CSE3020 - Data Visualization

Module 3 :Visual Analytics

Dr. K.P. Vijayakumar, VIT Chennai

Topics to be covered

- Visual Variables
- Networks and Trees
- Map Color and Other Channels
- Manipulate View

Visual Variables

- **Introduction**
  - an aspect of a graphical object that can visually differentiate it from other objects, and can be controlled during the design process
  - Used to communicate visually

Visual Variables

- **Introduction**
  - Visual Variables
    - Position
    - Size
    - Shape
    - Hue
    - Saturation / Lightness (Value)
    - Orientation
    - Spacing/arrangement
    - Perspective Height
    - Texture

Source: Wikipedia, https://www.youtube.com/watch?v=ey8-X6sq4-U

Visual Variables

- **Introduction**

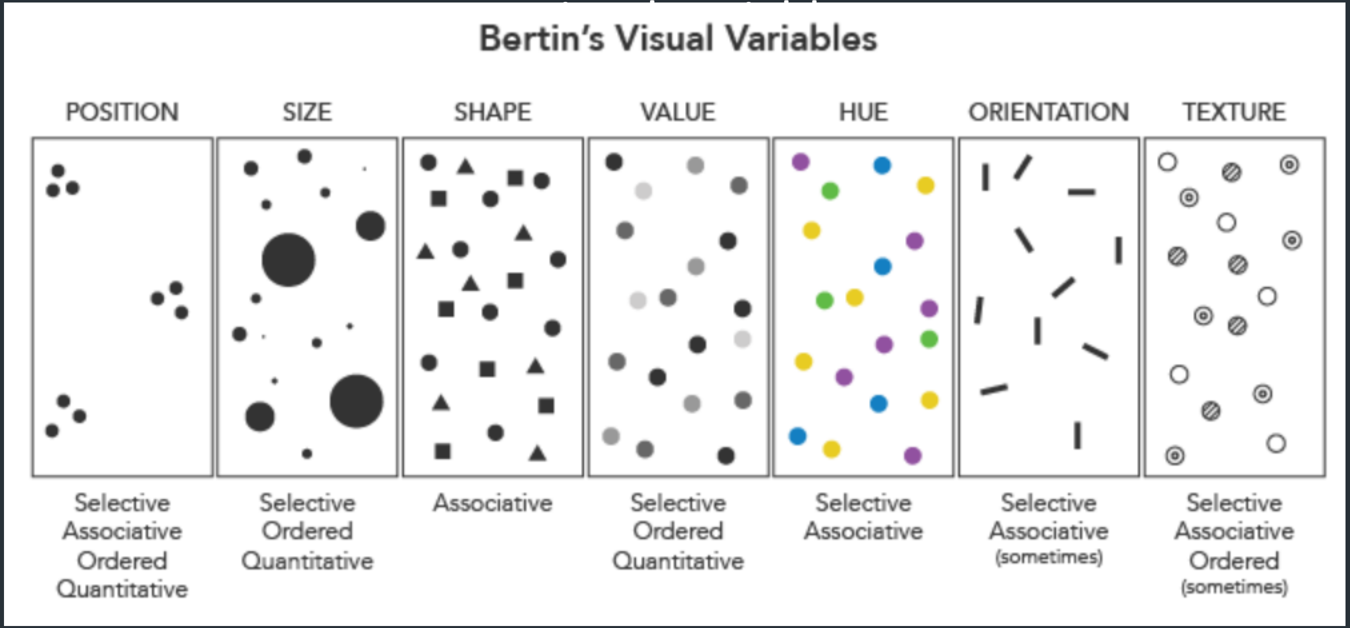
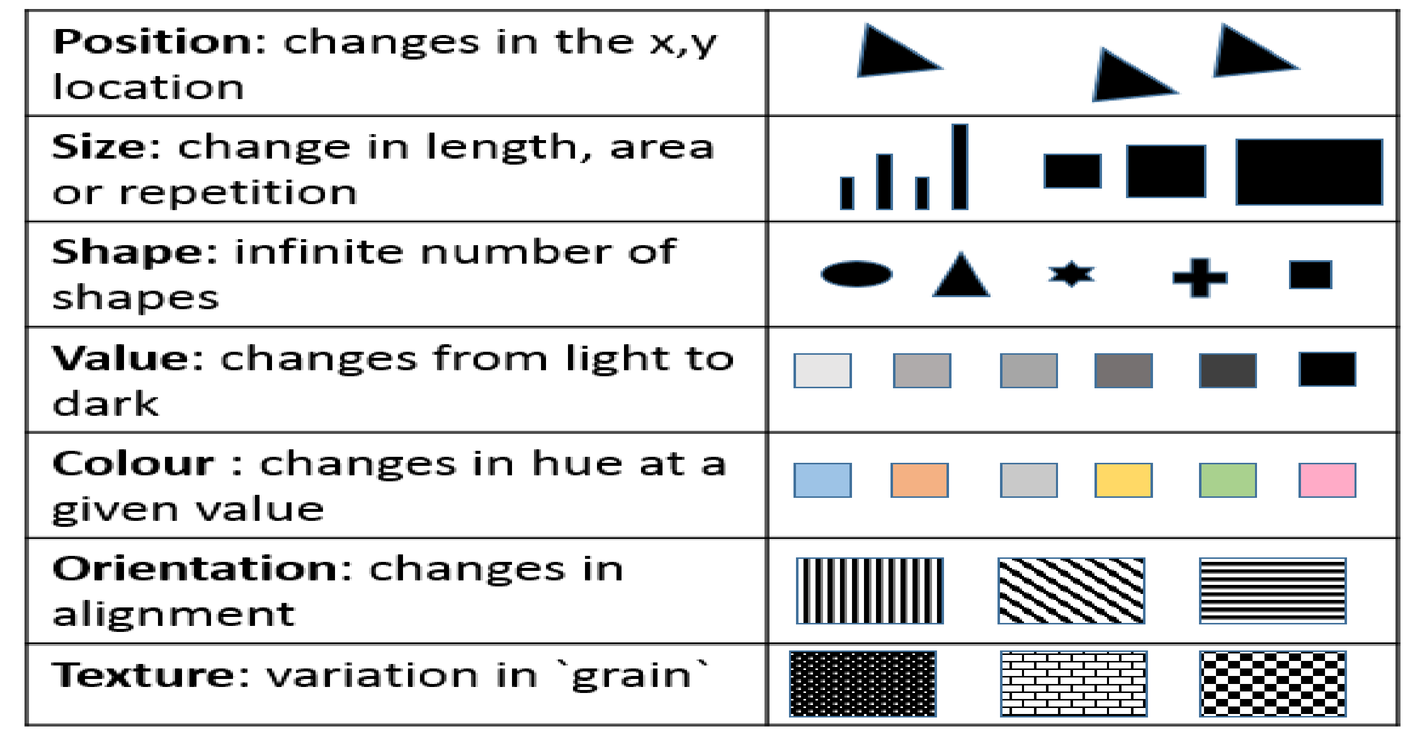


Image Source: https://www.axismaps.com/guide/visual-variables

Visual Variables

- Bertin’s Visual Variables



Position: changes in the x,y location	
Size: change in length, area or repetition	
Shape: infinite number of shapes	
Value: changes from light to dark	
Colour : changes in hue at a given value	
Orientation: changes in alignment	
Texture: variation in `grain`	

Image Source: https://www.researchgate.net/figure/Bertins-Visual-Variables-

Visual Variables

- **Size**

Image Source: https://en.wikipedia.org/wiki/Visual_variable

Visual Variables

- **Shape**

Image Source: https://en.wikipedia.org/wiki/Visual_variable

Visual Variables

- **Size and Shape**

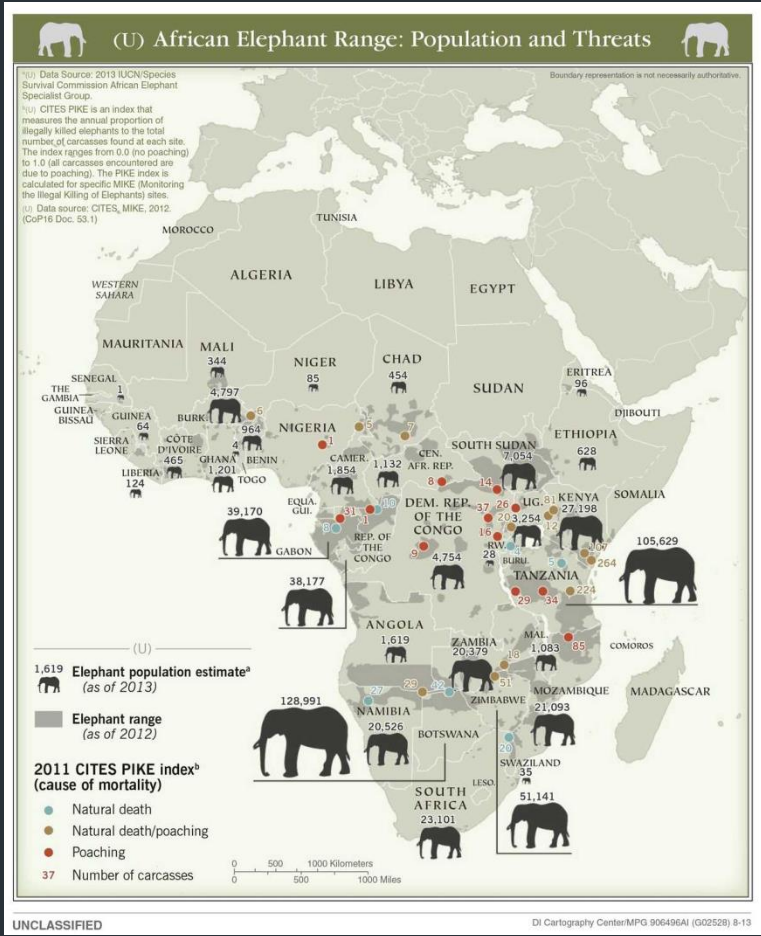
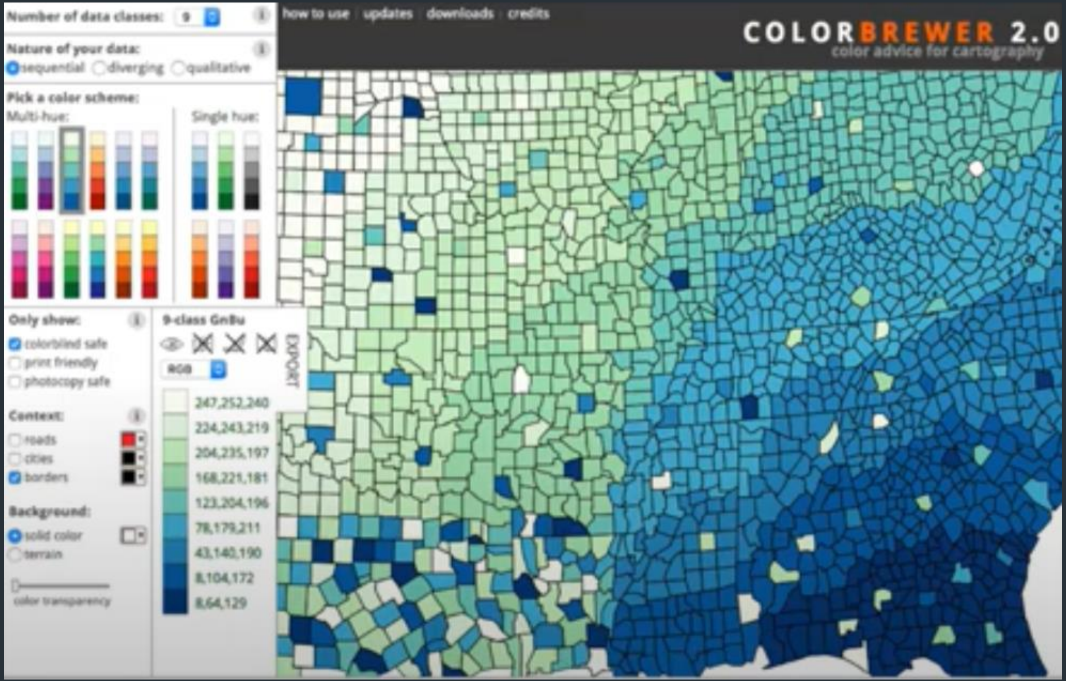


Image Source: https://cia.gov

Visual Variables

- **Hue**



Source: https://www.youtube.com/watch?v=ey8-X6sq4-U

Visual Variables

- **Saturation**

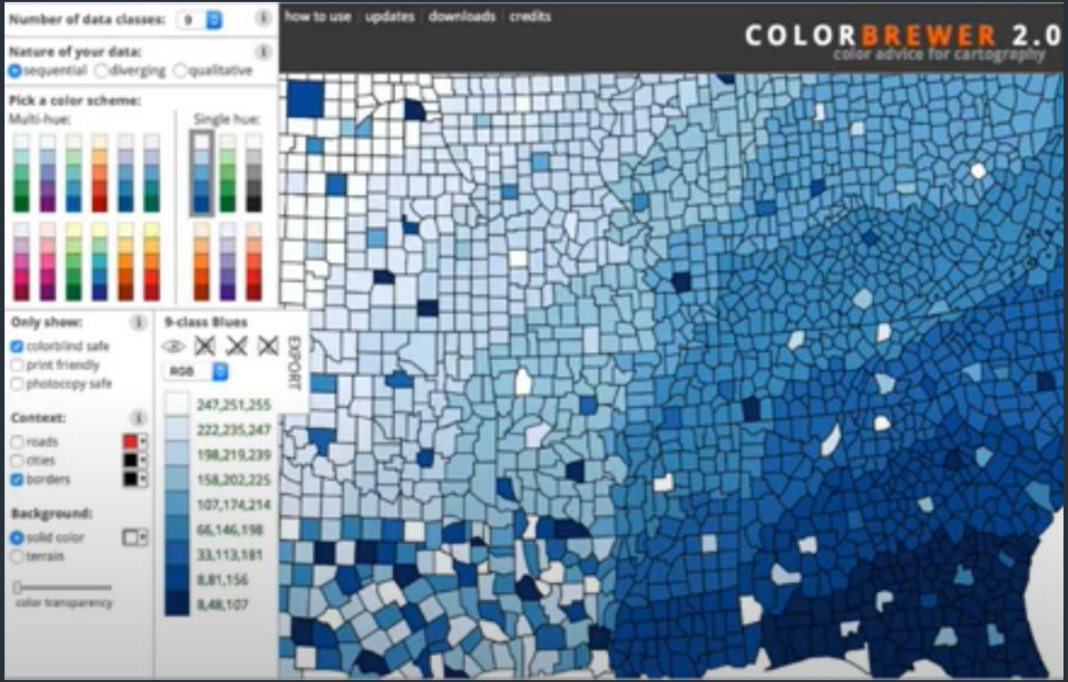
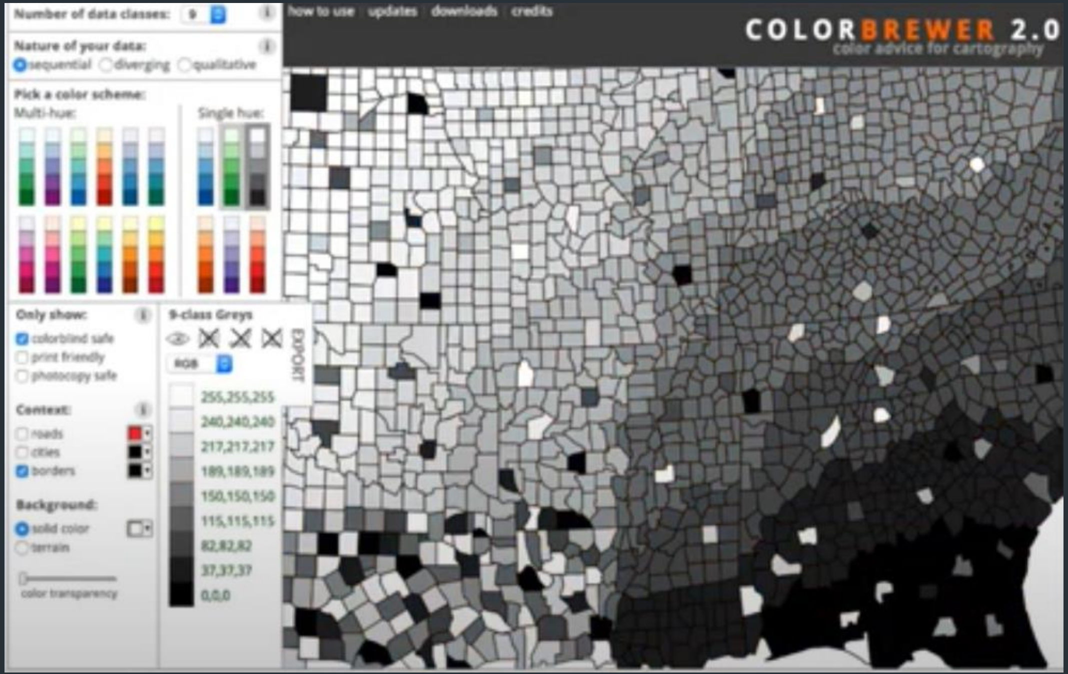


Image Source: https://en.wikipedia.org/wiki/Visual_variable , www.youtube.com/watch?v=ey8-X6sq4-U

Visual Variables

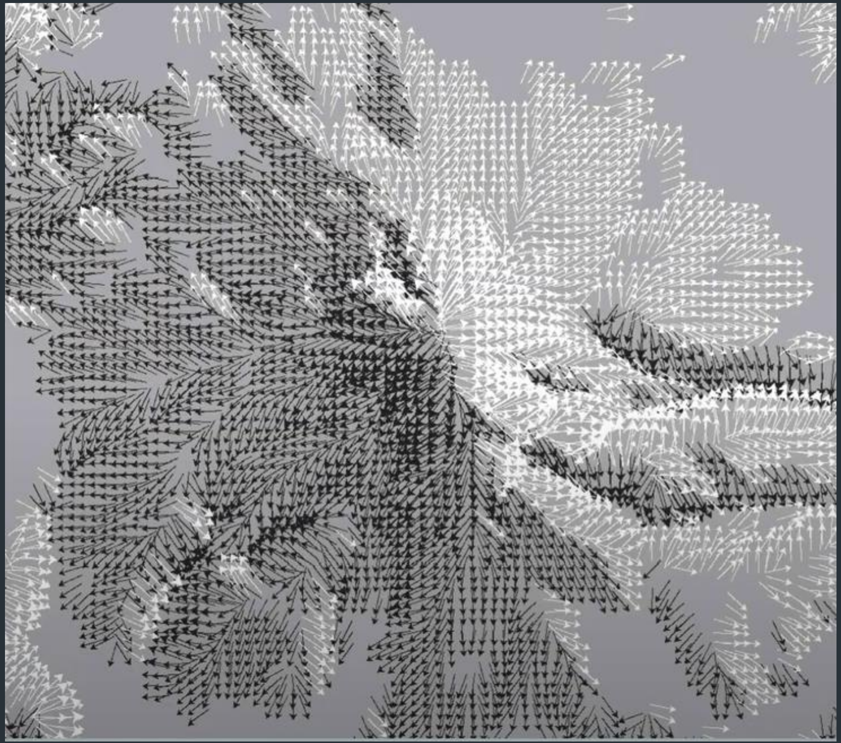
- **Lightness (grayness)**



Source: https://www.youtube.com/watch?v=ey8-X6sq4-U

Visual Variables

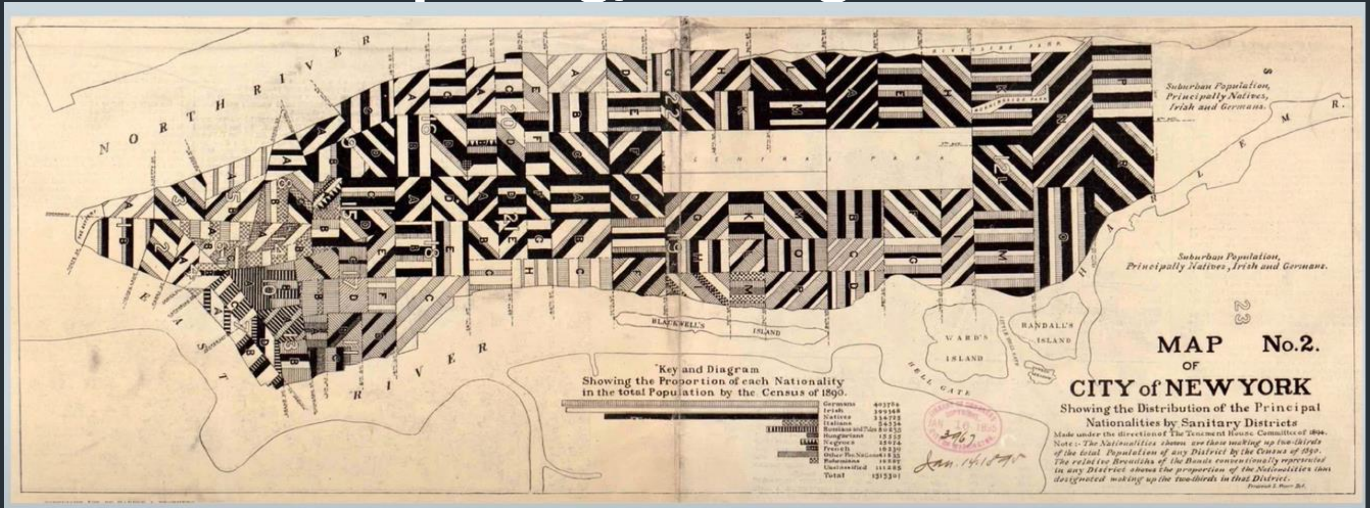
- **Orientation**

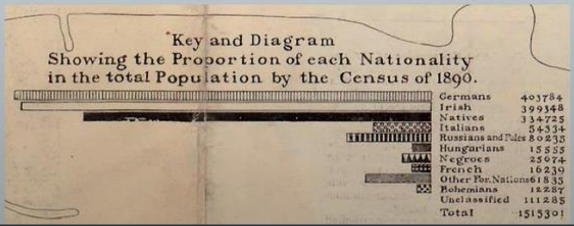


Source: https://www.youtube.com/watch?v=ey8-X6sq4-U

Visual Variables

- **Spacing/Arrangement**





Map of New York City, Showing the Distribution of the Principal Nationalities by Sanitary Districts published in Harper's Weekly (January 5, 1895, pp. 60–61) using 1890 U.S. Census data.

Source: https://www.youtube.com/watch?v=e

Visual Variables

- **Perspective Height**

METRICS

HANNAH FAIRFIELD

In the Shadow Of Foreclosures

IN the subprime mortgage squeeze, some regions are feeling the pain more acutely than others.

Although many Southern metropolitan areas have high percentages of subprime mortgages, homeowners in those areas have largely been able to pay their bills, so subprime foreclosure rates are low.

Not so in the Rust Belt, where subprime mortgages are less common but foreclosure rates are sky-high, mostly a result of rising unemployment.

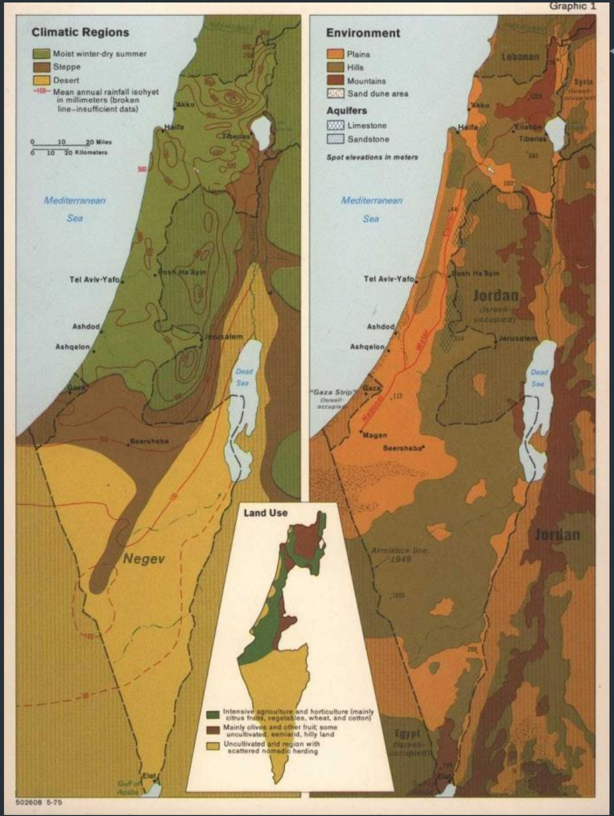
And overbuilding in regions of Florida, California and other states with housing bubbles lured overeager residents to become speculators, buying up several homes with the expectation that their values would rise. Getting subprime loans was all too easy.

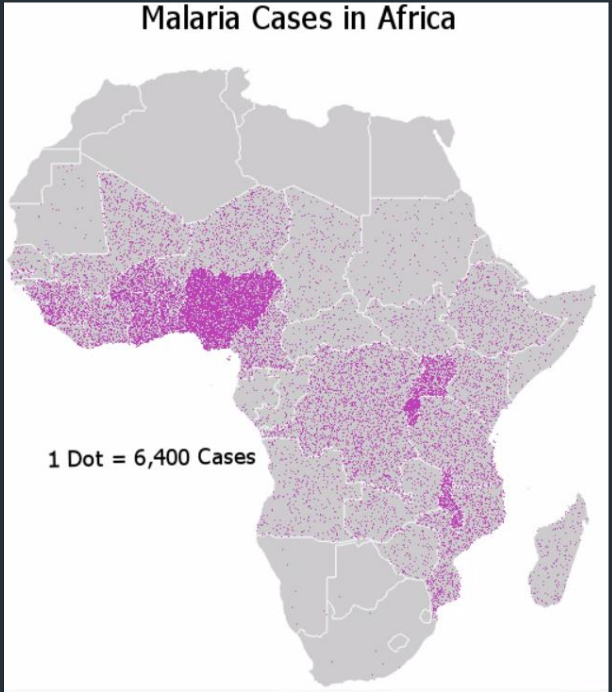
But paying the loans as housing prices fall is all too hard, and many economists believe that foreclosures will continue to rise.

"The collapse will affect other markets, like New York, Boston and

Visual Variables

- **Texture**





Source: https://www.youtube.com/watch?v=ey8-X6sq4-U

Visual Variables

- **Properties of Visual Variable**
  - Selective
    - allows us to immediately isolate a group of signs based on a change in the variable.
    - Ex :Hue
  - Associative
    - allows grouping across changes in the variable
    - Ex : Hue and Shape
  - Ordered
  - Qunatitative

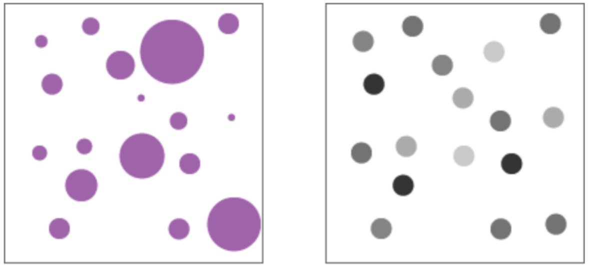
We can quickly and easily perceive a group of symbols based on color hue, e.g., the yellow symbols appear as a group. Therefore hue is *selective*.

Symbols can be perceived as a group despite variation in shape, e.g., these similarly sized purple symbols appear as a group. Shape is therefore *associative*.

Source: https://www.axismaps.com/guide/visual-variables

Visual Variables

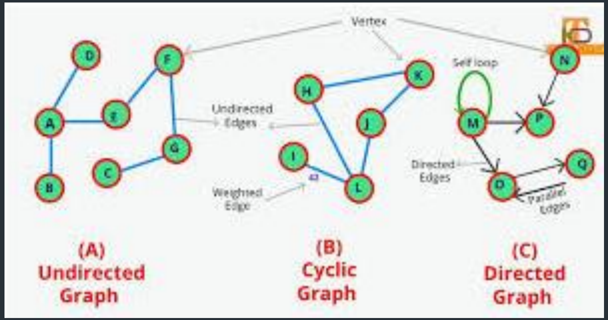
- **Properties of Visual Variable**
  - Ordered
    - immediately recognizable sequence
    - Ex: Size, Value, Position, Symbol
  - Quantitative
    - allow an estimation of the actual numerical difference between symbols
    - Ex: Size

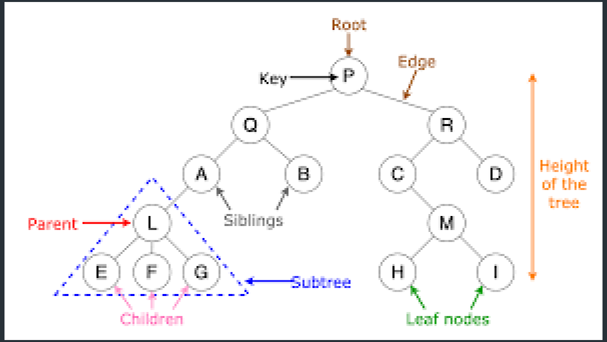


With both size and value, it is immediately obvious that there is some sequence to the symbols (small to large, light to dark). Size and value or *ordered*.

Source: https://www.axismaps.com/guide/visual-variables

Networks and Trees





- **Graphs and Trees**
- Graph
  - Vertex/node with one or more edges connecting it to another node.
  - Cyclic or acyclic
  - Edge can be weighted (value) or categorized
- Tree
  - Undirected graph where two nodes are connected by only one edge
  - used for hierarchy
  - Edge can be weighted (value) or categorized

Image Source: https://droidtechknow.com/programming/data-structure/an-introduction-to-graph-data-structure/, towardsdatascience

Networks and Trees

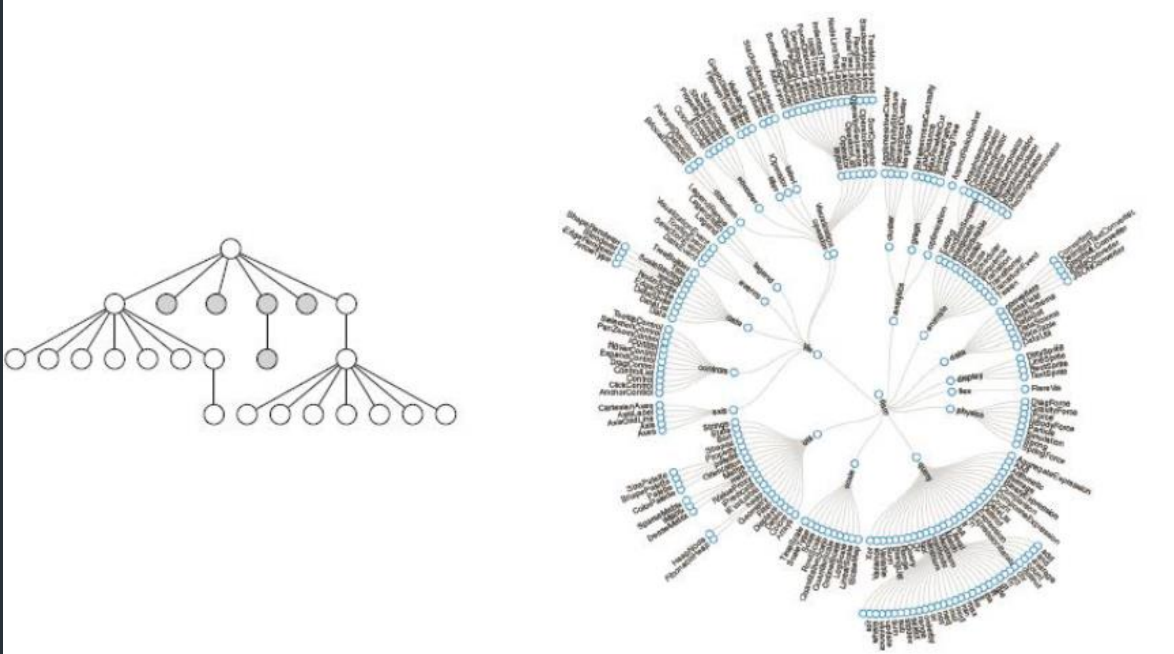
- **Design Choices**
- Connectivity
  - Node-link graphs
  - Good for finding pairwise/multiway relations
  - Good for following paths through structure
  - Force-directed placement
- Containment
  - Effective at showing hierarchical structure
  - Good for finding attributes of leaf nodes
  - Treemaps, nested views
- Matrices

Networks and Trees

- ## Node Link Diagrams
- Visual encoding idiom for tree and network data is with node–link diagrams, where nodes are drawn as point marks and the links connecting them are drawn as line marks

Networks and Trees

Node Link Diagrams

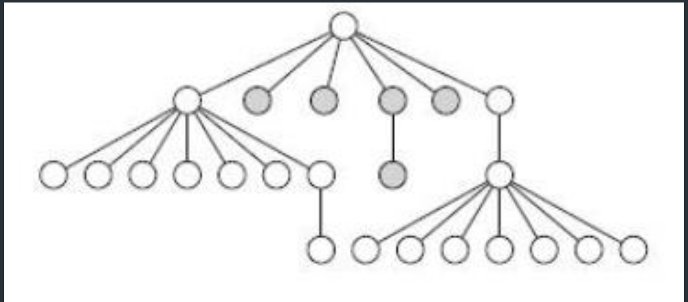


(a) Triangular vertical for tiny tree
(b) Spline radial layout for small tree

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

Node Link Diagrams

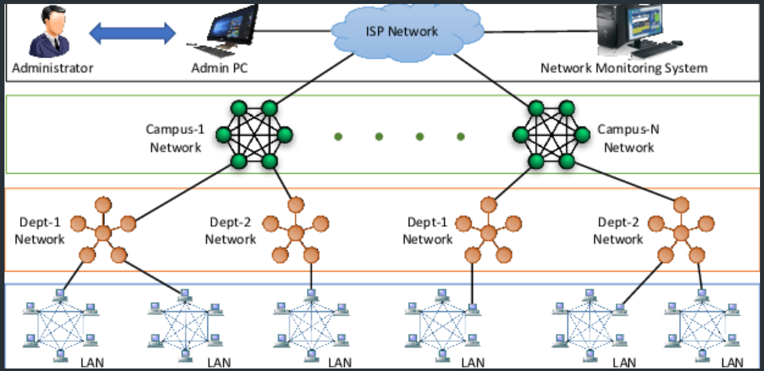


- directly connected by a single link are perceived as having the tightest grouping
- nodes with a long path of multiple hops between them are less closely grouped.
- The number of hops within a path - the number of individual links that must be traversed to get from one node to another - is a network-oriented way to measure distances.

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

Node Link Diagrams

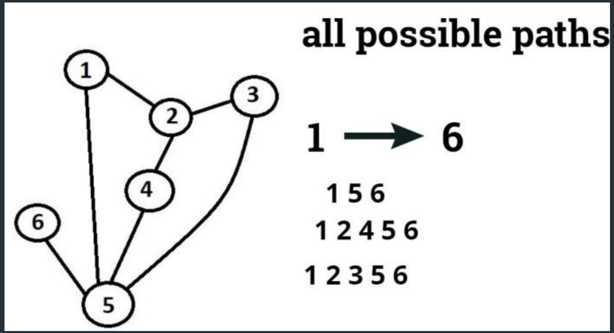


- It is well suited for tasks that involve understanding the network topology
  - direct and indirect connections between nodes in terms of the number of hops between them through the set of links.
- Examples of topology tasks include
  - finding all possible paths from one node to another,
  - finding the shortest path between two nodes,
  - finding all the adjacent nodes one hop away from a target node,
  - finding nodes that act as a bridge between two components of the network that would otherwise be disconnected.

Source: Visualization Analysis and Design, Tamara Munzner
Image Source : https://www.researchgate.net/figure/Basic-network-topology-of-a-

Networks and Trees

Node Link Diagrams



- Examples of topology tasks include
  - finding all possible paths from one node to another,
  - finding the shortest path between two nodes,
  - finding all the adjacent nodes one hop away from a target node,
  - finding nodes that act as a bridge between two components of the network that would otherwise be disconnected.

Source: Visualization Analysis and Design, Tamara Munzner
Image Source : https://www.researchgate.net/figure/Basic-network-topology-of-a-

Networks and Trees

- Example: Force-Directed Placement
  - One of the most widely used idioms for node-link layout using connection marks is force directed placement
  - Application areas:
    - network visualization, large graph visualization, knowledge representation, system management, or mesh visualization
  - Minimize the number of distracting artifacts such as edge crossings and node overlaps

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

- Example: Force-Directed Placement
  - Force-directed graph drawing algorithms assign forces among the set of edges and the set of nodes of a graph drawing.
  - Spring-like attractive forces based on Hooke's law are used to attract pairs of endpoints of the graph's edges towards each other.
  - simultaneously repulsive forces based on Coulomb's law are used to separate all pairs of nodes

Source: Visualization Analysis and Design, Tamara Munzner

(https://www.youtube.com/watch?v=T6tS0HIZbMA)

Networks and Trees

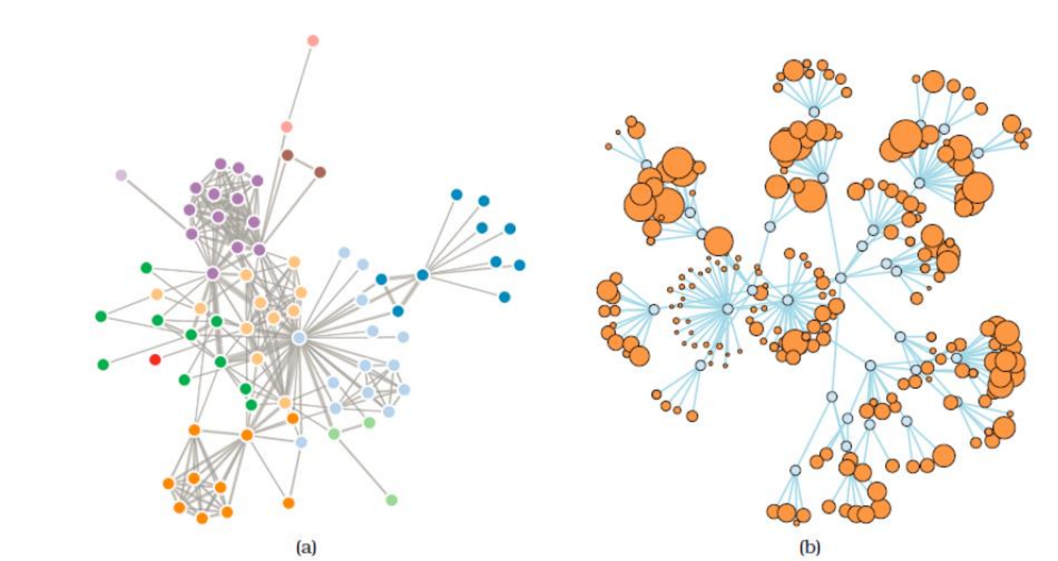
- Example: Force-Directed Placement

Idiom	**Force-Directed Placement**
What: Data	Network.
How: Encode	Point marks for nodes, connection marks for links.
Why: Tasks	Explore topology, locate paths.
Scale	Nodes: dozens/hundreds. Links: hundreds. Node/link density: $L < 4N$

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

- Example: Force-Directed Placement



(a) with size coding for link attributes. (b) with size coding for node attributes.

Source: Visualization Analysis and Design, Tamara Munzner

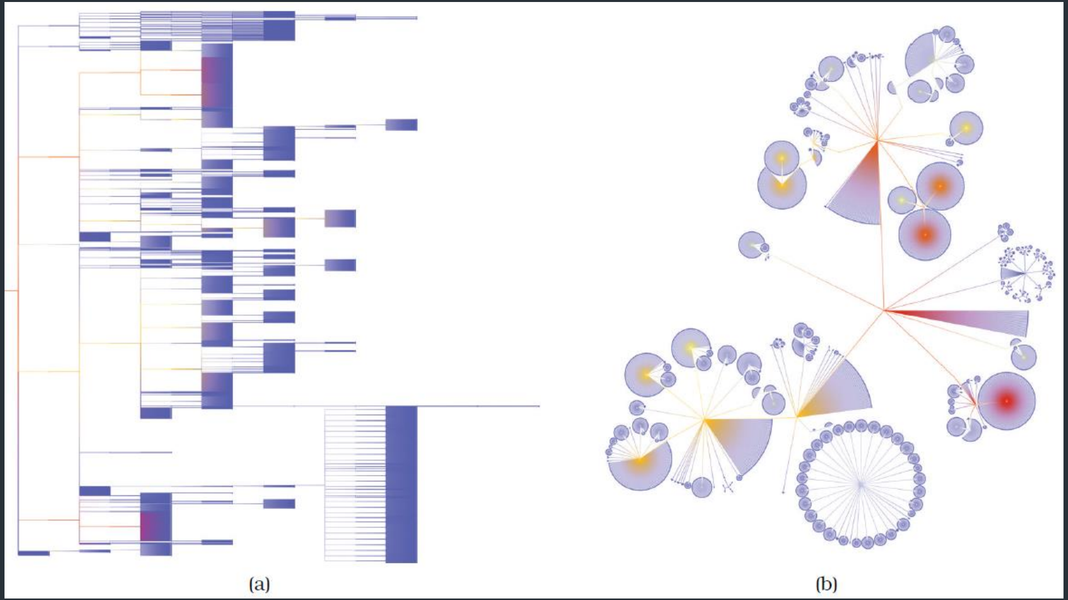
Networks and Trees

- Example: Force-Directed Placement

Two layouts of a 5161-node tree.

(a) Rectangular horizontal node–link layout.

(b) BubbleTree node–link layout.



Source: Visualization Analysis and Design, Tamara Munzner

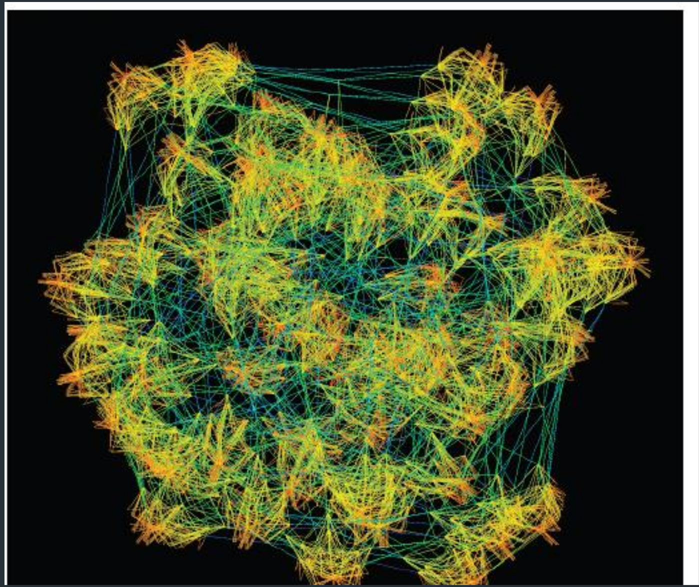
Networks and Trees

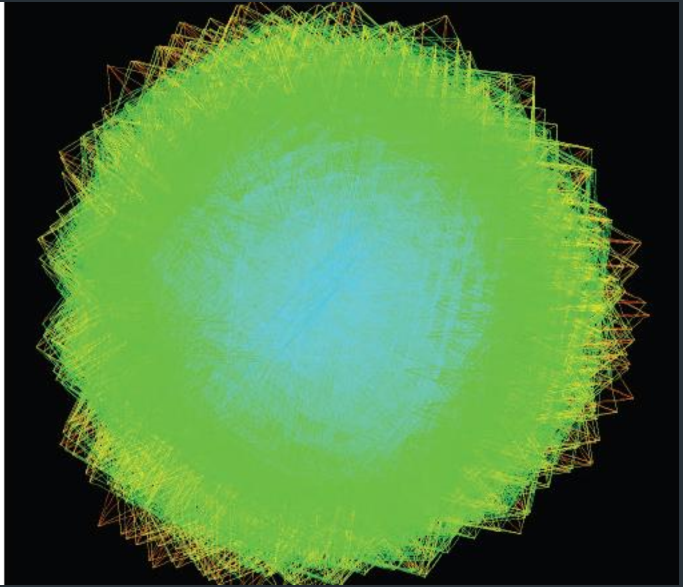
- Example: Force-Directed Placement
- Strength
  - Very easy to implement. Relatively easy to understand and explain at a conceptual level
- Weakness
  - Nondeterministic layout
    - different each time the algorithm is run
  - Scalability
    - visual and time complexity
  - Very brittle
    - based on optimization algorithms - get stuck in local minimum

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

- Scalable Force-Directed Placement
  - edges are colored by length





Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

- Multilevel Force-Directed Placement

Idiom	Multilevel Force-Directed Placement (sfdp)
What: Data	Network.
What: Derived	Cluster hierarchy atop original network.
What: Encode	Point marks for nodes, connection marks for links.
Why: Tasks	Explore topology, locate paths and clusters.
Scale	Nodes: 1000–10,000. Links: 1000–10,000. Node/link density: $L < 4N$.

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

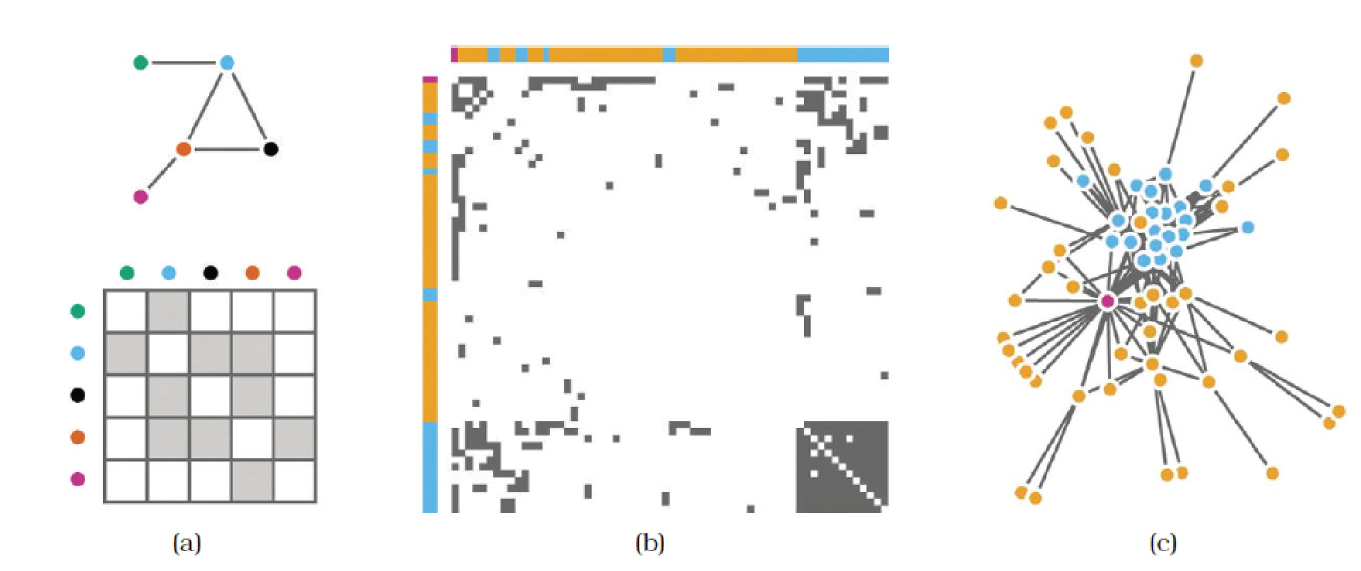
Adjacency Matrix View

- network can be visually encoded as an **adjacency matrix** view
- network is transformed into the derived dataset of a table with two key attributes that are separate
- full lists of every node in the network, and one value attribute for each cell records whether a link exists between the nodes that index the cell

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

- Adjacency Matrix View



(a) Node–link and matrix views of small network.
(b) Matrix view of larger network. (c) Node–link view of larger network

Source: Visualization Analysis and Design, Tamara Munzner

Networks and Trees

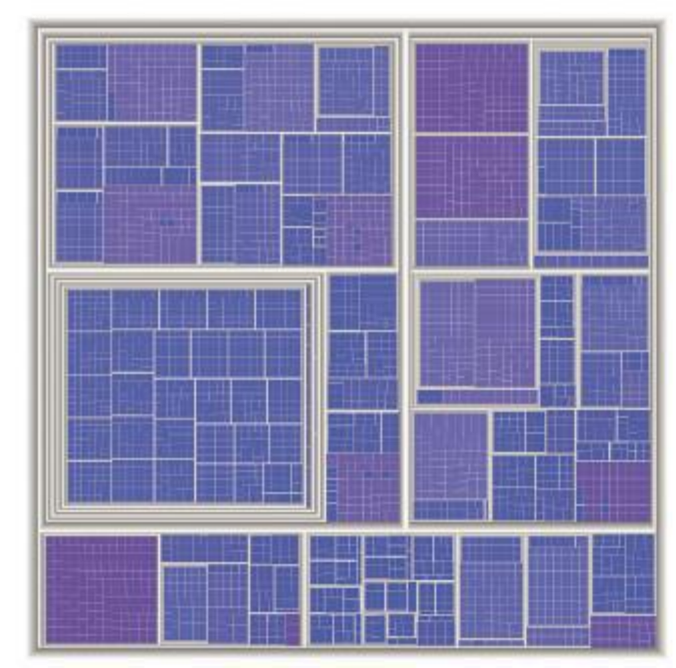
- Adjacency Matrix View

Idiom	Adjacency Matrix View
What: Data	Network.
What: Derived	Table: network nodes as keys, link status between two nodes as values.
How: Encode	Area marks in 2D matrix alignment.
Scale	Nodes: 1000. Links: one milllion.

Source: Visualization Analysis and Design, Tamara Munzner

Refer : https://people.revoledu.com/kardi/tutorial/GraphTheory/Adjacency-Matrix.html

Networks and Trees

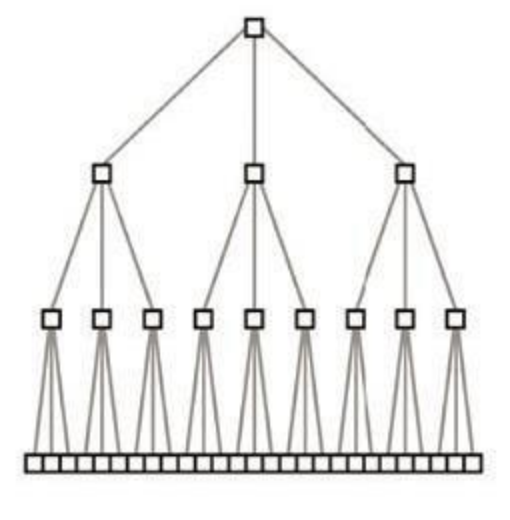


Containment

- Containment marks are very effective at showing complete information about hierarchical structure, in contrast to connection marks that only show pairwise relationships between two items at once.
- Tree Maps: The idiom of tree maps is an alternative to node–link tree drawings, where the hierarchical relationships are shown with containment rather than connection.

Source: Visualization Analysis and Design, Tamara Munzner

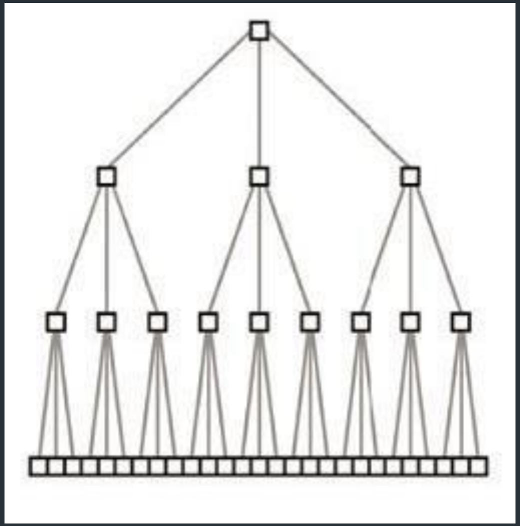
Seven different visual encoding idioms for tree data



- **Using different combinations of visual channels.**
- (a) Rectilinear vertical node–link
- Connection to show link relationships
- Vertical spatial position
  - tree depth
- Horizontal spatial position
  - sibling order.

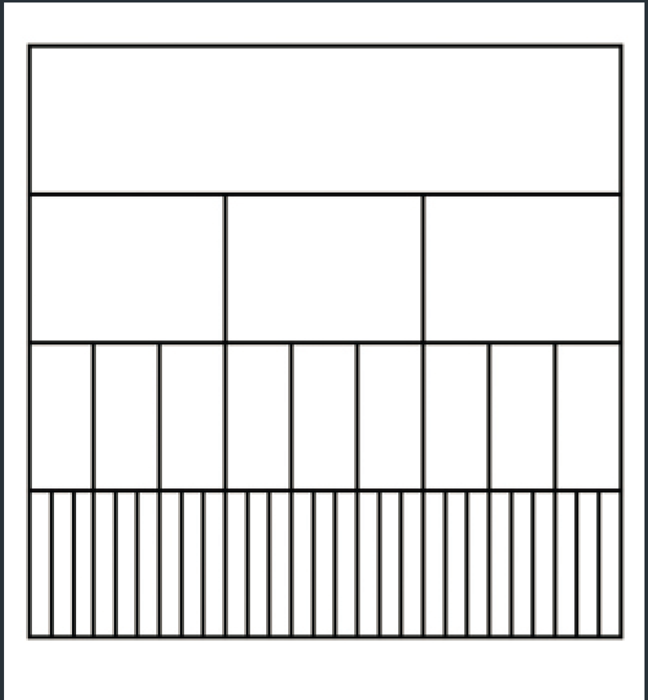
Seven different visual encoding idioms for tree data

- (b) Icicle



Source: Visualization Analysis and Design, Tamara Munzner

Seven different visual encoding idioms for tree data

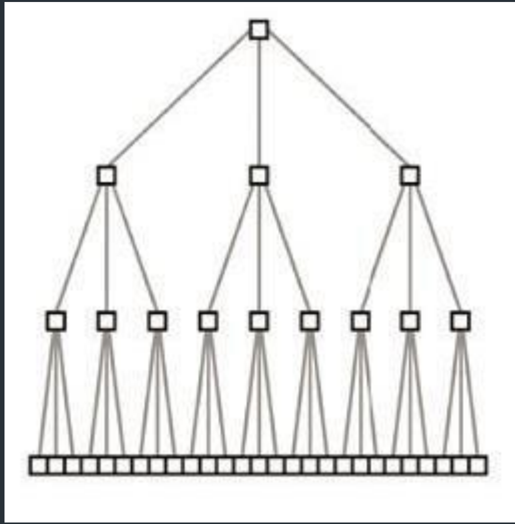


- **Using different combinations of visual channels.**
- (b) Icicle
- Vertical spatial position and size
  - tree depth
- Horizontal spatial position
  - link relationships and sibling order

Source: Visualization Analysis and Design, Tamara Munzner

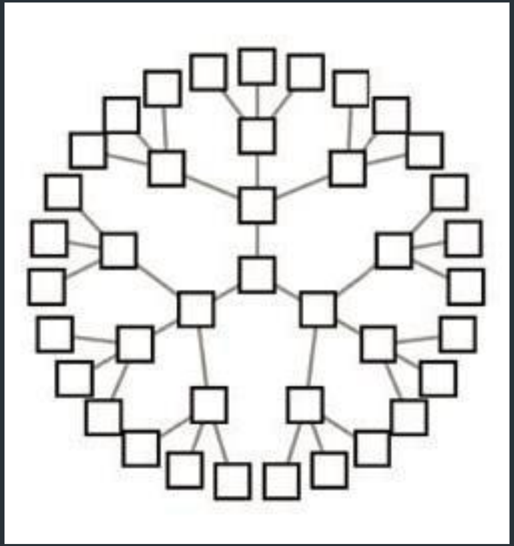
Seven different visual encoding idioms for tree data

- © Radial Node link



Source: Visualization Analysis and Design, Tamara Munzner

Seven different visual encoding idioms for tree data

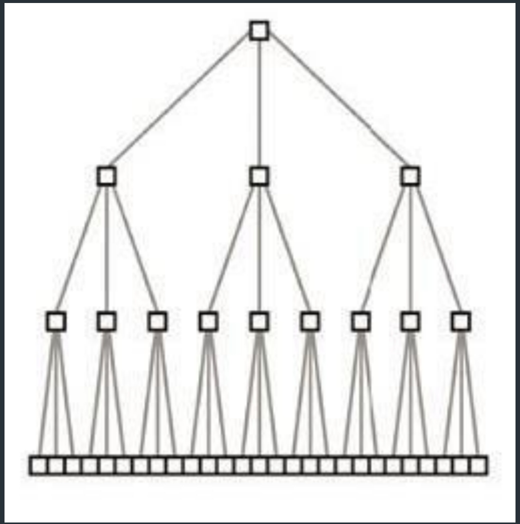


- **Using different combinations of visual channels.**
- (c) Radial node–link
- connection to show link relationships
- Radial depth spatial position
  - tree depth
- radial angular position
  - sibling order

Source: Visualization Analysis and Design, Tamara Munzner

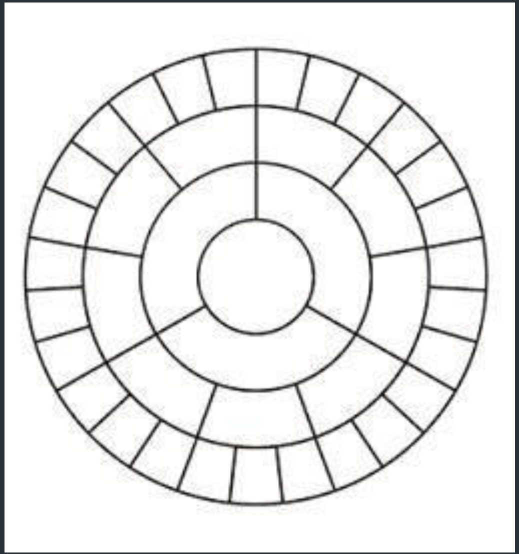
Seven different visual encoding idioms for tree data

- (d) Concentric circles



Source: Visualization Analysis and Design, Tamara Munzner

Seven different visual encoding idioms for tree data

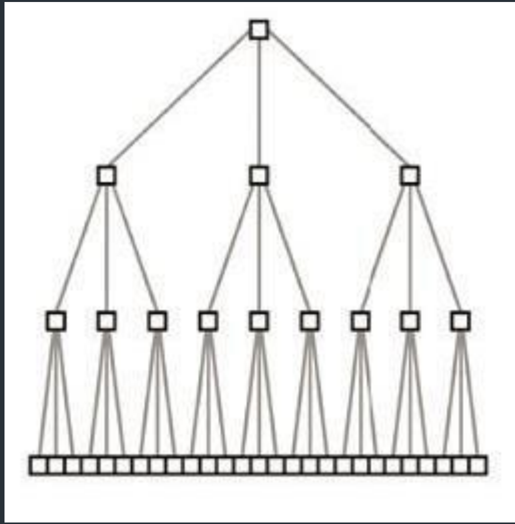


- **Using different combinations of visual channels.**
- (d) Concentric circles, with radial depth spatial position and size showing tree depth and radial angular spatial position showing link relationships and sibling order.

Source: Visualization Analysis and Design, Tamara Munzner

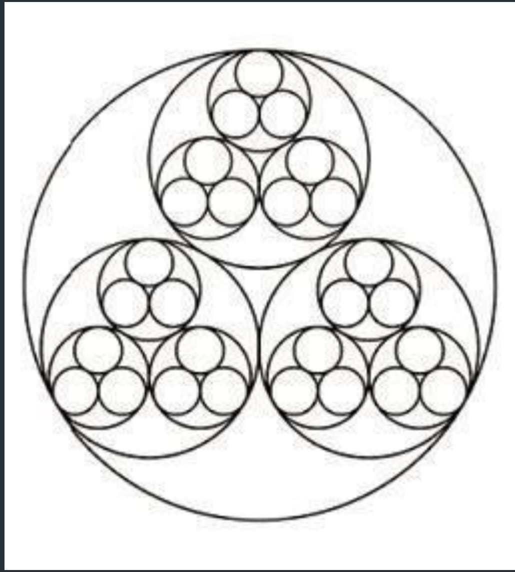
Seven different visual encoding idioms for tree data

- (e) Nested circles



Source: Visualization Analysis and Design, Tamara Munzner

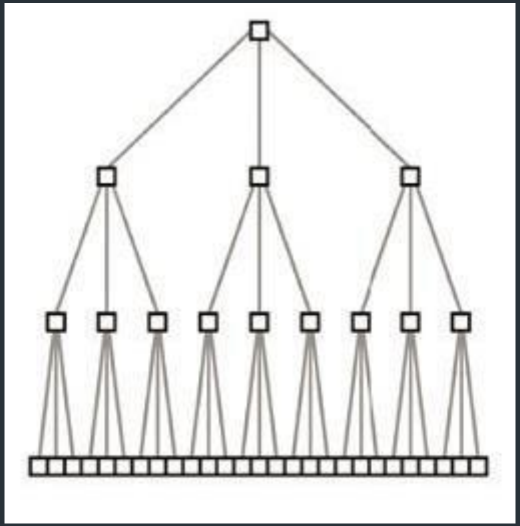
Seven different visual encoding idioms for tree data



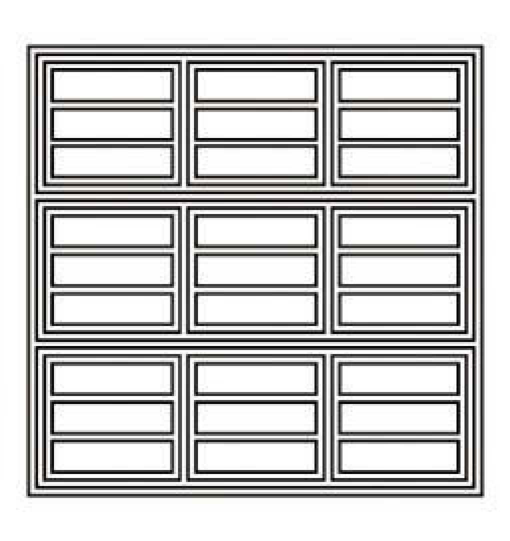
- **Using different combinations of visual channels.**
- (e) Nested circles, using radial containment, with nesting level and size showing tree depth

Seven different visual encoding idioms for tree data

- (f) Treemap



Seven different visual encoding idioms for tree data

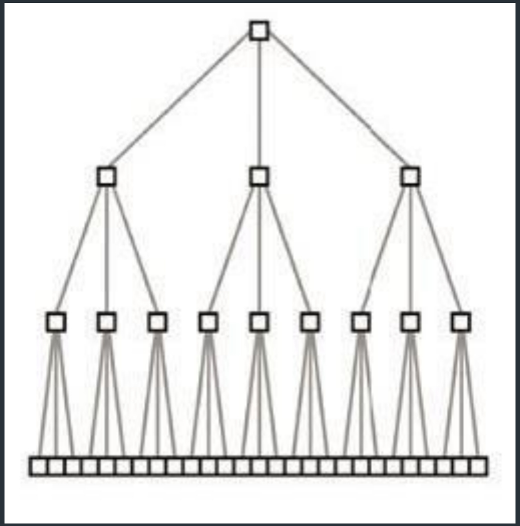


- **Using different combinations of visual channels.**
- (f) Treemap, using rectilinear containment, with nesting level and size showing tree depth

Source: Visualization Analysis and Design, Tamara Munzner

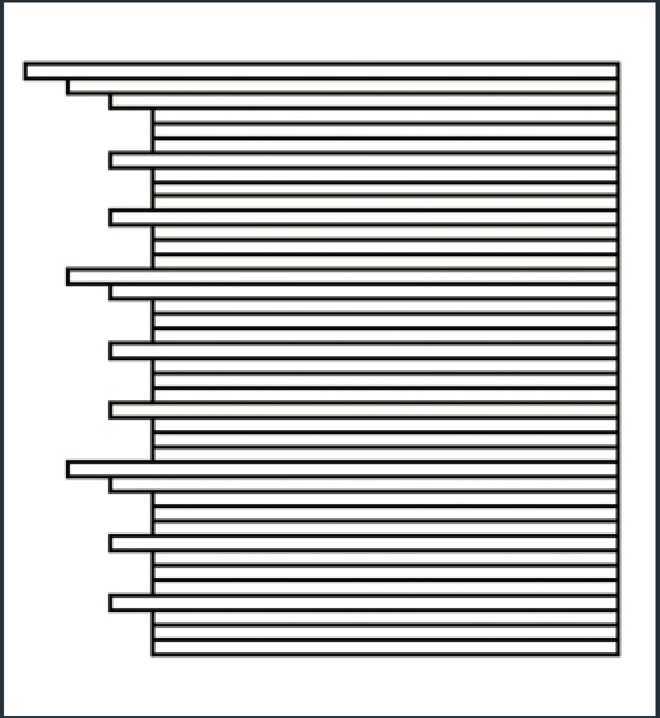
Seven different visual encoding idioms for tree data

- (g) Indented outline



Source: Visualization Analysis and Design, Tamara Munzner

Seven different visual encoding idioms for tree data

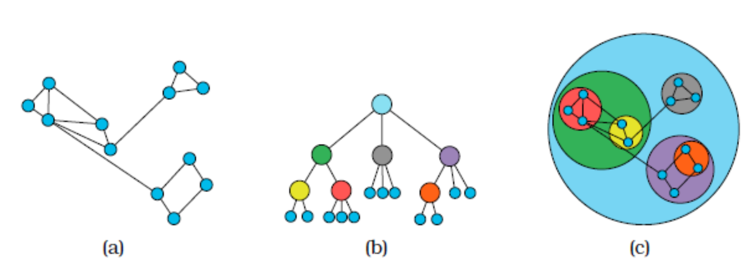


- **Using different combinations of visual channels.**
- (g) Indented outline, with horizontal spatial position showing tree depth and link relationships and vertical spatial position showing sibling order

Source: Visualization Analysis and Design, Tamara Munzner

Compound Network

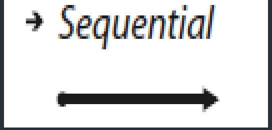
- A compound network is a combination of a **network** and a **tree** on top of it, where the nodes in the network are the leaves of the tree.
- Thus, the interior nodes of the tree encompass multiple network nodes.
- (a) shows a network (b) shows a cluster hierarchy built on top of it. (c) shows a combined view using of containment marks for the associated hierarchy and connection marks for the original network links

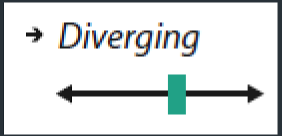


Source: Visualization Analysis and Design, Tamara Munzner

Map Color and Other Channels

Introduction - Color





- The ***color*** is best understood in terms of three separate channels: luminance, hue, and saturation.
- Sequential ordered colormaps
  - show a progression of an attribute from a minimum to a maximum value
- Diverging ordered colormaps
  - have a visual indication of a zero point in the center where the attribute values diverge to negative on one side and positive on the other.
- Bivariate colormaps are designed to show two attributes simultaneously using carefully designed combinations of luminance, hue, and saturation

Source: Visualization Analysis and Design, Tamara Munzner

Color Space

- The **color space** of what colors the human visual system can detect is three dimensional. That is, it can be adequately described using three separate axes.
- There are many ways to mathematically describe color as a space and to transform colors from one such space into another.
- Some of these are extremely convenient for computer manipulation, while others are a better match with the characteristics of human vision.

Source: Visualization Analysis and Design, Tamara Munzner

Color Space

- **RGB System**
  - The most common color space in computer graphics is the system where colors are specified as triples of red, green, and blue values
  - not useful as separable channels

Source: Visualization Analysis and Design, Tamara Munzner

Color Space

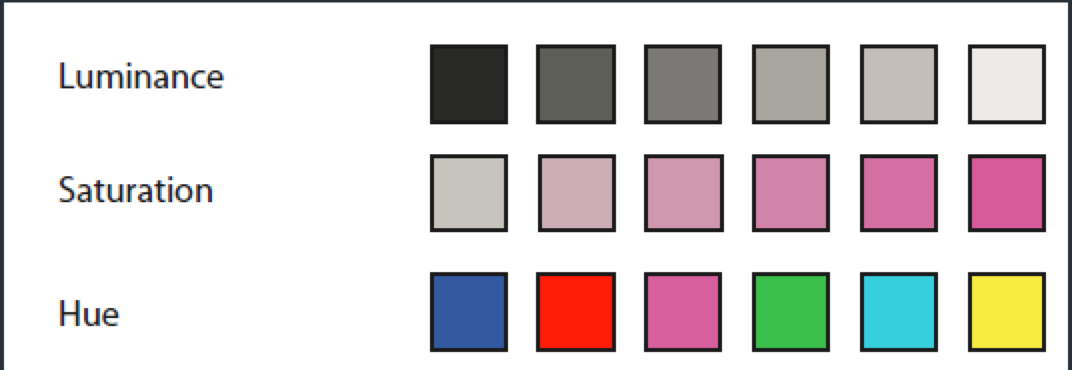
HSL System

- The **hue–saturation–lightness** or **HSL** system is more intuitive and is heavily used by artists and designers.
- The **hue** axis captures what we normally think of as pure colors that are not mixed with white or black: red, blue, green, yellow, purple, and so on.
- The **saturation** axis is the amount of white mixed with that pure color. For instance, pink is a partially desaturated red.
- The **lightness** axis is the amount of black mixed with a color.

Source: Visualization Analysis and Design, Tamara Munzner

Luminance, Saturation and Hue

- Color can be confusing in visual analysis because it is sometimes used as a magnitude channel and sometimes as an identity channel.
- *Luminance* and *saturation* are **magnitude** channels, while *hue* is a **identity** channel.



Source: Visualization Analysis and Design, Tamara Munzner

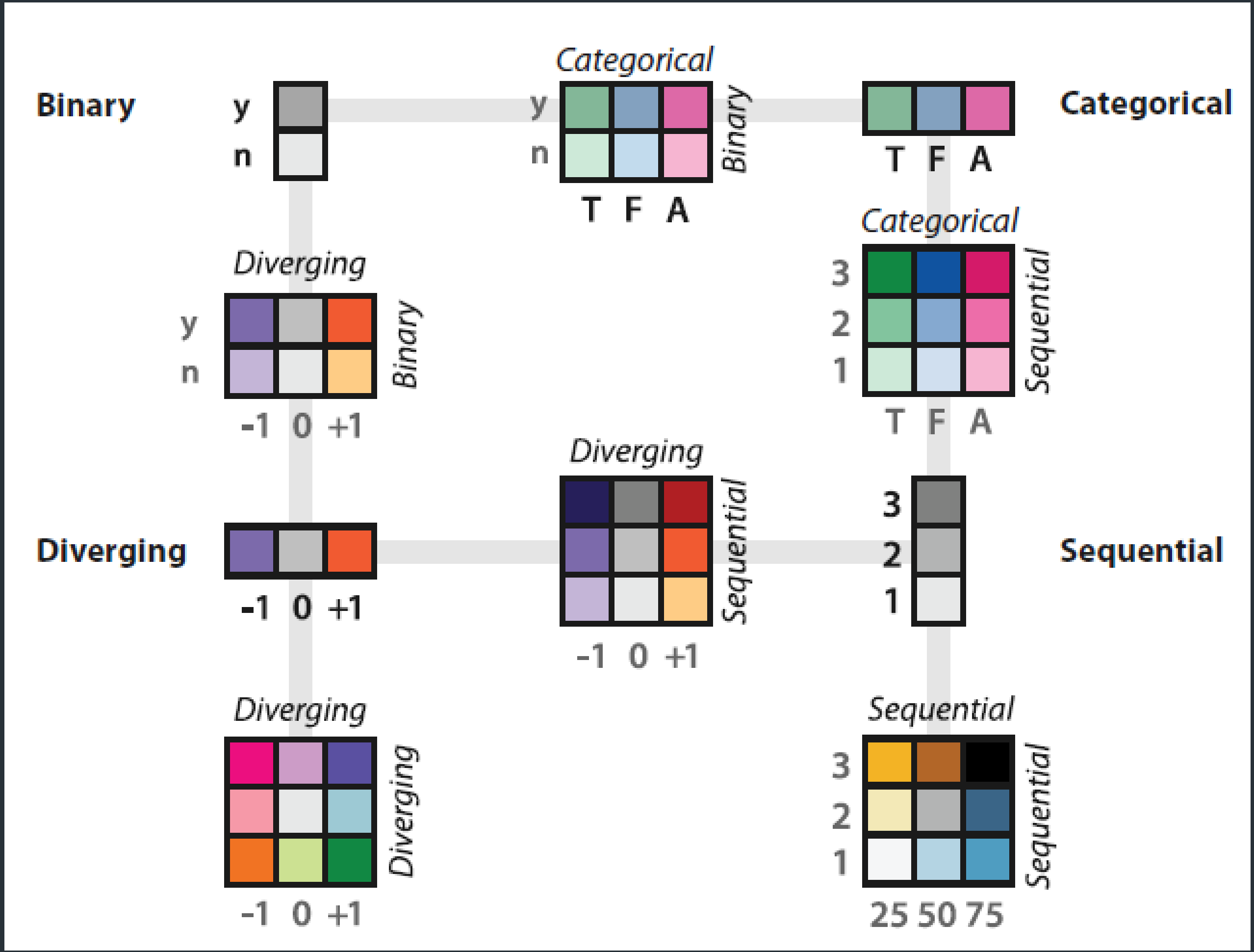
Transparency

- A fourth channel strongly related to the other three color channels is **transparency**
- Transparency is used most often with superimposed layers, to create a foreground layer that is distinguishable from the background layer

Source: Visualization Analysis and Design, Tamara Munzner

Colormaps

- A **colormap** specifies a mapping between colors and data values; that is, a visual encoding with color
- Colormaps can be **categorical** or **ordered**, and ordered colormaps can be either **sequential** or **diverging**.



Source: Visualization Analysis and Design, Tamara Munzner

Ordered Colormaps

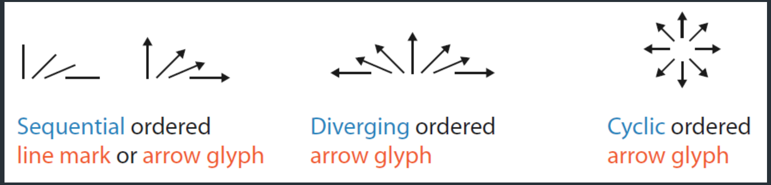
- An **ordered** colormap is appropriate for encoding ordinal or quantitative attributes.
- A **sequential** colormap ranges from a minimum value to a maximum value.
- A **diverging** colormap has two hues at the endpoints and a neutral color as a midpoint, such as white, gray, or black, or a high-luminance color such as yellow

Other Channels

- **Size Channels**
  - Size is a magnitude channel suitable for ordered data.
  - Length is one-dimensional (1D) size; more specifically, height is vertical size and width is horizontal size. Area is two-dimensional (2D) size, and volume is three-dimensional (3D) size.

Other Channels

- **Angle Channels**
  - The *angle* channel encodes magnitude information based on the **orientation** of a mark: the direction that it points.
  - With **angle**, the orientation of one line is judged with respect to another line.
  - With **tilt**, an orientation is judged against the global frame of the display.

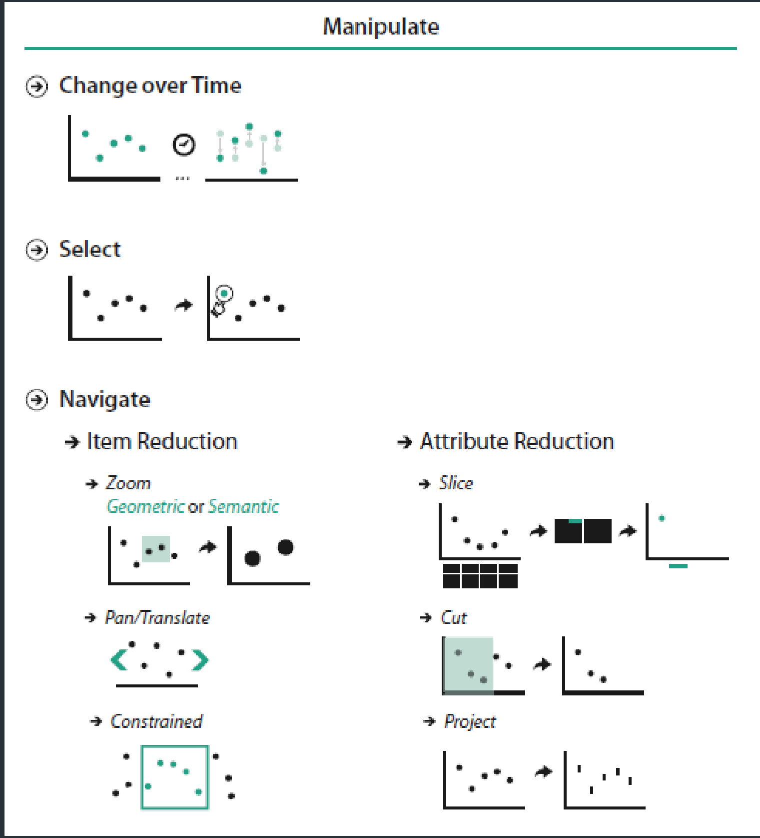


Source: Visualization Analysis and Design, Tamara Munzner

Other Channels

- **Shape Channels**
  - Shape as a identity channel that can be used with point and line marks.
  - Applying the shape channel to line marks results in stipple patterns such as dotted and dashed lines.
- **Motion Channels**
  - Several kinds of **motion** are also visual channels, including **direction** of motion, **velocity** of motion, and flicker **frequency**.
  - Motion is less studied than other channels

Manipulate View



- Introduction
  - A change could be made from **one choice to another** to change idioms, and any of the **parameters** for a particular idiom can be changed.
  - Any **aspect** of visual encoding can be changed, including the **ordering**, any other choice pertaining to the **spatial arrangement**, and the use of other visual channels such as **color**.

Source: Visualization Analysis and Design, Tamara Munzner

Manipulate View

- Why change?
- Datasets are often sufficiently large and complex that showing everything at once in a single static view would lead to overwhelming visual clutter

Manipulate View

- **Change View over Time**
- **The possibilities for how the view changes can be**
  - Change the encoding
  - Change the arrangement
  - Change the order
  - Change the viewpoint
  - Change which attributes are filtered
  - Change the aggregation level, and so on
  - Ex: Lineup

Source: Visualization Analysis and Design, Tamara Munzner

Manipulate View

- Select Elements
  - Allowing users to select one or more elements of interest in a vis is a fundamental action that supports nearly every interactive idiom.
  - Selection Design Choices
  - Highlighting
  - Selection Outcomes

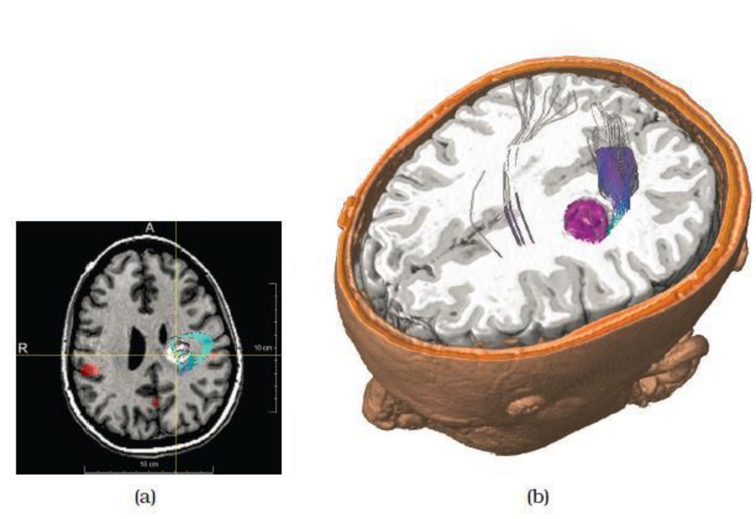
Manipulate View

- **Navigate: Changing Viewpoint**
  - Large and complex datasets often cannot be understood from only a single point of view.
  - Many interactive vis systems support a metaphor of navigation, analogous to navigation through the physical world.
  - In these, the spatial layout is fixed and navigation acts as a change of the viewpoint.
  - **Zooming** (Geometric, Semantic)
  - Panning
  - Rotating

Source: Visualization Analysis and Design, Tamara Munzner

Manipulate View

- **Navigate: Reducing Attributes**
- The geometric intuitions that underlie the metaphor of navigation with a virtual camera also lead to a set of design choices for reducing the number of attributes:
- Slice - Axis-aligned slice
- Cut - Axis-aligned slice
- Project
  - orthographic projection
  - perspective projection
  - map projections



Source: Visualization Analysis and Design, Tamara Munzner