

## A Review on Comparative Study of Popular Data Visualization Tools

Dr. Anilkumar J. Kadam<sup>1</sup>, Kishorkumar Akhade<sup>2</sup>

<sup>1</sup>Professor, Department of Computer Engineering (AI&DS), AISSMS COE, Pune, India

<sup>2</sup>M.E. Student, Department of Computer Engineering (AI&DS), AISSMS COE, Pune, India

### Abstract

*Data visualization tools have become essential for businesses to efficiently process, analyse, and communicate vast amounts of data. They help identify trends and aid in decision-making by presenting large datasets in graphical formats. This paper reviews the most widely used and distributed visualization tools for massive data sets, highlighting their features, strengths, and weaknesses. The paper introduces the concept of data visualization, its evolution from static charts to interactive and dynamic visualizations, and discusses available tools like scatter plots, heat maps, tree maps. It also discusses the features and functionalities of popular data visualization tools like Tableau, Power BI and Python libraries like Matplotlib and Seaborn. The paper emphasizes the importance of choosing the right visualization tool based on data type, audience, and purpose, and discusses best practices for creating effective data visualizations. Future trends include augmented and virtual reality, machine learning, and artificial intelligence.*

**Keywords:** Data Visualization, Comparative Analysis, Tableau, Power BI, Matplotlib, Seaborn, Business Intelligence, Interactive Dashboards.

## 1. INTRODUCTION

Data visualization is a crucial skill in today's business world, as the amount of available data becomes increasingly blurred. It helps present complex information in a more understandable and interpretable format, enabling analysts to quickly identify trends and patterns. Tools like Tableau and Power BI are essential for this purpose. These tools help present complex data visually appealing, allowing for better understanding and identification of potential anomalies or outliers.

They also enable communication of complex information to a broader audience, making data-driven decision-making more important. They aid in exploratory data analysis, allowing analysts to quickly identify areas of interest and perform further analysis. Furthermore, data visualization tools can unify data from different sources and apply AI-

driven analytics to reveal relationships between different data sets, providing businesses with new insights that would be difficult to obtain by looking at individual data sets separately.

## 2. LITERATURE SURVEY

**Aryamaan Pandey, Ishita Sharma, Akshat Sachan and Dr. P Madhavan (2022)**

This study, examines various Data Visualization applications in the Big Data field, including Power BI, Tableau, Zoho Analytic, Celonis, MS Excel, and Dataiku DSS. It highlights the challenges of choosing the right tool for different needs, such as licensing costs for small and medium businesses, hybrid environments with limited computers, and the need for a user-friendly interface. Power BI's drag-and-drop functionality and access to over 85 data visualization programs provide an aesthetically pleasing and useful report.

**Fehmi Skender (2022)**

In this study, the analysis of tools for visual data analysis and visualization of big data suggests they are suitable for organizations to solve dynamic problems and advance production and services. Visual data processing and visualization are crucial for users, consumers, manufacturers, and managers. The number of visualization tools is rapidly growing, and they are constantly improving their capabilities, especially in the field of big data. The study focuses on big data visualization tools and their key performance indicators. The analysis supports the necessity of Big Data, Visualization, Machine Learning, and Artificial Intelligence for success in all fields.

**Addepalli Lavanya, Sakinam Sindhuja, Lokhande Gaurav, Waqas Ali (2023)**

This study of data visualization tools highlights the importance of visual representation in data analysis

and communication. It covers various visualization types, such as bar charts, line graphs, and heat maps, and evaluates popular tools like Tableau, Power BI, Google Data Studio, D3.js, and Python libraries. The review also discusses best practices for creating effective visualizations, such as choosing the right colour scheme and designing for accessibility. Future trends include augmented and virtual reality, machine learning, and artificial intelligence, which could further enhance the capabilities of data visualization tools.

**Abhishek Chavan (2023)**

Data visualization tools have made significant progress, but there's still room for improvement. Integrating AI and machine learning can automate the process of creating visualizations, while Natural Language Processing (NLP) can help users create interactive visualizations. Virtual and Augmented Reality can create immersive, interactive visualizations. Improving cross-platform integration between tools and other business intelligence tools can enhance data strategy. Enhancing collaboration and accessibility can help teams work more effectively and drive business success. The future of data visualization is promising, with opportunities for improvement.

**3. METHODOLOGY**

The evaluation of a data visualization tool involves assessing its ease of use, flexibility, customization, interactivity, scalability, data connectivity, community support, and representative use cases. It involves evaluating the user interface, learning experience, and adaptability to specific needs. The tool should also assess its ability to connect to various data sources and formats. It is crucial to consider the availability of online communities and resources for each tool.

**3.1 Data visualization tools**

**3.1.1 Tableau**

Tableau is a widely used data visualization tool in the data analytics industry due to its ability to effectively manage massive datasets and generate eye-catching representations that reveal insights that cannot be gleaned simply by gazing at a spreadsheet. It is highly rated for its overall performance, making it easy to create different types of visualizations and handle millions of rows of data. Tableau's quick drag-

and-drop functions help users create interactive reports and dashboards within minutes, making it easy to implement.

However, one strong disadvantage of Tableau is its high cost, which is considered high for small to medium-sized businesses. The license cost is higher than other data visualization tools, such as open-source software like R or Python, which can be used at no cost. In conclusion, Tableau is a popular choice for data visualization, alongside other tools like Python libraries, and has been extensively studied for its advantages, uses, and importance in data analytics.

**Pros**

1. The use of scripting languages like R and Python can expedite the execution of complex table calculations.
2. Can efficiently handle large amounts of data.

**Cons**

1. The process of importing custom visualizations has been found to be challenging.
2. Reports are not easily integrated into other applications.



Figure 1: Interactive visualization by Tableau

**3.1.2 Power BI**

Power BI is a suite of corporate analytics tools that enables self-service data analysis and knowledge dissemination. It transforms disparate data sets into unified, interactive

visualizations and insights. The platform is scalable for enterprise-level use and user-friendly for self-service BI, allowing users to connect to any data source and seamlessly visualize it into their everyday apps. Power BI Desktop allows users to manage relationships between data tables and create histograms, while Power BI Service, also known as Power BI Workspace, Power BI Web Portal, or Power BI Site, is an on-cloud service that shares and publishes reports created on Power BI Desktop.

Power BI can be customized to fit specific needs, such as building Power Apps to enable users to enter notes for specific data points and use the Power Apps custom visual on Power BI reports. It is a powerful and versatile platform for data analytics and visualization, allowing users to connect to any data source and share insights with others. Power BI combines powerful analytics with a user-friendly interface and excellent data visualization capabilities, allowing users to create and manipulate data visualizations using the user-friendly interface.

Power BI can process large amounts of data efficiently, allowing users to drill down into granular details, providing better insights and strategic decision-making abilities. It manipulates the system's transactional, interaction, and behavioral data and external sources, such as API. Overall, Power BI is a powerful data visualization and analytics tool with a user-friendly interface and vast visualization capabilities, enabling users to process large amounts of data efficiently and make strategic decisions based on the insights gained from their data.

#### Pros

1. Reports and dashboards pre-built for SaaS solutions.
2. The data can be automatically refreshed after publishing to the Power BI web service without any manual intervention.

#### Cons

1. Users with the same email domain can only access the dashboard and report sharing feature.

2. A Power BI report's file size is limited to 1 GB.



Figure 2: Interactive visualization by Power BI

### 3.1.3 Matplotlib

Matplotlib is a Python-based data visualization library that generates 2D plots and graphs for scientific, engineering, and financial data. Its features include versatility in generating different types of plots, customization options for plot aesthetics, and compatibility with other libraries and frameworks. Matplotlib's advantage lies in its ability to provide interactive and dynamic visualizations, making it suitable for exploratory data analysis. However, its default plots may not be aesthetically pleasing, requiring further customization. Some sources suggest that Matplotlib may not be the most efficient library for generating complex and large-scale visualizations. Despite its popularity, Matplotlib may not be the most efficient for generating complex and large-scale visualizations. Its use cases include geographic data, time series data, and statistical analyses. Overall, Matplotlib is a powerful and versatile library for data visualization in Python.

#### Pros

The platform offers a diverse array of customizable visualization packages, cross-platform support, interactive plotting, and a vast user community for support.

#### Cons

The learning curve is challenging due to its verbose syntax, limitations in 3D and animation plotting, and limited interactivity between other visualization Python packages.

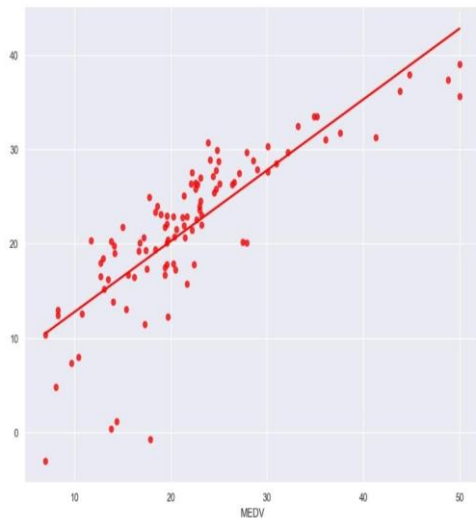


Figure 3: Scatter plot using matplotlib

### 3.1.4 Seaborn

Seaborn is a Python library that creates visually informative charts and graphs using Matplotlib and pandas' package for data analysis. Its declarative, dataset-focused API makes it easy to turn data inquiries into charts. Seaborn has been extensively studied and well-documented in research papers and online resources, providing a comprehensive literature review of its capabilities, syntax, and plot types. Research papers like Michael Waskom's "Seaborn: Statistical Data Visualization" and The Journal of Open-Source Software's "seaborn: statistical data visualization. Python" provide detailed descriptions of the library's architecture, plot types, integration with Pandas, and customization options. Online articles and tutorials also provide an introduction to Seaborn and its capabilities. The official website offers a getting started guide, tutorials, and API reference, as well as examples of using the library to visualize different types of data.

#### Pros

The tool offers a wide range of default multiple plotting options, strong integration with Pandas and NumPy, and inbuilt statistical tests.

#### Cons

R is frequently accessed out of memory and not fully embedded into Python.

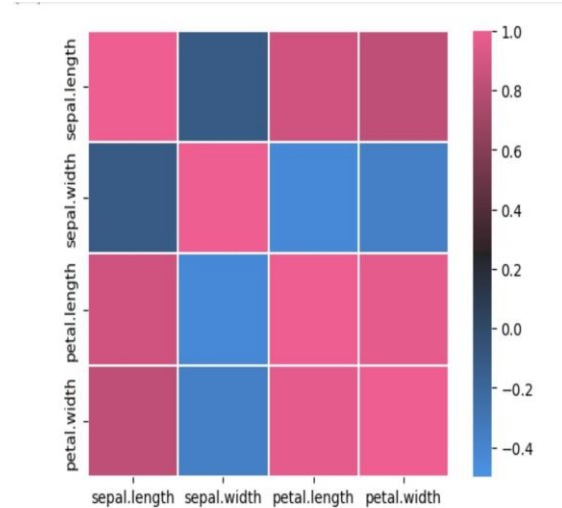


Figure 4: Heatmap using seaborn

Table 1. This summary provides a comprehensive overview of the advantages and disadvantages of data visualization tools.

Tools	Advantages	Disadvantages
1. Tableau	<p>Ease of use: Tableau is renowned for its user-friendly interface, enabling users with minimal technical knowledge to swiftly generate visualizations and dashboards.</p> <p>Data Connectivity: Tableau offers a diverse range of native connectors and supports various data sources, making it easy to integrate with various databases and file formats.</p> <p>Performance: Tableau is renowned for</p>	<p>Cost: Tableau can be costly, particularly for larger organizations or those scaling up to enterprise-level deployments.</p> <p>Limited Advanced Analytics: Tableau excels in visualization but may lack the robustness of advanced analytics and predictive modelling compared to other tools.</p>

	its exceptional performance and speed in handling large datasets, enabling efficient processing and visualization of vast amounts of data.	
<b>2. Power BI</b>	<p>Integration with Microsoft: Power BI seamlessly integrates with Microsoft's Excel, Azure, and SQL Server, offering a comprehensive experience for users already using these tools.</p> <p>Cost effective: Power BI is a cost-effective option for Microsoft ecosystem-invested organizations, offering various pricing plans, including a free version with limited capabilities.</p> <p>Natural Language Query: Power BI offers a user-friendly natural language query feature, enabling users to ask questions in plain language and receive</p>	<p>Limited Visualization Customization: Power BI offers a diverse array of visualization options, but some users find it less customizable than Tableau.</p> <p>Data Source Compatibility: Power BI, though capable of supporting various data sources, may not have a comprehensive list of data sources like Tableau for certain niche databases.</p>

	visualizations as answers.	
<b>3. Matplotlib</b>	<p>Versatility: Matplotlib is a versatile library that enables users to create diverse plots and charts, offering fine-grained control over their appearance.</p> <p>Integration: Matplotlib is a popular choice among developers due to its seamless integration with other libraries and frameworks, making it suitable for diverse environments.</p> <p>Wide Range of plots: Matplotlib offers a variety of plot types, including scatter plots, line plots, bar plots, histograms, and more.</p>	<p>Steep Learning Curve: Matplotlib's learning curve is steeper for beginners, and creating complex plots may require more code than other libraries.</p> <p>Default Aesthetics: The default visualizations in libraries like Seaborn may not be as visually appealing, necessitating users to invest time in customizing plots for a polished appearance.</p>
<b>3. Seaborn</b>	<p>Statistical Plotting: Seaborn, a statistical graphics tool, is a high-level interface built on Matplotlib, simplifying the creation of complex visualizations with concise code.</p>	<p>Limited plot Types: Seaborn excels in statistical plotting but may not offer the same level of flexibility and versatility as Matplotlib for creating highly customized plots.</p>

	<p><b>Default Aesthetics:</b> Seaborn offers a user-friendly interface with default themes and colour palettes, allowing for easy creation of visually appealing plots without extensive customization.</p> <p><b>Concise Syntax:</b> Seaborn simplifies visualization tasks and reduces boilerplate code, making it a more efficient tool for producing complex plots compared to Matplotlib.</p>	<p><b>Less Customization:</b> Seaborn is suitable for common use cases but may not be suitable for users who require extensive customization and control over their plots.</p>
--	--	--

#### 4. Conclusion

Data visualization is crucial for organizations to identify patterns, learn from historical data, and find success. Traditional visualization techniques struggle to keep up with the volume and velocity of data. Big data visualization software must address performance, complexity, and lag time. Power BI and Tableau are popular options, but it's important to consider their requirements and functionality before choosing. Tableau offers data mining and visualization solutions for business intelligence with horizontal interaction scalability, making it ideal for hybrid environments with limited desktop access. For this project, Tableau software will be chosen. Organizations should carefully consider their requirements and choose the right visualization tool for their specific situation.

The future of data visualization is promising, with opportunities to enhance its power, accessibility, and usability. Utilizing AI, NLP, and virtual and augmented reality technologies, these tools can help

organizations effectively utilize data for business success.

The study evaluates popular data visualization tools like Tableau, Power BI, and Python libraries Matplotlib and Seaborn, highlighting their strengths and weaknesses, and providing real-world examples of their use in data presentation and analysis.

#### References

- [1] Addepalli Lavanya, Sakinam Sindhuja, Lokhande Gaurav and Waqas Ali "A Comprehensive Review of Data Visualization Tools: Features, Strengths, and Weaknesses" International Journal of Computer Engineering in Research Trends Multidisciplinary, Review Paper Volume-10, Issue-1, 2023 Regular Edition
- [2] Fehmi Skender "Data Visualization Tools - Preview and Comparison" Journal of Emerging Computer Technologies
- [3] Abhishek Chavan "Data Visualization Tools Used for Decision Making" International Research Journal of Modernization in Engineering Technology and Science (Peer-Reviewed, Open Access, Fully Refereed International Journal) Volume:05/Issue:02/February-2023
- [4] Michael Diamond and Angela Mattia, Jacksonville University, "Data Visualization: an exploratory study into the software tools used by business", Journal of Instructional Pedagogies, vol. 17 pp. 1-3, November 2015.
- [5] Tvrdikova, M. (2007), 'Support of Decision Making by Business Intelligence Tools', Computer Information Systems and Industrial Management Applications, 2007. CISIM '07. 6th International Conference, pp. 368
- [6] K. Kalluri, 2022. [Online]. Available: <https://splashbi.com/importance-purposebenefit-of-data-visualization-tools/>. [Accessed 1 Dec 2023].
- [7] U. S. Junghare, Shri Shivaji Science college, "Analytical Evaluation towards Big Data Visualization Applications, Tools and It's Issues", The International journal of analytical and experimental modal analysis, September 2021.

[8] Leena Gokhale and Kirti Nilesh Mahajan, Bharati Vidyapeeth Deemed University, "Comparative Study of Data Visualization Tools" Research gate., pp. 48-49, May 2020.

[9] Soly Mathew Biju and Alex Mathew, "Comparative Analysis of Selected Big Data Analytics Tools", Journal of International Technology and Information Management, Volume 26, December 2017.

[10] Matthew N. O. Sadiku<sup>1</sup>, Adebowale E. Shadare, Sarhan M. Musa and Cajetan M. Akujuobi, Prairie View A&M University, "Data Visualization", Volume. 02, December 2016.

[11] S. Aybeyan, "Systematic Review of Big Data, Digital Transformation Areas and Industry 4.0 Trends In 2021," International Scientific Journal Vision, vol. 6, no. 2, pp. 27-41, 2021.

[12] M. Vinicius, M. M. V. Franco, I. S. Rampasso and R. Anholon, "Operationalizing Business Model Innovation through Big Data Analytics for Sustainable Organizations," Business model Innovation Research - Sustainability 12(1):277, no. 12, p. 277, 2022.

[13] Venkateswarlu Pynam, R Roje Spanadna, and Kolli Srikanth, "An Extensive Study of Data Analysis Tools (Rapid Miner, Weka, R Tool, Knime, Orange)," SSRG International Journal of Computer Science and Engineering, vol. 5, no. 9, pp. 4-11, 2018.

[14] Sameer Shukla "Unlocking the Power of Data: An Introduction to Data Analysis in Healthcare," International Journal of Computer Sciences and Engineering, vol. 11, no. 3, pp. 1-9, 2023.

[15] Samuel Soma Ajibade, and Anthonia Adediran, "An Overview of Big Data Visualization Techniques in Data Mining," International Journal of Computer Science and Information Technology Research, vol. 4, no. 3, pp. 105-113, 2016.