Perception and Cognitive Implications of Logarithmic Scales for Exponentially Increasing Data: Perceptual Sensitivity Tested with Statistical Lineups

Emily A. Robinson 1
Department of Statistics, California Polytechnic State University - San Luis Obispo
and
Reka Howard 2
Department of Statistics, University of Nebraska - Lincoln
and
Susan VanderPlas 3
Department of Statistics, University of Nebraska - Lincoln

January 12, 2024

Abstract

Logarithmic transformations are a standard solution to displaying data that span several magnitudes within a single graph. This paper investigates the impact of log scales on perceptual sensitivity through a visual inference experiment using statistical lineups. Our study evaluated participant's ability to detect differences between exponentially increasing data, characterized by varying levels of curvature, using both linear and logarithmic scales. Participants were presented with a series of plots and asked to identify the panel that appeared most different from the others. Due to the choice of scale altering the contextual appearance of the data, the results revealed slight perceptual advantages for both scales depending on the curvatures of the compared data. This study serves as the initial part of a three-paper series dedicated to understanding the perceptual and cognitive implications of using logarithmic scales for visualizing exponentially increasing data. These studies serve as an example of multi-modal graphical testing, examining different levels of engagement and interaction with graphics to establish nuanced and specific guidelines for graphical design.

Keywords: log scales, visual inference, graphical testing

1 Introduction

Effective communication of data is critical in influencing people's opinions and actions. This consideration was particularly true during the COVID-19 pandemic, where data visualizations and dashboards were vital in informing the public and policymakers about the outbreak's status. Local governments relied on graphics to inform their decisions about shutdowns and mask mandates, while residents were presented with data visualizations to encourage compliance with these regulations. A major issue designers encountered when creating COVID-19 plots was how to display data from a wide range of values (Fagen-Ulmschneider, 2020, Burn-Murdoch et al. (2020)). When faced with data that span several orders of magnitude, we must decide whether to show the data on its original scale (compressing the smaller magnitudes into a relatively small area) or to transform the scale and alter the contextual appearance of the data. Log axis transformations have emerged as a standard solution to this challenge, as they allow for the display of data over several orders of magnitude within a single graph.

Exponential data is one such example of a function that compresses smaller magnitudes into a smaller area; Fig. 1 presents hard drive capacity over the past forty years on both the linear and log scales to demonstrate the usefulness of log scales when dealing with data spanning multiple magnitudes. Logarithms facilitate the conversion of multiplicative relationships (displaying 1 & 10 with a distance of 10 units apart and displaying 10 & 100 with a distance of 90 units apart) to additive relationships (displaying 1 & 10 and 10 & 100 an equal distance apart), highlighting proportional relationships and linearizing power functions (Menge et al., 2018). Logarithms also have practical applications, simplifying the computation of small numbers such as likelihoods and transforming data to conform to statistical assumptions. Although log scales have a long history of use in fields such
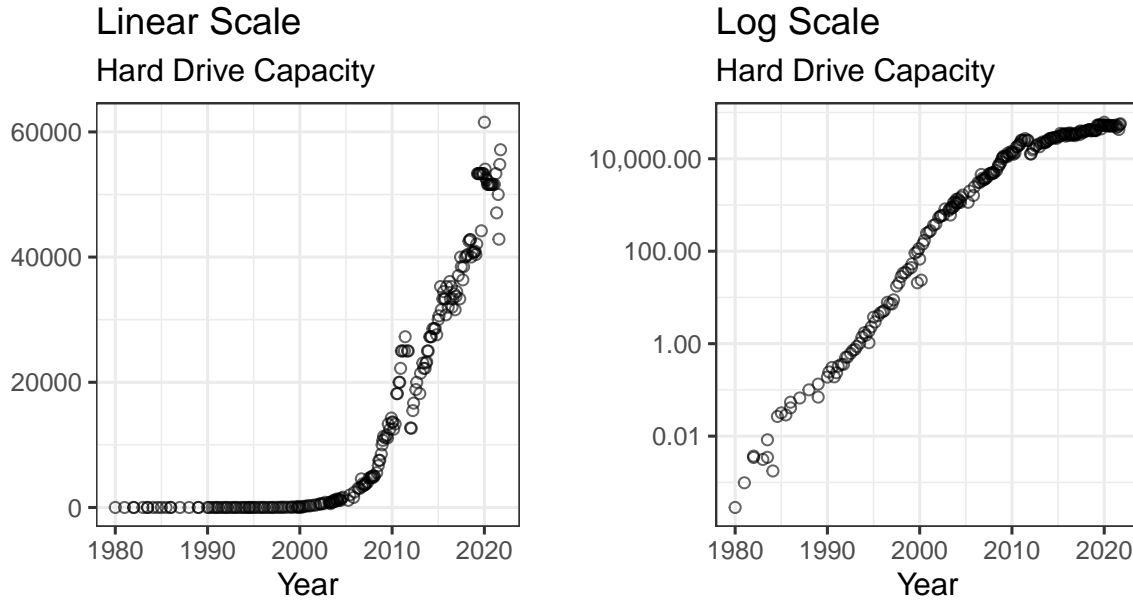
Figure 1: These plots present hard drive capacity over the past forty years on both the linear and log scale and illustrate the use of the log scale when displaying data which spans several magnitudes.

as ecology, psychophysics, engineering, and physics (Heckler et al., 2013; Waddell, 2005), there is still a need to understand the implications of their use and provide best practices for their implementation.

Apart from the biases resulting from using log scales, there is a general misinterpretation of exponential growth. Early stages of exponential growth often appear to have a small growth rate, while the middle stage seems to exhibit more quadratic growth. It is only in the later stages that the exponential growth becomes apparent. Fig. 2 highlights the three stages and associated appearances of exponential growth at each stage (Von Bergmann, 2021). This misinterpretation can lead to decisions made under inaccurate understanding,
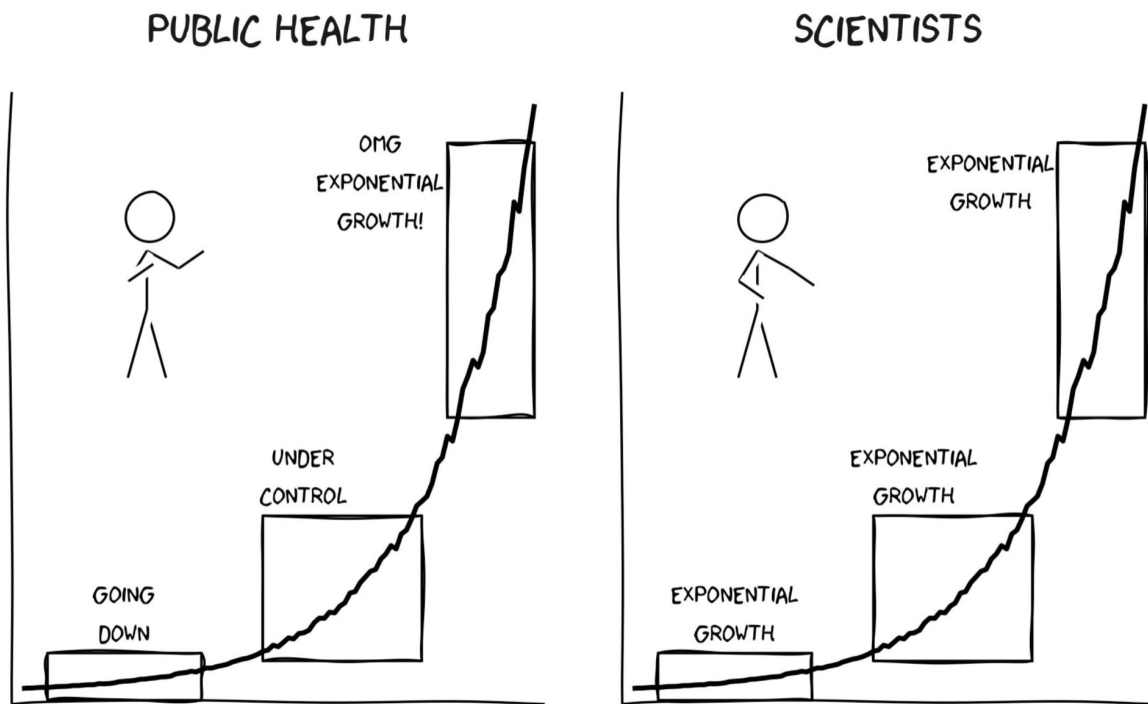
Figure 2: This figure highlights the three stages and associated appearances of exponential growth at each stage. Early stages of exponential growth often appear to have a small growth rate, while the middle stage seems to exhibit more quadratic growth. It is only in the later stages that the exponential growth becomes apparent.

resulting in potential consequences.

Previous studies have explored the estimation and prediction of exponential growth and found that individuals often underestimate exponential growth when presented values numerically and graphically (Wagenaar and Sagaria, 1975). The hierarchy of plot objects, such as lengths and angles, as described by Cleveland and McGill (1985), offers a possible explanation for the underestimation observed in exponentially increasing trends. Experiments conducted by Wagenaar and Sagaria (1975), Jones (1977), and MacKinnon and Wearing (1991) aimed to improve estimation accuracy for exponential growth. While contextual knowledge or experience did not enhance estimation, instruction on exponential growth

4

reduced underestimation by prompting participants to adjust their initial starting value (Wagenaar and Sagaria, 1975; Jones, 1977). Furthermore, providing immediate feedback to participants about the accuracy of their predictions improved estimation (MacKinnon and Wearing, 1991).

Log transforming the data may address our inability to predict exponential growth accurately. However, this transformation introduces new complexities, as most readers may need to be mathematically sophisticated enough to intuitively understand logarithmic math and translate it back into real-world effects. Despite the transformative power of logarithmic scales in facilitating accurate data representation, Menge et al. (2018)'s survey of ecologists highlights the challenges associated with the widespread comprehension of log-scaled data. Notably, the study identifies prevalent misconceptions arising from linear extrapolation assumptions in log-log space, a factor that often leads to neglect of the underlying exponential relationships in linear-linear space.

Building upon the need for a nuanced understanding of data representation, Buja et al. (2009) introduced statistical lineups as a framework for statistical inference and graphical tests. Statistical lineups treat a data plot as a visual statistic, summarizing the data as a numerical function or mapping. Evaluation of a panel in a statistical lineup requires visual inspection by a person, and if visual evaluations lead to different results, two visualization methods are deemed significantly different. Recent studies have utilized statistical lineups to quantify the perception of graphical design choices (Hofmann et al., 2012; Loy et al., 2017, 2016; VanderPlas and Hofmann, 2017). Statistical lineups provide an elegant way of combining perception and statistical hypothesis testing through graphical experiments (Majumder et al., 2013; Vanderplas et al., 2020; Wickham et al., 2010).

The term 'lineup' is an analogy to police lineups in criminal investigations, where wit-

nesses identify the criminal from a group of individuals. Similarly, researchers present a statistical lineup plot consisting of smaller panels and ask the viewer to identify the panel that contains the actual data from a set of decoy null plots. Researchers generate null plots containing data generated according to a prespecified hypothesis using permutation or simulation. Typically, a statistical lineup consists of 20 panels, with one target panel and 19 null panels. If the viewer can identify the target panel from the null panels, it suggests that the actual data is visually distinct from the data generated under the null model.

While explicit graphical tests direct the participant to a specific feature of a plot to answer a particular question, implicit graphical tests require the user to identify both the purpose and function of the plot in order to evaluate the plots shown. Furthermore, implicit graphical tests, such as lineups, simultaneously test for multiple visual features, including outliers, clusters, and linear and nonlinear relationships (VanderPlas and Hofmann, 2015). Researchers can collect responses from multiple viewers using crowd-sourcing websites such as Prolific and Amazon Mechanical Turk.

In this paper, our primary focus is to evaluate the benefits and drawbacks of using log scales, specifically delving into their impact on perceptual sensitivity towards the degree of curvature. To address this, we conducted a visual inference experiment employing statistical lineups (Buja et al., 2009). Although our findings could have broad applications to various functions resulting in curvature, our experiment deliberately centered on participants' ability to identify differences in the curvature of exponentially increasing curves when presented with both linear and log scales. We discuss the nuances and challenges of testing the perception of exponential growth in the appendix. Importantly, this investigation did not necessitate participants to undergo mathematical training or possess

a prior understanding of exponential growth or logarithmic scales. Instead, it aimed to unravel the inherent ability to identify differences in curvature within charts, focusing on the fundamental nature of visual perception.

In Section 2 we describe the participant sample, the graphical task, data generation process, and study design. Section 3 describes the participant data collected and shares results from the statistical analyses of the data using a generalized linear mixed model. We present overall conclusions and discussion of the results in Section 4, and provide an overview of future related papers. The Supplementary Material includes a link to the RShiny data collection applet, participant data used for analysis, and code to replicate the analysis. The results of this study lay the groundwork for further exploration of the implications of using log scales in data visualization.

2   Study Development and Methods

2.1   Data Generation

In this study, we simulated data from an exponential model to generate the target and null data sets; the models between panels differ in the parameter values selected for the null and target panels. In order to guarantee the simulated data spans the same domain and range of values for each statistical lineup panel, we began with a domain constraint of $x \in [0, 20]$ and a range constraint of $y \in [10, 100]$ with $N = 50$ points randomly assigned throughout the domain. We mapped the randomly generated x values to a corresponding y value based on an exponential model with predetermined parameter values and multiplicative random errors to simulate the response. These constraints assure that participants who select the target panel are doing so because of their visual perception differentiating between

7

curvature or growth rate rather than different starting or ending values.

We simulated data based on a three-parameter exponential model with multiplicative errors:

$$y_i = \alpha \cdot e^{\beta \cdot x_i + \epsilon_i} + \theta \tag{1}$$

with $\epsilon_i \sim N(0, \sigma^2)$.

The parameters α and θ were adjusted based on β and σ^2 to guarantee the range and domain constraints are met. The model generated $N = 50$ points $(x_i, y_i), i = 1, ..., N$ where x and y have an increasing exponential relationship. The heuristic data generation procedure is described in Algorithm 1 and Algorithm 2.

Algorithm 1 Lineup Parameter Estimation

- **Input Parameters:** domain $x \in [0, 20]$, range $y \in [10, 100]$, midpoint x_{mid}.

- **Output Parameters:** estimated model parameters $\hat{\alpha}, \hat{\beta}, \hat{\theta}$.

1: In order to obtain the two middle points (total of four points for estimating three parameters), determine the $y = -x$ line scaled to fit the assigned domain and range.

2: Map the values $x_{mid} - 0.1$ and $x_{mid} + 0.1$ to the $y = -x$ line for the two additional points.

3: From the set of points (x_k, y_k) for $k = 1, 2, 3, 4$, calculate the coefficients from the linear regression model $\ln(y_k) = b_0 + b_1 x_k$ to obtain starting values for $\alpha_0 = e^{b_0}, \beta_0 = b_1, \theta_0 = 0.5 \cdot \min(y)$

4: Using the `nls` function from the base `stats` package in Rstudio [@Rstudio] and the starting parameter values - $\alpha_0, \beta_0, \theta_0$ - fit the nonlinear model, $y_k = \alpha \cdot e^{\beta \cdot x_k} + \theta$ to get estimated parameter values for $\hat{\alpha}, \hat{\beta}, \hat{\theta}$.

Algorithm 2 Lineup Exponential Data Simulation

- **Input Parameters:** sample size $N = 50$, estimated parameters $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\theta}$, from Algorithm 1, and standard deviation σ from the exponential curve.

- **Output Parameters:** N points, in the form of vectors $\mathbf{x}$ and $\mathbf{y}$.

1: Generate $\tilde{x}_j, j = 1, ..., \frac{3}{4}N$ as a sequence of evenly spaced points in $[0, 20]$. This ensures the full domain of x is used, fulfilling the constraints of spanning the same domain and range for each parameter combination.

2: Obtain $\tilde{x}_i, i = 1, ..., N$ by sampling $N = 50$ values from the set of $\tilde{x}_j$ values. This guarantees some variability and potential clustering in the exponential growth curve disrupting the perception due to continuity of points.

3: Obtain the final x_i values by jittering $\tilde{x}_i$.

4: Calculate $\tilde{\alpha} = \frac{\hat{\alpha}}{e^{\sigma^2/2}}$. This ensures that the range of simulated values for different standard deviation parameters has an equal expected value for a given rate of change due to the non-constant variance across the domain.

5: Generate $y_i = \tilde{\alpha} \cdot e^{\hat{\beta}x_i + e_i} + \hat{\theta}$ where $e_i \sim N(0, \sigma^2)$.

Comparison of Densities for the Lack of Fit Statistic
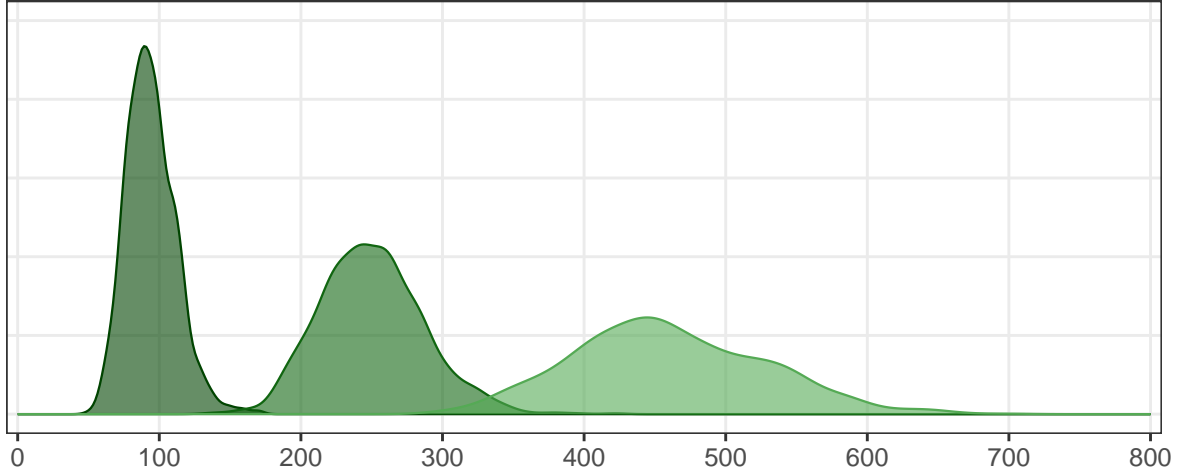between **High**, **Medium**, and **Low** Curvature Levels

Figure 3: Density plot of the lack of fit statistic showing separation of difficulty levels: obvious curvature, noticable curvature, and almost linear.

2.2 Parameter Selection

We chose three levels of trend curvature (low curvature, medium curvature, and high curvature). For each curvature level, we simulated 1,000 data sets of (x_{ij}, y_{ij}) points for $i = 1, ..., 50$ increments of x-values and replicated $j = 1, ..., 10$ corresponding y-values per x-value. Each generated x_i point from Algorithm 2 was replicated ten times. We fit a linear regression model on each of the individual data sets and computed the lack of fit statistic (LOF) which measures the deviation of the data from the linear regression model. After obtaining the LOF statistic for each level of curvature, we evaluated the density plots (Fig. 3) to provide a metric for differentiating between the curvature levels and thus detecting the target plot. While the LOF statistic provides a numerical value for discriminating between the difficulty levels, it cannot be directly related to the perceptual discriminability; it serves primarily as an approximation to ensure that we are testing parameters at several distinct curvature levels. Table 1 lists the final parameters used for data simulation.

Table 1: Lineup data simulation final parameters

Curvature Level	x_{mid}	$\hat{\alpha}$	$\tilde{\alpha}$	$\hat{\beta}$	$\hat{\theta}$	$\hat{\sigma}$
High	14.5	0.91	0.88	0.23	9.10	0.25
Medium	13.0	6.86	6.82	0.13	3.14	0.12
Low	11.5	37.26	37.22	0.06	-27.26	0.05

2.3 Lineup Setup

To generate the small multiple scatter plots for the statistical lineups shown to participants in the study, we simulated a single data set corresponding to curvature level A for the target plot and multiple data sets corresponding to curvature level B for the null plots. The `nullabor` package in R (Buja et al., 2009) randomly assigned the target plot to one of the panels surrounded by panels containing null plots. The target and null panels span a similar domain and range due to the implemented constraints when simulating the data; the rationale for this decision is based on preattentive feature perception (Wolfe and Utochkin, 2019) and is discussed in detail in the appendix.

There were a total of six lineup curvature combinations; Fig. 4 illustrates the six lineup curvature combinations (top: linear scale; bottom: log scale) where the solid line indicates the curvature level designated to the target plot while the dashed line indicates the curvature level assigned to the null plots. Two sets of each lineup curvature combination were simulated (a total of twelve test data sets) and plotted on both the linear scale and the log scale (24 test lineup plots). In addition, three curvature combinations generated homogeneous "Rorschach" lineups, where all panels were from the same distribution. Each participant evaluated one "Rorschach" lineup. Results from the "Rorschach" evaluations
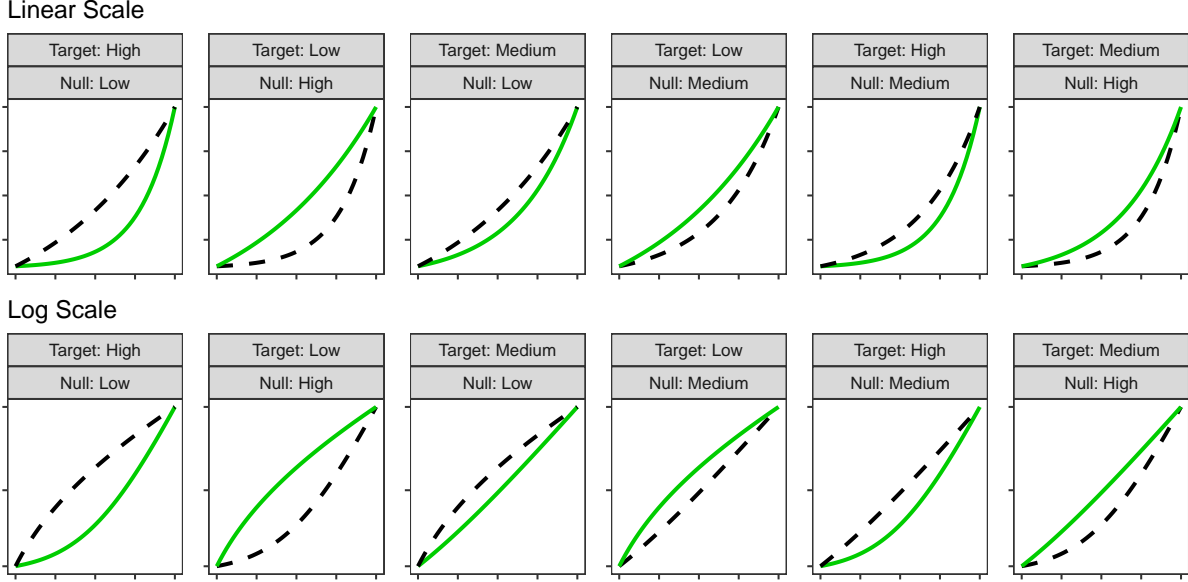
Linear Scale

Log Scale

Figure 4: Thumbnail plots illustrating the six curvature combinations displayed on both scales (linear and log). The solid line indicates the curvature level to be identified as the target plot from amongst a set of null plots with the curvature level indicated by the dashed line.

indicate null panel selections were distributed relatively evenly with multiple candidates for the most interesting panel. We display and further discuss the "Rorschach" evaluation results in the appendix.

Fig. 5 presents examples of statistical lineups with the target data simulated with exponential parameters corresponding high curvature and the surrounding null panels simulated with parameters for low curvature. The statistical lineup on the left presents increasing exponential data with displayed on a linear scale with panel 13 as the target panel. The lineup on the right shows increasing exponential data plotted on a log scale with panel 4
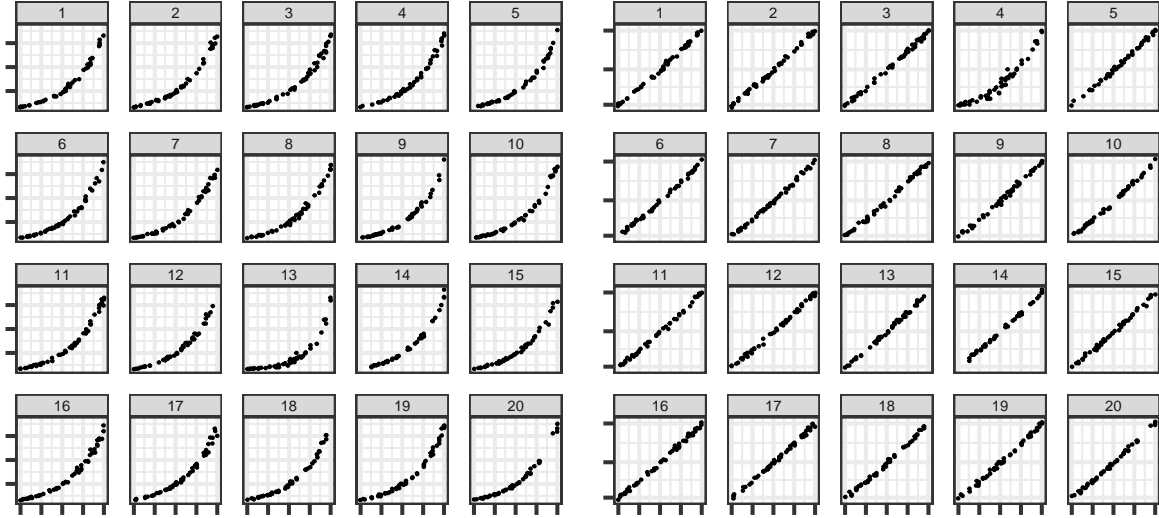
Figure 5: The lineup plot on the left displays increasing exponential data on a linear scale with panel (13 as the target. The lineup plot on the right displays increasing exponential data on the log scale with panel 4 as the target.

as the target panel.

References

Buja, A., Cook, D., Hofmann, H., Lawrence, M., Lee, E.-K., Swayne, D., and Wickham, H. (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.

Burn-Murdoch, J., Nevitt, C., Tilford, C., Rininsland, A., Kao, J., Elliott, O., Lewis, E., Fox, B., and Stabe, M. (2020). Coronavirus tracked: Has the epidemic peaked near you?

Cleveland, W. and McGill, R. (1985). Graphical perception and graphical methods for analyzing scientific data. *Science*, 229(4716):828–833.

Fagen-Ulmschneider, W. (2020). 91-divoc.

Heckler, A., Mikula, B., and Rosenblatt, R. (2013). Student accuracy in reading logarithmic plots: The problem and how to fix it. In *2013 IEEE Frontiers in Education Conference (FIE)*, pages 1066–1071. IEEE.

Hofmann, H., Follett, L., Majumder, M., and Cook, D. (2012). Graphical tests for power comparison of competing designs. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2441–2448.

Jones, G. (1977). Polynomial perception of exponential growth. *Perception & Psychophysics*.

Loy, A., Follett, L., and Hofmann, H. (2016). Variations of q–q plots: The power of our eyes! *The American Statistician*, 70(2):202–214.

Loy, A., Hofmann, H., and Cook, D. (2017). Model choice and diagnostics for linear mixed-effects models using statistics on street corners. *Journal of Computational and Graphical Statistics*, 26(3):478–492.

MacKinnon, A. and Wearing, A. (1991). Feedback and the forecasting of exponential change. *Acta Psychologica*, 76(2):177–191.

Majumder, M., Hofmann, H., and Cook, D. (2013). Validation of visual statistical inference, applied to linear models. *Journal of the American Statistical Association*, 108(503):942–956.

Menge, D., MacPherson, A., Bytnerowicz, T., Quebbeman, A., Schwartz, N., Taylor, B., and Wolf, A. (2018). Logarithmic scales in ecological data presentation may cause misinterpretation. *Nature Ecology & Evolution*, 2(9):1393–1402.

Vanderplas, S., Cook, D., and Hofmann, H. (2020). Testing statistical charts: What makes a good graph? *Annual Review of Statistics and Its Application*, 7:61–88.

VanderPlas, S. and Hofmann, H. (2015). Spatial reasoning and data displays. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):459–468.

VanderPlas, S. and Hofmann, H. (2017). Clusters beat trend!? testing feature hierarchy in statistical graphics. *Journal of Computational and Graphical Statistics*, 26(2):231–242.

Von Bergmann, J. (2021). xkcd exponential: Public health vs scientists.

Waddell, W. (2005). Comparisons of thresholds for carcinogenicity on linear and logarithmic dosage scales. *Human & Experimental Toxicology*, 24(6):325–332.

Wagenaar, W. and Sagaria, S. (1975). Misperception of exponential growth. *Perception & Psychophysics*, 18(6):416–422.

Wickham, H., Cook, D., Hofmann, H., and Buja, A. (2010). Graphical inference for infovis. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):973–979.

Wolfe, J. M. and Utochkin, I. S. (2019). What is a preattentive feature? *Current opinion in psychology*, 29:19–26.