'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 effectively? Through experimentation, graphical testing methods allow researchers to conduct studies designed to understand how we perceive graphics and perform graphical tasks such as differentiation, prediction, estimation, and extrapolation. Each of these levels of interaction with a graph require a different method of engagement with and use of the information presented in a chart. In this paper, we describe the adaptation of an old tool for graphical testing and evaluation, eye-fitting, for use in modern web-applications suitable for testing statistical graphics. We present an empirical evaluation of this testing method for linear regression, and briefly discuss an extension of this method to non-linear applications.

1.1 Measuring Patterns and Trends

One of the most common charts created is a scatter plot of points over time; these charts show up regularly in news articles and in scientific publications alike. These charts rely on our ability to identify and detect trends in data. Our visual system is naturally built to look for structure and identify patterns, including patterns and trends over time; many times we do not even notice this process happening subconsciously. As shown in Fig. 1, a viewer engaging with a plot of weekly average gas prices over time may perform several cognitive operations. First, they scan the plot and assesses points to determine if there are any outliers or otherwise remarkable
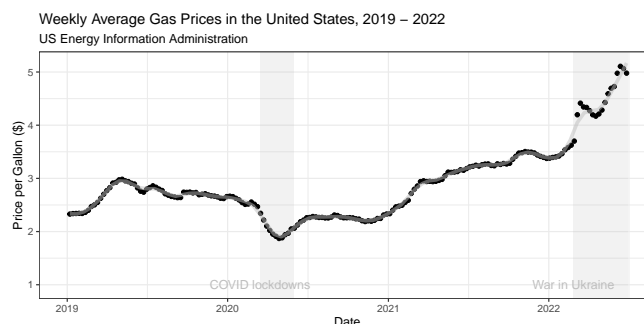

Figure 1: Weekly average gas prices in the United States, 2019-June 2022. Additional mental operations a viewer might perform while looking at the plot are annotated in grey. (US Energy Information Administration, 2022).

points. Then, they may fit a rough mental smooth/trend line to the points to summarize the useful information and remove variability. Finally, an interested and engaged viewer may pull in additional contextual information from long-term memory, seeking to explain variation in the mental 'trend' with supplemental information, such as COVID lock downs which reduced gasoline demand and the beginning of the war in Ukraine, which wreaked havoc on the supply and demand for global energy sources.

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 was 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. XXX and found what? Describe the results XXX

While psychologists and statisticians have been using eye-fitting techniques to assess our innate perceptual statistical modeling abilities, news organizations have used similar techniques in order to draw readers in and demonstrate the difference between readers' assumptions and reality. 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 of news interest and compare these assumptions to reality. The New York Times team used Data Driven Documents (D3), a JavaScript library that allows readers to interact with a chart directly by drawing a line on their computer screen with a mouse. Despite the

somewhat different purpose behind this feature, the D3 driven method used by the NYT is wonderfully intuitive, and does not require the assumption of linearity, making it much more adaptable to testing how viewers perceive and predict when presented with non-linear data.

We set out 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(R Core Team, 2022). We then share results from our study which validates 'You Draw It' as a method for graphical testing on linear trends and apply an appropriate data analysis method to the participant data. We also briefly demonstrate the use of the 'You Draw It' method and analysis on nonlinear 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. [1]
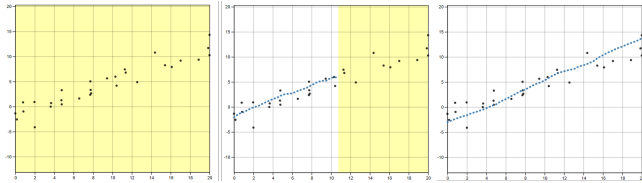

Figure 2: You Draw It' task plot as shown to user.
left: illustrates what user first sees with the prompt *Use your mouse to fill in the trend in the yellow box region.*
middle: illustrates what the user sees while completing the task.
right: illustrates the users finished trend line.

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 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.
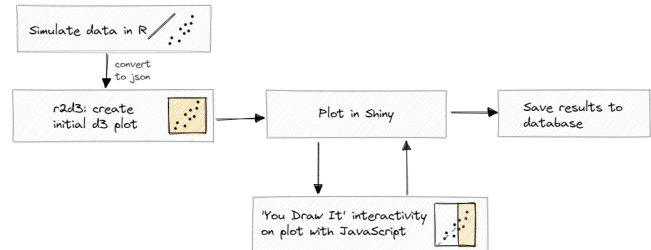

Figure 3: Sketch of underlying code for 'You Draw It', illustrating the data simulation conducted in R, the initial setup of the visual stimuli with D3 source code, along with the iterative process between the user interaction and plotting in Shiny. Once the user is done drawing the line, we saved the results of the drawn line to a database for analysis.

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. [2]

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

---

[1]Visit `emily-robinson.shinyapps.io/can-you-draw-it` for a test applet.

[2]For D3.js source code, visit `https://github.com/earobinson95/presentations/blob/master/can-you-draw-it/www/main-d3v5.js`.

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 reinforced 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.
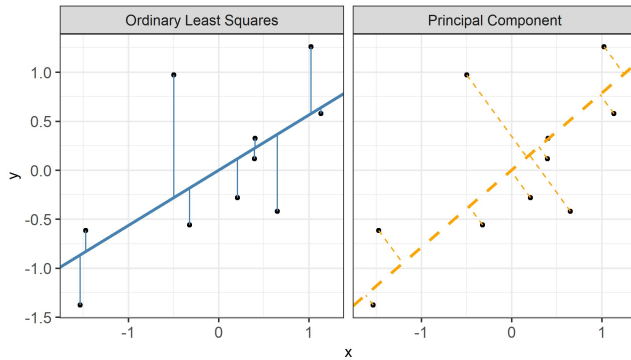

Figure 4: Comparison between an OLS regression equation which minimizes the vertical distance of points from the line and a regression equation with a slope calculated by the first principal component which minimizes the smallest distance of points from the line.

3.2   Data Analysis

Feedback data from conducted studies were collected and stored in a database for analysis (Figure 4). Within the collected feedback data, we knew 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
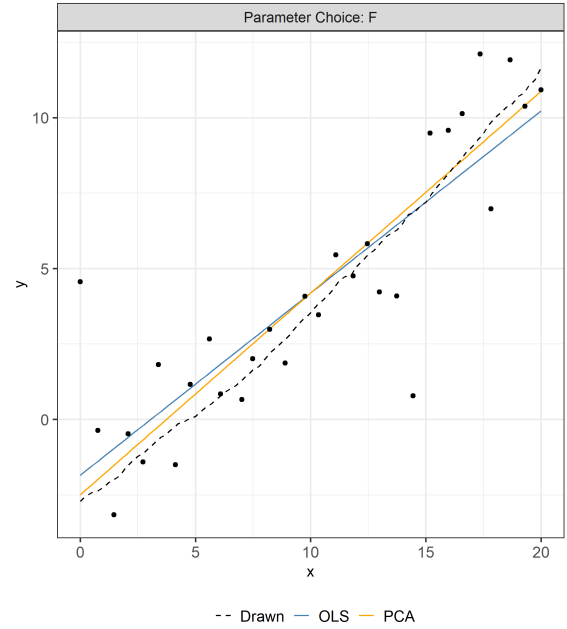

Figure 5: Example of validation feedback data where three trend lines show the the OLS fitted, PCA fitted, and participant drawn values overlaid on the simulated data points

line by observing the deviation, vertical residuals, between the user drawn line and the predicted values from the statistical model. Due to participant variation, we use mixed models to evaluate the vertical residuals in order to statistically compare visually fitted trends to actual metrics, simulated data models, or statistical regression results. In our validation study, we first analyzed the residual trend using a linear mixed model (LMM), thus constraining the fit to a linear trend (Figure 5). In addition, we fit a Generalized Additive Mixed Model (GAMM) to allow for flexibility in the residual trend and to generalized the method to nonlinear models (Figure 6).

After validating the 'You Draw It' method, we conducted a study utilizing the new method to evaluate participants ability to make forecasts for exponentially increasing data on log and linear scales. Along with analyzing the feedback data with the GAMM method for flexibility due to nonlinear data, we used spaghetti plots to visualize participants forecasts compared to the nonlinear least squares statistical model and make comparisons between two chart design features (Figure 6).

4   Future Work

In this work, we implemented and validated 'You Draw It' as a new method to measure the patterns we see in data. We demonstrated the use of generalized additive models to statistically model participant feedback data and generalized the analysis method to nonlinear trends. 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 with real data in order to facilitate scientific
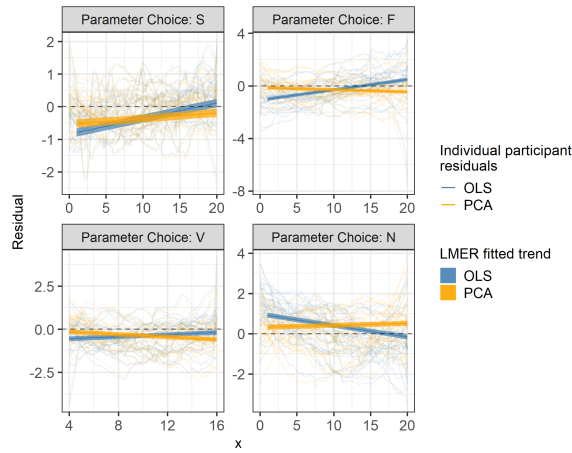
Figure 6: Validation study estimated trends of residuals (vertical deviation of participant drawn points from both the OLS (blue) and PCA (orange) fitted points) as fit by a linear mixed model.
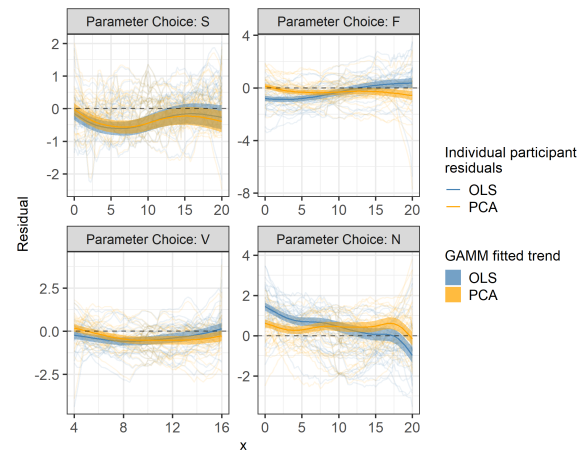

Figure 7: Validation study estimated trends of residuals (vertical deviation of participant drawn points from both the OLS (blue) and PCA (orange) fitted points) as fit by a generalized additive mixed model.
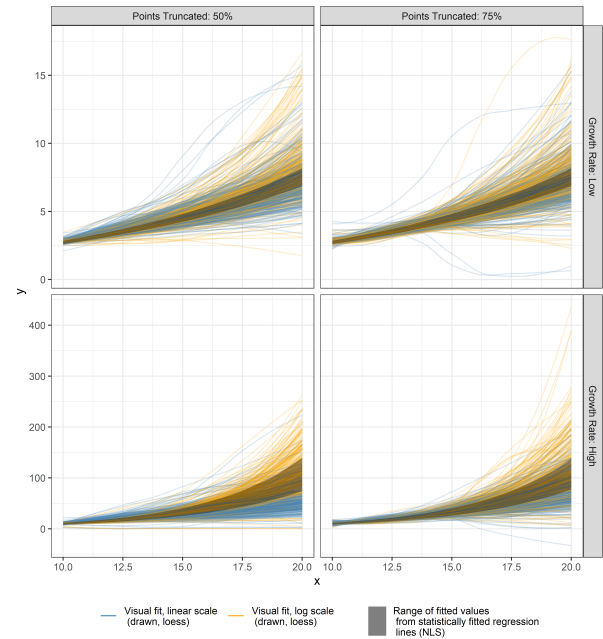
communication.

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.

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.

R Core Team, 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

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

US Energy Information Administration. 2022. Weekly U.S. All Grades All Formulations Retail Gasoline Prices (Dollars per Gallon), June.

Figure 8: 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.