

Charting Essentials

Ashwin Malshe

Table 2-1. Types of variables encountered in typical data visualization scenarios.

Type of variable	Examples	Appropriate scale	Description
Quantitative/ numerical continuous	1.3, 5.7, 83, 1.5×10^{-2}	Continuous	Arbitrary numerical values. These can be integers, rational numbers, or real numbers.
Quantitative/ numerical discrete	1, 2, 3, 4	Discrete	Numbers in discrete units. These are most commonly but not necessarily integers. For example, the numbers 0.5, 1.0, 1.5 could also be treated as discrete if intermediate values cannot exist in the given dataset.
Qualitative/ categorical unordered	dog, cat, fish	Discrete	Categories without order. These are discrete and unique categories that have no inherent order. These variables are also called <i>factors</i> .
Qualitative/ categorical ordered	good, fair, poor	Discrete	Categories with order. These are discrete and unique categories with an order. For example, “fair” always lies between “good” and “poor.” These variables are also called <i>ordered factors</i> .
Date or time	Jan. 5 2018, 8:03am	Continuous or discrete	Specific days and/or times. Also generic dates, such as July 4 or Dec. 25 (without year).
Text	The quick brown fox jumps over the lazy dog.	None, or discrete	Free-form text. Can be treated as categorical if needed.

- All data visualizations map data values into quantifiable features of the resulting graphic. We refer to these features as *aesthetics*.

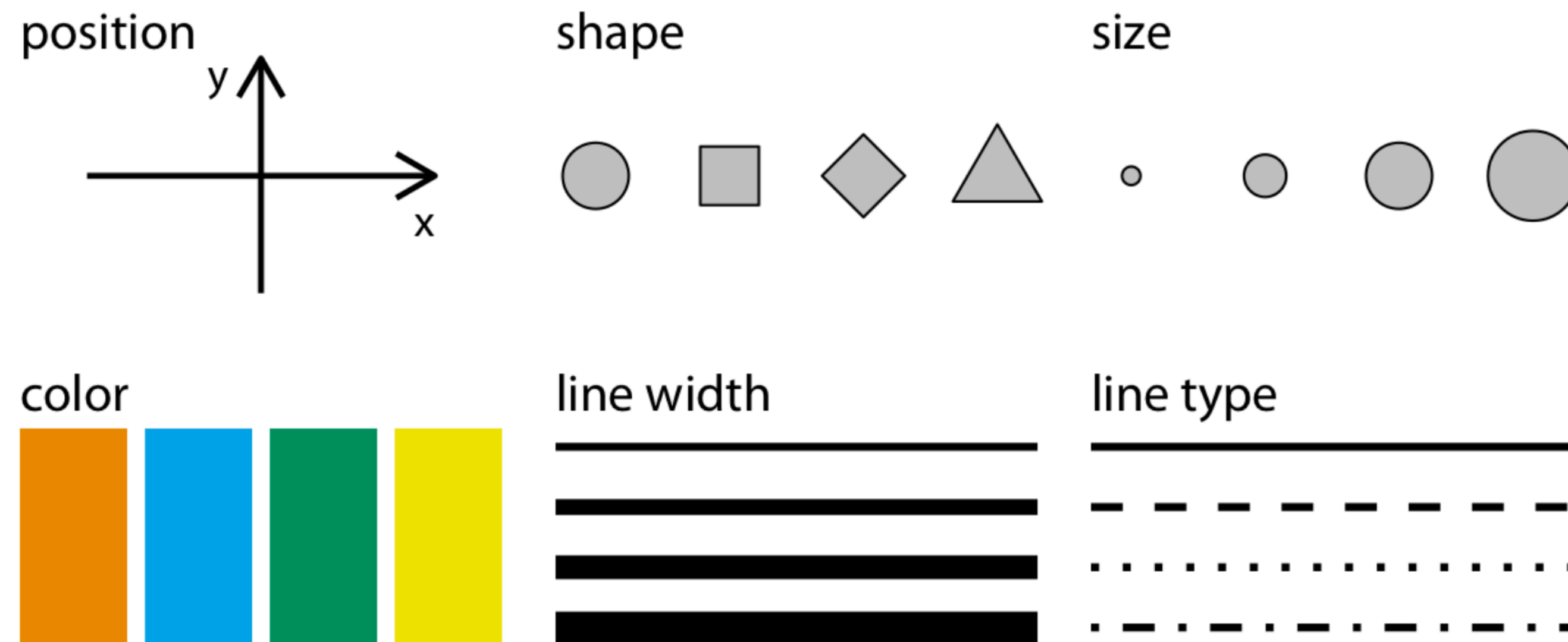
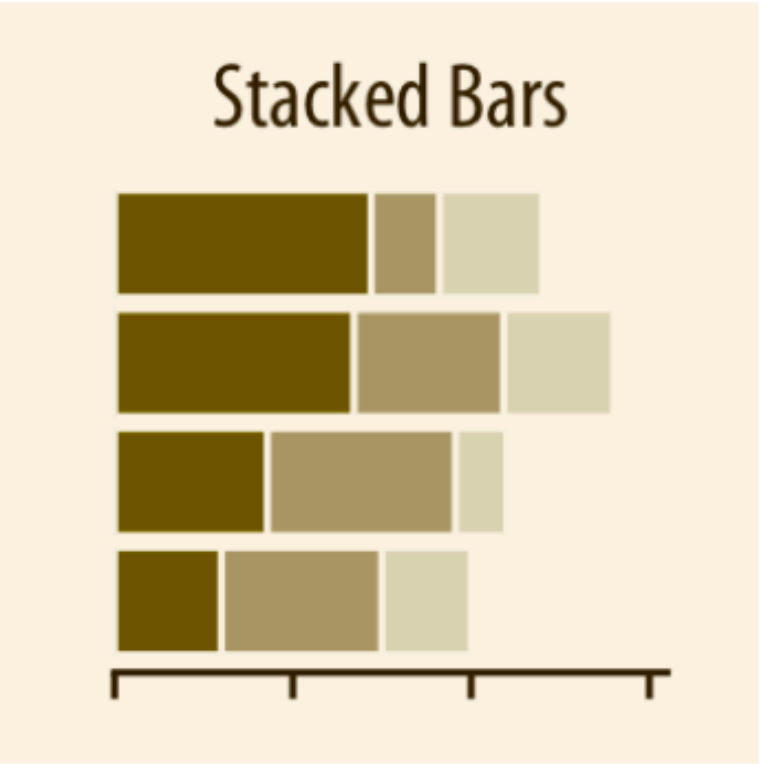
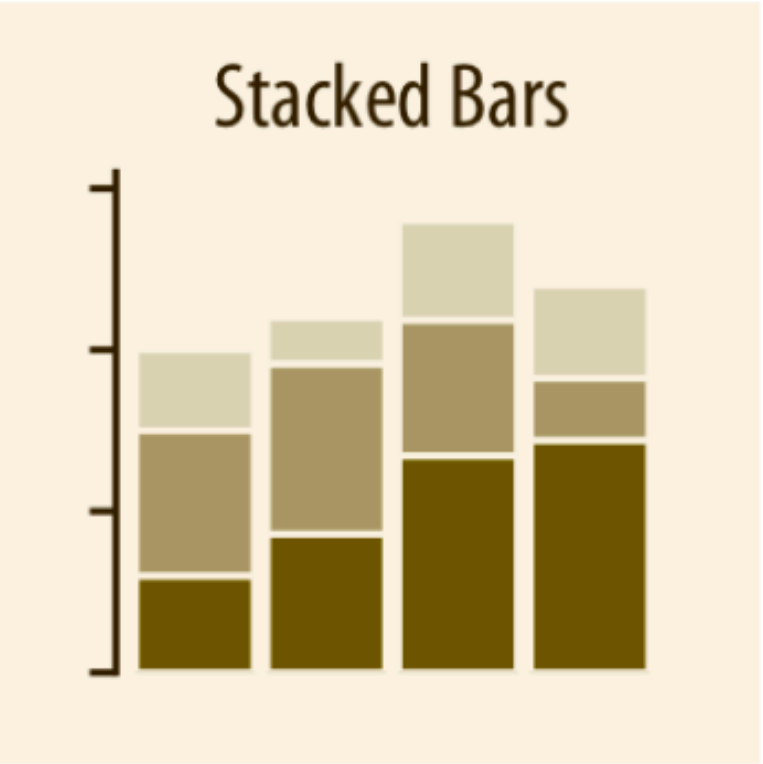
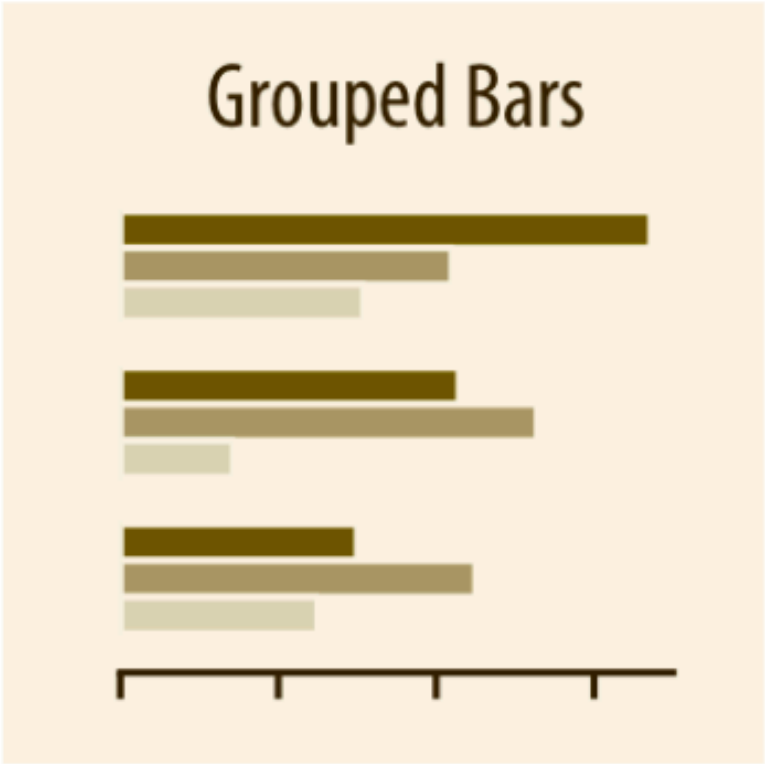
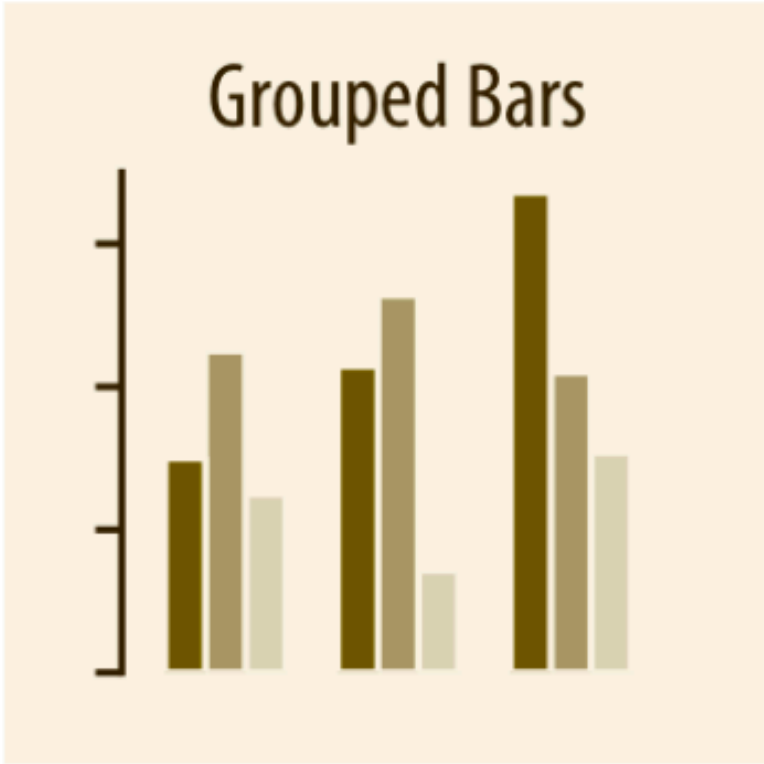
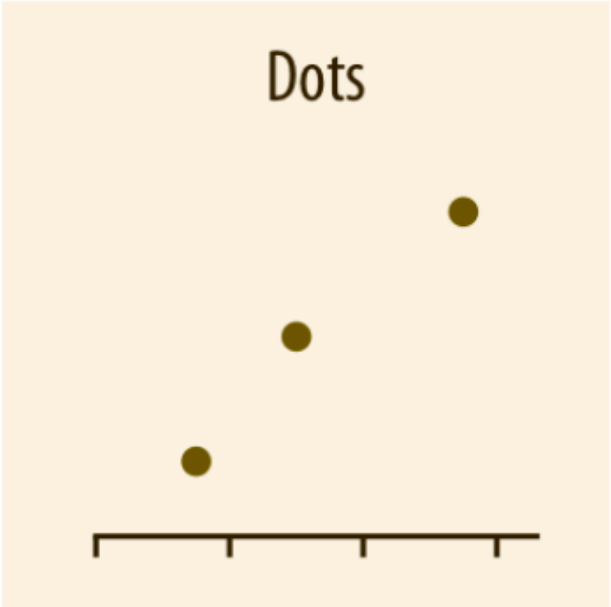
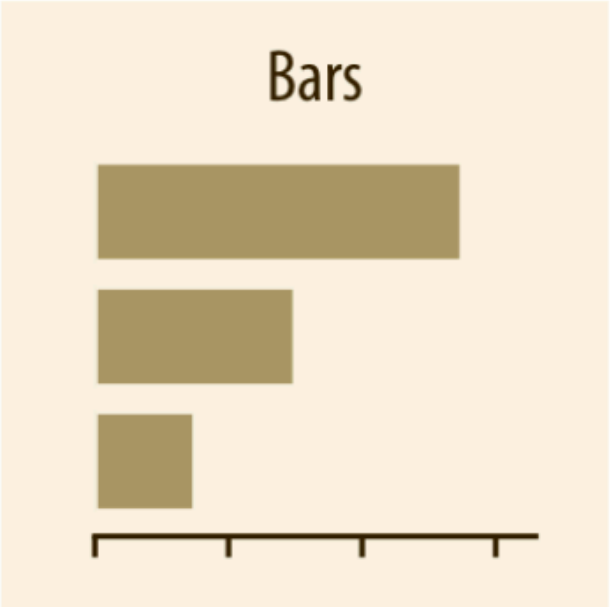
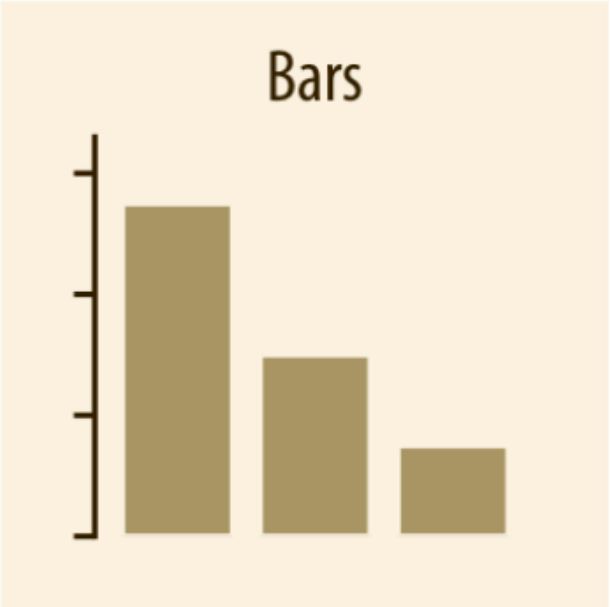


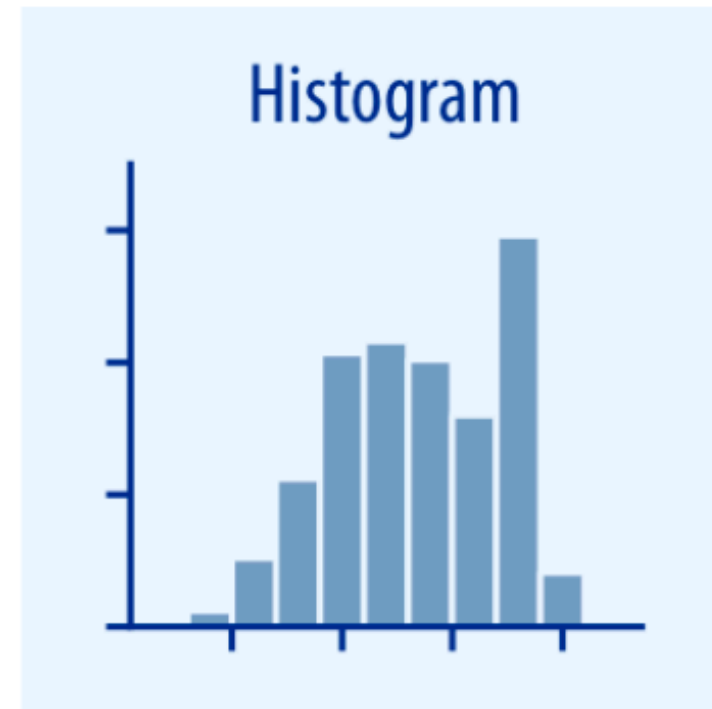
Figure 2-1. Commonly used aesthetics in data visualization: position, shape, size, color, line width, line type. Some of these aesthetics can represent both continuous and discrete data (position, size, line width, color), while others can usually only represent discrete data (shape, line type).

Chart types

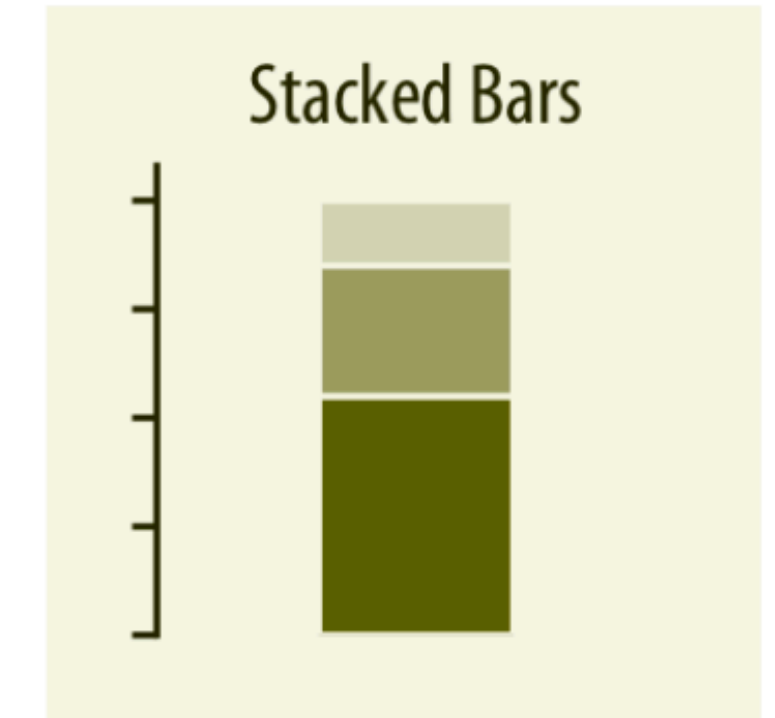
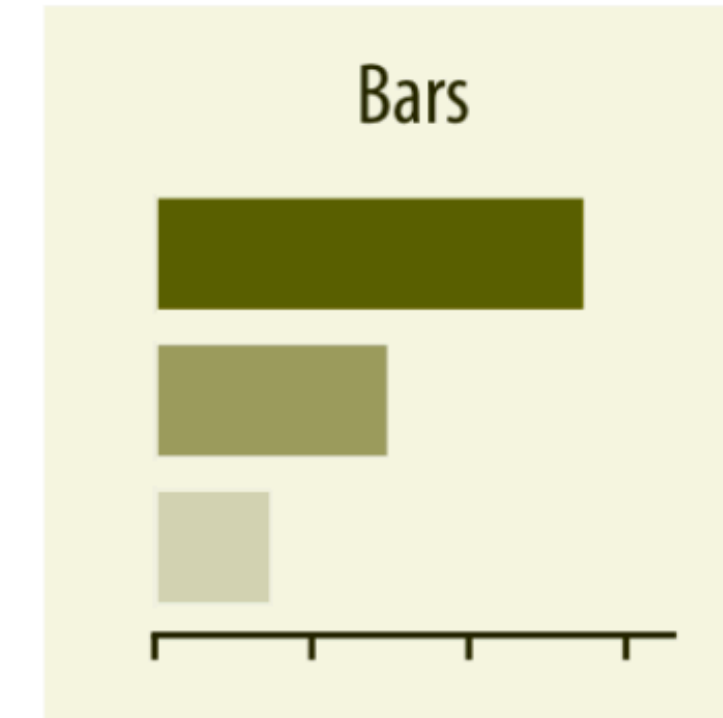
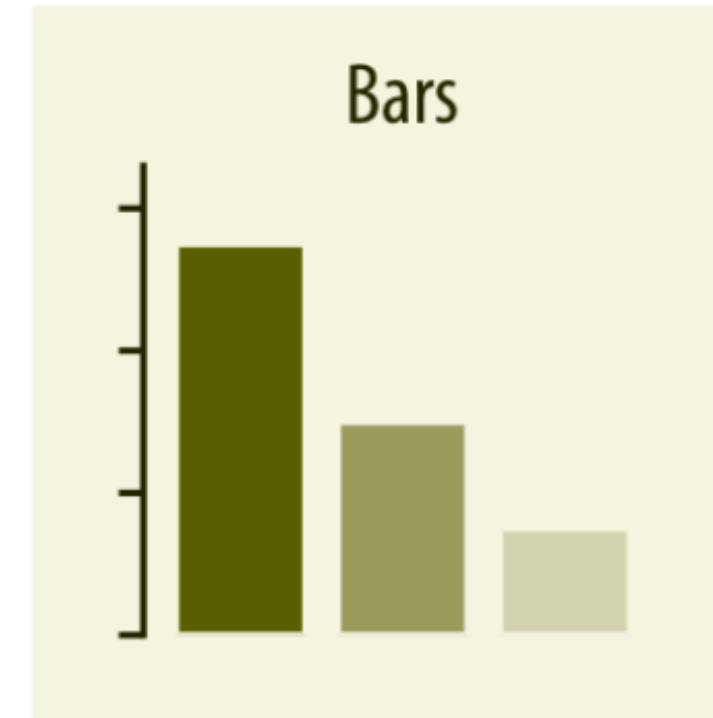
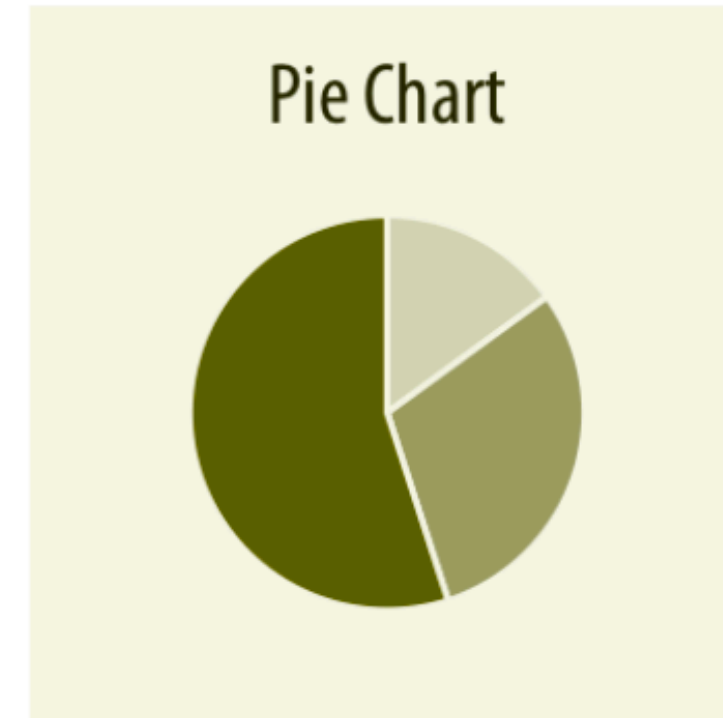
Amounts



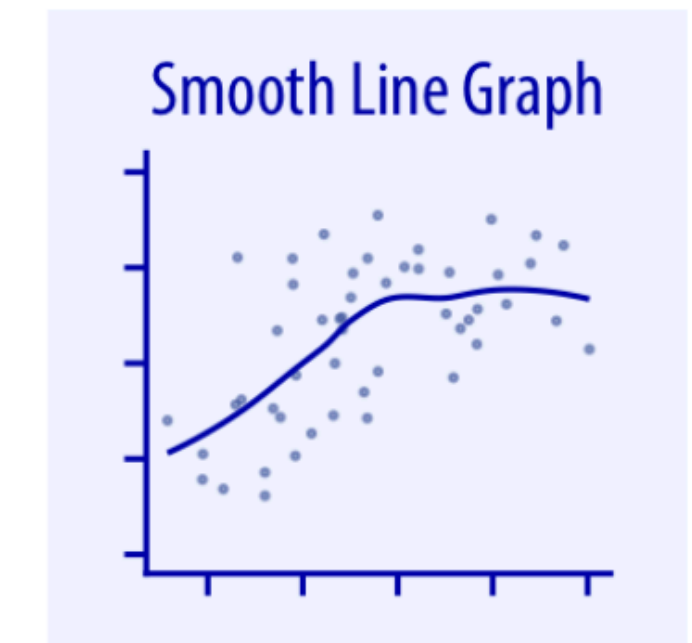
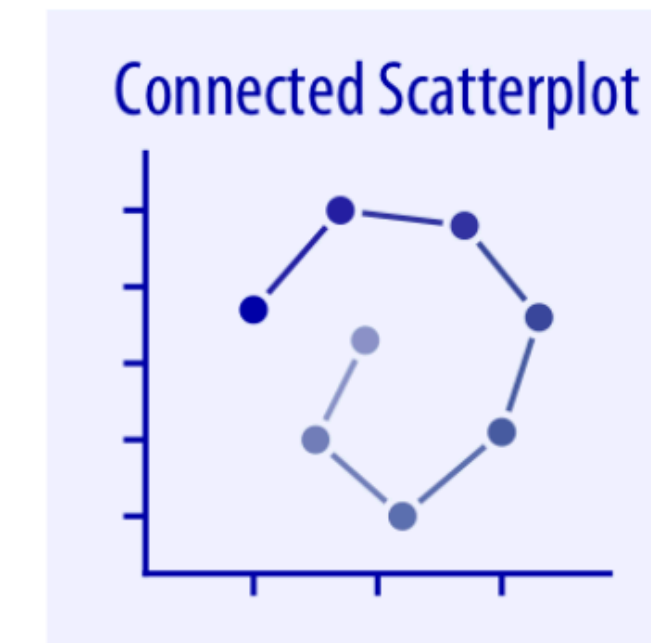
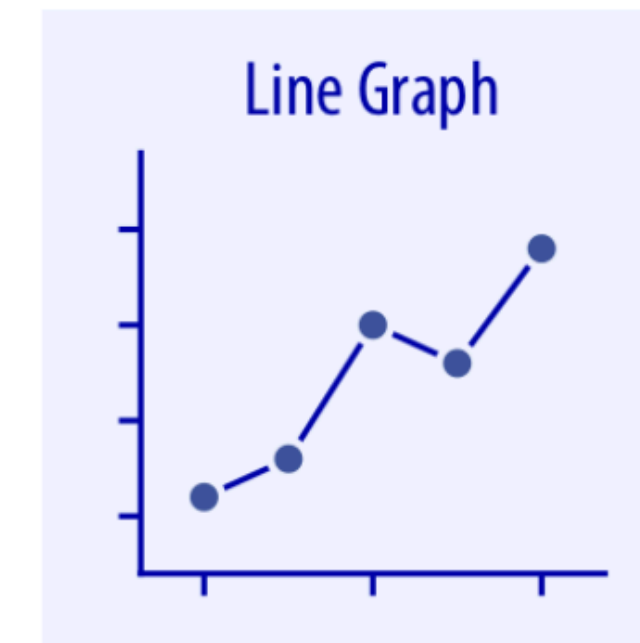
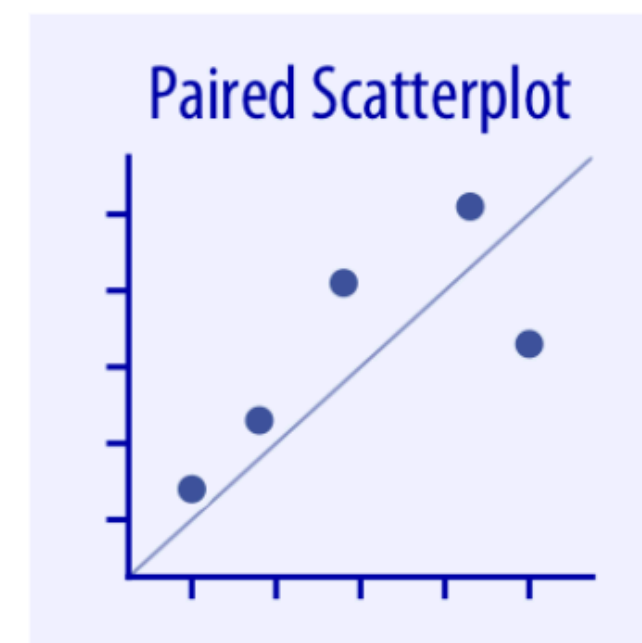
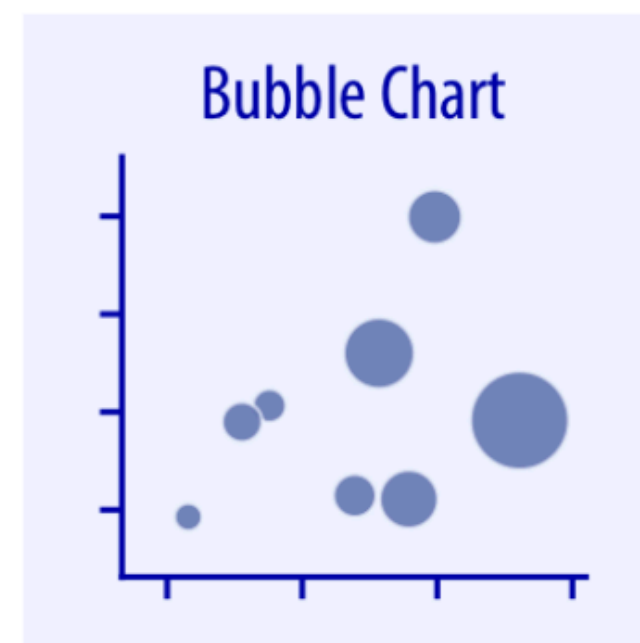
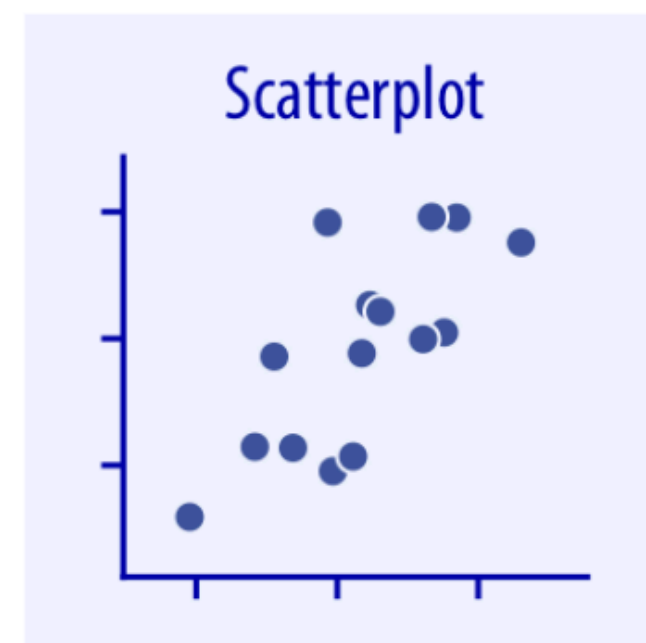
Distributions



Proportions



x–y relationships



Geospatial Data

Map



Choropleth



Color Palettes

Okabe Ito



ColorBrewer Dark2



ggplot2 hue



Discrete

ColorBrewer Blues



Heat



Viridis

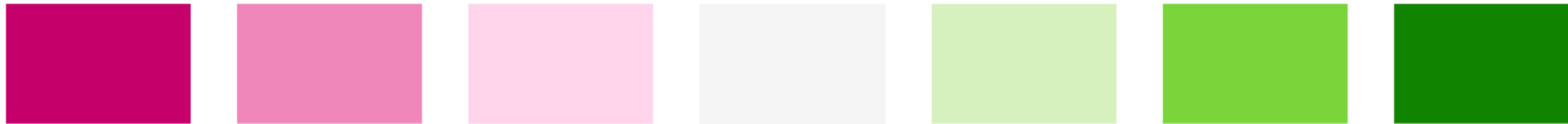


Sequential

CARTO Earth



ColorBrewer PiYG

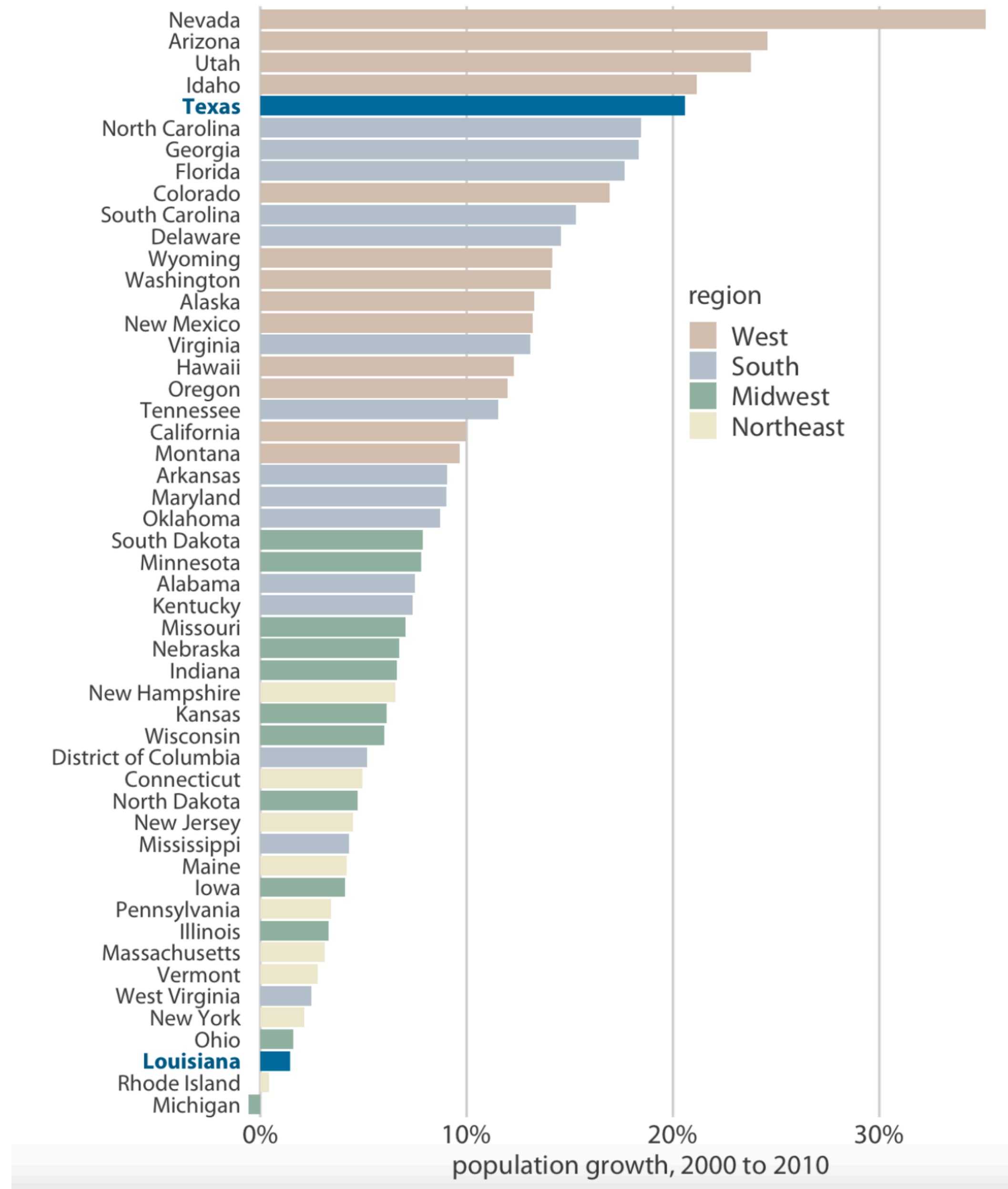


Blue-Red



Figure 4-5. Example diverging color scales. Diverging scales can be thought of as two sequential scales stitched together at a common midpoint color. Common color choices for diverging scales include brown to greenish blue, pink to yellow-green, and blue to red.

What will you change?



Relation between two variables

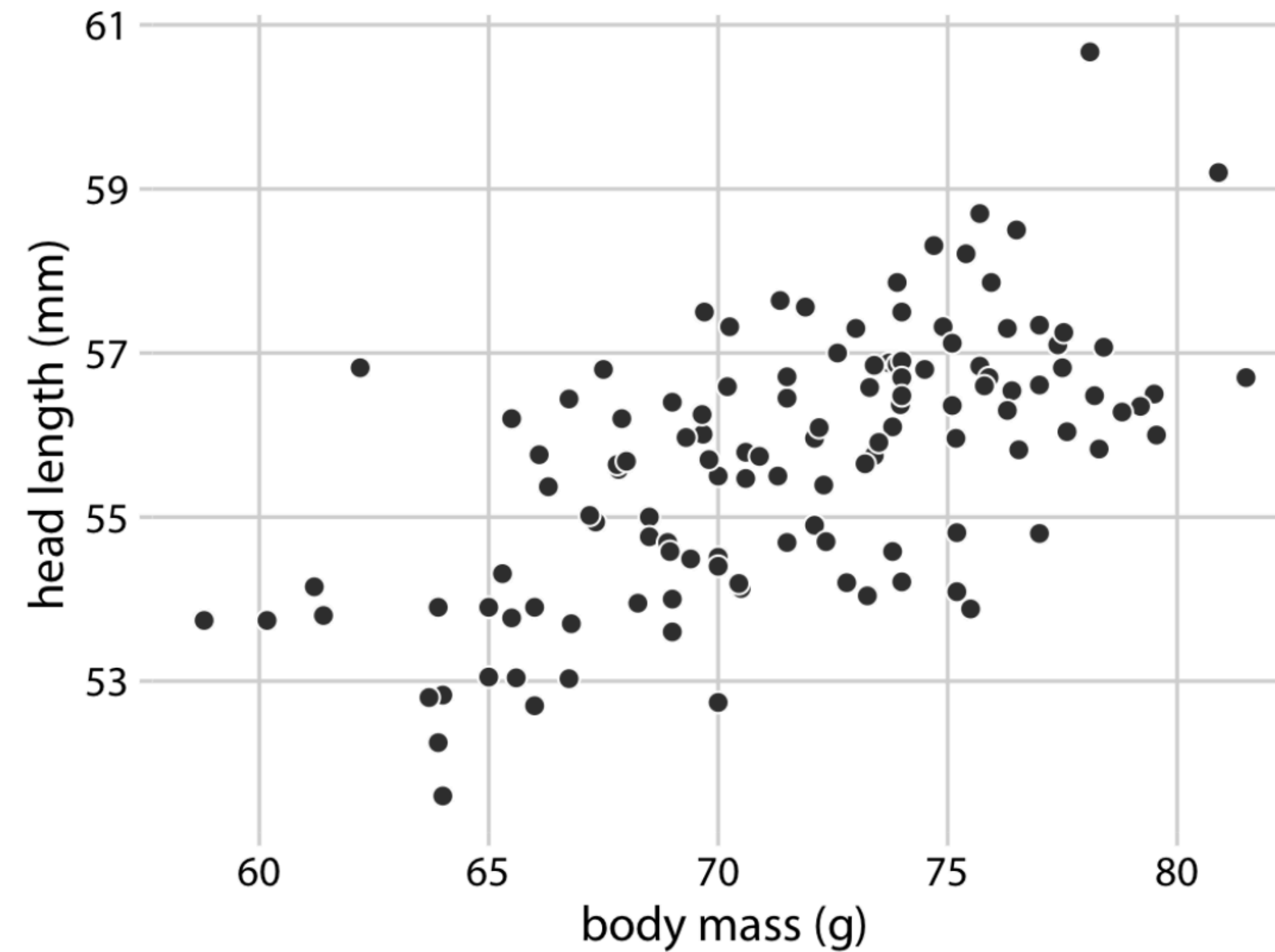


Figure 12-1. Head length (measured from the tip of the bill to the back of the head, in mm) versus body mass (in grams), for 123 blue jays. Each dot corresponds to one bird. There is a moderate tendency for heavier birds to have longer heads. Data source: Keith Tarvin, Oberlin College.

