

Editorial: Speaking the language of diagrams, or How to correctly visualise data in a scientific paper?

Tatiana Baklashova¹, Aydar Kalimullin²

¹ Kazan (Volga Region) Federal University, Kazan, Russia

E-mail: ptatyana2011@mail.ru

ORCID: <https://orcid.org/0000-0002-3038-6990>

² Kazan (Volga Region) Federal University, Kazan, Russia;

Samarkand State University named after Sharof Rashidov, Samarkand, Uzbekistan

E-mail: kalimullin@yandex.ru

ORCID: <https://orcid.org/0000-0001-7788-7728>

DOI: 10.26907/esd.20.1.01

EDN: BKDXJW

Nowadays we live in a world full of data, that surrounds us everywhere. Data helps us to test hypotheses, recognise patterns, and make discoveries, whether it is primitive or complex; heterogeneous or homogeneous (Larose & Larose, 2014). Research and data are intertwined. They are interconnected with the most important stages of academic pursuit, i.e., collection and analysis of information. The data analysis is designed to recognise the patterns and achieve clarity in the phenomenon under study.

Data plays a specific role in any science. In education, for instance, it helps to enhance the quality of teaching (Bienkowski et al., 2012). Personalisation of education becomes more accessible. Qualitative and quantitative metrics of a student's academic progress and preferences allow teachers to choose the content, tempo, and methods of teaching, depending on the student's individual needs. Moreover, those metrics provide the means for the academic progress prognosis, and, when in university, academic attrition. The content, namely, the educational programme can be improved with the help of the gathered data. A collection of learning assessment materials can be updated to ensure the quality of education. The data is also significant in the management of an educational institution. The administration is able to track the finances, plan the budget, and effectively coordinate the whole system on the basis of the data. Likewise, it is interconnected with innovations in education, helping in the development of new online courses, e-textbooks, and apps focused on ensuring availability and learning efficiency.

Working with data implies its visual representation. There are many guidelines and papers on the graphical representation of the data (Maaten & Hinton, 2008; Glazer, 2011; Yau, 2024). The most significant aspects will be covered hereinafter.

A right visual representation of the data in research papers helps to better understand the complex information on the research conducted. One of the means of visual representation and structuring is graphs. What advantages do they have? Firstly, they are the means of simplification (and not overcomplication!) of the information. Graphs help to visualise the data in the form of easy-to-understand images. Secondly, the type of data is considered when constructing a graph as otherwise the reader might be puzzled by unreliable data. Thirdly, the graphs should not capture the reasons behind the data alteration. Given that it is the exact main goal, the author should instead choose another means of visual representation. And at last, it is easier to work with graphs than facts when setting out the results of a research, as the graphs present a bulk data set, facilitating the reader's comprehension of the rates' tendencies.

What role does the graphical representation play in the exemplification of the research data? There are a few points to highlight. First of all, a graph helps the researcher to prove a point. It is much more decisive when there is a graphical representation of the tabular results. That will help the reader to understand the author's results and conclusion and not get confused in the deciphering of the estimates. Then, a graph allows us to analyse a bulk data set, while only taking up a fraction of the paper length. There is a notion in the publishing industry that having as many details in a text of as few words and as small length as possible is an art. That art could and should be practiced with graphs. Moreover, the graphs help the author to simplify the complicated data, upscaling the right perception of the author's ideas by the reader. Finally, the graphs allow us to visualise the data comparison. If a researcher compares two or more sets of data in the paper, that are comprised of a great number of metrics and criteria, it is viable to use a bar chart for the reader to comprehend the difference between those data sets.

It is important to remember, that graphs are recommended against for presenting a small amount of data. If that is the case, it is better to summarise it in the text. Likewise, the data that is given in the graph should not be duplicated in a text – otherwise, the purpose of using the graph, outlined above, i.e., to present a bulk of data in a concise format, is invalidated. A graph should not be used if the data show no real tendency, that is supposed to enable the author to prove a hypothesis, to address a research question.

The graph is primarily constructed so that the horizontal axis (OX-axis) is the independent variable and the vertical axis (OY-axis) is the dependent variable. Therefore, the time periods should be placed on the OX-axis. The graph should contain at least a title, axis name, scale, data field description, and, if necessary, a character representation. Graphic elements should be clear and legible, with a contrast to the background and the foreground design. Each caption should concisely share as much information about what the graph visualises as possible. It should be noted that the caption should not contain a concise interpretation of the results or experiment details (Slutsky, 2014).

It is vital to choose the right form of visualisation in accordance with the data type. Pie chart and bar chart describe an immediate situation and its result. A line chart is used to demonstrate the metrics' dynamics and it should invariably contain a time axis. If the author conducts a correlation analysis, it is viable to use a scatterogram to visualise the connection between the two metrics. A histogram is usually used to substantiate the metrics' tendencies.

One of the mistakes of the graph's authors is its overcomplication. Sometimes the authors try to merge a lot of different data into a single graph. That hinders the information reception and the reader might not understand the idea behind the visual representation. The same complications might happen if a lot of words are used in the visual representation. Hence, it should be minimised while ensuring that the visual message is intact.

The characters used must be clear so as not to facilitate confusion for the readers. The footnotes might be used for extra information on the graphs, whilst eliminating all unnecessary data from it. When thinking about the graph's layout one should ensure its legibility, e.g., the type of font, its size, symbols used. The colours and shades used, graph's, and axis' names should also be paid attention to (Slutsky, 2014).

Compliance with the guidelines on the visual data representation facilitates clearer communication between the author and the readers. Compacting bulk data into a graph makes the research paper structured and easier to comprehend. This makes the researcher's point much more well-reasoned and convincing.

References

- Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing Teaching and Learning through Educational Data Mining and Learning Analytics: An Issue Brief. *Office of Educational Technology, US Department of Education*.
- Glazer, N. (2011). Challenges with graph interpretation: A review of the literature. *Studies in science education*, 47(2), 183–210.
- Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining*. John Wiley & Sons.
- Maaten, L. V. D., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(Nov), 2579–2605.
- Slutsky, D. J. (2014). The effective use of graphs. *Journal of wrist surgery*, 3(02), 067–068.
- Yau, N. (2024). *Visualize this: the FlowingData guide to design, visualization, and statistics*. John Wiley & Sons.