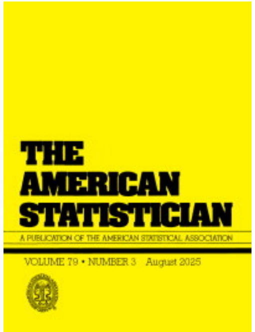

The American Statistician

ISSN: 0003-1305 (Print) 1537-2731 (Online) Journal homepage: www.tandfonline.com/journals/utas20

The Storyboard: A Tool to Synthesize, Reflect On, and Write About Data Investigations

Sara Stoudt & Deborah Nolan

THE AMERICAN STATISTICIAN
2025, VOL. 00, NO. 0, 1–9: Teacher's Corner
https://doi.org/10.1080/00031305.2025.2539997

The Storyboard: A Tool to Synthesize, Reflect On, and Write About Data Investigations

Sara Stoudt[a] and Deborah Nolan[b]

[a]Department of Mathematics, Bucknell University, Lewisburg, PA; [b]Department of Statistics, University of California, Berkeley, Berkeley, CA

ABSTRACT
Statistics and data science require both analytical and communication skills, and educators are faced with training students in both. These skills intersect when students work on data investigations and need to communicate their findings, such as in a class project. The transition from analysis to communication can be a challenge for students. We propose a storyboarding framework that helps students find the narrative in their investigation, choose appropriate details to present, organize their findings, and iterate to further strengthen both their data story and their statistical analysis. We discuss different implementations of this framework and provide guidance for approaching the activity in a variety of classroom settings.

ARTICLE HISTORY
Received April 2024
Accepted July 2025

KEYWORDS
Data communication; Investigative data analysis; Writing pedagogy

1. Introduction

The statistics and data science fields require both analytical and communication skills. Whether or not these skills are learned separately, they must eventually come together when students investigate data and communicate their findings. Nolan and Stoudt take a broad approach to teaching communication with data (Nolan and Stoudt 2022). They cover a spectrum from reading to prepare to write to editing and refining an argument, and they provide specific exercises to help students develop and practice communication skills. Of particular importance is teaching students how to develop and practice communication and analytical skills together. Storyboarding is a tool that we have found particularly effective at both making sense of findings and figuring out how best to present them.

We describe the concept of a storyboard, which we have adapted, building on an activity in Sara ElShafie's science communication workshop (ElShafie 2017), to the data analysis setting. Storyboarding can address the difficult task of stepping back from "doing" a data investigation and moving toward "sharing" the findings of the investigation. The storyboard provides a tool for students to synthesize and reflect on their findings and, as Pfannkuch et al. describe, "unlock the data story" (Pfannkuch et al. 2010). In the process, students also might find limitations or gaps they need to address in the next iteration of an analysis. This tool supports students in the writing process and provides a concrete set of steps to help them find and organize their story before they start to write.

In this article, we explain the pedagogical approach to storyboarding, provide examples, and discuss variations on the use of the tool in different classroom settings.

2. Storyboarding Basics

Traditionally, a storyboard for a movie visually displays the path of a story with a series of panels. These panels can depict characters, action, and scenery, and they are ordered according to how the plot unfolds. Storyboards are typically used in the entertainment industry to sketch out scenes on the way to developing a compelling story (Carnahan 2013; (The Credits) 2016). However, storyboarding has also been repurposed in design fields to represent different user experiences and as a helpful tool for planning "scrollytelling" journalistic pieces, an approach that walks readers through a story, zooming in and out of text and other media (Quesenbery and Brooks 2010; Seyser and Zeiller 2018; Asgari and Hurtut 2024).

Here we borrow the storyboard framework and place it in the context of a data story where each panel contains a graph or statistical table, panels are ordered in such a way that the findings of an exploratory or more formal analysis are organized and synthesized. Note that a table may hold the summary of a fitted model, such as parameter values and standard errors, results of hypothesis tests, model prediction comparisons, or simple summary statistics. We first outline the strategy for creating a storyboard that has proved useful in our classrooms.

We present storyboarding through a series of six concrete steps to help students gain aptitude in translating their statistical investigation into a compelling narrative. At this point, the students have completed, or nearly completed, an exploratory or formal analysis, and the storyboard offers an opportunity to reflect and organize their thoughts. Our directions to the students follow:

CONTACT Sara Stoudt sas072@bucknell.edu 380 Olin Science Center, Department of Mathematics and Statistics, Bucknell University, Lewisburg, PA 17837.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/TAS.

1. *Collect tables and plots.* Gather all of the tables and plots you have made as you explored and analyzed the data. Note that these tables and plots do not yet need to be fully polished (e.g, with descriptive titles and optimized font and theme settings). This is an opportunity to organize the work you have done so far in a physical or electronic notebook. Place the tables and plots in a format that you can easily move around. For example, you can print them so you can move them around on a desk or tape them on a wall; or you can put each on a separate slide so you can easily shuffle them about on a digital whiteboard. Either way, the code for your analysis should be organized so that you can easily reproduce your results and figures and make changes as needed, including polishing elements that you plan to use in your final story.
2. *Group related output*. Identify the tables and plots that contain similar messages, group them together, and provide a brief written summary of each group. Some groups will naturally form, but it can be helpful to try a random rearrangement to see if any new groupings appear that were not obvious originally. It may be helpful to duplicate some plots and tables so that you can place them in more than one group.
3. *Make an argument; unlock the story*. Now that you have a more manageable number of elements to consider (groups rather than individual plots and tables) it can be easier to identify high-level ideas. What do your groups of findings tell you? Are there connections between groups? For example, does one group build on another, provide an alternative to another, or supplement another group? Decide which groups best support your argument and best emphasize the more important findings. It's okay if you find this step challenging. It takes practice!

As you consider each group, anchor your findings to the original question, the data source, and the modeling techniques used in the investigation. Ideally, these considerations have been taken into account as you carried out your analysis, before the point of summarizing the findings for a report. However, it's a good idea to double check because it can be easy for an analysis to veer from the original question, misuse the data, or overreach and make unsubstantiated claims. If you uncover an issue of this sort, now is a good time to address it.

4. *Choose representative tables and plots.* An important part of the communication process is to winnow down what is most relevant to your argument and story. Select the essential plots and tables parsimoniously. Remove redundant plots and tables and any that follow unnecessary paths. Can your core message be seen in the plots and tables that remain? It's okay to drop a group entirely or replace it with a short description of what you found, tried, or checked.
5. *Sequence the chosen tables and plots.* It can be tempting to present findings in the order that we completed them in, but a compelling data story is more often about the findings themselves and not the process. Examine the groups altogether and consider how each group relates to the others. The goal here is to identify a flow. Is there a temporal component between them that needs to be preserved? Does one group motivate another? Does a group offer supplemental information that is not core to your argument and can be placed out of sequence from when it was done? Be sure to also place in the sequence any brief description that has replaced a group.
6. *Add captions and transitions.* For each plot or table that remains, write a caption that explains the message you want your audience to take away (Stoudt and Nolan 2022). Captions should describe the findings. For example, a caption that reads "this is a barplot of ages," is uninformative compared to the more analytical statement like "this barplot of ages shows that our sample has many more older than younger people." Captions and transitions are needed to provide context to your audience. You have spent much more time with the material than your audience, so reconsider what you think is obvious in the context of a newcomer.

The diagram in Figure 1 demonstrates these six steps. To begin, there are many tables and plots (18 are shown in the figure), but only 5 remain at the end. The original collection is split into six groupings. After reflection, one group is dropped, a representative is chosen for each group, and the results are placed in a new order augmented with captions and transitions.

Students often begin a formal analysis by looking at summary statistics and simple plots to get a sense for what their data contains, or an Exploratory Data Analysis (EDA) might make up the entire data investigation. Either way, a storyboard can be a helpful tool for students to organize their thoughts at this exploratory stage. As an example, Figure 2 shows a storyboard made from an exploration of various characteristics of dog breeds (McCandless 2014; American Kennel Club 2021). Here, and elsewhere in this article, the data exploration and analyses were done in R while the creation of storyboards themselves were made in Google Slides, although more sophisticated interactive project-planning tools like Mural, Miro, or Figma could also be used. The complete collection of figures and tables that were produced during the data investigation but that did not appear in the final storyboard can be found in the supplementary materials. An overview of these, their groupings, their eventual inclusion/exclusion in the storyboard, and their sequencing is described in Table 1.

When students dig in further beyond an exploratory phase and build models and test hypotheses, they can use the storyboard as they finish their formal investigation and transition to communicating their work. At this stage, we want them to be parsimonious. Summarizing a formal analysis requires synthesis and reflection and can be challenging; students in a variety of course contexts can benefit by going through the storyboarding process before beginning to write a summary or report. We address the benefits of this process next.

3. The Storyboard and Core Competencies in Statistics

A data investigation project is an important part of many statistics and data science courses; it gives students an opportunity to apply their newly learned skills to a real-world dataset and answer a question that personally interests them (Ledolter 1995; Holcomb and Ruffer 2000; White 2019; Davidson 2023). There are many aspects to a data analysis project that range from

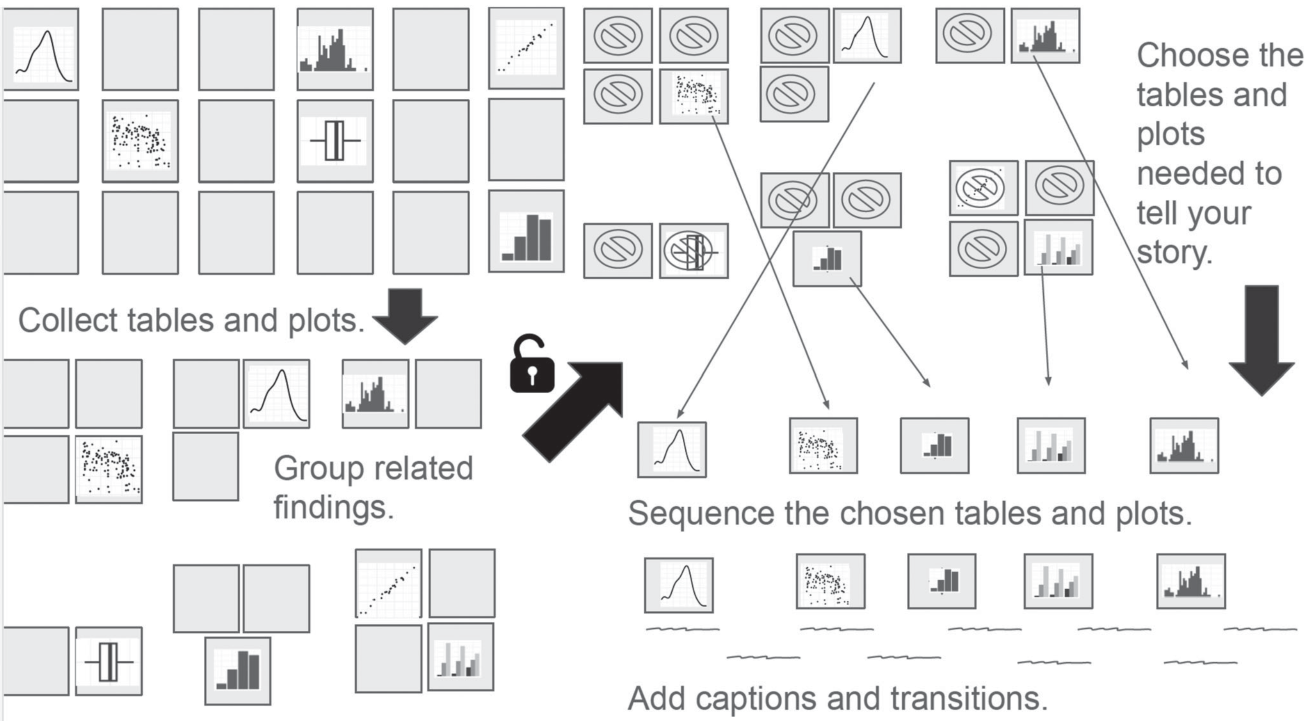

Figure 1. Diagram of the storyboarding process. Start in the top left corner where a storyboarder collects all tables and plots (step 1), continue downward to the bottom left corner where related findings are grouped together (step 2), unlock the data story (step 3) and continue to the upper right corner where the storyboarder chooses the elements needed to tell the story (step 4), and conclude in the bottom right corner where the chosen subset of tables and plots are sequenced (step 5) and captions and transitions are added to connect them (step 6).

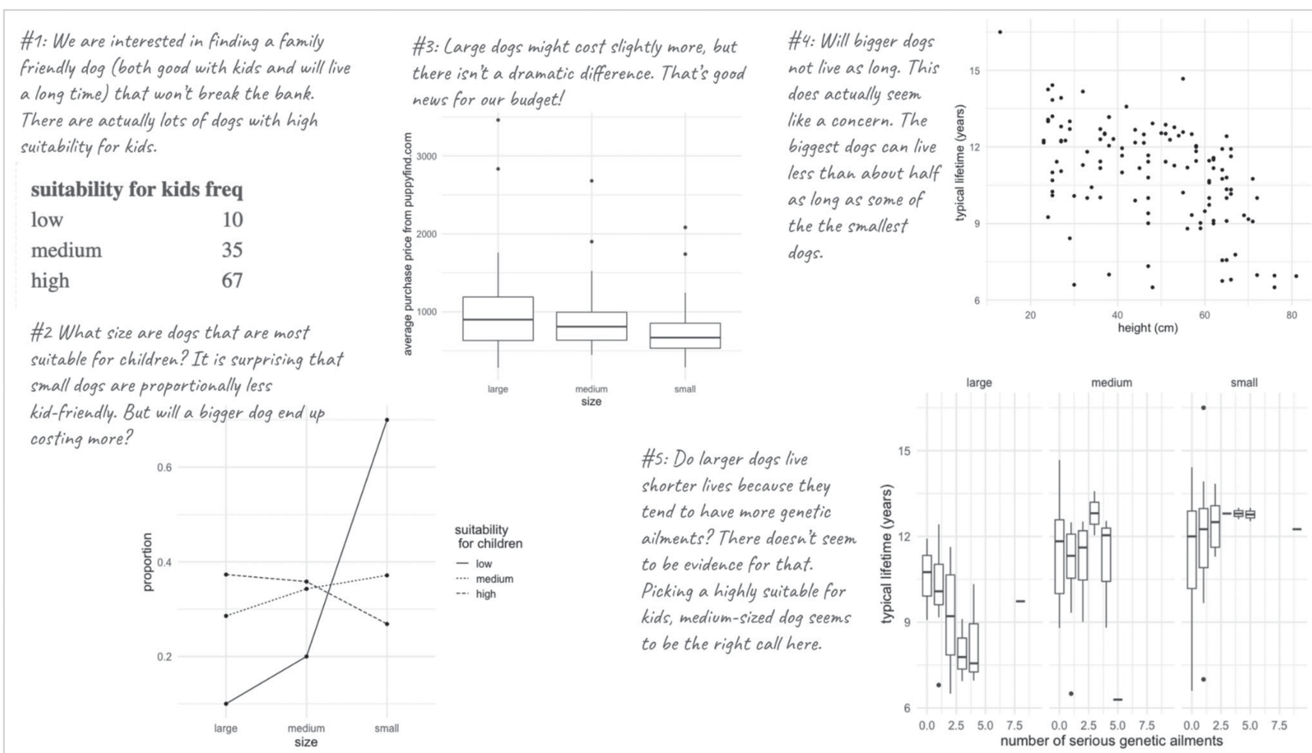

Figure 2. Storyboard for an exploratory analysis. This analysis is aimed at understanding what dog might be best for someone looking for a family-friendly, not too pricey, dog with a reasonable life span. The storyboard starts in the upper left corner with a summary table of a breeds' suitability for kids, continues to the bottom left corner showing the breakdown of suitability for kids by size, then goes to the upper middle with the relationship between size and price, proceeds to the upper right corner with the relationship between height and typical lifespan, and concludes in the bottom right corner with the relationship between lifespan and genetic ailments, broken down by size.

Table 1. Overview of plots and tables that were available to be included in the storyboard, how each was grouped, whether or not each was included in the storyboard, and the order in which it was included in the storyboard (if applicable).

Plot/Table	Group	Include	Order
Frequency table of suitability for kids	A	Y	1
Frequency table of number of repetitions for new command	B		
Frequency table of size	C		
Frequency table of group (dog type)	D		
Frequency table of number of genetic ailments	E		
Barchart of suitability for kids	A		
Barchart of suitability for kids for toy dogs	A, D		
Barchart of suitability for kids faceted by dog type	A, D		
Histogram of number of genetic ailments	E		
Density plot of height, grouped by dog size	C		
Density plot of weight, grouped by dog size	C		
Boxplot of height, grouped by dog size	C		
Boxplot of weight, grouped by dog size	C		
Scatterplot of height v. weight	C		
Clustered barchart of size and suitability for children	A, C		
Line chart of size and suitability for children	A, C	Y	2
Scatterplot of height v. typical lifetime	C, E	Y	4
Scatterplot of height v. typical lifetime with shape denoting repetition group	B, C, E		
Histogram of typical lifetime	E		
Density plot of typical lifetime	E		
Scatterplot of height v. typical lifetime, faceted by repetition group with lines of best fit and uncertainty bands	B, C, E		
Scatterplot of weight v. typical lifetime, faceted by size	C, E		
Scatterplot of average purchase price v. typical lifetime, faceted by size	C, E, F		
Histogram of typical lifetime per # of genetic ailments faceted by size	C, E	Y	5
Boxplot of average purchase price per size	C, F	Y	3

NOTE: Groups are: A = suitability for kids; B = number of repetitions to learn a new trick; C= size/height/weight; D= type of breed; E = number of genetic ailments/typical lifetime; F = price.

understanding the question of interest to data acquisition and cleaning to exploration and formal modeling. At some point, a data analysis project, whether it's for an introductory or upper-level course, transitions to a writing project. The storyboard aims at that transition.

Writing as a way for students to learn has been identified in a variety of disciplines and data communication has specifically been recognized as relevant and important to data-related pedagogy (Graham et al. 2020; Radke-Sharpe 1991). Approaches to teaching writing about data and statistics have taken many forms including asking students to read and critique scientific literature (Gelman and Nolan 2017; Samsa and Oddone 1994; Ye and Jin 2024), work through and present findings in case studies (Bennie and Erickson 2023; Khachatryan and Karst 2017; Kuiper and Sklar 2011; Nolan and Speed 2000), complete frequent, smaller writing assignments connected to course topics (Woodard et al. 2020; Stromberg and Ramanathan 1996), and complete writing workshops tailored to skills required of the profession (Hildreth et al. 2023).

However, the transition from analyzing data to communicating findings can be challenging for students and can often be put off until the last minute. In our experience, technical reports can be stilted and formulaic, especially if written by novices. The storyboard can scaffold a writing project; it is an organizational tool that neither follows the original sequence of investigations nor is overly prescriptive in format.

A storyboard activity helps break the barrier between the "numbers" and the "words" parts of the investigation. Statisticians and data scientists are also communicators and need to acquire this skill in addition to the technical coding and analysis ones. The more that students can gain fluency in organizing their work and structuring a narrative, the easier it will be for them to explain their findings accessibly and precisely to a variety of audiences in a variety of forms.

The storyboarding tool facilitates writing instruction as well as supports core competencies advocated for by the American Statistical Association's updated Guidelines for Assessment and Instruction in Statistics Education (GAISE) such as "statistics as an investigative process of problem-solving and decision making," "experience with multivariable thinking," and an emphasis on communication including being able to make and interpret graphs (Carver et al. 2016). Storyboarding addresses the GAISE core competencies and encourages students to go deeper as they work through a statistical investigation. Through storyboarding students grapple with all the work they have done and synthesize what that work means in the context of a particular problem. Additionally, the storyboard can be used to create intermediate checkpoints throughout the semester and to revise an analysis. Next we describe this reflection and revision process.

4. Reflection and Iteration with a Storyboard

Storyboarding can be effective as a tool for students to reflect on their data analysis, identify gaps, and propose a path forward to address them. This process matches the real-world one where a project is not one-and-done; there are often multiple stages of iteration and revision in a workflow before the project is completed. By showing that a storyboard is not a fixed object and that an initial storyboard often doesn't tell the final story, students experience firsthand that iteration is an important part of the statistical investigative process. This is especially useful as many students complete course projects at the end of a course, and so they rarely get the opportunity to revise their work based on instructor feedback.

The formal outline that students might be more familiar with at the first draft stage assumes you already know what you are going to say. Alternatively, storyboarding fosters reflection. It deviates from writing an outline and requires students to first grapple with their findings, decide what the interesting and defensible storyline is, and choose which material is crucial to telling that story. We have found that going through this process more than once before writing a first draft helps students avoid a formulaic report that details only the statistical analysis process in the order it was carried out.

Storyboarding is part of a cycle of discovery. When we ask students to reflect on their storyboard and revisit the steps of making a storyboard before launching into writing, they may come up with new ideas or identify missing pieces in their analysis. In our experience, students tend to "editorialize" as they are learning to go from basic descriptions to more sophisticated analyses, and it can help if they iterate on their storyboard. On a second pass through they might realize that some of their

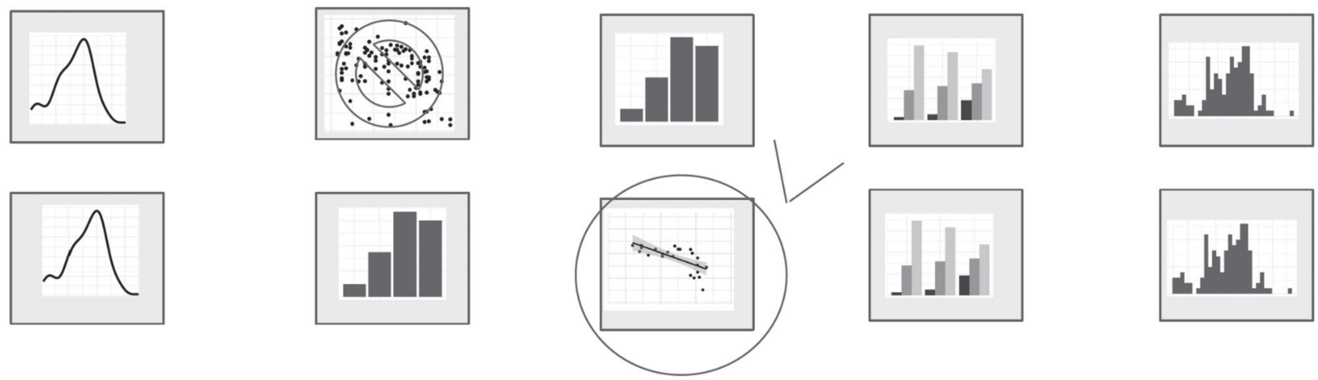

Figure 3. Diagram of the revision workflow for a storyboard. The top row of plots matches the ending point of Figure 1. The reflection has led to the removal of the second plot and insertion of a new figure in the middle. The revision appears in the bottom row. Captions and transitions also need to be updated.

statements aren't actually supported by their analysis. Some potential issues to check for in the second pass are tied to the connection to initial question, data design and collection, and model choice, as follows:

- Incorporation of the data design into the plots, summary statistics, and model fits (e.g., with sample weights)
- Appropriateness of the fitted model(s) for the kind of data collected
- Potential impact of data cleaning and exploration on findings (including initial mistakes or omissions in data cleaning steps)
- Interpretation of model in the context of the initial question
- Alignment of hypothesis test(s) with the data design and the scientific question
- Appropriateness of the model for prediction
- Generalizability of the findings beyond the observed data

When issues are identified, students often find they need to update their analysis and add to the storyboard. There are three basic approaches that we advise students to take to address newly found gaps in their story, as follows:

1. *Acknowledgement.* You don't always have to fill a gap, but at a minimum it needs to be acknowledged. This might take the form of a particular assumption, a point of discussion, or a topic of future work, and it might require updates to existing captions and transitions.
2. *Patchwork.* You may be able to fill the gap quickly by updating a plot, tweaking the analysis, or reading additional literature.
3. *Rework.* Some gaps take more time to address. You may need to try a different model or dig into a special case.

Whether a quick fix or a larger update, it's important to keep the code for the latest version of your findings (and storyboard) up to date. Here again, reproducibility is crucial. For example, one strategy is to organize the code needed to recreate and save each plot eligible for the storyboard (as we provide in the supplementary materials for the sample storyboards shown in this article) to streamline the process of revising as the story develops. With the iterative nature of storyboarding, incorporating version control tools like git can be particularly useful whether for individual assignments or when groups of students collaborate (Beckman et al. 2021).

Figure 3 provides a diagram that picks up from Figure 1 to show how a gap might be filled. Notice that addressing a gap may lead to the elimination of a figure as well as the addition of one. Some of the captions and transitions need to be rewritten as well.

As an example, the storyboard in Figure 4 aims to predict upward mobility using a single covariate (Chetty et al. 2014; Chetty et al. 2022). The model that is chosen as "best," using the fraction of 14- to 16-year-olds who work, makes a compromise between predictive accuracy and better meeting the condition for inference. However, some conditions for inference may still not be completely met, leading the storyboarder to acknowledge this and temper their conclusions accordingly.

An example of the patchwork approach to addressing a gap can be found in the supplemental materials. In that storyboard, the statistical investigator decided to find a second covariate to add to the model in order to account for some of the structure remaining in the residuals and thereby partially alleviate some of the concerns about meeting the conditions for inference.

Figure 5 shows a storyboard that is a result of a considerable rework of the statistical investigation. The storyboarder adds a third variable to the model, which identifies what region of the United States each observational unit is in, after determining that there is spatial structure in the residuals.

We have shown two examples of storyboards in different phases of an investigation. To help students first learn how to make a storyboard, we introduce it through an activity with prepared materials. We describe this approach next.

5. Learning How to Make a Storyboard Using Pre-Constructed Materials

Many educators advocate that students can investigate data without having made the graphs themselves (Hudiburgh and Garbinsky 2020; Engledowl and Weiland 2021; MacKay 2022; American Statistical Association and the New York Times 2024). Similarly, students don't need to make graphs to participate in an exploratory storyboarding activity. They can be provided with graphs and tables to investigate early on in a course and gain practice using the storyboarding tool.

Figure 4. Storyboard for a statistical analysis. This analysis is aimed at predicting absolute upward mobility using a single variable regression model. The storyboard starts in the upper left corner with a correlation matrix and a histogram of the response variable's distribution. It continues to the bottom left corner with a table of root mean squared errors of three, single variable models. Next, in the bottom middle are two scatterplots that show the multicollinearity of potential covariates. The storyboard concludes in the top right corner with a residual plot from the final model.

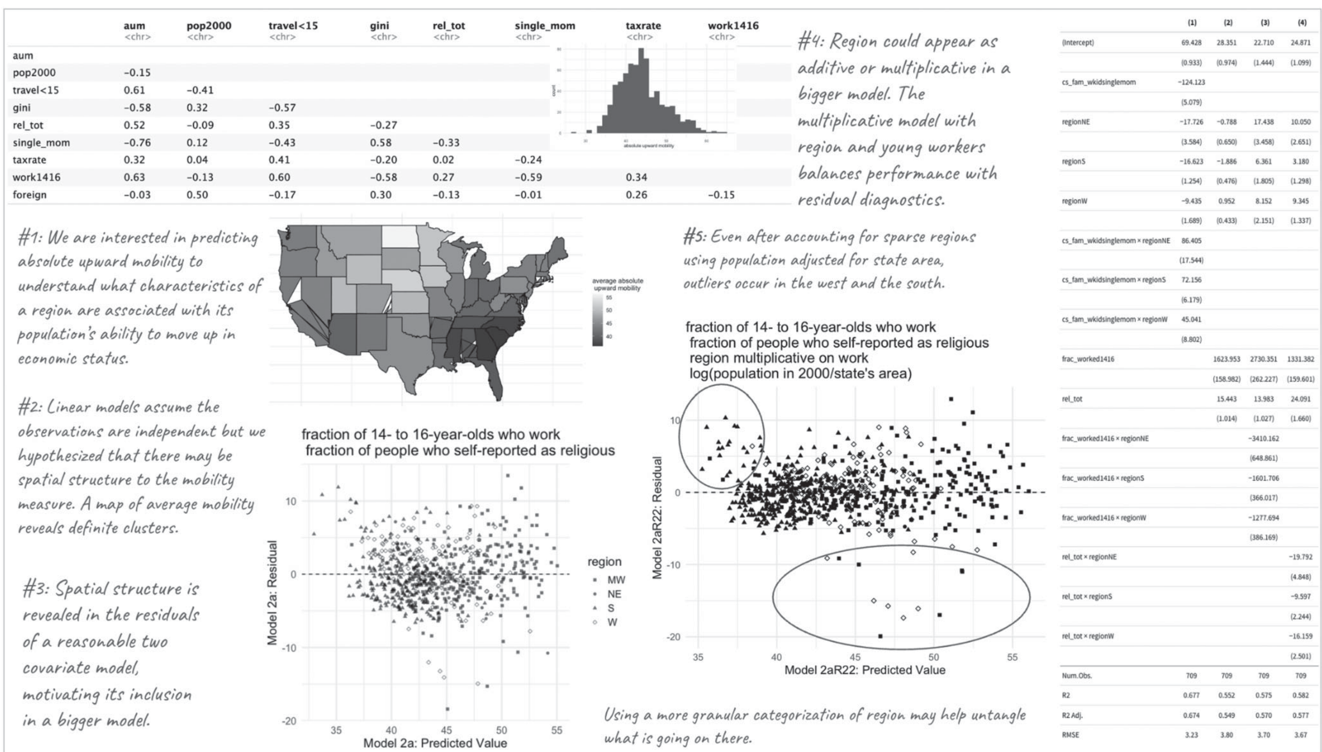

Figure 5. Storyboard for a statistical analysis. This analysis is aimed at predicting absolute upward mobility that reworked the analysis after finding spatial patterns in the residuals of the initially chosen model (Figure 4). The storyboard starts in the upper left corner with a correlation matrix and a histogram of the response variable's distribution. It continues from top to bottom on the left hand side with a map that shows spatial clustering of the response and the regional structure in the residuals of a model. The storyboard continues on the top right with summaries of models that include the region. The storyboard concludes in the bottom right corner with a residual plot with some annotations that motivate future work.

Providing pre-constructed materials for storyboarding has the advantage of making the activity accessible to students from many different backgrounds. We have used storyboarding in this way with high school and undergraduate students as well as with educators. By focusing on the interpretation of graphs rather than the logistics of creating them, students can learn to determine which graphs do and do not support a statistical argument or further the data's story. Pre-constructed materials can also be useful when students are first learning about different types of data and the visualization approaches that are most effective for the type (e.g., bar chart for categorical data, histogram for quantitative data). Alternatively, to help students through the struggle in uncovering a good story, they can be presented with materials and a story line and asked to choose which of the materials support that story. Later, they can move on to finding their own story lines.

Just as the choice of dataset for a class example is important, the choice of dataset for a storyboard activity is also important. Singer and Willett outline characteristics that make datasets particularly good for instructional use (Singer and Willett 1990). We especially echo characteristics such as "availability of multiple analyses," "background information," and "interest and relevance." Weiland and Williams advocate for including culturally relevant data to help students see themselves in the data and provide strategies for finding and using these datasets (Weiland and Williams 2024).

Ideally, when we use a common dataset for all students, we have them work in groups with a dataset that can reveal a variety of storylines. In this situation, we present students with different guiding themes to investigate, which then results in a diversity of stories. Like how Grimshaw advocates for moving away from curated datasets and empowering students to do the data wrangling on their own, we use the storyboard activity to move away from a curated point of view or storyline revealed by a dataset (Grimshaw 2015). Instead, students are challenged to go digging for their story.

We provide, in the supplementary materials, some sample datasets and pre-constructed materials that can be used in this way to help instructors get started. The building blocks for an exploratory storyboard include summary statistics and distribution plots, like histograms and barplots, of individual variables and scatterplots showing relationships between pairs of quantitative variables.

A storyboard activity where students use preexisting plots is also a good exercise in not overthinking and just jumping in, especially if the activity is time-limited. In this case, students often won't know the nuances of the data since they didn't produce the graphs themselves, but they can still make a first pass at a high-level story. If they are getting bogged down in the details, they can be reminded that further refinement would happen in the iterative stage of an analysis.

As students move toward multivariate thinking, more sophisticated plot types like time-series and subplots, or more complicated plots that display more than two variables can appear in storyboards. Here, pre-constructed materials may not be suitable for the task. Next, we discuss a variety of ways that storyboarding can be implemented in the classroom.

6. Storyboarding Throughout the Curriculum

A storyboarding activity is versatile and can be used throughout the statistics and data science curriculum. We have found it to be useful in classes of various sizes, heterogeneity of student experience, and levels of statistical content.

6.1. Class Size

In larger classes, storyboarding can help scale feedback. Peer review of a full written draft can seem daunting, but students can instead trade storyboards and be prompted to write a brief summary of the other student's storyboard. Then the pair can discuss if the summaries matched what the storyboard makers had in mind. Peer review of one another's storyboards can reveal gaps in logic before a formal report is written. This conversation also sparks a round of iteration. What could be changed in the storyboard to make an argument more clear? Feedback on storyboards can save students time by helping them get their story organized early on without requiring massive overhauls in writing later when the pressures of the end of the semester loom.

6.2. Student Background

Classes where the students come in with heterogeneous amounts of experience with data or code can be supported by the storyboarding process. Storyboarding of pre-made materials focused on exploration and description can be done early on while students get comfortable with coding and formal analysis. Then the class can transition toward more analytical storyboards using materials that are student-generated.

6.3. Placement of the Storyboard Activity in Data Projects

At its simplest, the storyboard helps students prepare the final written report for their data analysis project. In this situation, the storyboard starts in the middle, in the results section. After the key parts of an analysis to include in the report have been settled, the storyboard can be an effective tool to draft other sections. While a storyboard is an organizational strategy to distill and prioritize the results of an analysis, their connections, limitations, and assumptions, we can also use it to identify the key parts of a data description, recognize the ideas on which the findings rely (for the background material), prioritize main and secondary findings, and identify important assumptions and shortcomings (for the discussion).

Students can also go through the storyboarding process at multiple checkpoints throughout the semester, for example, once after the exploratory phase of the project and once after the initial formal analysis has been completed. To write their final report, students can build an outline that blends the two intermediate storyboards they have created, revising accordingly. This structure has the added benefit of helping students think about and make progress toward the communication phase of the project early on while getting feedback from the instructor.

6.4. Using the Storyboard for Other Assignments

The storyboard can also be useful in other course assignments that may not necessarily conclude with a formal report. For example, data visualization is increasingly emphasized early in the statistics curriculum, including the idea of bringing data visualization to "day one" of an introductory course (Wang et al. 2017; Gelman and Nolan 2017; Çetinkaya-Rundel and Ellison 2021). As students learn to make their own data visualizations, storyboarding can give them practice in moving from basic descriptions of what has been graphed to addressing what the figures are telling them about a motivating question (Datawrapper 2024; Hoogeveen et al. 2023). Sample instructions for using storyboarding in a course with a focus on visualization can be found in the supporting materials.

As another example, students and early career researchers in general are often given opportunities to participate in a poster session where they create a visual representation of their work that walks an observer through text and figures, accompanied by an "elevator pitch" to present to people who stop by their poster (Griffiths and Sheppard 2010; Brown 2020; Ghosh 2023). Since a poster is a mixture of limited text and compelling figures, a storyboard acts as a natural first draft.

Similarly, early career researchers are often given opportunities to give "speed," "snap," or "lightning" talks where they have a brief amount of time to talk about their findings (see, e.g., Lasser et al. 2020). The slide-based version of the storyboard lends itself naturally to an oral presentation. In fact, we routinely find ourselves storyboarding to write our own talks. We have all seen talks that fail to hook us and keep us engaged. We have found that iterative pruning of a storyboard can help us focus on one or two plots that convey the key finding, and those plots can form the foundation of an engaging talk. We do note that differing time constraints can affect the approach to writing a talk, including its level of detail (Stoudt 2023).

Pfannkuch et al. note a lack of opportunity for students to go beyond the description of their analysis process and "unlock" data stories (Pfannkuch et al. 2010). Whether a formal report, poster, lightning talk, or other written format such as a press release or blog, storyboarding can provide students with crucial opportunities to focus on crafting a narrative before they write a first draft.

7. Conclusion

Faculty who teach statistics and data science courses can benefit from teaching activities that support both analytical and communication skills at the same time, especially when faced with time constraints in covering content as well as skill-building in their classes. Storyboarding, in all its flexible variations of implementation across classroom settings, provides a way to both promote strong data investigation practices and strong writing practices. The activity pushes students to zoom out after being caught up in the details of a data investigation and provides them a framework to synthesize what they learned, reflect on the choices they have made in the analysis, and iterate on their approach to further refine their outputs.

Regardless of the statistical content of a course, storyboarding supports the writing process by cultivating good writing habits that are relevant for both writing in the discipline and for transferring to writing in general. These habits include parsimony in choosing what to write about, intentionality in the order and emphasis of details, descriptions of what is seen as well as what it means, strong transitions to support the flow of a narrative, and the value of revision.

The storyboarding activity can be used for both training data analysts *and* data communicators. With the storyboard, the instructor can support one skill and further hone the other.

Supplementary Materials

The R code and data used to produce the plots and tables used throughout the storyboard process for both the exploratory analysis and statistical analysis (the chosen plots and tables in Figures 4 and 5 as well as others that were discarded as part of the process) are included in the supplementary material. A variety of other sample storyboarding classroom materials are also provided.

Acknowledgments

The authors are grateful to the referee and associate editor for comments that led to an improved article.

Disclosure Statement

The authors report there are no competing interests to declare.

Data Availability Statement

The data used to create plots and tables for the storyboards in this article can be found on GitHub as part of the materials for the book Learning Data Science by Sam Lau, Joseph Gonzalez and Debroah Nolan (O'Reilly Media 2023). *https://github.com/DS-100/textbook/blob/master/content/datasets/akc.csv* (EDA) *https://github.com/DS-100/textbook/blob/master/content/ch/15/mobility.csv* (formal inference).

A supplementary dataset used in the formal inference storyboard is included in the supplementary materials.

ORCID

Sara Stoudt http://orcid.org/0000-0002-1693-8058

References

American Kennel Club. (2021), "American Kennel Club [dataset]. Available at *https://github.com/DS-100/textbook/blob/master/content/datasets/akc.csv* [2]

American Statistical Association & The New York Times. (2024), "What's Going On In This Graph?" available at *https://www.amstat.org/whats-going-on-in-this-graph* [5]

Asgari, M., and Hurtut, T. (2024), "A Design Language for Prototyping and Storyboarding Data-Driven Stories," *Applied Sciences*, 14, 1387. DOI:10.3390/app14041387 [1]

Beckman, M. D., Çetinkaya-Rundel, M., Horton, N. J., Rundel, C. W., Sullivan, A. J., and Tackett, M. (2021), "Implementing Version Control With Git and GitHub as a Learning Objective in Statistics and Data Science Courses," *Journal of Statistics and Data Science Education*, 29, S132–S144. DOI:10.1080/10691898.2020.1848485 [5]

Brown, J. A. L. (2020), "Producing Scientific Posters, Using Online Scientific Resources, Improves Applied Scientific Skills in Undergraduates," *Journal of Biological Education*, 54, 77–87. DOI:10.1080/00219266.2018.1546758 [8]

Carnahan, C. (2013), "Open Studio: Storyboards," available at *https://www.waltdisney.org/blog/open-studio-storyboards* [1]

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Holmes Rowell, G., Velleman, P., Witmer, J., and Wood, B.

(2016), "Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016." Available at *http://www.amstat.org/education/gaise* [4]

Çetinkaya-Rundel, M., and Ellison, V. (2021), "A Fresh Look at Introductory Data Science," *Journal of Statistics and Data Science Education*, 29, S16–S26. DOI:10.1080/10691898.2020.1804497 [8]

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014), "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 129, 1553–1623. DOI:10.1093/qje/qju022 [5]

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2022), " United States Mobility" [dataset]. *https://github.com/DS-100/textbook/blob/master/content/datasets/opportunity/mobility.csv* [5]

Datawrapper. (2024), "Enrich your stories with charts, maps, and tables." Available at *https://www.datawrapper.de/* [8]

ElShafie, S. (2017), "Data Science for the 21st Century Science Communication Short Course." Available at *https://www.sciencethroughstory.org/workshopsx* [1]

Engledowl, C., and Weiland, T. (2021), "Data (Mis)Representation and COVID-19: Leveraging Misleading Data Visualizations For Developing Statistical Literacy Across Grades 6–16," *Journal of Statistics and Data Science Education*, 29, 160–164. DOI:10.1080/26939169.2021.1915215 [5]

Gelman, A., and Nolan, D. (2017), *Teaching Statistics: A Bag of Tricks* (2nd ed.), Oxford: Oxford University Press. [4,8]

Ghosh, D. (2023), "Getting on the Program at JSM 2024." Available at *https://imstat.org/2023/12/15/getting-on-the-program-jsm2024/* [8]

Griffiths, P., and Sheppard, Z. (2010), "Assessing Statistical Thinking and Data Presentation Skills through the Use of a Poster Assignment with Real-World Data," in *Assessment Methods in Statistical Education: An International Perspective*, pp. 47–56, Chichester: Wiley. [8]

Grimshaw, S. D. (2015), "A Framework for Infusing Authentic Data Experiences Within Statistics Courses," *The American Statistician*, 69, 307–314. DOI:10.1080/00031305.2015.1081106 [7]

Hoogeveen, S., Sarafoglou, A., Aczel, B., Aditya, Y., Alayan, A. J., Allen, P. J., Altay, S., Alzahawi, S., Amir, Y., Anthony, F.-V., Kwame Appiah, O., Atkinson, Q. D., Baimel, A., Balkaya-Ince, M., Balsamo, M., Banker, S., Bartoš, F., Becerra, M., Beffara, B., Beitner, J., Bendixen, T., Berkessel, J. B., Berniûnas, R., Billet, M. I., Billingsley, J., Bortolini, T., Breitsohl, H., Bret, A., Brown, F. L., Brown, J., Brumbaugh, C. C., Buczny, J., Bulbulia, J., Caballero, S., Carlucci, L., Carmichael, C. L., Cattaneo, M. E. G. V., Charles, S. J., Claessens, S., Panagopoulos, M. C., Costa, A. B., Crone, D. L., Czoschke, S., Czymara, C., D'Urso, E. D., Dahlström, Ö., Rosa, A. D., Danielsson, H., De Ron, J., de Vries, Y. A., Dean, K. K., Dik, B. J., Disabato, D. J., Doherty, J. K., Draws, T., Drouhot, L., Dujmovic, M., Dunham, Y., Ebert, T., Edelsbrunner, P. A., Eerland, A., Elbaek, C. T., Farahmand, S., Farahmand, H., Farias, M., Feliccia, A. A., Fischer, K., Fischer, R., Fisher-Thompson, D., Francis, Z., Frick, S., Frisch, L. K., Geraldes, D., Gerdin, E., Geven, L., Ghasemi, O., Gielens, E., Gligoriæ, V., Hagel, K., Hajdu, N., Hamilton, H. R., Hamzah, I., Hanel, P. H. P., Hawk, C. E., K. Himawan, K., Holding, B. C., Homman, L. E., Ingendahl, M., Inkilä, H., Inman, M. L., Islam, C.-G., Isler, O., Izydorczyk, D., Jaeger, B., Johnson, K. A., Jong, J., Karl, J. A., Kaszubowski, E., Katz, B. A., Keefer, L. A., Kelchtermans, S., Kelly, J. M., Klein, R. A., Kleinberg, B., Knowles, M. L., Kołczyñska, M., Koller, D., Krasko, J., Kritzler, S., Krypotos, A.-M., Kyritsis, T., L. Landes, T., Laukenmann, R., Forsyth, G. A. L., Lazar, A., Lehman, B. J., Levy, N., Lo, R. F., Lodder, P., Lorenz, J., Łowicki, P., Ly, A. L., Maassen, E., Magyar-Russell, G. M., Maier, M., Marsh, D. R., Martinez, N., Martinie, M., Martoyo, I., Mason, S. E., Mauritsen, A. L., McAleer, P., McCauley, T., McCullough, M., McKay, R., McMahon, C. M., McNamara, A. A., Means, K. K., Mercier, B., Mitkidis, P., Monin, B., Moon, J. W., Moreau, D., Morgan, J., Murphy, J., Muscatt, G., Nägel, C., Nagy, T., Nalborczyk, L., Nilsonne, G., Noack, P., Norenzayan, A., Nuijten, M. B., Olsson-Collentine, A., Oviedo, L., Pavlov, Y. G., Pawelski, J. O., Pearson, H. I., Pedder, H., Peetz, H. K., Pinus, M., Pirutinsky, S., Polito, V., Porubanova, M., Poulin, M. J., Prenoveau, J. M., Prince, M. A., Protzko, J., Pryor, C., Purzycki, B. G., Qiu, L., Pütter, J. Q., Rabelo, A., Radell, M. L., Ramsay, J. E., Reid, G., J. Roberts, A., Luna, L. M. R., Ross, R. M., Roszak, P., Roy, N., Saarelainen, S.-M. K., Sasaki, J. Y., Schaumans, C., Schivinski, B., Schmitt, M. C., Schnitker, S. A., Schnuerch, M., Schreiner, M. R., Schüttengruber, V., Sebben, S., Segerstrom, S. C., Seryczyñska, B., Shjoedt, U., Simsek, M., Sleegers, W. W. A., Smith, E. R., Sowden, W. J., Späth, M., Spörlein, C., Stedden, W., Stoevenbelt, A. H., Stuber, S., Sulik, J., Suwartono, C., Syropoulos, S., Szaszi, B., Szecsi, P., Tappin, B. M., Tay, L., Thibault, R. T., Thompson, B., Thurn, C. M., Torralba, J., Tuthill, S. D., Ullein, A.-M., Van Aert, R. C. M., van Assen, M. A. L. M., Van Cappellen, P., van den Akker, O. R., Van der Cruyssen, I., Van der Noll, J., van Dongen, N. N. N., Van Lissa, C. J., van Mulukom, V., van Ravenzwaaij, D., van Zyl, C. J. J., Ann Vaughn, L., Veækalov, B., Verschuere, B., Vianello, M., Vilanova, F., Vishkin, A., Vogel, V., Vogelsmeier, L. V. D. E., Watanabe, S., White, C. J. M., Wiebels, K., Wiechert, S., Willett, Z. Z., Witkowiak, M., Witvliet, C. V. O., Wiwad, D., Wuyts, R., Xygalatas, D., Yang, X., Yeo, D. J., Yilmaz, O., Zarzeczna, N., Zhao, Y., Zijlmans, J., van Elk, M., and Wagenmakers, E.-J. (2023), "A Many-Analysts Approach to the Relation between Religiosity and Well-Being," *Religion, Brain & Behavior*, 13, 237–283. DOI:10.1080/2153599X.2022.2070255 [8]

Hudiburgh, L. M., and Garbinsky, D. (2020), "Data Visualization: Bringing Data to Life in an Introductory Statistics Course," *Journal of Statistics Education*, 28, 262–279. DOI:10.1080/10691898.2020.1796399 [5]

Kuiper, S., and Sklar, J. (2011), *Practicing Statistics: Guided Investigations for the Second Course*, Boston, MA: Pearson. [4]

MacKay, J. (2022), "Data Discovery Challenge Using the COVID-19 Data Portal from New Zealand," *Journal of Statistics and Data Science Education*, 30, 187–190. DOI:10.1080/26939169.2022.2058656 [5]

McCandless, D. (2014), "Best in Show: The Ultimate Data Dog," available at *https://informationisbeautiful.net/visualizations/best-in-show-whats-the-top-data-dog/* [2]

Nolan, D., and Speed, T. (2000), *Stat Labs: Mathematical Statistics through Applications*, New York: Springer. [4]

Nolan, D., and Stoudt, S. (2022), *Communicating with Data: The Art of Writing for Data Science*, Oxford: Oxford University Press. [1]

Pfannkuch, M., Regan, M., Wild, C., and Horton, N. J. (2010), "Telling Data Stories: Essential Dialogues for Comparative Reasoning," *Journal of Statistics Education*, 18, 9. DOI:10.1080/10691898.2010.11889479 [1,8]

Quesenbery, W., and Brooks, K. (2010), "Ways to Tell Stories," in *Storytelling for User Experience: Crafting Stories for Better Design*, New York: Rosenfeld Media. [1]

Seyser, D., and Zeiller, M. (2018), "Scrollytelling – An Analysis of Visual Storytelling in Online Journalism," in *22nd International Conference Information Visualisation (IV)*, pp. 401–406. DOI:10.1109/iV.2018.00075 [1]

Singer, J. D., and Willett, J. B. (1990), "Improving the Teaching of Applied Statistics: Putting the Data Back into Data Analysis," *The American Statistician*, 44, 223–230. DOI:10.1080/00031305.1990.10475726 [7]

Stoudt, S. (2023), "Can TV Make You a Better Stats Communicator?" *Significance*, 20, 28–31. [8]

Stoudt, S., and Nolan, D. (2022), "Captions: The Unsung Heroes of Data Communication," in *Proceedings of the IASE 2021 Satellite Conference. IASE 2021 Satellite Conference: Statistics Education in the Era of Data Science*. DOI:10.52041/iase.qythj [2]

(The Credits). (2016), "From Storyboard to Animation: See how Pixar Created Toy Story 2," available at *https://www.motionpictures.org/2016/10/storyboard-animation-see-how-pixar-created-toy-story-2/* [1]

Wang, X., Rush, C., and Horton, N. J. (2017), "Data Visualization on Day One: Bringing Big Ideas into Intro Stats Early and Often" (arXiv:1705.08544). arXiv. *http://arxiv.org/abs/1705.08544* [8]

Weiland, T., and Williams, I. (2024), "Culturally Relevant Data in Teaching Statistics and Data Science Courses," *Journal of Statistics and Data Science Education*, 32, 256–271. [7]