Do Defaults Matter? Evaluating the Effect of Defaults on User Preference for Multi-Class Scatterplots

Casey Haber, Lyndon Ong Yiu, Alark Joshi, Sophie Engle
University of San Francisco
2130 Fulton Street, San Francisco, CA
[cahaber,lcongyiu,apjoshi,sjengle]@usfca.edu

ABSTRACT

With the increasing availability and popularity of visualization tools, it is easier than ever to create visual representations of data. The available tools and libraries work for a range of users from non-programmers to those with significant programming experience. A major challenge, however, is that a majority of users frequently stick with the default settings when using software.

In this paper, we evaluate the effect of using defaults when visualizing the same data in four widely-used visualization tools: Tableau Desktop, Microsoft Excel, the ggplot2 R library, and the matplotlib Python library. We used the default settings in these tools to create multi-class scatterplots for several synthetic datasets generated using the scikit-learn package in Python.

We conducted a within-subjects pilot study with 39 users and a follow-up study with 202 users to explore whether users have strong preferences for different default settings. We found that computer science students prefer ggplot2, females preferred Tableau, young users or those with some college preferred Excel, and users in most other categories preferred matplotlib.

CCS Concepts

•**Human-centered computing** → **Information visualization;** *Visualization design and evaluation methods;*

Keywords

Defaults, aesthetics, user study, usability

1. INTRODUCTION

With the growing popularity of data visualization, there is an increased interest in using interactive visualizations for data exploration and decision making. The democratization of data visualization [21, 22, 27] has resulted in the wider population using visualization tools such as Tableau Desktop and Microsoft Excel. In the statistic and scientific communities, the R-based graphing package ggplot2 and Python-based plotting library matplotlib are widely used for rapid prototyping and creating publication-quality visualizations.

Alongside the growing interest in data visualization, there is increasing attention on the default settings in software. When Microsoft released Office 2007, they changed the default font from Times New Roman to Calibri because a majority of the users do not change the font and it had better on-screen readability [9]. Similarly, research has highlighted the disadvantages of having the rainbow colormap as the default [4, 16]. Additionally, Evergreen and Metzer [8] provide guidelines for designing clutter free, informative infographics and note that "default settings create too much visual noise" in Excel.

To learn more about the effect that the default visual style of a visualization tool can have, we studied user preferences for multi-class scatterplots generated using the defaults in Tableau Desktop [26], Microsoft Excel [25], ggplot2 [23], and matplotlib [13]. The main findings of our work are as follows:

1. *Defaults matter.* There are clear preferences for different default settings across tools. The settings most preferred users seem to depend on their background but not their familiarity with that tool.

2. *Preferences may be inconsistent.* Users are often inconsistent with their preferences. The dataset type and size does not explain this inconsistency, and more study is required to understand this phenomenon.

In the subsequent sections, we present related work, our approach and results from our user studies, a discussion of the results, and finally future directions for this work.

2. RELATED WORK

Gaviria [11] introduced *functional* information visualization versus *aesthetic* information visualization. According to the Gaviria, functional visualization aims to inform the viewer as quickly and as accurately as possible, whereas aesthetic visualization uses visually attractive forms to attract a viewer on an emotive level to increase the viewer's interest, draw the viewer's attention, and potentially lead to enjoyment [17]. Aesthetics [6, 7] and style [14, 24] in data visualization have been investigated to enhance the overall appeal of a visual representation.

Aesthetics and style of a visualization have been frequently discussed in the visualization community for user engagement in casual or ambient visualization settings [18]. Harri-

VINCI '16, September 24-26, 2016, Dallas, TX, USA

DOI: http://dx.doi.org/10.1145/2968220.2968241

son et al. [12] find that infographics aesthetics have a significant effect on engagement and memorability. With respect to engagement and memorability, Bateman et al. [1] have shed light on the effect of chartjunk on memorability and found significant recall for graphs with chartjunk. Borkin et al. [3] found that color and the "inclusion of a human recognizable object" can have a significant impact on memorability. In a recent study, Borkin et al. [2] further analyzed memorability and found that most memorable visualizations are able to convey their central message to the viewer.

Cawthon [7] evaluated the effect of aesthetic on the performance of users when exploring hierarchical data using techniques such as Treemap, IcicleTree, SpaceTree, Windows Explorer, BeamTrees, StarTree, Dendogram Tree, Polar View, StepTree, Botanical Viewer, and SunBurst. They found that the most familiar technique (Windows Explorer) did not actually perform the best. The SunBurst technique led to the highest accuracy (fewest errors). They also measured the "Latency in Erroneous Response" that measures the amount of time a participant spent with a visualization technique to eventually get an incorrect answer. This measure along with the "Rate of Abandonment" measure captures an aspect of user engagement.

Vande Moere [14] evaluated three specific styles for information visualization that consisted of an analytical style, magazine style, and an artistic style to represent the same data. They found that users perceived an analytical style visualization as being more usable and easy to understand. Interestingly, the style did seem to affect user perception of the usability and end goal of the visualization even though the styles were all representing the same data.

Artistic principles have been explored in the visualization community to convey information. Tateosian [19] discuss a painterly-rendering style to convey multi-attribute data to the viewer in the form of an engaging painting. They showed aesthetically pleasing visualizations of scientific phenomena such as a simulated supernova collapse, weather condition in a geographic region and so on. While the representations were aesthetically pleasing, there was no way of quantifying the benefit of using the painterly-rendering style as compared to traditional visualizations produced by software used by domain experts. Viegas and Wattenberg [20] encourage the visualization community to work more closely with artists to increase the aesthetic appeal and overall engagement of individuals to eventually "change attitudes" using the visualization. Wood et al. [24] present a sketchy style for presenting information visualizations in a prototype stage or to indicate uncertainty in data [5].

3. EVALUATING USER PREFERENCE

We decided to evaluate user preference for scatterplots due to their ubiquity in presentation graphics and scientific literature [10], and unlike simple line or bar charts, scatterplots have known issues with overdraw that may make them more sensitive to the default settings in many visualization tools. We focused on multi-class scatterplots so that the default colors of each tool would also play a role.

We studied the defaults for multi-class scatterplots in the following tools: Tableau Desktop for Mac v9.2.4, Microsoft Excel for Mac 2016, ggplot2 v2.0.0 package in R v3.2.3 [23], and the matplotlib v1.5 package in Python v3.5 [13]. We selected these tools since they are in wide use, have different user bases, and distinct default styles.

3.1 Pilot User Study

We conducted a within-subjects pilot study using Qualtrics and found that users overwhelmingly preferred ggplot2, with Tableau coming in second.

Setup. We received 46 responses. We removed 7 invalid responses resulting in 39 valid responses. The survey had 12 questions and took between 1–5 minutes to complete. Students were recruited from undergraduate and graduate data visualization classes at the University of San Francisco. Nearly all were 18 to 34 years old and familiar with scatterplots and common visualization tools like Tableau.

We generated a synthetic dataset with 500 rows, x and y values between 0 and 1, and 3 possible classes using the scikit-learn package in Python [15]. For each tool, we specified a 450 by 450 pixel size, upper-right legend placement, and consistent axis titles. We used the defaults for all other settings. We showed each user two images at a time and asked them to specify their preference. We showed every pair twice in opposite orders to identify consistent preferences. For example, users would see the matplotlib scatterplot next to ggplot2 scatterplot, and later ggplot2 followed by matplotlib. The order was counterbalanced.

Results. We had 12 questions and 39 users for a total of 468 responses. We looked at the frequency tools were selected and used the chi-squared test for count data to determine if the frequencies were statistically different from chance. **We found that ggplot2 was most preferred at 39% of the time or 184 times total.** The second was Tableau at 23% or 109 times followed by matplotlib at 21% or 97 times. Excel came in last at 17% or 78 times.

We also found that 44% of users were inconsistent at least once with their choices. An inconsistent response occurs when a user had different preferences for the same pair of tools. For example, suppose someone selected ggplot2 when shown ggplot2 and matplotlib, but selected matplotlib when the order was reversed.

This study gave us insight that users appear to have clear preferences for different default styles, but may not always be consistent. However, this study looked at a specific demographic and only one example scatterplot. We decided to run a larger follow-up study to explore this question further.

3.2 Amazon Mechanical Turk Study

We followed up our pilot study with a within-subject user study using Qualtrics and users from Amazon Mechanical Turk. We found the opposite conclusion from our pilot study—with ggplot2 being the *least* preferred. Instead, matplotlib was preferred but the results differed by background.

Setup. We received 256 responses. We filtered out 52 invalid responses resulting in 202 valid responses. Of these, we had 51% females 49% males. Most were 25 to 44 years old and attended some college or more. About 45% percent were 25 to 34 years old, 30% were 35 to 44, and 12% were 45 to 54 years old. The remaining 13% were outside of the 25 to 54 age bracket or declined to indicate. About 46% had a 4 year college degree and another 33% had some college or a 2 year college degree. High school graduates made up 10% and those with a professional or doctorate degree made up the remaining 10% of users.

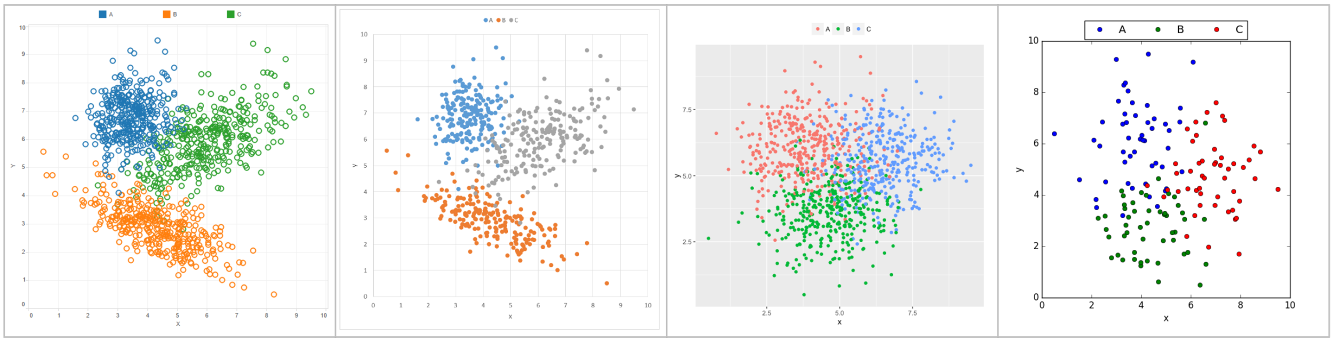

Figure 1: Example scatterplots used in the Amazon Mechanical Turk study. From left to right: (a) 1050 rows with mixed shape generated in Tableau, (b) 600 rows with mixed shape generated in Excel, (c) 1050 rows with round shape generated in ggplot2, (d) 150 rows with round shape generated in matplotlib.

We also found that most users were comfortable with scatterplots. About 54% were extremely or somewhat comfortable, 40% were neither comfortable or uncomfortable, and only 6% were somewhat or extremely uncomfortable with scatterplots. Most of our users were familiar with Excel, but not Tableau, ggplot2, or matplotlib. Of those that responded, 26% were very or extremely familiar with Excel, 31% were moderately familiar, and 12% were slightly to not familiar at all. About 58% were not familiar with Tableau, 62% were not familiar with ggplot2, and 60% were not familiar with matplotlib. Due to a glitch, we did not collect familiarity feedback from 31% of users.

We generated 6 datasets using the sci-kit learn package in Python [15]. Each had x and y values between 0 and 10 and three classes. We generated datasets with 150, 600, and 1050 rows and two different "shapes" of blobs (round and mixed). For each tool, we specified an output area of 600 by 600 pixels, an upper-center legend placement above the plot, and consistent axis titles. We used defaults for the colors, axis breaks, grid lines, point shape, and size. When the images were displayed to the user, they were scaled down to 300 by 300 pixels to allow for all four tools to fit on the same screen. See Figure 1 for examples.

We asked users several questions about their background (gender, age bracket, education level, familiarity with scatterplots, as well as familiarity with Excel, Tableau, ggplot2, and matplotlib). We then asked users to choose their preference from four visualizations (Tableau, Excel, ggplot2, and matplotlib) 6 times (once for each dataset). We did not label which visualization was generated by which tool, and the order of the datasets and tools were counterbalanced.

Results. We again looked at the frequency each tool was selected and used the chi-squared test to determine the probability the observed frequencies occurred by chance. If the frequency fell below 5, we combined together frequencies. For example, "Professional degree" and "Doctorate degree" were combined into a single "Graduate degree" category since very few users held doctorates. The overall observed frequencies were found to be significantly different from chance. **Users preferred ggplot2 with the *least* frequency, and many groups of users preferred matplotlib.**

We repeated our analysis for different subsets of users. We found that the distribution broken down by age bracket and education level were significant. Those 25 to 34 years old and with some college preferred Excel. Older users, high school graduates, and those with 4 year college degrees preferred matplotlib. Tableau was preferred by younger users and those with 2 year degrees. None of these groups preferred ggplot2. Gender was also highly significant. Females preferred Tableau did not prefer ggplot2, both by decent margins. Males preferred matplotlib and Excel.

We also found statistical significance in the distribution broken down by level of comfort with scatterplots. Users that were neutral to somewhat comfortable preferred matplotlib (but by a small margin) and did not prefer ggplot2. Users extremely comfortable with scatterplots preferred Tableau and those somewhat or extremely uncomfortable preferred Excel. However, the frequencies were very close.

Familiarity with a tool did not seem to track with user preference. The distribution was not statistically significant for most tools, although matplotlib did have borderline significance. Surprisingly, users familiar with matplotlib just barely preferred Tableau over matplotlib.

We finally observed that 53% of users (108 of 202 users) were inconsistent at least once. The other 47% of users always choose the same tool no matter the dataset. However, the distributions of preference by dataset type or size were not statistically different from chance.

4. DISCUSSION

Both our pilot study and follow-up Amazon Mechanical Turk study show that defaults matter. However, the studies had opposite conclusions: ggplot2 was strongly preferred in the pilot but never the top preference in the follow-up study. In fact, matplotlib was most preferred overall. This leads us to an interesting question—why? What factors might explain how well matplotlib performed in the Amazon Mechanical Turk study versus ggplot2 in the pilot study? There are distinct differences in the user demographics. Could this be explained by a bias for or against more "designed" tools in these two user pools?

Finally, we found that 44% to 53% of users were inconsistent with their preferences at least once—beyond what could be explained by noise. In our pilot study, the exact same images were used but in different orders. Users saw images for different datasets in our follow-up study, but we did not find any statistical significance in preference by dataset type or size. Does this indicate that these users did not have a

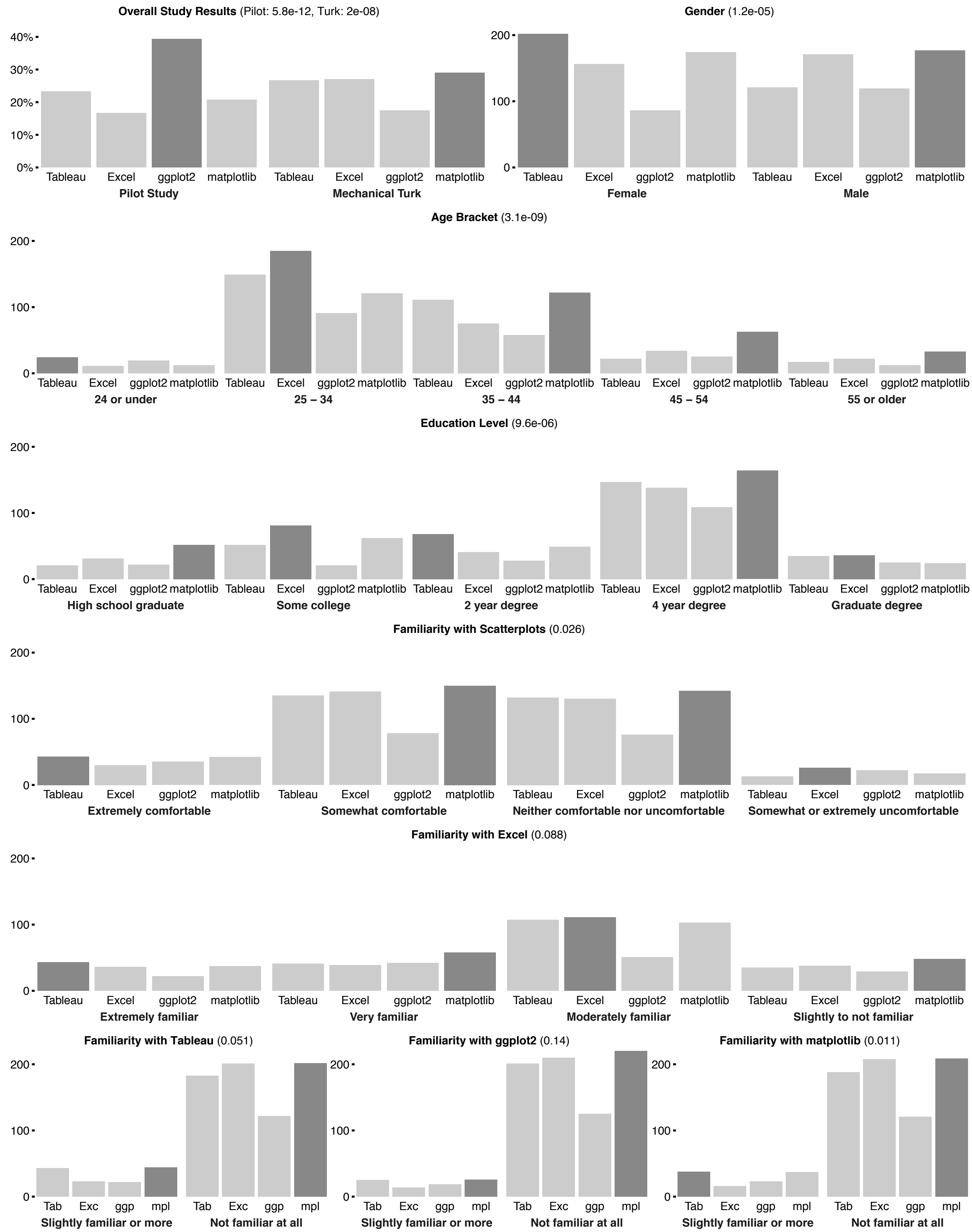

Figure 2: Preference frequencies from the pilot study and follow-up Amazon Mechanical Turk study. The p-value from each chi-squared test is given on each panel.

strong preference, preferred two visualizations equally, were not taking the study seriously, or is this phenomenon explained by other factors?

5. CONCLUSIONS

In this paper, we present the results of our 39 user pilot study with computer science students and our 202 user followup study using Amazon Mechanical Turk. We found that there are clear preferences for different default settings across tools. However, these preferences are not always consistent and depend on a user's background. Interestingly enough, familiarity with a tool did not seem to track with user preference. Some of these findings did not meet our expectations, and produced several questions as to why these differences occurred. It is clear more study is required, and that user background and preference should be considered when studying the impact of defaults on user performance.

6. REFERENCES

[1] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful junk?: The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2573–2582, 2010.

[2] M. Borkin, Z. Bylinskii, N. W. Kim, C. M. Bainbridge, C. Yeh, D. Borkin, H. Pfister, and A. Oliva. Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, 2016.

[3] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.

[4] D. Borland and R. M. Taylor II. Rainbow color map (still) considered harmful. *IEEE Computer Graphics and Applications*, (2):14–17, 2007.

[5] N. Boukhelifa, A. Bezerianos, T. Isenberg, and J.-D. Fekete. Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2769–2778, 2012.

[6] L. Byron and M. Wattenberg. Stacked graphs–geometry & aesthetics. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1245–1252, 2008.

[7] N. Cawthon and A. V. Moere. The effect of aesthetic on the usability of data visualization. In *Proceedings of the 11th International Conference on Information Visualization*, pages 637–648, 2007.

[8] S. Evergreen and C. Metzner. Design principles for data visualization in evaluation. *New Directions for Evaluation*, 140:5–20, 2013.

[9] J. Friend. Why did Microsoft change the default font to Calibri? URL: http://qr.ae/1KPYLj. Accessed: 2016-03-18.

[10] M. Friendly and D. Denis. The early origins and development of the scatterplot. *Journal of the History of the Behavioral Sciences*, 41(2):103–130, 2005.

[11] A. R. Gaviria. When is information visualization art? determining the critical criteria. *Leonardo*, 41(5):479–482, 2008.

[12] L. Harrison, K. Reinecke, and R. Chang. Infographic aesthetics: Designing for the first impression. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1187–1190, 2015.

[13] J. D. Hunter et al. Matplotlib: A 2D graphics environment. *Computing in Science and Engineering*, 9(3):90–95, 2007.

[14] A. V. Moere, M. Tomitsch, C. Wimmer, B. Christoph, and T. Grechenig. Evaluating the effect of style in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2739–2748, 2012.

[15] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*, 12:2825–2830, 2011.

[16] B. E. Rogowitz and L. A. Treinish. Data visualization: The end of the rainbow. *IEEE Spectrum*, 35(12):52–59, 1998.

[17] B. Saket, C. Scheidegger, and S. Kobourov. Towards understanding enjoyment and flow in information visualization. In *Eurographics Conference on Visualization (EuroVis) Short Papers*, 2015.

[18] T. Skog, S. Ljungblad, and L. E. Holmquist. Between aesthetics and utility: Designing ambient information visualizations. In *IEEE Symposium on Information Visualization*, pages 233–240, 2003.

[19] L. G. Tateosian, C. G. Healey, and J. T. Enns. Engaging viewers through nonphotorealistic visualizations. In *Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering*, pages 93–102, 2007.

[20] F. B. Viégas and M. Wattenberg. Artistic data visualization: Beyond visual analytics. In *Online Communities and Social Computing*, pages 182–191. Springer, 2007.

[21] F. B. Viegas, M. Wattenberg, and J. Feinberg. Participatory visualization with Wordle. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1137–1144, 2009.

[22] F. B. Viegas, M. Wattenberg, F. Van Ham, J. Kriss, and M. McKeon. ManyEyes: A site for visualization at internet scale. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1121–1128, 2007.

[23] H. Wickham. *ggplot2: Elegant Graphics for Data Analysis*. Springer Science & Business Media, 2009.

[24] J. Wood, P. Isenberg, T. Isenberg, J. Dykes, N. Boukhelifa, and A. Slingsby. Sketchy rendering for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2749–2758, 2012.

[25] Microsoft Excel. products.office.com/en-us/excel (Accessed 2016-03-18).

[26] Tableau Desktop. tableau.com/products/desktop (Accessed 2016-03-13).

[27] Tableau Public. public.tableau.com (Accessed 2016-03-13).