Multivariate Visual Representations 1

CS 4460 – Intro. to Information Visualization
Sep. 18, 2017
John Stasko

Learning Objectives

- For the following visualization techniques/systems, be able to describe each and its visual encoding, know what type of data it's best for, know its strengths and limitations, and understand how to apply it
  - Iconic representatons (Chernoff faces), Table Lens, InfoZoom, EZChooser, Mosaic plot, Star plots
- Explain the visual encoding and design issues of Parallel Coordinates, as well as their utility and limitations
- Understand how the different types of variables in a multivariate data set influence the visualization technique that should be chosen to represent the data
- Be able to apply any of these techniques to a data set that is an appropriate match for them

Revisit

How Many Variables?

- Data sets of dimensions 1, 2, 3 are common
- Number of variables per class
  - 1 - Univariate data
  - 2 - Bivariate data
  - 3 - Trivariate data
  - >3 – Hyper/Multivariate data Focus This Week

Earlier

- We examined a number of tried-and-true techniques/visualizations for presenting multivariate (typically <=3) data sets
  - Bar graph, line graph, pie chart, scatterplot, box plot, trellis display, crosstab, radar graph, heatmap
- Hinted at how to go above 3 dimensions

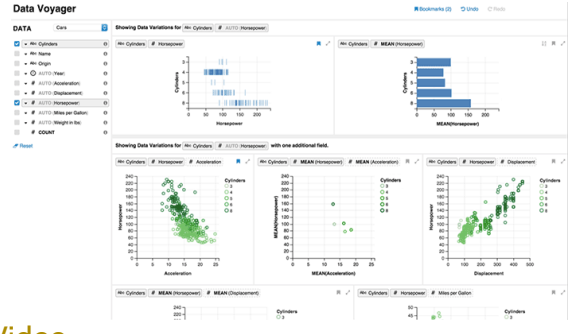
Hypervariate Data

- How about 4 to 20 or so variables (for instance)?
  - Lower-dimensional hypervariate data
  - Many data sets fall into this category

How would you handle that?

One Approach

- Stay with standard views, but use lots and lots of them



Voyager

Video

Wongsuphasawat et al
TVCG (InfoVis) '16

More Dimensions

- Fundamentally, we have 2 geometric (position) display dimensions
- For data sets with >2 variables, we must project data down to 2D
- Come up with visual mapping that locates each dimension into 2D plane

- Computer graphics: 3D->2D projections

Wait a Second

- A spreadsheet already does that
  - Each variable is positioned into a column
  - Data cases in rows
  - This is a projection (mapping)

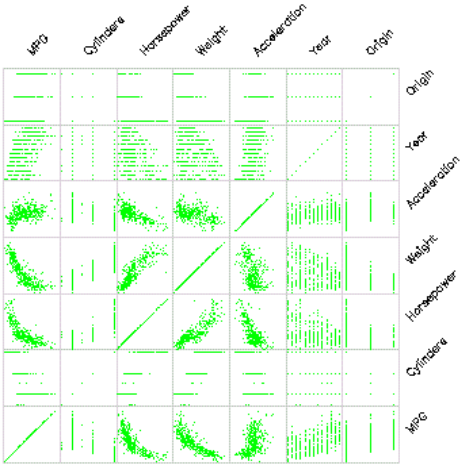
- What about some other techniques?
  - Already seen a couple

Revisit

Scatterplot Matrix

Represent each possible pair of variables in their own 2-D scatterplot

If pairwise correlation is key



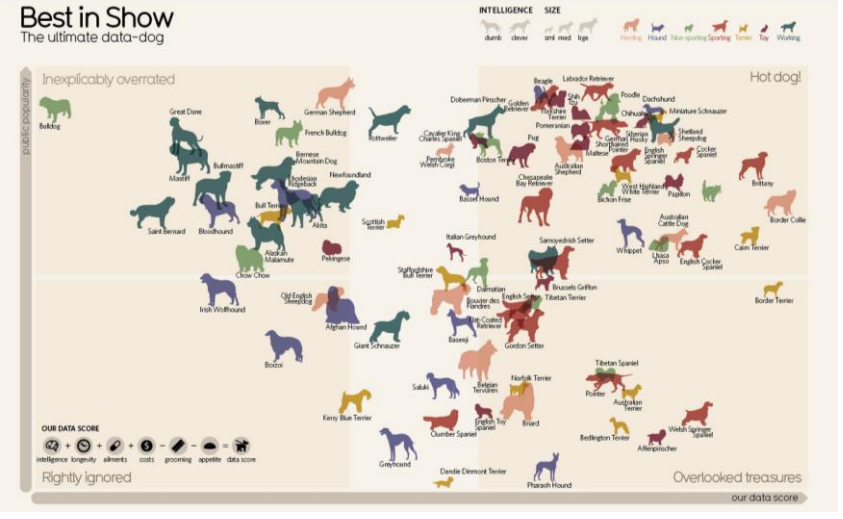
Key Principle (today)

- Handle all data sets generically
  - Examine techniques not specific to some data or domain
  - Technique can generally handle all data sets

Iconic Representations

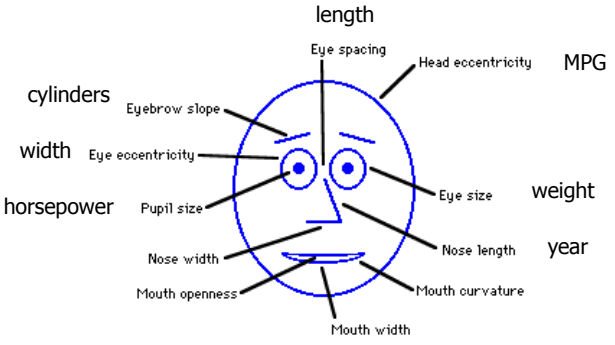
- Glyph (graphical object) represents a data case
- Visual properties of glyph represent different variables

Remember?

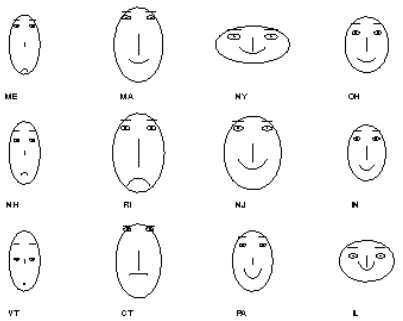


Chernoff Faces

Encode different variables' values in characteristics of human face



Examples



Cute applet: `http://www.cs.uchicago.edu/~wiseman/chernoff/`

Table Lens

- Spreadsheet is certainly one hypervariate data presentation
- Idea: Make the text more visual and symbolic
- Just leverage basic bar chart idea

Rao & Card
CHI '94

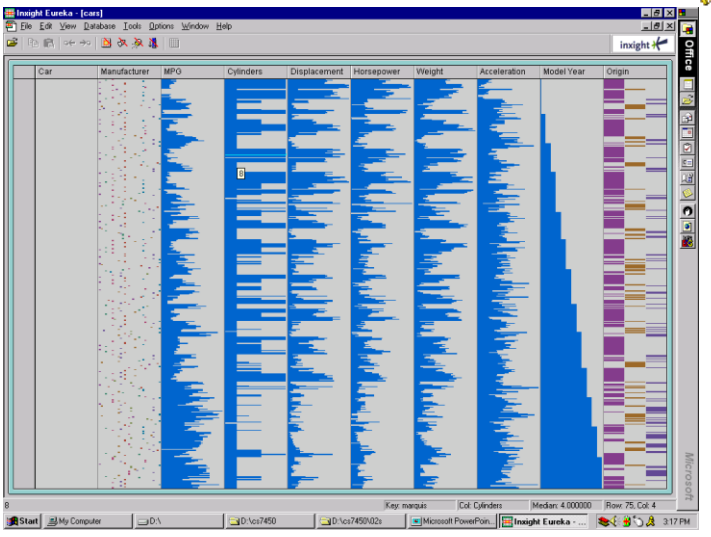
Visual Mapping

	A	B	C	D	E	F
1	Sales rep	Quota	Variance to quota	% of quota	Forecast	Actual bookings
2	Albright, Gary	200,000	-16,062	92	205,000	183,938
3	Brown, Sheryll	150,000	84,983	157	260,000	234,983
4	Cartwright, Bonnie	100,000	-56,125	44	50,000	43,875
5	Caruthers, Michael	300,000	-25,125	92	324,000	274,875
6	Garibaldi, John	250,000	143,774	158	410,000	393,774
7	Girard, Jean	75,000	-48,117	36	50,000	26,883
8	Jones, Suzanne	140,000	-5,204	96	149,000	134,796
9	Larson, Terri	350,000	238,388	168	600,000	588,388
10	LeShan, George	200,000	-75,126	62	132,000	124,874
11	Levenson, Bernard	175,000	-9,267	95	193,000	165,733
12	Mulligan, Robert	225,000	34,383	115	275,000	259,383
13	Tetracelli, Sheila	50,000	-1,263	97	50,000	48,737
14	Wotisek, Gillian	190,000	-3,648	98	210,000	186,352
15						

Change quantitative values to bars

How did they handle nominal data?

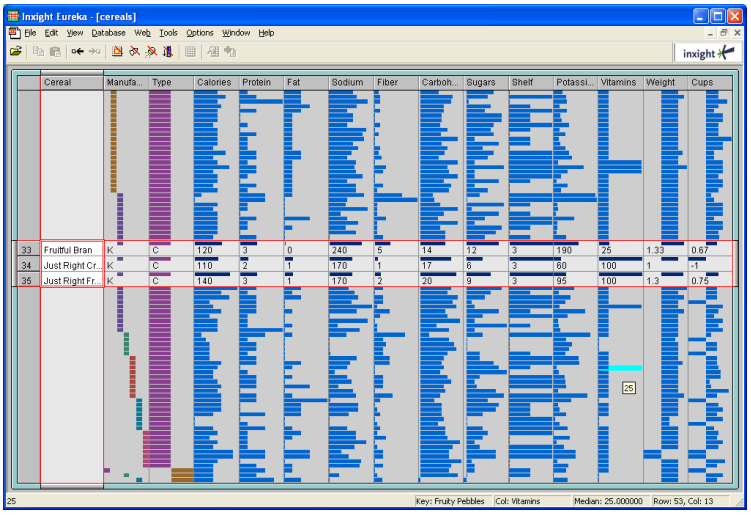
Instantiation



Details

Video

Focus on item(s) while showing the context

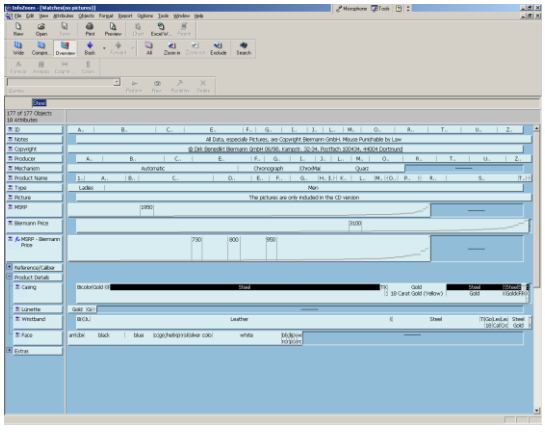


Alternative

What if you
- Flipped rows and columns
- Sorted each row

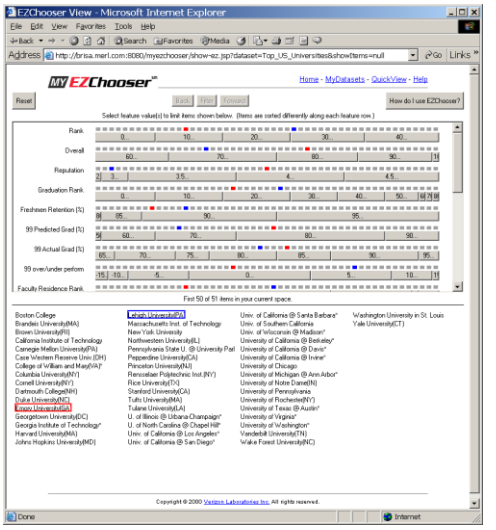
Result

InfoZoom



Commercial product
Demo/video

Similar Idea



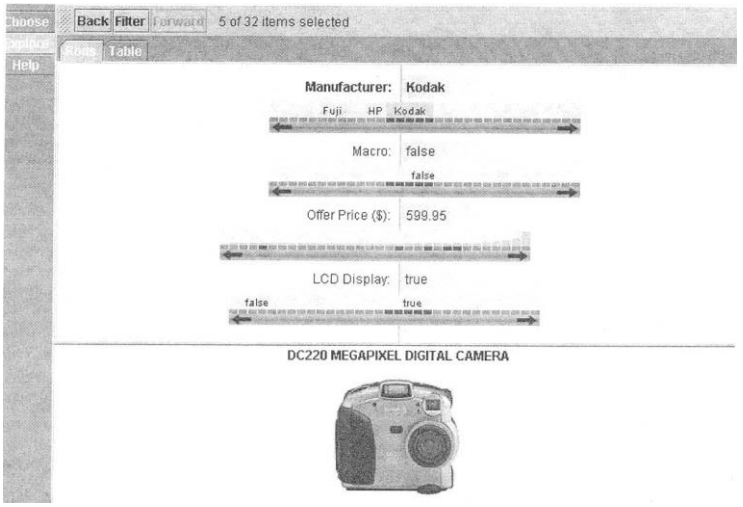
EZChooser

Video

Alternate UI

- Can slide the values in a row horizontally
- A particular data case then can be lined up in one column, but the rows are pushed unequally left and right

Attributes as Sliding Rods



Limitations?

- Number of cases (horizontal space)
- Nominal & textual attributes don't work quite as well

An Application

- What if you cared about ranking items?
  - Think of the attributes per item as contributing to some score or value for it
- Apply the Table Lens representation we saw earlier

LineUp

Video

Gratzl et al
TVCG (InfoVis) '13

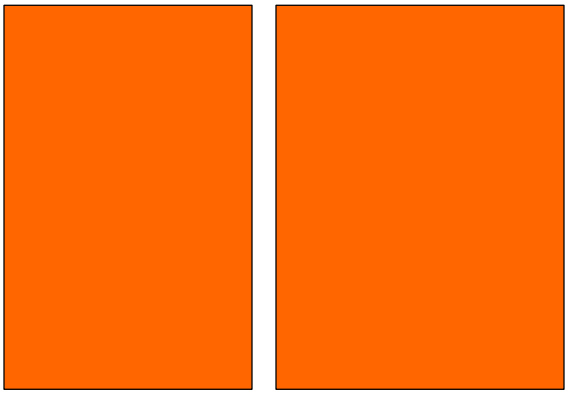
Categorical data?

- How about multivariate categorical data?

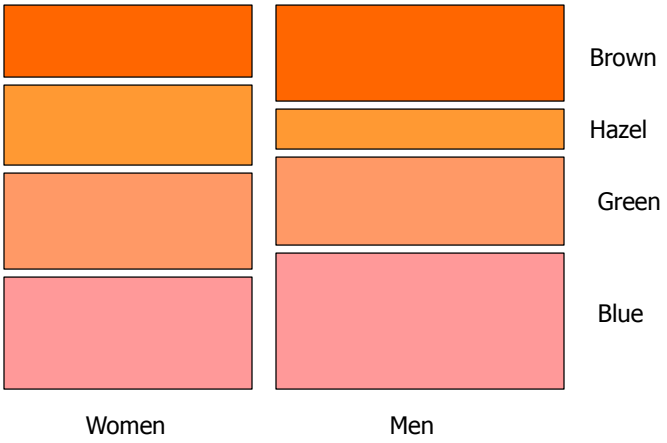
- Students
  - Gender: Female, male
  - Eye color: Brown, blue, green, hazel
  - Hair color: Black, red, brown, blonde, gray
  - Home country: USA, China, Italy, India, …

Mosaic Plot

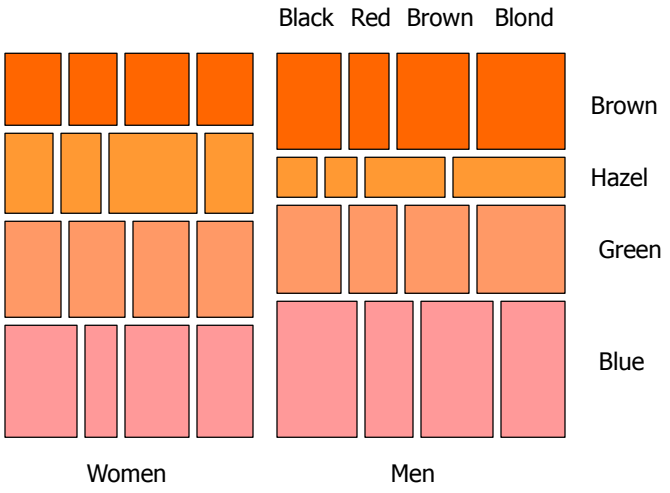
Mosaic Plot



Mosaic Plot



Mosaic Plot



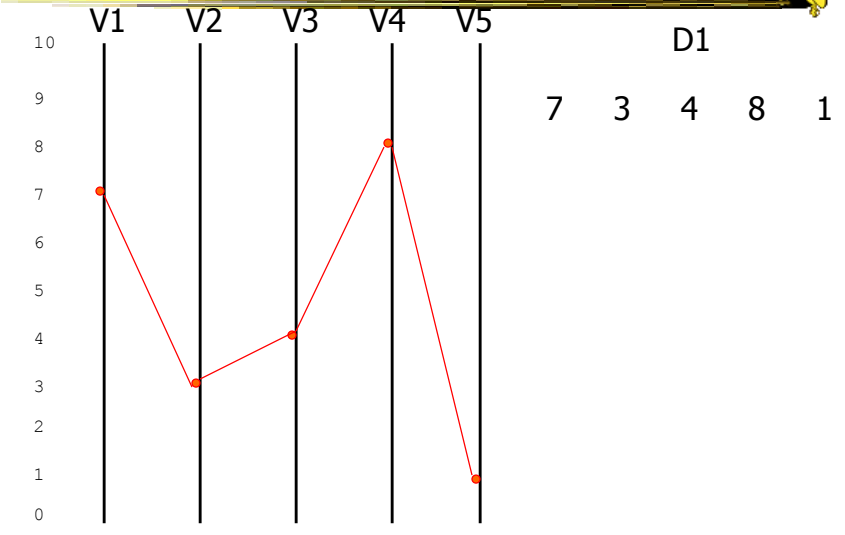
Fundamental Limitation

- Run out of rows/columns for lots of data cases
- How about an alternative generic representation?

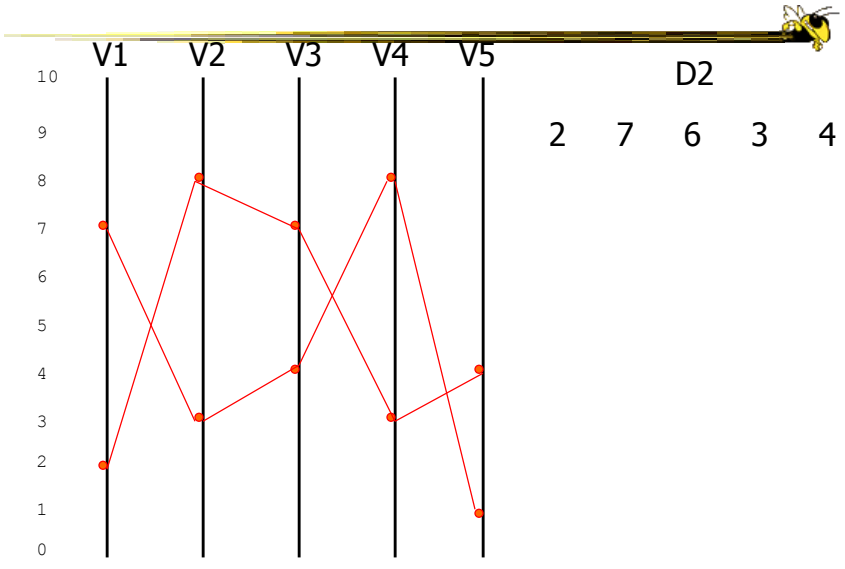
Parallel Coordinates

		variables				
		V1	V2	V3	V4	V5
	D1	7	3	4	8	1
data cases	D2	2	7	6	3	4
	D3	9	8	1	4	2

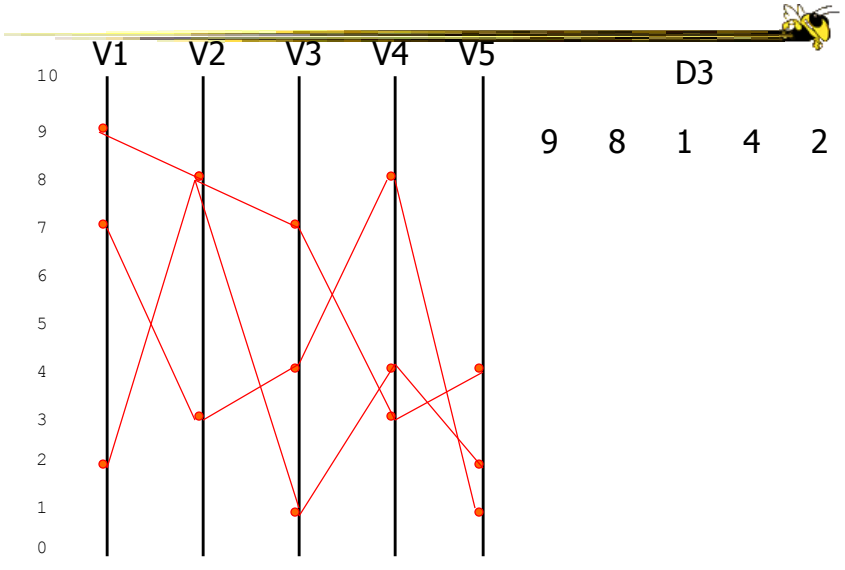
Parallel Coordinates



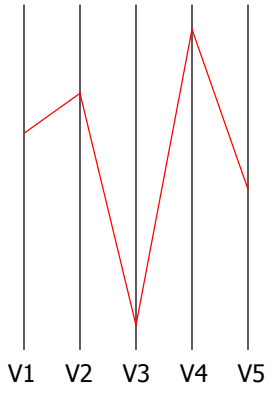
Parallel Coordinates



Parallel Coordinates



Parallel Coordinates



Encode variables along a horizontal row

Vertical line specifies different values that variable can take

Data point represented as a polyline

Questions

What do two correlated variables look like?

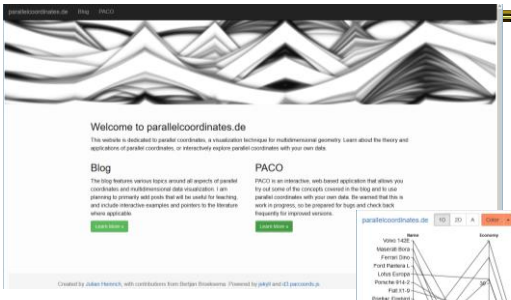
What do two inversely correlated variables look like?

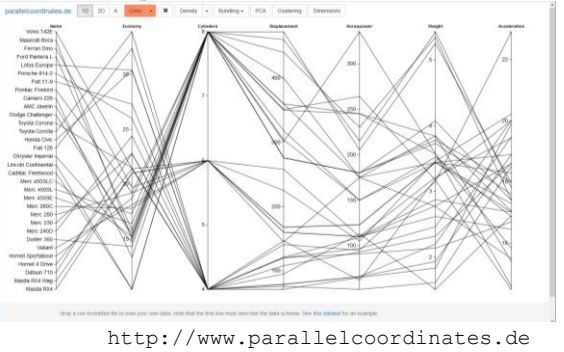
Issue

- Different variables can have values taking on quite different ranges
- Must normalize all down (e.g., 0->1)

To Learn More

Great site
that's all about
parallel coords

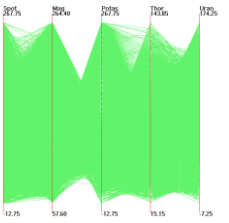




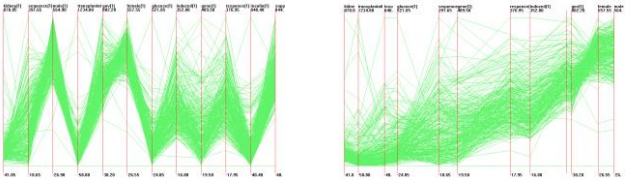
http://www.parallelcoordinates.de

Challenges

- Too much data

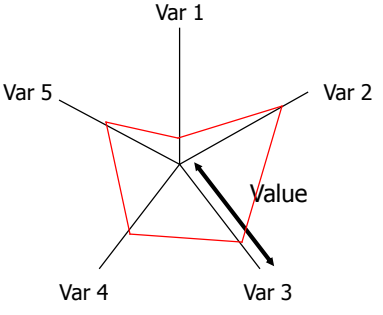


- Order of dimensions really matters



Yang et al
InfoVis '03

Star Plots (Radar Chart)



Space out the n variables at equal angles around a circle

Each "spoke" encodes a variable's value

Alternative Rep.

Data point is now a "shape"

Example

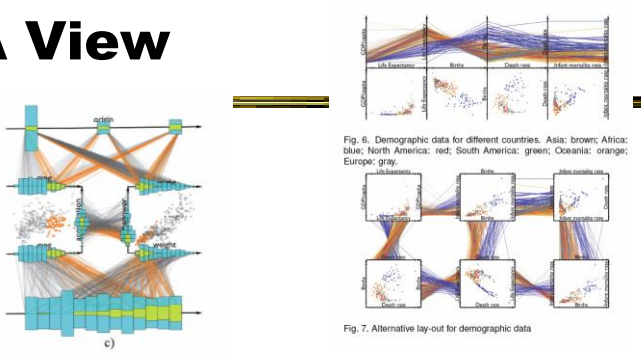
Time
April 16, 2012

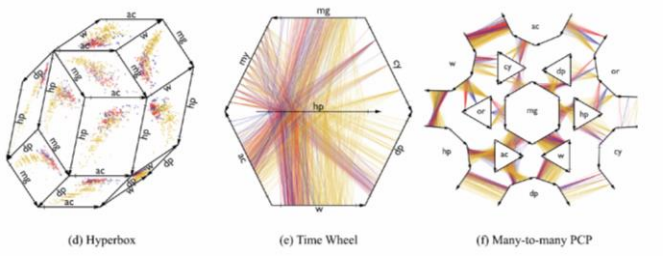
Generalizing the Principles

- General & flexible framework for axis-based visualizations
  - Scatterplots, par coords, etc.
- User can position, orient, and stretch axes
- Axes can be linked

Claessen & van Wijk
TVCG (InfoVis) '11

FLINA View





Video

Learning Objectives

- For the following visualization techniques/systems, be able to describe each and its visual encoding, know what type of data it's best for, know its strengths and limitations, and understand how to apply it
  - Iconic representatons (Chernoff faces), Table Lens, InfoZoom, EZChooser, Mosaic plot, Star plots
- Explain the visual encoding and design issues of Parallel Coordinates, as well as their utility and limitations
- Understand how the different types of variables in a multivariate data set influence the visualization technique that should be chosen to represent the data
- Be able to apply any of these techniques to a data set that is an appropriate match for them

HW 3

- Due Friday
- Questions?

Upcoming

- Multivariate Visual Representations 2
  - Prep: Dust & Magnet video
- Lab 2: SVG