

Information Visualization

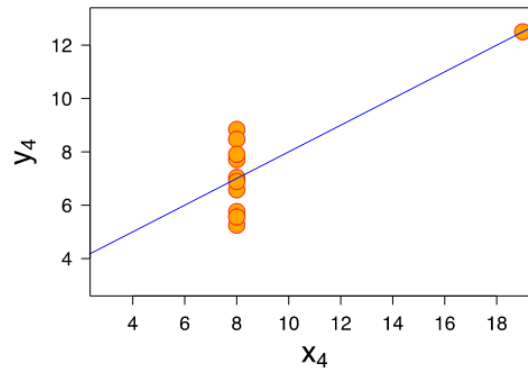
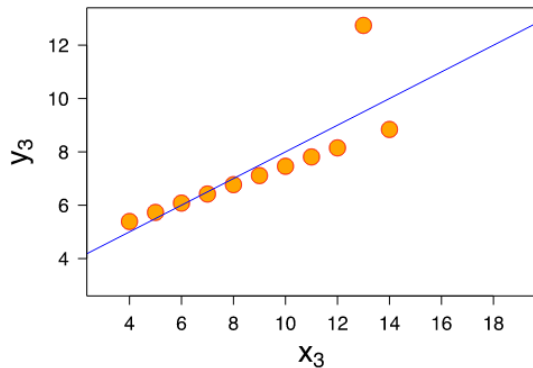
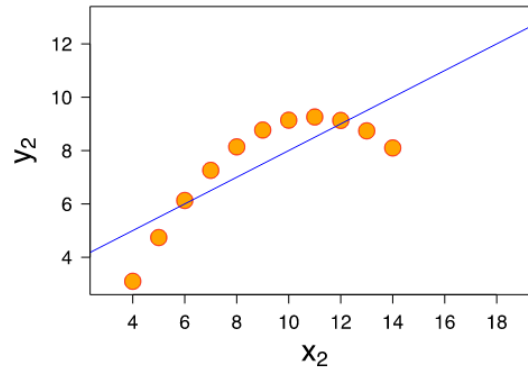
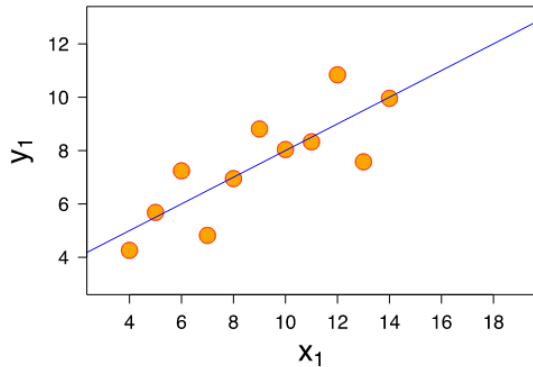
EECS6414 – Data Analytics and Visualization

Agenda*

- What is data visualization?
- Jacques Bertin's visual variables (semiotics)
- Cognition and perception
- Gestalt principles
- Principles of graphical excellence and best practices
- Beyond digital visualization: multimodal data expression

what is data visualization?

Why visualize data? Anscombe's Quartet



Summary statistics for all four datasets

- $\text{avg}(x) = 9$
- $\text{avg}(y) = 7.50$
- $\text{Var}(x) = 11$
- $\text{Var}(y) = 4.12$
- $\text{Correlation}(x,y) = 0.816$
- A linear regression line:
 $y = 0.5x + 3$

Always plot your data!

Anscombe's Quartet

Anscombe, F. (1973). Graphs in statistical analysis. American Statistician, 27:17--21.

What is data visualization?



Use of **visual elements** like **charts**, **graphs**, and **maps** to see and understand **trends**, **outliers**, and **patterns** in data

Some examples

The beauty of visualization

<https://www.youtube.com/watch?v=5Zg-C8AAlGg&feature=youtu.be&t=621>

Information is beautiful

<https://informationisbeautiful.net/>

On Broadway

<http://www.on-broadway.nyc/>

Network visualization

<https://www.youtube.com/watch?v=GW4s58u8PZo>

Gapminder

<https://www.gapminder.org/>

London whereabouts

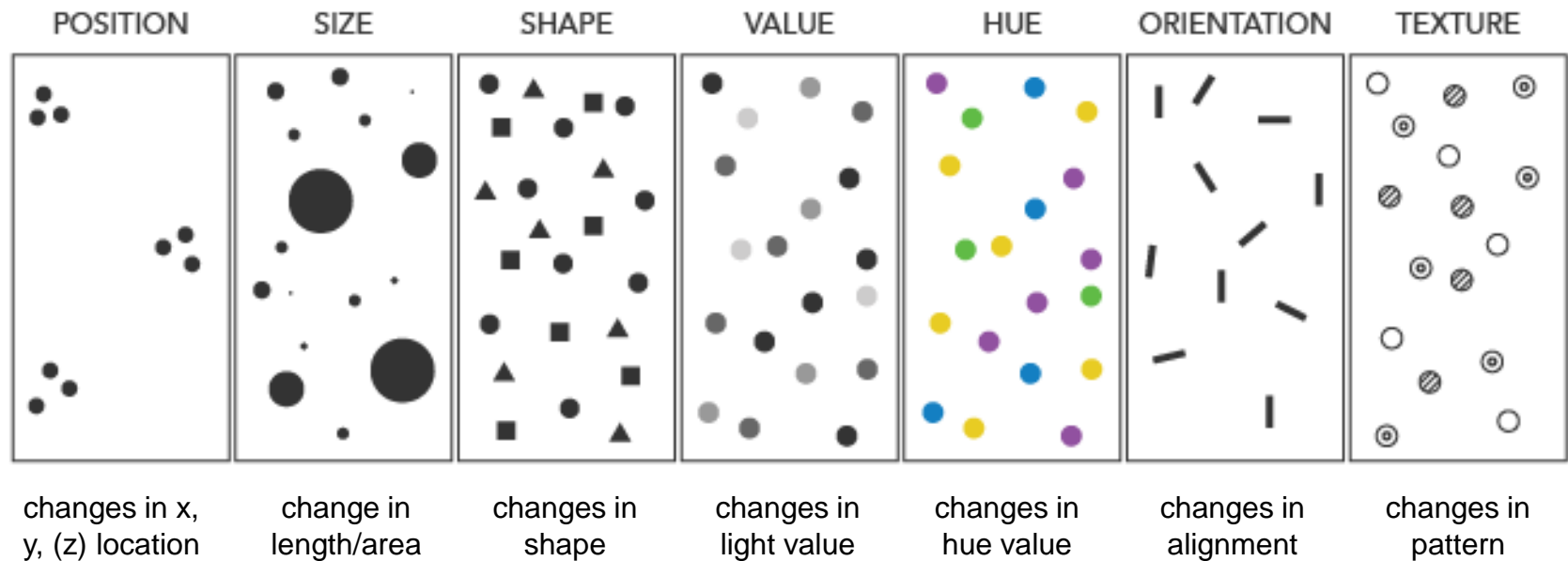
<http://whereaboutslondon.org>

D3 library

<https://observablehq.com/@d3/gallery>

Jacques Bertin's visual variables

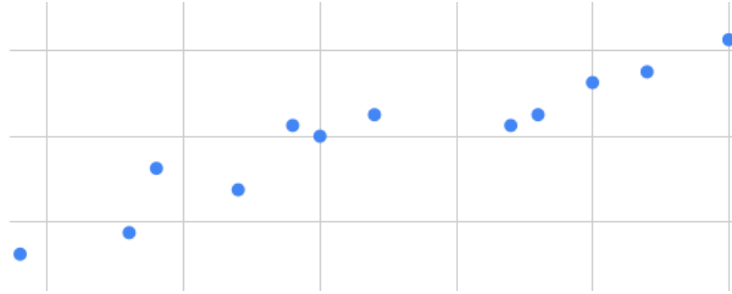
Jacques Bertin's visual variables (vv)



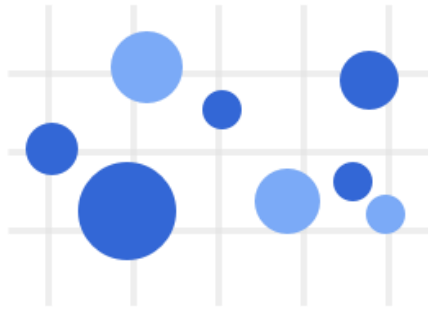
Jacques Bertin proposed an original set of “retinal variables” in Semiology of Graphics (1967)

Visual variables for quantitative data

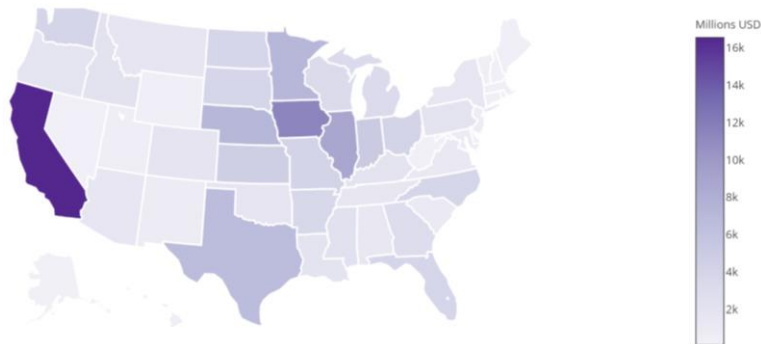
position



size



value

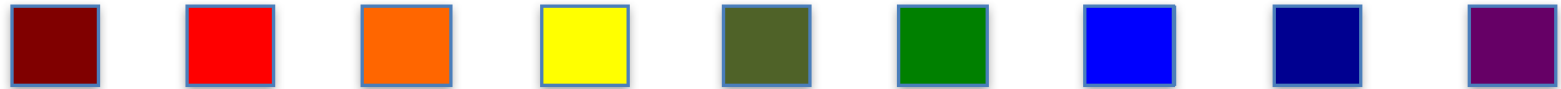


Visual variables for qualitative data

texture

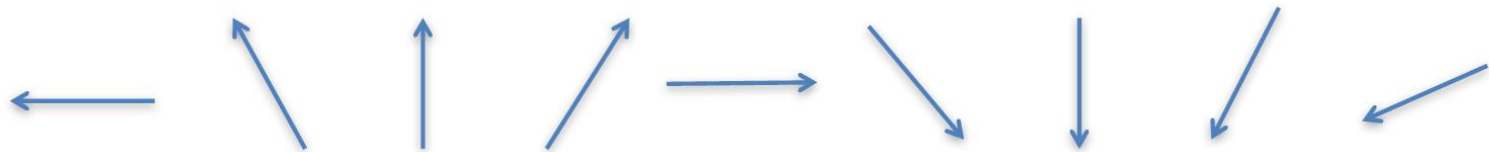


colour



Permanently delete sensitive text

orientation



shape



Motion: new visual variable

Motion as a new visual variable

changes in motion can include direction, speed, frequency, rhythm, flicker, trails, and style

example

100 years of rock music (<http://www.concerthotels.com/100-years-of-rock>)

Cognition and perception

Perception & cognition

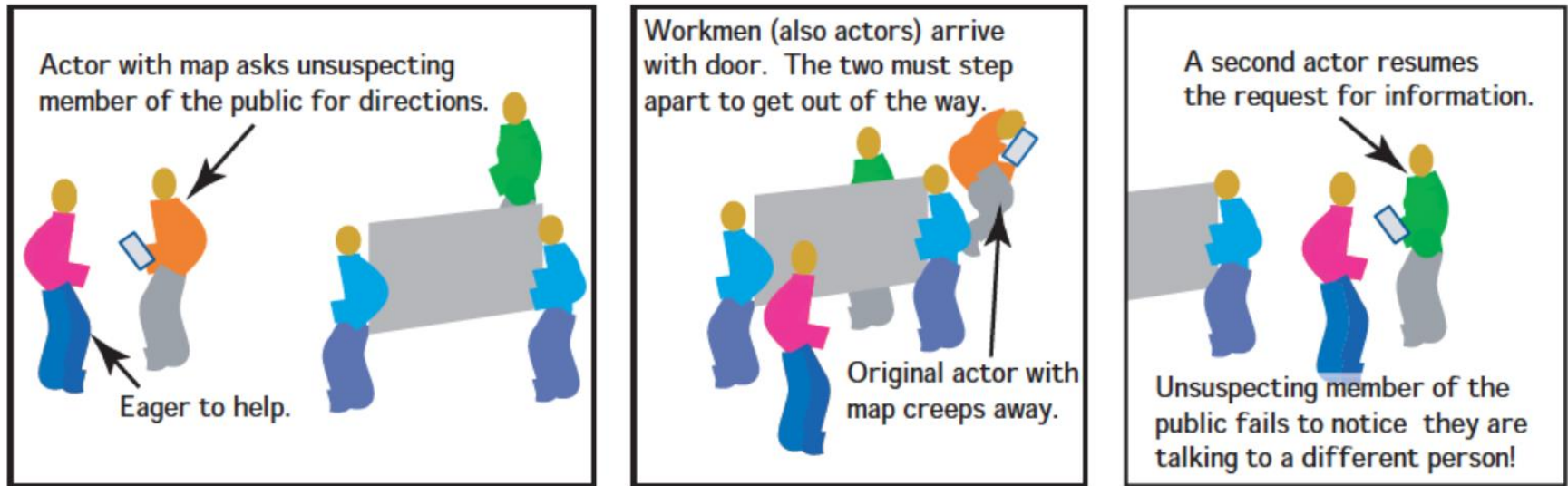


Image: Ware, Colin. Visual thinking: For design. Morgan Kaufmann, 2010

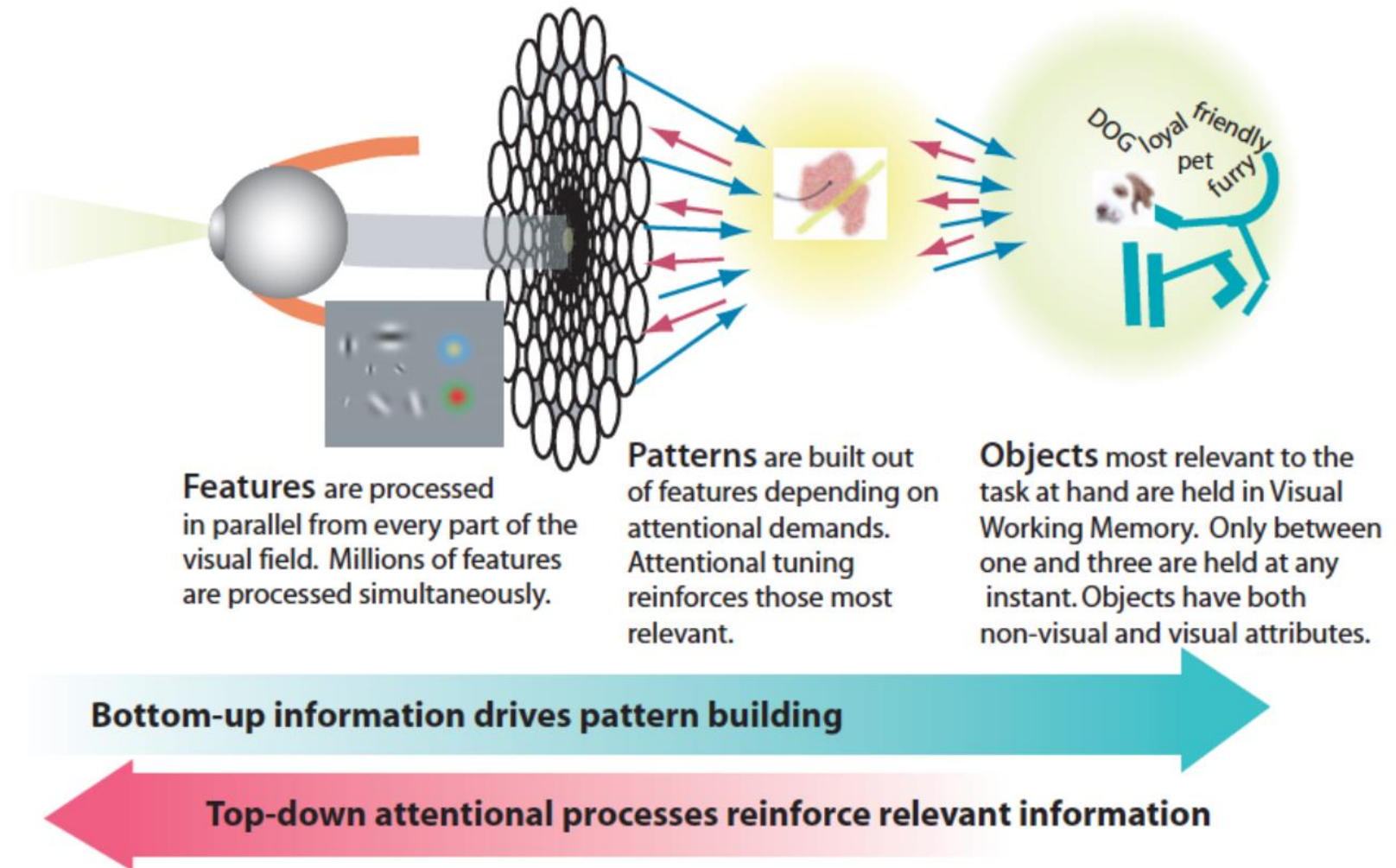
- perception is fragmented
- eyes are constantly scanning and constructing reality

The “Door Study”*

<https://www.youtube.com/embed/FWSxSQsspiQ>

* Daniel J. Simons and Daniel T. Levin. 1998. “Failure to detect changes to people during a real world interaction.” Psychonomic Bulletin and Review. 5: 644–669.

Visual thinking is about finding patterns



Pre-attentive vs attentive processing

Pre-attentive Processing

- bottom-up
- fast, automatic
- instinctive
- efficient
- multitasks

Attentive Processing

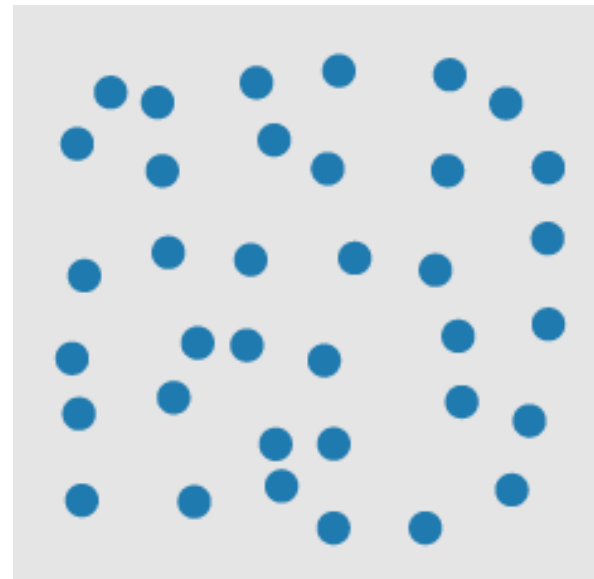
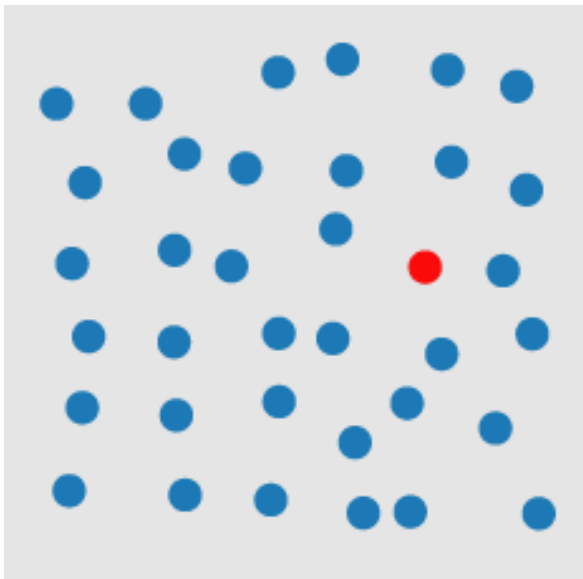
- top-down
- slow, deliberate
- focused
- single-task

goal of information design

- help humans process information as efficiently as possible
- make as much use of pre-attentive processing as possible

Pre-attentive proc. example (color selection)

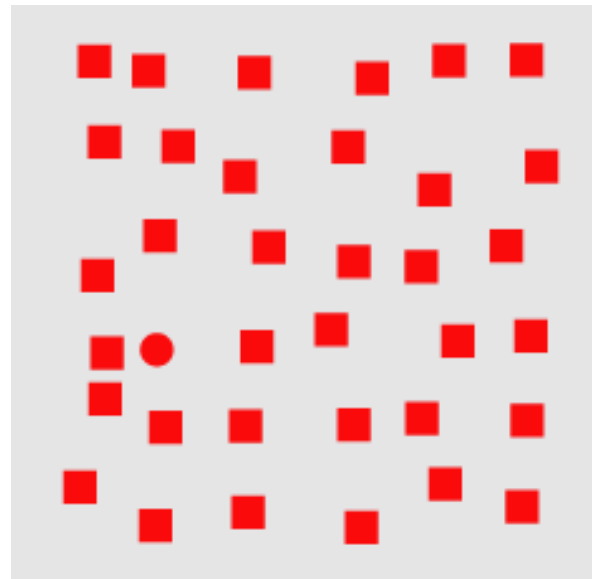
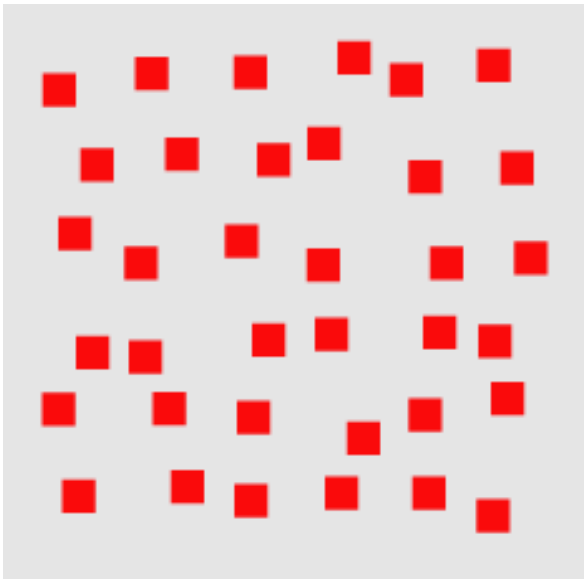
Is there a **red circle**?



- < 200-250ms qualifies as pre-attentive
 - eye movements take at least 200ms
- if a decision takes a fixed amount of time regardless of # of distractors (blue circles), it is considered to be pre-attentive

Pre-attentive proc. example (shape selection)

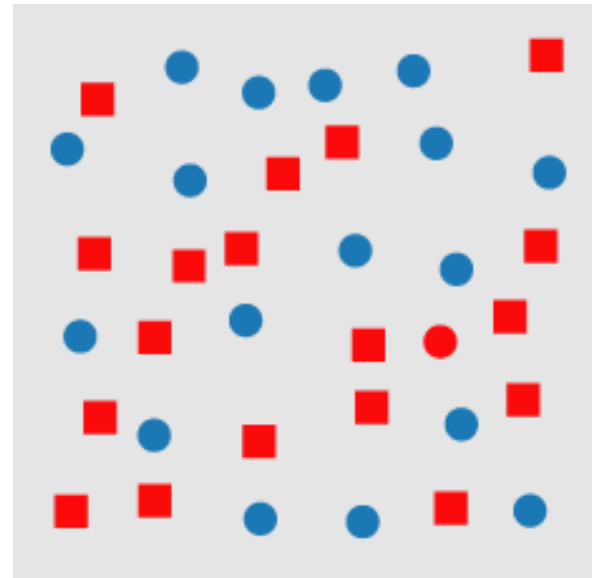
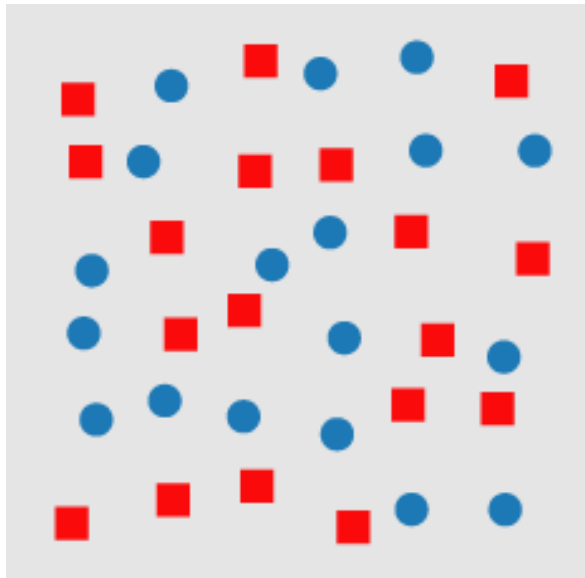
Is there a **red circle**?



As before, a viewer can rapidly and accurately determine whether the target is present or absent. Here, the visual system identifies the target through a difference in curvature (or form)

Attentive proc. example (conjunction search)

Is there a **red circle**?



A target made up of a combination of non-unique features (a conjunction target) normally cannot be detected pre-attentively. Figures shows an example of **conjunction search**

Demo: target detection search (Healey & Enns'11)

https://www.csc2.ncsu.edu/faculty/healey/PP/index.html#jscript_search

(three different target detection searches: colour, shape, and conjunction)

Demo: count the 7s

13579345978274055
24937916478254137
23876597277103866
19874367259047362
95637283649105676
32543787954836754
56840378465785690

time proportional to
the number of digits

13579345978274055
24937916478254137
23876597277103866
19874367259047362
95637283649105676
32543787954836754
56840378465785690

time proportional to
the number of 7's

13579345978274055
24937916478254137
23876597277103866
19874367259047362
95637283649105676
32543787954836754
56840378465785690

both 3's and 7's
seen pre-attentively

Demo: selective attention test (Simons & Chabris 1999)

<https://www.youtube.com/watch?v=vJG698U2Mvo>

(count how many times the players wearing white pass the basketball)

Goal of information design

- A limited set of visual properties are processed pre-attentively (without need for focusing attention)
- This is important for design of visualizations
 - what can be perceived immediately?
 - which properties are good discriminators?
 - what can mislead viewers?

goal of information design

- help humans process information as efficiently as possible
- make as much use of pre-attentive processing as possible

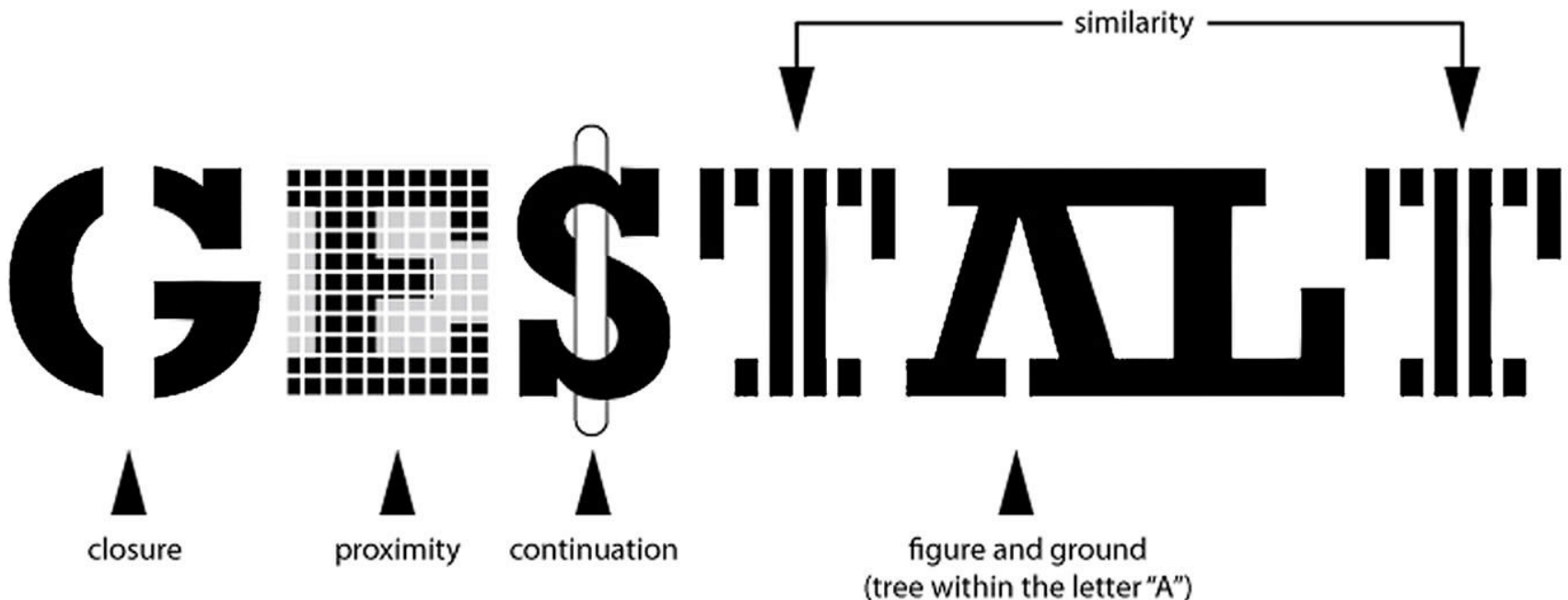
Pre-attentive visual properties (Healey 97)

Property	Literature
<i>length</i>	Triesman & Gormican [1988]
<i>width</i>	Julesz [1985]
<i>size</i>	Triesman & Gelade [1980]
<i>curvature</i>	Triesman & Gormican [1988]
<i>number</i>	Julesz [1985]; Trick & Pylyshyn [1994]
<i>terminators</i>	Julesz & Bergen [1983]
<i>intersection</i>	Julesz & Bergen [1983]
<i>closure</i>	Enns [1986]; Triesman & Souther [1985]
<i>colour (hue)</i>	Nagy & Sanchez [1990, 1992]; D'Zmura [1991]; Kawai et al. [1995]; Bauer et al. [1996]
<i>intensity</i>	Beck et al. [1983]; Triesman & Gormican [1988]
<i>flicker</i>	Julesz [1971]
<i>direction of motion</i>	Nakayama & Silverman [1986]; Driver & McLeod [1992]
<i>binocular lustre</i>	Wolfe & Franzel [1988]
<i>stereoscopic depth</i>	Nakayama & Silverman [1986]
<i>3D cues</i>	Enns [1990]
<i>lighting direction</i>	Enns [1990]

Gestalt principles

Gestalt Principles (Princ. of Visual Grouping)

Visual grouping principles have been invented by Gestalt psychologists and are based on human's innate ability to perceive objects as organized patterns and simplify complex compositions



Gestalt Principles (Princ. of Visual Grouping)

- Figure/Ground
- Proximity
- Similarity
- Symmetry
- Continuity
- Closure

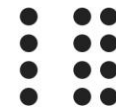
Gestalt Principles

UX CHEAT.com



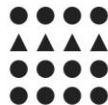
Good Figure

Objects grouped together tend to be perceived as a single figure. Tendency to simplify.



Proximity

Objects tend to be grouped together if they are close to each other.



Similarity

Objects tend to be grouped together if they are similar.



Continuation

When there is an intersection between two or more objects, people tend to perceive each object as a single uninterrupted object.



Closure

Visual connection or continuity between sets of elements which do not actually touch each other in a composition.

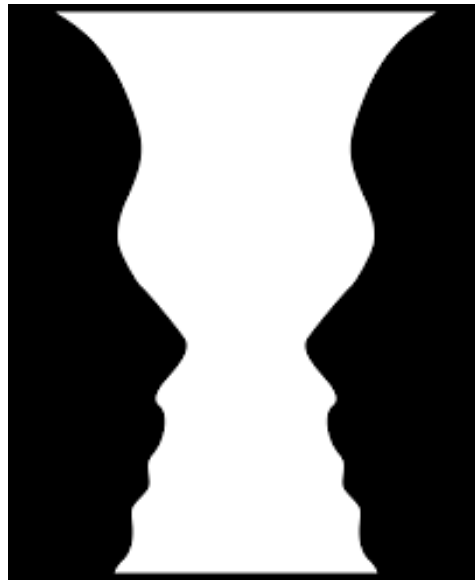


Symmetry

The object tend to be perceived as symmetrical shapes that form around their center.

Gestalt Properties: Figure/Ground

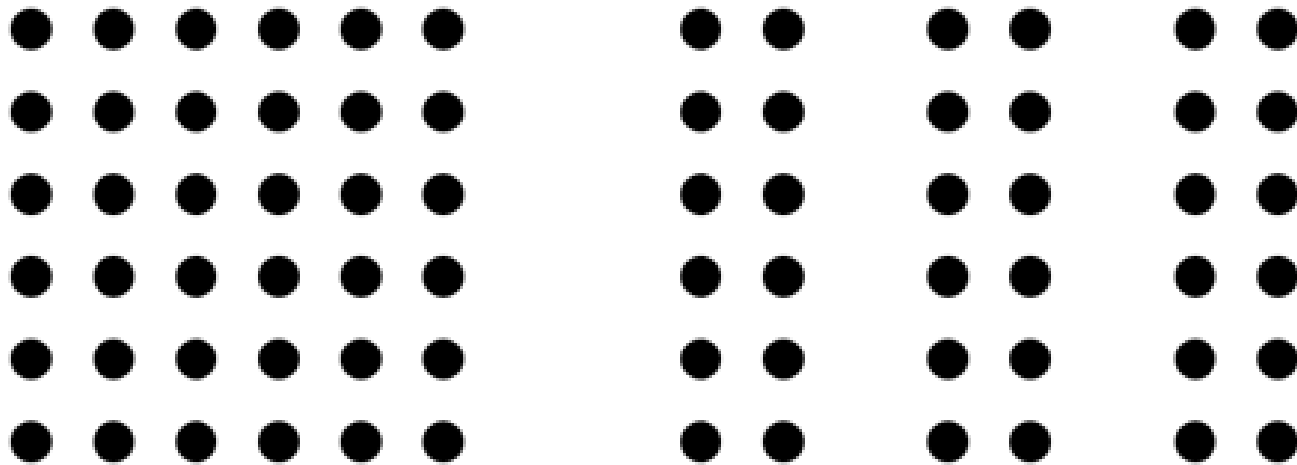
our mind separates the visual field into figure (foreground)
and ground (background)



you either see **two faces** as the figure and **a vase** as
the ground or the other way around

Gestalt Properties: Proximity

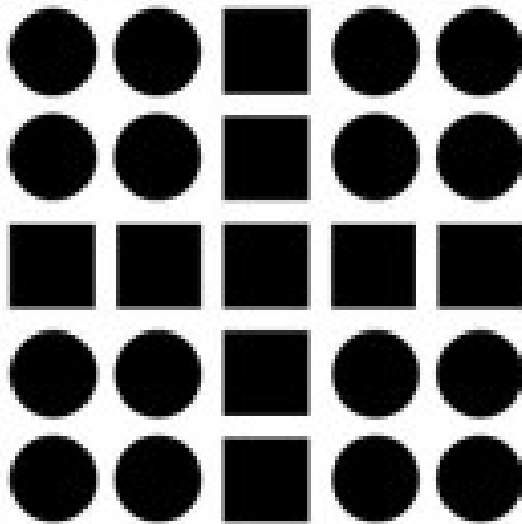
elements that are placed close to each other will often be perceived as one group



when the dots are placed close to each other in groups, we tend to perceive 3 columns rather than a whole group of dots

Gestalt Properties: Similarity

elements that have similarities will often be viewed together as a group or a pattern



similar shapes form different groups

Gestalt Properties: Continuity

continuity refers to when the viewer's eye is compelled to move beyond an object and continue through



the horizontal 'swipe' defines the figure as the letter H,
but also draws the viewers eye towards the leaf

Gestalt Properties: Closure

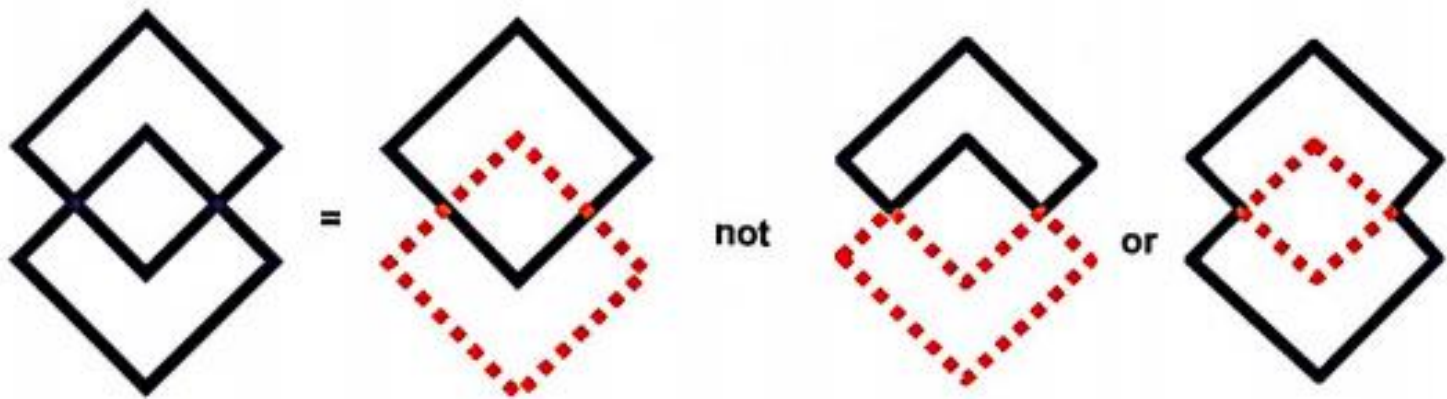
closure is about giving (the brain) enough information so that it can finish forming figures even if they don't explicitly exist



a series of horizontal lines arranged in such a way that our minds see the famous IBM logo

Gestalt Properties: Symmetry

symmetry is about our tendency to parse complex scenes in a way that reduces complexity



the human visual system tries to resolve complex scenes
into combinations of simple, symmetrical shapes

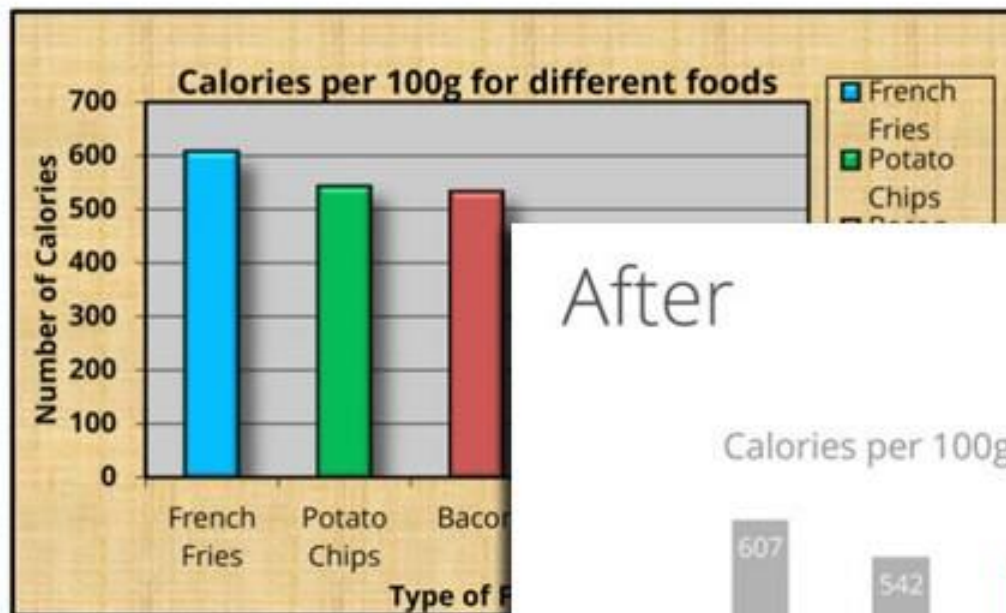
Principles of graphical excellence

Principles of Graphical Excellence (Tufte' 01)

- Show the data
- Induce the viewer to think about the substance of the findings rather than the methodology, the graphical design, or other aspects
- Avoid distorting what the data have to say
- Present many numbers in a small space, i.e, efficiently
- Make large data sets coherent
- Encourage the eye to compare different pieces of data
- Reveal the data at several levels of detail, from a broad overview to the fine structure
- Serve a clear purpose: description, exploration, tabulation, decoration
- Be closely integrated with the statistical and verbal descriptions of the data set

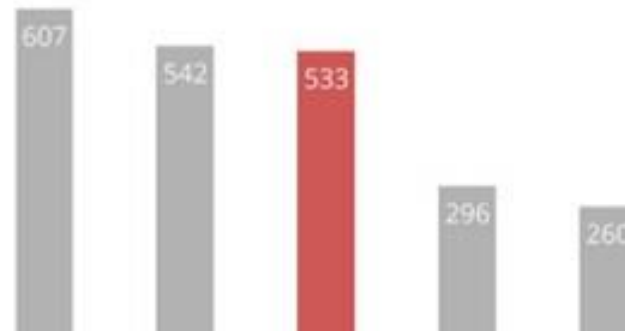
Show the data means high data to ink ratio

Before



After

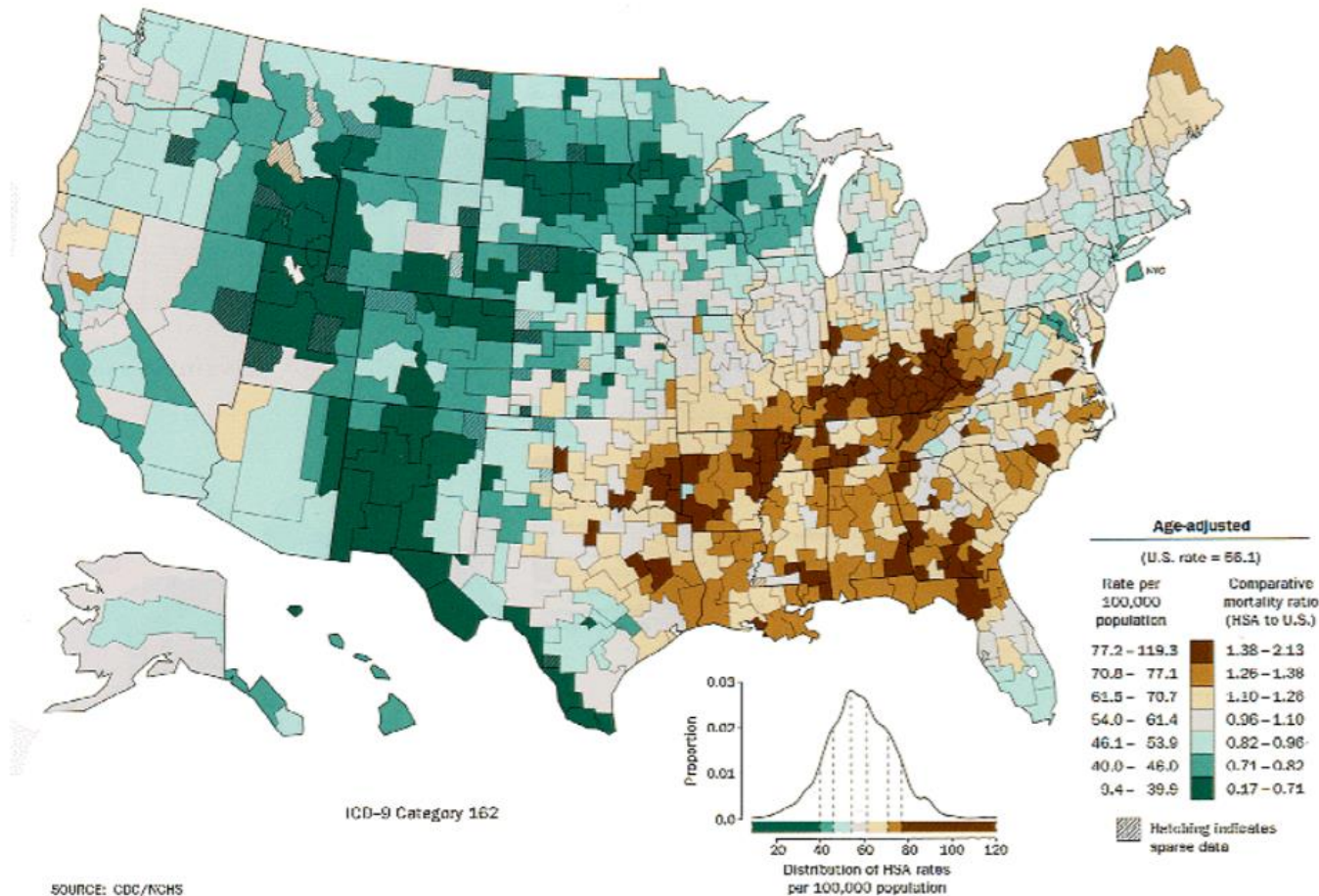
Calories per 100g



High data to ink ratio (demo)

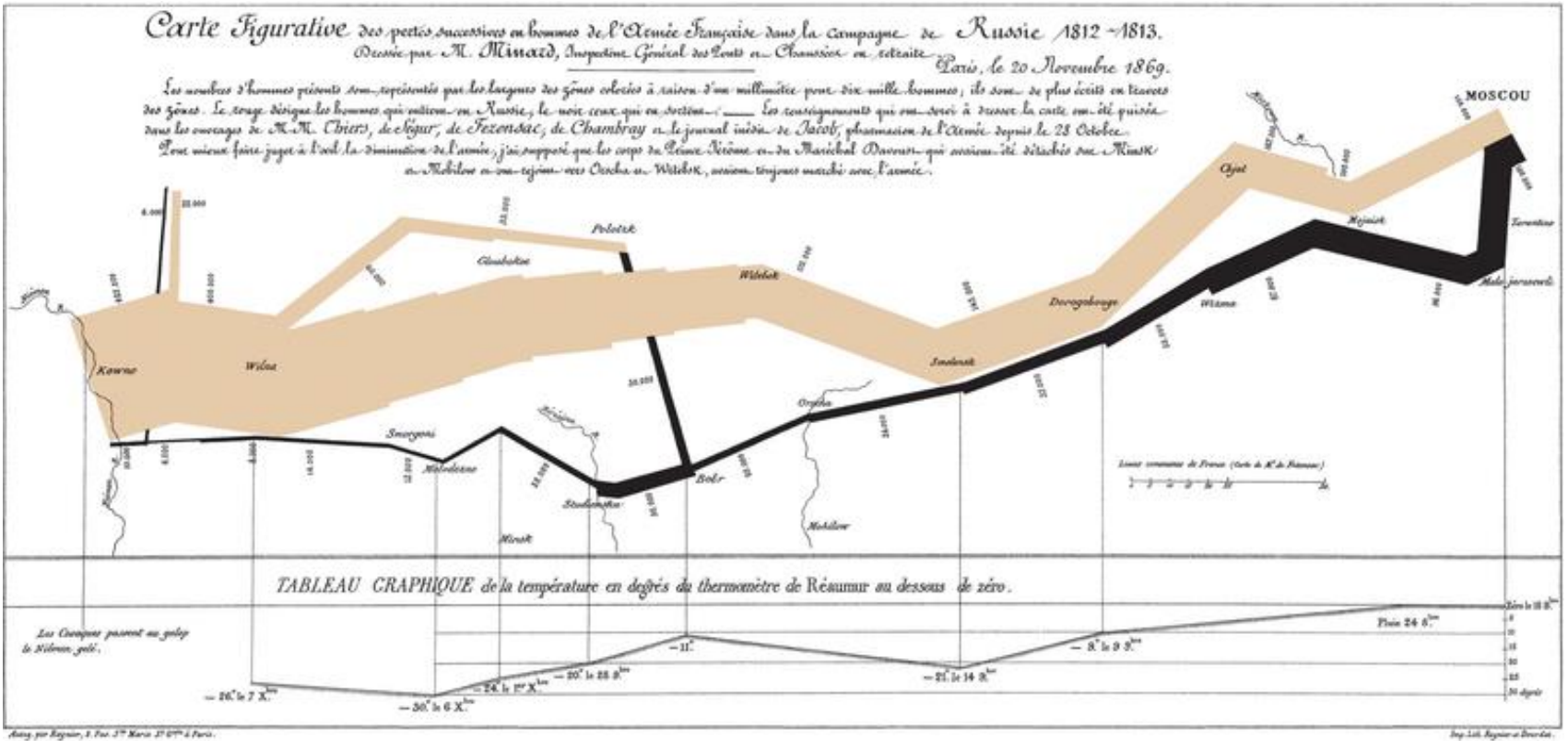
Remove
to improve
(the **data-ink** ratio)

Present many numbers in a small space



efficient visualization of large volume of data in a small space

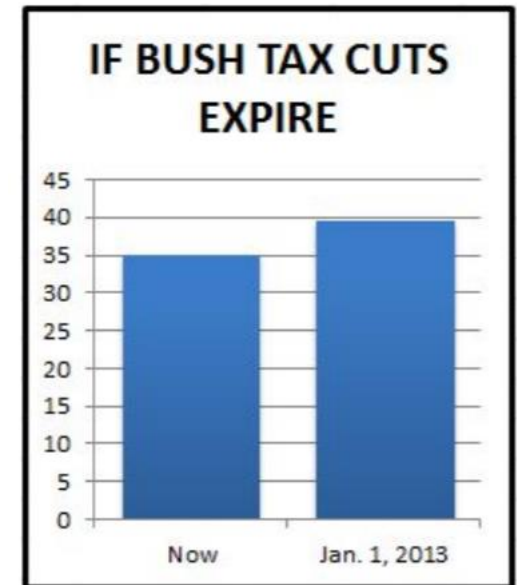
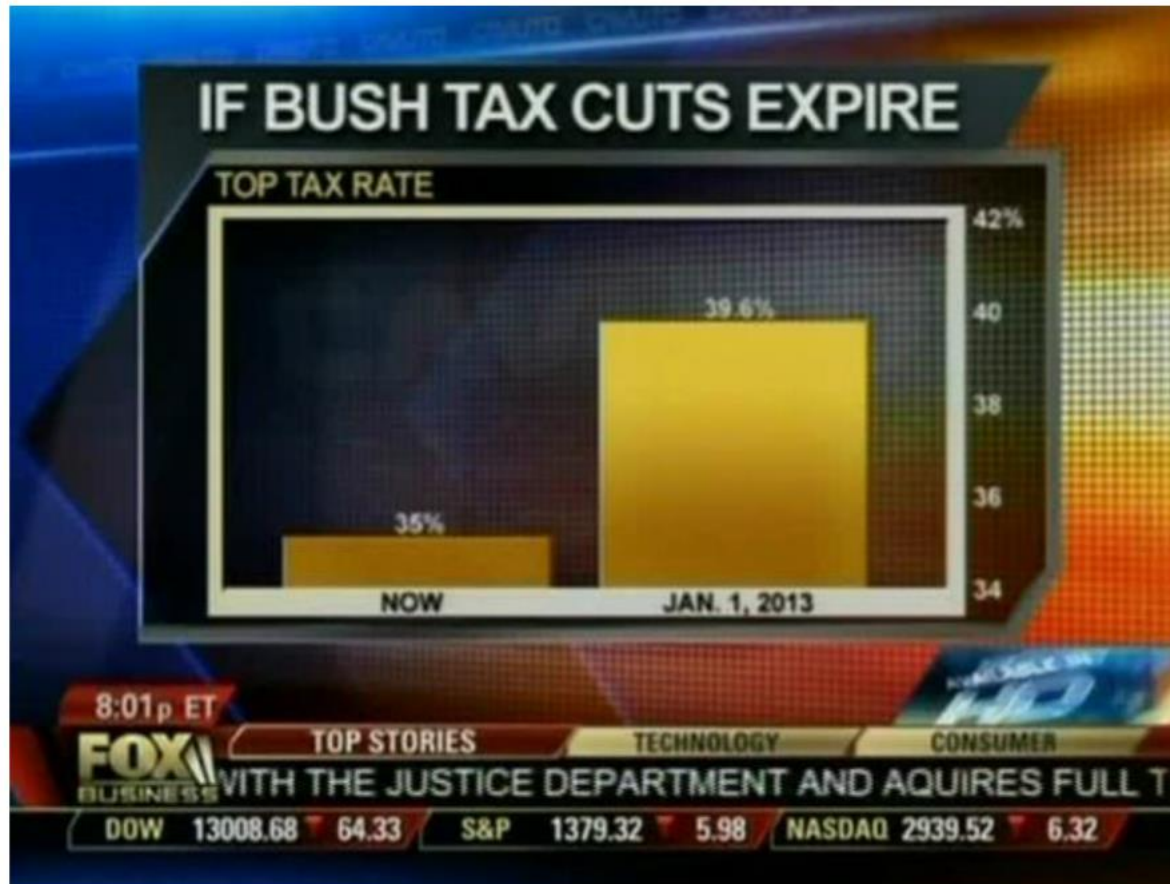
Reveal the data at several levels of detail



Charles Minard's map of **Napoleon's disastrous Russian campaign of 1812**. The graphic is notable for its representation in two dimensions of **six types of data**: *the number of Napoleon's troops; distance; temperature; the latitude and longitude; direction of travel; and location relative to specific dates*

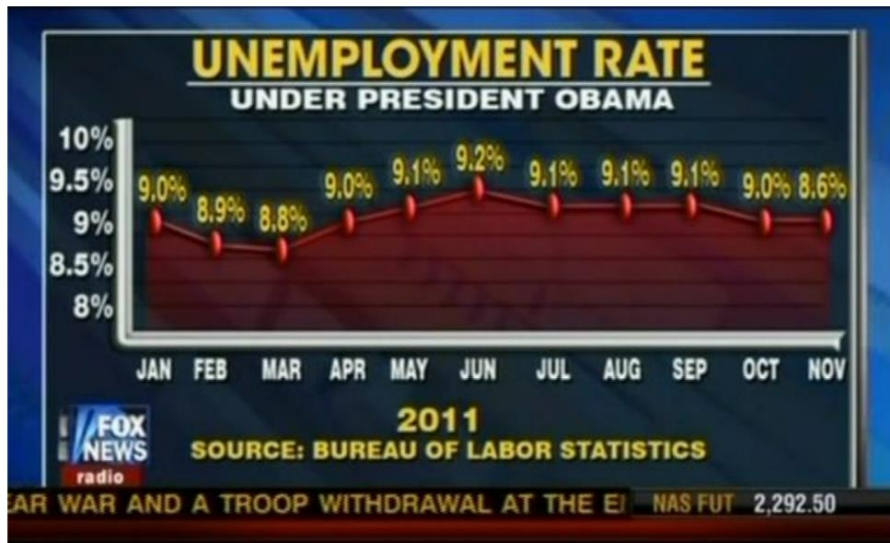
Avoid distorting what the data have to say

Misleading graphics



Avoid distorting what the data have to say

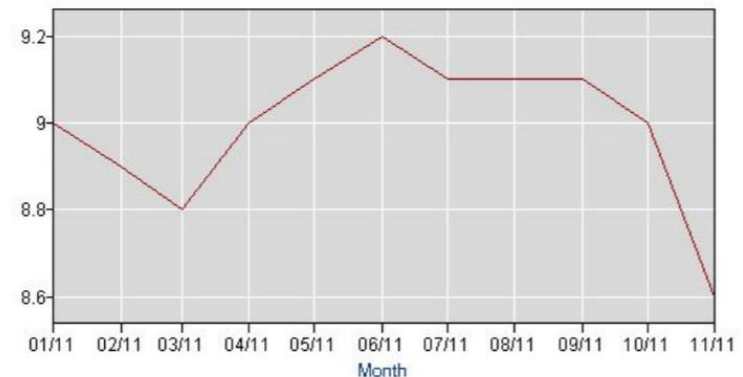
Misleading graphics



Data extracted on: December 12, 2011 (9:50:59 AM)

Labor Force Statistics from the Current Population Survey

Series Id: LNS14000000
Seasonally Adjusted
Series title: (Seas) Unemployment Rate
Labor force status: Unemployment rate
Type of data: Percent or rate
Age: 16 years and over



Graphical integrity

The representation of numbers, as physically measured on the surface of the graphic itself, should be directly proportional to the quantities represented

$$\text{lie factor} = \frac{\text{size of effect shown in graphic}}{\text{size of effect in data}}$$

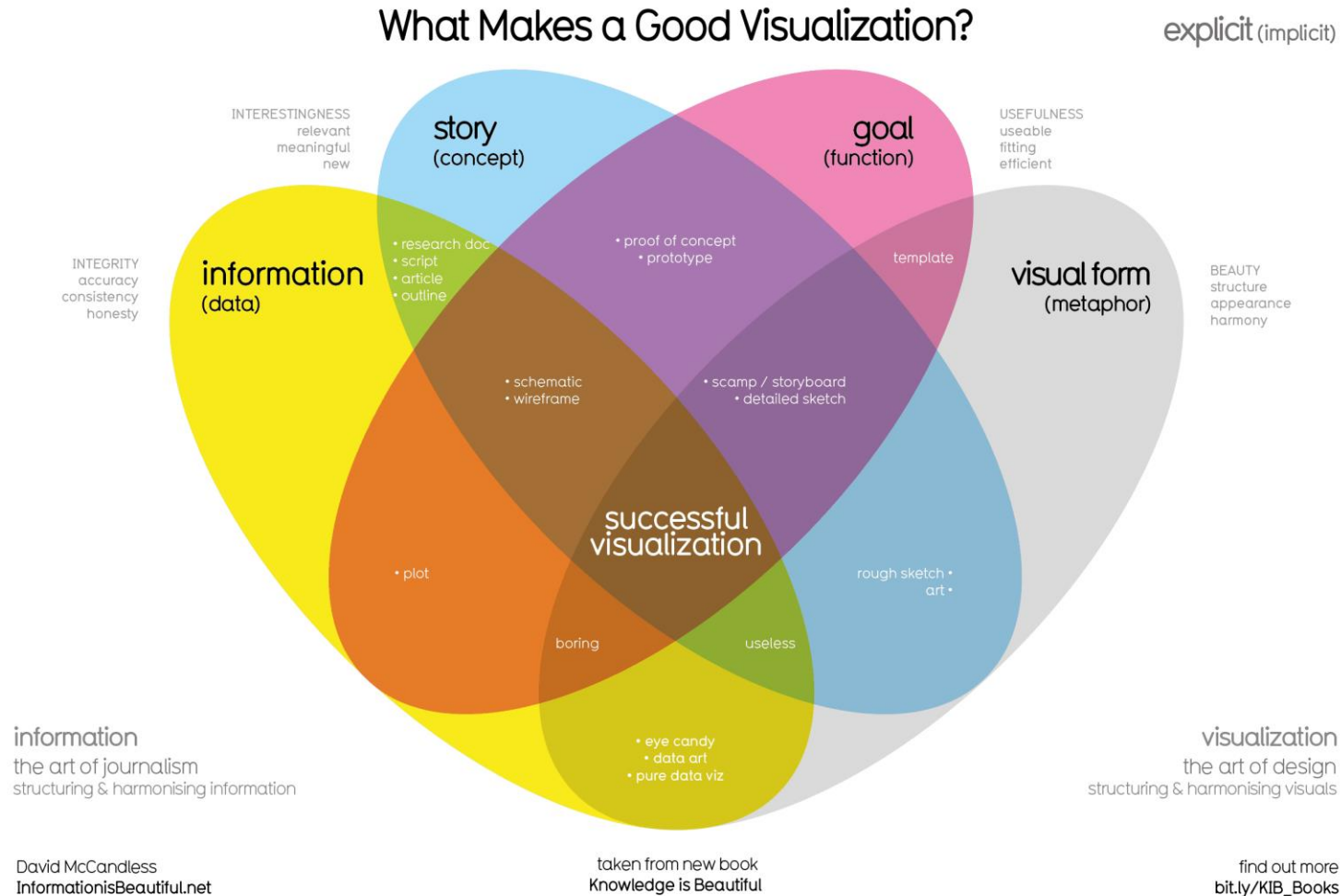
More examples (along with corrections) of misleading graphics

<http://qz.com/580859/the-most-misleading-charts-of-2015-fixed/>

See this research study on deceptive graphics

<https://dl.acm.org/doi/10.1145/2702123.2702608>

What makes a visualization beautiful?



<https://informationisbeautiful.net/visualizations/what-makes-a-good-data-visualization/>

Beyond digital visualizations

Physical visualizations (data sculpture)

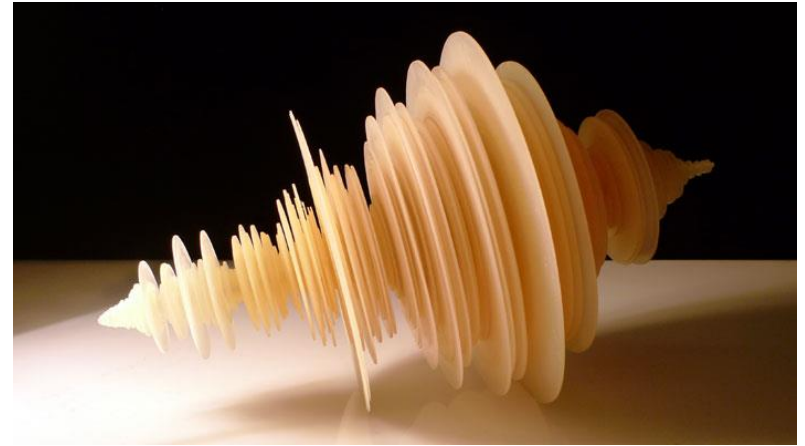
Keyboard Frequency Sculpture



A 3D bar chart on top of a keyboard which shows the frequency of each letter in the alphabet

Source: Michael Knuepfel

2011 – Tōhoku Japanese Earthquake Sculpture



A data sculpture by Luke Jerram that depicts nine minutes of seismographic readings during the 9.0 earthquake.

Source: Gizmodo

Physical visualizations



Manifest Justice Exhibition, Los Angeles, May 2015

<http://www.afropunk.com/profiles/blogs/feature-manifestjustice-art-exhibit-in-los-angeles>

Participatory visualization



<https://www.youtube.com/watch?v=hD5f8GuNuGQ>
(what is privilege?)