Multivariate visualization

C. Andrews

2016-03-25

Univariate



Univariate: Histogram



Univariate: Stem and Leaf

Stem	Leaf
4	12
5	7
6	34458
7	1222566666778999
8	0111122339
9	1235555678899
10	00



Adding a variable...



Univariate questions

Which is the biggest?

Which is the smallest?

What is the "center"?

What is the "shape" of the data?

What are the outliers?

Multivariate questions

Which items are most alike?

Which items are most exceptional?

How can these items be combined into logical groups based on similarity?

Bivariate



Trivariate



d

Hypervariate



d

MultiD Scatterplots



or just add another axis

Scatterplot matrix



http://ericksondata.com/wp/2012/150-varieties-of-hops/

Scatterplot matrix



http://www.statsoft.com/support/blog/entryid/212/finding-the-right-pieces-to-the-puzzle



http://junkcharts.typepad.com/junk_charts/2010/06/the-scatterplot-matrix-a-great-tool.html

Trellis plot



Trellis plot



Trellis graph - not just scatterplots



Small multiples



Small multiples



Tufte. The Visual Display of Quantitative Information

Chaiyya Chaiyya Dance Moves

1	Å	X	7	ł	ĥ	1
2	Å	ŕ	ŕ	Å	እ	ን
3	Å	ť	Ŷ	ŕ	ች	∱
4	Å	Å	ġ	þ	Ŕ	∱
5	Å	ħ	ş	∱	Ŗ	ĥ

Small multiples



Embedded visualization



Multidimensional scaling

Calculate the similarity of all pairs of records using some distance function

Create a map that maps each record into our 2 (or 3) dimensional space

Calculate the similarity of all pairs of points

Compute the stress on the system as function of the difference between the similarity of the points and the similarity of the original records

If the stress is above some threshold, move points to reduce stress and repeat

Multidimensional scaling



Х