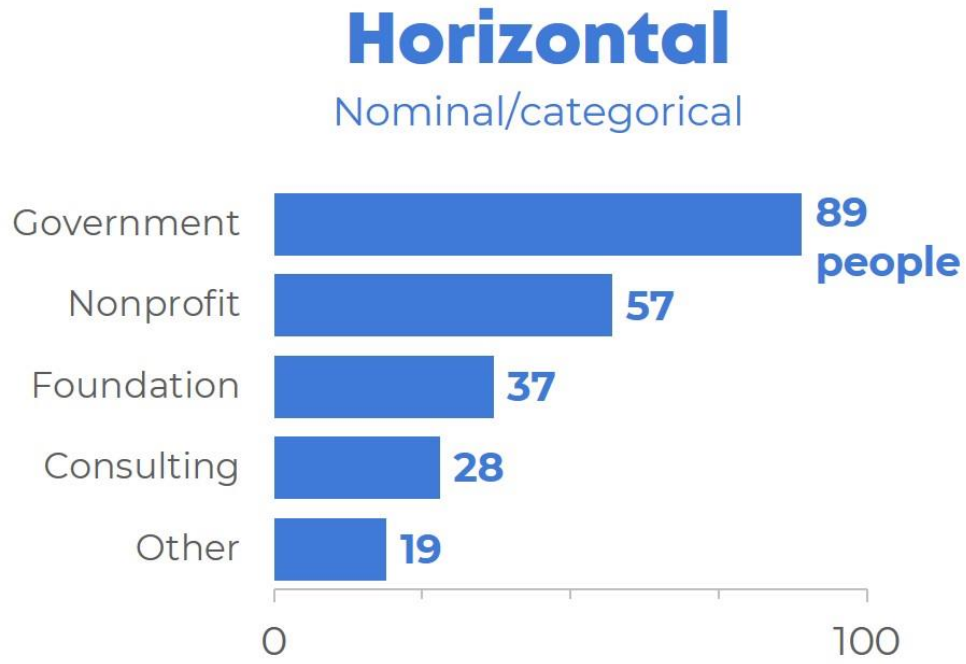
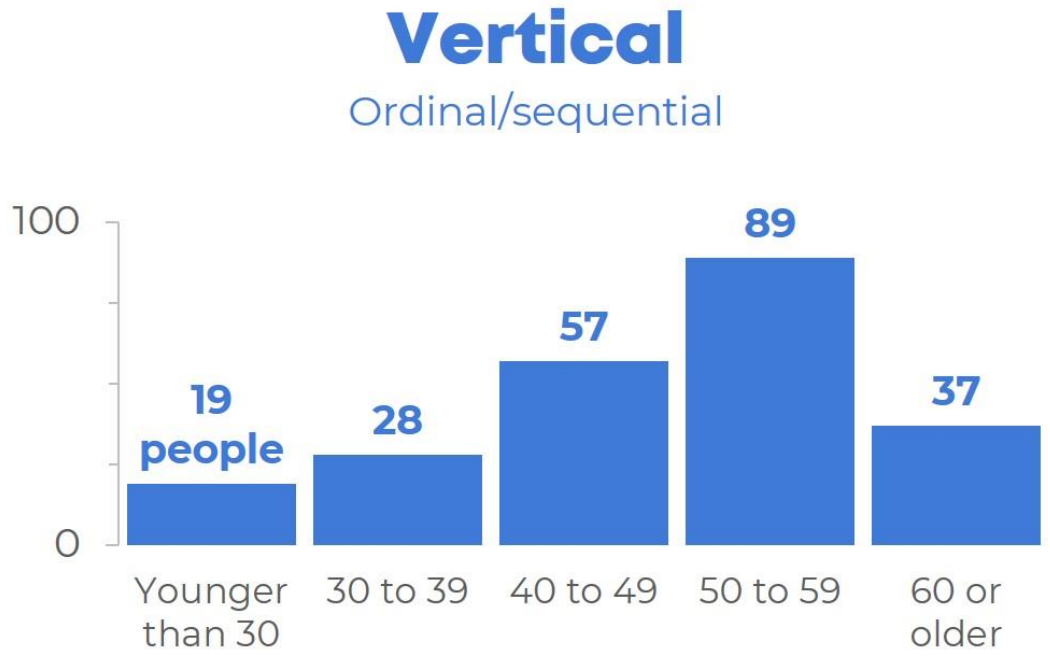
Common Visualization Idioms

Common Visualization Idioms

- Comparison
  - Bar Chart
  - Column Chart
- Composition
  - Pie Chart
  - Waterfall Chart
- Distribution
  - 3D Area Chart
  - Line Histogram
- Relationship
  - Bubble Chart
  - Scatter Chart
- Spatial Map

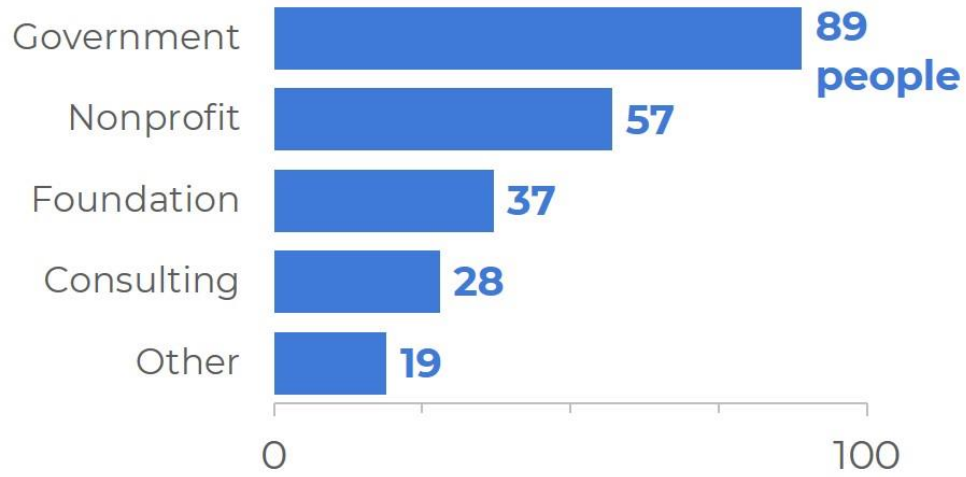
- Both bar and column charts display discrete categorical data and answer the question of ‘how many?’ or ‘how much?’ in each category.





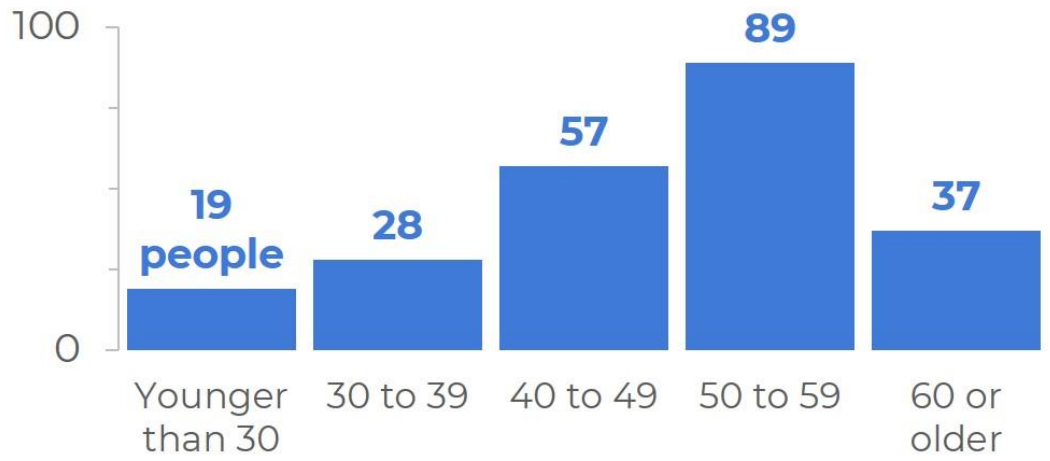
Horizontal

Nominal/categorical



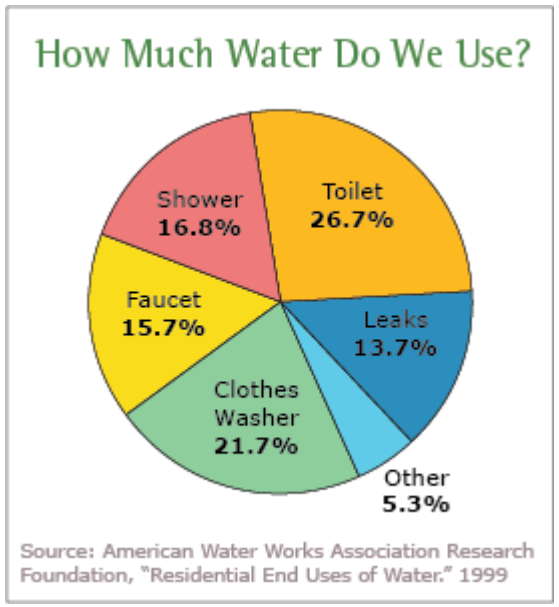
Vertical

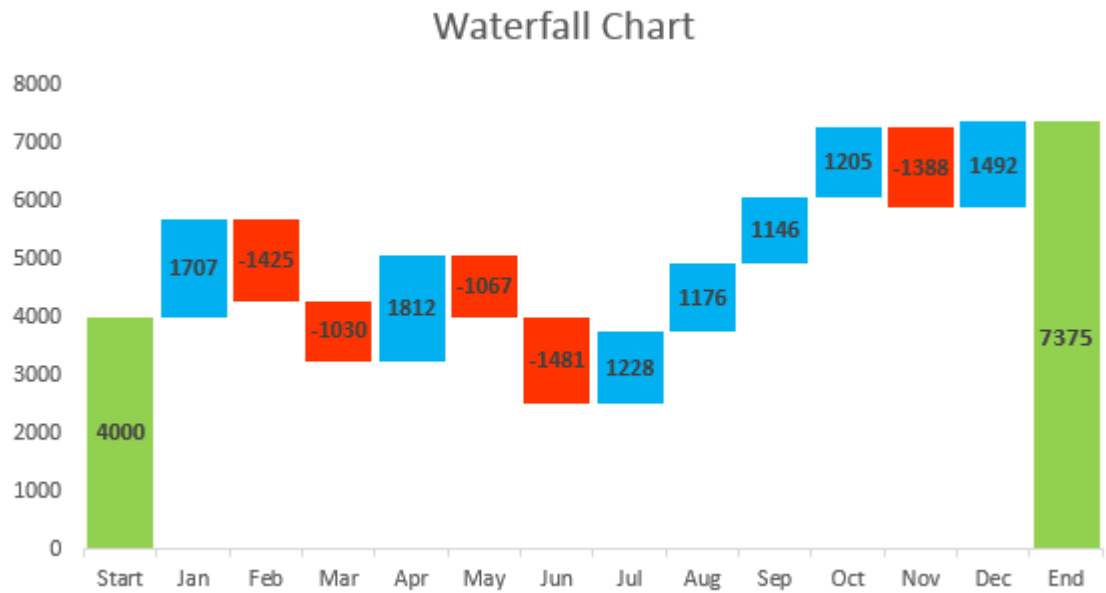
Ordinal/sequential



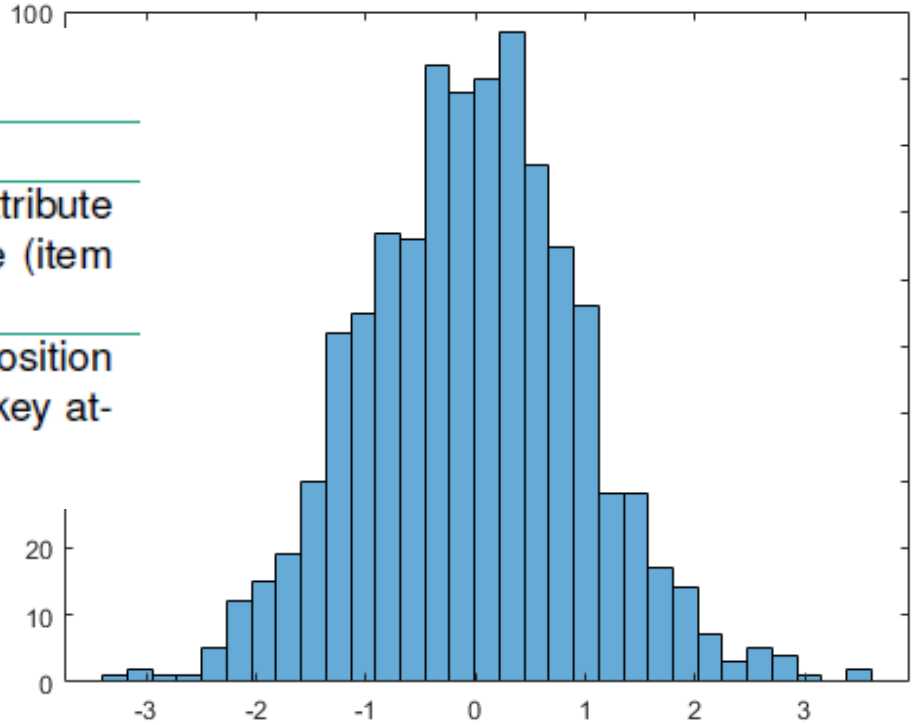
Idiom	**Bar Charts**
What: Data	Table: one quantitative value attribute, one categorical key attribute.
How: Encode	Line marks, express value attribute with aligned vertical position, separate key attribute with horizontal position.
Why: Task	Lookup and compare values.
Scale	Key attribute: dozens to hundreds of levels.

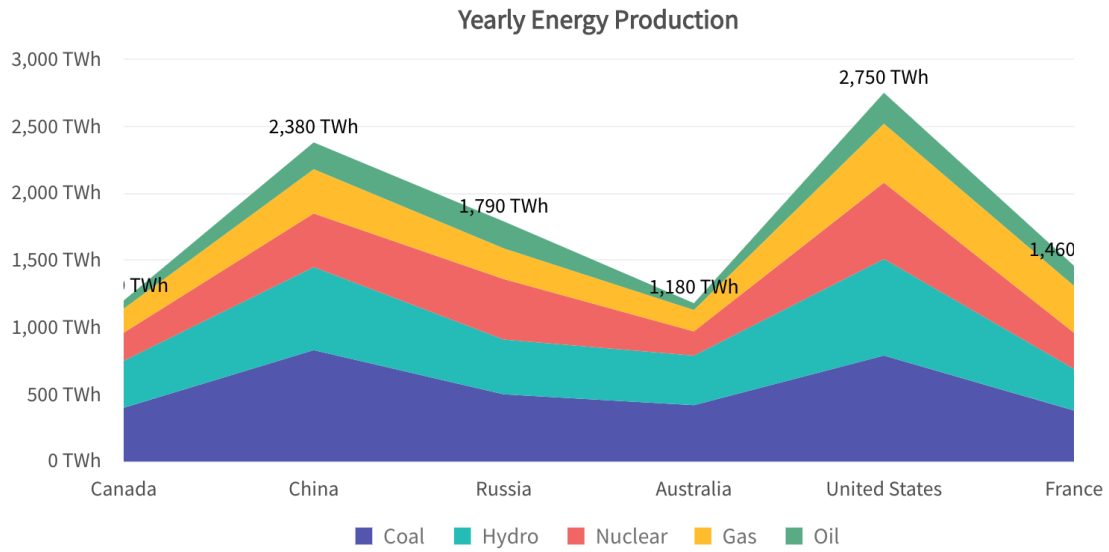
Idiom	**Pie Charts**
What: Data	Table: one quantitative attribute, one categorical attribute.
Why: Task	Part–whole relationship.
How: Encode	Area marks (wedges) with angle channel; radial layout.
Scale	One dozen categories.

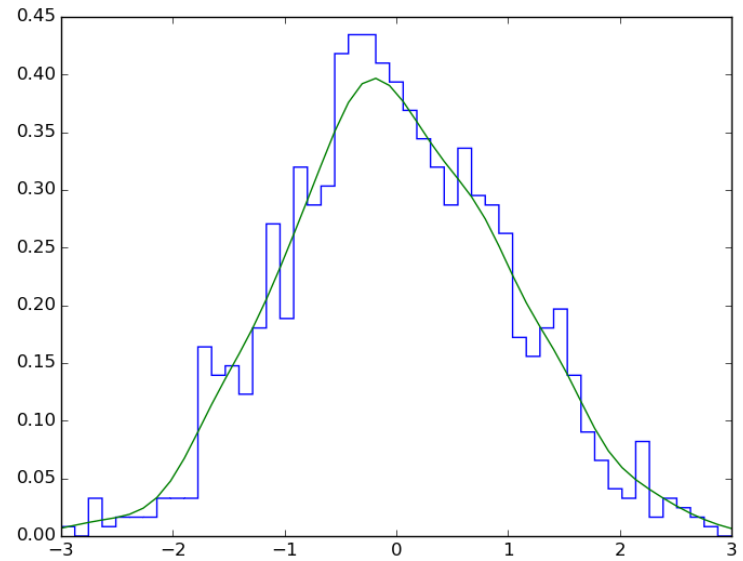




Idiom	Histograms
What: Data	Table: one quantitative value attribute.
What: Derived	Derived table: one derived ordered key attribute (bin), one derived quantitative value attribute (item count per bin).
How: Encode	Rectilinear layout. Line mark with aligned position to express derived value attribute. Position: key attribute.

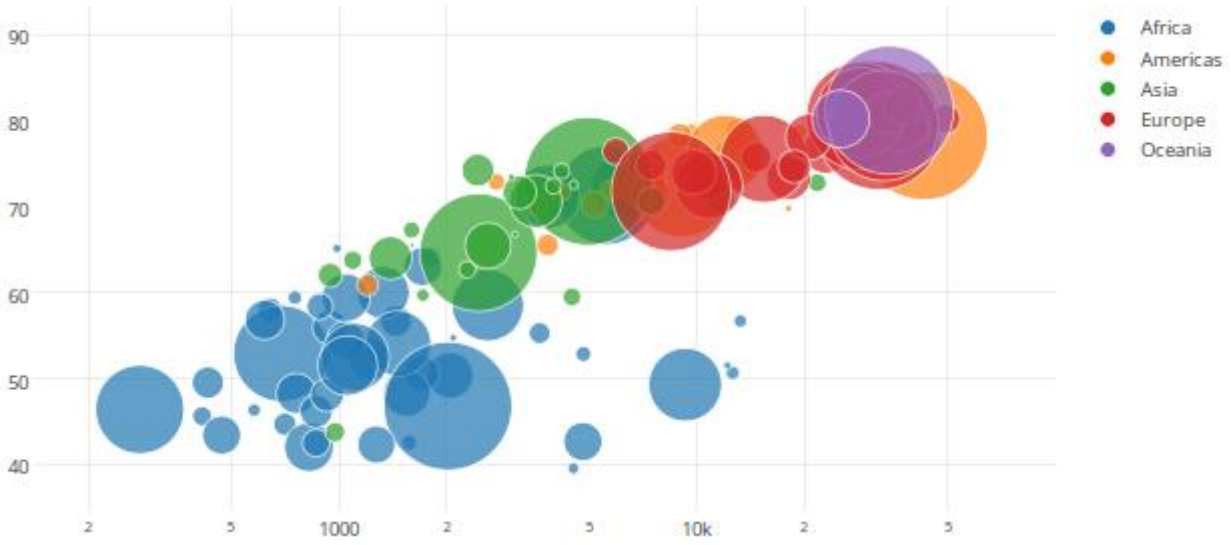


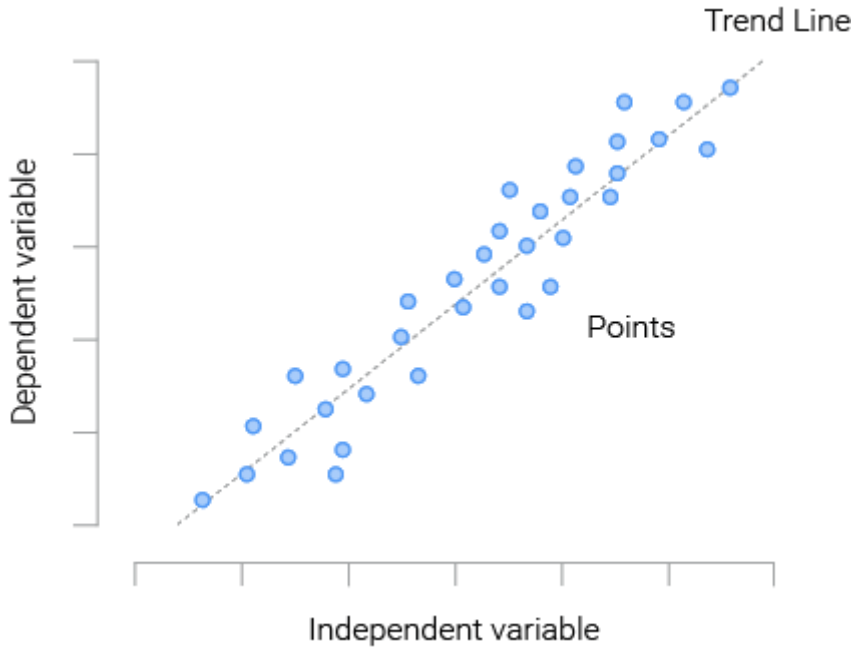




Area Chart

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.



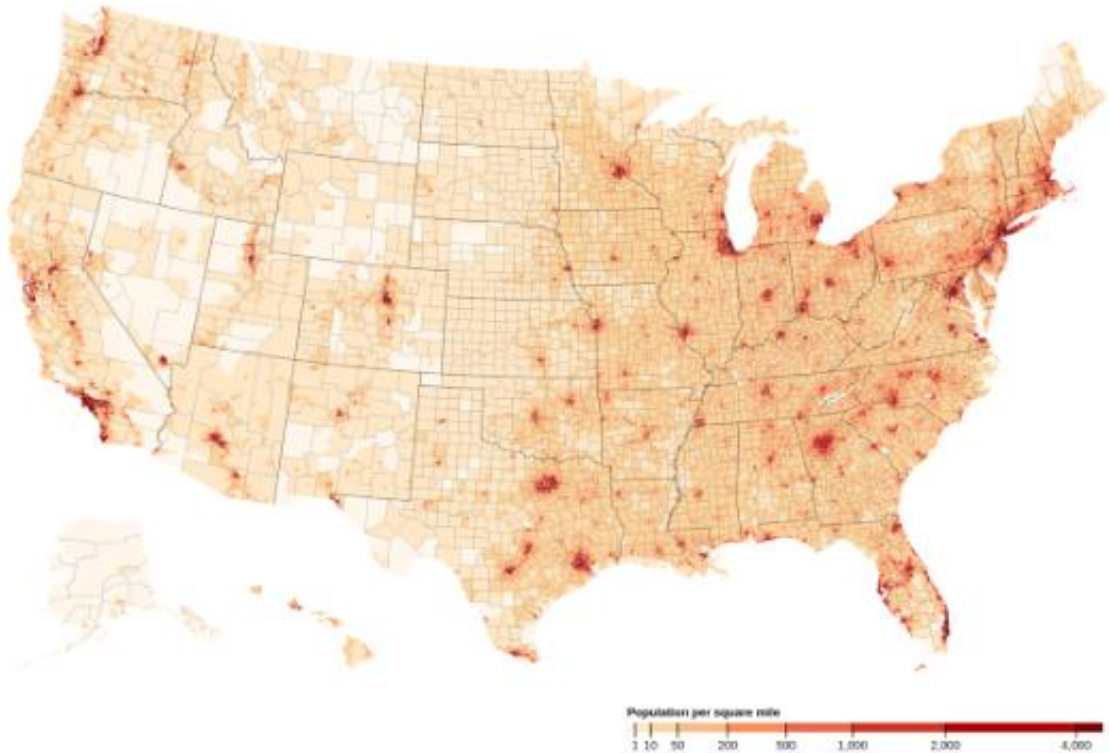


Idiom	**Dot Charts**
What: Data	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Express value attribute with aligned vertical position and point marks. Separate/order into horizontal regions by key attribute.

Idiom	**Line Charts**
What: Data	Table: one quantitative value attribute, one ordered key attribute.
How: Encode	Dot chart with connection marks between dots.
Why	Show trend.
Scale	Key attribute: hundreds of levels.

Figure 7.8. Line charts versus dot charts. (a) Dot charts use a point mark to show the value for each item. (b) Line charts use point marks connected by lines between them.

U.S. Population Density

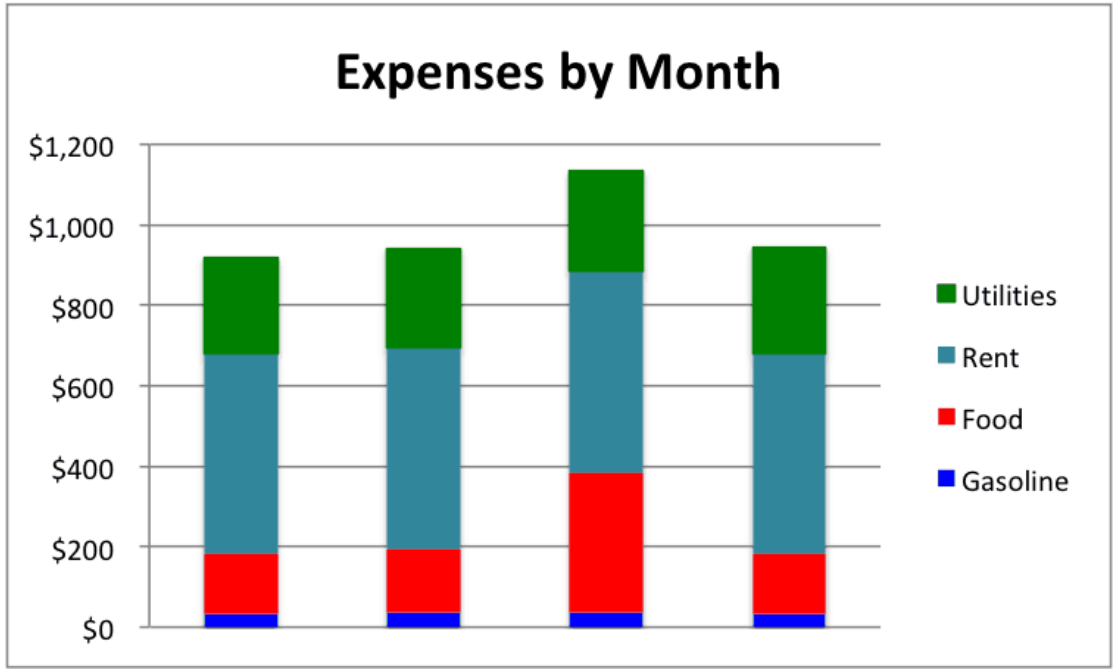


Source: American Community Survey, 2014 5-Year Estimate

Open

- To ascertain which chart is best for each type listed above, you should know first:
- How many variables you want to show
- How many data points you will display for each variable
- How you want to present values over a period of time, or among other values or data set

Idiom	**Stacked Bar Charts**
What: Data	Multidimensional table: one quantitative value attribute, two categorical key attributes.
How: Encode	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	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



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

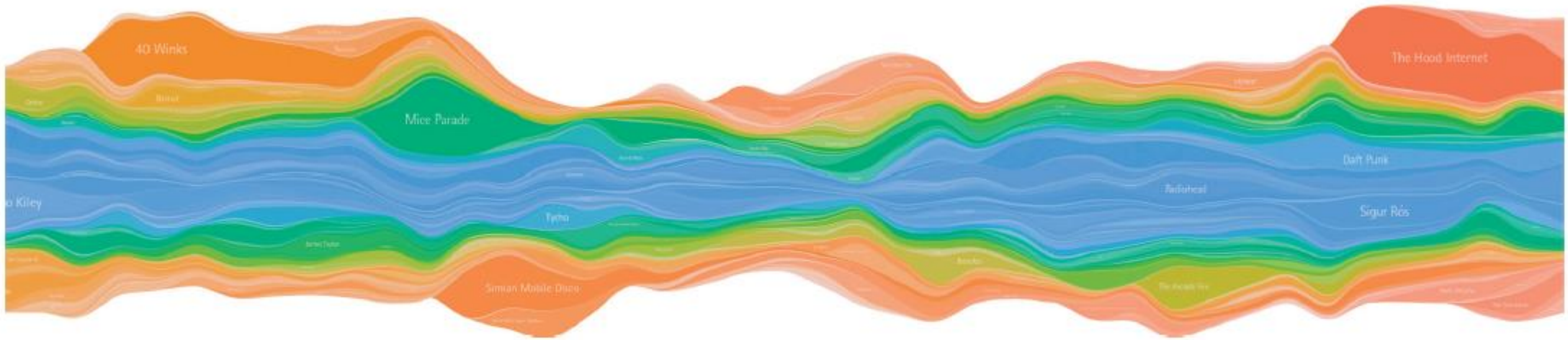


Figure 7.6. Streamgraph of music listening history. From [Byron and Wattenberg 08, Figure 0].

Idiom	**Boxplot Charts**
What: Data	Table: many quantitative value attributes.
What: Derived	Five quantitative attributes for each original attribute, representing its distribution.
Why: Tasks	Characterize distribution; find outliers, extremes, averages; identify skew.
How: Encode	One glyph per original attribute expressing derived attribute values using vertical spatial position, with 1D list alignment of glyphs into separated with horizontal spatial position.
How: Reduce	Item aggregation.
Scale	Items: unlimited. Attributes: dozens.

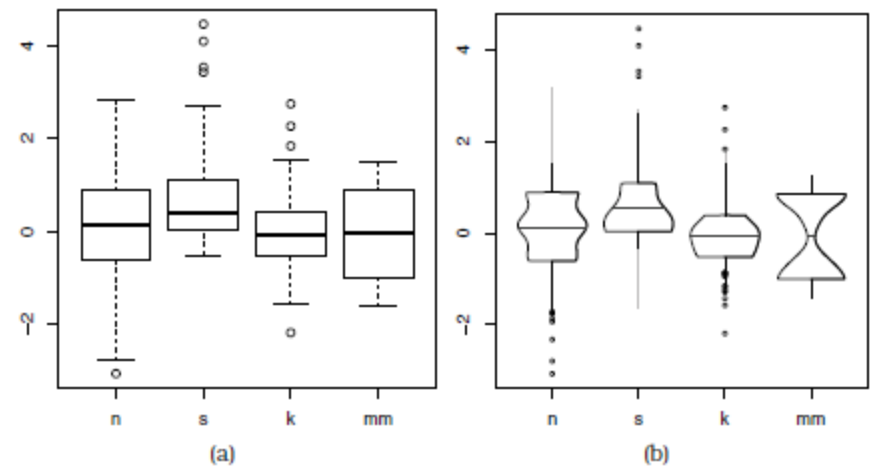
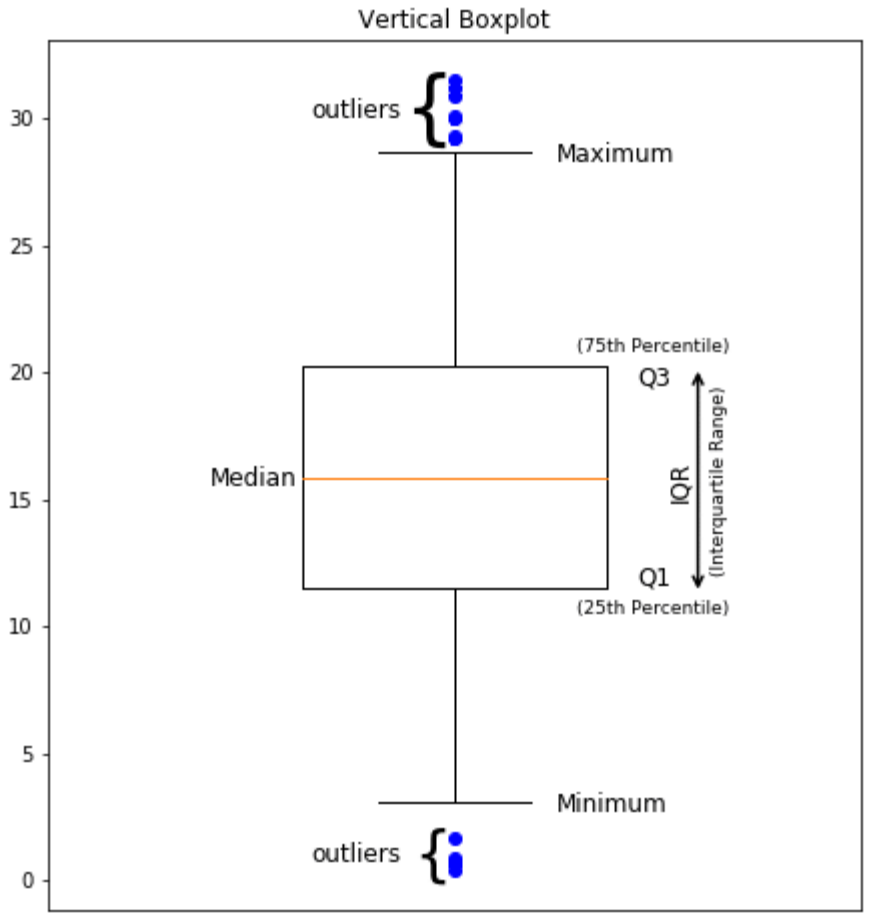
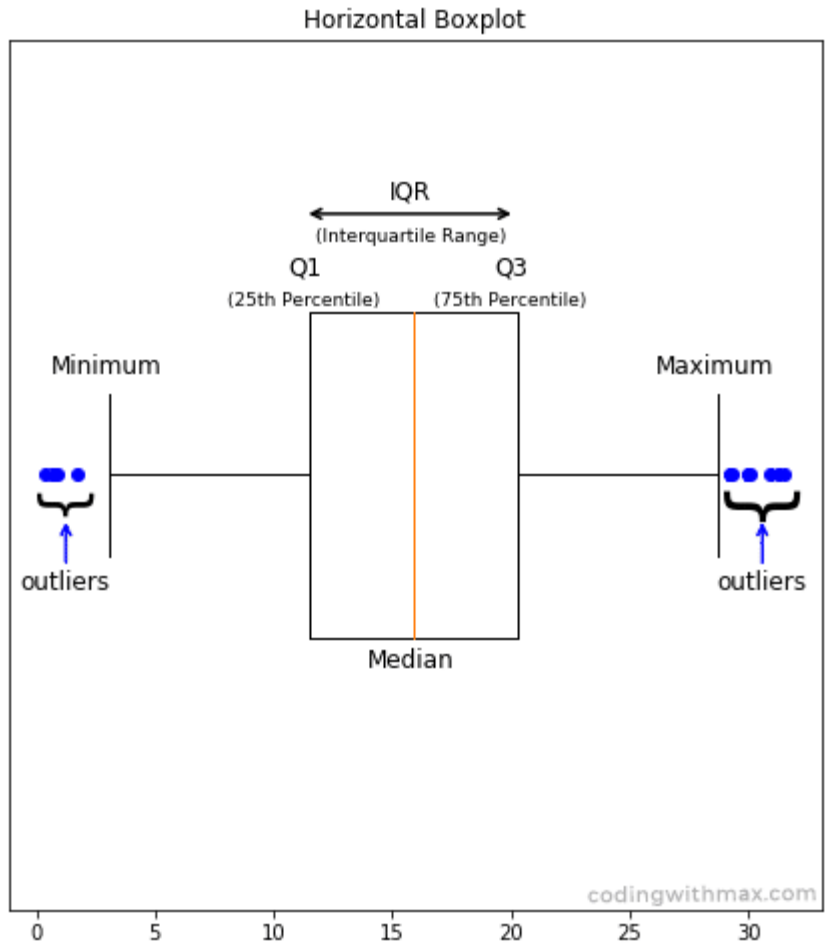


Figure 13.7. The boxplot is an idiom presenting summary statistics for the distribution of a quantitative attribute, using five derived values. These plots illustrate four kinds of distributions: normal (n), skewed (s), peaked (k), and multimodal (mm). (a) Standard box plots. (b) Vase plots, which use horizontal spatial position to show density directly. From [Wickham and Stryjewski 12, Figure 5].


Vertical Boxplot
outliers
Maximum
(75th Percentile)
Q3
Median
IQR
(Interquartile Range)
Q1
(25th Percentile)
Minimum
outliers
30
25
20
15
10
5
0


Horizontal Boxplot
IQR
(Interquartile Range)
Q1
(25th Percentile)
Q3
(75th Percentile)
Minimum
Maximum
outliers
outliers
Median
codingwithmax.com
0
5
10
15
20
25
30

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.

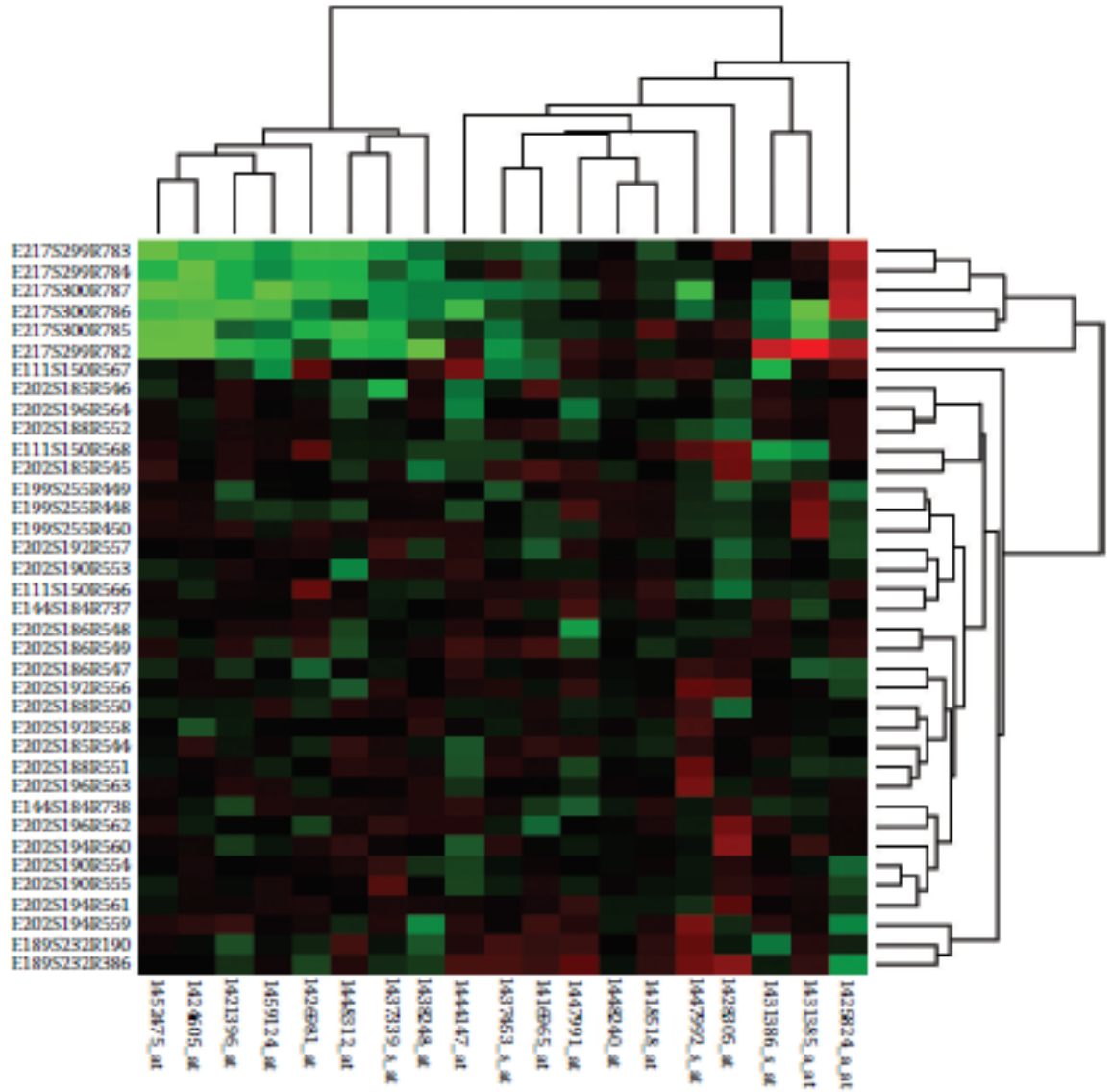


Figure 7.11. Cluster heatmap. A heatmap provides a compact summary of a quantitative value attribute with 2D matrix alignment by two key attributes and small area marks colored with a diverging colormap. The cluster heatmap includes trees drawn on the periphery showing how the matrix is ordered according to the derived data of hierarchical clusterings on its rows and columns.

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.

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.

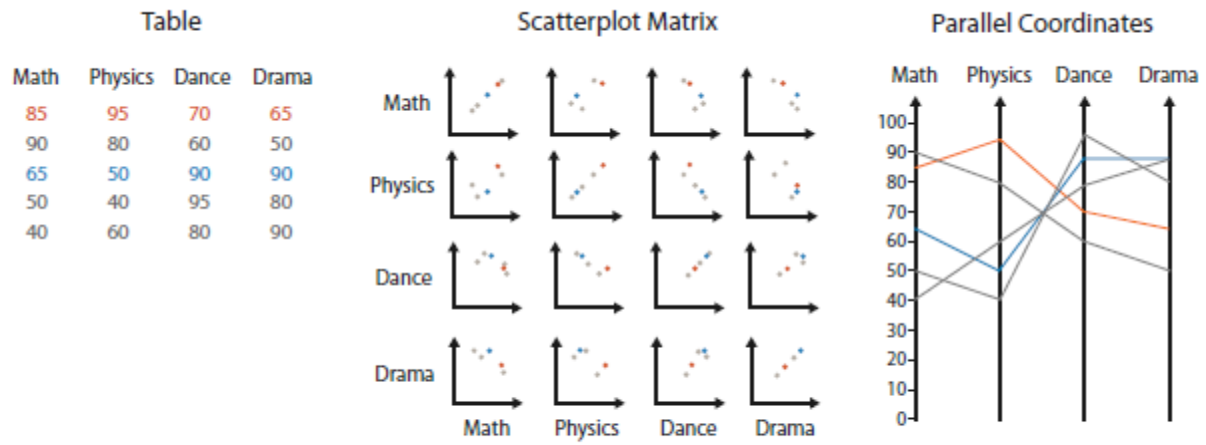


Figure 7.12. Comparison of scatterplot matrix and parallel coordinate idioms for a small data table. After [McGuffin 14].

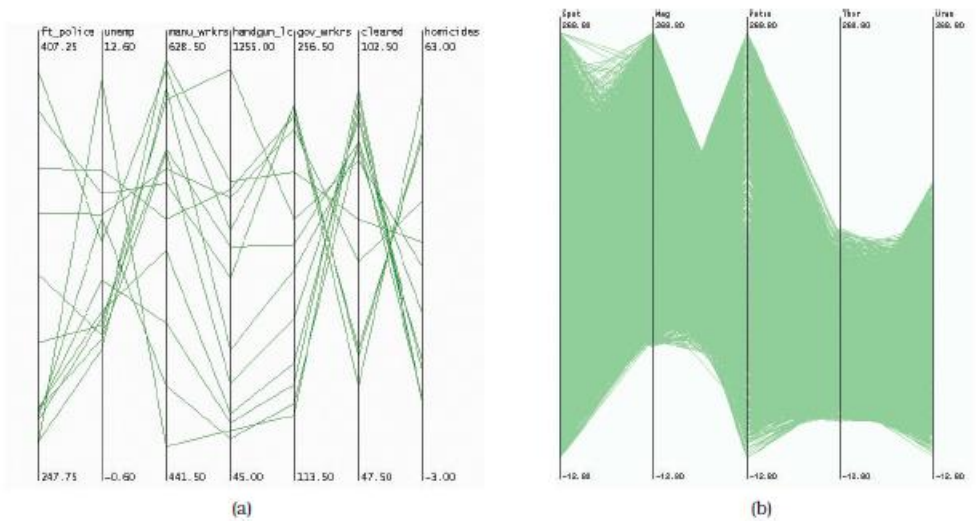


Figure 7.14. Parallel coordinates scale to dozens of attributes and hundreds of items, but not to thousands of items. (a) Effective use with 13 items and 7 attributes. (b) Ineffective use with over 16,000 items and 5 attributes. From [Fua et al. 99, Figures 1 and 2].

Why Analyze Tasks

Why Analyze Tasks

High level

Low level

Figure 3.1. *Why* people are using vis in terms of actions and targets.

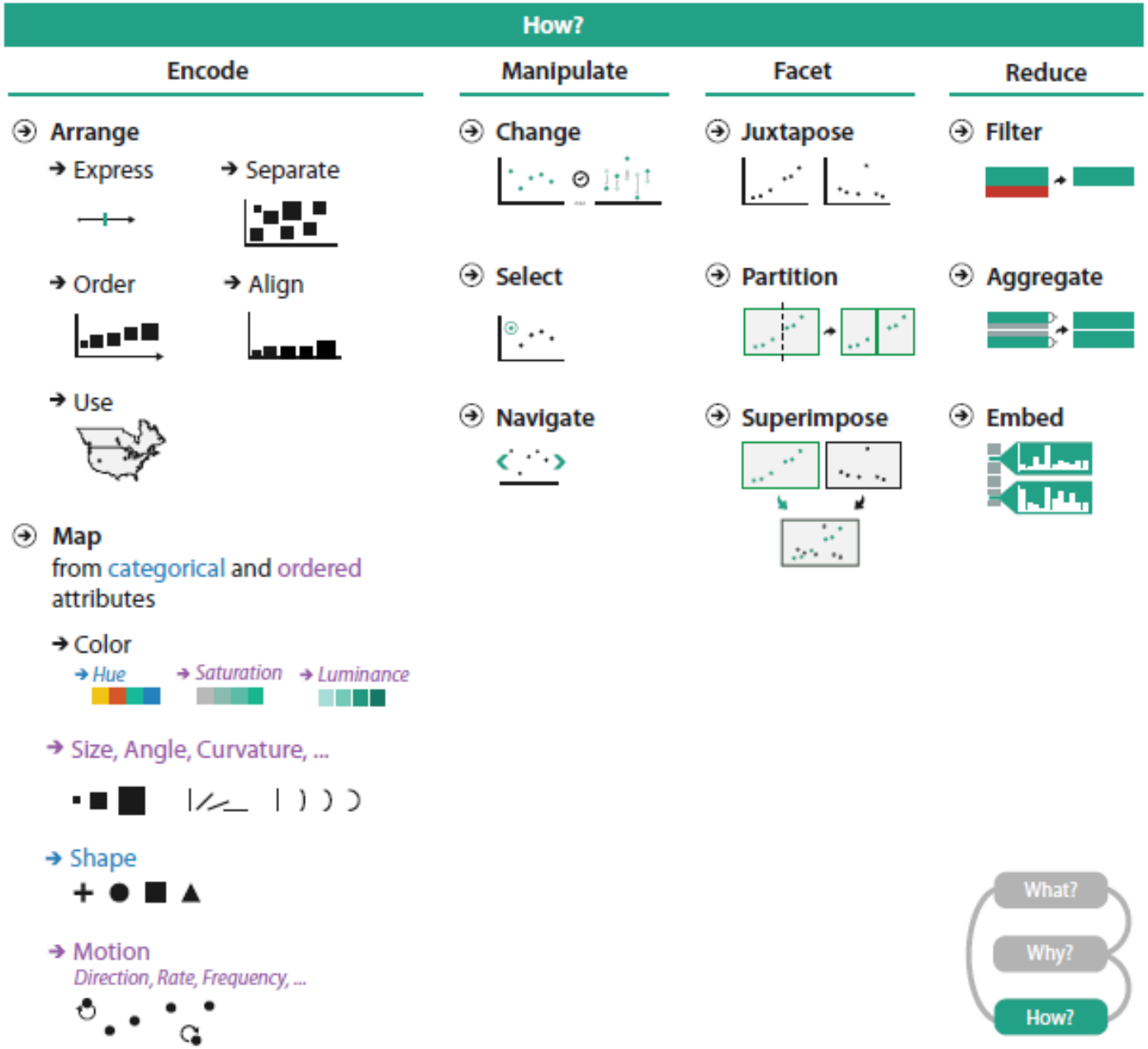
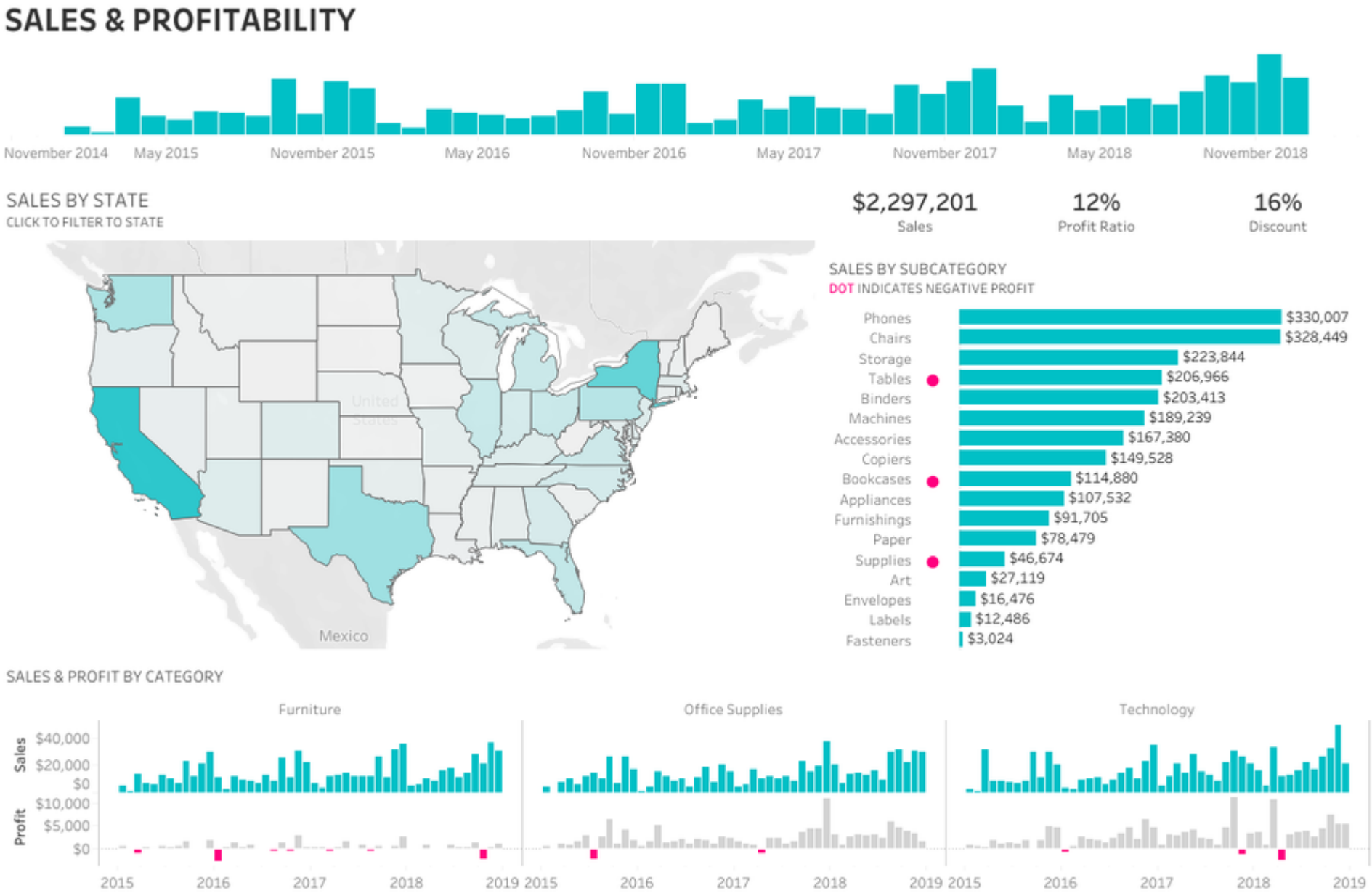


Figure 3.7. *How* to design vis idioms: encode, manipulate, facet, and reduce.



- Look at the below [Sales & Profitability dashboard] cre ated by [Ann Jackson].

Rules of Thumb

- No Unjustified 3D
  - The Power of the Plane
  - The Disparity of Depth
  - Occlusion Hides Information
  - Perspective Distortion Dangers
  - Tilted Text Isn't Legible
- No Unjustified 2D
- Eyes Beat Memory
- Resolution over Immersion
- Overview First, Zoom and Filter, Detail on Demand
- Responsiveness Is Required
- Get It Right in Black and White
- Function First, Form Next

Rules of Thumb

- No Unjustified 3D
  - The Power of the Plane
  - The Disparity of Depth
  - Occlusion Hides Information
  - Perspective Distortion Dangers
  - Tilted Text Isn't Legible
- No Unjustified 2D
- Eyes Beat Memory
- Resolution over Immersion
- Overview First, Zoom and Filter, Detail on Demand
- Responsiveness Is Required
- Get It Right in Black and White
- Function First, Form Next

Common Visualization Idioms

- To present the data you can use four basic data presentation types:
- Comparison
  - **Bar Chart**
  - **Column Chart**
- Composition
  - **Pie Chart**
  - Waterfall Chart
- Distribution
  - 3D Area Chart
  - Line Histogram
- Relationship
  - **Bubble Chart**
  - Scatter Chart
- Spatial Map

Box Plot