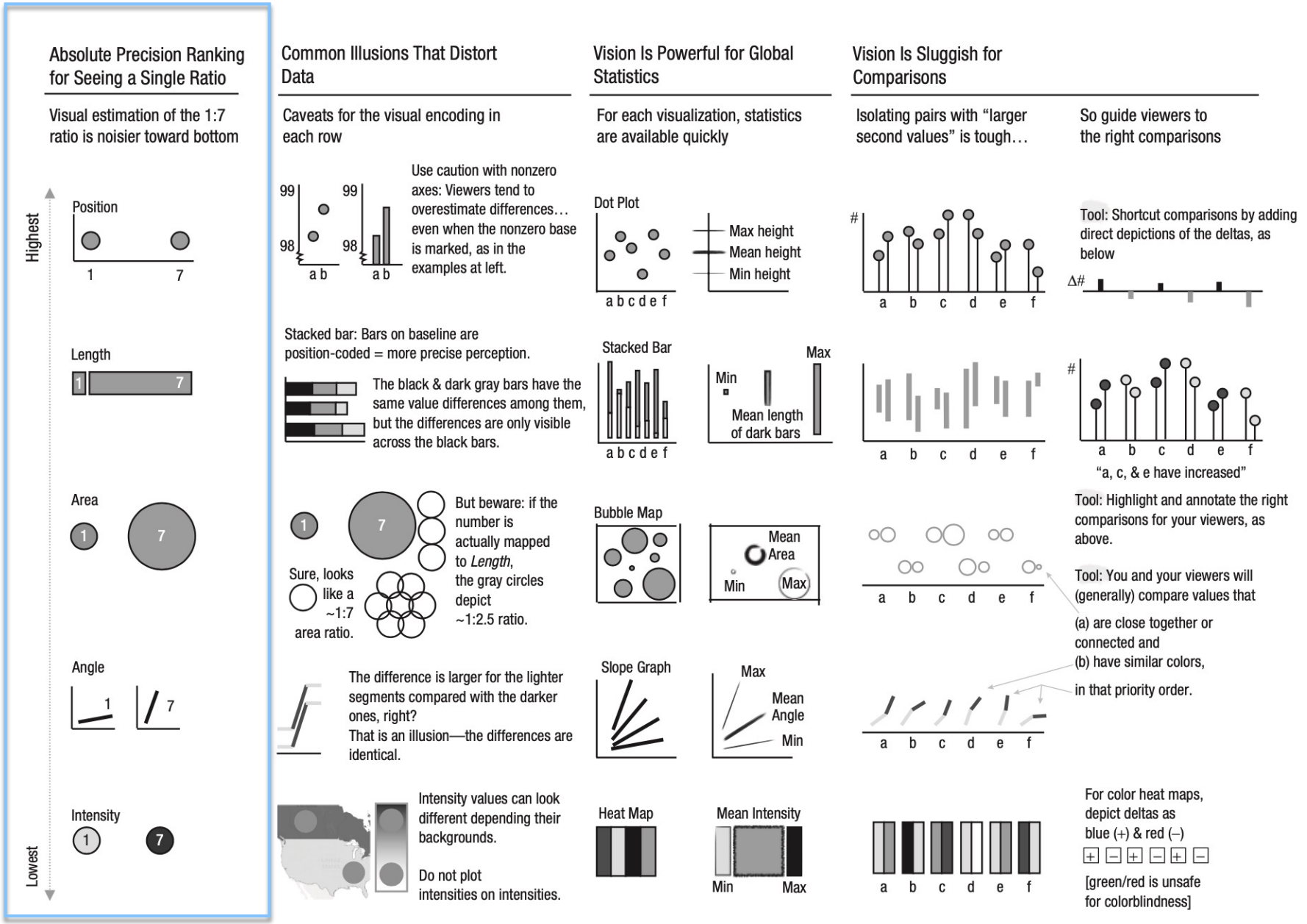
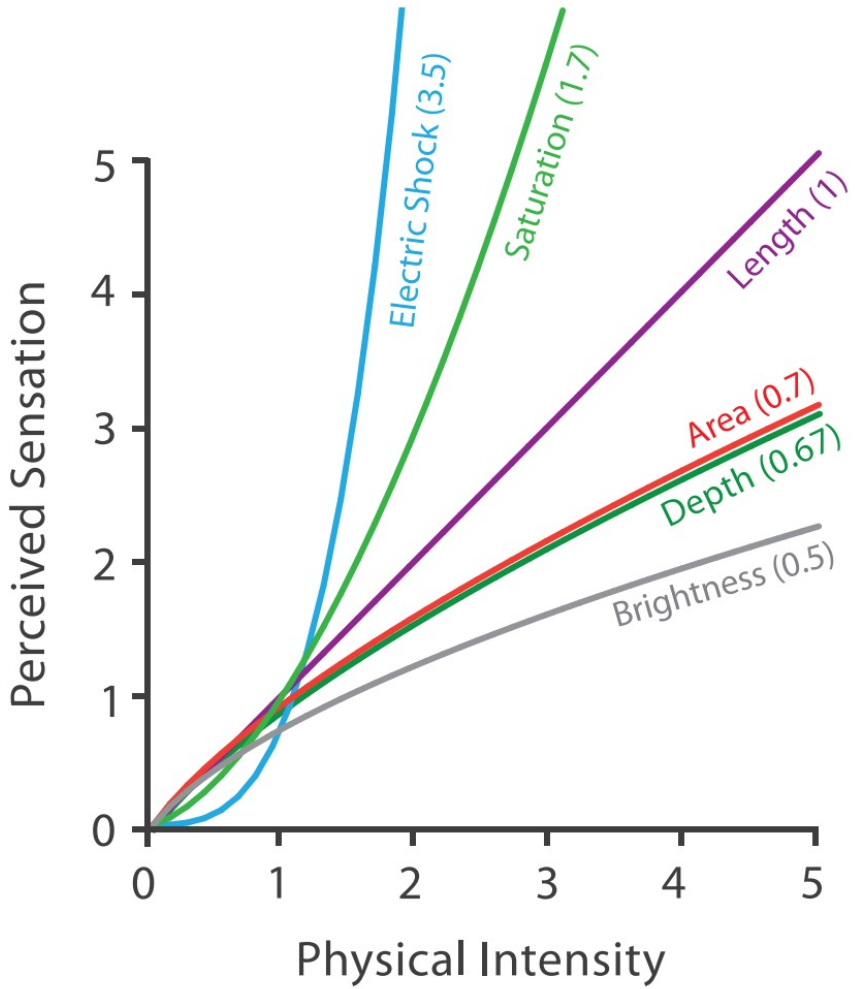


Fig. 2. A summary of the power and limits of visual data processing. The two columns on the left show a quick reference guide to channels that can depict data visually and common illusions for each channel. The column in the center presents a summary of how visual statistics are powerful. The two columns on the right illustrate how comparisons are severely limited and present a set of design techniques that focus viewers on the "right" ones.

From Franconeri et al., 2021. The Science of Visual Data Communication: What Works.

Effectiveness: Accuracy

- perceptual judgement vs. stimulus
- Weber's law: $S = I^n$



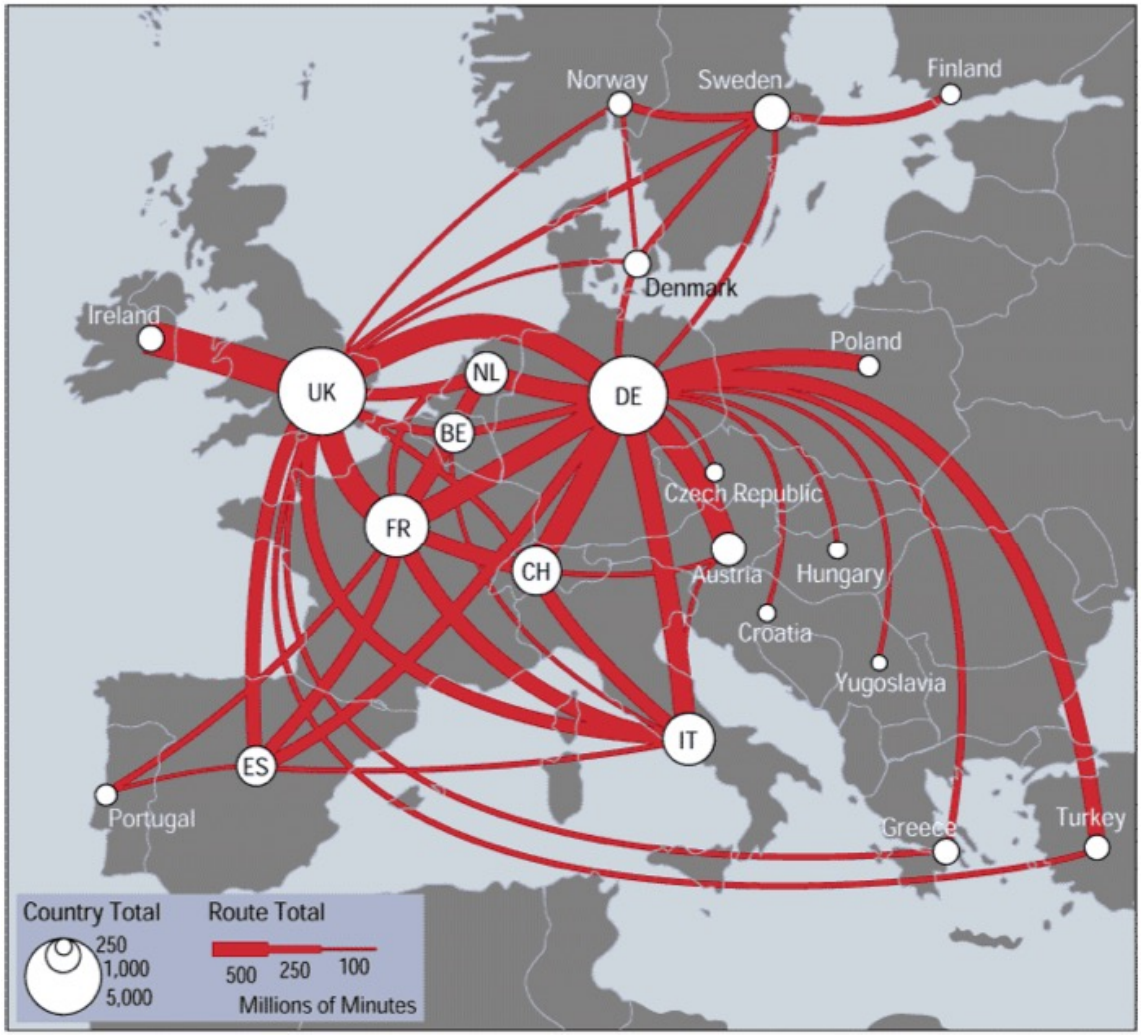
Data: Stevens, Psychophysics: introduction to its perceptual, neural, and social prospects, 1975
Figure: Munzner, Visualization Analysis and Design, 2014, p. 104 ff.

Effectiveness: Discriminability

- How many colors can I tell apart?
- How many levels of grey etc.
- Example: line width

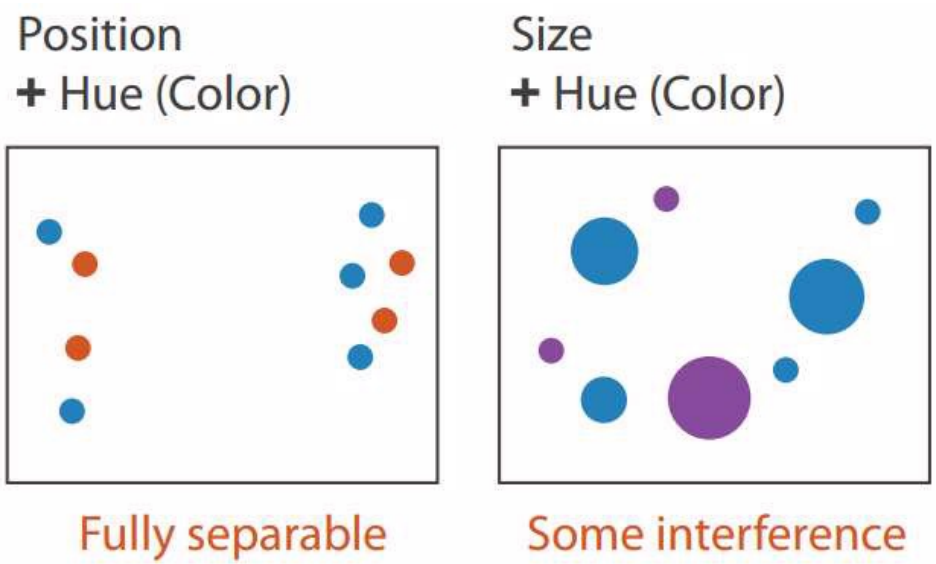
Effectiveness: Discriminability

- How many colors can I tell apart?
- How many levels of grey etc.
- Example: line width



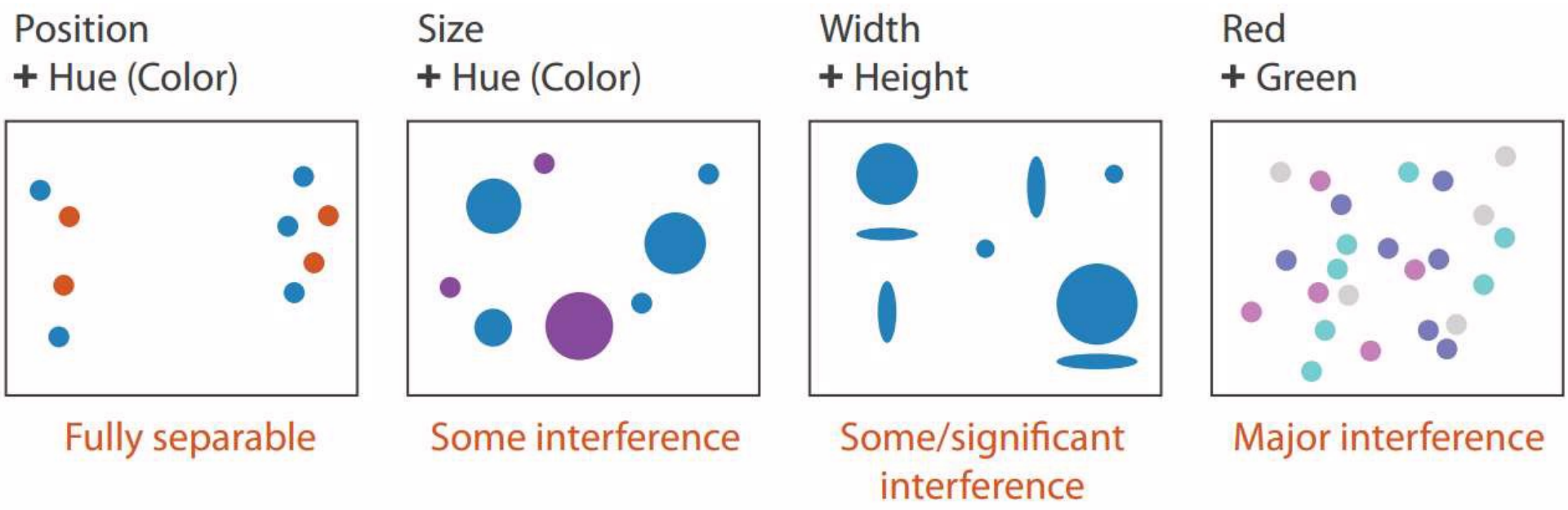
Effectiveness: Separability

- Separable vs. integral channels



Effectiveness: Separability

- Separable vs. integral channels



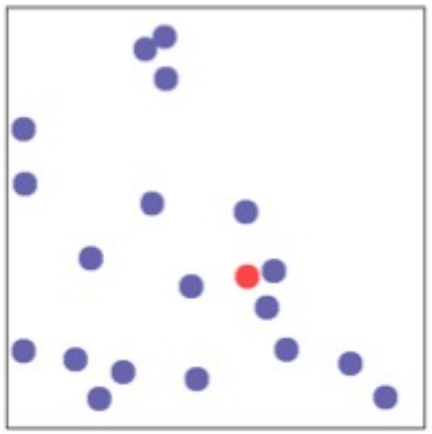
Effectiveness: Popout

Preattentive processing

Find the red dot.

Effectiveness: Popout

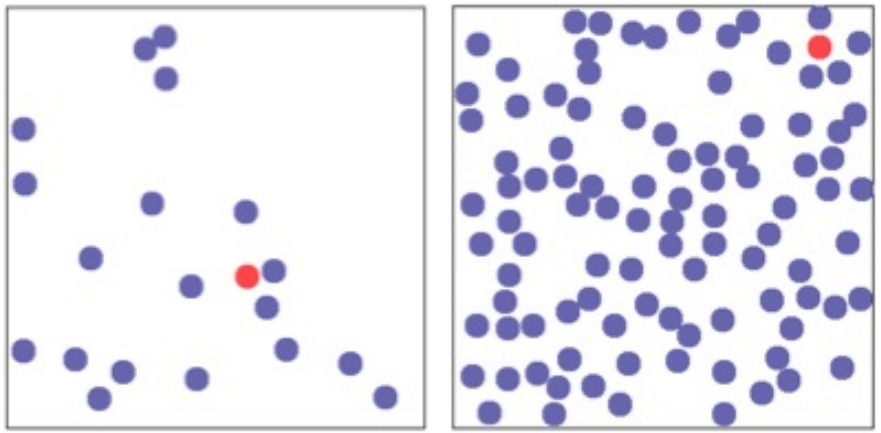
Preattentive processing



Find the red dot.

Effectiveness: Popout

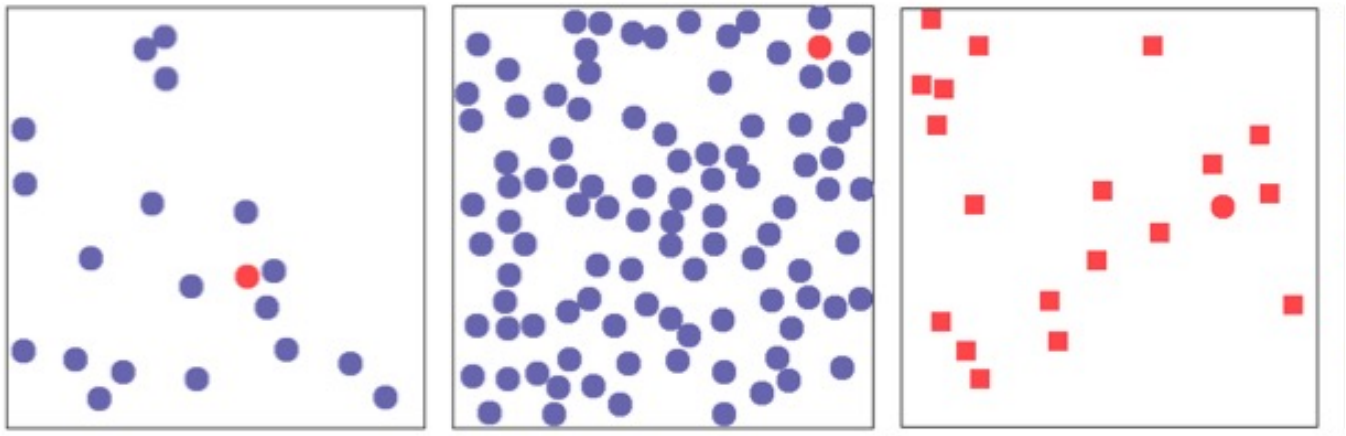
Preattentive processing



Find the red dot.

Effectiveness: Popout

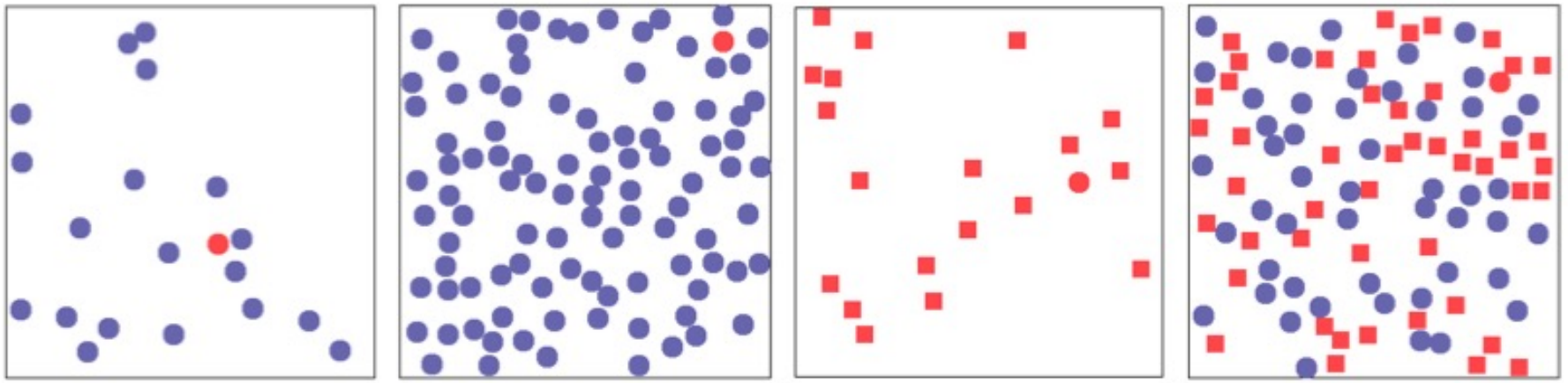
Preattentive processing



Find the red dot.

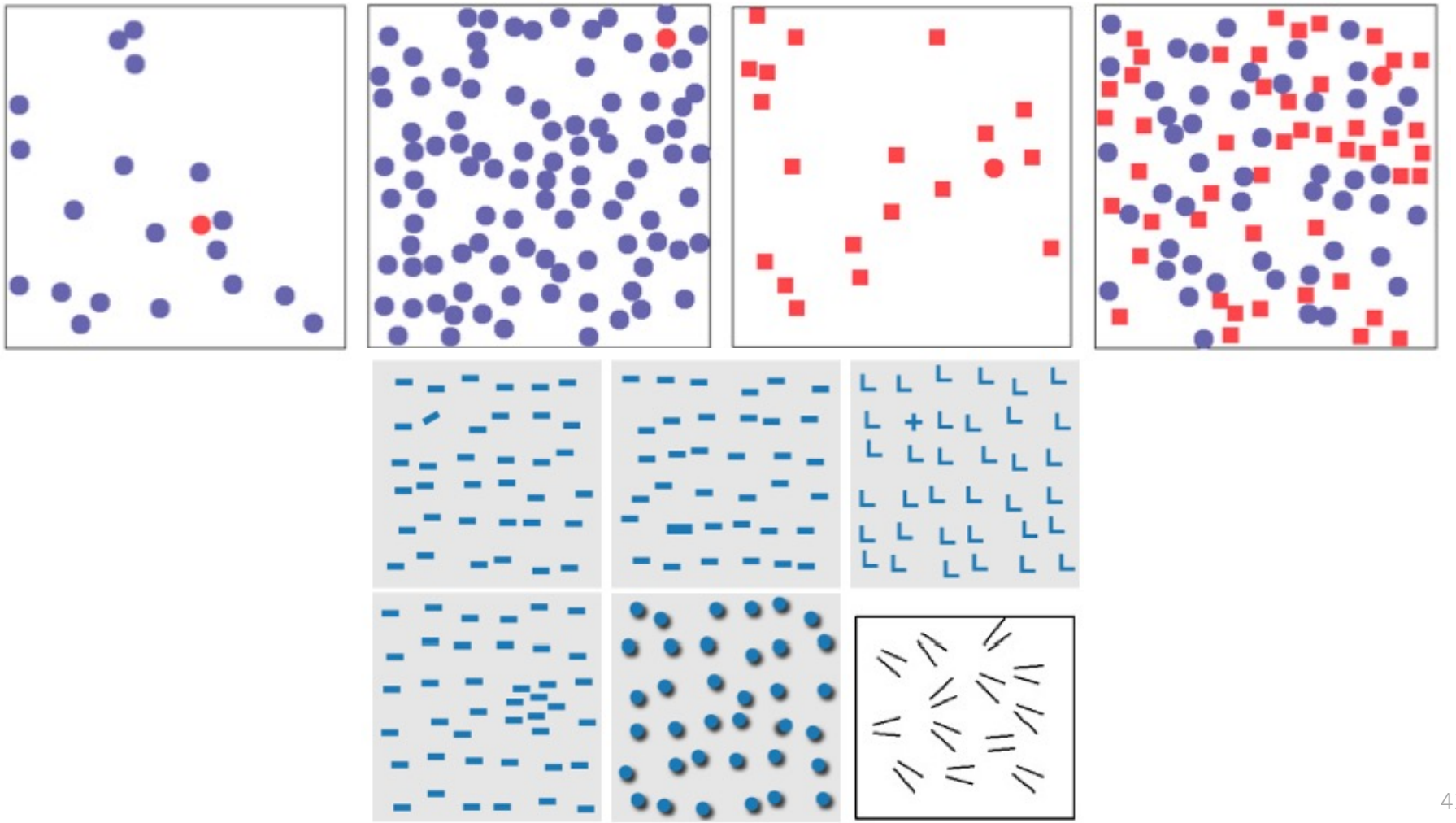
Effectiveness: Popout

Preattentive processing



Effectiveness: Popout

Preattentive processing

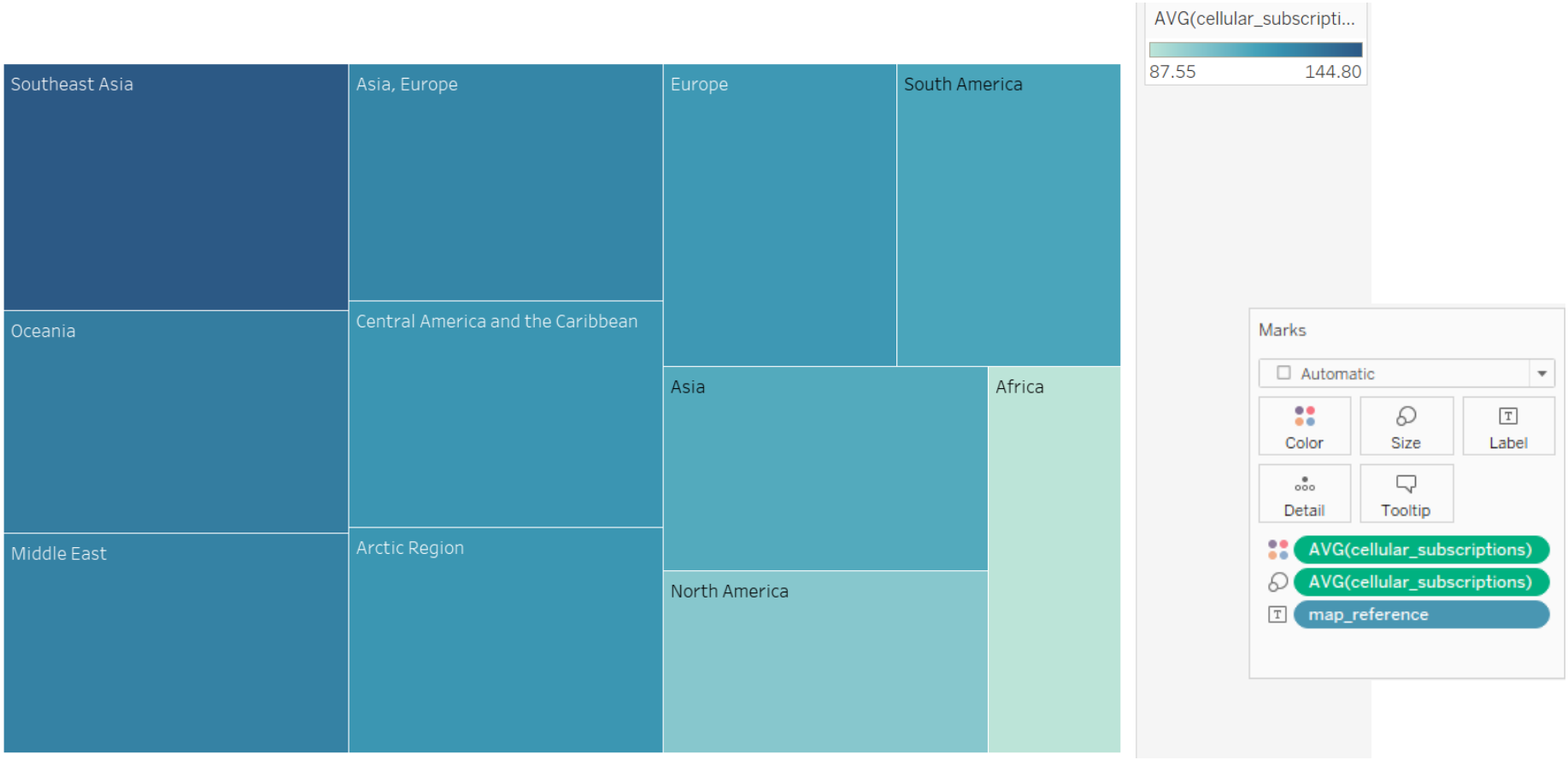


Effectiveness:

Tableau example

- Let us test such difference by comparing area and length as channels for a quantitative value
- Let's look at the **Cellular Subscriptions** per continent (**Map Reference**)
- We will use a Heat Map as area encoding and a bar chart as length encoding
- Which one is easier to read?

Effectiveness:
Tableau example



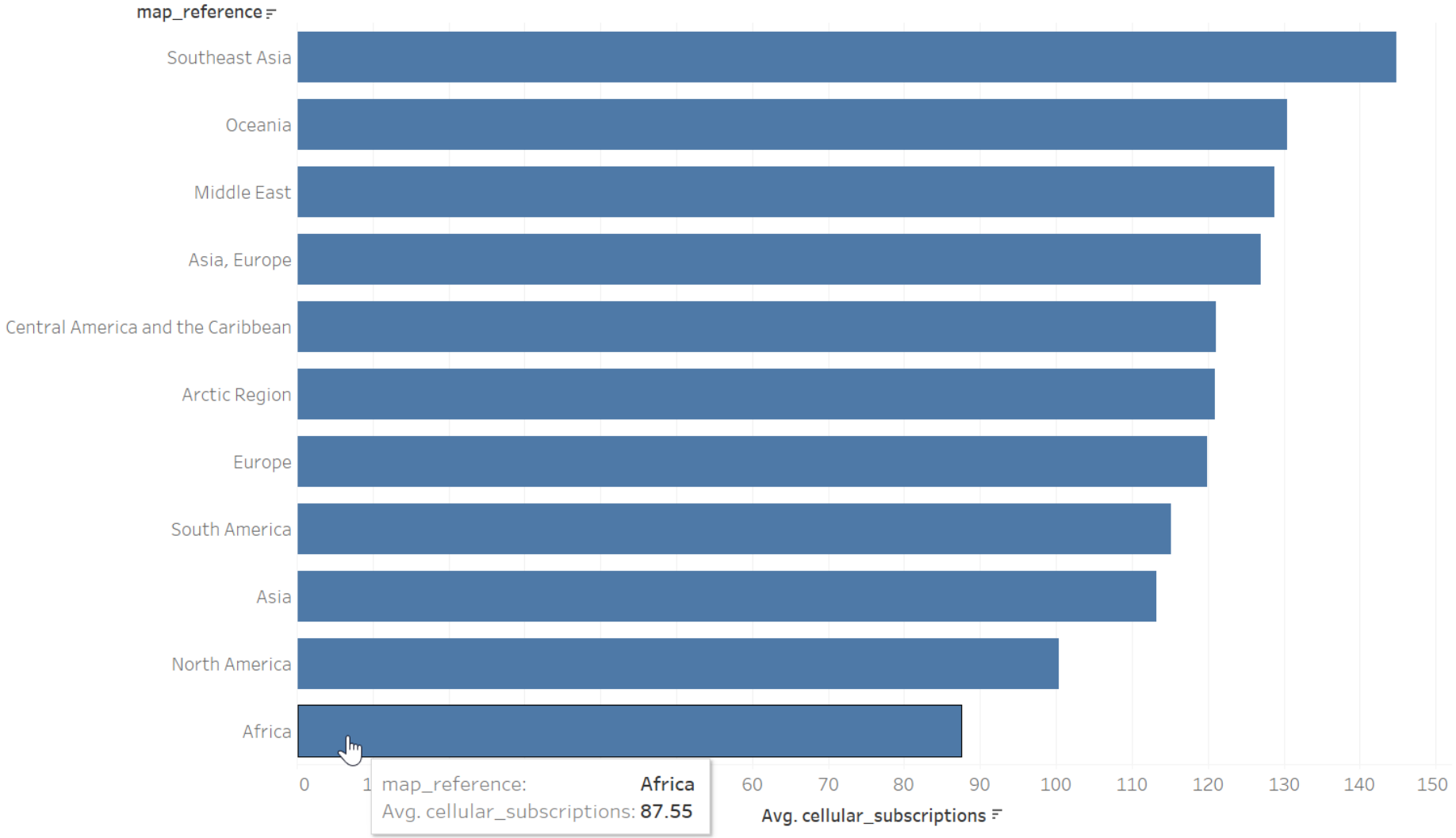
Effectiveness:

Tableau example

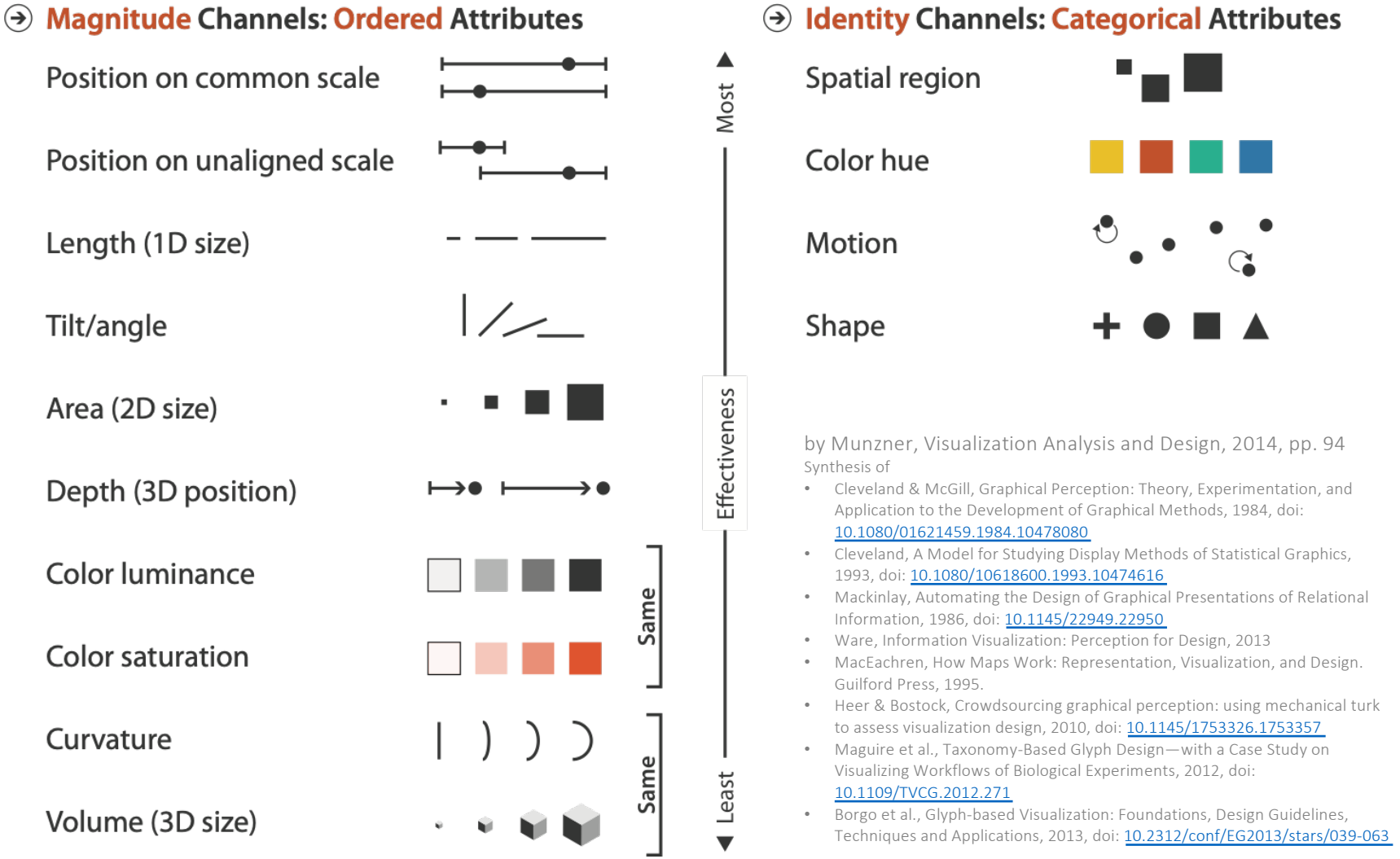
- On the right side is a **Show Me** button, there you can easily change the encoding

Effectiveness:

Tableau example



Effectiveness ranking



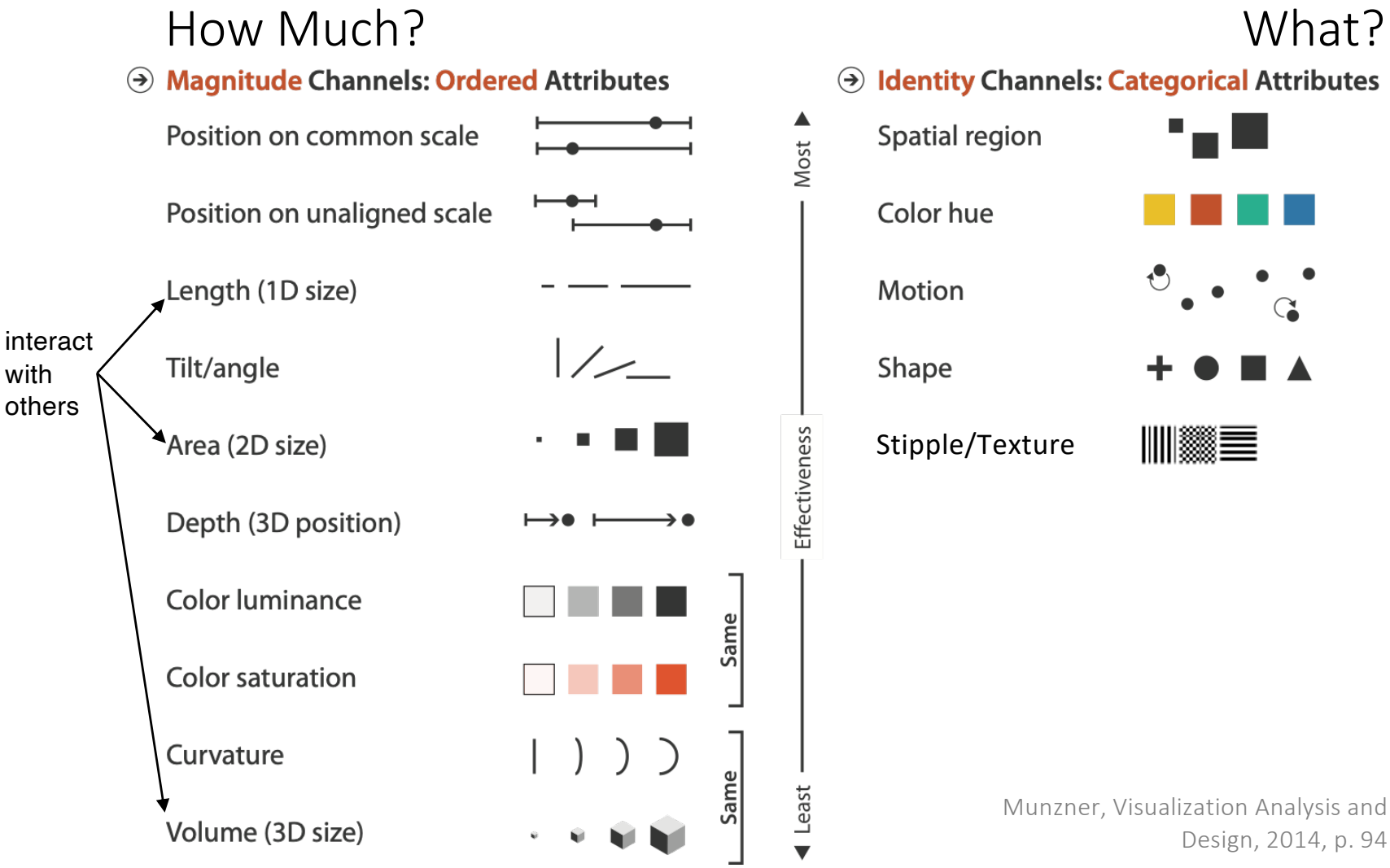
by Munzner, Visualization Analysis and Design, 2014, pp. 94

Synthesis of

- Cleveland & McGill, Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods, 1984, doi: 10.1080/01621459.1984.10478080
- Cleveland, A Model for Studying Display Methods of Statistical Graphics, 1993, doi: 10.1080/10618600.1993.10474616
- Mackinlay, Automating the Design of Graphical Presentations of Relational Information, 1986, doi: 10.1145/22949.22950
- Ware, Information Visualization: Perception for Design, 2013
- MacEachren, How Maps Work: Representation, Visualization, and Design. Guilford Press, 1995.
- Heer & Bostock, Crowdsourcing graphical perception: using mechanical turk to assess visualization design, 2010, doi: 10.1145/1753326.1753357
- Maguire et al., Taxonomy-Based Glyph Design—with a Case Study on Visualizing Workflows of Biological Experiments, 2012, doi: 10.1109/TVCG.2012.271
- Borgo et al., Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications, 2013, doi: 10.2312/conf/EG2013/stars/039-063

Channel characteristics

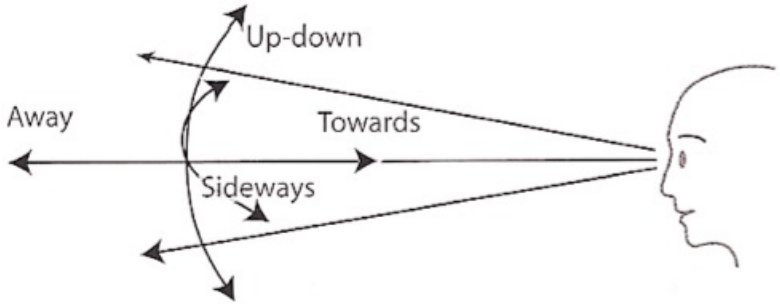
Available Channels

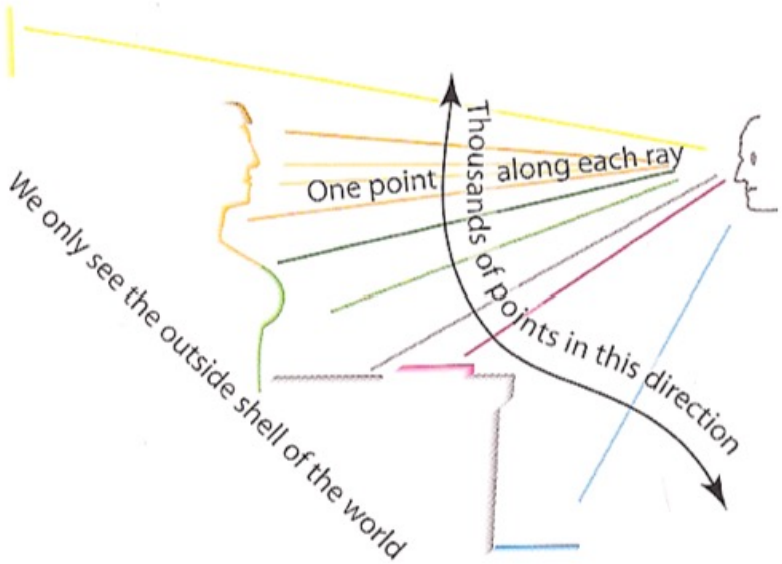


Munzner, Visualization Analysis and Design, 2014, p. 94

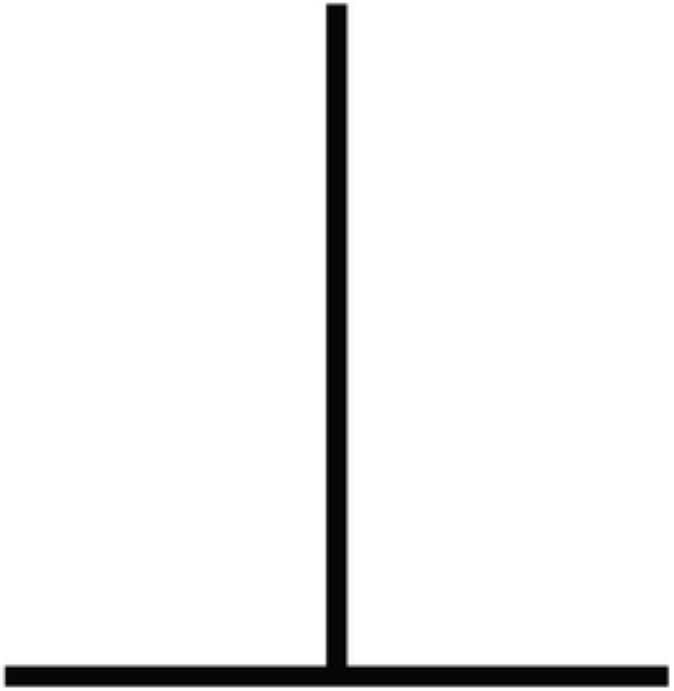
Spatial position

- Most effective for all data types
- But: only in left/right and up/down direction
- We only see in „2.05D“



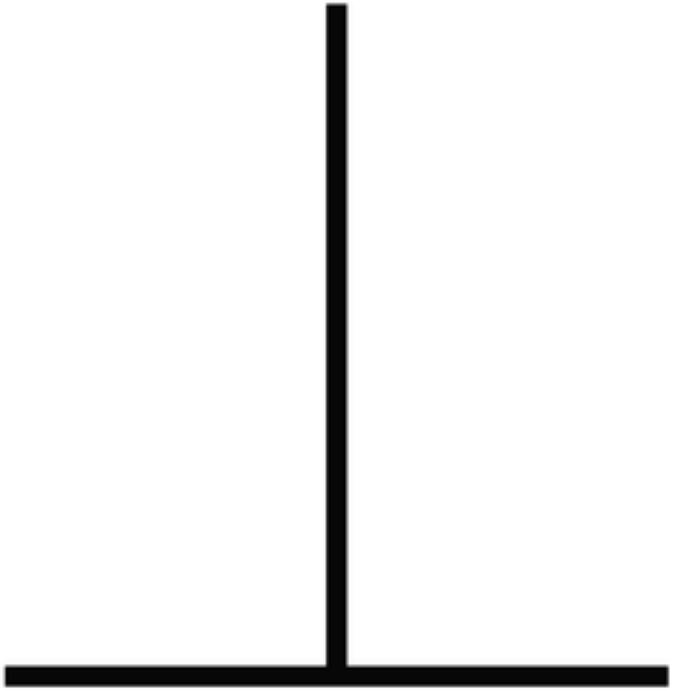


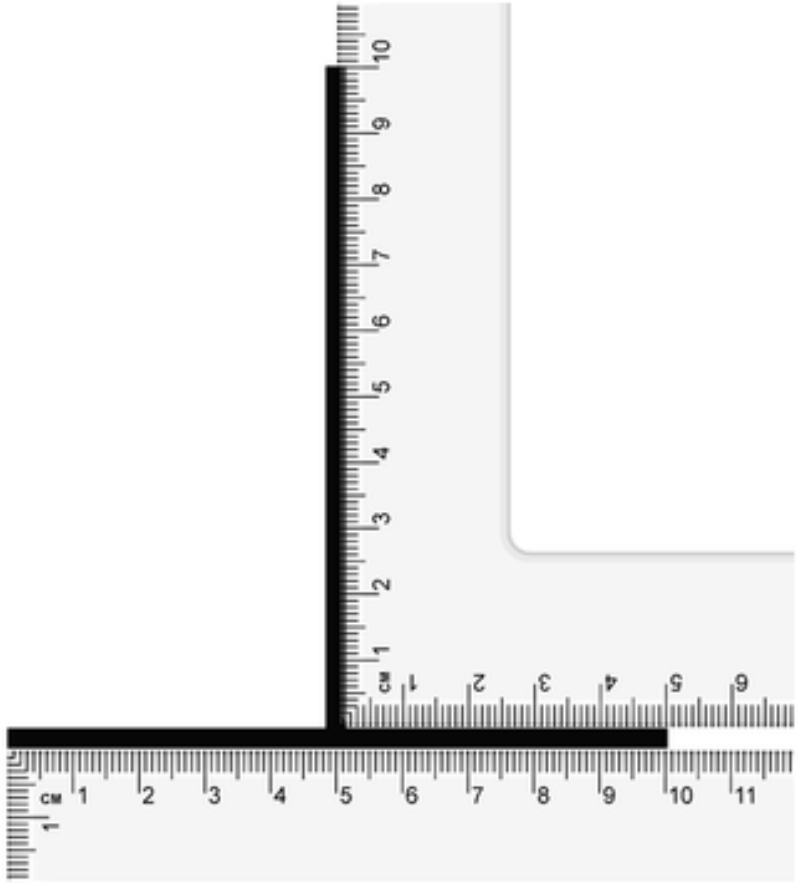
Which bar is longer?



- Horizontal
- Vertical
- Both have the same length

Vertical-horizontal illusion

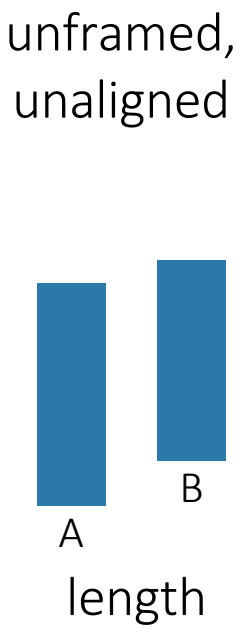




https://en.wikipedia.org/wiki/Vertical–horizontal_illusion

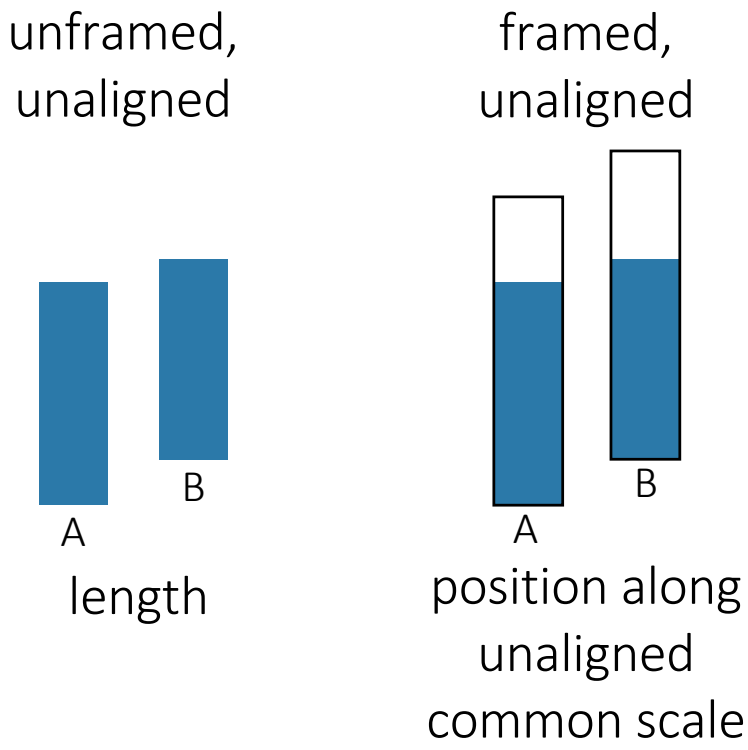
Relative vs. absolute judgement

- Weber's law says that everything is relative, i.e. the "intensity" depends on the background signal



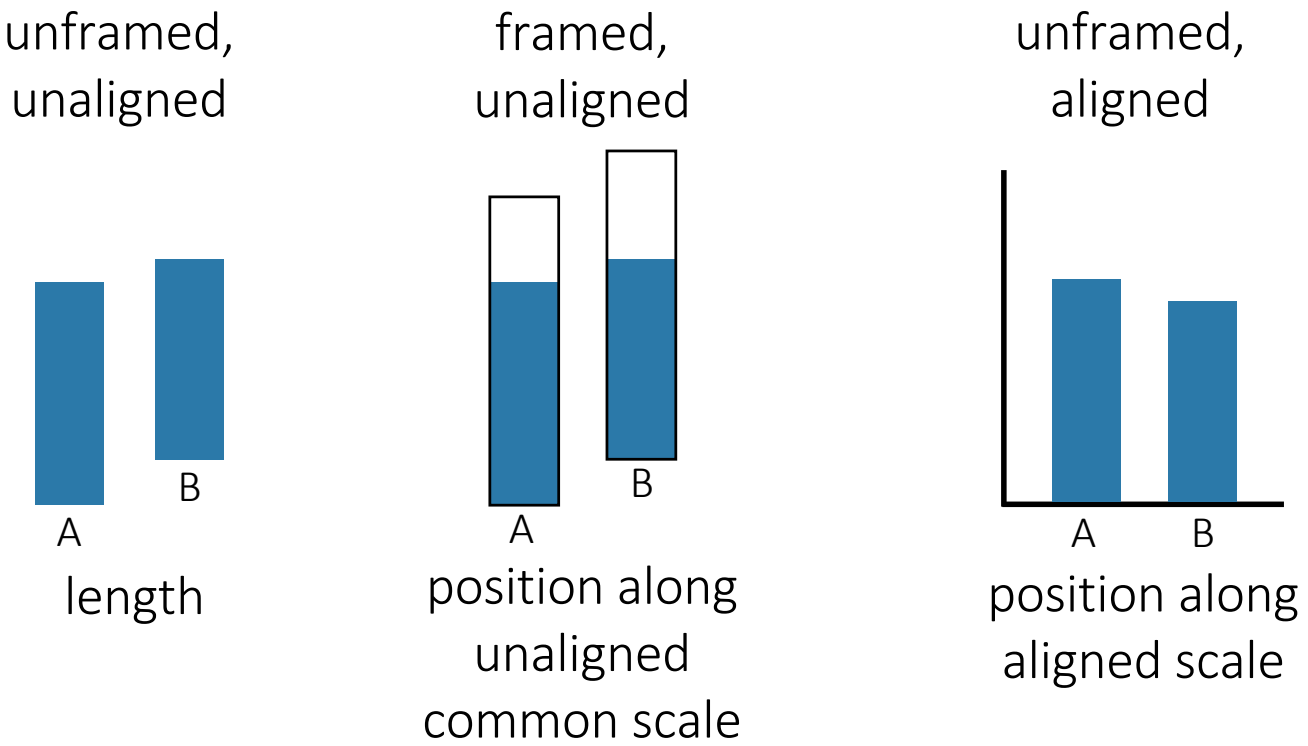
Relative vs. absolute judgement

- Weber's law says that everything is relative, i.e. the "intensity" depends on the background signal

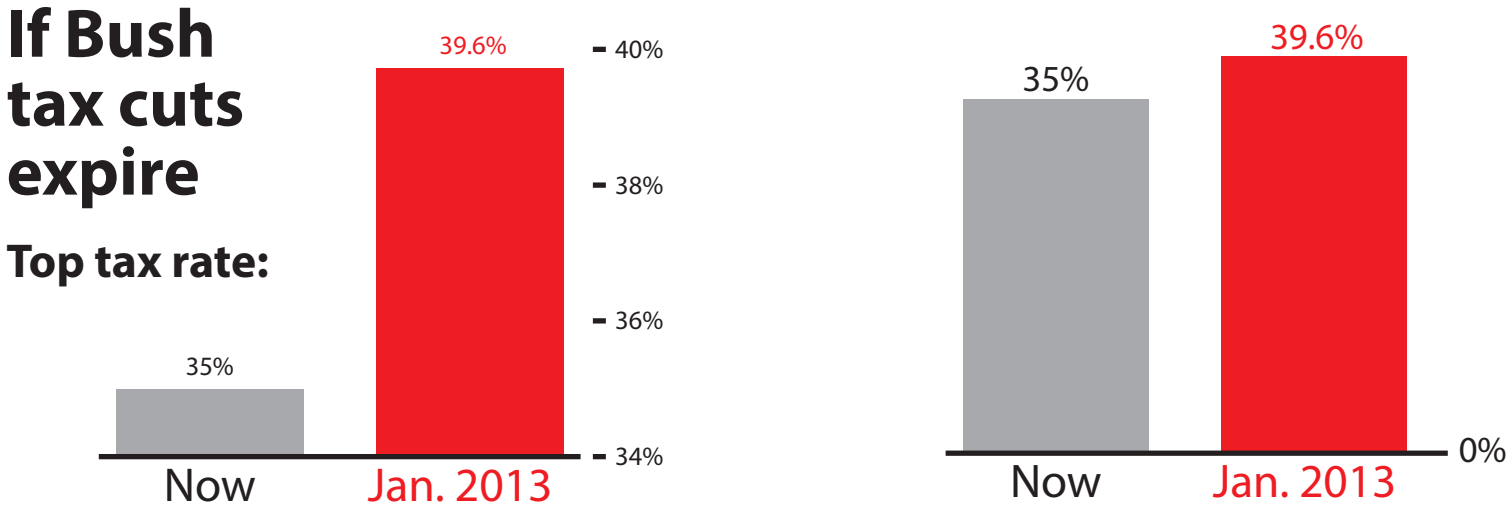


Relative vs. absolute judgement

- Weber's law says that everything is relative, i.e. the "intensity" depends on the background signal

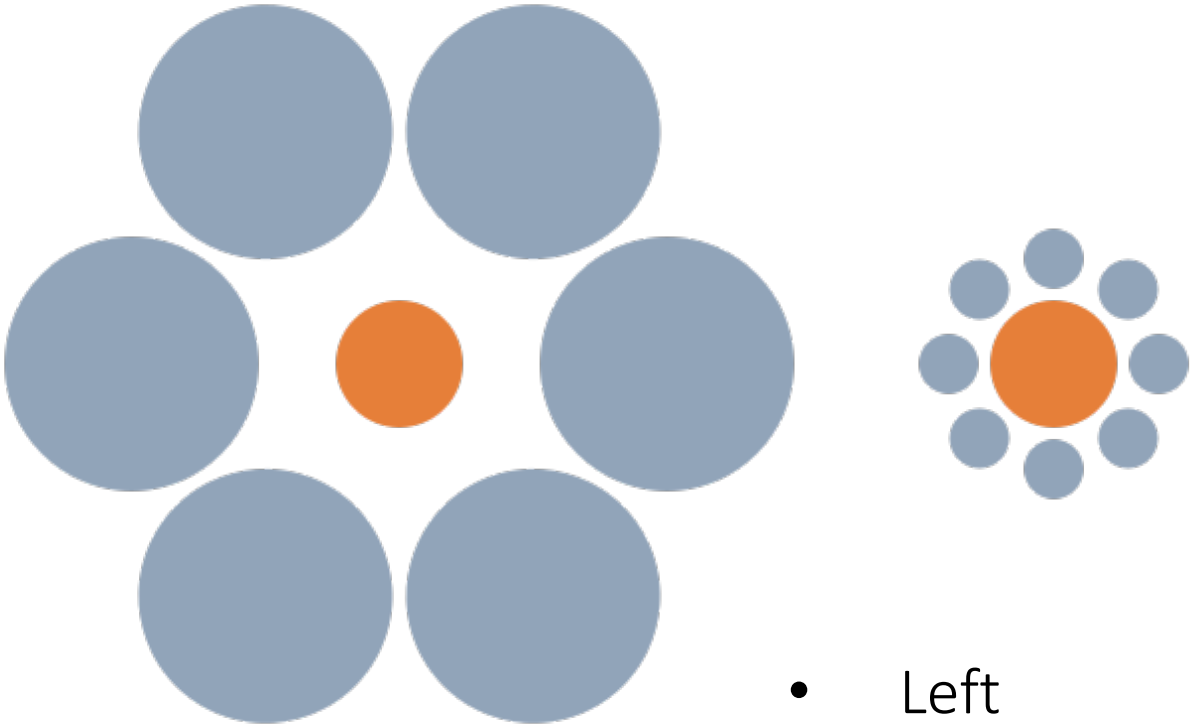


Relative vs. absolute judgement



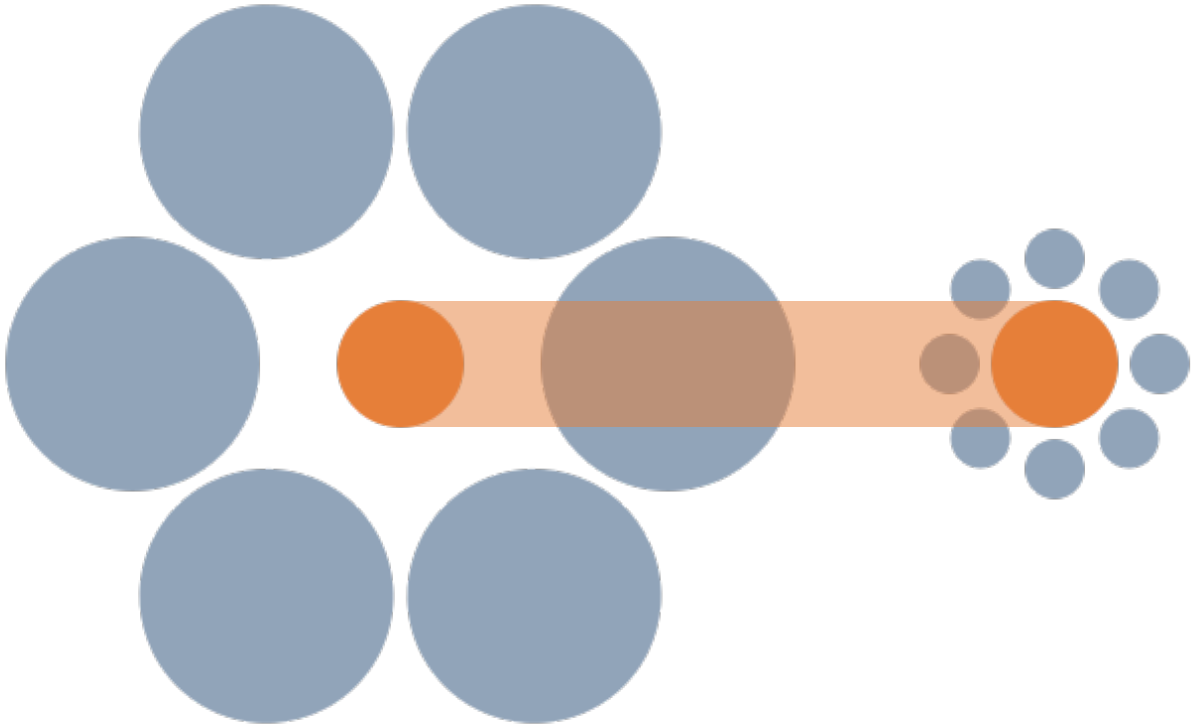
From Alberto Cairo, HOW CHARTS LIE

Which orange circle is larger?



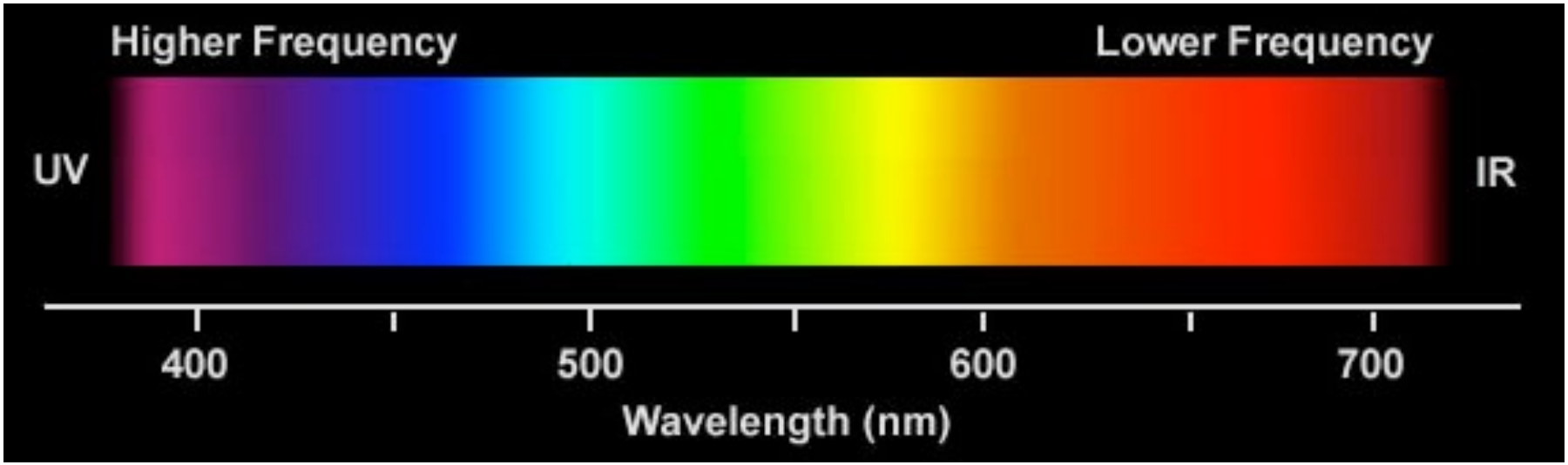
- Left
- Right
- Both are the same size

Ebbinghaus illusion



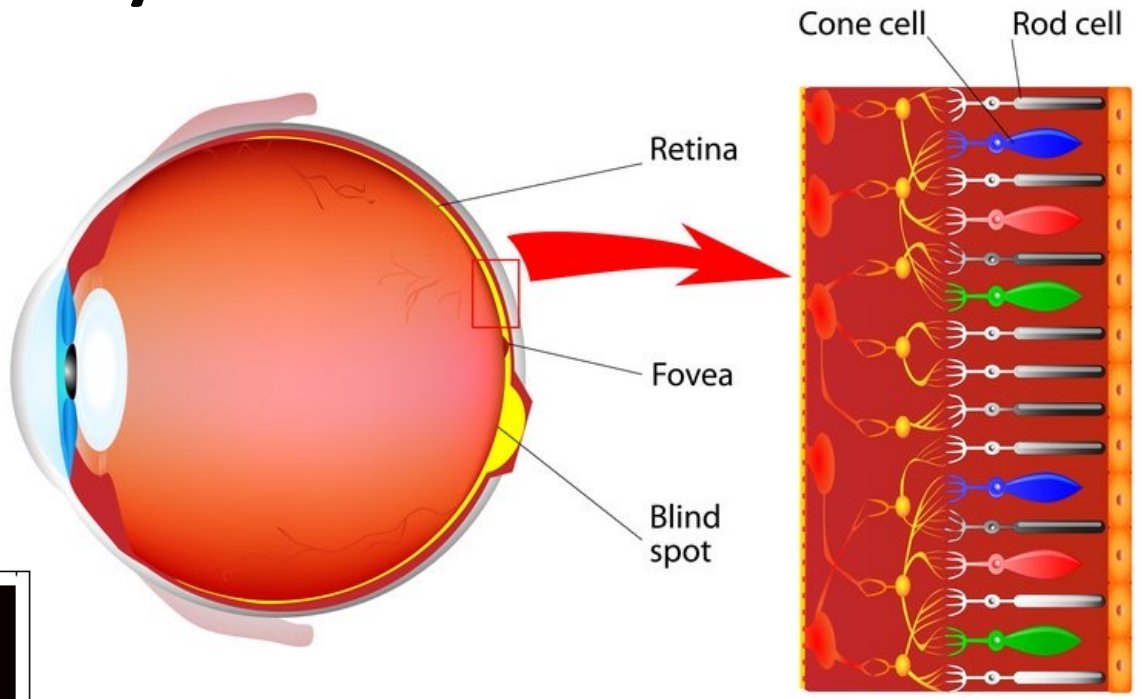
https://en.wikipedia.org/wiki/Ebbinghaus_illusion

Perception/Color theory basics

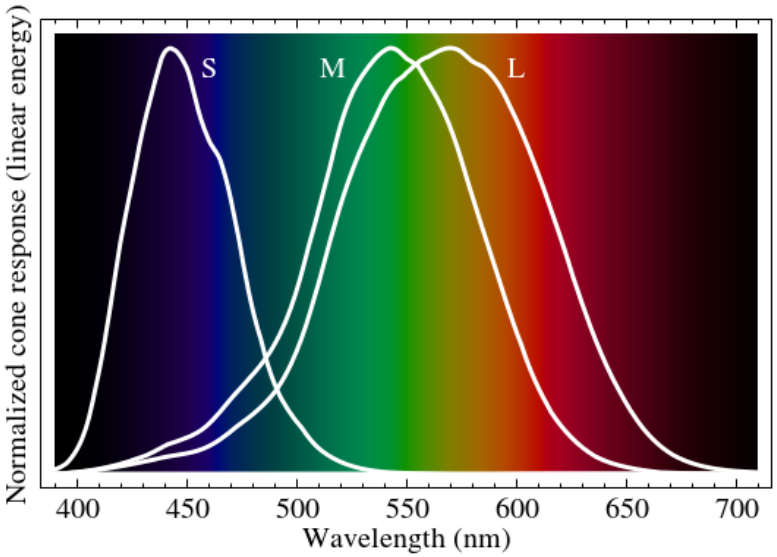


Perception/Color theory basics:

Human visual system



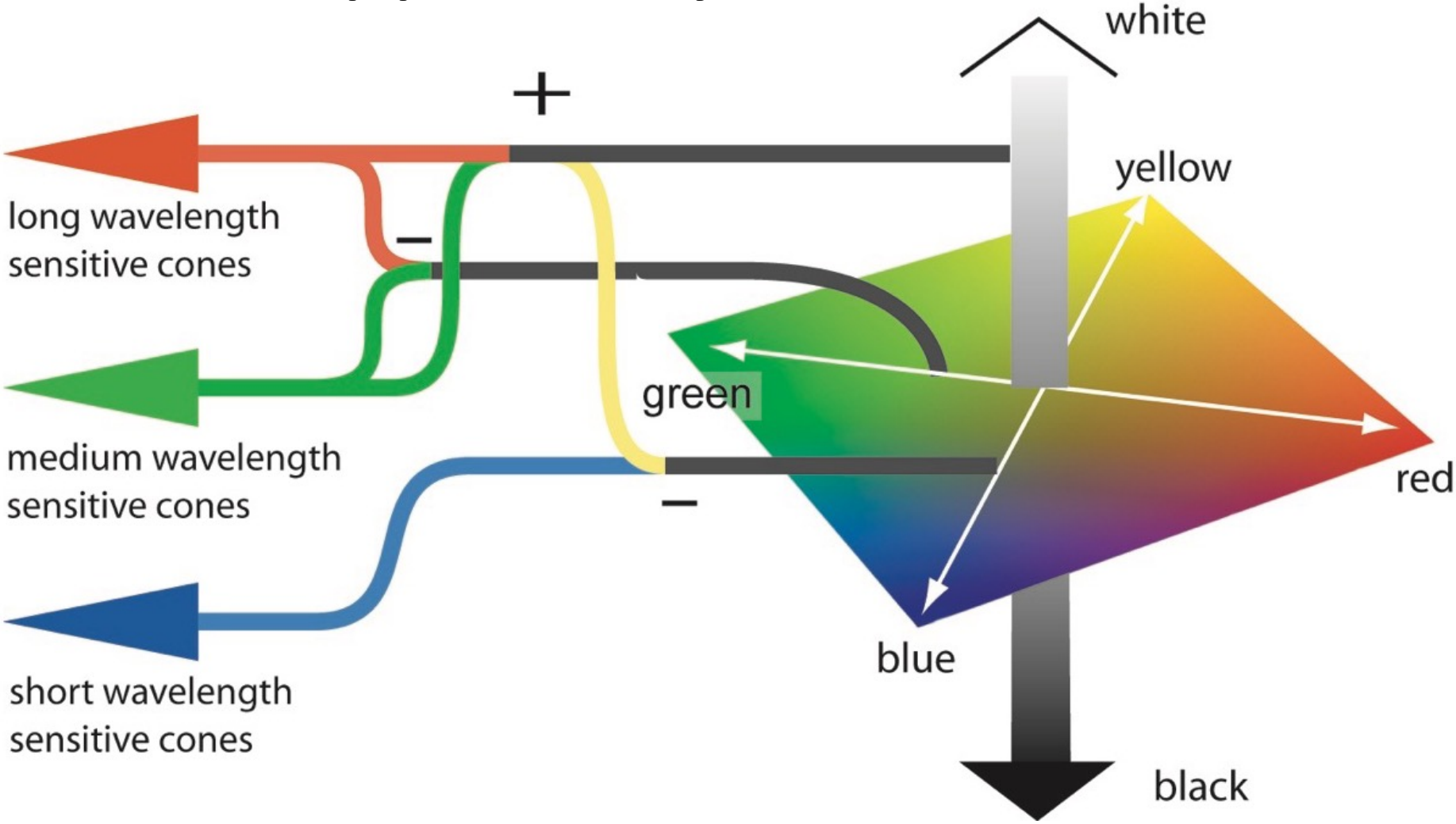
https://www.webrn-maculardegeneration.com/rods-and-cones.html



https://en.wikipedia.org/wiki/Color_vision

Perception/Color theory basics:

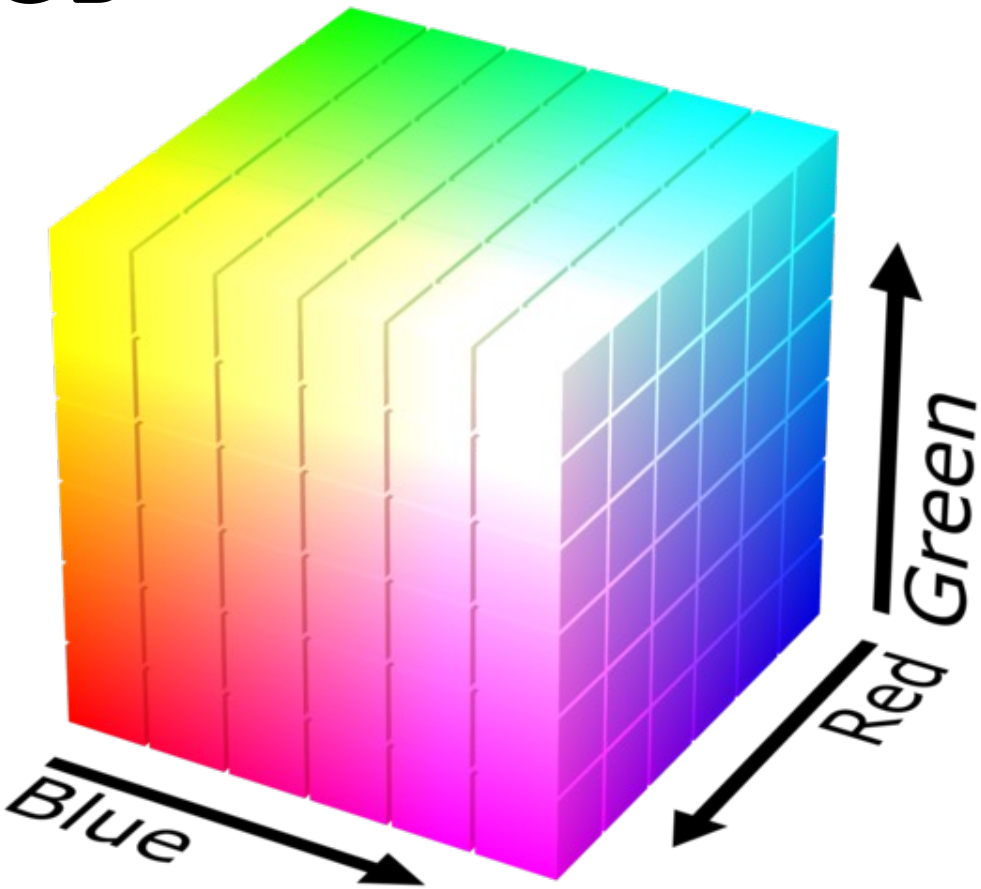
Color opponency



Ware, Visual thinking for design

Color models: RGB

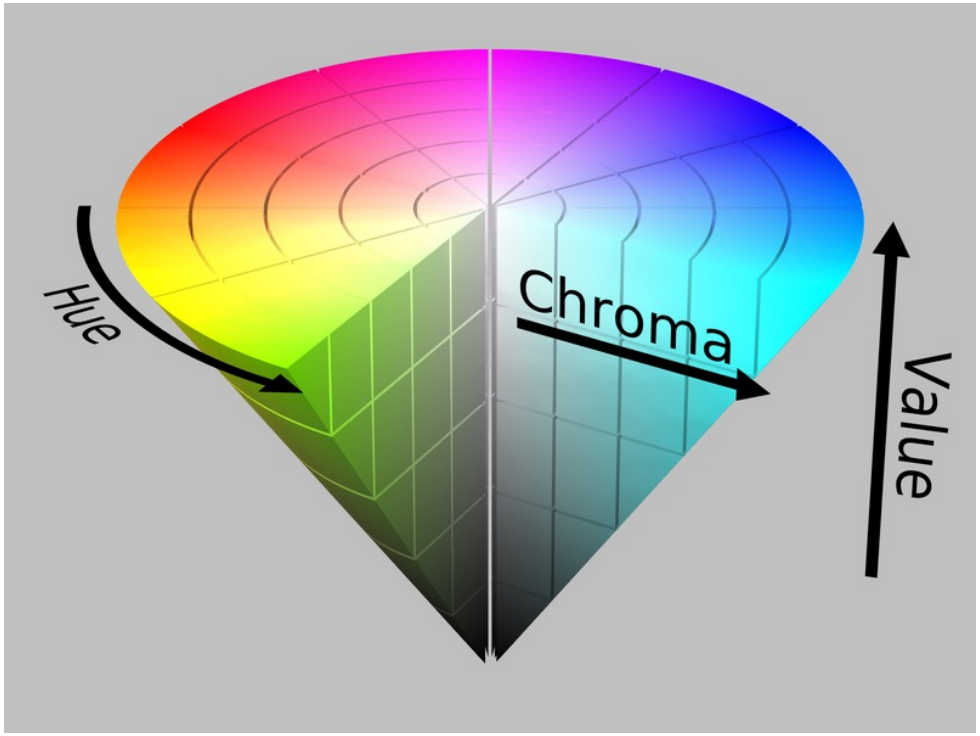
- Additive system
- Colors that can be represented by computer monitors
- Not perceptually uniform



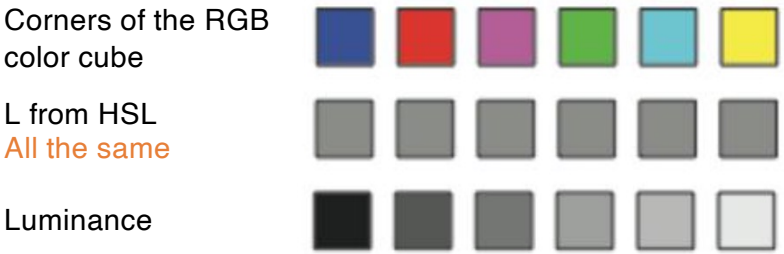
https://en.wikipedia.org/wiki/RGB_color_model

Color models: HSL/HSV

- Hue
- Saturation (Chroma)
- Lightness (Value)
- Not perceptually uniform



https://en.wikipedia.org/wiki/HSL_and_HSV



Munzner, Visualization Analysis and Design, p. 221

Luminance, Saturation, Hue

- Luminance
  - *How Much* channel
  - discriminability: ~2-4 bins
  - contrast important

- Saturation
  - *How Much* channel
  - discriminability: ~3 bins

- Hue
  - *What* channel
  - discriminability: ~6-12

Color: Perception issues + deficiencies

Order of the squares?

Thanks to Moritz Wustinger

Do all boxes have the same color?

- Left is different
- Center is different
- Right is different
- All are exactly the same

Thanks to Moritz Wustinger

Color: Perception issues + deficiencies

What do you see?

Thanks to Moritz Wustinger

Thanks to Moritz Wustinger
Smiley based on http://upload.wikimedia.org/wikipedia/commons/b/bd/A_Smiley.jpg

Color: Perception issues + deficiencies

Do all boxes have the same color?

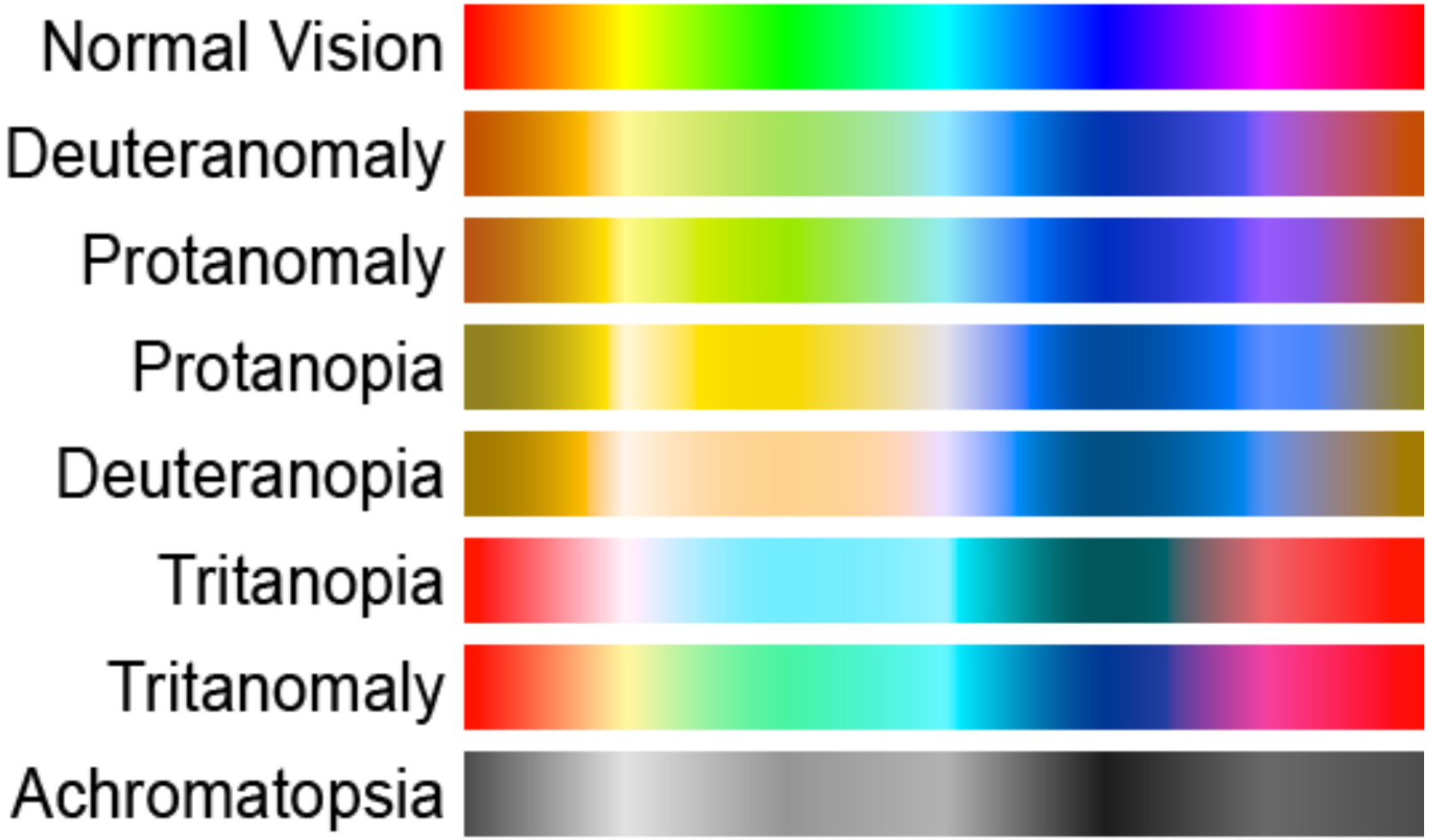
Left is different!

Thanks to Moritz Wustinger

Color blindness

- Affects opponent color processing
  - Red-green (deuteranopia, protanopia)
  - Blue-yellow possible (tritanopia)
  - Luminance channel (Monochromacy, Achromatopsia)
- Prevalence ~8% male, ~0.5% female
- Can be simulated
  - Brettel et al., Computerized simulation of color appearance for dichromats, 1997, doi: 10.1364/josaa.14.002647
  - Online Tools:
    - https://www.toptal.com/designers/colorfilter
    - https://www.color-blindness.com/coblis-color-blindness-simulator/
    - https://pilestone.com/pages/color-blindness-simulator-1

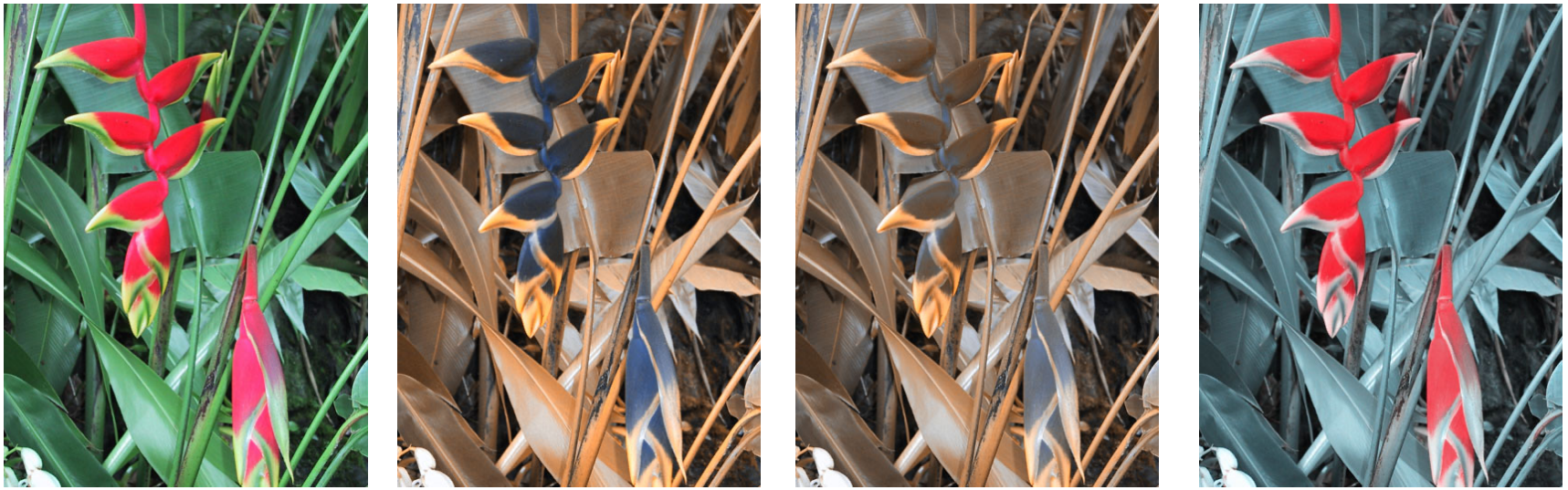
Color blindness



https://en.wikipedia.org/wiki/Color_blindness

Color blindness

Source: M. Stone



Protanope (no L-cones)

Deuteranope (no M-cones)

Tritanope (no S-cones)

red / green deficiencies

blue / yellow deficiency

Color: Perception issues + deficiencies

Small areas

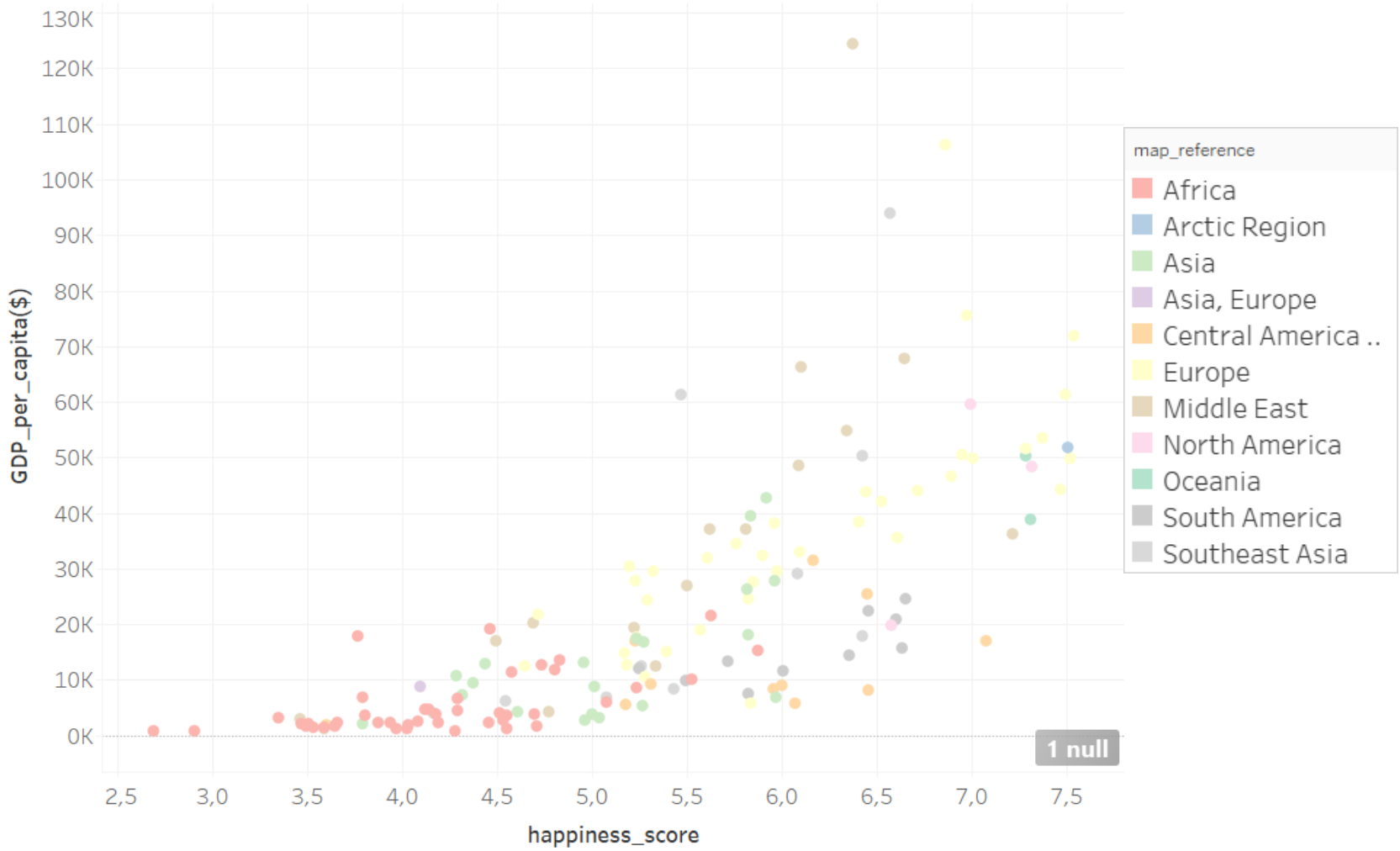


Tableau Software

Color: Perception issues + deficiencies

Small areas

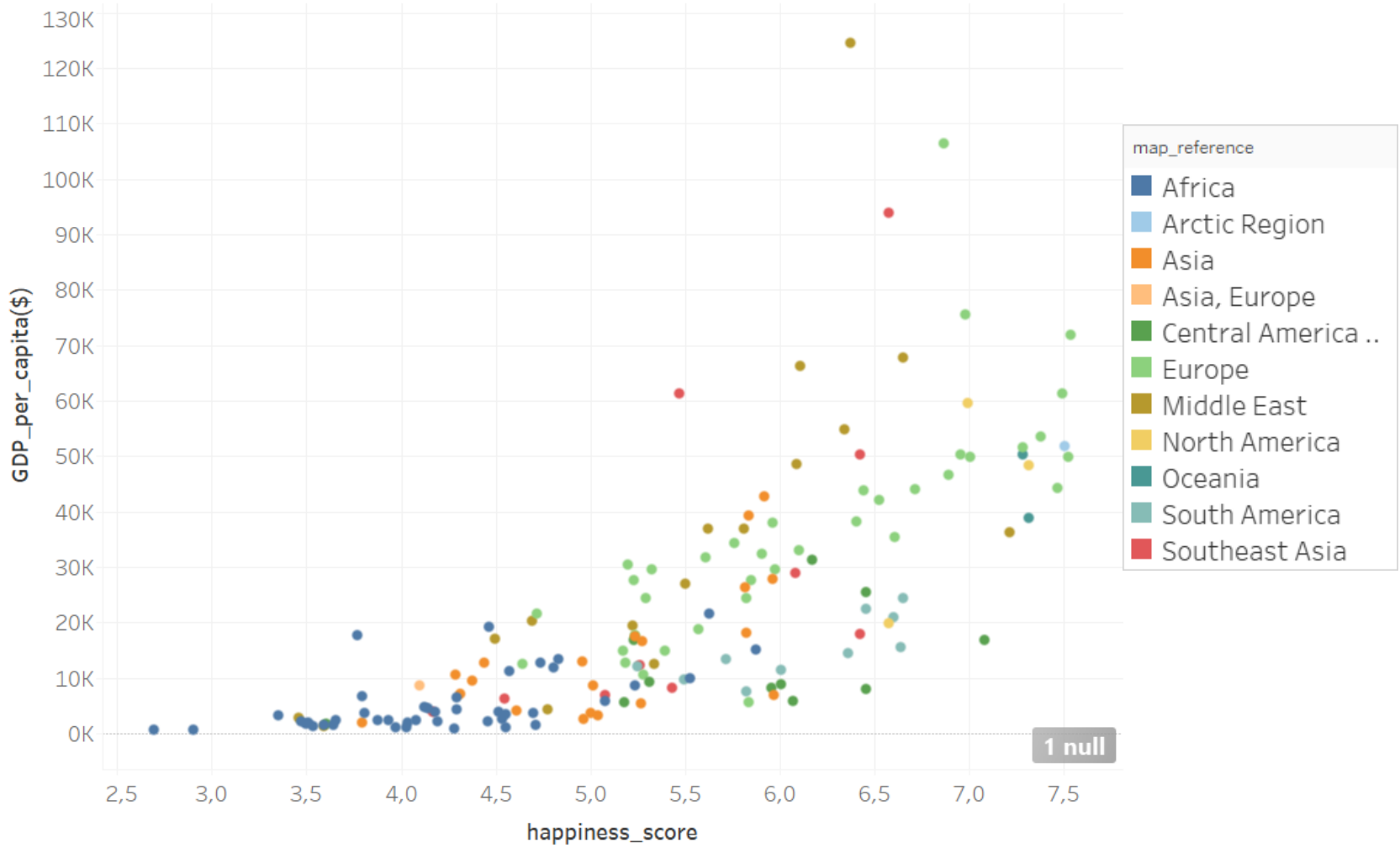
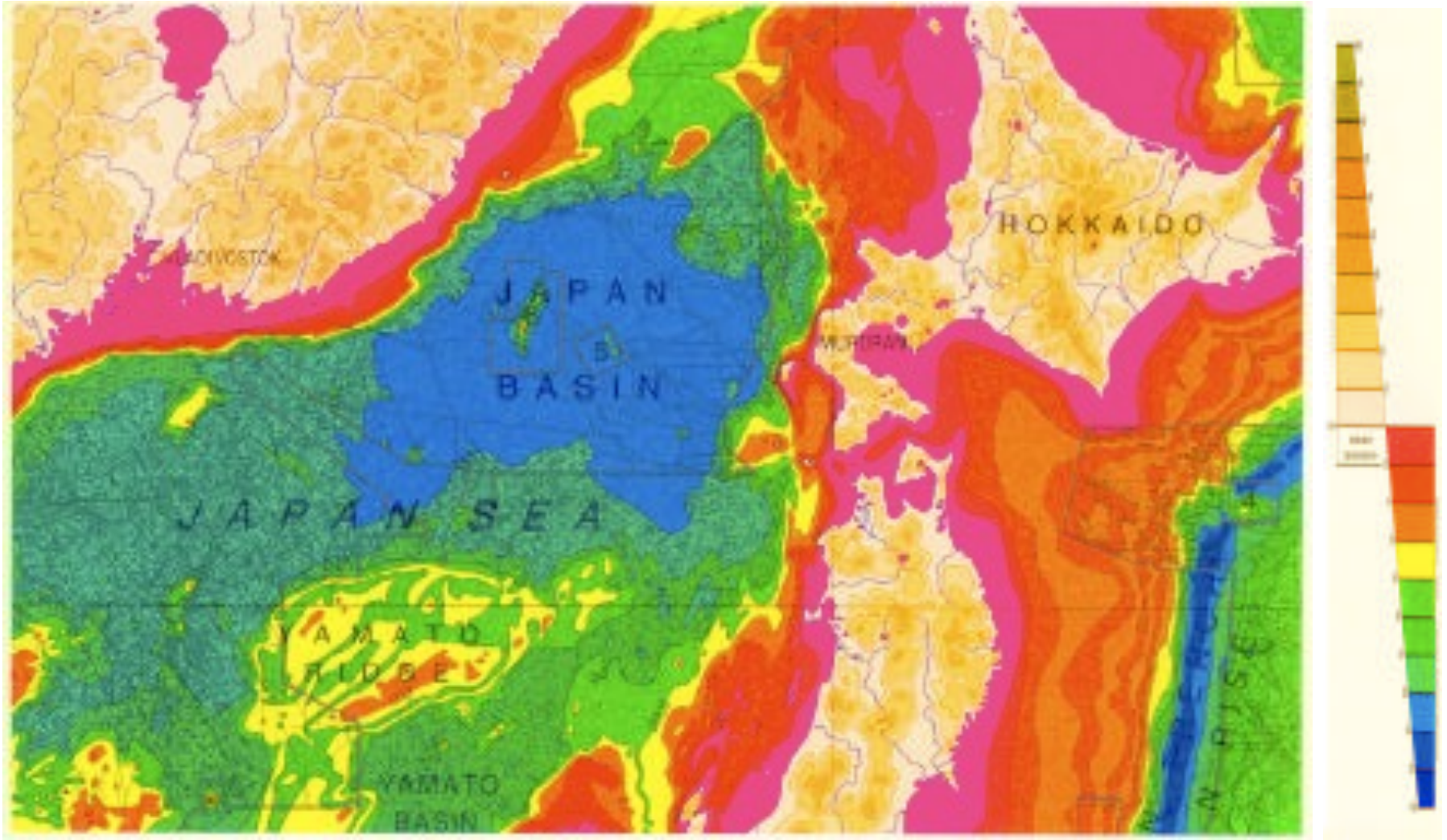


Tableau Software

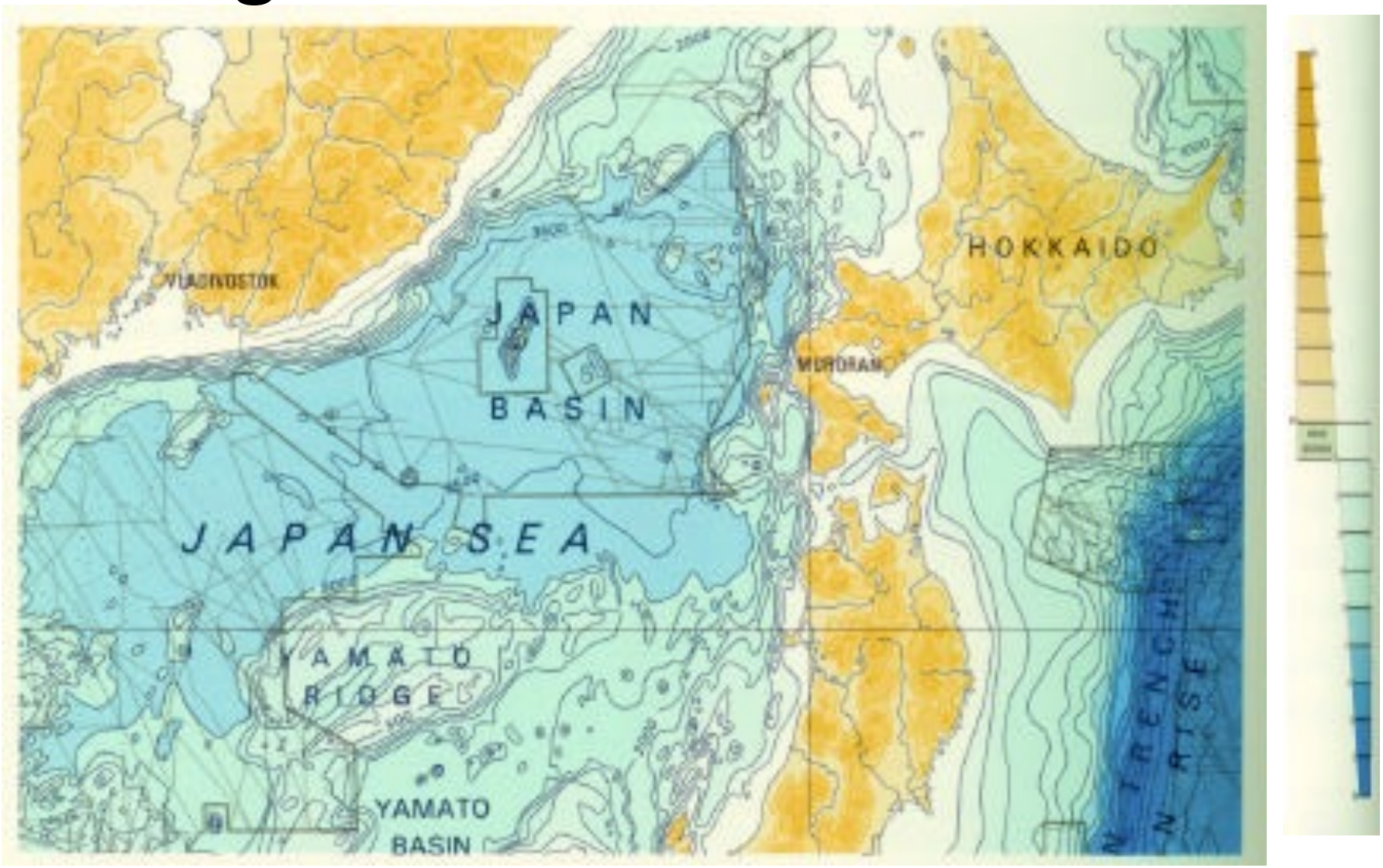
Color: Perception issues + deficiencies

Large areas



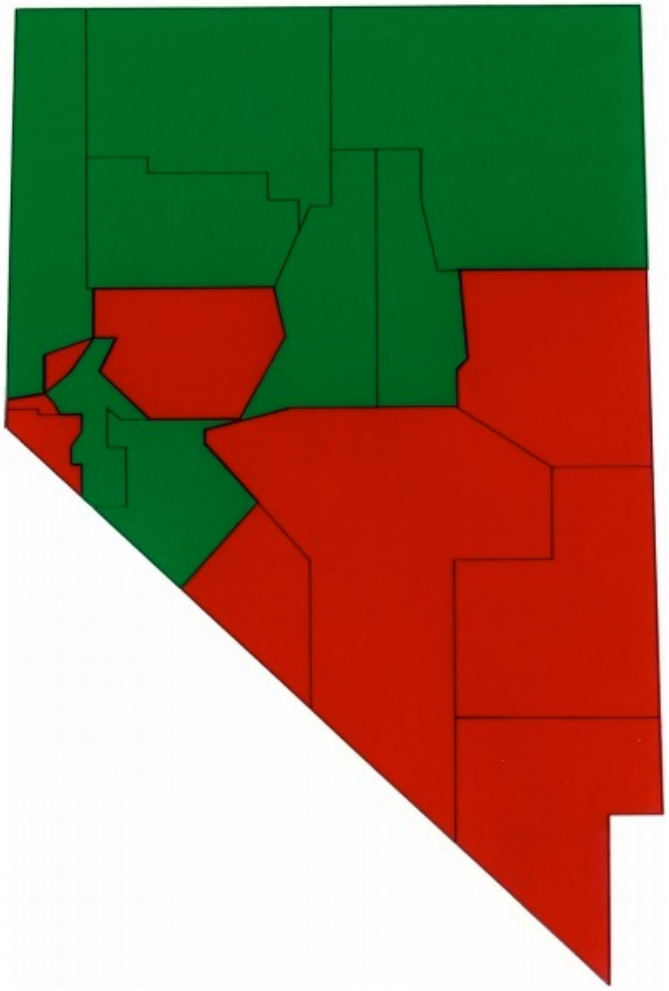
Tufte, The Visual Display of Quantative Information (Vol. 1), p. 77

Large areas



Tufte, The Visual Display of Quantitative Information (Vol. 1), p. 77

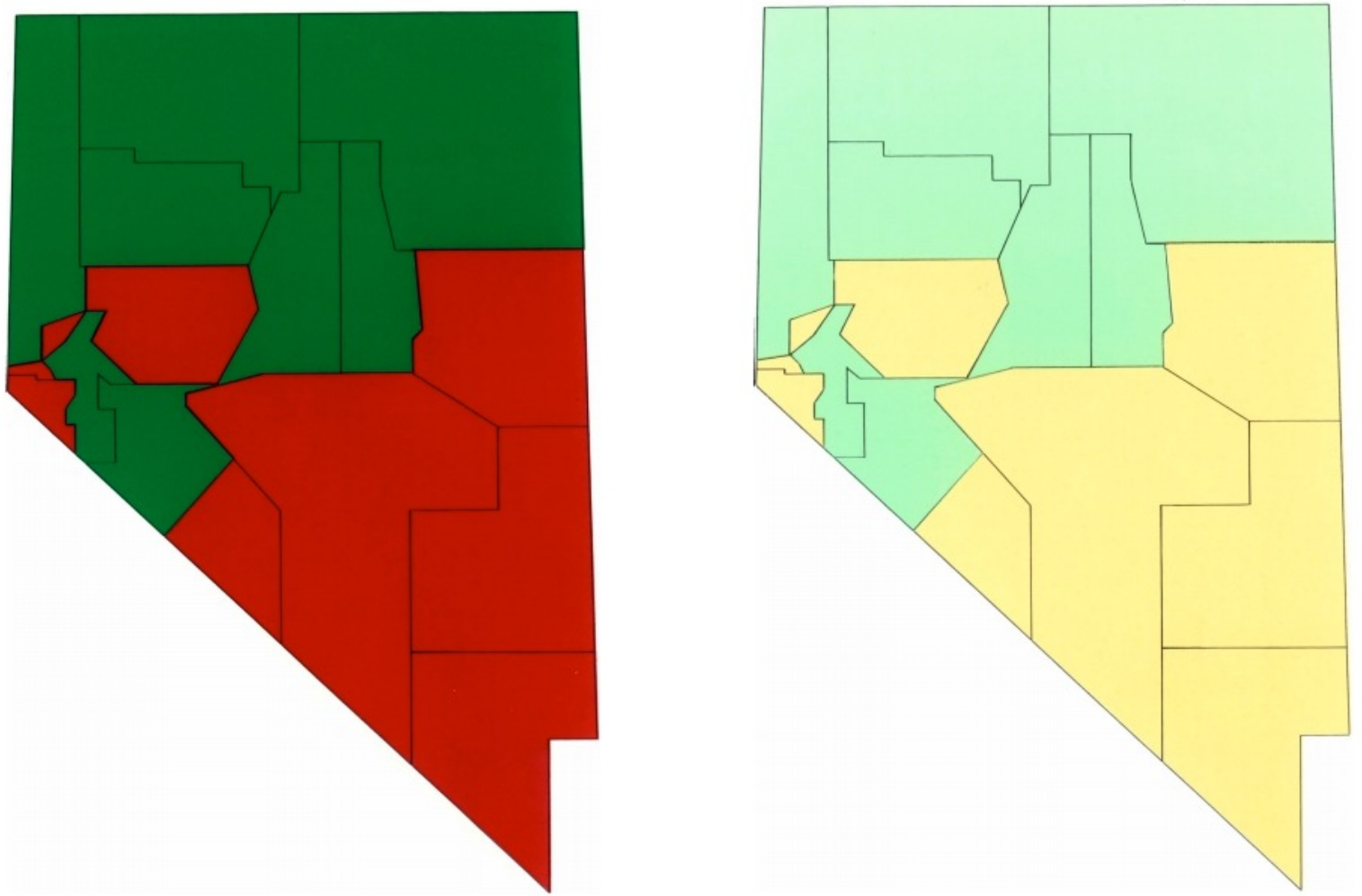
Which area is larger?



- Red is bigger
- Green is bigger
- Both are the same size

Color: Perception issues + deficiencies

Color size illusion



Cleveland & McGill, A Color-Caused Optical Illusion on a Statistical Graph, 1983,
doi: 10.1080/00031305.1983.10482720, link

Color: Perception issues + deficiencies

Maps

Ware, Visual Thinking for Design

Take-home message

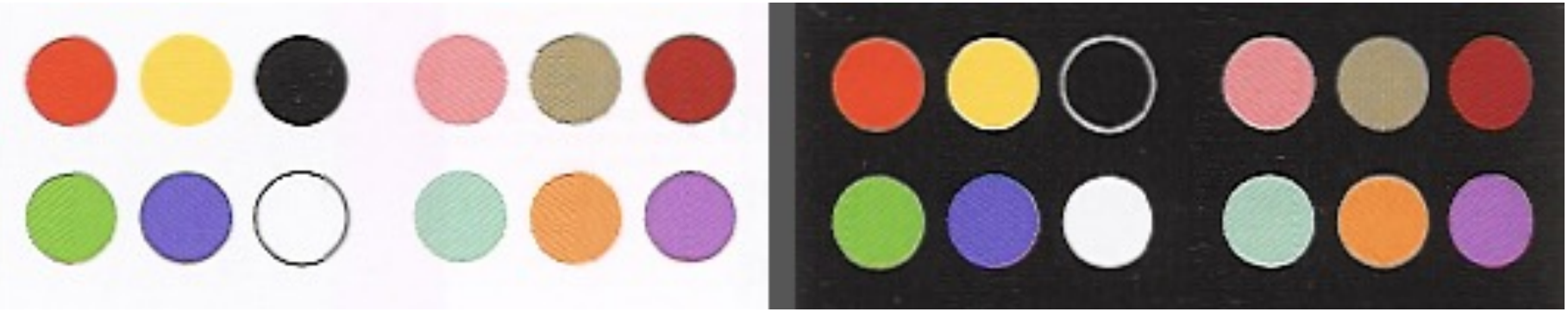
- Test charts for discoloration / color blindness issues

- Color in small regions difficult to perceive
  → Use bright, saturated colors for small regions

- Large areas in some saturated colors appear bigger
  → Use low saturation, pastel colors for large regions and backgrounds

Color: Categorical data

- *What* channel
- Limited distinguishability (8-14)
  - Best with Hue
  - Also vary Luminance and/or Saturation to address perception deficiencies

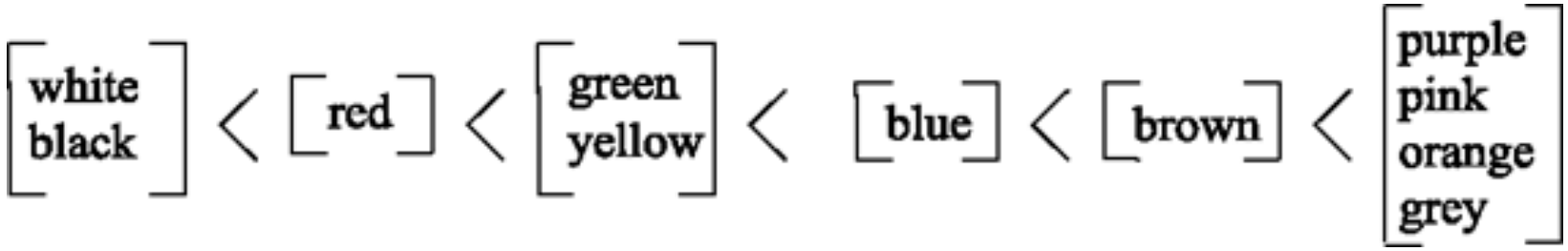


Ware, Information Visualization: Perception for Design

Basic Colors

The evolution of basic color terms is remarkably consistent across cultures

Berlin & Kay‘s (1969) implicational hierarchy:



Source: Dowman, Explaining Color Term Typology With an Evolutionary Model, 2010, doi: 10.1080/03640210709336986.

Take-home message

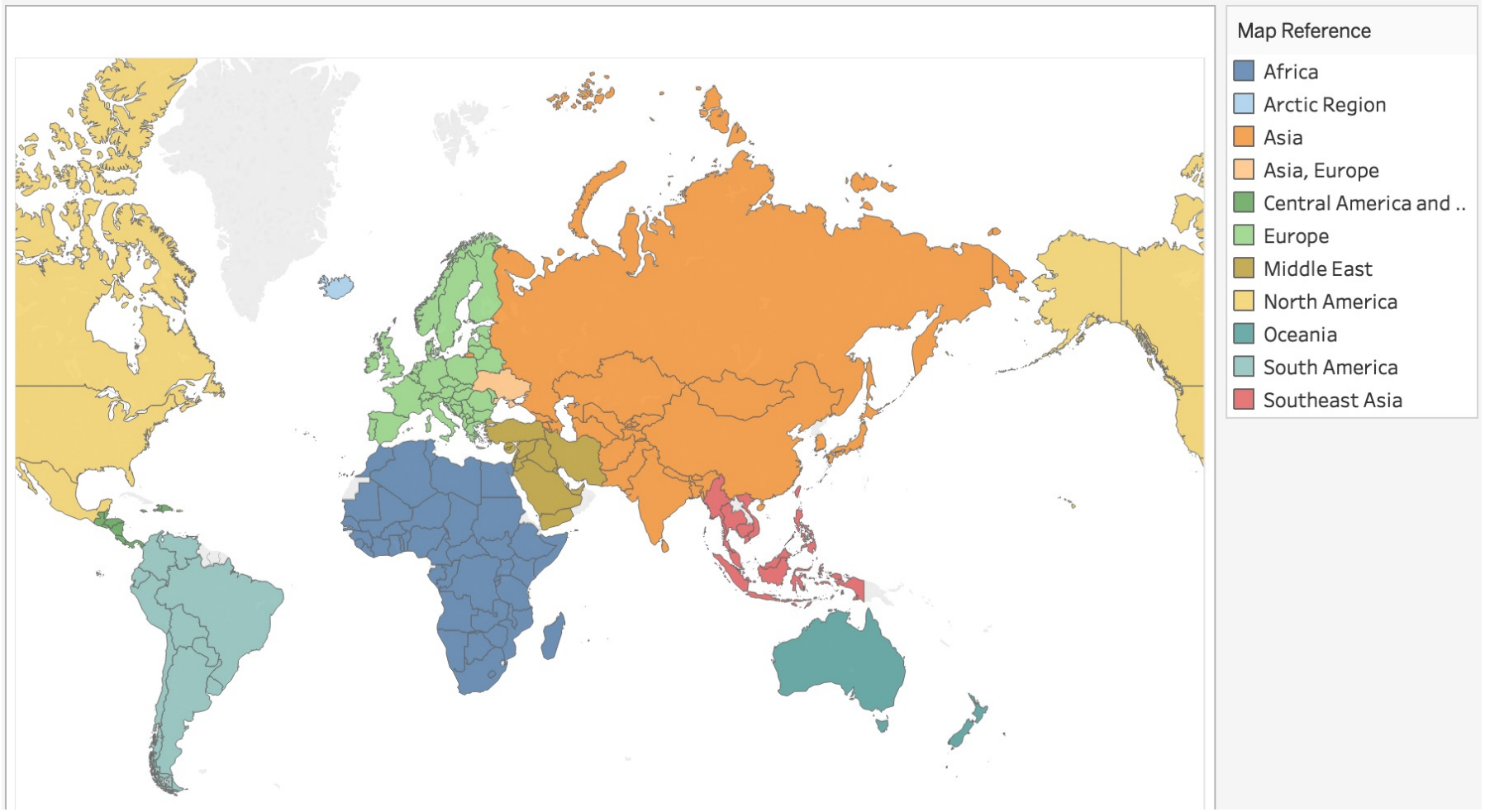
- Only a small number of colors can be used effectively as categorical labels
- Ideally use less than eight colors
- Also vary luminance/saturation to address deficiencies
- Use low saturation / gray backgrounds

Tableau example

- Let us use our map to test this!
- Try to use **Map Reference** or **Biggest official Language** as color for our map.
- What works and doesn't work?

Color: Categorical data

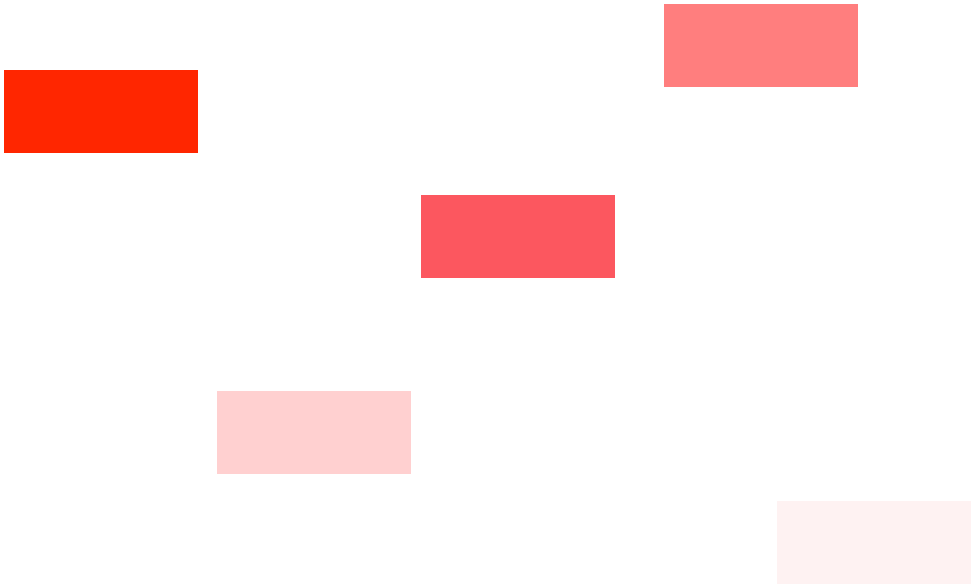
Tableau example



Source: J. Stasko

Color: Ordinal data

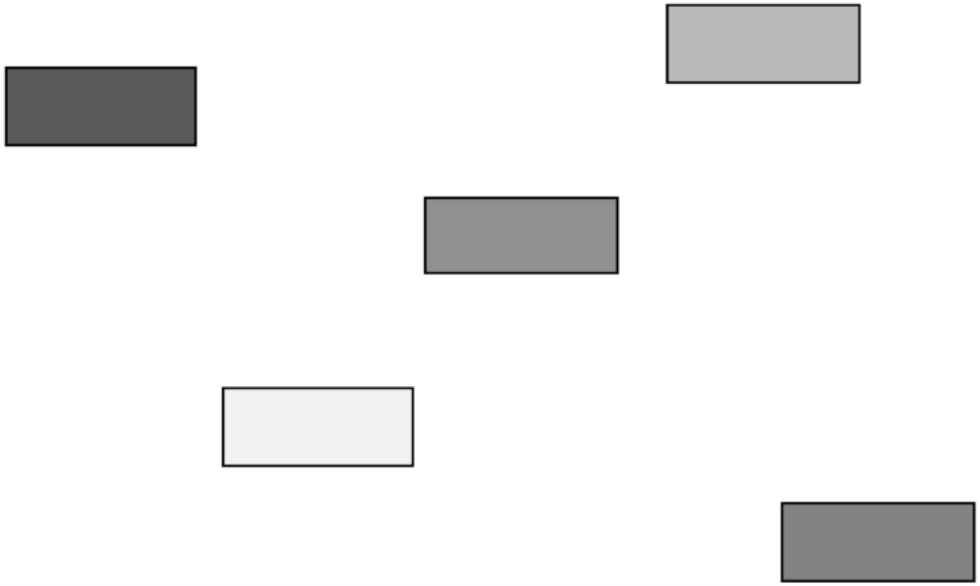
Order these colors



Source: J. Stasko

Color: Ordinal data

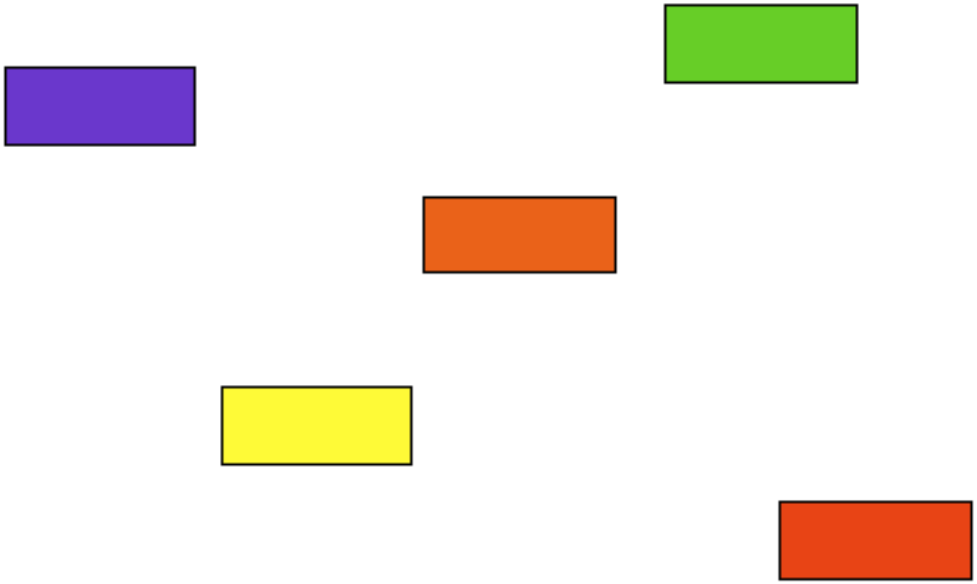
Order these colors



Source: J. Stasko

Color: Ordinal data

Order these colors



Take-home message

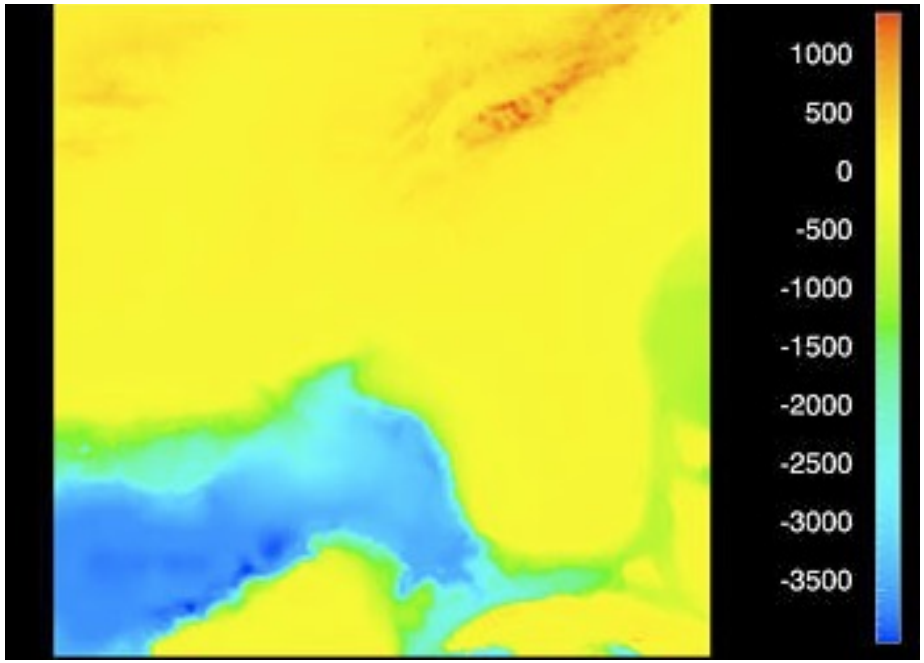
- Lightness and saturation are effective for ordinal data because they have an implicit perceptual ordering

- Show ordinal data with a discrete set of color values that change in lightness or saturation

Color: Quantitative data

Rainbow color map

How Much channel



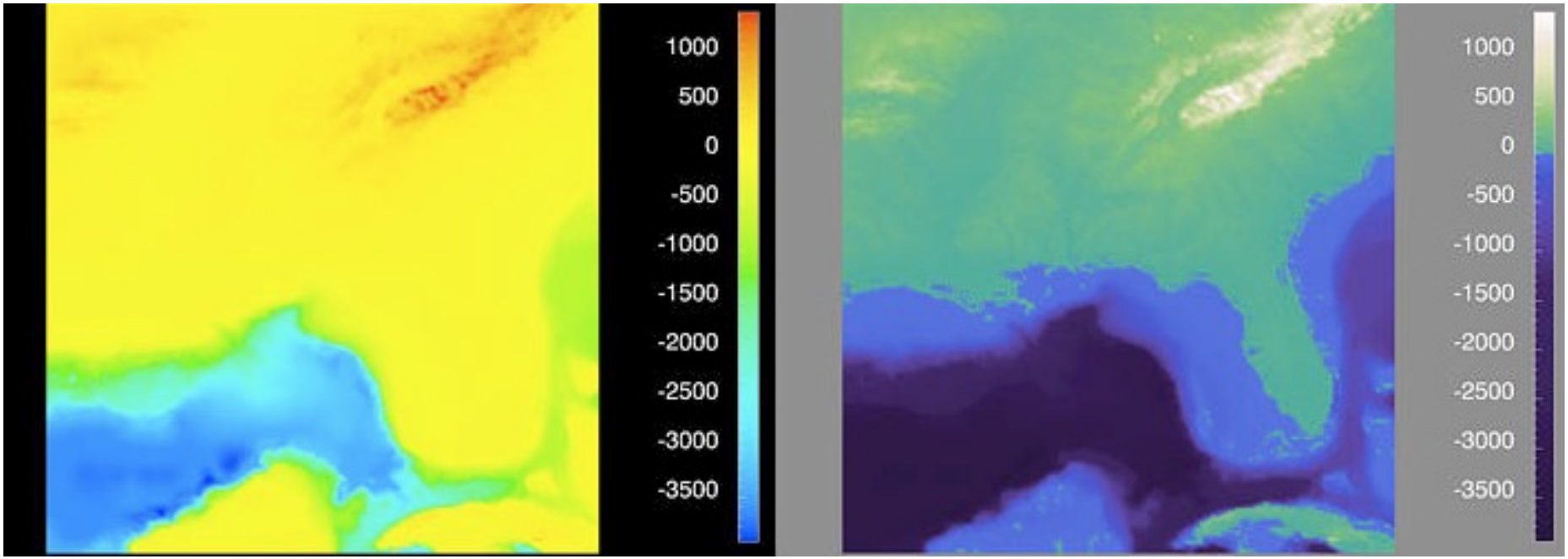
Which part of the world is visible here?

Rogowitz and Treinish, „Data Visualization: The End of the Rainbow" / "Why should engineers and scientists be worried about color?"

Color: Quantitative data

Rainbow color map

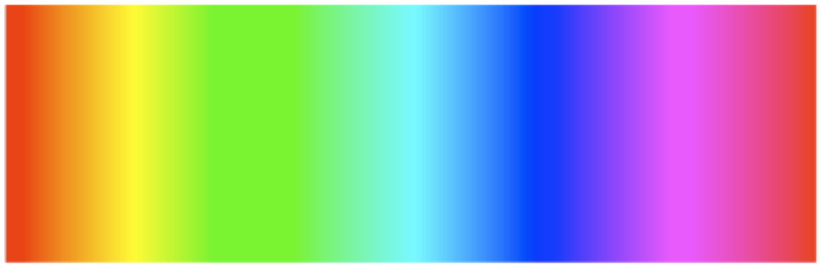
How Much channel



Rogowitz and Treinish, „Data Visualization: The End of the Rainbow" / "Why should engineers and scientists be worried about color?"

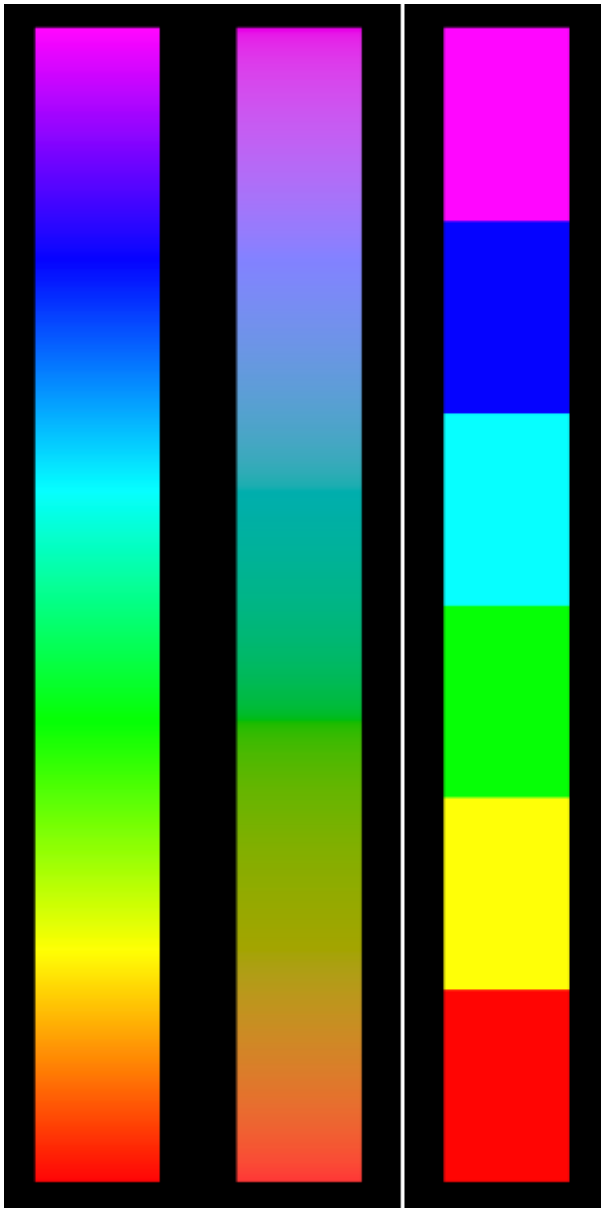
Rainbow color map

- Hue not ideal for ordinal data
- Not perceptually linear: Equal steps in the continuous range are not perceived as equal steps
- Not good for colorblind people

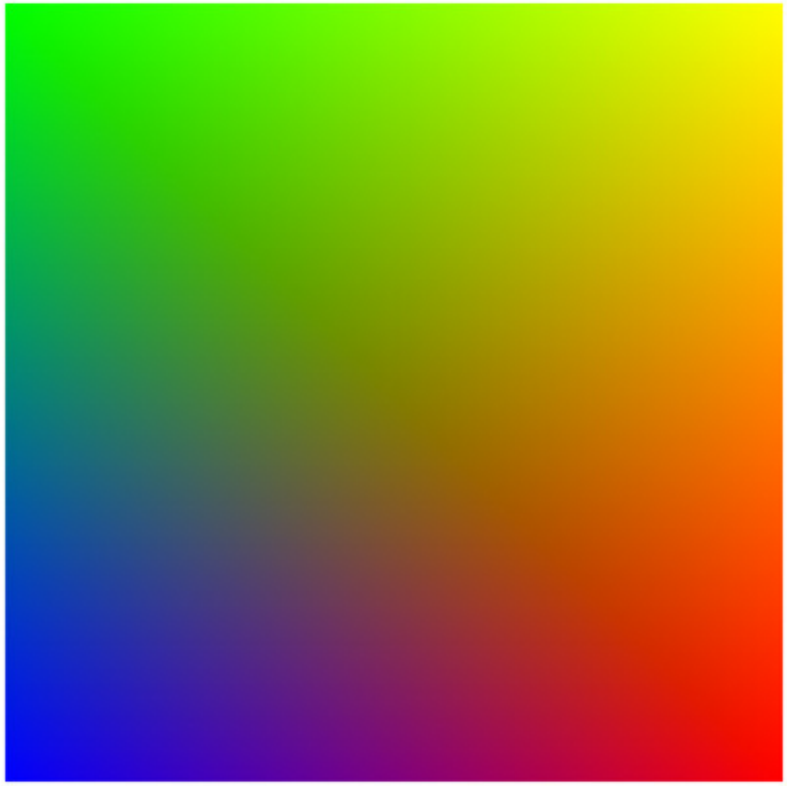


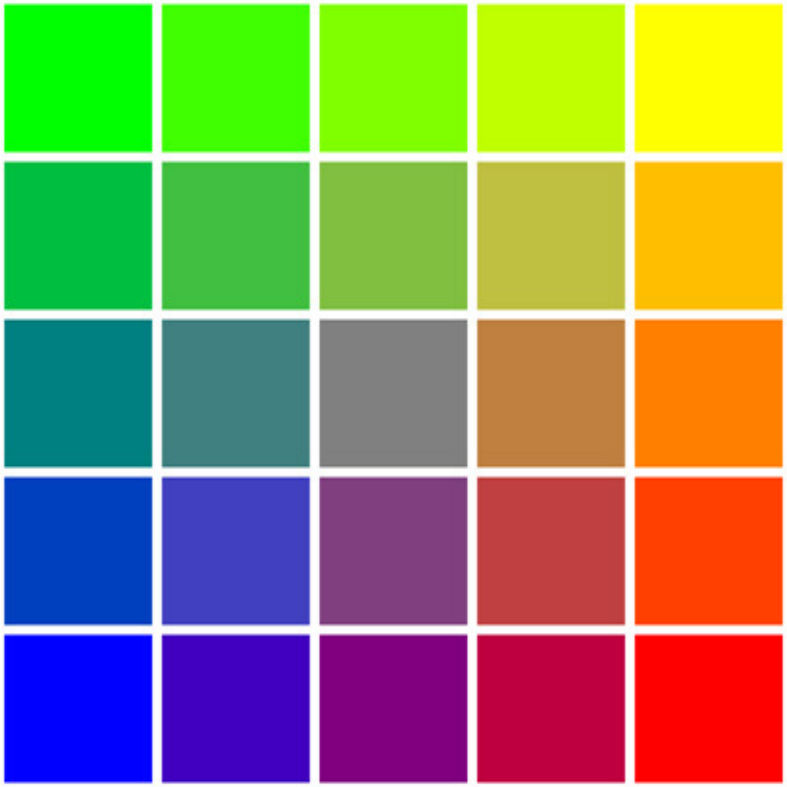
Rainbow color map

- Learned order
- Visually segmented
- Better:
  - Perceptually uniform colormaps (e.g. isoluminant rainbow)
  - Discretize colormap



Color segmentation





Ware, Visual Thinking for Design, 2010.

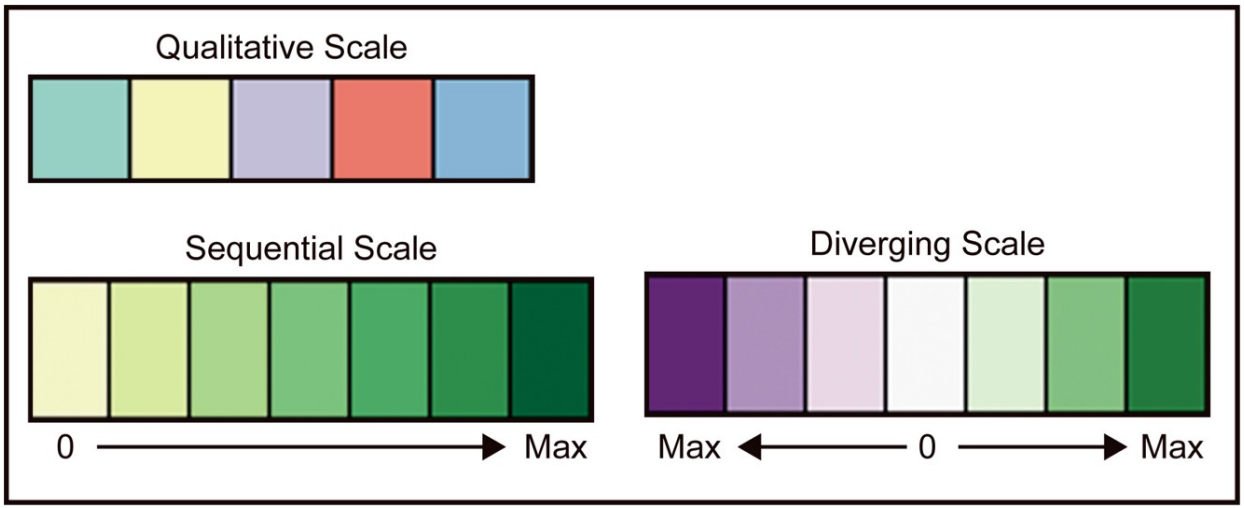
Take-home message

- Quantitative data can be shown with a discrete or continuous colormap
- Use colormaps with a limited hue palette and redundantly vary lightness and saturation
- Use discrete colormaps for accuracy

Brewer scales

Nominal (Categorical)

Ordinal

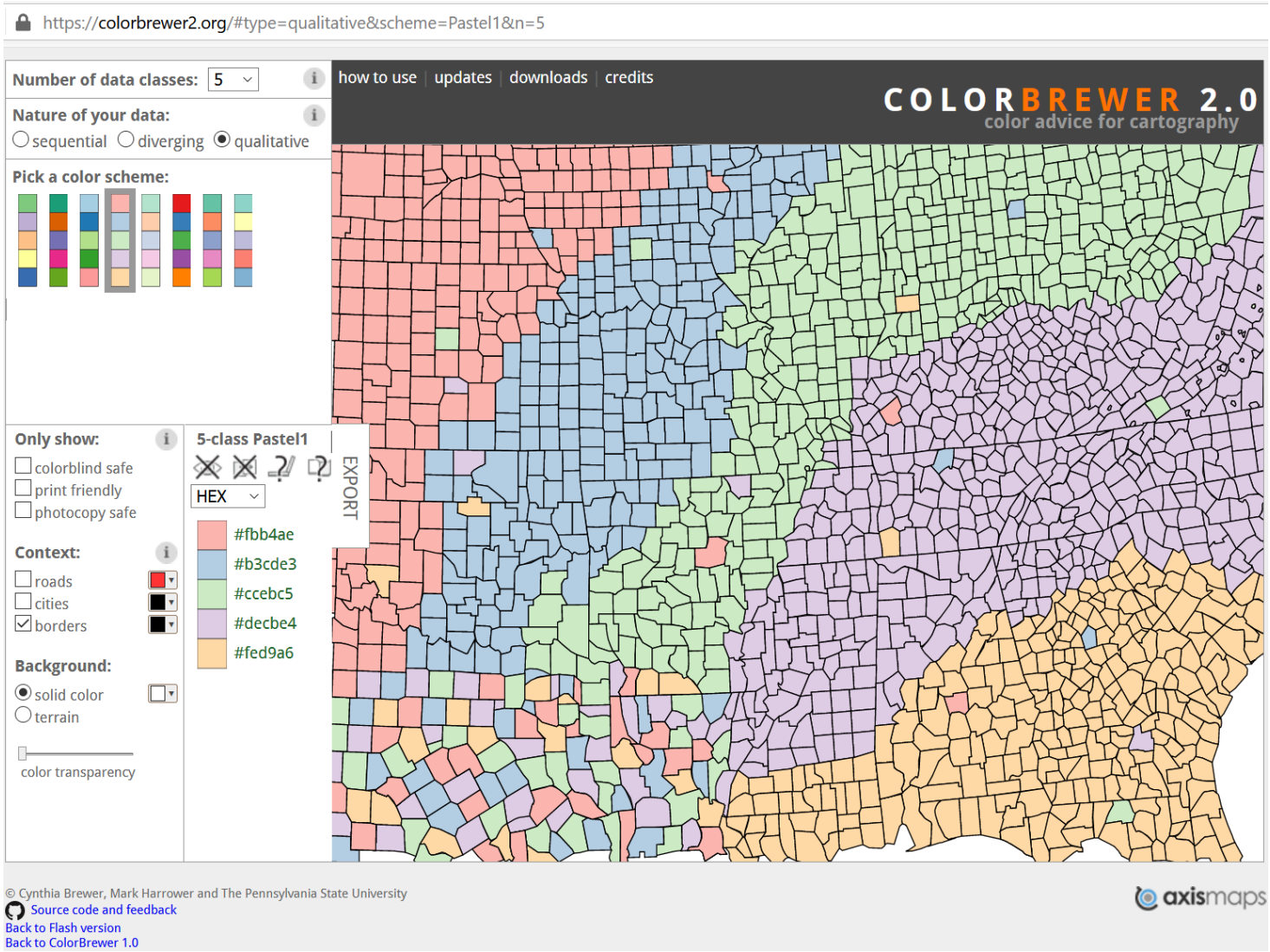


What?

How-much?

Brewer, Color Use Guidelines for Data Representation, 1999

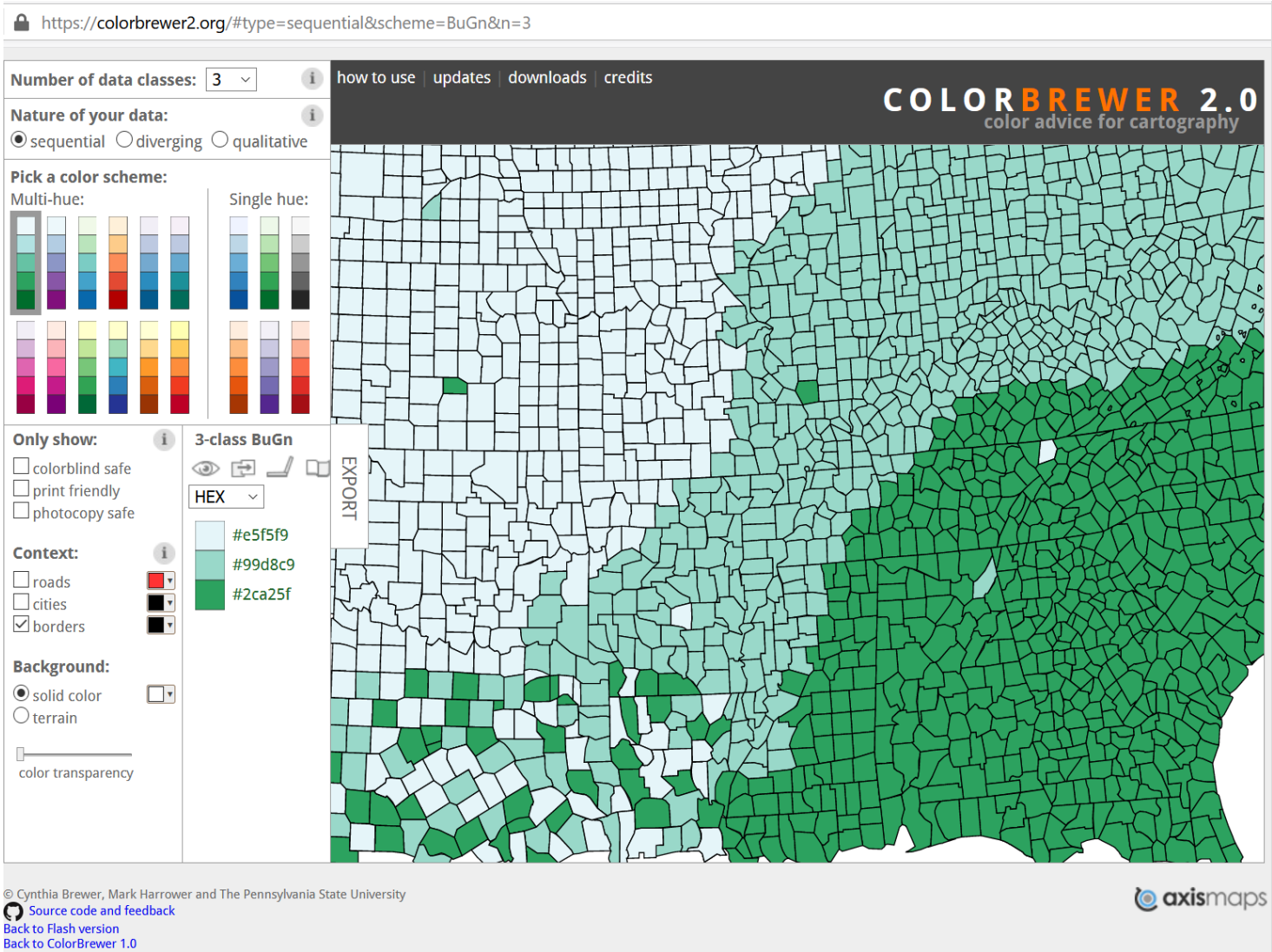
ColorBrewer



Qualitative color schemes for categorical data (*What* channel)

https://colorbrewer2.org

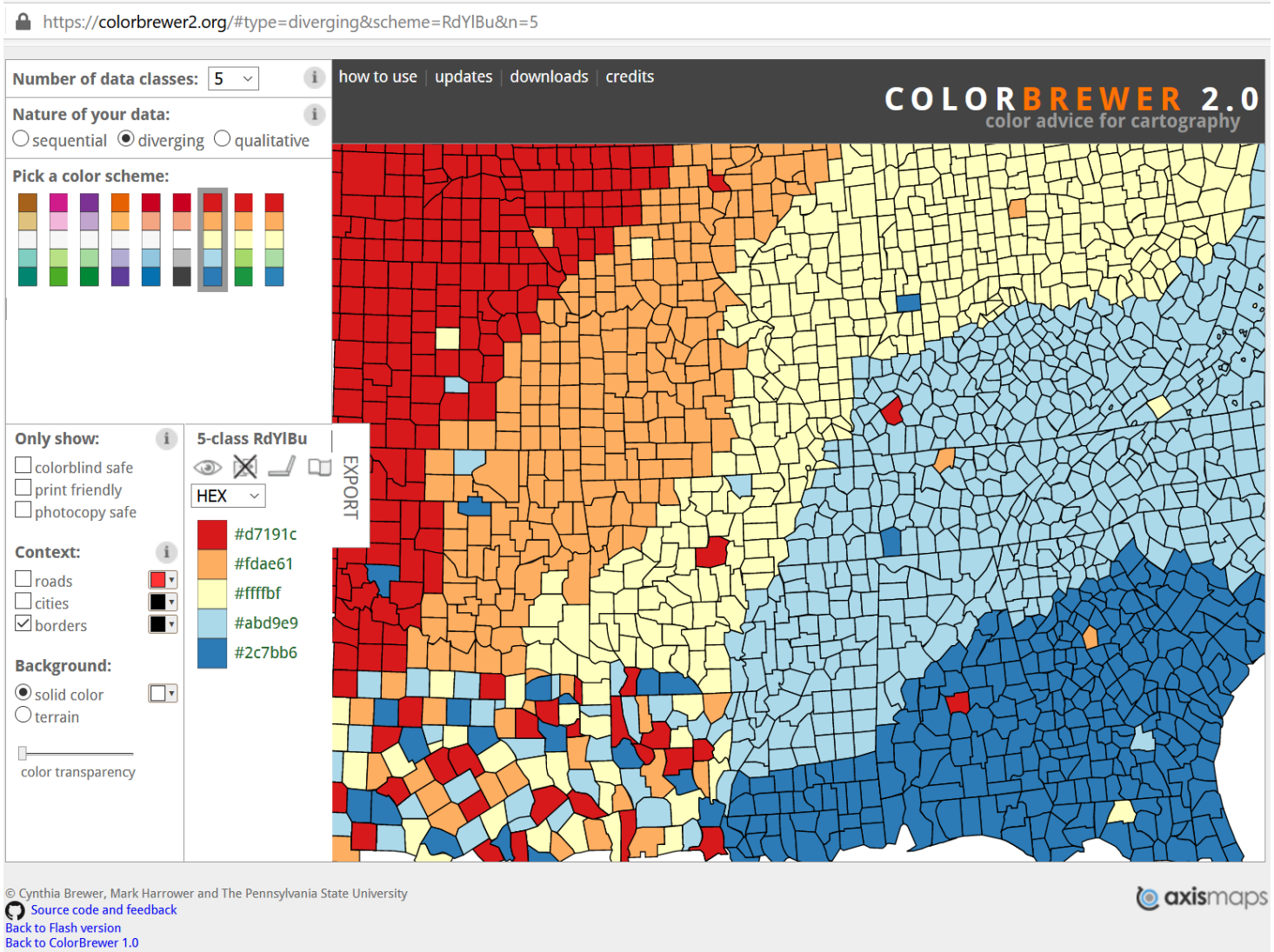
ColorBrewer



Sequential color schemes for ordinal/continuous data (*How-much* channel)

https://colorbrewer2.org

ColorBrewer



Diverging color schemes for ordinal/sequential data diverging from a central value (*How-much* channel)

https://colorbrewer2.org

Tableau example

- Let us use a map to test different color scales by using **Cellular Subscriptions** as color

- Compare
  - A continuous color scale
  - To a discrete (stepped) color scale

Color: Quantitative data

Tableau example

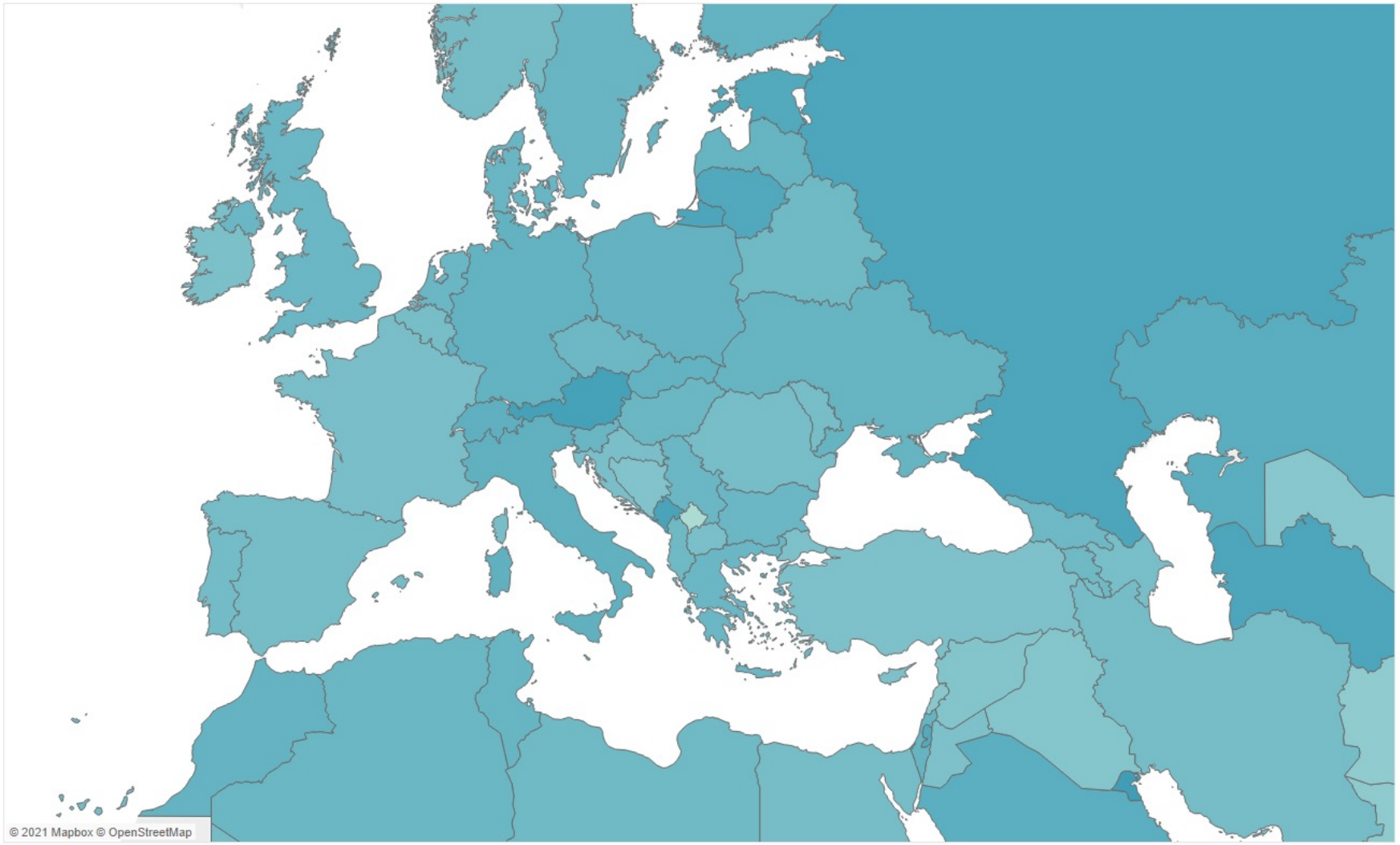


Tableau example

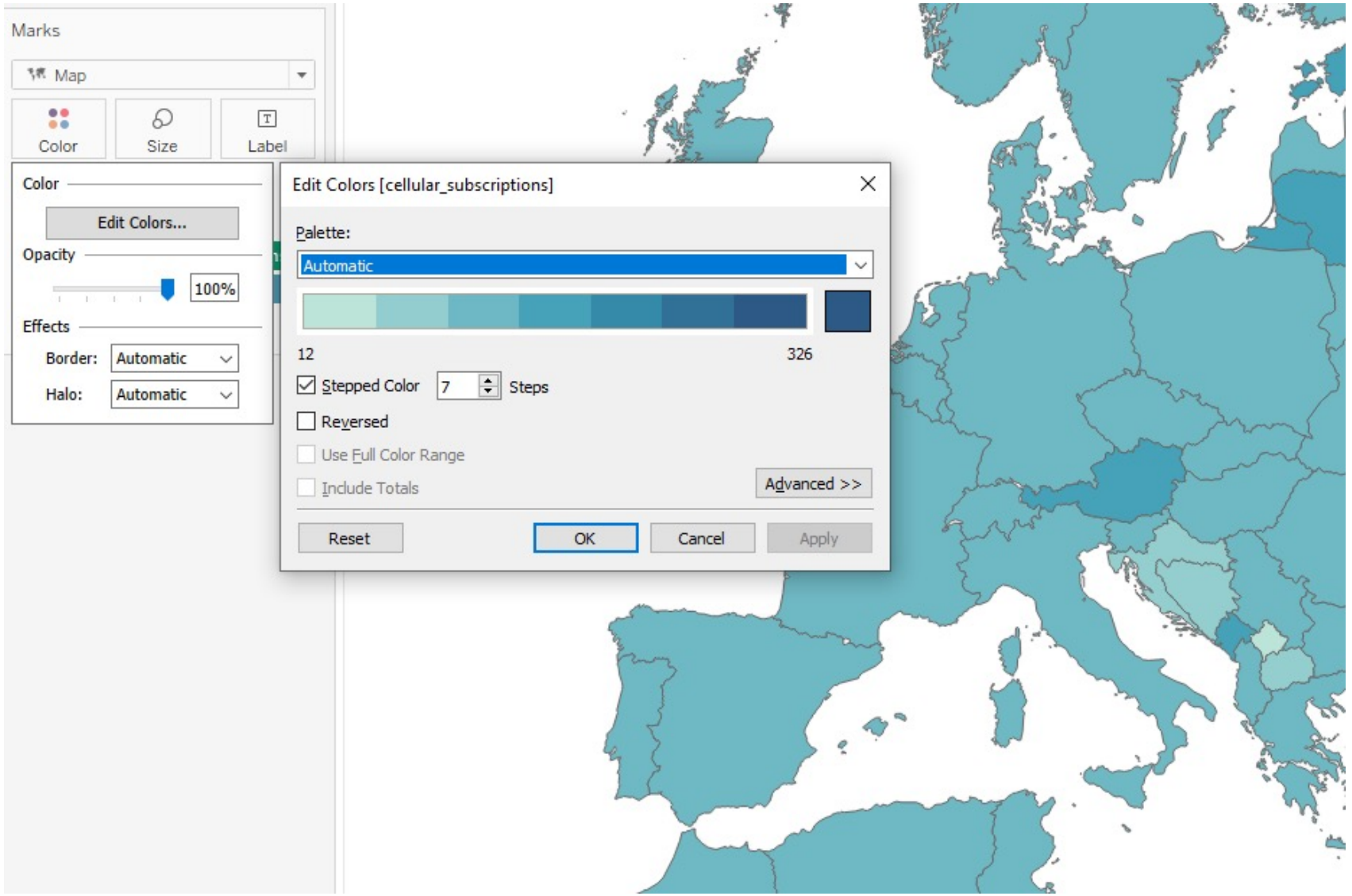
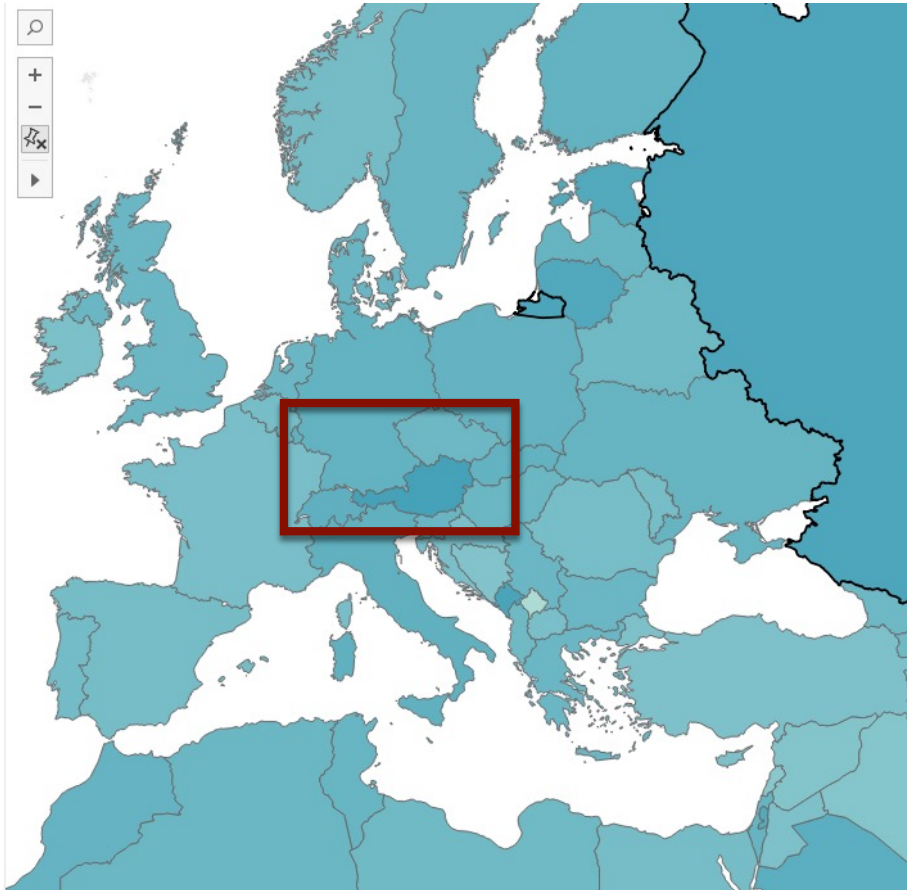
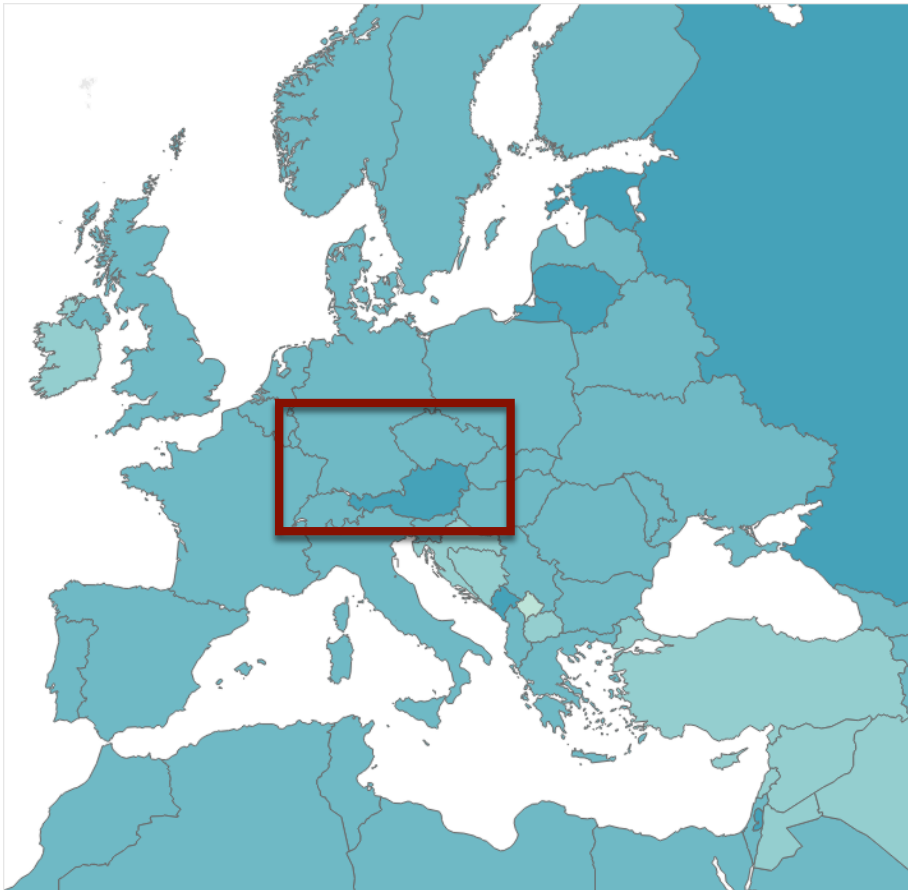


Tableau example

Continuous



Stepped



Color aesthetics:

DANGER!

- Inappropriate use of color can be disasterous to the application

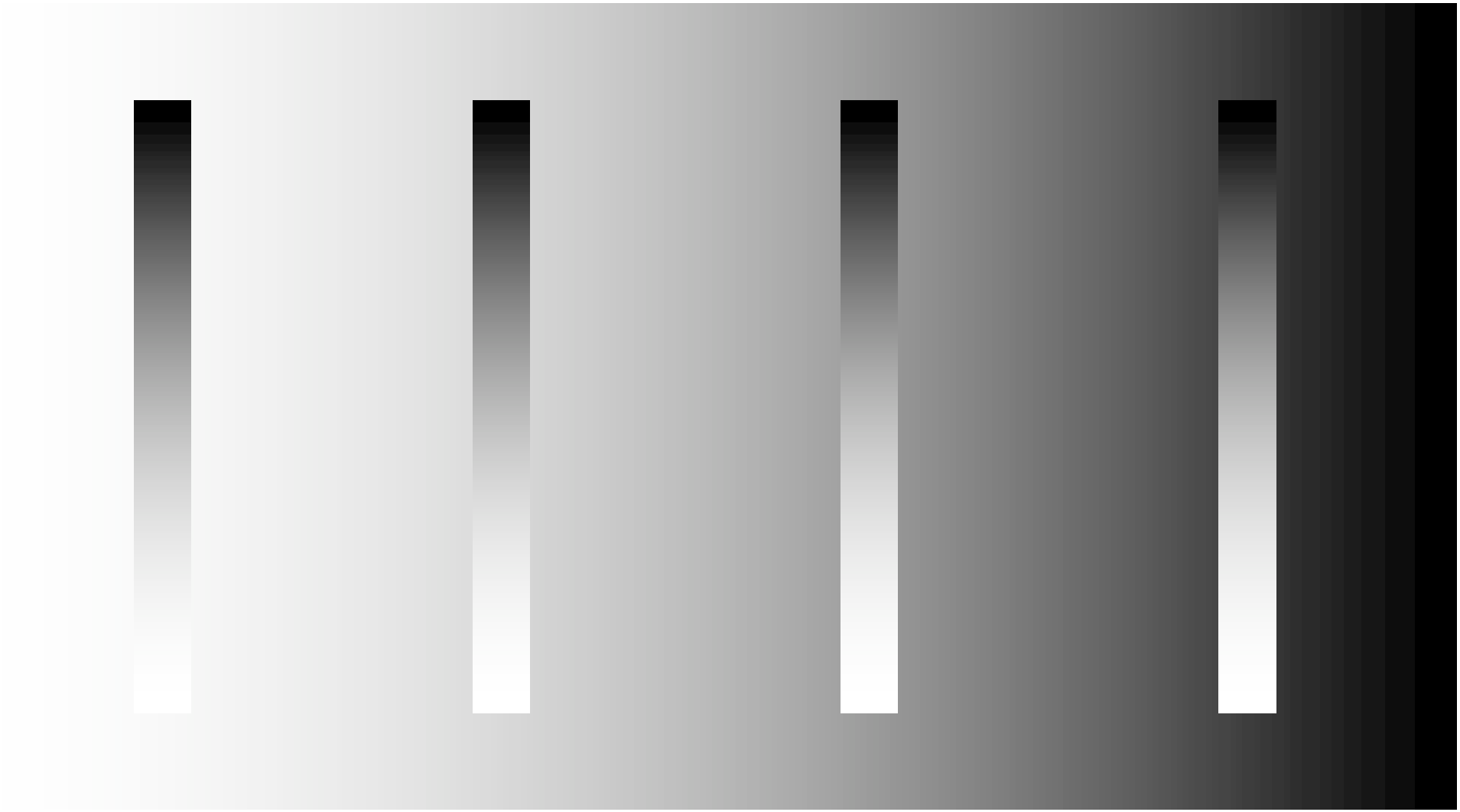
Color aesthetics:

Why should we care?

- Poorly designed color is confusing
  - Creates visual clutter
  - Misdirects attention
- Poor design devalues the information
  - Visual sophistication
  - Evolution of document and web design
- Don Norman:
  “Attractive things work better”

Optical illusions:

Contrast crispening



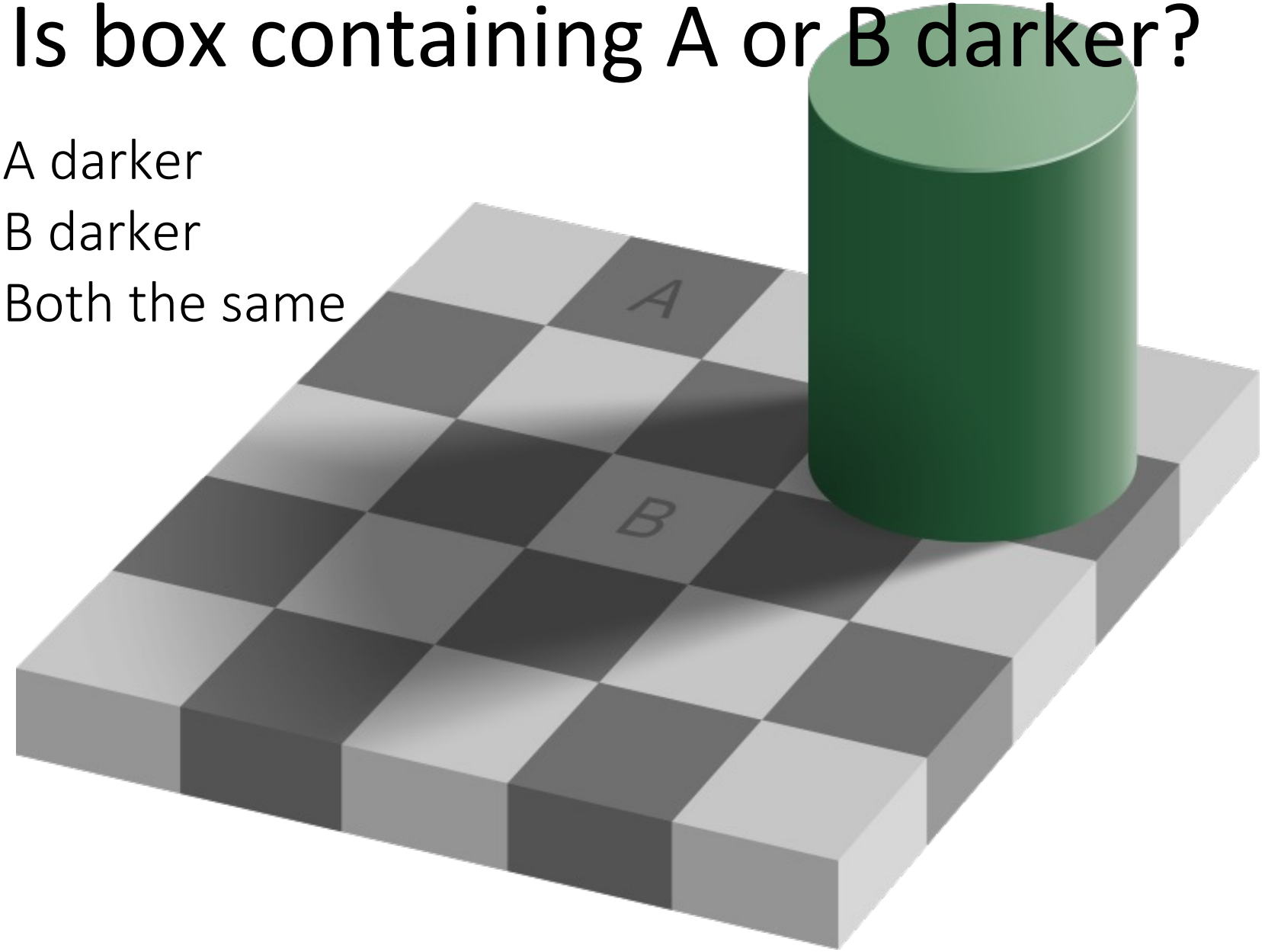
Ware, Information Visualization Perception for Design, p. 90

Is box containing A or B darker?

A darker

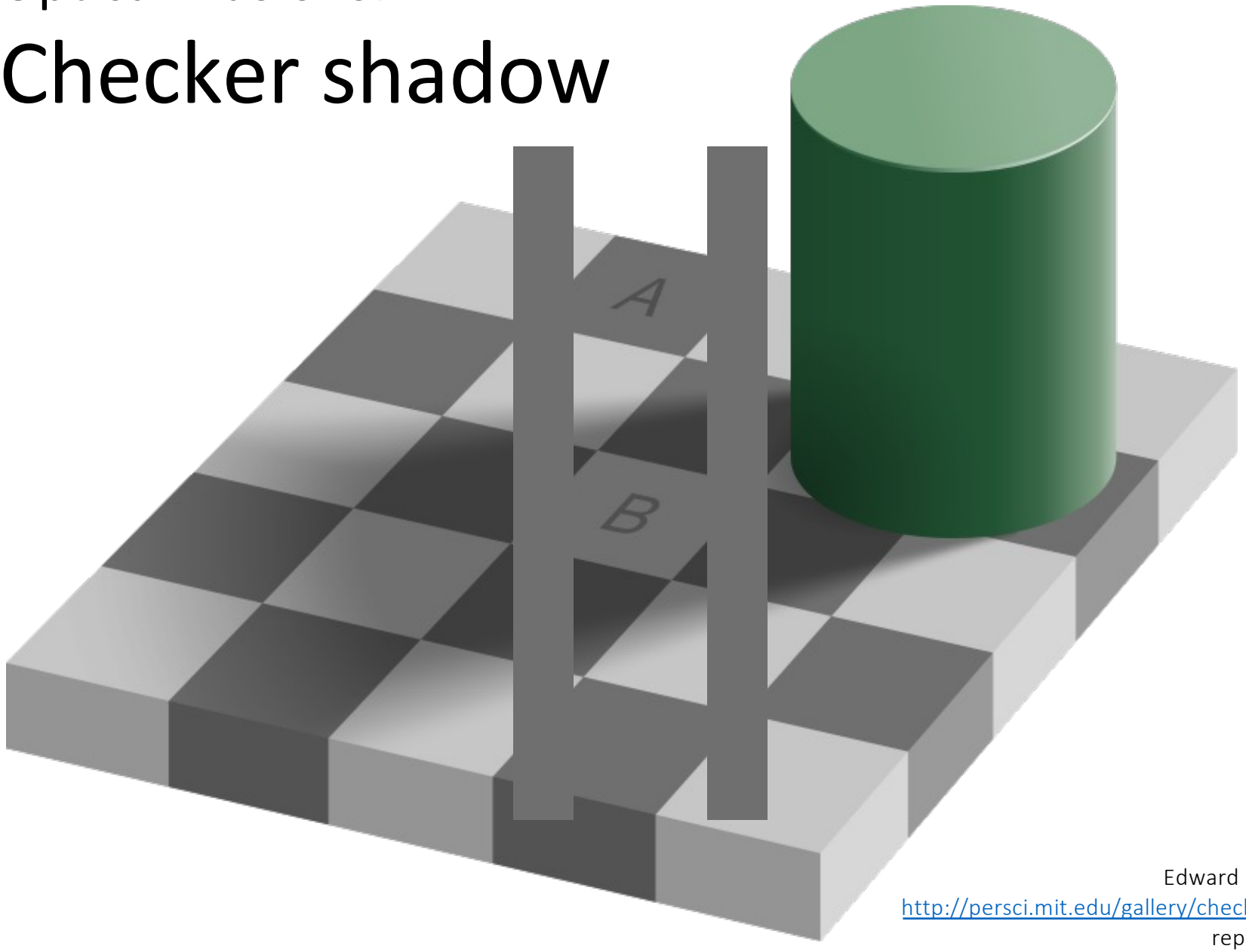
B darker

Both the same



Optical illusions:

Checker shadow



Edward H. Adelson,
http://persci.mit.edu/gallery/checkershadow,
reproduced by
https://en.wikipedia.org/wiki/Checker_shadow_illusion

In conclusion

- Visualizations are broken down into
  - Marks (elements), and
  - Channels (parameters)
  - Data is linked to the channels
- Take care in choosing channels:
  - Expressiveness: Encode all of, and only, information in dataset
  - Effectiveness: Most important data with most accurate channel
  - Which and how many channels (discriminability & separability)
- Be aware of channel characteristics, such as
  - Position
  - Relative vs. absolute judgement
  - Colors (Perceptual issues, e.g. color deficiencies, Aesthetics)