DeCART Summer School 2018

Data Visualization

Day 1 - Morning

16 July 2018

Nils Gehlenborg, PhD - Harvard Medical School

Instructor

Nils Gehlenborg, PhD

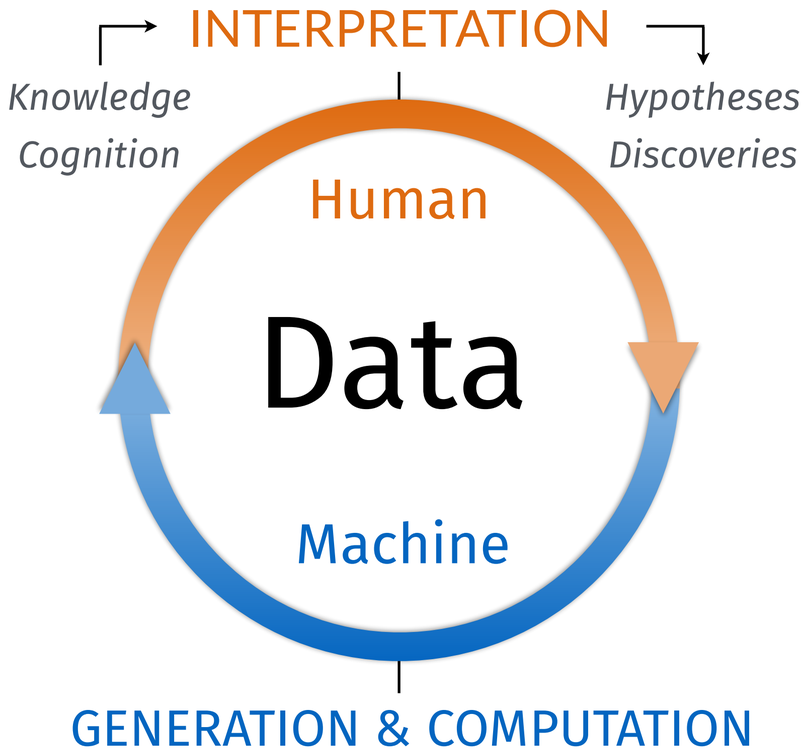
nils@hms.harvard.edu

Harvard Medical School

Department of Biomedical Informatics

http://gehlenborglab.org

Methods for Data Visualization and Exploration



Tools for Reproducible Research

Syllabus

- **Day 1**
  - Morning: Introduction to Data Visualization
  - Afternoon: Design Process, Evaluation, and Interaction
- **Day 2**
  - Morning: Introduction to Altair and visualization of high-dimensional and temporal data
  - Afternoon: Advanced Altair and visualization of genomic and network data

Syllabus - Day 1 (Morning)

- 9:00 - 9:15 | Introduction
- 9:15 - 9:45 | Design Exercise 1
- 9:45 - 10:15 | What? Why?
- 10:15 - 10:25 | Break
- 10:25 - 11:30 | How? How not!

Design your first visualization!

Laboratory-Confirmed Influenza Cases in the first ten weeks of 2010

	Asia	*S. America*	*Australia*	*N. America*	*Asia*	*S. America*	*Africa*	*Europe*	*Europe*	*Africa*	*N. America*
Week	Afghanistan	Argentina	Australia	Canada	China	Colombia	Egypt	Germany	Ireland	South Africa	USA
1	5	4	2	41	2179	36	739	26	23	0	366
2	13	21	1	15	2213	36	396	24	8	1	396
3	4	6	1	8	2228	14	192	18	4	0	447
4	0	1	0	14	2027	11	80	NA	8	0	402
5	0	4	1	12	1813	8	56	NA	4	0	404
6	0	0	1	6	1353	9	47	NA	0	0	361
7	1	3	0	6	799	7	32	NA	0	0	380
8	1	1	4	7	1218	5	16	NA	1	1	424
9	NA	0	3	3	1333	7	8	3	0	0	445
10	1	3	1	7	1614	5	8	7	0	0	475

Source: World Health Organization FluNet database (http://who.int/flunet)

What is data visualization?

What is data visualization?

The use of computer-supported, interactive, visual representations of data to amplify cognition.

—*Stu Card, Jock Mackinlay & Ben Shneiderman*

Computer-based visualization systems provide visual representations of datasets intended to help people carry out some task more effectively.

—*Tamara Munzner*

What is data visualization?

The use of computer-supported, interactive, visual representations of data to amplify cognition.

—*Stu Card, Jock Mackinlay & Ben Shneiderman*

Computer-based visualization systems provide visual representations of datasets intended to help **people** carry out some task **more effectively**.

—*Tamara Munzner*

What is data visualization?

What is data visualization?

> The purpose of computing is insight, not numbers.
>
> —*Richard Hamming*

What is data visualization?

The purpose of computing is insight, not numbers.

—*Richard Hamming*

The purpose of **visualization** is insight, **not pictures**.

—*Stu Card, Jock Mackinlay & Ben Shneiderman*

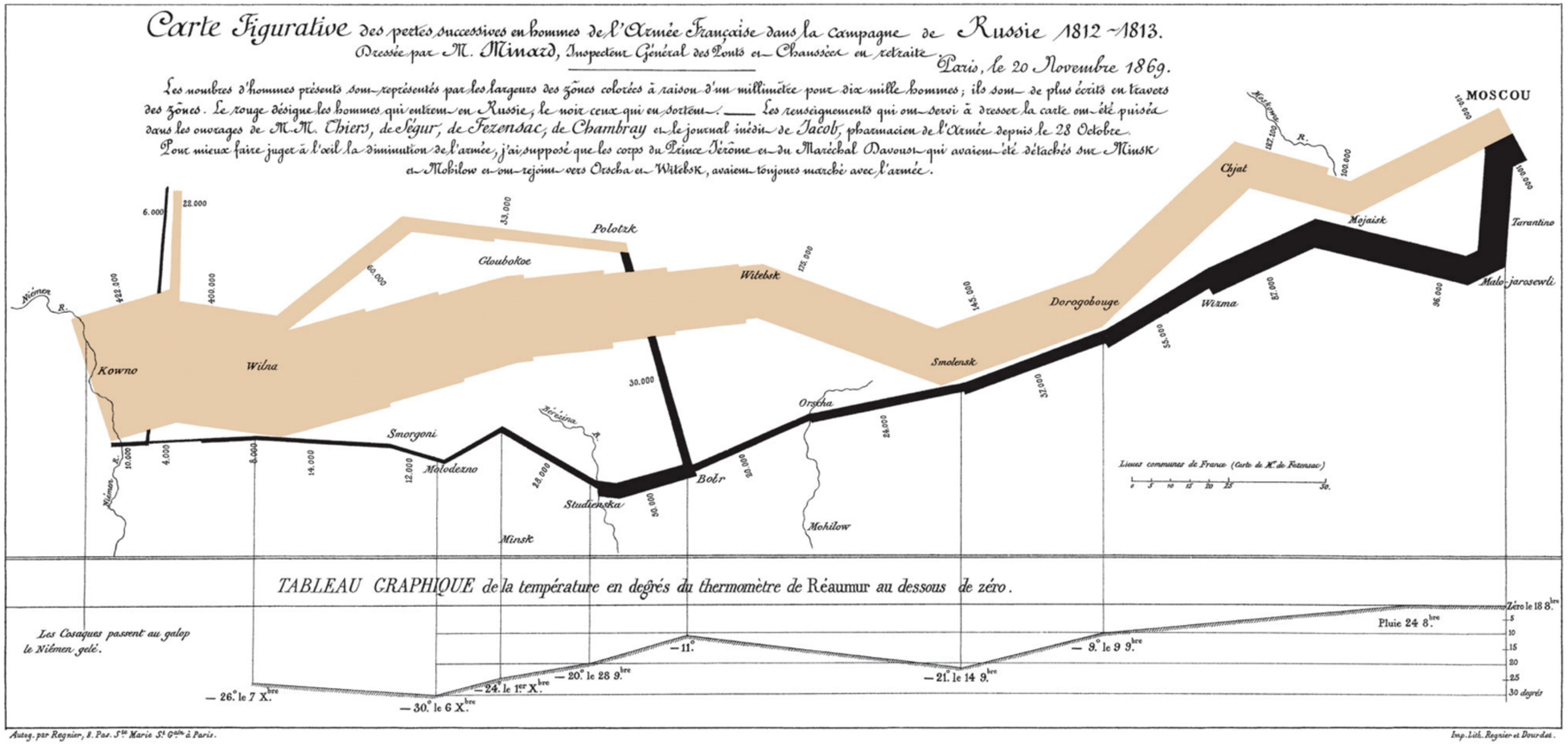
Why do we need data visualization?

Why do we need it?

A good sketch is better than a long speech.

—*Napoleon Bonaparte*

Charles Minard: Napoleon's March on Moscow



Minard 1869

Why do we need it?

I believe it when I see it.

—*Unknown*

Table 1.1: Anscombe's Quartet (Anscombe, 1973). In each of the four data sets mean $\mu_{X_i} = 9.0$, variance $\sigma^2_{X_i} = 11.0$, $\mu_{Y_i} = 7.5$, $\sigma^2_{Y_i} = 4.12$, correlation $cor(X_i, Y_i) = 0.816$ and the linear regression line is $Y_i = 3 + 0.5X_i$ for $i \in \{1,2,3,4\}$.

X_1	Y_1	X_2	Y_2	X_3	Y_3	X_4	Y_4
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

mean(X) = 9, var(X) = 11, mean(Y) = 7.5, var(Y) = 4.12,

cor(X,Y) = 0.816, linear regression line Y = 3 + 0.5*X

Anscombe 1973, The American Statistician

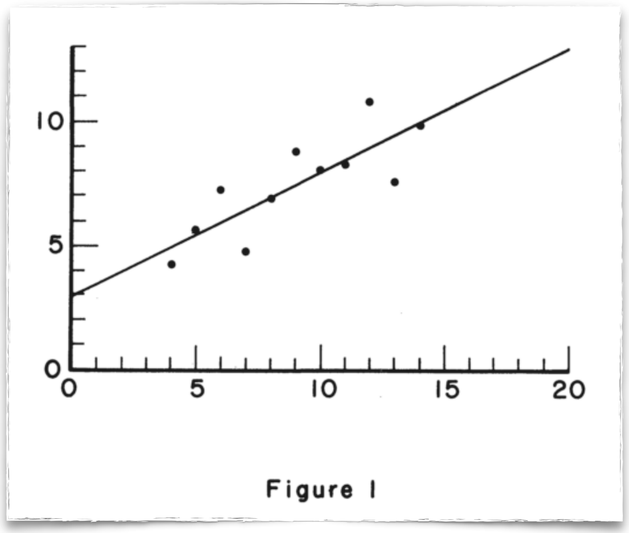


Figure 1

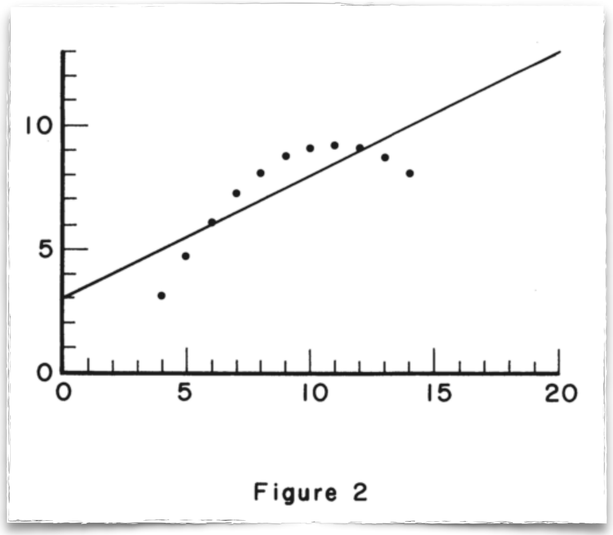


Figure 2

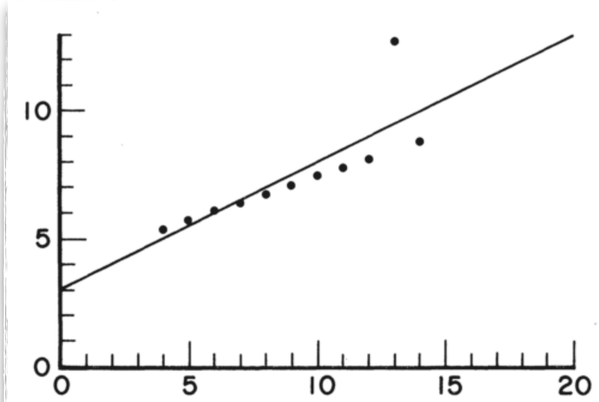


Figure 3

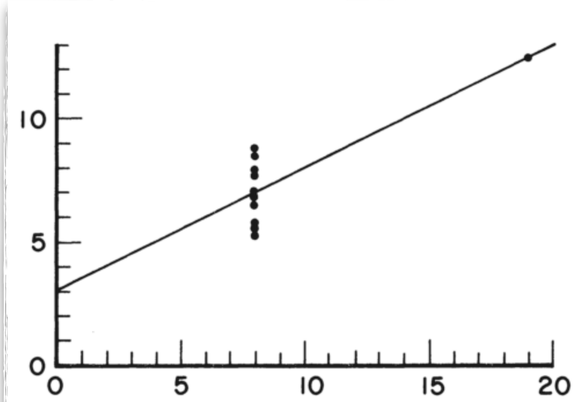


Figure 4

Anscombe 1973, The American Statistician

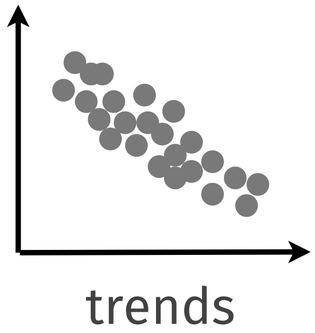
Why do we need it?

I’m wondering if there are any interesting patterns in my data.

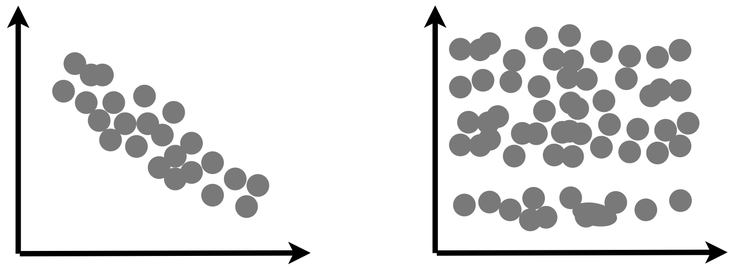
—*Almost Everyone*

Exploration: Hypothesis Generation

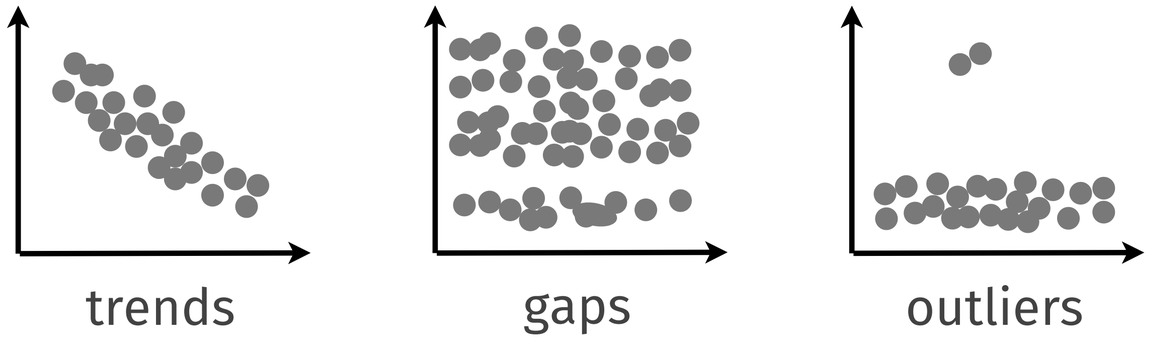
Exploration: Hypothesis Generation



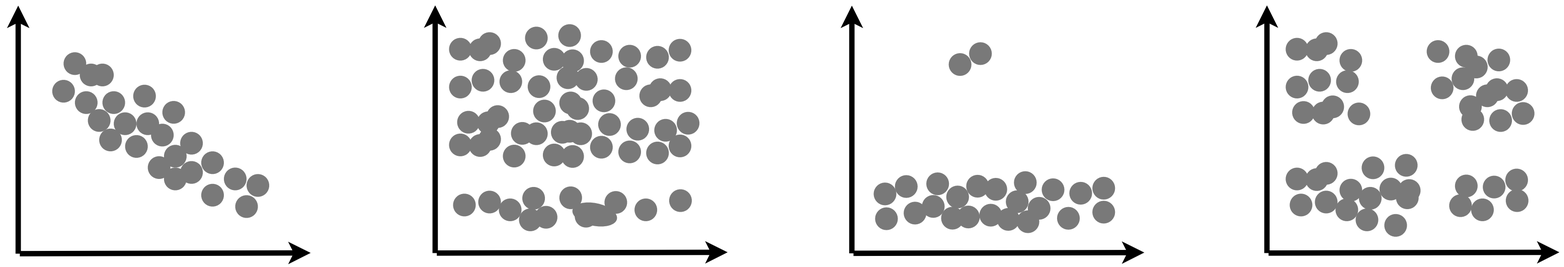
Exploration: Hypothesis Generation



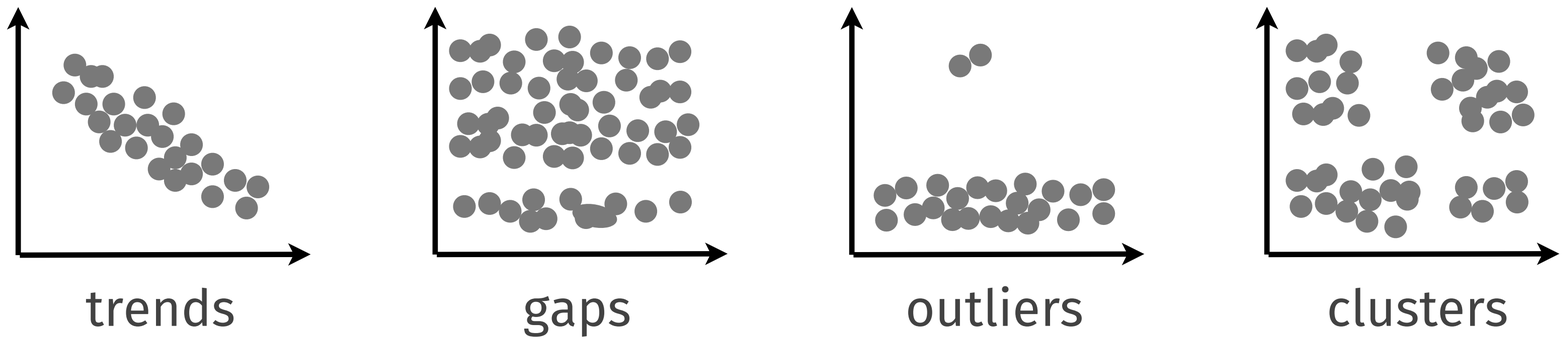
Exploration: Hypothesis Generation



Exploration: Hypothesis Generation



Exploration: Hypothesis Generation



Why? Generate hypotheses that can be tested with statistical methods or follow-up experiments.

How? Visualization is employed to perform pattern detection using the human visual system.

Visualization Use Cases

Visualization Use Cases

Exploration

Visualization Use Cases

Exploration

Confirmation

Visualization Use Cases

Exploration

Confirmation

Communication

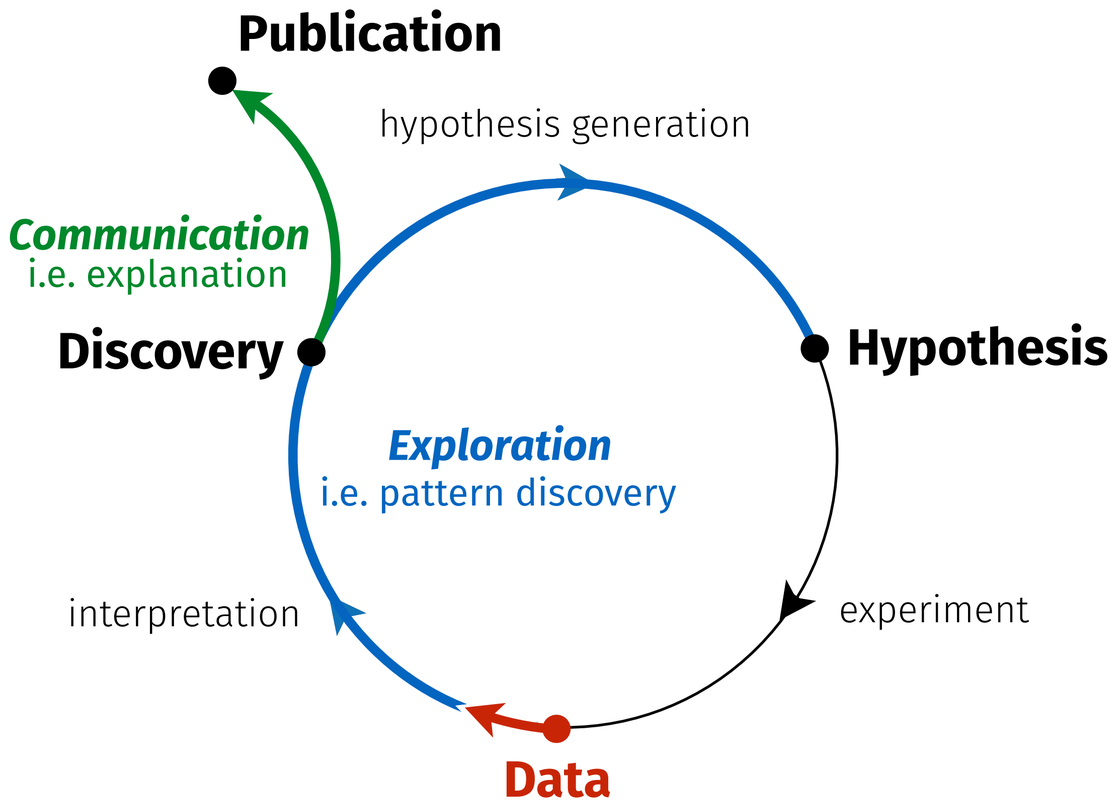
Visualization Use Cases

Exploration

Confirmation

Communication

Discovery Process



Discovery Process

Exploration

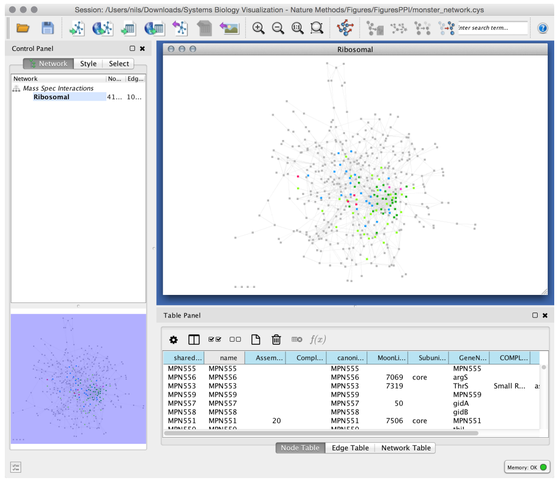
Communication

Discovery Process

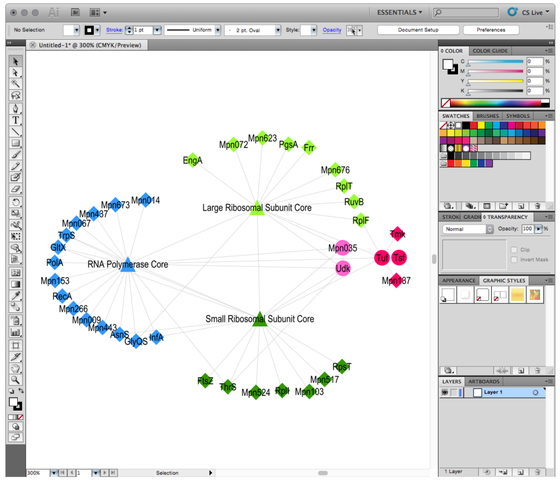
Discovery Process

Exploration ***Communication***

Insight



Cytoscape



Adobe Illustrator

Discovery Process

Discovery Process

Scientists

Discovery Process

Scientists

Scientists
Educators
Policy Makers
Citizens
...

Focus of the DeCART Course

Focus of the DeCART Course

Focus of the DeCART Course

How does it work?

> Visualization is really about *external cognition*, that is, how *resources outside the mind* can be used to *boost the cognitive capabilities* of the mind.
>
> —*Stu Card*

How does it work?

Why do we use the visual system and not other sensory systems?

Information bandwidth of visual system is much higher than of all other sensory system.

How does it work?

Visualization uses perception to free up cognition.

How does visualization work?

MALWMRLLPLLALLALWGPDPA

AAFVNQHLCGSHLVEALYLVCG

ERGFFYTPKTRREAEDLQVGQV

ELGGGPGAGSLQPLALEGSLQK

RGIVEQCCTSICSLYQLENYCN

http://www.uniprot.org/uniprot/P01308

How does visualization work?

MALWMRLLPLLALLALWGPDPA

AAFVNQHLCGSHLVEALYLVCG

ERGFFYTPKTRREAEDLQVGQV

ELGGGPGAGSLQPLALEGSLQK

RGIVEQCCTSICSLYQLENYCN

http://www.uniprot.org/uniprot/P01308

How does visualization work?

Visualization uses perception to free up cognition.

Visualization is an external cognitive aid
and augments working memory.

How does visualization work?

453 x 862 = ?

How does visualization work?

```
  453 x 862 = ?
-------------
      906
+  27,180
+ 362,400
-------------
  390,486
```

How does visualization work?

	f_1		f_2	
	x	y	x	y
1	10	9.14	10	7.46
2	8	8.14	8	6.77
3	13	8.74	13	12.74
4	9	8.77	9	7.11
5	11	9.26	11	7.81
6	14	8.10	14	8.84
7	6	6.13	6	6.08
8	4	3.10	4	5.39
9	12	9.13	12	8.15
10	7	7.26	7	6.42
11	5	4.74	5	5.73

Tasks

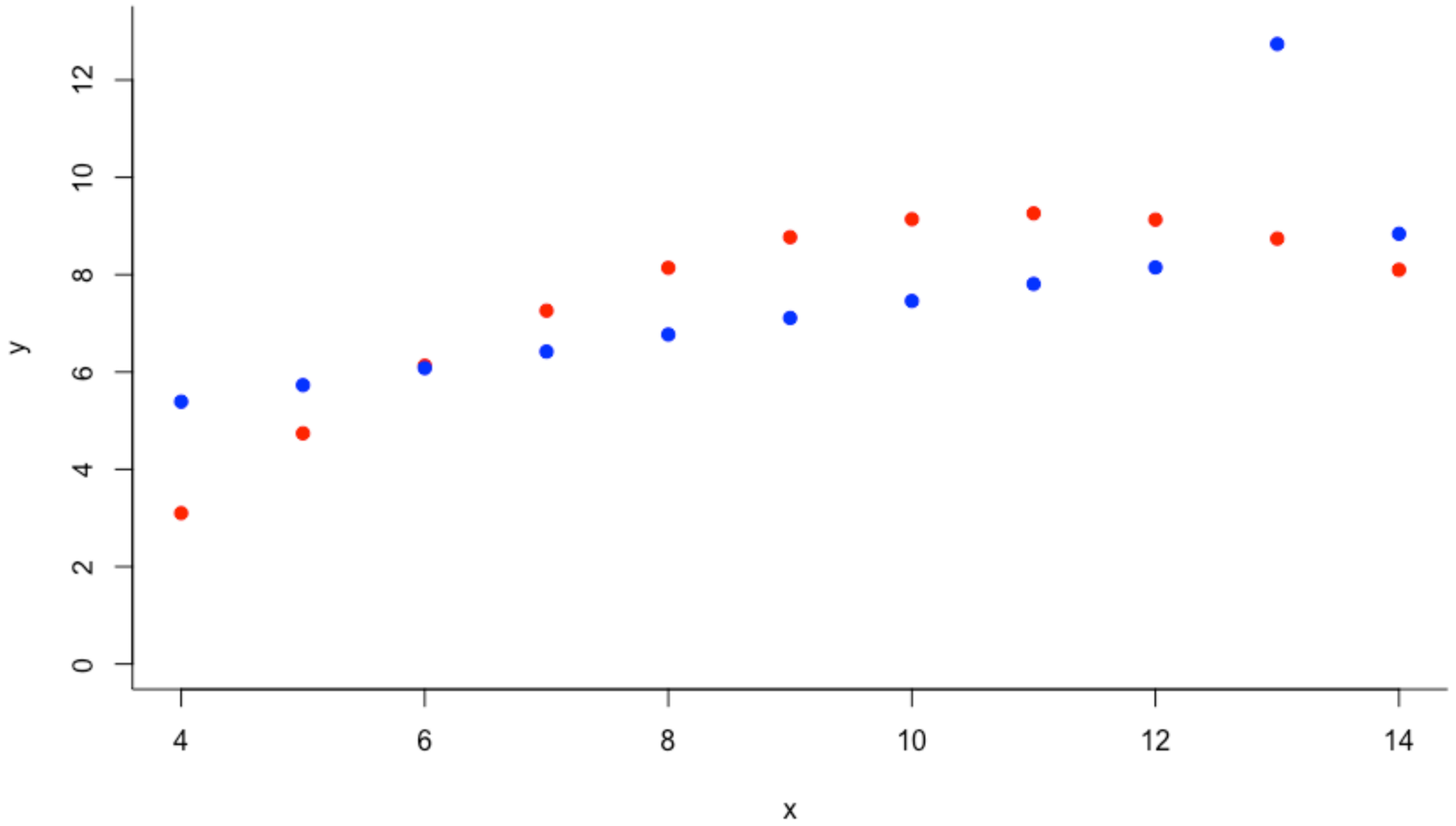
What is the shape of f_1? Of f_2?

How many times do f_1 and f_2 intersect?

Do they cross 0?

…

How does visualization work?

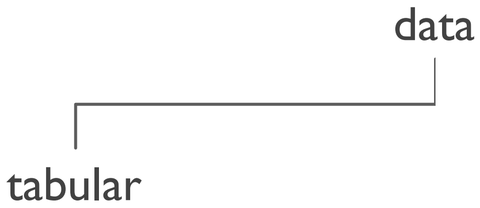


Design Critique

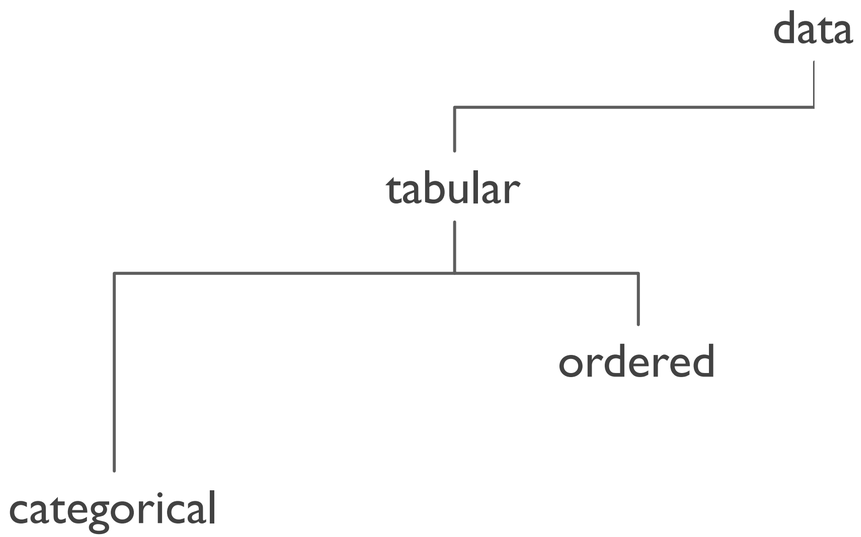
Visual Encoding of Data

data

Visual Encoding of Data

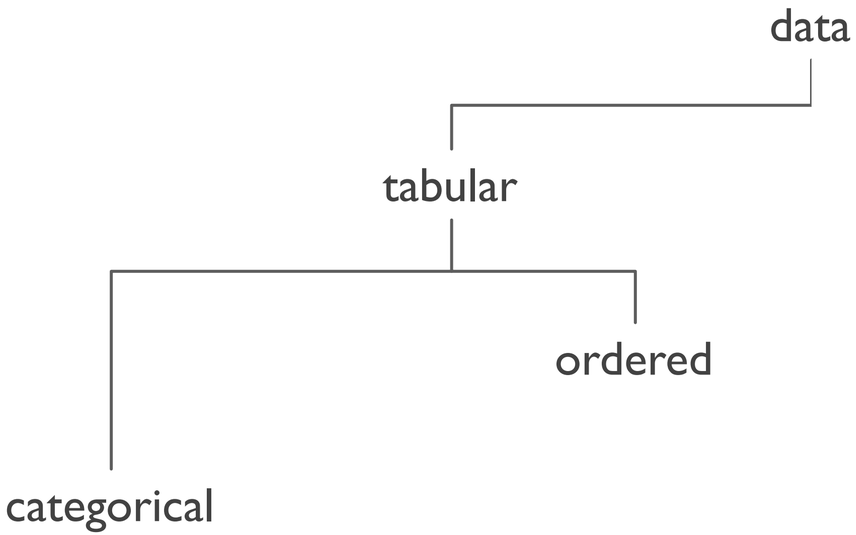


Visual Encoding of Data



slide adapted from Munzner 2011, Visualization Principles

Visual Encoding of Data

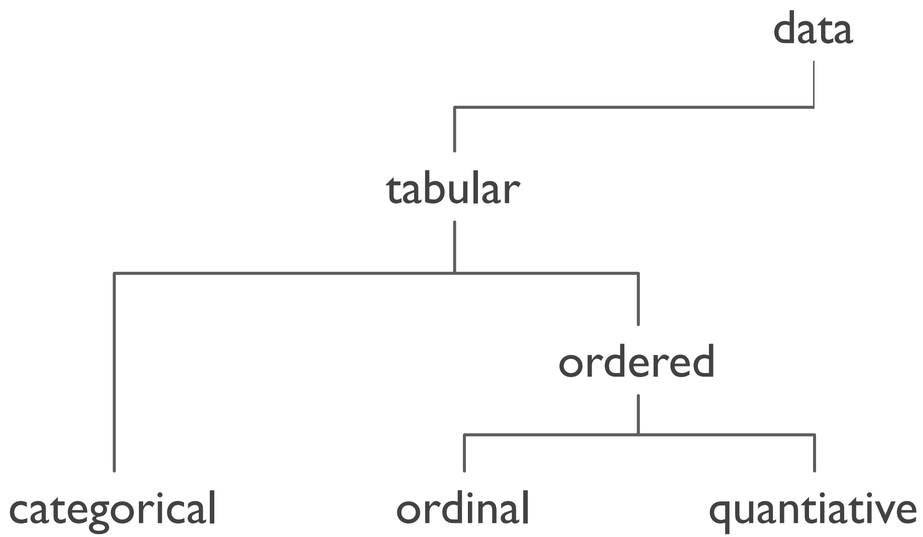


apples

oranges

bananas

Visual Encoding of Data

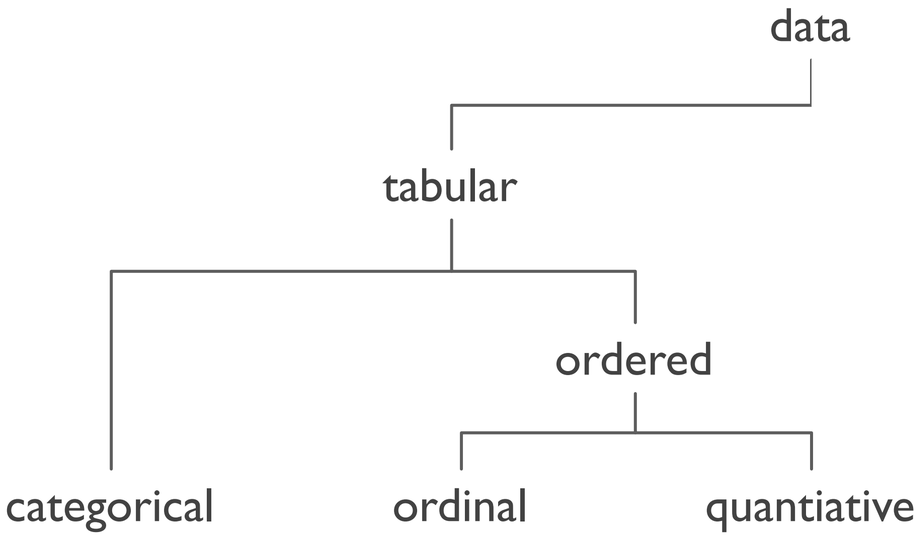


apples

oranges

bananas

Visual Encoding of Data

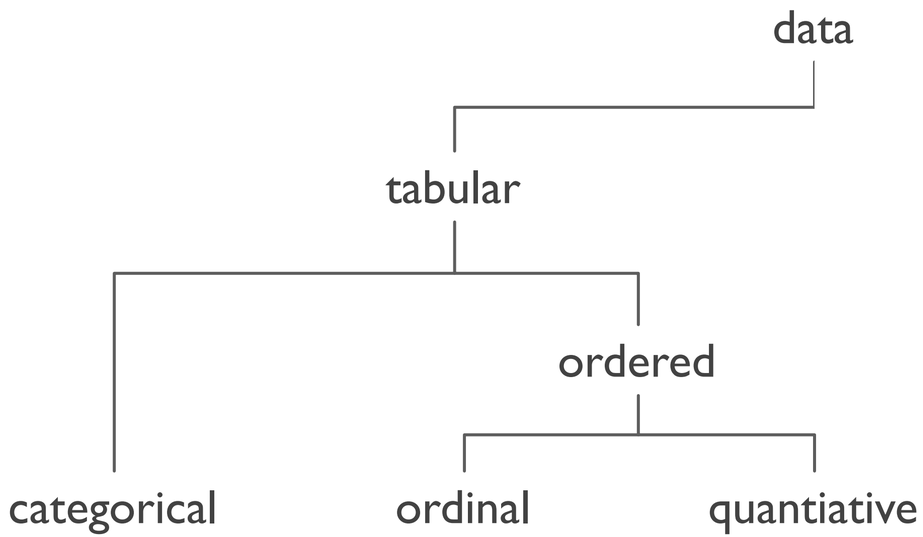


apples
oranges
bananas

small
medium
large

slide adapted from Munzner 2011, Visualization Principles

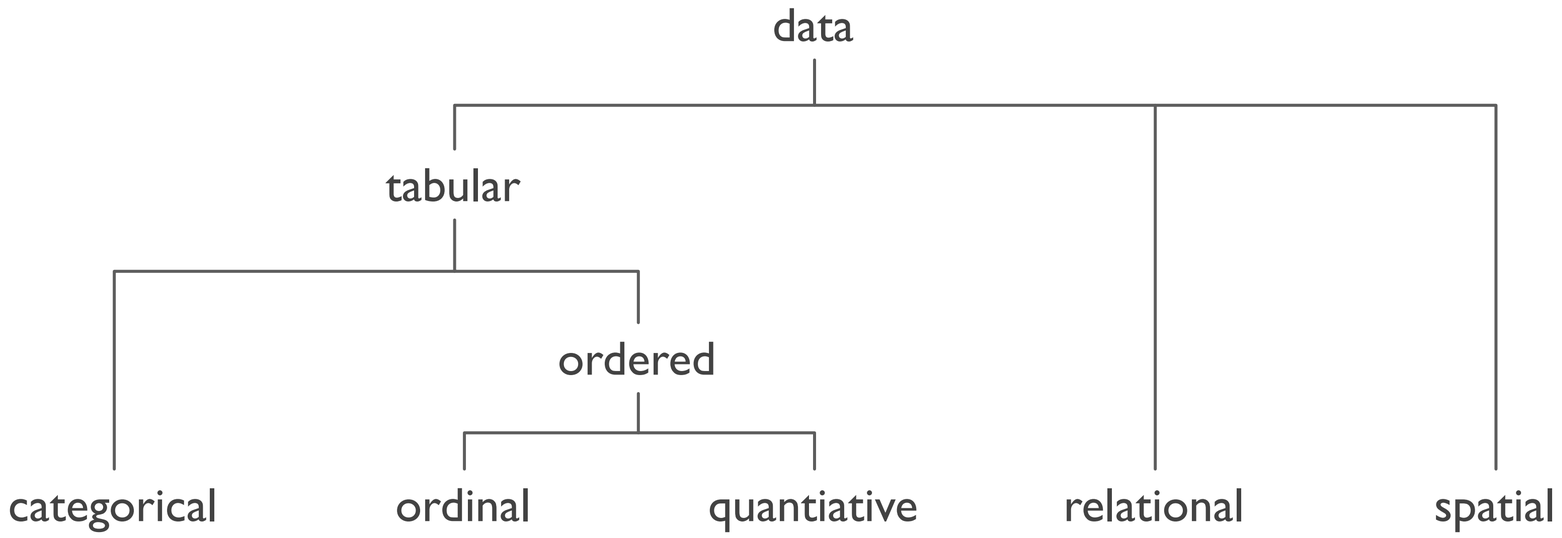
Visual Encoding of Data



apples	*small*	*10 inches*
oranges	*medium*	*13 inches*
bananas	*large*	*18.5 inches*

slide adapted from Munzner 2011, Visualization Principles

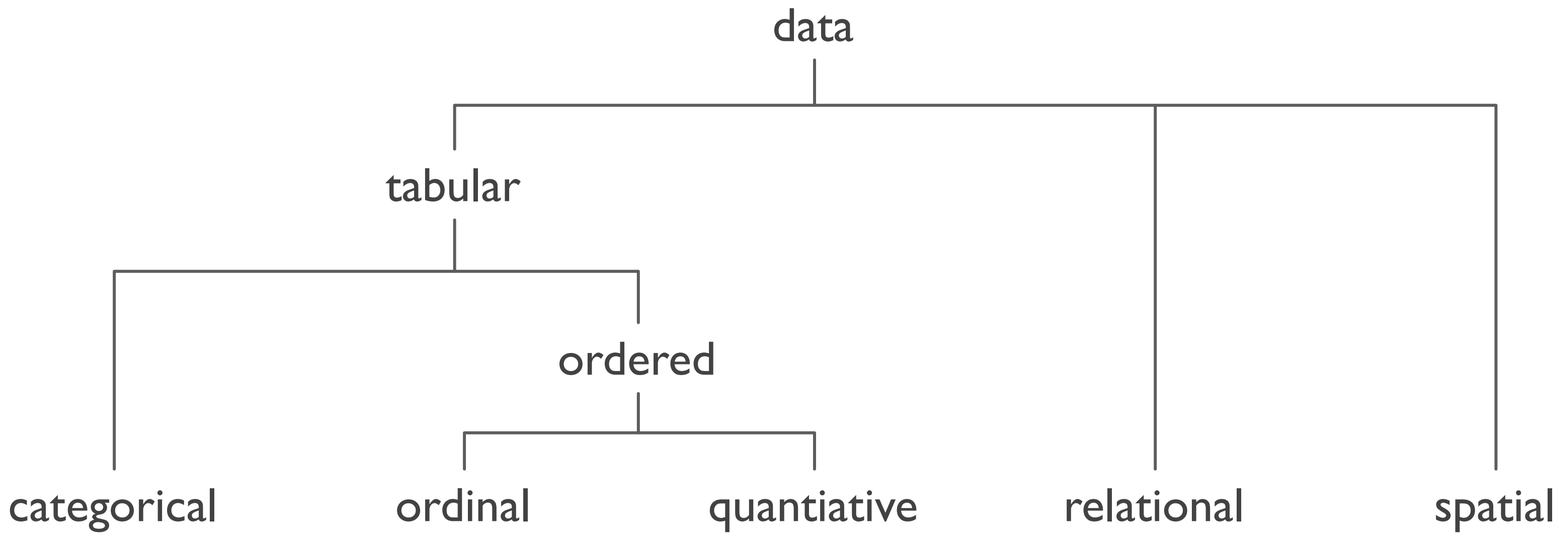
Visual Encoding of Data



apples	*small*	*10 inches*
oranges	*medium*	*13 inches*
bananas	*large*	*18.5 inches*

slide adapted from Munzner 2011, Visualization Principles

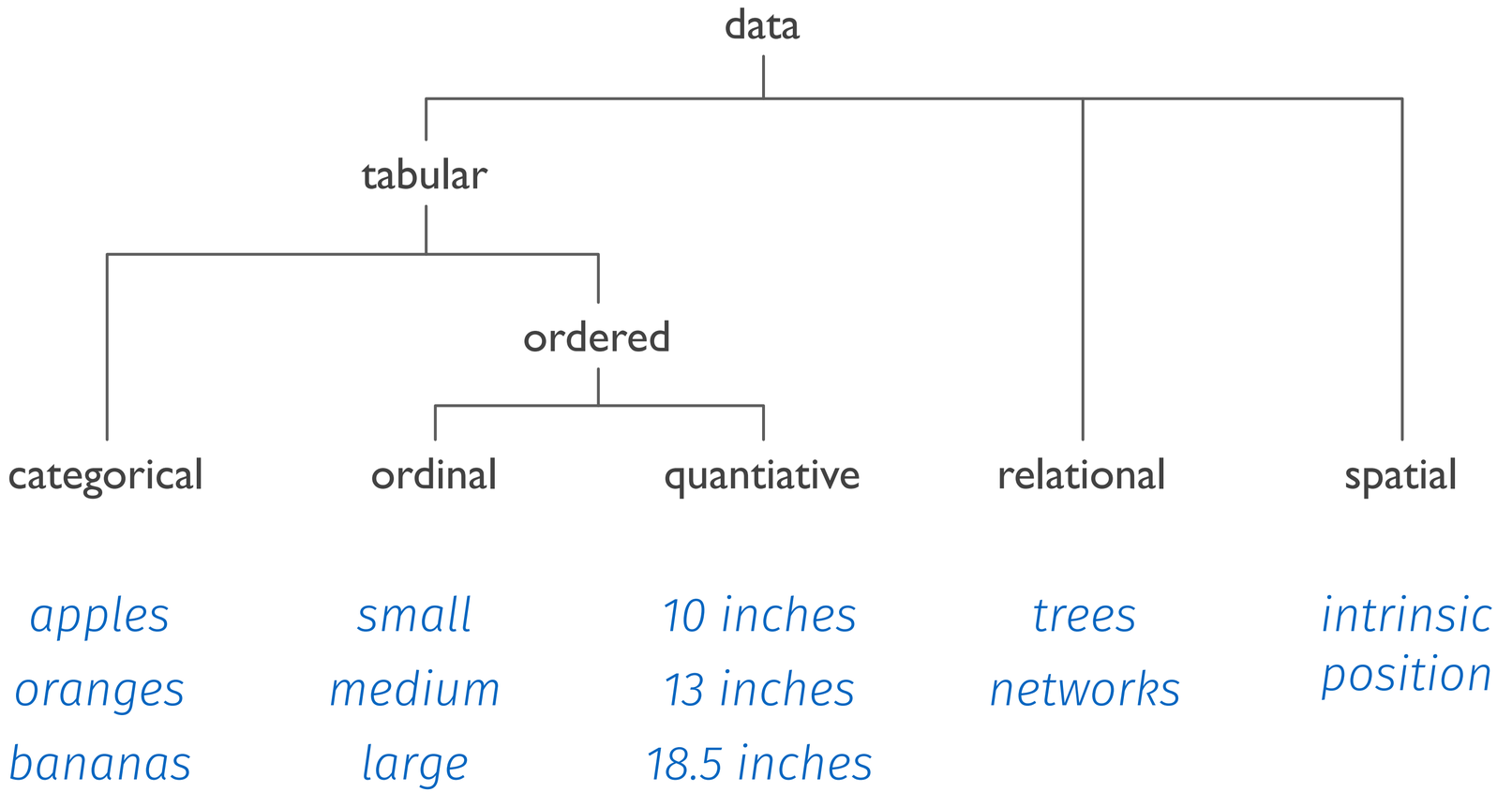
Visual Encoding of Data



apples	*small*	*10 inches*	*trees*
oranges	*medium*	*13 inches*	*networks*
bananas	*large*	*18.5 inches*	

slide adapted from Munzner 2011, Visualization Principles

Visual Encoding of Data



slide adapted from Munzner 2011, Visualization Principles

Exercise: Design Critique

Euler Diagram

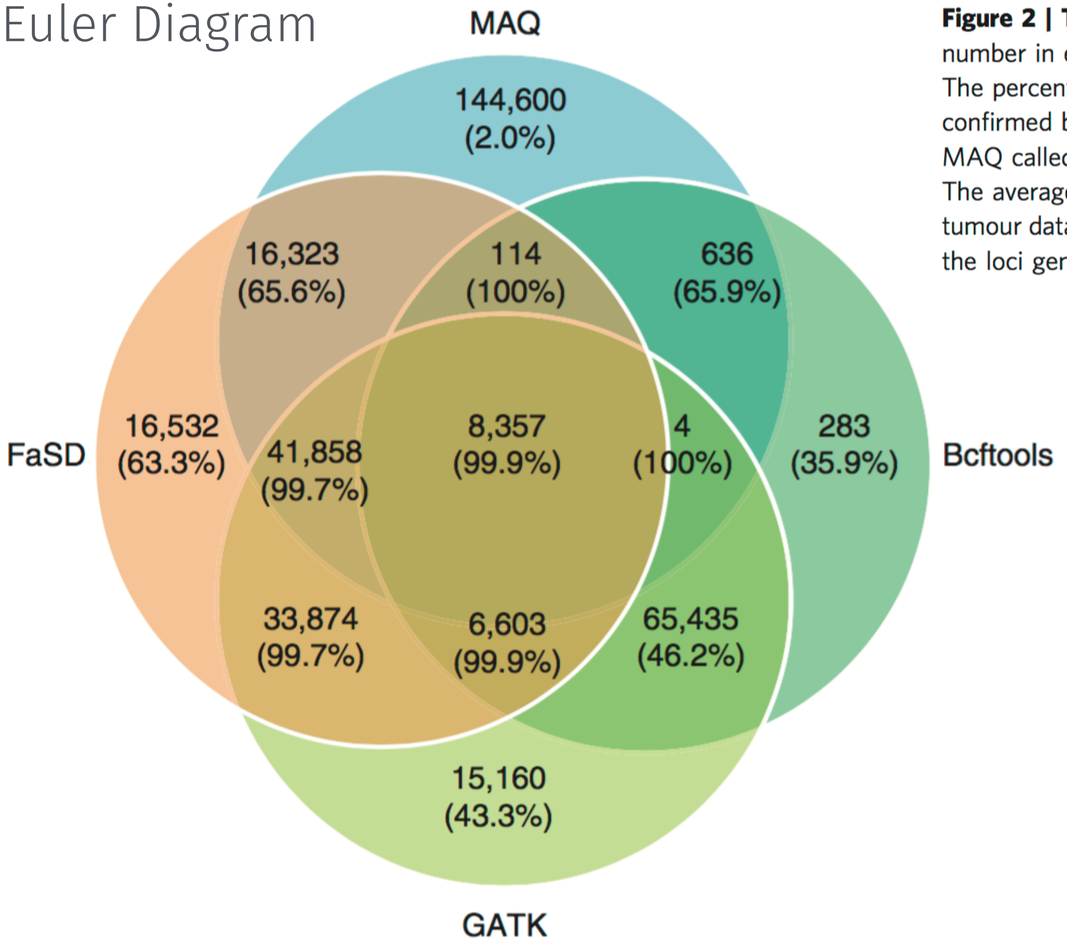


Figure 2 | The Venn diagram of SNPs detected by different tools. The number in each cell is the number of SNPs in the corresponding category. The percentage under the number is the proportion of SNPs that were confirmed by the Affymetrix SNP array. The FaSD, GATK, Bcftools and MAQ called 123661, 171291, 81432 and 211892 SNPs in total, respectively. The average depth of this data set was 10 ×. The figure is based on the tumour data set and Bowtie was used as aligner, and statistics are based on the loci genotyped by the Affymetrix SNP array.

F Xu et al., Nature Communications 3, 1258 doi:10.1038/ncomms2256

Visual Channels: Rankings

Visual Channels: Rankings

Categorical

What? Where?

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

color hue

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

color hue

shape

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

color hue

shape

slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

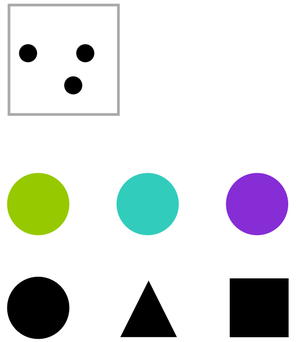
Categorical

What? Where?

position*
planar

color hue

shape



Relational

With whom?

slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

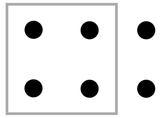
color hue

shape

Relational

With whom?

containment



Visual Channels: Rankings

Categorical

What? Where?

position*
planar

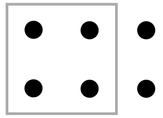
color hue

shape

Relational

With whom?

containment



connection

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

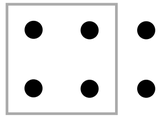
color hue

shape

Relational

With whom?

containment



connection

similarity

slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Categorical

What? Where?

position*
planar

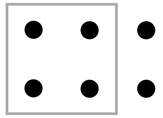
color hue

shape

Relational

With whom?

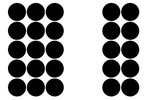
containment



connection

similarity

position*
proximity



slide adapted from Munzner 2011, Visualization Principles

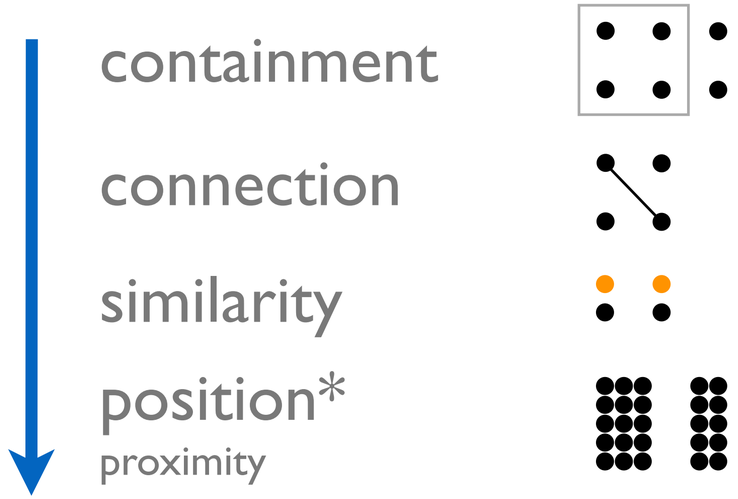
Visual Channels: Rankings

Categorical

What? Where?

Relational

With whom?



slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Visual Channels: Rankings

Ordinal &

Quantitative

How much?

Visual Channels: Rankings

Ordinal & Quantitative

How much?

position*
common scale

Visual Channels: Rankings

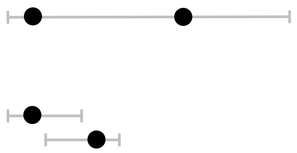
Ordinal &

Quantitative

How much?

position*
common scale

position*
unaligned scale



slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Ordinal &

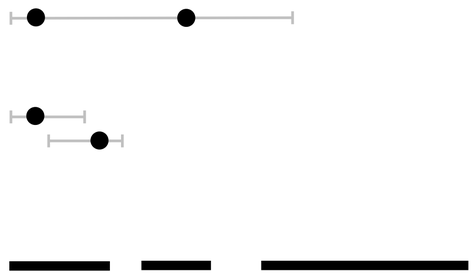
Quantitative

How much?

position*
common scale

position*
unaligned scale

length (1D)



slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Ordinal &

Quantitative

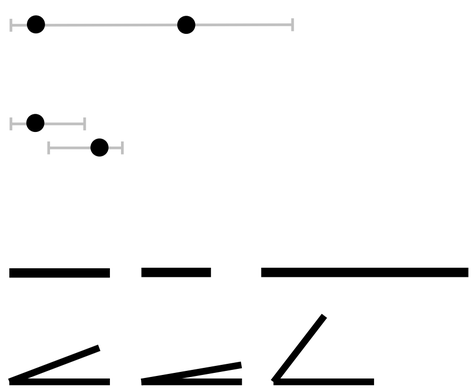
How much?

position*
common scale

position*
unaligned scale

length (1D)

angle/tilt



Visual Channels: Rankings

Ordinal &

Quantitative

How much?

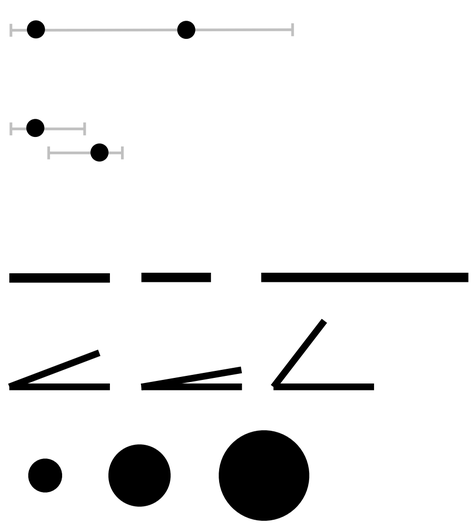
position*
common scale

position*
unaligned scale

length (1D)

angle/tilt

area (2D)



slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Ordinal & Quantitative

How much?

position*
common scale

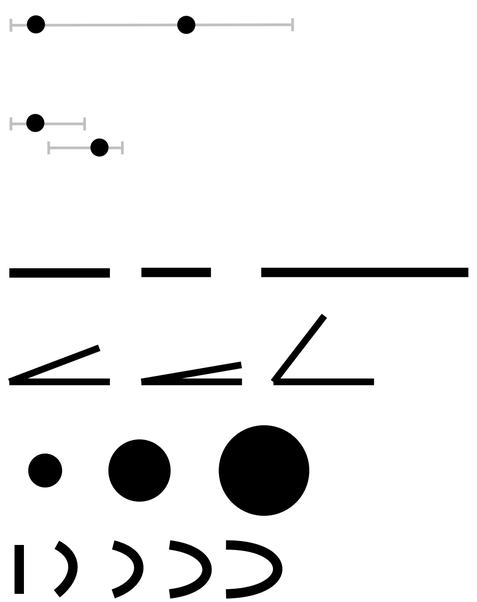
position*
unaligned scale

length (1D)

angle/tilt

area (2D)

curvature



slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Ordinal &

Quantitative

How much?

position*
common scale

position*
unaligned scale

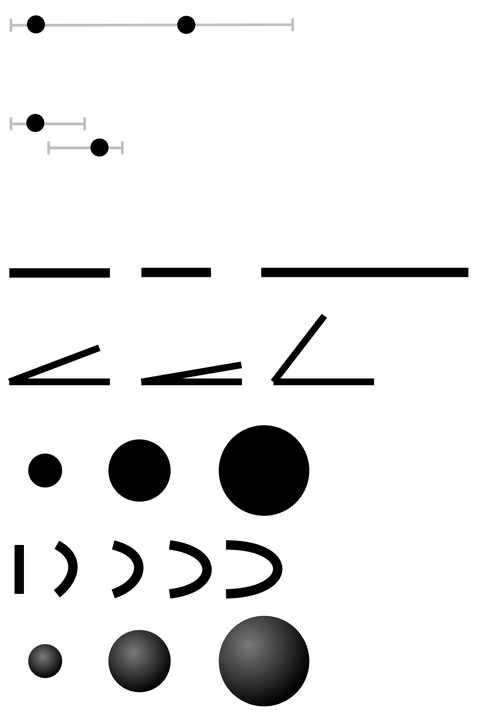
length (1D)

angle/tilt

area (2D)

curvature

volume (3D)



Visual Channels: Rankings

Ordinal &

Quantitative

How much?

position*
common scale

position*
unaligned scale

length (1D)

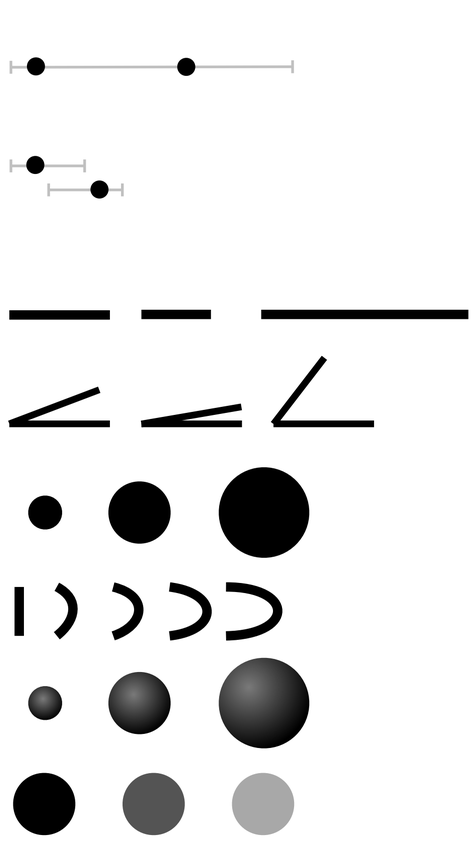
angle/tilt

area (2D)

curvature

volume (3D)

lightness



Visual Channels: Rankings

Ordinal &

Quantitative

How much?

position*
common scale

position*
unaligned scale

length (1D)

angle/tilt

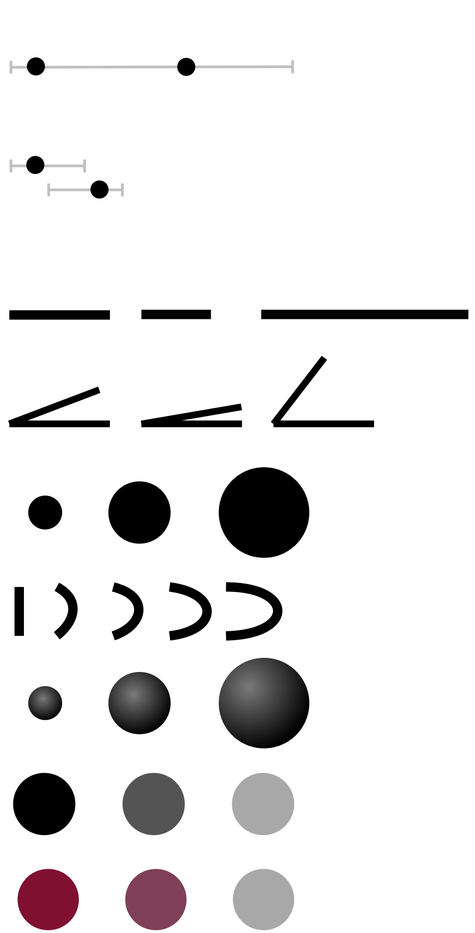
area (2D)

curvature

volume (3D)

lightness

color saturation



slide adapted from Munzner 2011, Visualization Principles

Visual Channels: Rankings

Ordinal &

Quantitative

How much?

position*
common scale

position*
unaligned scale

length (1D)

angle/tilt

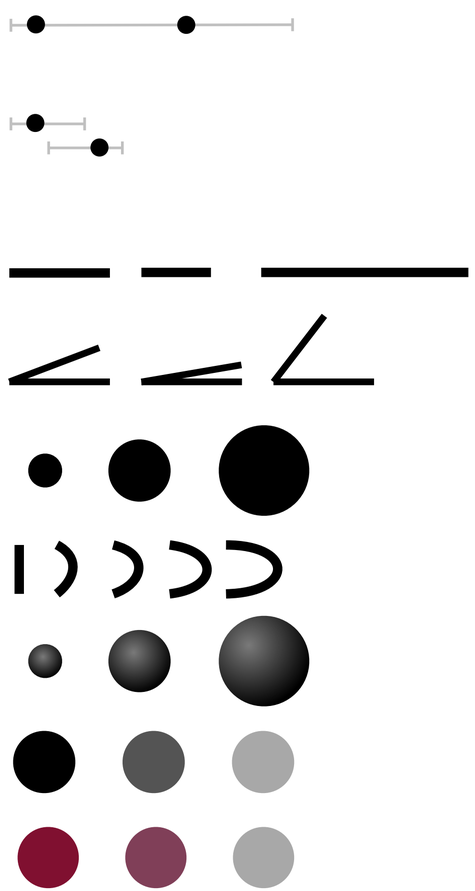
area (2D)

curvature

volume (3D)

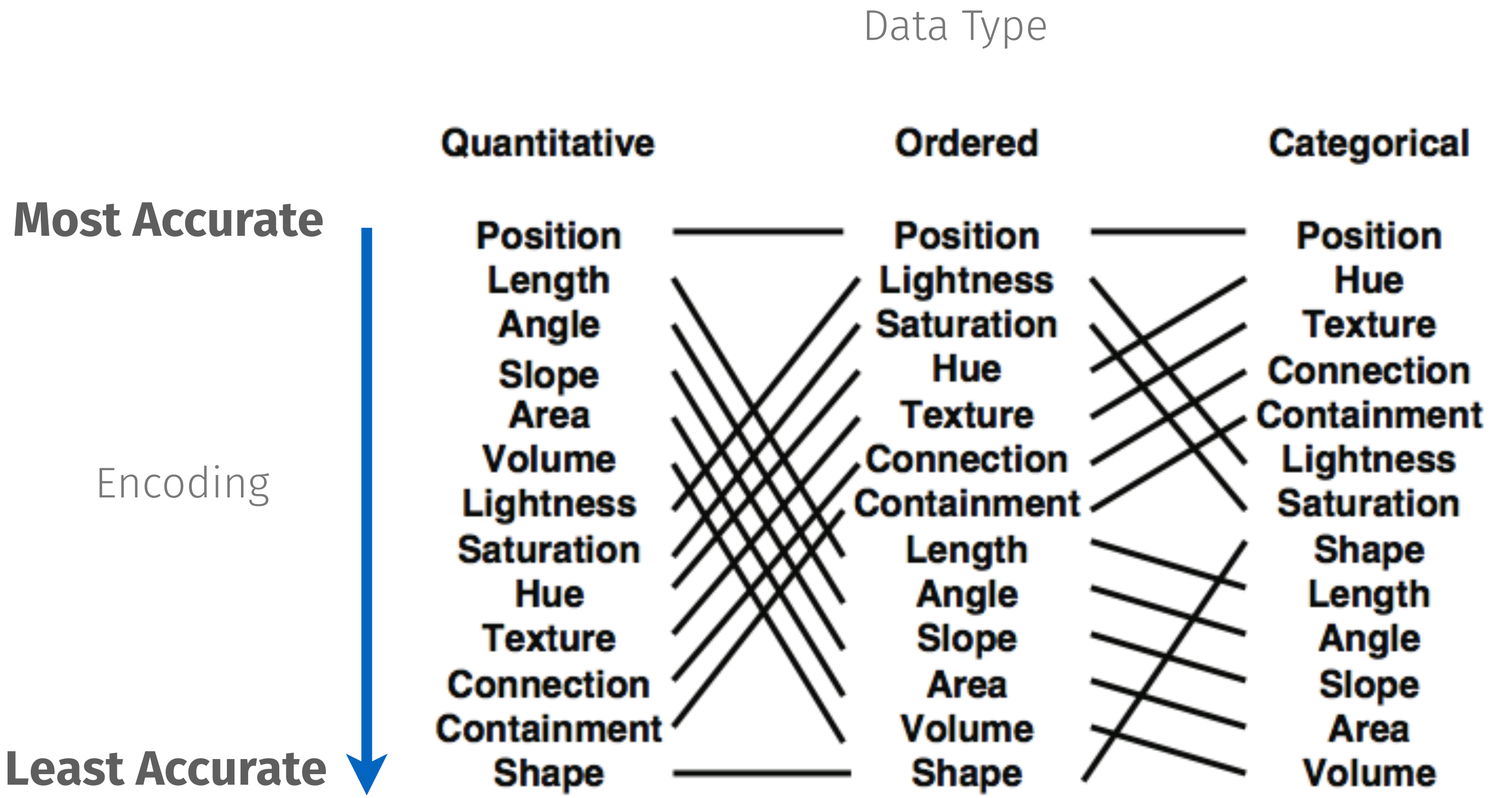
lightness

color saturation



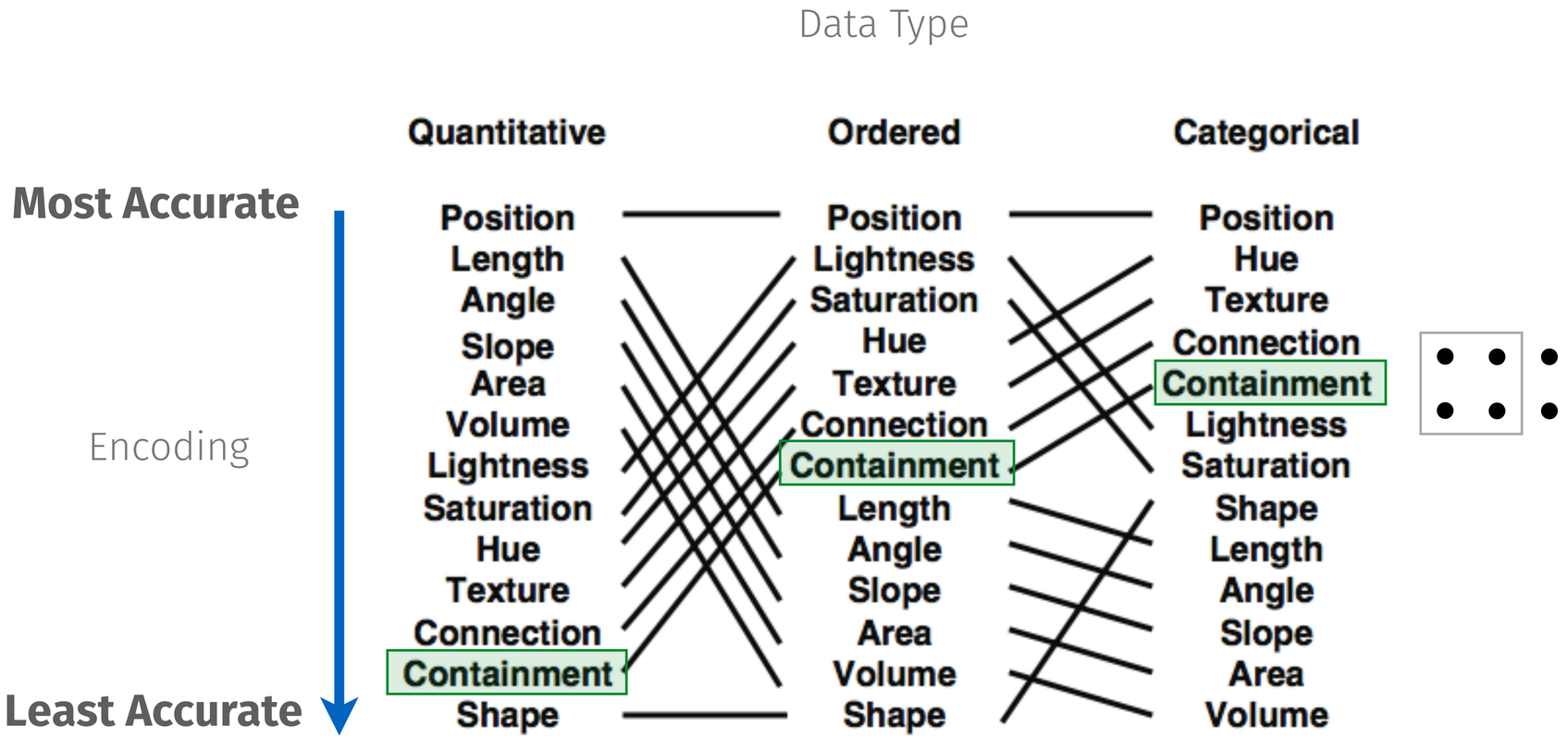
slide adapted from Munzner 2011, Visualization Principles

Ranking of Encodings



Munzner 2009, in Fundamentals of Computer Graphics (redrawn from Mackinlay 1986)

Ranking of Encodings



Munzner 2009, in Fundamentals of Computer Graphics (redrawn from Mackinlay 1986)

Back to our Design Critique …

Exercise: Design Critique

Euler Diagram

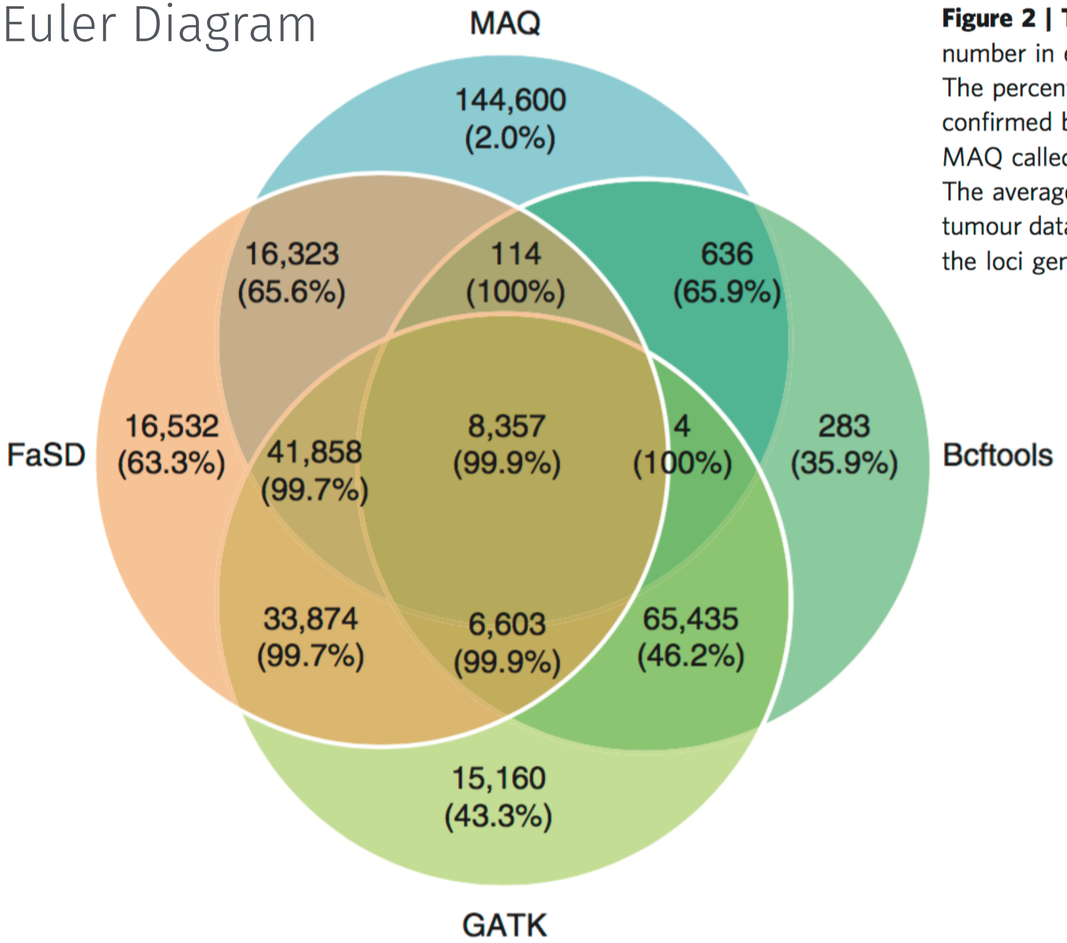


Figure 2 | The Venn diagram of SNPs detected by different tools. The number in each cell is the number of SNPs in the corresponding category. The percentage under the number is the proportion of SNPs that were confirmed by the Affymetrix SNP array. The FaSD, GATK, Bcftools and MAQ called 123661, 171291, 81432 and 211892 SNPs in total, respectively. The average depth of this data set was 10 ×. The figure is based on the tumour data set and Bowtie was used as aligner, and statistics are based on the loci genotyped by the Affymetrix SNP array.

F Xu et al., Nature Communications 3, 1258 doi:10.1038/ncomms2256

Exercise: Design Critique

Euler Diagram

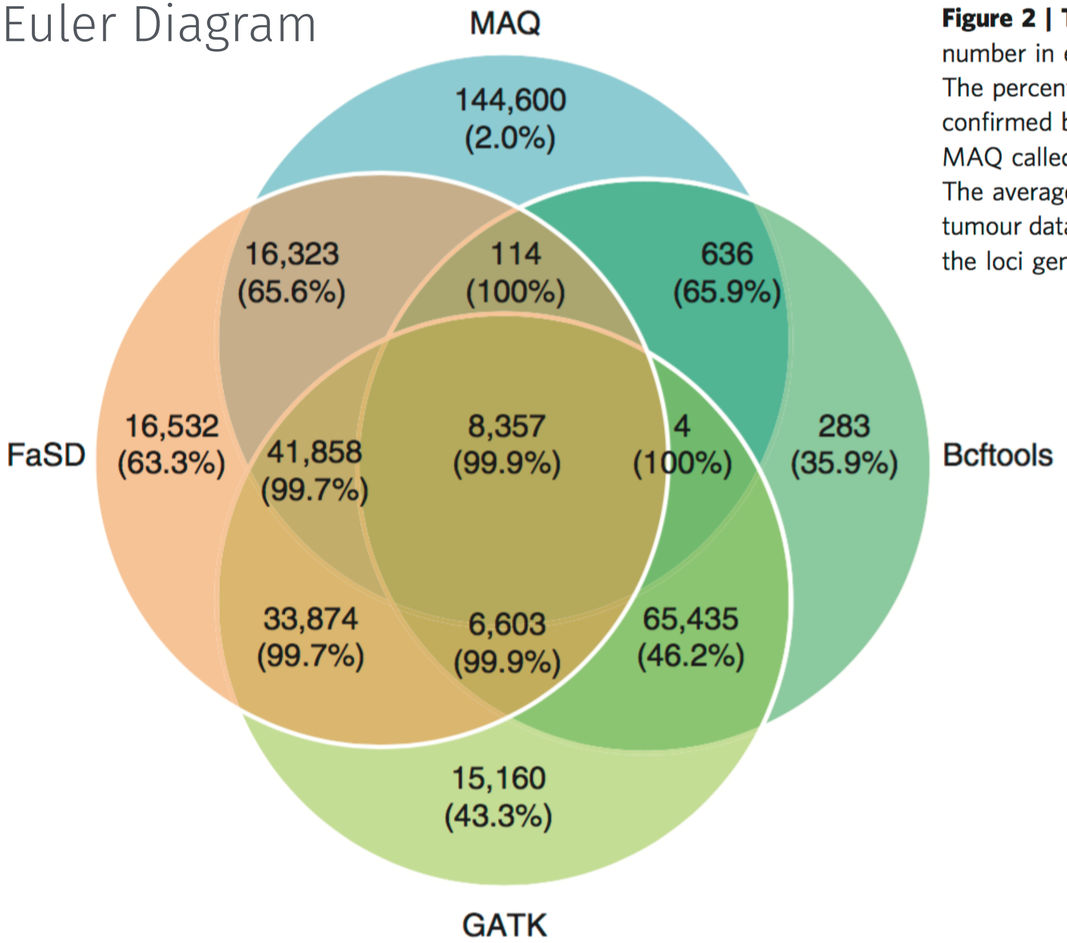
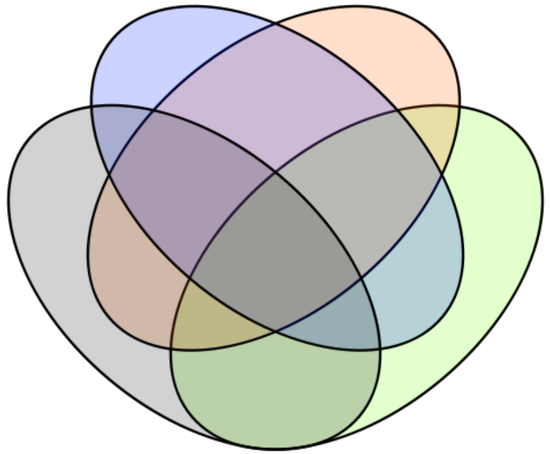


Figure 2 | The Venn diagram of SNPs detected by different tools. The number in each cell is the number of SNPs in the corresponding category. The percentage under the number is the proportion of SNPs that were confirmed by the Affymetrix SNP array. The FaSD, GATK, Bcftools and MAQ called 123661, 171291, 81432 and 211892 SNPs in total, respectively. The average depth of this data set was 10 ×. The figure is based on the tumour data set and Bowtie was used as aligner, and statistics are based on the loci genotyped by the Affymetrix SNP array.

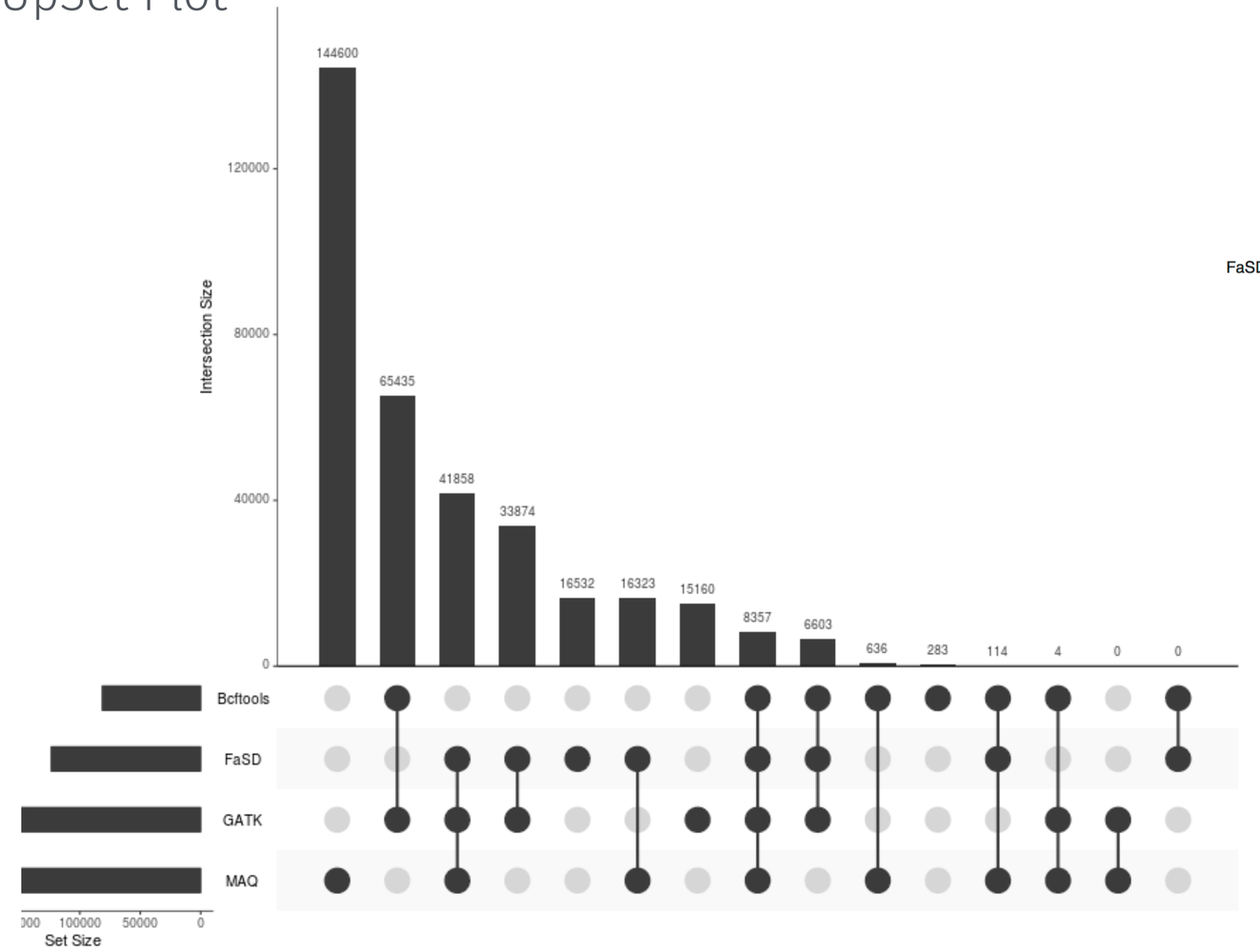
Venn Diagram



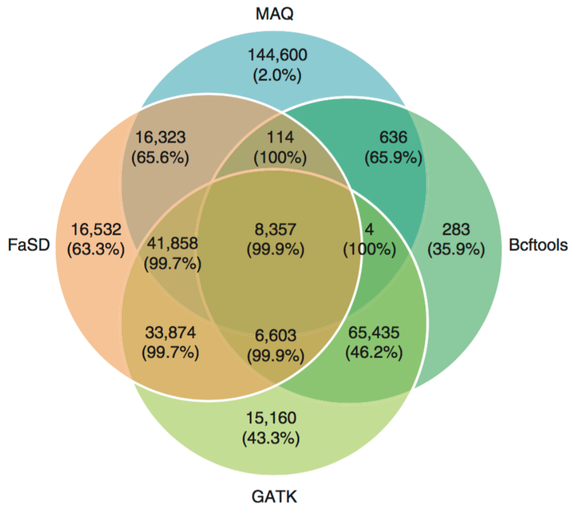
F Xu et al., Nature Communications 3, 1258 doi:10.1038/ncomms2256

Exercise: Design Critique

UpSet Plot



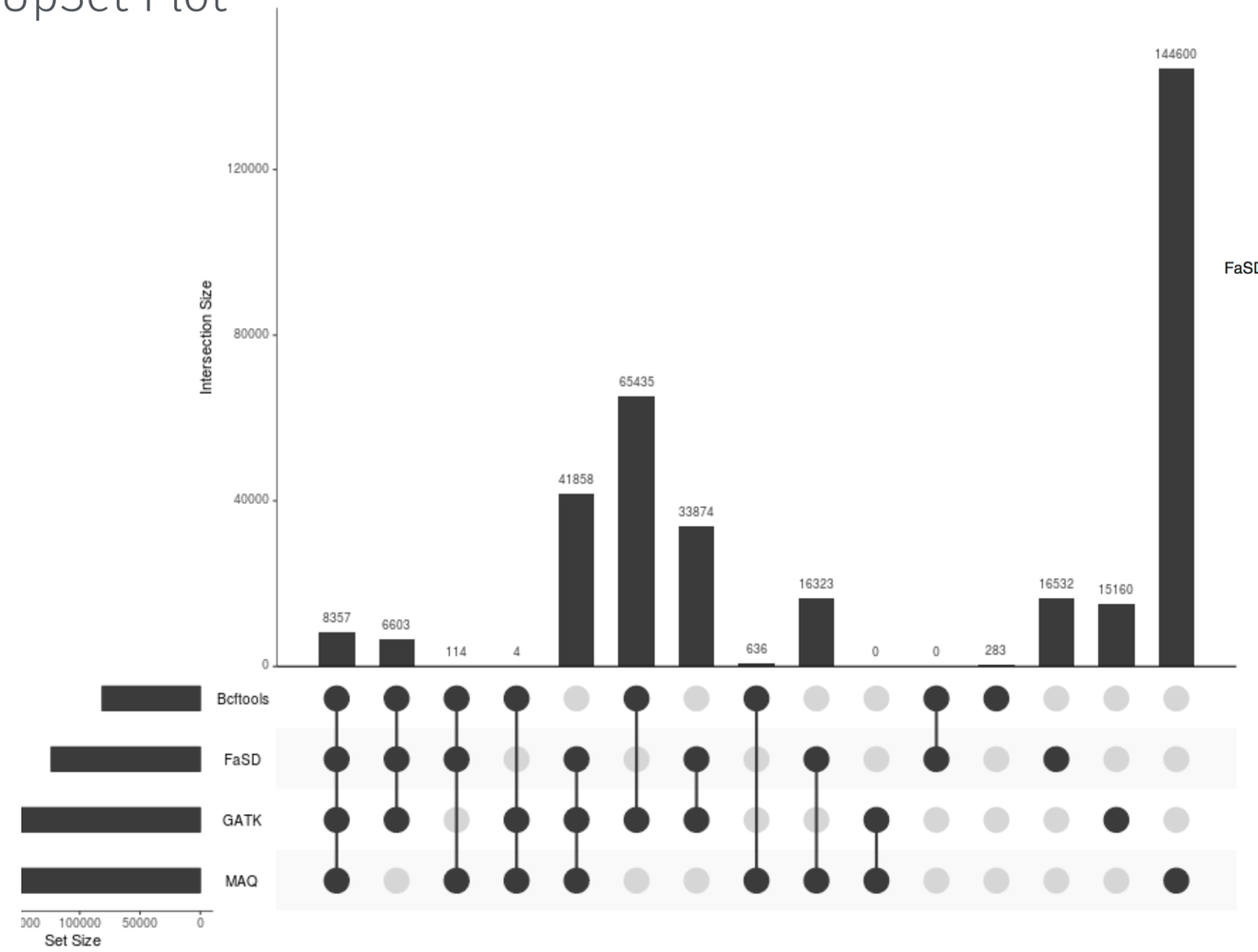
Euler Diagram



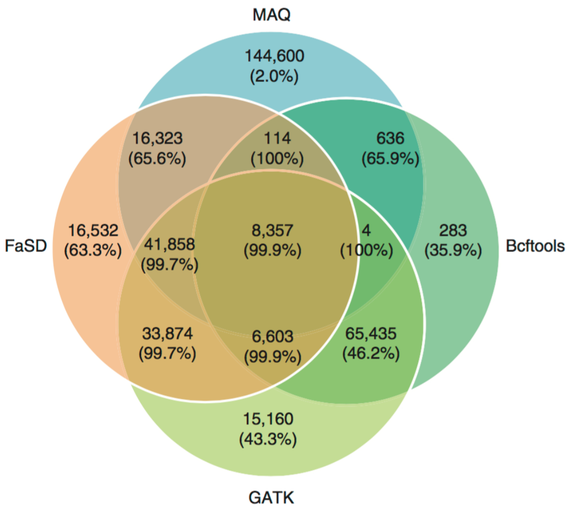
https://cran.rstudio.com/web/packages/UpSetR/ & https://gehlenborglab.shinyapps.io/upsetr/,

Exercise: Design Critique

UpSet Plot



Euler Diagram



https://cran.rstudio.com/web/packages/UpSetR/ & https://gehlenborglab.shinyapps.io/upsetr/,

Ranking of Encodings

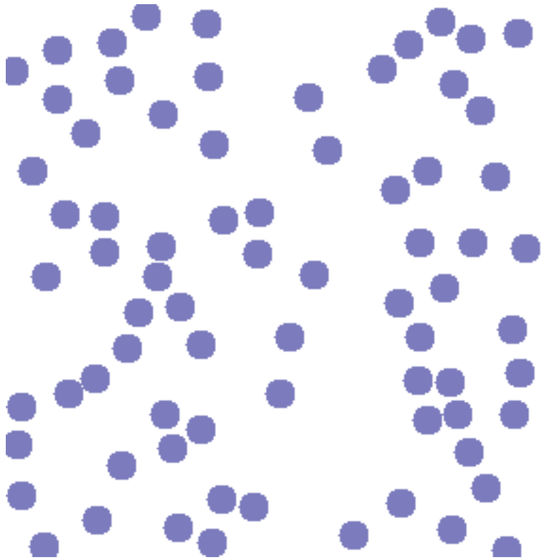
Principle of Importance Ordering (Mackinaly 1986):

Encode more important information more effectively.

- How accurately can the data be read from the visualization?
- How many classes can be distinguished?
- Can the channels be separated from each other?
- Which channels are processed preattentively?

Preattentive Processing: Color

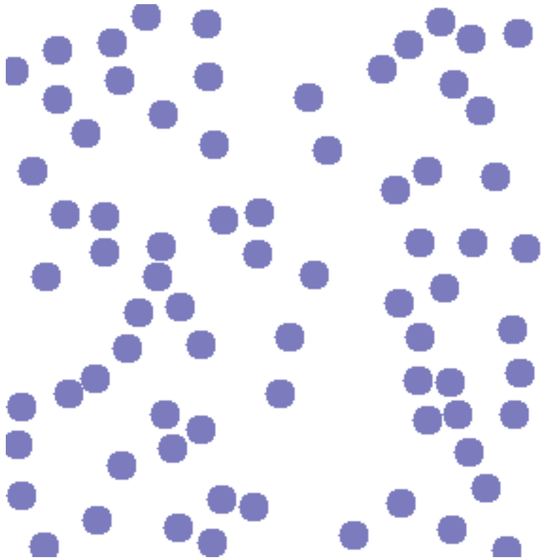
Can you spot this: ●?



Healey, http://www.csc.ncsu.edu/faculty/healey/PP/

Preattentive Processing: Color

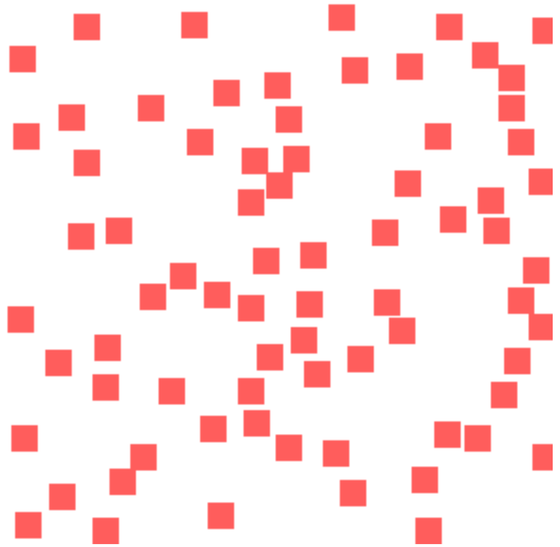
Can you spot this: ●?



Healey, http://www.csc.ncsu.edu/faculty/healey/PP/

Preattentive Processing: Shape

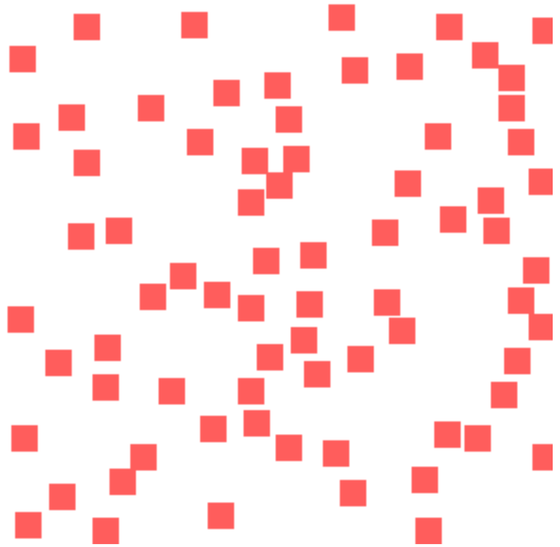
Can you spot this: ●?



Healey, http://www.csc.ncsu.edu/faculty/healey/PP/

Preattentive Processing: Shape

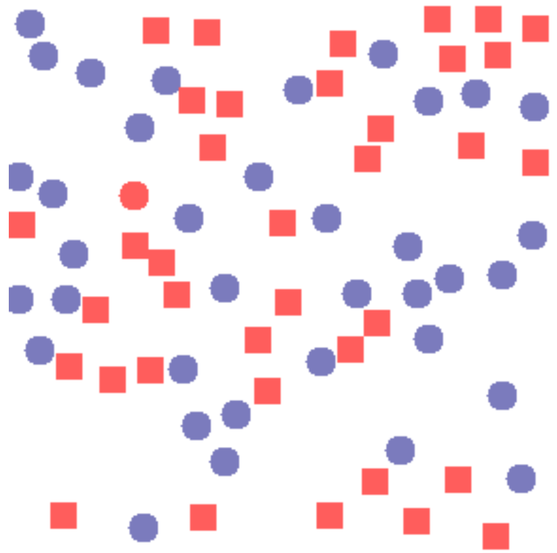
Can you spot this: ●?



Healey, http://www.csc.ncsu.edu/faculty/healey/PP/

Preattentive Processing: Shape & Color

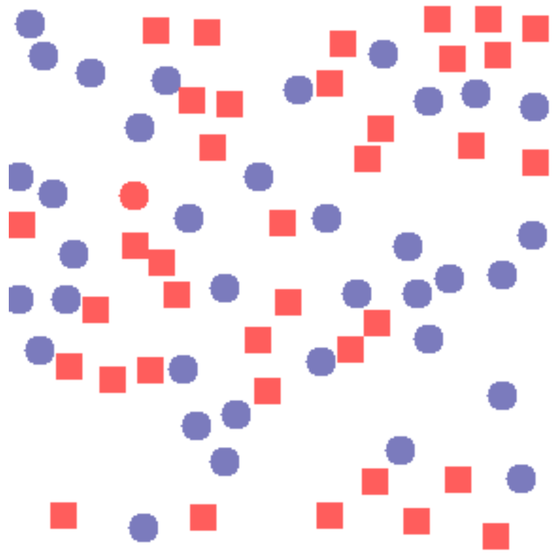
Can you spot this: ●?



Healey, http://www.csc.ncsu.edu/faculty/healey/PP/

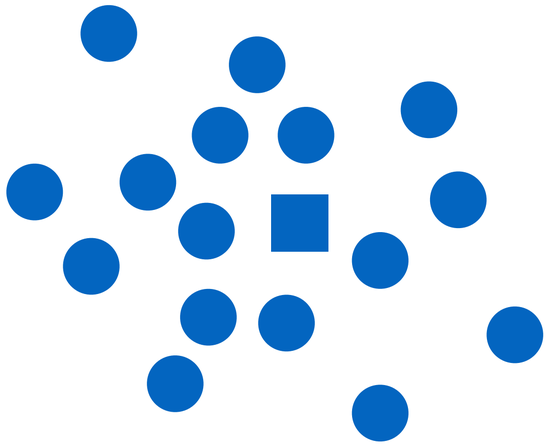
Preattentive Processing: Shape & Color

Can you spot this: ●?

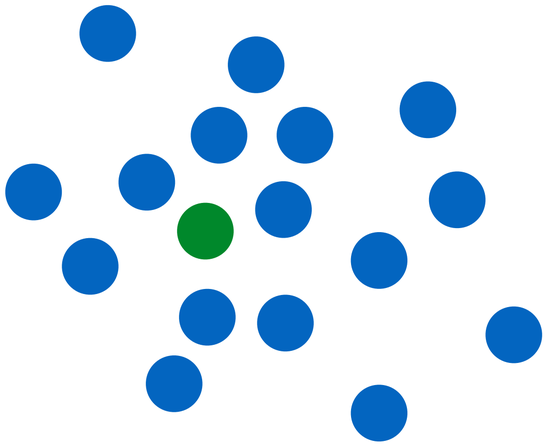


Healey, http://www.csc.ncsu.edu/faculty/healey/PP/

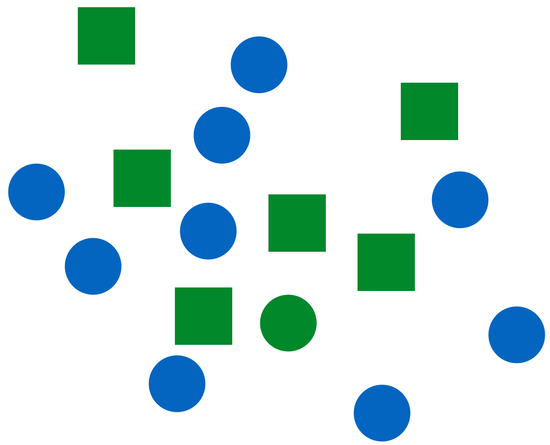
Preattentive Processing



Shape: Yes!



Color: Yes!



Shape + Color: No!

- visual properties that can be perceived in less than 250 ms
- no sequential scanning of the image required, unlike text or numbers
- examples for other visual properties that can be processed preattentively: orientation, curvature, direction of motion, size and others

after Healey et al. 1993, Healey and Enns 1999

Using Rankings

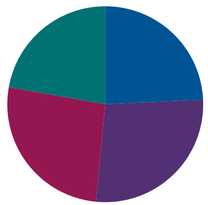
Year 1

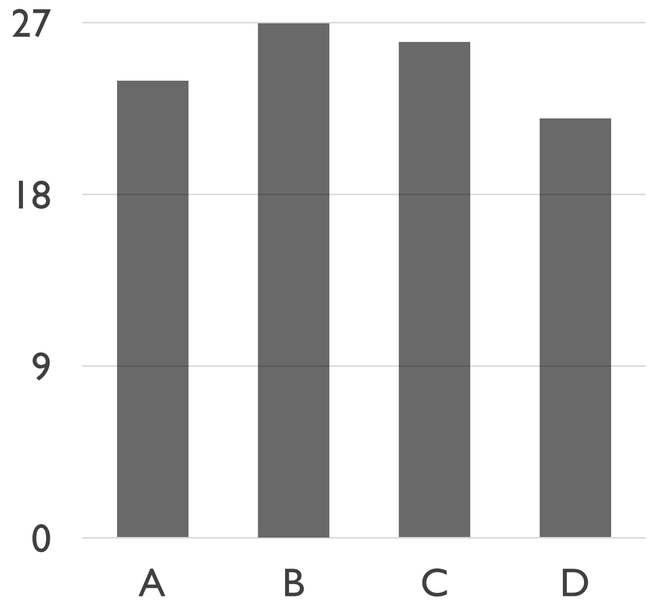
Year 2

A B C D

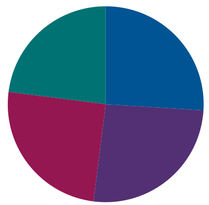
Using Rankings

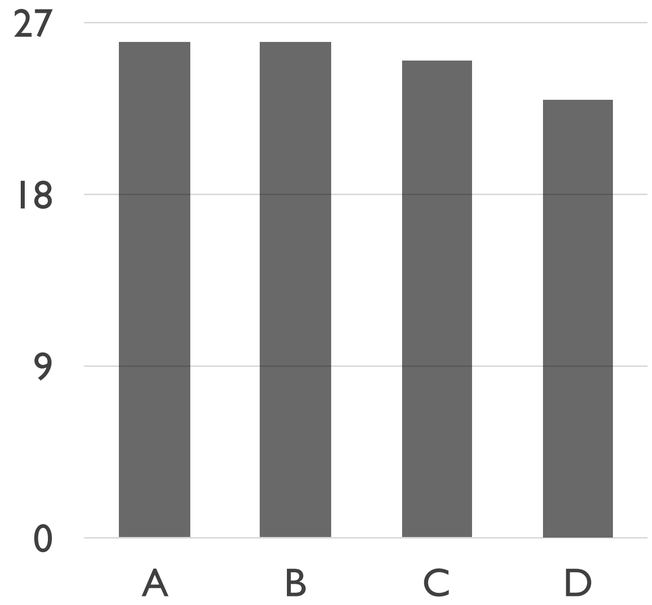
Year 1





Year 2





Bar Charts for Items & Categories

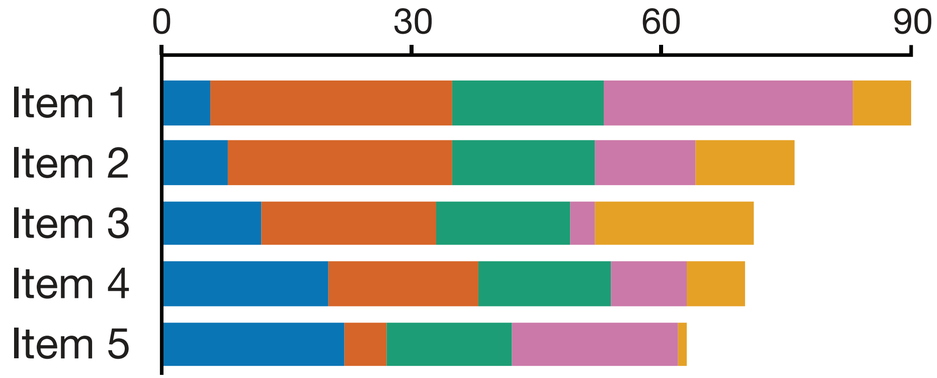
	A	B	C	D	E
Item 1	6	29	18	30	7
Item 2	8	27	17	12	12
Item 3	12	21	16	3	19
Item 4	20	18	16	9	7
Item 5	22	5	15	20	1

Streit and Gehlenborg, Nature Methods, 2014

Bar Charts for Items & Categories

	A	B	C	D	E
Item 1	6	29	18	30	7
Item 2	8	27	17	12	12
Item 3	12	21	16	3	19
Item 4	20	18	16	9	7
Item 5	22	5	15	20	1

Stacked Bar Chart

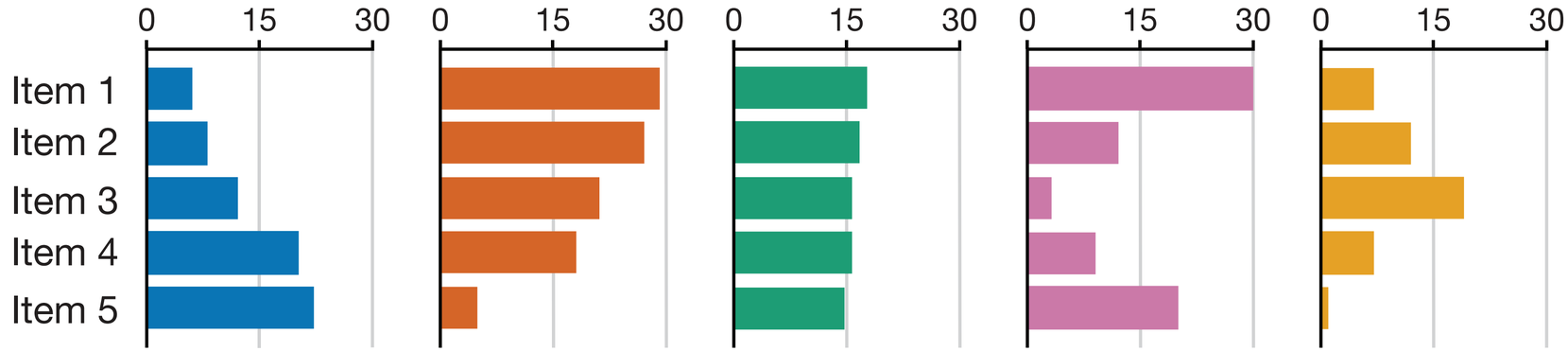


Streit and Gehlenborg, Nature Methods, 2014

Bar Charts for Items & Categories

	A	B	C	D	E
Item 1	6	29	18	30	7
Item 2	8	27	17	12	12
Item 3	12	21	16	3	19
Item 4	20	18	16	9	7
Item 5	22	5	15	20	1

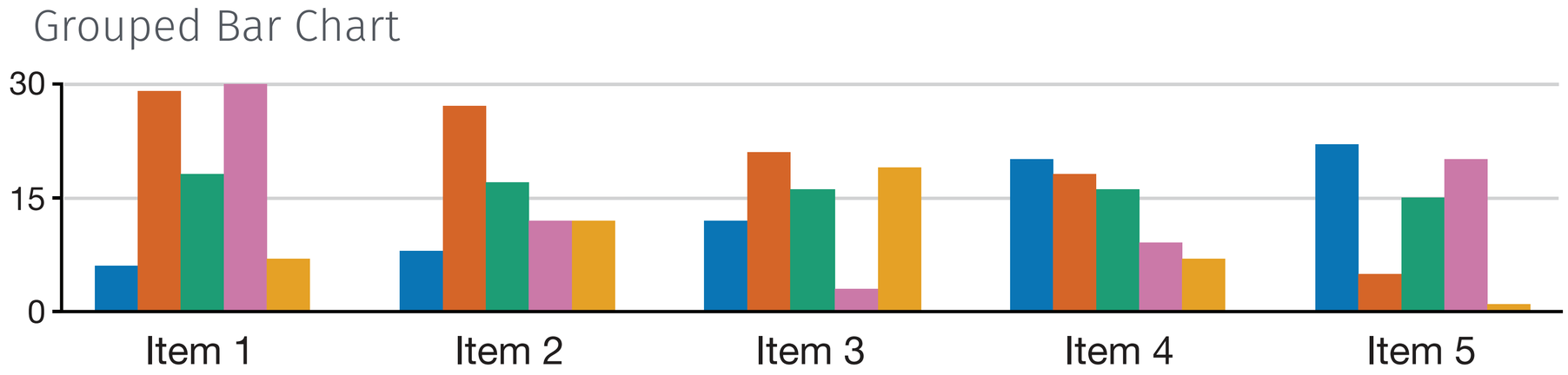
Layered Bar Chart



Streit and Gehlenborg, Nature Methods, 2014

Bar Charts for Items & Categories

	A	B	C	D	E
Item 1	6	29	18	30	7
Item 2	8	27	17	12	12
Item 3	12	21	16	3	19
Item 4	20	18	16	9	7
Item 5	22	5	15	20	1



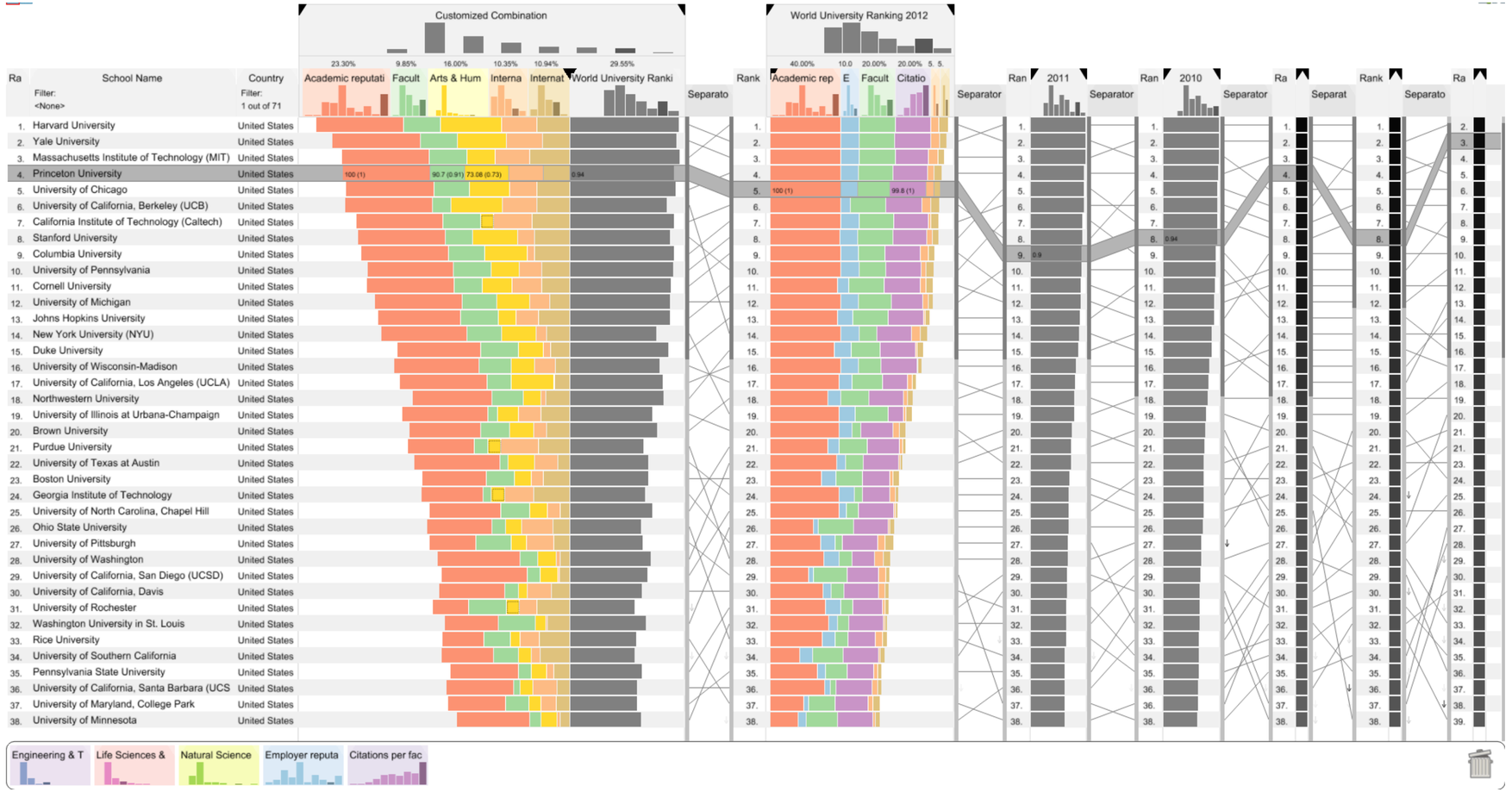
Streit and Gehlenborg, Nature Methods, 2014

Bar Charts for Items & Categories

- **Stacked Bar Chart**
  - if focus is on comparing the overall quantities across items but also need to illustrate contributions of each category to the total
- **Layered Bar Chart**
  - if focus is on distribution of values in each category across all items
  - comparisons within each category are more accurate than in stacked bar charts due to common baseline for the values in each category
- **Grouped Bar Chart**
  - if focus is on comparison of values across categories within each item while still enabling comparisons across items
  - if quantities add up to the same total for each item, then a grouped bar chart is equivalent to multiple pie charts, yet a grouped bar chart affords more accurate readings of values and comparisons

Streit and Gehlenborg, Nature Methods, 2014

LineUp: Ranking Visualization



http://lineup.caleydo.org, Gratzl, Lex, Gehlenborg, Pfister and Streit, IEEE Trans. Comput. Graph. Vis. (InfoVis 2013), 2013

Color Pitfalls: Rainbow Color Map

hard to order

after Borland & Taylor 2007, IEEE CG&A

Color Pitfalls: Rainbow Color Map

hard to order

easy to order

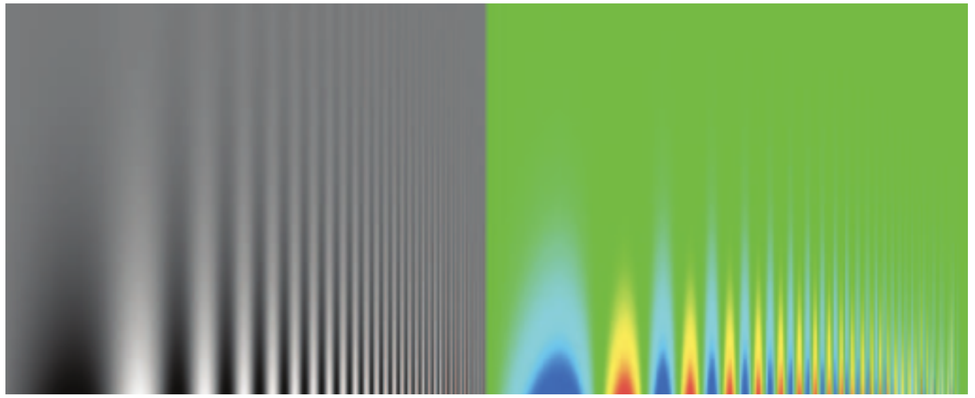
after Borland & Taylor 2007, IEEE CG&A

Color Pitfalls: Rainbow Color Map

hard to order

easy to order

lower resolution



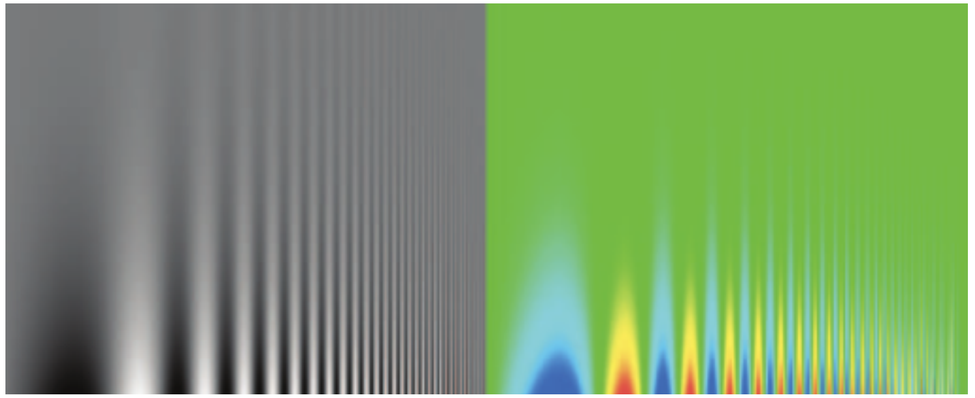
after Borland & Taylor 2007, IEEE CG&A

Color Pitfalls: Rainbow Color Map

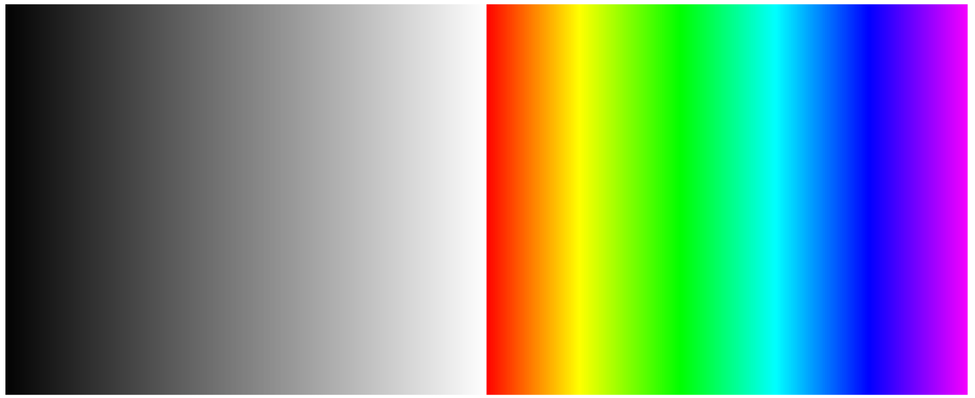
hard to order

easy to order

lower resolution



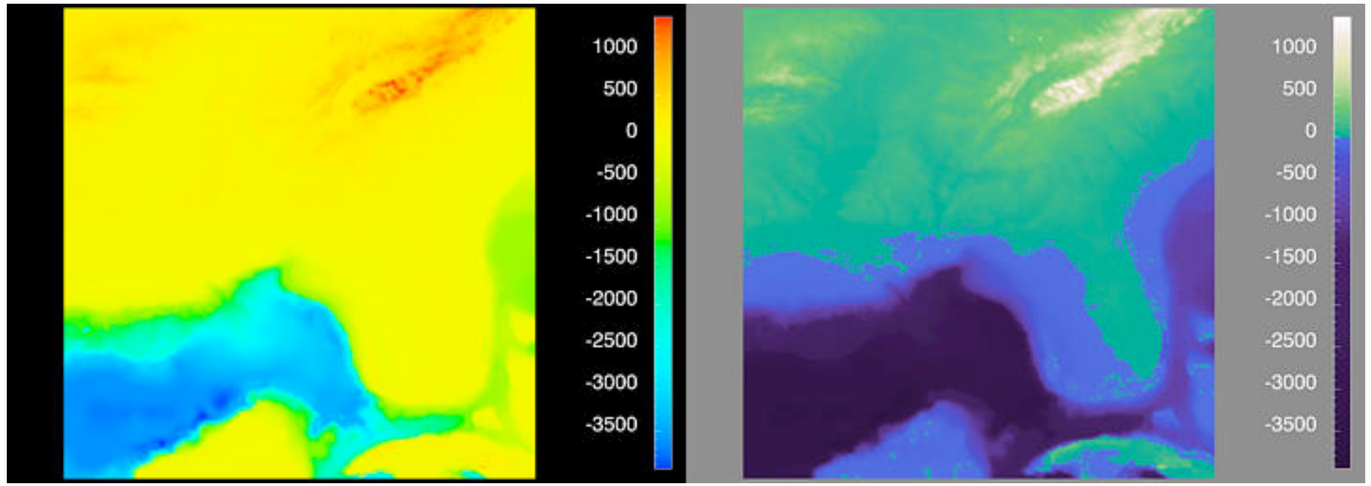
creates artifacts



after Borland & Taylor 2007, IEEE CG&A

Color Pitfalls: Rainbow Color Map

Southeastern United States and Gulf of Mexico



Problems

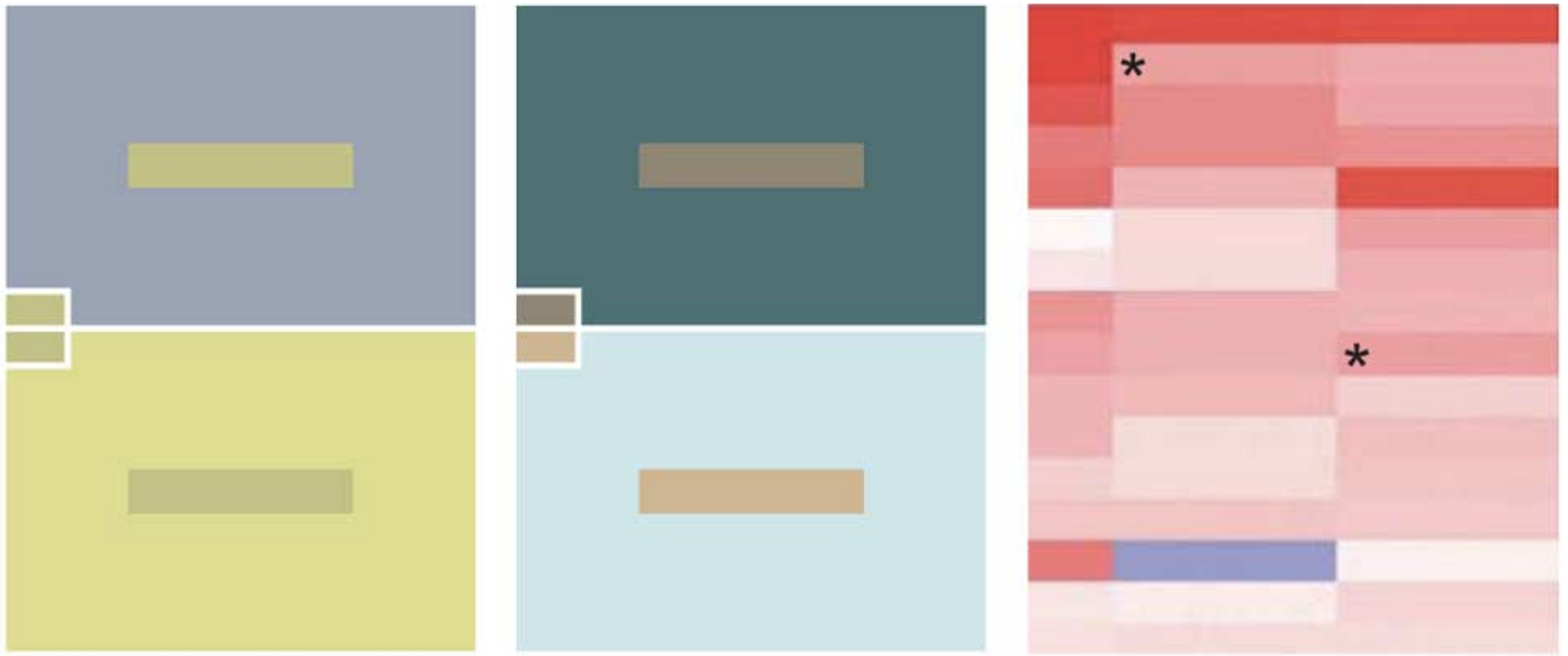
Zero crossing not explicit.

Lack of ordering of colors makes it hard to interpret the map.

Rogowitz & Treinish 1996, http://researchweb.watson.ibm.com/people/l/lloydt/color/color.HTM

Color Pitfalls: Relativity

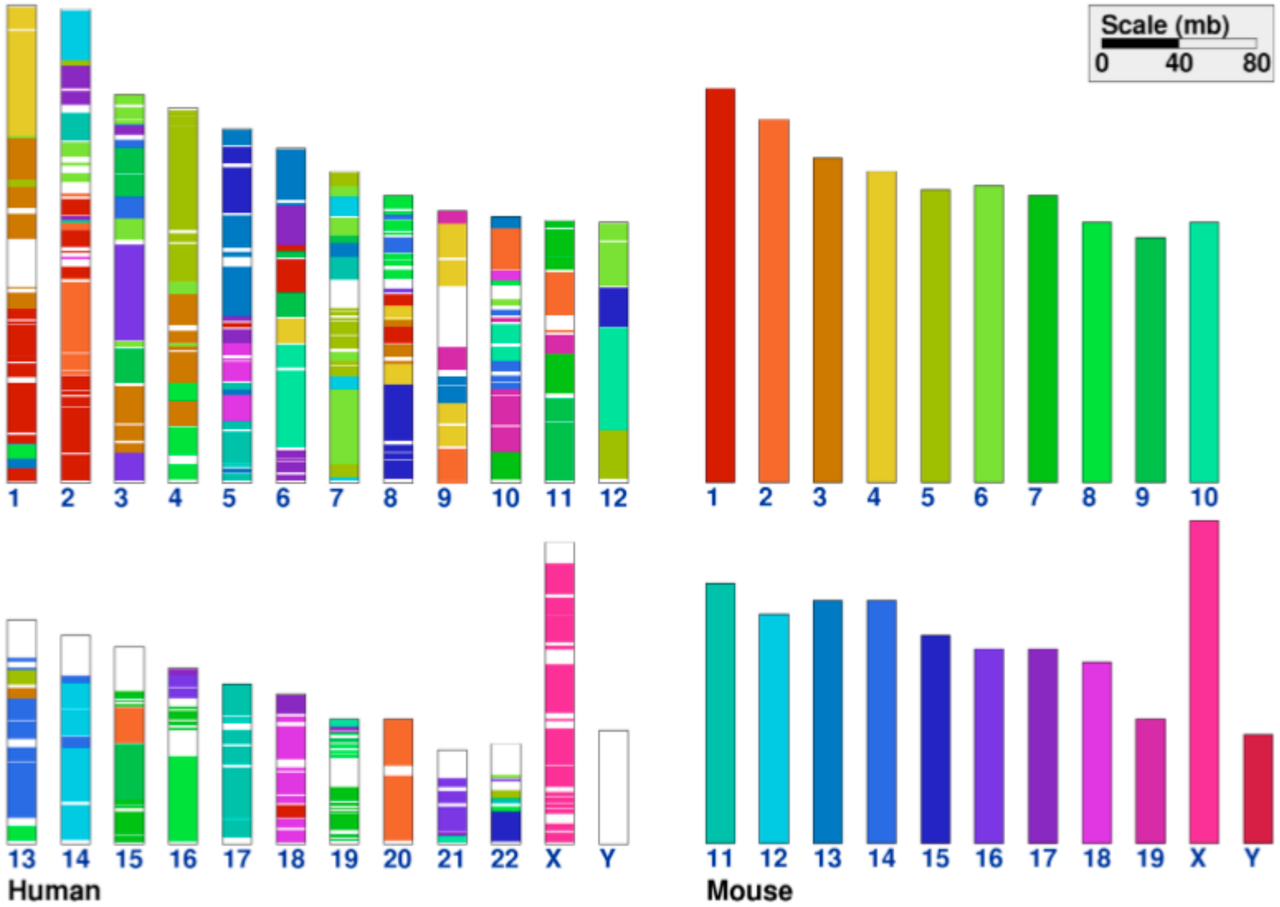
Color is a relative medium and context matters



Wong 2010, Nature Methods

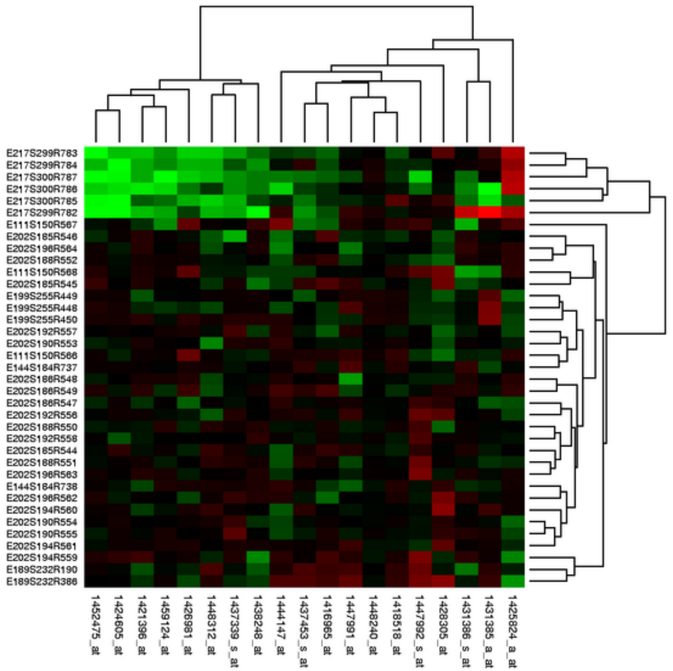
Color Pitfalls: Discriminability

Only 6-12 colors are visually discernible!

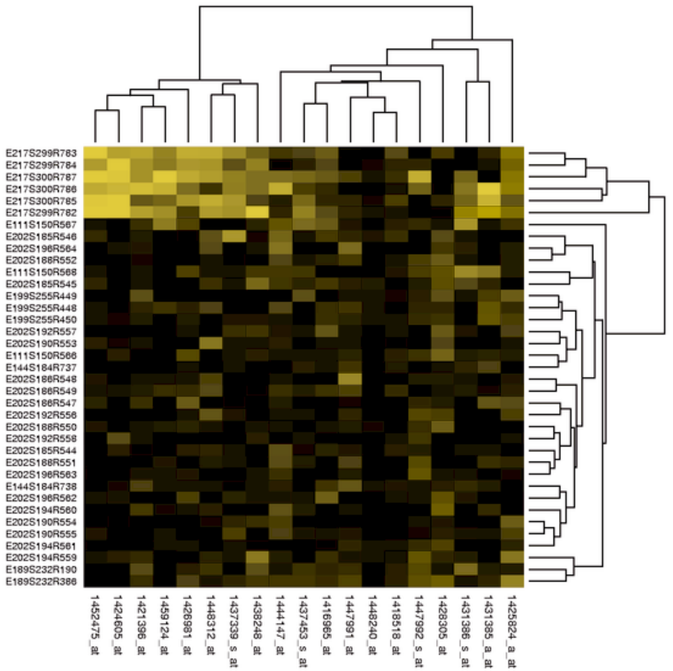


Sinha and Meller, Bioinformatics 2007

Color Pitfalls: Color Blindness



Normal Vision



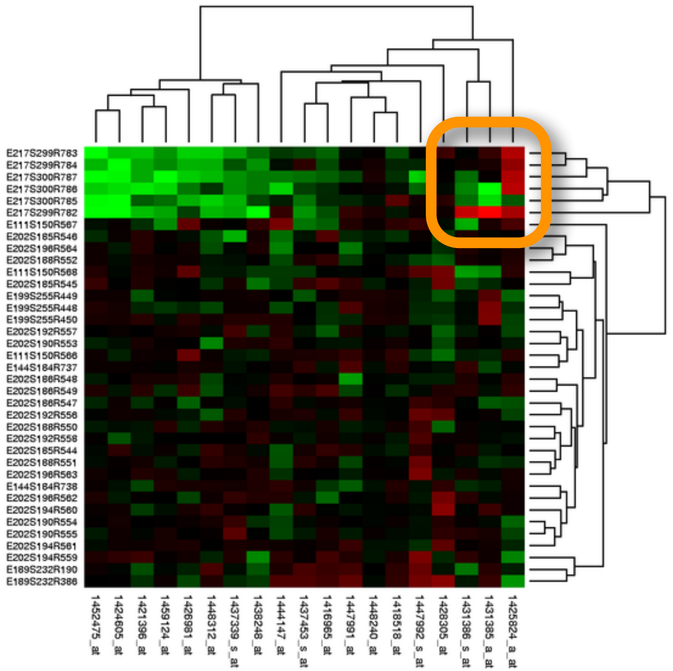
Deuteranope Vision

(“Red-Green Blindness”)

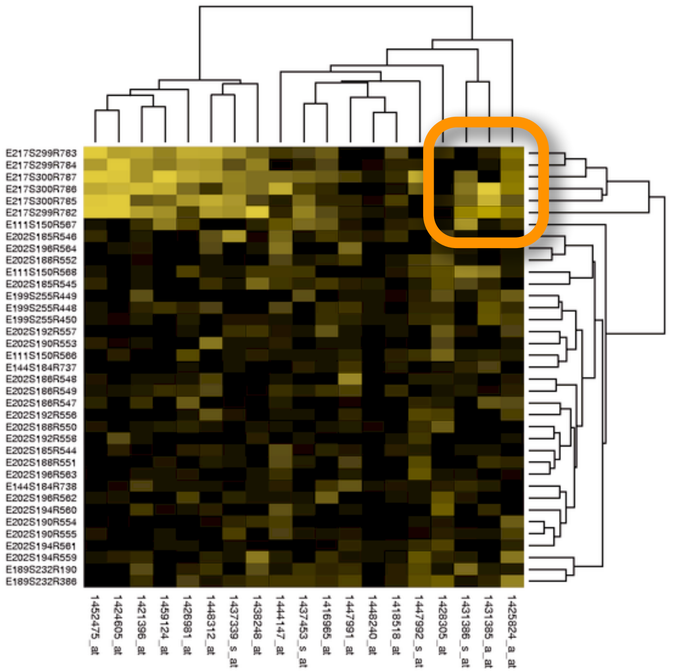
~ 7% of male population affected

estimate: Howard Hughes Medical Institute, http://www.hhmi.org/senses/b130.html

Color Pitfalls: Color Blindness



Normal Vision



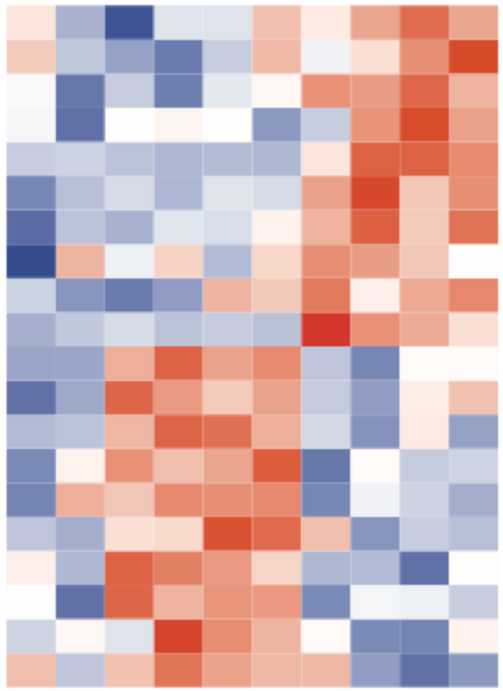
Deuteranope Vision

(“Red-Green Blindness”)

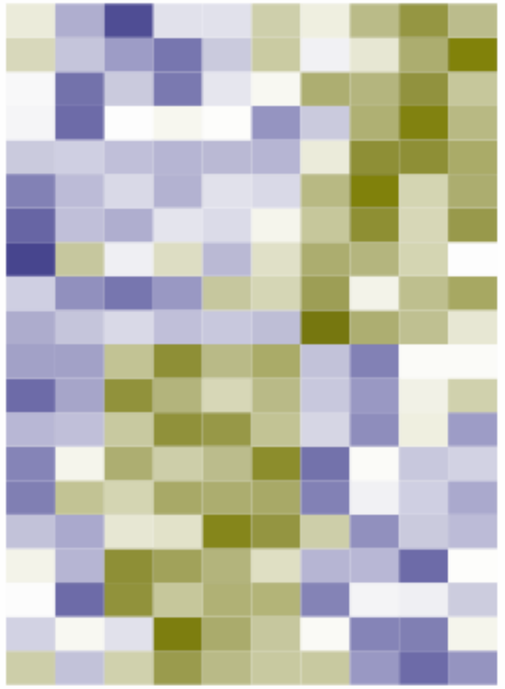
~ 7% of male population affected

estimate: Howard Hughes Medical Institute, http://www.hhmi.org/senses/b130.html

Color Pitfalls: Color Blindness



Normal Vision



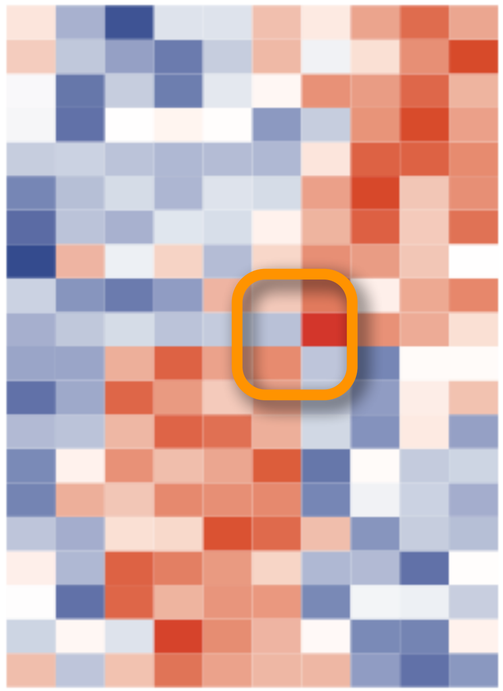
Deuteranope Vision

(“Red-Green Blindness”)

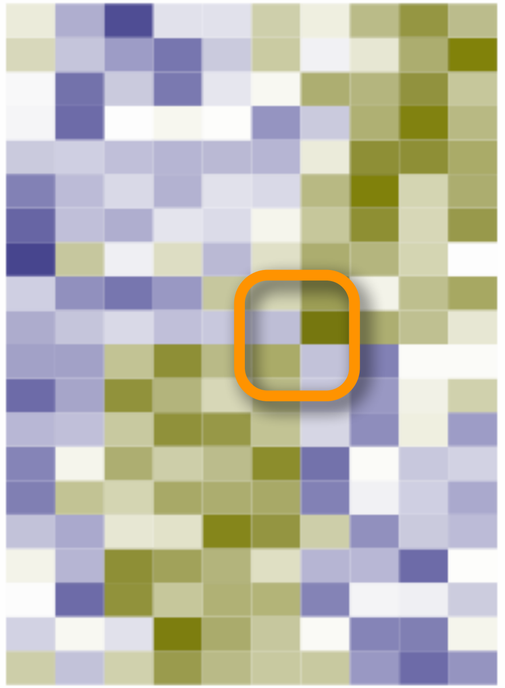
~ 7% of male population affected

estimate: Howard Hughes Medical Institute, http://www.hhmi.org/senses/b130.html

Color Pitfalls: Color Blindness



Normal Vision



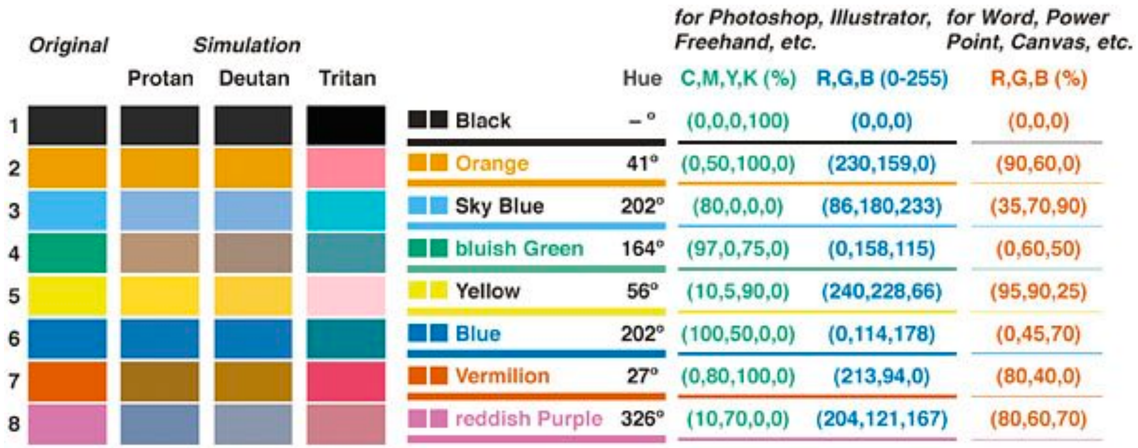
Deuteranope Vision

(“Red-Green Blindness”)

~ 7% of male population affected

estimate: Howard Hughes Medical Institute, http://www.hhmi.org/senses/b130.html

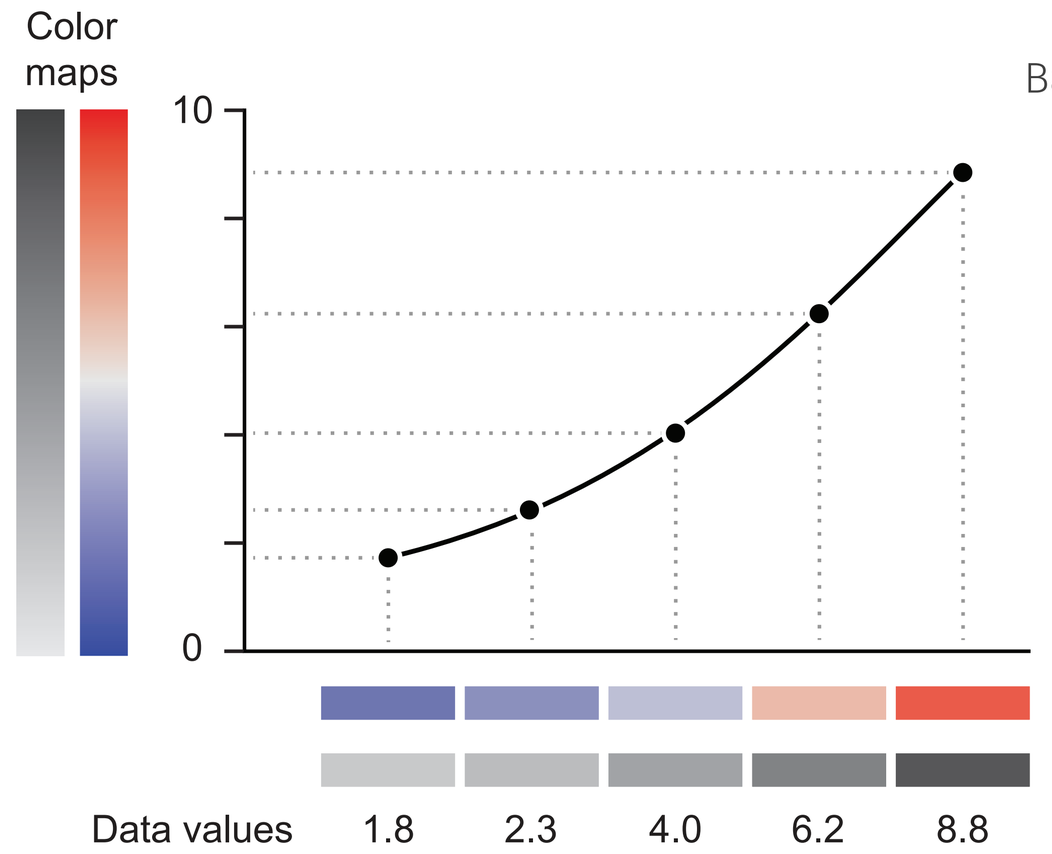
Hint: Color Blind Palette



	Original	Protan	Deutan	Tritan		Hue	for Photoshop, Illustrator, Freehand, etc. C,M,Y,K (%)	for Photoshop, Illustrator, Freehand, etc. R,G,B (0-255)	for Word, Power Point, Canvas, etc. R,G,B (%)
1					Black	– °	(0,0,0,100)	(0,0,0)	(0,0,0)
2					Orange	41°	(0,50,100,0)	(230,159,0)	(90,60,0)
3					Sky Blue	202°	(80,0,0,0)	(86,180,233)	(35,70,90)
4					bluish Green	164°	(97,0,75,0)	(0,158,115)	(0,60,50)
5					Yellow	56°	(10,5,90,0)	(240,228,66)	(95,90,25)
6					Blue	202°	(100,50,0,0)	(0,114,178)	(0,45,70)
7					Vermilion	27°	(0,80,100,0)	(213,94,0)	(80,40,0)
8					reddish Purple	326°	(10,70,0,0)	(204,121,167)	(80,60,70)

Okabe & Ito, http://jfly.iam.u-tokyo.ac.jp/color/

Color Pitfalls: Color Mapping

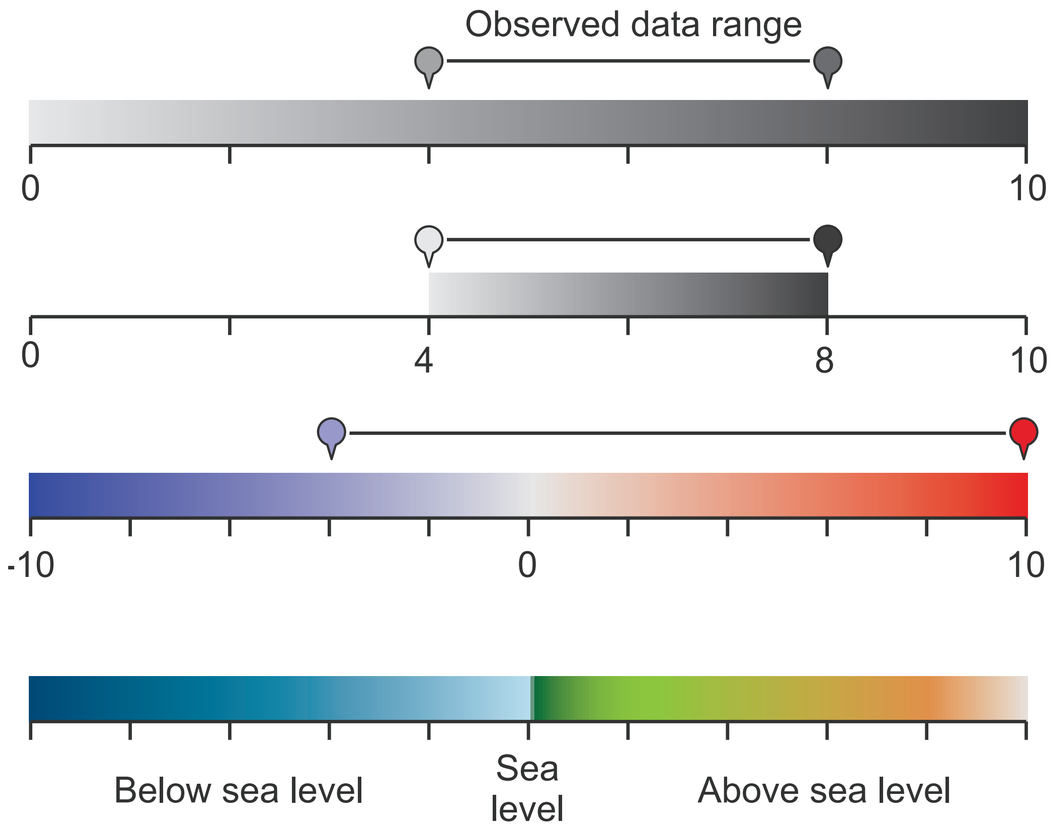


Bad color mapping!

Gehlenborg and Wong, Nature Methods, 2012

Color Pitfalls: Color Mapping

Good color mapping!



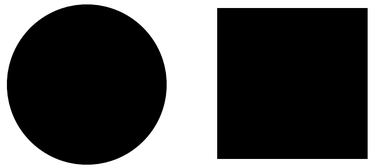
Gehlenborg and Wong, Nature Methods, 2012

Remember!

Color used poorly is worse than no color at all.

—*Edward Tufte*

Encoding Pitfalls: Interference



shape and size

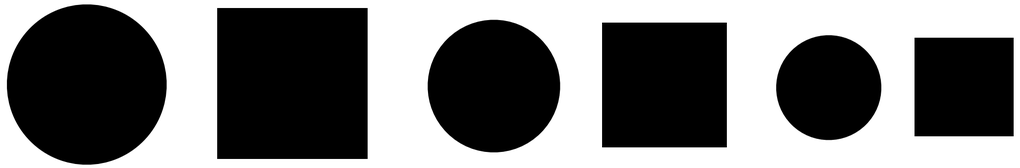
after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

shape and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference



shape and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

shape and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

shape and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

shape and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

shape and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

shape and size

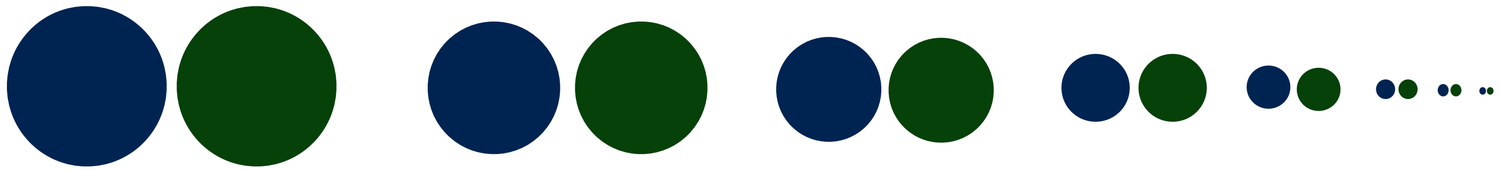
after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

Encoding Pitfalls: Interference

color and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

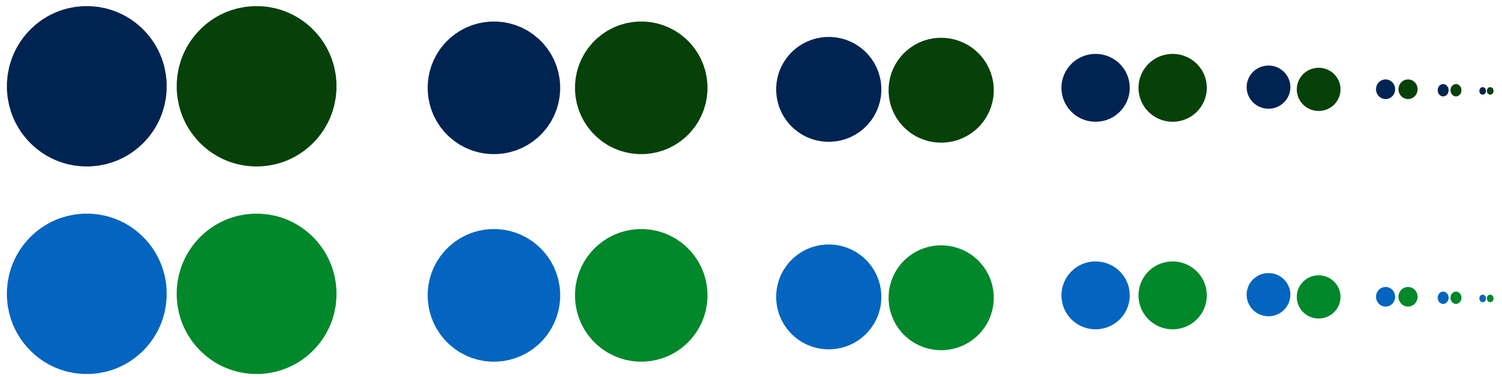
Encoding Pitfalls: Interference



color and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

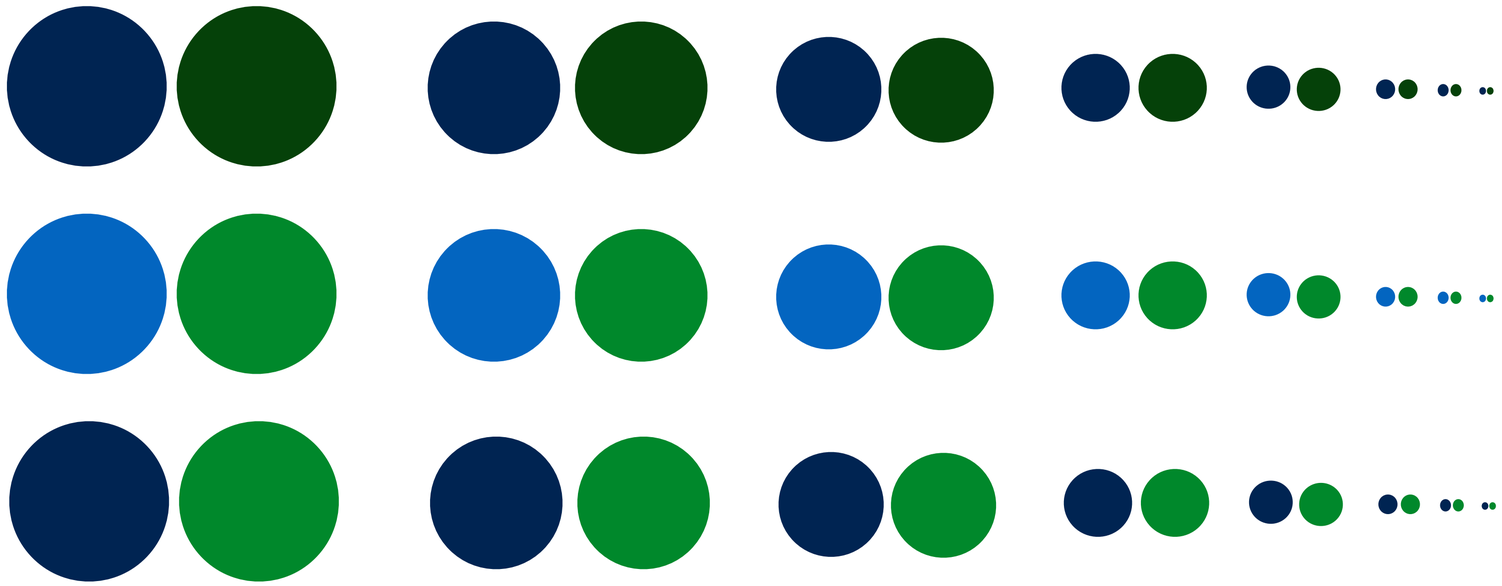
Encoding Pitfalls: Interference



color and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

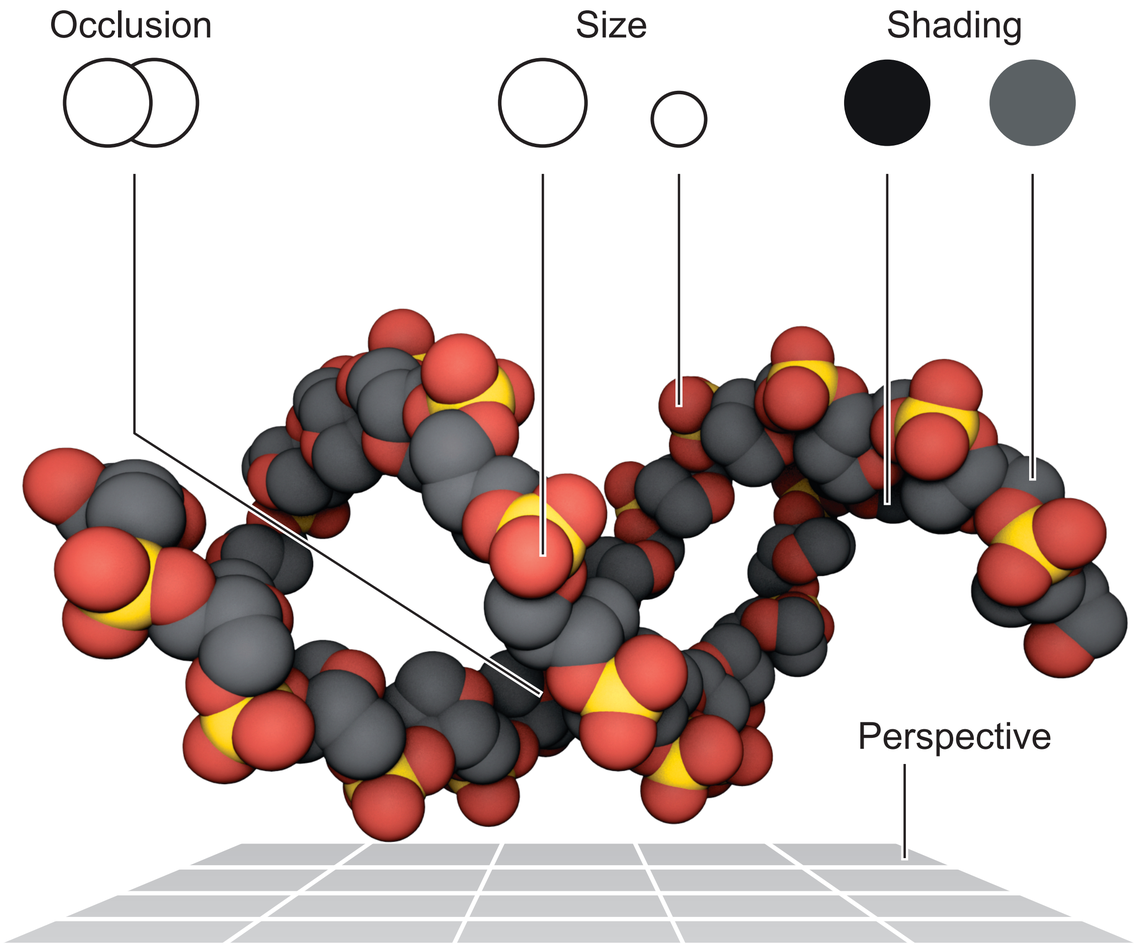
Encoding Pitfalls: Interference



color and size

after Mackinlay, ACM Transactions on Graphics 5, 110-141, 1986

3D: Depth Cues

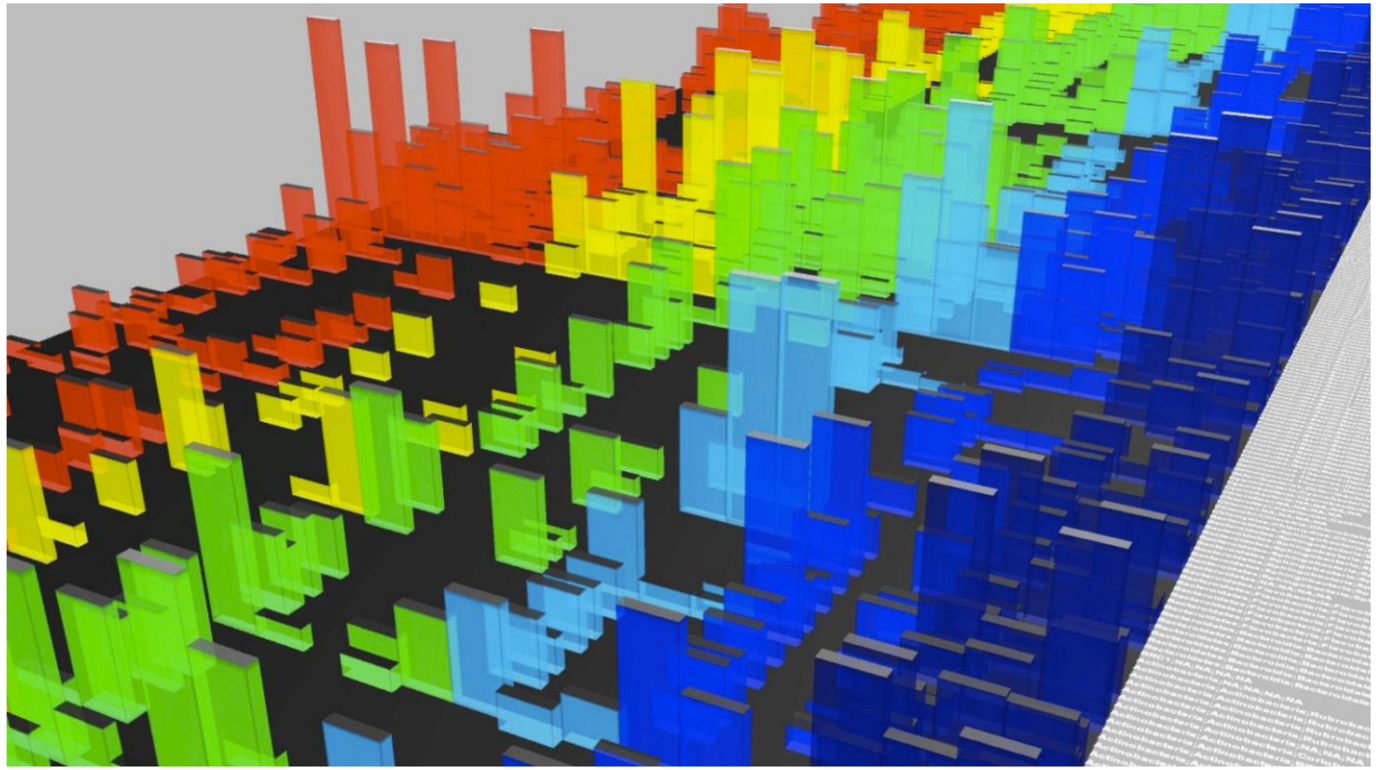


Gehlenborg and Wong, Nature Methods, 2012

3D Pitfalls: Perspective

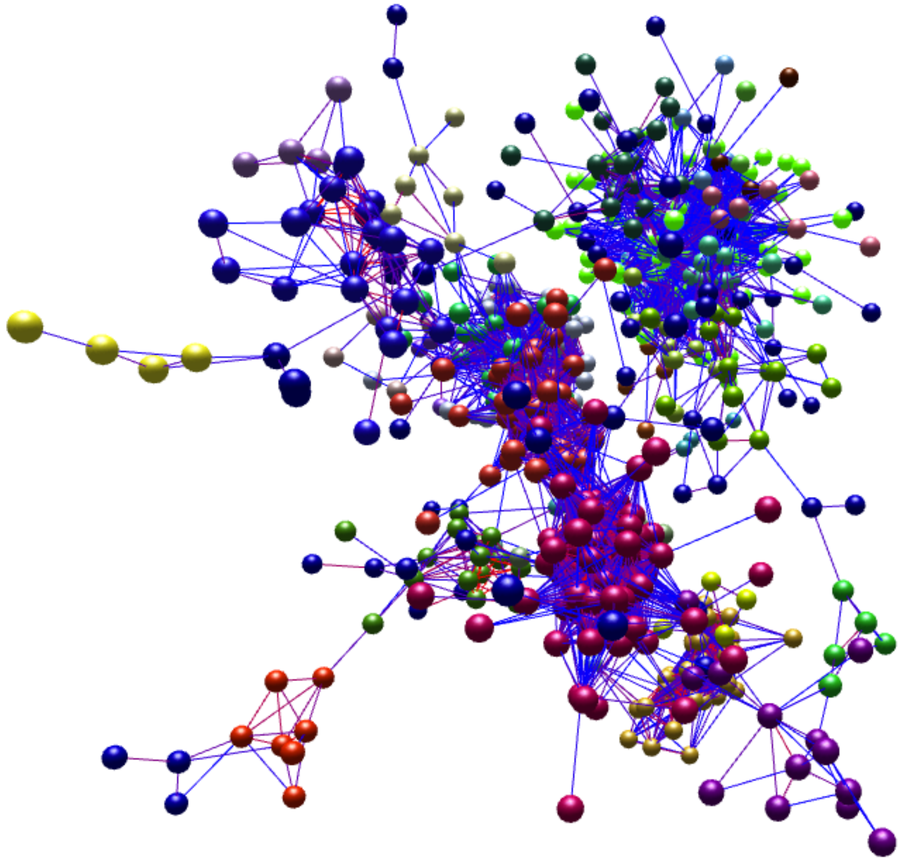
Perspective distortion: interferes with size channel encoding

Shading: interferes with color, lightness, and saturation channel encodings



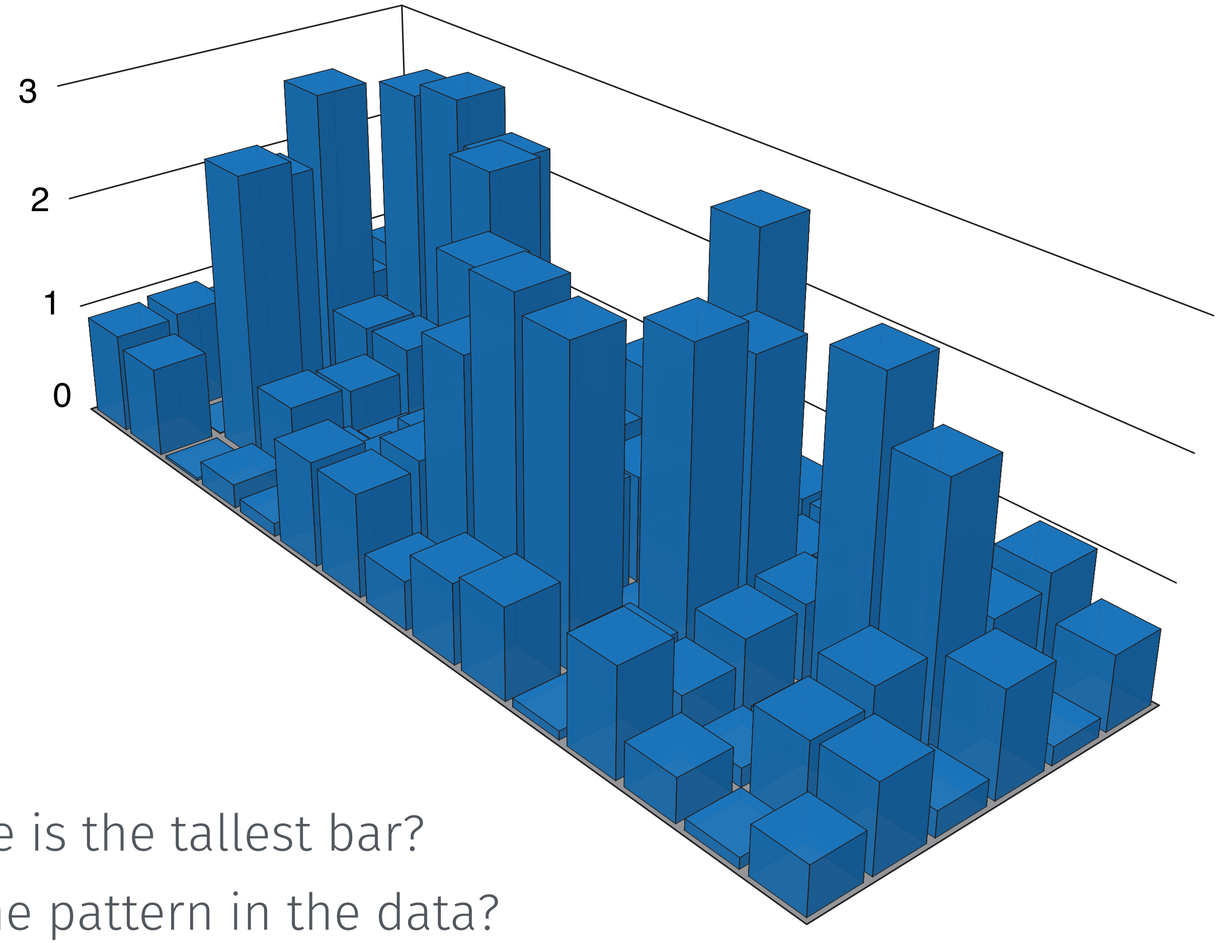
Moore et al. 2011, Proceedings of Pacific Symposium on Biocomputing 2011

3D Pitfalls: Occlusion



BioLayout 3D 2.0 sample dataset, http://www.biolayout.org

3D Pitfalls: Occlusion & Perspective

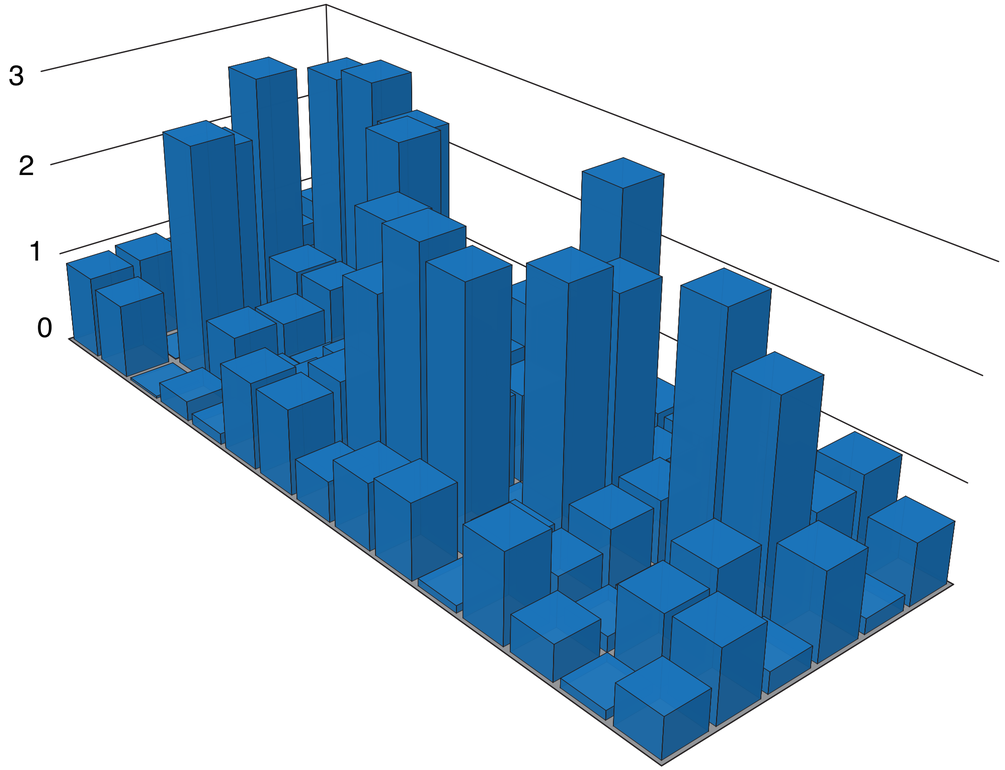


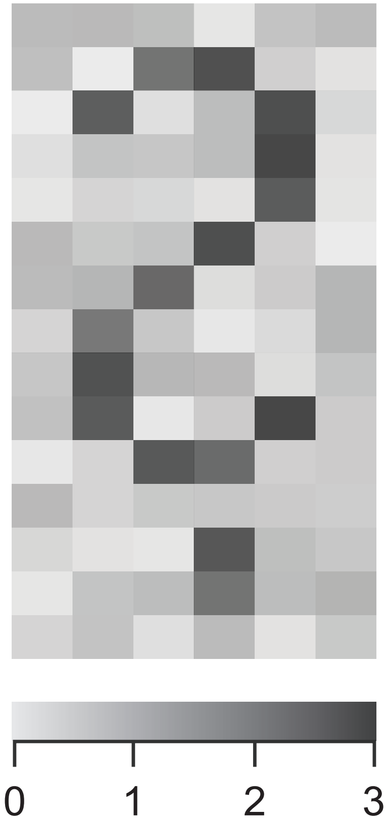
Which one is the tallest bar?

What is the pattern in the data?

Gehlenborg and Wong, Nature Methods, 2012

3D Pitfalls: Occlusion & Perspective





Which one is the tallest bar?

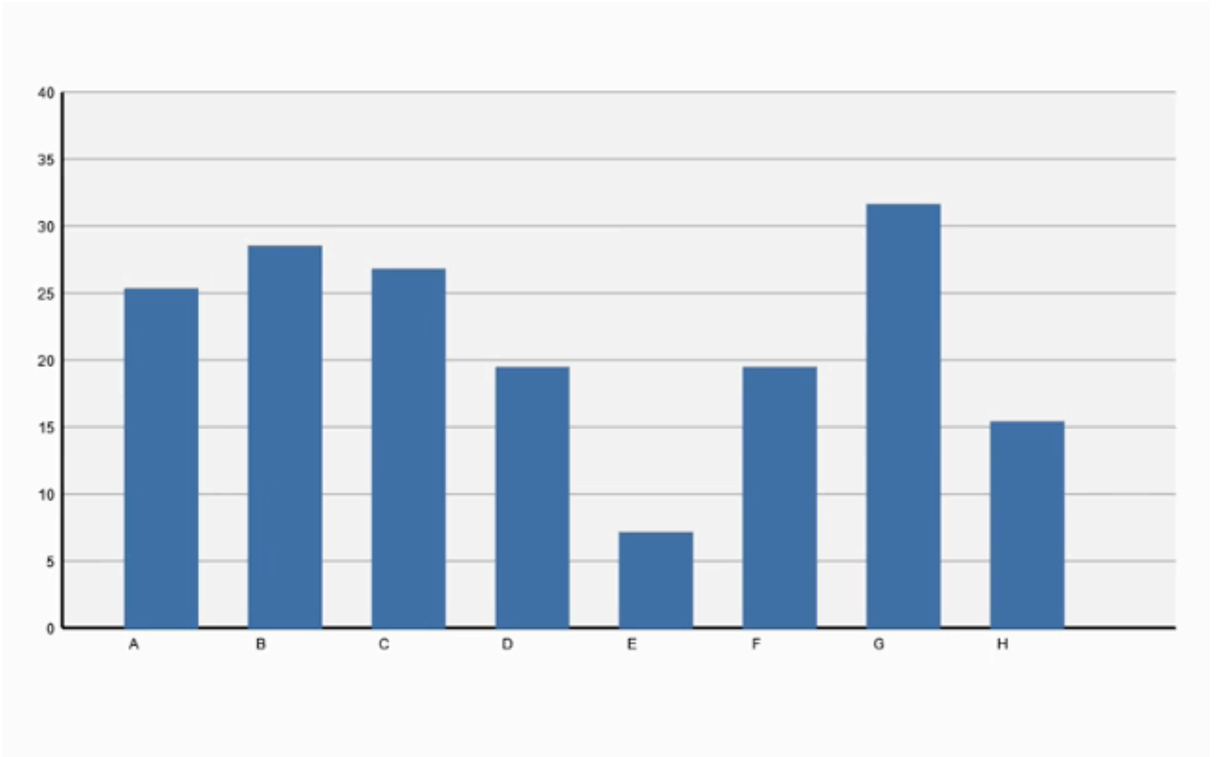
What is the pattern in the data?

Gehlenborg and Wong, Nature Methods, 2012

Animation

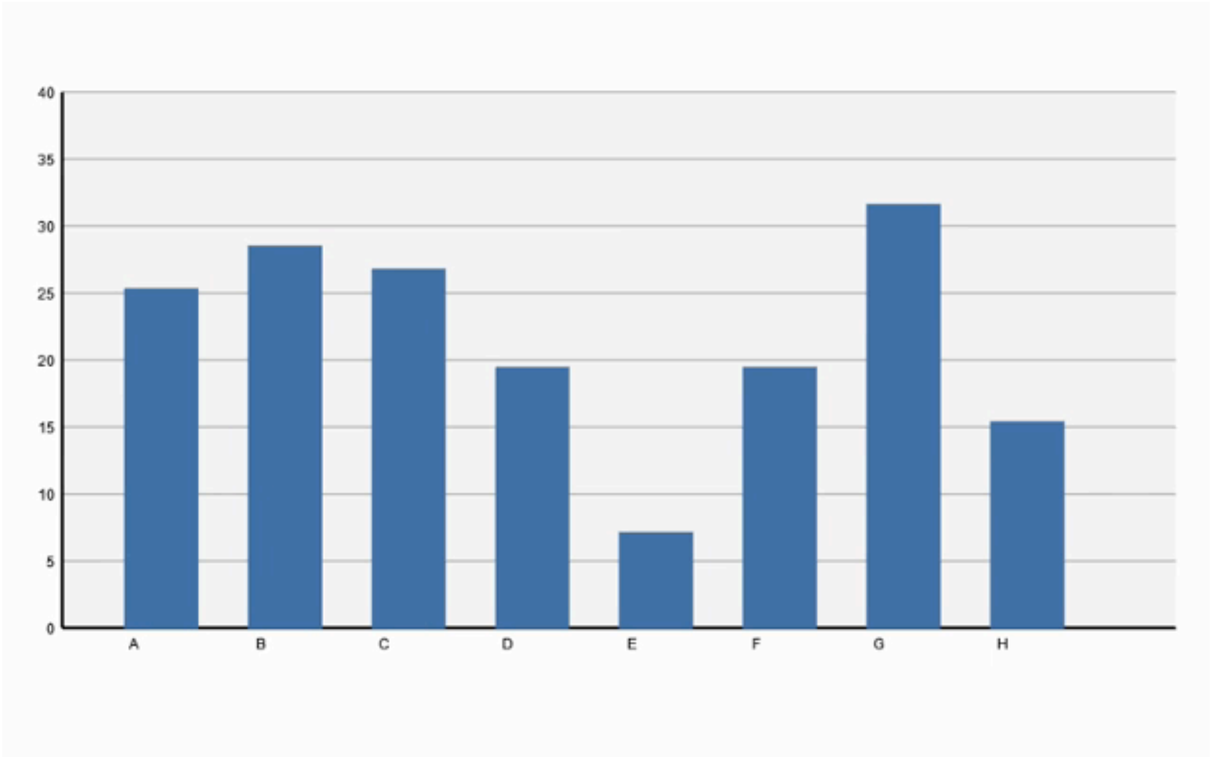
- external versus internal memory
  - easy to compare by moving eyes between views
  - hard to compare view to memory of what you saw
- when to use animation?
  - **good:** chronological storytelling
  - **good:** transition between states
  - **poor:** multiple states with multiple changes

Animation



Heer and Robertson, Animated Transitions in Statistical Data Graphics, 2007

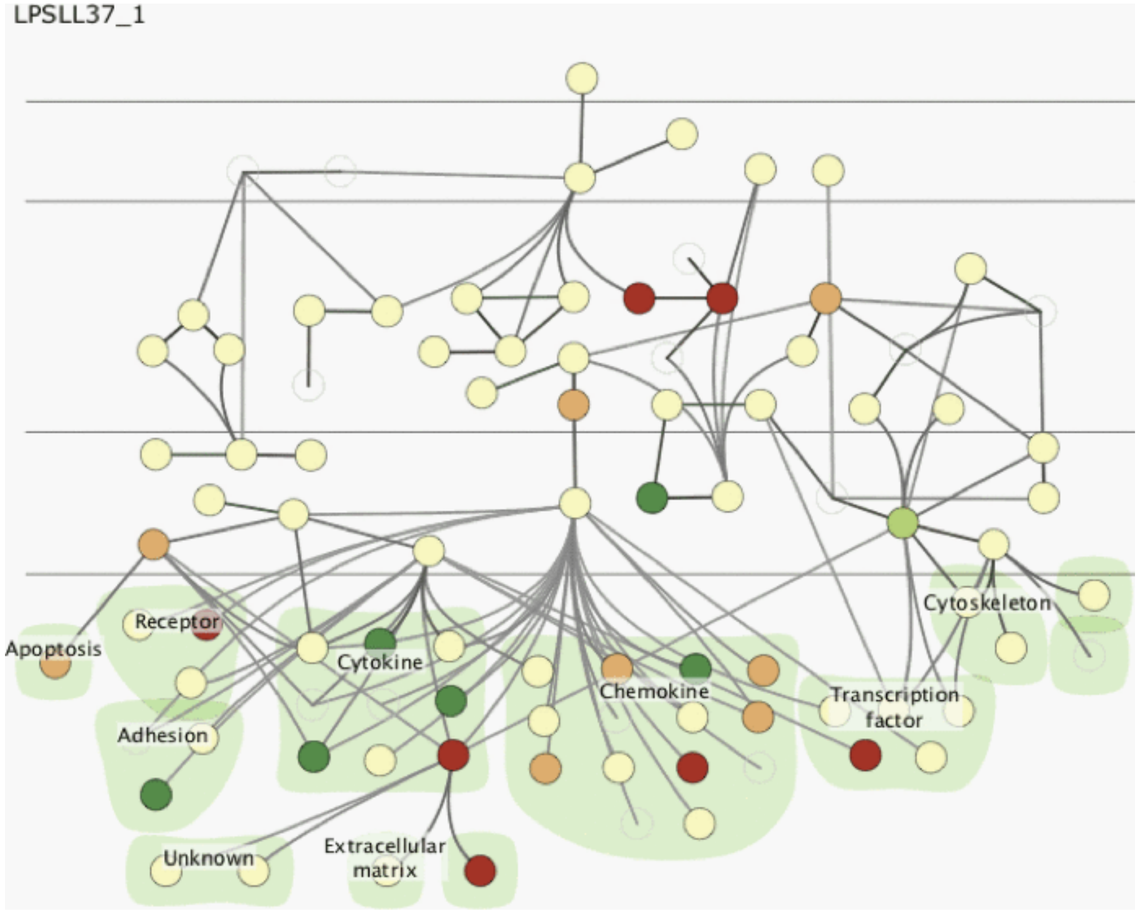
Animation



Heer and Robertson, Animated Transitions in Statistical Data Graphics, 2007

Animation Pitfall

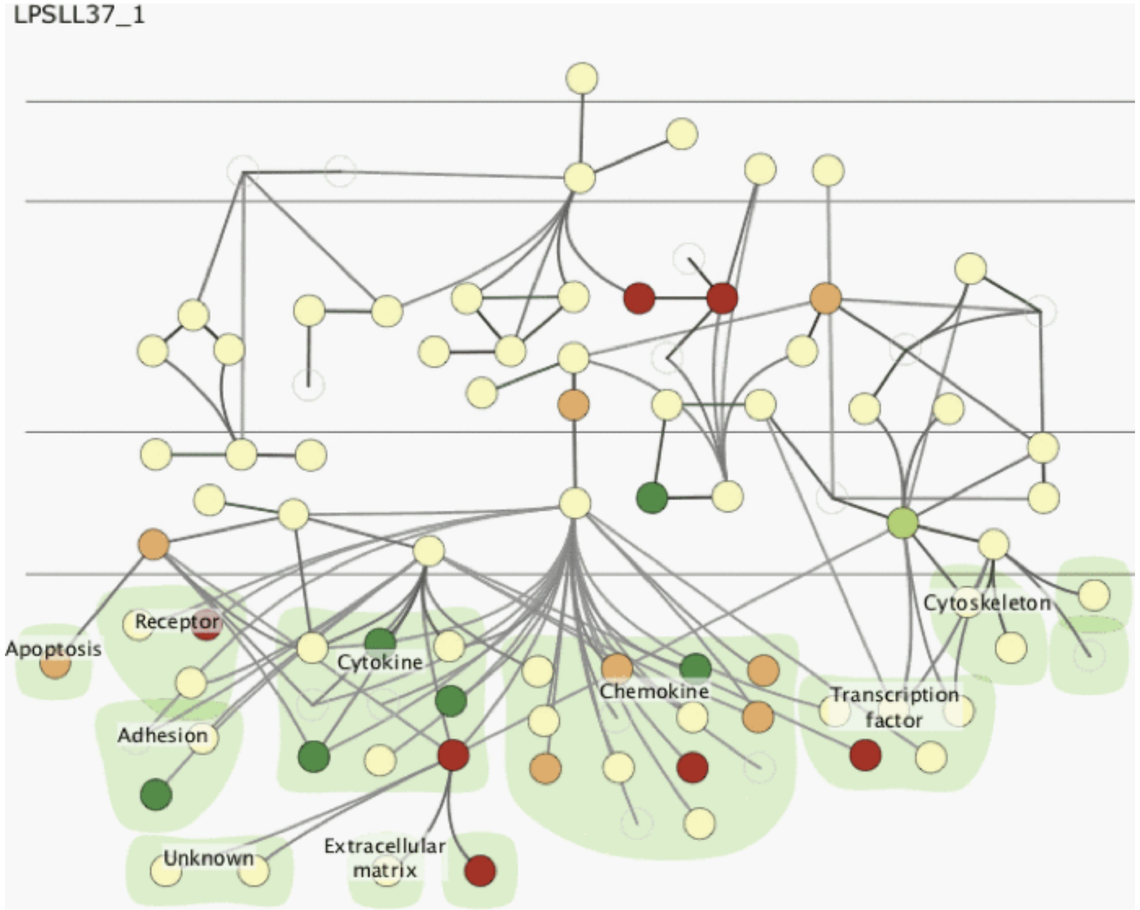
Global comparisons are difficult



Barsky et al. 2008, Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context

Animation Pitfall

Global comparisons are difficult



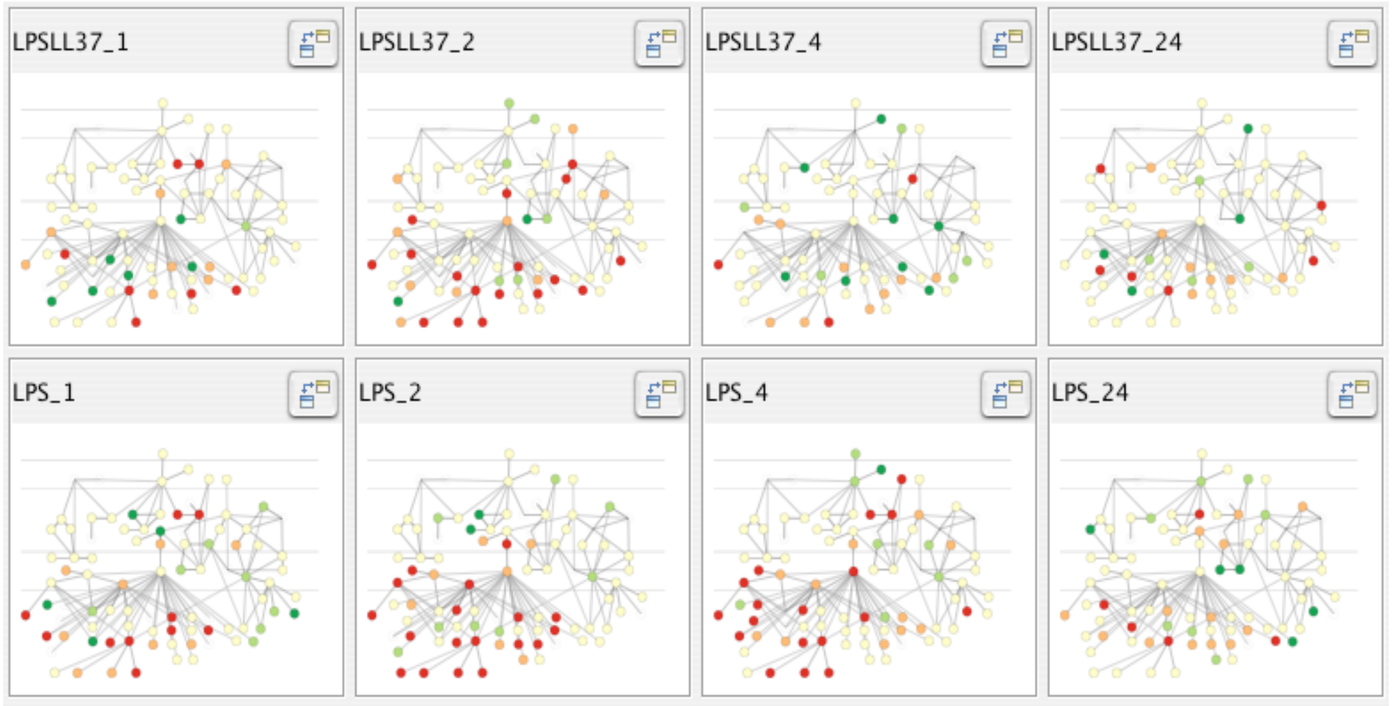
Barsky et al. 2008, Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context

Animation Pitfall

Small Multiples

one view per state

show time with space



Barsky et al. 2008, Cerebral: Visualizing Multiple Experimental Conditions on a Graph with Biological Context