

Information Visualization and Visual Analytics (M1522.000500)

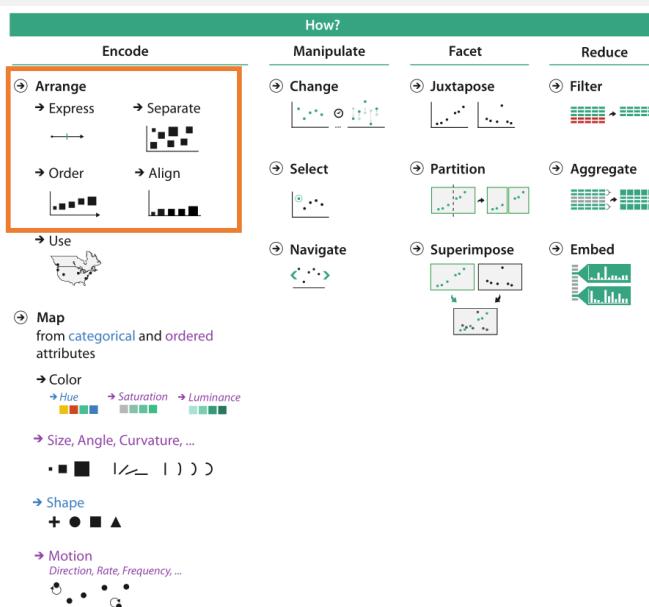
# Arrange Tables

Jinwook Seo, Ph. D.

Professor, Dept. of Computer Science and Engineering  
Seoul National University

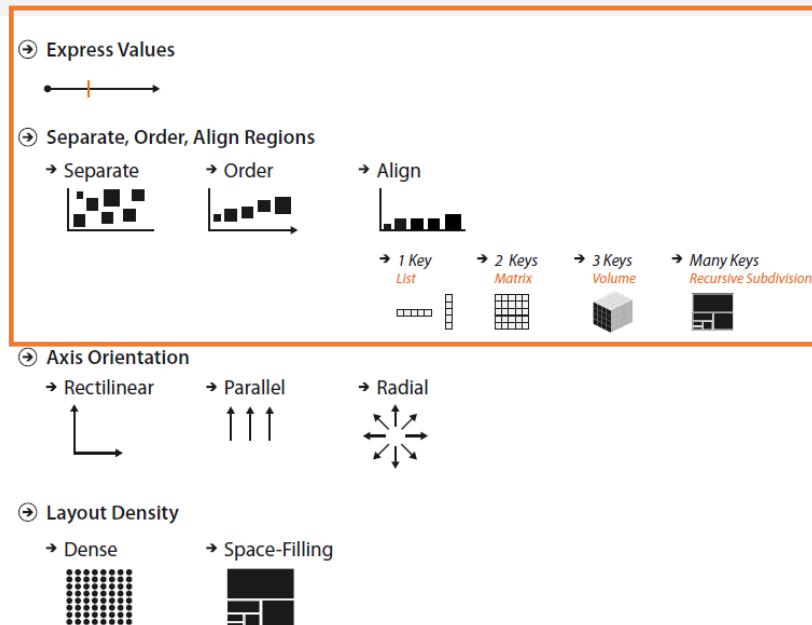
## How-Encode-Arrange

### How-Encode-Arrange



## Arrange Tables

### The Big Picture – Design Choices for Arranging Tables



Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables

### Why Arrange?

- **Arrange**

- Most crucial visual encoding choice
  - *Why?* The use of space dominates the user's mental model of the dataset

- **Arrange by Keys and Values**

**Key Attribute**  
Independent attribute

- Can be used as a unique index to look up items in a table
- Categorical or ordinal

**Value Attribute**  
Dependent attribute

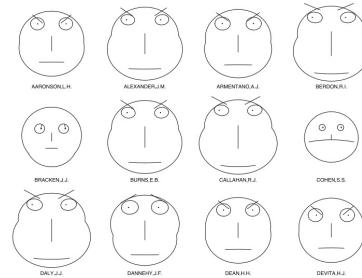
- The value of a cell in a table
- All three types: categorical, ordinal, or quantitative

Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables

### Express: Quantitative Values

- Use space to express quantitative attributes



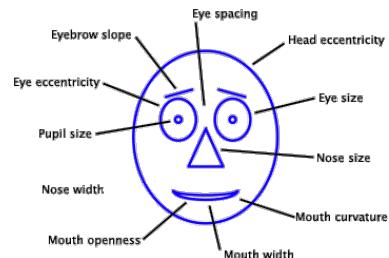
- Simple case of encoding value attributes

Encoding : a mark at some position along the axis

+ other non-spatial channels (color, size..)

- Complex case of encoding value attributes

Encoding : a composite glyph object showing multiple attributes

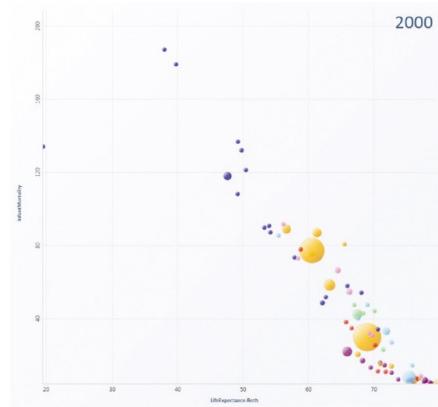
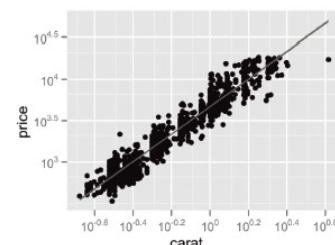
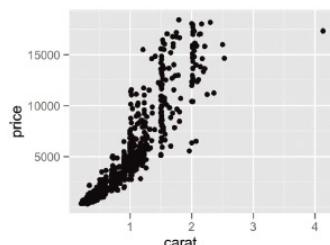


Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Express: Quantitative Values

### Idiom: scatterplot

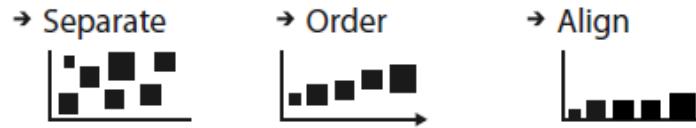
What: Data	Table: two quantitative value attributes
How: Encode (Mark & Channel)	Express values with <b>horizontal</b> and <b>vertical spatial position</b> and <b>point marks</b> .
Why: Task	Find trends, outliers, distribution, correlation; Locate clusters and outliers.
Scale	Items: hundreds.



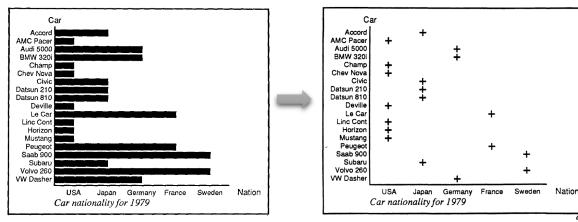
Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables

Some keys: Categorical regions



- **regions:** contiguous bounded areas distinct from each other
  - using **space** to **separate** (proximity)
  - following **expressiveness principle** for categorical attributes
    - identity channels for (**unordered**) categorical attributes

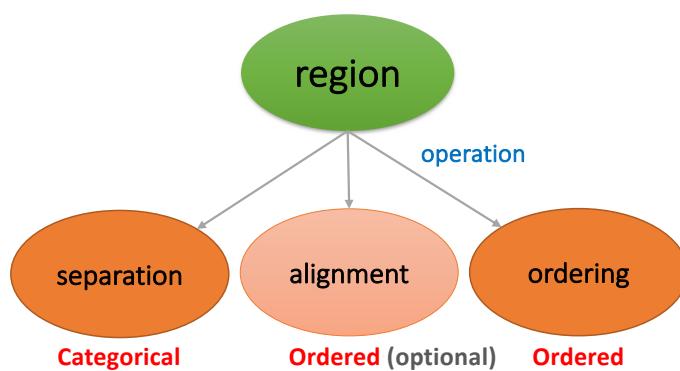
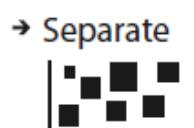


Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables

Separate, Order, and Align: Categorical Regions

- Categorical attribute → spatial region
  - each level is assigned to a distinct contiguous bounded area
- Use ordered attributes to **order** and **align** regions



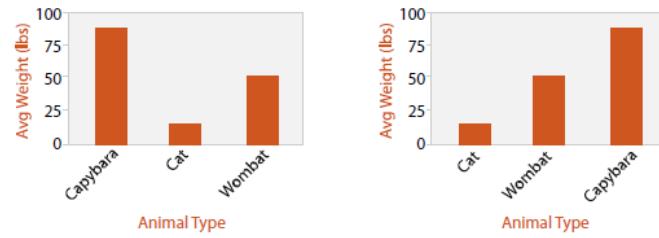
Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Idiom: bar chart

What: Data	Table: one quantitative value attribute, one categorical key attribute.
How: Encode (Mark & Channel)	<b>Line marks</b> , express value attribute with <b>aligned vertical position</b> , <b>separate</b> key attribute with <b>horizontal position</b> .
Why: Task	Lookup and compare values.
Scale	Key attribute: dozens to hundreds of levels.

- One key attribute
- channels
  - position to express quantitative value
  - spatial regions: one per mark
    - separated horizontally, aligned vertically
    - can be ordered by quantitative attribute
      - by label (alphabetical), by length attribute (data-driven)



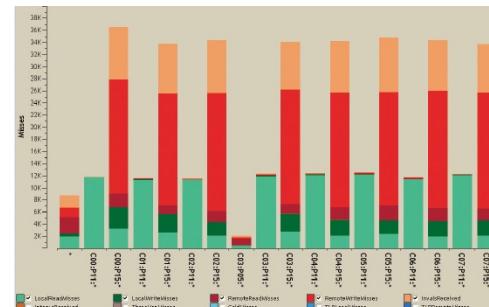
Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Idiom: stacked bar chart

What: Data	Multidimensional table: one quantitative value attribute, two categorical key attributes.
How: Encode (Mark & Channel)	<b>Bar glyph</b> with length-coded subcomponents of value attribute for each category of secondary key attribute. ( <b>vertical</b> ) <b>Separate bars</b> by category of primary key attribute. ( <b>horizontal</b> )
Why: Task	Part-to-whole relationship, lookup values, find trends.
Scale	Key attribute (main axis): dozens to hundreds of levels. Key attribute (stacked glyph axis): several to one dozen

- One more key
- channels
  - length and color hue
  - spatial regions: one per glyph
    - aligned: full glyph, lowest bar component
    - e.g., [dancing histograms](#)
    - unaligned: other bar components

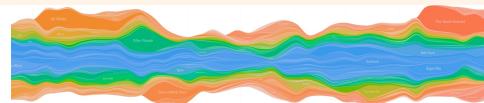


Information Visualization and Visual Analytics – Arrange Tables

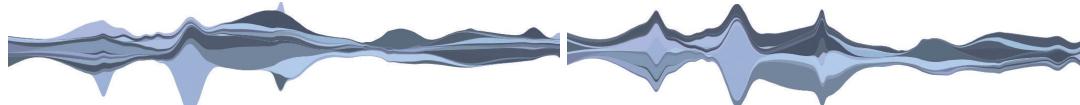
## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Idiom: streamgraph

What: Data	Multidimensional table: one quantitative value attribute (counts), one ordered key attribute (time), one categorical key attribute (artist). one derived attribute - quantitative attribute (for layer ordering).
How: Encode	Use derived geometry showing artist layers across time, layer height encodes counts.
Why: Task	Find temporal trends
Scale	Key attributes (time, main axis): hundreds of time points. Key attributes (artists, short axis): dozens to hundreds (better than stacked bars, since most layers don't extend across whole chart)



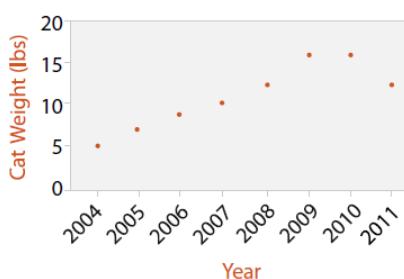
- [generalized stacked graph](#)
- emphasizes the **legibility of the individual streams** with a deliberately organic silhouette, rather than using the horizontal axis as the baseline



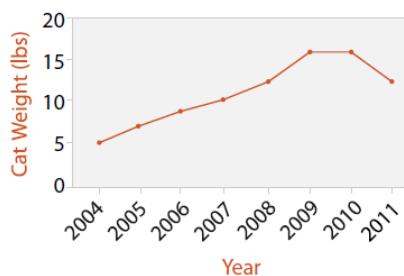
Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Idiom: dot chart and line chart



What: Data	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Express value attribute with aligned <b>vertical position</b> and <b>point marks</b> . Separate/order into horizontal regions by key attribute.
Why: Task	Find temporal trends
Scale	Key attribute: hundreds of levels

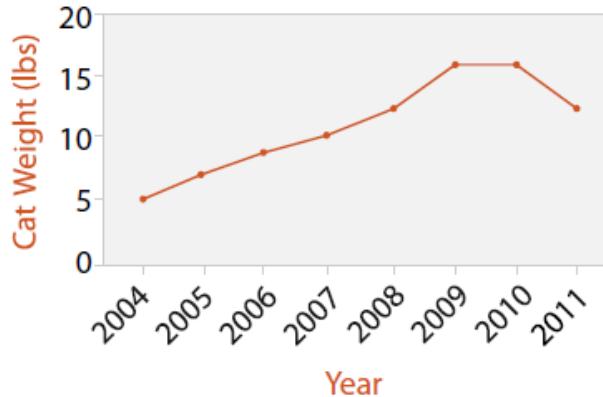


What: Data	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Dot chart with connection marks between dots.
Why: Task	Find temporal trends
Scale	Key attribute: hundreds of levels

Information Visualization and Visual Analytics – Arrange Tables

## Idiom: line chart

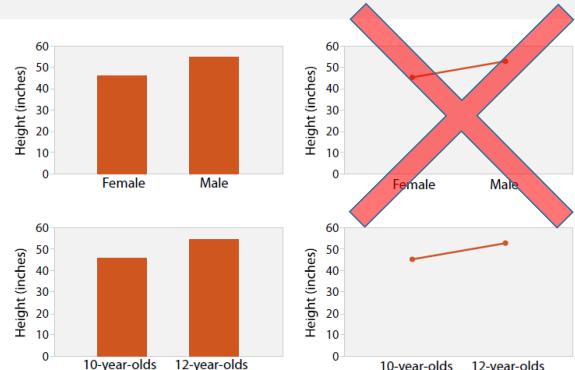
- one key, one value
- data
  - 2 quantitative attributes
- mark: points
  - line connection marks between them
- channels
  - aligned lengths (i.e., position) to express quantitative value
  - separated and ordered by key attribute into horizontal regions
- task
  - find trends
  - connection marks emphasize ordering of items along key axis by explicitly showing relationship between one item and the next



Information Visualization and Visual Analytics – Arrange Tables

## Choosing bar vs line charts

- depends on **type of key attribute**
  - **bar charts if categorical**
  - **line charts if ordered**
- do not use line charts for categorical key attributes
  - violates expressiveness principle
    - **implication of trend so strong that it overrides semantics!**
      - “The more male a person is, the taller he/she is”

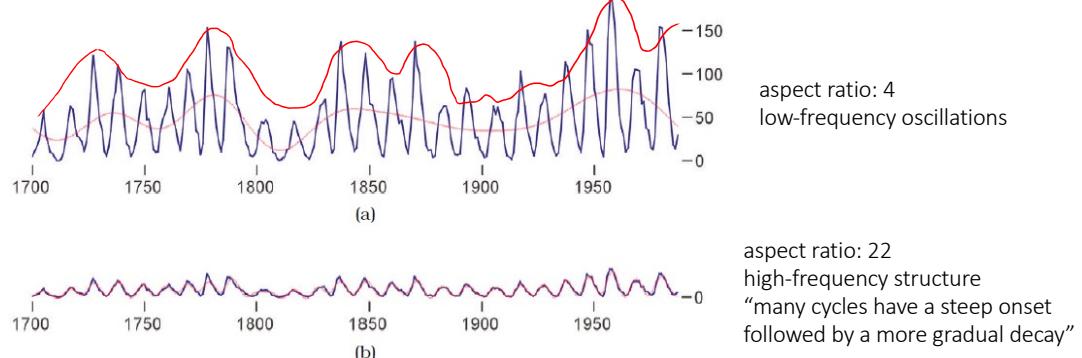


Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Idiom: banking to 45° (Line chart – consider aspect ratio)

- aspect ratio → reveal/hide “structures” in dataset
- our ability to judge angles is more accurate at exact diagonals than at arbitrary directions ( $45^\circ : 43^\circ$  vs  $22^\circ : 20^\circ$ )
- computes the best aspect ration for a line chart to maximize the number of line segments that fall close to the diagonal



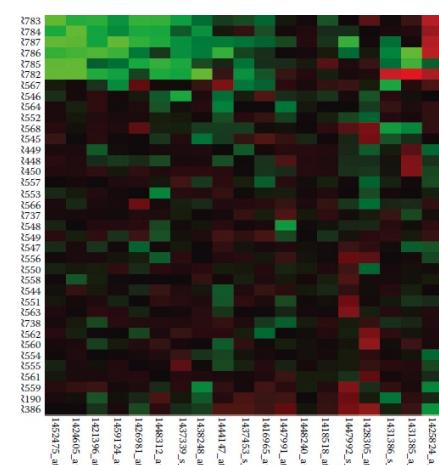
Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Idiom: heatmap

- Matrix Alignment: Two Keys

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.



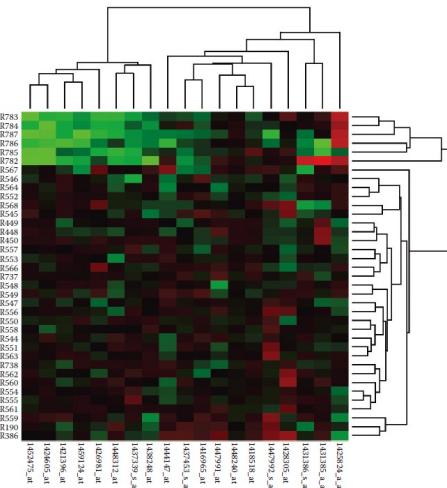
Information Visualization and Visual Analytics – Arrange Tables

## Arrange Tables - Separate, Order, and Align: Categorical Regions

## Idiom: cluster heatmap

- Derived attributes: two hierarchies

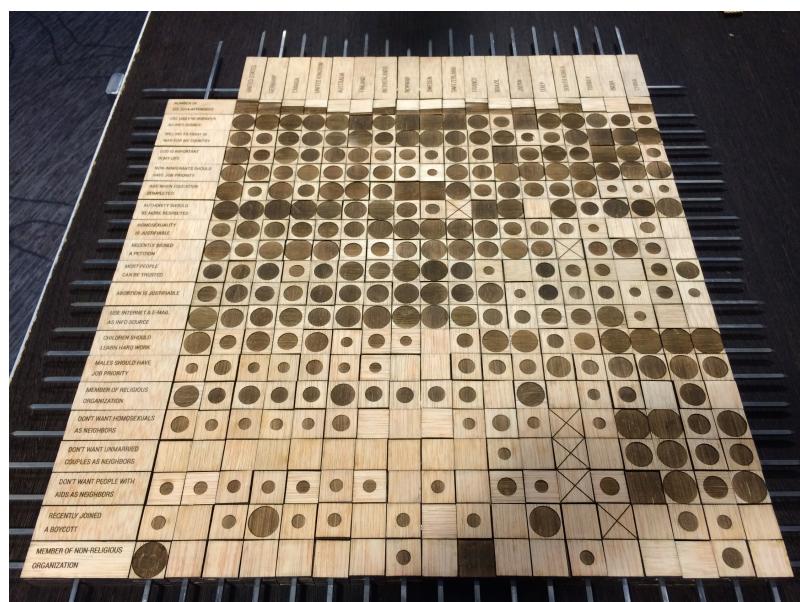
What: Data	<p>Table: two categorical key attributes (genes, conditions), one quantitative value attribute (activity level for gene in condition).</p> <p>Derived: Two cluster hierarchies for table rows and columns.</p>
How: Encode	<p>Heatmap: 2D matrix alignment, ordered by both cluster hierarchies.</p> <p>Dendrogram: connection line marks for parent–child relationships</p>



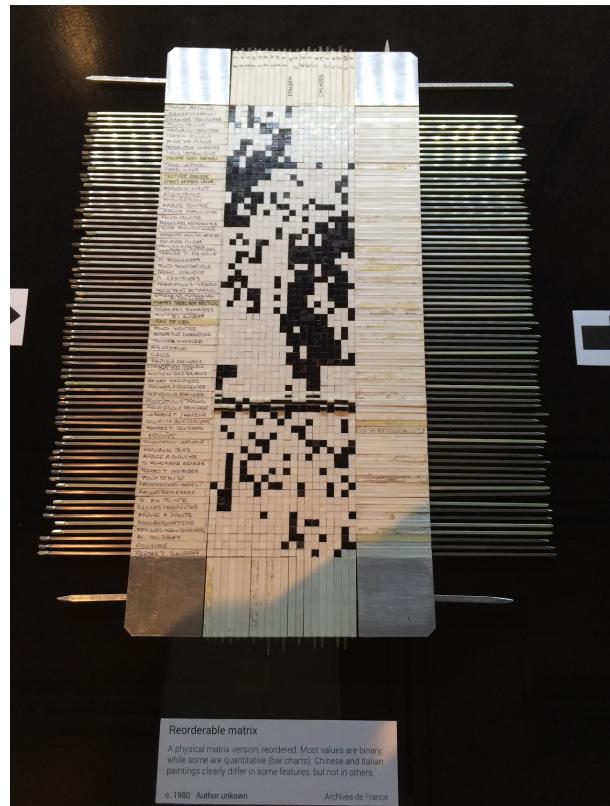
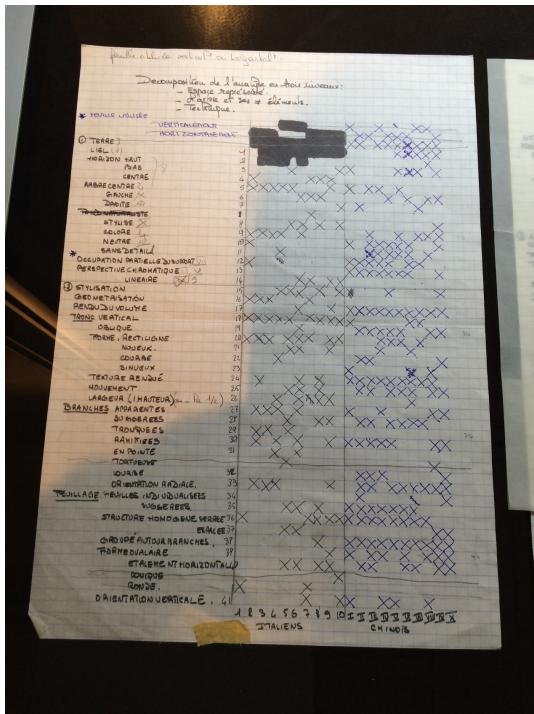
Information Visualization and Visual Analytics – Arrange Tables

## Bertin's Reorderable Matrix

## Bertin's Reorderable Matrix



Information Visualization and Visual Analytics – Arrange Tables



19

## Arrange Tables - Separate, Order, and Align: Categorical Regions

### Multiple Keys

- Volumetric Grid: Three Keys
  - 3D volumetric grid
  - Not recommended because of perceptual problems
- Recursive Subdivision: Multiple Keys
  - Possible to extend by recursively subdividing the cell within a list or matrix
  - Section 12.4

## Spatial Axis Orientation

- Design choice with the use of space

- How to orient the spatial axes

## Layout

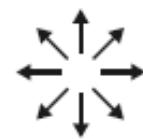
## → Rectilinear



## → Parallel



## → Radial

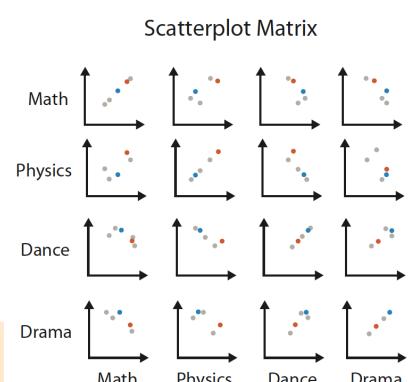


Information Visualization and Visual Analytics – Arrange Tables

## Idioms: scatterplot matrix (SPLOM)

- scatterplot matrix (SPLOM)

- rectilinear axes, point mark
- all possible pairs of axes
- scalability
  - one dozen attributes
  - dozens to hundreds of items



What: Data	Table 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.

Information Visualization and Visual Analytics – Arrange Tables

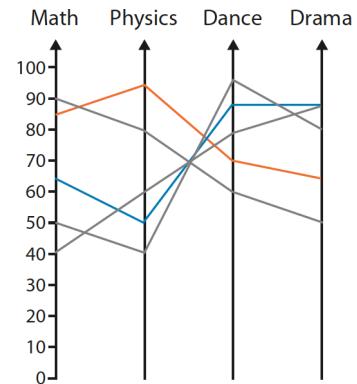
## Spatial Axis Orientation

### Idioms: parallel coordinates

- parallel coordinates
  - parallel axes, jagged line representing item
  - axis ordering is major challenge
- scalability
  - dozens of attributes
  - hundreds of items

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: Task	Find trends, outliers, extremes, correlation.
Scale	Attributes: dozens along secondary axis. Items: hundreds.

#### Parallel Coordinates



Information Visualization and Visual Analytics – Arrange Tables

## Task: Correlation

### scatterplot vs parallel coordinates

- scatterplot
  - positive correlation
    - diagonal low-to-high
  - negative correlation
    - diagonal high-to-low
  - uncorrelated
    - scattered points
- parallel coordinates
  - positive correlation
    - parallel line segments
  - negative correlation
    - all segments cross at halfway point
  - uncorrelated
    - scattered crossings

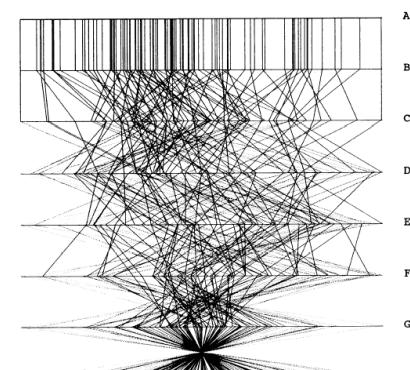
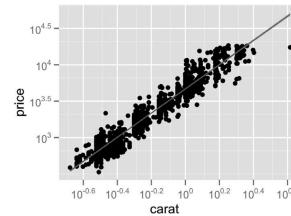
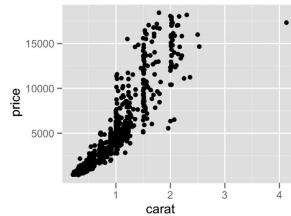
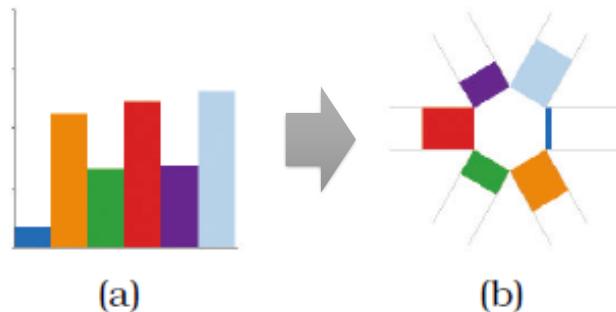


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of  $\rho = 1, .8, .2, 0, -.2, -.8$ , and  $-1$ .

Information Visualization and Visual Analytics – Arrange Tables

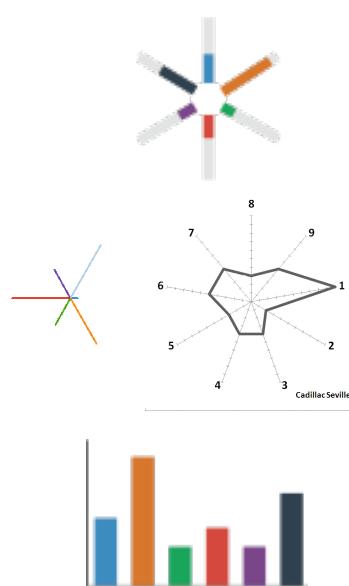
## Idiom: radial bar chart



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

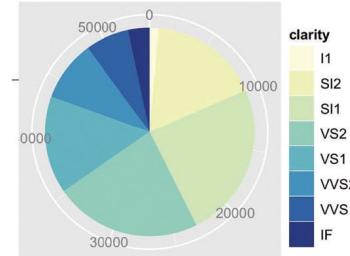
## Idioms: radial bar chart, star plot

- radial bar chart
  - radial axes meet at central ring, line mark
- star plot
  - radial axes, meet at central point, line mark
- bar chart
  - rectilinear axes, aligned vertically
- accuracy
  - length unaligned with radial
    - less accurate than aligned with rectilinear



## Idiom: pie chart, polar area chart

- pie chart
  - area marks with angle channel
  - accuracy: angle/area much less accurate than line length
- polar area chart
  - area marks with length channel
  - more direct analog to bar charts
- data
  - 1 categorical key attribute, 1 quantitative value attribute
- task
  - show relative contribution of **parts to a whole**

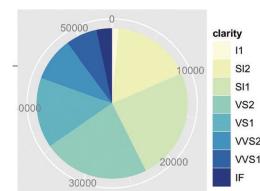


Information Visualization and Visual Analytics – Arrange Tables

## Idiom: pie chart, polar area chart

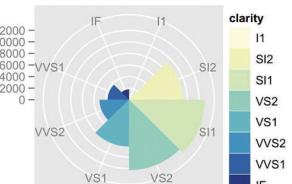
- pie chart

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



- polar area char

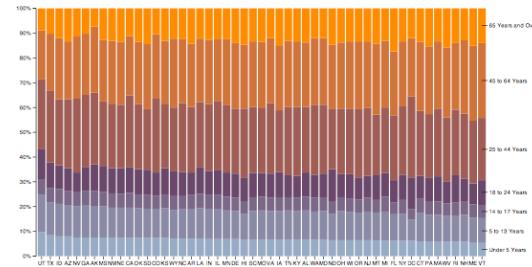
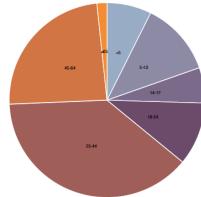
What: Data	Table: one quantitative attribute, one categorical attribute.
How: Encode	Area marks (wedges) with <b>length</b> channel; radial layout.
Why: Task	Part–whole relationship.
Scale	One dozen categories.



Information Visualization and Visual Analytics – Arrange Tables

## Idioms: normalized stacked bar chart

- task
  - part-to-whole judgements
- normalized stacked bar chart
  - stacked bar chart, normalized to full vert height
  - single stacked bar equivalent to full pie
  - high information density: requires narrow rectangle
- pie chart
  - information density: requires large circle

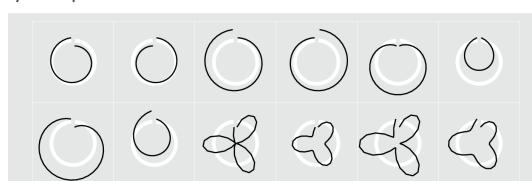
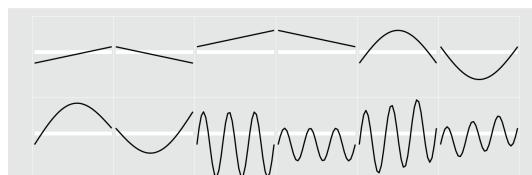


Idiom	Normalized Stacked Bar Charts
What: Data	Multidimensional table: one quantitative value attribute, two categorical key attributes.
What: Derived	One quantitative value attribute (normalized version of original attribute).
Why: Task	Part-whole relationship.
How: Encode	Line marks with length channel; rectilinear layout.
Scale	One dozen categories for stacked attribute. Several dozen categories for axis attribute.

Information Visualization and Visual Analytics – Arrange Tables

## Idiom: glyphmaps

- rectilinear - good for linear vs nonlinear trends
  - The rectilinear layouts are more effective at showing the differences between the linear and nonlinear trends
- the radial plots
  - more effective at showing cyclic patterns



[Wickham et al. 12] Hadley Wickham, Heike Hofmann, Charlotte Wickham, and Diane Cook. "Glyph-Maps for Visually Exploring Temporal Patterns in Climate Data and Models." Environmetrics 23:5 (2012), 382–393. (pages 170, 171)

Information Visualization and Visual Analytics – Arrange Tables

## Radial Layout

- **Channels**

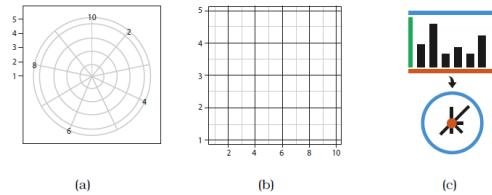
- Angle channel + one or more linear spatial channels.
- Rectilinear layout uses only two spatial channels.

- **Use Polar coordinates**

- Angle from a starting line, a distance from center point

- **Transforming**

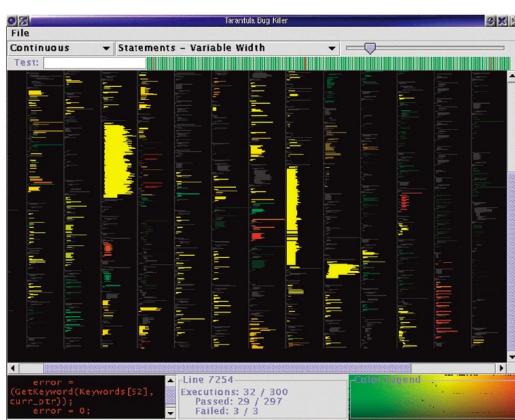
- Rectilinear layout = Radial layout ?
  - Mathematically YES
- Rectilinear layout  $\neq$  Radial layout ?
  - Perceptually NO



**Figure 7.15.** Layout coordinate systems. (a) Radial layouts use polar coordinates, with one spatial position and one angle channel. (b) Rectilinear layouts use two perpendicular spatial position channels. After [Wickham 10, Figure 8]. (c) Transforming rectilinear to radial layouts maps two parallel bounding lines to a point at the center and a circle at the perimeter.

## Spatial Layout Density

## Example: Dense Software Overviews

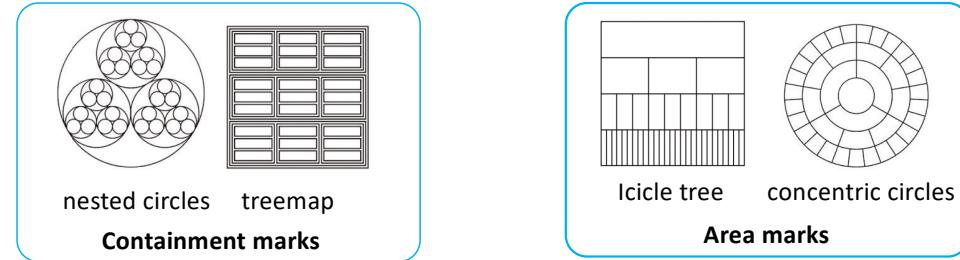


**Figure 7.20.** Tarantula shows a dense overview of source code with lines color coded by execution status of a software test suite. From [Jones et al. 02, Figure 4].

Idiom	Dense Software Overviews
What: Data	Text with numbered lines (source code, test results log).
What: Derived	Two quantitative attributes (test execution results).
How: Encode	Dense layout. Spatial position and line length from text ordering. Color channels of hue and brightness.
Why: Task	Locate faults, summarize results and coverage.
Scale	Lines of text: ten thousand.
(How: Facet)	Same encoding, same dataset, global overview with detail showing subset of data, different resolutions, linking with color.
(How: Reduce)	Detail: filter to local neighborhood of selection

### Space-Filling

- Fills all available space in the view
- Use
  - Area marks for items
  - Containment marks for relationships
- Example



- Advantage : maximize the amount of room available for color coding
- Disadvantage: designer cannot make use of white space in the layout

### Note

### Credits

- Many slides from Tamara Munzner's slide deck
- Many figures from Main Textbook by Tamara Munzner

Note

---

- Questions?