Visualizing Tables

Tables

- Rows of records (items), and each item has one or multiple columns
- A table has *keys* and *values*
  - Keys: independent attributes that can look up items in the table
  - Values: dependent attributes, i.e. ,the values of cells in the table

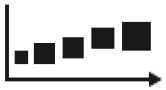
Arrange Tables

Use of spatial channels for visual encoding

- Express Values

- Separate, Order, Align Regions
  - Separate

  - Order



  - Align

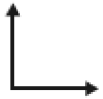
    - *1 Key* — *List*

    - *2 Keys* — *Matrix*

    - *3 Keys* — *Volume*

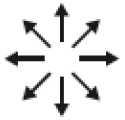
    - *Many Keys* — *Recursive Subdivision*

- Axis Orientation
  - Rectilinear

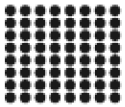


  - Parallel

  - Radial



- Layout Density
  - Dense



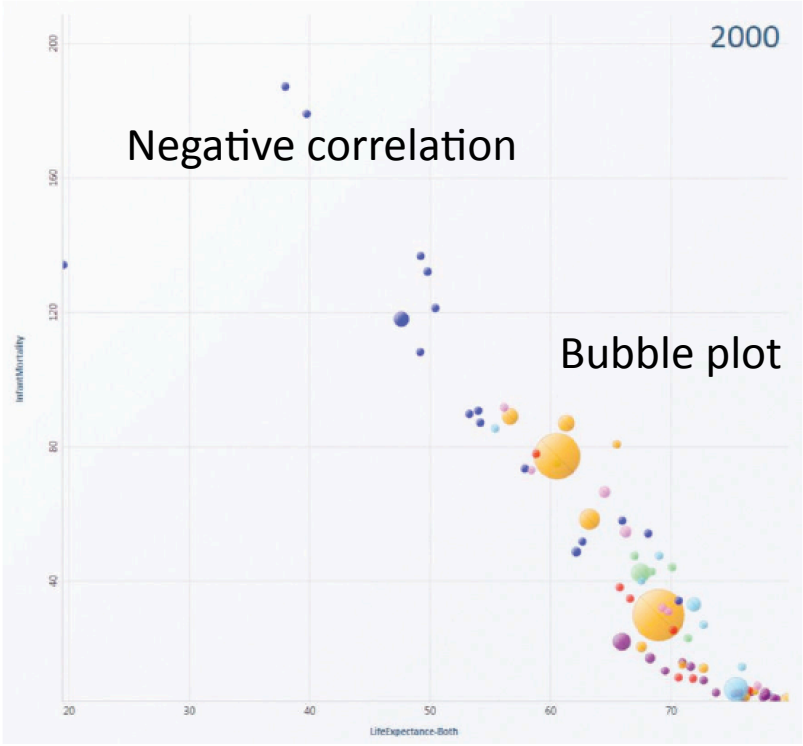
  - Space-Filling

Express Quantitative Values

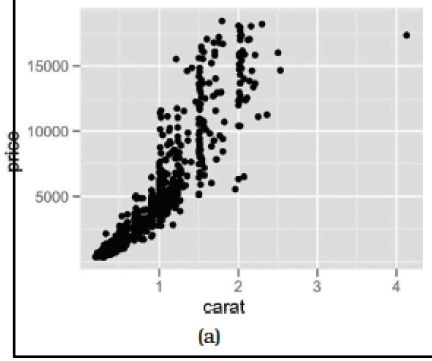
- Map the value to a spatial position along an axis
- Additional attributes can be encoded on the same mark with other nonspatial channels such as color and size.
- Example: scatterplots
  - Encode two quantitative attributes using the vertical and horizontal position channel
  - Mark: points
  - Tasks: finding outliners, extreme values, and correlation between the attributes

Example: Scatterplots

Idiom	Scatterplots
What: Data	Table: two quantitative value attributes.
How: Encode	Express values with horizontal and vertical spatial position and point marks.
Why: Task	Find trends, outliers, distribution, correlation; locate clusters.
Scale	Items: hundreds.



Derived quantities used to create Scatteplots sometimes can show attribute correlation



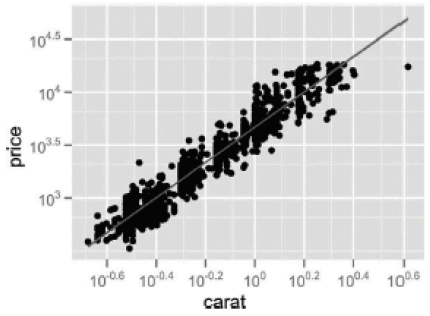


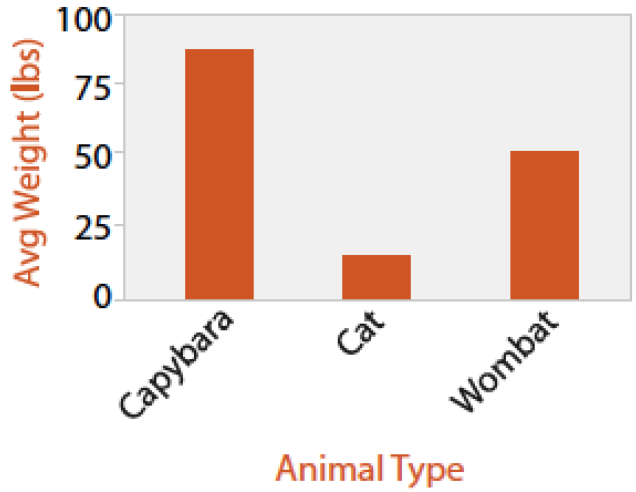
Figure 7.2. Scatterplot. Each point mark represents a country, with horizontal and vertical spatial position encoding the primary quantitative attributes of life expectancy and infant mortality. The color channel is used for the categorical country attribute and the size channel for quantitative population attribute. From [Robertson et al. 08, Figure 1c].

Encode Categorical Attributes Separate, Order and Align

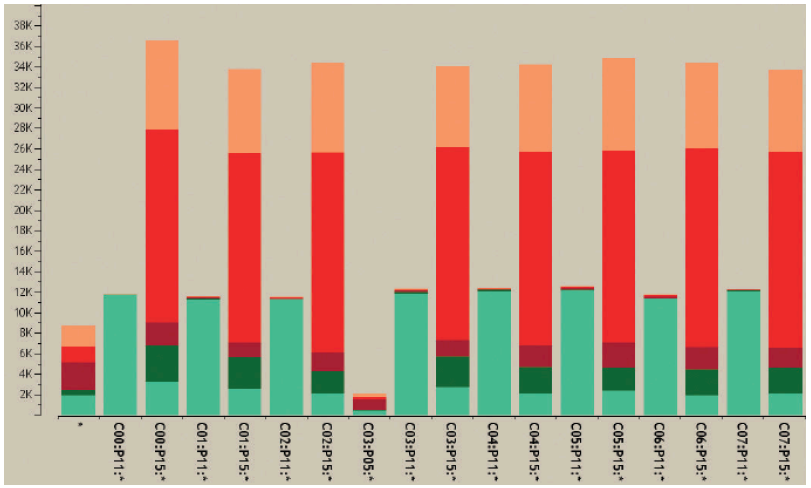
- Spatial position is not the best channel because it often implies ordered magnitude
- But spatial proximity or region matches well with categorical data
- *Separate* data with a categorical attribute into regions
- *Order* the regions based on an ordered attribute
- *Align* the regions and also items within a region to encode yet another ordered attributes

Example: Bar Charts

Idiom	Bar Charts	Stacked Bar Charts
What: Data	Table: one quantitative value attribute, one categorical key attribute.	Multidimensional table: one quantitative value attribute, two categorical key attributes.
How: Encode	Line marks, express value attribute with aligned vertical position, separate key attribute with horizontal position.	Bar glyph with length-coded subcomponents of value attribute for each category of secondary key attribute. Separate bars by category of primary key attribute.
Why: Task	Lookup and compare values.	Part-to-whole relationship, lookup values, find trends.
Scale	Key attribute: dozens to hundreds of levels.	Key attribute (main axis): dozens to hundreds of levels. Key attribute (stacked glyph axis): several to one dozen



Bar Charts



Stacked Bar Charts

Example: StreamGraphs

Idiom	Streamgraphs
What: Data	Multidimensional table: one quantitative value attribute (counts), one ordered key attribute (time), one categorical key attribute (artist).
What: Derived	One quantitative attribute (for layer ordering).
How: Encode	Use derived geometry showing artist layers across time, layer height encodes counts.
Scale	Key attributes (time, main axis): hundreds of time points. Key attributes (artists, short axis): dozens to hundreds

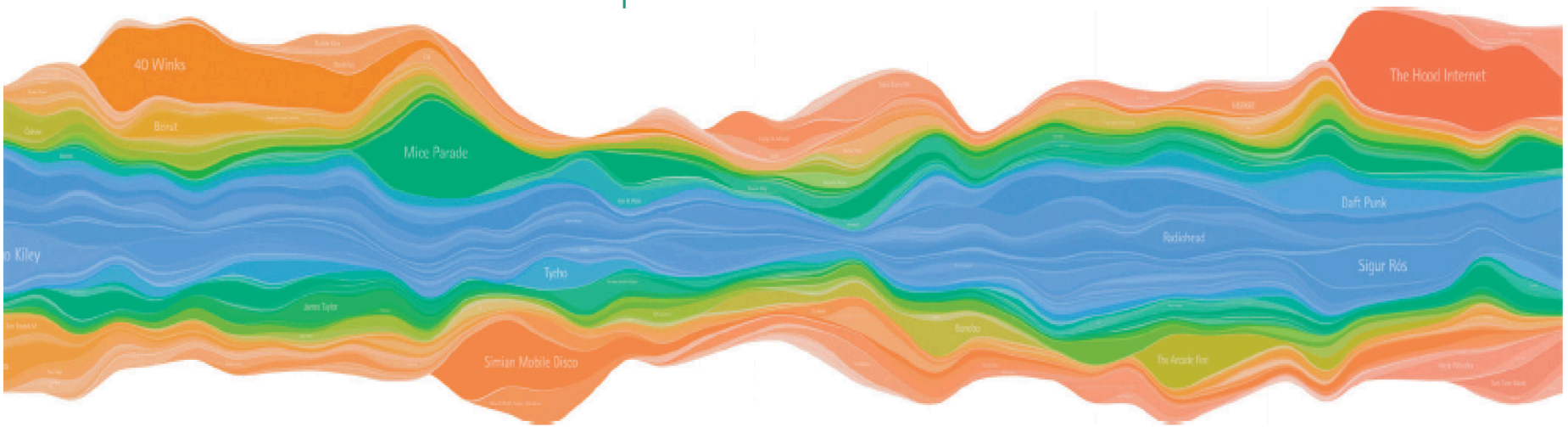
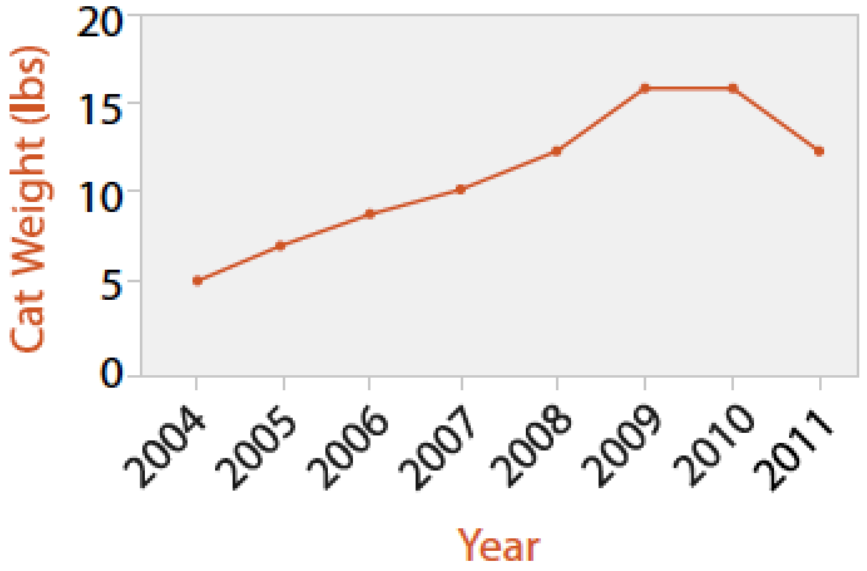


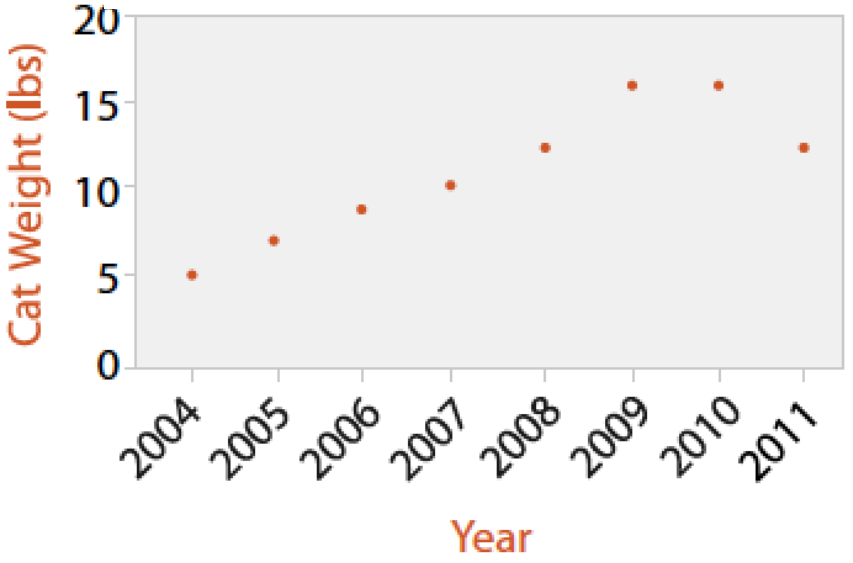
Figure 7.6. Streamgraph of music listening history. From [Byron and Wattenberg 08, Figure 0].

Dot Charts and Line Charts

- Line and dot charts should use ordered keys but not categorical keys
- Banking to the 45 - adjust the aspect ratio so that a majority of line segments are of 45 degrees

Idiom	Line Charts	Dot Charts
What: Data	Table: one quantitative value attribute, one ordered key attribute.	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Dot chart with connection marks between dots.	Express value attribute with aligned vertical position and point marks. Separate/order into horizontal regions by key attribute.
Why	Show trend.	
Scale	Key attribute: hundreds of levels.	



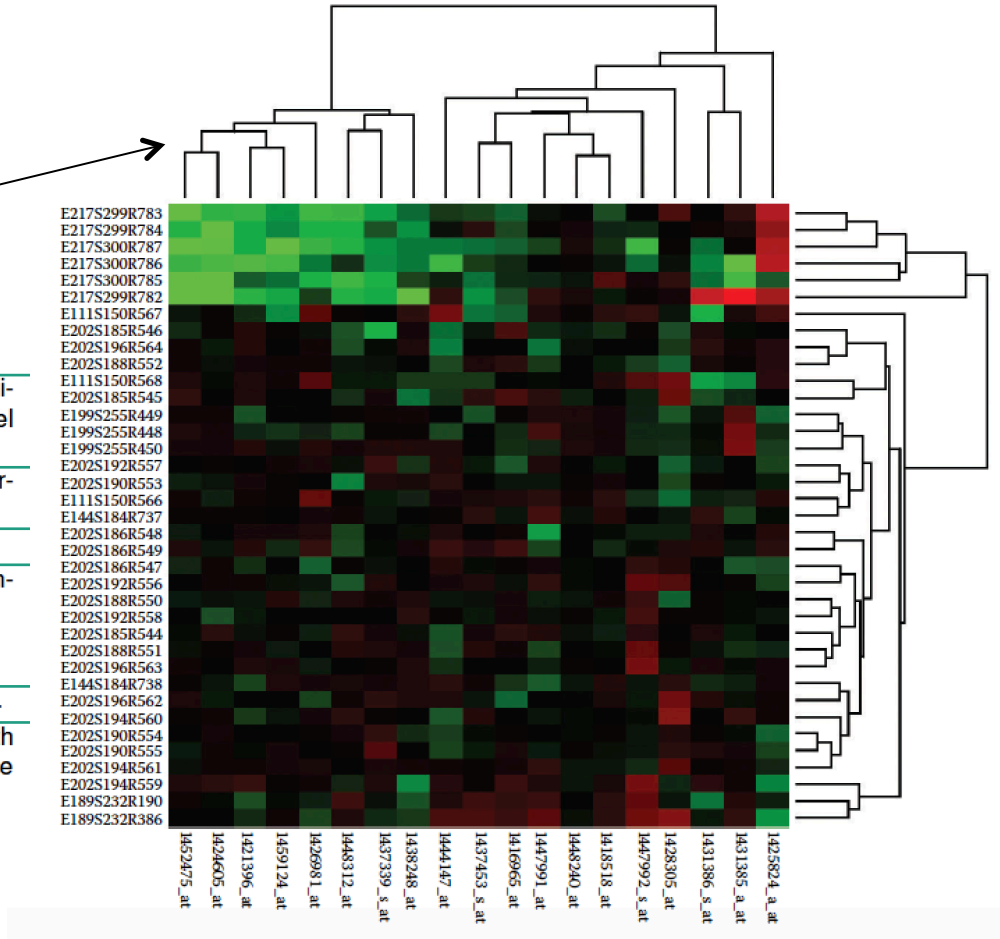


Align Two Keys: Matrix

- Arrange data with two keys
- Example: Heatmap

Cluster Heatmap

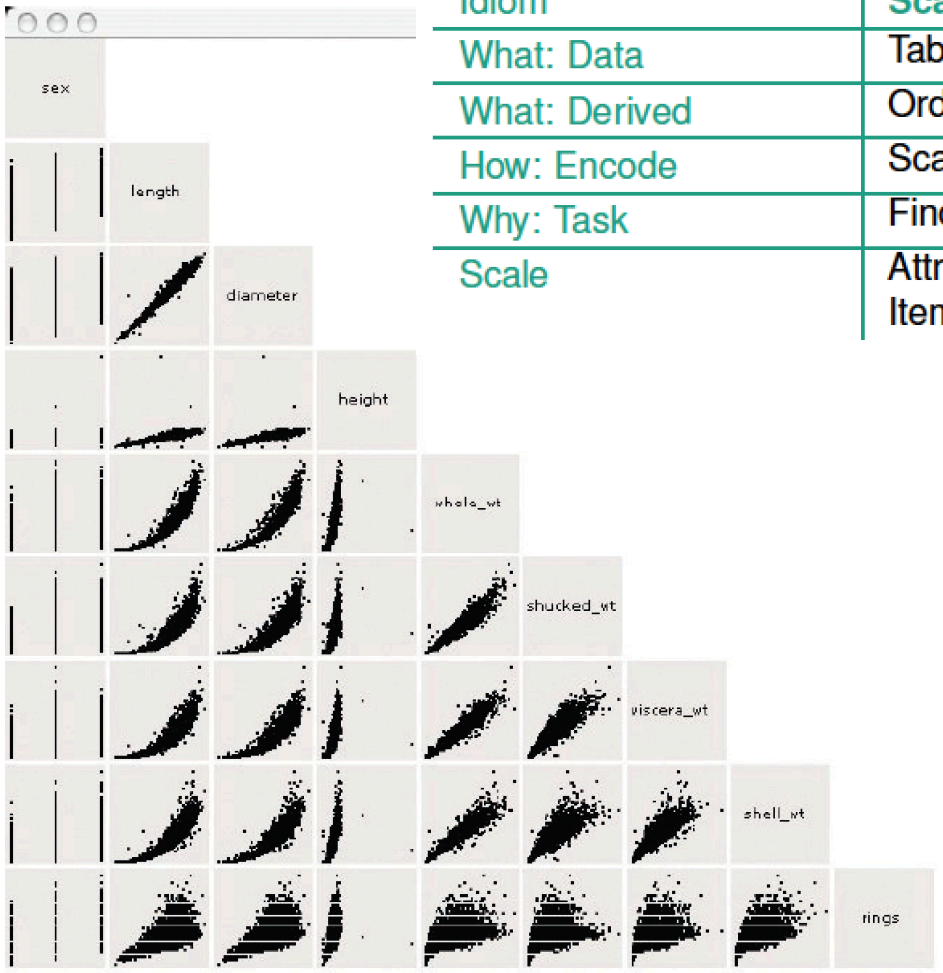
Dendrogram



Idiom	**Heatmaps**
What: Data	Table: two categorical key attributes (genes, conditions), one quantitative value attribute (activity level for gene in condition).
How: Encode	2D matrix alignment of area marks, diverging colormap.
Why: Task	Find clusters, outliers; summarize.
Scale	Items: one million. Categorical attribute levels: hundreds. Quantitative attribute levels: 3–11.

Idiom	**Cluster Heatmaps**
What: Derived	Two cluster hierarchies for table rows and columns.
How: Encode	Heatmap: 2D matrix alignment, ordered by both cluster hierarchies. Dendrogram: connection line marks for parent–child relationships in tree.

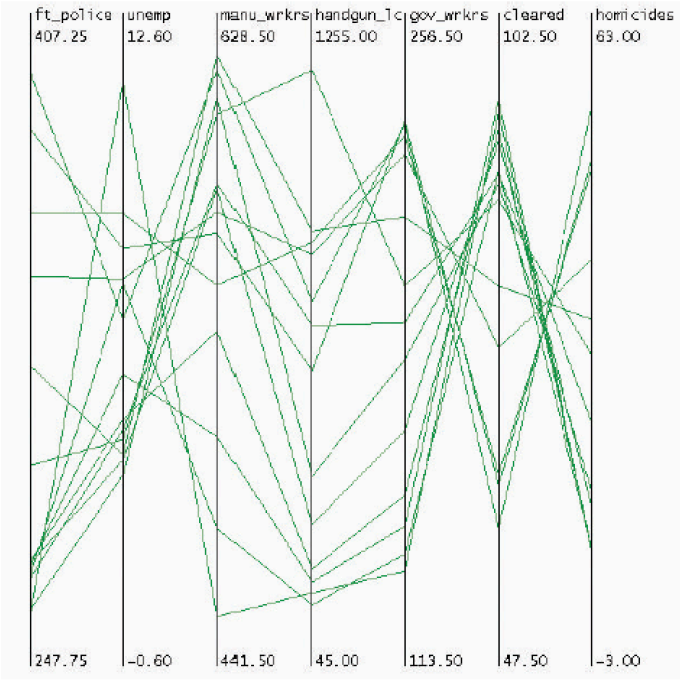
Example: Scatterplot Matrix (SPLOM)



Idiom	**Scatterplot Matrix (SPLOM)**
What: Data	Table.
What: Derived	Ordered key attribute: list of original attributes.
How: Encode	Scatterplots in 2D matrix alignment.
Why: Task	Find correlation, trends, outliers.
Scale	Attributes: one dozen. Items: dozens to hundreds.

Spatial Axis Orientation

- Rectilinear layout: horizontal and vertical axes
  - Everything we have seen so far
- Parallel layout: most useful when there are more than 2 axes
- Parallel Coordinates
  - Arrange axes vertically in parallel
  - A item with multiple attributes is plotted by connecting the values in the axes



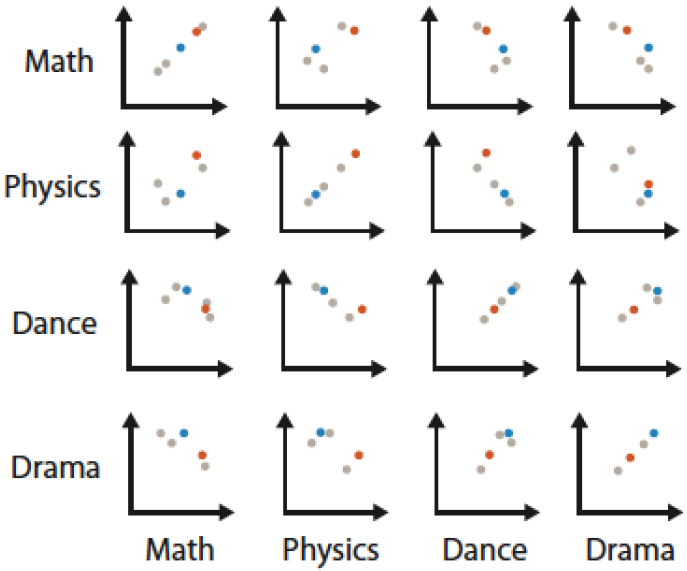
Parallel Coordinates Plots (PCP)

- Combine all attributes in a single plot

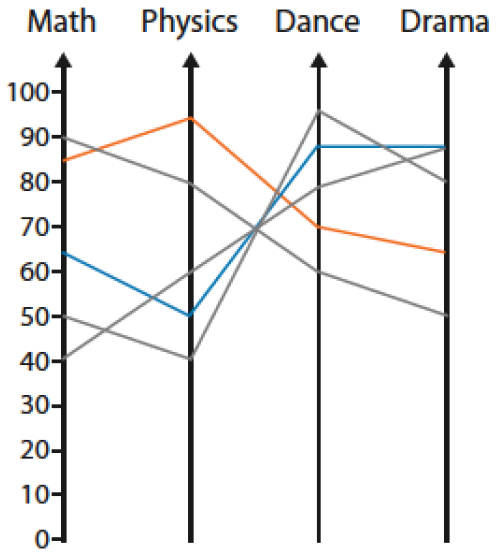
Table

Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

Scatterplot Matrix

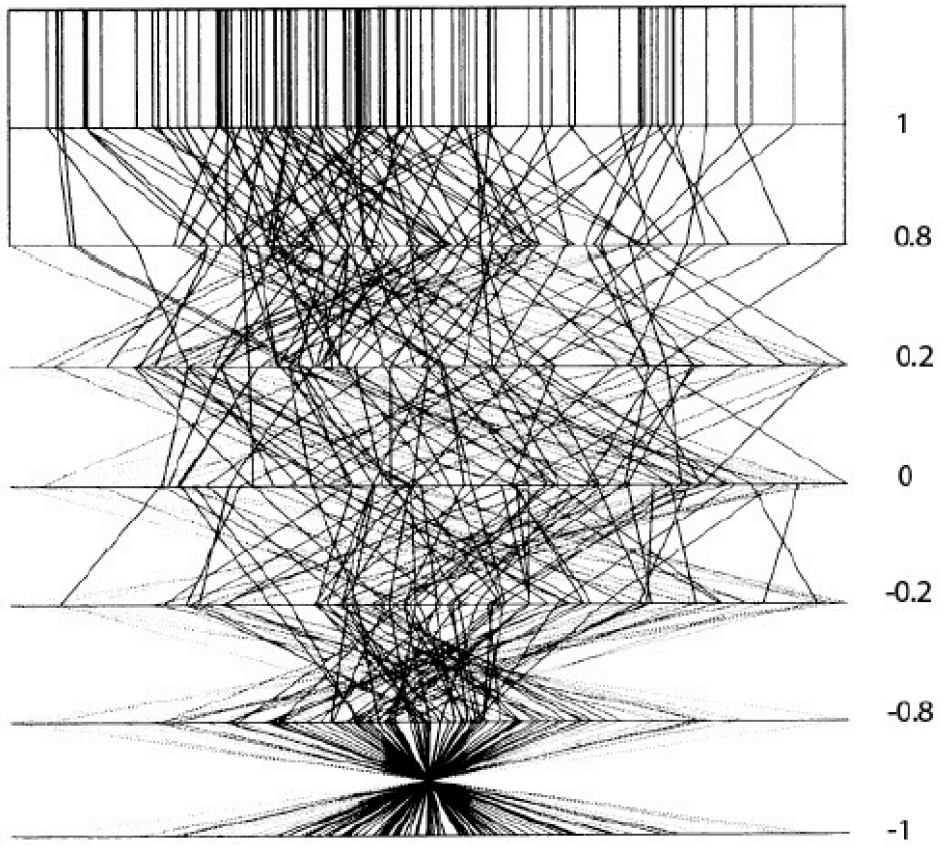


Parallel Coordinates



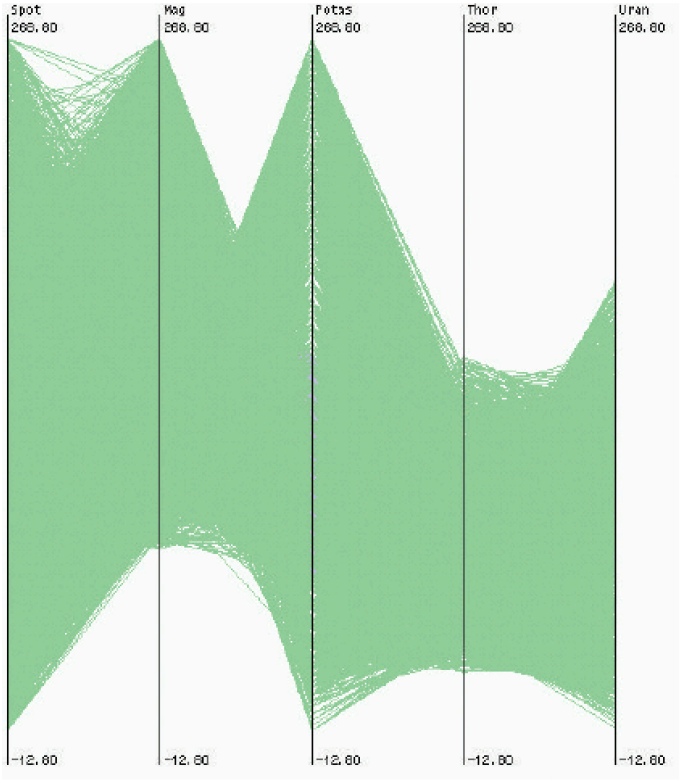
Parallel Coordinates Plots (PCP)

- Can be used to show correlations of variables



Parallel Coordinates Plots (PCP)

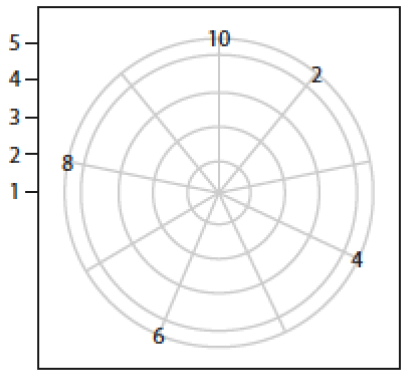
- Difficult to scale to thousands of items



Idiom	**Parallel Coordinates**
What: Data	Table: many value attributes.
How: Encode	Parallel layout: horizontal spatial position used to separate axes, vertical spatial position used to express value along each aligned axis with connection line marks as segments between them.
Why: Tasks	Find trends, outliers, extremes, correlation.
Scale	Attributes: dozens along secondary axis. Items: hundreds.

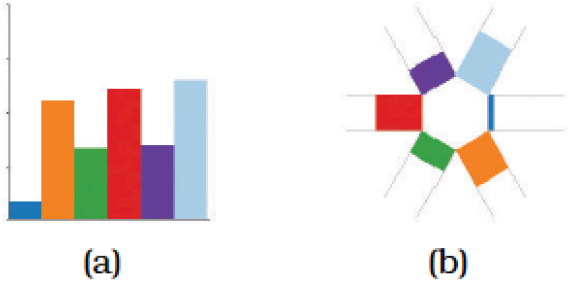
Radial Layouts

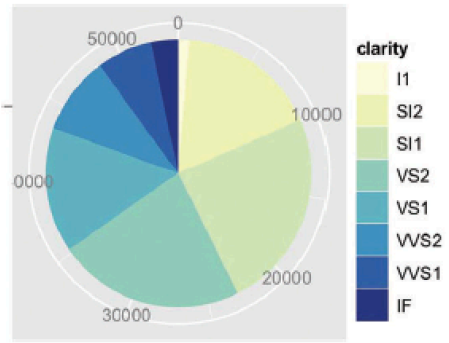
- Items are distributed around a circle using the angle channel
- Often use polar coordinate system (θ and r) angle around the circle and the radius
  - The angle channel is less accurately perceived than position
  - It is inherently cyclic
  - The two attributes are not symmetric
  - Good at showing periodic pattern

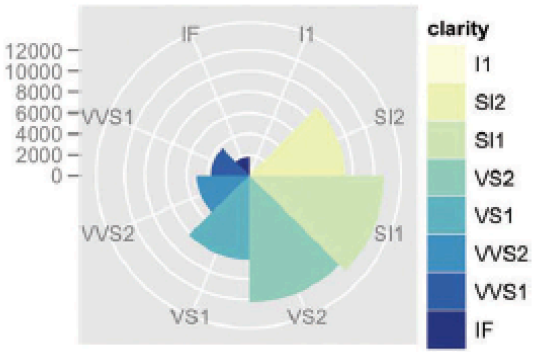


Polar coordinate system

Radial Bar Charts and Pie Charts







Idiom	**Radial Bar Charts**
What: Data	Table: one quantitative attribute, one categorical attribute.
How: Encode	Length coding of line marks; radial layout.

Idiom	**Pie Charts**
What: Data	Table: one quantitative attribute, one categorical attribute.
Why: Task	Part–whole relationship.
How: Encode	Area marks (wedges) with angle channel; radial layout.
Scale	One dozen categories.

Idiom	**Polar Area Charts**
What: Data	Table: one quantitative attribute, one categorical attribute.
Why: Task	Part–whole relationship.
How: Encode	Area marks (wedges) with length channel; radial layout.
Scale	One dozen categories.

- Pie Charts are problematic in the sense that area perception is not as sensitive as length perception
- Polar Pie Charts can fix this problem
- Both Pie Charts can show the part contribution to the sum but this is difficult for bar charts