

INTRODUCTION TO INFORMATION VISUALISATION TECHNIQUES FOR HUMANISTIC SCHOLARSHIP: *A SERIES*

CHAPTER 1:

Why visualize?

CHAPTER 2:

Getting to data

CHAPTER 3:

Choosing and creating a visualisation

CHAPTER 4:

Do's and don'ts of designing data visualisations



This introduction was developed as part of a series on information visualisation for the humanities developed for the KU Leuven Faculty of Arts. The series was authored by Houda Lamqaddam, under the supervision of Prof. dr. Margherita Fantoli.

Find more on the work done in DH by the KU Leuven Faculty of Arts on the following page:

<https://www.arts.kuleuven.be/digitalhumanities/english>.

Chapter 4

Do's and don'ts of designing data visualisations

Now that you know how to go from material to data, how to choose your chart types and how to create your very own visualizations, here are a few practical do's and don'ts to keep in mind. Remember, these are rules of thumb, and you may have good reasons to deviate from them. In any case, make sure to explore and have fun.

Do!

1. Limit the number of categories:

Research indicates that humans can hold about 7 (± 2) objects in our short term memory. This means that a chart that uses a large number of categories will most likely be a cognitive struggle for your readers and yourself. This is even more true for some chart types such as pie charts, which easily become cluttered if there is a large number of thin slices competing for attention.

👉 *Quick tip:* try not to go above 5 color or shape categories when possible.

🔗 ["The magical number seven, plus or minus two: Some limits on our capacity for processing information"](#)

2. Choose the right color scheme:

In visualization, colors matter a great deal. This is because color is one of the most effective ways to draw attention. So make sure you choose colors that readers will be able to distinguish, and that support your data well.

For instance, use different hues for data that is split in different categories; a continuous

saturation spectrum for sequential values; a spectrum between contrasting hues for a scale that has negative and positive values (diverging). It also helps to use colors that already have a meaningful association for users. For instance, blue for cold and red for hot is a good example of such 'affective colors'.

👉 *Quick tip:* use a resource like [Color Brewer](#) to pick an appropriate color scheme. Enter the number of your categories, and the nature of your data and get recommendations of color palettes you can copy into your data visualization tool.

🔗 Silva, S., Santos, B.S. and Madeira, J., 2011. [Using color in visualization: a survey](#). *Computers & Graphics*, 35(2), pp.320-333.

🔗 Bartram, L., Patra, A. and Stone, M., 2017, May. [Affective color in visualization](#). In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 1364-1374).



Source: <https://blog.datawrapper.de/which-color-scale-to-use-in-data-vis/>

3. Provide sources for the data:

It is very important to provide sources for the material you are presenting. Not doing so can undermine the credibility of your findings. In addition to the source of the material, you can also include any transformation you made or filters you applied to allow readers to critically engage with the final result.

👉 *Quick tip:* Add a link to the source of the material, or related publication.

🔗 Lamqaddam, H., Moere, A.V., Abeelee, V.V., Brosens, K. and Verbert, K., 2020. [Introducing layers of meaning \(LoM\): A framework to reduce semantic distance of visualization in humanistic research](#). *IEEE Transactions on Visualization and Computer Graphics*, 27(2), pp.1084-1094.

4. Write clear titles, labels and legends:

When we spend time designing a visualization, it can be easy to forget just how familiar with it we have become, and to feel that it is crystal clear and understandable at first sight. Keep in mind that readers may be processing your research context, data, and visual representation all at the same time. To make sure the visual encoding is clear, provide labels and legends to the elements that make up your visualization. For example, if you are making a bar chart, make sure to describe what each axis means, and what the color or size of any additional element refers to.

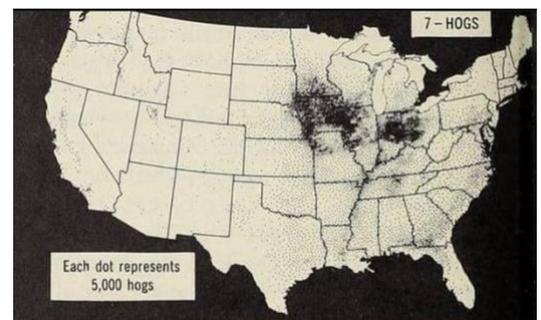
👉 *Quick tip:* Share your chart with a friend and ask them what they did or did not understand.

Don't!

1. Use inconsistent colors or shapes:

Each visual encoding (color, size, shape..) should map to one and only one meaning. Using inconsistent colors or shapes in a visualization can be misleading. See the image on the side for an example: The illustrators chose to show each hog with a black dot, and then used the same black color to draw the borders of states and of the country. So who is to say the US is not surrounded by a sea of hogs?

🔗 Qu, Z. and Hullman, J., 2017. [Keeping multiple views consistent: Constraints, validations, and exceptions in visualization authoring](#). *IEEE transactions on visualization and computer graphics*, 24(1), pp.468-477.



nemfrog

"Each dot represents 5,000 hogs." World Geography, 1948.

lemondemon

untapped infinite hog supply in the ocean

sketchmagetch

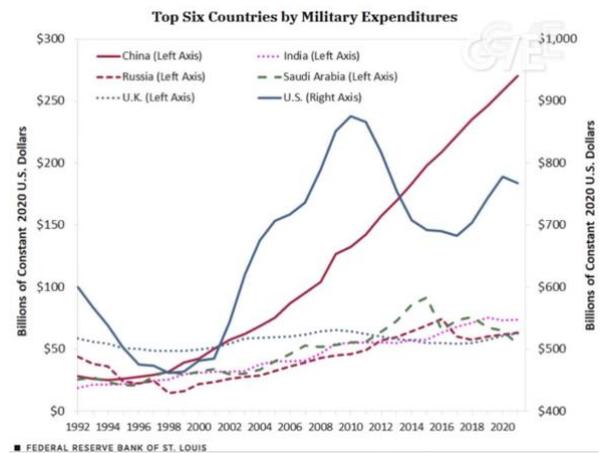
Each state is lined with an impenetrable wall of swine. We are trapped

2kittensinacup

We've lost canada and mexico to the hogs already

2. Use misleading scales:

One common error is using scales that are not proportional to the data. For example, using a truncated axis to exaggerate differences between data points or using a non-zero baseline to distort the appearance of changes in the data. Look at the chart on the right. Make sure you pause to read the legends on the top left. Who spends more on military expenditures? China or the USA? Such misleading scales can be chosen on purpose to create a false illusion.



Source: <https://twitter.com/MoriaMarmalade/status/1618200343623655424>

However, it can sometimes happen by mistake when you're designing visualisations of your own. Make sure to step back and look at the entire picture to make sure you are still representing data accurately.

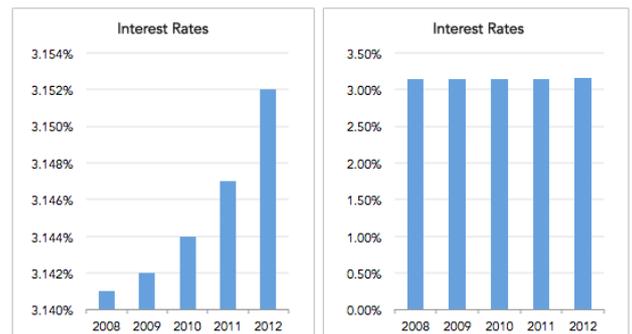


PolicyViz: [10 Ways to Mislead with Data Visualization](#)

3. Truncate your axes:

Similar to the previous point, a common error is to truncate the y axis so that the baseline is a chosen value instead of '0'. This can be unintuitive as sometimes, there is not much happening between 0 and 80% for examples, and most of the difference can be occurring in those last 20%. However, truncating the axis can make the difference between values exaggerated. Look at the example on the right. Starting the y axis at 3.140 makes the difference in years look drastic, while the same graph with a corrected baseline paints a much more subtle variation.

Same Data, Different Y-Axis



source: <https://www.datapine.com/blog/misleading-data-visualization-examples/>



Library Research Service: [Visualizing Data: a misleading y-axis](#)

4. Overload the visualization:

Sometimes the best story is a simple story. Resist the temptation to encode too much information in one single chart. This can leave you and other readers overwhelmed and discouraged. Make sure your visualization is clear before adding any additional encoding, and always check if it can still be read by your main audience. Similarly, if the visualization is cluttered because there are too many overlapping datapoints, consider using interaction to filter it, or proposing multiple charts each focusing on a single aspect of the data. Finally, while

embellishments and decoration can enhance the aesthetic value of your visualization, make sure it does not occlude any element, or distract away from your data story.



Tableau: [Data Visualization Tips For More Effective And Engaging Design](#)