Data Visualization Tools - Preview and Comparison

Fehmi Skender
PhD Candidate, International Vision University.
Faculty of Engineering and Architecture
Gostivar. North Macedonia.
fehmi.skender@vizyon.edu.mk
0000-0001-6046-5772

Violeta Manevska
Full Professor, University St Kliment Ohridski
The Faculty of Information and Communication Technologies
Bitola. North Macedonia.
violeta.manevska@uklo.edu.mk
0000-0003-4401-1230

Abstract— **Data visualization enables the display of information and data in a visual context. Data visualization tools and visual data analysis through visual elemQents make it possible to see trends and changes in data.**

Data visualization involves converting data into a visual context with the sole purpose of facilitating the human brain to understand and inspect the data, that is, to understand the results of their processing. The visual canal is the fastest human cognitive channel. Due to this, more people want to see data rather than hear or read it. In a time of digital data overproduction, it is usual for fast data processing, analysis, and display to become a priority increasingly in everyday life. On the one hand, developers and the other hand, managers and users are in constant communication, which leads to continuous updating and updating of processing tools and visual data analysis.

In this context, the paper will review the current visualization and visual data analysis tools and compare their performance.

Keywords: Big Data, Data Visualization, Visual Data Analysis, Data Visualization Tools.

I. INTRODUCTION

Visual data analysis (VDA) is a technique for preparing and displaying data to clarify the data for analysts and managers, as well as all those who deal with the management and use of big data in their management of activities. Big data analysts are particularly interested in data visualization tools, which will enable the interpretation of as much useful information as possible and support them in the decision-making process [1].

The rapid development of technologies is primarily based on the use of IT technologies but also on data analysis with the growing use of visualization, i.e., more accessible information in the visual form [2]. The visualization itself is based on the rapid perception of visual forms by man because the brain of an average person quickly memorizes visual representations and data that it will receive from the environment through the visual sense. Science is increasingly concerned with finding ways and procedures to facilitate daily human life. The competition between the tools available in the market for data visualization is based on the rule that they are "closer" to users, ie. User friendly.

At a time of increasing dynamism in all spheres of life, the process of making the right decision is becoming more and more relevant. If it is known that the wrong decision is better than indecision, then it is very clear why the tools for visual display of data are increasingly a factor for success in decision making. Because of all this, different data visualization and visual data analysis tools are increasingly present in the software markets today. While developers try to find simple, content-friendly, and easy-to-use tools, users have a more excellent choice with more user-friendly tools.

II. RELATED WORKS

The capabilities of the highest-ranking tools for data visualization and visual data analysis are continuously applied and, simultaneously, upgraded with new features by many researchers and analysts. To summarize some of their results, we will review the most commonly used techniques, methods, and tools used by big data analysts and researchers. Development in this area has been enhanced by the advent of powerful tools for interactive visual data analysis that is available through the many user-friendly tools available on the market. Minatogawa et al. used a design methodology to develop an artifact to assist in seeking business model innovation. The customer-driven artifact uses indicators to measure the performance of companies' business models, powered by Big Data analytics to enable business model innovations. The research results show the artifact's successful assistance in the proactive and continuous effort toward creating the business model. Although based on technical concepts, the artifact is accessible to the small business context, helping to democratize business model innovation practices and big data analytics outside large organizations. Big data and its visualization facilitate the results of large statistical operations and help decide and mediate innovations, especially in the business sector [3].

Alharthi et al. visually researched the organization of the Hajj religious rite in Mecca, Saudi Arabia. They emphasize the importance of using graphical visualization with graphs and tables to analyze big data. Their specific research is related to the organization of religious rites and required and offered services. It flows of believers for specific research on the necessary data for Hajj, which the organizers needed in the decision-making process. The paper shows how to effectively use big data visualization based on deeper details in service visualization, thus identifying needs for improvement of certain services. The research clarifies the fact that ordinary visual representations on a map do not provide details that focus on improving services by taking the right action at the right time. Advanced visual data analysis

techniques are needed to detect problems and focus the attention of decision-makers on proper assessment to improve services. [4].

Kennedy et al. researched an excellent approach to studying the factors influencing the general public and the media. Among other things, the research proposes bridging the paradigms for HCI (human-computer interaction) and media and communication studies to develop awareness among people about the effective use of data visualization as a method. The conclusion of the paper mentions the need for greater understanding and interpretation of the visual data of the general public, i.e., for different groups of society [1].

Jagadish et al. explored the technical challenges of visualizing large amounts of data. In addition to defining many concepts of big data and their visualization, it examines the challenges posed by big data visualization but also explains big data as a certainty in the modern way of life, in business, and their use by the wider public [5].

Ventartaman et al. explored the possibilities of using cloud computing in visualization, with predictions about the new data center constellation and the use of the VDA. Here, the big data visualization process analysis is based on providing business intelligence for timely and effective decision-making. The paper recognizes the value of creating a big data infrastructure and delivering higher performance and scalable business intelligence in different organizations. The use of state-of-the-art tools and technologies for big data infrastructure, as well as the NIST (National Institute for Technology Standards) framework, is demonstrated. The advantages of data visualization are illustrated with in-depth scenarios of various production examples. In short, the paper contributes greatly to providing valuable insights into the flow of big data from organizations to enable a scalable infrastructure to make more informed and better critical decisions [6].

Rushton, G. has taken rudimentary steps to establish the basis for spatial analysis for Geographic Information Systems - GIS, for future data processing and more effective data processing and decision-making primarily in the health sector. In short, the paper contains systems that contribute to general public health in the United States that are geographically mapped so that they can be considered the beginning of the visualization of modern data [7].

Kumar and Singh in their paper reach large amounts of structured, unstructured, and semi-structured data, in short, heterogeneous data concerning big data, their processing tools, analysis, and decision-making techniques. Their research is based on the impact of big data in the healthcare sector and various tools such as Hadoop in exploring the conceptual architecture of data analysis. The research is based on the use of the Hadoop and MapReduce tools [8].

Debrus R. presents detailed information and a range of capabilities on Power BI data visualization boards. Explains the reasons why processed data becomes the very final process visualization essentially. The paper presents the advantages of Power BI in research, such as user interface, Consolidation of multiple data sources (Excel, CSV, XML, Text), interactive reports and map mapping, R, Python, and SQL integration, through various examples of the benefits of data analysis in decision making have been demonstrated [9].

Der. G and Everitt B in SAS research papers, justify the justification of emphasizing the skills necessary to perform statistical analysis based on visualized results. The effectiveness and importance of the SAS tool as a basis for processing big data in the study are also explained. The paper also contains research done on divergent diagrams, correlation, simple regression, and operations offered by the SAS tool [10].

In the research of Mani. M and Fei. S data visualization is considered an integral part of big data analysis. The research examines effective ways of visualizing big data, focusing on visualizing interactive processes. During an interactive visualization session, the analyst can issue multiple visualization requests, with each visualization subsequently building on previous visualizations. The research covers integrating distributed data processing systems that can effectively process large data with a visualization system, even effective interactive visualization for smaller amounts of data. Emphasis is also placed on the search for alternatives so that the delay period of the visualization is minimized. All this is taken as a conclusion which is also a possible answer to future demands, as a basis for new experiences for analysts, which would increase their productivity [11].

Many other researchers have worked with tools that enable visual data analysis of big data and have obtained outstanding visual representations. Because of this, we will look at the possibilities of the tools most commonly used today for VDA and interactive visualization.

III. PREVIEW OF DATA VISUALIZATION TOOLS

Undoubtedly, data visualization is a rapidly growing scientific field that arouses great interest among data scientists and the scientific community that applies this science in practice to obtain more talkative, faster, and more efficient observations that will lead to data penetration. And will help make more efficient, quicker, and more useful decisions for their organizations. Therefore, presenting the data, which prefers using innovative techniques, primitives such as colors, elements, and dimensions, and analyses that affect the representativeness of the data, is very important. Because data visualization, especially in big data, requires not only knowledge of design and data but also basic statistical knowledge, a variety of visualization tools have been developed and used. We would single out a few of them.

a. Tableau Public

Tableau is one of the Gartner group's most highly rated visualization and visual data analysis tools. It is rated as one

of the best visualization tools today with several versions, including the version available to the general public but with limited capabilities, named Tableau Public. The tool stands out with its easy, intuitive user interface and simple use. With Tableau Public, all kinds of visualizations can be performed easily and quickly, without requiring code information. To get the full capabilities of the software, you need to pay for a license and get the full paid version. In the paid version of the Tableau tool, the possibilities are greater, and at the same time, different types of data inherent in Excel or PDF format can be combined. In short, with Tableau Public, data can be visualized simply by creating tables, lists, maps, and many tools that are particularly effective for visualizing big data, but also for interactive data analysis between the data and the visualizations themselves [12].

With Tableau Public, data and tables can be combined or linked together. Data can be edited using group and cluster properties. Tableau Public is used in many sectors, from private companies to public institutions. You can share your own visualized data with other people online. The data can also be shared with mobile devices compatible with the tool [13].

The architecture of the Tableau lives version makes it easy to connect to Hive (Big Data Business Intelligence) resources, enabling online data analysis without needing any data transfer (when data reaches an order of five bytes in size). Tableau live's architecture protects investments in both big data and data warehouses. Tableau can also visualize large data by connecting directly to the operating database in memory. Tableau enables the analysis of hundreds of millions of rows of data with tremendous speed and capacity, and streaming capabilities.

The main features of the Tableau tool are: Adding a comment to the table, changing the view, dragging and dropping, having a Tableau Reader for viewing data, converting data into visualizations, the ability to reduce the size of data at different intervals, creating interactive tables, data exchange, highlighting and filtering data, spreadsheet sharing, data notifications, automatic display of new features and security permissions at any level.

b.   Yellowfin Bi

Yellowfin BI is an analytics platform specializing in spreadsheets and visual analysis. As a modern tool for data visualization, it brings with it a rich library of pre-developed spreadsheets. The tool, with its structure, allows for greater performance of the users. The manufacturer offers a wide range of additional tools with the possibility of automated insight, built-in explanations, and contextual analysis. The tool easily connects to various resources and relational databases, such as Hadoop and NoSQL. With the features provided by the tool, faster responses to KPI based requests can also be found [14].

Features of the Yellowfin BI tool are: The answers are based on machine learning, as a result of automated insights,

and machine learning algorithms, providing on-demand visualization and quite understandable and easy-to-present visualizations. Getting focused results by selecting or excluding certain data points in a spreadsheet, as well as enabling complete control of data points in their visualization, is another advantage of this tool. The mobile application is based on HTML5 and replaces the source application in version 6.3 of the tool. The tool automatically detects changes in the data and gives the results at the same time [15].

c.   SAP ANALYTICS CLOUD

SAP ANALYTICS CLOUD is a tool for data visualization for business research, primarily designed for planning companies and analyzing their oversights and decision-making. Different data can be repeated and used at any time during processing. SAP ANALYTICS CLOUD supports hands-on collaboration based on simply sharing information in PDF format with SAP Analytics Cloud users as well as external collaborators [16].

Features of the SAP Cloud Analytics Story tool are: SAP Cloud Analytics Story software enables the use of a wide range of professional, interactive, important lists and other objects to display your data responsibly. By applying custom filters to widgets and spreadsheets, data can be personalized, and bookmarks can be entered into specific views for future research. The built-in calendar allows assigning, planning, and monitoring of the status, as well as temporary reminders for the same. Data can be shared on the spot with all dynamic elements when using the tool and in PDF format without dynamic functionality. At the same time, it provides security at the level of companies with access permits for owners, and users, individually or in groups.

d.   ORACLE ANALYTICS CLOUD

ORACLE ANALYTICS CLOUD is a data visualization tool for joint reporting with analytics for organizations of all data sizes, with the capabilities of the machine learning tool enabling the detection of critical information through intuitive data visualization.

One of the major advantages of ORACLE ANALYTICS CLOUD is the processing of data in the natural language, i.e., there is no need for good mastery of code to discover and analyze data that is of interest to the user. At the same time, ORACLE ANALYTICS CLOUD, with the help of multidimensional data analysis and scenario simulations, provides a much more detailed view of business data [17].

Features of the ORACLE ANALYTICS CLOUD tool are: The ability to automatically create visualizations by identifying algorithms based on machine learning in big data analysis is multiple features of the tool. The intuitive interface of the tool provides identical data analysis in various data visualization techniques and methods.

e.   DOMO

DOMO is a cloud-based Business Intelligence (BI) platform with thousands of built-in visualizations, including more than 1,500 list types and approximately 7,000 maps. Machine learning enables automated data detection by alerting and asking questions in natural language. The DOMO tool can also access, filter, sort, and group data deep after its data genesis. The tool lets you customize your layout by adding custom colors, images, and text [18].

Features of the DOMO tool are: Ability to combine individual views, for example, combining sales, return on investment, performance measurement, and all other key performance indicators that can be selected in one interface. With the help of the Pop-Up menus, the individual visualizations from the table can be separated. The tool provides the opportunity to improve the appearance of the visual display by selecting different types of lists, as well as data series and filters to enhance the visualization itself interactively.

f.   POWER BI

Power BI is a collaborative tool for software services, applications, and links that work together to enable the transformation of independent data sources into consistent, visual, and interactive insights. The data can be from Excel, a collection of cloud data, or hybrid systems in various formats. Power BI makes it easy to connect to current data sources, and visualize and share any part of the results in the cloud or between groups of users. Power BI consists of several different components that work together, including the three basic elements:

- Windows desktop application or Power BI Desktop,

- Online SaaS or Power BI service,

- Power BI mobile applications for Windows, iOS and Android devices.

Features of Power BI tool are: The biggest advantage of Power BI is its availability and being relatively inexpensive as a data visualization tool. Power BI Desktop version is free. It is easy to download, install and use.

The use of Power BI is with easy custom visualizations in reports and spreadsheets and allows interactivity in visualization. Features include many tools, such as visualizing key performance indicators, maps, charts, graphics, R image scripts, dashboards, and more. An important advantage of using Power BI as a data analysis tool is the ability to transfer data from a wide range of sources and link data to XML and JSON formats. In addition to the many advantages, one of the most important features of Power BI is the establishment of direct access to big data resources. It should also be noted that Excel integration includes the option to load Excel data into Power BI. The option to import unprocessed data into Excel allows it to be available for processing, to cut some of it as data from Power BI, and to easily transfer it back to Excel. Power BI can be used to add various visualizations to the report with the ability to establish interaction [19].

g.   Other Tools

Other Gartner-rated data visualization software applications include:

TIBCO SpotFire **-** is a software platform that assists in software integration and business visual analysis. The software is also used as software for data integration and analysis, as well as for rapid detection of concepts for better decision making. TIBCO Spotfire is a software platform that simply"animates" data, in various big data surveys [20].

SAS VDA - SAS (Statistical Analysis System) is a software package that provides statistical analysis that provides access to big data, data management, analysis, and presentation of data. It also offers a range of products that can be used in many areas related to data analysis, data cleansing, or data analysis operations. SAS is mainly used in big data processing and visualization [21].

Google Data Studio is a free online tool that converts data into customizable info reports and spreadsheets. Google Data Studio combines the marketing tools of Google or different data sources in one interface, all with the sole purpose of creating a spreadsheet. It is quite easy to use and has a simple interface [22].

Qlik Sense -is a visualization and data discovery tool that allows you to create flexible, interactive visualizations and make meaningful decisions. The tool is licensed, which answers questions that come after the inspection or visualization. Qlik Sense answers questions constantly with its relational model, allowing continuing insights [23].

Amazon QuickSight - is a licensed tool. Amazon QuickSight lets everyone in organizations understand the data by asking questions in natural language, browsing through interactive spreadsheets, or automated machine learning. QuickSight supports millions of spreadsheet views for customers all the time with the sole purpose of assisting customers in decision-making [24].

IV. DISCUSSION AND RESULTS

The research of several papers in the field of data and information visualization shows that the competition moves the software of the tools proportionally to the needs and opportunities. In summary, the current visualization tools in terms of their characteristics can be shown in the following table:

TABLE I. *STRENGTHS AND WEAKNESSES OF DATA VISUALIZATION TOOLS.*

DATA VISUALIZATION TOOLS	
ADVANTAGES	*WEAKNESSES*
TABLEAU PUBLIC	
• Manages big data and machine learning applications. • Salesforce can integrate with advanced database solutions like Hadoop, SAP, Teradata. • Effective graphics can be created.	o There is no option to refresh reports automatically. o The solution is not so comprehensive, and knowledge of SQL is inevitable.
YELLOWFIN BI	
• Ability to analyze operational and strategic planning of the organization by following the critical metrics through updated visualizations. • Expand spreadsheets by combining widgets, action keys, and codeless functionality with the JavaScript API. Create floating panels, filter controls with HTML, CSS, and JavaScript	o Does not allow change of hierarchies and data types inherited from OLAP. o Does not support financial planning and profit analysis
SAP ANALYTICS CLOUD	
• The tool allows deeper penetration into the data source through the Smart Predict, Smart Insights and Smart Discovery modules, allowing in-depth analysis. • Eliminate foresight in the decision-making process. • Save time, select data to find information when automatically generating relevant insights and descriptions	o Sometimes does not provide solid data validation for AI-based features. o Data modeling is not flexible enough.
ORACLE ANALYTICS CLOUD	
• Ability to automatically create submersible visualizations by identifying insights through machine learning algorithms. • Its intuitive interface provides the same data analysis in multiple visualizations. • Provides specific answers to specific business questions through personalized searches. • Regardless of technical skills, it provides opportunities for research and analysis.	o Does not provide enough graphics. o Not intuitive in terms of reinforcement and recovery. o Does not include visualization like its competitors.
DOMO	
• Ability to ask and answer questions from everyday life, the ability to lower the standards for in-depth analysis. • Working in teams with clients on visualization reduces the need for meetings. • Ability to record specific data points, leave comments and observations so they can be tracked by other clients.	o Does not allow export of more than 5 MB of additional data in the scheduled CSV reports. o Does not allow easy removal of data from the platform. o Not intuitive enough in terms of user interface.
POWER BI	
• The Power BI user receives data from various data sources, such as files, Azure Resources, online services, queries, or gateway resources. Then they work with this data in a development tool like Power BI Desktop. Here, the imported data is cleaned and converted according to the needs of the user. • Reports created on the Power BI Desktop can then be published on two types of platforms: Power BI Service and Power BI Report Server. Power BI Service is a public cloud-based platform, and Power BI Report Server is a firewall-protected platform. • Spreadsheets and reports can be shared using web browsers, tablets, laptops, phones, etc.	o Provides AutoML only in Power BI DataFlows. This means that it is not possible to run AutoML on a Power BI Desktop. o It is also not possible to autom after creating the data model. o There is no intelligence for an advanced position. o Lack of opportunity for graphical analysis. o PowerQuery & m is the most powerful tool in Power BI, while DAX and the data model are an obstacle that weighs equal to Power BI.

V. CONCLUSION

As a result of the analysis of the mentioned tools for visual data analysis and visualization of big data, we conclude that they are suitable for use in organizations. These tools have been developed for solving dynamic problems such as creating losses or illiquidity. Modern organizations use visualization tools effectively, thus advancing production and services much more confidently and with better quality.

Big data processing and its visualization greatly help all users, consumers, manufacturers, and managers. More and more attention is being paid to using visual facts when deciding and investing. Using tables and graphs is a certainty in today's dynamic time in such dynamic environment.

The number of tools for data visualization is growing rapidly, and they are constantly improving and increasing their capabilities for analysis, especially in the field of Big data. If we know that the human eye is attracted by visual representations, colors, and patterns, i.e., 90% of the information presented in the brain is visual, it can be easily concluded that data visualization is the most desirable technique for business and, in general for everyday needs for visual data analysis. in any area of life. As a result, the trend of increasing the use of visualization tools is increasing, and the use of tools is becoming more than necessary.

Therefore, the paper analyzed research on previously published articles that use big data visualization for various tools, including tool analysis, video data analysis, and interaction. The study focuses on big data using visualization tools where their key performance indicators are analyzed. Finally, it can be rightly noted that Big Data, Visualization, Machine Learning, and Artificial Intelligence are not only a challenge for analysts, managers, managers, and professionals but are increasingly becoming a necessity and inevitability for every human being. The analysis supported by the possibility of interaction in the processing and visualization of big data results in a solid basis for obtaining useful and insightful information as a prerequisite for success in all fields.

REFERENCES

[1] H. Kennedy, "Engaging with data visualisation," http://journals.uic.edu/ojs/index.php/fm/rt/printerFriendly/6389/5652, 2016.

[2] 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.

[3] 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.

[4] A. Alharthi, V. Krotov and M. Bowman, "Addressing barriers to big data," p. https://www.researchgate.net/publication/317279522, 2017.

[5] Jagadish, J. Gehrke, A. Labrinidis, Y. Papakonstantinou and J. Patel, "Big Data and Its Technical Challenges," Communications of the ACM, pp. 86-94, 2014.

[6] VENKATRAMAN, "Cloud Computing A Research Roadmap In Coalesence With Software Engineering," An International Journal (SEIJ), vol. 2, no. 2, 2012.

[7] G. Rushton, "Public health, GİS and spatial analytic tools," Annu.Rev. Pub.Health 2003 - Department of geography - Unıversıty of Iowa, USA, no. 24, 2003.

[8] Maninder and S. S. Kumar , "Big Data Analytics for Healthcare Industry," Big Data Mining And Analytics, vol. 2, no. ISSN222096-0654 1 l05/06l, pp. 48-57, 2019.

[9] D. Roxane, "Power BI in Clinical Data exploration,," PHUSE Virtual EU Connect 2020, 2020.

[10] G. Der and B. S. Everitt, Statistical Analysis of Medical Data Using SAS, vol. 16, 2006, p. http://ftp.sas.com/samples/A60928.

[11] S. F. Murali Mani, "Effective Big Data Visualization," In Proceedings of IDEAS '17, Bristol, United Kingdom, July 12-14, 2017, 6 pages., 2017.

[12] "https://www.tableau.com/products/public," TABLEAU SOFTWARE, 2022. [Online]. Available: https://www.tableau.com/products/public.

[13] A. Ohmann and M. Floyd, Creating Data Stories with Tableau Public, Packt Publishing, ISBN: 9781849694766, 2015.

[14] "https://wiki.yellowfinbi.com," Yellowfin, 12 6 2021. [Online]. Available: https://wiki.yellowfinbi.com/display/yfcurrent/REST+API.

[15] www.yellowfinbi.com, "https://www.yellowfinbi.com/," 2022. [Online]. Available: https://www.yellowfinbi.com/.

[16] https://community.sap.com, "https://community.sap.com," 2022. [Online]. Available: https://community.sap.com/topics/cloud-analytics#:~:text=SAP%20Analytics%20Cloud%20is%20a,you%20with%20your%20learning%20journey.

[17] https://docs.oracle.com, "https://docs.oracle.com/en/cloud/paas/analytics-cloud/visualize-data.html," 6 2022. [Online]. Available: https://docs.oracle.com/en/cloud/paas/analytics-cloud/visualize-data.html.

[18] https://www.domo.com/, "https://www.domo.com/solution/data-visualization-software#:~:text=Domo's%20data%20visualization%20software%20lets,faster%2C%20better%2Dinformed%20decisions.," [Online]. Available: https://www.domo.com/.

[19] A. S. Rob Collie, "Power Pivot and Power BI," in Power Pivot and Power BI, Yayınevi: Holy Macro! Books / ISBN: 9781615473496, 2015.

[20] sistembul.com, "https://www.sistembul.com/cozumler/tibco-spotfire," 21 3 2022. [Online]. Available: https://www.sistembul.com/cozumler/tibco-spotfire.

[21] globaltechmagazine.com, "https://www.globaltechmagazine.com," 2022. [Online]. Available: https://www.globaltechmagazine.com/2016/06/06/sas-viya-devrim-yaratacak-yeni-analitik-mimarisi/.

[22] newgenapps.com, "https://www.newgenapps.com/mk/blogs/what-is-the-google-data-studio-2/," 18 12 2019. [Online]. Available: https://www.newgenapps.com/mk/blogs/what-is-the-google-data-studio-2/.

[23] "https://home.vizlib.com/vizlib-library-for-qlik-sense/," vizlib, 2022. [Online]. Available: https://home.vizlib.com/vizlib-library-for-qlik-sense/.

[24] "https://aws.amazon.com/quicksight/," aws.amazon.com, 2022. [Online]. Available: https://aws.amazon.com/quicksight/.