Tackling tables

- homogeneity
  - same data type? same scales?

- need different approaches based on scale
  - how many attributes?
  - up to >50 translatable with direct visual encoding
  - thousands need transformations / analytical methods
  - how many items?
  - up to 1K or scalable with direct visual encoding
  - >1K: need transformations / analytical methods

0 Keys: Express values (magnitudes)

Idiom: scatterplot

- express values
  - quantitative attributes
  - no keys, only values
  - axes
    - 2 quant axes
    - mark points
    - channels
    - horiz + ver position
    - tasks
    - size trends, outliers, distribution, correlation, clusters

Scatterplots: Encoding more channels

- additional channels for point marks
  - color
  - size (bubbleplot)
  - square root of size

Keys and values

- key
  - independent attribute
  - used as unique index to look up items

- simple tables: 1 key
- multidimensional tables: multiple keys

- value
  - dependent attribute, value of cell

- classify arrangements by key count
  - ~0, 1, 2, many...

Some keys: Categorical regions

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

- use ordered attribute to order and align regions

Scatterplot tasks

- correlation
  - clusters/groups, and clusters vs classes

Some keys: Express Values

- one key, one value
  - marks
  - channels
  - length to express quant value

- spatial regions one per mark
  - separated horizontally aligned vertically
  - ordered by quant value
    - by label (alphabetically), by length (data-driven)

- task
  - compare, looking values

- scalability
  - dozens to hundreds of levels for key mark

LIMITATION: Hard to know rank. What's the 4th most? The 7th?

Exercise: Sketch 2 ways to visualize each table

1 Key 2 Keys 3 Keys Many Keys
List Recursive Subdivision
Volume Matrix
Rectilinear Parallel Radial
**Pie chart perception**

- Same empirical evidence that people respond to arc length
- Not angles
- Maybe arc length?
- Donut charts no worse than pie charts

**Pie chart best practices**

- Not bad for (two or) few levels, for part-to-whole task
- Dubious for several levels if details matter
- Terrible for many levels

**Idioms:-spacefilling**

- Rectilinear good for linear vs nonlinear trends
- Radial good for cyclic patterns

**Idioms: cluster heatmap**

- In addition
  - Derived data
  - 2 cluster hierarchies
- dendrogram
  - Parent-child relationships in tree with connection line marks
  - Lewis graph or interior branch heights easy to compare
- Heatmap
  - Marks (e.g., bordered by color) hierarchy traversal
  - Scale: sense quality of clusters hand by systematic methods

**Idioms: normalized stacked bar chart**

- Task
  - Part-to-whole judgements
- Normalized stacked bar chart
  - Stacked bar chart, normalized to full-vf height
  - Single stacked bar equivalent to full pie
  - High information density: requires narrow rectangles
- Pie chart
  - Information density: requires large rectangle

**Idioms: parallel coordinates**

- Scatterplot matrix (SPLOM)
  - Rectilinear axes, point mark
  - All possible pairs of axes
  - Scalability
  - One dozen axes
  - Dozens to hundreds of items

**Idioms: glyphmaps**

- Rectilinear good for linear vs nonlinear trends
- Radial good for cyclic patterns

**Idioms: heatmap**

- Two keys, one value
  - Data
  - 2-cm scales (gene, experimental condition)
- Metrics
  - Area
  - Color by quartile
- Channels
  - Color by quartile
  - (Centered diverging color map)
  - Task
  - Find clusters, outliers
  - Scalability
  - ~19 turns, 100 of levels, ~10 quartile levels

**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 aligned, axes aligned vertically
  - Accuracy
  - Length aligned with radial
  - Less accurate than aligned with rectilinear

**Idioms: pie chart, polar area chart**

- Pie chart
  - Line marks with angle channel variable (sector) width
- Polar area chart
  - Line marks with length channel variable (sector) width
  - Accuracy: arc less accurate than sector length
### Arrange Tables

- **Task: Correlation**
  - scatterplot matrix
  - diagonal-low-to-high correlation
  - diagonal-high-to-low correlation
  - uncorrelated spread out

- **Parallel coordinates**
  - positive correlation
  - negative correlation
  - not sure
  - horsepower to acceleration
  - weight to mileage

- **Parallel coordinates quiz: car data**
  - What correlations do you see?
    - positive?
    - negative?
    - none?
    - not sure?
  - horsepower to acceleration
  - weight to mileage?

- **Parallel coordinates, limitations**
  - visible patterns only between neighboring axis pairs
  - how to pick axis order?
  - some weaknesses as many other techniques
  - downside of interaction: human-powered search
  - some algorithms proposed, none fully solved

- **Orientation limitations**
  - rectilinear: scalability wrt #axes
  - 2 axes best
  - 3 problematic
  - 4+ impossible
  - parallel: unfamiliarity, training time

### Orientations

- **Radial orientation**
  - perceptual limits
  - polar coordinate asymmetry
  - angles lower precision than length
  - nonuniform sector width/size depending on radial distance
  - frequently problematic
  - sometimes can be deliberately exploited
    - for 3 attrs of very unequal importance

- **Layout density**
  - polar coordinate asymmetry
  - angles lower precision than length
  - nonuniform sector width/size depending on radial distance
  - frequently problematic
  - sometimes can be deliberately exploited
    - for 3 attrs of very unequal importance

- **Idiom: Dense software overviews**
  - data: text
  - text = 1 quad attr per line
  - derived data:
    - one pixel high line
  - length according to original
  - color line by attrib
  - scalability
    - 10K+ lines

### Encode: tables: Arrange space

- **Encode tables: Arrange space**
  - **Encode**
    - **Arrange**
      - **Express**
      - **Separate**
      - **Order**
      - **Align**

- **Design critique & redesign: NZ**
  - Consider the following questions:
    - 1 What could be the goals of the designer for your design?
    - 2 What is the visualization scheme (domain-specific & abstract)?
    - 3 How is the data type visually encoded (meta/attributes)?
    - 4 Can you read the data precisely? Is the visual encoding appropriately chosen?
    - 5 How would you make this work without numeric labels?
  - Develop two alternative designs to visualize this data.
    - Fine to discuss with your peers, but draw your own solution.
    - Mark your best design, briefly note why you think it's best.

### Upcoming

- **D3 videos week 3**
  - Making a Bar Chart with D3 and SVG [30 min]
  - Quiz 3, due by Fri Jan 24, 8am
  - Reading: Reusable D3 Components
  - Friday: Pattern of D3.js [60 min]
  - Interaction with Unidirectional Data Flow [16 min]

### Design critique & redesign: NZ

- Consider the following questions:
  - 1 What could be the goals of the designer for questions that the visualization answers (domain-specific & abstract)?
  - 2 What data is represented in this visualization? Be specific.
  - 3 How is each data type visually encoded (meta/attributes)?
  - 4 Can you read the data precisely? Is the visual encoding appropriately chosen?
  - 5 How would you make this work without numeric labels?
  - Develop two alternative designs to visualize this data.
    - Fine to discuss with your peers, but draw your own solution.
    - Mark your best design, briefly note why you think it's best.

### Credits

- Visualization Analysis and Design (Ch 7)
- Alex Lex & Miriah Meyer: http://datavisualization.net/
- Ben Jones, UW/Tableau