

# Information Visualization for the Construction Industry



Michele Mauri 

**Abstract** The chapter introduces the topic by contextualizing, historically and conceptually, the use of visual properties to convey data and information for the built environment. It presents how design affects the visualization process, highlighting the authorial choices that influence how readers perceives information, showing why visualizations are not the result of a mechanical process of data encoding, but rather the outcome of an authorial process of representation. This is done by presenting the elements composing a visualization: the visual variables, the layout elements, and contextual information. Visual variables can be seen as “construction blocks” used to encode data dimensions, and strengths and weaknesses of their usage are exposed. The chapter then discusses visual models as “blueprints” featuring specific mappings of data onto visual variables, presenting a selection of useful ones and highlighting their context of use as well as issues in their usage. Finally, it is highlighted the relevance of contextual elements (such as titles, legends, and annotations) in the process of encoding/decoding of information and their role in providing readers the implicit knowledge needed to understand the subject and to get insights from the visualizations.

**Keywords** Information visualization • Data visualization • Visual variables • Visual models • Visualization layout • Legends • Annotations

## 1 Introduction

The visual communication of data has recently gained a huge popularity. Many software and tools are today available for the creation of data visualizations, and the risk is that given a dataset, the user feels overwhelmed with all these possibilities, without having a clear understanding on how to choose a visual model rather than another one.

---

M. Mauri (✉)

Politecnico di Milano, Via Durando 10, 20158 Milano, Italy  
e-mail: [michele.mauri@polimi.it](mailto:michele.mauri@polimi.it)

The risk is that the visualization process becomes driven by the software: The only envisioned solutions are the ones given by it, instead of starting from the communicative goal.

The visual communication of data is not something new: It is from centuries ago, the first known representation on a coordinate system visible in Fig. 1.

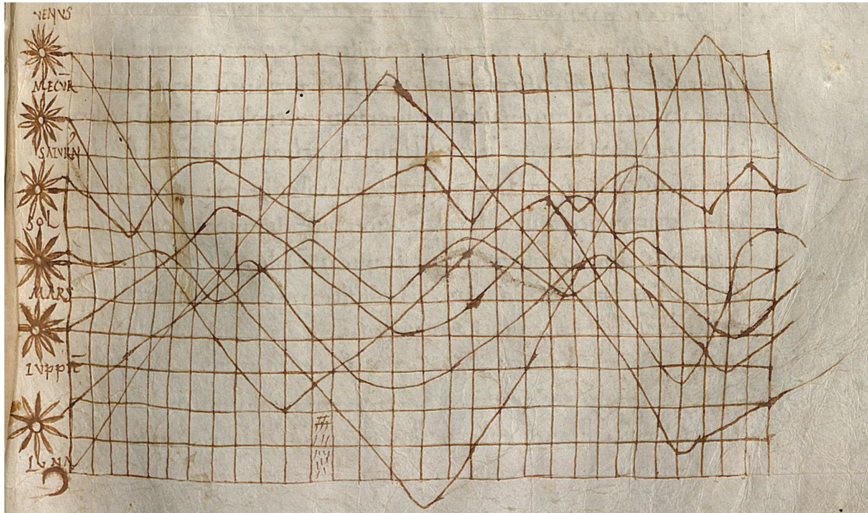
A huge corpus of literature has been developed in the last two centuries focusing on this topic, exploring how visualizations can support the analysis process, how they can help domain expert to develop strategies, and finally how they can help in communicating with the larger public.

Today visualizations are used on a vast array of supports: from static and printed medium, such as newspaper, books, and posters, to interactive and animated supports, such as dashboards, explorative Web sites, and applications.

Visualizations are often used for their appeal, their ability to communicate idea of complexity, precision, and knowledge on the represented topic.

When dealing with Industry 4.0 in the built environment, a great amount of data are created. Data and information visualization principles and techniques become essential to communicate with colleagues, partners, and clients and enable effective decision-making processes.

In this chapter, it will be provided an overview on visualization strategies, good practices, and approaches. However, before starting to deal with the practice, it is important to provide an understanding on the underlying reasons for which visualizations are used, or, in other words, why data visualization should be used instead of representing data in tables or texts.



**Fig. 1** Earliest-known 2-dimensional visualization dated to late tenth century. *Source* Bayerische Staatsbibliothek München, Clm 14436, fol. 61r

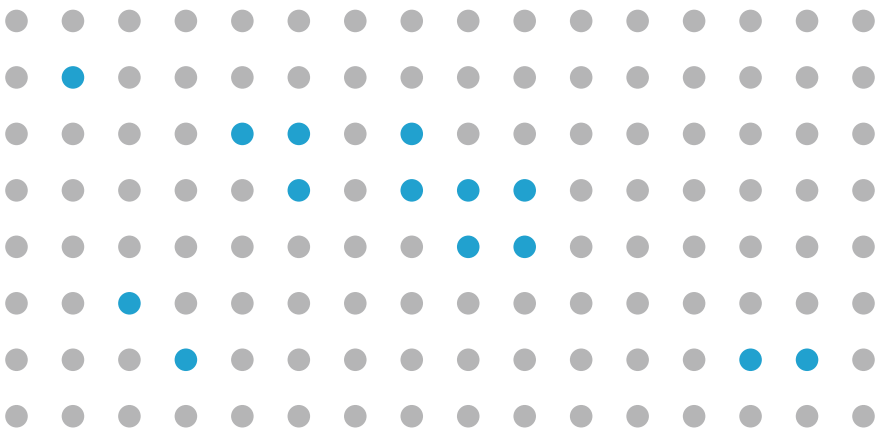
### 1.1 Perception and Data Visualization

The posed question could seem trivial, but it is fundamental. To provide an example, in 1913, the employees of the city of New York organized a statistical exhibit featuring large-scale visualization to educate the population on the results achieved by the different departments of the city administration. A contemporary observer wrote: “The Health Department, in particular, made excellent use of graphic methods, showing in most convincing manner how the death rate is being reduced by modern methods of sanitation and nursing [...] There can be no doubt that many of the thousands who saw the parade came away with the feeling that much is being accomplished to improve the conditions of municipal management” [1].

This example is useful to underline the words used: *showing*, *saw*, and *feeling*. Basically, visualization is powerful since it leverages on our natural abilities: visually identify patterns and make sense of them. Humans in their everyday life are used to visually compare the objects around us, identifying pale and bright colors and evaluating the size and the distance of the different objects.

Alberto Cairo, in his book “The Functional Art,” explains that “what we commonly call seeing is not a single phenomenon, but a group of at least three operations: sight, perception, and cognition” [2]. This means that not all the visual stimuli gathered by the eyes are treated in the same way by our brain, and that not all the information processed by the brain reaches our consciousness.

These operations made without the need of consciousness are called “pre-attentive.” As noted by Spence [3], “these low-level operations are performed automatically without conscious awareness; these processes are sometimes referred to as ‘pre-attentive’ because they occur without conscious intervention and control.” A common example is the image represented in Fig. 2: Our attention focuses on the blue dots effortlessly, isolating them from all the other gray dots.



**Fig. 2** Example of pre-attentive processing: Most of the people will immediately focus on the blue dots

Psychologists from the Gestalt school observed that humans naturally perceive objects as organized structures, in particular, as aggregation of items. They identified several grouping principles that can come at hand when dealing with visualization [4].

- **Proximity:** closer visual items are perceived as part of the same group.
- **Similarity:** items with similar visual properties (e.g., color or shape) are perceived as part of the same group.
- **Connectedness:** objects that are visually joined are perceived as connected (e.g., in a node-link network diagram).
- **Continuity:** when lines or areas intersect, people tend to identify as separate visual items the ones that follow an established direction.
- **Symmetry:** symmetrical lines are perceived as forming a whole.
- **Closure:** close lines are perceived as objects, and people have a perceptual tendency to see complete figures even if the contour is incomplete or partially hidden.
- **Relative Size:** smaller visual items tend to be perceived as objects and bigger as background.

Given these pre-attentive automatisms, the emerging question is how to use efficiently these abilities to convey data to our readers.

This question hides two other ones, which are: more efficient compared to what, and what do we mean with “efficiency.”

Often, it is a synonym of effortlessly or quickly, but reading information in the shortest amount of time is only one of the possible goals. It is possible to identify at least four ways, in which a visualization could be considered efficient:

- **Amplify Cognition:** the visualization improves the accuracy and completeness in understanding the information [5].
- **Gain Insights:** generates new ideas that are seeds for further analysis and research [6].
- **Increase Memorability:** helps the reader to remember the insights on the topic over time [7, 8].
- **Engage the Reader:** the reader does not gain only information, but also a pleasing experience engaging with it, with the final goal of fascinating the reader with the subject [9].

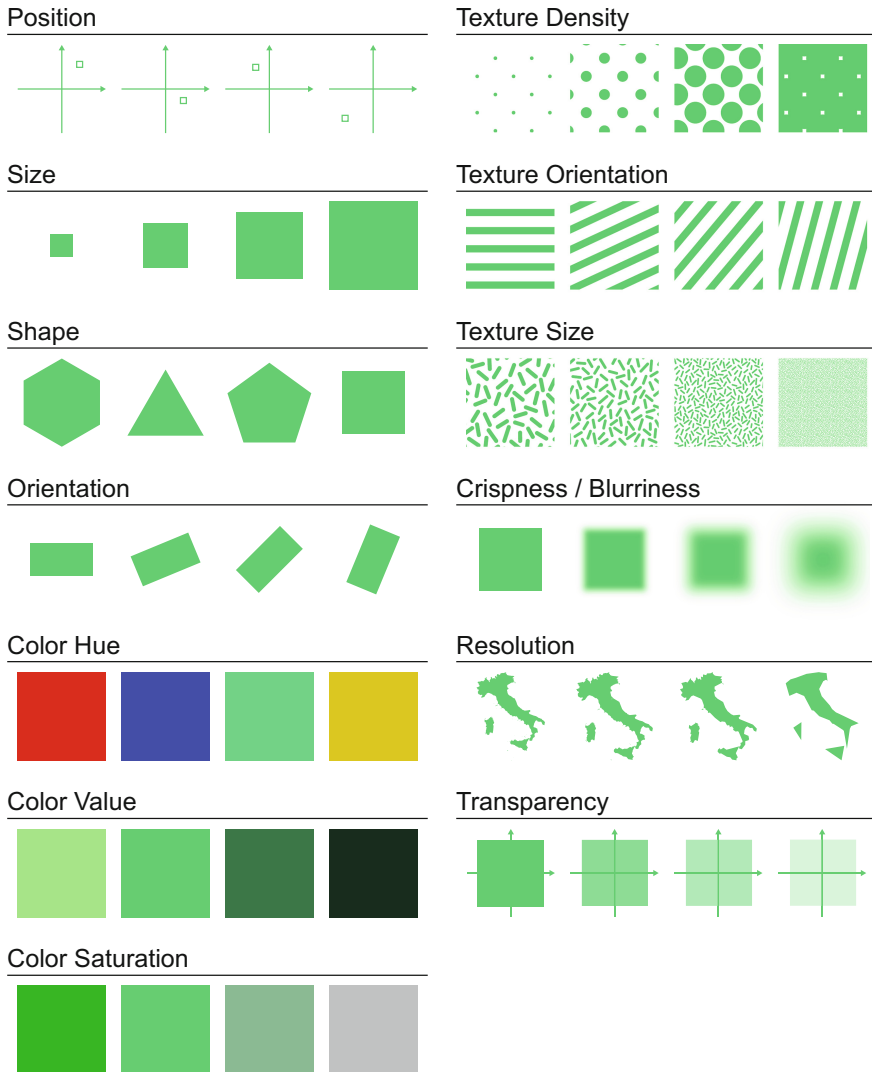
## 1.2 Visual Variables

After having defined the communicative goal, it becomes important to define how to leverage on these perceptive automatisms to visually encode the data. A common approach is the one originally developed by the French cartographer and theorist Jacques Bertin. In his book “*Semiologie Graphique*” [10], the author provides an

in-depth analysis of the practices in cartography, identifying the basic elements that can be used to encode data and comparing their effectiveness in enabling visual operations. While its reflection was originally meant to be applied in cartographic representations, it is, however, fundamental also in approaching the visual communication of data in every field, including the built environment.

The author starts from the very basic elements used to build a visual representation of data: shapes, or, as he calls them, visual *marks*. Marks are graphic primitives and can be points, lines, or areas. Each mark can have multiple visual features that define its appearance, such as position, color, shape, or rotation that can be used to encode or “map” data values. In this chapter, these features will be called as “visual variables.” The visual variables originally identified by Bertin has been expanded and systematized over time [10–14], and the most relevant according to literature are the ones depicted in Fig. 3.

- **Position.** A mark can be placed in a two-dimensional space. The horizontal and vertical positions can be used to encode information, as it happens in a scatterplot. Several studies [14] identified position as the most readable visual variable and therefore the most relevant in the encoding.
- **Size.** The amount of space occupied by the mark. This is one of the most versatile dimensions that can be used to enable different kind of operations.
- **Shape.** A mark can have different shapes, for example, circles, squares, triangles, or more complex ones. Shapes can be efficiently used to identify groups of marks.
- **Orientation.** The rotation of the mark can be used to encode data even if is rarely used. Remarkable examples are wind maps or vector fields. In wind maps, for example, is possible to follow the direction of the wind in different areas thanks to marks orientation.
- **Color.** Color is a complex variable since it can be described with several mathematical models. For the sake of simplicity, it can be sub-divided into three variables:
  - **Hue.** The perceived “pure” color (e.g., red, green, yellow, and brown).
  - **Color Value.** The perceived tone of the color (as dark or light). An extremely simplified description is the amount of black or white added to the “pure” color.
  - **Saturation.** The perceived brightness of the color. An extremely simplified description is the amount of gray added to the pure color (from gray to pastel colors and to the bold ones).
- **Texture.** A pattern that can be used to fill the mark. It provides some sub-variables as well:
  - **Orientation.** The orientation variable defines the rotation of the texture.
  - **Size.** The same pattern can be scaled according to a data variable.
  - **Density.** The percentage of colored space in the pattern.



**Fig. 3** Four examples for each visual variable

- **Crispness/Blurriness.** The sharpness or the fuzziness of the mark contours. It is often used to encode uncertainty of data points.
- **Resolution.** The spatial precision in the display of the mark or, in other words, the level of details.
- **Transparency.** The visual blending of the mark with the underlying colors.

The visual variables can allow the reader to perform different operations, namely: the reading of groups, the reading of orders, and the reading of quantities [10, 12].

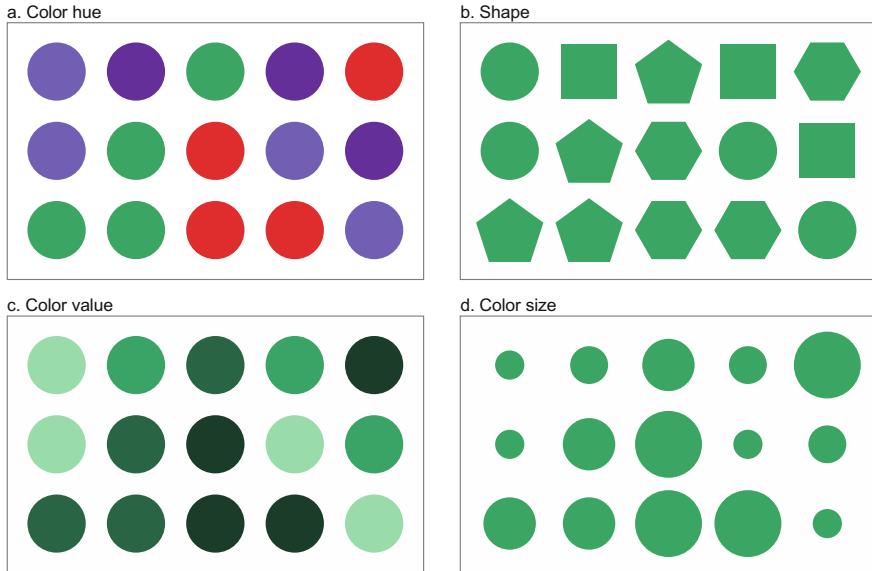
- **Grouping.** Visual variables are powerful in helping the user in the reading of groups by associating all the marks featuring the same visual variable. On the other hand, it is possible to isolate a mark from a group thanks to a different visual variable.
- **Orders.** Visual variables can be used to enable the reading of an order of the marks even if without quantifying them. An example could be the usage of color value to allow the reading from the palest to the brightest value.
- **Quantities.** Finally, visual variables can be used to allow the reading of exact quantities (e.g., being able to say that a value is two times bigger than another one). This is the most complex operation, and few visual variables are suitable for it: In the literature, only position and size are recognized as suitable [14].

Visual variables do not have neither the same efficiency nor flexibility in conveying information. Some of them are very powerful on one kind of operation, but not the others, few of them enable all of them. To make an example, space is one of the most powerful and versatile visual variables: It can be used for reading groups, showing orders (as in a ranked list) or quantities (as in a scatterplot).

In Fig. 4, it shows the same dataset mapped with different visual variables. Even without knowing the context of use of this data, or what it represents, different information can be perceived. In the first one, using color hue, the reader will likely see groups of marks, but not orders neither quantities, since there is not a shared known order of color hues. Similar is the case for shapes even if less powerful than color hue. On the other hand, using transparency or color value, it is possible to communicate an order from the darkest to the palest color even if it is difficult to evaluate the exact value. The last example encodes the information by using size: This variable is again very powerful in communicating orders, since it is possible to read marks from the smallest to the biggest, but it can also convey the reading of quantities.

Three reflections are relevant, following the approach based on the encoding of data into visual variables. First, not all the variables are equal. Some of them are perceptually stronger. This means that if several variables will be encoded in the visualization, it must be defined an order of relevance. Even if the communicative goal is to tell that three or four variables are equally important, some of them will be mapped using a less powerful visual variable and therefore will be perceived at a second level. For example, if the goal is to show how many objects are in each discipline in different building typologies, then a choice has to be made: if to use more powerful aspects to represent the different building typologies, the discipline, or the object type.

Second, if particular data dimension should be underlined, it is possible to use a redundant mapping, for example, using color and size to represent the same set of values: In this way, its relevance is stressed and reinforce its reading. To give an



**Fig. 4** Same dataset mapped with four visual variables: color hue, shape, color value, and size. Adapted from the work by Roth [14]

example, it is possible to visualize the spaces in a schematic way using the same colors and size for representing their square meters. In this way, it would be easier to detect the same spaces and the variable in dimension.

Finally, it is possible to encode dozens of variables, but this does not mean that our reader will be able to decode them.

A final consideration should be made on the relevance of the support in the creation of data visualization. The device on which the visualization will be shown (can be a book, a poster, a computer/tablet/mobile monitor, or a projector) will heavily influence its reading. To make an example, when Bertin introduced his work, he stated that it was valid under specific conditions [10]:

- representable or printable;
- on a sheet of white paper;
- of a standard size, visible “at a glance”;
- at a distance of vision corresponding to the reading of a book or an atlas;
- under normal and constant lighting.

These conditions could seem trivial, but they are not: Today visualization is often seen on computers that have completely different properties than paper. On printed artifacts, such as book and posters, the physical bounds are quite evident. It is, therefore, straightforward also the evaluation of the amount of information that is possible to encode, as well as the minimum level of details that can be achieved,



like the smallest size for a mark or a text. When moving to digital supports, the risk is to imagine that one solution will fit any use—from a computer monitor, to a tablet one, to a mobile one, or to a projector. It is, therefore, relevant, before crating any visualization, to define the device(s) where the visualization will be presented as this aspect can influence the creation of the visualization itself.

## 2 Visual Models

Visual models can be seen as blueprints featuring specific mappings of data onto visual variables (see Sect. 1.2). In this chapter, the term “visual model” will be used to define such reusable solutions for representing data.

They could be called “charts” even if this term does not cover all the possible representation of data, like maps or graphs. Since there could be infinite variations and hybridizations of such models, in this chapter, the most common solutions will be highlighted, explaining for what they are useful.

In the publication “A Tour through the Visualization Zoo” [16], the authors identify five main families of visualizations based on the underlying data structure:

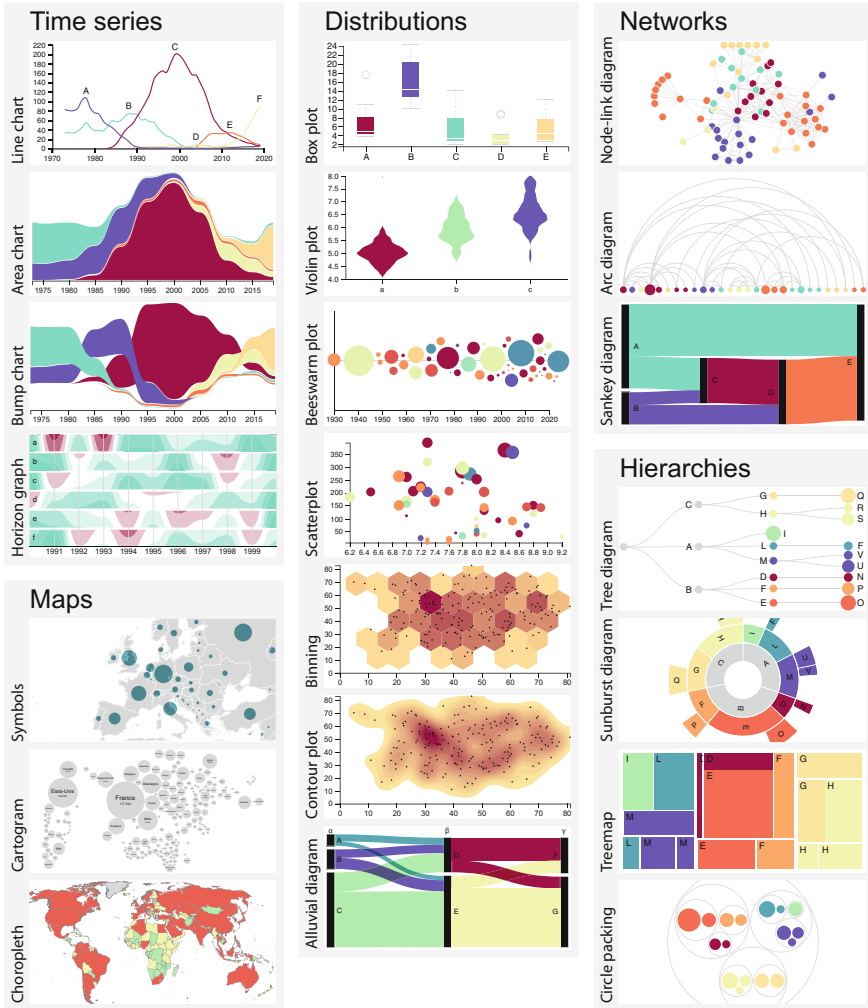
- Time-series data
- Statistical distribution
- Maps
- Hierarchies
- Networks

These categories are only indicative, and many visual models could belong to more than one category, and furthermore, some models could not fall in any of those categories. They are, however, useful to identify the main areas, in which the visual representation of data can ease the reading of information (Fig. 5).

When dealing with visualizations for the built environment, the first idea that come into mind is a geographical or spatial representation. However, also the other presented visualization “families” can be useful in this field. Below, some examples will be provided to illustrate how those approaches can communicate different kind of data structures.

### 2.1 *Time Series*

When the main point of the analysis is on the temporal evolution of a phenomenon, it is a good idea to choose a time-based visualization. For example, if a company would like to compare the number of workers in a construction site in the last three months as in Fig. 6. Common solutions are the line charts and area graphs and since they are able to convey the temporal continuity on the horizontal axis. A less



**Fig. 5** Selection of visual models for representing data grouped by underlying data structure. Time series, distributions, and hierarchies thumbnails have been adapted from RAWGraphs [15]

common model, the so-called bump chart, is useful for visualizing the temporal evolution of a ranking. Another less known option is the horizon chart [17], which allows to minimize the vertical space occupied by the chart. When dealing with time-based charts, it is important to pay attention to the temporal grain: A common problem is when in the dataset, there are daily values but no values for few days. One way is to count them as zero, and another one is to interpolate the values. Depending on the visualization context, the choice may vary.

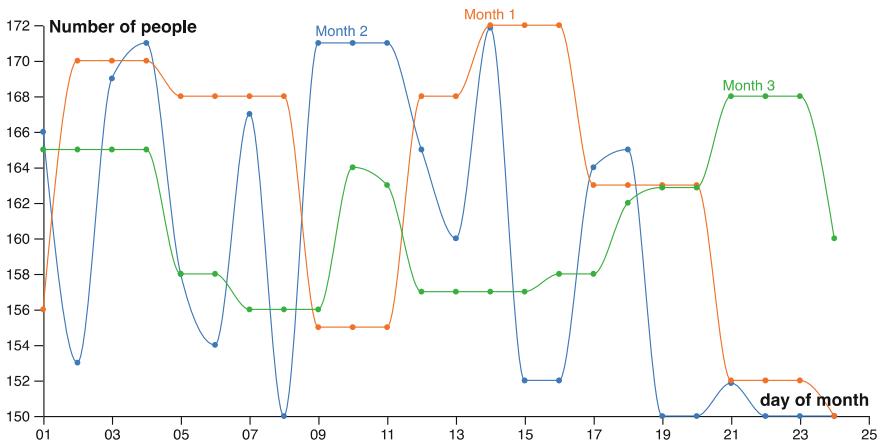


Fig. 6 Line chart comparing the number of workers in a construction site in three different months

## 2.2 Maps

When dealing with geo-located data, it is possible to map it over a cartographic representation. For example, if an asset owner would like to visualize where their assets are located across countries as in Fig. 7.

This kind of representation has its own complexities related to the projection of the globe on a two-dimensional space and would fall out the scope of the chapter the in-depth analysis of such problem. It is, however, useful to highlight three possible ways to display values over a geographic map.

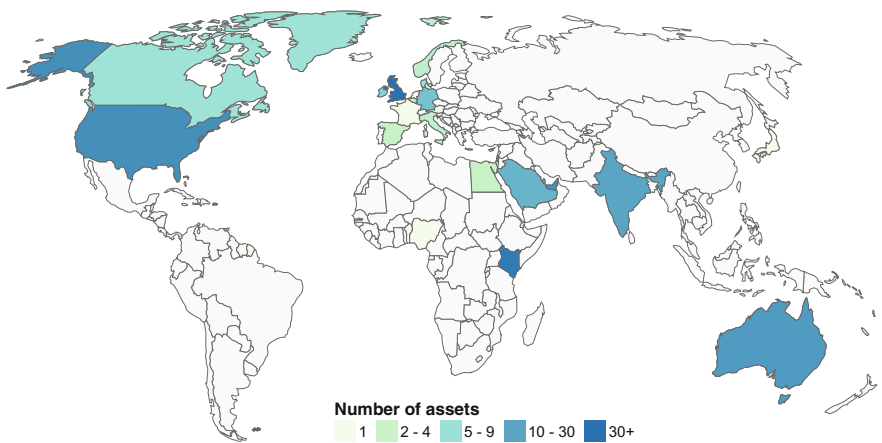


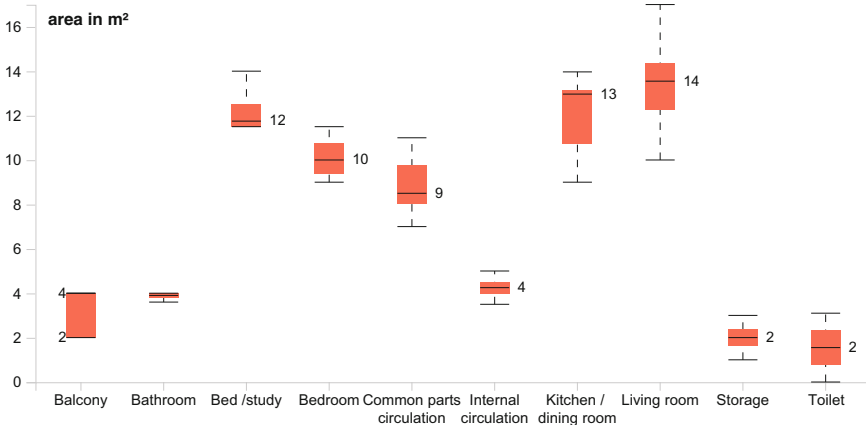
Fig. 7 Choropleth map showing the amount of owned assets across countries

The most straightforward approach is to overlying symbols (e.g., circles) on a map. The risk in this approach is that since symbols have their own area, it could be difficult to relate them to geographical space. A second solution is the usage of choropleths, which means the encoding of color on geographical areas. While this solution avoids the problem related to the previous approach, it has the disadvantage that the visibility of data is related to the size of geographical areas, and smaller territories could become invisible. A third solution is cartograms, in which geographical areas are distorted according to the mapped value.

### 2.3 Statistical Distributions

If the goal is to show the distribution of values on one dimension, a common solution is the boxplot that allows to show the quartiles (the four breaking points in the data series, excluded the outliers). For example, to visualize how the space per type of room varies across a group of building typologies, as in Fig. 8. With the graph, it is possible to see that while kitchens span among 9 and 14 m<sup>2</sup>, the median value is 13 m.

This solution, however, can be not at all immediate for readers not familiar with those concepts [18]. A possible alternative is the violin plot, in which an area is used to represent the span of the distribution over an axis, and its width represents the density over the axis. A third solution is the beeswarm plot or dot plot, in which each value is visualized as a dot and disposed on one axis to represent the distribution and on the other to avoid overlapping. This model is very intuitive even if works with a lower amount of data in respect to the two previously cited.



**Fig. 8** Boxplot showing the distribution of minimum area per room type

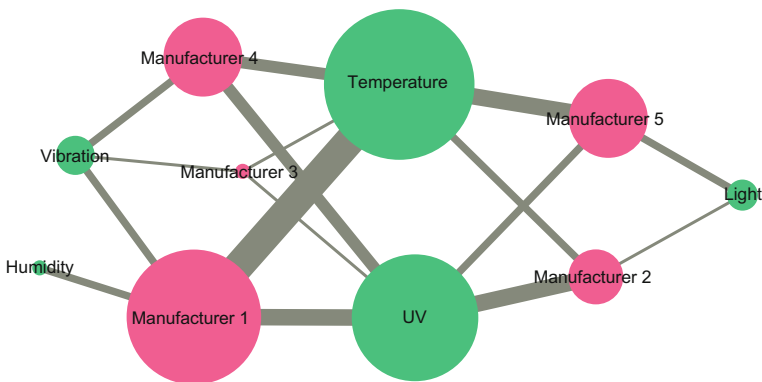
If the goal is instead to show the distribution on two dimensions, the most common solution is the usage of a scatterplot, sometimes called bubble chart if a further dimension is mapped as circles area.

This visual model shows its limit when dealing with large amount of records, or when multiple values are disposed over a small area and is impossible to read the density of values [19]. A solution is the binning approach by dividing the space in equal areas and using color to show the number of values falling into each one. A second solution, less known and less immediate, is the contour chart, in which a third dimension is calculated evaluating the amount of record for a given point and then visualized as projection of a three-dimensional space.

If the goal is to show the distribution of three or more dimensions, the parallel coordinates allow to compare them in pairs. An example is to visualize the maintenance date for each type of maintainable asset per floor. The downside of this approach is that by changing the order of dimensions, it will also change the kind of reading. If instead of dealing with continuous data, the goal is to show distribution among multiple categories, the alluvial diagram is the correspondent solution [20].

### 2.4 Networks

If the goal is to represent connections between elements, several visual solutions have been developed, for example, if you want to visualize each sensor installed in an estate, which performance is measuring (e.g., humidity and temperature) and who is maintaining them as in Fig. 9. An intuitive approach is the node-link diagram, although its use brings with it a complexity, which is the disposition of nodes in a way that enables the reading of their connection patterns. Several algorithms have been developed for this task, and the most used are the force-based



**Fig. 9** Network diagram showing the number of sensors (green nodes) across an estate and companies maintaining them (red nodes)

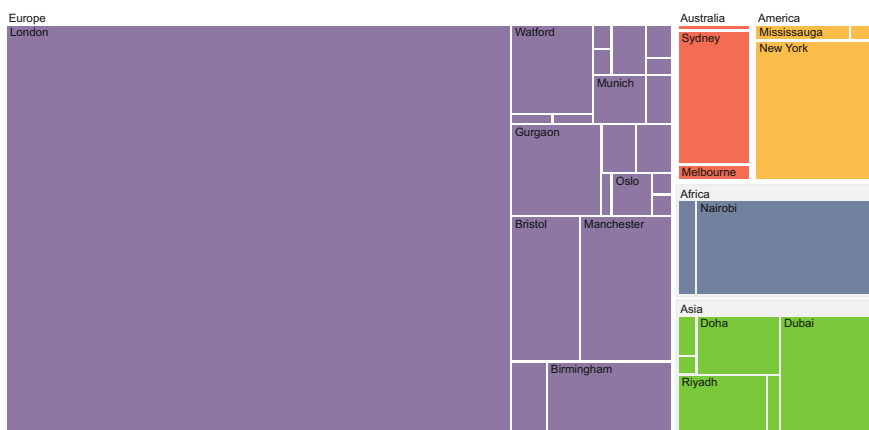
ones, in which two main forces define the disposition: The first one is the repulsion among nodes, and the second one is to create links to keep nodes close.

Another solution is the arc diagram, in which nodes are disposed in line, and links are represented as arcs connecting them. Also in this case, node order can be algorithmically computed to minimize links overlapping. If the goal is to show flows among nodes, an interesting solution is the Sankey diagram, in which nodes are represented as bars representing the total flow passing through them and links' width is proportional to their value.

## 2.5 Hierarchies

When the goal is to show tree-like structures, such as structure of project team or company structure, this group of visualization comes at hand. The most straightforward solution is the usage of node-link diagrams that visually represents the structure.

If the goal is also to represent values for the different items in the hierarchy, a common solution is the icicle diagram [21] or its radial variant called sunburst diagram. If the communicative goal is more on the leaf values rather on the hierarchical structure, a good solution is the usage of a treemap, which divides the available space in areas proportional to the values arranged by the underlying hierarchical structure. An example is Fig. 10, in which a treemap is used to show the amount of assets owned by an organization across the administrative hierarchy. A variation is the circle packing, in which values are represented with circles grouped according to hierarchical structure.



**Fig. 10** Treemap showing the amount of assets owned divided by country and city

### 3 Good Practices

As stated in the previous section, visualizations are not the result of a mechanical operation, but rather the result of a series of choices that affects the final result [22]. There are no golden rules [23], however, both from practice and from research on human perception some good practices emerged over time. It is possible to identify four key points that should be evaluate critically when dealing with information visualization: the visual hierarchy, the legend design, the use of annotations, and the juxtaposition of visualizations.

#### 3.1 *Building the Visual Hierarchy*

In the theory related to visual communication, visual hierarchy is the design of the order of importance of elements disposed in the space [24].

The visual hierarchy is related to the whole layout from the title to the visualization itself and to legends and annotations. Visualizations indeed are not isolated objects, but they are always placed inside a specific container such as a web page, a newspaper spread, and the page of a book. Furthermore, visual marks are just one of the elements constituting a visualization, and on their own, they would not be enough to read the encoded information. Text strings are fundamental to decode a visualization, understand its parts, and how to read them.

A brief list of items that should be kept in mind while designing a visualization are:

- **Title and Subtitle.** After looking at the visualization, the first question in the mind of the reader is “what am I looking to.” A good title clearly related to the visualized information can help to provide a good access point.
- **Axes.** If information is encoded on the horizontal or vertical axis, it should be clear which data it represents and the used scale.
- **Labels.** When marks are used to represent unique items in our dataset, their meaning should be provided clearly to the user.
- **Key Map (Legend).** The legend is fundamental in any data visualization, since it provides the keys to decode encoded information. A common mistake is to forget to add this information in the visualization. More details will be provided in the next section.
- **Annotations.** By highlighting part of the visualization, it is possible to provide context and insights on the presented data to the reader.

An example is the work by Simon Scarr for the South China Morning Post (Fig. 11), which features a clear visual hierarchy. It represents Hong Kong’s power consumption in 2013. The attention is immediately caught by the main visualization, and the immediate second level of reading is the title. The reader can then understand the meaning of colors (e.g., lighting, refrigeration, and cooking)

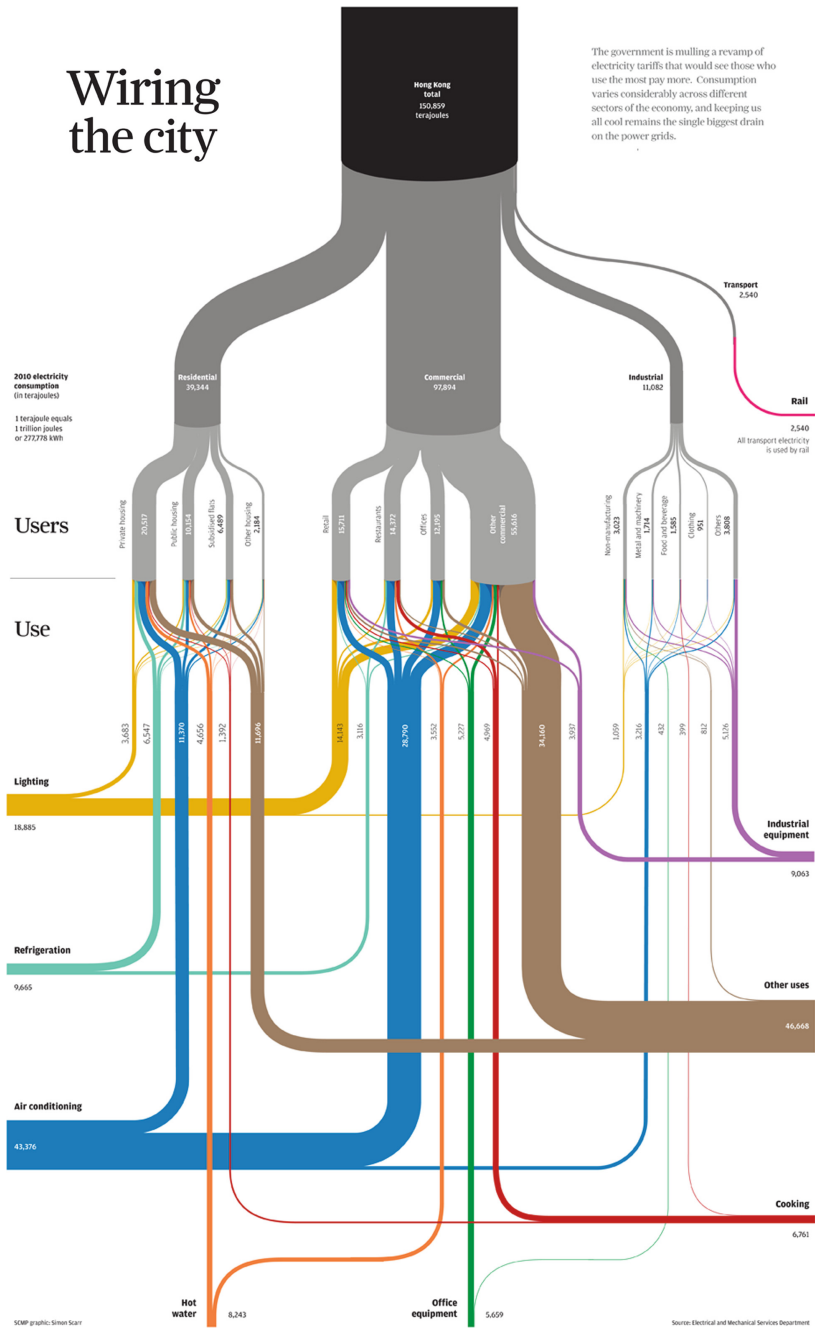


Fig. 11 “Wiring the city” designed by Simon Scarr for the “South China Morning Post” [25]



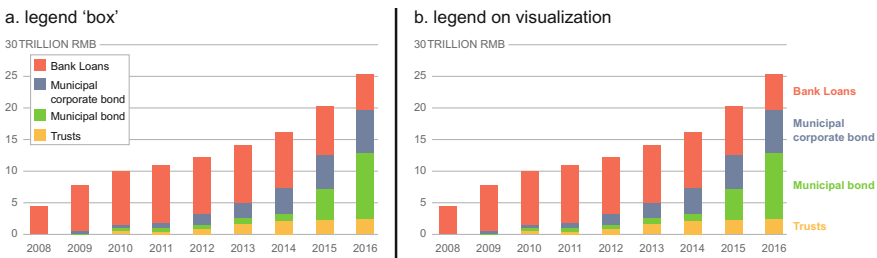
identifying the main area of consumption, and with the vertical axis, they can understand that the upper part is about users and the lower part about kind of usage. The use of typography, varying size, and weight of the font face helps the user to follow the intended hierarchy of the information.

### 3.2 Key Map (Legend)

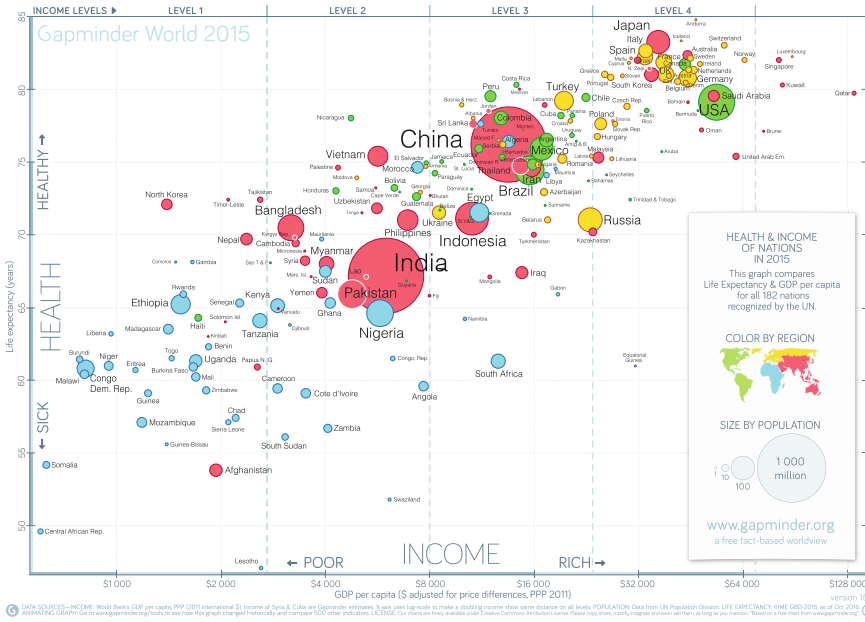
The map keys are the graphical and textual elements that describe how to interpret color and symbols used in the visualization. In general, each visual variable used in the visualization needs to be explained in the legend. Size of items and their scale, color schemes and textures, and the position of the elements: All these variables need to be explained. Well-designed key maps make the process of decoding a visualization easier for the reader.

Many available software that enable the creation of visualizations solve the problem by enclosing all the keys in a box: The reader can, therefore, move between visualization and the legend area and decode the information. However, this is not the best practice: The UK Office for National Statistics guidelines openly discourage this approach stating that “a legend or key should not be used, instead label the data directly” [26]. Taking as example the following graph (Fig. 12), it is possible to see that by placing the labels directly near the values, the reading is simpler not having to move continuously from the visualization to the legend; or by taking again the example shown in Fig. 11, it is possible to see that there is any “legend” box: All the required information is directly explained on the visualization itself.

Disposing legend keys directly on the visualization is a great approach, but it is not always possible to adopt this solution. When the visualization is dense, for example, it could create even more noise. To avoid visual clutter, an accepted solution is to move part of the keys to a dedicated space [27]. In this case, it is important to leverage on the visual hierarchy to make it findable by the reader, making as simple as possible the decoding.



**Fig. 12** On the left, (a) Keys are enclosed in a box, making the reading slower. On the right, (b) Keys are placed directly on the visualization



**Fig. 13** Countries by life expectancy, GDP per capita and population in 2015. Credits [www.gapminder.org](http://www.gapminder.org)

A good example is the chart in Fig. 13. Labels are used to provide the name of single country. Size and color are then provided in a box in the bottom-right corner. Colors are provided on a world map, helping the user to understand which countries belong to the same geographical group.

If there is a need of relying on the legend box, it should be kept small, synthetic, and easy to read. Legends that are too big and complex require too much cognitive effort and make heavier the process of understanding a visualization rather than making it easier. If the legend is too long and complex, it could also mean that too many visual variables have been used in the visualization, and a possible solution could be to reduce the amount of data visualized.

### 3.3 Annotations

Another way to engage users with your visualizations is to enrich them with textual and graphical annotations. They provide context to the data and add depth and richness to the insights that users can have with the visualization [28].

Annotations are an additional interpretative layer that the designer can add on top of the simple visual encoding of the data. They are usually added as the last step in the process.

For example, in a scatterplot, it can be used to highlight the most relevant items. In the same way, in a network graph, clusters of similar closer nodes can be highlighted with a circle and described with a short text. Or, in a trendline showing data over time, the highest peak can be annotated with a sentence about the events that sparked the rise of the data. Annotations can be useful to provide the meaning of the represented data: In Fig. 13, the axes have a double description, the first by data dimension (life expectancy, GDP per capita), then by the interpretation suggested by the author (health, income).

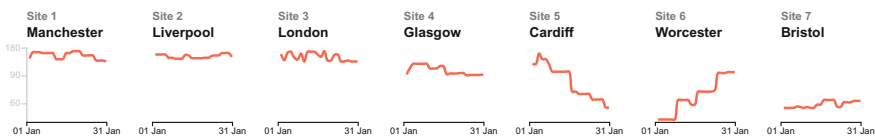
### 3.4 Small Multiples

As stated in Sect. 2, visual models should be seen as blueprints that can be hybridized and combined according to the communicative aim of the visualization. A good practice is the juxtaposition of more than one graph to enable the visual comparison among sets of data [29].

Small multiples are series of similar visualizations that allow their comparison among dominant dimensions, e.g., time or space. An example is the usage of the same visual model repeated over an interval of time to see its evolution. Another usage of small multiples is to compare the same graph over different locations. Figure 14 shows the number of workers in different construction sites of the same construction company.

When adopting this approach, it is fundamental that at least one dimension is coherent among the visualizations to allow a visual comparison. In this case, both the temporal scale and the vertical axis are fixed across the charts, enabling a comparison of absolute values. A second solution would be to use different scales on the vertical axis, to compare the trajectories.

The small multiples approach can be applied to any visual model even if could be difficult when dealing with models that have more complexity.



**Fig. 14** Small multiples line charts the number of workers in different building sites over the same month

## 4 Conclusions

Information visualization plays a key role for Industry 4.0 in construction. As more and more data is produced by digital and automatic processes, as described in Chapter “[Shaping the Future of Construction Professionals](#)”, it is essential to effectively communicate to enable informed decisions. This chapter highlighted the relevance of identifying good practices and the possible errors that could bring a failure in the visual communication.

Today, much software is available for the creation of data visualizations even for the most unusual visual models and therefore available for everyone. The problem is indeed communicative rather than technical: When communicating data, it is important to clearly identify the communicative goal, the device that will be used, and the kind of public will engage with it. Visualization is one of the possible tools that can be used to convey information to our public and are not isolated and self-contained. In this chapter, the encoding/decoding process has been presented as the result of many choices made by the visualization author.

In this chapter, the focus has been on static visualization, without dealing with animation or interaction. These two features would require a chapter on their own, since brings with them a vast array of complexity. Today, with the shift to digital devices, it is possible to create very complex web-based artifact using ad-hoc coding libraries (such as d3.js) or with different services. While interactions are useful and powerful, it is often overlooked. Sometime, the need for interaction comes from the inability of the author of making clear visual and communicative choices, hoping to be able to “show everything.” As a general rule of thumb, the suggestion is that if an information is relevant, it should be immediately visible and not hidden using interaction features such as rollover: Therefore, the design effort should be primarily on the encoding of data.

**Acknowledgements** The approach exposed in this publication is the result of the joint experience of the people at DensityDesign Lab (Politecnico di Milano). A special thank goes to Gabriele Colombo and Angeles Briones, who over the years helped the author in identifying and test approaches and sources on the topic.

## References

1. Brinton, W.C.: *Graphic Methods for Presenting Facts*. The Engineering Magazine Company, New York (1914)
2. Cairo, A.: *Functional Art* (2012)
3. Spence, I.: William playfair and the psychology of graphs. In: American Statistical Association; Papers Presented at the Joint Statistical Meetings, pp. 2426–2436 (2006) (Online). Available: <https://www.tib.eu/en/search/id/BLCP:CN064191962/William-Playfair-and-the-Psychology-of-Graphs?cHash=cad10a6d21c18ff38295d5483f53c0f1>. Last accessed 7 Dec 2021
4. Ware, C.: *Information Visualization: Perception for Design*, 2nd edn (2004)

5. Card, S.K., Mackinlay, J.D., Shneiderman, B., (eds.): *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1999)
6. Fekete, J.D., Van Wijk, J.J., Stasko, J.T., North, C.: The value of information visualization. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence in Lecture Notes Bioinformatics)*, LNCS, vol. 4950, pp. 1–18 (2008). [https://doi.org/10.1007/978-3-540-70956-5\\_1](https://doi.org/10.1007/978-3-540-70956-5_1)
7. Bateman, S., Mandryk, R.L., Gutwin, C., Genest, A., McDine, D., Brooks, C.: *Useful Junk? The Effects of visual Embellishment on Comprehension and Memorability of Charts* (2010). <https://doi.org/10.1145/1753326.1753716>
8. Borkin, M.A., et al.: What makes a visualization memorable. *IEEE Trans. Vis. Comput. Graph.* **19**(12), 2306–2315 (2013). <https://doi.org/10.1109/TVCG.2013.234>
9. Hung, Y.H., Parsons, P.: Assessing user engagement in information visualization. *Conf. Hum. Factors Comput. Syst. Proc. Part F1276*, 1708–1717 (2017). <https://doi.org/10.1145/3027063.3053113>
10. Bertin, J., Barbut, M.: *Sémiologie graphique. Les diagrammes, les réseaux, les cartes*. La Haye, Paris; Gauthier-Villars, Mouton (1967)
11. Morrison, J.L.: A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. *Int. Yearb. Cartogr.* **14**(1974), 115–127 (1974)
12. MacEachren, A.M.: How maps work: representation, visualization and design. *How maps Work Represent. Vis. Des.* (1995)
13. Halik, Ł.: The analysis of visual variables for use in the cartographic design of point symbols for mobile augmented reality applications. *Geod. Cartogr.* **61**(1), 19–30 (2018). <https://doi.org/10.2478/v10277-012-0019-4>
14. Roth, R.E.: Visual variables. *Int. Encycl. Geogr. People, Earth, Environ. Technol.* 1–11 (2017). <https://doi.org/10.1002/9781118786352.wbieg0761>
15. Mauri, M., Elli, T., Caviglia, G., Uboldi, G., Azzi, M.: RAWGraphs. In: *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter—CHIItaly'17*, pp. 1–5 (2017). <https://doi.org/10.1145/3125571.3125585>
16. Heer, J., Bostock, M., Ogievetsky, V.: A tour through the visualization zoo. *Commun ACM* (2010). <https://doi.org/10.1145/1743546.1743567>
17. Heer, J., Kong, N., Agrawala, M.: Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In: *Proceedings of the 27th International Conference on Human Factors in Computing Systems—CHI 09*, p. 1303 (2009). <https://doi.org/10.1145/1518701.1518897>
18. Allen, M., Poggiali, D., Whitaker, K., Marshall, T.R., Kievit, R.A.: Raincloud plots: a multi-platform tool for robust data visualization. *Wellcome Open Res.* **4**, 1–41 (2019). <https://doi.org/10.12688/wellcomeopenres.15191.1>
19. Yan, H.: *2D Density Plot—From Data to Viz* (2020). <https://www.data-to-viz.com/graph/density2d.html>. Last accessed 7 Dec 2021
20. Rosvall, M., Bergstrom, C.T.: Mapping change in large networks. *PLoS One* **5**(1) (2010). <https://doi.org/10.1371/journal.pone.0008694>
21. Kruskal, A.J.B., Landwehr, J.M.: *Icicle Plots: Better Displays for Hierarchical Clustering*. Published by: Taylor & Francis, Ltd. on Behalf of the American Statistical Association Stable. *Icicle Plots: Better Displays for Hierarchical Clusterin*, vol. 37, no. 2, pp. 162–168. <http://www.jstor.org/stable/2685881> (1983)
22. Kitchin, R.: *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences* (2014)
23. Pettersson, R.: *Information Design—Principles and Guidelines*, vol. 29, no. 2, pp. 167–182 (2010)
24. White, A.: *The Elements of Graphic Design*. Allworth Press, New York (2011)
25. Scarr, S.: *Wiring the City*. South China Morning Post (2013)
26. Office for National Statistics.: *Chart Titles and Text—Style.ONS* (2020). <https://style.ons.gov.uk/category/data-visualisation/titles-and-text/#legends-or-keys>. Last accessed 7 Dec 2021

27. U.S. Census Bureau.: Data Visualization Standards—Legends (2019). <https://xdgov.github.io/data-design-standards/components/legends>. Last accessed 7 Dec 2021
28. Brock, T.: How to Improve Your Data Visualizations with Annotations|Infragistics Blog (2015). [https://www.infragistics.com/community/blogs/b/tim\\_brock/posts/annotate-wisely](https://www.infragistics.com/community/blogs/b/tim_brock/posts/annotate-wisely). Last accessed 7 Dec 2021
29. Tufte, E.R.: Envisioning Information. Graphics Press (1990)