Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Principles of Information Visualization Tutorial – Part 1
Design Principles

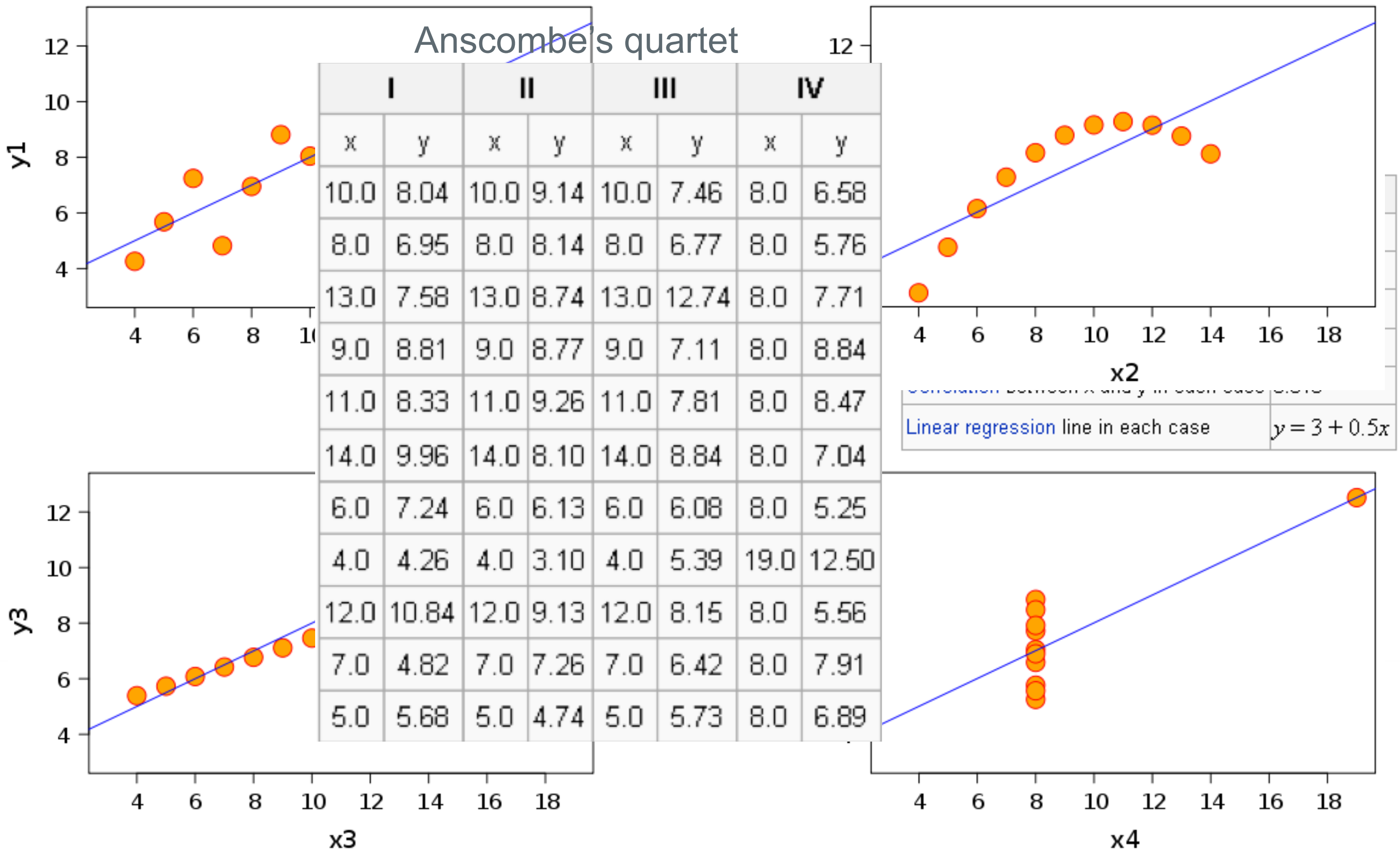
Prof Jessie Kennedy
Institute for Informatics & Digital Innovation

Edinburgh Napier
UNIVERSITY

Overview

- Fundamental principles of graphic design and visual communication
  - help you create more effective information visualizations.
- Use of salience, colour, consistency and layout
  - communicate large data sets and complex ideas with greater immediacy and clarity.

Why Visualise? To see what's in the data



Anscombe's quartet

I		II		III		IV	
x	y	x	y	x	y	x	y
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

Information Visualization

- 2 main objectives
- Data analysis
  - understand the data
  - derive information from them
  - involves comprehensivity
- Communication
  - of information
  - involves simplification

J Bertin, Semiology of Graphics, - Brief Presentation of Graphics, 2004

How do we get from Data to Visualization?

- Need to understand
  - the properties of the data or information
  - the properties of the image
  - the rules mapping data to images

How do we get from Data to Visualization?

- Need to understand
  - the properties of the data or information
  - the properties of the image
  - the rules mapping data to images

Types of Data

- Nominal (labels or types)
  - Sex: Male, Female,,
  - Genotype: AA, AT, AG…
- Ordinal
  - Days: Mon, Tue, Wed, Thu, Fri, Sat, Sun
  - Abundance: abundant - common – rare
- Quantitative
  - Physical measurements: temperature, expression level

S. S. Stevens, On the theory of scales of measurements, 1946

Data Type Taxonomy

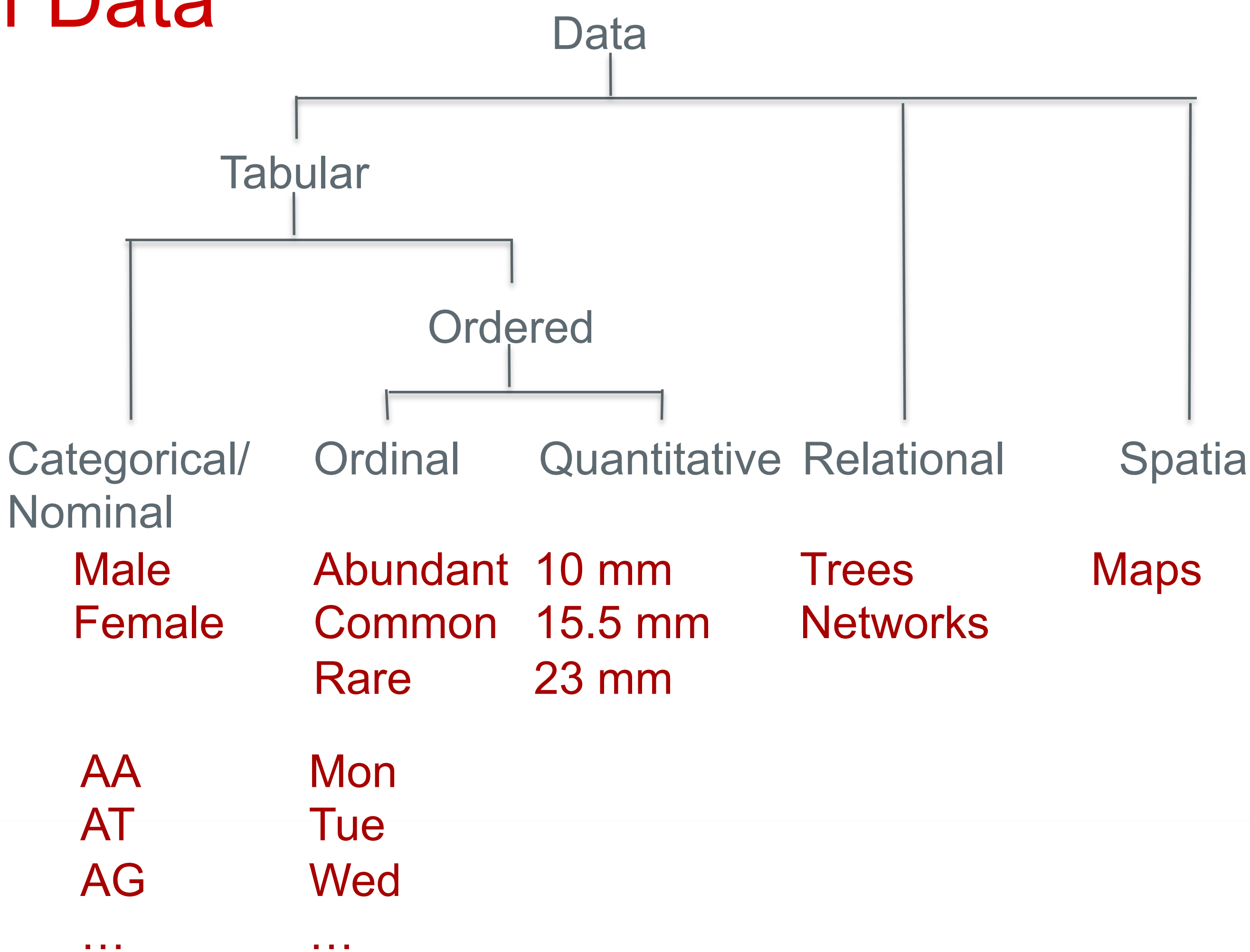
- 1D e.g. DNA sequences
- Temporal e.g. time series microarray expression
- 2D e.g. distribution maps
- 3D e.g. Anatomical structures
- nD e.g. Fisher's Iris data set
- Trees e.g Linnean taxonomies, phylogenies
- Networks e.g. Metabolic pathways
- Text and documents e.g. publications

B. Shneiderman, The eyes have it: A task by data type taxonomy for information visualization, 1996

Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Types of Data



slide adapted from Munzner 2011, Visualization Principles

How do we get from Data to Visualization?

- Need to understand
  - the properties of the data or information
  - the properties of the image
  - the rules mapping data to images

Theory of Graphics

- Application of human perception
  - understand and memorize forms in an image
  - XY dimensions of the plane and variation in Z dimension
- Correspondence between data and image
- Level of perception required by objective
- Mobility or immobility of the image

J Bertin, Semiology of Graphics, - Brief Presentation of Graphics, 2004

Semiology of Graphics

- visual encoding
  - points, lines, areas

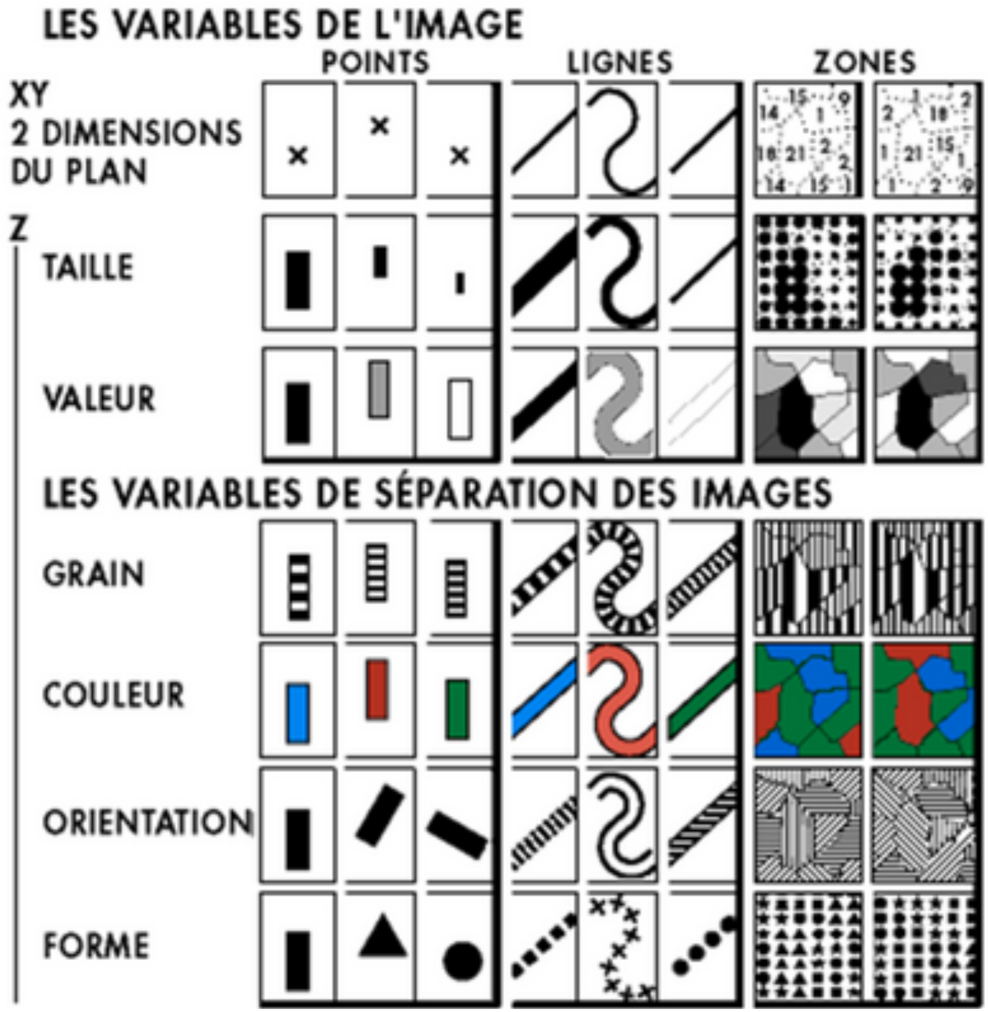
    patterns, trees/networks, grids
  - positional: XY

    1D, 2D, 3D
  - retinal: Z

    size, lightness, texture,

    colour, orientation, shape,
  - temporal:

    animation

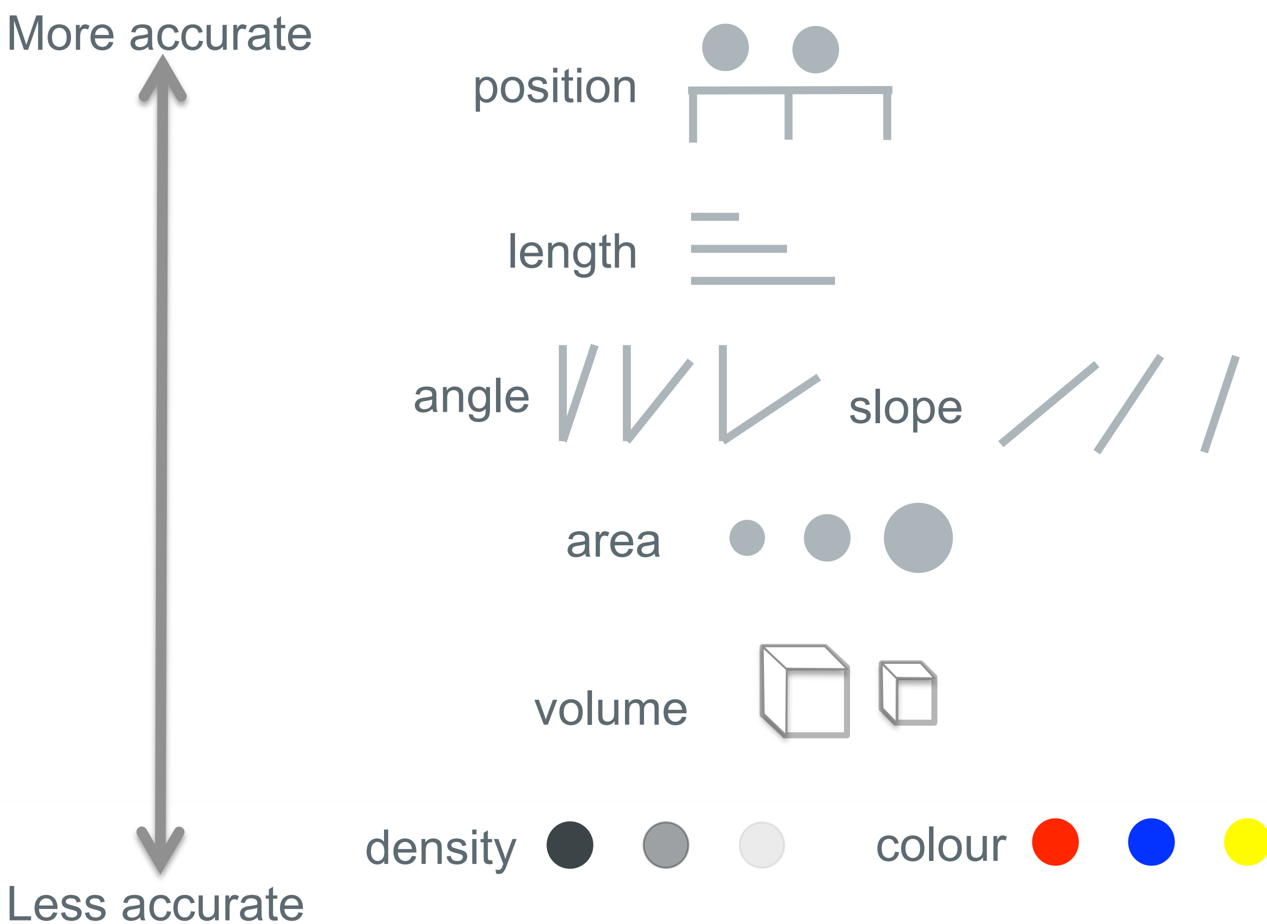


Semiology of Graphics. Jacques Bertin, Gauthier-Villars 1967, EHESS 1998

Language of Graphics

- Graphics can be thought of as forming a sign system:
  - Each mark (point, line, or area) represents a data element.
  - Choose visual variables to encode relationships between data elements
    - difference, similarity, order, proportion
    - only position supports all relationships
- Huge range of alternatives for data with many attributes
  - find images that express and effectively convey the information.

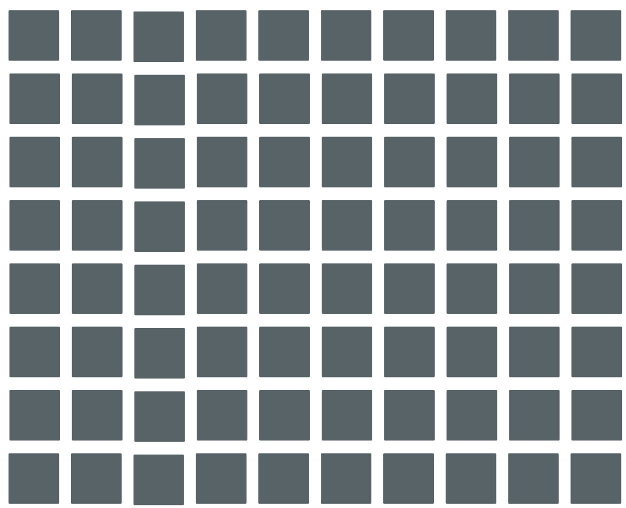
Accuracy of Quantitative Perceptual Tasks



Cleveland, W.S. & McGill, R. Science 229, 828–833 (1985).

Gestalt Effects

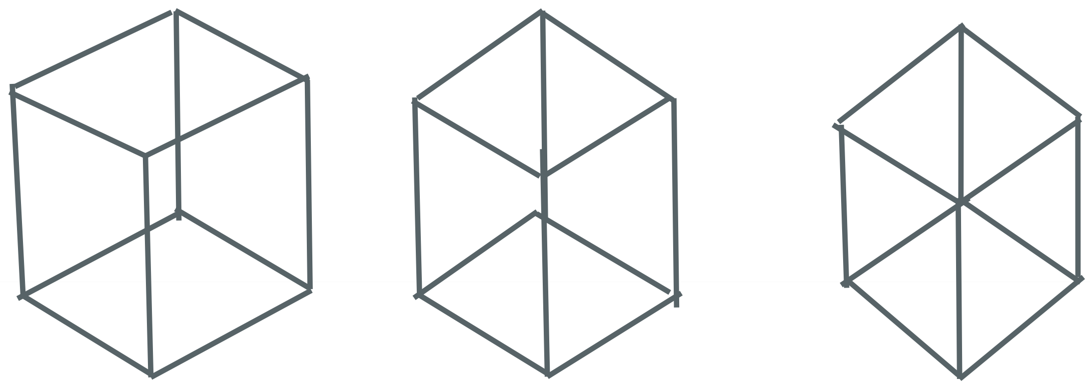
- Visual system tries to structure what we see into patterns
- Gestalt is the interplay between the parts and the whole
  - "The whole is 'other' than the sum of its parts." – Kurt Koffka



- Gestalt Laws/Principles

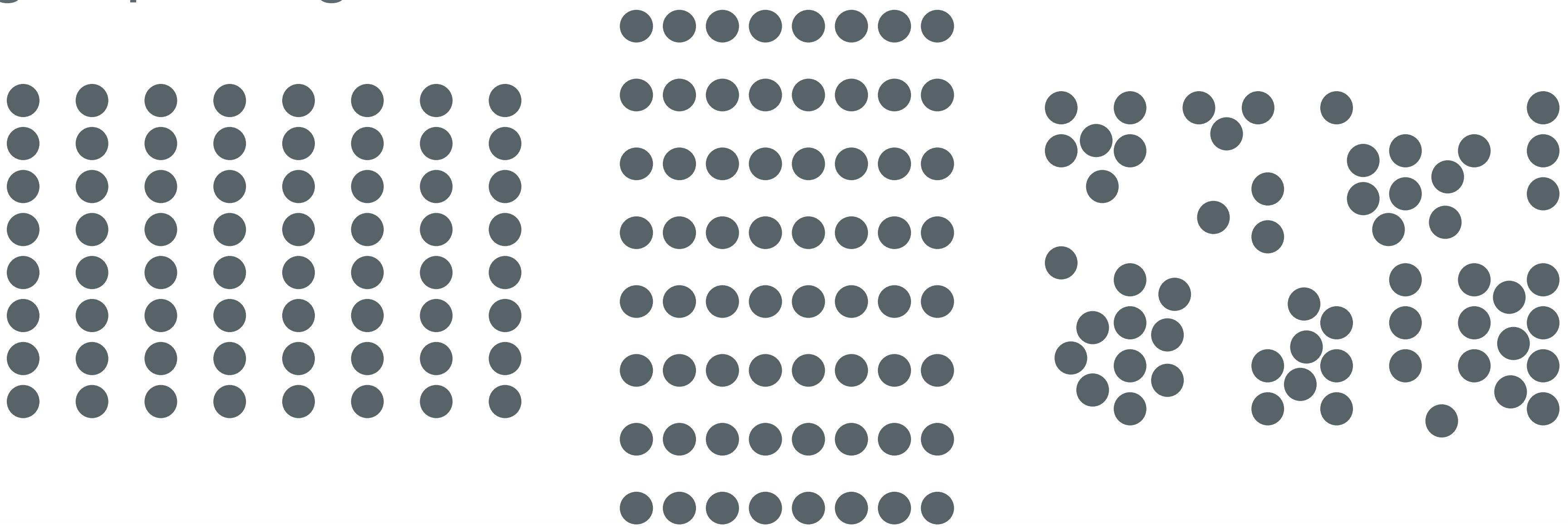
Principle of Simplicity

- Every pattern is seen such that the resulting structure is as simple as possible
  - Different projections of same cube
  - Perceived as 2 or 3 D
  - Depending on the simpler interpretation



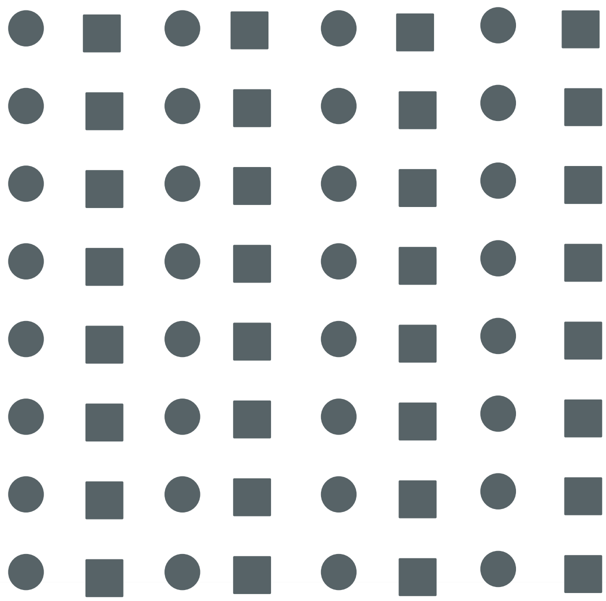
Principle of Proximity

- Things that are near to each other appear to be grouped together



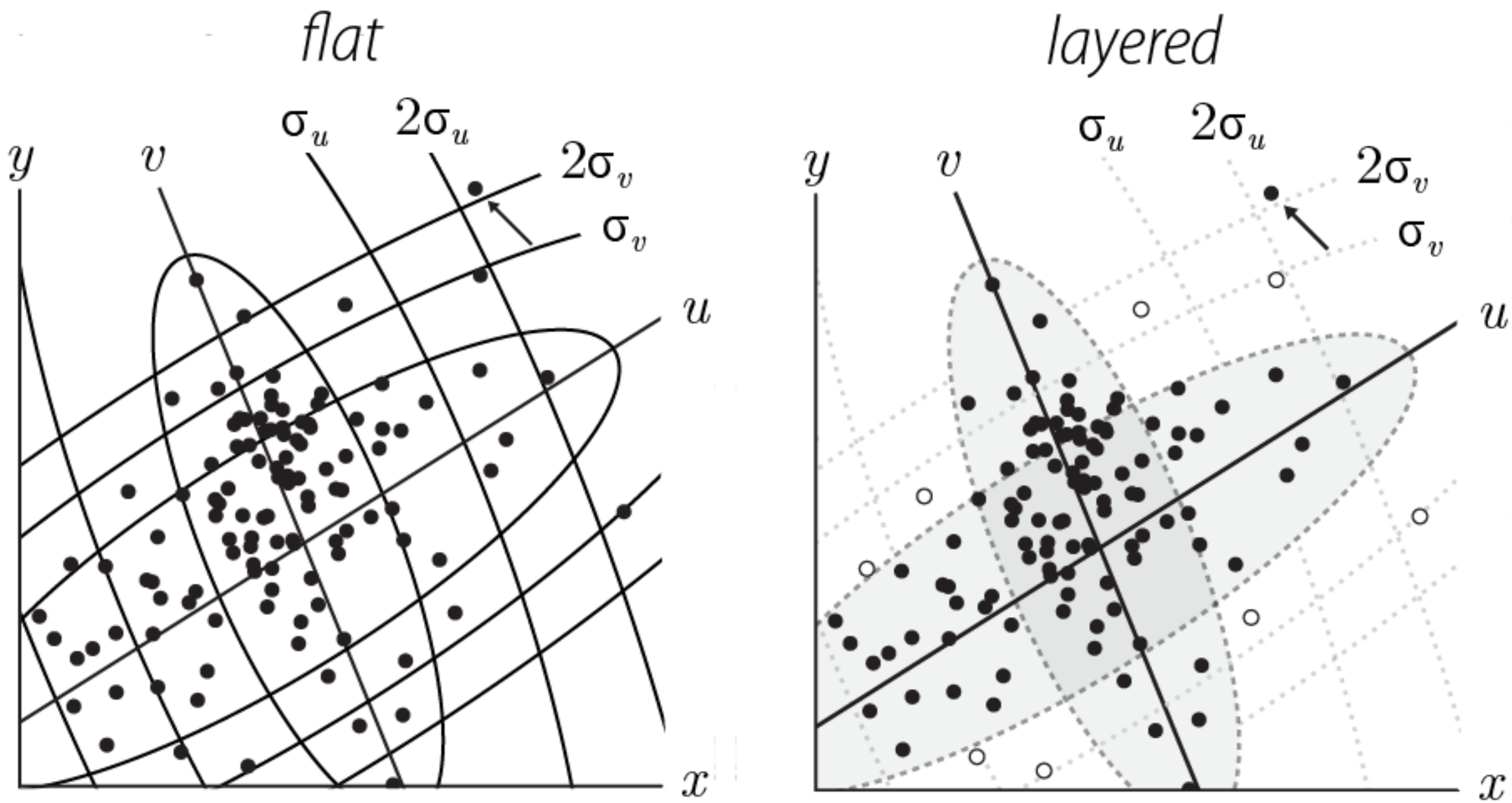
Principle of Similarity

- Similar things appear to be grouped together



Variable Opacity for Clarity

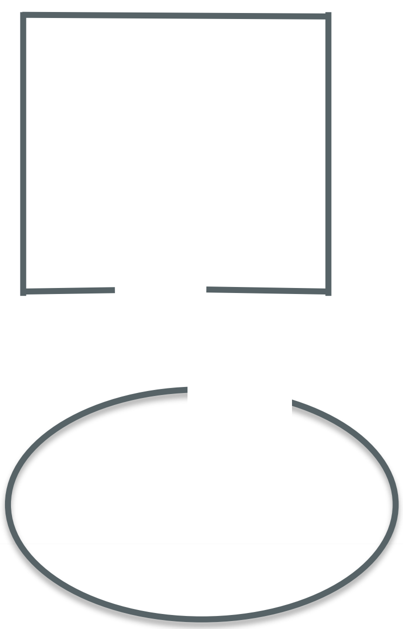
- Use of similarity of stroke and opacity to clarify image
  - Layers in the image

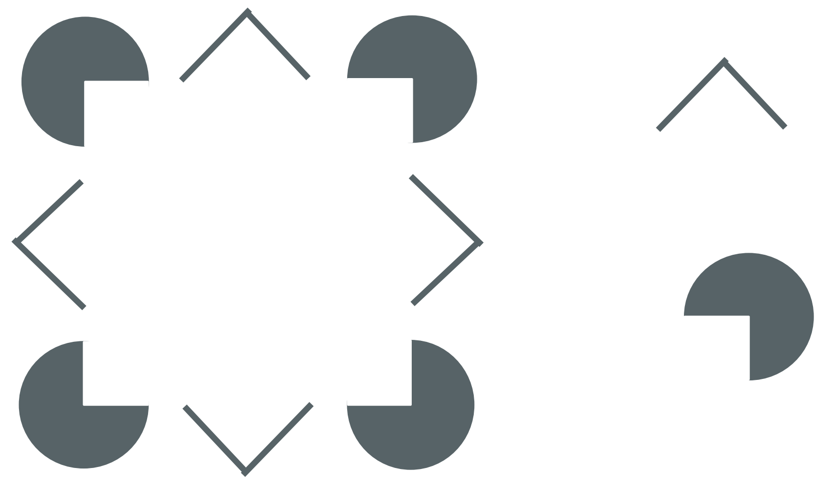


M. Krzwinski, behind every great visualization is a design principle, 2012

Principle of Closure

- The law of closure posits that we perceptually close up, or complete, objects that are not, in fact, complete

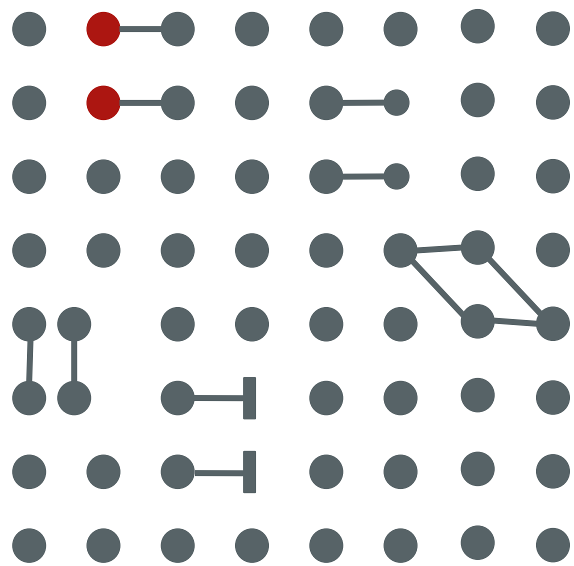




Illusory

Principle of Connectedness

- Things that are physically connected are perceived as a unit



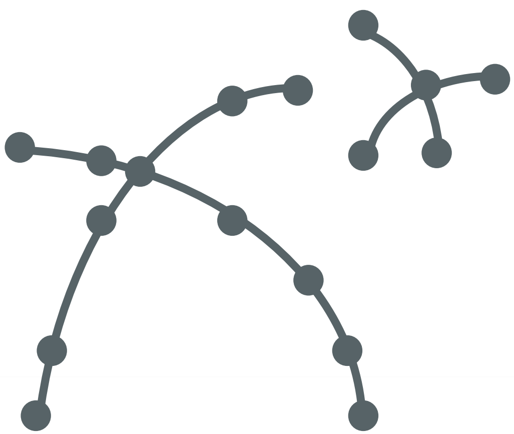
- Stronger than colour, shape, proximity, size

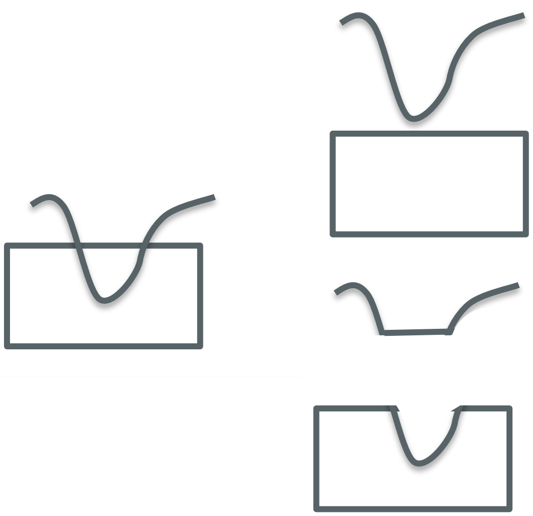
Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Principle of Good Continuation

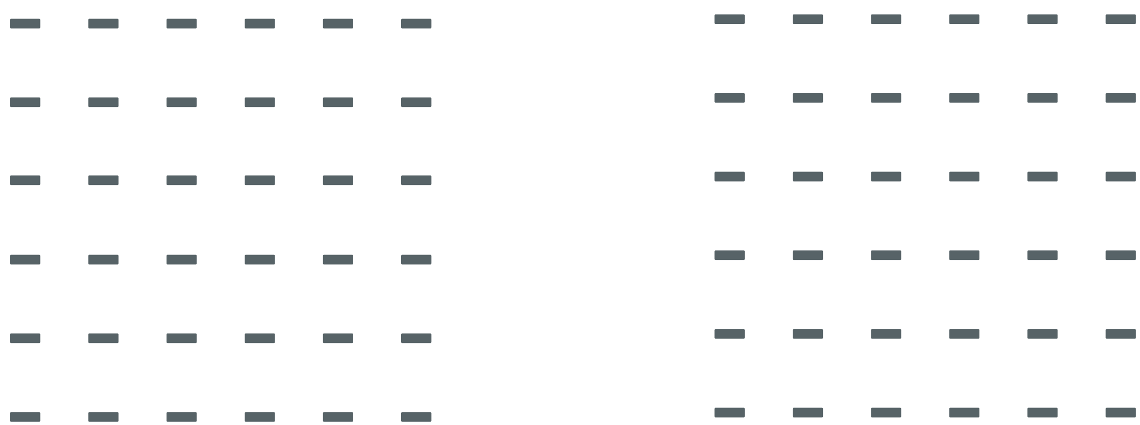
- Points connected in a straight or smoothly curving line are seen as belonging together
  - lines tend to be seen as to follow the smoothest path





Principle of Common Fate

- Things that are moving in the same direction appear to be grouped together



Principle of Familiarity

- Things are more likely to form groups if the groups appear familiar or meaningful

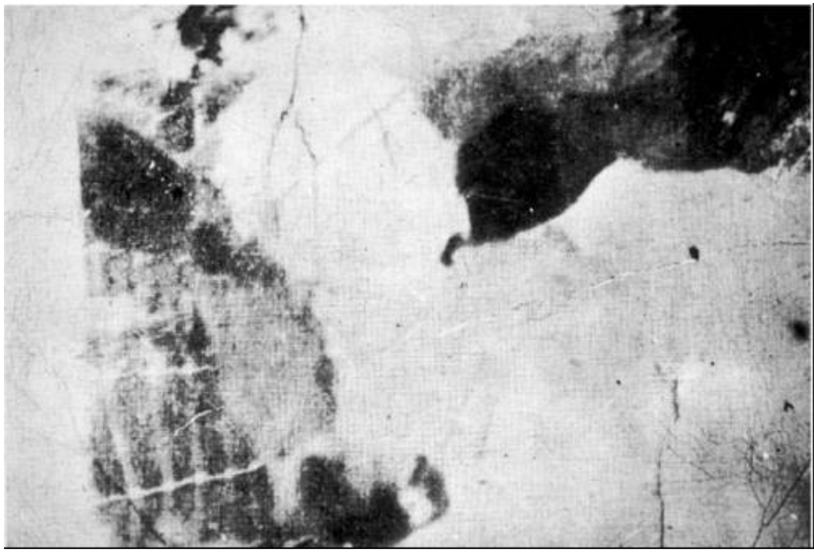
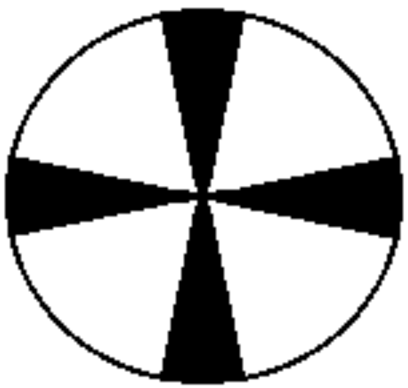
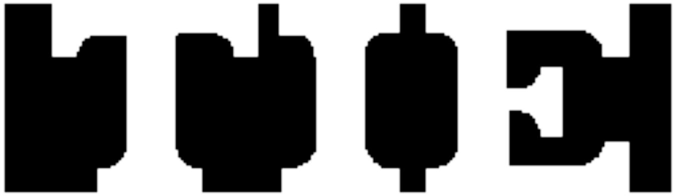


Figure-Ground & Smallness

- Smaller areas seen as figures against larger background



- Surroundedness

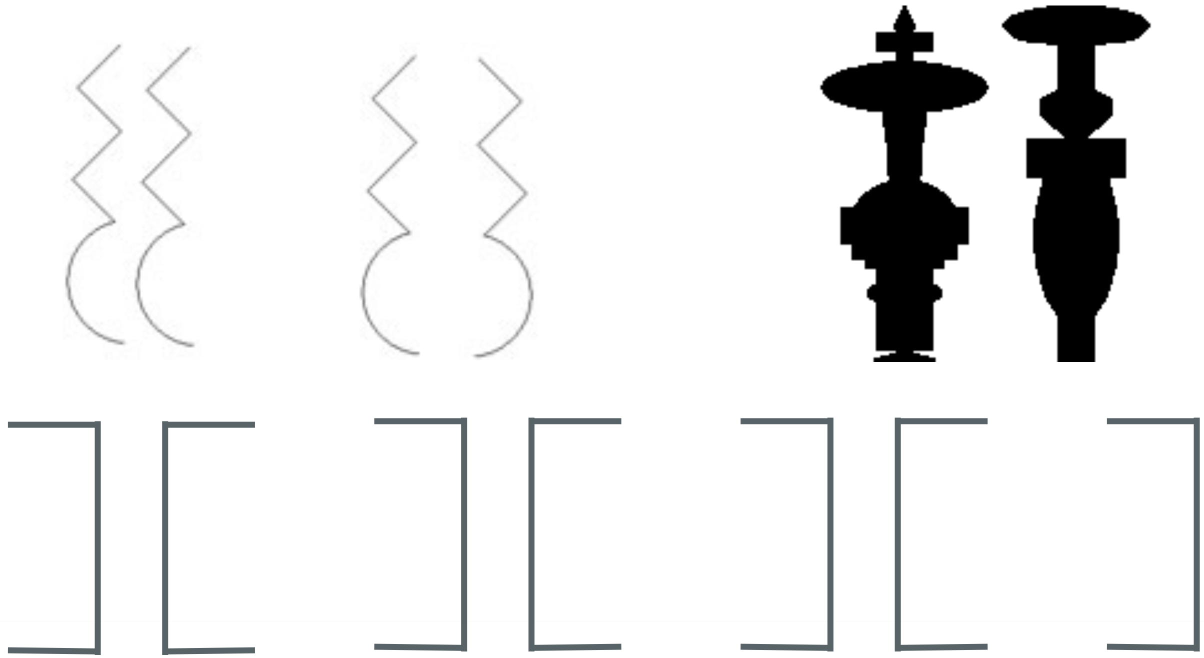


Shaping our
digital future

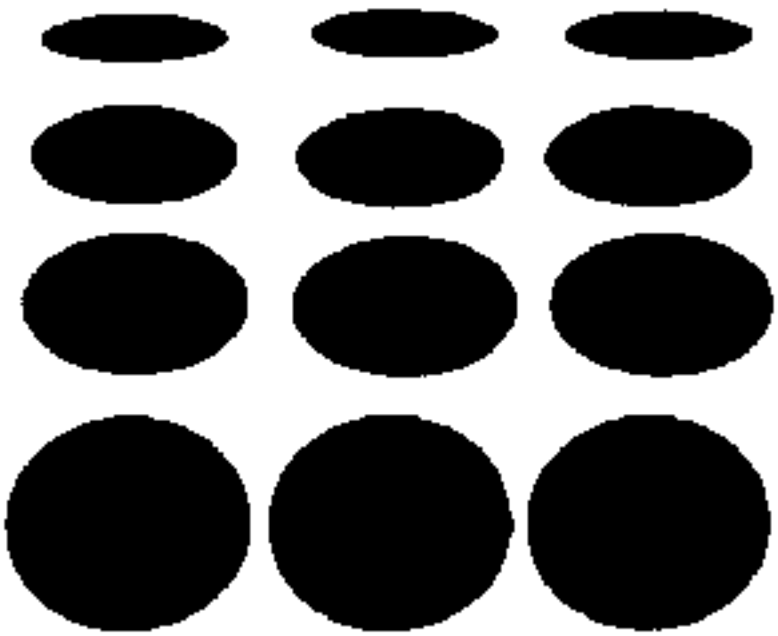
Institute for
Informatics
& Digital
Innovation

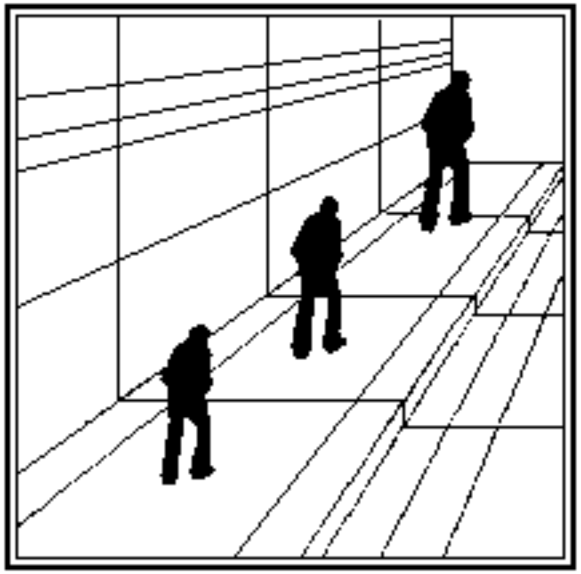
Principle of Symmetry

- The principle of symmetry is that, the symmetrical areas tend to be seen as figures against the asymmetrical background.



3D Effect

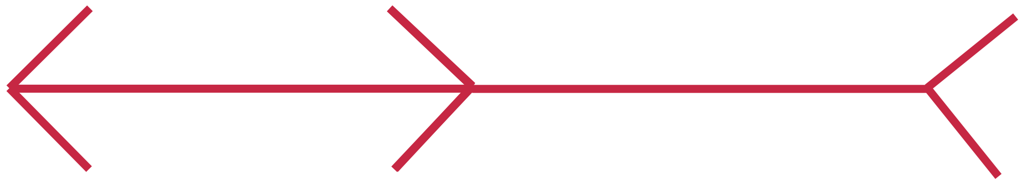




Context affects perceptual tasks

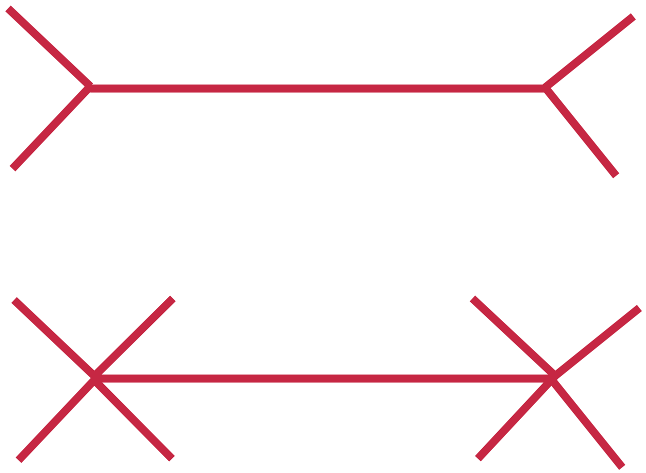
- Comparing values
  - Length
  - Curvature
  - Area
  - 2.5D shape
  - Position in 2.5D

Ambiguous Information: Length



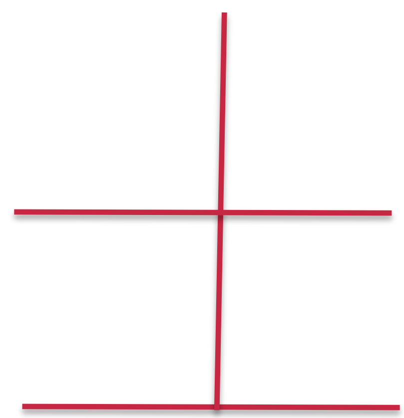
Muller-Lyer

Ambiguous Information: Length



Muller-Lyer

Horizontal-Vertical Illusion



Shaping our digital future

Institute for Informatics & Digital Innovation

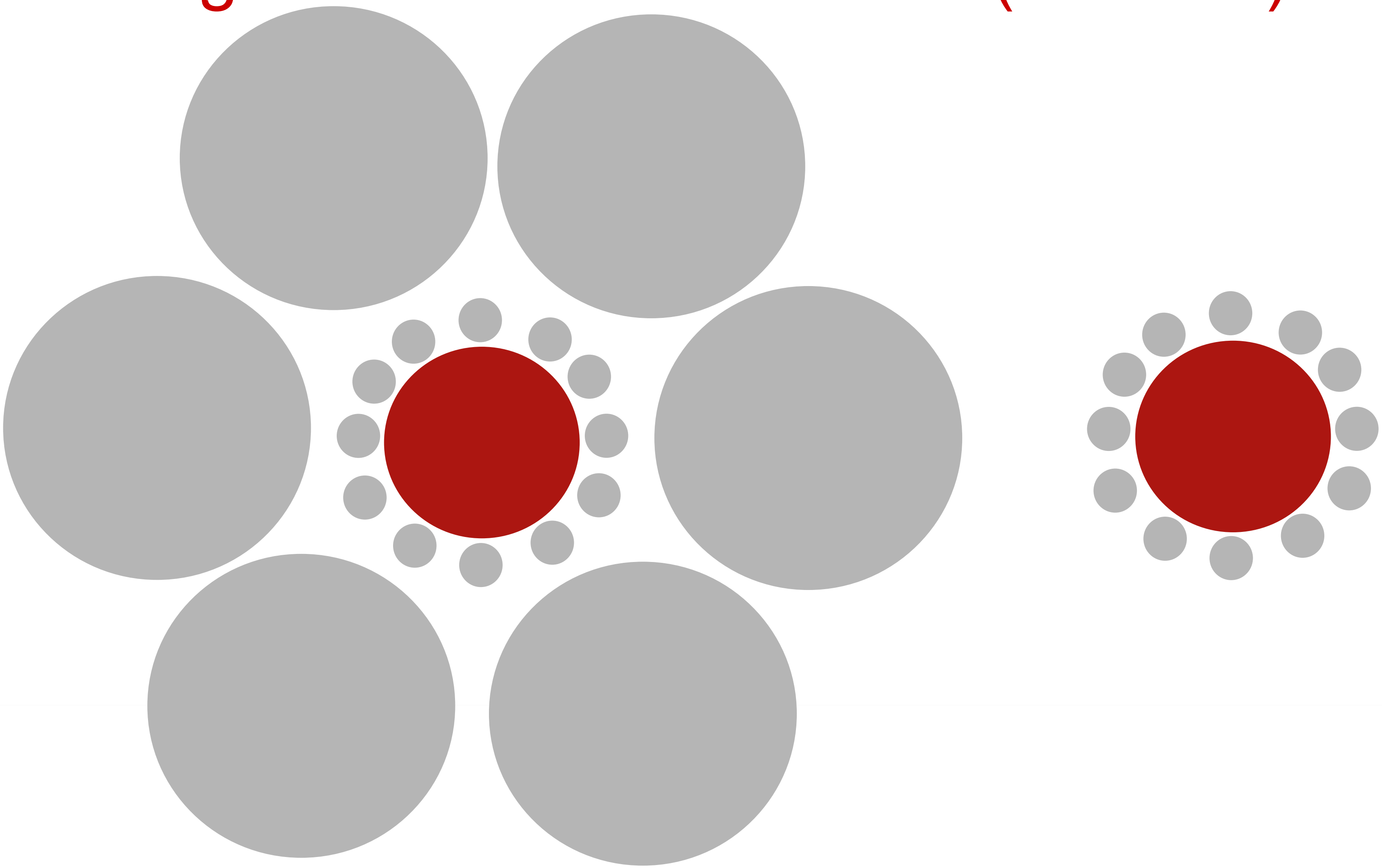
Ambiguous Information: Curvature

Tollansky

Shaping our digital future

Institute for Informatics & Digital Innovation

Ambiguous Information: Area (Context)

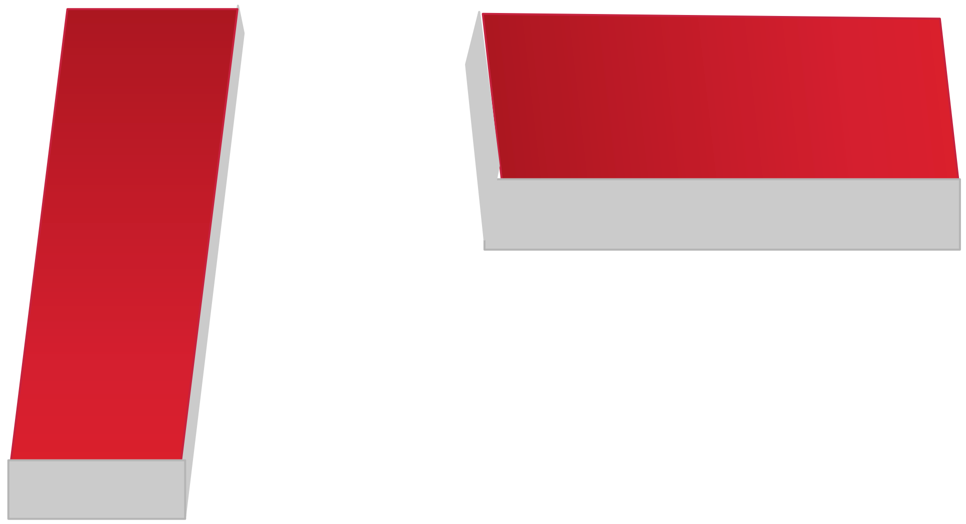


Ebbinghaus

Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

2.5D Shape

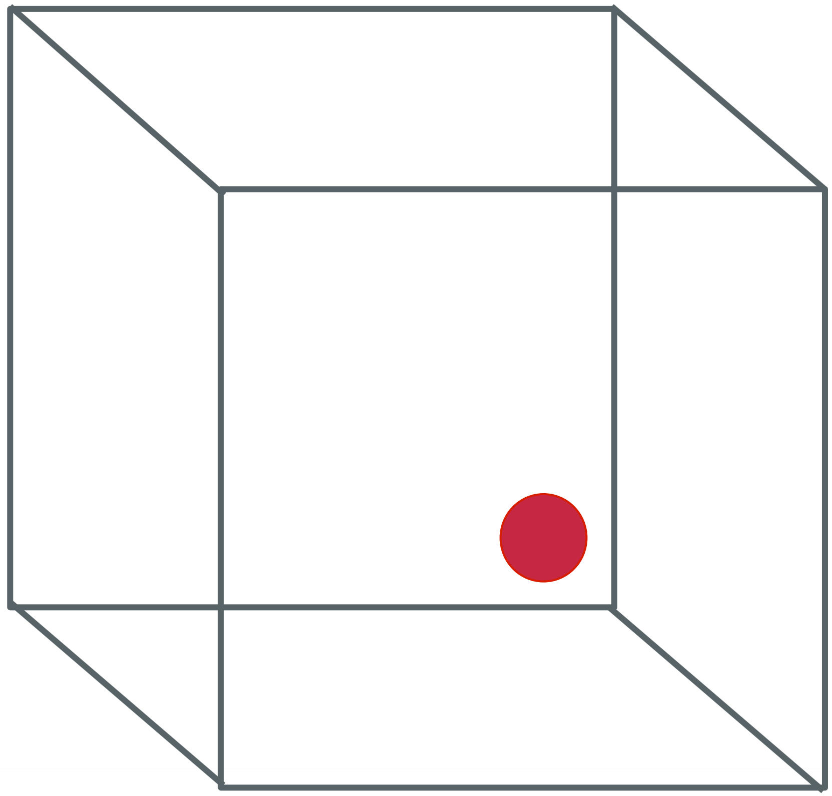


Adapted from Shepard R N, 1990 Mind Sights: Original Visual Illusions, Ambiguities, and other Anomalies (

2.5D Shape

Adapted from Shepard R N, 1990 Mind Sights: Original Visual Illusions, Ambiguities, and other Anomalies (

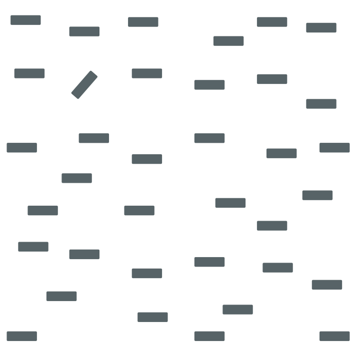
Ambiguous Information: Position in 2.5D space



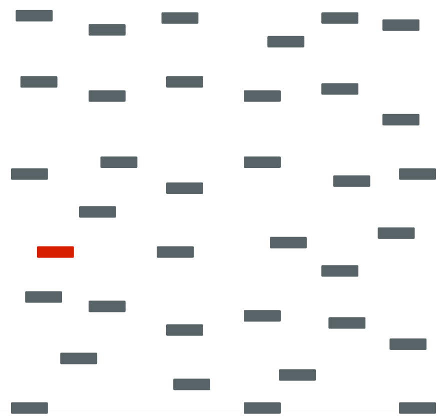
Necker Cube

Preattentive Visual Features

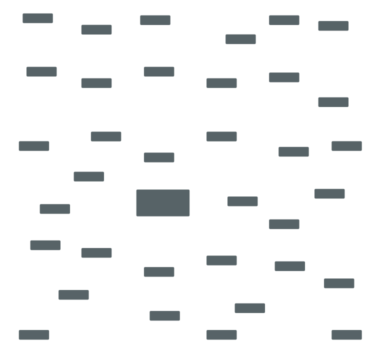
- the ability of the low-level human visual system to rapidly identify certain basic visual properties
- a unique visual property e.g., colour red allows it to "pop out"
- aids visual searching



orientation



colour

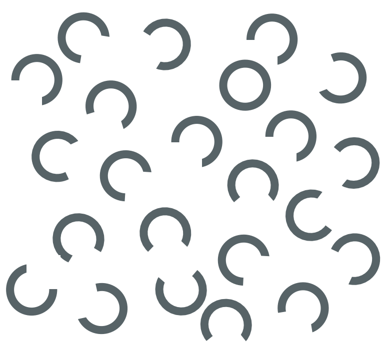


size

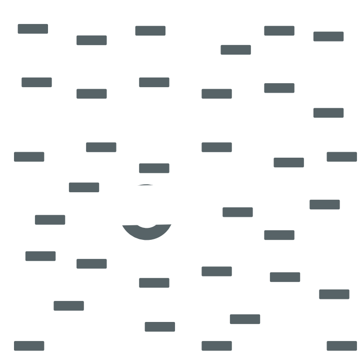
Adapted from Perception in visualization C. Healey, : http://www.csc.ncsu.edu/faculty/healey/PP/

Preattentive Visual Features

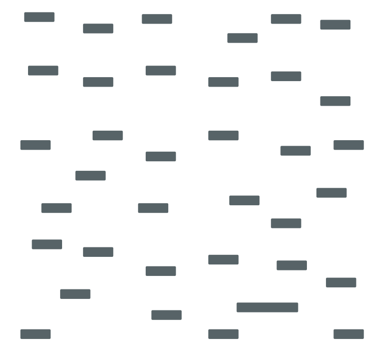
- Some more effective than others



closure



curvature



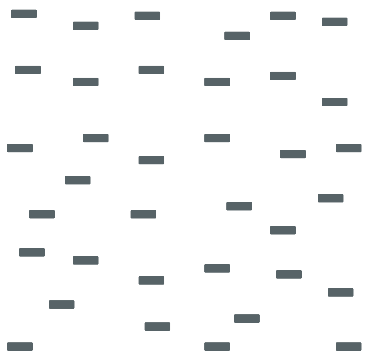
length

Adapted from Perception in visualization C. Healey, : http://www.csc.ncsu.edu/faculty/healey/PP/

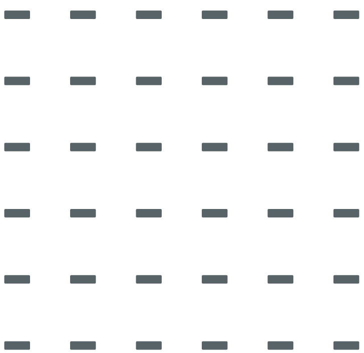
Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

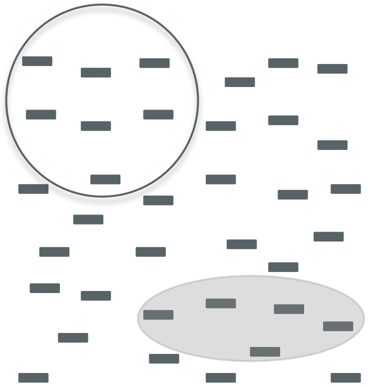
Preattentive Visual Features



flicker



direction of movement

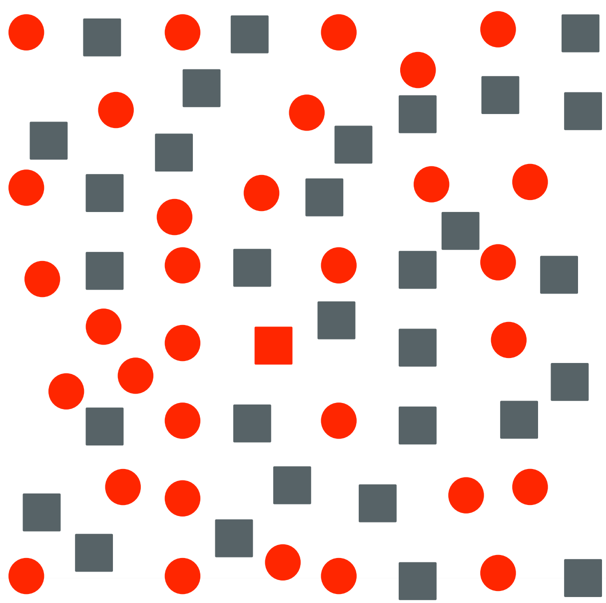


enclosure/containment

Adapted from Perception in visualization C. Healey, : http://www.csc.ncsu.edu/faculty/healey/PP/

More than 2 Preattentive visual features

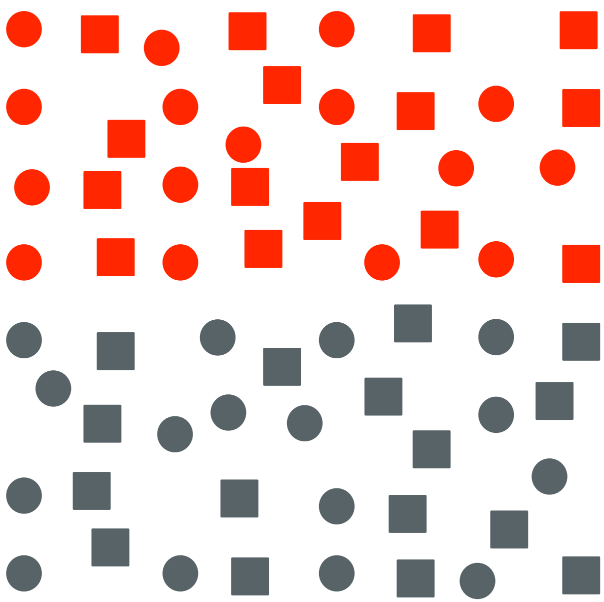
- A target made up of a combination of non-unique features normally cannot be detected preattentively



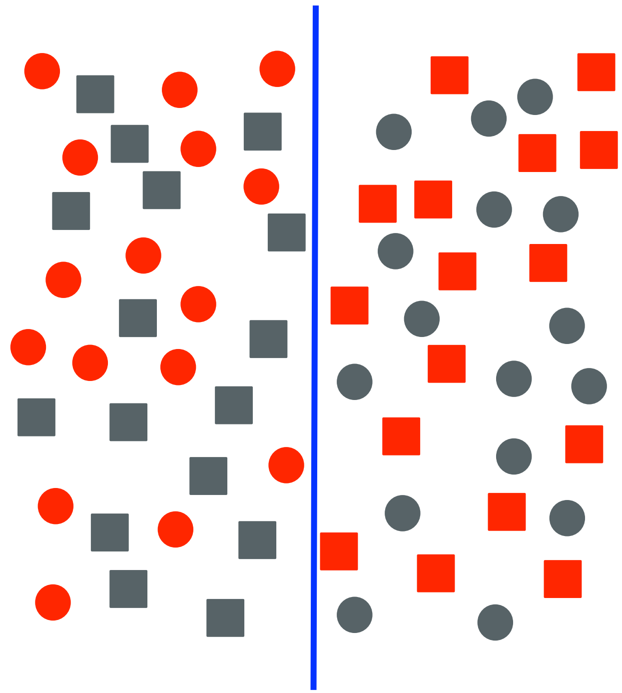
- spot the red square
- difficult to detect
- serial search required

Adapted from Perception in visualization C. Healey, : http://www.csc.ncsu.edu/faculty/healey/PP/

Boundary detection



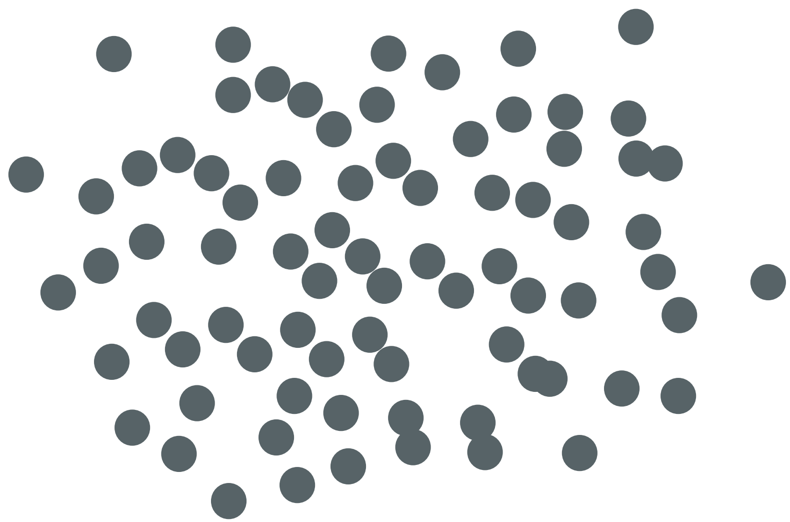
Horizontal boundary



Vertical boundary

Adapted from Perception in visualization C. Healey, : http://www.csc.ncsu.edu/faculty/healey/PP/

Region tracking



Use of preattentive features

- target detection:
  - users rapidly and accurately detect the presence or absence of a "target" element with a unique visual feature within a field of distractor elements
- boundary detection:
  - users rapidly and accurately detect a texture boundary between two groups of elements, where all of the elements in each group have a common visual property
- region tracking:
  - users track one or more elements with a unique visual feature as they move in time and space, and
- counting and estimation:
  - users count or estimate the number of elements with a unique visual feature.

COLOURCO

LOURCOLO

URCOLOUR

COLOURCO

Colour

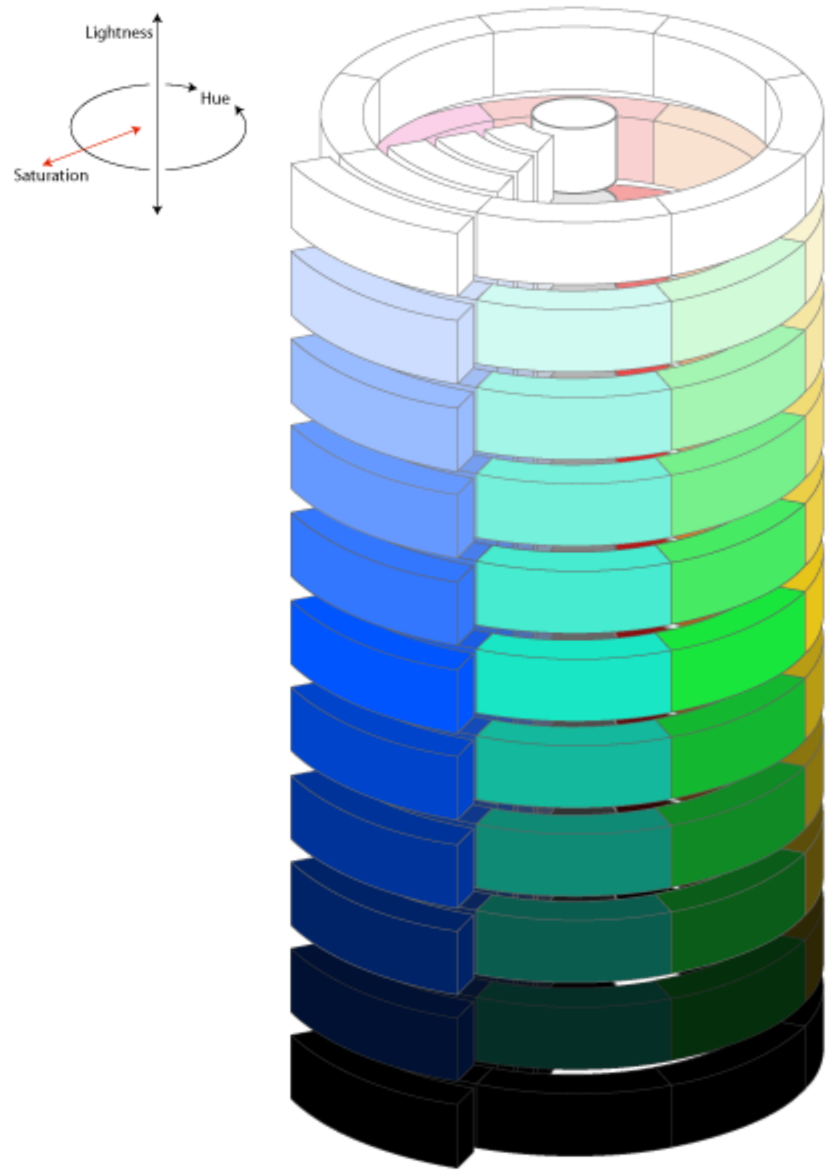
- “Colour used poorly is worse than no colour at all” - Edward Tufte
  - “Above all, do no harm”
  - colour can cause the wrong information to stand out and
  - make meaningful information difficult to see.

Colour space

- A *colour space* is mathematical model for describing colour.
  - RGB, HSB, HSL, Lab and LCH
- RGB is the most common in computer use,
  - but least useful for design
  - our eyes do not decompose colours into RGB constituents
- HSV, describes a colour in terms of its hue, saturation and value (lightness),
  - models colour based on intuitive parameters
  - more useful.

Colourimetry

- Hue (colour)
  - around the circle
- Saturation
  - Inside to outside
  - Colour to grey scale
- Lightness (value)
  - top to bottom



Figs. Courtesy of S Rogers, ONS

Brewer Palettes

- Brewer palettes (colorbrewer.org) provide a range of palettes based on HSV model which make life easier for us….

Avoid the use of hue to encode quantitative variables

Quantitative encoding e.g. heat maps

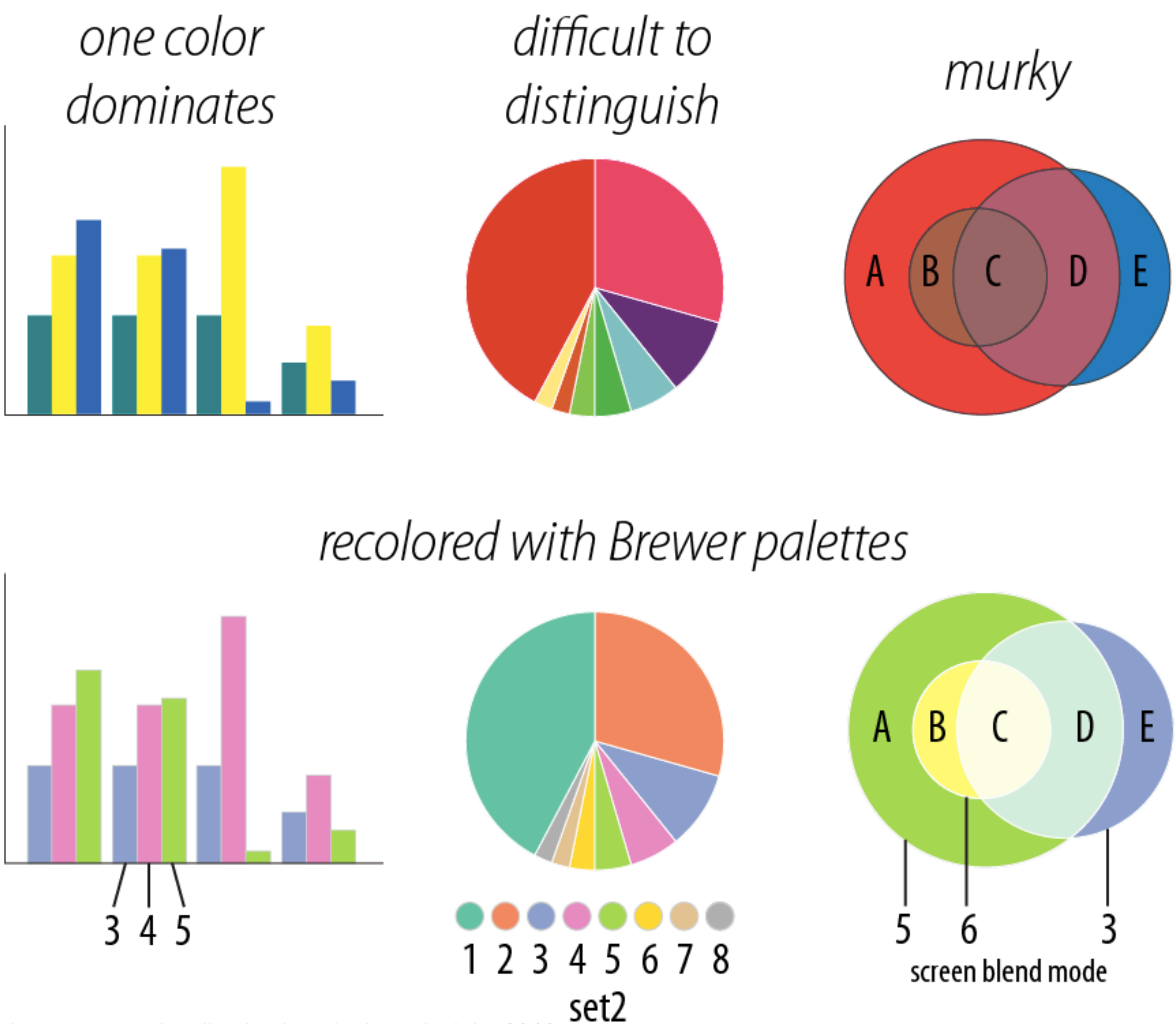
Two-sided quantitative encodings

Fig. Courtesy of M. Krzwinski,

Examples

Poor use of colour

Brewer colours



M. Krzwinski, behind every great visualization is a design principle, 2012

Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Conversion to Grey scale

- Ensure chosen colour set works well in grey scale
  - Sequential palette works well here

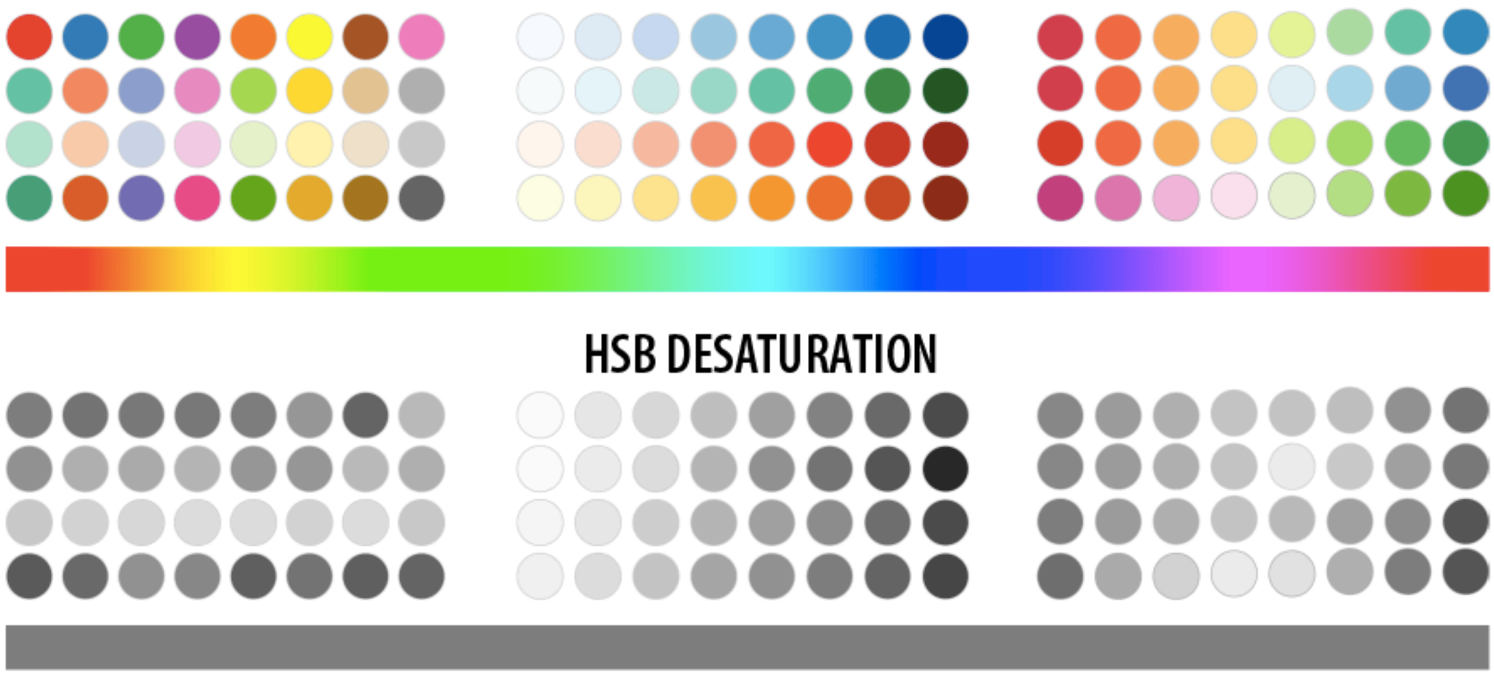
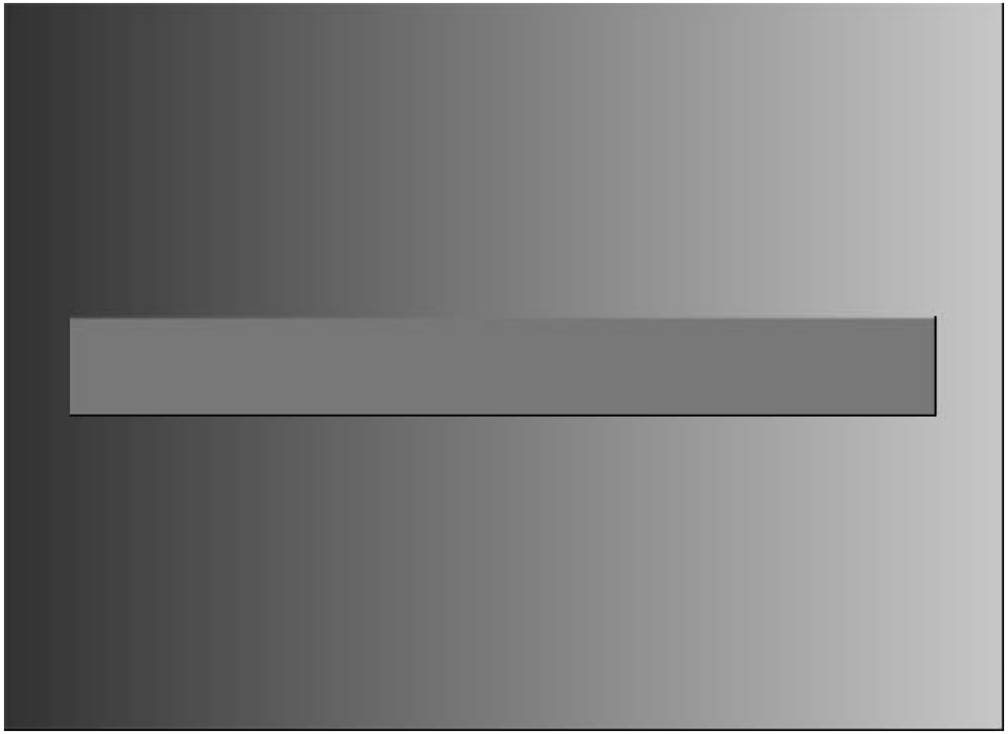


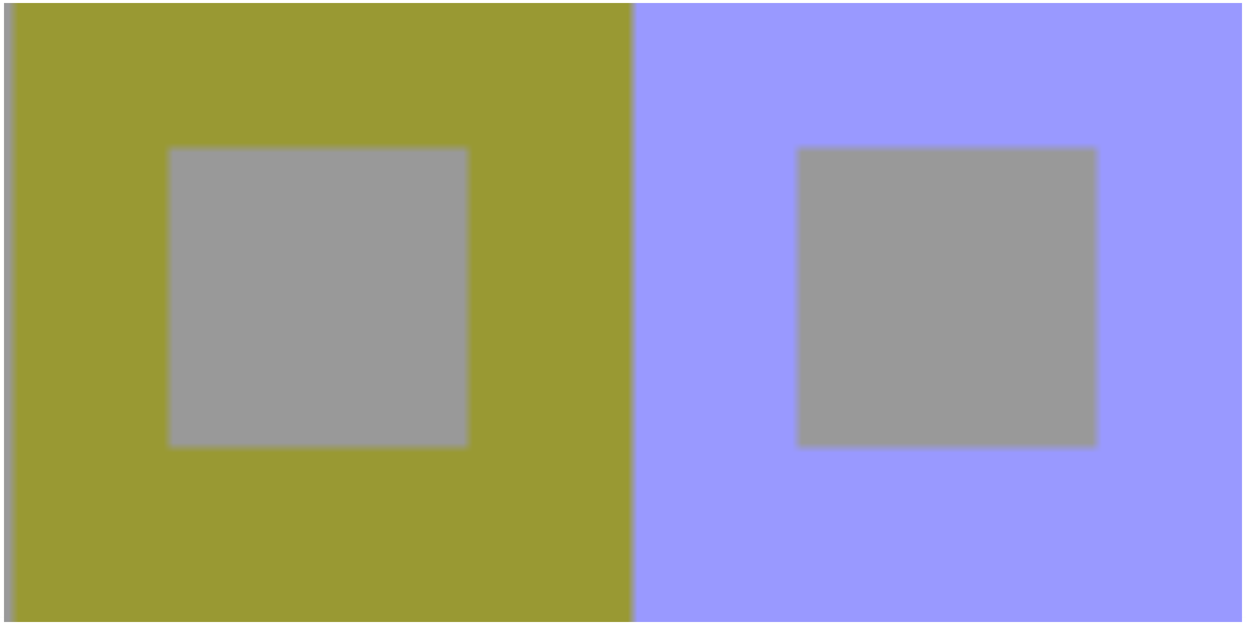
Fig. Courtesy of M Krzywinski

Trouble with perceptual colour….



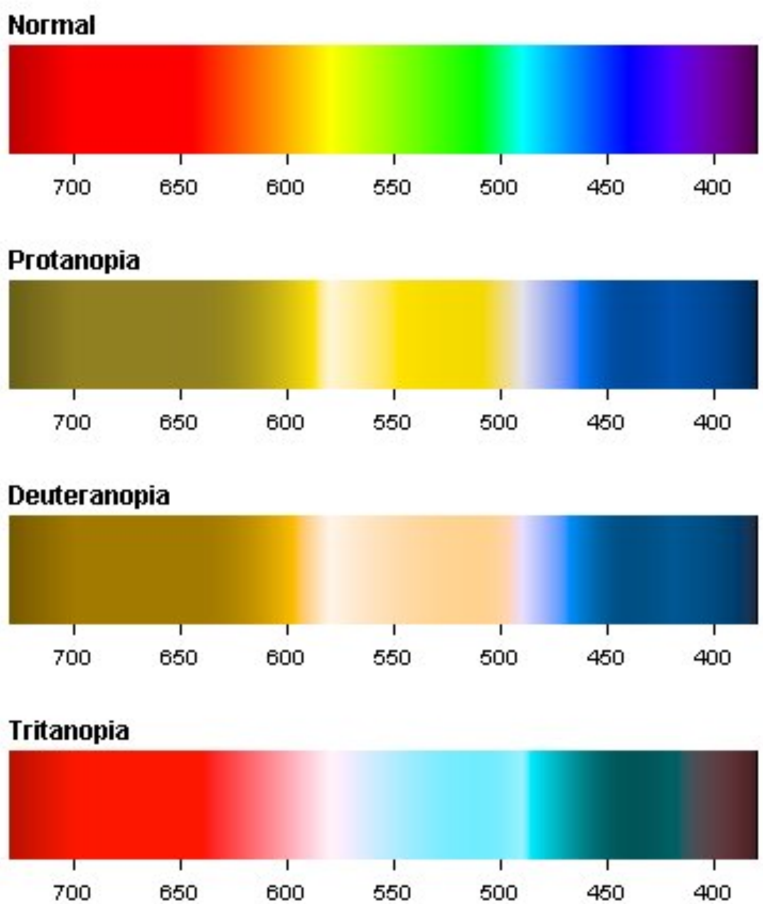
Figs. Courtesy of S Rogers, ONS

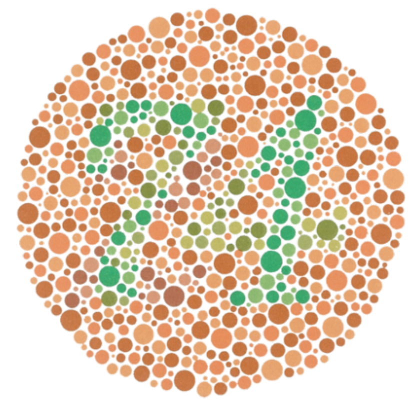
Context Affects Perceived Colour



Figs. Courtesy of S Rogers, ONS

Colour & Accessibility….





Accessibility (W3C):

10-20% of population are red/green colour blind.

(74? 21? No number at all?)….

Figs. Courtesy of S Rogers, ONS

Colour Blindness

8% males of USA descent

Red-green

Red-green

Blue-yellow

Fig. Courtesy of M Krzywinski

Shaping our digital future

BioVis Example: Immunofluorescence images

red-green image of P2Y1 receptor and migrating granule neurons,

effectively remapped to magenta-green using the channel mixing method.

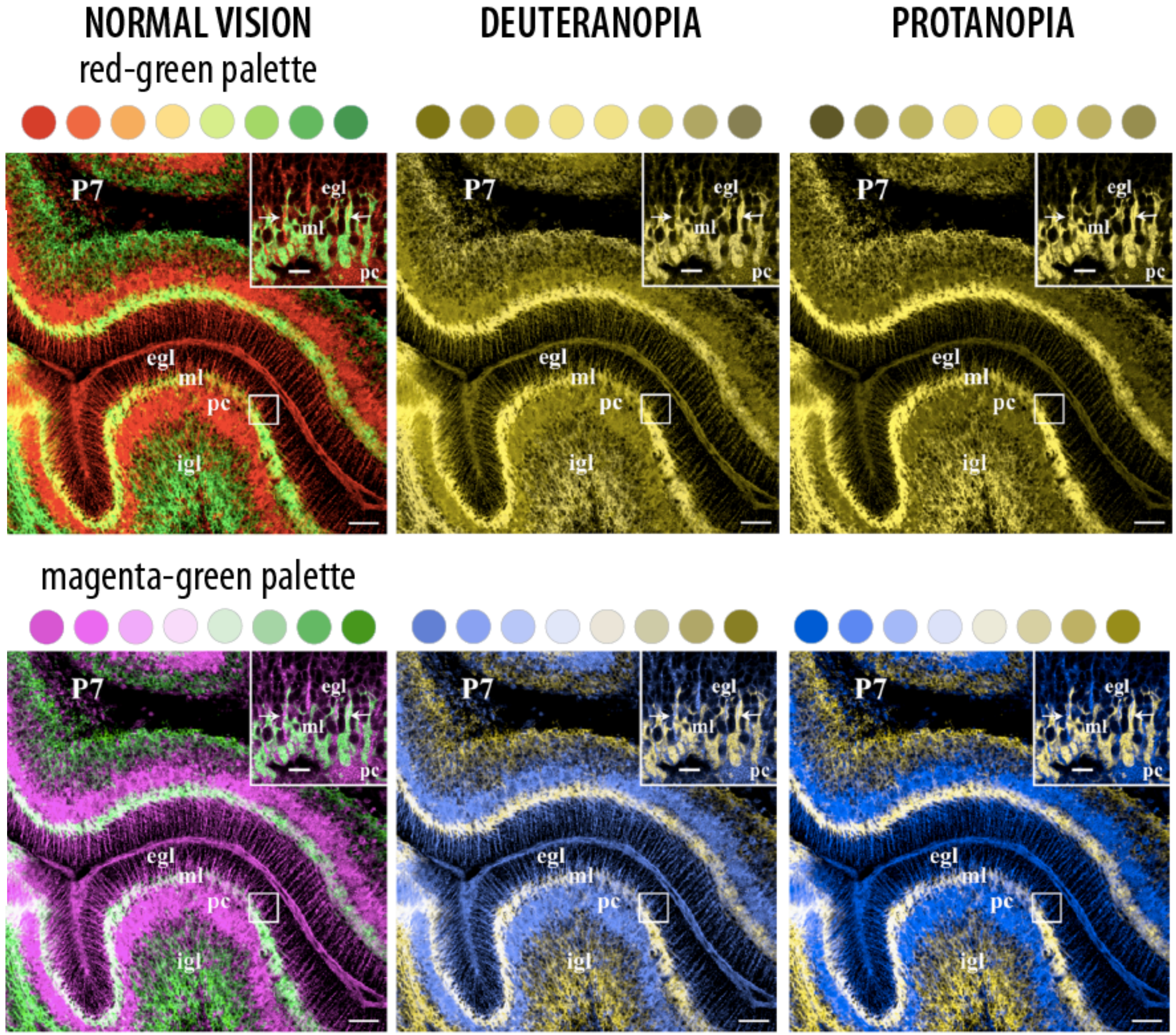
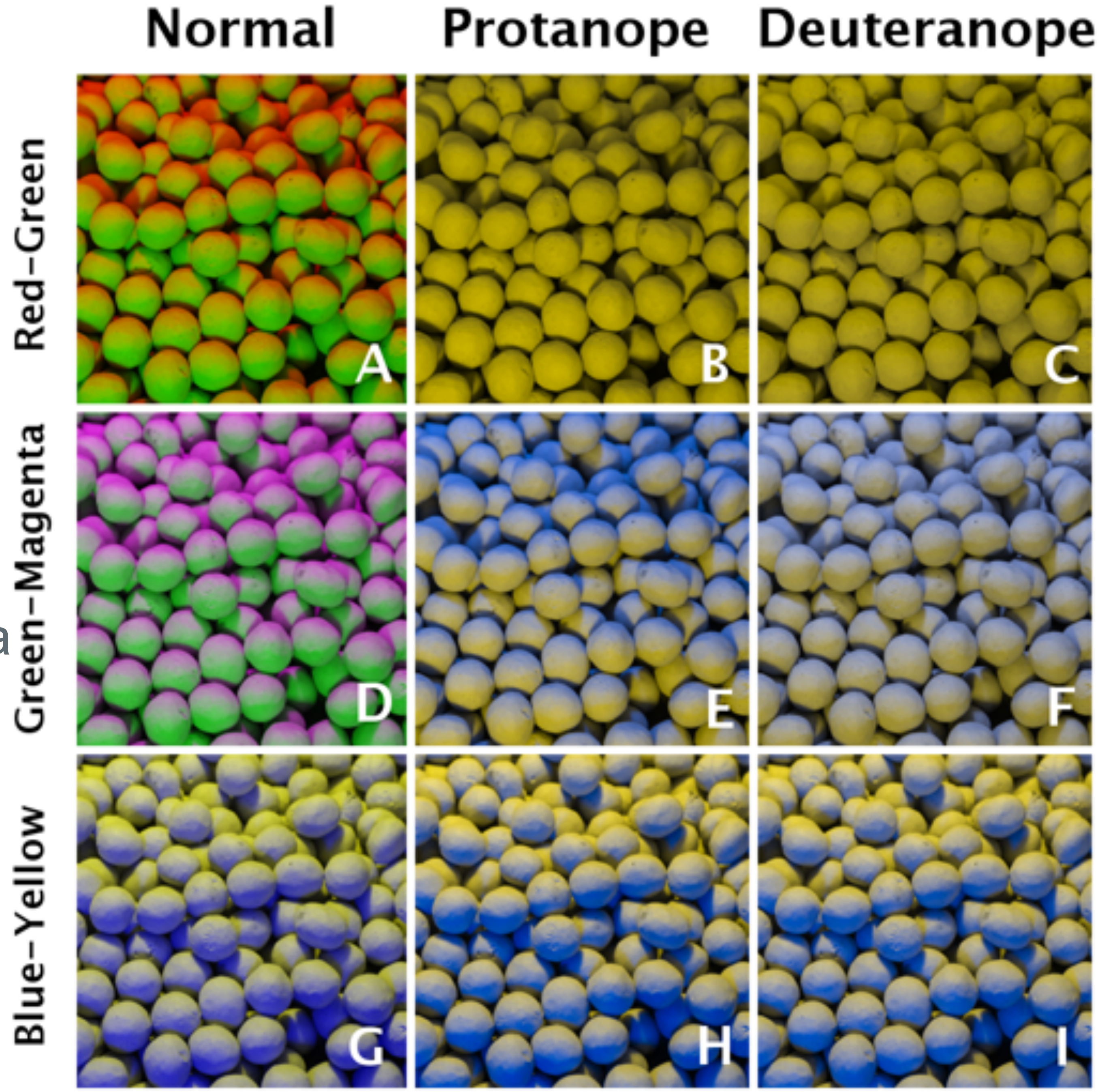


Fig. Courtesy of M Krzywinski

Shaping our
digital future

- Blue-Yellow
  - might be better than
- Green-Magenta
  - talk about same colours



Gabriel Landini & D Giles Perryer, Image recolouring for colour blind observers

From Data to Visualization…

- The properties of the data or information
- The properties of the image
- The rules mapping data to images

Encoding Schemes

Adapted from Mackinlay J (1986) Automating the design of graphical presentations of relational information.

Mapping data types to encoding

Quantitative	Ordinal	Nominal
Position	Position	Position
Length	Density	Hue
Angle	Saturation	Texture
Slope	Hue	Connection
Area	Texture	Containment
Volume	Connection	Density
Density	Containment	Saturation
Saturation	Length	Shape
Hue	Angle	Length
Texture	Slope	Angle
Connection	Area	Slope
Containment	Volume	Area
Shape	Shape	Volume

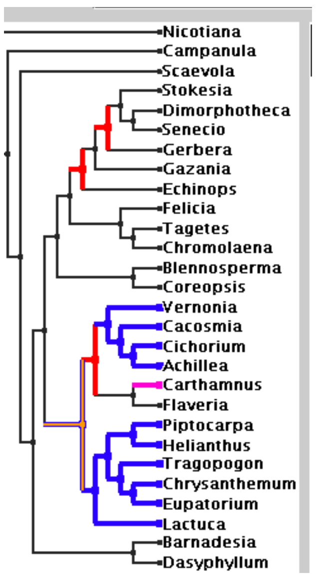
Mackinlay J (1986) Automating the design of graphical presentations of relational information.

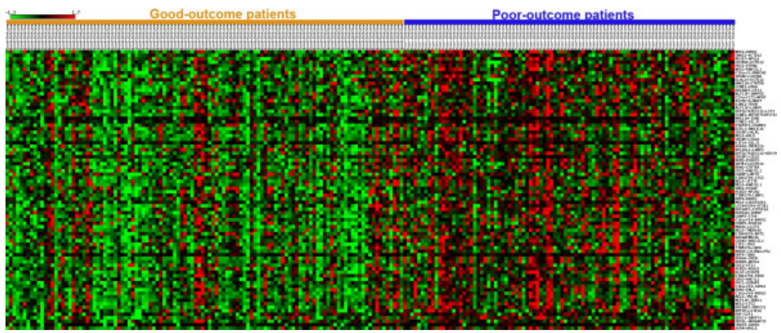
Don’t forget Salience…

- Physical properties that set an object apart from its surroundings
  - Distinct features have high salience
- Encodings have differences in discrimination and accuracy
- Context affects salience
- Choose salient encodings for primary navigation
  - Colour is good for categories - salience decreases with more hues.
- Focus attention by increasing salience of interesting patterns
- Unexpected or bad things can happen when unimportant elements in a figure are salient.
  - The reader will use salience to suggest what is important.

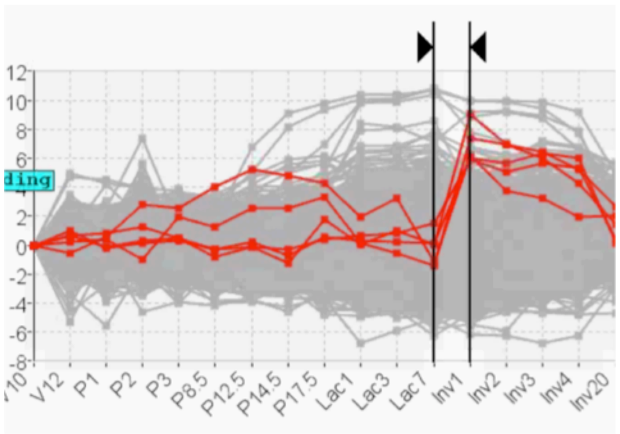
Example Encodings in BioVis…

DNA sequence – 1D, Nominal data, colour





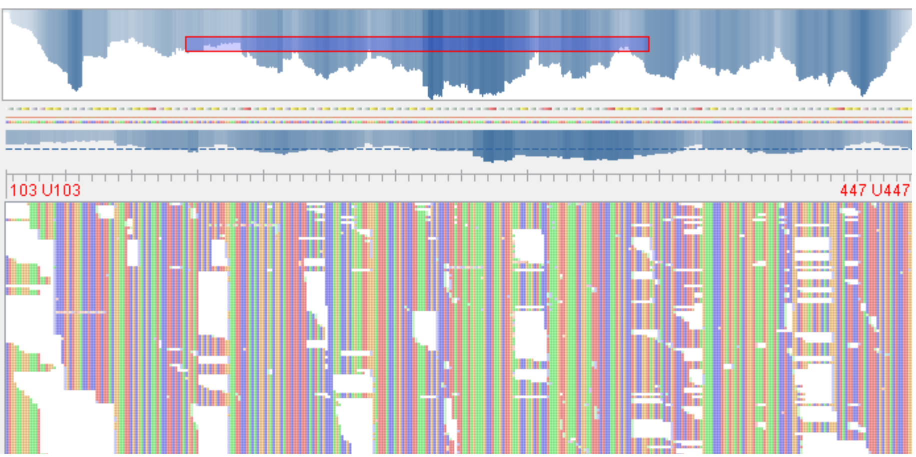
Microarray gene expression –
2D, Ordinal data, colour, position



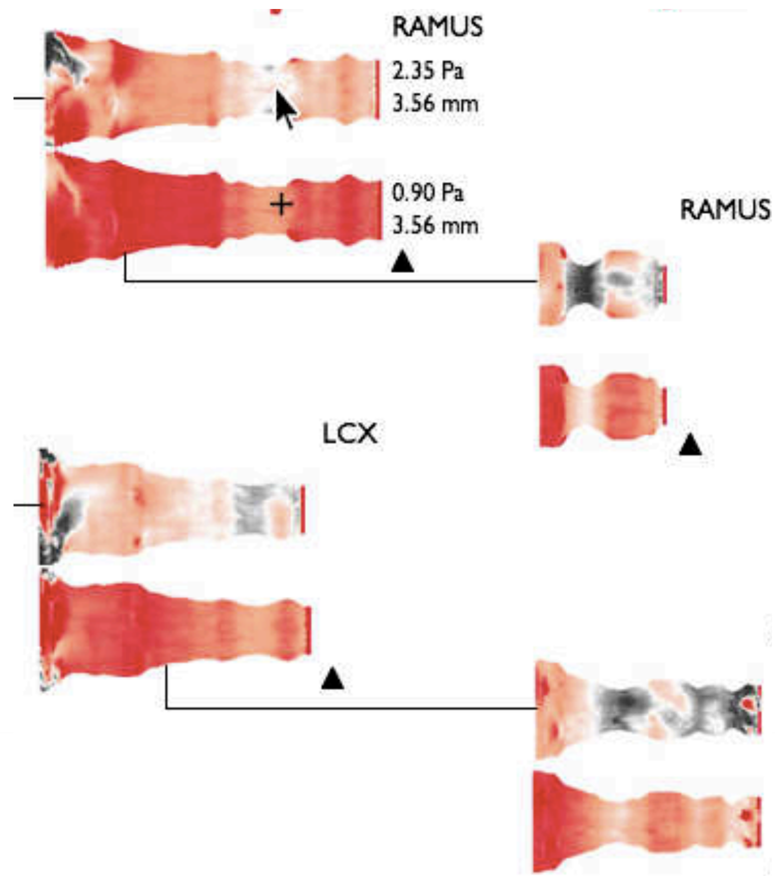
Microarray time-series –
temporal data, quantitative data,
Position, height, colour

Phylogeny –
Tree, Nominal data, position, colour

Examples



Sequence alignments – matrix, colour, position, length



Use of Symmetry, hue, saturation, length

Marshall et al http://bioinf.hutton.ac.uk/tablet/screenshots.shtml

Borkin et al, 2011 Evaluation of artery visualizations for heart disease diagnosis

Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Examples

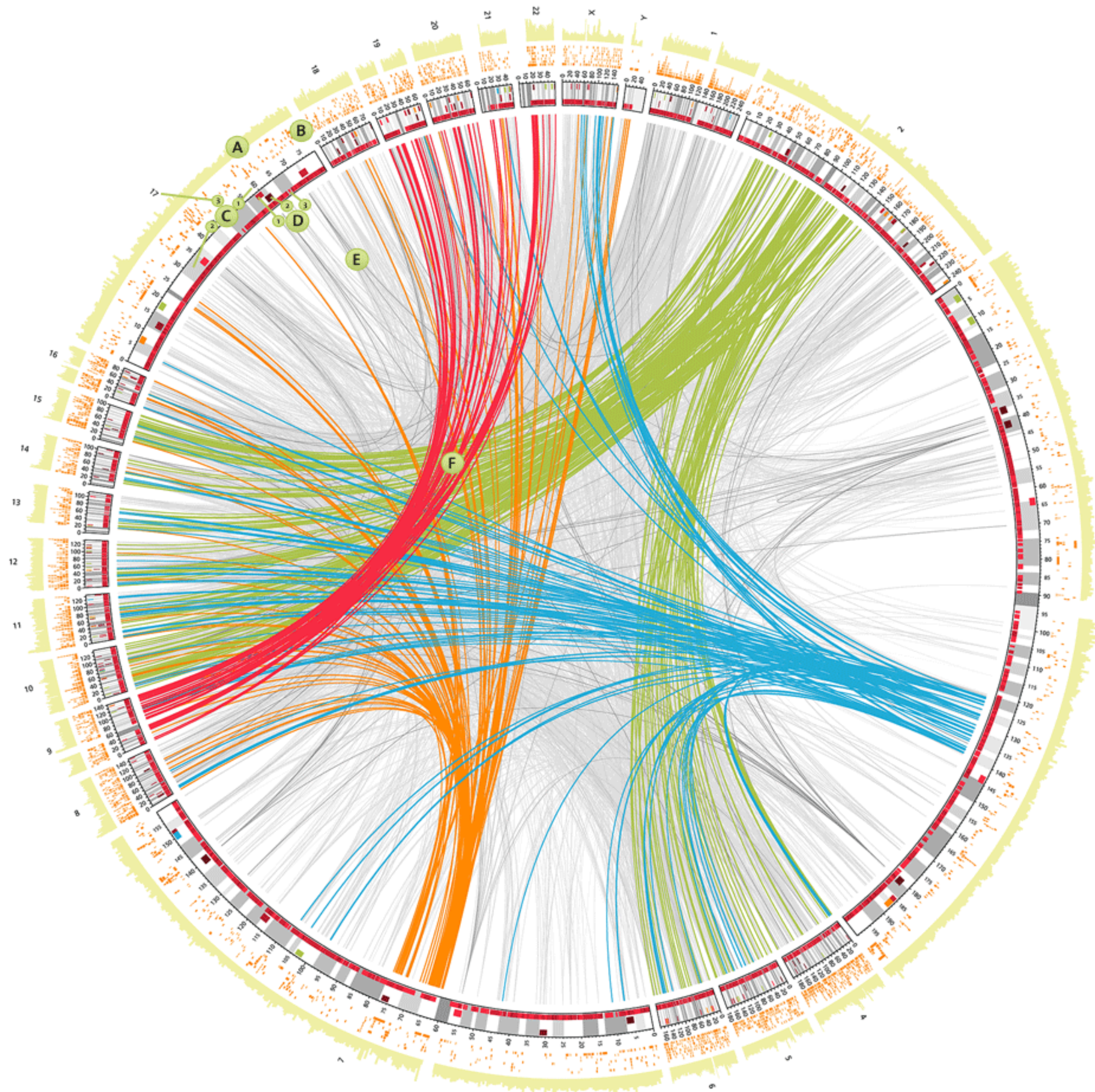
- Circos
  - human genome
  - location of genes implicated in disease
  - regions of

    self-similarity

    structural variation within populations
  - Uses:

    links, heat maps, tiles, histograms

    Use of colour, good continuity, length, transparency, ..



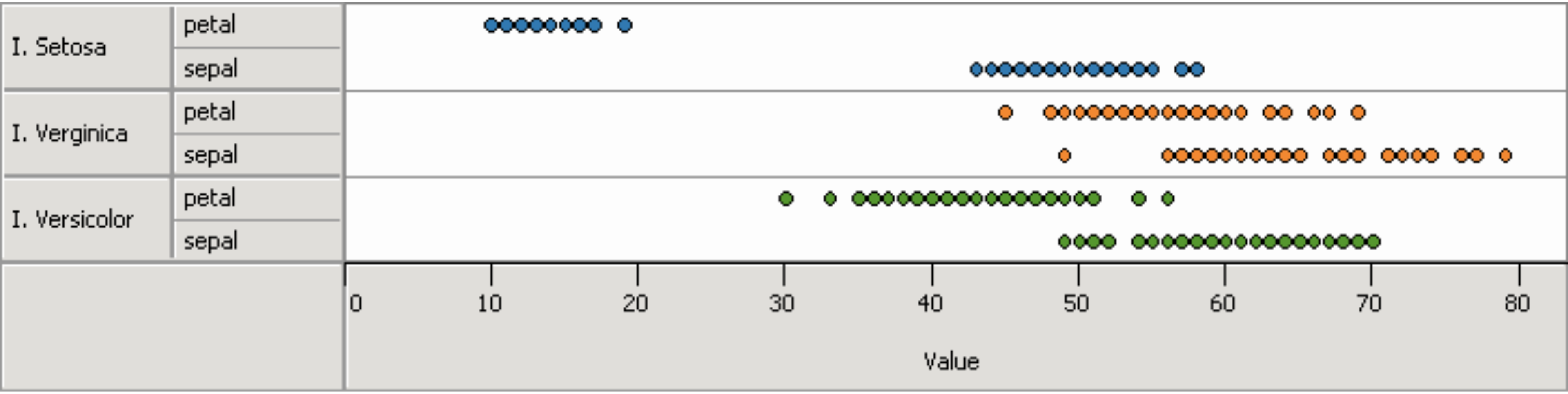
Krzywinski http://circos.ca/intro/genomic_data/

Which Encoding?

- Challenge:
  - Pick the best encoding from the large number of possibilities.
  - Wrong visual encoding can mislead or confuse user
- Visual Representation should be expressive
- Principle of Consistency:
  - The properties of the representation should match the properties of the data.
- Principle of Importance Ordering:
  - Encode the most important information in the most “effective” way.

Expressiveness

- Visual Representation encodes all the facts
  - An nD (or 1:N) data set e.g. Iris data set



  - cannot simply be expressed in a single horizontal dot plot because multiple cases are mapped to the same position

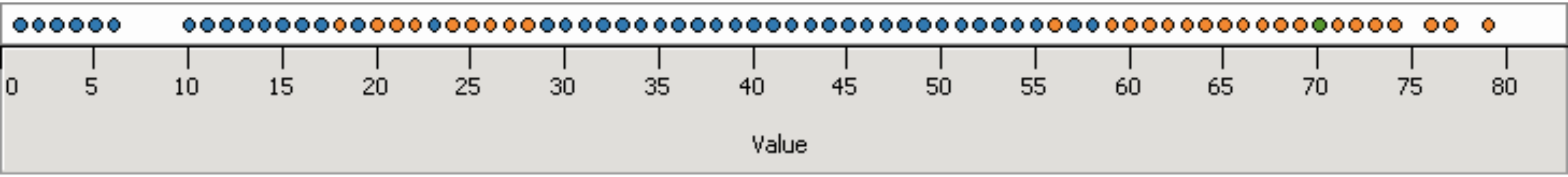
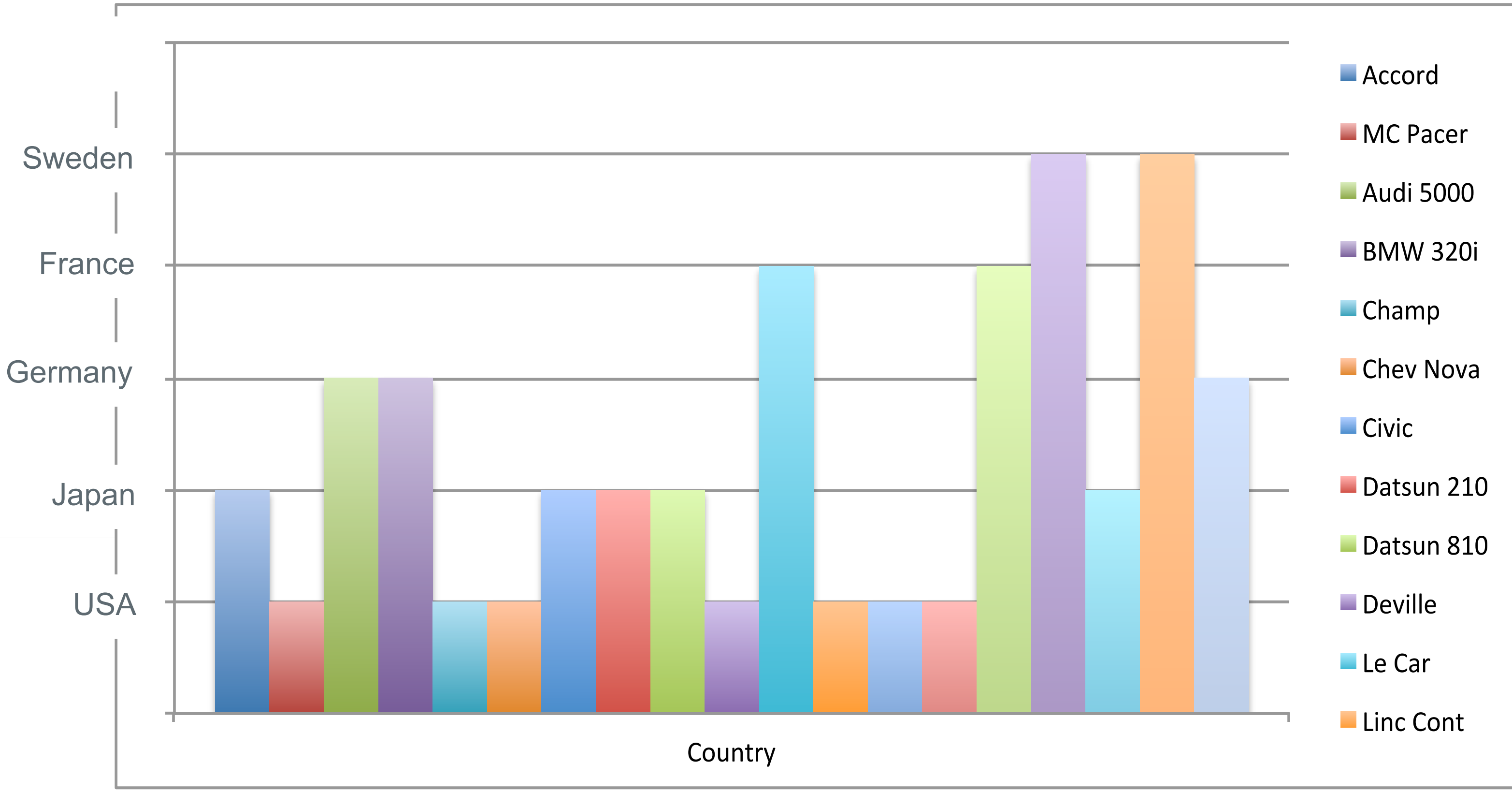


Fig. Courtesy of M Krzywinski

Expressiveness

- Encodes only the facts
  - Wrong use of a bar chart implies something better about Swedish cars than USA ones…



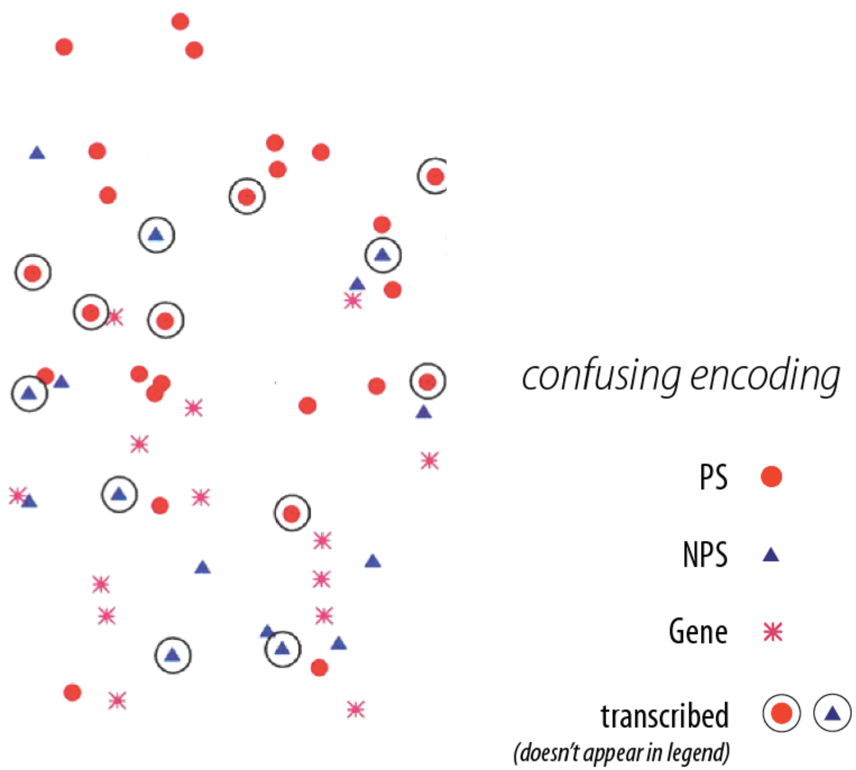
Adapted from Mackinlay J (1986) Automating the design of graphical presentations of relational information.

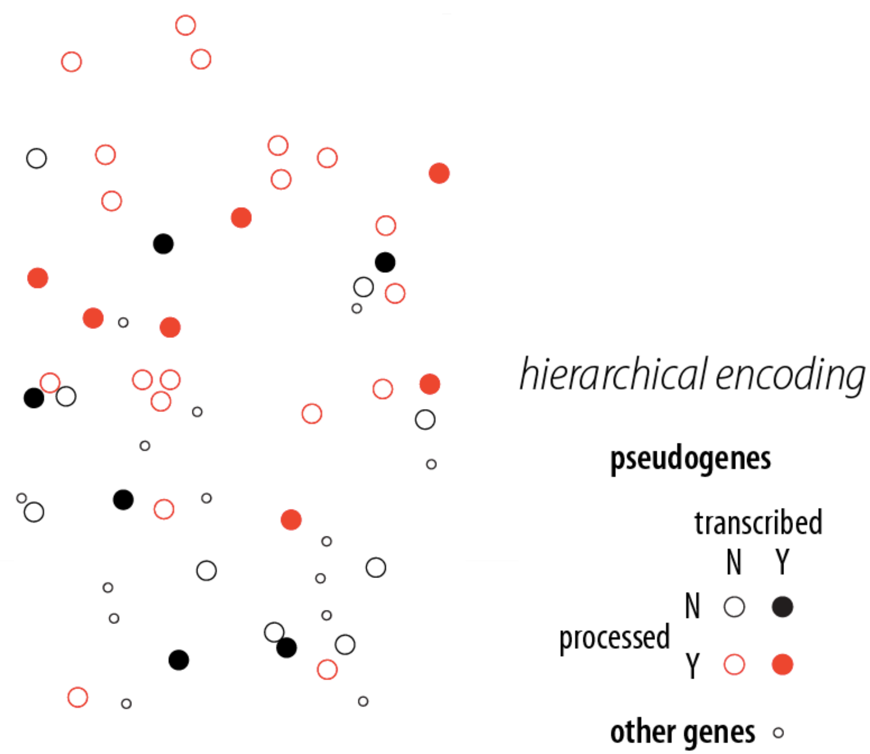
Consistency

- Visual variation in a figure should always reflect and enhance any underlying variation in the data.
- Avoid using more than one encoding to communicate the same information.
- Do not use visually similar encodings for independent variables

Consistency

- red - processed genes, but salience attenuated
  - other genes encoded with competing glyph - red star.

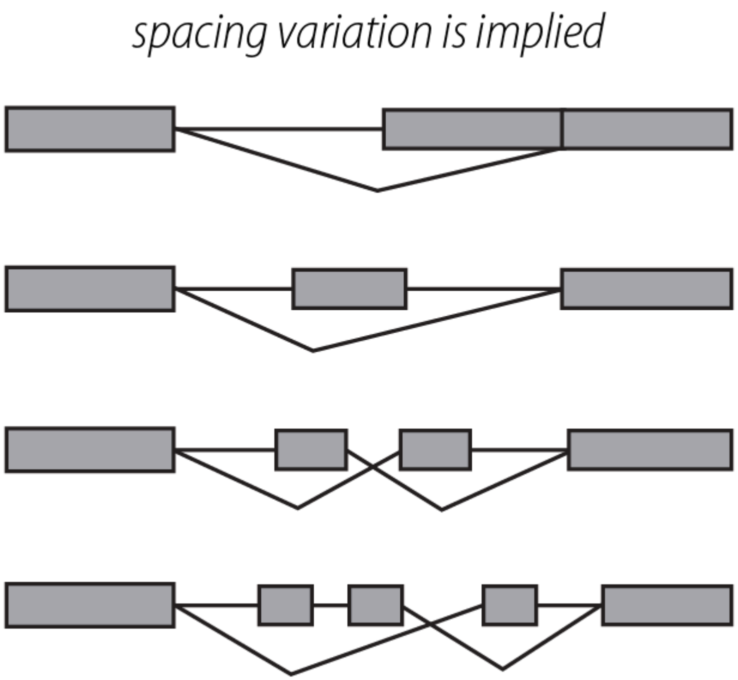


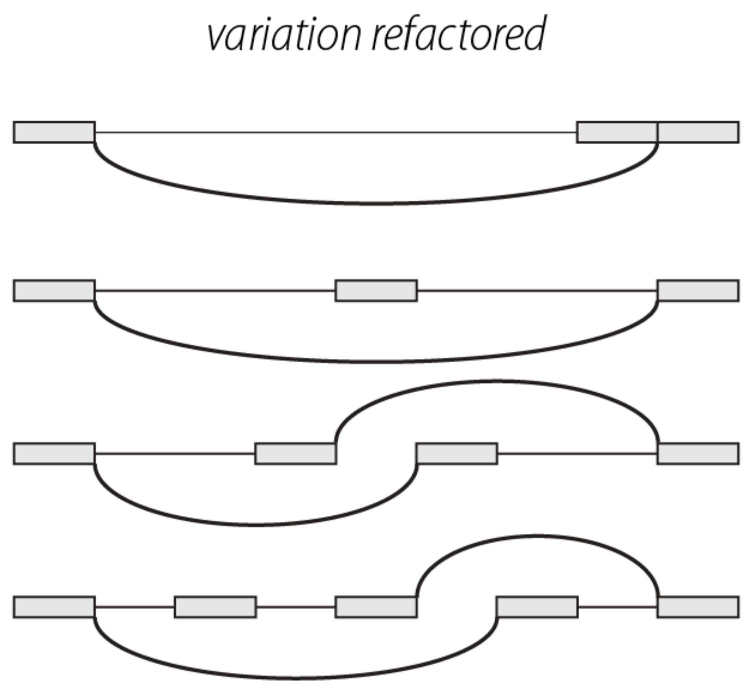


M. Krzwinski, behind every great visualization is a design principle, 2012

Consistency

- Uniform size and alignment of exons and introns reduces complexity and aids interpreting their complex arrangement.



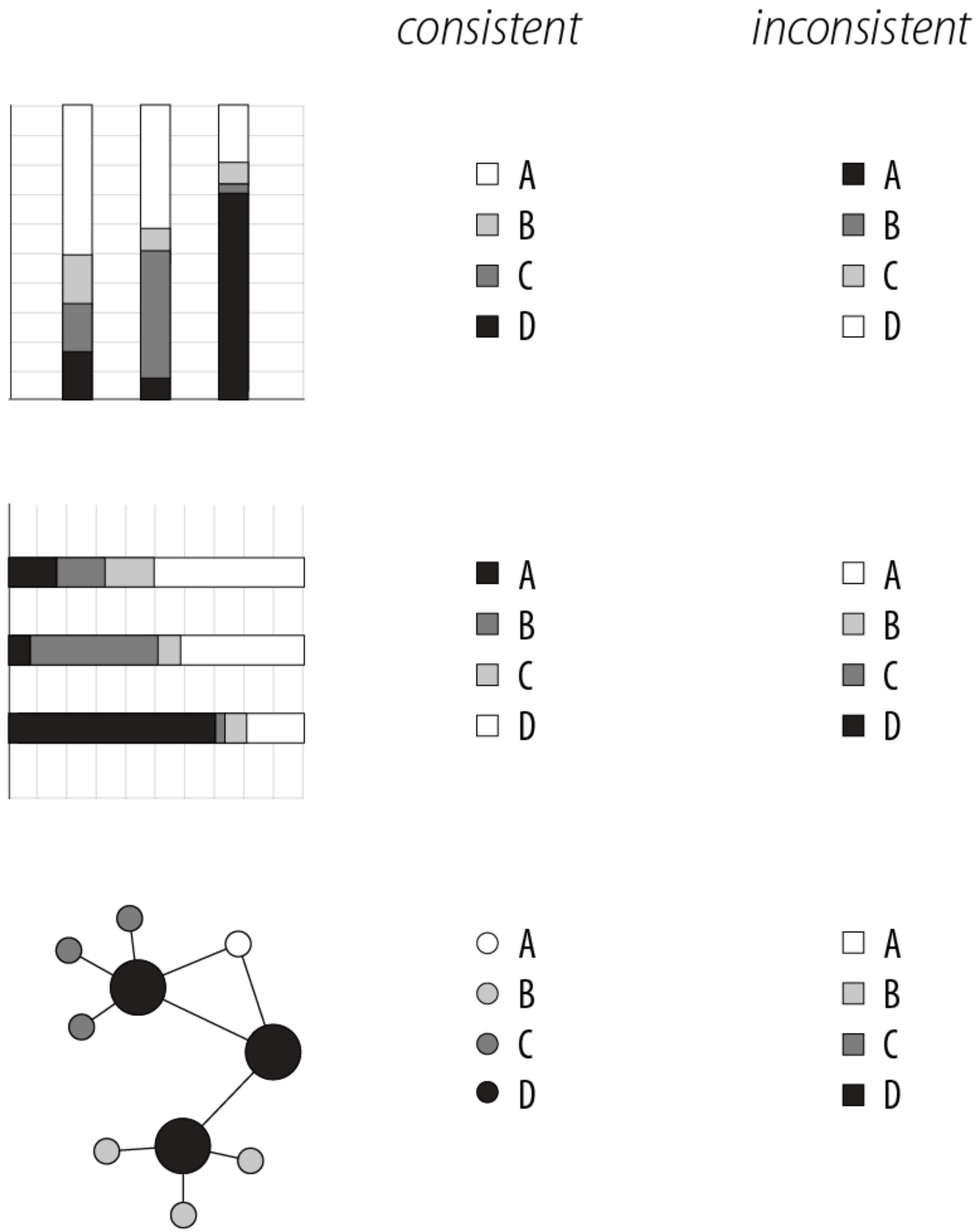


Sharov AA, Dudekula DB, Ko MS (2005) Genome-wide assembly and analysis of alternative transcripts in mouse. Genome Res 15: 748-754.

M. Krzwinski, behind every great visualization is a design principle, 2012

Consistency

- Order items in a legend according to order of appearance in the plot



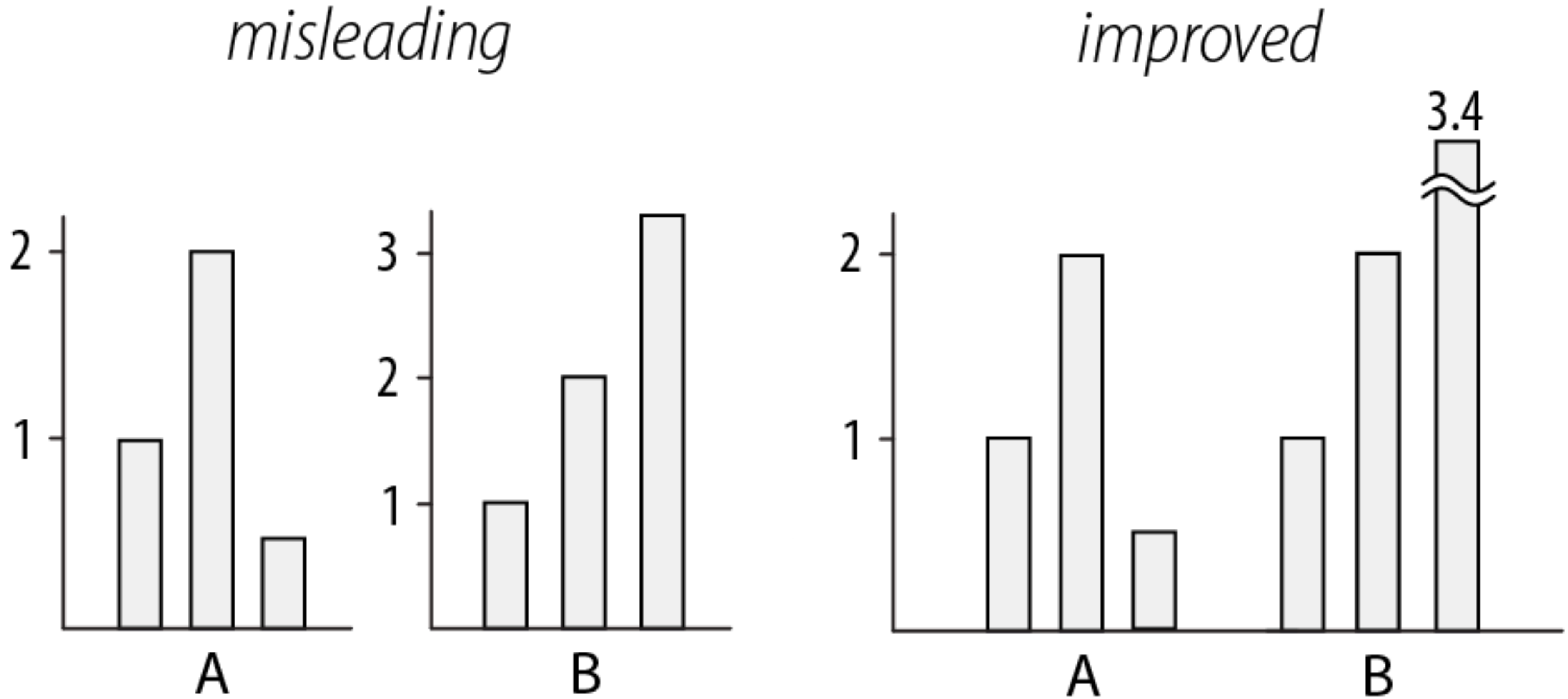
M. Krzwinski, behind every great visualization is a design principle, 2012

Shaping our digital future

Institute for Informatics & Digital Innovation

Consistency - Navigational aids

- Use consistent axes when comparing charts

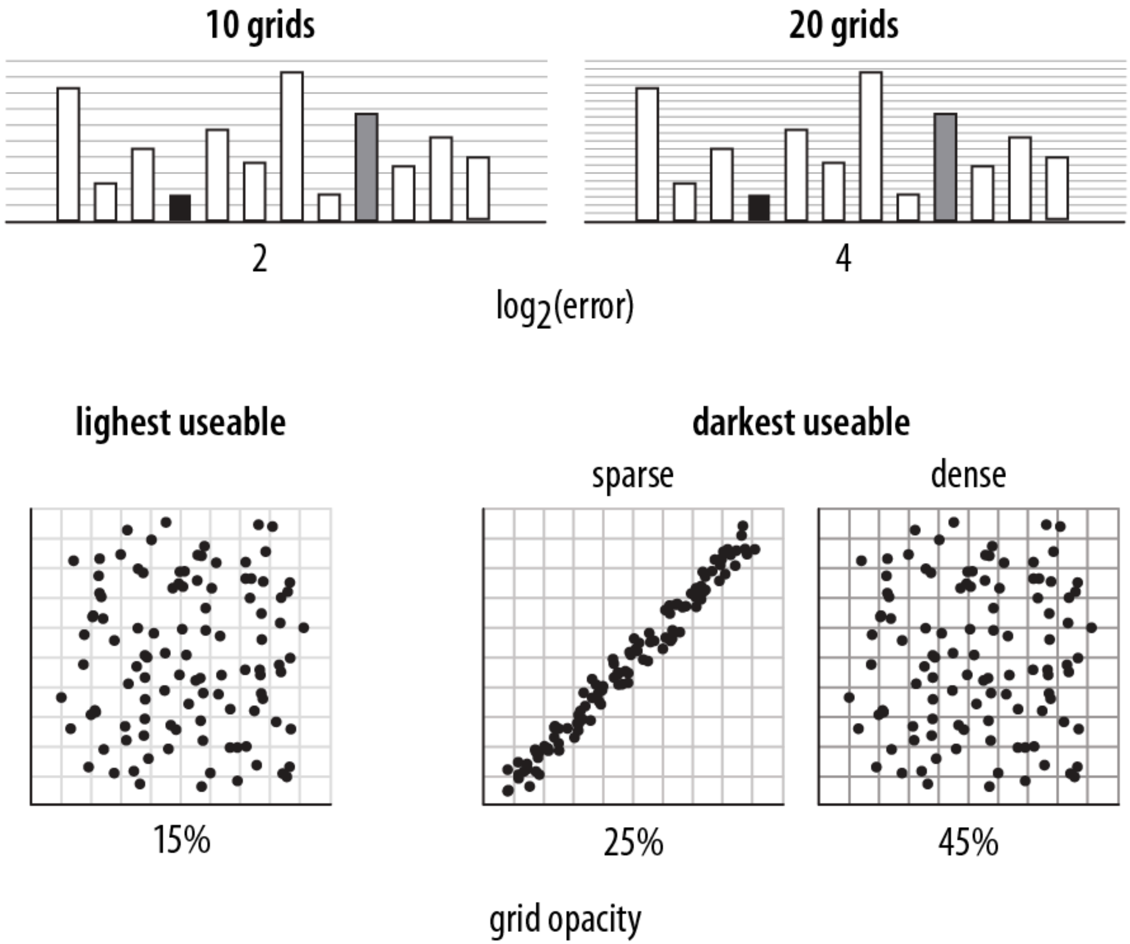


Raina SZ, et al. (2005) Evolution of base-substitution gradients in primate mitochondrial genomes. Genome Res 15: 665-673.

M. Krzwinski, behind every great visualization is a design principle, 2012

Increase data:ink ratio

- Navigational aids
  - should not compete with the data for salience.
- Avoid
  - heavy axes,
  - error bars and
  - glyphs



Heer J, Bostock M (2010) Crowdsourcing graphical perception: using mechanical turk to assess visualization design. Proceedings of the 28th international conference on Human factors in computing systems. Atlanta, Georgia, USA: ACM. pp. 203-212.

Increase data:ink ratio

- Avoid unnecessary containment

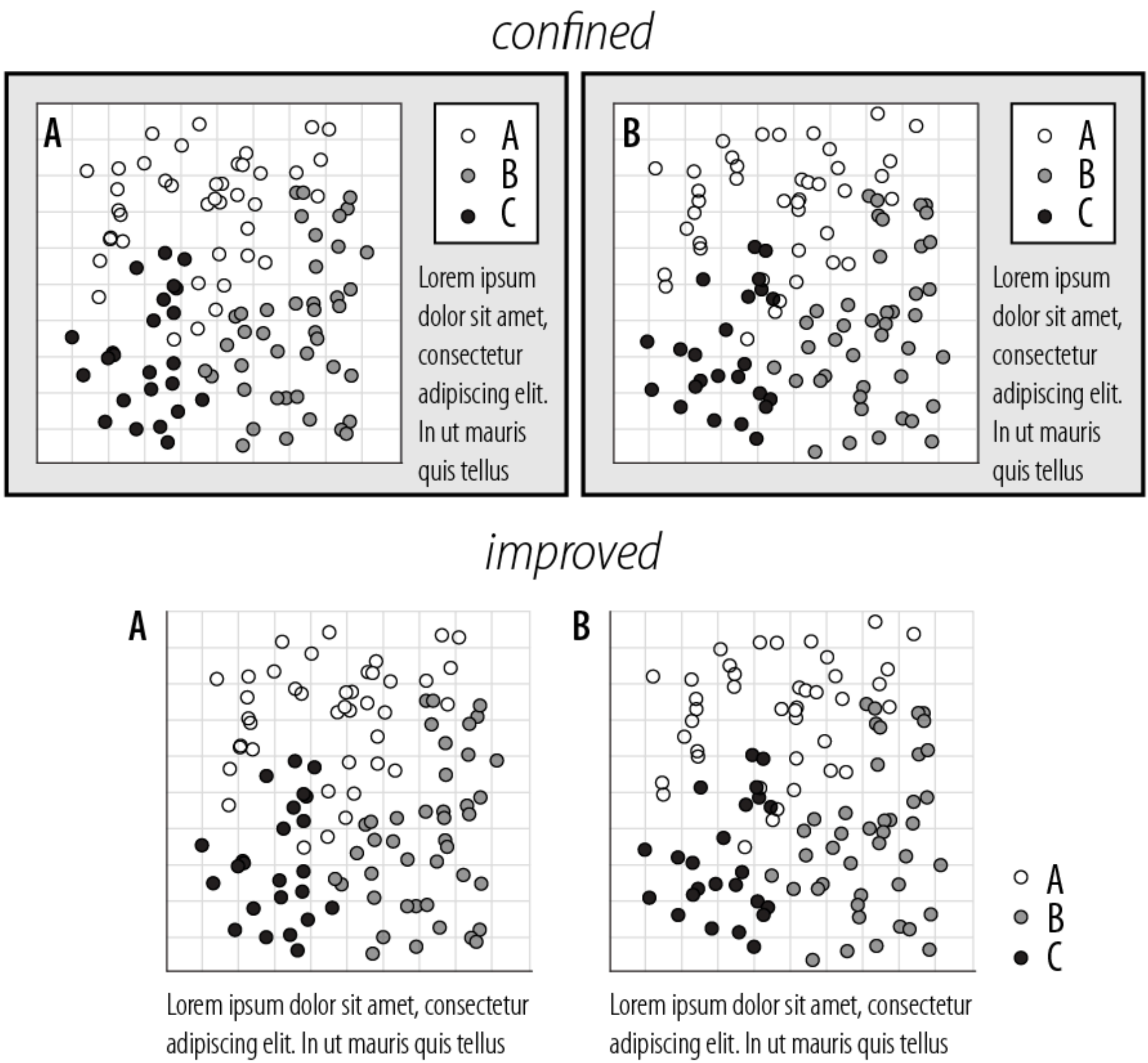
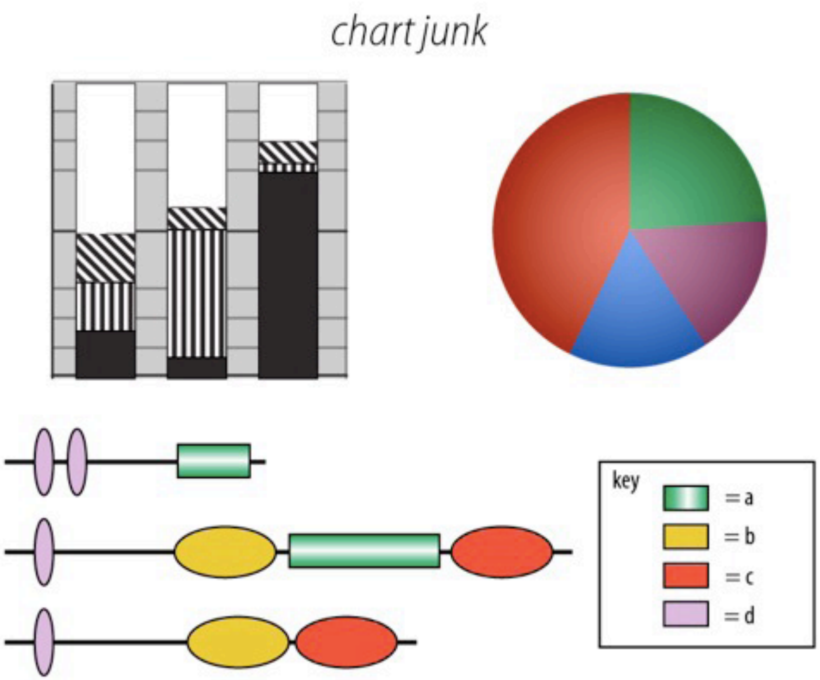
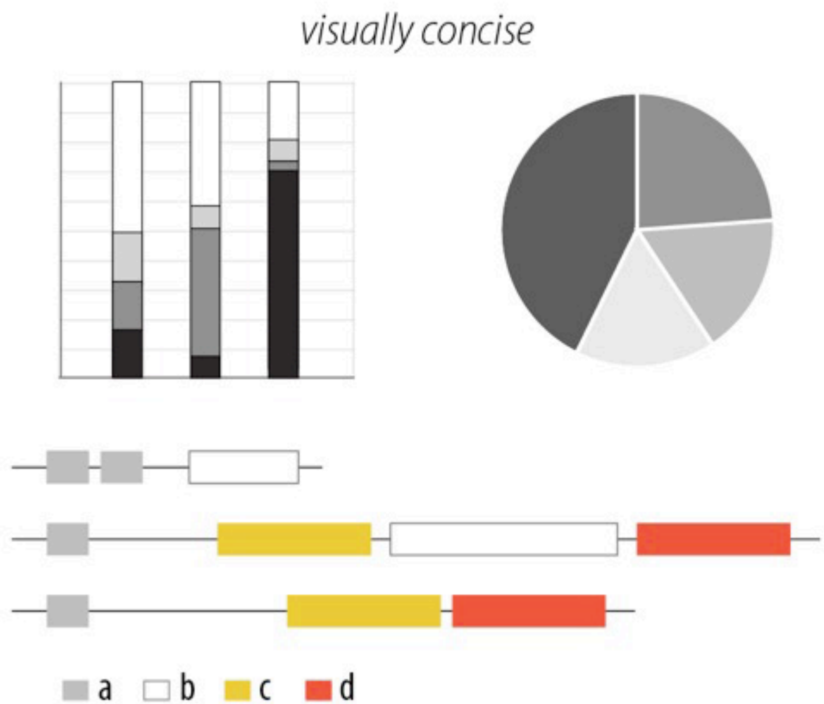


Fig. Courtesy of M Krzywinski

Increase data:ink ratio

- Avoid “Chart junk”





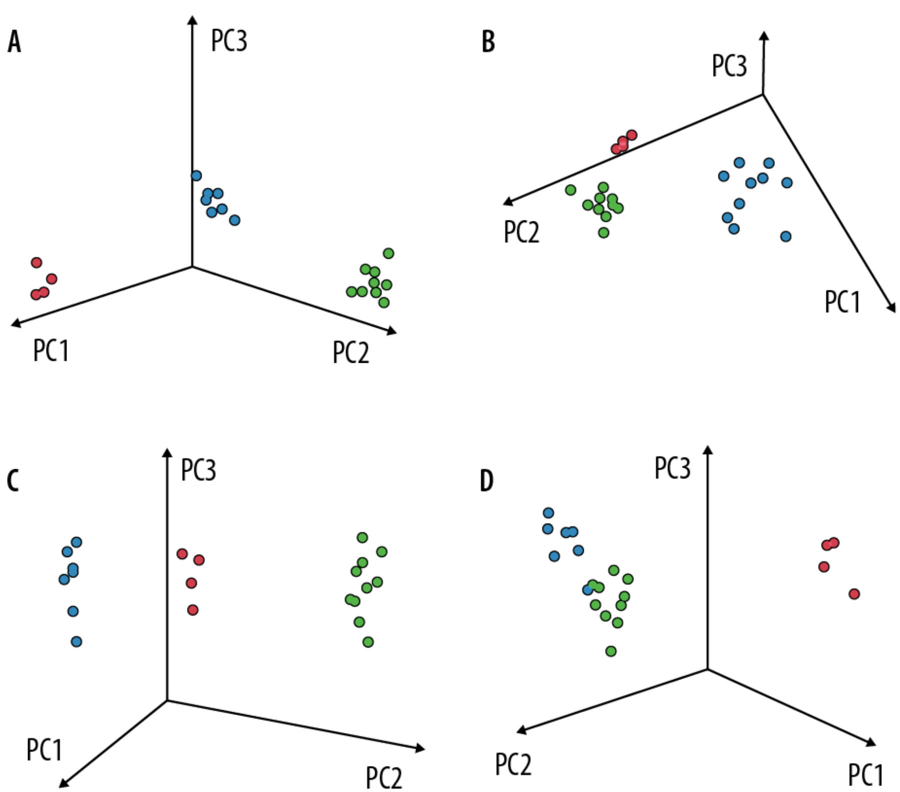
Sharov AA, et al (2006) Genome Res 16: 505-509.
Peterson J, et al. (2009) Genome Res 19: 2317-2323.
Thomson NR, et al. (2005) Genome Res 15: 629-640.
DB, Ko MS (2005) Genome Res 15: 748-754.

M. Krzwinski, behind every great visualization is a design principle, 2012

Keep things simple - Avoid 3D

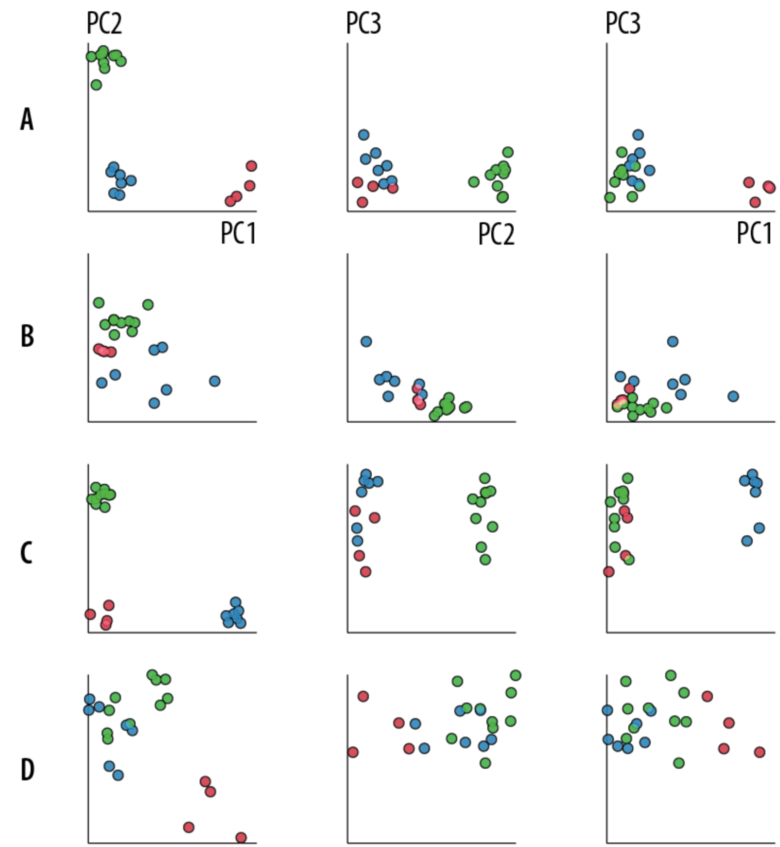
- 3D scatter plots are better as a series of 2D projections.

confusing



Son CG, et al. (2005) Database of mRNA gene expression profiles of multiple human organs. Genome Res 15: 443-450.

improved



M. Krzwinski, behind every great visualization is a design principle, 2012

Beyond Basic Design: Interaction

- The potential to overcome well known problems with static imagery....

Change Blindness….

SPOT THE DIFFERENCE

Fig. courtesy of S Rogers, ONS, UK

Interaction

- Supports the user in exploring data

- Shneiderman's Information Seeking Mantra:
  - Overview first, zoom and filter, then details on demand

Interaction: operations on the data

- sorting
- filtering
- browsing / exploring
- comparison
- characterizing trends and distributions
- finding anomalies and outliers
- finding correlation
- following path

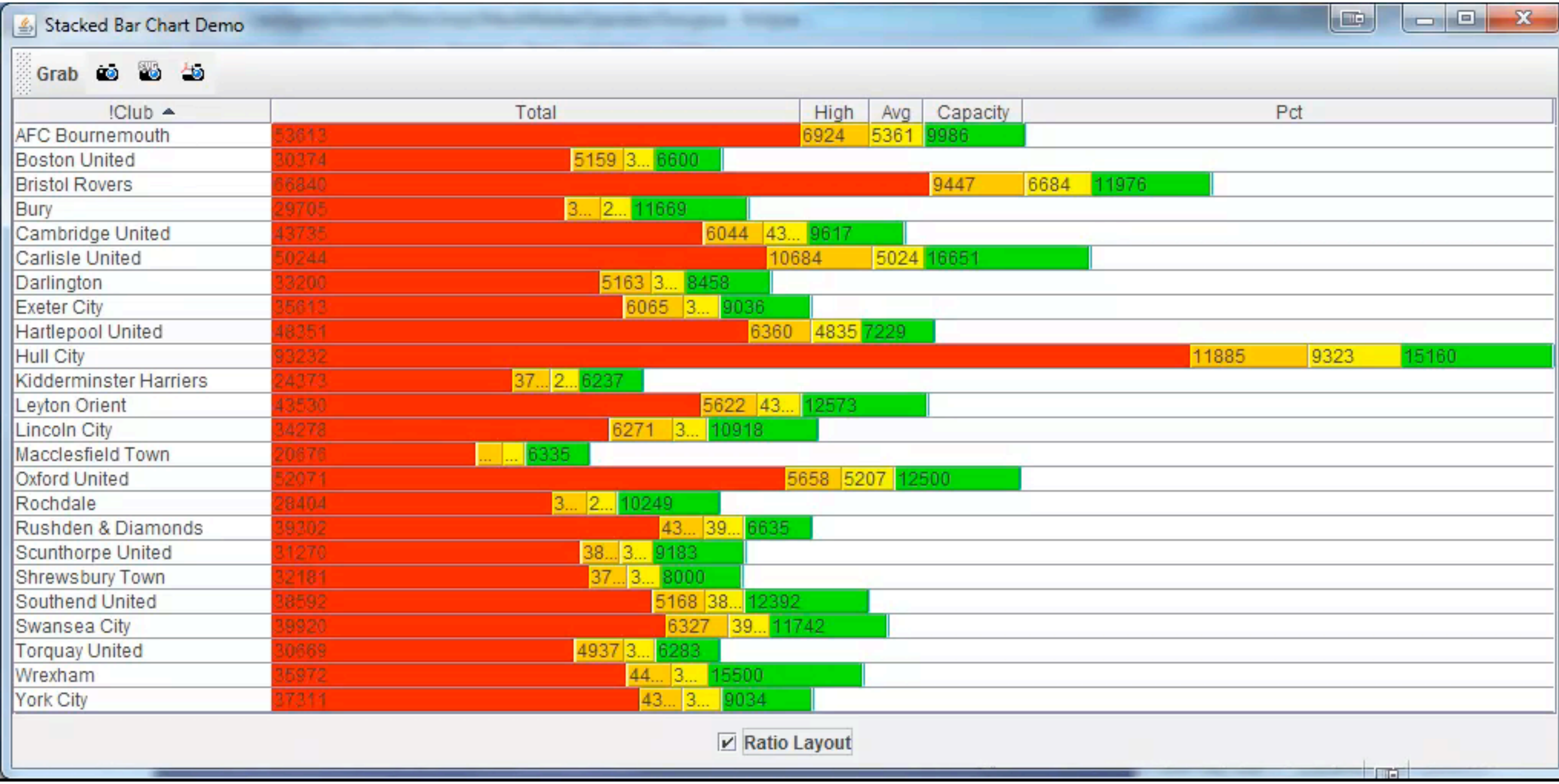
Interaction: Techniques to support operations

- Re-orderable matrices - sorting
- Brushing - browsing
- Linked views – comparison, correlation, different perspectives
  - Linking
- Overview and detail -
  - Excentric labelling
- Zooming – dealing with complexity/amount of data
- Focus & context - dealing with complexity/amount of data
  - Fisheye….
  - Hyperbolic
- Animated transitions - keeping context
- Dynamic queries - exploring

Shaping our
digital future

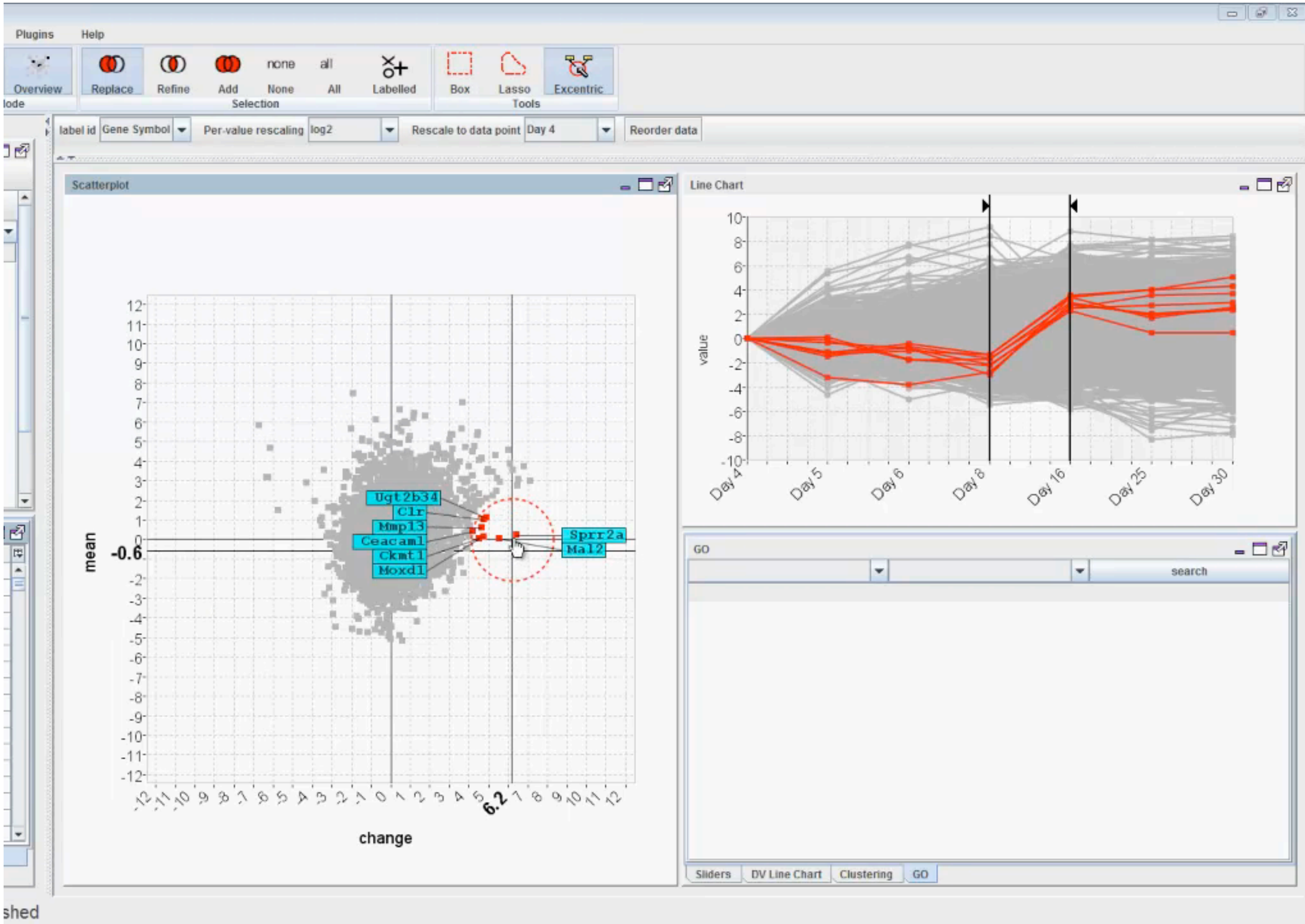
Institute for
Informatics
& Digital
Innovation

Sortable tables, matrices



Linked Views, Linking, Brushing, Excentric Labels

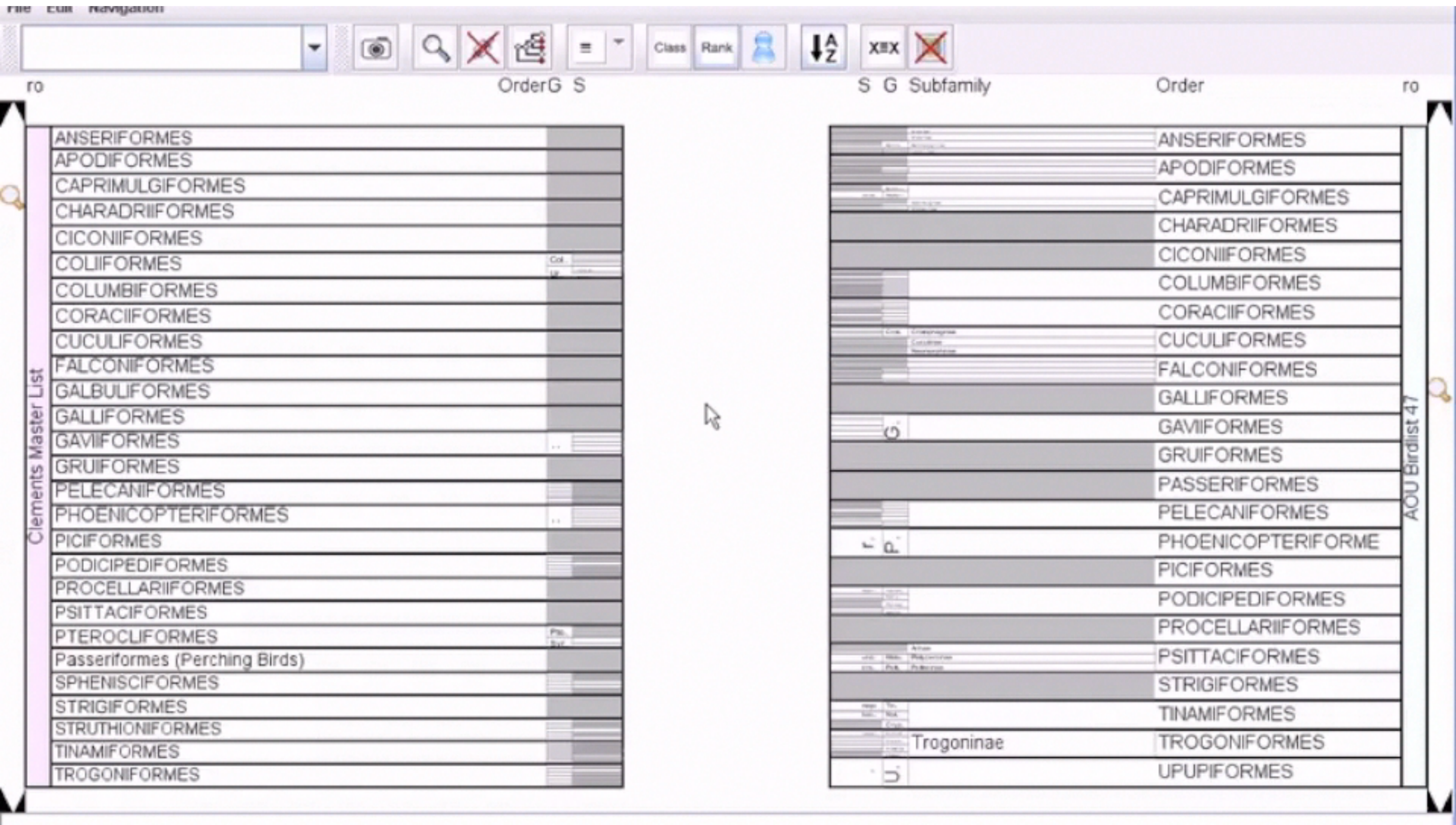
- Time-series Microarray Data



Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

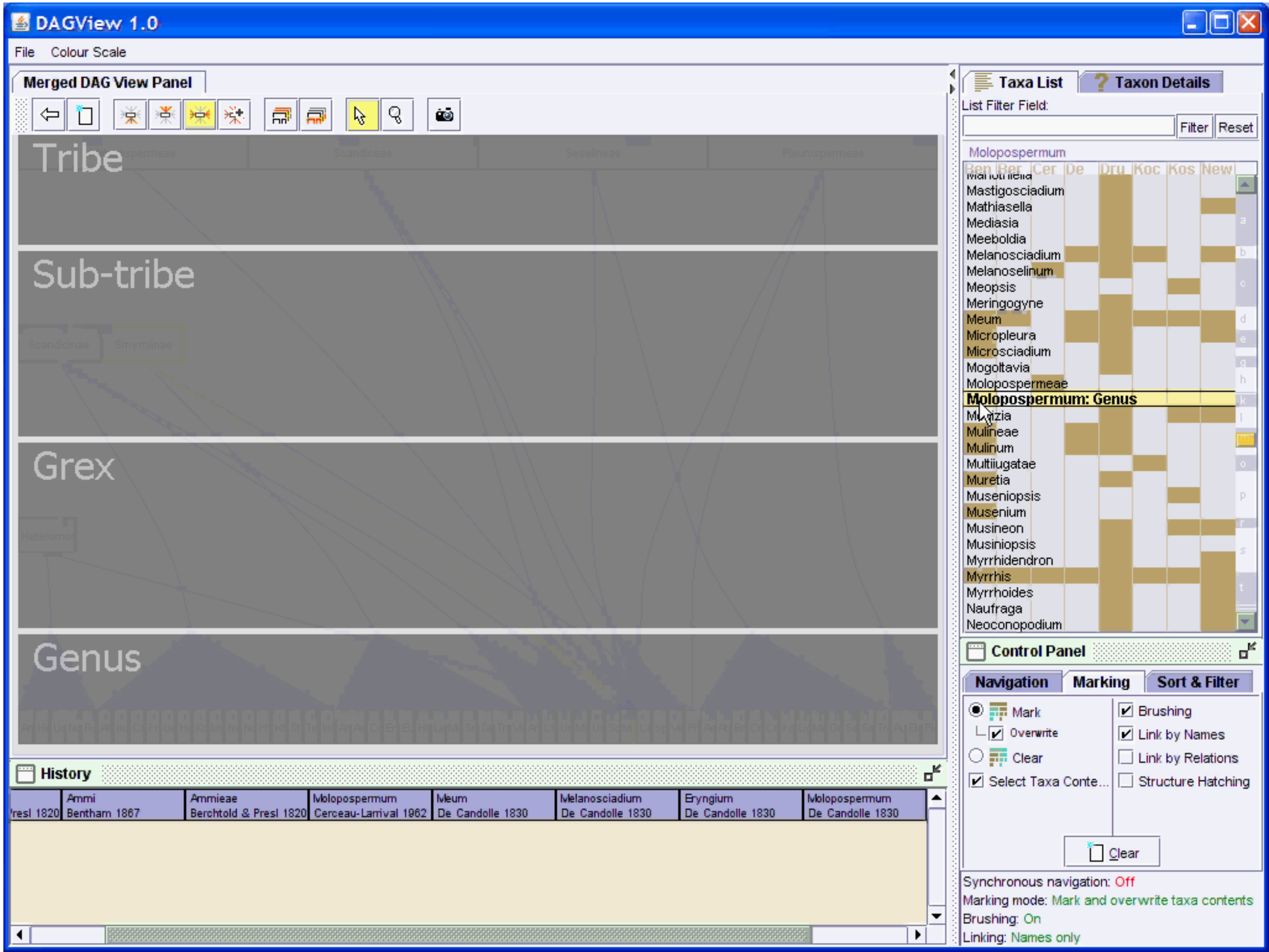
Focus and Context – Fisheye tree



Shaping our digital future

Institute for Informatics & Digital Innovation

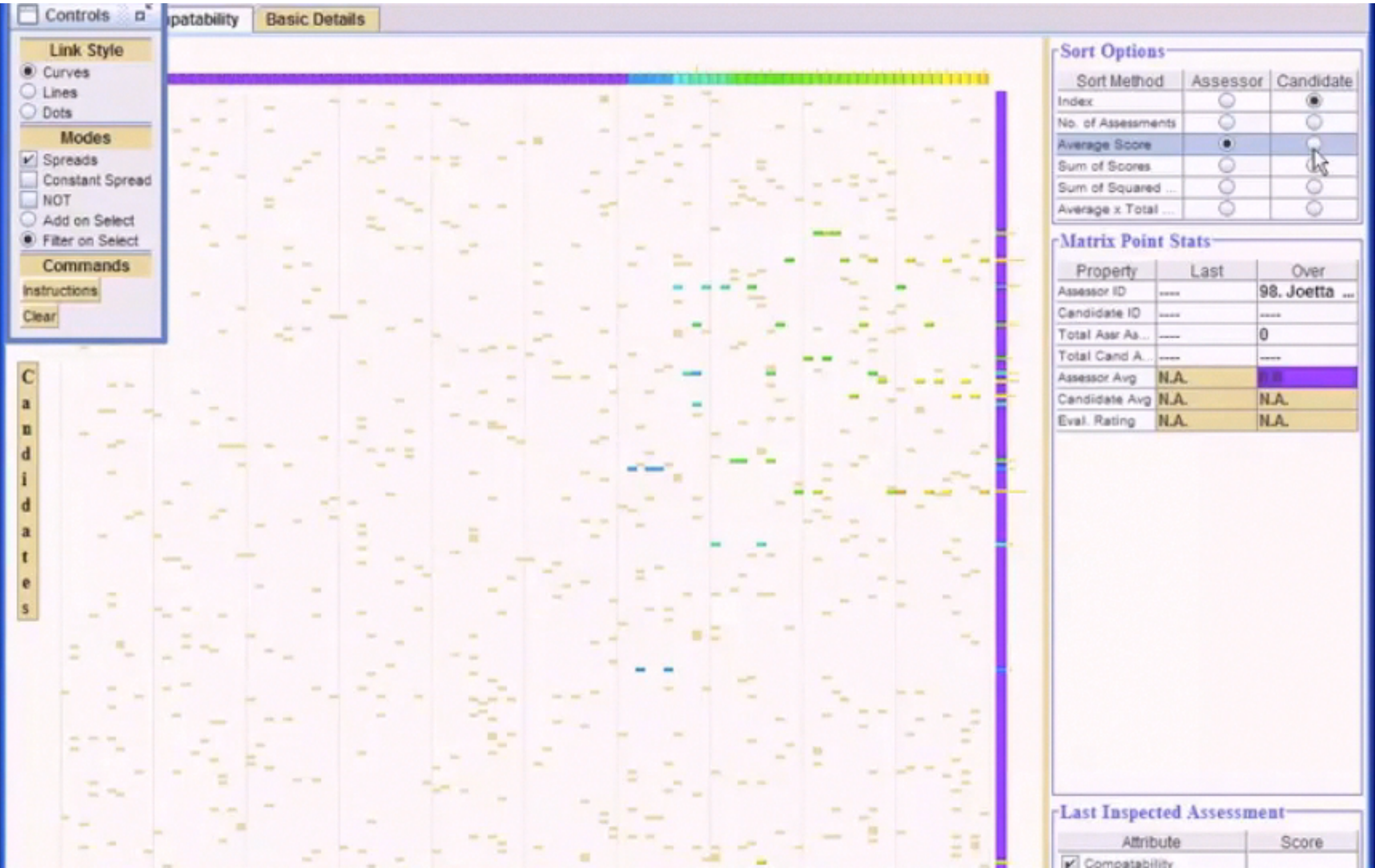
Animated transitions, brushing



Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

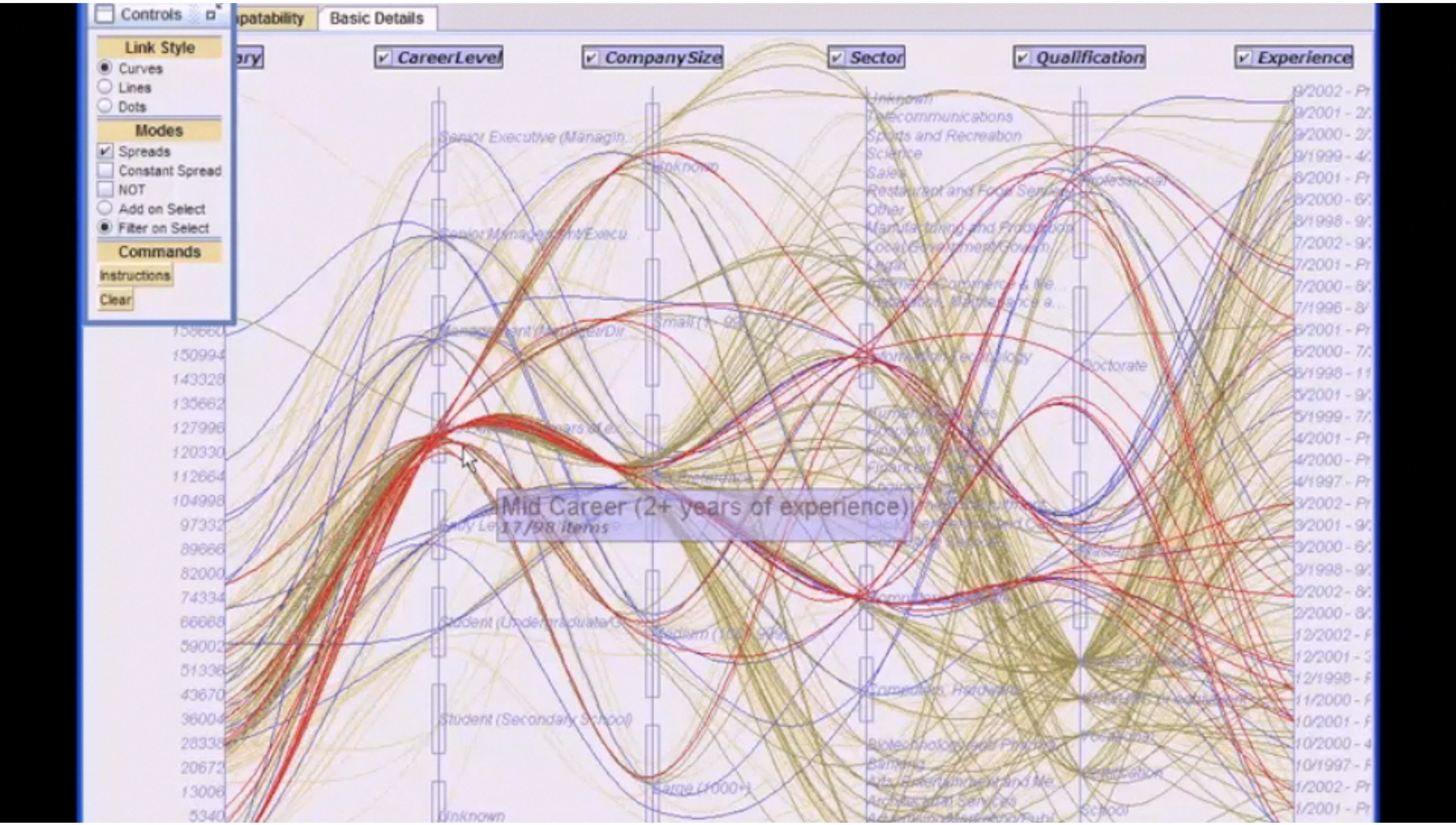
Drill Down, Select, Zooming, Focus & Context



Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Brushing, Filtering, Principle of Continuity



Shaping our
digital future

Institute for Informatics & Digital Innovation

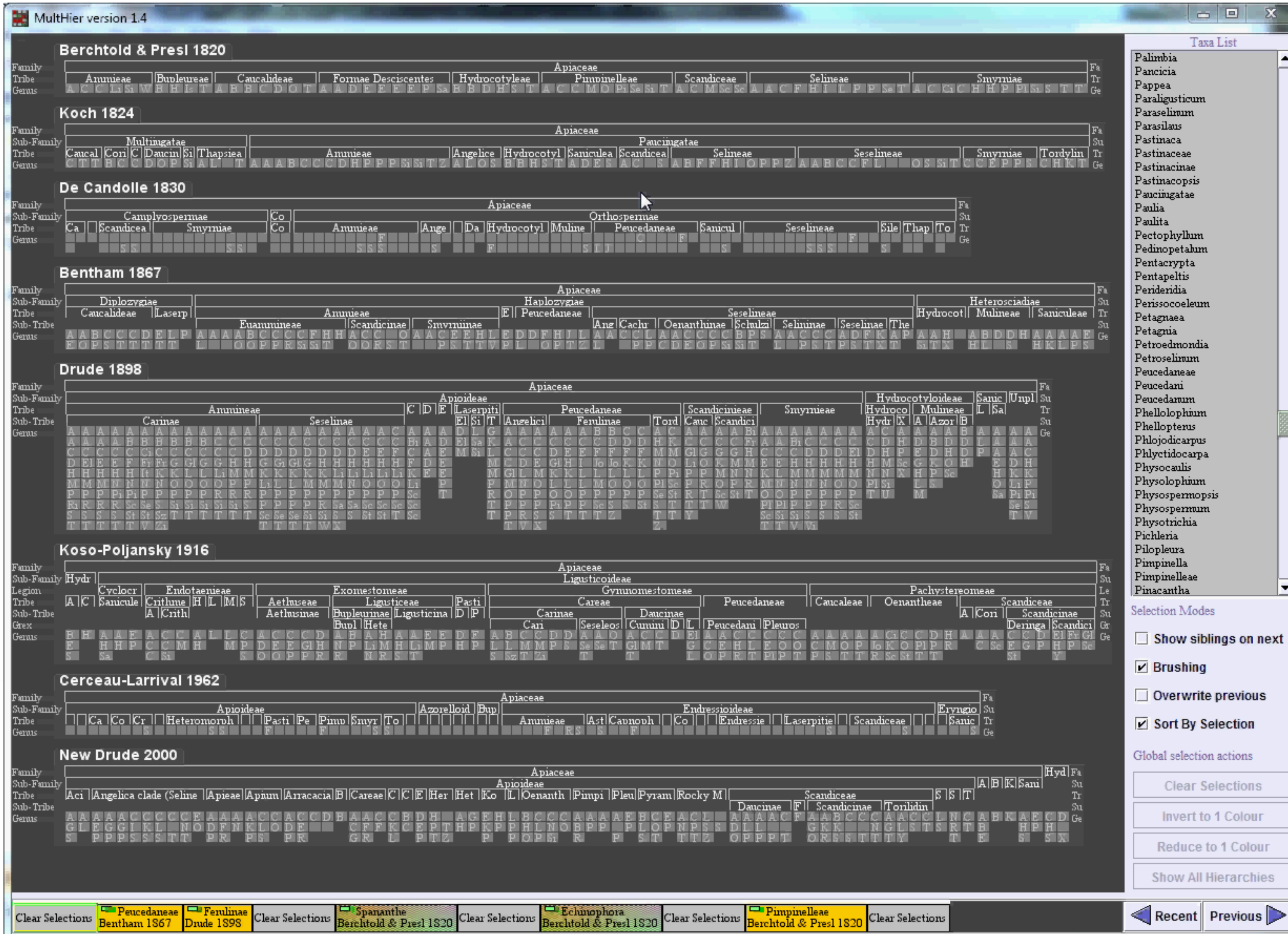
Linked Views, Filtering

Shaping our digital future

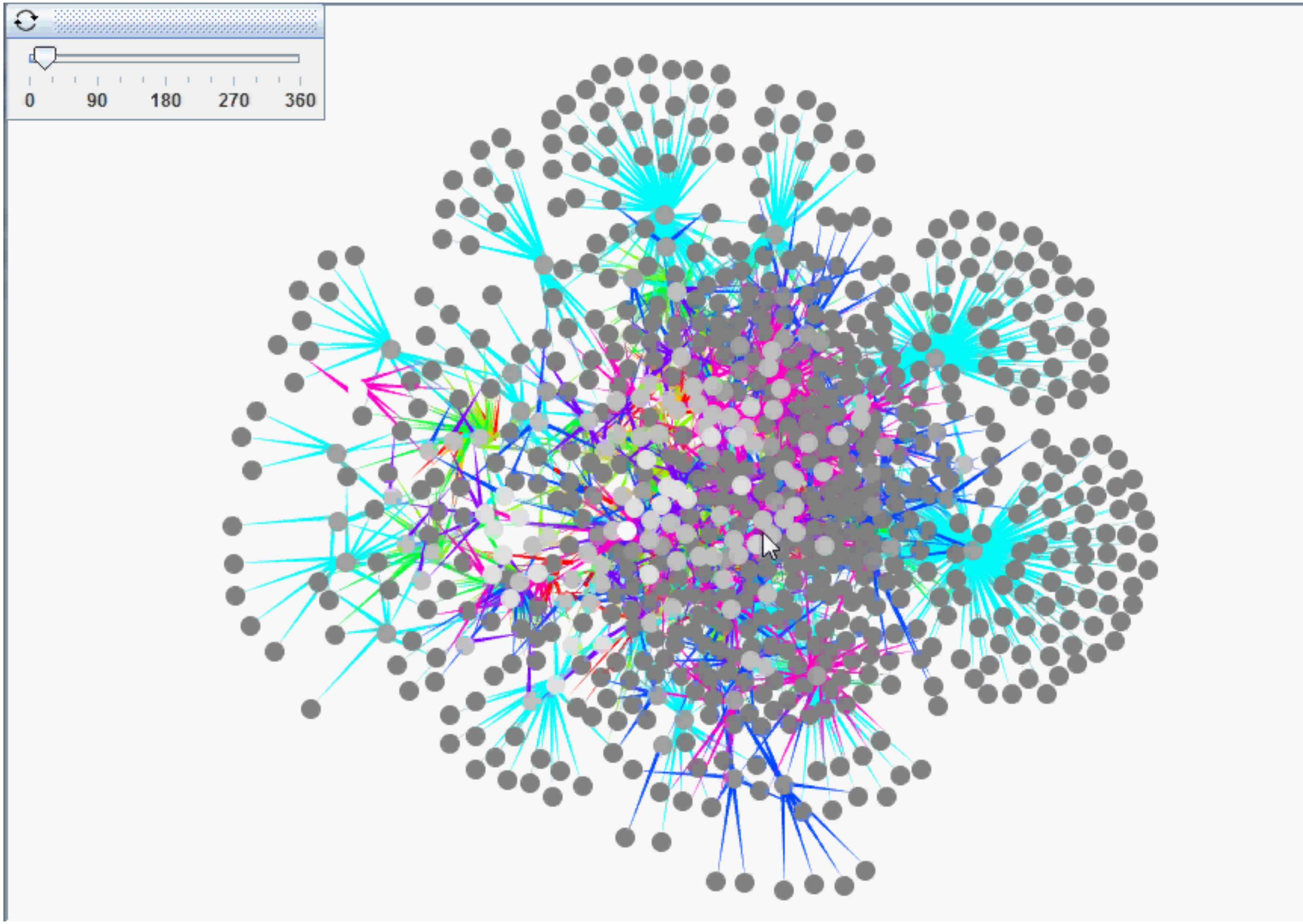
Institute for Informatics & Digital Innovation

Multiple Trees Comparison

- Brushing
- Focus & context



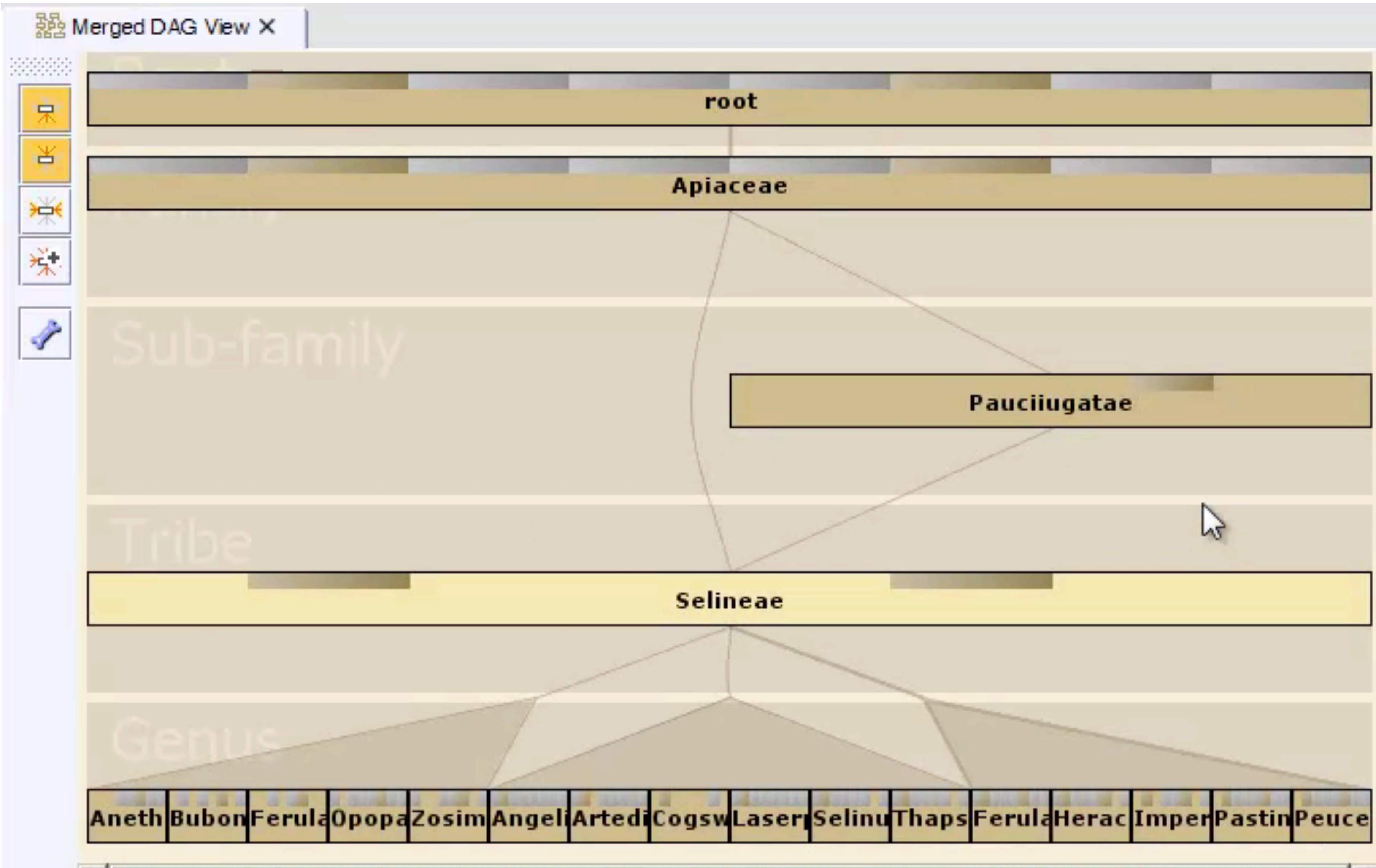
Zooming, filter



Shaping our
digital future

Institute for
Informatics
& Digital
Innovation

Smoothly Animated Transitions



Shaping our digital future

Institute for Informatics & Digital Innovation

Principles of Informaiton Visualisation Tutorial : Part 2 - Design Process

Cydney Nielsen

Edinburgh Napier UNIVERSITY