Data Visualization

VISUAL PERCEPTION (1)

Outline

Motivation

Memory

Visual encoding
- o Channel accuracy
- o Channel discriminability
- o Channel salience (pop-out)
- o Channel separability
- o Grouping

Visual order

Motivation
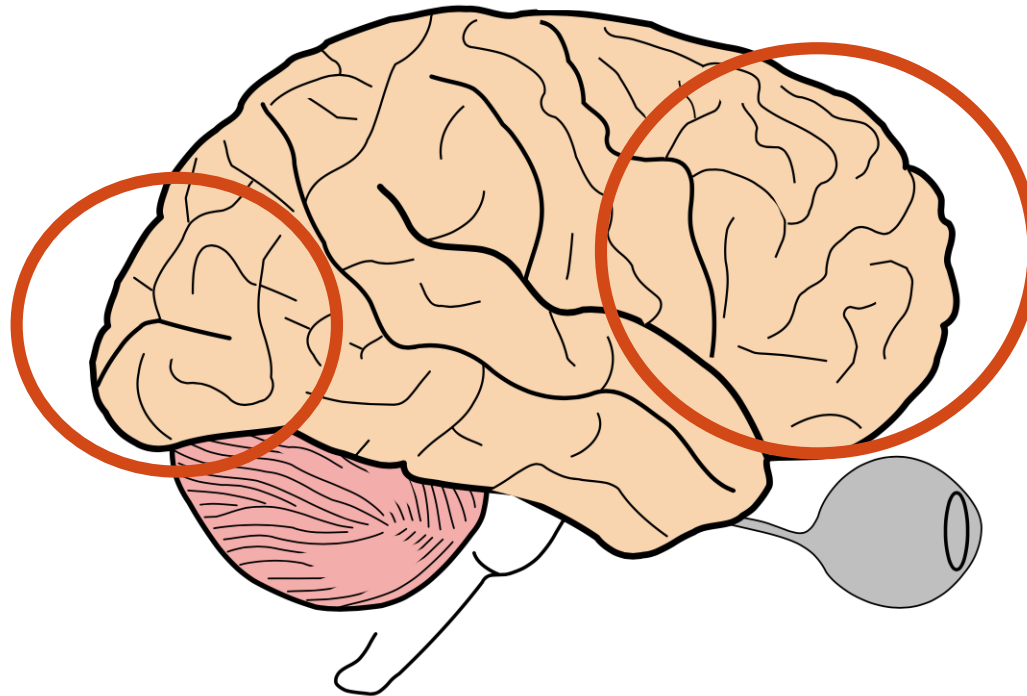
Why does visualization work?

Perception vs. cognition

Seeing (visual perception) is extremely fast and efficient

Thinking (cognition) is much slower and less efficient

Data visualization is effective because it shifts the balance between cognition and perception to take fuller advantage of the brain's abilities
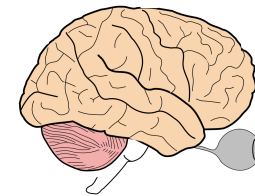
4

Understanding visual perception

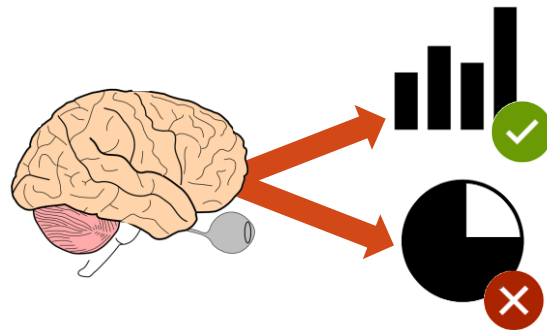One might think that the quality of a visualization is a matter of subjective taste

But visual perception follows specific rules derived from how the brain works

Understanding visual perception

Understanding visual perception enables to make informed decisions about visualization design

The space of possibilities is huge – you need something to guide you in the choices you make
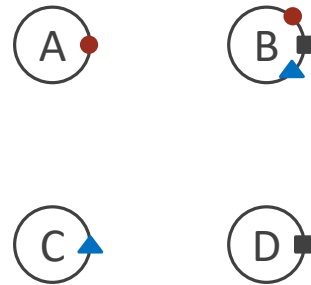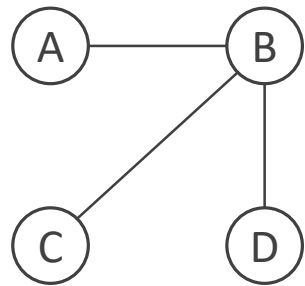
And you will read this last

You will read this first

And then you will read this
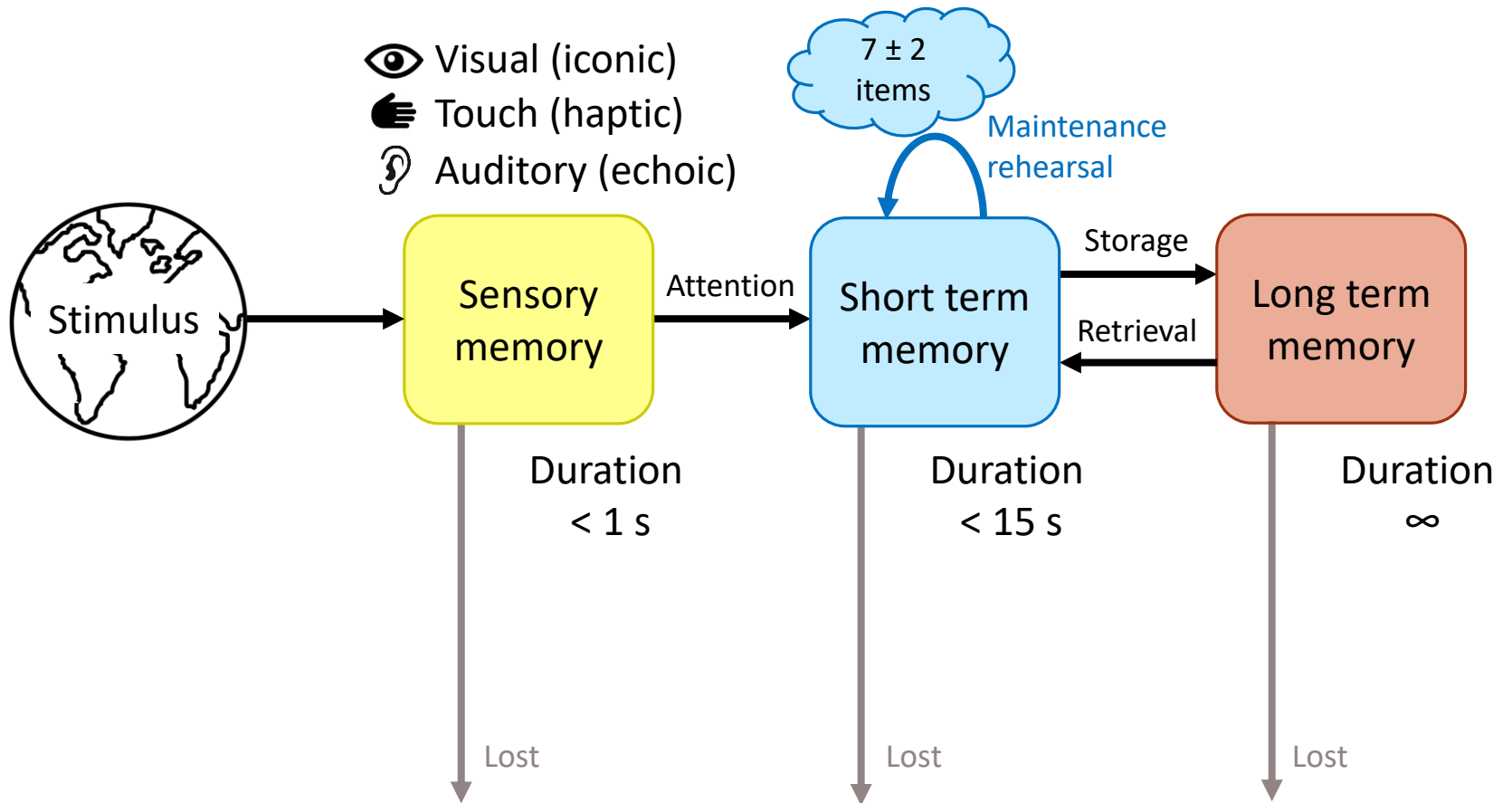
Then this one

Which is easier to understand?

Memory

Memory

Visual (iconic)
Touch (haptic)
Auditory (echoic)

7 ± 2 items

Maintenance rehearsal

Stimulus

Sensory memory

Short term memory

Long term memory

Attention

Storage

Retrieval

Duration < 1 s

Duration < 15 s

Duration ∞

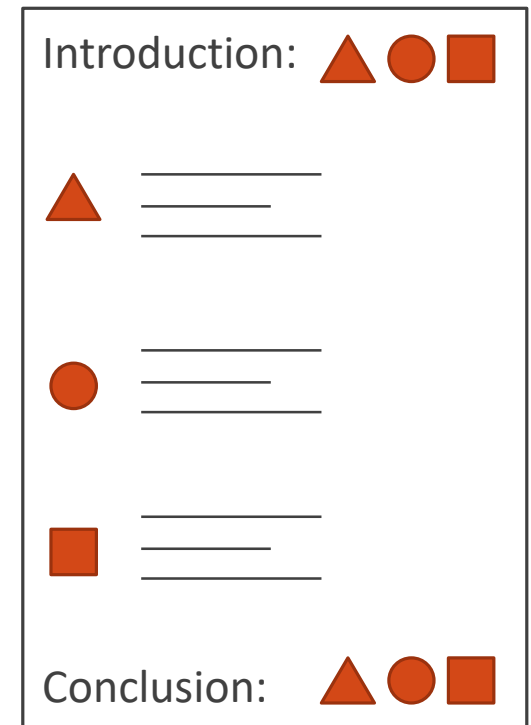Lost

Lost

Lost

Memory: Implications for design/presentation

Do not display more than 7 ± 2 items/categories

The power of repetition: Bing, Bang, Bongo
- o Introduce what you are going to tell the audience (Bing)
- o Tell the audience (Bang)
- o Summarize what you just told them (Bongo)

Be consistent (always follow the same order) ▲●■

Introduction: ▲●■

▲ ═══

● ═══

■ ═══

Conclusion: ▲●■

Visual encoding

Visual encoding

Mapping between data properties and graphical properties
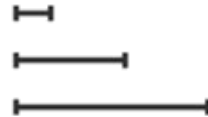
Data attributes ➡ Visual channels

Data attributes

→ Ordered

   → Ordinal   → Quantitative

→ Categorical

Visual channels

length (1D size)		colour hue	
angle		texture density	
curvature		texture pattern	
shape		position (2D)	
area (2D size)		depth (3D position)	
volume (3D size)		motion	
lightness black/white		blur/sharpness	
colour saturation		containment	
transparency		connection	
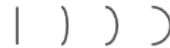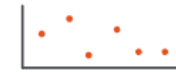
E. J. Maguire. *Systematising Glyph Design for Visualization*, PhD Thesis, University of Oxford, 2014.

Visual channels
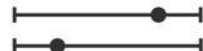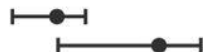
Channel properties
- o Expressiveness – what can be expressed with a channel
- o Effectiveness – how well it can be expressed

Channels that can express order (magnitude channels)

→ Ordered

→ Ordinal   → Quantitative

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)

Channels that can express categories (identity channels)

→ Categorical

╋ ● ■ ▲

Spatial region

Color hue

Motion

Shape ╋ ● ■ ▲

Channel effectiveness

o Single channel

   o Accuracy (estimating magnitude)

   o Discriminability (number of values that can be distinguished)

o Multiple channels

   o Salience or pop-out (attracting attention)

   o Separability (interference between channels)

   o Grouping (pattern formation)

Channel accuracy

Channel accuracy

Steven's Psychophysical Power Law: $S = I^N$

Perceived Sensation vs. Physical Intensity

Electric Shock (3.5)
Saturation (1.7)
Length (1)
Area (0.7)
Depth (0.67)
Brightness (0.5)
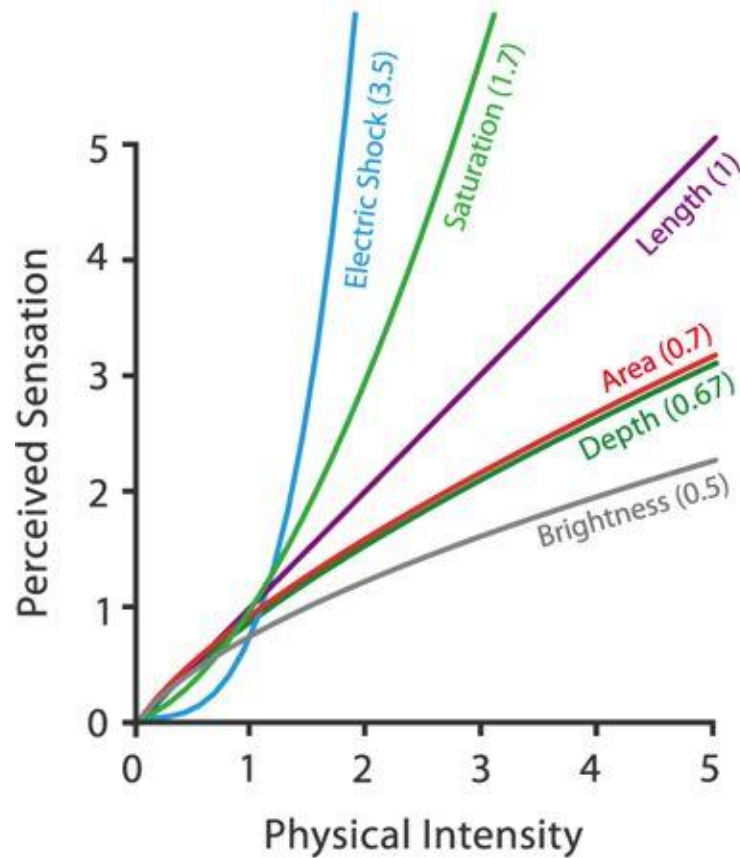
Published in 1957

Channel accuracy

Experiments in graphical perception by Cleveland and McGill in 1983
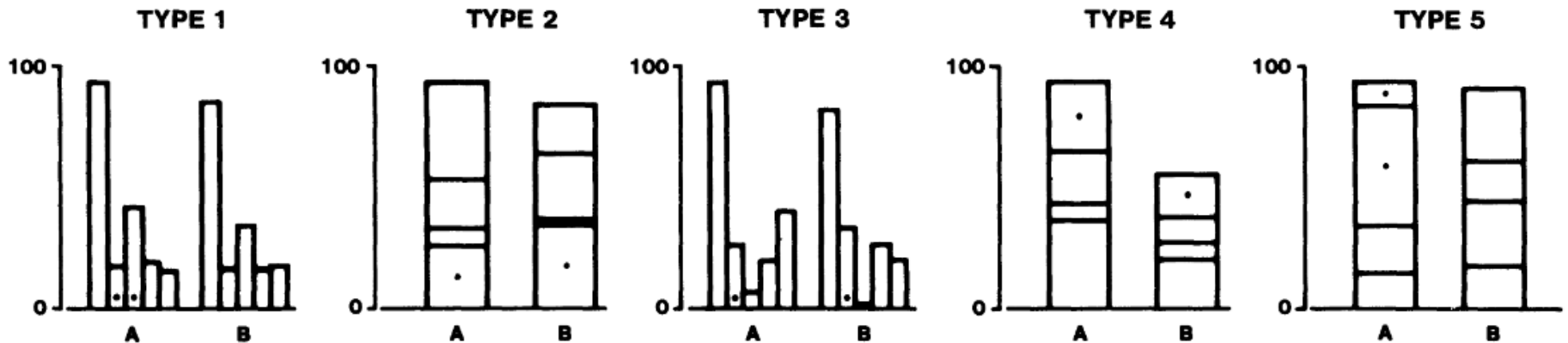
Which is smaller?
How much smaller is it?
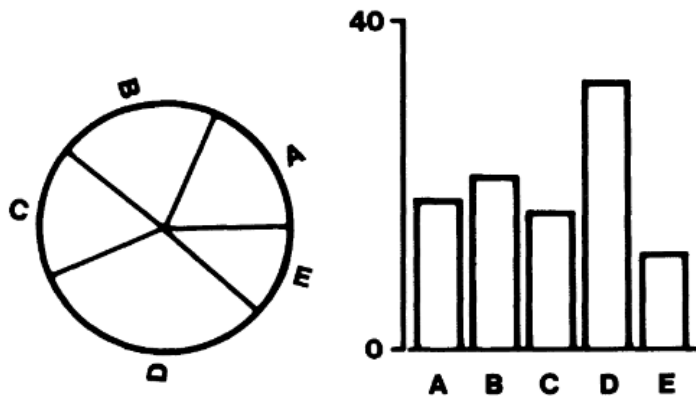
W. S. Cleveland, R. McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 1984

22

Channel accuracy

How much bigger is B compared to A?

W. S. Cleveland, R. McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. Journal of the American Statistical Association, 1984

Channel accuracy

Cleveland & McGill's Results (1984)

Heer and Bostock Crowdsourced Results (2010)
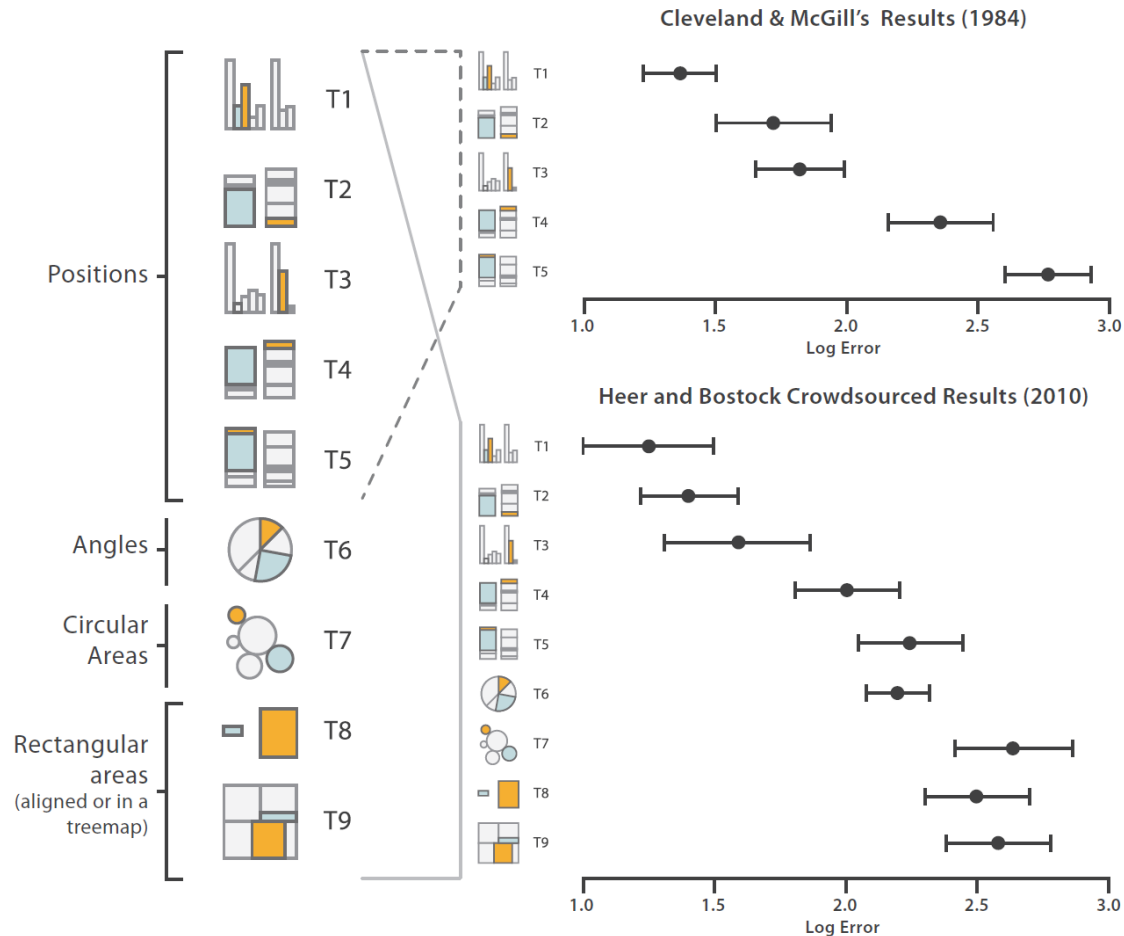
Channel accuracy

Recent experiments by Skau and Kosara show that pie charts are not read by angle

Channel accuracy: Implications for design

Magnitude Channels: Ordered Attributes
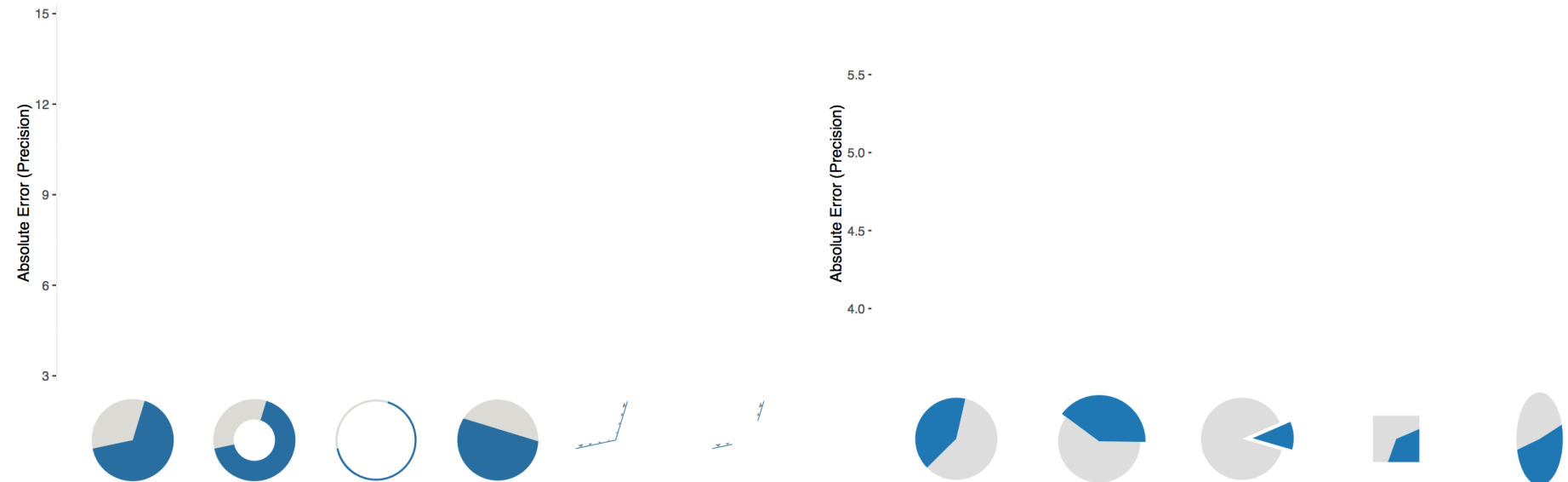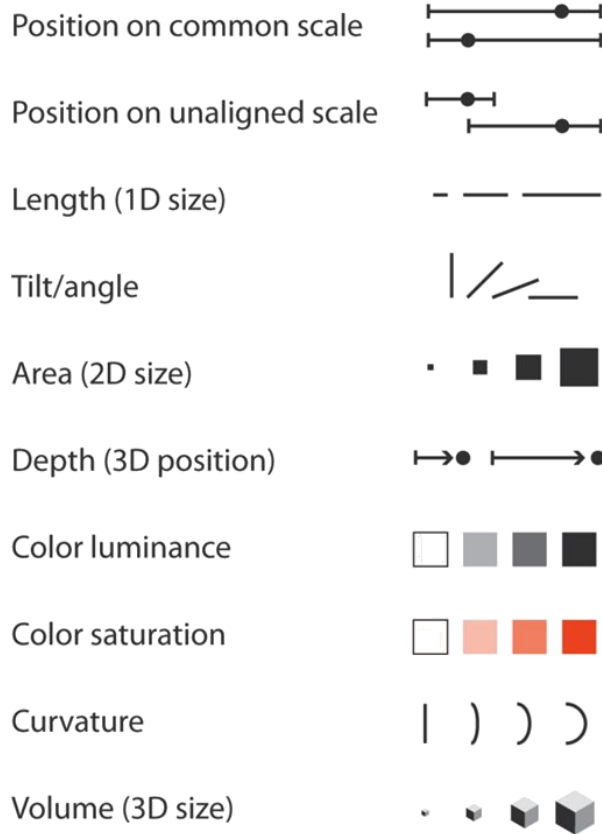- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance
- Color saturation
- Curvature
- Volume (3D size)

Identity Channels: Categorical Attributes
- Spatial region
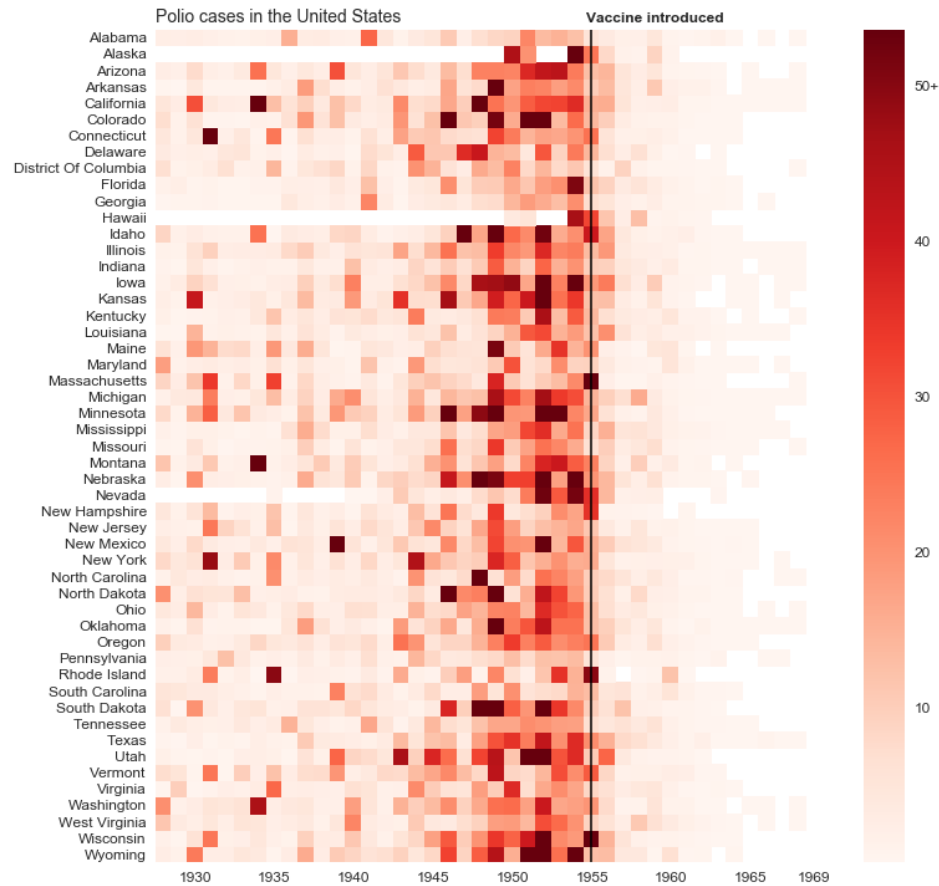- Color hue
- Motion
- Shape

Effectiveness: Most ▲ / Least ▼

Channel accuracy: Limitations

Specific to comparing and estimating magnitudes – not everything in data visualization is about magnitudes

Trade accuracy for something else, for example, scalability

Polio cases in the US

Data source: Project TYCHO (tycho.pitt.edu) | Author: Randy Olson (randalolson.com / @randal_olson)

Channel discriminability

Channel discriminability

How many distinct values can be distinguished within a channel
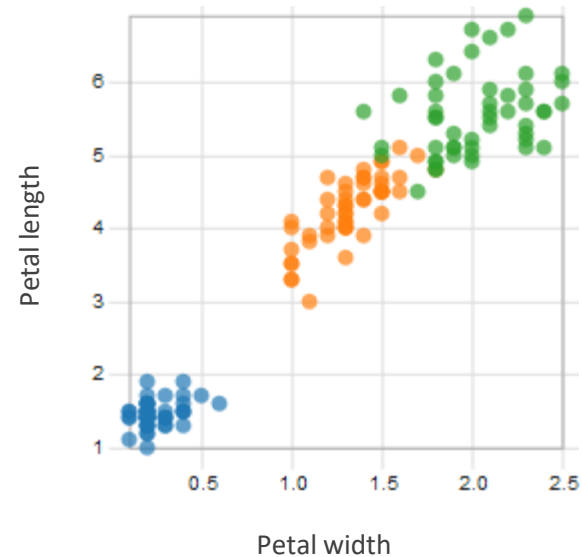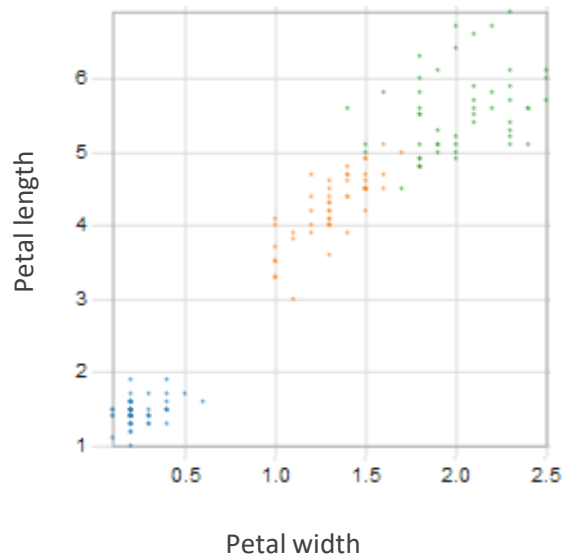
Discriminability depends on
- Channel properties (similar to accuracy)
- Size
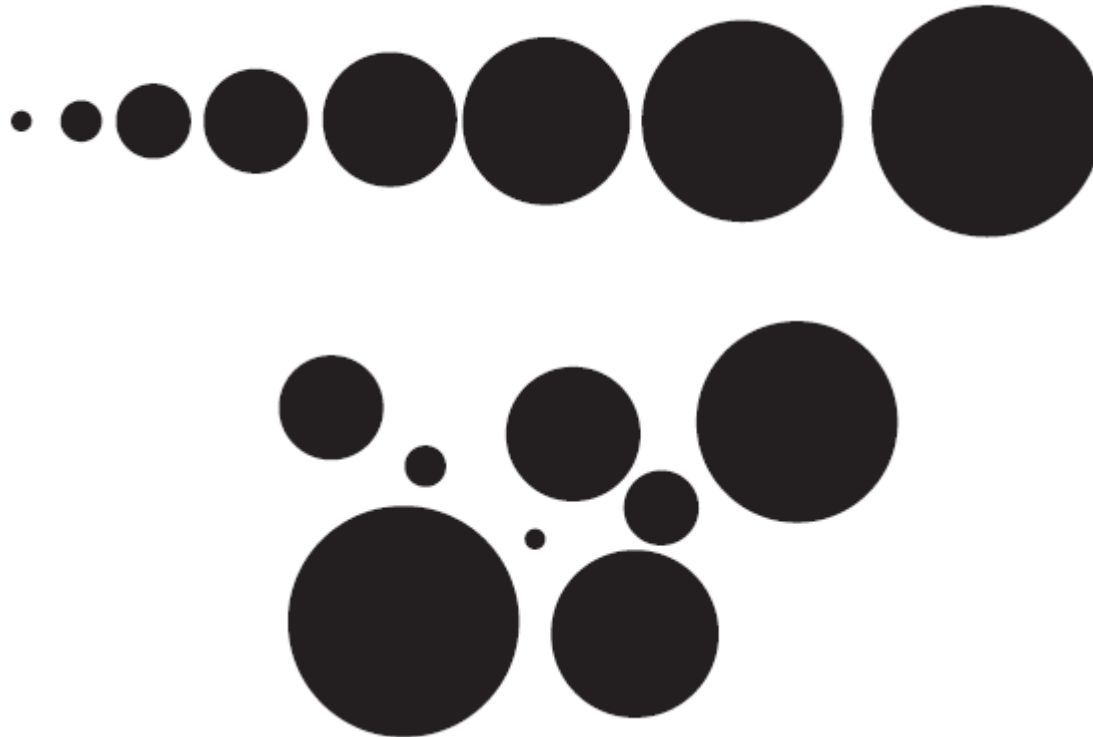- Spatial arrangement
- Cardinality

Channel discriminability

The effect of size

Channel discriminability

The effect of spatial arrangement

Channel discriminability

The effect of cardinality

Channel discriminability: Implications for design

Do not overestimate the number of values viewers can perceive/discriminate

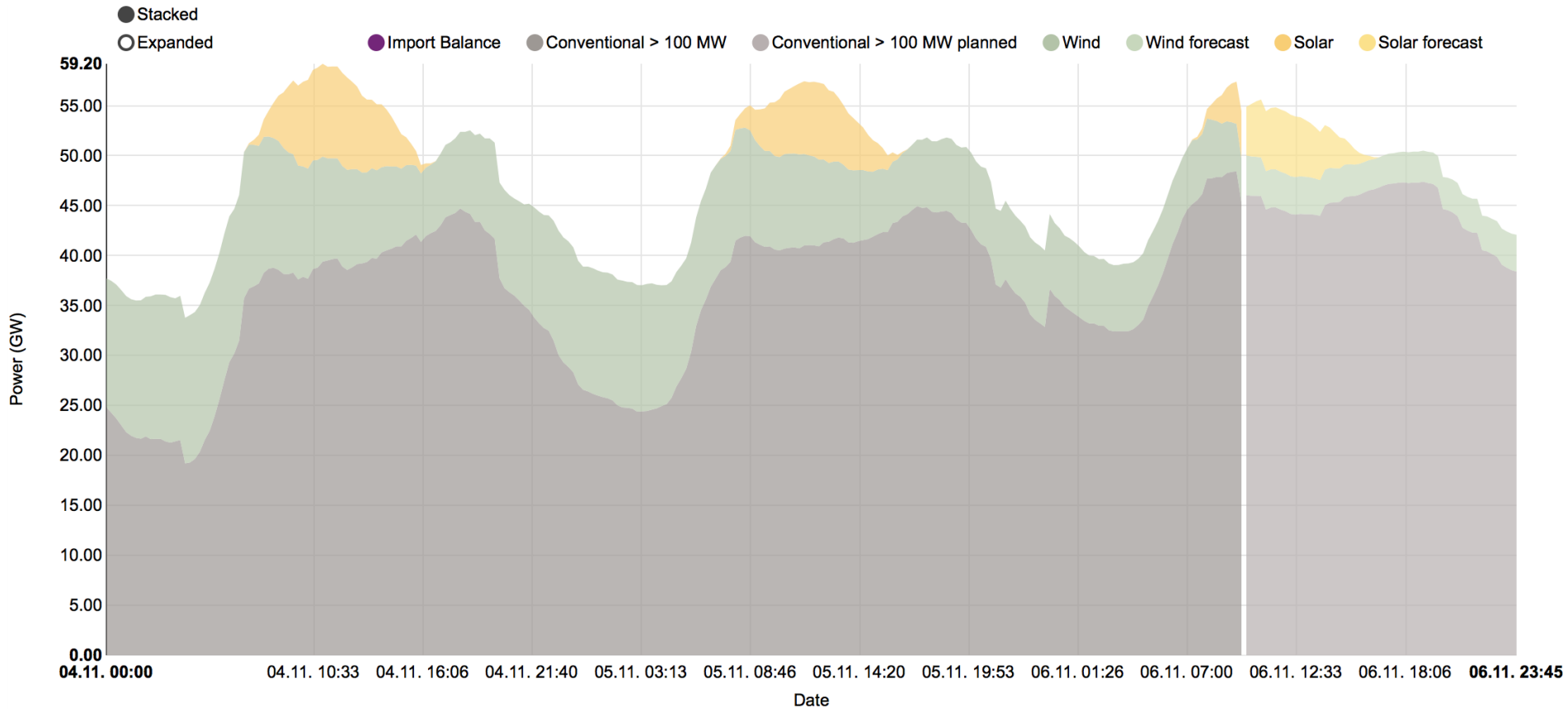Short term memory limitation: 7 ± 2 items (rather 5 than 9)

What to do in case of a large number of categories?
- Grouping (show groups of categories)
- Filtering (show only selected few)
- Faceting (use small multiples)

Electricity production in Germany – all sources
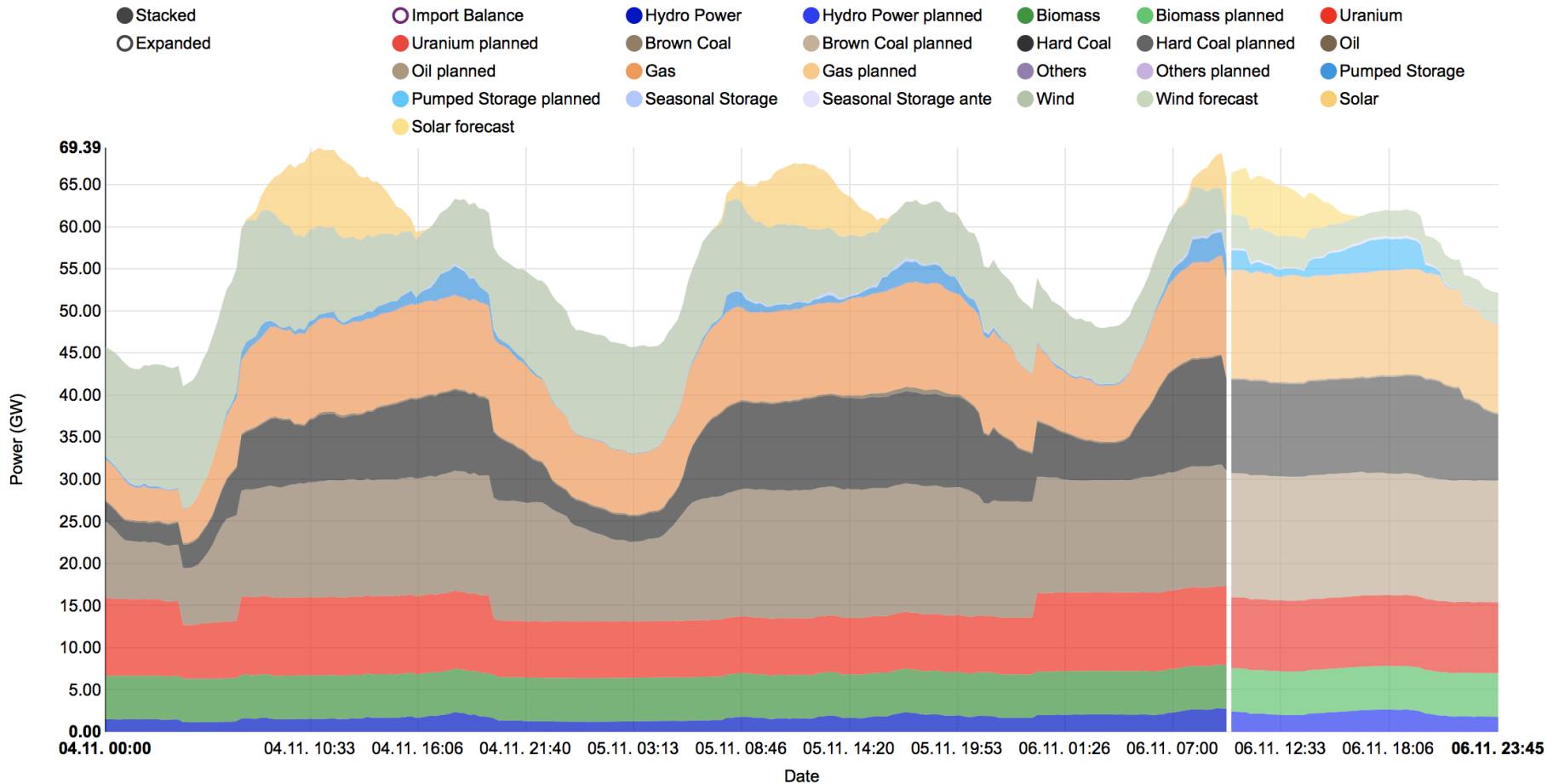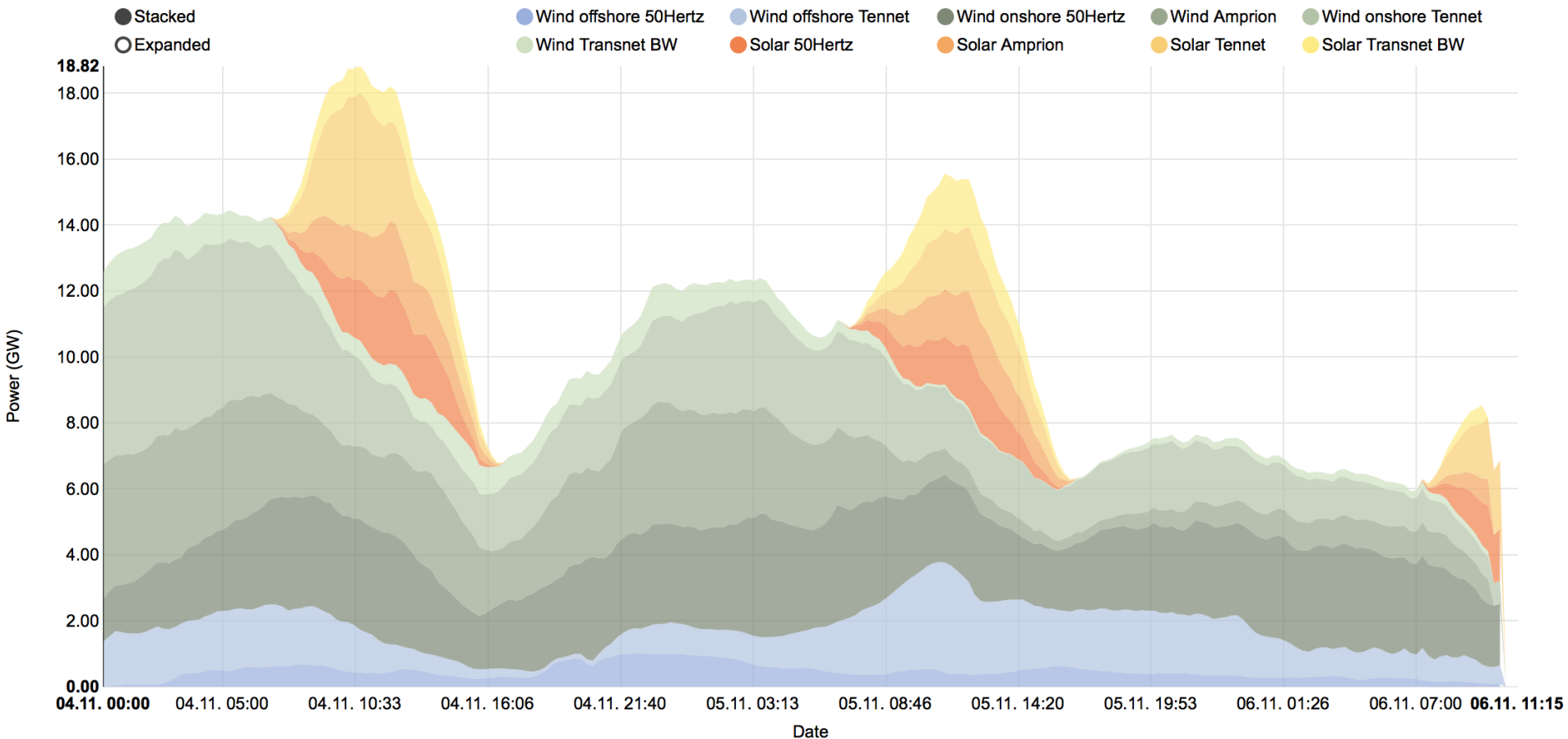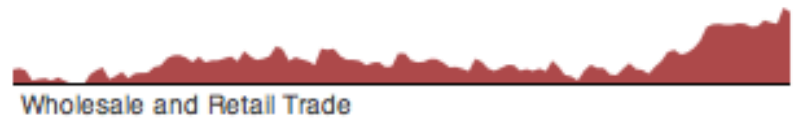
Electricity production in Germany – grouped sources

Electricity production in Germany – filtered sources

Small multiples

Agriculture

Business services

Construction

Education and Health

Finance

Government

Information

Leisure and hospitality

Manufacturing

Mining and Extraction

Other

Self-employed

Transportation and Utilities

Wholesale and Retail Trade
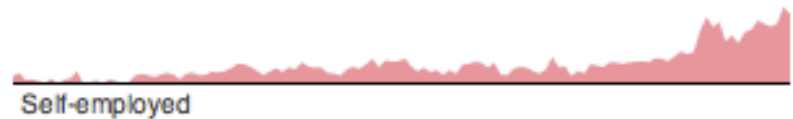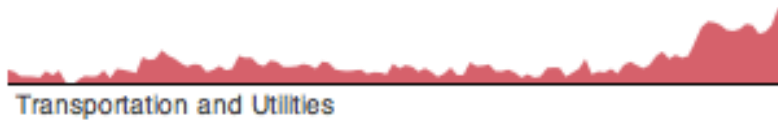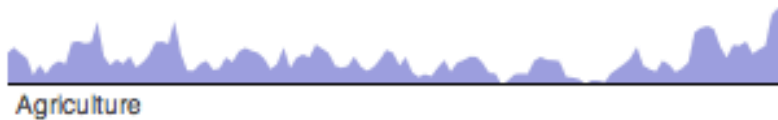
Channel salience

Channel salience (pop-out)

Ability to stand out in a scene

Highly related to preattentive processing

o Uses sensory memory

o Happens automatically

o Tasks performed in less than 250 ms (faster than eye movement initiation)

Neurons in the brain are tuned to specific properties, called preattentive attributes

An example

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

An example

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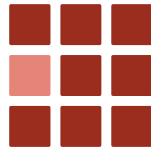7 2 4 6 1 7 **5** 9 7 3 2 4 9 1

Preattentive attributes

Color hue

Color intensity

Line length

Line width

Enclosure
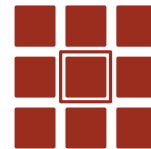
Shape

Size

Curvature

Orientation

Added marks

Blur

Position

Motion

Flicker

Preattentive attributes

Many attributes are asymmetric

Preattentive attributes

Some attributes are stronger than others
- In boundary detection, color hue is stronger than shape
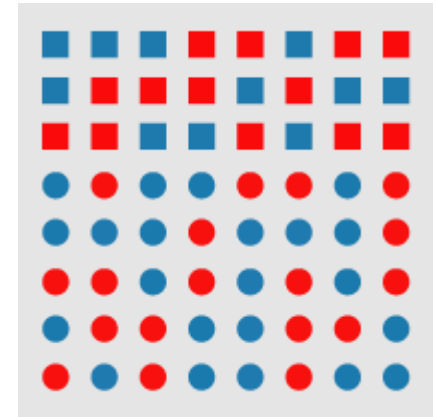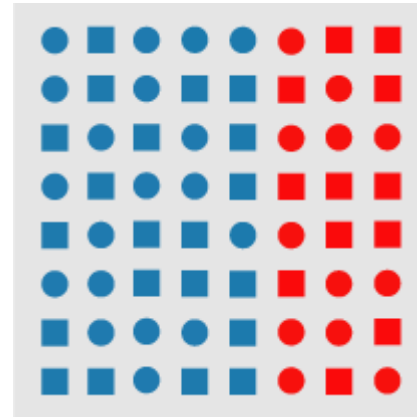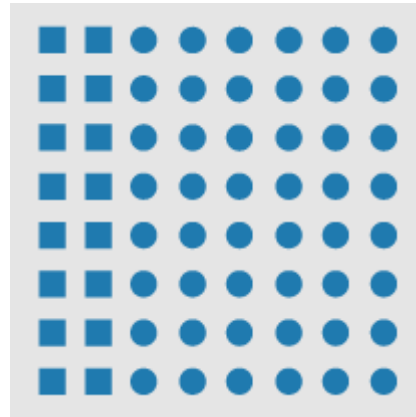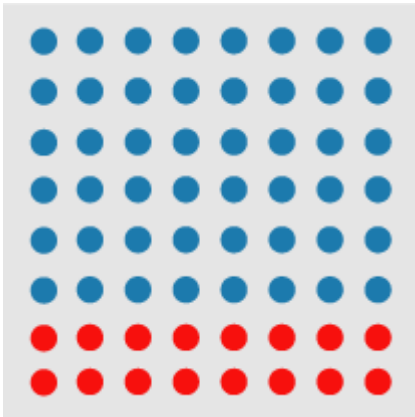
Preattentive attributes

Conjunctions of two attributes often not preattentive
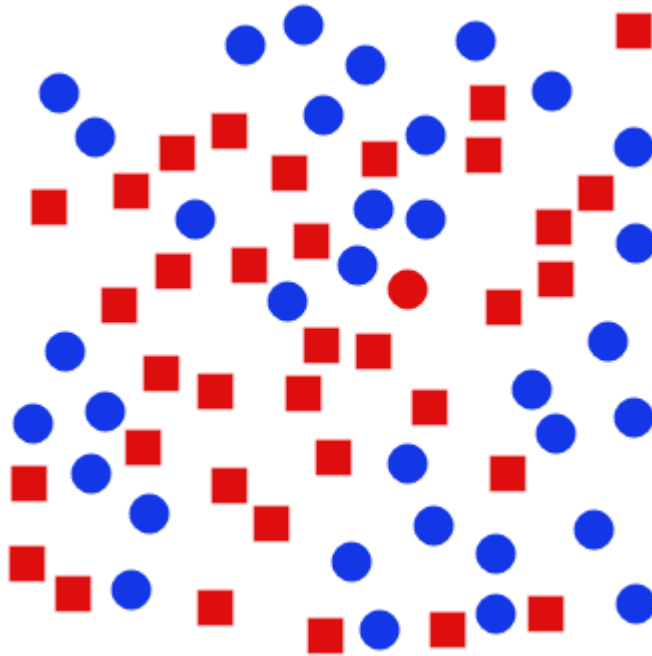
Preattentive conjunctions
◦ Space and color
◦ Motion and shape

Channel salience: Implications for design
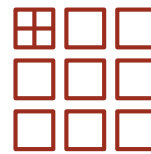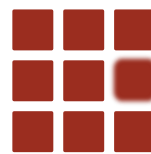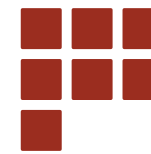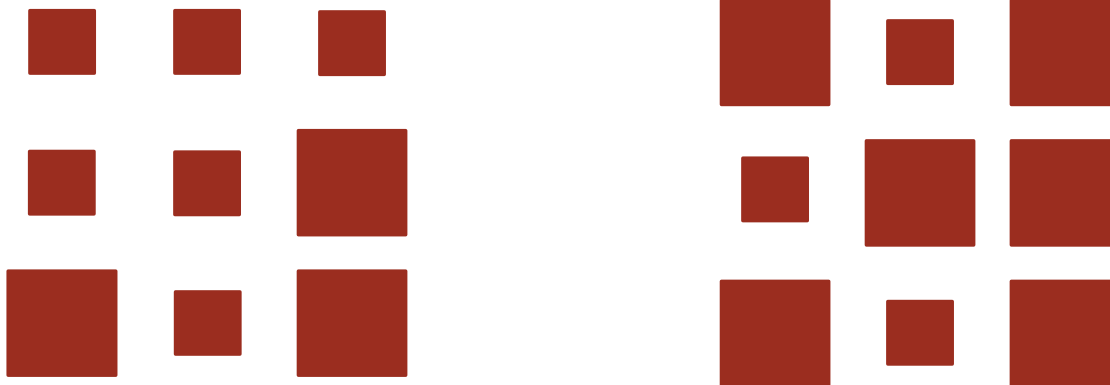
Preattentive attributes can be used to draw attention

7 of the top 10 complaints have 10 or more concerns per 1,000

concerns per 1,000

Engine power is less than expected	12.9
Tires make excessive noise while driving	12.3
Engine makes abnormal/excessive noise	11.6
Seat material concerns	11.6
Excessive wind noise	11.0
Hesitation or delay when shifting	10.3
Bluetooth system has poor sound quality	10.0
Steering system/wheel has too much play	8.8
Bluetooth system is difficult to use	8.6
Front seat audio/entertainment/navigation controls	8.2
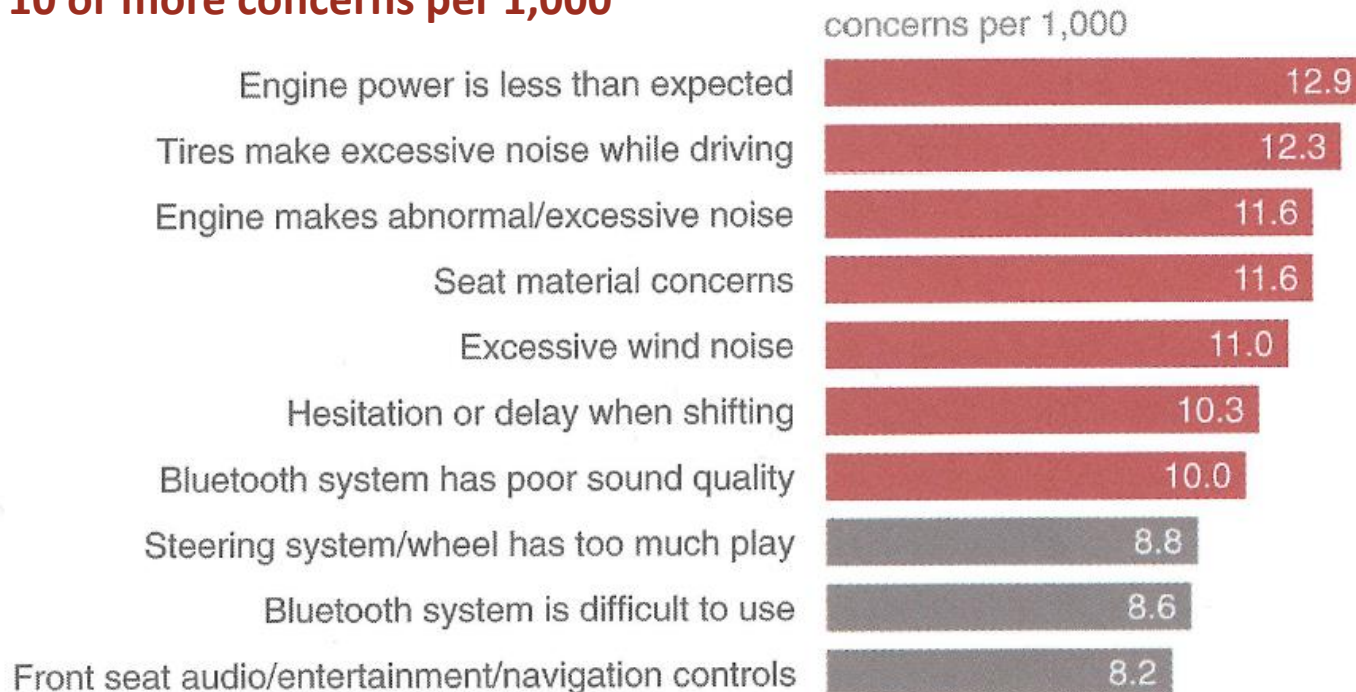
Channel salience: Implications for design

Preattentive attributes can be used create a visual hierarchy of information

3 of the top 10 complaints are noise related

concerns per 1,000

Engine power is less than expected	12.9
Tires make excessive noise while driving	12.3
Engine makes abnormal/excessive noise	11.6
Seat material concerns	11.6
Excessive wind noise	11.0
Hesitation or delay when shifting	10.3
Bluetooth system has poor sound quality	10.0
Steering system/wheel has too much play	8.8
Bluetooth system is difficult to use	8.6
Front seat audio/entertainment/navigation controls	8.2

Comments indicate that **noisy tire issues** are most apparent **in the rain**.

Complaints about **engine noise** commonly cited **after the car had not been driven for a while**.

Excessive **wind noise** is noted primarily in **freeway driving at high speeds**.

Channel salience: Implications for design

Use color sparingly

Use gray for elements that are not that important

When you highlight one point, you make the other points harder to see

Do not use preattentive attributes in exploratory data analysis
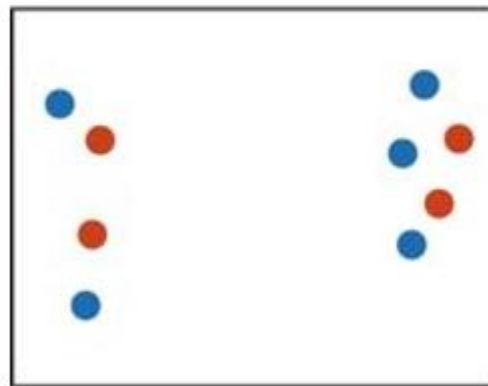
Channel separability

Channel separability

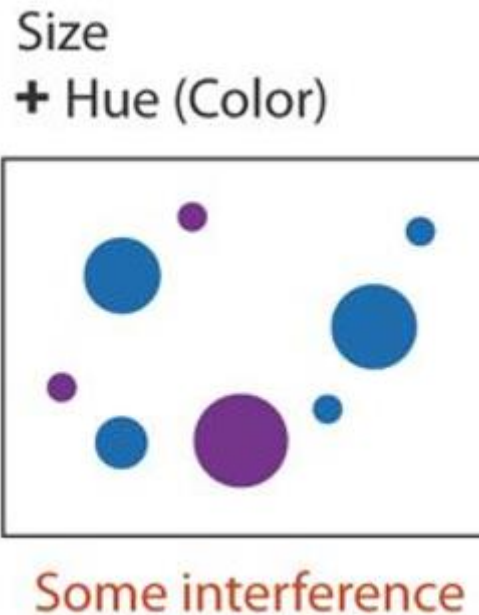Amount of interference between channels

Position
+ Hue (Color)

An example of
separable channels

Fully separable

Channel separability

Amount of interference between channels

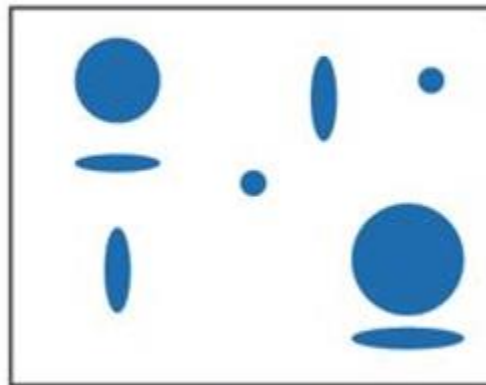
Size
+ Hue (Color)

Some interference

Channel separability

Amount of interference between channels

Width
+ Height

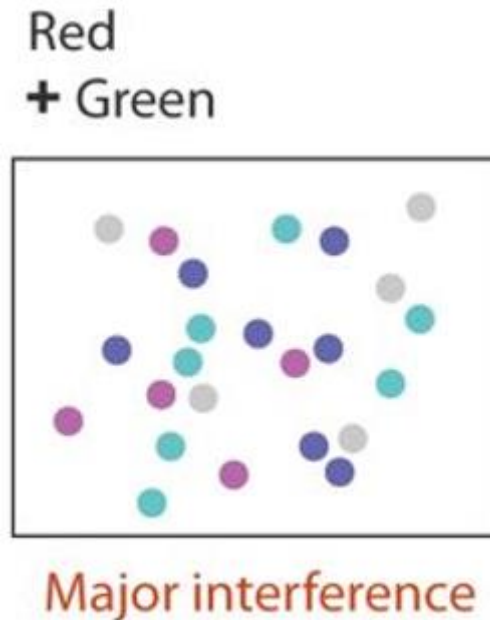Some/significant
interference

An example of
integral channels

Channel separability

Amount of interference between channels

Red
+ Green

Major interference

An example of
integral channels

Are the richest Americans also the best educated?

Channel separability: Implications for design

Use separable channels when the audience should perceive one variable at a time

Use integral channels when you want a holistic effect

Grouping

GESTALT LAWS

Gestalt laws

Gestalt (German) = shape, form

Gestalt psychology aims to understand how individual visual objects are grouped to form a pattern

The whole is other than the sum of its parts

Kurt Koffka, Gestalt psychologist

Gestalt laws

Proximity

Similarity

Connection

Enclosure

Closure

Figure/Ground

Gestalt law of Proximity

We perceive objects close to each other as belonging to a group
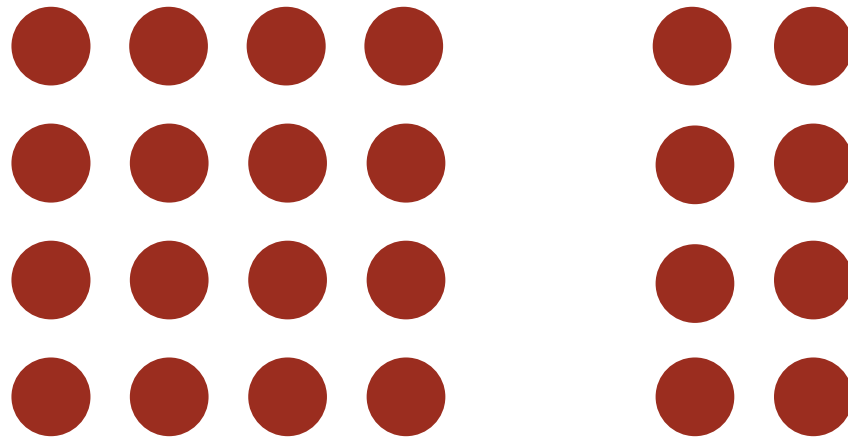
Gestalt law of Proximity: Implications for design

Place annotations close to the data
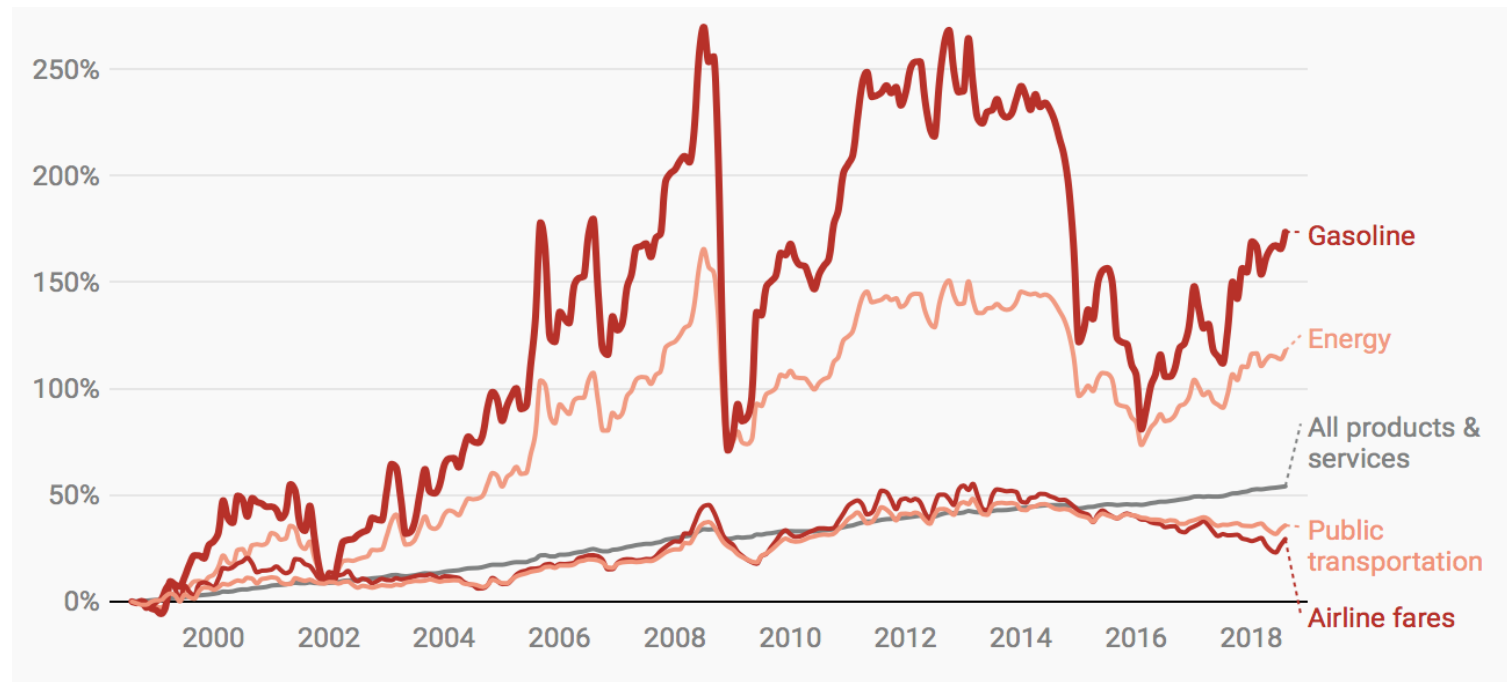
Gestalt law of Similarity

We perceive similar objects as belonging to a group
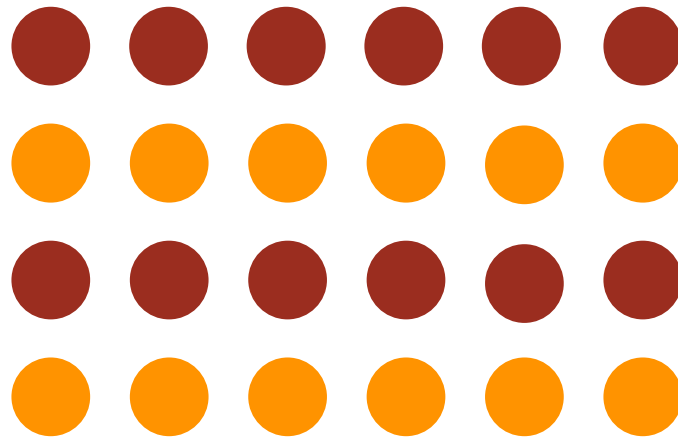
Gestalt law of Similarity: Implications for design

Russians Are Turning Their Backs On Vodka
Liters of pure alcohol consumed per capita in Russia by beverage

● Beer   ● Wine   ● Spirits

4.68

1.47
0.86

3.29
3.25

1.08

'90 '91 '92 '93 '94 '95 '96 '97 '98 '99 '00 '01 '02 '03 '04 '05 '06 '07 '08 '09 '10 '11 '12 '13 '14 '15 '16

@StatistaCharts   Source: World Health Organization

statista
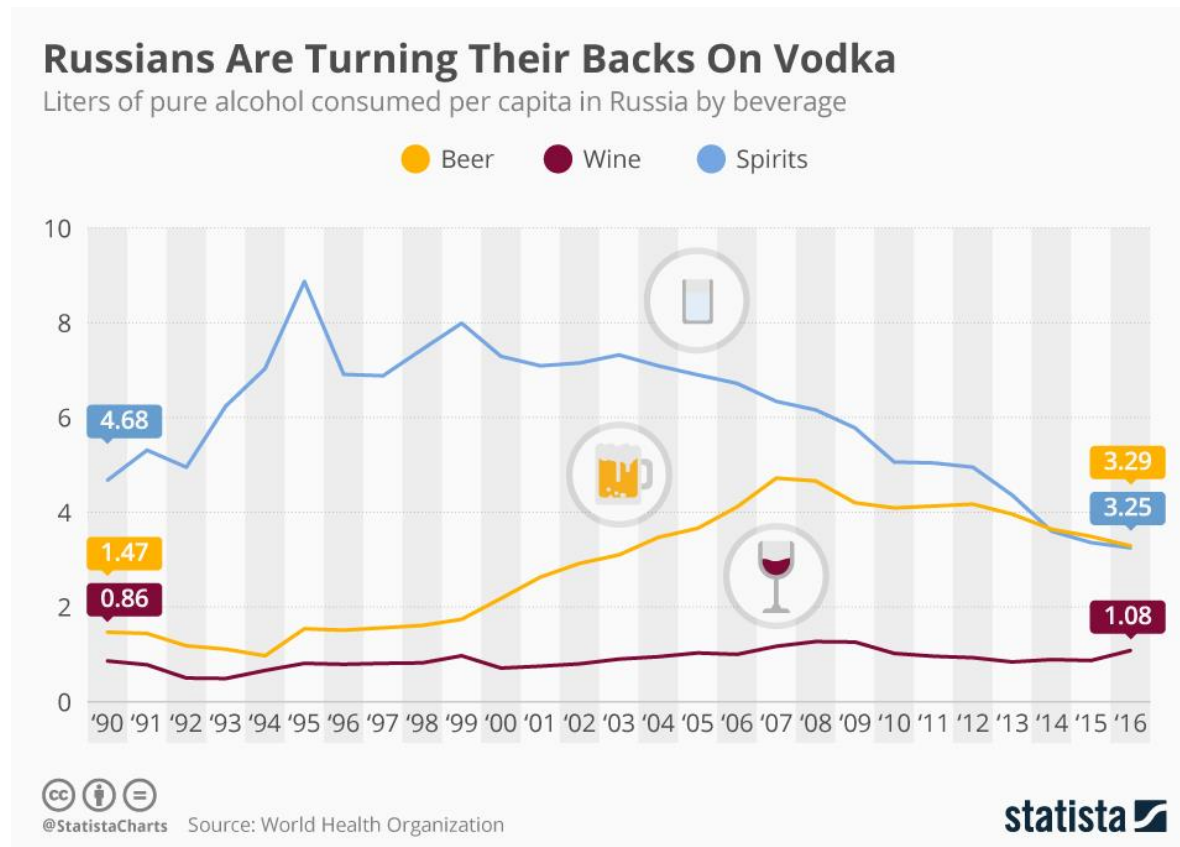
Gestalt law of Connection

We perceive objects connected to each other as a single group
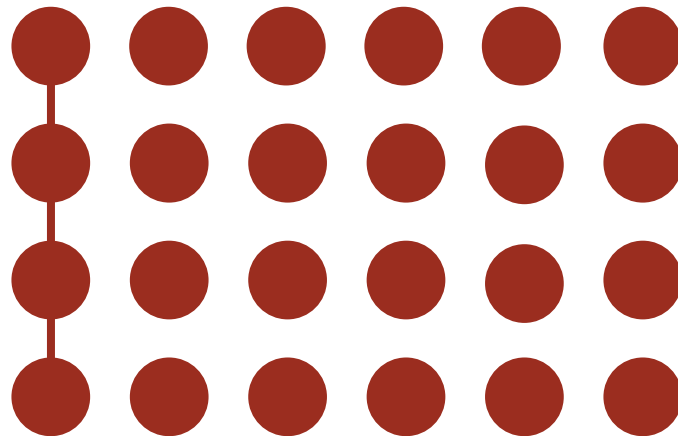
Gestalt law of Connection: Implications for design

Use lines to show the data is in the same group
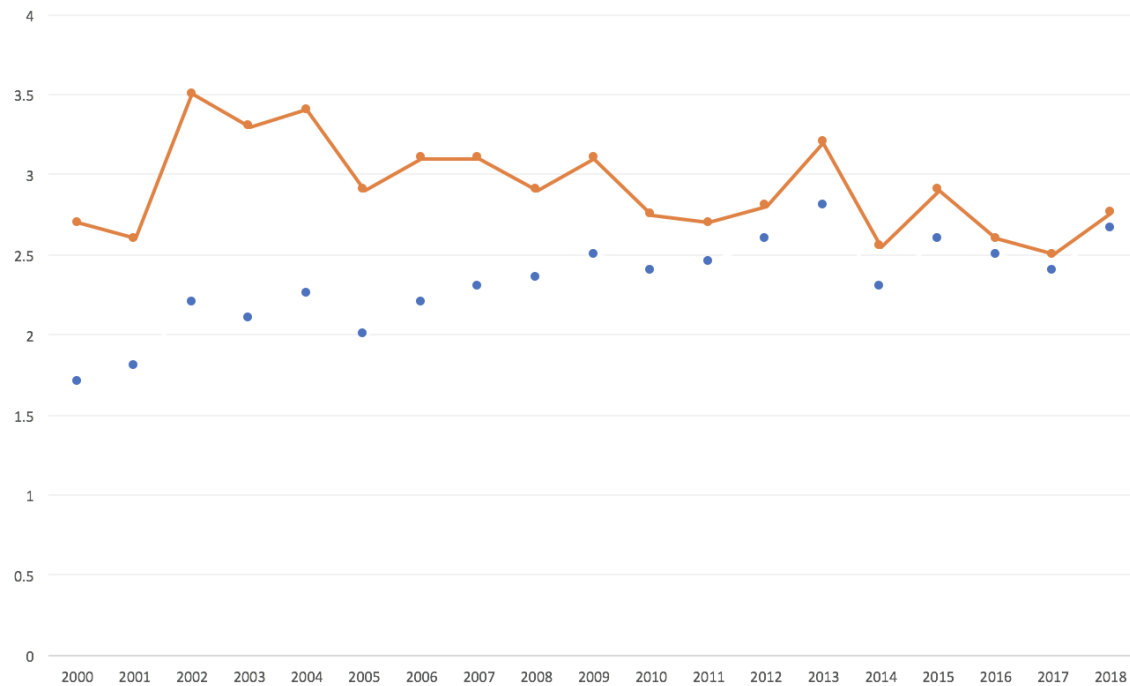
Gestalt law of Enclosure

We perceive physically enclosed objects as part of a group
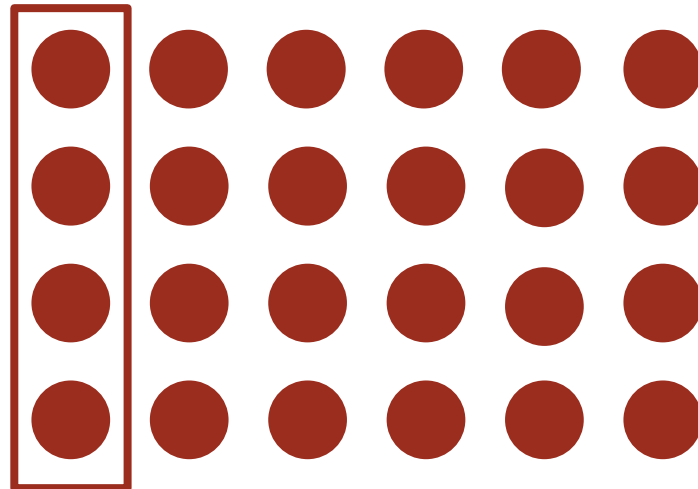
Gestalt law of Enclosure: Implications for design

Use enclosures to show groups

Gestalt law of Enclosure: Implications for design

Bubble sets visualization

Gestalt law of Enclosure: Implications for design

Winners and Losers: Job Gains and Losses Jump to National Unemployment

Track the number of sectors gaining or losing jobs each month. Boxes are shaded based on percentage change from the previous month in each sector's payrolls.
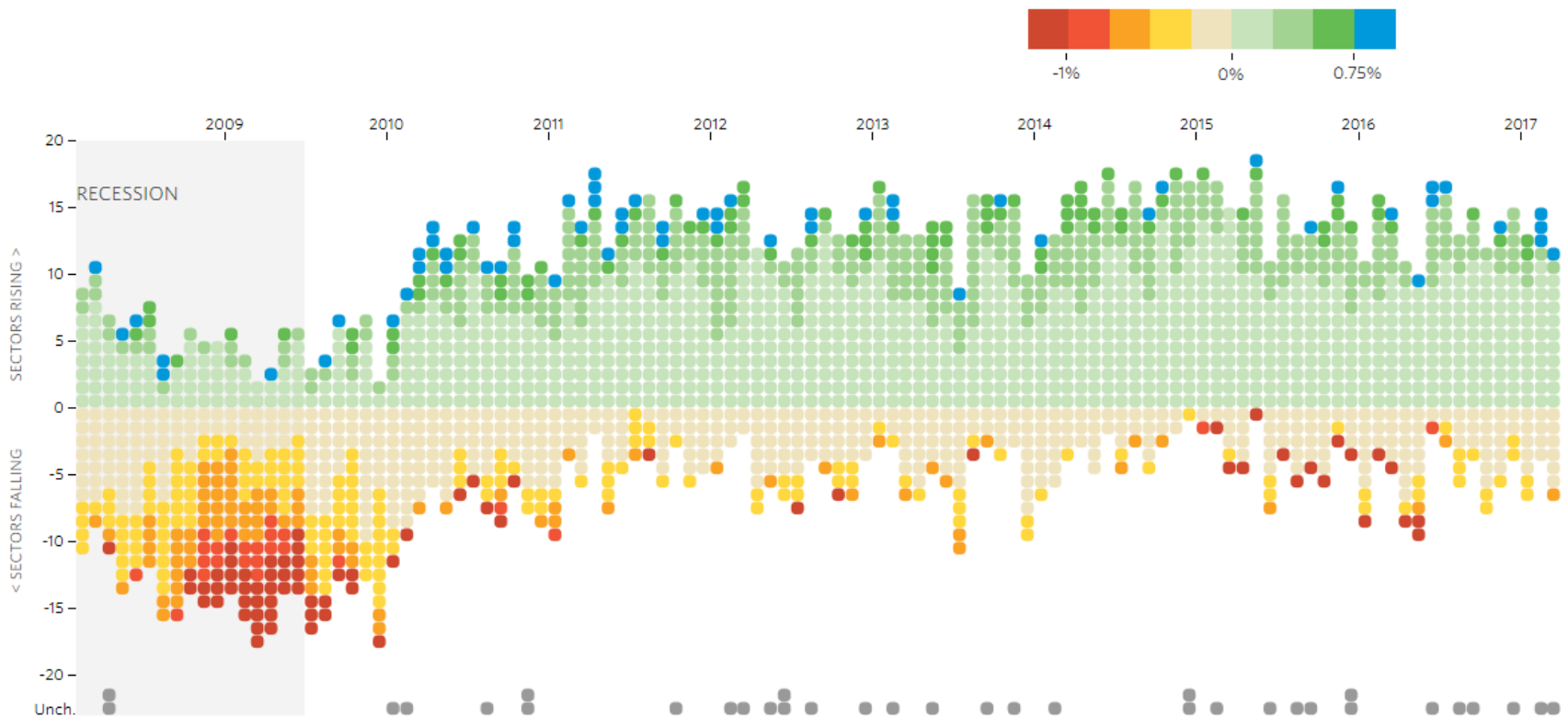
Hierarchy within grouping

Similarity < Proximity < Connection & Enclosure
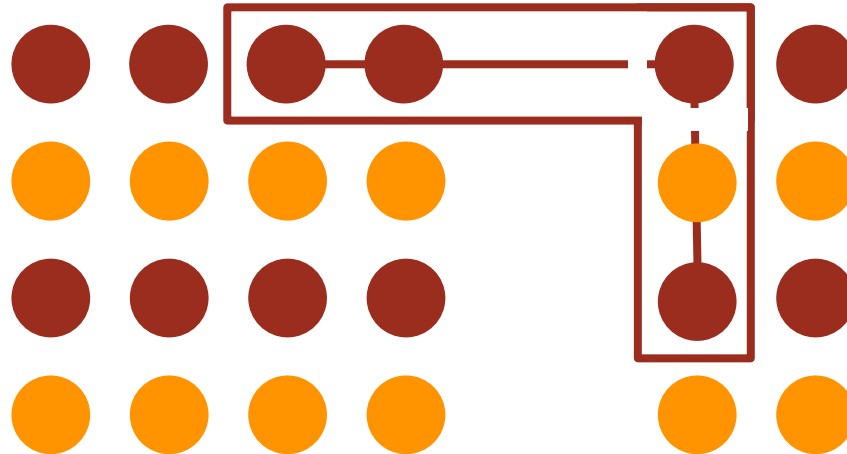
Gestalt law of Closure

We perceive objects as being whole even when they are not complete
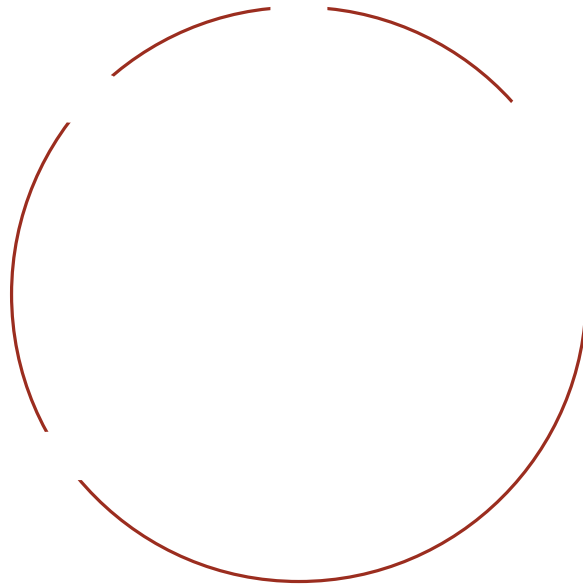
Gestalt law of Closure: Implications for design

No need to draw chart borders

Calories per 100g
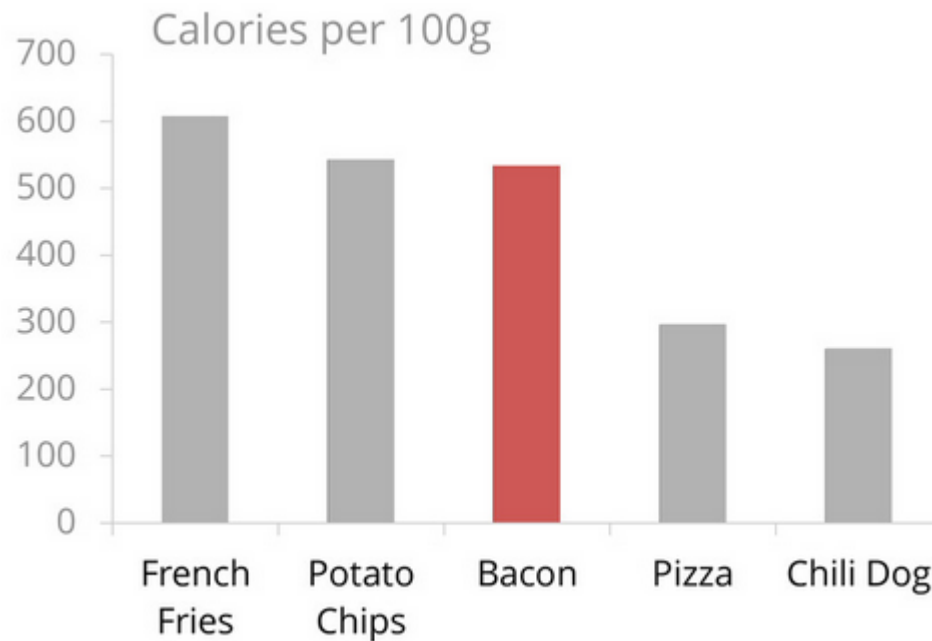
Gestalt law of Closure: Implications for design

Be careful in case of missing values
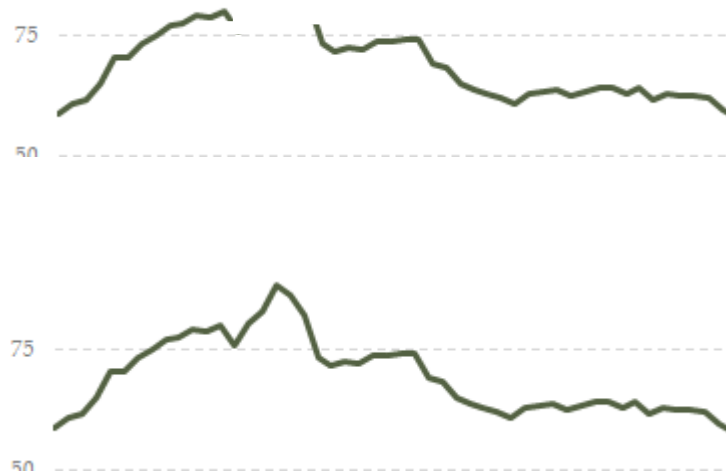
Gestalt law of Figure/Ground

We perceive elements as either figure (element of focus) or ground (background)
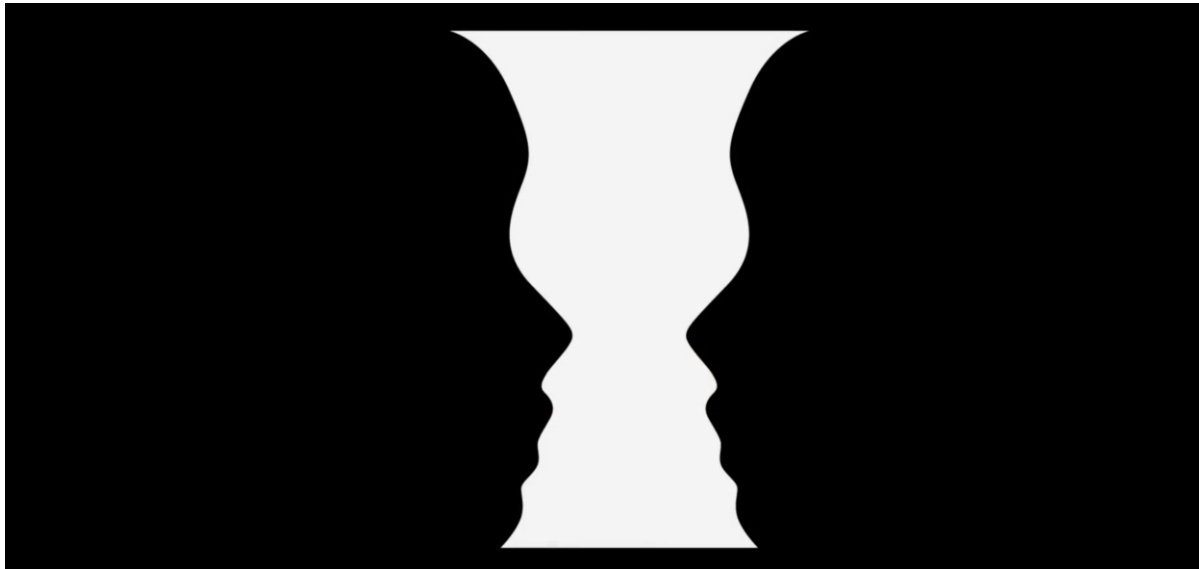
Gestalt law of Figure/Ground: Implications for design

Color contrast and overlays can be used to discern the figure from the background

International Migrants Day: Mediterranean Death Trap
Migrant deaths worldwide by region in 2016

- 53 Europe
- · 3 East Asia
- 4,812 Mediterranean
- 409 U.S./Mexico Border
- 96 Middle East
- 52 Caribbean
- 1,089 Sahara & North Africa
- 208 South-East Asia
- 159 Central America & Mexico
- 195 Horn of Africa
- 27 South America
- 77 Sub-Saharan Africa

@StatistaCharts International Migrants Day on Dec 18, 2016
Source: IOM

statista

Channel efficiency summary

Accuracy
o Prioritize high ranking channels

Discriminability
o Do not use more than 5-7 colors

Salience (pop-out)
o Be mindful with how you direct attention

Separability
o Use separable channels to perceive one variable at a time
o Use integral channels to obtain a holistic effect

Grouping
o Be mindful of how visual elements form groups

Visual order

Visual order

The attention of people follows the Z shape
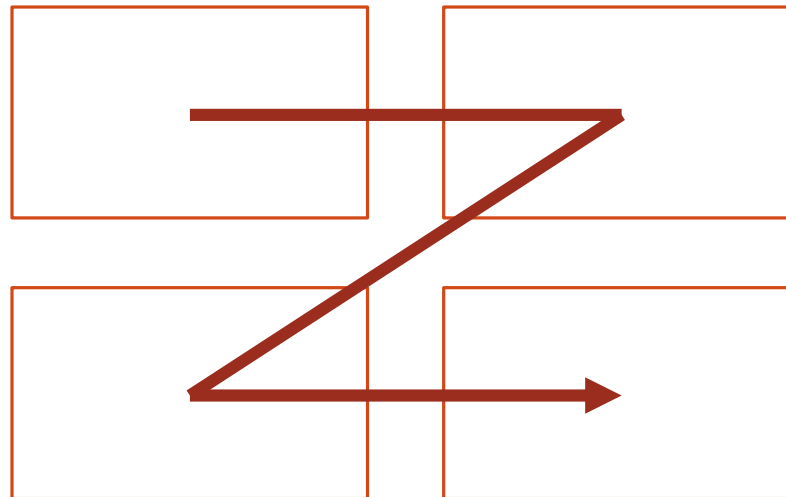
You should place the important things on the top (left) of the display

Visual order

All elements should be aligned – create clean vertical and horizontal 'lines' to establish a sense of unity and cohesion

Do not be afraid of white (empty) space – do not add more data (or stretch the graphics) to get rid of it

Stay away of diagonal components (especially text)

Visual order

Stay away of diagonal components (especially text)

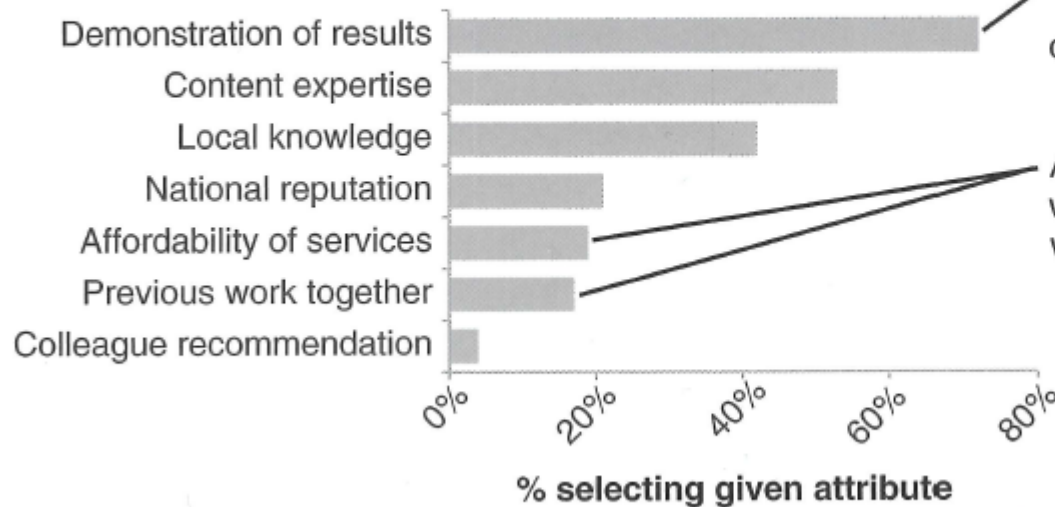Text rotated by 45 degrees (in either direction) is 52% slower to read than normally oriented text

Text rotated by 90 degrees (in either direction) is 205% slower to read than normally oriented text

Visual order – an example

Demonstrating effectiveness is most important consideration when selecting a provider

In general, what attributes are the most important to you in selecting a service provider?
(Choose up to 3)

- Demonstration of results
- Content expertise
- Local knowledge
- National reputation
- Affordability of services
- Previous work together
- Colleague recommendation

0%   20%   40%   60%   80%

% selecting given attribute

Survey shows that demonstration of results is the single most important dimension when choosing a service provider.

Affordability and experience working together previously, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.
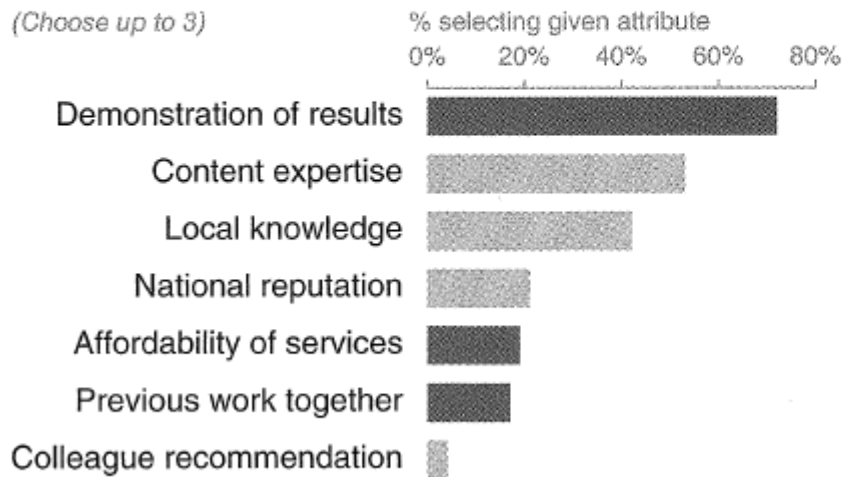
Data source: xyz; includes N number of survey respondents. Note that respondents were able to choose up to 3 options.

Visual order – an example

Demonstrating effectiveness is most important consideration when selecting a provider

In general, **what attributes are the most important** to you in selecting a service provider?

(Choose up to 3)

% selecting given attribute

Survey shows that **demonstration of results** is the single most important dimension when choosing a service provider.

Affordability and **experience working together previously**, which were hypothesized to be very important in the decision making process, were both cited less frequently as important attributes.

Data source: xyz; includes N number of survey respondents.
Note that respondents were able to choose up to 3 options.
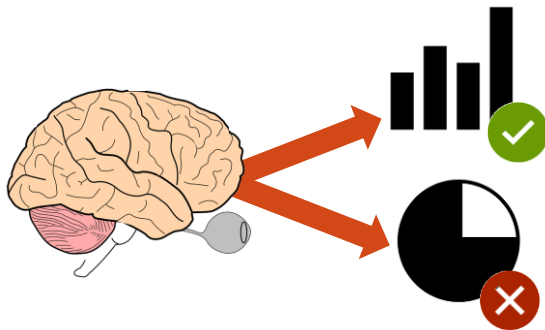
Visual order

Pay attention to details

Avoid
- Too much centered text
- Diagonal components, especially text
- Too many things on a single display

Summary

How data visualization works

The quality of visualizations
is mostly not subjective

How data visualization works

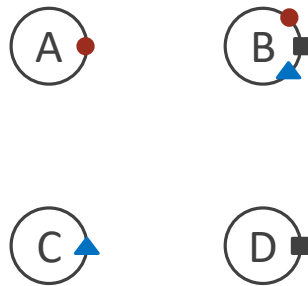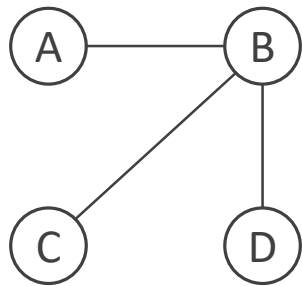And you will read this last

You will read
this first

And then you will read this

Then this one

Pop-out and

visual order

How data visualization works

A —— B

A (•) B (•■▲)

C —— D   C (▲) D (■)

Connection more powerful than similarity