Radial Visualization of Multidimensional and Multivariate Data

Gabriel Zhou

MULTIDIMENSIONAL/MULTIVARIATE

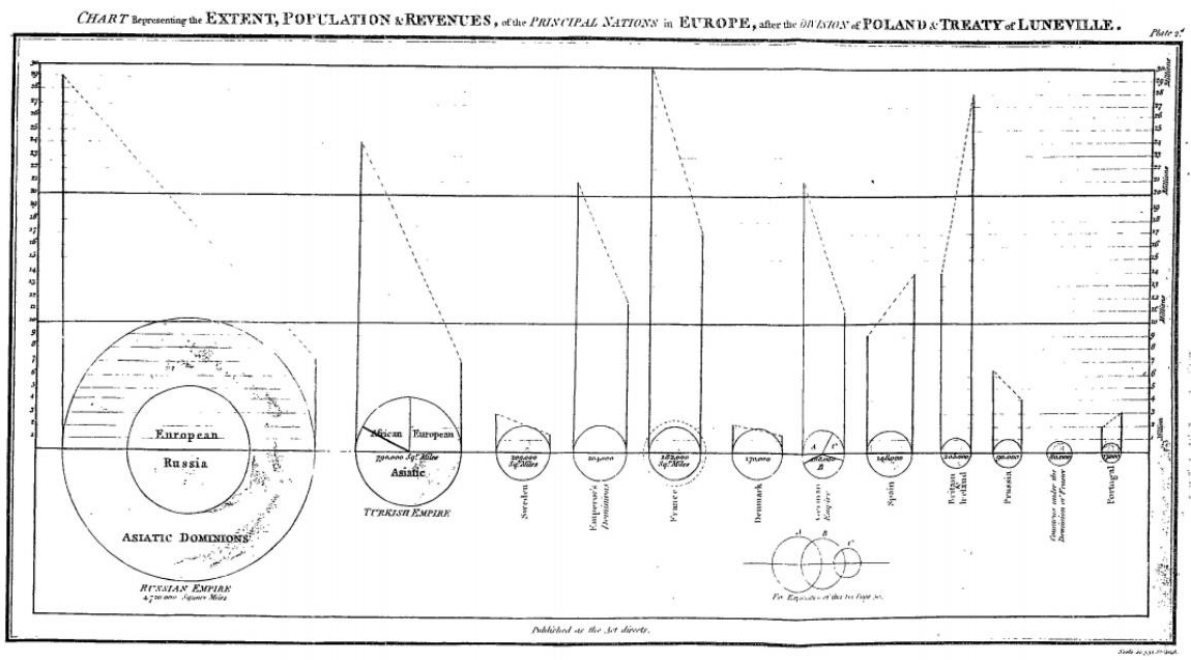
- Think of data as a sample from k-variate function F(x) defined over an n-dimensional domain D. Thus $F = (f_1, f_2, \ldots, f_k)$ has k components, and $X = (x_1, x_2, \ldots, x_n)$ is a point in D. We shall allow k to be zero, in which case we just have a point in D, and we allow n to be zero, in which case we just have a value of F. We shall talk in terms of dependent variables F and independent variables X. In this case, X represents multidimensional data and F represents multivariate data.

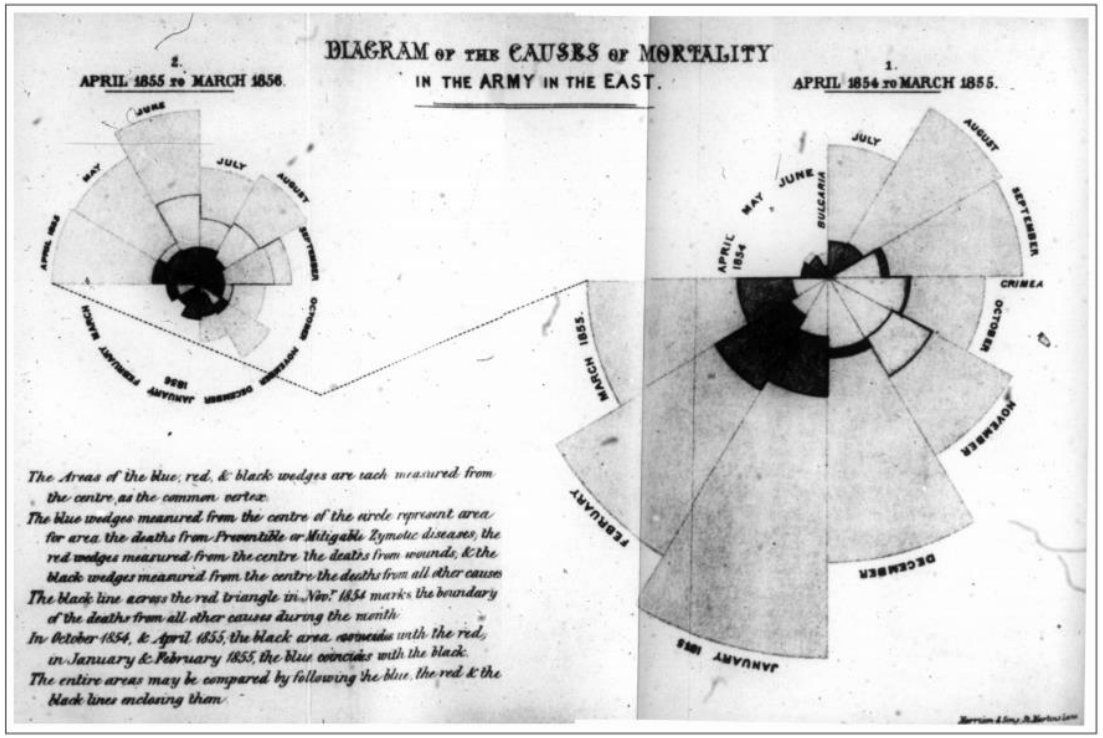
MULTIDIMENSIONAL/MULTIVARIATE

- Key attributes and value attributes
- Key attributes act as index to look up value attributes
- Multidimensional – multiple key attributes
- Multivariate – multiple value attributes

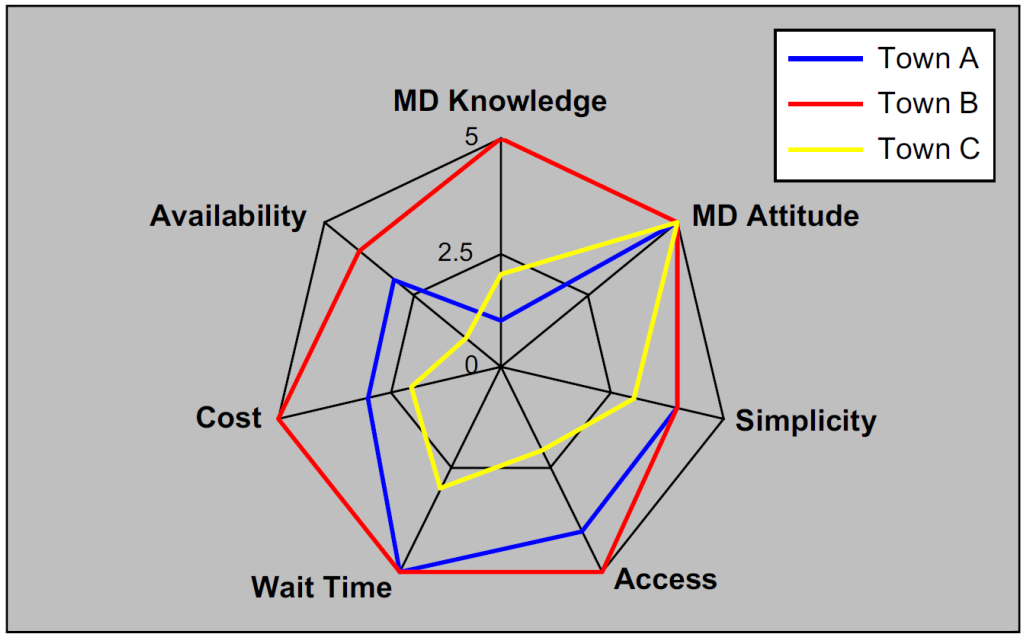
RADIAL VISUALIZATION

- Radial Visualization: information is rendered in a circular or elliptical fashion
- "Radial Visualization" was first mentioned by Hoffman in 1997
- Earliest use of radial display: pie chart by William Playfair in 1801
- Florence Nightingale in 1850s: polar area chart





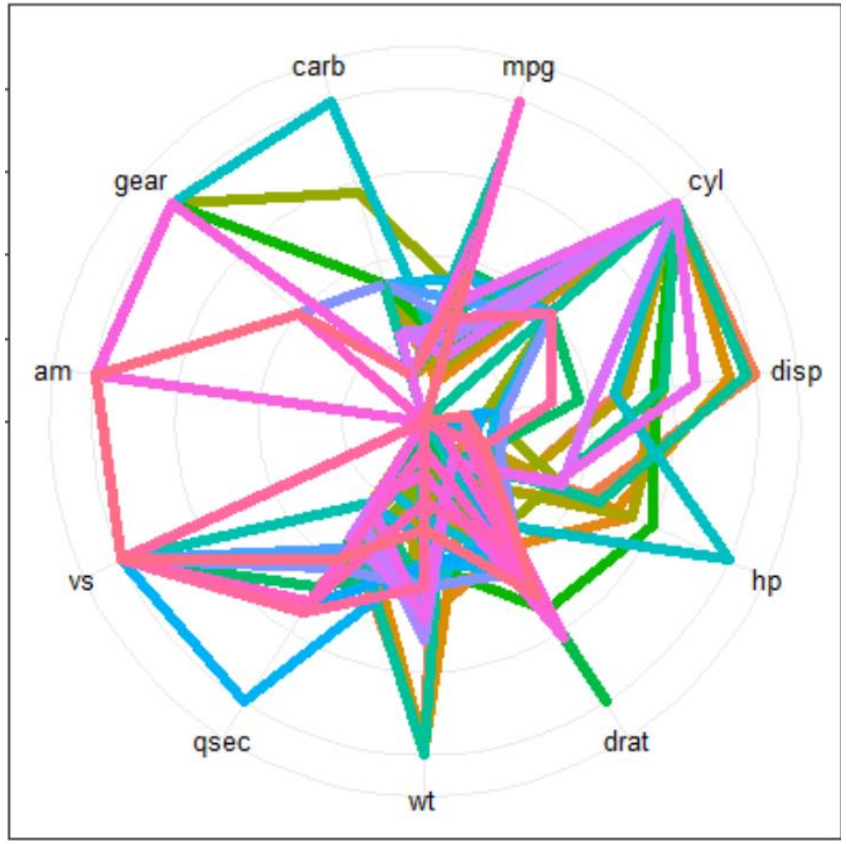
Radar Chart



- A series of spokes or rays projecting from the center point, with each ray representing a different attribute
- The value of the variables are encoded with the lengths of the rays, the values so plotted are sometimes connected to form an enclosed figure
- Dominant perceptual properties often include size and shape of resulting figure
- Parallel coordinates in a radial layout

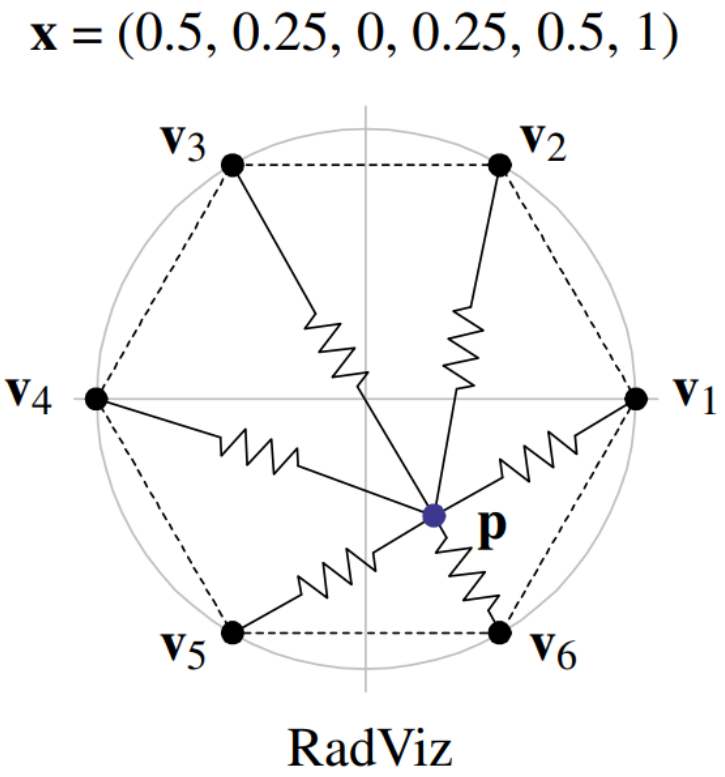
Radar Chart Limitation

- Scalability issue (just like parallel coordinates)
- Comparison across axes
- Misleading to cross-axes comparison when its meaningless
- Size and shape depend on axis sorting



RadViz

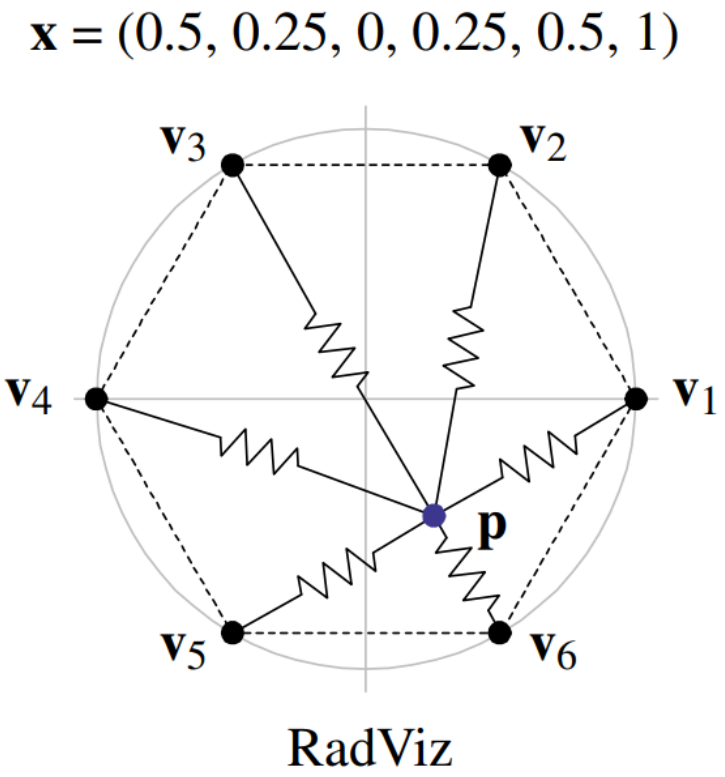
- 1997 by Hoffman
- N-dimensional data points are laid out as points equally spaced around the perimeter of a circle
- Imaginary springs (invisible) connecting perimeter points and the data point
- Spring constant of each spring equals to the value of the data in that coordinate
- All data points values are usually normalized to have values between 0 and 1



RadViz

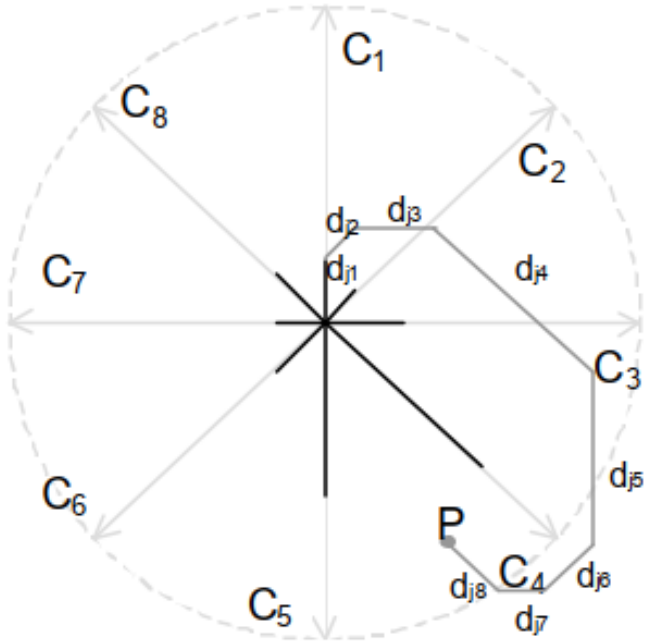
RadViz Limitations

- All dimensional points with difference by a scale number that map into the same position, all points separated on the original space cannot distinguished on the visual space
- Ordering of dimensional anchors is vital, difficult to find the optimal order for a good visualization view



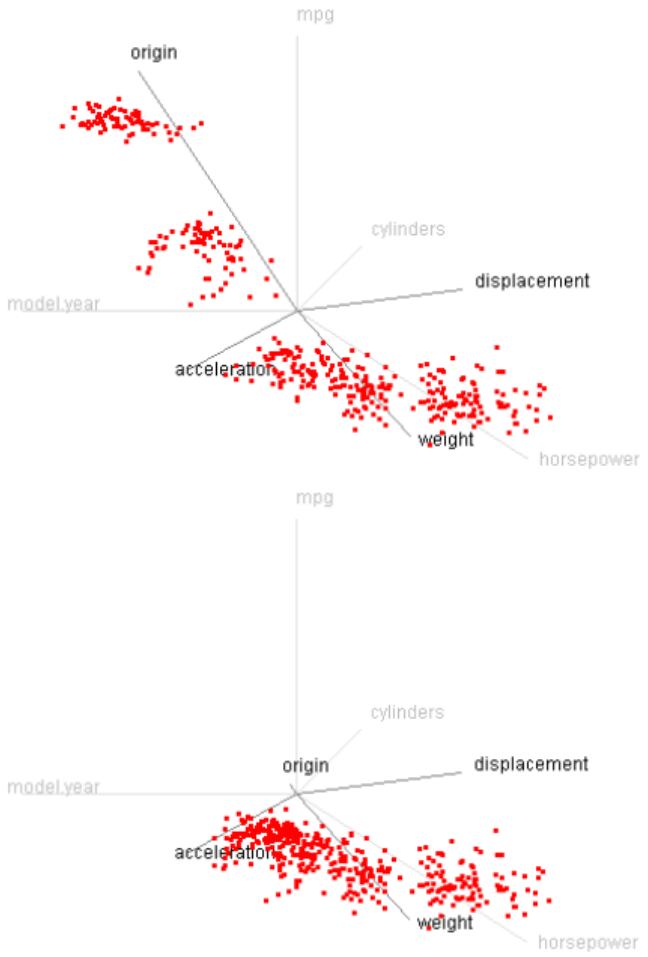
Star Coordinates

- An extension of typical scatterplots to higher dimensions
- Arrange the coordinate axes on a circle with equal angles and equal length (initially)
- The mapping of a data element is determined by the sum of all unit vectors on each coordinate multiplied by the value of the data element for that coordinate



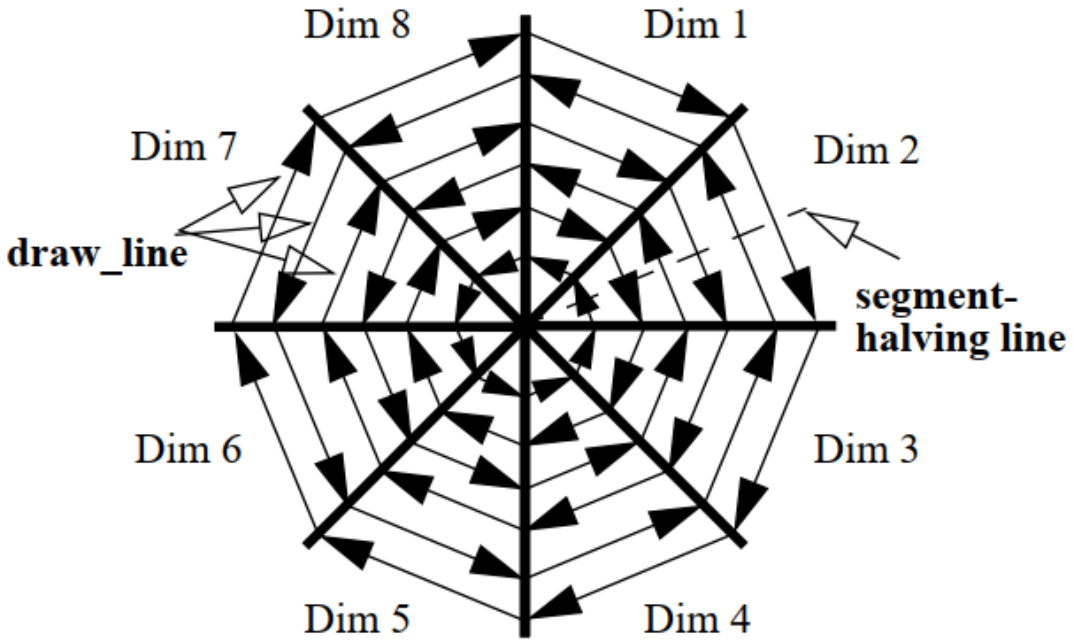
Star Coordinates Limitations

- Ambiguity (similar issue as RadViz)
- Can be solved (reduced) by interactions
- Scaling:
  - changing the length of axes
  - Increase or decrease the contribution of a particular axis on the resultant visualization
- Rotation:
  - changing the direction of axes
  - Changing the correlations of a particular data axis to other axes



Circle Segment

- Pixel per value technique
- Display the data dimensions as segments of a circle
- The data items within one segment are arranged in back and forth manner along “draw line”
- Coloring maps high data values to light colors and low data value to dark colors



Circle Segment Limitations

- Outer data may receive undue emphasis due to their greater radius (arc length is derived directly from the radius)