The Stacked-Stacked Bar Graph: A new twist on an old visualization

Steve Szigeti, Joana Patrasc, David Schnitman, Sara Diamond
OCAD University, Toronto, Canada

Abstract

Stacked-stacked bar graph is the working title of a visualization which builds on the strengths of a stacked bar graph. Where a stacked bar graph allows for a visual comparison of the parts to the whole, our proposed visualization further divides the parts to allow for additional points of comparison. A pilot study compared our visualization with a stacked bar graph using an identical data set and found that while users expressed some uncertainty regarding how to read the stacked-stacked bar graph, they indicated a preference for this visualization. The study also highlighted the need for additional qualitative and quantitative analysis.

Key Words: Visualization techniques, visual analytics

Index Terms: H.5 [Information interfaces and presentation], H.5.2 [User interfaces: prototyping]

1 Background

Bar graphs, which provide a visual representation of categorical data, may be found in publications as early as the 14th century, although it is generally acknowledged that William Playfair proposed this visualization in 1786 [4]. Bar graphs provide a visual representation of categorical data, and subdividing each of the bars allows for a visual comparison of the parts of a category which comprise a larger category. While Cleveland [2] argued that bar graphs are more effective than stacked bar graphs since the alignment of the baseline was a key factor in judging comparisons, the effectiveness of bar graphs and stacked bar graphs is arguably task dependant [3]. In other words, stacked bar graphs are better for part-to-whole comparisons, while bar graphs are better for part-to-part comparisons.

This research emerges from collaboration between OCAD University's Visual Analytics Lab and BBM Analytics, a Canadian company which tracks and analyses media consumption.

2 Design Process and Results

The stacked-stacked bar graph emerged from a series of iterative design sessions, where eight graphic designers considered a data set and proposed visualizations. The data set was compiled by BBM Canada, consisting of detailed demographic information on radio listeners and the media consumption habits of market research participants collected on April 1, 2013 for the Toronto, Canada area. The industrial partner was involved in the design process, and upon viewing the initial designs they stressed the importance of age and gender demographics in relation to station listenership (these represented three variables in the data).

Inspired by our industry partner's current visualization methods, which include bar graphs (but not stacked bar graphs) and tables (containing numerical data), we designed a visualization that added demographic information (age and gender) to the traditional bar graph design.

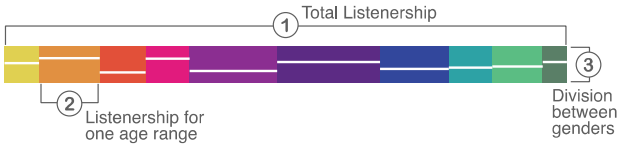

Figure 1: Example of stacked-stacked bar graph with legend

Figure 1 shows a visualization of three characteristics of data related to a radio station; (i) total audience, (ii) total audience divided by age range, and (iii) the gender per age range. In this example, we incorporated a stacked bar graph to divide total listenership into relevant age categories. Then we bisected the vertical area of the bar graph in order to display a third variable of data; the binary information of the station's audience by gender within each age range. This is illustrated using a white line, creating disparate areas above and below the line. In this example, female listenership is represented on top of the line and male listenership is represented below the line.

3 Pilot Study

We conducted a pilot study to test the efficiency, clarity and response to our design. Participants were tasked with answering a series of questions by consulting visual representations of the data. The visualizations, in the form of jpeg images, appeared on a computer screen. The representations alternated between (i) a collection of bar graphs and (ii) the stacked-stacked bar graph. The collection of bar graphs represents how similar data is currently viewed by clients of BBM Analytics, whereas the stacked-stacked bar graph represented an alternative.

The 17 participants (eight women and nine men) were all students with varying levels of exposure to information visualizations. The ten tasks were divided equally between the two representations of data and participants were asked to respond to questions regarding the total listenership of a radio station, as well as age demographics and gender of a radio station audience. For example, participants were asked to rank radio stations based on total audience, to find the largest age group for a particular station and to find the predominant gender from within a specific age category. Participants were asked identical questions for each type of visualization and the order of questions was alternated between participants to address potential order bias.

Answers were provided not verbally, but via a participant's interaction with the visualizations. This approach (using a mouse or track pad to select an answer) was intended to allow for the measurement of response times, via a screen recording. Finally the efficacy of our visualization, in terms of user sentiment and engagement, was captured through an online survey administered after the tasks were completed.

Regarding the usefulness of the proposed visualization, one participant noted that where, "data is related to each other and where decisions need to be made based on multiple [data] sources… it would be easier to consult one stacked-stacked bar graph than 3 individual ones." However another participant noted, the advantage of the stacked-stacked bar graph has limits, particularly where data, "involves a limited number of subsets." Dividing the areas within the stacked bar charts into two was potentially useful, but dividing the space into three or more areas would make comparisons across bars very difficult. For example, "if two groups have similar gender proportions but are not positioned right against each other," noted a participant, "it can be difficult to compare the groups to discern their ranking." As another participant noted, it may be necessary for, "some delineation on the third variable if it's right around the 50/50 split".

Participants were asked to rate the use of the stacked-stacked bar graph to complete the tasks. Using the System Usability Scale (SUS) for measurement, the average score of the 17 respondents was 80.9. A SUS score above 68 is considered above average [1]. It is worth noting that three participants gave the system a score below this benchmark (with scores of 47.5, 52.5 and 55). However an average score of 80.9, supported by positive commentary from participants, suggests that the stacked-stacked bar graph has potential value for the analysis of data.

4 Discussion and Future Work

We argue that the strength of a stacked-stacked bar graph lies primarily in the ability to increase the density of information displayed within space. Instead of asking users to search for demographic information on different screens, or having the user navigate through the interface to display one variable of information at a time, the stacked-stacked bar graph represents a possible improvement by presenting all the relevant data in one image.

The pilot study provided an opportunity to share the visualization with potential users to collect feedback and to test some of the assumptions we had regarding this visual approach. However, the study was problematic in not representing an actual use case. If, as we believe, one of the strengths of the stacked-stacked bar graph is the ability to display an additional variable within a data set (which currently requires a separate visualization), then an evaluative study would ideally allow the user to step through a similar process of generating visualizations from the data. The pilot study did clarify the importance of both (i) the use of an interactive tool similar to the application our industrial partner currently uses and (ii) qualitative feedback. .

The positive impression of the participants suggests that the stacked-stacked bar graph represents a more visually efficient and engaging comparison of data over existing methods. Current tools employed by our industrial partner require the user to toggle back and forth between visualizations or to generate multiple bar graphs, where a single visualization may provide the insight required by the user from the data. We welcome input from the research community regarding this visualization and different methodologies that may help us better understand the strengths, weaknesses and implications of this approach. Further study will also identify types of data that are best suited for this approach, as well as the limits of subdividing the data for visualization.

Acknowledgements

The authors wish to thank our research colleagues for their contributions to this project: Alex Gebhardt, Ana Jofre, Fanny Chevalier, Anne Stevens, Borzu Talaie and Robert Tu. We also thank Jonathan Lee and David Phillips at BBM Analytics. Research occurred in the Visual Analytics Laboratory, OCAD University and was funded through the Centre for Information Visualization and Data Driven Design (ORF-E) and NSERC.

References

[1] Brook, J. SUS – A Quick And Dirty Usability Scale. In *Usability Evaluation In Industry*. Taylor & Francis Ltd, London, Pages 189-194. 1996.

[2] Cleveland, W. S. *The elements of graphing data*. Pacific Grove, CA: Wadsworth.1985.

[3] Follette, L.J. and Hollands, J. G. Bars vs Stacked Bars: Effectiveness is task dependent. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* '98 42. Page 1652. 1998.

[4] Friendly, M.. A Brief History of Data Visualization. In *The Handbook of Data Visualization.* Springer Berlin Heidelberg. Pages 15-56. 2008

.