Integration of Data Visualization into the Legal Field

By Tahmina Ahmad

University of Colorado Boulder, College of Engineering & Applied Sciences

Department of Computer Science

Advised by Danielle Szafir, Casey Fiesler, and Sandra Ristovska

In partial fulfillment of the requirements for the degree of

Bachelor of Science in Computer Science, May 2022

Table of Contents

Abstract

While certain aspects of information and data science are rapidly being integrated into the field of law, the role that data visualization could play remains largely unexplored. Data visualization is a powerful tool that is not currently leveraged to its fullest potential in the legal sphere; it could be used to increase the understanding of data in various matters and lead individuals in legal settings towards improved decision making. In this paper, I explore how data visualization can aid data comprehension and improve the use of data in law. To do so, I investigate the inner workings of different cognitive biases, framing bias, in particular, to understand what enhances and diminishes decision-making processes and analytical skills. I also explore how it can act as a method of nonverbal persuasion and cover the advantages and potential disadvantages of using data vis. I apply this learned information about data visualization, framing bias, and their relationship to explore how it can be leveraged (and why it *should* be utilized) in the legal field by lawyers, attorneys, juries, judges, and others. Finally, I propose recommendations for creating effective data visualizations that can be used in the legal field for increased fairness and more confident court verdicts.

Introduction

Data visualization is the visual representation of information, and in the 21st century, it is essential for understanding data. When humans are presented with any format or arrangement of data, research shows that data visualizations assist them in understanding it much better. Studies by scientists like Lurie and Mason show that data visualization (otherwise known as data vis) and the associated tools help make decision-makers understand data much more comprehensively [1]. There are countless methods of taking information and converting it into a graphic, and each method can play a role in precisely how a data visualization is understood or interpreted. Beyond this, it is also possible for data visualization to add or remove bias, and this can present complications in legal settings such as courtrooms. Again, the way data is represented in a visualization can influence how misleading it is (or isn't).

Examples of data misrepresentation are ubiquitous. For example, during the COVID-19 pandemic, the rate of misinformation about COVID drastically increased [61]. To organize and cope with the large amounts of data produced as a result of COVID, data visualizations were created by different media and organizations. A study conducted by Lee et al. looked into over 41,000 visualizations and found that certain groups, such as anti-mask groups, often utilize data visualization to their advantage to convince people of certain perspectives, for example, that masks are ineffective in preventing the spread of COVID. Going a step further, they found that the same data from the CDC was taken by different groups and represented differently in visualization depending on the beliefs their group valued. For example, where mainstream American media and politics showed high rates of mortality across the United States (due to COVID), networks of anti-maskers circulated "their graphs to show that the pandemic is on its last legs (or that it was never a problem, to begin with)" [60]. Francesco Cafaro, an assistant professor at the University of Illinois says that people use visualizations to make decisions about their lives concerning COVID, like for example, "whether or not it's safe to send their kids back to school, whether or not it's safe to take a vacation, and where to go" [62]. Understanding how data visualization can be interpreted in different ways, and understanding that it can be biased and misleading is integral because visualizations are often used to make decisions concerning data. Bias can emerge in visualizations in many forms, such as through cherry-picking of data, strategic placement of text and embellishment, and various other ways. Anything that pushes the user to think differently about a visualization, whether it pushes them to understand the data better or to be persuaded of something within the data, is some sort of bias. As a result, different types of biases arise in these visualizations, such as framing bias.

It is important to note that data visualization is currently used heavily in litigation, to a point where it is overused to a certain extent [87]. For example, "computer-generated graphics can be overused to the extent where the witness, whose testimony is the paramount issue in jury decision-making, is overshadowed by videos and graphics" [87]. In some areas of law, such as law enforcement, data visualization is commonly used but it is often created without careful analysis and learned experience of how elements of its design can influence the ways in which people use it and interpret the data from it. Data visualization that is currently used may have any number of design features that could be considered mistakes. Or, they may not do their best to ensure maximum understanding of the data represented. For instance, data is not always presented in the form that best suits the type of data. Where one type of chart would be better suited for a dataset, an entirely different chart may be used [88] and this can hinder how well the data is understood and how sound the decisions made from it are. Lack of attention to such details can lead to biases that mislead people. However, manipulating the design and details of a data visualization can provide an advantage in how legal decisions are made and influenced. For example, adding saliency through color helps certain aspects of data to stand out; this could lead to something called anchoring bias, leading a viewer to depend on and believe the first piece of information they see [13], which would be the salient, colored details. The anchoring bias can then be used to sway one's decision making process. The intentionally incorporated saliency adds bias and utilizes it. Is this ethical? This question leads us to a necessary discussion about the ethics of the use and design of legal data visualization. In addition to the investigational aspect of the advantages and disadvantages of legal data visualization, a main aspect of this thesis is also making recommendations on how to use and design data visualization well. This will answer some of our motivating questions, such as what biases can arise as a result of data visualization in a legal setting and what ethical issues the use of data visualization can pose.

Importance of Data Visualization

Understanding

Research shows that people can "process more information when it is presented graphically than when it is presented in text form" [55]. When data, especially large amounts of it, is represented graphically, the human brain can understand more information more quickly than when it is written or conveyed verbally [1]. Data visualization allows viewers to visualize the data in a way that lets them not only understand it better but also identify patterns and trends within it when applicable [2]. Some of these detectable patterns would never have been picked up with standard statistical methods [3]. These representations are a great way to enhance trend detection, pattern comparison, and value interpolation [1]. This is because data visualization "supports simultaneous processing and is likely to lead to more intuitive and holistic, rather than piecemeal processing"; a piecemeal process happens gradually and at irregular intervals, and is likely inefficient [4] [5].

Visualization tools increase the rate of data analysis and make it easy for people to see correlations, outliers, and trends and to make comparisons [1]. In research conducted by Vessey in 1991, he explains that "although the same information is presented, graphic presentations enhance the evaluability of spatial information, whereas tables (of numbers) enhance the evaluability of symbolic information" [2]. They allow a person to process more information at once and thus enhance their problem-solving capacities [6]. We can observe that when data analysis is necessary, exploratory data visualization methods are used to represent what a data set is conveying. In exploratory data visualization, "visualization tools are used before and during the process of gathering and evaluating audit evidence to explore data relationships from various perspectives, discover new and meaningful patterns, and detect discontinuities, exceptions, and outliers…" [7]. Similarly, when in need of explaining the conclusions of analyses, explanatory data visualization techniques are utilized to explain certain trends, patterns, results, and more [7]. It is used to "synthesize and communicate auditors' main findings to convince the viewers of the auditors' conclusion" [1]. This could lead to changes in how data and visualizations are evaluated in ways like changing the speed and accuracy of decision processes. For example, in research studies done by Simkin and Hastie, they found that "discrimination was faster and comparison judgments were more accurate when bar charts were used, whereas proportion judgments were more accurate when pie charts were used" [10]. It could also affect their explanation skills when it comes to analysis processes; Lurie and Mason in their studies found that "decision-makers visualization tools may be less able to explain their choices than those who use text-based tools, for which particularly desirable or undesirable aspects are more easily identified" [1, 4]. Data visualization is an easy-to-integrate tool that enables easier pattern detection and analysis, and this could be leveraged in a variety of settings including legal ones where decision making processes are abundant and decisions are constantly made.

Vividness is another aspect that is important in terms of decision-making facilitated by data visualization. Vividness in visualization refers to the saliency or availability of specific information [46]. By looking at research from various papers, we see that "visualization tools are likely to affect vividness simply by presenting data in a form that uses preattentive graphic features, such as line orientation, width, length, and color" [48, 49, 50, 51]. Studies show more vivid visual information is, the "more likely it is to be acquired and processed before less vivid visual information" [45]. It also "leads to greater attention" [43, 44]. For example, the coloring of specific aspects makes certain parts of a bar graph salient and stand out; parts that would maybe otherwise be given less attention are highlighted. Thus vividness can aid in understanding data visualization better by allowing for more cognitive data processing and better attention towards data analysis through emphasizing specific parts of data. It can be leveraged in a legal setting by allowing individuals such as lawyers to push forward evidence and data points that more closely align with their desired outcome. On the other hand, the heightened focus that comes from increased vividness in a graphic may mean that other information gets ignored due to more weight being given to the salient aspects of a visualization [47]. They can often be misleading because they may lack context. Added text and captions are an example of a potential salient feature in a data vis. If there are captions regarding statistics without explanations of what they mean, they will stand out due to their saliency but they may mislead or confuse the viewer. Although saliency could be a great addition in legal data vis in order to boost understanding of the data presented, there could be potential drawbacks that it is important to be aware of.

Current Use of Data Visualization in Law

Use of Data in Legal Cases

Data is present in every sphere of the world today, including in the legal field. Whether it be in court cases, negotiations, litigation consultancy, or in other realms of law, data is crucial since law has so much to do with facts. People are driven and influenced in decision making by quantitative values. For example, statistics are regarded as particularly important in criminal justice, where for decades “police departments, states, and the U.S. Department of Justice have gathered and reported statistics on a wide variety of crimes” [83]. Furthermore, “an increasing number of jurisdictions are turning to data analytics…in 2016, 120 jurisdictions—covering over 91 million people—had signed onto the Data-Driven Justice Initiative, an ambitious U.S. Department of Justice program designed to increase the use of data in multiple areas of criminal justice” [83]. The introduction of programs like these exhibits how widely data is used in legal settings. In solely criminal justice alone, data is used for “response planning, crime prevention, criminal assessment, and risk assessment” [83]. These statistics can easily be put into data vis to boost data comprehension. Data is also heavily used in court settings although data visualization isn’t (yet). Former New Jersey attorney general Anne Milgram uses data and has even designed a data system “for determining whether a convicted criminal is likely or unlikely to pose a threat to public safety if released from jail or given probation [84]”. It is also reported that this data system “can help judges and parole boards make better assessments” about criminals and convicted individuals, like when they should be released from jail or put on probation [83]. Data is already used in courtrooms, like when forensic data is used to defend clients about their whereabouts, emails, physical location, and more. Since we know for a fact about the existence of legal data and the extensive use of data in law, it is inevitable that data visualization is already integrated into law to some extent albeit small.

Case Studies, Research, and Investigations

While data visualization has actively been used in many fields for decades, its presence in the legal field has been somewhat lacking. However, in recent years there has been a slight surge in its use of it in different areas within legal matters. David Waterfield and Mark Doughty of AlixPartners’ Law Firm recently incorporated data visualization into their firm, just as many other firms have. One example of the use of data visualization in law is for case evaluation and research [19]. Before a case is fought, there is a lot of research that goes into it. Vast amounts of data are often looked at by litigation attorneys and analysts as well as lawyers, case specialists, and others to truly study cases and legal matters. Investigations and disputes involve sifting through information by a legal team, and sometimes something noteworthy is only found when all facts are put together. Waterfield and Doughty note that when there is a data visualization that pieces the facts together, it is "transformative in terms of understanding the case history and where a particular item does or doesn't support the currently accepted narrative of the case" [19]. This means that when data vis is utilized in cases, it can help with pattern identification and analytics, as mentioned before.

Case evaluation and research also include interviews conducted by people involved in the legal matter. There are moments of human recollection that can be recorded from interviews, and digital footprints and traces recorded from devices such as times on security cameras or traffic cams. "Understanding where human recollections and the digital footprint diverge can be a useful proofing point" and solidifies the chronology of the case, meaning that testifying experts, lawyers, and others use "dashboards as an anchor for the conversation" to "prevent unnecessary arguments about the facts and chronology of the case" [19]. Chronology is very important in legal matters and timeline visualizations are often produced using emails and documents containing information such as timestamps, recipients, and senders to envision the order of events and increase case comprehension. The intense research done by legal teams for legal matters uses client databases and "subscription services like LexisNexis and WestLaw…[these] services offer indispensable access to vast databases of case documents" [21]). Due to the vast amounts of data, Waterfield and Doughty highlight that the best data for visualization is when it is " structured in a tabulated format…to filter the underlying records across particular dimensions – such as dates, specific people, geographic locations, business units" [19]. This filtering allows for easier evaluation and comparison.

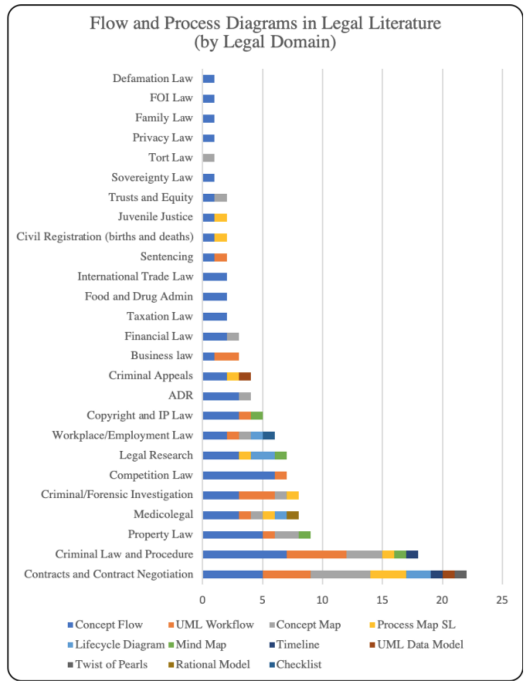
Courtrooms, Juries, and Judges

Another use of data visualization in law is in courtrooms. In courtroom settings, there are juries and judges present to listen to different cases and they are involved in various decision-making processes. Thus, they play a large role in evaluating evidence. A jury that is "bombarded with overly technical information is likely to tune out or become angry" during court proceedings, leading to clouded judgments and increased bias. This means that "in a courtroom, the ability to present evidence in a way that is accessible for a jury or judge is critical" [22]. Sometimes data visualization is presented to them to offer clarification; this is not a common practice right now but is increasing more as lawyers and attorneys become aware of the use of technical tools in law. Presenting information visualization to jurors and judges allows for the "courtroom narrative to be far more compelling and readily understood" [22]. During court proceedings, juries are often given data and even minimal charts and graphs as part of evidence presented to them [85]. However, more often, data is presented to them verbally and in an unstructured way.

Courtrooms are one of the most promising places to introduce and integrate data visualization. An article on courtroom forensic data explains that "the ability to present the data can make or break a case. Presenting…data in a clear way that explains complex terminology in an easy to understand way is crucial" [86]. Giving juries (and judges) intentionally designed data visualization can be so impactful. They would be able to understand the case and any data present in a clearer way, have visual aids to support their decision-making processes in the midst of long oral arguments and rebuttals in court and could be persuaded toward the desired outcome of the legal team. It is not just helpful in terms of understanding the case, but also can be empowering for the legal team and their client; where traditional court proceedings heavily rely on oral arguments and verbal persuasion, introducing data visualization to courtrooms can serve as a modern-day approach to court.

Importance

To reiterate, data visualization is currently used in a multitude of ways in the legal field, such as case building, evaluation, and investigation, as well as courtroom evidence for juries and judges as mentioned above. Its use goes beyond this and extends to contract matters, negotiations, forensic examinations, and more. The figure to the right from McLachlan & Webley [23] differentiates different sects of law and shows us which type of data visualization is used most often in that sect. At a glance, we see that contracts and contract negotiation involve the most data visualization out of a sample of 574 carefully picked legal documents; these documents were chosen by filtering a database for certain keywords, like “law”, “legal”, “process flow”, “process map” or “flowchart” [23]. Out of types of information visualization, we see that concept flows are the most common type of visualization seen in McLachlan and Webley’s study. It is important to look at how data visualization currently plays a role in law so that we can take a closer look at it and identify how it can play a role in legal decision-making processes. Beyond that, it is important because it can help us identify where and how data visualization can fit into the legal field in the future. Whether that be in case research, court, or any other legal setting, data visualization and the knowledge of framing and designing it has the potential to be greatly beneficial.



Bias in Data Visualization

Biases Caused By Data Visualization

Now, by taking a deeper look we see that data visualization can specifically affect different kinds of cognitive biases. Especially when employing data visualization in a setting as important as law, it's important to know how it could affect the decision making processes of those viewing the data vis. Dimara et al. outline 154 different cognitive biases in their paper that they categorize into seven different groups. Overconfidence bias is included because Dimara et al. discussed that “most often participants exhibit overconfidence, i.e., their self-rating is higher than their accuracy” [58]. The prevalence of overconfidence in most studies makes it a very common bias among people, thus making it integral to relate to the understanding of data visualization. Anchoring bias is one of the “seven biases that could affect visual analysis” according to Ellis et al., thus making it an important part of data visualization discussions [59]. Finally, we focus on framing bias as well because of its importance in data visualization in general. Dimaria et al. use framing effects as their first example to describe cognitive biases in their paper due to their prevalence and vast influence on data vis.

Overconfidence bias is when people are more confident in their abilities compared to the true extent of their abilities. Data visualization is easier to work with than plain text or data, and thus it can “increase decision makers’ confidence without increasing decision accuracy and potentially aggravate decision makers’ overconfidence bias” [11]. Anchoring bias, on the other hand, causes us to depend on and believe the first piece of information we are given on a topic simply because it was given first. Anchoring bias comes into play with data visualization because a visualization might have a certain aspect that stands out or there may be some initial decision about the visualization made upon looking at it; this is called saliency. Chang and Luo highlight in their work that “decision-makers tend to fixate on their initial estimate or expectation instead of sufficiently adjusting away from their initial anchor as they progress through the evidence-collection process” [12][13][14][15]. Since visual representations simplify pattern detection and have the ability to increase or decrease salience, the increase in focal information may improve decision-making quality to be better, faster, and more confident [52, 53, 54].

Finally, we take a closer look at framing bias, which is one of the most important and impactful biases when it comes to data visualization. Frames are “mental structures that decision-makers use, usually subconsciously, to simplify, organize and guide their understanding of a situation; these frames shape their perspectives and determine what information they see as relevant or irrelevant, important or unimportant” [18]. Data can be framed in a lot of different ways, and how a single visualization takes data and portrays it can change a lot about how it is interpreted. There are many studies done about framing bias in data visualization, such as one by Synovate in 2005 [20]. By looking into pie charts of healthy and unhealthy food, they learned that the green section on a pie chart will be perceived as healthy and the red section as unhealthy, even if it is not accurate. They also saw that the perceived differences are drastic between pie charts that are 75% green and 25% red, 25% green and 75% red, and black and white versions of the same information [1]. Hullman and Diakopoulos in their studies suggest visualization techniques (such as the use of color, saliency, type of vis, etc) that “prioritize particular interpretations…that tell a story [which] can significantly affect end-user interpretation” [81]. These examples show that the same data can be manipulated visually to be framed negatively or positively. Small choices like the type of graph, colors in the charts, sizes, shapes of visualization features, and more can influence a user's judgment.

Data vis is framed when you take the same data and present it in different ways, either to highlight something or take away attention from it, whether to increase understanding or decrease it. It can either be framed in a certain way intentionally or unintentionally by those who are unaware of how delicate the process of creating data visualization is because it can produce framing effects so easily. There are many choices that can be made about how data is visually represented; these can affect the decision making processes of those viewing it, whether they're to analyze the data, find extrema, or perform any number of other tasks that will be discussed in future sections. Through various examples, we can see how the same data can be framed in different ways to influence decision-making processes. For these examples, we'll be looking into several different studies that compare data representation in bar graphs, line graphs, pie charts, and more, as well as framing that occurs due to the use of colors, and embellishments, and more. In a study conducted by Saket et al [24], they analyzed the effects of different kinds of data visualization on different tasks like finding anomalies, clusters, correlations, exact values, extremums, trends, and more. Wu et al. also did similar work in analyzing types of data visualizations in terms of how well they are understood by people with and without intellectual and developmental disabilities. These studies, as well as a few others, will be used to build an argument and show how framing effects affect the interpretation of data visualization, and how this could play a role in law.

How Framing Bias Manifests in Data Visualization

Type of Visualization

The type of visualization chosen to represent data can affect exactly how it is framed, and thus how it is interpreted. Many studies commonly look at the interpretation differences between a bar chart and a line graph. Saket et al [24] found that the bar chart was the fastest and most accurate type of visualization because "people can decode values encoded with length faster than other encodings such as angle or volume". Especially when it comes to identifying extrema, most people tend to be significantly more accurate with bar charts compared to line charts [27]. Line charts were found to have the highest accuracy for correlation and distribution tasks because they are effective for trend finds, but overall for all other types of tasks line charts are the least effective and accurate [24]. Although some studies have shown that there is not a huge difference in accuracy between bar charts and line graphs for trend estimations [24], in general, there is a seeming consensus in the data visualization community that data expressed in bar charts is better for certain types of tasks like identification of extremum, while line charts are best to be avoided in most cases besides trend finding tasks like correlation and distribution analysis. Next, we look at pie charts, where one study showed that a "pie chart is comparably as accurate and fast as bar chart and table for retrieve, range, order, filter, extremum, derived and cluster tasks, [but] it is less accurate for correlation, anomalies and distribution tasks" [24]. Pie charts are easy to embellish and make even more usable by adding colors and text labels to show data values.

Another study found that in comparison to stacked bar charts and treemaps, pie charts are preferred for their ease of use, even though people seem to be more accurate when dealing with stacked bars over pie charts for proportion data.

When discussing types of data visualization and the accuracy that comes from each type of task and visualization, studies show that there is a "positive correlation between accuracy and user preference, indicating people have a preference for visualizations that allow them to accurately complete a task" [24]. This means that as much as the portrayal of the data and type of task may have to do with understanding the visualization, the user's preference may also play a key role in how accurately they interact with it. It is important to note that some contradictory studies found that if individuals have a preference for a certain type of visualization, it doesn't mean their results will always be more accurate using it. In a study that compared users' preferences between 2D and 3D data visualizations, they found that although the majority of the participants preferred 2D visualizations, for certain tasks they observed "participants being nearly twice as slow" in task performance. However, the individuals who preferred 3D visualizations did openly admit that the 2D vis would be much more efficient and easy to understand [63]. Thus, the format and way that data is visually represented plays a heavy role in how it is interpreted, and how decisions are made based on it.

Cherry Picking

In addition to data being framed to fit different types of charts like bar charts, pie charts, line graphs, and more, we must consider the case where data is being framed by exempting important portions from the visualization. This is called cherry-picking data. Sanchez [28] states that cherry-picking is a "common way of attempting to deceive audiences into thinking a trend exists (or not). If you need to drop values or observations, make sure this is mentioned and justified." Cherry-picking is often seen in the media, and it is dangerous because it is dishonest, unethical, and a deliberate attempt to leave information out. In a legal setting, regardless of how badly someone needs to be convinced of something, cherry-picking would be considered wrong. For example it could lead to a completely inaccurate narrative being presented in court and result in unjust decisions made by juries and judges. When dealing with cherry-picked data and data visualization, it is a form of framing bias since important information is exempted for some intent and ends up affecting the decision-making process of the user.

Color, Embellishments, and Text

The use of color, embellishments, and text could add framing effects to any data visualization. For example, the most basic example is that the color red is "often attached with an aggressive connotation" [28]. This was covered briefly in a previous section of this paper when we looked at Lurie and Mason's research and how they found that "by looking into pie charts of healthy and unhealthy food, they learned that the green section on a pie chart will be perceived as healthy and the red section as unhealthy, even if it is not accurate" [1]. This example shows that the same data can be manipulated visually to be framed negatively and positively, or to highlight certain information pieces over others.

Bright, bold colors are often paid more attention to and can be used to highlight certain pieces of information in visualization as opposed to data that is colored in more subtle colors or black and white. This is referred to as saliency. Healey et al. say that "one of the most important considerations for a visualization designer is deciding how to present information in a display without producing visual confusion" and they analyze how certain visual features may add saliency to data visualizations [56]. They learned that "feature hierarchies suggest that the most important data attributes should be displayed with the most salient visual features, to avoid situations where secondary data values mask the information the viewer wants to see" [56]. Through this information, we can see potential framing effects that can be brought on by color due to increased saliency.

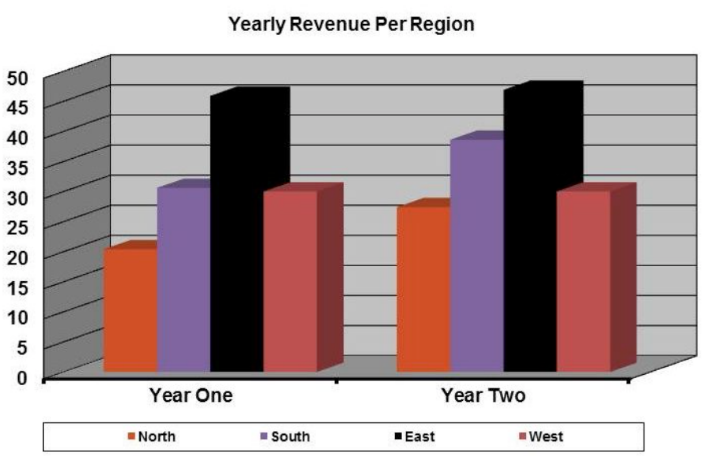
In another study (which was referred earlier) by Wu et al. [27], they look into "semantically meaningful embellishments (i.e., chart junk and icons) to enhance semantic reasoning and discretization to support working memory". By adding in embellishments, it is possible to add in framing bias through the selection and placement of pictorial cues, for example. These pictorial cues or extra images to embellish a given data vis would lead a reader to pay attention to certain details over others. Embellishments may include elements that "are not essential to understanding the data" [29], or more extremely, "chartjunk", which "seeks to attract and divert attention using display apparatus and ornament" [30]. This is most definitely a framing effect and affects the decision-making process of an individual. Finally, the placement of text is another detail to be aware of. Depending on where text is placed, what data is labeled versus not labeled, and whether the text is even present or not, this could frame a visualization in different ways. This pertains to the type of framing bias called attribute framing, which was previously mentioned. An apt example of this would be using the word "collision" instead of "accident" to frame it as more violent and intentional [26], or labeling only aspects of a company's growth and not labeling any company decline data in a company annual review chart.

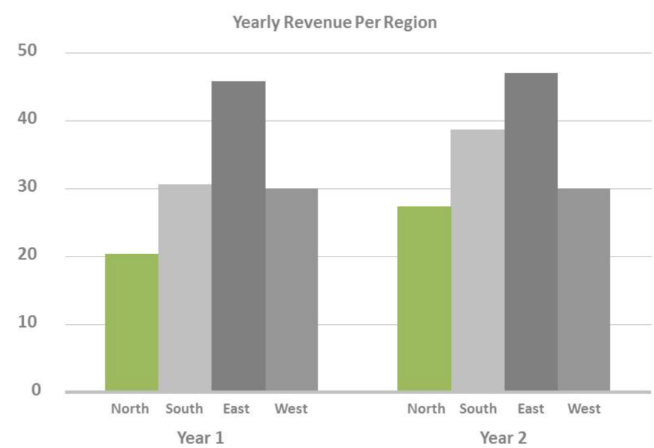
The negatives of using embellishments in data vis are supported by various literature, especially when it comes to chart junk. For example, Tufte, who popularized the term, says that the addition "seeks to attract and divert attention using display apparatus and ornament" [30]. Tufte believed that data vis should be extremely minimalist, and argued that they were unnecessary and complicated by making it more difficult to understand. The overcompensation can also lead to the introduction of different biases within the data vis.

If data vis creators are not mindful when adding in chart junk, it can throw off how the vis is interpreted and could come off as "polluting it with a whole bunch of nonsense" [33]. However, Parsons and Shukla describe that while in some contexts chart junk can be "unnecessary and can have costs [66], it can also have significant value for comprehension, especially with complex data" [33]. Color, embellishments, and text are essential to be aware of when working with data visualization as they can be counterproductive if not used well, but could serve as useful visual aids in certain circumstances.

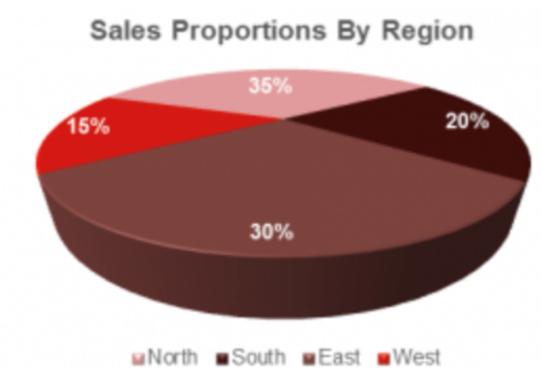
Unnecessary Details, Added Confusion, and 3D charts

The use of added unnecessary details can lead to misconstrued understanding. In an article from Harvard Business Review, Duarte reviews common data visualization mistakes and how data is sometimes simply framed poorly [31]. The figure below on the left shows a 3D chart that shows that the Year One North data point is a value under 20, like 19 or 18, when in fact the data says it represents 20.4. This graph is misleading because of how it is designed. By flattening that chart as Duarte does in the figure below to the right, we see that the conclusions can be drawn much more easily without being misleading. Oftentimes, individuals aren't even aware of the fact that 3D visualizations, for example, can have negative consequences for those viewing them. The figure to the left shows an example of how the use of a 3D chart can make things ambiguous [31].





Individuals viewing 3D data visualizations often aren’t aware of the negative consequences it can bring. In addition to adding ambiguity to visualization, it has additional drawbacks such as being directly misleading. The graphic to the right is a 3D pie chart, which completely misleads the data represented. At first glance, we can see that the category labeled “East " seems to be the largest part of the graph, but when we take a closer look we see that it's only 30%. This can be compared with the value towards the top portion of the graphic, which is the “North” category. This “North” category represents 35%, which is a larger percentage than the “East” category. However, due to its placement on the graphic, it looks significantly smaller than the “East” category; this is directly misleading. This can happen with various types of 3D charts, including pie charts, bar charts, and more. If your intent as someone in a legal setting was to make reading the chart confusing, however unethical that may be, you could consider using techniques like this to decrease understanding of the



graphic and attempt to persuade the reader towards something that would require hard to read data.

Summary

By analyzing these different ways to frame data, we see that specific details of visualization can affect what the user makes of the information. The effects of understanding data due to different framings in visualization hereby lead to influence decision making. However, it is important to keep in mind that in the context of this paper, we are looking into the use of visualization in legal matters. Data that would be used in legal settings would be transformed into data vis and presented in a way that a lawyer, attorney, expert witness, or anyone else may want to use for their purposes. In an ideal world, these purposes would just be for better data comprehension to achieve more just decisions and judgments. However, these purposes may include nefarious ones like persuasion, where in a court an attorney may want to use framing effects to their advantage to persuade a court to win the case. This could potentially lead to ethical concerns. Regardless, the important takeaway is that different framing techniques lead to different effects in data vis viewers, and by understanding what those are, certain standards can be created and recommendations are made to aid people in the field of law to utilize data visualization effectively..

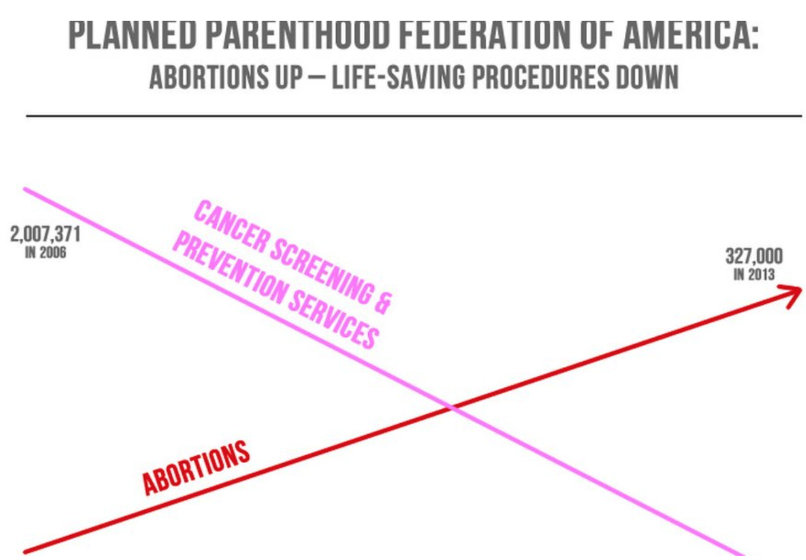
Framing Bias in Legal Settings

Ethical Tensions of Data Visualization Due to Framing Bias

It is important to talk about framing bias in legal settings because of how much the bias can influence decision-making. Legal work is deeply immersed in decision making, whether it be the decisions made by lawyers building cases, litigation experts doing legal research, attorneys overseeing deals and negotiations, jurors and judges in court, or anything else. These groups should be using data vis to improve their work. This way, they can anticipate certain framed narratives from the opposition and create data vis to support their counterarguments. The disadvantages of framing bias as they pertain to data visualization are that it can lead to biased and unfair decision-making. Some might consider this aspect of framing bias in data vis an advantage if they have nefarious intentions and are actively trying to perpetuate biased and unfair decision-making. For example, in their study of data visualizations via American Twitter users, Lee et al. saw that certain groups, like anti maskers and certain right-wing conservative groups in the United States, actively utilized cognitive biases (including framing bias) in the data visualization they tweeted and retweeted to persuade individuals of their particular beliefs [60]. This facilitated the spread of misinformation and shows how framing bias in the wrong hands can be dangerous and inaccurate. When attempting to avoid framing bias in data visualization and not use it to your advantage, the bias can be hard to eliminate. Thus, it can be hard to find neutrality when needed. In addition, it is necessary to be careful how something is framed because different framing can have different comprehension rates. The disadvantages could ultimately lead to less just decision-making. Anyone involved in the legal field, or any other field, is susceptible to framing bias. Whether we're looking at case studies, research, investigations, or juries and judges in courtrooms, everyone involved in these processes should be aware of the framing biases they can cause or fall prey to. By being aware of this, the disadvantages of framing bias can potentially be avoided. It can also be decided whether data visualization should even be used.

A Closer Look: Misleading Data Vis in Congress

By adding framing bias, visualization can be used persuasively. However, in adding in framing bias intentionally through any of the aforementioned ways, we are faced with an ethical dilemma: how far is too far? How extensively can data be framed a certain way? Where is the line between framing versus misleading and data misrepresentation? At what point does intentional framing become fabrication instead of data representation for comprehension/persuasion purposes? The ethics of using framing bias intentionally in data visualization can get tricky. An example of this is seen through a data visualization projected by Rep. Jason Chaffetz during a high-profile congressional hearing investigating Planned Parenthood [76] (pictured to the right). The visualization depicts a seemingly drastic decrease in cancer screenings/prevention services and a large increase in abortions. However, this graphic was created by Americans United for Life, which is an anti-abortion group that is trying to push anti-abortion agendas into Congress through Rep. Chaffetz. The graphic has no y axis, and, as data vis designer Stephen Tracy puts it, Chaffetz "took the data points for 2 nonadjacent years and stretched the data across an 8-year time frame making it appear that there has been a consistent downward/upward trend in either data set" and also used a "multi-axis chart and didn't convey this anywhere in the chart labeling" [80]. Upon viewing this visualization, we can circle back to the question of whether it is ethical for this graphic to be created and displayed in a legal



proceeding. The fact is that the visualization is entirely misleading and does nothing for data comprehension. It uses bad visualization techniques, where the creator "manipulated the presentation of the data to tell his own story" [80]. However, there is no doubt that the graphic has the purpose and potential of being persuasive. We can consider that our aim in this paper is not to discuss how to portray data as accurately as possible, but rather to make individuals in legal settings cognizant of the effects of framing bias in data vis so they can make an informed decision of *if* and *how* they want to use data vis for their purposes.

Disadvantages of Data Visualization Due to Framing Bias

The main disadvantage to using data vis is framing bias. There are instances where you'd rather not frame something a particular way but rather want to eliminate as much bias as possible. In these cases, where you are looking for a very unbiased piece of data visualization, it may be hard to take specific data and present it in a way that does not provoke bias. It is also necessary to be careful with framing something a certain way intentionally, because different framing can affect viewers in different ways which may backfire, leading the viewer to pay attention to or believe something other than what you wanted them to. For example, jurors are often common people without a law background, so intentional framing directed towards them needs to be done very carefully so data and information are not misconstrued [64]. A visualization could be framed in such a way that it unintentionally emphasizes irrelevant information. Curley and Munro stated in their paper that "previous research has highlighted that task-irrelevant contextual information influences judgments" [25]. This could easily cause a potentially unbiased and fair decision to be ruined because of accidental framing. In general "the ramifications of biased and unfair decision-making by jurors can result in injustice" so it is integral to remain cognizant of what bias a piece of data visualization can cause to whoever viewing it, especially to those who are making decisions [25].

Advantages of Data Visualization Due to Framing Bias

While framing bias in data visualization can potentially be unfair and disadvantageous, the possible bias created can be used positively to create and increase understanding and serve as a mechanism for persuasion. For instance, framing can highlight the important parts of data, cases, and more, which aid in increasing the understanding of whoever is looking at it. Emphasizing important features of visualization has great potential to improve the quality of evidence presented to jurors or the quality of an argument pieced together by attorneys. Thus, rather than being used to emphasize irrelevant information, the data visualization would be able to properly frame the important and relevant information, allowing jurors or judges to make decisions based solely on facts of the case. In addition to increased understanding, framing bias can play to advantage by being used for persuasion. For example, in a court setting, this could be used to sway a jury one way or another. Jurors' perceptions are influenced by several different things, such as “the type and complexity of evidence they are giving testimony on (e.g., eyewitness, footwear, DNA), the type of expertise they are communicating (e.g., clinical or actuarial), the characteristics of the expert (e.g., gender, appearance, attractiveness)”, and more [25]. Thus, utilizing the effects of framing bias through data visualization can add a push to jurors, for example, to make the decision you want them to amid all the other factors pushing them in different ways. Framing bias in data visualization that is used in this manner can be portrayed as positive. However, with this persuasion and intentional framing of data, a grey area can appear; it is important to consider the ethical implications of using framing bias in data visualizations.

Is Using Framing Bias in Data Vis Unethical?

Decision-making processes can be heavily impacted by framing bias and framing effects, and this can be used as an advantage when trying to persuade or dissuade a group. However, there are ways to limit framing bias in visualization, and it can also be advantageous to portray data as neutrally as possible. The intentional framing of data in a persuasive, dissuasive, or even neutral way presents an interesting ethical dilemma. In Michael Correll’s work on visualization ethics, Correll states that visualization could technically be considered “an ethically neutral activity…provided that we did not introduce bias or intentionally deceive when presenting our data” [77]. This implies that introducing bias is not ethically neutral. Later on, in his work he argues that “data [itself] is not neutral and [are] objective facts about the world” because data is always presented in one way or another, and even when data is presented as neutrally as possible, it has still been represented in some way that can produce bias, whether it be due to included context or something else. Correll poses another ethical concern about decision-making biases being employed intentionally. He states that “we are enabling bad behavior (unjustified, incorrect, and potentially damaging conclusions from data) without adequate care for the people that can be harmed by decisions based on these conclusions, or adequate understanding about the literacy and capabilities of the people who we are empowering” [77]. Here he states his belief that it is unethical to intentionally attempt to add in bias to manipulate someone.

Correll’s claims about the unethical news of purposefully adding in bias can be disputed. To justify this he claims that the supposed “bad behavior” is enacted “without adequate care for the people that can be harmed by decisions based on these conclusions” [77]. However, it is important to note that in the context of this paper, data visualization effectively *in law* is what is being discussed. In law, individuals, whether they are lawyers, attorneys, or others, are thinking about their clients, not the people that can be harmed by any decisions made. They are not actively thinking about how to create the most ethical visualization possible, but rather thinking first and foremost about their clients’ best interests [79]. Thus, since it is their job to support their clients, adding data visualization can help; it acts as an extension of a lawyer's case. For example, in a court setting, a lawyer's main role is to convince a jury and a judge toward the desired outcome. A lawyer verbally argues a case built with facts, evidence, and data; this data that could hypothetically be used to create a visualization used in court would have to be honest, true data due to the fact that it is illegal to use fabricated data in most if not all legal settings such as courtrooms [78]. While American law prohibits lawyers from "making false statements of material fact or law… and from failing to disclose material facts when necessary" [78], there are no rules as far as how a lawyer can appeal to the emotions and cognitive processes of juries. Lawyers are there to build a convincing case using persuasion, and using data visualization that portrays true data, whether it is framed or not, is nothing but a nonverbal method of persuasion. Additionally, the idea of using data visualization in law cannot be abandoned entirely due to the possibility of someone using framing bias to their advantage; the definite benefits of data visualization in law outweigh the potential misuse of it. However, I propose that standards can be set in order to minimize data misrepresentation and avoid excessive accidental (or accidental) use of framing bias as an advantage.

Standardization of Data Visualization

Thus far, we've established in this paper that data visualization is a powerful tool that is widely missing from the legal community. We know that it can be used as a tool for two main purposes: to increase understanding for those viewing it, and to persuade viewers towards a specific point or conclusion. We also know that framing bias is a part of data visualization and it directs us to look closely at *what* information is visually represented and *how* it is graphically displayed. This is because there are many various ways to represent given data, and how it is done so can affect the decisions made using it. Aspects of data visualization that influence visualization framing bias to affect understanding and/or persuasion include but are not limited to: the type of chart, cherry-picking of data, and use of color and text. To wrap up this proof of concept, I propose a set of recommendations to guide those in the legal field towards leveraging data visualization effectively in their work. These recommendations are made based on several different research studies and build on the concepts and topics aforementioned in this paper.

Data Visualization Recommendations

I. Multiple Visualized Presentations of Data

If your goal is to increase understanding and create a fair judgment process for data visualization viewers and analyzers, one recommendation is to represent data in dashboards instead of single data visualizations. By using a dashboard, it is possible to have "multiple visualized presentations to examine different dimensions of the data set… to apply various views/dimensions to the data set to reveal possible different stories" [32]. In a study from 2012, it is noted that "the quality of problem-solving decisions is enhanced when decision-makers are required to identify multiple explanations for the source or cause of a given problem" [15]. Essentially, having the same or similar data represented in different ways will provide diverse visual aids for viewers to make decisions about. They would not have to be reliant on a single graphic to influence their decision-making; it allows them to “gain the insights from data in various contexts through graphical presentations” [34]. This would ultimately be beneficial to understanding information visualization better, especially in legal settings. Legal data is often multi-dimensional and it's important to highlight the necessary aspects of it, so creating multiple visualizations from a particular dataset gives the viewers, such as jurors, more content on which to base their decisions. However, it is advisable to use caution when creating multiple visualizations for various reasons. First, not every dataset is complex enough or dense enough to need multiple visualizations. Second, having too many visualizations can give users too much to look at, create extra confusion, and make the data harder to use. A study by Reitsma and Marks found that overcomplicating data visualizations caused participants to take longer to comprehend the complex versions and made larger errors interpreting their data [65]. Having significantly more visualizations to view is not beneficial for any decision-maker, so additional visualizations should only be created when they can enhance a user's experience and decision-making process. In a courtroom setting,

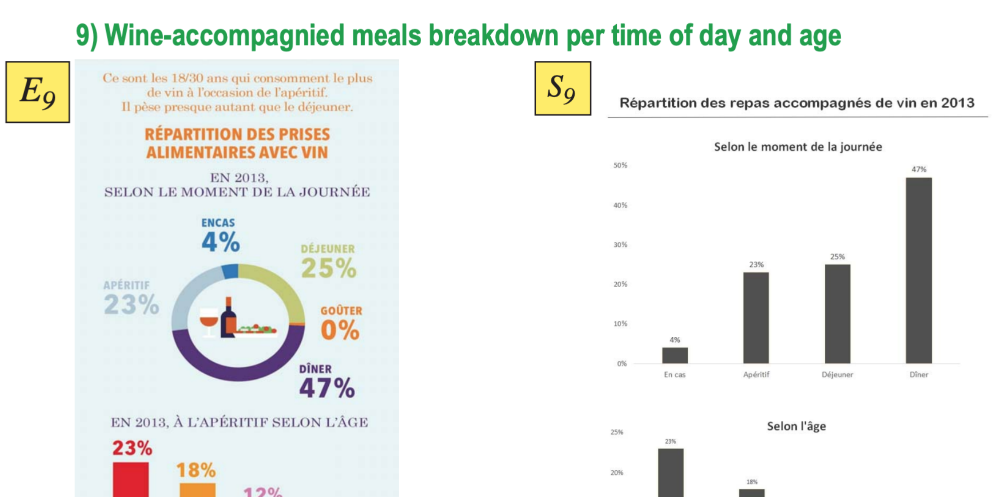
II. Analyze Relationships between Task Type and Data Visualization Type

For data visualization to be used to its full potential, it is important to look into the relationship between the type of task that will be done using the visualization and the type of data visualization being used. As Saket et al. [24] say, “the effectiveness of a visualization depends on several factors including task at hand--while one chart might be suitable for answering a specific type of question it might not be appropriate for other types”. Saket et al. give a comprehensive overview of the relationships between task types and data visualizations types and show that it would be an important aspect for anyone in the legal field looking to begin using visualizations to understand. Notably, Saket et al. find that bar charts are best for finding clusters; pie charts are very effective as well, but in their studies, users preferred bar charts, and additional studies have found that there is a “positive correlation between accuracy and user preference, indicating people have a preference for visualizations that allow them to accurately complete a task” [24]. Thus, bar charts are deemed better than pie charts, but for our purposes, we can say both are significantly better for clustering tasks than other types of vis. Next, line charts are the best visualization for tasks that involve finding correlations. Again, scatter plots are almost equally effective, but users report a preference for line charts [24]. Finally, scatter plots are the best visualization for finding anomalies in data sets. Since the most common tasks associated with data visualization are cluster, correlation, and anomaly tasks, these observations should be taken under consideration when deciding which visualization type to use. However, in general, the recommendation is that extensive thought is given to the type of task the visualization viewers will be ensuing in before picking and designing a visualization. Whether the goal is for viewers to understand data better or to be swayed in some particular direction, different types of visualization will allow for different levels of salience, allow users to see data, and more. A legal use case for this recommendation could be a jury example. A jury in court has limited time to evaluate a case and make decisions on it, so it is best if their time is used efficiently and meaningfully. Depending on the nature of the case and the case data, the visualization they view will have certain tasks inevitably associated with it, such as detecting an anomaly. Since studies show that scatter plots are the best visualization for finding anomalies [24], it would thus be advisable to use a scatter plot as opposed to a bar chart to represent this data. By presenting the jury with the right type of visualization given the case, the data, and the task that needs to be performed to evaluate the data, the jury will be able to make a better decision.

III. Add Meaningful Embellishments

Adding a meaningful embellishment has many benefits that could aid those in the legal field in working with data visualization. Embellishments can come in several different forms, including small pictorial figures, chart junk (which is unnecessary details and made to be purposefully distracting), and color-coding. Research shows that numerous cognitive benefits come from embellishment [27]. For one, adding it has added value, e.g., in memorization and recall [6, 14, 16, 49] and in engaging users to study and understand the infographic [24, 58, 102]. In addition, Haroz et al. found that the use of pictorial icons in visualizations has positive effects on working memory and increases the speed at which information can be found [35]. Hullman et al. back these claims and voice that introducing extra visual images “may better engage the user to read information and improve their comprehension and recall” [36]. Finally, Andry et al. state that using embellishment “reduces the adverse feeling of effort needed to read the data and engages the reader with the visualization topic” [37]. In addition, the embellishments can also be used to increase the persuasiveness of data visualization because they can attenuate certain details. The attenuation could lead viewers to focus on the embellishments or corresponding pieces of information more, leading them to make a decision that is more heavily influenced. As Parsons and Shukla say in their studies, “visual embellishments can be useful depending on the context” [33].

Adding embellishments to legal data vis could look like adding in supporting text, captions, images, icons, and deviating from classical chart design [37]. In the figure below from Andry et al. and their studies, we see an example of a data set presented with embellishment (left) and without (right). They found that the embellished visualizations were overall liked more than standard ones. They also found that there was a “feeling of effort needed to read overloaded

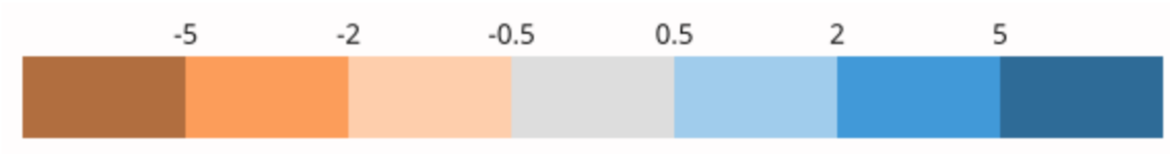


visualizations, i.e., highly embellished ones but also standard ones that show much data; yet, the feeling that a visualization is beautiful creates a pleasant sensation that reduces this effort" [37]. The users also reported that "embellishments, in particular icons (for their suggestive simplicity), remind them of the visualization's topic, increasing their feeling of being 'immersed'" [37]. Andry et al. conclude their study by stating that "embellishments have a positive impact on reading a visualization, by (1) reducing the adverse feeling of effort needed to read the data and (2) engaging the reader with the visualization topic" [37]. Using embellished visualizations in the legal field will allow for increased understanding between lawyers, juries, attorneys, and anyone else in need to understand cases, arguments, and more. For example, given everything we know about the perception of embellishments on vis, jurors presented with embellished visualizations could likely benefit from them.

IV. Use Color Effectively

In regards to color, several intentional actions should be taken, including intentional decisions being made about the choice of colors, called a color palette. Studies show that poor color choices can "distort data" [70]; a study from 2011 by Borkin et al. found that "physicians were significantly worse at diagnosing heart disease from arterial scans that used a rainbow scale than from scans designed for improved perception" [63]. This goes to show the importance of color in visual aids. The color palettes used should be accessible. This means allowing for the colors included in the plots to maximize their effectiveness by being inclusive to those who are colorblind, for example [28]. Red-green color blindness is the most prevalent type of color blindness, and it makes it hard to tell the difference between red and green. Less prevalent color blindnesses are blue-yellow color blindness and monochromacy, which is when you cannot identify color at all [69]. Modern technology and websites can and should be utilized to quickly find the color palettes that are most accessible to the most groups of people; this information is readily available on the internet. For example, the Coblis tool (https://www.color-blindness.com/coblis-color-blindness-simulator/) "allows for testing color deficiencies of specific color palettes by simulating select information visualization examples" [71]. People who are colorblind or have visual impairments work better with specific color palettes, so being mindful of this could lead to better visualizations because everyone, whether they have visual impairments or not, can better read and understand the given vis.

There are three categories of color palettes, which are qualitative, sequential palettes, and diverging palettes. Mike Yi says that "the type of color palette that you use in a visualization depends on the nature of the data mapped to color" [68]. The qualitative palette consists of ten or fewer colors and is used when the variables of data are categorical; this helps differentiate categories and "each possible value of the variable is assigned one color from a qualitative palette". The sequential palette should be used "when the variable assigned to be colored is numeric or has inherently ordered values"; this involves using one color but depicting different aspects of the data in different hues or lightness. Traditionally, this means that "lower values are associated with lighter colors, and higher values with darker colors", although this depends on the background color of the visualization. Finally, diverging palettes are used when the "numeric variable has a meaningful central value, like zero". An example of a diverging palette is shown in the figure below [68].



Color should be used meaningfully, which means using each color in visualization with intention. Studies have found that the more colors used in a visualization, the more difficult it is for individuals to perform tasks using them. In a study by Christopher Healey, they found that "observers had very little difficulty identifying targets during three-colour and five-colour studies" but "target identification became significantly more difficult for certain colours during the seven-colour and nine-colour studies" [67]. This isn't to say that various colors should not be used but rather shows that the use of too many colors can have adverse effects on individuals and color should be used intentionally and not excessively. Color should also be used consistently across visualizations as much as possible to avoid confusion between variables in different vis [68].

Another aspect that is important regarding color is color semantics. This refers to the "meanings imparted by colors and their effects on emotions and moods" [72]. A basic example of color semantics that was explained earlier in this paper is the use of red as "bad" and green as "good", like when the green section on a health pie chart is perceived as healthy and a red section as unhealthy [1]. There are various cultural significance and implied meanings behind different colors and how they are perceived by people. Thus, being mindful of color semantics could be helpful in using color in data vis effectively.

V. Use Supplementary Text

Supplemental text can be added where clarity might be needed in the visualization, where something is a bit confusing or misleading. This text could serve as a clarifier and lead the viewer to have a better understanding of the data and visualization. By adding supplemental text to data visualization, it is possible to not only increase understanding but also push the viewer towards a certain conclusion. This is seen commonly in news media, where journalists will often phrase text in certain ways depending on the point they want to push forward [73]. A prominent example that Kong et al. use in their studies is how "the same news story was titled "Israeli police shoot man in east Jerusalem," "Jerusalem driver shot after ramming pedestrians: police," and "Jerusalem car 'attack' kills baby at rail station" in three different news sources" [74]. Thus, the way text is phrased in a data vis is a crucial part of the user's experience interacting with it, as the way something is phrased can sometimes change their interpretation of the entire visualization itself. However, Kong et al. found that “titles with a contradictory slant triggered more people to identify bias compared to titles with a miscued slant, [but] visualizations were persistently perceived as impartial by the majority” [74]. Thus, the text in data visualization is important and the text slant is important to be aware of, but it is not drastically bias-adding and influential in most cases. Regardless, recommendations should still be set in some capacity for text and its phrasing in data vis. It is difficult to set recommendations about how to exactly phrase text, but there are two methods to go about it. When trying to persuade a group towards something using your vis, you could opt to use persuasive language in your text and frame it in a biased way that pushes them towards what you want them to see and believe. For vis to be more neutral, it would be useful to eliminate persuasive language in the text and opt for “matter-of-fact” language that is simply descriptive of what is being portrayed.

Second, supplemental text can be added in places where it is not necessarily needed but it can highlight different features of the data that the creator wants to bring attention to. By bringing attention to certain details and hence labeling them, this could lead the viewer towards a decision that the creator is guiding them to. This, like the other proposed recommendations, has some tradeoffs to address. A study by Xiong et al. describes that “when people are primed to see one pattern in the data as visually salient, they believe that naive viewers will experience the same visual salience” [75]. However, not all viewers experience the same visual salience, and we only believe that because experts often struggle to “recreate the state of mind of a novice”; this is called “the curse of knowledge bias” [75]. In the context of legal data vis, supplementary text on a data vis may inform a jury about a certain law. This text can be framed in a way such that the litigation specialist creating the data vis intends to direct a jury to think about the law a certain way. However, because they’re considered an “expert” and more aware of the law, they may simply have a curse of knowledge bias and be under the false impression that the jury will experience the same visual salience as them. Thus, although the supplemental text to bring salience to a data vis might work in some instances (and would still be recommended for its potential benefits), it is not a guaranteed method.

Conclusions and Looking Ahead

Summary

This paper encompasses a thorough analysis of data visualization and biases and how they might be leveraged in the legal field. Extensive research into the benefits of utilizing data visualization showed that it allows for an increased understanding of data and can cause bias in a viewer. Then the paper moves into looking at data visualization in legal settings, such as its uses, areas for potential growth, and the importance in decision making. Next, framing bias is analyzed more in-depth by looking at the different ways it can manifest in data visualization, including through the type of visualization used, the cherry-picking (or not) of data, any colors, embellishments, 3D features, or text used within the vis, and the addition of unnecessary and confusing details. The scope of data visualization then narrows down to specifically framing biases within data vis in the legal field, such as the advantages and disadvantages of it. The content concludes with the idea of standardizing certain aspects of data visualization, where I propose several recommendations to guide those in the legal field towards leveraging data visualization in their work.

Future Work

There are many specific spheres of work that were touched on in this paper, and every sphere has vast amounts of research already conducted. However, while information about each separate topic is readily available, there are gaps in connecting these various topics. Where there is information about data visualization, framing bias, cognitive biases, bias in the legal field, bias in data vis, and data vis in law, there are no studies about the overlap of these topics; this was an identifiable limitation for this paper. There is further scope for improvement, including expanding on certain sections, such as the recommendation to use color effectively. For example, there are extensive studies on color theory, color accessibility, color palette building, effective color use in data vis, and so on that were not used or referred to in this paper; future work could include additional modern studies regarding color to make a more specific recommendation regarding the use of color in legal data vis. Another limitation comes from the limited current use of data vis in law today, which was the partial inspiration for the creation of this paper. Due to information being used sparingly, there are even fewer studies and data about it, making certain features of this paper limited in scope. In their studies about sources of bias in juror decision making, Curley and Munro narrate that "before recommendations can be implemented, more research is needed to assess the effectiveness of certain bias reducing strategies (both independently and when interacting with other strategies) [and] governments and legal bodies need to take the effects of bias more seriously and fund high-quality jury studies aimed at tackling the effects of bias" [25]. For future work, as the use of data vis expands in the world of law, I could look more closely at the pain points lawyers, attorneys, juries, and others experience when using data vis and make more precise recommendations for their use cases as well.

Impact

The widespread use of data visualization in legal settings, especially in courtrooms, would be nothing short of game changing and impactful. Court proceedings have consistently been based on oral arguments with minimal data and little to no data visualization, where juries are given large volumes of information verbally that they must physically take notes on or rely on their memory for. The addition of data visualization would offer juries a visual component to rely on and be influenced by in the midst of verbal defenses, cross-examinations, rebuttals, and more. It can act as an extension of a lawyer's argument and offer a modern change in court proceedings. It can be impactful not only for lawyers looking to improve and solidify their cases and arguments but also for the clients who are positively impacted by a fair and just trial made fairer and more just due to the addition of data and data visualization in their case. Court proceedings have remained relatively the same over the decades, with few changes besides the introduction of technological devices. Using visualizations can aid decision making and the understanding of data; crucial, legal decisions will be more informed than ever. In legal settings such as courtrooms where decisions are intensely important and often life altering, the addition of data visualizations could support fairer court outcomes and empower juries and judges to produce more confident verdicts; ideally, this could improve the justice system entirely.

Acknowledgments

I'd like to thank my advisor Danielle Szafir for providing me with incredible support and guidance throughout my thesis writing process. I'd also like to thank Casey Fiesler and Sandra Ristovska for reviewing and evaluating my work and providing me with feedback, as well as Reiland Rabaka for teaching the class that sparked my inspiration for the topic of this paper. Last, I'd like to thank my family and friends for serving as my support system through every shade and color of this paper.

References

[1] Lurie, N.H. and Mason, C.H. (2007), "Visual representation: implications for decision making", Journal of Marketing, Vol. 71 No. 1, pp. 160-177.
[2] Vessey, Iris (1991), "Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature," Decision Sciences, 22 (Spring), 219–40.
[3] Kolata, Gina (1982), "Computer Graphics Comes to Statistics," Science, (September 3), 919–20.
[4] Holbrook, Morris B. and William L. Moore (1981), "Feature Interactions in Consumer Judgments of Verbal Versus Pictorial Presentations," Journal of Consumer Research, 8 (June), 103–113.
[5] Sloman, S.A. (1996), "The empirical case for two systems of reasoning", Psychological Bulletin, Vol. 119 No. 1, pp. 3-22.
[6] Tegarden, David P. (1999), "Business Information Visualization," Communications of the Association for Information Systems, 1 (4), (accessed March 21, 2005), [available at http://cais. isworld.org/articles/default.asp?vol=1&art=4].
[7] Chang, C. J., & Luo, Y. (2018, December 31). Data visualization and cognitive biases in audits. https://www.emerald.com/insight/content/doi/10.1108/MAJ-08-2017-1637/full/pdf
[8] Eppler, M., & Aeschimann, M. (2009). A Systematic Framework for Risk Visualization in Risk Management and Communication. Risk Management, 11(2), 67-89. Retrieved from https://www.researchgate.net/publication/228533068_A_systematic_framework_for_risk_visualization_in_risk_management_and_communication
[9] Brath, Richard and Mike Peters (2005), "Information Visualization for Business: Past & Future," DM Review, (January), 40–43.

[10] Simkin, David and Reid Hastie (1987), “An Information- Processing Analysis of Graph Perception,” Journal of the American Statistical Association, 82 (398), 454–65.
[11] Tang, F., Hess, T.J., Valacich, J.S. and Sweeney, J.T. (2013), “The effects of visualization and interactivity on calibration in financial decision-making”, Behavioral Research in Accounting, Vol. 26 No. 1, pp. 25-58.
[12] Joyce, E.J. and Biddle, G.C. (1981), “Anchoring and adjustment in probabilistic inference in auditing”, Journal of Accounting Research, Vol. 19 No. 1, pp. 120-145.
[13] Kinney, W.R. Jr, and Uecker, W.C. (1982), “Mitigating the consequences of anchoring in auditor judgments”, The Accounting Review, Vol. 57 No. 1, pp. 55-69.
[14] Northcraft, G.B. and Neale, M.A. (1987), “Experts, amateurs, and real estate: an anchoring-and adjustment perspective on property pricing decisions”, Organizational Behavior and Human Decision Processes, Vol. 39 No. 1, pp. 84-97.
[15] Knapp, M.C. and Knapp, C.A. (2012), “Cognitive biases in audit engagements: errors in judgment and strategies for prevention”, The CPA Journal, Vol. 82 No. 6, pp. 40-45.
[16] Valdez, A.C., Ziefle, M. and Sedlmair, M. (2018), “Priming and anchoring effects in visualization”, IEEE Transactions on Visualization and Computer Graphics, Vol. 24 No. 1, pp. 584-594.
[17] Hutchinson, J. Wesley, Joseph W. Alba, and Eric M. Einstein (2004), “Numerical Inferences: The Effects of Prior Beliefs, Semantic Frames, and Graphical Formats on Budget Allocation Decisions,” working paper, Marketing Department, The Whar- ton School, University of Pennsylvania.
[18] KPMG (2011), “Elevating professional judgment in auditing and accounting: the KPMG professional judgment framework”, available at: www.drlillie.com/a544/kpmg/jdgmt/ KPMG_ProfJudgment_Monograph.pdf (accessed 23 May 2018).
[19] Waterfield, D., & Doughty, M. (2019, December 23). Data Visualization Tools Are Poised to Revolutionize the Legal Industry. Corporate Counsel Business Journal.
https://ccbjournal.com/articles/data-visualization-tools-are-poised-to-revolutionize-the-the-legal-industry
[20] Synovate (2005), “Quantitative Evaluation of Alternative Food Signposting Concepts of Work,” U.K. Food Standards Agency, (November), (accessed November 20, 2005), [available at
http://www.food.gov.uk/multimedia/pdfs/signpostquan research.pdf].
[21] VanHemert, K. (2014, June 11). *A New Way to Look at Law, With Data Viz and Machine Learning*. WIRED. Retrieved April 2022, from https://www.wired.com/2014/06/ravel-law/
[22] Gill, L. H. (2021, June 11). Forensic Examination | The Benefits Of Data Visualization. Forensic Strategic Solutions. Retrieved April 12, 2022, from
https://www.forensicstrategic.com/forensic-examinations-how-data-visualization-can-help-lawyers-and-investigators-tell-a-compelling-story/
[23] “Visualization of Law and Legal Process” by Scott McLachlan and Lisa C Webley
https://journals.sagepub.com/doi/full/10.1177/14738716211012608
[24] Bahador Saket, Alex Endert, and Çağatay Demiralp. 2018. Task-based effectiveness of basic visualizations. IEEE Transactions on Visualization and Computer
Graphics 25, 7 (2018), 2505–2512.
[25] Curley, L. J., Munro, J., & Dror, I. (2022, February 17). Cognitive and human factors in legal layperson decision making: Sources of bias in juror decision making. https://doi.org/10.1177/00258024221080655
[26] Chopra, S. (2020, January). The psychology of framing and jury decision-making. Advocate Magazine. Retrieved April 3, 2022, from
https://www.advocatemagazine.com/article/2020-january/the-psychology-of-framing-and-jury-decision-making
[27] Wu, K., Peterson, E., Ahmad, T., Burlinson, D., Tanis, S., & Szafir, D. (2021). Understanding Data Accessibility for People with Intellectual and Developmental Disabilities. https://doi.org/10.1145/3411764.3445743
[28] Sanchez, A. H. (2021). Data Visualization for International Relations.
https://bookdown.org/alhdzsz/data_viz_ir/

[29] S. Bateman, R. L. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. 2010. Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. In Proc. ACM CHI. 2573–2582. https://doi.org/10.1145/1753326.1753716
[30] E. R. Tufte. 1990. Envisioning information. Graphics Press.
[31] Duarte, N. (2014, April 16). The Quick and Dirty on Data Visualization. Harvard Business Review. Retrieved April 12, 2022, from https://hbr.org/2014/04/the-quick-and-dirty-on-data-visualization
[32] KPMG (2015), "Seeing beyond the numbers: improving revenue cycle results through data visualization", is available at: www.kpmg-institutes.com/institutes/healthcare-lifesciences-institute/articles/2015/01/ visualizing-rev-cycle-issue-brief.html (accessed 23 May 2018).
[33] Parsons, P., & Shukla, P. (2020, September 6). *Data Visualization Practitioners' Perspectives on Chartjunk*. https://doi.org/10.48550/arXiv.2009.02634
[34] Singh, K. and Best, P. (2016), "Interactive visual analysis of anomalous accounts payable transactions in SAP enterprise systems", Managerial Auditing Journal, Vol. 31 No. 1, pp. 35-63.
[35] Steve Haroz, Robert Kosara, and Steven L. Franconeri. 2015. ISOTYPE Visualization: Working Memory, Performance, and Engagement with Pictographs. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 1191–1200. https://doi.org/10.1145/2702123.2702275
[36] J. Hullman, E. Adar, and P. Shah. 2011. Benefitting InfoVis with Visual Difficulties. IEEE Transactions on Visualization and Computer Graphics 17, 12 (2011), 2213–2222.
[37] Andry, T., Hurter, C., Lambotte, F., Fastrez, P., & Telea, A. (2021). Interpreting the Effect of Embellishment on Chart Visualizations. https://doi.org/10.1145/3411764.3445739
[38] Cleveland, William S. and Robert McGill (1984), "Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods," Journal of the American Statistical Association, 79 (September), 531–54.
[39] Cleveland, William S. and Robert McGill (1985), "Graphical Perception and Graphical Methods for Analyzing Scientific Data," Science, (August 30), 828–33.
[40] Krider, Robert E. Priya Raghubir, and Aradhna Krishna (2001), "Pizzas: Π or Square? Psychophysical Biases in Area Comparisons," Marketing Science, 20 (4), 405–425.
[41] Raghubir, Priya and Aradhna Krishna (1996), "As the Crow Flies: Biases in Consumers' Map-Based Distance Judgments," Journal of Consumer Research, 23 (June), 26–39.
[42] Raghubir, Priya and Aradhna Krishna(1999), "Vital Dimensions in Volume Perceptions: Can the Eye Fool the Stomach," Journal of Marketing Research, 36 (August), 313–26.
[43] Hoffman, Donna L. and Thomas P. Novak (1996), "Marketing in Hypermedia Computer-Mediated Environments: Conceptual Foundations," Journal of Marketing, 60 (July), 50–68.
[44] Steuer, Jonathan (1992), "Defining Virtual Reality: Dimensions Determining Telepresence," Journal of Communication, 42 (4), 73-93.
[45] Jarvenpaa, Sirkka L. (1989), "The Effect of Task Demands and Graphical Format on Information Processing Strategies," Management Science, 35 (March), 285–303.
[46] Nisbett, Richard E. and Lee Ross (1980), Human Inference: Strategies and Shortcomings of Social Judgment. Englewood Cliffs, NJ: Prentice-Hall.
[47] Mandel, Naomi and Eric J. Johnson (2002), "When Web Pages Influence Choice: Effects of Visual Primes on Experts and Novices," Journal of Consumer Research, 29 (2), 235–45.
[48] Bederson, Benjamin B. and Ben Shneiderman (2003), The Craft of Information Visualization: Readings and Reflections. San Francisco: Morgan Kaufman.
[49] Healey, C.G., K.S. Booth, and J. Enns (1995), "Visualizing Real-Time Multivariate Data Using Preattentive Processing," ACM Transactions of Modeling and Computer Simulation, 5 (3), 190–221.

[50] Julesz, Bela (1981), "Textons, the Elements of Texture Perception, and Their Interactions," Nature, (July), 91–97.
[51] Treisman, Anne (1985), "Preattentive Processing in Vision," Computer Vision, Graphics, and Image Processing, 31 (August), 156–77.
[52] Borzo, Jeanette (2004), "Get the Picture: In the Age of Information Overload, Visualization Software Promises to Cut Through the Clutter," The Wall Street Journal, (January 12), R4.
[53] Esfahani, Elizabeth (2005), "Why 'Buy' Buttons Are Booming," Business 2.0, (July 1), 36.
[54] Miller, Paul (2004), "Lands' End Upgrades 'Virtual Model,'" Catalog Age, (February 10), (accessed November 11, 2005), [available at http://multichannelmerchant.com/catalogage/ ecommerce/marketing_lands_end_upgrades/].
[55] Miller, George A. (1956), "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information," Psychological Review, 63, 81–97.
[56] Healey, C. G., & Enns, J. T. (2012, July). https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5963660&tag=1. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, *18*(7). 10.1109/TVCG.2011.127
[57] A Task-Based Taxonomy of Cognitive Biases for Information Visualization. (2020, February). *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, *26*(2). https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8476234
[58] J. Klayman, J. B. Soll, C. Gonzalez-Vallejo, and S. Barlas, "Overconfidence: It depends on how, what, and whom you ask," Organizational Behavior Human Decision Processes, vol. 79, no. 3, pp. 216–247, 1999.
[59] G. Ellis and A. Dix, "Decision making under uncertainty in visualization?" in Proc. IEEE Workshop Vis. Decision Making Under Uncertainty, 2015.
[60] Lee, C., Tanya Yang, Gabrielle Inchoco, Graham M. Jones, & Arvind Satyanarayan. (2021, March 1). *The Data Visualizations Behind COVID-19 Skepticism*. MIT Visualization Group. Retrieved April 2022, from http://vis.mit.edu/covid-story/
[61] *COVID-19 Misinformation is Ubiquitous: 78% of the Public Believes or is Unsure About At Least One False Statement, and Nearly a Third Believe At Least Four of Eight False Statements Tested*. (2021, November 8). KFF. Retrieved April 2022, from https://www.kff.org/coronavirus-covid-19/press-release/covid-19-misinformation-is-ubiquitous-78-of-the-public-believes-or-is-unsure-about-at-least-one-false-statement-and-nearly-at-third-believe-at-least-four-of-eight-false-statements-tested/
[62] Kent, J. (2020, November 10). *COVID-19 Data Visualizations Shared on Twitter May Lack Reliability*. Health IT Analytics. Retrieved April, 2022, from https://healthitanalytics.com/news/covid-19-data-visualizations-shared-on-twitter-may-lack-reliability
[63] Borkin MA, Gajos KZ, Peters A, Mitsouras D, Melchionna S, Rybicki FJ, Feldman CL, Pfister H. Evaluation of artery visualizations for heart disease diagnosis. IEEE Trans Vis Comput Graph. 2011 Dec;17(12):2479-88. doi: 10.1109/TVCG.2011.192. PMID: 22034369.
[64] *Introduction II. The Roots of the Jury System*. (n.d.). American Bar Association. Retrieved April 20, 2022, from https://www.americanbar.org/content/dam/aba/administrative/public_education/publication_documents/LawCourts_intro_roots_of_jury_system.pdf
[65] Reitsma, R., & Marks, A. (2009, Summer). The Future of Data: Too Much Visualization, Too Little Understanding? *Dialectic*, *2*(2). https://doi.org/10.3998/dialectic.14932326.0002.207
[66] A. Mosca, A. Ottley, and R. Chang. Does interaction improve bayesian reasoning with visualization? In IEEE VIS Workshop on Visualization for Communication (VisComm '20), 2020.
[67] C. G. Healey. Choosing effective colours for data visualization. In Proc. IEEE VIS, pages 263–270, 1996.
[68] Yi, M. (n.d.). *How to Choose Colors for Data Visualizations | Tutorial by Chartio*. Chartio. Retrieved April 19, 2022, from https://chartio.com/learn/charts/how-to-choose-colors-data-visualization/

[69] *Types of Color Blindness*. (n.d.). National Eye Institute. Retrieved April 19, 2022, from https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness/types-color-blindness
[70] Katsnelson, A. (2021, October 4). *Colour me better: fixing figures for colour blindness*. Nature. https://www.nature.com/articles/d41586-021-02696-z
[71] Rhyne, T., & van Haagen, E. (2020, October 15). *Ten simple rules to colorize biological data visualization*. PLOS. Retrieved April 19, 2022, from https://journals.plos.org/ploscompbiol/article/file?id=10.1371/journal.pcbi.1008259&type=printable
[72] Niazi, Mahwish & Ahmad, Nawaz & Aftab, Maria. (2015). A Psycho-Linguistic Exploration of Color Semantics. Journal of Literature, Languages and Linguistics. 16. 1-6.
[73] Blake C Andrew. 2007. Media-generated shortcuts: Do newspaper headlines present another roadblock for low-information rationality? Harvard International Journal of Press/Politics 12, 2 (2007), 24–43.
[74] Student News Daily. 2014. 3 headlines that mislead readers. (Oct. 2014). https://www.studentnewsdaily.com/example-of-media-bias/3-headlines-that-mislead-readers/
[75] C. Xiong, L. Van Weelden and S. Franconeri, "The Curse of Knowledge in Visual Data Communication," in IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 10, pp. 3051-3062, 1 Oct. 2020, doi: 10.1109/TVCG.2019.2917689.
[76] Qiu, L. (2015, October 1). *Chart shown at the Planned Parenthood hearing is misleading and 'ethically wrong'*. PolitiFact. Retrieved April 2022, from https://www.politifact.com/factchecks/2015/oct/01/jason-chaffetz/chart-shown-planned-parenthood-hearing-misleading-/
[77] Michael Correll. 2019. Ethical Dimensions of Visualization Research. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 188, 1–13. DOI:https://doi.org/10.1145/3290605.3300418
[78] *Can Lawyers Lie in Negotiations?* (2015, June 17). St Francis School of Law. Retrieved April, 2022, from https://stfrancislaw.com/blog/can-lawyers-lie-negotiations/
[79] *The Client's Best Interest: A Lawyer's Obligation to their Client*. (2014, July 4). Dale Streiman Law. Retrieved April, 2022, from https://www.dalestreimanlaw.com/clients-best-interest-lawyers-obligation-client/
[80] Tracy, S. (2015, October 2). *Jason Chaffetz's Impossible Chart & Some DataViz Sins You Should Never Commit*. Analythical by Stephen Tracy - An Analytics & Research Blog. Retrieved April 2022, from https://analythical.com/blog/data-visualization-sins
[81] J. Hullman and N. Diakopoulos, "Visualization Rhetoric: Framing Effects in Narrative Visualization," in IEEE Transactions on Visualization and Computer Graphics, vol. 17, no. 12, pp. 2231-2240, Dec. 2011, doi: 10.1109/TVCG.2011.255.
[82] Miller, R. (2022, January 12). With more data available than ever, are companies making smarter decisions? TechCrunch. Retrieved April 25, 2022, from https://techcrunch.com/2022/01/07/with-more-data-available-than-ever-are-companies-making-smarter-decisions/
[83] Walden University. (2021, September 15). Why-data-matters-in-criminal-justice. Walden University. Retrieved April 25, 2022, from https://www.waldenu.edu/programs/criminal-justice/resource/why-data-matters-in-criminal-justice
[84] Milgram, A. (2013, October). Why smart statistics are the key to fighting crime. Anne Milgram: Why smart statistics are the key to fighting crime | TED Talk. from https://www.ted.com/talks/anne_milgram_why_smart_statistics_are_the_key_to_fighting_crime#t-502254
[85] Data in the Court. The Practice. (2022, March 30). from https://thepractice.law.harvard.edu/article/data-in-the-court
[86] How to ensure your forensic data can be used in court. Eide Bailly LLP. (2021, January 19). from https://www.eidebailly.com/insights/articles/2017/12/how-to-ensure-your-forensic-data-can-be-used-in-court
[87] "Technology In Trial." *Aronberg Law*, The Legal Intelligencer, 28 July 2008, https://aronberglaw.com/technology-in-trial-71608/.

[88] Mohan, Ronita. “7 Data Visualization Mistakes and How Marketers Can Avoid Them.” *60 Second Marketer*, 11 June 2021,
https://60secondmarketer.com/2021/06/11/7-data-visualization-mistakes-and-how-marketers-can-avoid-them.