

DATA VISUALIZATION BEST PRACTICES

Galina Manahilova¹

Received: 28.03.2023, Accepted: 25.04.2023

Abstract

Today nearly all organizations use data visualization to review their financial and business results, drive conclusions and make decisions. There are many benefits of data visualization. These include but are not limited to presenting many numbers in a small space, a synthesized view of the data, making comparisons easy, and the users thinking of the data rather than methodology, etc. Essentially, the overall graphics reveal the data.

This paper will review three best practices graphs: voting results, categories performance and comparison, and marketing performance. While looking into best data visualization practices, we will also examine dangers and tainted influence using graphs.

Keywords: data visualization; graphs; geographical map; spider web map; funnel graph

JEL Codes: B12; C61; Y10

1. Introduction

There is evidence that “seeing” is a good way of understanding and generating knowledge. Humans have a well-developed sense of sight, as more than 50% of human brain neurons are used in vision (Mahapattanakul & Puttatida, 2019).

According to Fauvelle (2020) “Data Visualization, or DataViz, is the discipline that focuses on the graphical representation of raw data. The purpose of data visualization is to create value”.

Human tendency makes visualization necessary. But, thru visualization, we turn data into insights.

The history of data visualization dates back to ancient times when maps were used to illustrate geographical information. Data visualization was further developed in the 17th century by cartographers who used statistical graphics to represent population data. In the 19th century, scientists and statisticians began using graphical representations to analyze data sets, and in the 20th century, computer-generated graphics became the norm. Today, data visualization is used in various industries, from business intelligence to healthcare and education. We are all accustomed to

¹ SWU “Neofit Rilski” - Blagoevgrad, Department of Finance, PhD Student; e-mail: manahilova@yahoo.com, ORCID ID: 0000-0002-2390-1050

different sights helping us daily in giving advice and warnings. That dictates the need to explain business results via visuals and tell the story with a picture.

Edward Tufte, an American statistician and Professor Emeritus at Yale University, is widely regarded as the "father of data visualization". He authorizes several books on the subject, including *The Visual Display of Quantitative Information*, *Envisioning Information*, and *Beautiful Evidence*. He is also the founder of the Center for Human Information Processing at Yale. Other notable data visualization experts include Stephen Few, Alberto Cairo, and Dona Wong.

The first famous data visualization example is the Mendeleev table. This table presents multiple types of complex information in a small space (Mendeleev Periodic Table Image).

Other great examples of data visualizations are:

- Gapminder's World Map - This data visualization uses a world map to show different countries' population and economic trends over time.

- Flight Patterns - This data visualization shows the real-time flight patterns of airplanes around the world.

- NBA Shot Chart - This data visualization showcases the shooting performance of NBA players in different areas of the court.

- The Covid-19 Dashboard - This data visualization shows the coronavirus pandemic in real-time, displaying the number of confirmed cases, deaths, and recoveries worldwide.

- Donut Chart - This data visualization displays the composition of a whole in the form of a donut chart, showing the contribution of each component to the total.

There are eight types of qualitative messages, according to Stephen Few (2004):

The first one is time series. There is a single variable in the time series, such as the death rate over several years. A line chart must be used to reveal the trend.

The second is ranking: This message incorporates categories ranked in ascending or descending order, such as ranking employee performance (the measure) by a person during a single period. Again, the best representation is via a bar chart showing a comparison among the employees.

The third one is part-to-whole: Some categories are revised as part of the whole (percentage out of 100%). Finally, the recommended graph is a pie chart, for example, employee representation by race.

The fourth is deviation: In this one, specific metrics are compared to a reference number and are usually used in finance, comparing actuals to plans. Again, a bar chart is recommended to show the variance visually.

The fifth is frequency distribution: Used when the number of observations of a particular variable for an assumed interval, such as the number of years in which the stock market return is between intervals such as 0-10%, 11-20%, etc. Again, the preferred graph is a histogram or boxplot, as it will help visualize vital distribution statistics, such as median, quartiles, outliers, etc.

The sixth is a correlation: In correlation, the dependencies between two variables and whether they move in the same or opposite direction are explored. The graph that best represents correlation is a scatter plot.

The event is a nominal comparison. It is used when absolute numbers organized in categories need to be compared. For example, the preferred graph is a bar chart.

The last one is geographic or geospatial. It is used when the comparison is needed based on a geographical basis. An example is GDP by country. A typical graphic used is a cartogram.

Digitalization changes are happening fast, and there needs to be more certainty for the organization in the implementation of new techniques. According to Zlateva (2020), digital transformation “provides the conditions for a more efficient process of creating economic value”. At the same time, the platforms provide an opportunity to present manipulative information and influence consumer behavior negatively.

One has to be mindful when choosing the best graph for presenting different types of information. Simultaneously the message they want to communicate must be clear. There are many great examples of doing just that, but also instances, where the type of graphs selected can share misleading messages.

2. Data Visualization Best Practices Examples

Data visualization is used in many fields, including business, science, government, and journalism. It can be used to explore patterns and trends in data, identify relationships between different variables, communicate findings to stakeholders, and help people make informed decisions based on data.

Many data visualization techniques and tools range from simple bar charts and line graphs to more complex visualizations such as heat maps, scatter plots, and network diagrams. The choice of visualization depends on the type of data being analyzed and the goals of the analysis. Therefore, multiple considerations and strategies must be followed when deciding which visualization to use. The goals are for the graphs to be easily readable, relevant, and depict the correct message.

Data visualization tools such as Tableau and PowerBi can make presentations look more excellent and professional, allowing slice and dice of the data as needed.

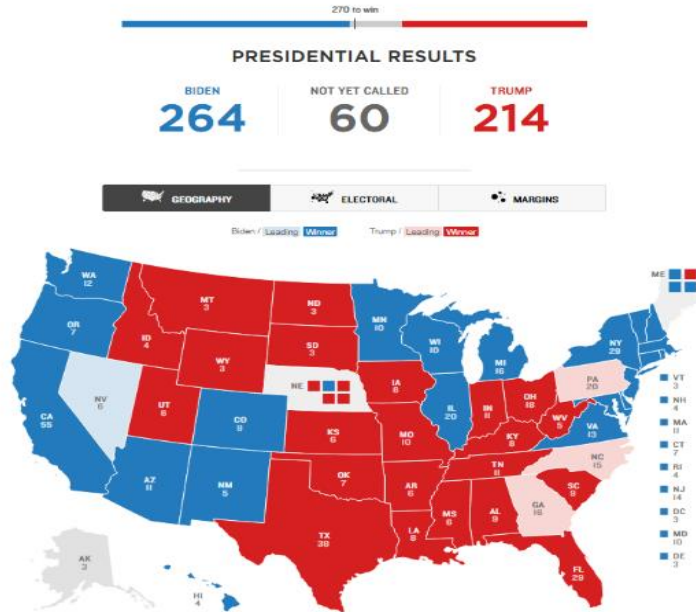
A. Presenting voting results

The first graph was used by the Republican party (red) to show the results of the 2020 Presidential election.

The second graph shows the actual votes for both parties correctly. This visual illustrates that blue is predominant, with few but larger dots (people voted), and red has many tiny dots.

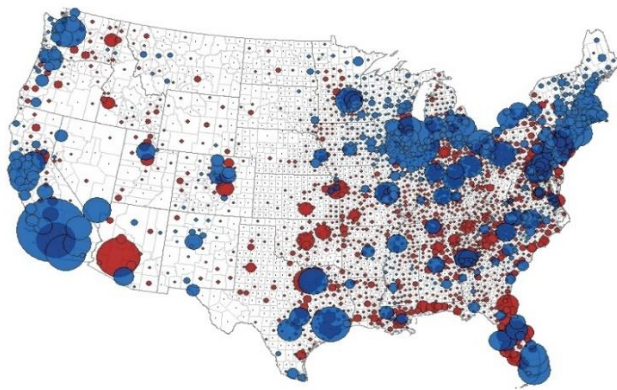
An example of the voting results can be seen in Figure 1 and Figure 2.

Figure 1. Visualizing Voting Results – Wrong



Source: [NPR](#), 03.06.2021

Figure 2. Visualizing Voting Results – Correct



Source: <https://www.core77.com/posts/90771/A-Great-Example-of-Better-Data-Visualization-This-Voting-Map-GIF>, 04.25.2023

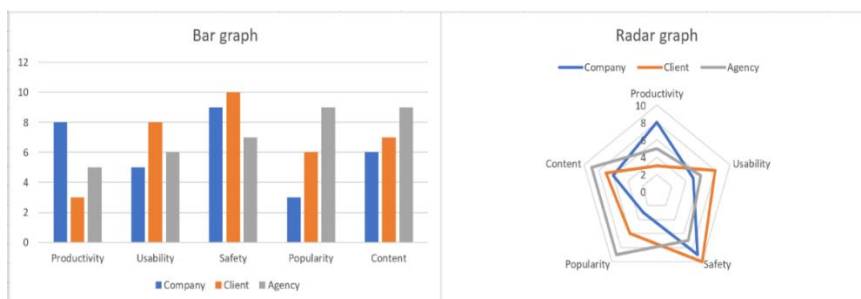
Choosing the first graph is inaccurate because this misrepresentation might lead to the improper conclusion that the Republican party won the majority vote. On the

other hand, the second graph is an accurate representation because the land doesn't vote; people do.

B. Category results and comparison

The recommendation is to use the Radar chart to compare three groups- user, consumer, and agency- regarding category importance. The Radar chart is more accessible for the eye to read and for the brain to comprehend. However, the caveat is that people are used to reading and learning bar charts and need to become more accustomed to the Radar chart. The logic for recommending the Radar graph is that when comparing categories, the typical brain will remember 3-4 comparisons in sets of 2's and lose sight afterward (Figure 3).

Figure 3. Showing ratings by Department by Category via bar graph and spider web graph



Source: Author

Similarly, the story from the graph will improve if the benchmark is added to the data. According to Madgerova and Atanasova (2013), benchmarking is used as one of the methods for “improving the competitiveness of the organizations”.

C. Marketing report

Marketing reporting involves collecting, analyzing, and presenting marketing activities and outcomes data. The purpose of marketing reporting is to provide insights into the effectiveness of marketing efforts, help identify improvement areas, and inform decision-making related to future marketing strategies.

The best practice in marketing reporting is to include the whole sales cycle from lead initiation until the order is completed. That decision is driven by the quality of the leads that can be attracted using various channels. Therefore, the quality of the information changes at all steps, from the initial impression and click to completing/canceling the order. A suitable way to represent the sales funnel results is shown in Figure 4.

Figure 4. Marketing funnel by Channel and all steps in the funnel

	Channel 1	Channel 2	Channel 3
Marketing expenses	\$ 100,000	\$ 120,000	\$ 145,000
Impressions	123,456,789	145,678,912	151,213,146
Conversion	0.02%	0.02%	0.01%
Clicks/Calls	21,000	22,123	13,178
Conversion	58%	59%	34%
Leads	12,098	13,000	4,500
Conversion	65%	35%	77%
Orders	7,890	4,567	3,456
Conversion	95%	92%	90%
Completed	7,496	4,202	3,110
Revenue per completed	\$ 132	\$ 156	\$ 210
Cost per completed	\$ 121	\$ 120	\$ 183

Source: Author

The numbers should be presented in funnel form since the visuals in the length of the boxes will help the user to understand the story based on visuals vs. doing math. Adding colors to compare the three channels will add an understanding of how one channel performs compared to another. The colors are not added to this visual because they can only be judged effectively by direct numbers with a valid attribution model.

3. Conclusion

Choosing the correct type of graph is essential for visualization because it can significantly affect how easily it is to interpret and understand the presented data. Different types of graphs are suited for different types of data and different types of analysis.

Choosing the wrong type of graph can lead to confusion, misinterpretation, and incorrect conclusions.

Here are some recommendations for creating compelling and informative data visualizations:

- ✓ Start with a clear goal in mind. Before making data visualization, ensure you understand what you hope to communicate with the visualization.
- ✓ Use the right chart type. Choosing the correct type of chart is essential for effectively sharing your data.
- ✓ Please keep it simple. Data visualizations should be easy to interpret and understand. Ensure to complete them with enough information.
- ✓ Focus on the data. Make sure the data is the star of the show. Choose colors and fonts that don't distract from the data.

- ✓ Use labels. Labels are essential for helping viewers interpret data visualizations. Make sure to include labels on all axes and any data points or trend lines.
- ✓ Make it interactive. Interactive data visualizations allow viewers to explore the data in more depth. Consider adding interactive features such as hover-over effects or tooltips.
- ✓ Consider animation. Animating data visualizations can help bring the data to life and make it more engaging.
- ✓ Make it shareable. Make sure your data visualization can be easily shared online. For example, consider exporting it as an image or interactive HTML file.

REFERENCES

- Fauvelle, L. (2020). Data visualization: definition, examples, tools, advice [guide 2020], Market research consulting, <https://www.intotheminds.com/blog/en/data-visualization/>
- Cambridge Dictionary, <https://dictionary.cambridge.org/us/dictionary/english/data-visualization>. Retrieved on 2021-06-03.
- Few, St. (2004). *Tapping the power of Visual Perception*. Perpetual Edge. http://www.perceptualedge.com/articles/ie/visual_perception.pdf
- Madgerova, R. & Atanasova, A. (2013). Definition, princips and precondition for implementation of the benchmarking to improve the enterprises' activity. *Entrepreneurship*, Vol. 1, Issue 1&2.
- Mahapattanakul, P. (2019). From Human Vision to Computer Vision- how far off are we? *Becoming Human: Artificial Intelligence Magazine*.
- NPR, <https://www.anychart.com/blog/wp-content/uploads/2020/11/npr-election-results-map-data-visualization.png>
- Tufte, E. R. (1990) *Envisioning Information*. Graphics Press.
- Core77, <https://www.core77.com/posts/90771/A-Great-Example-of-Better-Data-Visualization-This-Voting-Map-GIF>
- Zlateva, D. (2020). Digital transformation of Marketing Communications. *Economics and Management*, Vol. XVII, Issue 1, 171-181.