

**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 1

Why visualise

In this series, we present an introduction to information visualization as a tool for humanistic research. We introduce the concepts, techniques, and tools of infovis, and provide extra readings and references for readers who want a little bit extra.

We start with a quick intro of what infovis is and how it can help.

Information visualization, or infovis, is the process of taking a large amount of data – or research material - and representing it visually to make it easier to understand. It includes a variety of tools you have probably encountered before such as Excel charts, web infographics, 3D maps, interactive dashboards. In essence, there are many possibilities, and part of the challenge (or the fun?) is deciding which tool will be most helpful for your particular research subject.

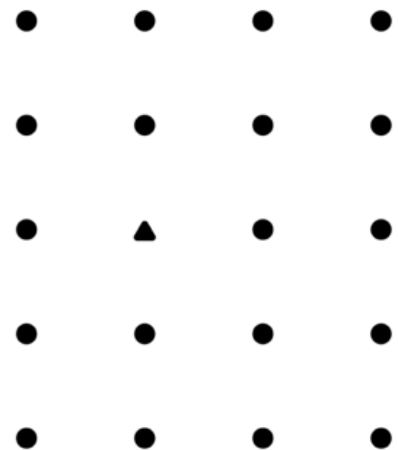
One of the main strengths of visualization is presenting readers with a global representation of the data. This global representation helps give an overview, and is often referred to as the first step in understanding or analysing material. Having an overview gives the opportunity to our brains to do some tasks they are very good at. Namely: pre-attentive processing, and pattern recognition.

What is pre-attentive processing?

Pre-attentive processing roughly means the type of analysis our brain makes before we start consciously processing any information. It is what allows information to ‘jump out’ to us without requiring us to deploy conscious cognitive effort. In the figure on the right, the triangle is particularly salient. It attracts our attention extremely fast because the visual contrast with the circles is identified by our visual pre-attentive processing. Now why does this matter? Because when the task is to identify a triangle among a sea of circles, this can seem trivial. However, this is the principle we want to make use of when creating data visualizations because it turns a repetitive tedious task (going through thousands of files to find the one outlier) into a near-immediate visual summary.



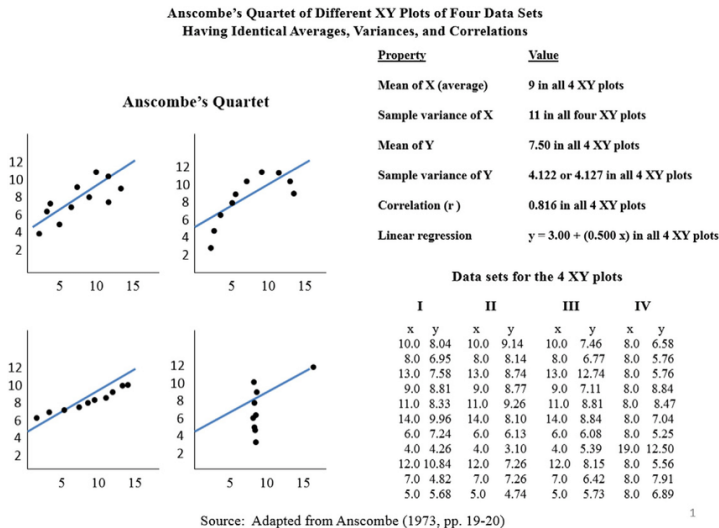
Figure 1 Information visualization works in the same way that if you open your fridge, cupboard or closet. Before trying to do anything, it helps to see what you're working with.



Infovis in practice: Anscombe's quartet

Statistician Francis Anscombe produced the 'Anscombe quartet' in 1973 to show that summary statistics alone do not tell the whole story, and that information visualization can be of crucial help.

These are four datasets consisting of eleven (x,y) coordinates, which, while consisting of different values, share the same statistical characteristics. Indeed, they have the same mean, variance, correlation and



linear regression. If we plot them, however, very different patterns emerge and gives us a much better view of each datasets. We learn whether they have linear relationships or not, or if there is an outlier that skews the entire statistical summary. These relationships that we hope to figure out from data analysis, can, in this case, be seen immediately through visualization. They also complete the statistical summary which does not give us the full picture.

Infovis for exploration and Infovis for communication

Visualisations can be useful for you as a researcher looking into a collection, for example, and trying to learn new things about it. Or it can be used as support for communicating a result to a wider public, at a conference for instance. We therefore say that the two main goals of visualization are exploration and communication. These two goals have important implications on the design and features required in the visualization. Tools that aim to support exploration need to support interactive, comprehensive and accurate representations. Whereas for communication, we mostly want the representation to be clear, simplified (or perhaps summarized), and accurate.

vis for exploration	vis for communication
comprehensive	can be selective or summarised
needs to be legible to you	needs to be accessible for the audience
usually interactive	usually static
must be accurate	must be accurate

Examples of DH findings enabled by vis

This all sounds good in theory, but you may still be wondering: would data visualization really help in my project? how does this fit into a humanistic research question? We include references to relevant articles from the DH literature that use visualization as a method to draw conclusions, elicit new research questions or devise novel hypothesis.

- 🔗 Derungs, C., Sieber, C., Glaser, E. and Weibel, R., 2020. Dialect borders—political regions are better predictors than economy or religion. *Digital Scholarship in the Humanities*, 35(2), pp.276-295.
- 🔗 Brown, D.M., Soto-Corominas, A. and Suárez, J.L., 2017. The preliminaries project: Geography, networks, and publication in the Spanish Golden Age. *Digital Scholarship in the Humanities*, 32(4), pp.709-732.
- 🔗 Crissaff, L., Ruby, L.W., Deutch, S., DuBois, R.L., Fekete, J.D., Freire, J. and Silva, C., 2017. ARIES: enabling visual exploration and organization of art image collections. *IEEE computer graphics and applications*, 38(1), pp.91-108.