

Significance of Digital Data Visualization Tools in Big Data Analysis for Business Decisions

Kirti Mahajan, PhD
Bharati Vidyapeeth Deemed University (BVDU)
IMED
Pune, Maharashtra, India

Leena Ajay Gokhale
Bharati Vidyapeeth Deemed University (BVDU)
IMED
Pune, Maharashtra, India

ABSTRACT

Big Data is the large volume of data collected through various resources which integrates variety of structured, semi structured and unstructured data making it complex to analyze by using conventional methods and tools. Computer based visualization is the way of communicating the analyzed information to the user for the purpose of taking data driven business decisions. Visualization of the data is the way of presenting the large data graphically. It is a technique of displaying the information in such a graphical format that the information becomes easy to understand to the user. This interaction between the computer and the user is called graphical user interface. The entire process involves storing, processing and presenting the huge data for which different tools are being used. The tools used for storing and processing the data are considered as backend tools while visualization tools are called frontend tools. Visualization aims to improve the precision and manifestation of the information discovered. There are many conventional methods used for visualization of data. However they miss the potential of the Big Data. Images present the content in varied formats including graphs, charts, moving images, animation. Appropriate visualization of big data can have a major impact on the business in discovering hidden insights, improving decisions and automating business processes. Therefore specific framework has to be designed which will give precise idea to the user about the entire process. This paper shows the panorama of the significance of visualization of big data, the protocols to be followed while selecting the visualization tool and the classification of these tools based on various factors.

General Terms

Data analysis, Visualization, Visualization Tools, Graphical User Interface (GUI).

Keywords

Big data, big data analysis, digital data visualization, digital data visualization tools and techniques.

1. INTRODUCTION

Big Data is the large volume of data collected through various resources integrating a variety of structured, semi structured and unstructured data. The size of big data is increasing due to the ubiquitous sources of digital data such as computers, smart mobile phones, digital cameras, social networking websites, radio frequency identification readers and other machine generated sensor data. Due to advancement of digital technologies, the growth in the volume and variety of big data will keep on growing speedily. [7] The conventional digital data analyzing tools are not adequate to analyze and get significant information from this complex data. [3]

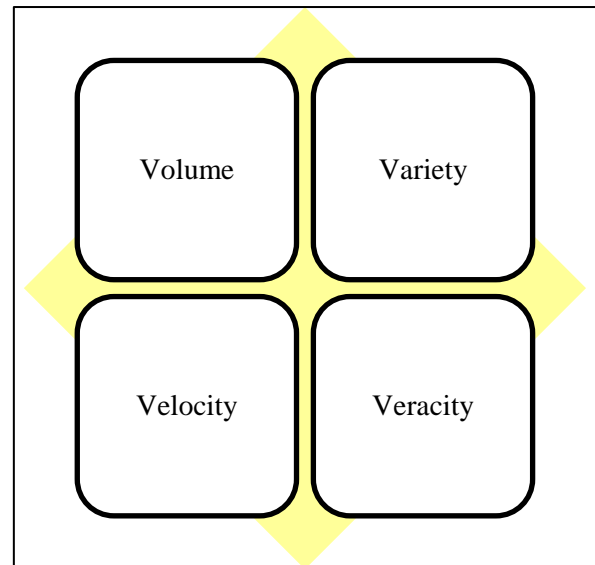


Fig 1: Features of Big Data

- 1) Volume:** The digital data in many businesses has reached to the volume of Petabyte (PB), Exabyte (EB), Zettabyte (ZB) and Yottabyte (YB) [4].
- 2) Variety:** The digital data is in the varied formats such as structured tables, semi structured emails, unstructured chat records, text files, audio files, video files and pictures.
- 3) Velocity:** It is the speedy analysis of big data which is essential to get the timely information.
- 4) Veracity:** Authenticity and accuracy of this data is essential as numerous decisions rely on the information presented after analysis.

2. BACKGROUND

A pilot study was being conducted to understand the significance of digital data visualization tools that are in practice in the manufacturing companies in Pune, India. The data was collected through questionnaires and personal interviews of the focused group of managers.

- 1) The accuracy, simplicity, speed and the cost of the visualization tool are vital factors that are considered while selecting the visualization tool.
- 2) According to the pilot study, managers find visualized information easy to understand and explain.
- 3) Various business decisions depend on the information presented through processes such as analyzing market

trends, analyzing customer requirements, inventory management, production planning and control that can be enhanced if these processes are visualized.

- 4) Different data such as product data, customer data, financial data and inventory data is there for visualization.

After the pilot study, it was being observed that digital data visualization plays a vital role in the business for making data driven decisions.

3. BIG DATA ANALYSIS

The process of big data analysis involves collection, storage, analysis and visualization of the big data. [5] Each of these processes is equally significant for the businesses. However the purpose of each process is different. While big data features (Diagram 1) describe what logical asset businesses have, big data analysis describes how this big data can be useful for gaining benefits for organization. Big data analysis helps in numerous ways new product development, smart decision making, understanding customer behavior and predicting sales. Different tools are used for this analysis. The tools that are used for storing and processing the data are considered as backend tools while the tools used for visualizing the analyzed information tools are called frontend tools.

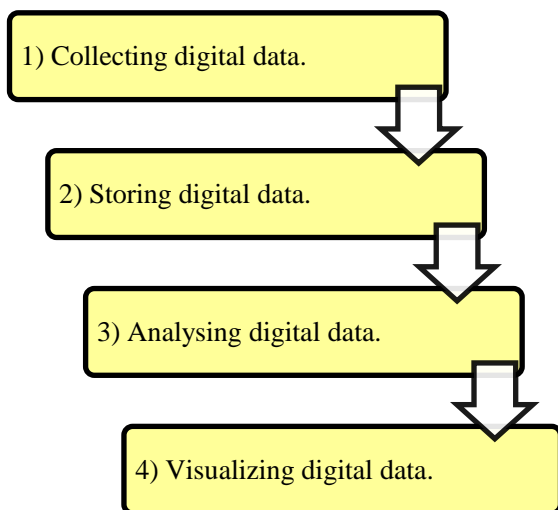


Fig 2: Steps in Big Data Analysis

1) The businesses are collecting the digital big data through the varied sources. Each business may need assorted data depending upon the type of business. Organizations may integrate data about their students, customers, products, services, sales, finance and patients. [9]

2) Due to the varied sources of collection, the formats of digital data are also different making it difficult to store in uniformed way.

3) Analysis of digital data is vital to extract the meaningful information from the huge datasets.

4) The graphical presentation of the analyzed information is easy to communicate.

4. DIGITAL DATA VISUALIZATION TOOLS

Digital data visualization tools provide an immense opportunity to explore significance from digital big data. Visualization tools have the potential to create business value. Big data visualization leads to many advantages such as

smarter business decision making, new insights, new business opportunities and increased profitability. Visualization is a vital tool for understanding the information quickly. It is a technique in which the information is presented in a graphical format to the user for the purpose of enhancing data driven decisions. It is presumed that the humans will be using eyes and brain for understanding this graphical information. Accurate visualization follows a specific format to display the information. Before visualizing the data, it is essential to understand the data types and the possible relations among them.

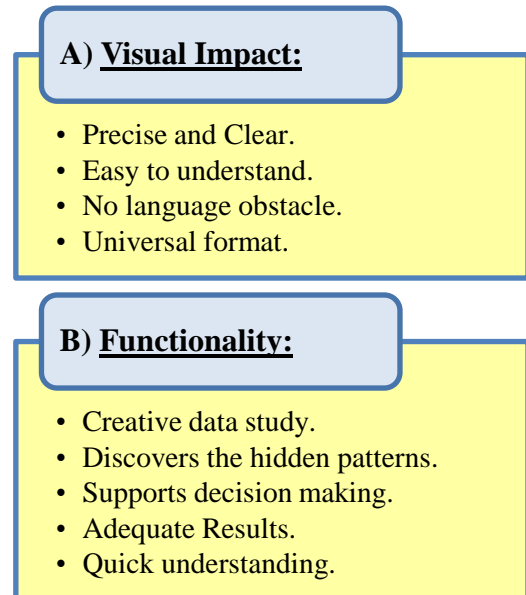


Fig 3: Features of Digital Data Visualization Tools

4.1 Classification of digital data visualization tools

These tools may be further classified as per the dimensions used to display the data, interactions made through the visualization, coding required for using them and whether the tool is available online for the user or it has to be used offline.

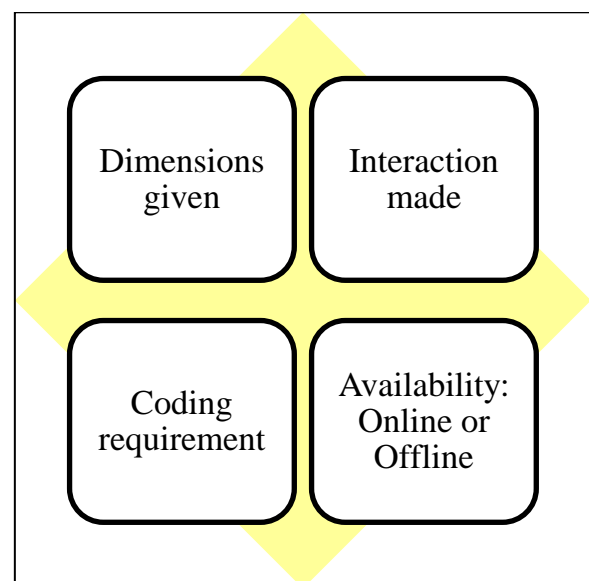


Fig 4: Classification of Digital Data Visualization Tools

1) **Dimensions used:** The visualized information may be one dimensional, two dimensional or multidimensional. The dimension requirement may vary depending upon the type of information that is to be presented.

2) **Interaction made:** The visualization may be static visualization in which the information presented is in flat format and does not allow the user to interact with the graphic further. Interactive visualizations allow the user to interact with mouse click or other events which display further information after the event is fired. Animated visualization has moving objects for explaining the information. Animated visualizations are good when some process or procedure is to be explained to the user.

3) **Coding requirement:** The visualization tools that require coding may be good for the developers and the tools that do not require any coding are suggested for the non developers.

4) **Availability Online or Offline:** Some tools such as web based visualization may be available online for the users. Other tools may have to be used offline.

Visualization consists of comparison among the datasets and understanding relationship among these datasets. Numerous methods are available for visualizing the digital data. However selecting the appropriate chart that will suit the requirement of the business is vital. Bar Chart, Pie Chart, Column Chart, Line Chart, Area Chart, Scatter Plot are the conventional methods of displaying the data graphically. Bubble Chart, Heat Map are the modern methods of visualizing the data.

4.2 Protocols for selecting digital data visualization tool

Both physical and logical appearance of the visualized information is equally important. While physical appearance takes care of the creative sense of the design, logical appearance is responsible for displaying the accurate information required. The precise framework is essential before beginning the visualization process. Various factors should be taken into consideration at the initial planning stage as different tool may be used for varied purposes. Understanding the prerequisite and the desired output should be clear as the first step of selecting the visualization tool.

Visualization consists of comparison among the datasets and understanding relationship among these datasets. Numerous methods are available for visualizing the digital data. However selecting the appropriate chart that will suit the requirement of the business is vital. Bar Chart, Pie Chart, Column Chart, Line Chart, Area Chart, Scatter Plot are the conventional methods of displaying the data graphically. Bubble Chart, Heat Map are the modern methods of visualizing the data. Visualization designs should follow some fundamental protocols so as to make it more understandable and effective. Different factors influencing the physical and logical appearance of the graphic should be taken into consideration as they make an impact on the viewer. Since lot of decisions are taken after viewing the graphic, utmost care has to be taken while designing and selecting the tool.

The physical appearance of the graphic is measured by the fonts, colors, size of the graphic, graphic colors, orientation of the graphic. The logical appearance of the graphic can be set with the tool and data selected depending upon the audience of the presentation.

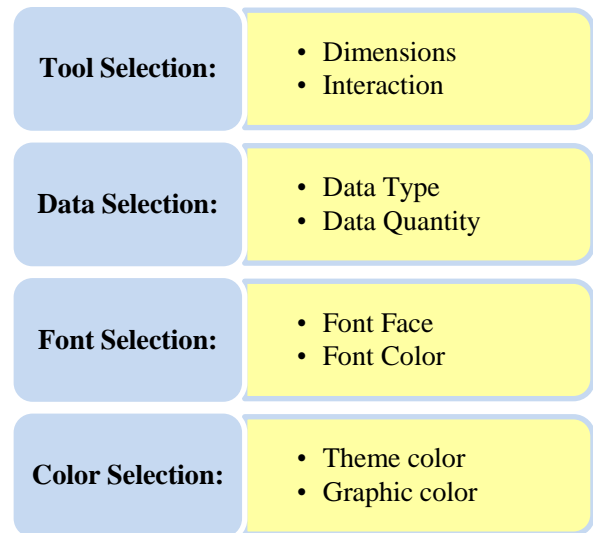


Fig 5: Protocols for selecting Digital Data Visualization Tool.

Appropriate tool should be selected for the precise data as different type of data may be represented by different visualization tool. The color scheme should be chosen carefully as per the type of information to be presented and the formal or informal type of audience. It has to express the specific information that is to be communicated. Using appropriate labels help the user to understand the information. The font has to match the purpose of the visualization. The purpose of the visualization and audience that will be viewing the information should be understood before selecting the design. Determining what type of information is to be shared is essential. Knowing the audience to whom this information is to be presented as depending upon the type of audience, the simplicity level of presentation can be taken into consideration. The expected impact of the information presented on the audience. The communication should be clear, precise and easy to understand.

5. SIGNIFICANCE OF DIGITAL DATA VISUALIZATION

Digital data visualization is a graphical representation for data that help users to understand the data. Understanding the data and extracting the meaningful information is indispensable part of any business. Discovery of the hidden insights is reliant on the computer based data visualization. The abstract information can be conveyed in an enhanced way if presented graphically. The emphasis on visualization should be given as the volume and variety of big data is increasing rapidly. The output of these tools is significantly interactive and dynamic. Visualization has the power to identify the areas that need attention or improvement. It addresses issues at the right time before they become complex problems. By comprehending information quickly, user gets support for data driven decisions. It identifies relationships and patterns and communicates the message quickly. Since the quality of the information depends on accuracy, reliability, timeliness, precise, completeness and presentation of information which are taken care of by digital data visualization tools.

The main purpose of the visualization is to present the information in a graphic format that is easy to understand conveys the hidden insights to take user. Visualization aims to improve the precision and manifestation of the information. Computer based visualization is useful for taking precise

data-driven decisions. An aim is to display the meaningful results. Making sense of facts, numbers and measurements is a kind of art called data visualization. It is an extremely effective communication from the complex information.

5.1 Merits of digital data visualization

- 1) Graphically displayed information is easy to understand.
- 2) It allows communicating the information irrespective of the languages known to the user.
- 3) When visualized, the large information can be understood at a glance reducing the time taken in reading the information in text or table formats otherwise.
- 4) Such information helps in enhancing the business decisions.
- 5) It allows a quick view towards the complex information that would have taken long time to be understood otherwise.
- 6) The user may not depend on others to help him understand the information as the images are self explanatory.
- 7) The user may not require specific skills for interpreting the visualized information.
- 8) It gives more accurate impression of the data than mere words.
- 9) Interaction between the user and the image becomes more user-friendly.

5.1 Opportunities and challenges of digital data visualization

5.1.1 Opportunities of digital data visualization

All the businesses have the opportunities to get the benefits of big data. [1] Digital data visualization offers numerous opportunities to several sectors including manufacturing, retail, telecommunication, research and healthcare. Big Data is going to be more complex, diverse, larger and faster. Since big data is the logical asset of the business, each data has to be visualized so that the potential will not be missed. Various digital data visualization tools are available by which information can be visualized in numerous formats. Any data that influences decision making has to be visualized. To create more accurate and effective data visualizations, knowing the data to be visualized is vital and some fundamental designing protocols should be followed.

5.1.2 Challenges of digital data visualization

Due to the large size, complexity and varied formats, big data is facing many challenges. [2] The key challenges are lack of trained employees, skills required, accurate analysis and selecting the appropriate visualization tool.

Presenting the information in an effective way is essential while handling the large data. If the data is misrepresented or presented ineffectively, the key information may be lost. Since the organizations are investing big amount in processing

the data, the returns in terms of accurate information has to be done. [6]

Selecting the right visualization tool for the specific type of data is the challenge. [8] If the accurate tool is not selected, the visualization may present the improper information which may result into incorrect business decisions that are based on the information displayed by these tools. Perfect visualization of the data may have a major impact on the business in discovering hidden insights, improving decisions and automating business processes.

6. ACKNOWLEDGMENTS

We would like to thank the respondents of the pilot study, experts for their advice and the authors whose research papers are being referred for writing this research paper.

7. REFERENCES

- [1] C.L. Philip Chen, C.-Y. Zhang 2014. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences* 275 (2014) 314–347.
- [2] Nasser T, Tariq RS 2015. Big Data Challenges. *J ComputEngInfTechnol*
- [3] Sam Madden 2012. From Databases to Big Data. *IEEE Computer Society*, 1089-7801/12, 2012.
- [4] Prof. Ray Harris 2012. ICSU and the challenges of Big Data in Science. Page 11, Section 4, Research Trends, Special Issue on Big Data.
- [5] Gerry George 2014. Big Data and Management. Martine Haas and Alex “Sandy” Pentland, *Academy of Management Journal* 2014, Vol. 57, No. 2, 321–326.
- [6] Ekaterina Olshannikova, Aleksandr Ometov, Yevgeni Koucheryav and Thomas Olsson October 2015. Visualizing Big Data with augmented and virtual reality: challenges and research Agenda. Olshannikova et al. *Journal of Big Data* 2:22 DOI 10.1186/s40537-015-0031-2, 1 October 2015.
- [7] Peter O’Donovan, Kevin Leahy, Ken Bruton and Dominic T. J. O’Sullivan 2015. Big data in manufacturing: a systematic mapping study. O’Donovan et al. *Journal of Big Data* (2015) 2:20 DOI 10.1186/s40537-015-0028-x.
- [8] Lijun Cai, Xiangqing Guan, Peng Chi, Lei Chen and Jianting Luo 2015. Big Data Visualization Collaborative Filtering Algorithm Based on RHadoop. *Hindawi Publishing Corporation International Journal of Distributed Sensor Networks*. Volume 2015, Article ID 271253, 9 pages <http://dx.doi.org/10.1155/2015/271253>.
- [9] Thomas H. Davenport and Jill Dyché May 2013. Big Data in Big Companies. *International Institute for Analytics*.