A Systematic Review of Data Visualization

Shilpa Balan
California State University, Los Angeles

Data visualization involves presenting data in graphical or pictorial form that in turn helps with decision support. This study addresses the importance of data visualizations in today's data world, and its role in effective communication in various industry sectors. This study applied the PRISMA methodology to conduct the literature review. Four major themes are identified in this paper: data visualization principles, methods, contemporary chart types, and data visualization challenges. Popularly applied data visualization principles such as Gestalt Principles that help make data visualizations more effective are examined. Advanced chart types include doughnut chart, chord diagram, sankey diagram, and violin plot. Some challenges associated with data visualizations are data accuracy, complexity and uncertainty. Given the challenges, visualizations still benefit any field of study that requires interpreting and presenting complex information.

Keywords: data visualization, Gestalt Principles, charts

INTRODUCTION

Data visualization is the graphic representation of data. Visualization is "a mental picture" (Cayne, 1987, p. 1100). According to *the Random House Dictionary of the English Language* (Flexner, 1987, p. 2127) to visualize is "to recall or form mental images or pictures; to make visual or visible; to form a mental image of; to make perceptible to the mind or imagination." These definitions suggest that a visualization is a picture of some phenomena.

Visualizing data helps with decision support (Viégas and Wattenberg, 2007) and for storytelling. Data visualization techniques provide a pictorial overview of complex data sets to identify patterns, relationships, and trends. Visualization enables faster cognitive load of information processing leading to better recall of data (Borkin et al., 2013). Charts provide cues that draw people's attention to focus on areas of interest (Tegarden, 1999).

Data visualizations have a long history and have been frequently used in daily newspapers. Many of the chart types we currently use date back many years. Graphs such as bar charts, column charts, doughnut charts and pie charts are commonly applied in data visualization. These chart types are presented in the books '*The Commercial and Political Atlas*' and 'Statistical Breviary' published by William Playfair, respectively in 1786 and 1801. Other commonly used chart types are histogram, line chart, table, scatter plot, bubble plot, area chart, and tree map.

Visualization trends change over time. The research question examined in this study is: What are the current data visualization practices directing to emerging visualization trends? This paper examines a literature review of visualization literacy to convey visualization trends. Following are the main contributions of this paper:

- An updated source of current and emerging trends in visualization practices and charts
- Frequently applied data visualization principles in the present time.
- A review of industries that commonly apply data visualizations at the present time.
- Frequently observed challenges in data visualization at the present time.

Various techniques, along with their representative works have been summarized in a table for readers' reference, serving as recommendations based on the data being visualized. Further, a discussion of challenges and problems in the data visualization field and suggestions of possible future research directions have been presented.

BACKGROUND

Data is easier to interpret when represented as visualizations (Dur, 2014). John Snow, a physician used mapping to identify the source of the cholera outbreak that occurred in London in 1854. He created a map based on the locations of cholera deaths, enabling him to identify a pattern that no one had noticed previously. Snow identified the outbreak's source as the contaminated public water pump on Broad Street that is presently Broadwick Street. Snow noted that they were mostly people who had access to water from the Broad Street pump (Begum, 2016). It was later discovered that the water for the pump was polluted by sewage contaminated with cholera. His studies of the disease pattern led to disabling the well pump.

Over the years, visualization methods have been advanced to illustrate enormous amounts of information for further analysis (Ajibade et al., 2016). Figure 1 describes the data visualization process. As illustrated in Figure 1, data can be quantitative or qualitative. In most cases, visualization works best for answering concrete quantitative questions. Distributions, correlations, and trends of a value set are well understood when depicted visually. Questions of a qualitative nature can be represented by visualizations showing patterns with characteristics in terms of shape, position, or data. An example of a qualitative question is, "given a medical scan of a patient, are there any anomalies that may indicate clinical problems?" (Telea, 2015).

Visuals can be static or interactive. We can create dashboards that are interactive. A dashboard is a visual display of consolidated information. Visuals can be planned using wireframes or storyboards. A wireframe is a plan of a user interface. Using wireframes, it is easy to picture what will be developed. With this technique, user suggestions can be incorporated early, increasing the end product's quality (Visual Paradigm, 2023). Using a storyboard, one can illustrate the data narrative in a way that makes sense to the audience (Dougherty and Ilyankou, 2024). Gestalt Principles, which are principles of perception, describe how the audience for a visualization group similar elements and recognizes patterns when perceiving objects (Entango, 2023). Data visualization techniques and chart types applied are discussed in the later sections.

METHODOLOGY

Articles on data visualization were searched using keywords such as 'data visualization' and 'visual analytics.' Articles were then filtered based on their significance and relevance. After filtering, a total of forty-seven peer-reviewed articles were selected in all and reviewed from journal sources such as Visual Informatics, International Journal of Computer Science and Information Technology Research, Circ Cardiovasc Qual Outcomes, International Journal for Research in Applied Science & Engineering Technology and IEEE to name a few, in addition to other data visualization sources such as 'Our World in Data' (*https://ourworldindata.org/*) , Nightingale- Journal of the Data Visualization Society and several conference proceedings.

Figure 2 displays the systematic review's data collection and screening process using the PRISMA methodology. The PRISMA methodology is a guideline for reporting systematic reviews (Page et al., 2021).

The inclusion and exclusion criteria resulted in the identification of 550 publications. 370 records were screened of which 260 articles were suitable for the next step. 110 articles were eliminated for being off-topic, leaving 260 publications. Out of these, 142 were accessible for full-text extraction, while the remaining 118 were assessed by abstract. Of the full-text 142 articles, 35 were found to be irrelevant. Of the 118 articles reviewed by abstract, 18 were not retrievable. The remaining publications were thoroughly examined manually. In the end, 47 articles were considered for this research. Further, Figure 3 shows the number of papers collected on data visualization by publication year. These articles are further reviewed in this manuscript.

FIGURE 1
DATA VISUALIZATION PROCESS

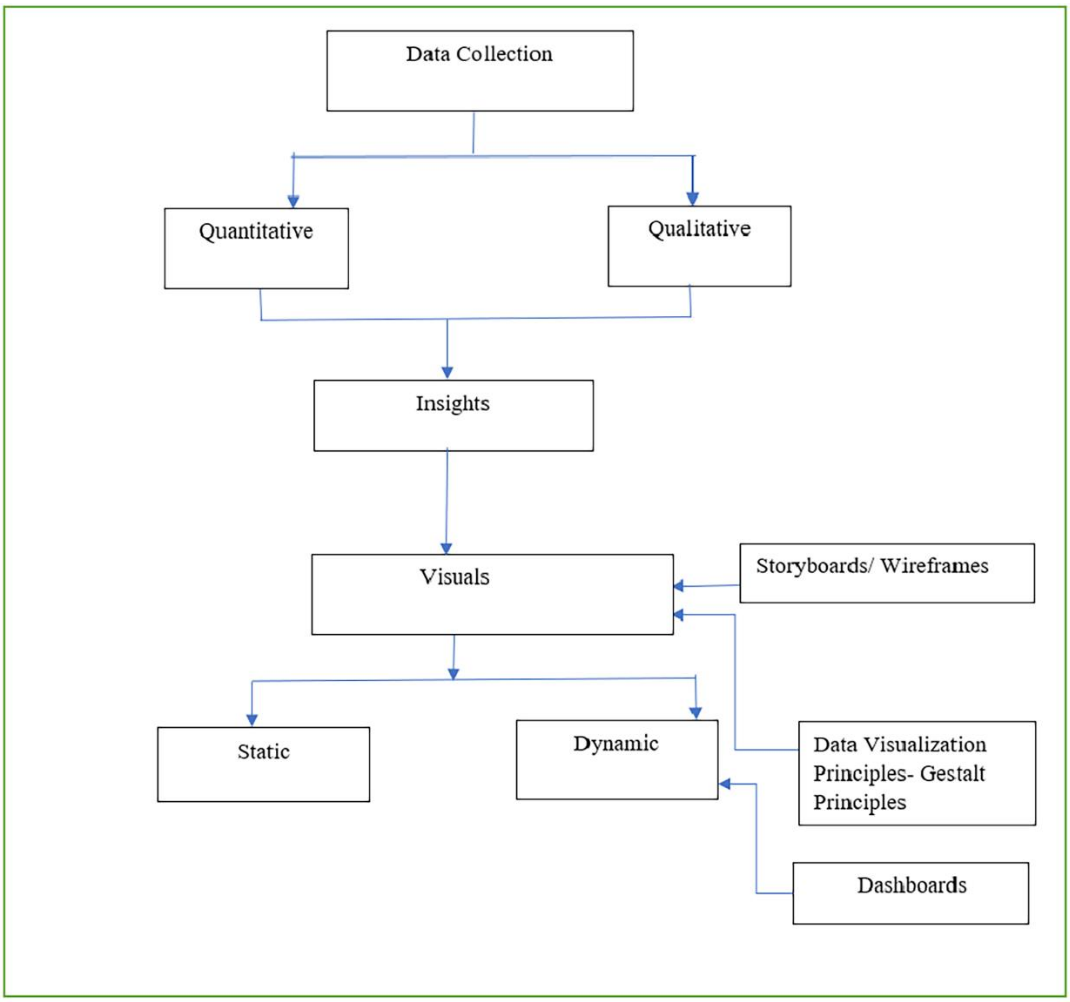
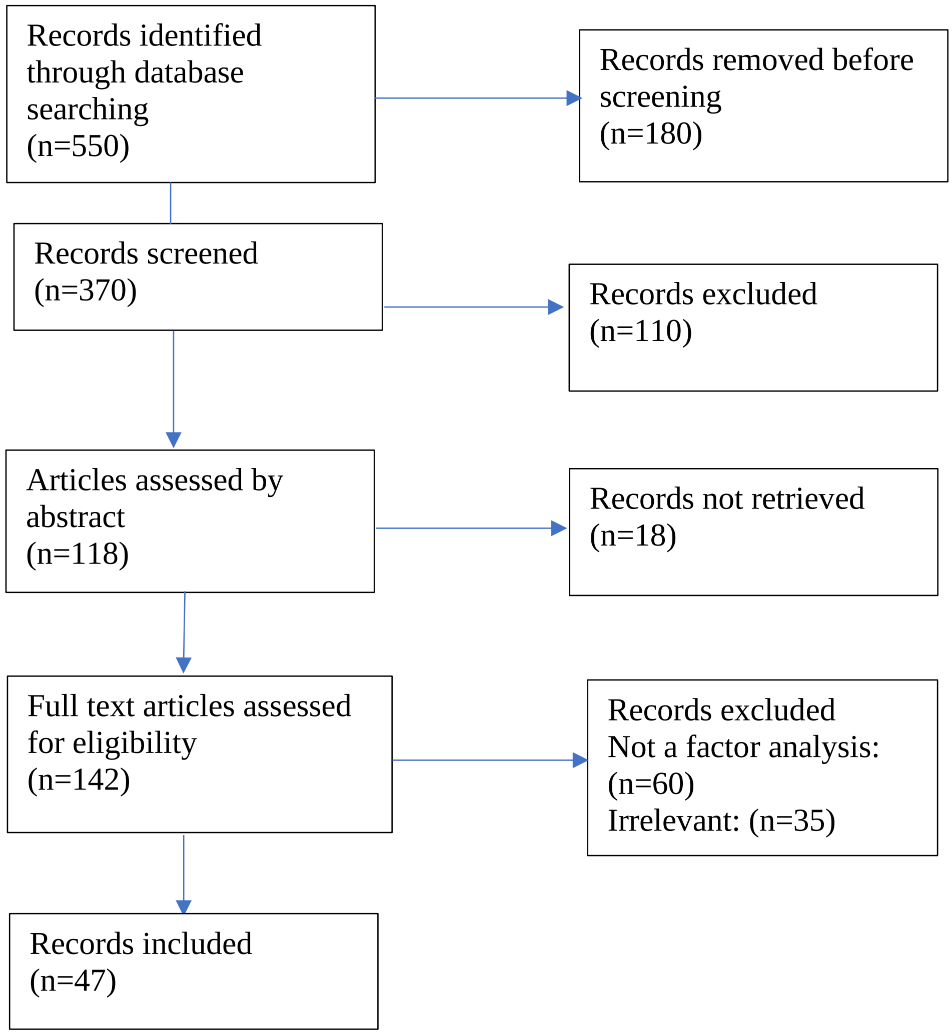


FIGURE 2
FLOW DIAGRAM OF SYSTEMATIC REVIEW OF DATA VISUALIZATION



Data visualization is classified concerning various themes and applications where advanced visualization techniques are still required to make better informed decisions. Table 1 lists four different themes. Theme 1 discusses the data visualization principles in handling various types of data, graphs, and colors. In Theme 2, the focus is on data visualization tools and popularly used chart types. In Theme 3, emerging trends in data visualization are analyzed. In the final Theme 4, some major challenges related to data visualization are reviewed.

FIGURE 3
NUMBER OF PAPERS BY PUBLICATION YEAR

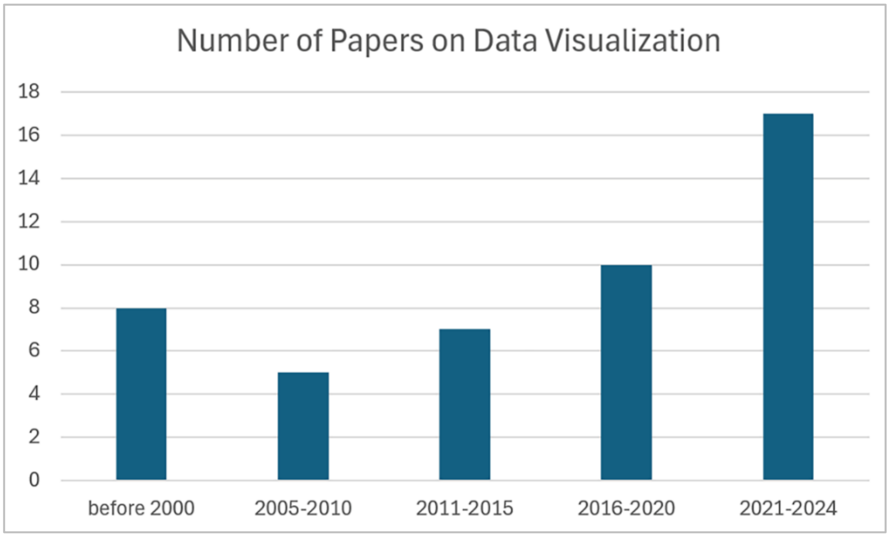


TABLE 1
DATA VISUALIZATION THEMES

Theme	Topic
Theme 1: Data Visualization Principles	• Gestalt Principles • Grammar of Graphics • Cognitive Fit Theory
Theme 2: Data Visualization Methods	• Popular chart types used • Data visualization tools
Theme 3: Contemporary Chart Types	• Reviewing emerging trends in data visualization
Theme 4: Data Visualization Challenges	• Data quality, complexity • Data uncertainty • Misleading visuals

THEME 1: DATA VISUALIZATION PRINCIPLES

Some popularly applied data visualization principles are Grammar of Graphics, Cognitive Fit Theory and Gestalt Principles which are further discussed in this section. Table 2 describes the top data visualization principles frequently applied: Grammar of Graphics, Cognitive Fit Theory and Gestalt Principles. While Gestalt Principles help break down complex images, Grammar of Graphics and Cognitive Fit Theory guide designing and evaluating visuals for effective decision making.

TABLE 2
DATA VISUALIZATION PROCESS

Data Visualization Principle	Purpose	Source
Grammar of Graphics	Guides the design of graphics for data visualization	Wilkinson, 1999
Cognitive Fit Theory	Evaluate the effectiveness of information visualization technologies	Teets et al., 2010
Gestalt Principles	Helps to recognize patterns and break down complex images	Kuznetsova, 2023

Grammar of Graphics

Grammar of Graphics is a theory that guides graphics design for data visualization. Leland Wilkinson develops this theory for visualizing and communicating data in a structured way (Wilkinson, 1999). The Grammar of Graphics is based on the idea that every graphic can be broken down into a series of components. These components include the data, the aesthetic mapping, the geometric shapes, the statistical transformation, and the scales. The data component denotes the raw data that is being visualized while the aesthetic mapping refers to the way that data is mapped to visual properties such as color, size, and shape. A variant of the Grammar of Graphics is known as the *layered grammar of graphics framework*, proposed by Hadley Wickham, the creator of the famous R visualization package ggplot2 (Wickham, 2010).

Cognitive Fit Theory

Cognitive fit theory is a basis for evaluating information visualization technologies' effectiveness (Teets et al., 2010). Cognitive Fit Theory (Vessey, 1991) has been used to explain why graphs are sometimes better than tables for supporting decision making. The better the fit is between two constructs, the more effective the problem-solving process.

Gestalt Principles

When creating a visualization such as a dashboard, there are two effective philosophies to develop storytelling and interaction skills: *Cognitive Load Theory* and *Gestalt Principles*. Cognitive Load means the amount of information processing resources. An effective visualization aims to minimize the cognitive load. Gestalt Principles helps with this. Gestalt Principles, Gestaltism or Gestalt Psychology all of them describe how people's brain and eyes perceive the world (Kuznetsova, 2023). Gestalt means 'form' in German.

Gestalt principles are a set of principles that describe how humans perceive and organize visual information (Porras, 2023). These principles can be applied to data visualization to help make visualizations effective. For example, we perceive objects as a unified entity if they are connected. When objects are in close proximity, we infer a connection between them as shown in Figure 5.

Using NLTK, a widely used open-source NLP (Natural Language Processing) library in Python, the important themes related to Gestalt principles were generated as described in Figure 4. Gestalt Principles helps to recognize patterns and break down complex images. Using Gestalt principles, the data visualization elements are enhanced to an audience. These elements are vital in data visualizations because they enable better understanding.

The Gestalt principles further described in Figure 5 illustrate that similarity refers to objects sharing the same color, shape, or size that are perceived as related or part of the same group (Kuznetsova, 2023). In visualization, the closure property aids in eliminating unnecessary elements from charts, ensuring clarity and simplicity.

FIGURE 4
NLP THEMES RELATED TO GESTALT PRINCIPLES

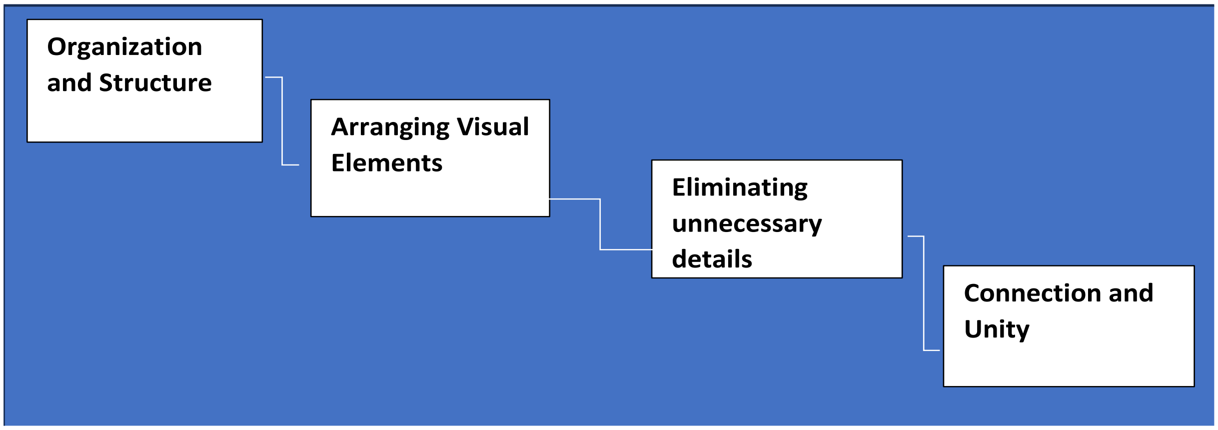
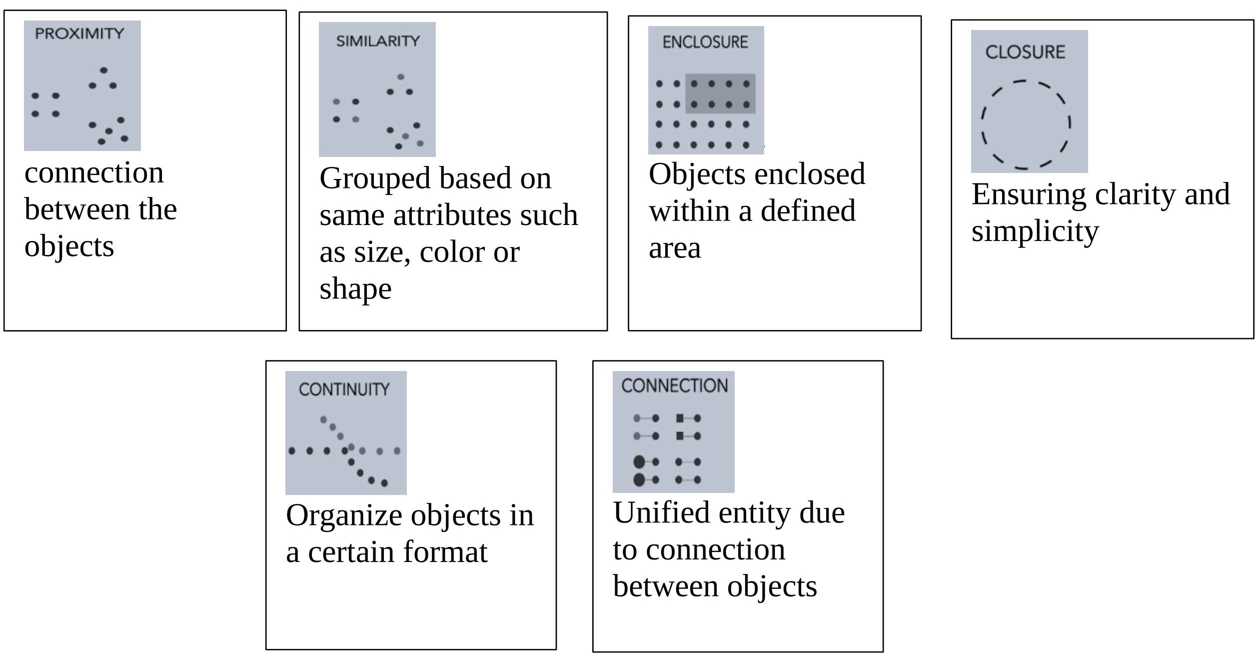


FIGURE 5
GESTALT PRINCIPLES



THEME 2: DATA VISUALIZATION METHODS

This section discusses data visualization methods, chart types and applications. Table 3 lists popular data visualization methods such as bar chart, line chart, pie chart, scatter plot, bubble chart, tree map and stacked bar chart.

TABLE 3
DATA VISUALIZATION METHODS

Data Visualization Method	Source
Node diagram	Liu et al. (2022)
Bar chart	Cleveland and McGill (1984) Kosara (2016)
Line Chart	Ajibade, Adediran (2016) Kim, Setlur, Agrawala (2021)
Pie Chart	Khan, M., Khan, S. (2011) Spence (2005) Ajibade, Adediran (2016)
Table	Khan, M., Khan, S. (2011)
Scatter Plot	Khan, M., Khan, S. (2011) Keim et al. (2010)
Bubble Chart	Khan, M., Khan, S. (2011) Chien T-W., Lin Y-F., Chang C-H., Tsai M-T. & Uen Y-H. (2012).
Tree Map	Khan, M., Khan, S. (2011) Keahey (2013)
Stacked Bar	Keahey (2013)

Some commonly used chart types are discussed below (SAS, 2023) which are further described in Table 4.

- *Line graph*: This chart type can be used to compare changes over a period of time.
- *Bar chart*: This chart type is used to compare quantities of different categories.
- *Scatter plot*: This chart type is a two-dimensional plot showing variation of two items.
- *Pie chart*: This chart type is used to compare parts of a whole.

The format of frequently used graphs and charts take the form of bar chart, pie chart, and line graph to name a few. It is important to understand which chart or graph to use for your data. Table 4 illustrates the use and benefits of these chart types.

With data visualization, insights can be extracted faster. A massive spreadsheet of data would be less meaningful in comparison to an effective visualization. Something as simple as displaying data in a graphic format may seem to have no shortcomings. However, data can sometimes be misrepresented when determined incorrectly in the data visualization. Data visualizations will not be useful if biased or inaccurate information is shown (Tableau, 2023).

Data visualization is storytelling with a purpose. Data visualization tools and software generate a variety of charts, graphs, and map types. While all visualization tools provide the capability to create chart types, the complexity of the product can vary (Power BI, 2023). Most use templates to create simple graphics. Some tools help with creating interactive visualizations which could have greater impact with the audience. Some of the popularly used visualization software are further discussed. Table 5 shows the popular used data visualization tools along with the compatible data sources.

Tableau is a business intelligence platform which is easy to navigate (Gour, 2019). It offers a platform to create meaningful visualizations to represent the data set pictorially. **Tableau is popular for** its drag-and-drop functionality to create visualizations (Gour, 2019). Tableau provides capabilities for creating and sharing reports and dashboards.

With Tableau, one can connect to several data sources such as Oracle, Teradata, SAP HANA, MongoDB, Excel, Text files, JSON, Google Cloud, SQL, Hadoop, and Amazon Redshift. In addition to

Tableau Desktop, Tableau offers a version called Tableau Public. The main difference between Tableau Desktop and Tableau Public is that all visualizations created with Tableau Public are available to see for anyone.

TABLE 4
COMPARISON OF CHART TYPES

Chart Type	Function	Application Purpose
Line Chart	Line charts are frequently used to make comparison between lots of items at the same time.	Time series trend
Bar Chart	Bar chart is as well referred to as column chart and they are used to for comparison of items of different groups.	Comparison of data points
Pie Chart	A pie chart shows information statistics and data in a way that is not difficult to read called “pie-slice” form and the various sizes of slice shows how much of an element is in existence.	Part-to-whole comparisons
Scatter plot	Scatter plot shows how data items are related	Relationship of data items

TABLE 5
DATA VISUALIZATION TOOLS

Tool	Frequently applied Data Sources
Tableau	Oracle, Teradata, SAP HANA, MongoDB, Excel, Text files, JSON, Google Cloud, SQL, Hadoop, and Amazon Redshift
Power BI	Oracle, IBM, SQL Server, Salesforce, Google analytics, Azure DevOps, Excel, text files, and JSON
R	capable of reading data from most formats Excel (in CSV, XLSX, or TXT format), SAS, Stata, SPSS, or others,
Python	Csv, json, html, Excel, Stata, SAS, SQL, Google Big Query
D3.js	Array, CSV, TSV, JSON, XML

Microsoft Power BI is another popular data visualization tool. It is a cloud-based software which is available in the desktop (*Power BI Desktop*) and mobile (*Power BI Mobile)*. A Power BI report may be a single page with one visual, or it could be multiple pages of visuals. Microsoft Power BI provides an easy-to-use functionality for data preparation and visualizations (Gour, 2019). Similar to Tableau, this tool enables connectivity to a variety of data sources. Using Power BI, data sources such as Oracle, IBM, SQL Server, Salesforce, Google analytics, Azure DevOps, Excel, text files, and JSON can be connected to.

R programming is a popular data analytics tool when it comes to visual analytics (Khaled, 2020). ‘R’ is an open-source tool that can provide a more thorough way to create useful graphs for audiences. R

Studio is a user-friendly IDE (Integrated Development Environment) for R. There is an associated community for R programming that makes learning and sharing effective visualizations easier.

ggplot2 is a data visualization package in the R community (Khaled, 2020) that was created by Hadley Wickham in 2005. It was implemented based on Leland Wilkinson's *Grammar of Graphics*, a guideline for data visualization that breaks up graphs into semantic components containing layers. ggplot2 can improve the quality and aesthetics of the graphics (R Graph Gallery, 2018) by customizing the chart with a title, annotation, or using faceting.

D3.js is a JavaScript library used to create custom-made, interactive charts and maps on the web (Cook, 2023). Creating a chart with D3.js requires experience with JavaScript, HTML, SVG and CSS. This is because D3's features include data-driven modification of HTML and SVG elements. D3.js supports loading and transforming data (e.g. CSV data) and generating complex charts such as treemaps and networks.

It is seen that increasingly many applications of programming using Python involve large amounts of data (Gruppetta, 2024). Matplotlib and seaborn are some popular Python packages used for data visualizations. Matplotlib is a library for creating static, animated, and interactive visualizations in Python. Seaborn is a library built on top of matplotlib for making statistical graphics in Python. Seaborn helps to explore and understand the data.

Figure 6 demonstrates a word cloud analysis of data visualization tools indicating that Tableau is a popular data visualization software followed by packages in Python such as seaborn and matplotlib and closely followed by D3.js.

FIGURE 6
WORD CLOUD ANALYSIS OF DATA VISUALIZATION TOOLS

THEME 3: CONTEMPORARY CHART TYPES

Advanced chart types have been developed beyond line and bar charts. Some of the contemporary chart types are doughnut charts, chord diagram, sankey diagram, and violin plot. Table 6 illustrates the use of these chart types.

Doughnut charts are similar to pie charts. They show the relationship of parts to a whole (R Graph Gallery, 2018). Chord diagrams are ideal for comparing similarities within a dataset (Severino, 2023). This type of diagram visualizes the inter-relationships. In a chord diagram, nodes are arranged along a circle, with relationships between points connected either using arcs or Bézier curves (Severino, 2023). Each connection is represented proportionally by the size of each arc. However, over-cluttering can become an issue with chord diagrams when there are too many connections.

A Sankey diagram is a flowing diagram that shows the flow of things within a given scenario. Sankey diagrams represent the flow and use of resources, materials, or costs (Stafford, 2019). The Sankey diagram enables the discovery of insights from moving data within the shortest time possible. Flow arrows or lines can combine or split apart at each stage of a process. Color can be used to show the transition from one state of the process to another.

Violin plots are used when you want to observe the distribution of numeric data (Yi, 2021). According to Severino (2023), this chart is a combination of a box plot and a density plot that shows the distribution shape of the data. Severino (2023) further states that although box plots tend to hide significant details about how values in the data are distributed, violin plots display more information. This can make violin plots visually noisier than box plots. While violin plots may look overwhelming at first, they are easy to read. The widest part of the violin shows the highest probability. The tails of each group's density curve can be compared to see where groups are similar or different (Yi, 2021). Unlike bar charts, violin charts can be plotted starting with a number higher than zero.

Network Visualization, which is also called Network Graph, is often used to visualize complex relationships between a huge number of elements. This type of visualization illustrates relationships between entities displayed as round nodes with lines showing the relationships between them. The exhibit of network nodes can emphasize data discrepancies.

TABLE 6
COMPARISON OF CONTEMPORARY CHART TYPES

Chart Type	**Function**
Doughnut charts	show the relationship of parts to a whole.
Chord Diagram	ideal for comparing similarities within a dataset.
Sankey diagram	enables the discovery of insights from moving data within the shortest time possible.
Violin Plot	Violin plots are helpful if you want to look at a set of data values for a category and analyze the highest, lowest, and most probable value
Network Visualization	used to visualize complex relationships between a huge number of elements.

THEME 4: DATA VISUALIZATION CHALLENGES

Figure 7 illustrates major challenges in data visualization such as data complexity and quality, uncertainty and misleading visualizations. For instance, large amounts of data can make it difficult to visualize. Data quality is another vital issue. Moreover, uncertainty and misleading visualizations can impose challenges to decision making.

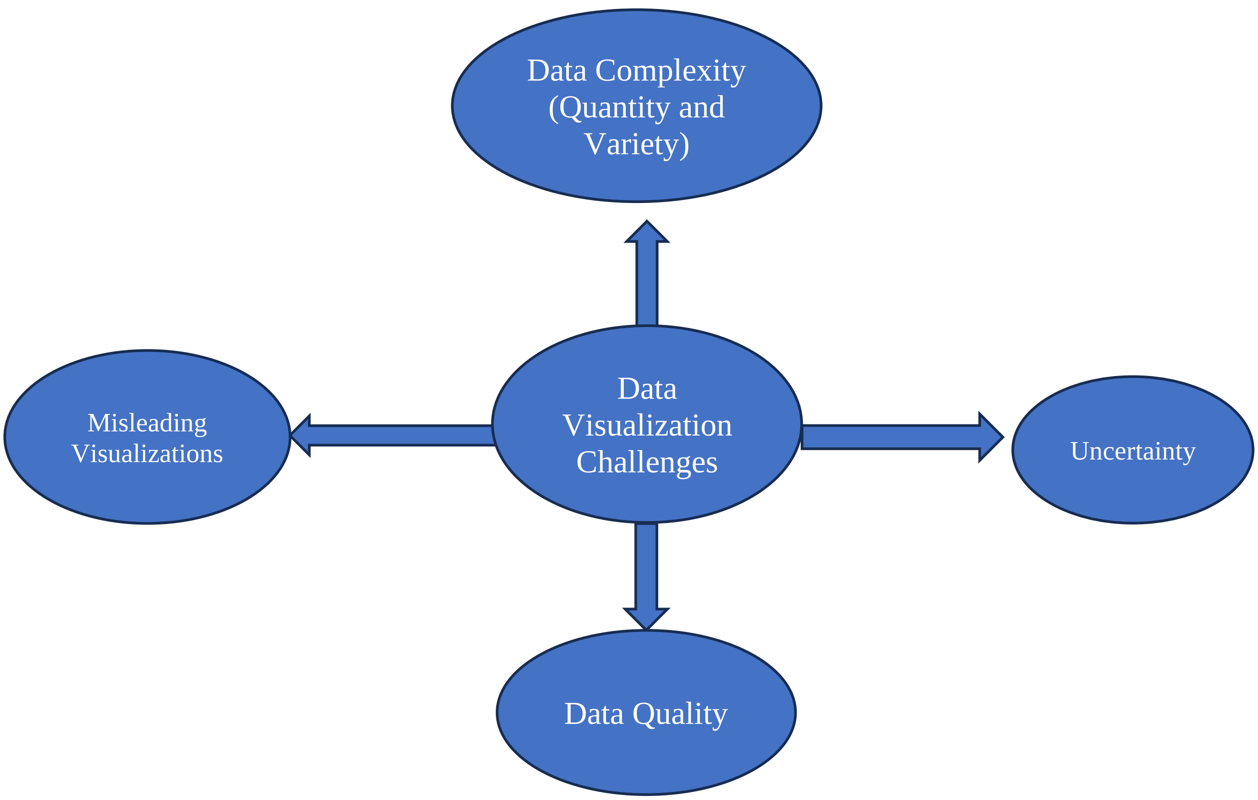
Although the purpose of visualization is to make sense of the enormous amount data, this can impose a problem for data exploration and analysis. Charts should not overwhelm the viewer or hide significant details behind a wall of clutter. This problem has received a great deal of attention by the data visualization research community. Methods are being explored and developed to fix this problem by reducing the amount of data to avoid loss of meaning. Novel approaches are applied to reduce clutter in the visualizations by positioning of data objects, better uses of color, or other visual attributes such as transparency. Tufte (1997) introduces the concept of maximizing the data-ink ratio, where every piece of ink in a chart should represent data. This is a good area for future research in data visualization.

Another challenging area is misleading visualizations. (Geidner et al., 2017) found that bar charts with a truncated y-axis were found to be significantly less credible. Visualizations are only as good as inputting data to create them (Stobierski, 2023). Cairo (2015) states that reasons for misleading visualizations can include lack of information or skills by the visual designer. When data is incorrect, it can lead to misleading visualizations. This in turn can have serious business implications. Misleading data visualization might lead to erroneous conclusions and poor business choices that may not be in the best interest of the organization.

According to Tufte (1997), we must aim for some uniformity in graphics with assurance that perceivers can get the numbers right. It is important to avoid misleading data visualizations to make better decisions based on reliable sources of information. Although previous authors have examined reasons for misleading visualizations, for future, exploring ways to construct visualizations that are both useful and trustworthy can be examined.

Data quality is another critical issue. Missing or invalid data is commonly produced by improper data entry or a lack of system interoperability (Gotz and Borland, 2016). Data structuring is particularly impactful if it changes the data characterization to generate views in the visualization design. Even when a data change is seemingly innocuous and does not change the overall data mapping, it may affect the implementation stage. For example, changing a column name or unit symbol may disrupt data parsers.

FIGURE 7
DATA VISUALIZATION CHALLENGES



Lastly, uncertainty is a challenging problem that cannot be ignored; otherwise, it may lead to erroneous or ambiguous decision-making. It is a challenging and complex concept (MacEachren et al., 2012) and has emerged as an essential component of data science and analytics. Uncertainty can arise at any stage of the visualization process in collection, transformation, or visualization (Pang et al., 1997). Uncertainty can result from data sampling, modeling, and visualization (Bonneau et al., 2014). It is vital to establish the source of uncertainty to represent uncertainty accurately (Masalonis et al., 2004).

DISCUSSION

Table 7 lists some industry applications of data visualizations (Rawat, 2021). More commonly seen industry applications are in retail, healthcare, finance, marketing, education and e-commerce. For example, in the retail industry, structured data visualizations can help businesses deliver a superior customer experience, improve inventory planning, and exceed revenue targets.

In the healthcare industry, data visualizations help deliver patient outcomes. Data visualization helps to perform tasks such as analyzing patient demographics, monitoring various healthcare facilities and evaluating how accessible pharmacies were. Another industry that benefits from the data visualization process is retail. Companies use data visualization to gain insight into customer behavior to better understand their needs and plan the product supply accordingly.

The finance industry comprises of huge volumes of data. These organizations can use visualization tools to notice industry trends, perform efficiently, and offer competent services. However, simply collecting and processing information is not enough. It is also important to provide actionable insights to the audience so that they can understand the information being presented. Financial data visualization allows the presentation of data insights to stakeholders in a way they can access.

TABLE 7
DATA VISUALIZATION APPLICATIONS BY INDUSTRY

Industry	Applications
Retail	Analyze sales performance for product categories and plan inventory
Healthcare	Visualizing treatment trends of a patient for a disease
Finance	Stock trend analysis and financial decision making
Marketing	sales analysis, market research analysis, customer analysis, defect analysis, cost analysis, and forecasting.
Education	The ability to monitor students' progress throughout the semester, allowing advisers to act quickly with outreach to failing students.
Ecommerce	Consumer trends, providing you with knowledge to help you attract new clients and close sales

CONCLUSION

Data visualization helps analyze and interpret large and complex data (Sadiku et al., 2016). This study demonstrates that data visualization plays a significant role in decision-making. The literature review provided an in-depth analysis of the data visualization techniques.

Due to its benefits, data visualization is now applied in many fields of study such as healthcare, finance, retail, and hospitality. Different activities will require different data visualization tools. For example, patterns of customer behavior, tracking logistical information and formulating schedules will require different tools. Some data visualization tools for instance focus on specific types of charts such as a map.

Communication of data is effective using data visualization, and various tools are available to create valuable graphics (Stobierski, 2023). There are several criteria to consider when choosing data visualization tools. Before narrowing down the data visualization tool choices, organizations and decision-makers must decide what needs to be pulled from the available data (Kaelin, 2022).

For future research, one can look at the topics presented in the data visualization challenges section in this paper. The main challenges seen with data visualization are misleading visuals and uncertainty followed by data quality and complexity.

REFERENCES

Ajibade, S., & Adediran, A. (2016). An overview of big data visualization techniques in data mining. *International Journal of Computer Science and Information Technology Research*, *4*(3), (105–113).

Begum, F. (2016). *Mapping disease: John Snow and Cholera*. Royal College of Surgeons of England Library and Publications. Retrieved from https://www.rcseng.ac.uk/library-and-publications/library/blog/mapping-disease-john-snow-and-cholera/

Bonneau, G.P., Hege, H.C., Johnson, C.R., Oliveira, M.M., Potter, K., Rheingans, P., & Schultz, T. (2014). Overview and state-of-the-art of uncertainty visualization. In *Scientific Visualization*, Springer, pp. 3–27

Borkin, M.A., Vo, A.A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2306–2315.

Cairo, A. (2015). Chapter 5. Graphics Lies, Misleading Visuals Reflections on the Challenges and Pitfalls of Evidence-Driven Visual Communication. In D. Bihanic (Ed.), *New Challenges for Data Design*. Springer-Verlag, London.

Cayne, B. (1987). *The New Lexicon Webster's Dictionary of the English Language*. Lexicon Publications, New York, NY, 1149 pages.

Chien, T.W., Lin, Y.F., Chang, C.H., Tsai, M.T., & Uen, Y.H. (2012). Using a bubble chart to enhance adherence to quality-of-care guidelines for colorectal cancer patients. *European Journal of Cancer Care*, *21*(6), 712–721.

Cleveland, W., McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, *79*(387), 531–554.

Cook, P. (2023). *Introduction to D3.js. D3 in Depth*. Retrieved from https://www.d3indepth.com/introduction/

Dougherty, J., & Ilyankou, I. (2024). *Hands-On Data Visualization. Interactive Storytelling from Spreadsheets to Code.* O'Reilly.

Dur, B. (2014). Data visualization and infographics in visual communication design education at the age of information. *Journal of Arts and Humanities (JAH)*, *3*(5).

Entango. (2023). *Improving Data Visualization with Gestalt Principles*. Medium. Retrieved from https://medium.com/@entango/improving-data-visualization-with-gestalt-principles-b316f474ea07

Flexner, S. (1987). *The Random House Dictionary of the English Language* (2nd Ed.). Unabridged. Random House, New York, NY.

Geidner, N., & Cameron, J. (2017). Readers perceive deceptive graphics as less credible. *Newspaper Research Journal*, *38*(4), 473–83.

Gotz, D., & Borland, D. (2016). Data-Driven Healthcare: Challenges and Opportunities for Interactive Visualization. *IEEE*, pp. 90–96

Gour, R. (2019). Top 5 BI Tools widely used for Data Visualization. *Towards Data Science Medium*. Retrieved from https://towardsdatascience.com/top-5-bi-tools-that-you-must-use-for-data-visualization-7ccc2a852bd3

Gruppetta, S. (2024). *The Python Coding Book: A relaxed and friendly programming textbook for beginners*.

Kaelin, M. (2022). *Best data visualization tools and software 2022*. ITBlendHub. Retrieved from https://itblendhub.com/article/data-visualization-tools/

Keahey, T.A (2013). Using Visualization to understand big data. *IBM*, pp. 1–15.

Keim, D.A., Hao, M.C., Dayal, U., Janetzko, H., & Bak, P. (2010). Generalized Scatter Plots. *Information Visualization*, *9*(4), 301–311.

Khaled, A. (2020). Data Visualization with ggplot2. *Towards Data Science Medium*. Retrieved from https://towardsdatascience.com/data-visualization-with-ggplot2-db04c4956236
Khan, M., & Khan, S. (2011). Data and Information Visualization Methods and Interactive Mechanisms: A Survey. *International Journal of Computer Applications*, (0975–8887), *34*(1), 1–14.
Kim, D., Setlur, V., & Agrawala, M. (2021). Towards Understanding How Readers Integrate Charts and Captions: A Case Study with Line Charts. *CHI '21: CHI Conference on Human Factors in Computing Systems.*
Kosara, R. (2016). Presentation-Oriented Visualization Techniques. *IEEE Computer Graphics and Applications*, *36*, 80–85.
Kuznetsova, A. (2023). *Gestalt Principles in Data Visualization*. Medium. Retrieved from https://nastengraph.medium.com/gestalt-principles-in-data-visualization-a4e56e6074b5
Liu, Y., Hu, M., Zhang, R., Xu, T., Wang, Y., & Zhou, Z. (2022). Visual aggregation of large multivariate networks with attribute-enhanced representation learning, *Neurocomputing*, *494*, 320–335.
MacEachren, A.M., Roth, R.E., O'Brien, J., Li, B., Swingley, D., & Gahegan, M. (2012) Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2496–2505.
Masalonis, A., Mulgund, S., Song, L., Wanke, C., & Zobell, S. (2004). Using probabilistic demand predictions for traffic flow management decision support. In *AIAA Guidance, Navigation, and Control Conference and Exhibit*, p.5231.
Page, M., Mckenzie, J., Bossuyt, P., Boutron, I., Hoffmann, T., Mulrow, C., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372.
Pang, A.T., Wittenbrink, C.M., & Lodha, S.K. (1997). Approaches to uncertainty visualization. *The Visual Computer*, *13*(8), 370–390.
Porras, G. (2023). *Visualizing data with purpose: A guide to applying gestalt principles and color theory*. Medium. Retrieved from https://medium.com/@giovannisolanoporras/visualizing-data-with-purpose-a-guide-to-applying-gestalt-principles-and-color-theory-ba9ccbe104c1
Power BI. (2023). *Tips for evaluating Data Visualization tools*. Microsoft Power BI. Retrieved from https://powerbi.microsoft.com/en-us/data-visualization-tools/
Rawat, A. (2021). *10 Applications of Data Visualization*. Analytic Steps. Retrieved from https://www.analyticssteps.com/blogs/10-applications-data-visualization
R Graph Gallery. (2018). *Ggplot 2*. R Graph Gallery. Retrieved from https://r-graph-gallery.com/ggplot2-package.html
Sadiku, M., Shadare, A., Musa, S., & Akujuobi, C. (2016). Data Visualization. *International Journal of Engineering Research and Advanced Technology*, *2*(12), 11–16.
SAS. (2023). *Data Visualization Techniques*. SAS. Retrieved from https://www.sas.com/content/dam/SAS/bp_de/doc/whitepaper1/ba-wp-data-visualization-techniques-2252344.pdf
Severino, R. (2023). *The Data Visualization Catalogue*. Retrieved from https://datavizcatalogue.com/methods
Spence, I. (2005). No Humble Pie: The Origins and Usage of a Statistical Chart. *Journal of Educational and Behavioral Statistics*, *30*.
Stafford, A. (2019). The What, Why, and How of Sankey Diagrams. *Towards Data Science Medium*. Retrieved from https://towardsdatascience.com/the-what-why-and-how-of-sankey-diagrams-430cbd4980b5
Stobierski, T. (2023). *Top data visualization tools for business professionals*. HBS. Retrieved from https://online.hbs.edu/blog/post/data-visualization-tools
Tableau. (2023). *What is Data Visualization? Definition, Examples, and Learning Resources.* Retrieved from https://www.tableau.com/learn/articles/data-visualization

Teets, J., Tegarden, D., & Russell, R. (2010). Using Cognitive Fit Theory to Evaluate the Effectiveness of Information Visualizations: An Example Using Quality Assurance Data. *IEEE Transactions on Visualization and Computer Graphics*, *16*, 841–853.

Tegarden, D.P. (1999). Business information visualization. *Communications of the AIS, 1*(4).

Telea, A. (2015). *Data Visualization Principles and Practice* (Second Edition). CRC Press.

Tufte, E. (1997). *The Visual Display of Quantitative Information* (2nd Edition). Graphics Pr.

Vessey, I. (1991). Cognitive Fit: A Theory-Based Analysis of the Graphs Versus Tables Literature. *Decision Sciences*, *22*(2), 219–240

Viégas, F.B., & Wattenberg, M. (2007). Artistic data visualization: Beyond visual analytics. In *Proceedings of the 2nd International Conference on Online Communities and Social Computing* (pp. 182–191). Berlin, Heidelberg: Springer-Verlag.

Visual Paradigm. (2023). How to Create a Storyboard of Wireframes? *Visual Paradigm User's Guide*. Retrieved from https://www.visual-paradigm.com/support/documents/vpuserguide/2822/3406/85137_creatingawir.html

Wickham, H. (2010). A Layered Grammar of Graphics. *Journal of Computational and Graphical Statistics*, *19*(1), 3–28.

Wilkinson, L. (1999). *The Grammar of Graphics.* New York: Springer.

Yi, M. (2021). *A Complete Guide to Violin Plots*. ChartIO. Retrieved from https://chartio.com/learn/charts/violin-plot-complete-guide/