What?

Datasets

- Data Types
  - Items → Attributes → Links → Positions → Grids
- Data and Dataset Types

Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

- Dataset Types
  - Tables
  - Networks
  - Fields (Continuous)
  - Multidimensional Table
  - Trees
- Geometry (Spatial)
- Dataset Availability
  - Static
  - Dynamic

Attributes

- Attribute Types
  - Categorical
  - Ordered
    - Ordinal
    - Quantitative
- Ordering Direction
  - Sequential
  - Diverging
  - Cyclic

Why?

Actions

- Analyze
  - Consume
    - Discover → Present → Enjoy
  - Produce
    - Annotate → Record → Derive
- Search

	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

- Query
  - Identify → Compare → Summarize

Targets

- All Data
  - Trends → Outliers → Features
- Attributes
  - One
    - Distribution
    - Extremes
  - Many
    - Dependency → Correlation → Similarity
- Network Data
  - Topology
  - Paths
- Spatial Data
  - Shape

What? Why? How?

Domain situation

Observe target users using existing tools

Data/task abstraction

Visual encoding/interaction idiom

Justify design with respect to alternatives

Algorithm

Measure system time/memory

Analyze computational complexity

Analyze results qualitatively

Measure human time with lab experiment (*lab study*)

Observe target users after deployment (*field study*)

Measure adoption

Marks

- Points (0D) • Lines (1D) • Areas (2D)

Channels

- Position
  - Horizontal → Vertical → Both
- Color
- Shape
- Tilt
- Size
  - Length → Area → Volume

How?

Encode

- Arrange
  - Express → Separate
  - Order → Align
  - Use
- Map from categorical and ordered attributes
  - Color
    - Hue → Saturation → Luminance
  - Size, Angle, Curvature, ...
  - Shape
  - Motion
    - Direction, Rate, Frequency, ...

Manipulate

- Change
- Select
- Navigate

Facet

- Juxtapose
- Partition
- Superimpose

Reduce

- Filter
- Aggregate
- Embed

What? Why? How?

Summary

Visual encoding: attribute -> visual channel

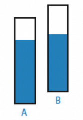
Expressiveness and effectiveness

Discriminability

Separability

Popout effect

Relative vs. Absolute Judgements



A B

Framed Unaligned (b)

Grouping

- Containment
- Connection

48

Magnitude Channels: Ordered Attributes

Most → Effectiveness → Least

- Position on common scale
- Position on unaligned scale
- Length (1D size)
- Tilt/angle
- Area (2D size)
- Depth (3D position)
- Color luminance (Same)
- Color saturation (Same)
- Curvature
- Volume (3D size) (Same)

Identity Channels: Categorical Attributes

- Spatial region
- Color hue
- Motion
- Shape

Idiom: scatterplot

- Express values
  - Quantitative attribute
- No keys, only values
  - Data
    - 2 quantitative attributes
  - Mark:
    - points
  - Channels
    - Horizontal + vertical position
  - Tasks
    - Find trends, outliers, distribution, correlation, clusters
  - Scalability
    - Hundreds of items

you do not care which date of t... But correlation, outlier......

Idiom: bar chart

- One key, one value
  - Data
    - 1 categorical attribute (key), 1 quantitative attribute
  - Mark: lines
  - Channels
    - Length to express quantitative value
    - Spatial region: one per mark
      - Separated horizontally, aligned vertically
      - Ordered by quantitative attribute
        - By label (alphabetical) , by length attribute(data-driven)
  - Task
    - Compare, lookup values
  - Scalability
    - Dozens to hundreds of key attribute

Idiom: line chart/ dot plot

- One key, one value
  - Data
    - **2 quantitative attributes**
  - Mark:
    - points
    - Line connection marks between them
  - Channels
    - Aligned lengths to express quantitative value
    - Separated and ordered by key attribute into horizontal regions
  - Task
    - Find trend: connection marks (perceptually) **emphasize** ordering of items along key axis by explicitly showing relationship one item and the next
  - Scalability
    - Hundreds of key levels, hundreds of value levels

Idiom: Stacked bar chart

- One more key
  - Data
    - 2 categorical/ordinal attr, 1 quantitative attr
  - Mark: vertical stack of line marks
    - Glyph: composite object, internal structure from multiple marks
  - Channels
    - Length and color hue
    - Spatial regions: one per glyph
      - Aligned: full glyph, lowest bar component
      - Unaligned: other bar components
  - Task
    - Part-to-whole relationship
    - Compare bars at bottom (close to the baseline)
  - Scalability
    - Several to one dozen levels for stacked attribute

Idiom: Streamgraph

- Generalized stacked graph
  - Emphasizing horizontal continuity (create block size for comparison)
    - vs vertical item
  - Data
    - 1 categorical key attribute
    - 1 ordered key attribute
    - 1 quantitative value attribute
  - Task
    - Part-to-whole relationship over x-axis
  - Scalability
    - Hundreds of keys (x-axis direction)
    - Dozens to hundreds of keys (y-axis direction)
      - More than stacked bars, since most layers do not extend across whole chart

Connect to ... to visualize ...

Idiom: Gantt charts

- One key, two (related) values
  - Data
    - 1 categorical attr, 2 quantitative attr
  - Mark: line
    - Length: duration
  - Channels
    - Horizontal position: start/end time
    - Horizontal length: duration
  - Task
    - Emphasize **temporal overlaps**, start/end dependencies between items
  - Scalability
    - Dozens of key levels
    - Hundreds of value levels

D3: http://bl...

Pros and Cons

- Scatterplot matrix
  - Pros: easy to understand
  - Cons: scalability
    - The number of the sub-scatter plots grows exponentially
- Parallel coordinates
  - Pros: scalability
    - It can show much more variables
  - Cons:
    - Most people know how to interpret parallel coordinate plot.
    - People need to be trained to interpret it.
    - Not good to use it on news paper (interaction)

Idiom: Slopegraph

- Two values
  - Data
    - 2 quantitative values attributes
    - 1 derived attribute: change magnitude
  - Mark: point, line
    - Line connecting mark between points
  - Channels
    - 2 vertical position: express attribute value
    - linewidth/size, color
  - Task
    - Use the "slope" to emphasize changes in rank/value
  - Scalability
    - Hundreds of value levels

Idiom: heatmap

- Two keys, one value
  - Data
    - 2 categorical attributes
    - 1 quantitative attribute
  - marks: area
    - Separate and align in 2D matrix
    - Indexed by 2 categorical attributes
  - Channels
    - Color by quantitative value
  - Task
    - Find patterns, clusters, outlier
  - Scalability
    - 1M items, 100s of categorical levels, ~10 quantitative attribute levels

Idiom: Adjacency Matrix View

- Data: network
  - Transform into same data/encoding as heatmap
- Derived data: table from network
  - 1 quantitative attribute
    - Weighted edge between nodes
  - 2 categorical attributes:
    - Node list * 2
- Visual encoding
  - Cell shows presence/absence of edge (or weight of the edge)
- Scalability
  - 1K nodes, 1M edges

Idiom: Radial Layout

Pie chart · Radial bar · Radius pie chart · Star plot

D3: https://www.d3-graph-gallery.com/circular_barplot

- Pie chart (classic/controversial)
  - Area mark with angle channel
  - Data: 1 categorical key attribute, 1 quantitative value attribute
  - Task: Part-to-whole judgements
- Advantage: visually appealing and ok for Part-to-whole judgements
- Many alternative visualization is better than radial layout in terms of data understanding purpose
  - **Accuracy: angle/area less accurate than line length**
    - **Arc length also less accurate than line length**

Axis Orientation: → Rectilinear → Parallel → Radial

Idiom: Animated Transition (Important!!!)

- Smooth interpolation from one state to another
  - Alternative to jump cut, supports item tracking
    - Best case for animation
  - Staging to reduce "**cognitive load**"
  - Use can know what happen between two view without explanation

Measuring our planets' distances

Geometric zooming

Measuring our planets' distances

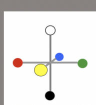
Sematic zooming

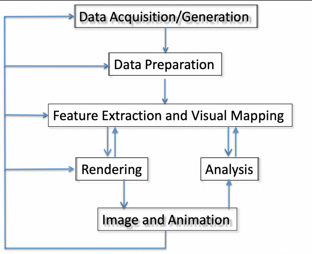
Summary of Cartograms

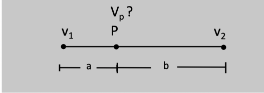
	Continuous cartogram	Dorling cartogram	Grid cartogram	Geographical map
Easy to recognize by **position?**	●●	●	●	●●●
Easy to recognize by **size?**	×	×	×	●●●
Easy to recognize by **shape?**	●	×	×	●●●
Easy to compare shape **size?**	●●	●●●	×	×

Color Space

- RGB:
  - good for display hardware
  - poor for visual encoding
- CIE L*a*b*: good for visual design
  - L* intuitive: perceptually linear luminance
  - a*b* axes: perceptually linear but nonintuitive
- HSL (HSV/HSB): better (but beware)
  - HSL: (Hue/Saturation/Lightness)
  - Lightness ≠ luminance





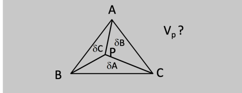


- Parametric coordinate of P: α = a / (a+b)
- Linearly interpolated value of P:

$$\boxed{V_p = (1-\alpha) * V_1 + \alpha * V_2}$$

lerp(v1,v2, α)

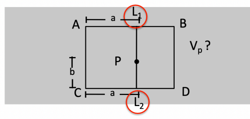
Lerp in Triangle



- Parametric coordinates of P: (α,β,γ)
  - $\alpha = \delta A / (\delta A + \delta B + \delta C)$
  - $\beta = \delta B / (\delta A + \delta B + \delta C)$
  - $\gamma = \delta C / (\delta A + \delta B + \delta C)$

  Baricentric Coordinates
- Linearly interpolated value of P: $\boxed{V_A * \alpha + V_B * \beta + V_C * \gamma}$

Lerp in Rectangle

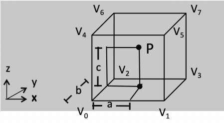


- Parametric coordinates of P: (α,β)
  - α = a / width; β = b / height

Bi-linear interpolation
Bi-Lerp(V_A,V_B, V_C,V_D)

- Linearly interpolated value of P: $\boxed{Lerp(V_{L1}, V_{L2}, \beta)}$

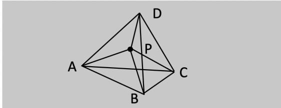
Lerp in Cube



Another way to perform calculate the value at P:

- Parametric coordinates of P: (α,β,γ)
  - α = a / width; β = b / depth (along y);
  - γ = c / height
- Value at P =
  $(1-\alpha)(1-\beta)(1-\gamma)V_0 + \alpha(1-\beta)(1-\gamma)V_1 +$
  $(1-\alpha)\beta(1-\gamma)V_2 + \alpha\beta(1-\gamma)V_3 +$
  $(1-\alpha)(1-\beta)\gamma V_4 + \alpha(1-\beta)\gamma V_5 +$
  $(1-\alpha)\beta\gamma V_6 + \alpha\beta\gamma V_7$

Lerp in Tetrahedron

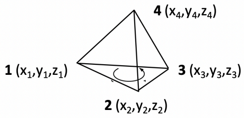


- Tetrahedral coordinates of P: $(\alpha,\beta,\gamma,\delta)$

$\alpha = V_{BDCP} / V_{ABCD}$
$\beta = V_{ACDP} / V_{ABCD}$
$\gamma = V_{ADBP} / V_{ABCD}$
$\delta = V_{ABCP} / V_{ABCD}$

Linearly interpolated value of P: $\boxed{V_A * \alpha + V_B * \beta + V_C * \gamma + V_D * \delta}$

Volume of Tetrahedron



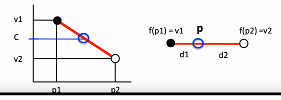
$$V = \frac{1}{6}\det\begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{bmatrix} = \frac{1}{6}\det(\mathbf{J}) = \frac{1}{6}J.$$

V will be positive if when you look at the triangle $_{123}$ from vertex 4, vertex 1 2 3 are In a counter clockwise order

- The intersection point **f(p) = c** on the edge can be computed by linear interpolation

d1/d2 = (v1-C) / (C-v2) $\Rightarrow$ **(p – p1)/(p2-p1)** = (v1-C) / (v1-v2)

p = (v1-C)/(v1-v2) * **(p2-p1) + p1**



$$\boxed{I(D) = I_0 \times e^{-\int_0^D \tau(t)dt} + \int_0^D g(s)e^{-\int_s^D \tau(t)dt}ds}$$

I_0 : background light

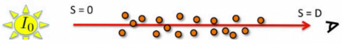
$\tau(t)$: extinction coefficient at t , related to the rate that light is occluded

D : total distance light will travel

$e^{-\int_0^D \tau(t)dt}$: transparency of medium between 0 and D

$1 - e^{-\int_0^D \tau(t)dt} = \alpha$: opacity of medium between 0 and D

$g(s)$: source term at point s, typically derived from the data value

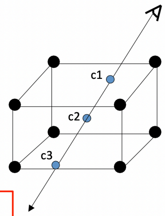


Back-to-Front Compositing

The initial pixel color = Black

Back-to-Front compositing:
use 'under' operator

C = C1 'under' background
C = C2 'under C
C = C3 'under C
...

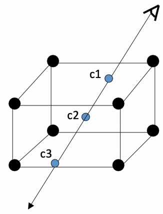


$\boxed{C_{out} = C_{in} * (1-\alpha(x)) + C(x)*\alpha(x)}$
(this is the alpha blending formula)

Front-to-Back Compositing

Front-to-Back compositing:
use 'over' operator

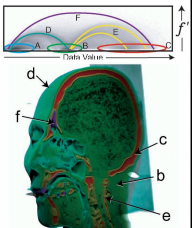
C = background 'over' C1
C = C 'over' C2
C = C 'over' C3
...



$C_{out} = C_{in} + C(x)\ \alpha(x) * (1-\alpha_{in})$;
$\alpha_{out} = \alpha_{in} + \alpha(x) * (1-\alpha_{in})$

Multi-Dimensional Transfer Functions

- 1D histogram can capture homogeneous region only
  - A : air
  - B : tissue
  - C : bone
- 2D histogram can capture
  - D : air and tissue boundary
  - E : tissue and bone boundary
  - F : air and bone boundary



Type of flow lines:

- **Streamline:** a field line tangent to the velocity field at an instant in time
- **Pathline:** the trajectory of a massless particle released from a seed point over a period of time
- **Streakline:** a line joining the positions, at an instant in time, of particles released from a seed point
- **Timeline:** a line connecting a row of particle that are released simultaneously

切線方向
丟一粒子流動軌跡

丟n次一粒子，連起來

丟一次N粒子，連起來

$$\boldsymbol{p_{k+1} = p_k + v(p_k) \times \Delta t}$$

$\boldsymbol{p^* = p_k + v(p_k)\ \Delta t}$
$\boldsymbol{p_{k+1} = p_k + (v(p_k) + v(p^*)\ ) \times \Delta t/2}$

$\boldsymbol{a = 2\Delta t\ v(p_k)}$,
$\boldsymbol{b = 2\Delta t\ v(p_k+a/2)}$,
$\boldsymbol{c = 2\Delta t\ v(p_k+b/2)}$,
$\boldsymbol{d = 2\Delta t\ v(p_k+c/2)}$,

$$\boxed{\boldsymbol{p_{k+1} = p_k + (a+2b+2c+d)/6}}$$