‘You Draw It’: Implementation of visually fitted trends with `r2d3`

Emily A. Robinson
University of Nebraska - Lincoln
Lincoln, NE
`emily.robinson@huskers.unl.edu`

Reka Howard
University of Nebraska - Lincoln
Lincoln, NE
`rekahoward@unl.edu`

Susan VanderPlas
University of Nebraska - Lincoln
Lincoln, NE
`susan.vanderplas@unl.edu`

Abstract

How do statistical regression results compare to intuitive, visually fitted results? Fitting lines by eye through a set of points has been explored since the 20th century. Common methods of fitting trends by eye involve maneuvering a string, black thread, or ruler until the fit is suitable, then drawing the line through the set of points. In 2015, the New York Times introduced an interactive feature, called ‘You Draw It,’ where readers are asked to input their own assumptions about various metrics and compare how these assumptions relate to reality. This research is intended to implement ‘You Draw It,’ adapted from the New York Times, as a way to measure the patterns we see in data.

Keywords: graphics, user interaction, regression

1 Introduction

We all use statistical graphics, but how do we know that the graphics we use are communicating properly? Through experimentation, graphical testing methods allow researchers to conduct studies geared at understanding human ability to conduct tasks related to the perception of statistical charts such as differentiation, prediction, estimation, and extrapolation. All of these types of tests require different levels of use and manipulation of the information being presented in the chart. Efforts in the field of statistical graphics have developed graphical testing tools and methods, such as the lineup protocol (Buja et al., 2009). The advancement of graphing software provides the tools necessary to develop new methods of testing statistical graphics.

1.1 Measuring Patterns and Trends

One such aspect of interest is the ability to identify and detect trends in data. Our visual system is naturally built to look for structure and identify patterns. For instance, points going down from left to right indicates a negative correlation between the x and y variables.

Initial studies in the 20th century explored the use of fitting lines by eye through a set of points (Finney, 1951; Mosteller et al., 1981). Common methods of fitting trends by eye involved maneuvering a string, black thread, or ruler until the fit is suitable, then drawing the line through the set of points. Recently, Ciccione and Dehaene (2021) conducted a comprehensive set of studies investigating human ability to detect trends in graphical representations from a psychophysical approach.

In 2015, the New York Times introduced an interactive feature, called ‘You Draw It’ (Aisch et al., 2015), where readers input their own assumptions about various metrics and compare how these assumptions relate to reality. The New York Times team utilizes Data Driven Documents (D3) that allows readers to predict these metrics through the use of drawing a line on their computer screen with their mouse.

What remains to be determined is how we can compare our intuitive visual sense of patterns to those determined by statistical methods.

1.2 Research Objectives

The goal of this research is to implement ‘You Draw It,’ adapted from the New York Times feature, as a way to measure the patterns we see in data. Here, we provide technical details of the software development, utilizing interactive graphics in R. We then share results from our study which validates ‘You Draw It’ as a method for graphical testing and apply an appropriate data analysis method to the participant data.

2 Development

2.1 ‘You Draw It’ Task

When completing the graphical task, users are shown an interactive scatter-plot (Figure 1) along with the prompt, “Use your mouse to fill in the trend in the yellow box region.” The yellow box region moves along as the user draws their trend-line, providing a visual cue which indicates where the user still needs to complete a trend line. After the entire domain has been visually estimated or predicted, the yellow region disappears, indicating the participant has completed the task. Prior to study participation, example gifs are shown and users are asked to complete practice plots to ease the learning curve of the task. Visit `emily-robinson.shinyapps.io/can-you-draw-it` for a test applet.

2.2 Code Sketch

Using Shiny (Chang et al., 2021) and JavaScript, we modified the New York Times ‘You Draw It’ feature for the purpose of testing statistical graphics allowing us to

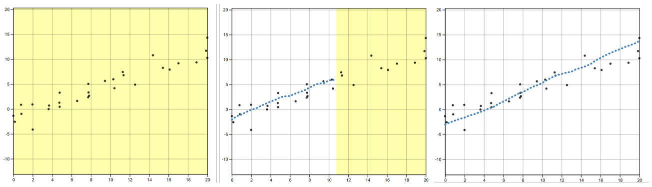

Figure 1: You Draw It' task plot as shown to user.
left: illustrates what user first sees with the prompt *Use your mouse to fill in the trend in the yellow box region.*
middle: illustrates what the user sees while completing the task.
right: illustrates the users finished trend line.

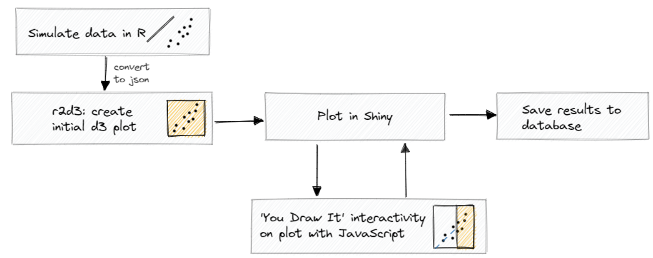

Figure 2: Sketch of underlying code for 'You Draw It', illustrating the data simulation conducted in R, the initial setup of the visual stimuli with D3 source code, along with the iterative process between the user interaction and plotting in Shiny. Once the user is done drawing the line, we saved the results of the drawn line to a database for analysis.

incorporate user interaction, conduct studies online, and store participant responses.

Data Driven Documents (D3) (Bostock et al., 2011), a JavaScript-based graphing framework that facilitates user interaction, was used to create the 'You Draw It' visual. Major news and research organizations such as the New York Times, FiveThirtyEight, Washington Post, and the pew Research Center use D3 to create and customize graphics. A challenge of working with D3 is the environment necessary to display the graphics and images. The `r2d3` package (Strayer et al., 2020) provides an efficient integration of D3 visuals into R HTML formats. We integrate the D3 visual source code into an R Shiny (Chang et al., 2021) application in order to allow for user interaction and data collection.

Figure 2 illustrates the initial setup of the visual stimuli along with the iterative process between the user interaction and plotting. We conducted all data simulation and processing in R and output two data sets - *point data* and *line data* - containing (x, y) coordinates corresponding to either a simulated point or fitted value predicted by a statistical model respectively. Then, the r2d3 package converts the data sets in R to JavaScript Object Notation (JSON) to be interpreted by the `D3.js` code. We define functions in `D3.js` to draw the initial plot and set up drawable points for the user drawn line. Drag events in `D3.js` are utilized to observe and react to user input. Shiny Messages are used to communicate the user interaction between the `D3.js` code and the R environment. The plot is then rendered and updated on user interaction into the R shiny application with the `RenderD3` and `d3Output` functions. Once the user is done drawing the line, we saved the results of the drawn line to a database for analysis.

Parameters for aesthetic design choices are defined in a list of options and `r2d3` passes these to the D3.js code. For instance, we can specify the buffer space allowed for the x and y axes to avoid users to anchor their lines to the axes limits. For `D3.js` source code, visit `https://github.com/earobinson95/presentations/blob/master/can-you-draw-it/www/main-d3v5.js`.

2.3 Challenges

During the development process, there were a few challenges we had to overcome. We documented these challenges with GitHub commits and outline places where frustration occurred. Data Driven documents uses Scalable Vector Graphics (SVG), thus requiring careful transformation between the pixels and plot coordinates to align the simulated data and the user drawn line appropriately. With the layered framework of SVG's, it was important to place the layers in the right order so certain features would appear where desired. For example, we wanted the opacity of the yellow box region to appear under the points while still showing the grid lines. As previously mentioned, `r2d3` automatically converts the data set to a JSON file to be interpreted by the underlying source code, however we provided two data sets (point data and line data) and converted the two data sets to a JSON file before passing them to the `r2d3` argument so that we could reference both data sets within the `D3.js` code. One constraint of the development code to date is that it is only built to work with one-to-one functions and does not allow for multiple estimates or predictions for one x-value.

3 Application

3.1 Tool Validation

We conducted a study in order to validate 'You Draw It' as a method for graphical testing, comparing results to the less technological method utilized in Mosteller et al. (1981). Results from our study were consistent with those found in the previous study; when shown points following a linear trend, participants tended to fit the slope of the first principal component over the slope of the ordinary least-squares regression line (Figure 3). This trend was most prominent when shown data simulated with larger variances. This study reinforces the differences between intuitive visual model fitting and statistical model fitting, providing information about human perception as it relates to the use of statistical graphics.

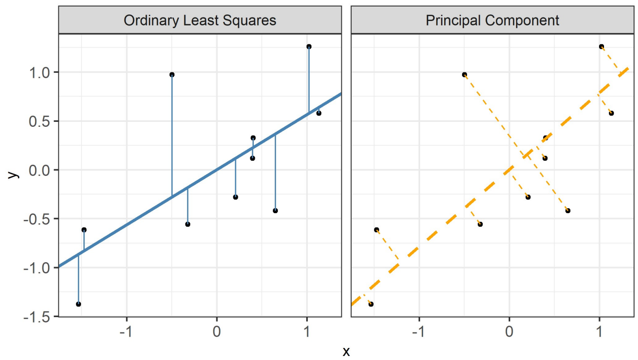

Figure 3: Comparison between an OLS regression equation which minimizes the vertical distance of points from the line and a regression equation with a slope calculated by the first principal component which minimizes the smallest distance of points from the line.

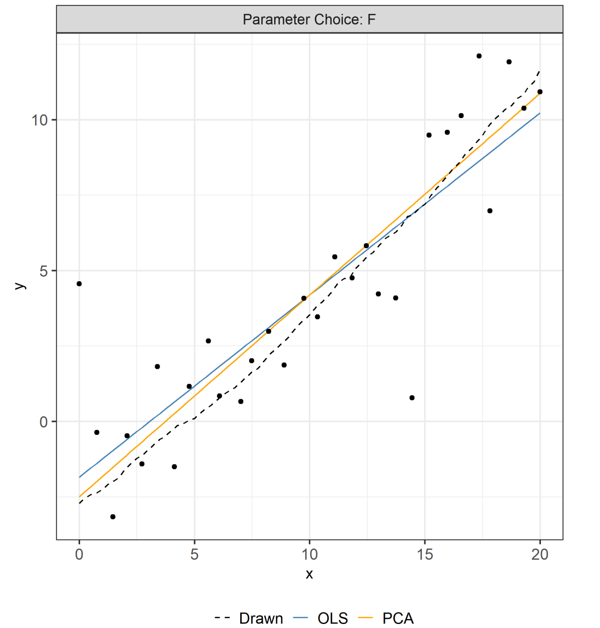

Figure 4: Example of three trend lines showing the the OLS fitted, PCA fitted, and participant drawn values overlaid on the simulated data points

3.2 Data Analysis

Feedback data from conducted studies are collected and stored in a database for analysis (Figure 4). Within the collected feedback data, we know the simulated data points, the predicted values from the statistical model, and the predicted values from the user drawn line. In our initial studies, a unique data set was simulated independently for each participant. Therefore, we evaluate the accuracy of the user drawn line by observing the deviation, vertical residuals, between the user drawn line and the predicted values from the statistical model. We use a Generalized Additive Mixed Model (GAMM) to model the vertical residuals in order to statistically compare visually fitted trends to actual metrics, simulated data models, or statistical regression results. A benefit of using a GAMM is the estimation of smoothing splines, allowing for flexibility in the residual trend.

4 Future Work

In this work, we implemented and validated 'You Draw It' as a way to measure the patterns we see in data. We demonstrated the use of generalized additive models to statistically model participant data. While technical details of the development process are presented here, we intend to create an R package designed for easy implementation of 'You Draw It' task plots in order to make this tool accessible to other researchers. Further investigation is necessary to implement this method real data in order to facilitate scientific communication.

References

Gregor Aisch, Amanda Cox, and Kevin Quealy. 2015. You draw it: How family income predicts children's college chances. *The New York Times*, 28.

Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. 2011. D^3 data-driven documents. *IEEE transactions on visualization and computer graphics*, 17(12):2301–2309.

Andreas Buja, Dianne Cook, Heike Hofmann, Michael Lawrence, Eun-Kyung Lee, Deborah F Swayne, and Hadley Wickham. 2009. Statistical inference for exploratory data analysis and model diagnostics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1906):4361–4383.

Winston Chang, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges, 2021. *shiny: Web Application Framework for R*. R package version 1.7.1.

Lorenzo Ciccione and Stanislas Dehaene. 2021. Can humans perform mental regression on a graph? accuracy and bias in the perception of scatterplots. *Cognitive Psychology*, 128:101406.

DJ Finney. 1951. Subjective judgment in statistical analysis: An experimental study. *Journal of the Royal Statistical Society: Series B (Methodological)*, 13(2):284–297.

Frederick Mosteller, Andrew F Siegel, Edward Trapido, and Cleo Youtz. 1981. Eye fitting straight lines. *The American Statistician*, 35(3):150–152.

Nick Strayer, Javier Luraschi, and JJ Allaire, 2020. *r2d3: Interface to 'D3' Visualizations*. R package version 0.2.5.

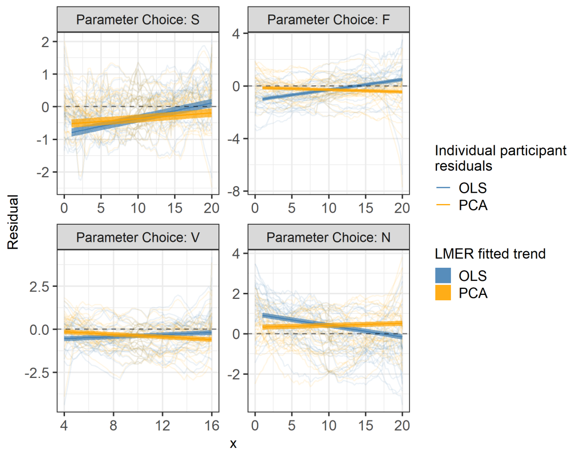

Figure 5: Estimated trends of residuals (vertical deviation of participant drawn points from both the OLS (blue) and PCA (orange) fitted points) as fit by a linear mixed model.

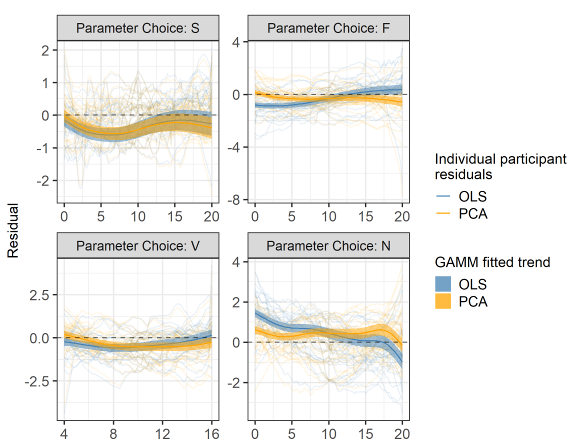

Figure 6: Estimated trends of residuals (vertical deviation of participant drawn points from both the OLS (blue) and PCA (orange) fitted points) as fit by a generalized additive mixed model.

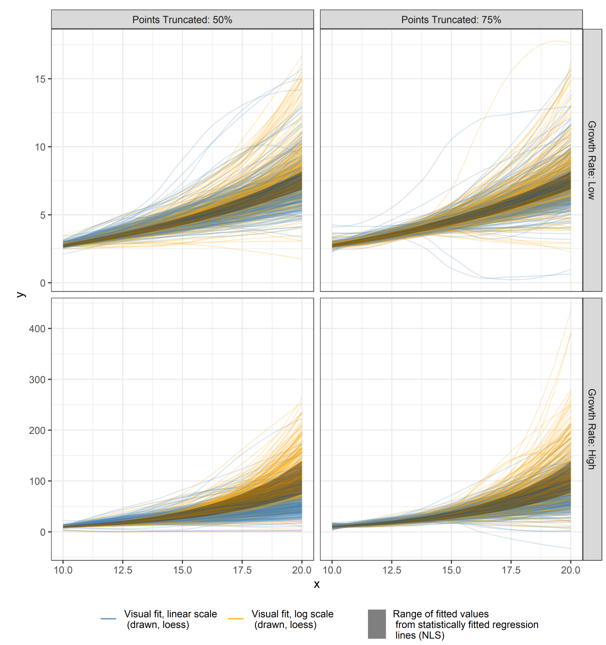

Figure 7: Spaghetti plot of results from a study which asked participants to forcast trends of exponentially increasing data. Participants drawn lines on the linear scale are shown in blue and the log scale are shown in orange. Variability in the statistically fitted regression lines occured due to a unique data set being simulated for each individual; the gray band shows the range fitted values from the statistically fitted regression lines.