REGIONAL COLLEGE OF MANAGEMENT AND ENTREPRENEURSHIP

IV SEMESTER PROJECT

DATA VISUALIZATION

Project Report submitted in partial fulfilment of the requirement

for the award of degree of

POST GRADUATE DIPLOMA IN MANAGEMENT

BY

AICTE

SUBMITED

BY

RATHNA KUMARI NARAYANAPARAPU

SPECIALIZATION: BUSINESS ANALYTICS

ROLL NO: RCME/2023/0130

REGISTRATION NO: 143680802404

Under the guidance of

PROF.TAPAS PANDA

Plot No 34/4 and 34/5 Devanahalli, Mudugurik

Bangalore Karnataka 562110, 2023-2025.

DECLARATION BY THE STUDENT

I**, RATHNAKUMARI NARAYANAPARAPU** the undersigned, here by declare that the project report entitled **"DATA VISUALIZATION"** is the result of the project work carried out by me under the guidance of Professor **"TAPAS PANDA"** in partial fulfilment for the award of **POST GRADUATE DIPLOMA IN MANAGEMENT** by AICTE.

I also declare that this project is the outcome of my own efforts and that it has not been submitted to any other university or Institute for the award of any other degree or Diploma or Certificate.

Date:26/03/2025 **Name: RATHNAKUMARI NARAYANAPARAPU**

Place: BANGALORE **Registration: 143680802404**

GUIDE CERTIFICATE

This is to certify that the project on **"DATA VISUALIZATION** submitted by **RATHNAKUMARI NARYANAPARAPU,** Registration No: **143680802404** to AICTE, for the award of diploma of **POST GRADUATE DIPLOMA IN MANAGEMENT** is a record of independent research work carried out by her under my guidance.

The report is the result of her efforts and endeavours. This has not been submitted to any other University or Institution for award of any degree/diploma certificate.

Place: BANGALORE **Signature of Guide:**

Date: **Name of the guide: TAPAS PANDA**

REGIONAL COLLEGE OF MANAGEMENT AND ENTREPRENEURSHIP (Approved by AICTE, Ministry of Education, Govt. of India)

CERTIFICATE OF ORIGINALITY

This is to certify that the 4th Semester Project report entitled **"DATA VISUALIZATION"** is an original work of **RATHNA KUMARI NARAYANAPARAPU** bearing Registration No **143680802404** and is being submitted in partial fulfilment for the award of Post Graduate Diploma in Management under **"REGIONAL COLLEGE OF MANAGEMENT AND ENTREPRENEURSHIP"** The report has not been submitted earlier either to this any University/institution for the fulfilment of the requirement of any course of the study. **RATHNA KUMARI NARAYANAPARAPU** is guided by Professor **TAPAS PANDA** who is the Guide as per the regulations of **REGIONAL COLLEGE OF MANAGEMENT AND ENTREPRENEURSHIP.**

SIGNATURE AND SEAL OF DIRECTOR/ PRINCIPAL/HOD

Regional College of Management and Entrepreneurship, Survey No. 34/4 34/5, Mudugurki, Devanahalli, Bangalore – 562110, Karnataka.

Website: www.rcmb.in

ACKNOWLEDGEMENT

I, **RATHNA KUMARI NARAYANAPARAPU**, take the opportunity to express my gratitude to all of them who in some way helped me to accomplish this project entitled "**DATA VISUALIZATION**". Without their encouragement, active guidance, and co-operation this project would not have been a success. I would like to gratefully acknowledge **AICTE** for this opportunity, **REGIONAL COLLEGE OF MANAGEMENT AND ENTREPRENUERSHIP**, for facilitating the completion of this project.

I am extremely thankful and pay my gratitude to my guide Professor **TAPAS PANDA** for his active guidance in successfully completing this project.

I also express my deep gratitude to my parents and those who have contributed for the timely completion of this project.

Place: BANGALORE **Name:** RATHNA KUMARI NARAYANAPARAPU

Date: **Registration no: 143680802404**

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Data visualization sits at the confluence of advances in technology, the study of human cognition and perception, graphical interfaces, widespread adoption of standards for rich Internet applications, and the continuing expansion of interest and experience in analytics and data discovery. Data visualization can contribute significantly to the fruitful interpretation and sharing of insights from analytics, enabling nontechnical subject matter experts to perform data discovery in a self-directed fashion. Implementation of chart engines and the growth in the number and variety of visualizations available in graphics libraries are supporting new sophistication in visual analysis, allowing users to go beyond simple bar and pie charts to express more advanced insights about quantitative information.

Users need data visualization for a variety of BI and analytics activities, including reporting, scorecards, operational alerting, and data discovery and analysis. Rather than just giving users “new toys” to play with, organizations should examine how they can match visualization technologies and practices to user requirements.

Across the board, however, a key element in the success of visualization is data interaction; users need broad capabilities for manipulating data, including to drill down, cross cut, slice, and dice data directly from graphical interfaces.

For many organizations, dashboards take the centre stage for data visualizations, especially for business intelligence reporting and performance management. Many users would like to consolidate views of multiple sources and types of information into their dashboard workspaces. One new source of interest is geographical information. Although using maps to enhance corporate data (and vice versa) is not yet widespread, organizations in a growing number of industries are interested in geospatial analysis as an addition to their visualization repertoire.

CHAPTER-01
INTRODUCTION

1.1 INTRODUCTION

Data visualization is the visual and graphic representation of data through charts, graphs, maps, interactive dashboards, and other visual imagery. Data visualization can help simplify large or complex datasets to encourage broader interest on an issue or topic. Common types of data visualization include pie charts, bar charts, maps, tables, and line charts. Dashboards or infographics may feature multiple data visualizations.

Tools like Cube play a significant role in enhancing data visualization by providing a unified semantic layer that ensures consistency across all data sources. Whether you're using BI tools, spreadsheets, or dashboards, Cube helps deliver accurate, secure, and reliable visualizations that empower your team to make informed decisions with confidence.

Data Visualization Software -Data visualization software is a specialized tool designed for data visualization and analysis. Examples of data visualization software include Tableau, QlikView, and Power BI. These tools provide advanced data visualization capabilities, including interactive dashboards, heat maps, and network diagrams.

The ability to visualize data is crucial to scientific research. Today, computers can be used to process large amounts of data. Data visualization is concerned with the design, development, and application of compute generated graphical representation of the data. It provides effective data representation of data originating from different sources. This enables decision makers to see analytics in visual form and makes it easy for them to make sense of the data. It helps them discover patterns, comprehend information, and form an opinion.

Data visualization is also regarded as information visualization or scientific visualization. Human beings have always employed visualizations to make messages or information last in time. What cannot be touched, smelled or tasted can be represented visually.

Data Visualization Made Simple is a practical guide to the fundamentals, strategies, and real-world cases for data visualization, an essential skill required in today's information-rich world. With foundations rooted in statistics, psychology, and computer science, data visualization offers practitioners in almost every field a coherent way to share findings from original research, big data, learning analytics, and more.

Imagine you are Nike, Netflix, Amazon, or Twitter. Your data helps these companies better understand you and other users like you. Companies utilize this information to target markets, develop new products, and ultimately outpace their competition by knowing their customers' habits and needs. However, such insights do not just "automagically" happen.

"Data visualization is the presentation of data in a pictorial or graphical format, and a data visualization tool is the software that generates this presentation. Data visualization provides users with intuitive means to interactively explore and analyse data, enabling them to effectively identify interesting patterns, infer correlations and causalities, and supports sense-making activities."

1.1 NEED OF THE STUDY

Data visualization means drawing graphic displays to show data. Sometimes every data point is drawn, as in a scatterplot, sometimes statistical summaries may be shown, as in a histogram. The displays are mainly descriptive, concentrating on 'raw' data and simple summaries. They can include displays of transformed data, sometimes

based on complicated transformations. One person's statistics may be another person's raw data. As with other aspects of working with graphics, it would be useful to have an agreed base of concepts and terminology to build on.

The main goal is to visualize data and statistics, interpreting the displays to gain information. Data visualization is useful for data cleaning, exploring data structure, detecting outliers and unusual groups, identifying trends and clusters, spotting local patterns, evaluating modelling output, and presenting results. It is essential for exploratory data analysis and data mining to check data quality and to help analysts become familiar with the structure and features of the data before them. This is a part of data analysis that is underplayed in textbooks, yet ever-present in actual investigations. Look, for instance, at the one-sided peaks in the distributions of marathon finishing times (marastats, 2019).

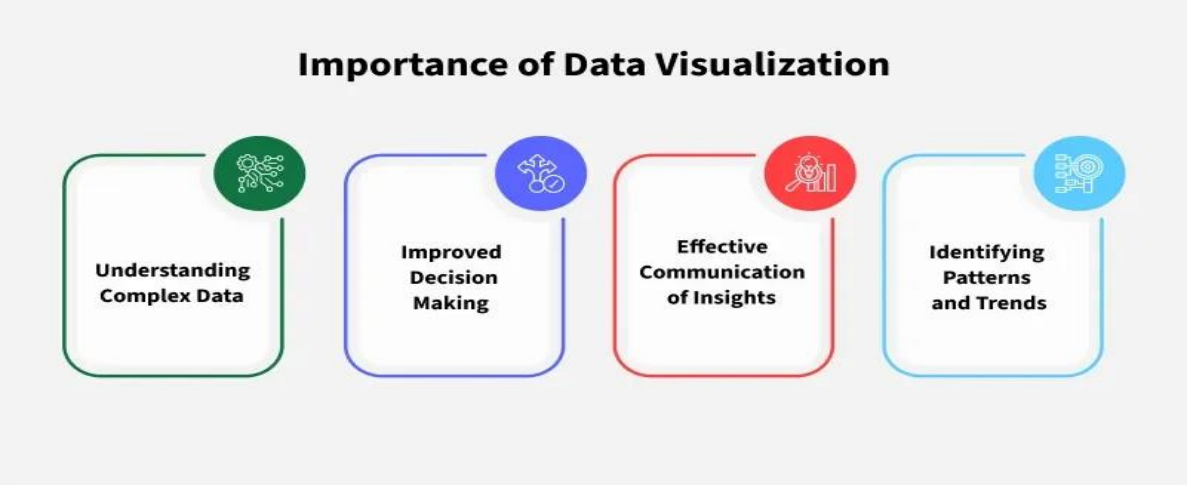
Graphics reveal data features that statistics and models may miss: unusual distributions of data, local patterns, clustering's, gaps, missing values, evidence of rounding or heaping, implicit boundaries, outliers, and so on. Graphics raise questions that stimulate research and suggest ideas. It sounds easy. In fact, interpreting graphics needs experience to identify potentially interesting features and statistical nous to guard against the dangers of overinterpretation. Just as graphics are useful for checking model results, models are useful for checking ideas derived from graphics (for more on models, see Hand, 2019).

This overview concentrates on static graphics. Dynamic graphics and, more especially, interactive graphics are in an exciting stage of development and have much to add. They require an article of their own. Superb examples include Human Terrain, a dynamic graphic showing the world's population in 3-D, and the interactive Name

Voyager. ‘A Picture Is Worth a Thousand Words’ Famous sayings have a way of developing a life of their own.

A picture is not a substitute for a thousand words; it needs a thousand words (or more). What Is Important in Data Visualization? 3 the data, how and why they were collected, whether more could be collected, the reasons for drawing the displays, and how people with the necessary background knowledge advise they might be interpreted. There is a story that M. G. Kendall reviewed a book of R.A. Fisher's with the words: "No one should read this book who has not read it already." It is like that with graphics. If you have read all the supporting text, the display is often memorable and readily understandable. If you have not, it is not. Graphics on their own are insufficient, they are part of a whole.

They complement text and are complemented by text. Student's reanalysis of the Lanarkshire Milk Experiment (Student, 1931) is an excellent example (and is also interesting as an early analysis of a large data set). The potential synergy of text and graphics can be appreciated by talking through your own graphics, explaining them to others. Why have you drawn those graphics? How have you drawn them? What can be seen? Are there interesting patterns? What could be changed and improved? Which other graphics might be drawn? How can conclusions be checked? There should be more talking about graphics and less relying on the graphics to speak for themselves. When it comes to graphics you have not drawn yourself, the same kinds of questions are still relevant, although they may be more difficult to answer.



1.2 OBJECTIVE

Effective data visualization is properly sourced, contextualized, simple and uncluttered. The underlying data is accurate and up-to-date to make sure that insights are reliable. Graphical items are well-chosen for the given datasets and aesthetically appealing, with shapes, colours and other visual elements used deliberately in a meaningful and non-distracting manner.

- The visuals are accompanied by supporting texts (labels and titles). These verbal and graphical components complement each other to ensure clear, quick and memorable understanding. Effective information visualization is aware of the needs and concerns and the level of expertise of the target audience, deliberately guiding them to the intended conclusion.

- Such effective visualization can be used not only for conveying specialized, complex, big data-driven ideas to a wider group of non-technical audience in a visually appealing, engaging and accessible manner, but also to domain experts and executives for making decisions, monitoring performance, generating new ideas and stimulating research.

- In addition, data scientists, data analysts and data mining specialists use data visualization to check the quality of data, find errors, unusual gaps and missing values in data, clean data, explore the structures and features of data and assess outputs of data-driven models.

- In business, data and information visualization can constitute a part of *data storytelling*, where they are paired with a coherent narrative structure or storyline to contextualize the analysed data and communicate the insights gained from analysing the data clearly and memorably with the goal of convincing the audience into making a decision or taking an action in order to create business value. This can be contrasted with the field of statistical graphics, where complex statistical data are communicated graphically in an accurate and precise manner among researchers and analysts with statistical expertise to help them perform exploratory data analysis or to convey the results of such analyses, where visual appeal, capturing attention to a certain issue and storytelling are not as important.

The field of data and information visualization is of interdisciplinary nature as it incorporates principles found in the disciplines of descriptive statistics (as early as the 18th century), visual communication, graphic design, cognitive science and, more recently, interactive computer graphics and human-computer interaction.

Since effective visualization requires design skills, statistical skills and computing skills, it is argued by authors such as Gershon and Page that it is both an art and a science.

The neighbouring field of visual analytics marries statistical data analysis, data and information visualization and human analytical reasoning through interactive visual interfaces to help human users

reach conclusions, gain actionable insights and make informed decisions which are otherwise difficult for computers to do.

Research into how people read and misread various types of visualizations is helping to determine what types and features of visualizations are most understandable and effective in conveying information.

Visualization is the first step to make sense of data. To translate and present complex information and relations in a simple way, analysts use different methods of data visualization — charts, diagrams, maps, etc.

Choosing the right technique and its setup is often the only way to make data understandable. Vice versa, poorly selected tactics won't let you unlock the full potential of data or even make it irrelevant.

"The success of data visualization is due to the soundness of the basic idea behind it: the use of computer-generated images to gain insight and knowledge from data and its inherent patterns and relationships. A second premise is the utilization of the broad bandwidth processes, and simulations involving data sets from diverse scientific disciplines and large collections of abstract data from many sources."

Data visualization can help tell the story, conveying complex issues clearly. It can play a key role for identifying the significant information from the noise, including outliers and anomalies.

It can help you with your growing volume of data. Visual interaction with large data sets can simplify analysis, revealing new business insights.

Data visualization can help you do all that—if you have the right tool. So, what should you look for? A number of factors should be considered.

A tool with that capability should have the power to help you with all the steps in analysing and conveying information, starting with data

preparation. Traditionally, preparing data for analysis has been a manual process, often time-consuming, frustrating, and prone to Data Visualization is a vast and interdisciplinary approach that incorporates elements from statistics, psychology, design, and computer science.

1.3 SCOPE OF THE STUDY

Data Visualization is a vast and interdisciplinary approach that incorporates elements from statistics, psychology, design, and computer science. According to Keim, Mansmann, and Thomas, visual analytics is a subset of Data Visualization that blends automated data analysis with the cognitive abilities of humans in order to make sense of large data sets (Keim et al., 2006).

This is an interaction between human reasoning with machine-generated data that can enable a business to manage their complex information. Visual analytics is important in areas like business intelligence (BI) and big data. This must deal with sheer volume and complexity of information that conducting traditional analysis would be inefficient.

The historical development of Data Visualization has influenced many of its current applications. As expressed by Few, early pioneers in the field such as Descartes and Playfair laid groundwork for modern visualization techniques by introducing coordinate systems and geographical forms. Today, Data Visualization no longer stays within the bounds of static representations. It now integrates interactive dashboards, real-time data feeds, and geospatial analysis.

Despite its growing promise, there are numerous challenges in the field of Data Visualization. One challenge is scalability. Data continues to grow exponentially so visualization tools must be able to handle larger datasets. Keim et al. (2006) notes that filtering, data reduction, and aggregation techniques can be essential when dealing with massive datasets. Another challenge that faces the field is ensuring accuracy and precision.

CHAPTER-02
LITERATURE REVIEW

2. LITERATURE REVIEW

Data visualization (Matthew N. O. Sadiku):

Data Visualization involves presenting data in graphical form which make information easy to understand. It explains the summary of any vast data. It will benefit any field of study that requires innovative ways of presenting large and complex information to understand the underlying data. Advanced computer graphics has reshaped data visualization.

Narrative Scientific Data Visualization in an Immersive Environment by: (Richen Liu, Hailong Wang, Chuyu Zhang, Xiaojian Chen, Lijun Wang, Genlin Ji, Bin Zhao, Zhiwei Mao, Dan Yang; (2021))

Narrative visualization for scientific data studies can help users better understand the domain knowledge because narrative visualizations frequently present a sequence of data and observations linked together by a unifying theme or argument. Narrative visualization in immersive surroundings can provide users with an intuitive experience to interactively explore scientific data because immersive surroundings give a brand-new strategy for interactive scientific data visualization and exploration.

DeepEye: Towards Automatic Data Visualization (Luo, Yuyu, Qin, Xuedi, Tang, Nan, Li, Guoliang; (2018)):

Data Visualization is extremely useful to explain the significance of data to people who are visually oriented. The main task of automatic data visualization is, given a dataset to visualize its storytelling by transforming the data (e.g. selecting attributes, grouping and binning values) and deciding which visualization (e.g. bar, pie, or line charts) is the right type to a particular data. In this paper, they present Deep Eye: A novel system for automatic data visualization that tackles three

problems Visualization Recognition Visualization Ranking Visualization Selection.

The Case for a Learned Sorting Algorithm by (Ani Kristo, Kapil Vaidya, Ugur Çetintemel, Sanchit Misra, Tim Kraska; (2020)):

Sorting is one of the most basic algorithms in the field of Computer Science. Its operation is not just for sorting the data, but also a part of joins. In this paper, they introduced a new type of distribution sort that leverages the learning model of the eCDF (i.e. empirical Cumulative Distribution Function). Algorithms which used in a model to effectively get an approximate value of the scaled empirical CDF for each record key and map.

Seaborn: statistical data visualization by (Michael L. Waskom; (2021)):

Seaborn is a high statistical graphical library in Python used for data visualization. It provides a high interface to the matplotlib library. It integrated with the panda's data structure. When plotting the dataset has to make, seaborn astronomically maps the data values to visualize the colour, size, and style.

Data Visualization Techniques for real-time information -A Custom and Dynamic Dashboard for Analyzing Surveys' Results by (Renato Toasa, Marisa Maximiano, Catarina Reis, David Guevara; (2020)):

Implement a generic and dynamic dashboard based on real-time information to assess the impact of the Data Visualization Techniques. The dashboard helps to interact with users based on initial set and existing set of Data Visualization Techniques.

Visualization of Big Data Text Analytics in the Financial Industry (Živko Krstić, Sanja Seljan, Jovana Zoroja; (2019)):

Textual data can decide new insights and bring precious business insights. Sources that are used for text analysis in financial industry.

CHAPTER -03
RESEARCH METHODOLOGY

3. RESEARCH METHODOLOGY

3.1 SOURCE OF DATA

Data visualization research methodology involves a systematic approach to transforming raw data into meaningful visual representations that enhance comprehension and decision-making. It begins with defining clear objectives—whether the visualization aims to explore trends, communicate insights, or support decision-making. The next step involves data collection and preparation, ensuring accuracy, completeness, and proper structuring for visualization. Data visualization relies on various sources to gather meaningful data for analysis and presentation. These sources can include official reports and datasets from government agencies and financial organizations like the World Bank and IMF, which provide extensive and reliable data for visualization. Additionally, tools such as Tableau, Power BI, Google Data Studio, and Excel enable users to generate insights through built-in visualization capabilities. Public APIs also serve as valuable sources, allowing access to real-time and historical data from platforms like Open Weather for weather trends or Yahoo Finance for stock market data. Moreover, businesses rely on intelligence platforms such as SAP Analytics Cloud, Looker, and Domo to visualize enterprise data for strategic decision-making.

There are different methods to gather data, all of which fall into two categories: **primary data source** and a **secondary data source**.

3.1.1 PRIMARY DATA

The term primary data refers to the data originated by a researcher himself, while secondary data is the already existing data collected by agencies and organizations for the purpose of conducting an analysis. Primary data sources can include surveys, observations, questionnaires, experiments, personal interviews, and more. The data from ERP (Enterprise Resource Planning) and CRM (Customer Relationship Management) systems can also be used as a primary source of data. In this project questionnaire method for survey is used for collection of primary data.

3.1.2 SECONDARY DATA

On the contrary, secondary data sources can be government publications, staging websites, publications from independent research labs, journal articles, etc. The transformed "raw" data set into another format, in the process of data wrangling, can also be seen as a secondary data source.

Secondary data can be a key concept in terms of data enrichment when the primary source data is not solid enough with information, and it can improve the precision of the analysis by adding more attributes and variables to the sampling.

3.2 SAMPLING

Data sampling is a statistical analysis technique used to select, process, and analyse a representative subset of a population. It is also used to identify patterns and extrapolate trends in an overall population.

With data sampling, researchers, data scientists, predictive modelers, and other data analysts can use a smaller, more manageable amount of data to build and run analytical models.

This allows them to more quickly produce accurate findings from a statistical population. For example, if a researcher wants to determine the most popular fruit in a country with a population of 100 million people, they would select a representative sample of N (e.g., 1000, 10,000).

3.2.1 SAMPLING TECHNIQUE

In Statistics the sampling method or sampling technique is the process of studying the population by gathering information and analysing that data. It is the basis of the data where the sample space is enormous. There are several different sampling techniques available, and they can be subdivided into two groups. All these methods of sampling may involve specifically targeting hard or approach to reach groups.

COMMON DATA VISUALIZATION USE CASES

- **Sales and marketing:** Market and consumer research firm eMarketer estimated $264 billion would be spent on U.S.-based digital advertising in 2023. That number is expected to cross the $390 billion mark by 2027. Given the size of the investment in advertising, marketing teams must pay close attention to their sources of web traffic and how their web properties generate revenue. Data visualization helps to illustrate how marketing efforts affect traffic trends over time.
- **Politics:** A common use of data visualization in politics is a geographic map that displays the candidates' states, counties or other geographic regions voted for.
- **Healthcare:** Healthcare professionals frequently use choropleth maps to visualize important health data. A choropleth map displays geographical areas or regions that are assigned a certain colour in relation to a numeric variable. Choropleth maps allow professionals to see how a variable, such as the mortality rate of heart disease, changes across specific geographic areas.
- **Scientists:** Scientific visualization, sometimes referred to as SciVi, allows scientists and researchers to gain greater insight from their experimental and other collected data.
- **Finance:** Finance professionals must track the performance of their investment decisions when choosing to buy or sell an asset. Candlestick charts are used as tools to help finance professionals analyse price movements over time, displaying information such as securities, derivatives, currencies, stocks, bonds and commodities. By analysing how prices have changed over time, data analysts and finance professionals can detect trends.
- **Logistics:** Shipping companies use visualization tools to create a data-driven supply chain and determine the most efficient shipping routes.
- **Data scientists and researchers:** Data professionals typically build visualization for their own use or to present the information to a select audience. They use visualization libraries of the chosen programming languages and tools. Data scientists and researchers frequently use opensource programming languages -- such as Python -- or proprietary tools designed for complex data analysis.

Data scientists and researchers use visualizations to get a greater understanding of data sets and to identify patterns or trends that might otherwise go unnoticed.

The science of data visualization

- The science of data visualization is based on an understanding of how humans gather and process information. Daniel Kahneman and Amos Tversky collaborated on research that defined two different methods for gathering and processing information.
- The first method focuses on thought processing that is fast, automatic and unconscious. This method is frequently used in day-to-day life and helps accomplish tasks including the following:
- Reading the text on a sign.
- Solving simple math problems, like 1+1.
- Identifying where a sound is coming from.
- Riding a bike.
- Determining the difference between colours.
- The second focuses on slow, logical, calculating and infrequent thought processing, as demonstrated by the following:
- Reciting a phone number.
- Solving complex math problems, like 132 x 154.
- Determining the difference in meaning between multiple signs standing side by side.
- Understanding complex social cues.

Why is Data Visualization Important?

- Let's take an example. Suppose you compile data of the company's profits from 2013 *to 2023* and create a line chart. It would be very easy to see the line going constantly up with a drop in just 2018. So, you can observe in a second that the company has had continuous profits in all the years except a loss in 2018. It would not be that easy to get this information so fast from a data table. This is just one demonstration of the usefulness of data

visualization. Let's see some more reasons why visualization of data is so important.

1. **Data Visualization Simplifies the Complex Data**
   Large and complex data sets can be challenging to understand. Data visualization helps break down complex information into simpler, visual formats making it easier for the audience to grasp. For example, in a scenario where sales data is visualized using a heat map on Tableau states that have suffered a net loss are coloured red. This visual makes it instantly obvious which states are underperforming.

2. Enhances Data Interpretation

Visualization highlights patterns, trends, and correlations in data that might be missed in raw data form. This enhanced interpretation helps in making informed decisions. Consider another Tableau visualization that demonstrates the relationship between sales and profit. It might show that higher sales do not necessarily equate to higher profits this trend that could be difficult to find from raw data alone. This perspective helps businesses adjust strategies to focus on profitability rather than just sales volume.

3. Data Visualization Saves Time

It is definitely faster to gather some insights from the data using data visualization rather than just studying a chart. In the screenshot below on Tableau it is very easy to identify the states that have suffered a net loss rather than a profit. This is because all the cells with a loss are coloured red using a heat map, so it is obvious states have suffered a loss. Compare this to a normal table where you would need to check each cell to see if it has a negative value to determine a loss. Visualizing Data can save a lot of time in this situation.

4. Improves Communication

Visual representations of data make it easier to share findings with others especially those who may not have a technical background. This is important in business where stakeholders need to understand data-driven insights quickly. Let see the below Tree Map visualization on Tableau showing the number of sales in each region of the United States with the largest rectangle representing California due to its high sales volume. This visual context is much easier to grasp rather than detailed table of numbers.

5. Data Visualization Tells a Data Story

Data visualization is also a medium to tell a data story to the viewers. The visualization can be used to present the data facts in an easy-to-understand form while telling a story and leading the viewers to an inevitable conclusion. This data story should have a good beginning, a basic plot, and an ending that it is leading towards. For example, if a data analyst has to craft a data visualization for company executives detailing the profits of various products, then the data story can start with the profits and losses of multiple products and move on to recommendations on how to tackle the losses.

The Power of Good Data Visualization

Data visualization involves the use of graphical representations of data, such as graphs, charts, and maps. Visuals allow data scientists to summarize thousands of rows and columns of complex data and put it in an understandable and accessible format. Compared to descriptive statistics or tables, visuals provide a more effective way to analyse data, including identifying patterns, distributions, correlations, and spotting outliers in complex datasets.

By bringing data to life with insightful plots and charts, data visualization is vital in decision-making processes. Whether it’s data analysts breaking down their findings to non-technical stakeholders,

data scientists performing A/B tests for marketing purposes, or machine learning engineers explaining potential bias in complex large language models like ChatGPT, data visualization is the key to moving from data insights to decision-making.

Despite the use of data visualization, many thorough and detailed data analyses still end up in the drawer for the simple reason that they didn't get to captivate the audience, whether decision-makers, stakeholders, or other members of the team.

Thanks to progress in disciplines like neuroscience, today, we know the way a data visualization is depicted can severely affect how people perceive it. The choices you make when designing a graph – for example, the colours, the layout, and the size– can make a big difference. Interested in the theory behind data visualization? Our Understanding Data Visualization course is a great place to get started.

While data visualization has an important role to play when communicating data insights, the recipe for successful communication is more complex. That's the idea behind data storytelling, an innovative approach that advocates for using visuals, narrative, and data to turn data insights into action. To know more about data storytelling, check out our Data Framed podcast, where we speak with Brent Dykes, Senior Director of Insights & Data Storytelling at Blast Analytics and author of Effective Data Storytelling.

Types of Data Visualization Analysis

Data visualization is used to analyse visually the behaviour of the different variables in a dataset, such as a relationship between data points in a variable or the distribution. Depending on the number of variables you want to study at once, you can distinguish three types of data visualization analysis.

- **Univariate analysis:** Used to summarize the behaviour of only one variable at a time.
- **Bivariate analysis:** Helps to study the relationship between two variables
- **Multivariate analysis:** Allows data practitioners to analyse more than two variables at once.

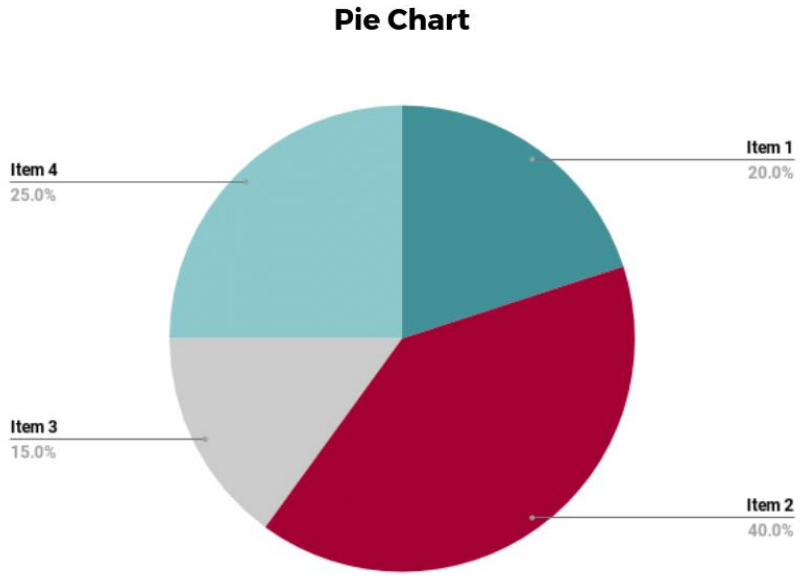
Data Visualization Techniques

The type of data visualization technique you leverage will vary based on the type of data you're working with, in addition to the story you're telling with your data.

Here are some important data visualization techniques to know:

- Pie Chart
- Bar Chart
- Histogram
- Heat Map
- Box and Whisker Plot
- Area Chart
- Scatter Plot
- Pictogram Chart
- Timeline
- Network Diagram
- Correlation Matrices

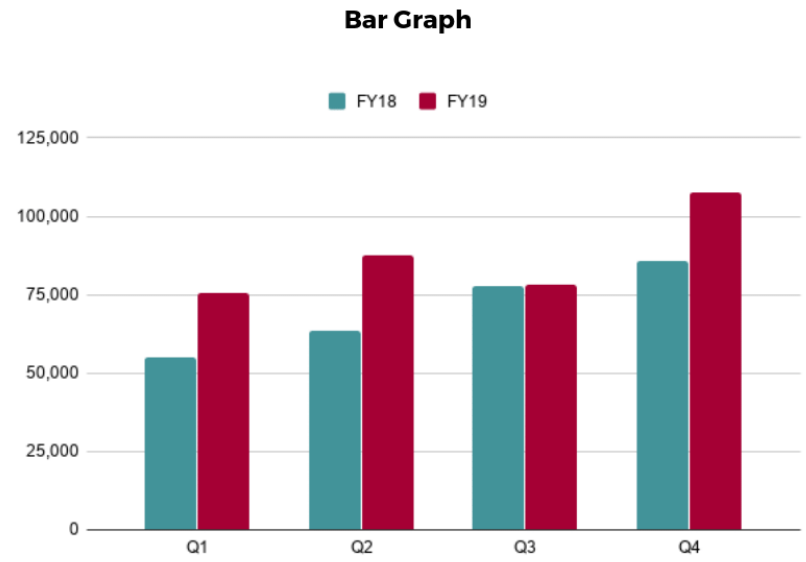
1. Pie Chart



Pie charts are one of the most common and basic data visualization techniques, used across a wide range of applications. Pie charts are ideal for illustrating proportions, or part-to-whole comparisons.

Because pie charts are relatively simple and easy to read, they're best suited for audiences who might be unfamiliar with the information or are only interested in the key takeaways. For viewers who require a more thorough explanation of the data, pie charts fall short in their ability to display complex information.

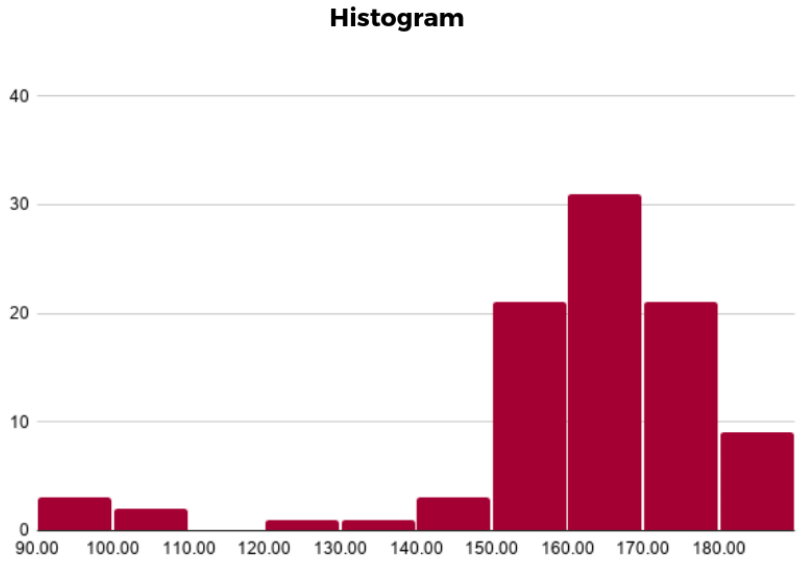
2. Bar Chart



The classic bar chart, or bar graph, is another common and easy-to-use method of data visualization. In this type of visualization, one axis of the chart shows the categories being compared, and the other, a measured value. The length of the bar indicates how each group measures according to the value.

One drawback is that labeling and clarity can become problematic when there are too many categories included. Like pie charts, they can also be too simple for more complex data sets.

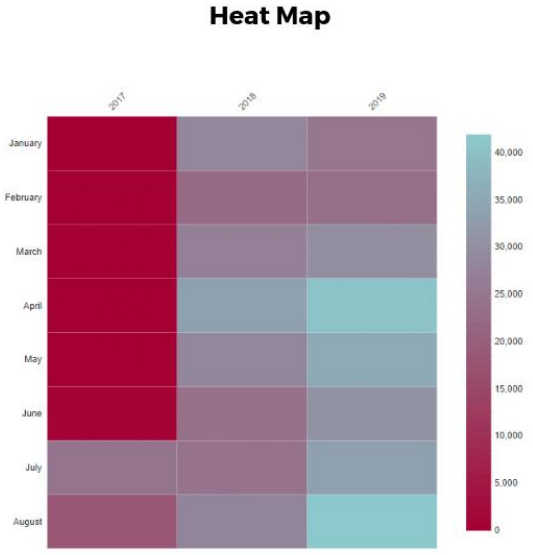
3. Histogram



Unlike bar charts, histograms illustrate the distribution of data over a continuous interval or defined period. These visualizations are helpful in identifying where values are concentrated, as well as where there are gaps or unusual values.

Histograms are especially useful for showing the frequency of a particular occurrence. For instance, if you'd like to show how many clicks your website received each day over the last week, you can use a histogram. From this visualization, you can quickly determine which days your website saw the greatest and fewest number of clicks.

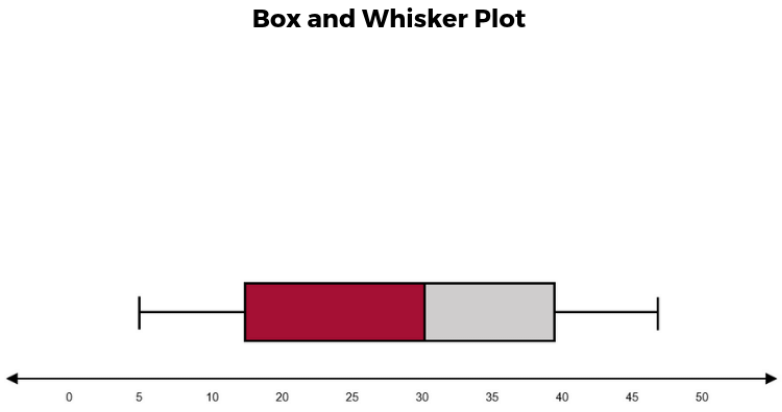
4.Heat Map



A heat map is a type of visualization used to show differences in data through variations in colour. These charts use colour to communicate values in a way that makes it easy for the viewer to quickly identify trends. Having a clear legend is necessary in order for a user to successfully read and interpret a heatmap.

There are many possible applications of heat maps. For example, if you want to analyse which time of day a retail store makes the most sales, you can use a heat map that shows the day of the week on the vertical axis and time of day on the horizontal axis. Then, by shading in the matrix with colours that correspond to the number of sales at each time of day, you can identify trends in the data that allow you to determine the exact times your store experiences the most sales.

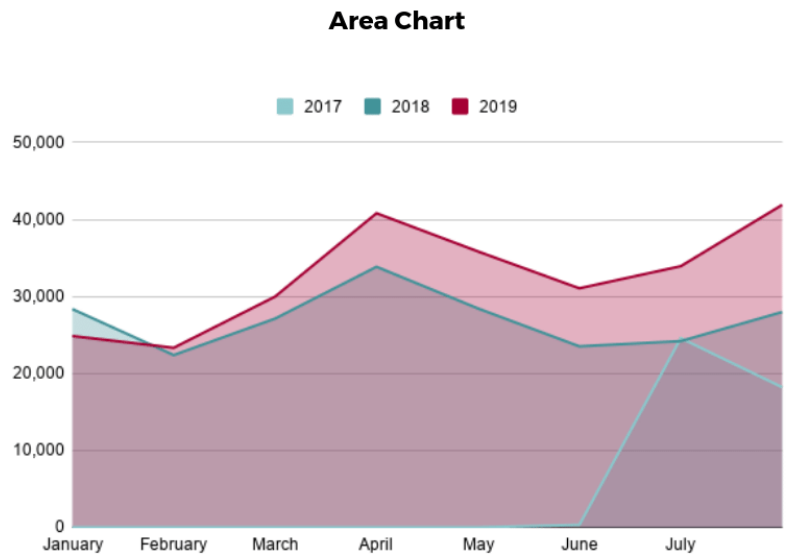
5.A Box and Whisker Plot



A box and whisker plot, or box plot, provides a visual summary of data through its quartiles. First, a box is drawn from the first quartile to the third of the data set. A line within the box represents the median. “Whiskers,” or lines, are then drawn extending from the box to the minimum (lower extreme) and maximum (upper extreme). Outliers are represented by individual points that are in-line with the whiskers.

This type of chart is helpful in quickly identifying whether or not the data is symmetrical or skewed, as well as providing a visual summary of the data set that can be easily interpreted.

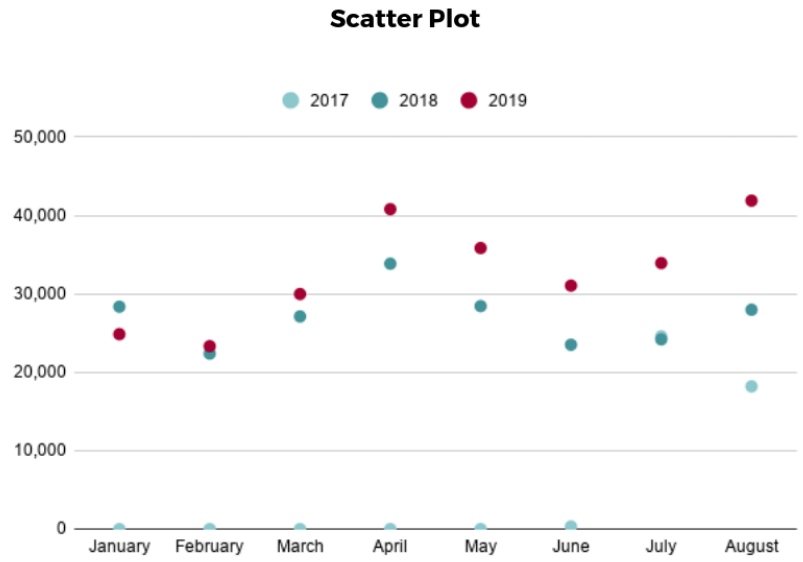
6. Area Chart



An area chart, or area graph, is a variation on a basic line graph in which the area underneath the line is shaded to represent the total value of each data point. When several data series must be compared on the same graph, stacked area charts are used.

This method of data visualization is useful for showing changes in one or more quantities over time, as well as showing how each quantity combines to make up the whole. Stacked area charts are effective in showing part-to-whole comparisons.

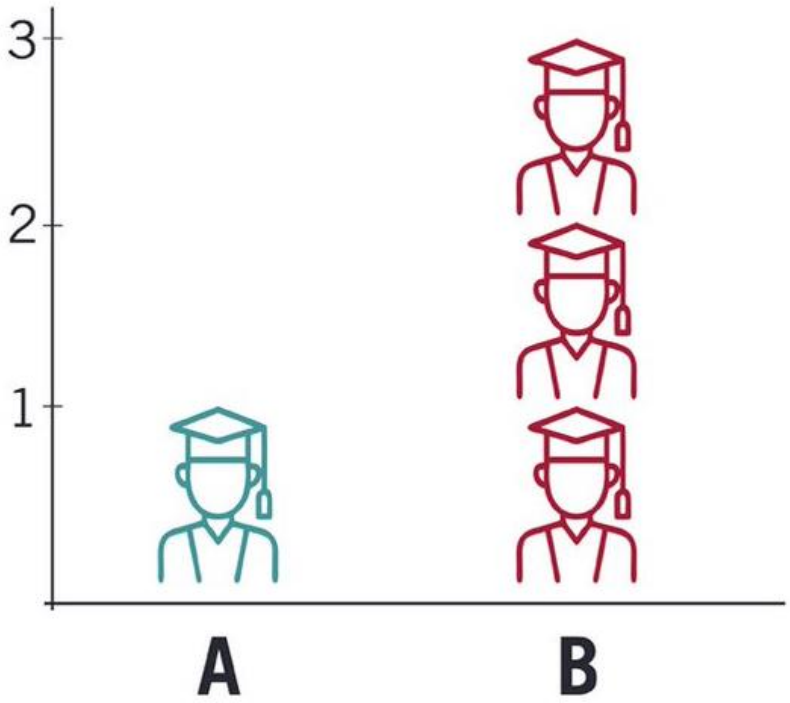
7. Scatter Plot



Another technique commonly used to display data is a scatter plot. A scatter plot displays data for two variables as represented by points plotted against the horizontal and vertical axis. This type of data visualization is useful in illustrating the relationships that exist between variables and can be used to identify trends or correlations in data.

Scatter plots are most effective for fairly large data sets, since it's often easier to identify trends when there are more data points present. Additionally, the closer the data points are grouped together, the stronger the correlation or trend tends to be.

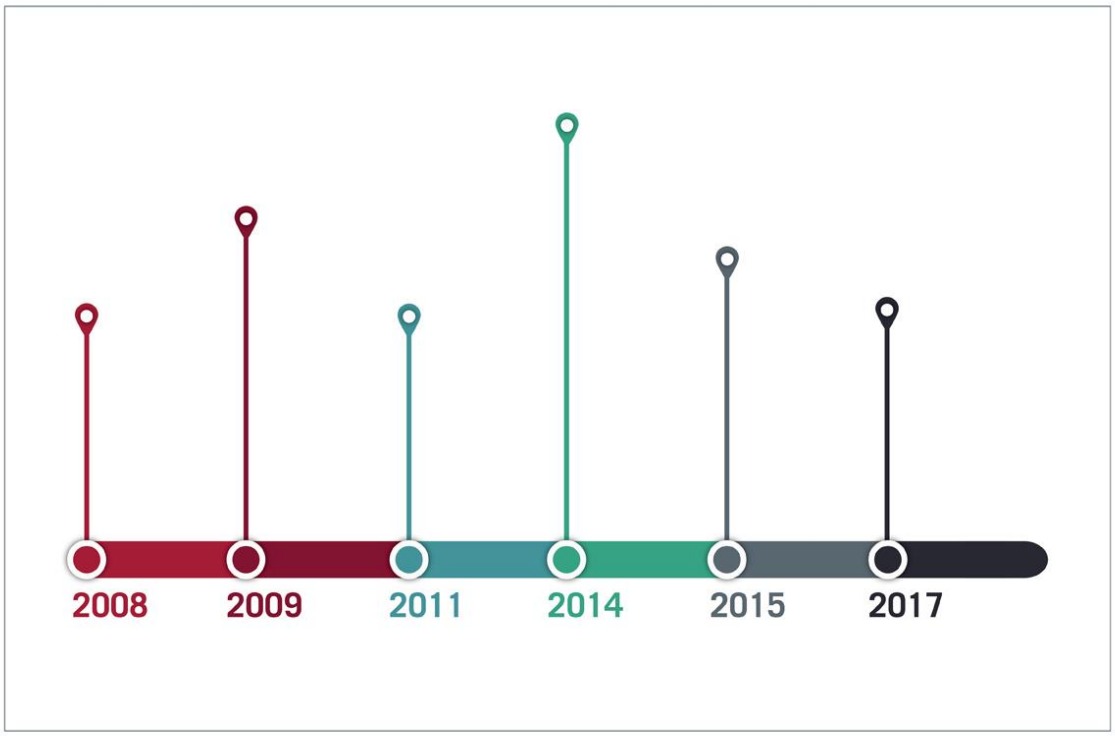
8. Pictogram Chart



Pictogram charts, or pictograph charts, are particularly useful for presenting simple data in a more visual and engaging way. These charts use icons to visualize data, with each icon representing a different value or category. For example, data about time might be represented by icons of clocks or watches. Each icon can correspond to either a single unit or a set number of units (for example, each icon represents 100 units).

In addition to making the data more engaging, pictogram charts are helpful in situations where language or cultural differences might be a barrier to the audience's understanding of the data.

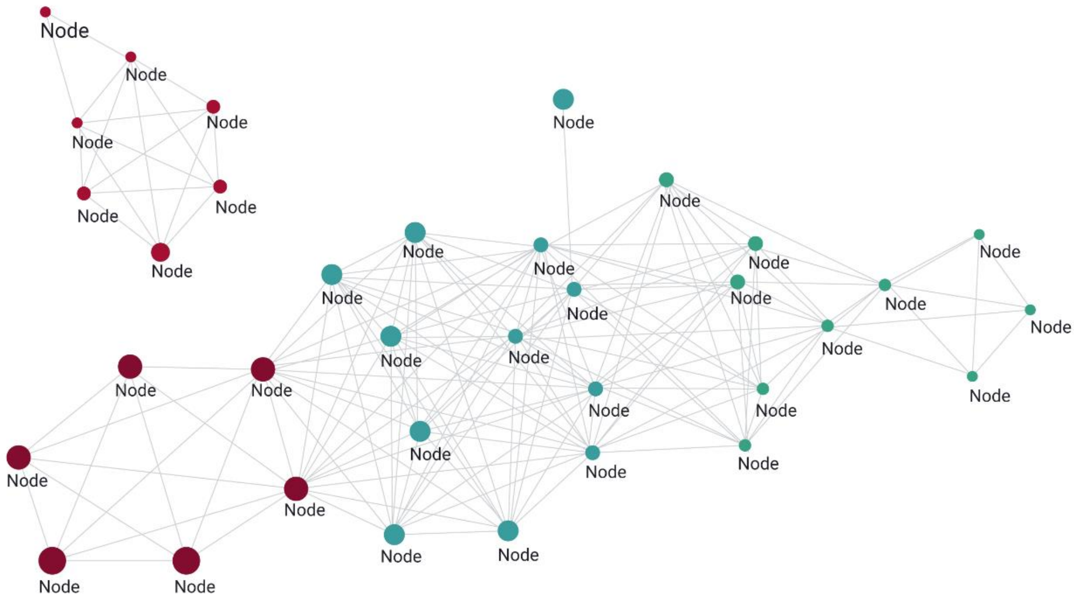
9. Timeline



Timelines are the most effective way to visualize a sequence of events in chronological order. They're typically linear, with key events outlined along the axis. Timelines are used to communicate time-related information and display historical data.

Timelines allow you to highlight the most important events that occurred, or need to occur in the future, and make it easy for the viewer to identify any patterns appearing within the selected time period. While timelines are often relatively simple linear visualizations, they can be made more visually appealing by adding images, colours, fonts, and decorative shapes.

10.Network Diagram



Network diagrams are a type of data visualization that represent relationships between qualitative data points. These visualizations are composed of nodes and links, also called edges. Nodes are singular data points that are connected to other nodes through edges, which show the relationship between multiple nodes.

There are many use cases for network diagrams, including depicting social networks, highlighting the relationships between employees at an organization, or visualizing product sales across geographic regions.

11. Correlation Matrix

A correlation matrix is a table that shows correlation coefficients between variables. Each cell represents the relationship between two variables, and a colour scale is used to communicate whether the variables are correlated and to what extent.

Correlation matrices are useful to summarize and find patterns in large data sets. In business, a correlation matrix might be used to analyse how different data points about a specific product might be related, such as price, advertising spend, launch date, etc.

The most important statistic to remember about data visualization is that 90% of information conveyed to the human brain is visual. At a fundamental level, we *see* the world. Being able to visualize data, therefore, is one of the most powerful and effective ways to express and communicate information.

Since the 1800s, data visualization has been used in scientific research, business endeavours, and mass media—such as newspapers and now the internet—to help explain the complexity of the world around us. Improved means of obtaining and storing data—first with hand-recorded observations, then with large computers, and now with smartphones—has meant an ongoing democratization of data as well. Today, it's never been easier to create, obtain, visualize, and distribute data.

So what does this mean in a practical sense? With seemingly *everything* having potential to be quantified and turned into data, it can be downright dizzying to determine how to engage with visualization. Paradoxically, figuring out how to visualize data, supposed to make data comprehension easier, can actually be a huge challenge in its own right!

From raw data to data visualization

Strictly speaking, the primary role of data visualization is to make the analysis of your data, if not the data itself, clear. Especially in social science research, data visualization makes it easy to see how data scientists collect and analyse data.

Think about a minute-by-minute description of a soccer match, for example. The sequence of events describing one play after another, who has the ball, who passes it to whom, who scores a goal, and how they scored it make up the raw data to be analysed. Pundits analyse the data by generating statistics such as goals, assists, key passes, possession ratios, and goalkeeper saves. These statistics are visualized in all sorts of ways in box scores and player profiles, which help fans understand the aspects of a particular game or player and make determinations

about the quality of the game they are watching. In the above example, the visualizations depend on the statistical analyses that pull the relevant information from the raw data. As a result, to rigorously visualize data, you need to have a well-intentioned strategy for analysing raw data. In other words, how do you look at the data? What are you looking for when you examine your data? The decisions you make to collect and analyse data will inform your data visualization.

Data visualization software

Of course, data scientists with a comprehensive understanding of their data can manually produce meaningful data visualization in platforms like Microsoft Excel and Google Drawings. However, it is the organization and coding of research that are essential to visualizing your data, which is where ATLAS. ti can help. As intuitive as the best data visualization tool might be, the preparation of collected data is essential to make the analysis and visualization processes easier and quicker, both for you and other users.

Organizing data

To facilitate the use of data visualization techniques, it is important to consider how the data in your research project should be organized. A market survey might produce hundreds or thousands of response records, so you might want to separate the responses of satisfied customers from the responses of unsatisfied customers to analyse the differences. Likewise, if you are conducting an interview study, you might benefit from separating documents by interview respondent characteristics.

However, what if you have conducted multiple interviews with respondents? Document groups in ATLAS. ti can help you

categorize documents that share a common characteristic. Document groups provide a convenient unit of analysis for many of the analysis tools in ATLAS. ti. You can think of a document group like a folder that holds a set of individual documents.

Unlike physical documents in file folders, documents in ATLAS. It can be assigned to multiple groups which can prove useful for later data visualization. In the example of the interview study, documents representing individual interviews can be categorized into multiple groups along lines of gender, age group, ethnicity, or customer satisfaction. In conjunction with coding and data visualization tools in ATLAS. ti document groups can help you create useful visualizations comparing broader sets of data. Instead of comparing individual interview respondents, for example, you can compare customer satisfaction across different age groups or ethnicities.

What *should* visualization research be about?

At base, visualization is a method for contextualizing data, enabling people to apply their prior experiences and perceptual and cognitive abilities to draw conclusions about phenomena in the real world. If you stop and think about it, this is a pretty ambitious target! Visualization research covers an impressive breadth of topics from perception to memorability, from complex system design to theory about what comprises a graph. But as in all research communities, biases what we think a research contribution *should* look like in our field can limit the types of questions we consider worth pursuing. For example, while research on perception has played a role in much of the visualization research field's history, we have not necessarily embraced cognition. Perhaps we are more comfortable with the

"bottom-up" nature of perception because its more clearly tied to the visual encodings. But what people take away from a visualization arises from "top-down" forces like what a person wants to believe, what they've learned about graphs in the past, or what they know about the domain as much as from patterns emerging as the visual system does its work.

A focus on performance keeps visualization research relevant to the world, where people want to know which chart to use. But, that shouldn't absolve researchers from trying to *explain why* a difference was found. This might take the form of proposing and testing for mechanisms in the visual system or cognitive strategies like heuristics. Considering the "why" in addition to the "what" can make it easier to reason about how a difference found between two specific visualizations might also be found under slightly different conditions.

Finally, we often draw a sharp line around representations of abstract data, considering other forms of diagrams or ways of contextualizing data to make it understandable outside our purview. But it is really productive to consider topics like interactive illustration , or satellite imagery, or sketching and analogical reasoning, or understanding aesthetics independent of analytical utility as not "core visualization" enough? Many people think of tables as "not visualization", but is this distinction useful? From a research perspective, thinking about visualization with less sharp boundaries might allow us to keep our field (and flagship conference!) relevant and exciting to the many people thinking about data presentation in the world. From a practical perspective, thinking of visualization as just one way of contextualizing data may help us realize when a task might be better served with a representation that doesn't involve

mapping data to visual encodings. When visualization is the only answer we see, we fail to acknowledge that sometimes information is better expressed in text or not at all.

Key Principles for Data Visualization and Reporting

- Do not clutter up the dashboard – it should be kept simple and contain easy-to-read information.

- Do not clutter up the dashboard – it should be kept simple and contain easy-to-read information.

- Make sure that you can interpret what each graph is saying quickly by eliminating anything from your graphs that confuses a reader or requires too much thought.
- All data points should be presented accurately using visualization techniques that do not distort them in any way.
- Remember who will view this material before you choose any type of chart – ask yourself what level of details is really needed here.

Techniques for Effective Reporting

- **Overview**: Explain shortly the central information and recommendations in the report.

- **Report Scope, Objectives, and Methodology:** Outline. Data, Analysis, Insights Transformation.
- Present data in a way that is logical and coherent; provide analysis that makes sense and creates insight.
- Structure your writing such that everything flows nicely together Key Points (Conclusion) And Actionable
- **Recommendations:** Highlight three points plus give examples from which people will leave with ideas of what they can do next time they encounter these scenarios.

Why data visualization is such a powerful tool:

- Intuitive: Presenting a graph as a node-link structure instantly makes sense, even to people who have never worked with graphs before.
- Fast: It is fast because our brains are great at identifying patterns, but only when data is presented in a tangible format.

Armed with visualization, we can spot trends and outliers very effectively.

- Flexible: The world is densely connected, so as long as there is an interesting relationship in your data somewhere, you will find value in graph visualization.
- Insightful: Exploring graph data interactively allows users to gain more in-depth knowledge, understand the context and ask more questions, compared to static visualization or raw data.

Importance of Data Visualization

According to the World Economic Forum, the world produces 2.5 quintillion bytes of data every day, and 90% of all data has been created in the last two years. With so much data, it's become increasingly difficult to manage and make sense of it all. It would be impossible for any single person to wade through data line-by-line and see distinct patterns and make observations. Data proliferation can be managed as part of the data science process, which includes data visualization("Data Visualization," n.d.). Data visualization is the presentation of data in a pictorial or graphical format. It enables decision-makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed("History of Data Visualization," n.d.).

Data visualization has become an indispensable part of the business world and an ever-increasing part of managing our daily life. Effective data visualization should be informative, efficient, appealing, and in some cases interactive and predictive. Patentor explains basic criteria that a data visualization should satisfy to be effective (Patentor 2018):

Criteria	Description
Informative	The visualization should be able to convey the desired information from the data to the reader.
Efficient	The visualization should not be ambiguous.
Appealing	The visualization should be captivating and visually pleasing.
Interactive and Predictive (Optional)	The visualizations can contain variables and filters with which the users may interact to predict results of different scenarios.

Design Principles

The role of data visualization in communicating the complex insights hidden inside data is vital. This is becoming more and more important since the audience for data visualizations is also expanding along with the size of data. Data visualizations are now consumed by people from all sorts of professional backgrounds. For the same reason, the ease of consumption is now a hot topic. While data scientists and analysts have an eye for digging out the key insights from even complex visualizations, a top business stakeholder or an average person might not be able to do the same.

And this is what makes effective data visualization the need of the hour. Communicating the data effectively is an art. However, many data scientists lag behind when it comes to the design and aesthetic aspects of visualizing data.

Here are some of the key design principles for creating beautiful and effective data visualizations for everyone.

Telling a Story with Insights.

Storytelling is an essential component of data visualization. The visualization must communicate complex ideas with precision and efficiency. The presenter must understand their audience's level of understanding and tailor their visualizations accordingly. An audience's level of analysis is key to creating and presenting a compelling story. Stikeleather's article outlined five key points to consider for telling a compelling story through a visualization(Jim Stikeleather 2013).

1. Find the Compelling Narrative
2. Understand your Audience
3. Be Objective and Offer Balance
4. Don't Censor
5. Edit

How to choose the best form of Visualization

Since just loading data into a table format could be a form of visualization, our focus should not be whether visualization is needed but on which form of data visualization is best for the situation.

Focus	Description
5 Second Rule	Research shows that the average modern attention span for viewing anything online is less than 5 seconds, so if you can't grab attention within 5 minutes, you've likely lost your viewer. Include clear titles and instructions, and tell people succinctly what the visualization shows and how to interact with it.
Design and layout matter	The design and layout should facilitate ease of understanding to convey your message to the viewer. Artists use design principles as the foundation of any visual work. If you want to take your data visualization from

Focus	Description
	an everyday dashboard to a compelling data story, incorporate graphic designer Melissa Anderson's principles of design: balance, emphasis, movement, pattern, repetition, proportion, rhythm, variety, and unity, discussed in more detail in the design principles section (Anderson 2017).
Keep it simple	Keep charts simple and easy to interpret. Instead of overloading viewers' brains with lots of information, keep only necessary elements in the chart and help the audience understand quickly what is going on.
Pretty doesn't mean effective	There is a misconception that aesthetically pleasing visualization is more effective. To draw attention, sometimes we want them to be pretty and eye-catching. But if it fails to communicate the data properly, you'll lose your audience's interest as quickly as you gained it.
Use colour purposely and effectively	Use of colour may be prettier and attractive but can be distracting too. Thus, the colour should be used only if it assists in conveying your message. Also, another thing to keep in mind is to be consistent with the colour scheme that the organization/consumer is used to and also try and follow the same colour across dashboards while communicating a story.

TRANSFORM DATA

Naturally, with the insights that you have gathered from the last visualization, you might have an idea of what you want to see next. You might have found some interesting pattern in the dataset which you now want to inspect in more detail. Possible transformations are the following.

Focusing the attention: What can be removed? Realize that consistency can help eliminate unnecessary distractions. There may be a trade-off between losing information but conveying the ultimate meaning more clearly. Label important things rather than relying on a legend, which requires the viewer to hold on to too much information at once.

Transformation	Description
Zooming	This allows us to have look at a certain detail in the visualization Aggregation To combine many data points into a single group
Filtering	This helps us to (temporarily) remove data points that are not in our major focus
Outlier handling	This allows us to get rid of single points that are not representative of 99% of the dataset.

IMPLICATIONS OF GOOD DATA VISUALIZATION

Raw data is often meaningless or at the very least is difficult to derive immediate meaning from. When people face a broad set of measurements and/or in large quantities, they are unable or unwilling to spend the time required to process it. Technological advances of the Digital Age contribute to an ever-growing pool of “big data” and have dramatically improved our ability to collect such large amounts of information. Thus, filtering, visualization, and interpretation of data becomes increasingly important.

We should understand how to best derive meaning from data, but first we should understand why its presentation in graphical format is so powerful. Furthermore, while the ideal purpose of data visualization is to facilitate understanding of data, visualization can also be used to mislead. Some of the main methods of doing so are omitting baselines, axis manipulation, omitting data, and ignoring graphing convention. Examples of these methods will be explored later in this chapter.

S No.	Principle	Description
1.	**Easy Recall**	People can process and remember images quicker than words. When data is transformed into images, the readability and cognition of the content greatly improves.
2.	**Providing Window for Perspective**	With infographics, you can pack a lot of information into a small space. Colours, shape, movement, the contrast in scale and weight, and even sound can be used to denote different aspects of the data allowing for multi-layered understanding (Mullis 2015).
3.	**Enable Qualitative Analysis**	Colour, shape, sounds, and size can make evident relationships within data very intuitive. When data points are represented as images or components of an entire scene, readers are able to see the correlation and analytical insights can be easily derived.

Data visualization tools are essential for transforming raw data into meaningful visual representations, making it easier to identify

patterns, trends, and insights. Here's an overview of some popular tools:

1. **Tableau**: Known for its user-friendly interface and powerful analytics, it supports drag-and-drop capabilities and offers interactive dashboards.
2. **Power BI**: A Microsoft tool that's great for business intelligence. It integrates seamlessly with other Microsoft products like Excel and Azure.
3. **Google Data Studio**: A free tool by Google that allows you to create reports and dashboards with real-time data from multiple sources.
4. **D3.js**: A JavaScript library that provides great flexibility for creating custom visualizations but requires some coding knowledge.
5. **QlikView and Qlik Sense**: These are robust tools for interactive data analysis, providing insights through smart search and data integration.
6. **R and Python Visualization Libraries (ggplot2, Matplotlib, Seaborn)**: If you're comfortable with coding, these libraries offer extensive customization for data visualizations.
7. **Charito (or similar tools)**: A cloud-based platform designed for business analytics with the capability to combine multiple data sources.

POWER BI

Power BI is a powerful business analytics tool developed by Microsoft. It helps users connect to various data sources, transform raw data into meaningful insights, and create visually appealing reports and dashboards.

Microsoft Power BI is a business intelligence (BI) platform that provides nontechnical business users with tools for aggregating, analysing, visualizing and sharing data. Power BI's user interface is fairly intuitive for users familiar with Excel, and its deep integration with other Microsoft products makes it a versatile self-service tool that requires little upfront training.

Users can download Power BI for Windows 10 or Windows 11, called Power BI Desktop, and native mobile apps for Windows, Android and Apple iOS devices. There is also Power BI Report Server for companies that must maintain their data and reports on premises. That version of Power BI requires a special version of the desktop app -- aptly called Power BI Desktop for Power BI Report Server.

Why is Microsoft Power BI popular?

A few factors make Power BI stand out as an option for businesses requiring BI capabilities. Power BI provides compelling and customizable visuals and requires little in the way of prior analytics or software experience. Many, if not all departments in an organization can use its capabilities, which include these popular ones:

- **Usability.** Power BI's user-friendly interface and navigation mean it has a simple learning curve for nontechnical business users, who are empowered to create data visuals with a few clicks and no programming experience.
- **Affordability.** Power BI is offered through different pricing plans so companies of all sizes can take advantage of its capabilities without worrying too much about budgetary constraints.

- **Visualizations.** Business leaders make use of data storytelling and visually enticing displays that help them communicate with clients, employees and other stakeholders.
- **Customization and interactivity.** Power BI lets users customize and interact with graphs, charts and other visualizations displaying their data. These features are useful in a number of ways to many different users.

Power BI use cases

Microsoft Power BI is used to find insights within an organization's data. It can help connect disparate data sets, transform and clean the data into a data model and create charts or graphs to provide visuals of the data. All of this can be shared with other Power BI users within the organization.

Data models created from Power BI can be used in several ways, including the following:

- Telling stories through charts and data visualizations.
- Examining what-if scenarios within the data.
- Creating reports that answer questions in real time and help with forecasting to make sure departments meet business metrics.

Power BI also provides executive dashboards for administrators and managers, giving management more insight into how departments are doing.

Who uses Power BI?

Power BI is a self-service BI tool that brings data analytics to employees. However, it's mostly data analysts and BI professionals who use it to create the data models before disseminating reports

throughout the organization. However, those without an analytical background can still navigate Power BI and create reports.

Both department staff members and management use Power BI to create reports and forecasts to aid sales and marketing reps. They also use it to provide data for management on how a department or individual employees are progressing toward their goals.

Advantages of Power BI

Multiple advantages make Power BI an attractive option for businesses looking to explore their data through BI visualizations, patterns and insights. These include the following:

- **Accessibility.** Both nontechnical and technical audiences can easily use Power BI. Its range of pricing options makes it available to businesses with smaller budgets as well as larger enterprises.
- **Integrations.** Power BI's ability to integrate with software such as Excel and other Microsoft applications means it can work with data from various sources.
- **Customer support.** Microsoft offers customer support for those in need of troubleshooting and assistance.
- **Customization.** The ability to create custom visuals adds value for businesses looking to capture and visualize important relationships or patterns found in complex data sets.
- **Collaboration.** Power BI is built to facilitate collaboration, for example, among different teams within an organization.
- **Scalability.** Power BI can support large-scale data sets, making it suitable for enterprises as well as small businesses.

Disadvantages of Power BI

While the benefits of the platform appear enticing, there are nonetheless caveats for business users to be aware of, as there are no perfect BI tools in existence yet. They include the following:

- **Pricing tiers.** While the free, Pro and Premium plans offer the essential capabilities for data reporting and analytics, the free version has a minimum number of features.
- **Resource-intensive.** Running the platform using an organization's existing infrastructure can consume compute and other resources faster than some organizations can tolerate.
- **Dense or cluttered user interfaces.** User-friendliness doesn't necessarily equate to neat dashboards as there are many visual components that appear cluttered.
- **Vast amounts of features.** Analytics and reporting can be done with easy-to-use basic features, but Power BI offers quite a few other features that are more complicated. For example, application programming interfaces (APIs) require complex knowledge and technical expertise.
- **Complex table relationships.** Power BI can take a few data tables and quickly find relationships among data points, but introducing more tables increases the risk of inaccuracies and slow processing times.

HOW TO USE POWER BI

Power BI Desktop is where analysts and other users can create data connections, data models and reports. The Power BI service is where those reports can be shared, so other users can view and interact with them.

To build a Power BI report, users must take the following steps:

1. Connect their data sources.
2. Query the data to create reports based on user needs.
3. Publish the report to the Power BI service.
4. Share the report, so cloud and mobile users can see and interact with it.
5. Add permissions to give or limit colleagues' ability to edit reports and create dashboards.

Power BI use case examples

There are a wide variety of uses for Power BI across different industries. The platform's ability to synthesize data from various sources makes it useful for analysing and reporting data that's culled using an array of tools and methods. Some of the ways different industry sectors use Power BI to their advantage include the following:

- **Healthcare.** Patient data is aggregated, synthesized and analysed to produce reports allowing healthcare providers to make new discoveries or detect new patterns.
- **Retail.** Reports based on customer purchase data give retailers insight into which products to prioritize and stock.
- **Manufacturing and engineering.** Power BI dashboards are used for process monitoring and reporting resource use. For instance, manufacturing data collected with IoT devices can be analysed using Power BI.
- **Education.** Tracking student performance with Power BI reports and dashboards helps school administrators and teachers identify where improvements are needed.
- **Finance and insurance.** Power BI is conducive to analysing the large-scale data sets the financial sector uses. The insurance

industry can use reports and dashboards to get insight into risk assessments.

- **Public sector.** Government agencies use Power BI to analyse intelligence or census data, as just two examples. Local government entities use it as well. Figure 3 shows a custom dashboard on bridge traffic that would be useful for planning transportation development and services.

TABLEAU

Tableau is a powerful tool used for data analysis and visualization. It allows the creation of amazing and interactive visualization and that too without coding. Tableau is very famous as it can take in data and produce the required data visualization output in a very short time. Basically, it can elevate your data into insights that can be used to drive your action in the future.

Tableau is a leading data visualization and business intelligence tool known for its simplicity and versatility. Here's what sets Tableau apart:

- **Ease of Use**: With its drag-and-drop interface, creating charts, graphs, and dashboards is incredibly user-friendly.
- **Rich Visualizations**: It offers a wide range of visuals, including bar charts, scatter plots, heat maps, and more, to present data effectively.
- **Data Connectivity**: Tableau can connect to multiple data sources, such as databases, spreadsheets, and cloud services, and blend data seamlessly.
- **Interactive Dashboards**: You can build dashboards that allow users to filter, drill down, and interact with data in real time.

- **Automation and Updates**: Tableau refreshes data automatically so your dashboards are always up to date.
- **Extensibility**: It integrates well with programming languages like Python and R for advanced analytics.

Tableau is a visual analytics platform that is revolutionizing the way we use data to solve problems by enabling individuals and organisations to make the most of their data. Tableau is a great data visualization and business intelligence application that can be used to report and analyse massive amounts of data. Salesforce purchased Tableau in June 2019, an American firm founded in 2003. It enables users to build various charts, graphs, maps, dashboards, and stories for visualising and analysing data in order to aid in business choices. Tableau offers several unique and fascinating features that make it one of the most popular business intelligences (BI) applications.

Why use Tableau?

Tableau is the fastest and most powerful visualization tool. It is very easy to use. There are no complex formulas like Excel and other visualization tools. It provides the features like cleaning, organizing, and visualizing data, it easier to create interactive visual analytics in the form of dashboards. These dashboards make it easier for non-technical analysts and end-users to convert data into understandable ones.

Tableau Features

- Tableau supports powerful data discovery and exploration that enables users to answer important questions in seconds
- No prior programming knowledge is needed; users without relevant experience can start immediately with creating visualizations using Tableau

- It can connect to several data sources that other BI tools do not support. Tableau enables users to create reports by joining and blending different datasets
- Tableau Server supports a centralized location to manage all published data sources within an organization

Values in Tableau

There are two types of values in the tableau:

- **Dimensions:** Values that are discrete (which cannot change with respect to time) in nature called Dimension in tableau. Example: city name, product name, country name.
- **Measures:** Values that are continuous (which can change with respect to time) in nature called Measure in tableau. Example: profit, sales, discount, population.

Advantages of Tableau

- **Quick calculation-** All the calculations on the tableau done by the backend, so it is relatively faster than any other tool.
- **Interactive dashboards**– Tableau dashboards are very interactive and easy to draw.

CHAPTER-04
FINDINGS, SUGGESTIONS& CONCLUSION

1.DATA QUALITY AND ACCURACY

One of the main limitations of data visualization tools is that they depend on the quality and accuracy of the data sources and the data preparation processes. Data architects need to ensure that the data is clean, consistent, reliable, and valid before feeding it to the visualization tools. Otherwise, the results may be misleading, incomplete, or erroneous. Data quality and accuracy issues can arise from various factors, such as data entry errors, missing values, duplicates, outliers, inconsistencies, biases, and format differences. Data architects need to apply rigorous data quality checks, data cleansing, data integration, and data validation techniques to ensure the data is ready for visualization.

2. DATA COMPLEXITY AND VOLUME

Another limitation of data visualization tools is that they may not be able to handle the complexity and volume of some data sets. Data architects need to deal with various types of data, such as structured, unstructured, semi-structured, streaming, spatial, temporal, and multidimensional data. Some of these data types may require special processing, transformation, or aggregation to make them suitable for visualization. Moreover, some data sets may be too large or too diverse to be visualized effectively. Data architects need to use appropriate data modelling, data compression, data sampling, data filtering, and data partitioning techniques to reduce the complexity and volume of the data and to select the most relevant and meaningful data for visualization.

3. DATA VISUALIZATION DESIGN AND SELECTION

A third limitation of data visualization tools is that they may not offer the best design and selection options for data visualization. Data architects need to choose the right visualization techniques, formats, colour, shapes, sizes, scales, labels, legends, and interactivity features to convey the data message clearly, accurately, and aesthetically. However, some data visualization tools may have limited or predefined options that may not suit the data type, the

audience, or the purpose of the visualization. Data architects need to use their creativity, judgment, and knowledge of data visualization principles and best practices to design and select the most effective and appropriate data visualizations for their data sets.

4. DATA VISUALIZATION INTERPRETATION AND COMMUNICATION

A final limitation of data visualization tools is that they may not guarantee the correct interpretation and communication of the data. Data architects need to ensure that the data visualizations are easy to understand, relevant, and actionable for the intended users and stakeholders. However, some data visualizations may be confusing, misleading, or ambiguous due to poor design, insufficient context, inappropriate comparison, or hidden assumptions. Data architects need to use their communication skills, data storytelling techniques, and data ethics principles to present and explain the data visualizations in a clear, concise, and trustworthy manner.

Data visualization tools are powerful and useful tools for data architecture, but they also have some limitations that data architects need to overcome. By paying attention to data quality and accuracy, data complexity and volume, data visualization design and selection, and data visualization interpretation and communication, data architects can leverage the full potential of data visualization tools and deliver impactful and meaningful data solutions.

ADVANTAGES AND DISADVANTAGES OF DATA VISUALIZATION

So what are the advantages and disadvantages of data visualization? Generally speaking, we at Tableau will say that advantages far outweigh the disadvantages, but it's always good to walk into creating a viz with your eyes open, and any possible issues in mind. So below, we've outlined the advantages and disadvantages of data visualizations in detail.

Advantages of data visualization

As we stated earlier, there are some very clear advantages to using data visualization, and as long as you're aware of and solving for the disadvantages, then the advantages are far larger than the disadvantages. Some of the advantages of data visualization include things like simple sharing, better quality analysis, and intuitive use. We dive into more of them in-depth below.

Intuitive

Many people find visuals much easier to understand than numbers or the written word. This means that most people find reading data visualizations much more intuitive than any other way of understanding data. That way even people who don't like math or say that numbers confuse them can easily scan and understand data, which makes it easy to keep everyone in an organization on the same page.

Simple data sharing

Similar to the point above, another benefit of data visualization is that it makes sharing data simple and easy because you can ensure that everyone is on the same page when they're viewing your visualization. Instead of risking people not being able to understand or interpret strings of numbers or raw data, a viz brings everyone onto the same page and is easy to share if you're using a BI platform.

Disadvantages of data visualization

With such great advantages, what disadvantages could possibly put a damper on it? Well, sometimes people can accidentally (or even purposefully) misrepresent data, or your message may get muddied. Or, if you're not careful in how you build your viz, you may end up with inexact conclusions or improper visualizations. We expand on those disadvantages below.

Improper visualization

The core of a lot of issues and disadvantages stems from this main one. If you're not careful in how you build your visualizations, you may end up with visualizations that don't properly convey your data. This can lead to confusion and issues down the line if you use that improper viz to do analysis and draw conclusions.

You can solve for this by ensuring that the people in your organization (and you) have the proper training to create good vizzes with your chosen BI software.

Incorrect conclusions

As talked about above, a risk of using data visualization is that your audience may draw incorrect conclusions. And that's not just because of improper visualizations. Sometimes a visual medium can lead to confusion in the viewer, so different people in your audience may walk away with drastically different conclusions after viewing the same viz.

You can solve for this by ensuring that you explain your vizzes, and provide a proper key for the viewer to understand.

Using data visualization in business

So after laying out the advantages and disadvantages, it's clear to everyone that data visualization is a powerful tool to use in business, as long as it's used properly. Ensure that your vizzes are built properly and carefully, and you'll find the tool of data visualization will help your business thrive. Use Tableau to help you visualize all your important business data in a meaningful and easy-to-understand way.

Findings:

The What and Why of Data Visualization

- Data visualization means drawing graphic displays to show data. Sometimes every data point is drawn, as in a scatterplot, sometimes statistical summaries may be shown, as in a histogram. The displays are mainly descriptive, concentrating on 'raw' data and simple summaries. They can include displays of transformed data, sometimes based on complicated transformations. One person's statistics may be another person's raw data. As with other aspects of working with graphics, it would be useful to have an agreed base of concepts and terminology to build on. The main goal is to visualize data and statistics, interpreting the displays to gain information.
- Data visualization is useful for data cleaning, exploring data structure, detecting outliers and unusual groups, identifying trends and clusters, spotting local patterns, evaluating modelling output, and presenting results. It is essential for exploratory data analysis and data mining to check data quality and to help analysts become familiar with the structure and features of the data before them. This is a part of data analysis that is underplayed in textbooks, yet ever-present in actual investigations. Look, for instance, at the one-sided peaks in the distributions of marathon finishing times (marastats, 2019).
- Graphics reveal data features that statistics and models may miss: unusual distributions of data, local patterns, clustering's, gaps, missing values, evidence of rounding or heaping, implicit boundaries, outliers, and so on. Graphics raise questions that stimulate research and suggest ideas. It sounds easy. In fact, interpreting graphics needs experience to identify potentially interesting features and statistical nous to guard against the dangers of overinterpretation. Just as graphics are useful for checking model results, models are useful for checking ideas derived from graphics (for more on models, see Hand, 2019).

This overview concentrates on static graphics. Dynamic graphics and, more especially, interactive graphics are in an exciting stage of development and have much to add. They require an article of their own. Superb examples include Human Terrain, a dynamic graphic showing the world's population in 3-D, and the interactive NameVoyager.

‘A Picture Is Worth a Thousand Words’

Famous sayings have a way of developing a life of their own. A picture is not a substitute for a thousand words; it needs a thousand words (or more). For data visualization you need to know the context, the source of the data, how and why they were collected, whether more could be collected, the reasons for drawing the displays, and how people with the necessary background knowledge advise they might be interpreted. There is a story that M. G. Kendall reviewed a book of R.A. Fisher's with the words: "No one should read this book who has not read it already." It is like that with graphics. If you have read all the supporting text, the display is often memorable and readily understandable. If you have not, it is not. Graphics on their own are insufficient, they are part of a whole. They complement text and are complemented by text. Student's reanalysis of the Lanarkshire Milk Experiment (Student, 1931) is an excellent example (and is also interesting as an early analysis of a large data set).

The potential synergy of text and graphics can be appreciated by talking through your own graphics, explaining them to others. Why have you drawn those graphics? How have you drawn them? What can be seen? Are there interesting patterns? What could be changed and improved? Which other graphics might be drawn? How can conclusions be checked? There should be more talking about graphics and less relying on the graphics to speak for themselves.

When it comes to graphics you have not drawn yourself, the same kinds of questions are still relevant, although they may be more difficult to answer. Edward Tufte described Charles Minard's display of Napoleon's Russian campaign as the best statistical graphic ever drawn (Tufte, 2001). It is a magnificent graphic, fully deserving of the praise heaped on it, yet as Lee Wilkinson has pointed out in his book *The Grammar of Graphics* (Wilkinson, 2005), there are inaccuracies and imprecisions in the display. Why did no one point them out before? We are too used to accepting graphics uncritically, not asking enough questions of them.

Research in Data Visualization

There are great opportunities for future research in data visualization. Principles are needed on how to decide which of many possible graphics to

draw. It is not a matter of drawing a single, 'optimal' graphic, if such a thing even existed; it is a matter of choosing a group of graphics that will provide more information. It is like taking photographs of a complicated object, a single one would not be enough, and taking pictures from every possible angle and distance would be far too many. Sets of graphics are useful for providing context, as the scatterplots in Klimek, Yegorov, Hanel, and Thurner (2012) demonstrate.

More understanding of combining and linking graphics is needed, whether in static ensembles or in interactive displays, just as better software is needed for these. The value of alignment and common scaling for making effective comparisons, for instance, with small multiples and faceting (displaying many graphics of the same form conditioning on other variables) is one part of this. It is a historical curiosity that the current exciting work on interactive graphics on the Web still lags behind standalone systems that were already available more than 30 years ago in linking multiple windows. Data Desk and JMP were commercial examples at the time (see Velleman, 2019, and Sall, 2019, for current versions).

Published graphics are sometimes attractive and beautifully produced. The content does not always match. That may be because authors and publishers do not expect the graphics to be examined in any detail. They may be added as illustrations to balance the layout and make it look more agreeable. If you do not have a suitable photograph, cartoon, or map, you could use a colourful statistical graphic. I have many times heard people say that they do not understand numbers and were bad at mathematics in school. No one has ever said to me they do not understand graphics, perhaps because they regard them as illustrations and not as central parts of an argument. There is work to be done in educating researchers and readers in the value of graphics.

Research into new and innovative graphics is exciting and productive. Simultaneously, it is essential to make the best use of known and well-understood graphics. There is a risk of emphasis on novelty at the expense of familiarity. New, innovative graphics need instruction and experience to interpret them. Their designers have spent much time developing them and reasonably enough believe that what is obvious to them should be obvious to everyone. Just think of the humble scatterplot. It is only in recent years that

scatterplots have appeared in the media, although they are one of the most important statistical graphics. If you have never seen one before, they can be intimidating, even more so when you are told 'It is clear that...' or 'You can easily see that...' We should build on the familiar to carry our readers along with us.

What Happens Now?

Educating people in choosing, drawing, and interpreting graphics is more difficult than you might think. Data visualization is not taught badly, it is just not taught very much at all. Ideally, there should be better theory, and consequently better graphics. That will take time. In the meantime, we should:

—discuss more graphics more;

—interpret more graphics more;

—teach more graphics more.

CONCLUSION:

In summary, there are a variety of obstacles associated with data visualization, from choosing the best visual representation to successfully communicating complex data. To limit these challenges, it takes a deep comprehension of the target audience and the intended message to solve them successfully. Hope that the sharing about the challenges of data visualization in this post is useful for you. Don't forget to visit our site frequently for more information about big data and technology for growth and innovation.

In today's data-rich environment, data visualization reigns supreme in unlocking insights. By transforming complex datasets into clear visuals, it unveils hidden patterns and trends. As data explodes, these tools become crucial across diverse fields. Whether you are a data pro or a beginner, numerous resources empower you to master this valuable skill. Remember, successful data visualization transcends aesthetics; it is about effectively narrating the data's story.

Are you considering a career in data visualization? Certifications like Big Data Hadoop and Spark Developer can equip you with valuable industry-recognized skills.

Explore more options at Sprintzeal's all-course page, and connect with their course experts to find the training best suited for your career goals. Stay updated by subscribing to the newsletters for insights and trends.

data visualization is a powerful tool that plays a crucial role in various domains, including healthcare, sports, the humanities, environmental sciences, and more. Using data visualization tools and techniques makes it possible to analyse and understand complex data, identify patterns and trends, and communicate insights to a wider audience. Effective data visualization requires choosing the right tool and technique for the data and message being conveyed and presenting the data in a clear and visually appealing manner. Data visualisation helps support informed decision-making, improve outcomes, and advance knowledge and understanding in various fields. The importance of data visualization cannot be overstated, and it is likely to continue to play an increasingly important role in various domains in the future. In addition to the benefits outlined above, data visualization also has the potential to democratize information and make it accessible to a wider audience. By presenting data in a clear and visually appealing manner, data visualization can help to communicate complex information and insights to the public, promoting transparency and understanding. This is especially important in fields such as healthcare, where understanding and communicating data is essential for promoting public health and well-being. Data visualization is crucial for analyzing and communicating data in various domains. Whether it's used to understand patient data in healthcare, player performance in sports, cultural trends in the humanities, environmental data in environmental sciences, or any other type of data, data visualization has the potential to support informed decision-making, improve outcomes, and advance knowledge and understanding. As technology continues to evolve, the possibilities for data visualization will continue to expand, and it will remain an important tool for data analysis and communication for years to come.

This work describes the results obtained from the analysis of two focus groups about data visualizations. Both groups had different backgrounds to, additionally, analyze differences between the interventions of the participants; one group was composed of data visualization experts, while the other was composed of PhD students.

The analysis of the interventions provided interesting insights. On the one hand, experts were more conscious of the technical aspects of data visualizations and more critical of the potential enhancements applied to the different figures employed in the experiment, which consisted of including more context through legends and complementary information.

On the other hand, doctoral students focused more on the lack of context and the necessity of including more information through textual elements. In this sense, doctoral students preferred the “enhanced version” of each data visualization, as they included more context, although the visual elements included were not the most effective or useful option.

BIBILOGRAPHY

Bibliography

- https://www.geeksforgeeks.org/data-visualization-and-its-importance/
- https://guides.library.jhu.edu/datavisualization
- https://digital.gov/resources/an-introduction-to-data-visualization/
- https://www.researchgate.net/publication/370593444_An_Introduction_to_Data_Visualization_Tools_and_Techniques_in_Various_Domains
- https://documents1.worldbank.org/curated/en/099010424125523914/pdf/P1758801e975c106e1a92e197e0627b10a3.pdf
- https://www.techtarget.com/searchbusinessanalytics/definition/data-visualization
- https://library.carleton.ca/sites/default/files/help/Intro%20to%20Data%20Viz%202016.pdf
- https://www.cs.put.poznan.pl/jstefanowski/sed/DM14-visualisation.pdf
- https://www.sciencedirect.com/science/article/pii/S0747563224000293#sec5
- https://www.sciencedirect.com/science/article/abs/pii/S0263786320300260
- https://www.sciencedirect.com/science/article/abs/pii/S0263786320300260
- https://www.cs.put.poznan.pl/jstefanowski/sed/DM14-visualisation.pdf
- https://library.carleton.ca/sites/default/files/help/Intro%20to%20Data%20Viz%202016.pdf

DATA VISUALIZATION

PLAGIARISM REPORT:

RATHNA KUMARI NARAYANAPARAPU SEM-IV PROJECT.

ORIGINALITY REPORT

14%	6%	3%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

1	*Google schloar.net* Internet Source	3%
2	*Submitted to Roehampton University* Student Paper	1%
3	*fastercapital.com* Internet Source	1%
4	*Submitted to University of Leicester* Student Paper	1%
5	*Submitted to UNITEC Institute of Technology* Student Paper	<1%
6	*en.wikipedia.org* Internet Source	<1%
7	*HDFC Bank.com* Internet Source	<1%
8	*Submitted to Purdue University* Student Paper	<1%
9	*Submitted to Toronto Business College* Student Paper	<1%

10	Submitted to pes university Student Paper	<1%
11	www.coursehero.com Internet Source	<1%
12	Submitted to ALC Engineering College Student Paper	<1%
13	*Submitted to University of Stirling* *Student Paper*	<1%
14	*Submitted to University of Arizona Global Campus (UAGC)* Student Paper	<1%
15	*Submitted to Mont Rose College* Student Paper	<1%
16	*Submitted to Republic of the Maldives* Student Paper	<1%
17	*Submitted to Park University* Student Paper	<1%
18	*docplayer.net* Internet Source	<1%
19	*eprajournals.com* Internet Source	<1%
20	*Submitted to University of Hertfordshire* Student Paper	<1%
21	*Submitted to American* university Student Paper	<1%

21	*Submitted to The University of the West of Scotland* Student Paper	<1%
22	*Submitted to Glyndwr University* Student Paper	<1%
23	*Submitted to University of Technology, Sydney* Student Paper	<1%
24	*Submitted to West Herts College* Student Paper	<1%
25	*RBI.GOV.IN* Internet Source	<1%
26	*Case study.com* Internet Source	<1%
27	*Submitted to Harrisburg University of Science and Technology* Student Paper	<1%
28	*Submitted to University of Wolverhampton* Student Paper	<1%
29	*Submitted to London School of Science & Technology* Student Paper	<1%
30	*Submitted to University of Hull* Student Paper	<1%

31	*bizmarksolutions.com* Internet Source	<1%
32	*Submitted to Bath Spa University College* Student Paper	<1%
33	*Submitted to London School of Commerce* Student Paper	<1%
34	*Submitted to Rider University* Student Paper	<1%
35	*www.icici bank.com.* Internet Source	<1%
36	*Submitted to CSU, San Jose State University* Student Paper	<1%
37	*Submitted to Isabela State University - Cabagan Campus* Student Paper	<1%
38	*Submitted to Xavier Institute of Management and Entrepreneurship, Bangalore* Student Paper	<1%
39	*S&P Capital IQ.com* Internet Source	<1%
40	*www.ecomcrew.com* Internet Source	<1%

41	*www.ift.org* Internet Source	<1%
42	*www.scirp.org* Internet Source	<1%
43	*Submitted to Curtin University of Technology* Student Paper	<1%
44	*www.quizsol.com* Internet Source	<1%
45	*Submitted to Brisbane Catholic Education* Student Paper	<1%
46	*en.everybodywiki.com* Internet Source	<1%
47	*www.bocsar.nsw.gov.au* Internet Source	<1%
48	*"Phantom Ex Machina", Springer Science and Business Media LLC, 2017* Publication	<1%
49	*Saba N. S., Ritesh Gandhi, Sindhu R. Rajendran, Natasha D. Abraham. "Chapter 1 Revolutionizing Digital Marketing Using Machine Learning", IGI Global, 2023* Publication	<1%
50	*Submitted to University of West Georgia* Student Paper	<1%

51	*Zovio. "Business and Film: A Reader", Zovio* Publication	<1%
52	*Christian Fuchs, Vincent Mosco. "Marx in the age of digital capitalism", brill, 2016* Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		