Hampton
February 1, 2018
Word account: 634

Statistical Graphs

Statistical graphs are used across many disciplines to summarize a huge amount of qualitative and quantitative data in a visual display. In order to achieve the aim of a given graph, essential data should be presented in a way that enables the viewer to achieve a rapid and accurate interpretation of that graph. A better understanding of perceptual mechanisms underlying visual processing is necessary for the designers to provide graphs that communicate essential information effectively. This essay covers different aspects of graphical displays as they relate to perceptual processing mechanisms.

Perhaps the most common types of statistical representation are either bar or pie charts. Pie charts display quantitative values by the slice area and the angle formed between the slice radiating around the center of the pie. Bar charts on the other hand display values in 2D bar diagrams, with categories on one axis and one variable on the other axis. The bar height or length is always proportional to the values represented. The human eye is excellent when it comes to comparing either locations or differences in height or length. But when it comes to 2D angles and areas they fail because we don’t estimate angles accurately. Research has found that when it comes to angle estimation we tend to either overestimate obtuse angles or underestimate acute angles (Nundy, Lotto, Coppola, Shimpi & Purves, 2000). In light of that, when it comes to making comparisons, bar charts are more appropriate than pie charts, which are more appropriate when estimating proportions (Simkin & Hastie, 1987).

Considering a scatter plot as a mean to visualize data, different kinds of symbols can be used to represent the data points. Symbols should be selected in a way that allow them to pop out of the display. One way to do that is by plotting dark circles to represent data points. The contrast between these dots and the white background makes the recognition and comprehension of the data easier. However, with a large amount of data points, other symbols would be more appropriate as dark circles could distort the display of information as they clutter up, creating visual noise. In such a case, using an ‘x’ sign is more appropriate because it would enable the viewers to detect the exact location of each point even for the points that are really close to each other.

The importance of the symbol type can be extended to cases when presenting multiple categories on a scatter plot, such as when presenting data of different ethnic groups. According to the gestalt law of similarity, similar items would be misperceived as a single group. To facilitate comparison, a presenter could include different symbols (such as circles, triangles, and diamonds) each representing a different category. The texture of these symbols should be distinct (Cleveland & Cleveland, 1985), and that can achieved by those having strong boundaries (e.g. □ and ○) rather than those having weak boundaries (e.g + and ×, ◊ and △) in order to ensure the accurate interpretation of the displayed data. Another technique in crafting an effective display is using colored data points. As colors are pre-attentive (they do not require focused attention), they are easy to pinpoint. Basic colors maximize the ability to perceive and remember figures effectively (Ware, 2012). However, designers should consider the viewers with color blindness when selecting colors. One way to do so is by printing the colored graph in black and white and checking if the colors are reflected in different shades of gray (Pasta, 2006).

Statistical data analysis and science communication rely heavily on graphs. To achieve the aim of a given graph, statistical graphs should be designed in a way that enables the viewer to process data within a short period of time and without confusion. An awareness of perceptual mechanisms not only allows the designer to create more effective graphs, but also keeps the viewer from being misled by poorly-designed ones.

References:

Cleveland, W. S., & Cleveland, W. S. (1985). *The elements of graphing data (Vol. 2).* Monterey, CA: Wadsworth Advanced Books and Software.

Nundy, S., Lotto, B., Coppola, D., Shimpi, A., & Purves, D. (2000). Why are angles misperceived?. *Proceedings of the National Academy of Sciences, 97*(10), 5592-5597.

Pasta, D. (2006). Using Statistical Graphics to Understand Your Data (Not Just to Present Results). In: *Proceedings of the Thirty-first Annual SAS® Users Group International Conference*, Cary, NC: SAS Institute Inc., [online] San Francisco, California, pp.191-31. Available at: http://www2.sas.com/proceedings/sugi31/191-31.pdf

Simkin, D., & Hastie, R. (1987). An information-processing analysis of graph perception. *Journal of the American Statistical Association, 82*(398), 454-465.

Ware, C. (2012). *Information visualization: perception for design*. Elsevier.