Data Visualization Method as the Facilitator for Business Intelligence

Megha Narayanan,   Sanil Shanker KP

Abstract: *Data visualization methods are used to support business analysis. This paper explores the study on the sophisticated data visualization methods for business inventiveness. These comprise Line and Bar Charts, Scatter Plots, Network Diagrams, Bubble plot, Correlation Matrices and Donut Diagram. The concepts of visual elements are explained in the context of business perception. The paper reviews the capabilities and prophecies of the visualization methods in business analysis. When more data has to be symbolized, the concentration increases and it leads to difficulty in understanding the information to be dillydallied. Data visualization methods for business decisions save the time and resources as well as provide better understanding. The study explores that there exist habituated data visualization methods that are useful in business intelligence. The methods serve as the elite outfits to epitomize Big Data effectively. These techniques are scouted through this literature review. We investigate the pros and cons of data visualization methods in business intelligence.*

Keywords: *Big Data, Business Intelligence, Data Visualization, Visual Elements, Visualization Methods*

I. INTRODUCTION

Data visualization is the process of graphically representing data with the intention of gaining insights from real world situations. Visual representation of data gives a quick understanding of what has to be convinced and interpreted. The application of data visualization spans over the area of science, education, entertainment, business, and so on [1, 2]. Among these, one of the major areas that make use of data visualization methods is the business world. Business Intelligence is a process for analyzing business data and presenting identified information which helps in making proper business decisions. Thus, data visualization roots for Business Intelligence. For a business to exist competently, the two things to keep up are: the management of time and better understanding of current status of the organization. This can be fulfilled by the proper use of data visualization methods. Numerous data visualization methods are used to support Business Intelligence. These include Line and Bar

* Correspondence Author
Megha Narayanan*, Department of Information Technology, Kannur University, Kerala, India, e-mail: meghanarayananofficial@gmail.com
Sanil Shanker KP, Department of Computer Science, Government College Thalassery, Kerala, India. e-mail: sanilshankerkp@gmail.com

Charts, Scatter Plots, Network Diagrams, and Correlation Matrices [3- 13]. Here, visual elements are the elementary edifice in a diagram to visualize data stuffs. Today, business deals with big data, which covers a large variety and volume of data. Data visualization method serves as the finest tool to represent Big Data effectively [14]. This paper explores the potentials and prospects of the data visualization methods in the realm of business intelligence.

II. DATA VISUALIZATION METHODS

For the better future of organizations, a better understanding of business activities is essential. To accomplish this, we can make use of data visualization methods which can handle the business data in an efficient manner. Business Intelligence is an efficient technique used to make good decisions for the business. The growth of business is tied up with better business decisions. There exist erudite data visualization methods that are useful in business intelligence. Those methods are reconnoitered through this literature review.

A. Existing methods

1. Table: The simplest visualization method used to represent data is table (Figure 1). Table is a structure which consists of rows and columns where each cell may or may not contain a value. Each column represents a set of similar values where each row is a record of set of different variables [3]. Table is a data visualization method that no business organization can get rid of.

Month	Expected Profit
April	50%
August	55%
October	65%
December	80%

Figure 1: Table representing the expected profit in an organization

2. Line Chart: Line chart connects data points over straight lines [4]. They are used to indicate the change of data over time (Figure 2). Often in business, analyzing the progress of an activity (for example, the yearly profit of business) to make business decisions can be accomplished with the help of a line chart.

3. Bar Chart: When line graphs often represent continuity of data points, bar chart suits well for representing discrete data [5]. Bar charts are also known as bar graphs. The bars can be horizontal as well as vertical (Figure 3).
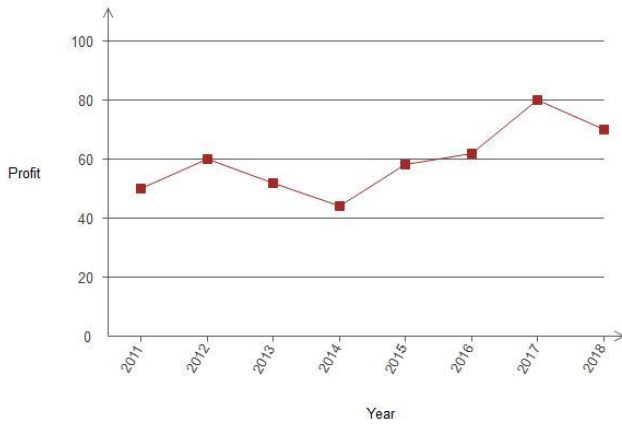
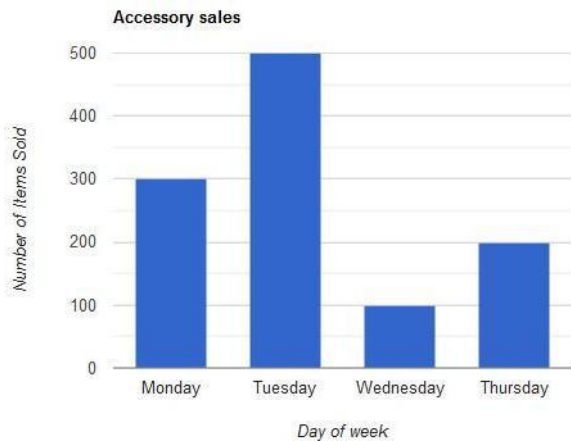**Figure 2: Line chart showing the yearly profit of an organization.**

Figure 3: Bar chart representing the sale of accessories on different days

4. Scatter plot: Scatter plot represents data in Cartesian coordinate system, and so known as X-Y plot. Since the data points lies in an X-Y plane, it is undisputable that scatter plot shows the relation between two variables (Figure 4). Other names for scatter plots are plot, plot chart, scatter chart, scatter gram, scatter graph or scatter diagram [6]. Using a scatter plot for visualizing data helps in identifying certain patterns and relations within the data points.
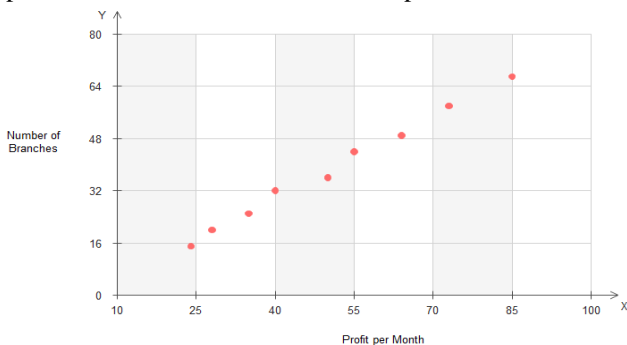

Figure 4: Scatter plot that depicts the relationship pattern of the profit percentage and number of branches of a business organization.

5. Bubble plot: Bubble plot is a variation of a scatter plot, also known as bubble chart [7]. It is represented within the X-Y plane where the size of the bubble determines the volume of corresponding data (Figure 5). We can use different colors for the bubbles to add additional meanings or information to

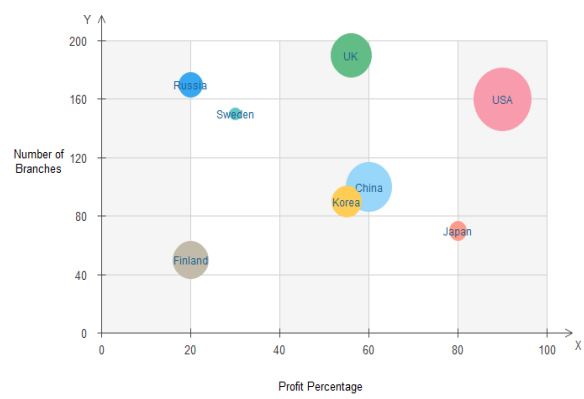the data. To indicate changes of data over time, bubbles can be animated.

Figure 5: Bubble chart depicts number of branches versus profit percentage in different countries of a business

6. Pie Chart: Pie charts are circles divided into a number of sectors, where each sector describes a proportion of the whole quantity [8]. The sectors are shown as percentage values (Figure 6). Pie charts are often used for making many business decisions. Pie charts are also known as pie diagrams.
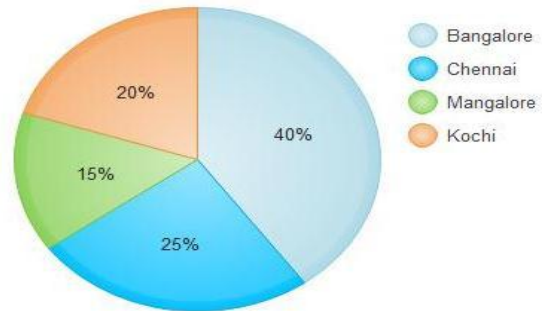

Figure 6: Pie chart representing the percentage of profit in different branches of the company

7. Donut charts: It is also known as donut diagram. The major difference with pie chart is that the donut charts have a hollow center [8]. Coloring/shading is used for differentiating the data in both pie and donut charts (Figure 7).

8. Box Plot: Like bubble plots, box plots can also show the volume of data being visualized. The bubble plot holds three measures namely, X- value, Y- value and the size (Figure 8). On the other hand, box plots can hold five values- minimum, lower quartile, median, upper quartile and maximum [9].
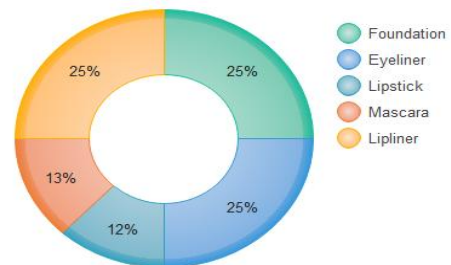

Figure 7: Donut chart representing the percentage of sales of products in a cosmetic company

Lower quartile is represented by the lower edge of the box, upper quartile is represented by the upper edge of the box, median is represented by a central line that divides the box into sections, and extreme values are represented by whiskers that extent out from the edges of the box. Box plots can be horizontal or vertical.
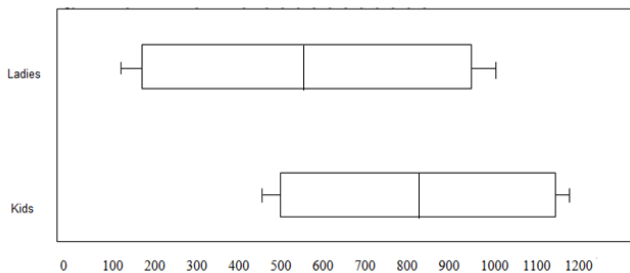

Figure 8: Box plot representing the number of items sold for a month

9. Word Cloud: Word clouds, as its name suggests, are the representation that contain a collection of words. The words will be in different size where size of the word represents its frequency within the text [10]. Sometimes, business has to deal with the importance of a word (example: to find out the most-selling product in a cosmetics company) (Figure 9). In such situations, word clouds are the most relevant data visualization method to be used.

Figure 9: Word cloud representing the demand of Lipstick brands that are sold in a company

10. Network Diagram: Similar to the computers in a network, the data points in a network diagram are spread across and related data are tied with lines. The data points are called nodes and the lines connecting them are ties. Network diagrams are one of the effective data visualization methods when dealing with Big Data [11]. Business happens across states, countries, and across the whole world. For a business to track its progress or condition over geographical areas, network diagrams can help a lot (Figure 10). For example, to represent the branches of a company over a country, or to know about the status of sales across a large area, network diagrams are the best for the interpretation of data and will result in better business decisions.
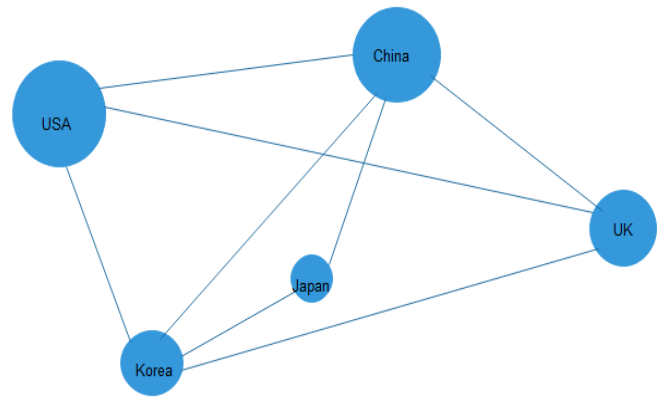

Figure 10: Network diagram showing the connection between branches of a company in different countries

11. Correlation Matrix: This is a laudable data visualization method that can be used when dealing with big data. They are used for knowing the relationship between variables (Figure 11). A correlation matrix is a square matrix in which the rows and columns have same variable names [13]. The correlation is indicated either by numerical values or by means of color. Larger values and darker colors indicate high correlation while smaller values and lighter colors indicate less correlation. So, all the diagonal values in a correlation matrix will be maximum because diagonal values show the relation of one variable to itself.

	LAKME	REVLON	DAZZLER	COVERGIRL
LAKME	1			
REVLON	0.9	1		
DAZZLER	0.7	0.8	1	
COVERGIRL	0.6	0.7	0.9	1

Figure 11: Correlation matrix to analyze the customer need of Lipstick brands in a cosmetic company

III. ADVANTAGES OF THE EXISTING METHODS

The study explores that each data visualization method has its own merits. Tables are simple to design and it provides easy understanding of the data. Line charts are better for analyzing the change of continuous data over a period of time [4]. When dealing with discrete data, bar diagrams are the best to convey the ideas. Scatter plot can show whether there exists a relation between two variables [6]. It is possible to identify the pattern of data changing in accordance with the variables. Bubble plot has a lot of advantages. They can represent the volume of data. By using coloring and animation in bubble plots, more attributes can be added to the data points. Pie and donut charts are most suitable when representing limited data components which contain percentage values [8].

Word clouds are also good in dealing with limited data items, and they can convince the idea within a very less amount of time. The word clouds are very easy to understand [10]. Network diagrams can be used for dealing with big data. They can be superimposed to geographic maps for better visualization [11]. To identify the pattern and the depth of relation between variables, correlation matrices are the finest choice. Here, numerical values or coloring strategies can be used to represent the perceptiveness of the relation [12].

IV. LIMITATION OF EXISTING METHODS

To know the appropriateness of data visualization methods in business intelligence, it is important to understand the limitation of those methods. Knowing the limitations helps to identify the context in which the method can or cannot be used. Table is a data visualization method that is commonly used in business. When there exists large volume of data, tables cannot be used. When it comes to line graph, it is a continuous line that goes with the specified time period. Since line graphs are drawn by connecting the data points by straight lines, each and every point in the line does not hold for the data point [15]. Here, we construe that, data holds for every point in the line; this sometimes may lead to fault conclusions. Scatter plots are used for identifying relation between variables. But it can only show whether two variables are related or not, and cannot show the depth or any other information about the relation. Pie and donut charts are suitable only when a limited number of data points are included. The human eyes are not capable of better identification of visual angles, which is considered as a drawback of pie charts and donut charts [8, 16]. In the case of bubble chart, it will be difficult to quantify the data decorously [17]. On the other hand, the lack of reminiscent information in the replicated result is the limitation of correlation matrix [18]. When more data has to be represented, the complexity of word cloud increases and it leads to the difficulty in understanding the information to be convinced. Network diagrams are well-suited for representing big data, but when the relation between data points increases, the number of ties (lines) increases thus making the interpretation difficult.

V. CONCLUSION

This paper is a scholastic attempt to explore and illustrate the role of data visualization methods in business intelligence. The growth and expansion of the business is determined by the decisions that the business organization flairs. Data visualization methods for business decisions save the time and resources as well as provide deepened understanding. Since the human brain always construes and memorizes data as images, data visualization is an ingenious idea to address the complexity of business environment. We explore the classy data visualization methods that are used in business intelligence, their advantages as well as limitations. Usage of data visualization techniques over business data helps to build a better foundation for business intelligence.

REFERENCES

1. Shixia Liu et al, (2014). A survey on information visualization. 30 (12): 1373–1393
2. Börner, et al, (2019). Data visualization literacy. Proc. of the Nat. Ac. of Scie, 116(6),: 1857-1864.
3. Milosevic et al, (2019). A framework for information extraction from tables in biomed literature. Int J on Doc. Anal and Recog. 22 (1): 55–78.
4. Friendly M, (2008). Milestones in history of thematic cartography, Stat. graphics and data visualization. (7): 13–14.
5. Daniel A. Keim, (2001). Visual Exploration of Large Data Set. Comm. of ACM 44 (8): 38- 44.
6. Utts JM, (2005). Seeing Thro. Stati. Thomson Books, 166-167. Tufte E, (2001). Visual Display of Quanti. Inf. Graphics Press.
7. Harris, R L, (1999). Information graphics, Oxford University Pre. Frigge, M et al, (1989). Some Impl. of the Boxplot. American Stati.. **43** (1): 50–54.
8. Collins et al, (2009). Parallel Tag Cloud to Exp & Anal Faceted Text Corpora. Proc. IEEE Sym. on Visul Anal. Sci. and Tech, pp. 91–98.
9. Herman, Ivan et al, (2000). Graph Visual & Navi in Info. Visualization- A Survey, IEEE Tra. on Visl. & Comput. Grap, **6** (1): 24– 43
10. Finlayson et al, (2001). Color by correlation, IEEE Trans. on Pat. Anal. and Mach. Intel, 23 (11) 1209–1221
11. Székely et al, (2007). Measuring & testing indep. By correl. of distances. Annals of Stat.. **35** (6): 2769–2794
12. McAfee and Brynjolfsson, (2012). Big Data The Management Revolution. Harvard Bus. Rev.90, 61-67
13. Nathan Yau, (2013). Data Points- Visualization that means something. John Wiley & Sons Inc.
14. John C. Russ and F. Brent Neal, (2016). Image Processing Handbook. CRC Press.
15. Cooper et al, (1999). New Product Portfolio Management 16(4): 333–351.
16. Ian Ford, (1986). Confounded Correlations: Stati. Limits in the Analysis of Interregional Relationship. of Cerebral Metabolic
17. Activity. Journal. of Cer. Blood Flow and Metabol. 6:385-388.

AUTHORS PROFILE

Megha Narayanan stood first in both BCA and MCA degree examinations of Kannur University in 2016 and 2018 respectively. Her areas of interest are AI techniques, algorithms for Big Data, and decision making under uncertainty. Presently, Megha is doing research in Data Visualization methods.

Sanil Shanker KP did his PhD in Computer Science from the University of Kerala in 2011. He was the recipient of European Research and Educational Collaboration with Asia Research Fellowship in 2009. Sanil successfully completed Kerala State Council for Science, Technology and Environment (KSCSTE), and UGC funded research projects in the year 2015 and 2016 respectively. He has more than a dozen published papers to his credit. His areas of interest include computational structures in Data Science, computational complexity, graph algorithm for complex network and Information Visualization. At present, he is working as the Assistant Professor in Computer Science under the Directorate of Collegiate Education, Government of Kerala.