Digitise, Optimise, Visualise: Data Visualization

Peter H. Gruber

July 1-5, 2019

Why data viz?

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

Translation

Color

Danish physicist Tor Nørretranders converted the bandwidth of the senses into computer terms



- □ Data Viz is the highway to the brain.
- \square Brain has incredible graphics processor. \rightarrow Viz as cognitive aid.

Data Visualization 2 / 32

Unique Skillset

Why data viz?

Dataviz

Distinctions

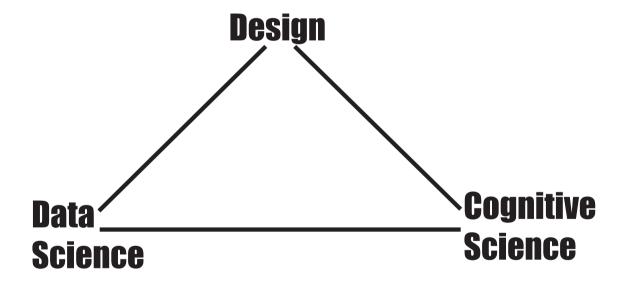
Example

Dataviz = science?

Perception

Translation

Color



Data Visualization 3 / 32

Dataviz

Why data viz? What is data viz? Unique Skillset Translation of data into graphic language to enable . . . Dataviz Distinctions Exploratory data analysis Example (i.e. formualte hypothesis, John Tukey, 1977) Dataviz = science?Confirmation Perception Translation Presentation Color Story telling How is it done? Mapping of data onto graphic objects **Context** is important Application, scientific discipline, Audience: background, data literatcy Culture (e.g. colors)

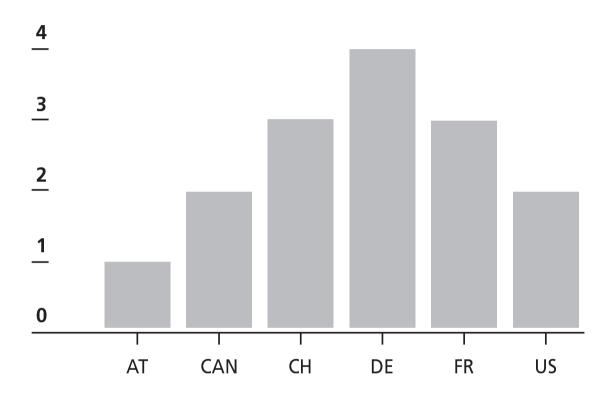
Data Visualization 4 / 32

Distinctions

Why data viz? **Infographics** Unique Skillset Explain a mechanism Dataviz Distinctions Information vizualization Example Dataviz = science?Translate facts into images Perception **Data visualization** \leftarrow this is what we do Translation Color Translate data into graphics Goals: exploration – confirmation – presentation Data art Translate data into graphics with some artistic freedom Goal: beauty Data decoration Add graphical elements to data that have no information content

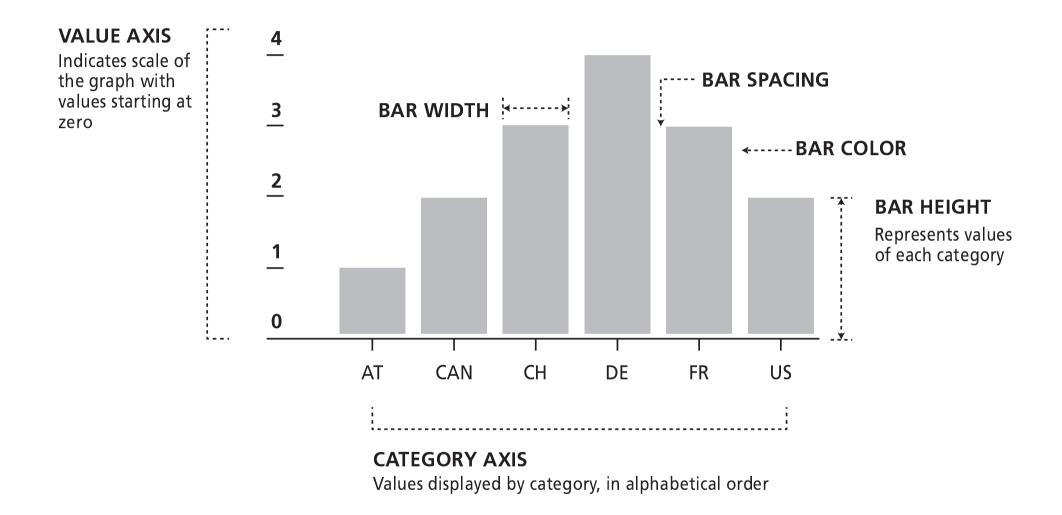
Data Visualization 5 / 32

Example



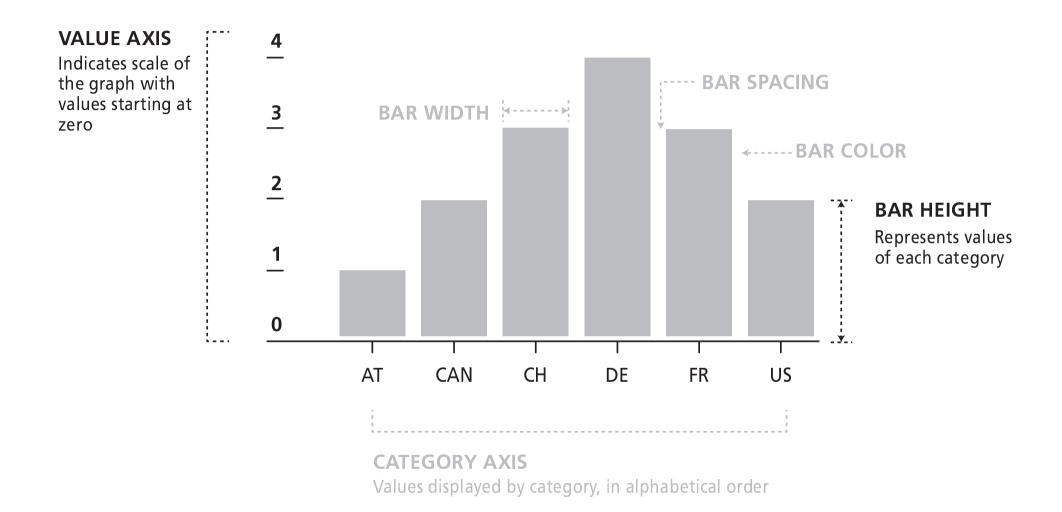
Data Visualization 6 / 32

Example with annotations



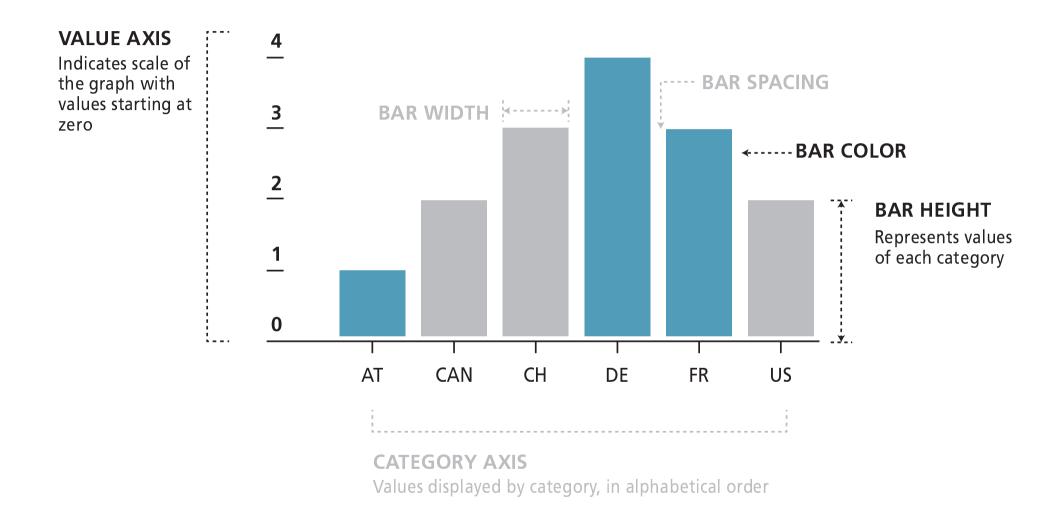
Data Visualization 7 / 32

Information content (1)



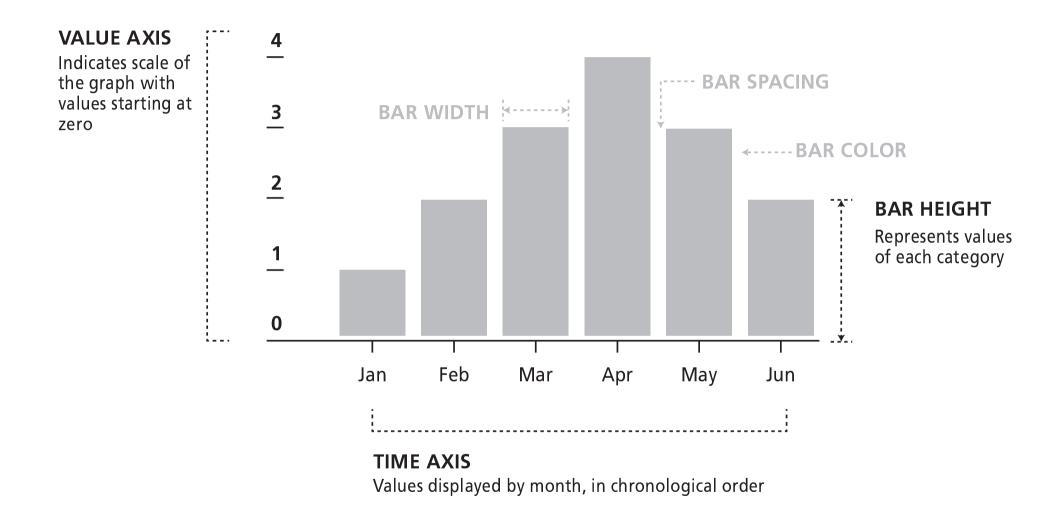
Data Visualization 8 / 32

Information content (2)



Data Visualization 9 / 32

Information content (3)



Data Visualization 10 / 32

Dataviz = science?

Why data viz? Def: A statistic is a calculable function of the data Unique Skillset \rightarrow every visualization is a statistic. Dataviz Distinctions Data visualizations are falisifiable. Example \triangleright Dataviz = science? Scientific foundation in cognitive science + statistics. Perception Translation Good viz. leads to statistically superior inference. Color Design = finding the best possible (most efficient) solution for a given problem.

Data Visualization 11 / 32

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

> Perception

The night sky ...

... and how we see it

Gestalt Principles

Translation

Color

Perception

Data Visualization 12 / 32

The night sky ...

Why data viz? Unique Skillset

Dataviz

Distinctions

Example

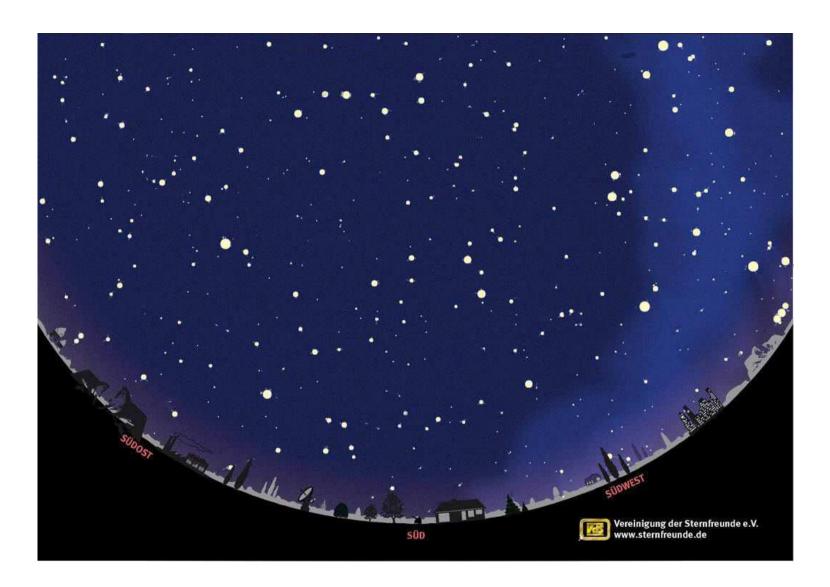
Dataviz = science?

Perception

▶ The night sky and how we see itGestalt Principles

Translation

Color



Data Visualization 13 / 32

... and how we see it

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

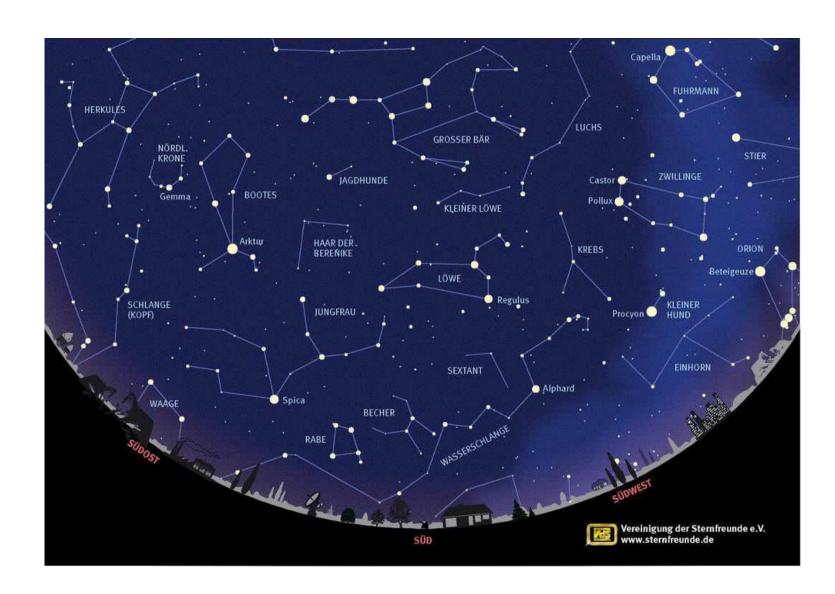
The night sky ...

... and how we see

Gestalt Principles

Translation

Color



14 / 32 Data Visualization

Gestalt Principles

Why data viz?
Unique Skillset
Dataviz
Distinctions
Example
Dataviz = science?

Perception
The night sky ...
... and how we see it
Gestalt Principles

Translation

Color

The brain wants to find sense in our environment

- \square Christian von Ehrenfels (1890)
- □ Law of Prägnanz:"We order our experience in a manner that is . . .
 - regular,
 - orderly,
 - symmetrical,
 - and simple."
- □ Pre-attentive processing
- → Exercise: seeing plots

Gestalt Principles

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

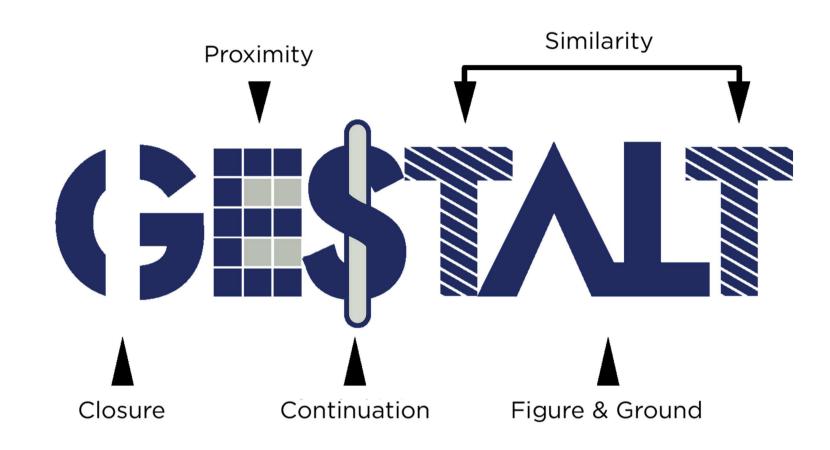
Perception

The night sky ...

... and how we see it

Translation

Color



More: symmetry, common fate, past experience, common region, connectedness

→ Exercise: how can we make use of the gestalt principles?

Data Visualization 16 / 32

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

▶ Translation

Aesthetics
The translation

process: simple case

The translation

process: advanced

case

Aesthetics

Visual Encodings

Color

The translation problem

Data Visualization 17 / 32

Aesthetics

Why data viz?
Unique Skillset
Dataviz
Distinctions
Example

Dataviz = science?

Perception

Translation

Color

Aesthetics
The translation
process: simple case
The translation
process: advanced
case
Aesthetics
Visual Encodings

Q: How to translate numbers into graphics?

- ☐ Jacques Bertin (cartographer, 1967), Semiology of Graphics List of "les variables de l'image"
- \square Leyland Wilkinson (2005), Grammar of graphics **Aesthetics** = mappings of the data

Data Visualization 18 / 32

The translation process: simple case

Why data viz?
Unique Skillset
Dataviz
Distinctions
Example
Dataviz = science?
Perception
Translation
Aesthetics

The translation process: simple

case

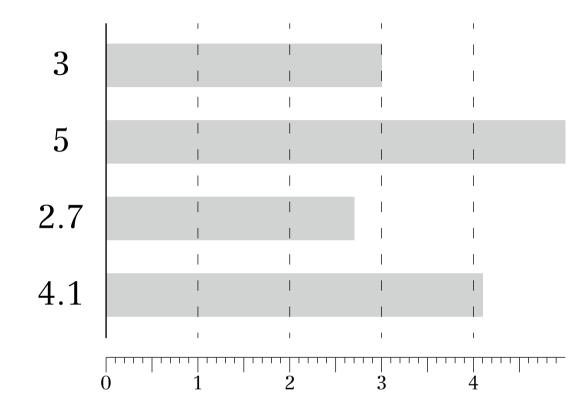
The translation process: advanced

case

Aesthetics

Visual Encodings

Color



 $Data \longrightarrow Length$

The translation process: advanced case

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

Translation

Aesthetics

The translation

process: simple case

The translation

process: advanced

Aesthetics

Visual Encodings

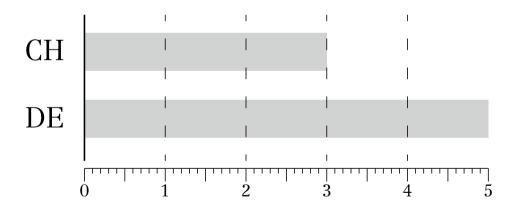
Color

Data DE, DE, CH, CH, DE, CH, DE, DE

Statistics 3xCH, 5xDE (count)

Aesthetics

length



Data \longrightarrow Statistics \longrightarrow Length

Aesthetics

Why data viz? Unique Skillset Dataviz Distinctions Length Width Orientation Intensity Example Dataviz = science?Perception Translation Aesthetics The translation process: simple case The translation process: advanced Shape Size **Enclosure** 2-D position case > Aesthetics Visual Encodings Color

Data Visualization 21 / 32

Visual Encodings

Why data viz? Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

Translation

Aesthetics

The translation

process: simple case
The translation

process: advanced

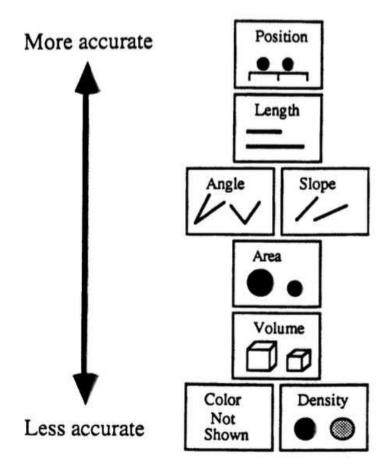
case

Aesthetics

Color

Which encodings work best?

Jock D. Mackinlay (1986): Automating the design of graphical presentations of relational information



Data Visualization 22 / 32

Visual Encodings (2a)

Example	Encoding	Ordered	Useful values	Quantitative	Ordinal	
000	position, placement	yes	infinite	Good	Good	
1, 2, 3; A, B, C	text labels	optional alpha or num	infinite	Good	Good	
	length	yes	many	Good	Good	
. 0	size, area	yes	many	Good	Good	
/_	angle	yes	medium	Good	Good	
	pattern density	yes	few	Good	Good	
	weight, boldness	yes	few		Good	
	saturation, brightness		few		Good	

Data Visualization 23 / 32

Visual Encodings (2b)

Example	Encoding	Ordered	Useful values	Quantitative	Ordinal	Categorical
	color	no	few (<20)			Good
	shape, icon	no	medium			Good
	pattern texture	no	medium			Good
	enclosure, connection	no	infinite			Good
====	line pattern	no	few			
5	line endings	no	few			
	line weight	yes	few		Good	

From: Julie Steele, Noah Ilinsky: Designing Data Visualizations, Safari books

Data Visualization 24 / 32

Why data viz?

Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

Translation

Color

Humans and color

RGB Color Wheel

Alternatives to RGB

Color scheme

Color Brewer

Final considerations

Color

Data Visualization 25 / 32

Humans and color

Why data viz?
Unique Skillset
Dataviz
Distinctions
Example
Dataviz = science?

Perception

Translation

Color

Humans and color
RGB Color Wheel
Alternatives to RGB
Color scheme
Color Brewer
Final considerations

Perception

- ☐ Human retina has 100-150 mio light-sensitive cells
 Only 7 mio are color-sensitive
- ☐ Approx. 8% of men and 0.5% of women are color-blind Online test: enchroma.com/test/instructions
- extstyle ext
- □ Discrete: Can distinguish, recognize and name 100s of colors

Communication

- □ Names: "blue", "sky blue", "steel blue", "navy blue"
- ☐ Formula: how much Red, Greed and Blue light is in a color?
- \square RGB is color mixing by light: (0,0,0) = black, (1,1,1) = white

RGB Color Wheel

Why data viz? Unique Skillset

Dataviz

Distinctions

Example

Dataviz = science?

Perception

Translation

Color

Humans and color

RGB Color Wheel

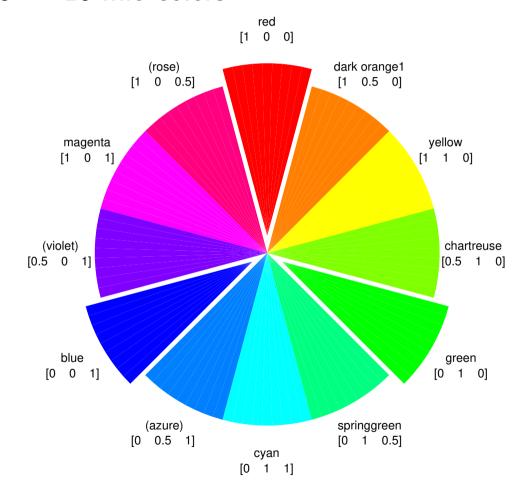
Alternatives to RGB

Color scheme

Color Brewer

Final considerations

 \square Most computers: 256 shades of R/G/B each $\rightarrow 256^3 \approx 16$ mio colors



Alternatives to RGB

Why data viz? Unique Skillset

Dataviz

Distinctions

Example

 $\mathsf{Dataviz} = \mathsf{science}?$

Perception

Translation

Color

Humans and color RGB Color Wheel Alternatives to

▶ RGB

Color scheme

Color Brewer

Final considerations

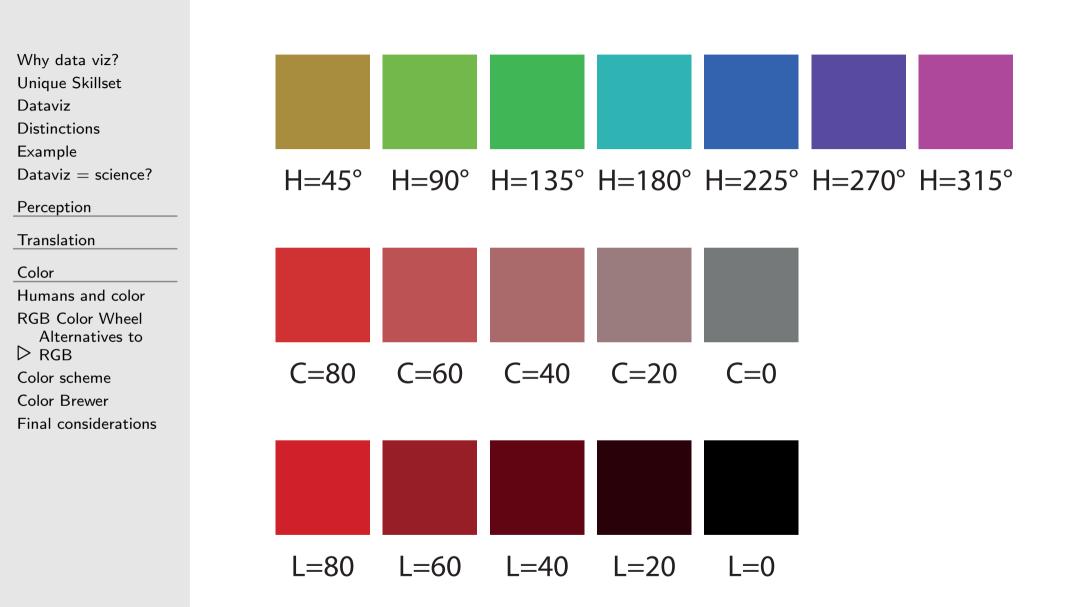
Is the RGB system the best?

- ☐ Technically motivated
- □ Color Space of the eye is(1) black-white, (2) blue-yellow, (3) red-green
- \Box Luminance (1): L = 0.31R + 0.59G + 0.1B
 - → RGB coloes with equal color sums have different brightness

HCL Model

- □ Perception-based color space HCL
 - \mathbf{H} ue [0-360] = angle in the color wheel
 - \mathbf{C} roma [0-100] = color intensity between grey and intense
 - Luminance [0-100] = perceived brightness

HCL Model



Data Visualization 29 / 32

How to choose a color scheme for visualization?

Why data viz? Default colors not a good idea Unique Skillset Dataviz Type of variable \rightarrow color scheme Distinctions Binary – contrast Example Dataviz = science? Qualitative – distinction Perception Diverging – continuous scale with neutral center Translation Sequential – continuous scale Color Humans and color Possible to combine 2 schemes RGB Color Wheel Alternatives to RGB Cultural connotations Color scheme Hot - cold Color Brewer Final considerations Countries, companies, continents Color blindness, printing process, limitations of beamers Predefined schemes: colorbrewer.py

See www.colorbrewer2.org

Data Visualization 30 / 32

Color Brewer Combinations

Why data viz? Unique Skillset

. Dataviz

Distinctions

Example

Dataviz = science?

Perception

Translation

Color

Humans and color

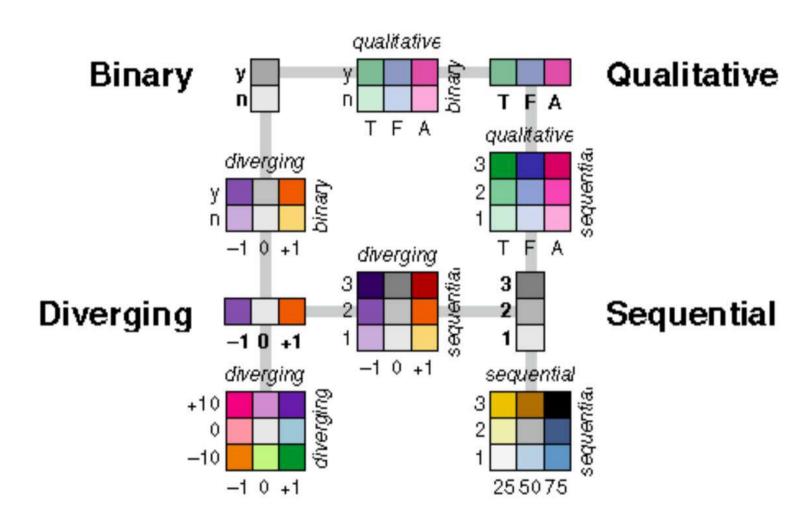
RGB Color Wheel

Alternatives to RGB

Color scheme

Color Brewer

Final considerations



See colorbrewer2.org

Source www.personal.psu.edu/faculty/c/a/cab38/ColorSch/Schemes.html

Final considerations

The graphical method has considerable superiority for the exposition of statistical facts over the tabular.

A heavy bank of figures is grievously wearisome to the eye, and the popular mind is as incapable of drawing any useful lessons from it as of extracting sunbeams from cucumbers.

Arthur & Henry Farquhar

in Economic and Industrial Delusions (1891)

Data Visualization 32 / 32