Advanced Data Visualization

CS 6965

Spring 2018

Prof. Bei Wang Phillips

University of Utah

Lecture 08

Today

- A project presentation by Prof. Melodie Weller from the School of Dentistry (a potential final project topic). Please contact the course instructor if your team is interested in the project; we will go through a selection process if there are more than one team.
- Visual mapping: glyphs, pixel-oriented, hierarchy-based, animation
- Visual mapping perception evaluation
- View transformation

Visual Mapping Continued

HD

Visualization pipeline for HD data

[LiuMaljovecWang2017]

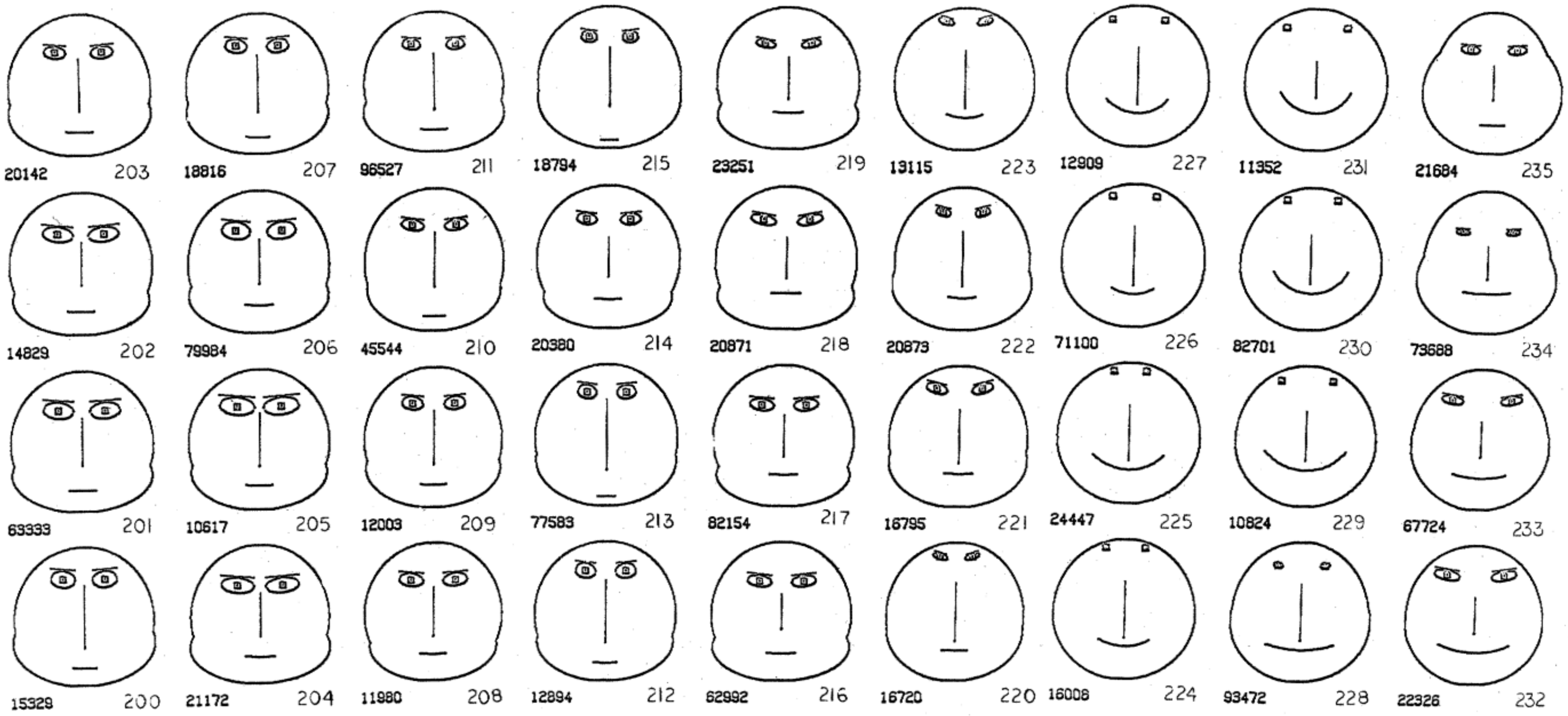
Visual mapping continued...

Glyphs

Glyphs: small graphical symbol

- Using shape, color, opacity, size, location to encode high-dim info

2. FACES FOR 53 GEOLOGICAL SPECIMENS OF EXAMPLE 2



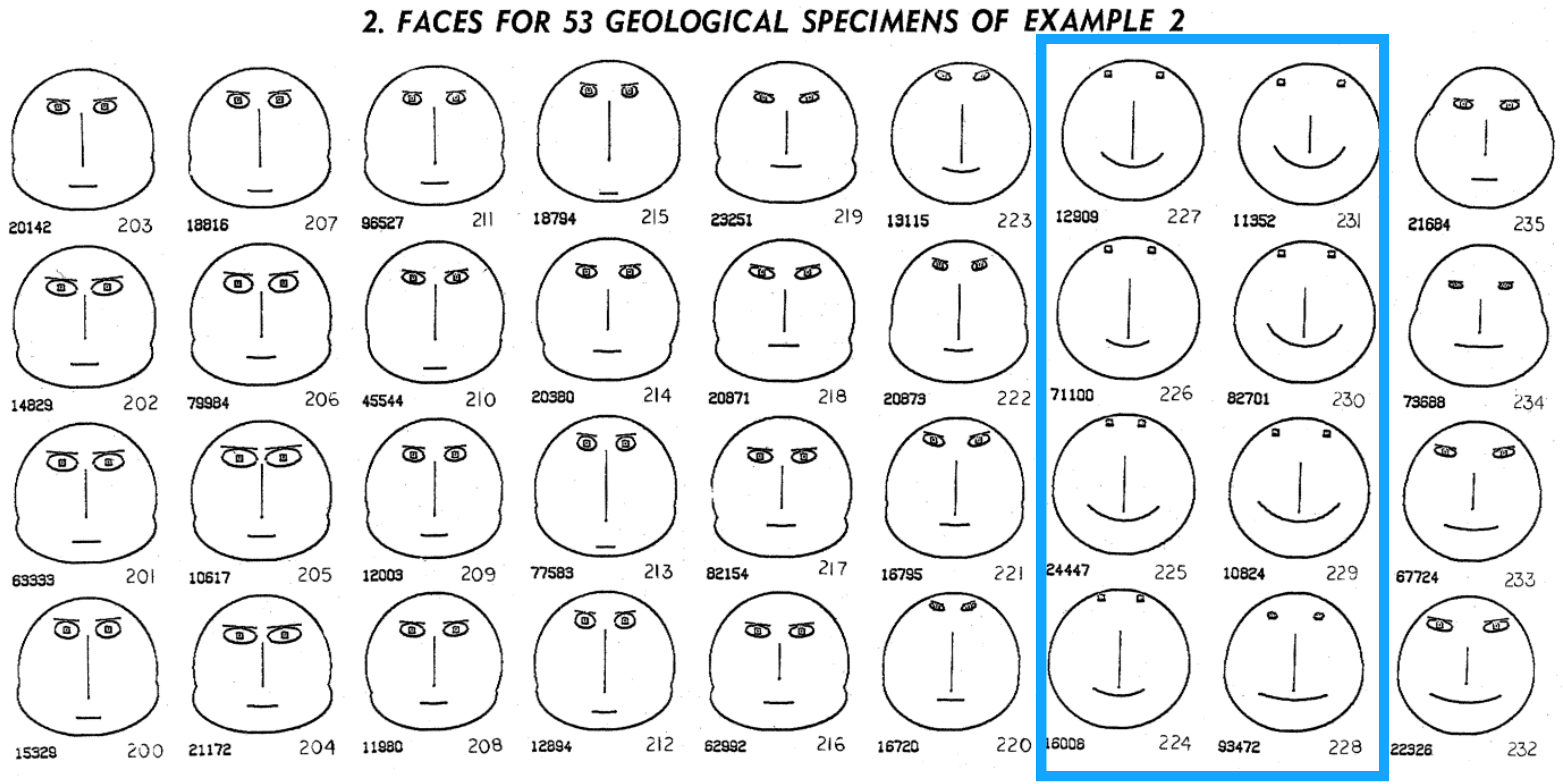
[Chernoff1973]

Chernoff faces

- One of the first attempts to map a high-dim data point to a single glyph
- Mapping different facial features to a separate dimension
- Each point in k-dimensional space ($k<18$) is represented by a cartoon of a face whose features, such as length of nose and curvature of mouth, correspond to components of the point.
- People grow up studying and reacting to faces all of the time. Small and barely measurable differences are easily detected and evoke emotional reactions from a long catalogue buried in the memory. …human mind subconsciously operates as a high-speed computer, filtering out insignificant visual phenomena and focusing on the potentially important…It is this flexibility which is lacking in standard computer programs.

[Chernoff1973]

Chernoff faces



[Chernoff1973]

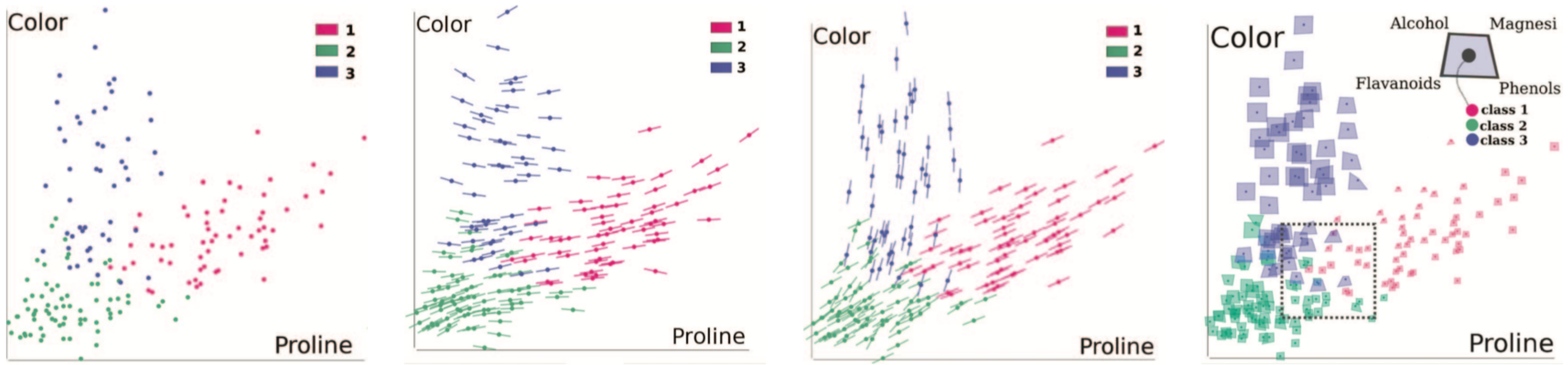
Why glyphs?

- Enhancing the user's ability to detect and comprehend important phenomena.
- Serving as a mnemonic device for remembering major conclusions
- Communicating major conclusions to others.
- Providing the facility for doing relatively accurate calculations informally.

- Recent efforts: provide statistical and sensitivity information to present trends in data

[Chernoff1973]

Generalizing scatterplot



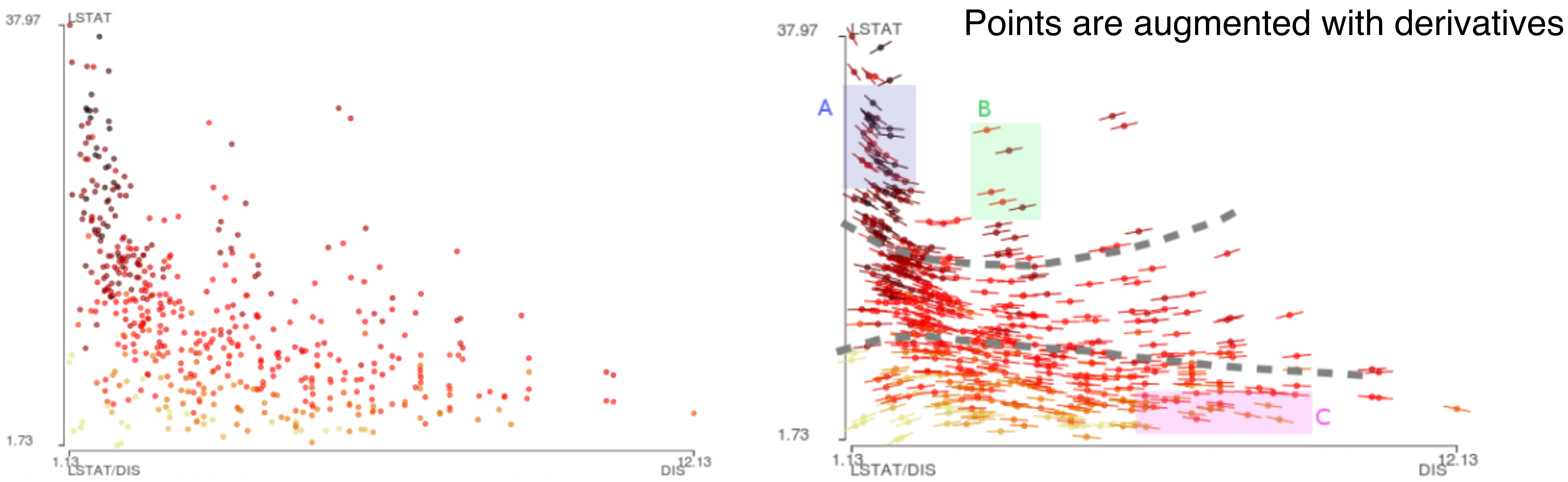
(a) Scatterplot (b) Flow-based scatterplot (c) GSS (d) Sensitivity Star Glyphs

Fig. 1. Visualization of two variables of the Wine data set [47], where points are color coded by class. (a) Traditional scatterplots suffer from overlap and different classes may appear to mix in arbitrary ways. (b) a FBS [10] reveals a positively correlated trend, but misrepresents the trend of the points in blue. (c) a GSS on a 3D subspace represents better the trends of points belonging to different classes; class 3 is distinguishable from others. (d) A star glyph plot summarizes the GSS across multiple 3D subspaces, where all of the three classes stand out distinctly, as evidenced by the shape and size of the glyphs.

By utilizing local linear regression to compute partial derivatives around sampled data points and representing the information in terms of glyph shape, sensitivity information can be encoded into scatterplots [ChanCorreaMa2013]

Flow-based scatterplot

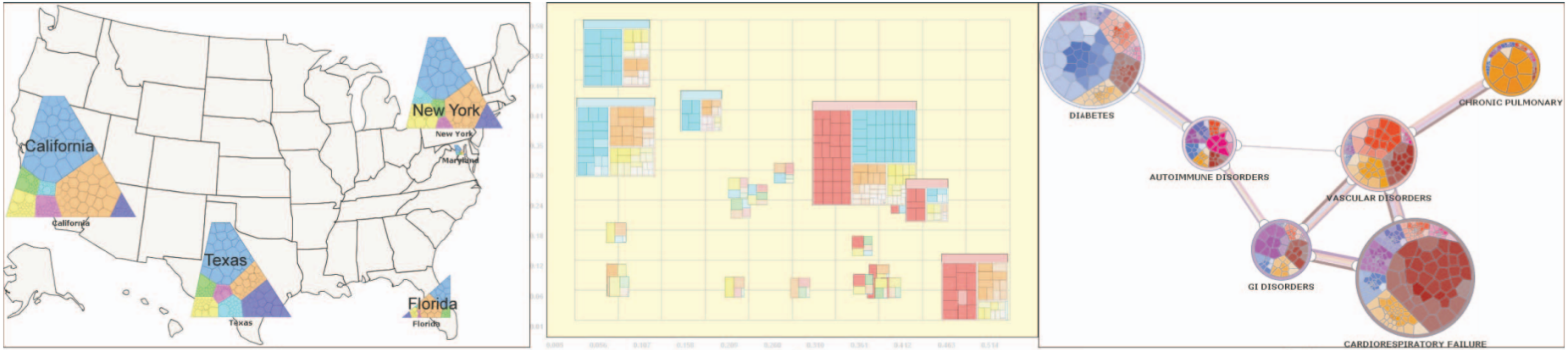
- Computing the sensitivity of one variable w.r.t. the other
- Sensitivity can be approximated by the partial derivative of one variable with respect to another.
- Approximate the partial derivative for point (x0,y0) in a neighborhood of N points: $\frac{\partial y}{\partial x} \approx \frac{\sum_{i=0}^{N}(y_i - y_0)(x_i - x_0)}{\sum_{i=0}^{N}(x_i - x_0)^2}$



(a) (b)

[ChanCorreaMa2010]

Icon-based visualization



(a) (b) (c)

Fig. 1. DICON is a dynamic icon-based visualization technique that helps users understand, evaluate, and adjust complex multidimensional clusters. It provides visual cues describing the quality of a cluster as well as its multiple attributes, and can be embedded within many kinds of visualizations such as maps, scatter plots, and graphs.

[CaoGotzSun2011]

Pixel-Oriented Approaches

Pixel based visualization

- Encode maximal amount of information (?)
- Dense pixel displays
- Encode data values as individual pixels and creating separate displays or subwindows for each dimension

Pixel bar charts

- Use the pixels within the bars to present the detailed information of the data records.

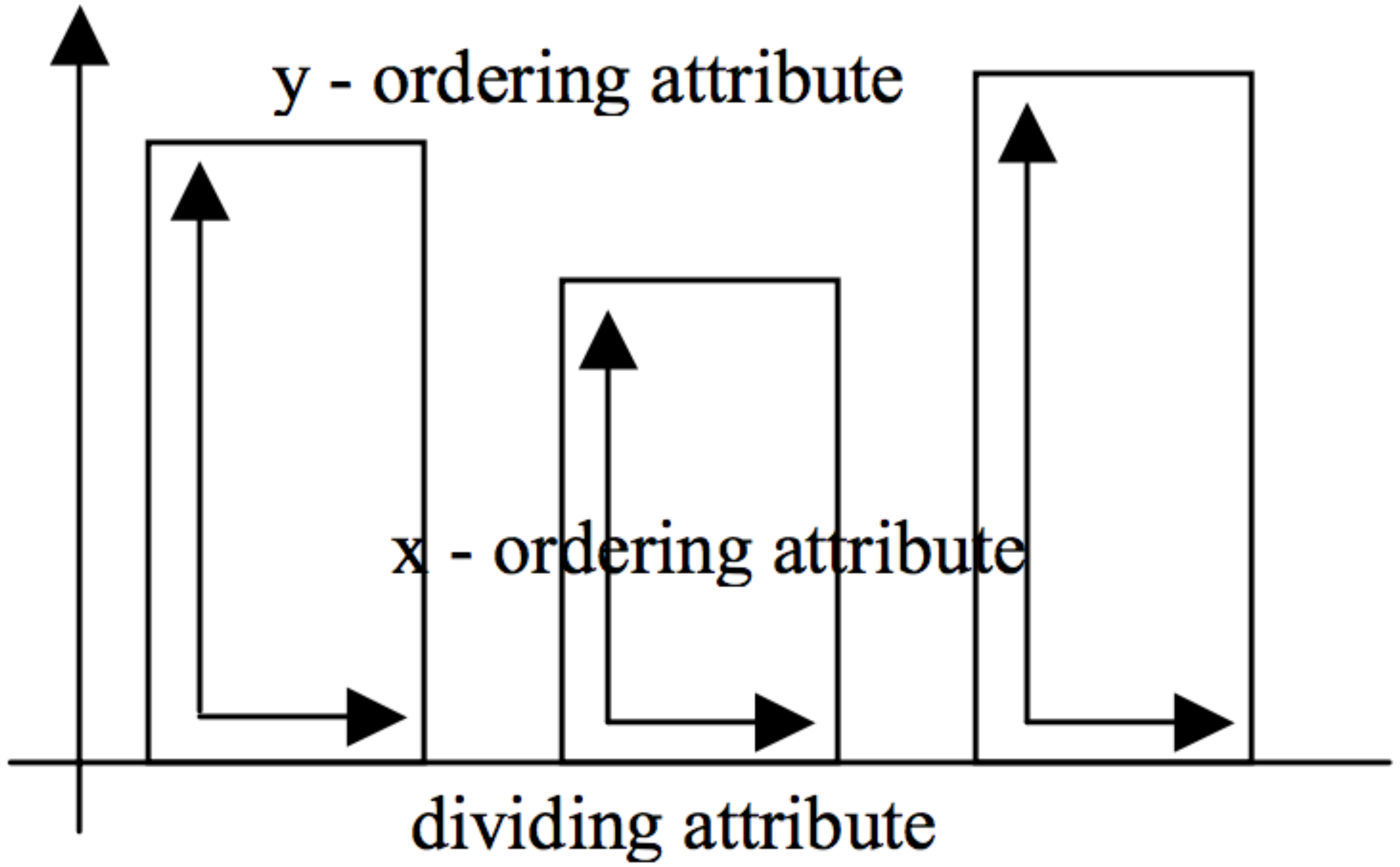
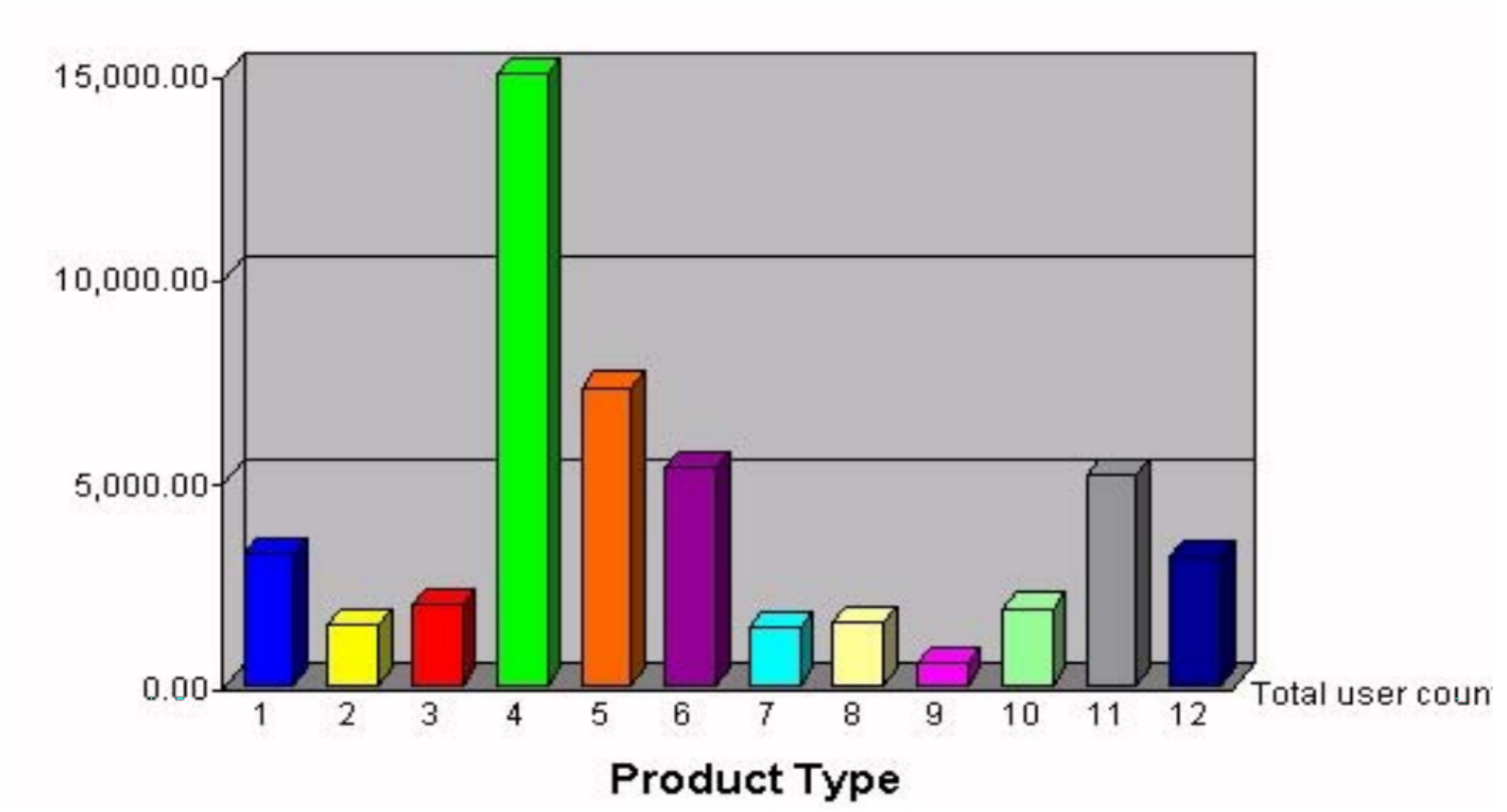


Figure 2: A Pixel Bar Chart

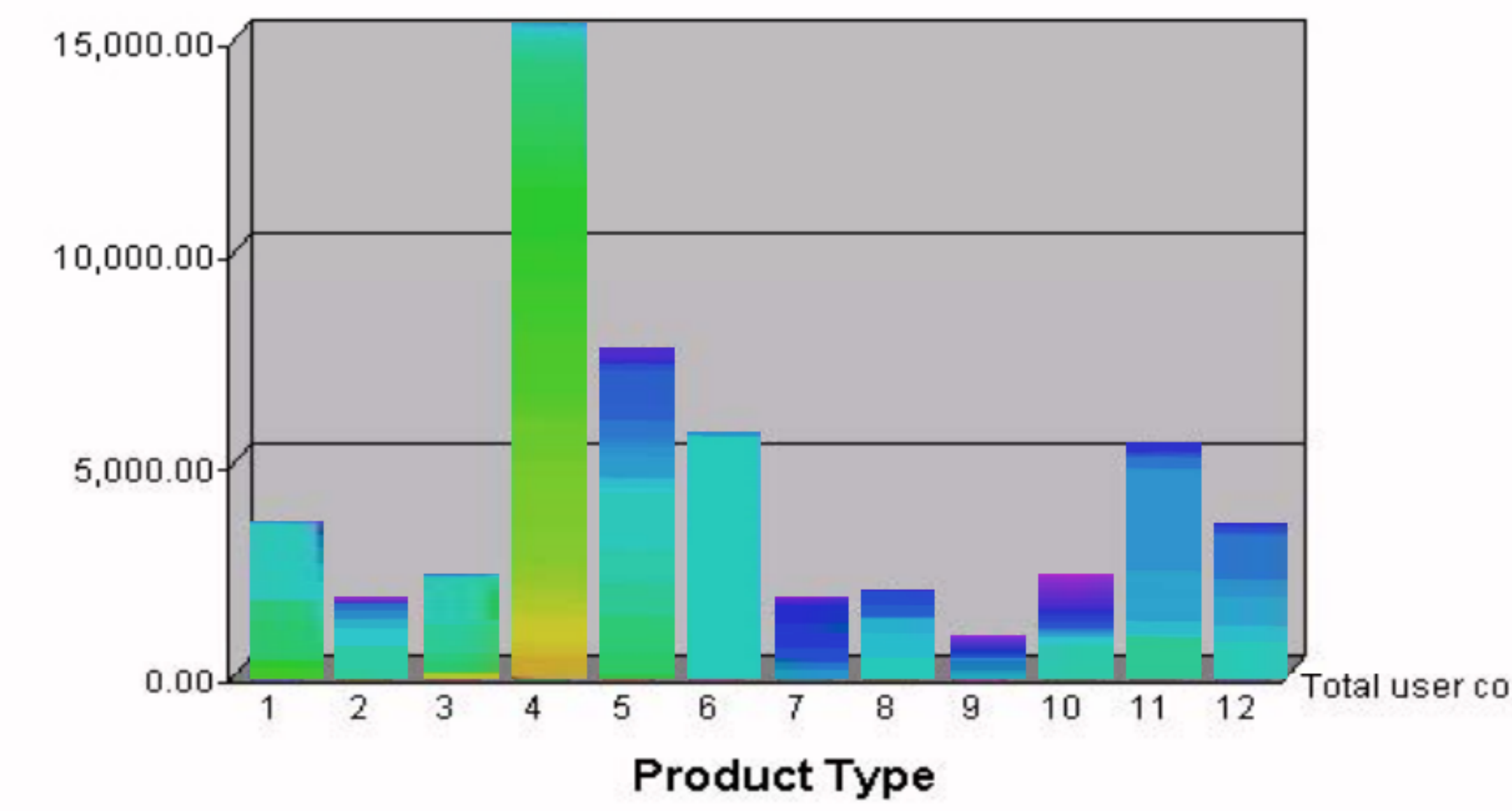
[KeimHaoLadisch2001]

Pixel bar charts



a) Equal-Width Bar Chart

Regular bar chart



a) Equal-Width Pixel Bar Chart

[KeimHaoLadisch2001]

Treemap

http://www.cs.umd.edu/hcil/treemap/

Space-constrained vis of hierarchical structures

Fig 3.

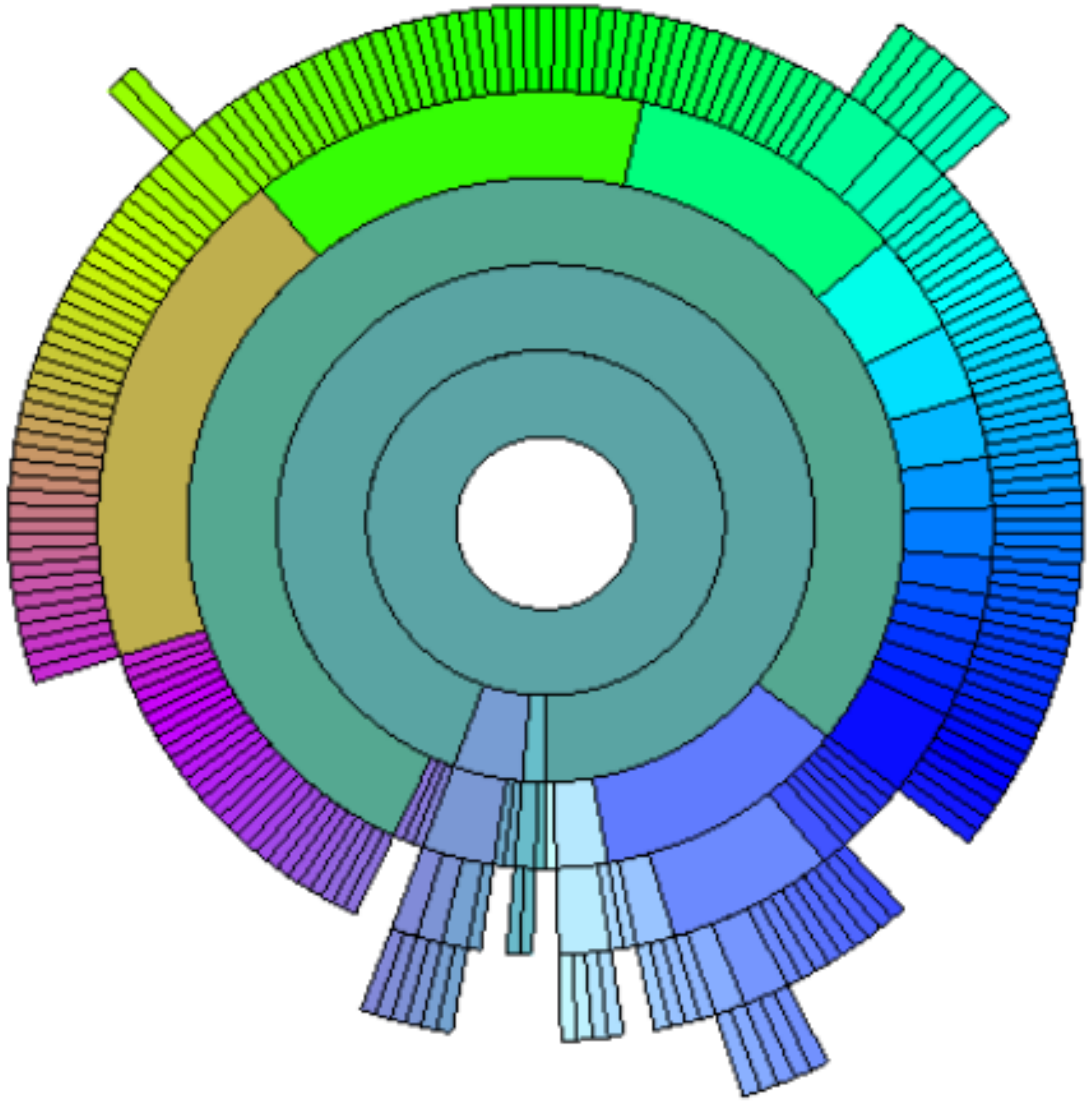
Hierarchical tree map contrasting the relative popularity of research th categories addressed with CS datasets to that of the wider UE literatur (a) and butterflies (b): the size of the boxes represents the relative popular category amongst CS datasets, while the shading represents the relative po each category out of the overall UE dataset.

https://doi.org/10.1371/journal.pone.0156425.g003

Hierarchy-Based Approaches

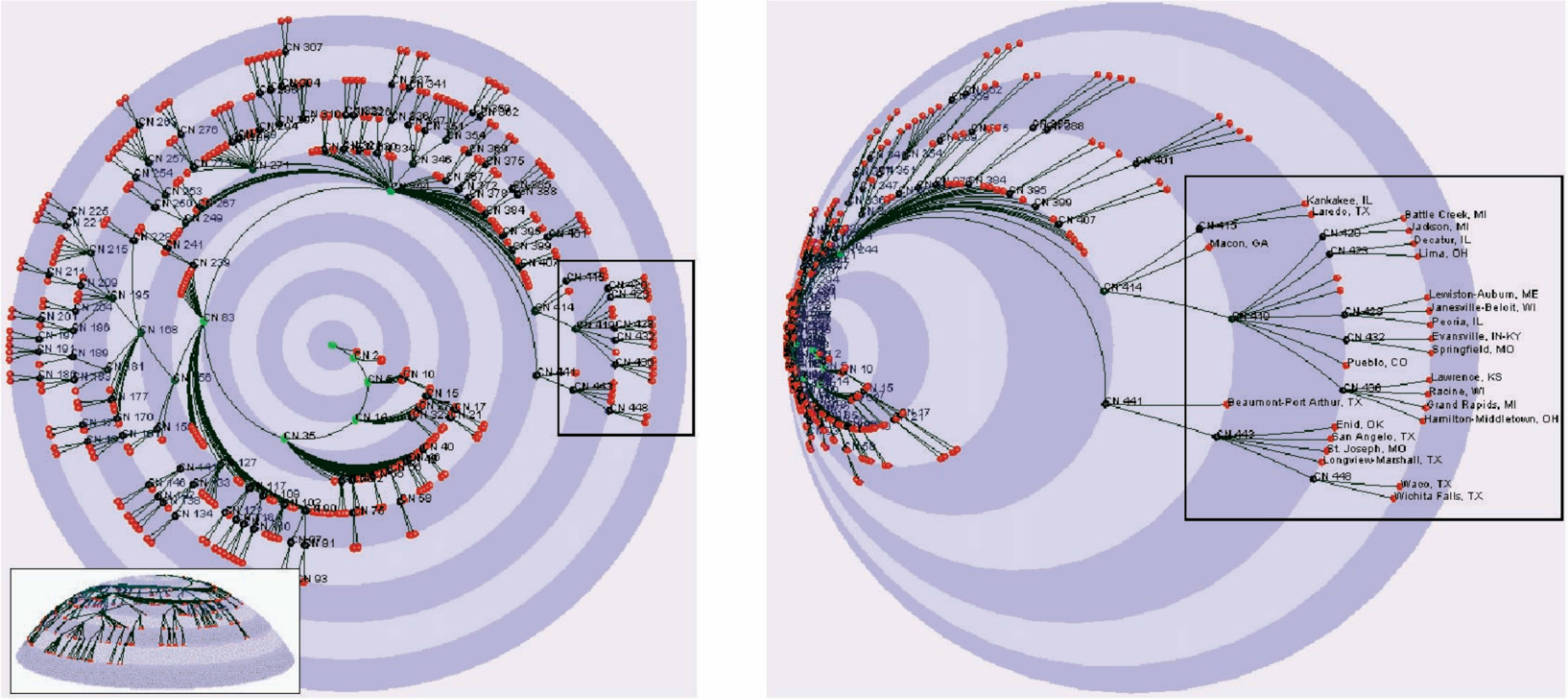
Dimension hierarchies

- Group similar clusters of dimensions together in a hierarchy
- Key: using a similarity measure of dimensions



YangPengWard2003

Focus+context: hierarchical graphs



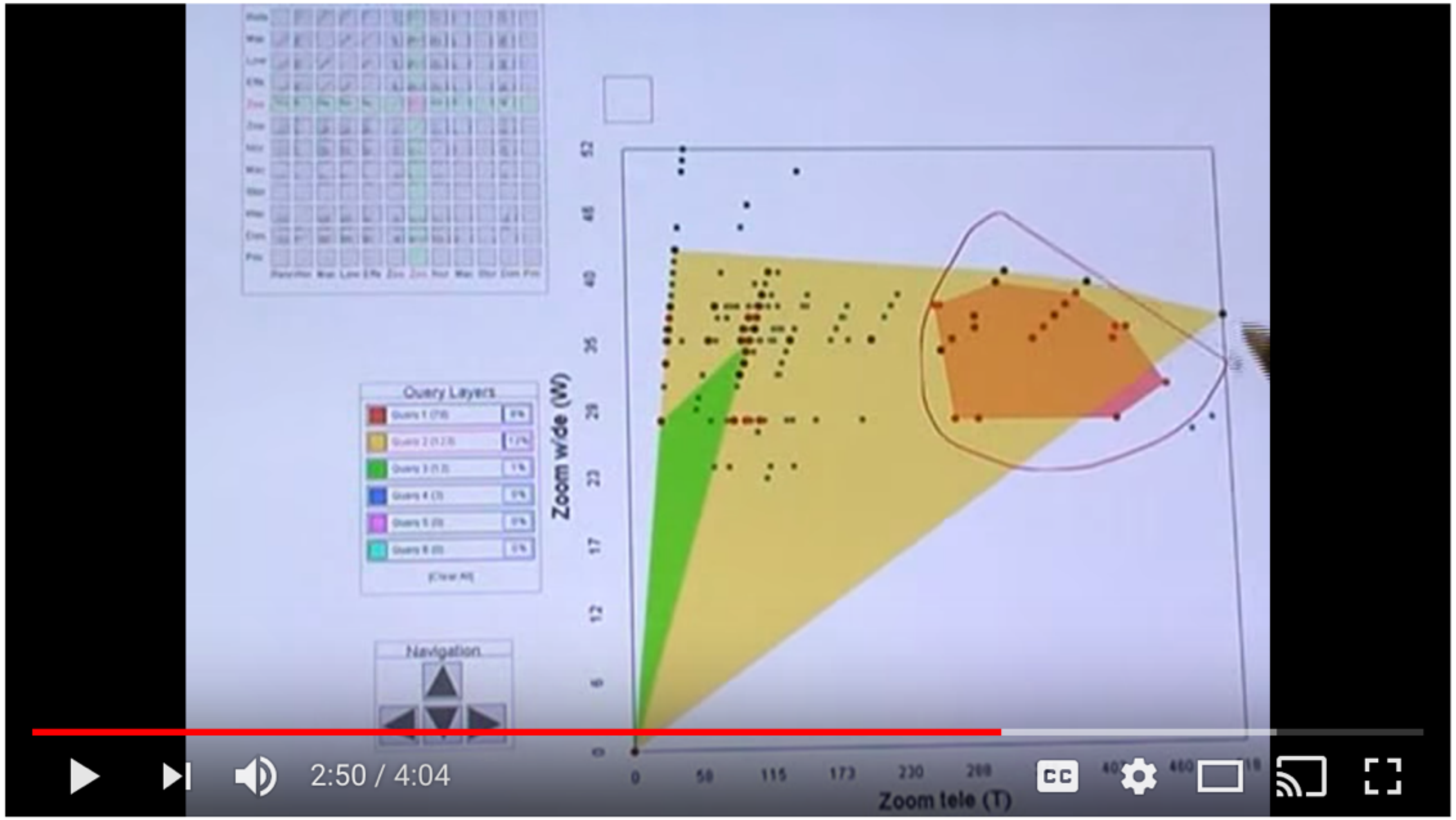
(a) (b)

[KreuselerSchumann2002]

Animation

Rolling the Dice

- Transition between any pair of scatterplots in a SPLOM
- Connecting a series of 3D transitions between scatterplots



Rolling the Dice: Scatterplot Matrix Navigation
https://www.youtube.com/watch?v=E1birsp9iYk

Other Examples

- Dynamic projection, continuous linear projection such as GGobi etc.
- TripAdvisor-ND

Perception Evaluation

Evaluation

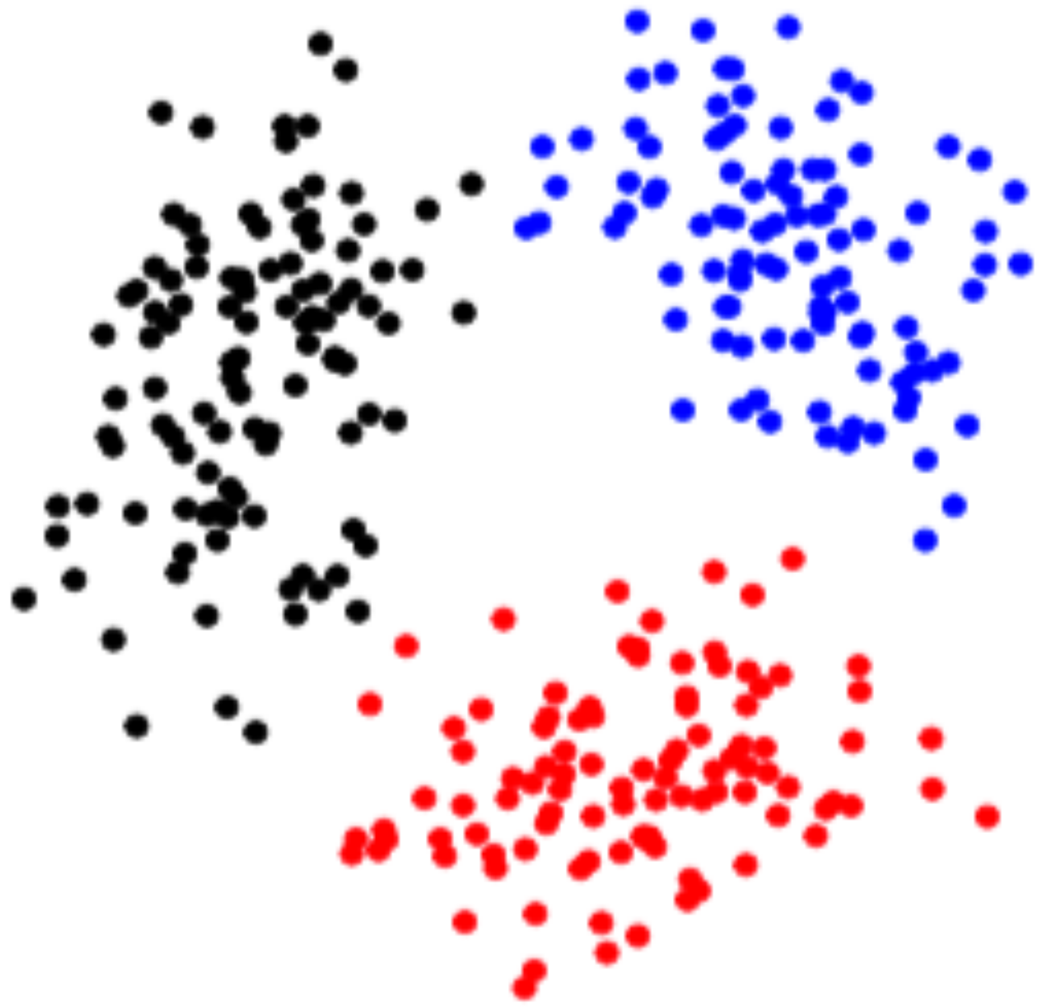
- Design goal of visual mapping: directly convey th information to the user through visual perception
- Determining the effectiveness of the overall visualization
- Some existing conclusions

Evaluate scatterplots and DR

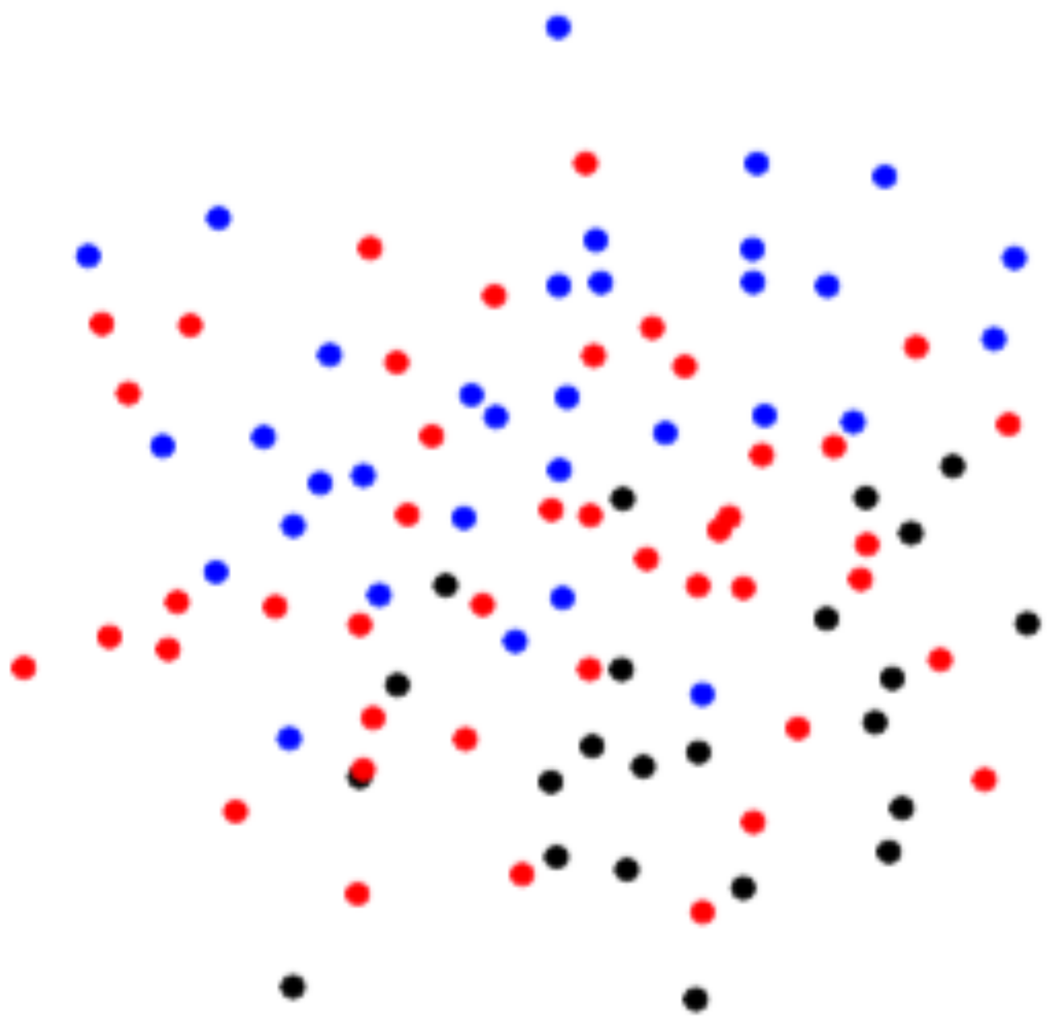
- An empirical data study.
- Two human coders manually inspected a broad set of 816 scatterplots derived from 75 datasets, 4 DR techniques, and the 3 scatterplot techniques.
- Each coder scored all color-coded classes in each scatterplot in terms of their separability from other classes.

[SedlmairMunznerTory2013]

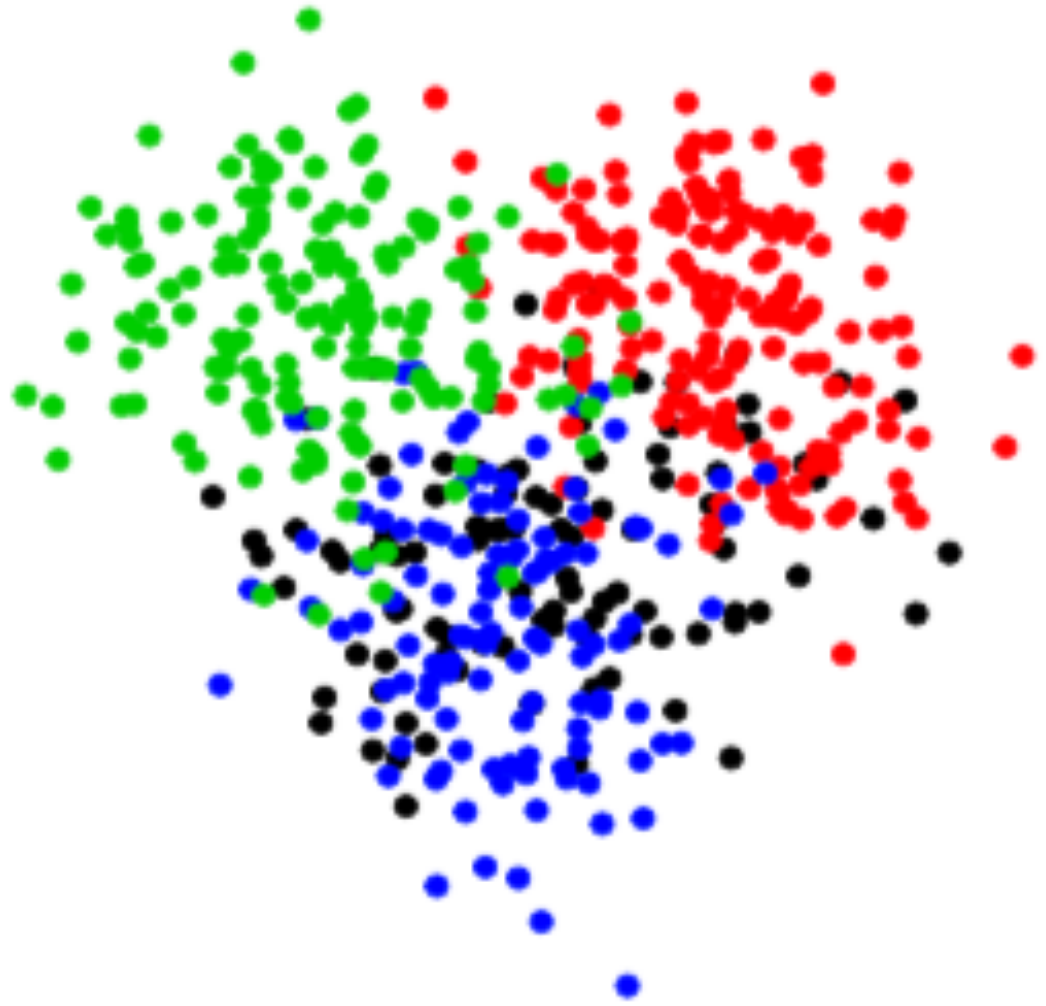
Verifying cluster separation



(a) max sep: (5,5,5)



(b) no sep: (1,1,1)



(c) mixed: (4,3,2,1)

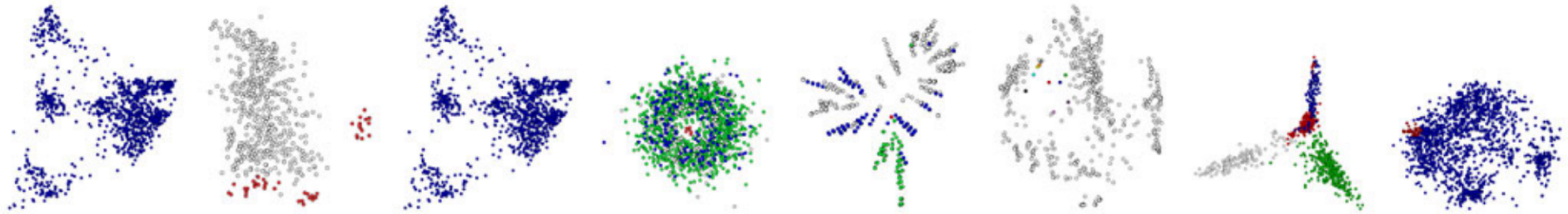
[SedlmairMunznerTory2013]

Some Conclusions

- 2D scatterplots are often ‘good enough’, that is, neither SPLOM nor interactive 3D adds notably more cluster separability with the chosen DR technique.
- If 2D is not good enough, the most promising approach is to use an alternative DR technique in 2D.
- SPLOM occasionally adds additional value, and interactive 3D rarely helps but often hurts in terms of poorer class separation and usability.

[SedlmairMunznerTory2013]

Perception evaluation of DR



(a) #Clu (b) #SClu (c) #Out (d) fCluClu (e) fCluObj (f) rKnn (g) rDens (h) #Obj

Fig. 1. Instances of task stimuli: (a) Estimate number of clusters, (b) estimate number of subclusters of red group, (c) estimate number of outliers, (d) determine whether green or blue cluster is closer to red object, (e) determine whether green or blue cluster is closer to red cluster, (f) find five closest objects to red object, (g) rank red, green, and blue clusters by density, and (h) estimate number of objects in red group.

[EtemadpourMottaPaiva2015]

Perception evaluation of DR

	#Clu	#SClu	#Out	fCluClu	fCluObj	rKnn	rDens	#Obj
Glimmer	3.67	2.80	4.31	3.13	3.65	3.17	3.30	4.31
Isomap	3.58	3.19	3.72	3.10	3.23	3.53	3.57	3.72
LSP	3.12	3.23	3.57	3.4	3.58	3.45	3.78	3.57
PCA	3.61	3.63	3.10	2.77	3.12	2.81	3.75	3.10
Tree	2.91	3.31	3.74	3.26	3.84	3.17	3.61	3.74

Fig. 4. Confidence: Comparing mean confidence values for completing tasks with different projection methods. Colors indicate groups of no significant pairwise differences in form of winners shown in dark red, loser in white, and the ones in between in light red.

[EtemadpourMottaPaiva2015]

Evaluating PCP

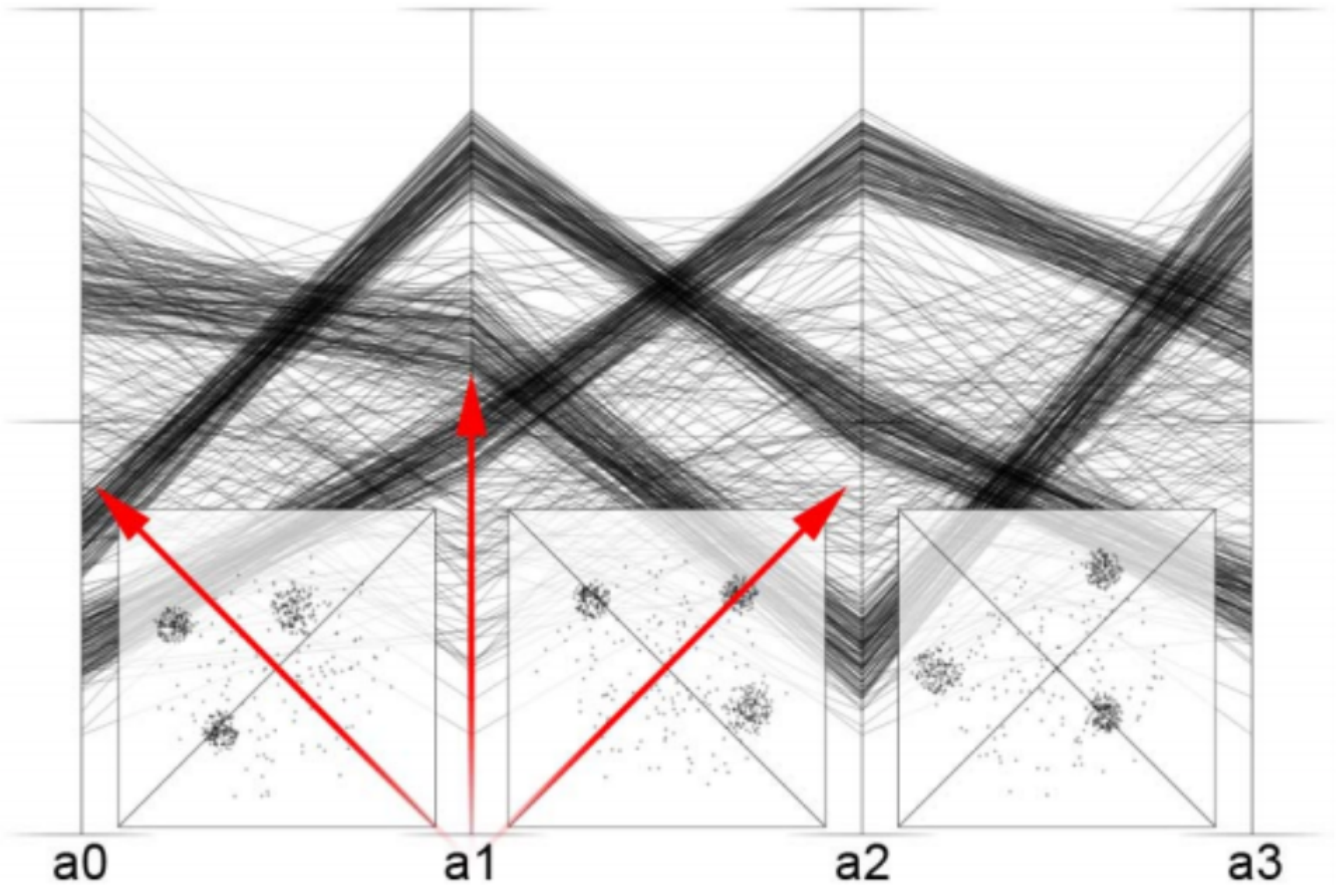


Figure 1: *Scatter plots embedded into a PCP. Red arrows denote the direction of the* a1 *axis in each visualization.*

- Evaluation of cluster identification performance for different PCP variants
- Conclusion: A fair number of the seemingly valid improvements, with the exception of scatter plots embedded into PCPs, do not result in significant performance gains.

[HoltenWijk2010]

Thanks!

Any questions?

You can find me at: **beiwang@sci.utah.edu**

CREDITS

Special thanks to all people who made and share these awesome resources for free:

- Presentation template designed by Slidesmash
- Photographs by unsplash.com and pexels.com
- Vector Icons by Matthew Skiles

Presentation Design

This presentation uses the following typographies and colors:

Free Fonts used:

http://www.1001fonts.com/oswald-font.html

https://www.fontsquirrel.com/fonts/open-sans

Colors used