

Literature Review

Ineffective Visualization: Hyperbox

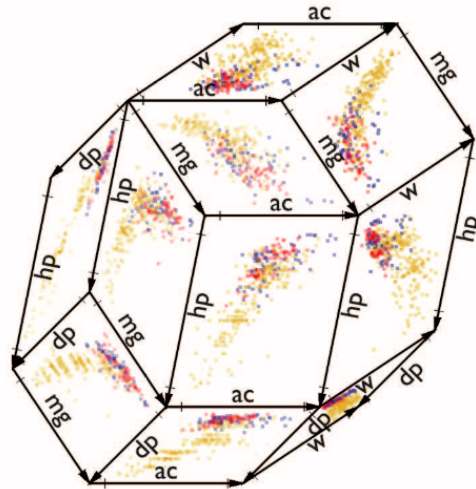


Fig. 1 (d) Example of method for multivariate data visualization, showing the cars dataset, with ac = acceleration, mg = miles per gallon, w = weight, hp = horsepower, yr = model year, or = origin, cy = cylinders, and dp = displacement. Colors denote the origin: yellow for USA; blue for Europe; and red for Japan.

Above visualization ‘hyperbox’ of Alpern and Bowers^[1] is used as an example in the literature cited for this review. This visualization is used for representing multivariate data. It is a representation of N-dimensional box in two dimensions. Due to this transformation to two dimensions, it looks like a multi-faceted polygonal surface. Each parallelogram contains a scatter plot representing relation between two attributes. Each edge of the parallelogram represents an axis. Hence, scatter plots contained within strangely shaped polygon is the most noticeable element of this plot. Axes are repeated over and over again to cover all relationships within the dataset. Colors are used to represent country of origin which is effective in that purpose as long as number countries remain low. Higher numbers will create problems with differentiation.

Representations of N-dimensional box in two dimensions creates major misrepresentation of data instead of creating a better multivariate visualization. The contained scatter plots are drawn within parallelogram on two dimensions, this distortion of data makes it hard to visualize exact relationships. A scatter plot matrix is much better representation of same data without the associated distortion of hyperbox.

Along with the distortion, this examples lacks a legend for the colors used and justifies them in the description of figure instead. It also lacks the information of domain extremes further hiding important piece of information.

Same dataset can be represented with scatterplot matrices, Parallel Coordinate Plots, Star Plots etc in much better form. It is particularly not advisable to represent multidimensional data in two-dimensions. We can improve this representation by adding the ability to switch perspective between different dimensions. This can solve the problem of excessive distortion but at the cost of excessive unnecessary perspective changes, which might hinder the process of correlation discover.

[1] B. Alpern and L. Carter. The hyperbox. In *Proceedings 2nd IEEE Conference on Visualization (Vis'91)*, pages 133–139, 1991