**Chapter 2: Data Visualization**

Cues:

| Table 2.1: Visual cues and what they signify. | | |
| --- | --- | --- |
| **Visual Cue** | **Variable Type** | **Question** |
| Position | numerical | where in relation to other things? |
| Length | numerical | how big (in one dimension)? |
| Angle | numerical | how wide? parallel to something else? |
| Direction | numerical | at what slope? in a time series, going up or down? |
| Shape | categorical | belonging to which group? |
| Area | numerical | how big (in two dimensions)? |
| Volume | numerical | how big (in three dimensions)? |
| Shade | either | to what extent? how severely? |
| Color | either | to what extent? how severely? |

\*Human’s ability to sense differences based on these cues decreases as you go down the list.

**Coordinate systems:**

**Cartesian:** x-y plot

**Polar:** Like a pie chart

**Geographic:** Locations on Earth’s curved surface as a 2-D plane/grid

**Scales:**

‘The translation of values into visual cues’

**Numeric:** linear, log, or % scales. Self-explanatory.

**Categorical:** grouping that may have no order (democrat vs. republican, grassland versus forest), or that may be ‘ordinal’ (current versus former smoker, previous versus post drought).

**Time:** A special numeric quantity, with specific properties rendering its differentiation from other numeric scales.

**Small multiple and layers:**

‘While 3D images are sometimes useful, oftentimes they are more confusing than helpful’. So we can use some ways to get multi-dimensional information in a 2-D plane.

**Small multiples:** facts. Small graphic embedded in a larger graphic, each ‘facet’ is a ‘multiple’ of the same plot, put a discrete variable is being changed.

**Layers:** Adding a new layer on an existing graphic.

**Animation:** One way to show time, can’t do this in print, best for a presentation.

**Color**

-One of the most misperceived and misused cues.

- ~8% of population are colorblind. Mostly men, many of which don’t know they are colorblind.

-RColorBrewer is a creation of Cynthia brewer and contains color palettes that are color blind safe. The package specifically deals with:

**Sequential:** Ordered values in one direction

**Diverging:** Ordered values in two directions

**Qualitative:** Simply need to differentiate between groups. No ordering.

‘Shading’ of color is what you do when you want to show continuous range of values using colors, like you would with a map of a continuous variable (e.g., NPP)

‘If an audience is reading your slides, they are not listening to you’

Chapter 3: A grammar for graphics

-Focus on using ggplot2

**Aesthetic:** a mapping between a variable and the visual cues that present its values, for example designating which variable map onto which axes in a bivariate plot.

**Glyph:** a basic graphical element that represents ‘one case’, such as a point/symbol.

**Facets:** Multiple side-by-side plots showing the same relationships but for different categories.

Table 3.3: Table of canonical data graphics and their corresponding **ggplot2** commands. Note that the mosaic plot function is not part of the **ggplot2** package.

|  |  |  |
| --- | --- | --- |
| numeric | histogram, density | geom\_histogram(), geom\_density() |
|  | categorical | stacked bar | geom\_bar() |
| numeric | numeric | scatter | geom\_point() |
| numeric | categorical | box | geom\_boxplot() |
| categorical | categorical | mosaic | geom\_mosaic() |