Article from
Predictive Analytics & Futurism
December 2018
Issue 19

The Psychology of Visual Data

By Dorothy Andrews

Accessing the information content of data, large or small, and making it accessible and meaningful is an art form, according to Boehnert.[1] Visualization artists attempt to reduce the abundance and abstract essence of big data to actionable, neutral and objective information. This goal is not easily achieved using data reduction techniques such as line graphs, and bar, bubble and donut charts, all which are subject to axes and scale and manipulation in 2D and 3D. Bergstrom & West[2] point out how graph axes can be used to magnify numerically insignificant heights and lengths of bars in bar charts by excluding the zero origin. By eliminating the zero origin, designers can zoom in on a section of a bar graph to create a visual difference the mind comprehends as more significant than it numerically is. Including zero allows the eye to see the actual heights of all bars and accurately perceive numerical differences. This is less of a concern in line graphs. However, a technique that amplifies insignificant differences in line graphs is changing the scale of graph axes from one where the increments are constant to one where the increments are logarithmic. In common logarithmic scales, each increment along one axis creates a change in the other in multiples of 10, as in measurements of earthquakes. A measure of two is 10 times worse than a measure of one, but a measure of three is 100 times worse than a measure of one and 10 times worse than a measure of two. Mixing such changing magnitudes with non-logarithmic units along the same axis should be avoided to prevent misleading the viewer into seeing a difference that does not materially exist, according to the authors.

Bergstrom & West[3] provide a heuristic to guide the construction of data graphics. "The Principle of Proportional Ink" simply states: "when a shaded region is used to represent a numerical value, the area of that shaded region should be directly proportional to the corresponding value (p.1)." In other words, the area of the ink used to represent a number should be a function of the magnitude of the number and the same function should be used consistently to represent all numbers. The easier the measurement lends itself to a mathematical function, the easier it is to identify and prevent violations of the principle proportional ink.

Filled line charts (line graphs with shading below the line) are akin to bar charts with zero separation between the bars. The authors contend filled line charts should never exclude the zero origin to prevent violation of the principle of proportional ink. Bubble charts can leverage size, color, and horizontal and vertical coordinates to "encode four different attributes (p.5)" of each data element, with some attributes more tractable than others in the design of the chart. Donut bar charts are challenged with conforming to the principle of proportional ink, as well as graphs constructed using changing denominators. The authors point out how easy it is to confound results when percentages rather than absolute numbers are used as the metric in determining the size of graph objects, as illustrated in the graph of causes of death by age groups. While there are more deaths in the 65+ group, the graph suggests (in percentage terms) more 1- to 4-year-olds die of accidental causes. The larger denominator of deaths in the 65+ group is causing this effect. One of these elements could be eliminated from the graph to eliminate the confusion.

Dimensionality is another technique subject to frequent violations of the principle of proportional ink and manipulating perspective to obscure relative sizes of chart objects. Despite a potential ink violation, adding a third dimension can be helpful in visualizing data outliers (data points that are quite distal from the others) that appear to cluster with other data elements in a 2D view, where all the data lie in the plane of a page. A cube view of data adds a height dimension. It becomes easy to perceive a single point that hovers far above a cluster of points as an outlier in 3D than it is in 2D. A purely aerial view can obscure outliers.

Data visualization tools are but one type of big data reduction technique just like the calculation of the mean of a large set of numbers is a data reduction technique. Animation is another and can heighten the experience of visualizing data beyond adding a third dimension (Koblin[4], Lizama[5], Thorp[6], Wright[7]). The "Stop and Frisk" video (Lizama[5]) produced by the Morris Justice Project (MJP) evokes visually as well as emotionally, as it illuminates the geography and frequency of stops and frisks by police in the name of community safety, incurring great costs with little benefit. The video reflects the three elements Wright[7] says are essential to data visualizations: 1) data mining, 2) programming and 3) design. The MPJ collected stop and frisk data on policing in an NYC area neighborhood, programmed the geography of the stops using a flashing light effect, visually simulating a rear-view mirror view of the flashing lights on a police car light bar. Thorp[6] provides several examples of how animating data can reduce its complexity to 2D and 3D to emphasize spatial differences

among data points, allowing for the comprehension of the data from several different "viewing angles." These perspectives are important in differentiating true data clusters from disparate data outliers.

Mack[8] discusses how inattention can distract from perception by interfering with how the brain interprets what the eyes see. The demonstration by Daniel Simon and Christopher Chabris discussed by Mack[8] is a perfect example of this distraction. Viewers are asked to count the number of times a ball is passed among a group of individuals dressed in white or black shorts and shirts. The attention to the task prevents the perception of a man in a gorilla suit weaving among them and exhibiting gorilla-like behaviors. The failure on the part of the observers to see the gorilla results in what the author calls "inattentional blindness (p.1269)." The implication is that attention is necessary in order for perception to occur. Understanding inattentional blindness is important in making sure animated data visualizations do not coerce viewers into seeing less than what is in an image as well as more than what is visually present. The author discusses how the mind will fill in any "blind spots" present in an image. Designers should be cognizant of the potential for the mind to fill in blind spots in such a way that detracts from the intended message of the visualization.

Paying attention to the psychology of visual data can help you use visualizations to inform; and not to mislead. ■

Dorothy L. Andrews, ASA, MAAA, CSPA (Certified Specialist in Predictive Analytics), is the chief behavioral data scientist at Insurance Strategies Consulting LLC. She can be reached at *dandrews@insurance-strat.com.*

ENDNOTES

1 Boehnert, J. (2016). Data Visualisation Does Political Things. Proceedings of DRS 2016, Design Research Society 50th Anniversary Conference. Brighton, UK, 27-30 June 2016.

2 Bergstrom, C., West, Jevin. (2017a). Visualization: Misleading Axes on Graphs. Retrieved from *https://callingbullshit.org/tools/tools_misleading_axes.html*

3 Bergstrom, C., West, Jevin. (2017b). Visualization: The Principle of Proportional Ink. Retrieved from *https://callingbullshit.org/tools/tools_proportional_ink.html*

4 Koblin, A. (n.d.). Flight Patterns. Retrieved from *http://www.aaronkoblin.com/work/flightpatterns/*

5 Lizama, S. (2013, March 19). This Is Our Home: Scars of Stop and Frisk. *Morris Justice: A Public Science Project* [Video File]. Retrieved from *https://www.youtube.com/watch?v=qLWWa2De2b4*

6 Thorp, J. (2013). Numbers that paint the picture [Online video]. National Geographic. Retrieved from *http://library.fora.tv/2013/06/12/Jer_Thorp_Numbers_That_Paint_the_Picture*

7 Wright, L. (2013, December 2). Research Report: Data visualization [Blog]. Retrieved from *http://blog.ocad.ca/wordpress/digf6l01-fw201302-01/files/2013/11/Laura-Wright-Research-Report-final.pdf*

8 Mack, A. (2011). The Image: Seeing More and Seeing Less Than is There. *Social Research*. 78(4). 1263-1274.