Unit III:

Visualization formats: Bar chart, Histograms, Pie charts, Scatter plots, Heat maps, Line charts, Bubble charts, Radar charts, Waterfall charts, Tree maps, Area charts

Basic principles for data visualization, Layout and design: communicative elements, Prioritize patterns in your visualizations: Gestalt

11 formats

There are two types of visualizations: static and interactive. Their use depends on the search and analysis dimension level. **Static visuals can only analyze data in one dimension, whereas interactive visuals can analyze it in several.**

As with any other form of communication, familiarity with the code and resources that are available to us is essential if we're going to use them successfully our goal. In this page, we present the different kinds of graphics that we can use to transform our data into information. This group of visualization types is listed in order of popularity in the "Visualization Universe" project by Google News Lab and Adioma, as of the publication of this report.

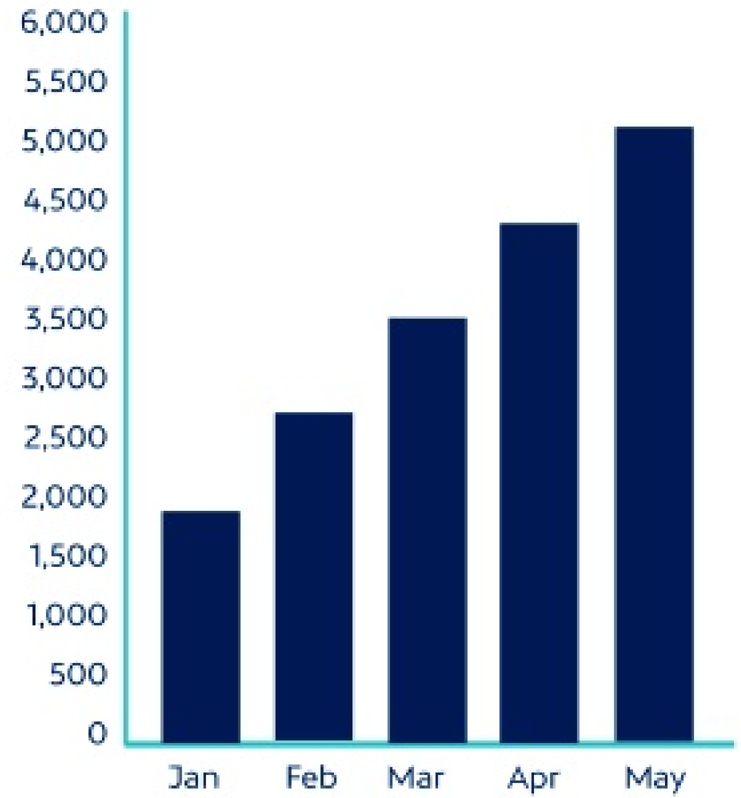
1. Bar chart

Bar charts are one of the most popular ways of visualizing data because they present a data set in a quickly understood format that enables viewers to identify highs and lows at a glance.

They are very versatile, and they are typically used to compare discrete categories, to analyze changes over time, or to compare parts of a whole.
The three variations on the bar chart are:

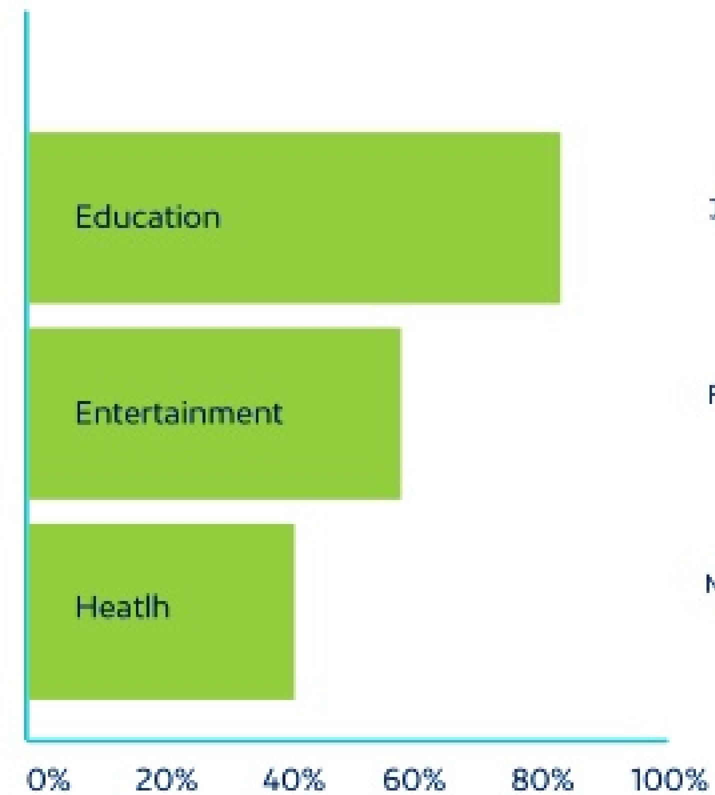
Vertical column

Used for chronological data, and it should be in left-to-right format.



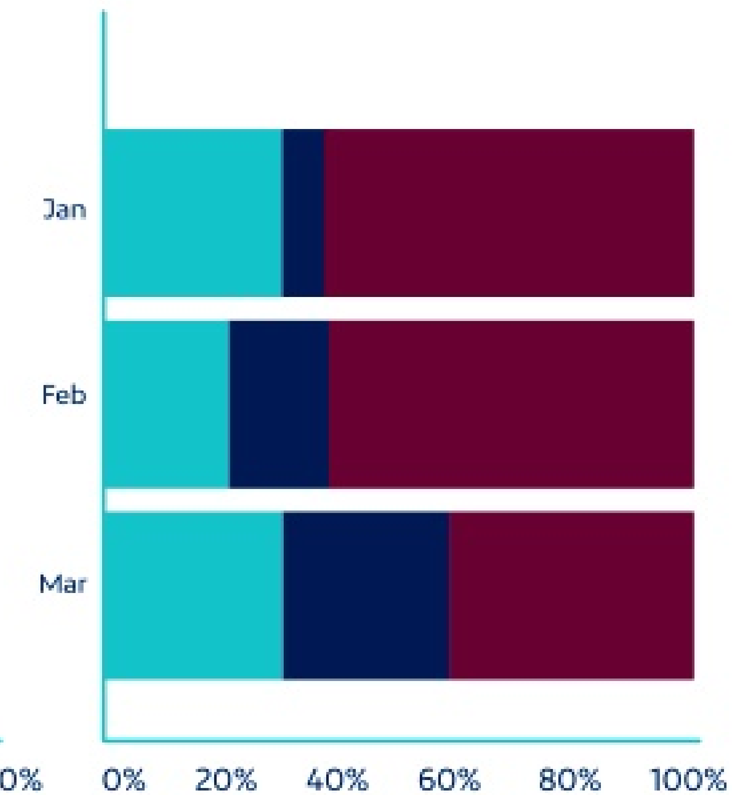
Horizontal column

Used to visualize categories.



Full stacked column

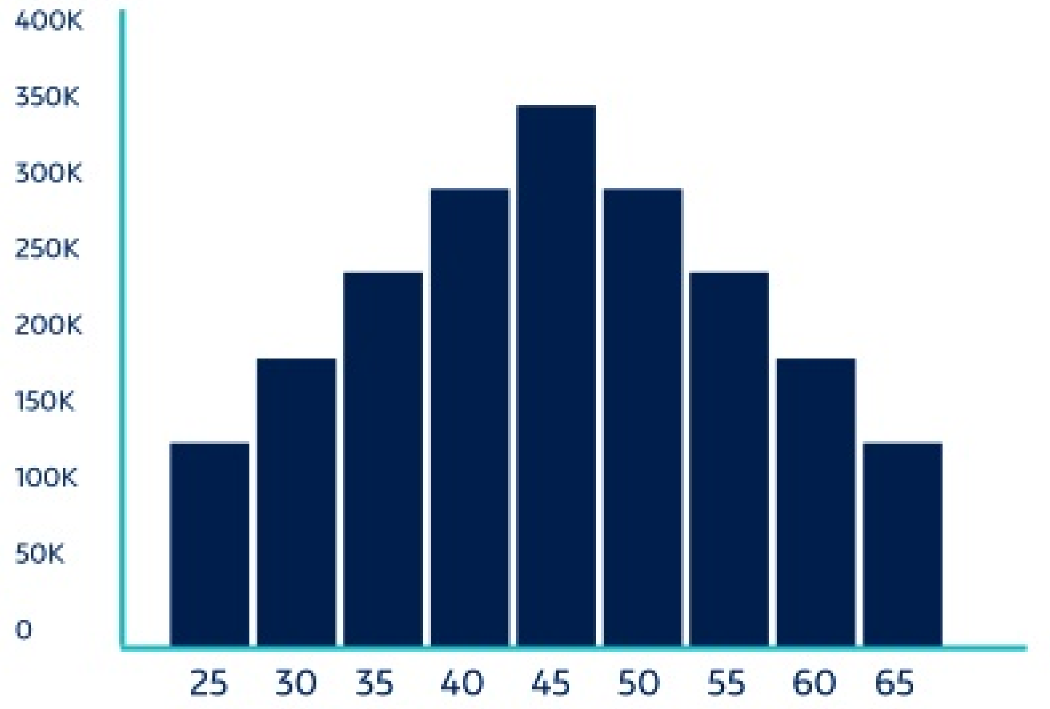
Used to visualize categories that collectively add up to 100%.



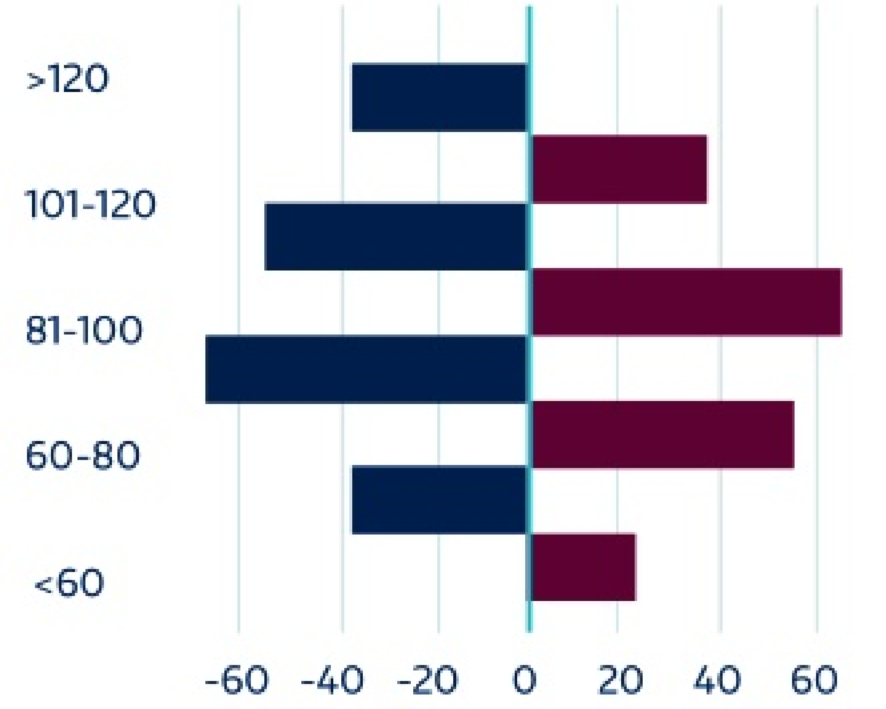
2. Histograms

Histograms represent a **variable in the form of bars, where the surface of each bar is proportional to the frequency of the values represented.** They offer an overview of the distribution of a population or sample with respect to a given characteristic.
The two variations on the histogram are:

- **Vertical columns**
- **Horizontal columns**



Vertical columns

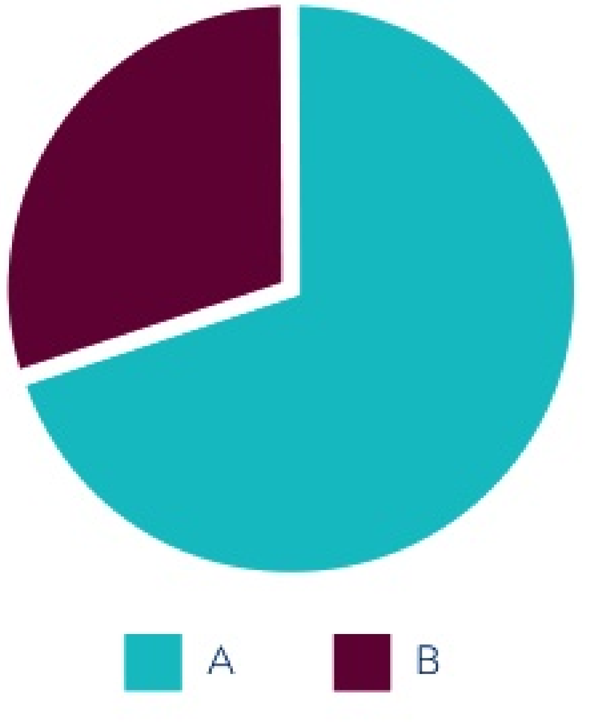


Horizontal columns

3. Pie charts

Pie charts consist of **a circle divided into sectors, each of which represents a portion of the total.** They can be subdivided into no more than five data groups. They can be useful for comparing discrete or continuous data.
The two variations on the pie chart are:

- **Standard:** Used to exhibit relationship between parts.
- **Donut:** A stylistic variation that facilitates the inclusion of a total value or a design element in the center.

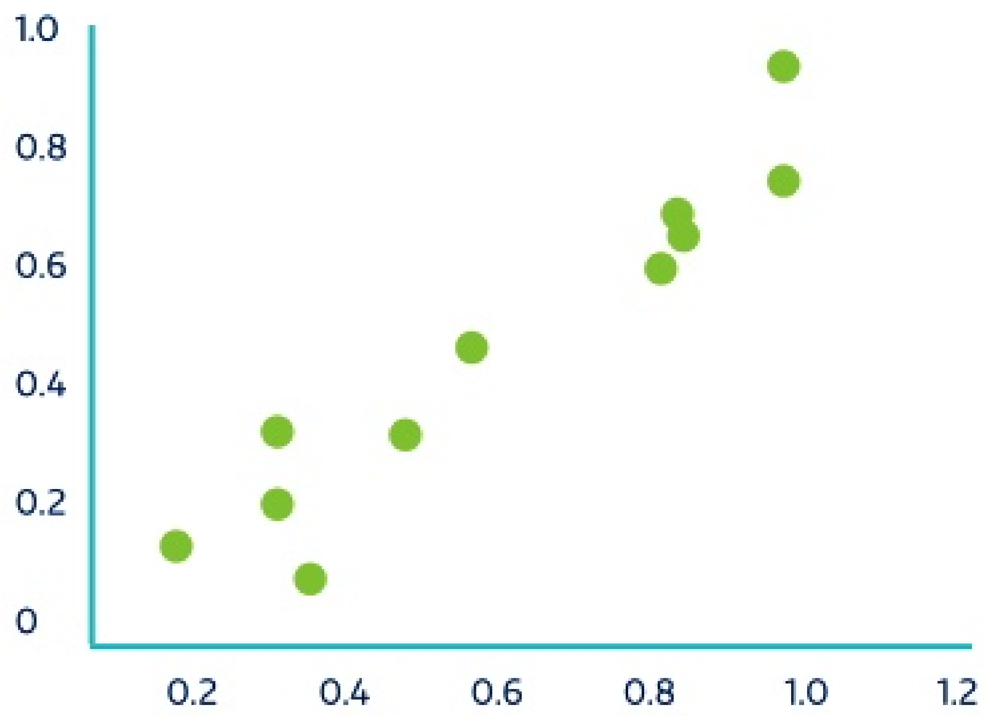


Standard pie chart

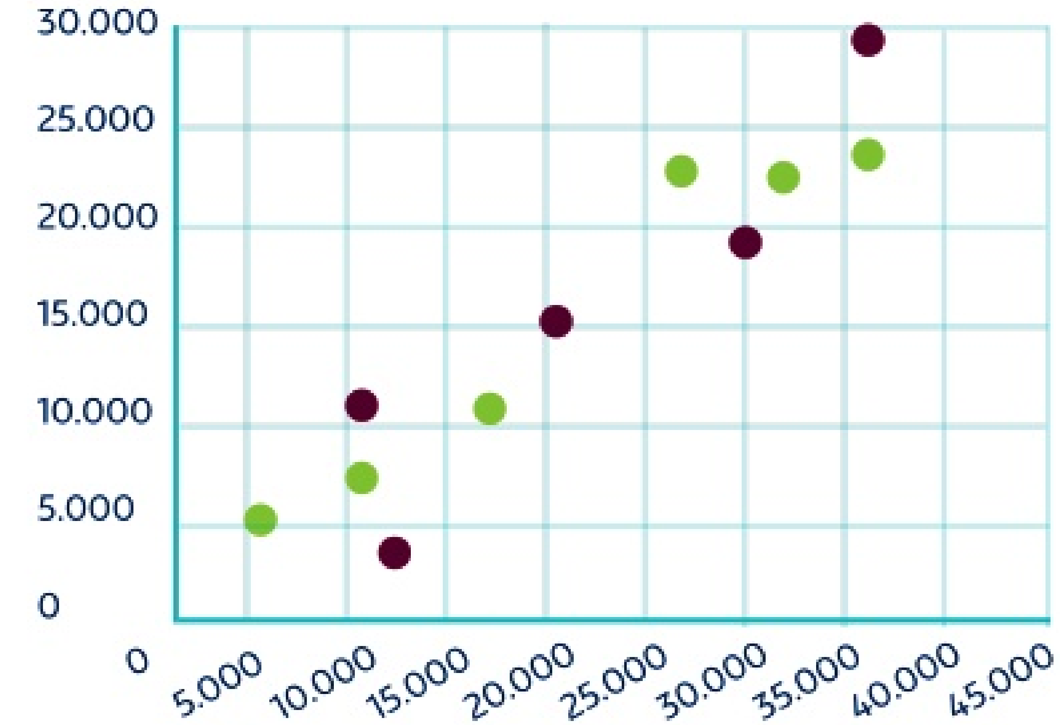
Donut pie chart

4. Scatter plots

Scatter plots use the spread of points over a Cartesian coordinate plane to show the relationship between two variables. They also help us determine whether or not different groups of data are correlated.



Scatter plot

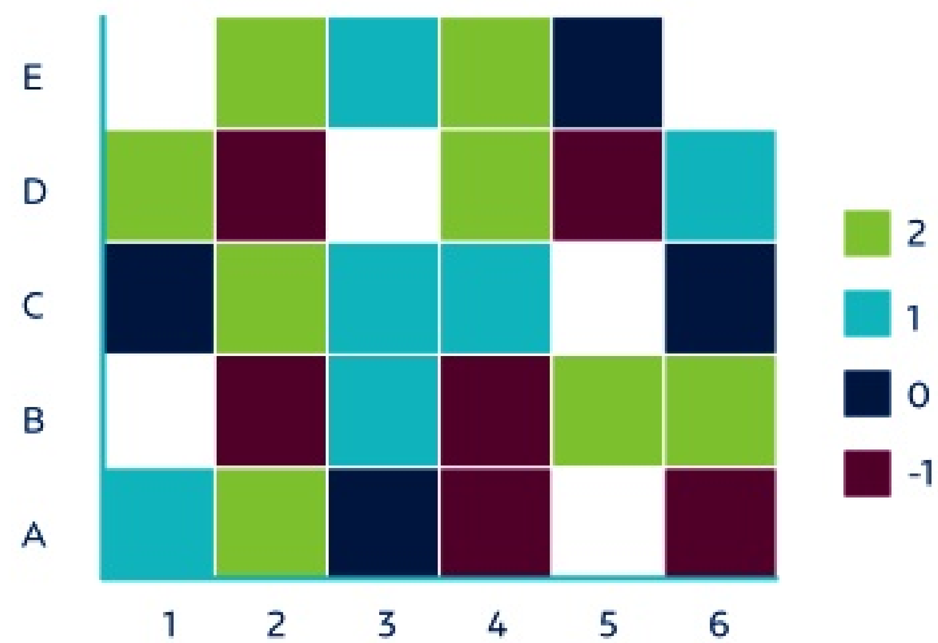


Scatter plot with grid

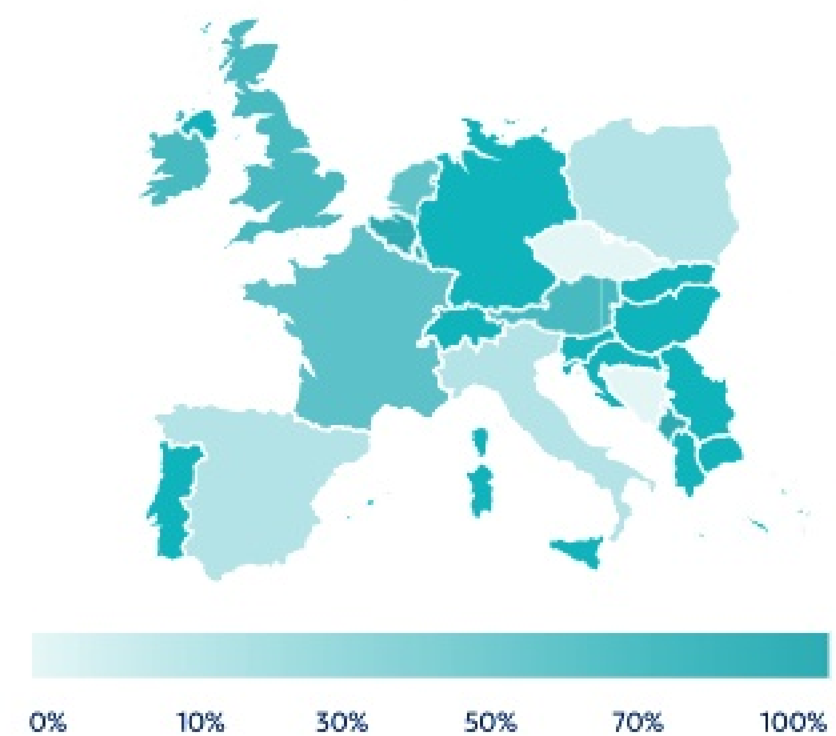
5. Heat maps

Heat maps represent individual values from a data set on a matrix using variations in color or color intensity. They often use color to help viewers compare and distinguish between data in two different categories at a glance. They are useful for visualizing webpages, where the areas that users interact with most are represented with "hot" colors, and the pages that receive the fewest clicks are presented in "cold" colors. The two variations on the heat map are:

- **Mosaic diagram**
- **Color map**



Mosaic diagram



Color map

6. Line charts

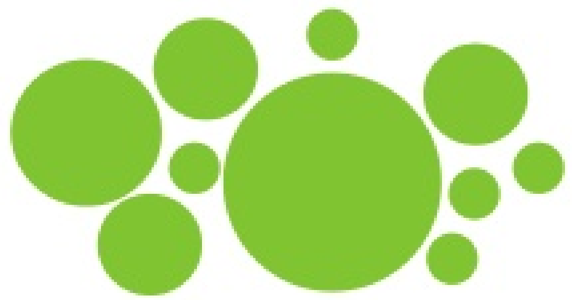
These are **used to display changes or trends in data over a period of time.** They are especially useful for showcasing relationships, acceleration, deceleration, and volatility in a data set.

Line chart

7. Bubble charts

These graphics **display three-dimensional data and accentuate data in dispersion diagrams and maps.** Their purpose is to highlight nominal comparisons and classification relationships. The size and color of the bubbles represent a dimension that, along with the data, is very useful for visually stressing specific values. The two variations on the bubble chart are:

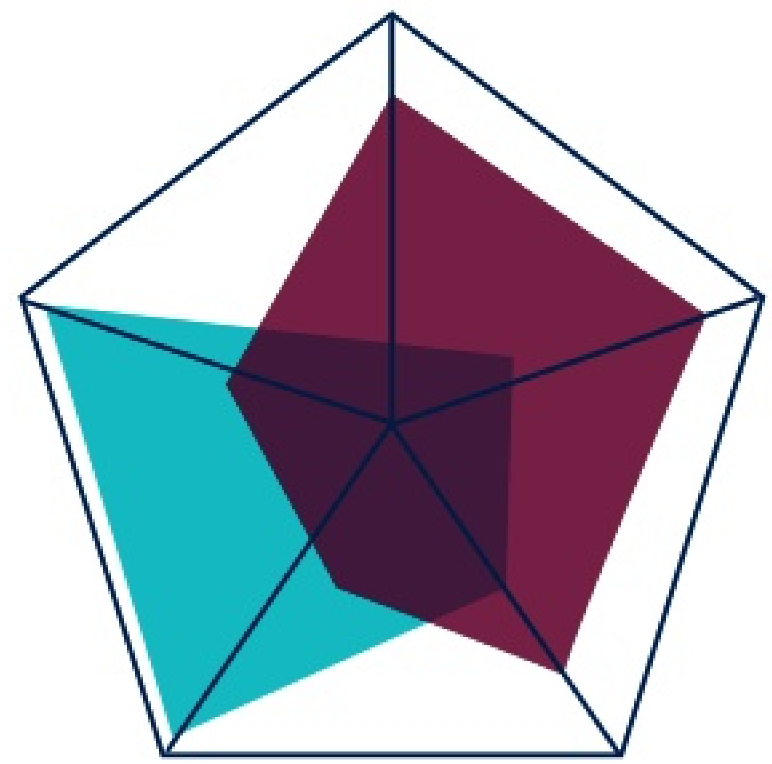
- **The bubble plot:** used to show a variable in three dimensions, position coordinates (x, y) and size.



- **Bubble map**: used to visualize three-dimensional values for geographic regions.

8. Radar charts

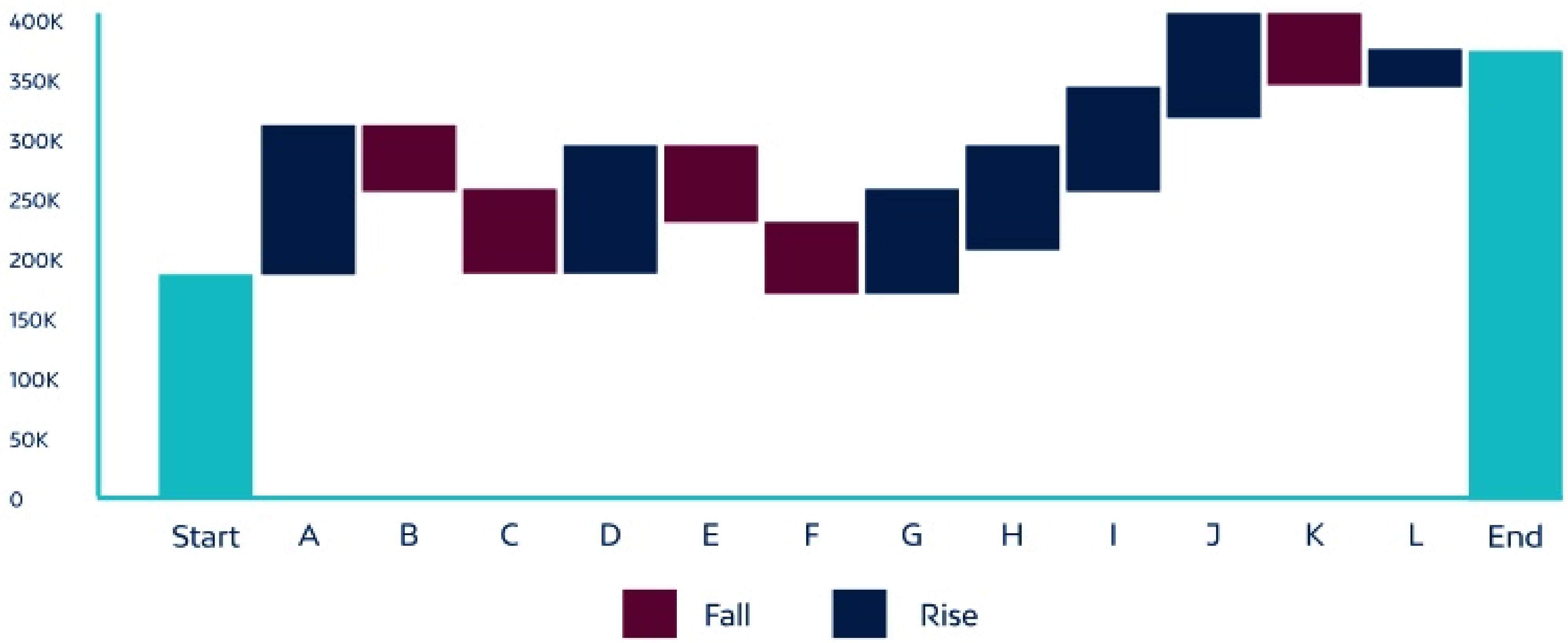
These are **a form of representation built around a regular polygon that is contained within a circle, where the radii that guide the vertices are the axes over which the values are represented.** They are equivalent to graphics with parallel coordinates on polar coordinates. Typically, they are used to represent the behavior of a metric over the course of a set time cycle, such as the hours of the day, months of the year, or days of the week.



Radar chart

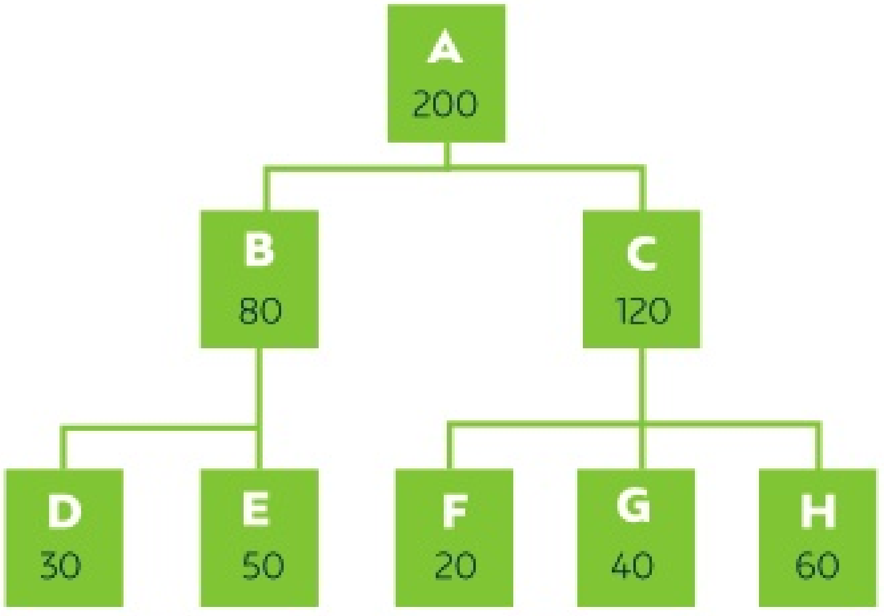
9. Waterfall charts

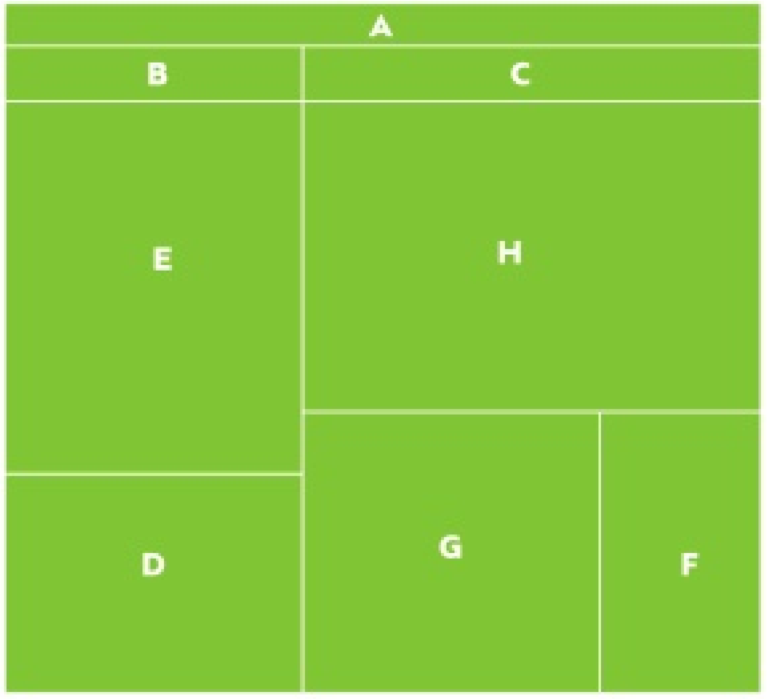
These **help us understand the cumulative effect of positive and negative values on variables in a sequential fashion.**



10. Tree maps

Tree maps **display hierarchical data (in a tree structure) as a set of nested rectangles that occupy surface areas proportional to the value of the variable they represent.** Each tree branch is given a rectangle, which is later placed in a mosaic with smaller rectangles that represent secondary branches. The finished product is an intuitive, dynamic visual of a plane divided into areas that are proportional to hierarchical data, which has been sorted by size and given a color key.

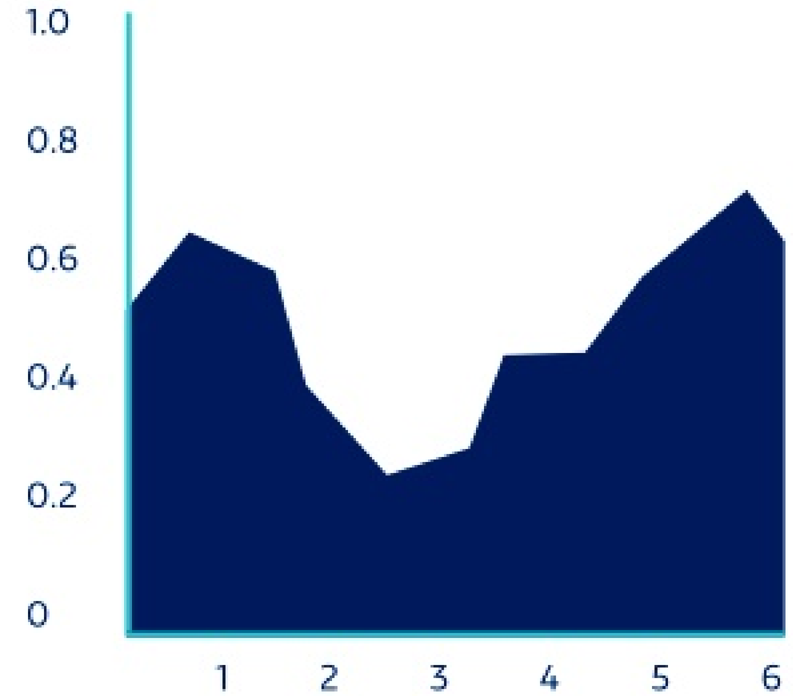




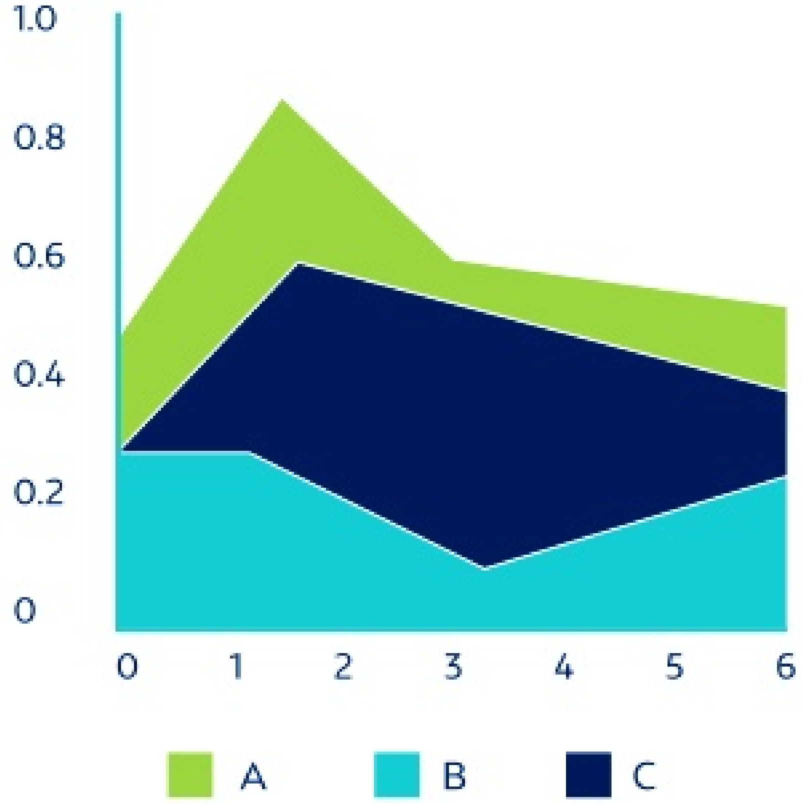
11. Area charts

These **represent the relationship of a series over time, but unlike line charts, they can represent volume.** The three variations on the area chart are:

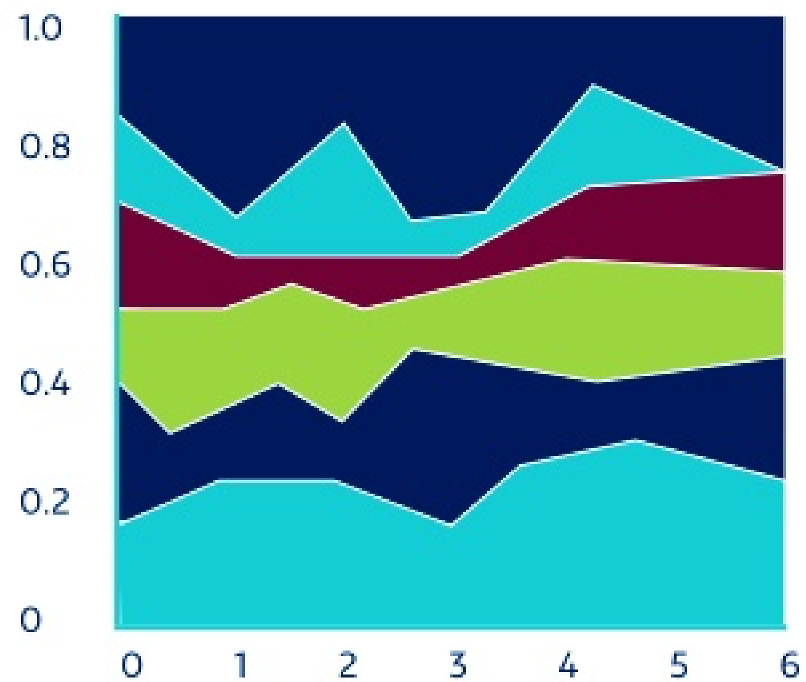
- **Standard area:** used to display or compare a progression over time.
- **Stacked area:** used to visualize relationships as part of the whole, thus demonstrating the contribution of each category to the cumulative total.
- **100% stacked area:** used to communicate the distribution of categories as part of a whole, where the cumulative total does not matter.



Standard area



Stacked area



100% stacked area

Selecting the right graphic to effectively communicate through our visualizations is no easy task. Stephen Few (2009), a specialist in data visualization, proposes taking a practical approach to selecting and using an appropriate graphic:

- **Choose a graphic** that will capture the viewer's attention for sure.
- **Represent the information** in a simple, clear, and precise way (avoid unnecessary flourishes).
- Make it easy to compare data; **highlight trends and differences.**
- **Establish an order** for the elements based on the quantity that they represent; that is, detect maximums and minimums.
- **Give the viewer a clear way to explore the graphic** and understand its goals; make use of guide tags.

Basic principles for data visualization

Graphics with an objective: seeking your mantra

The goal of data visualizations is to help us understand the object they represent. They are a medium for communicating stories and the results of research, as well as a platform for analyzing and exploring data. Therefore, having a sound understanding of how to create data visualizations will help us create meaningful and easy-to-remember reports, infographics, and dashboards. Creating suitable visuals helps us solve problems and analyze a study's objects in greater detail.

The first step in representing information is trying to understand that data visualization.

Ben Shneiderman gave us a useful starting point in his text "The Visual Information-Seeking Mantra" (1996), which remains a touchstone work in the field. This author suggests a simple methodology for novice users to delve into the world of data visualization and experiment with basic visual representation tasks.[5]

5 Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Visual Information Seeking Mantra (p. 336). Available at: https://www.cs.umd.edu/~ben/papers/Shneiderman1996eyes.pdf

Shneiderman introduces his famous mantra on how to approach the quest for visual information, which he breaks down into three tasks:

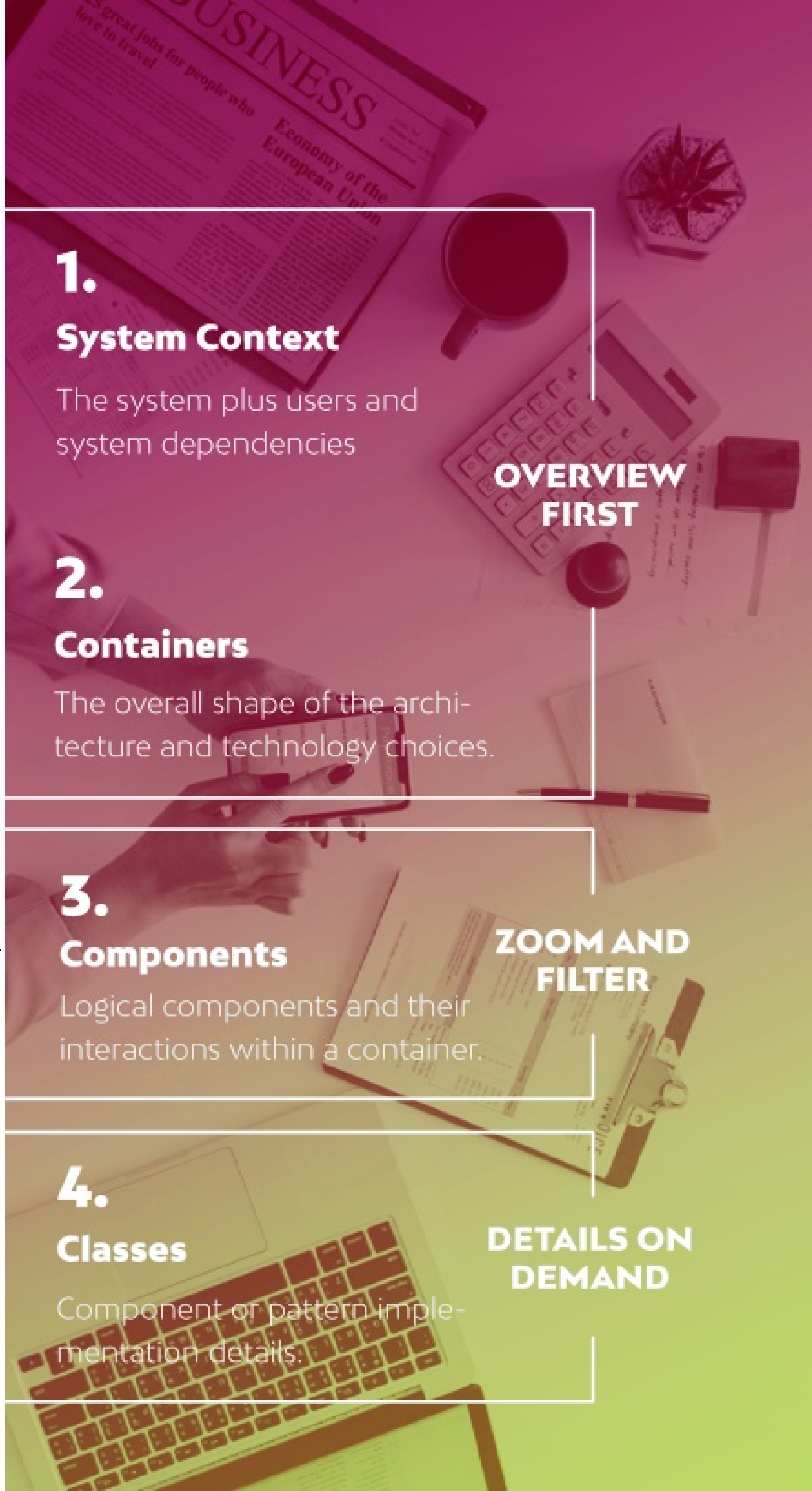
1. Overview first: This ensures viewers have a general understanding of the data set, as their starting point for exploration. This means offering them a visual snapshot of the different kinds of data, explaining their relationship in a single glance. This strategy helps us visualize the data, at all its different levels, at one time.

2. Zoom and filter: The second step involves supplementing the first so that viewers understand the data's underlying structure. The zoom in/zoom out mechanism enables us to select interesting subsets of data that meet certain criteria while maintaining the sense of position and context.

3. Details on demand: This makes it possible to select a narrower subset of data, enabling the user to interact with the information and use filters by hovering or clicking on the data to pull up additional information.

The chart on the right side summarizes the key points to designing such a graphic, with an eye to human visual perception, so that users can translate an idea into a set of physical attributes.

These attributes are: **structure, position, form size, and color.** When properly applied, these attributes can present information effectively and memorably.



Layout and design: communicative elements

In order to begin designing our reports and statements, it is essential to understand that visual representations are cognitive tools that complement and strengthen our mental ability to encode and decode information[6]. Meirelles (2014) notes that: **"All graphic representation affects our visual perception, because the elements of transmission utilized act as external stimuli, which activate our emotional state and knowledge."**

Thus, when our mind visualizes a representation, it transforms the information, merges it, and applies a hierarchical structure to it to facilitate interpretation.

For this reason, in order to have an efficient perceptive impact, it is important to adhere to a series of best practices when creating reports and infographics. As with any other form of communication, success depends largely on the business's familiarity with the established code and the resources available. Space, shapes, color, icons, and typography are a few of the essential elements of a striking visual with communicative power.

6 Meirelles, I (2014). "La información en el diseño," (p.21-22). Barcelona: Parramón.

Structuring: the importance of layout

All visual representations begin with a blank dimensional space that will eventually hold the information which will be communicated. The process of spatial coding is a fundamental part of visual representation because it is the medium in which the results of our compositional decisions and the meaning of our visual statement will be visualized, thereby having an impact on the user.

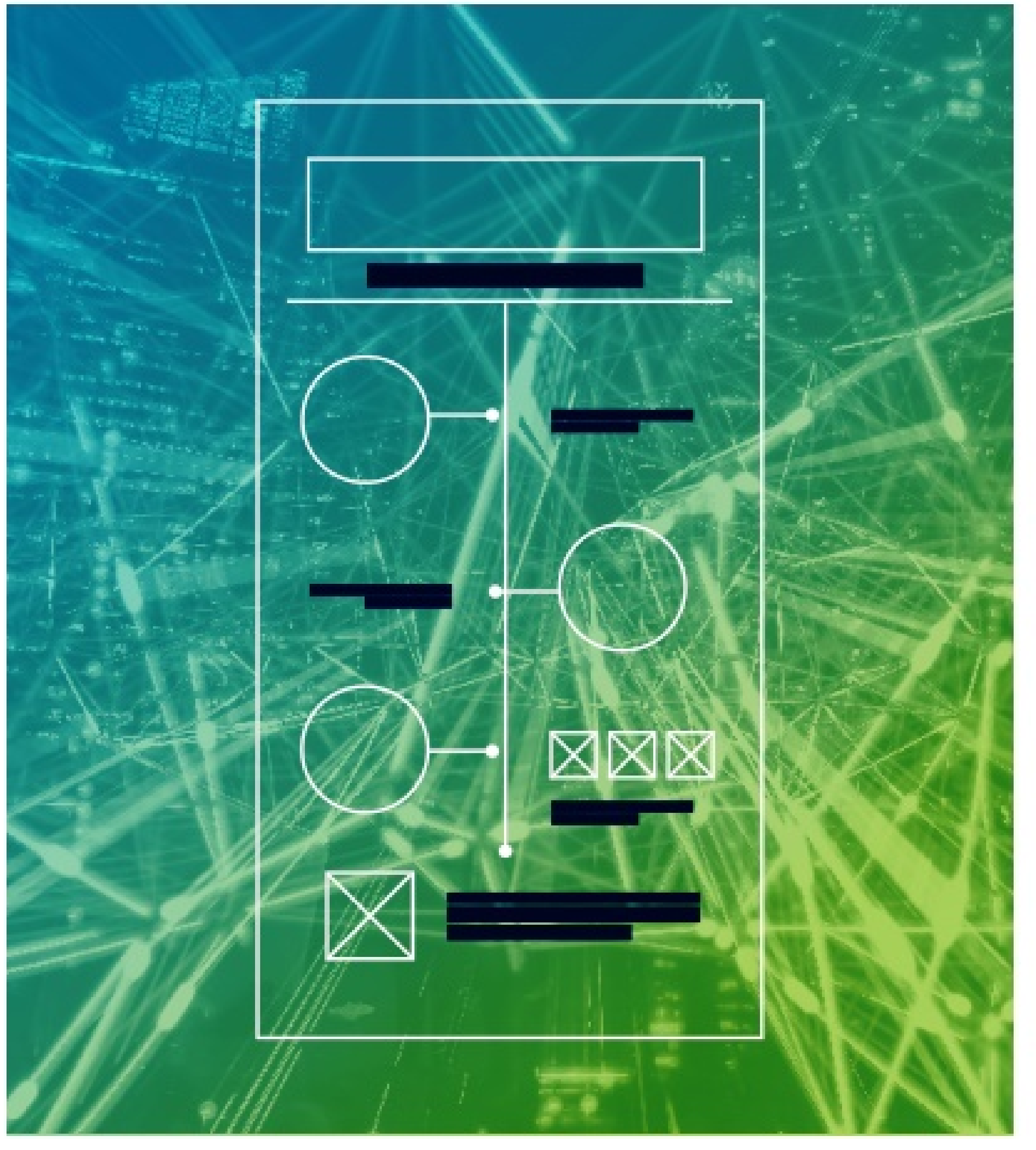
Edward Tufte (1990) defines "layout" as a scheme for distributing visual elements in order to achieve organization and harmony in the final composition. Layout planning and design serve as a template for applying hierarchy and control to information at varying levels of detail.[7] In his book *Envisioning Information*, Tufte offers several guidelines for information design:

- Have a properly chosen format.
- Give a broad visual tour and offer a focused reading at different detail levels.
- Use words, numbers, and drawings.
- Reflect a balance, a proportion, a sense of relevant scale, and a context.

Spatial encoding requires processing spatial proportions (position and size), which have a determining role in the organization of perception and memory.

7 Tufte, E. (1990). Envisioning Information. Cheshire: Graphics Press.

Furthermore, the visual hierarchy of elements plays a role in this encoding process, because the elements' organization and distribution must have a well-defined hierarchical system in order to communicate effectively (Meirelles: 2014). **In a sense, visualizations are paragraphs about data, and they should be treated as such.** Words, images, and numbers are part of the information that will be visualized. When all of the elements are integrated in a single structure and visual hierarchy, the infographic or report will organize space properly and communicate effectively, according to your user's needs.



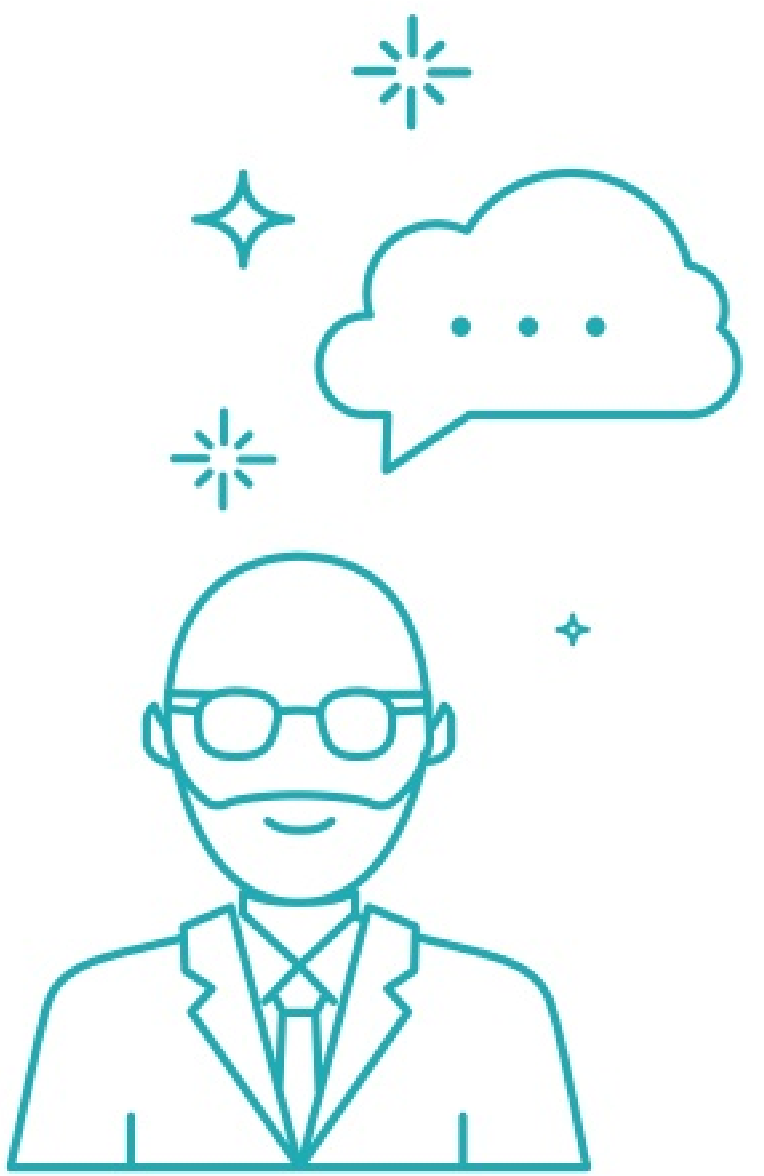
Prioritize patterns in your visualizations: Gestalt

The basic elements of the visualization process also involve preattentive attributes. Preattentive attributes are visual features that facilitate the rapid visual perception of a graphic in a space. Designers use these characteristics to better uncover relevant information in visuals, because these characteristics attract the eye.

Colin Ware, Director of the Data Visualization Research Lab at the University of New Hampshire, has highlighted that **preattentive attributes can be used as resources for drawing viewers' immediate attention to certain parts of visual representations** (2004). According to Ware, preattentive processing happens very quickly—typically in the first 10 milliseconds. This process is the mind's attempt to rapidly extract basic visual characteristics from the graphic (stage 1). These characteristics are then consciously processed, along with the perception of the object, so that the mind can extract patterns (stage 2), ultimately enabling the information to move to the highest level of perception (stage 3). This makes it possible to find answers to the initial visual question, utilizing the information saved in our minds. Colin Ware, cited in Meirelles (2014), explains it as follows:

Bottom up information contributes to the pattern creation process →

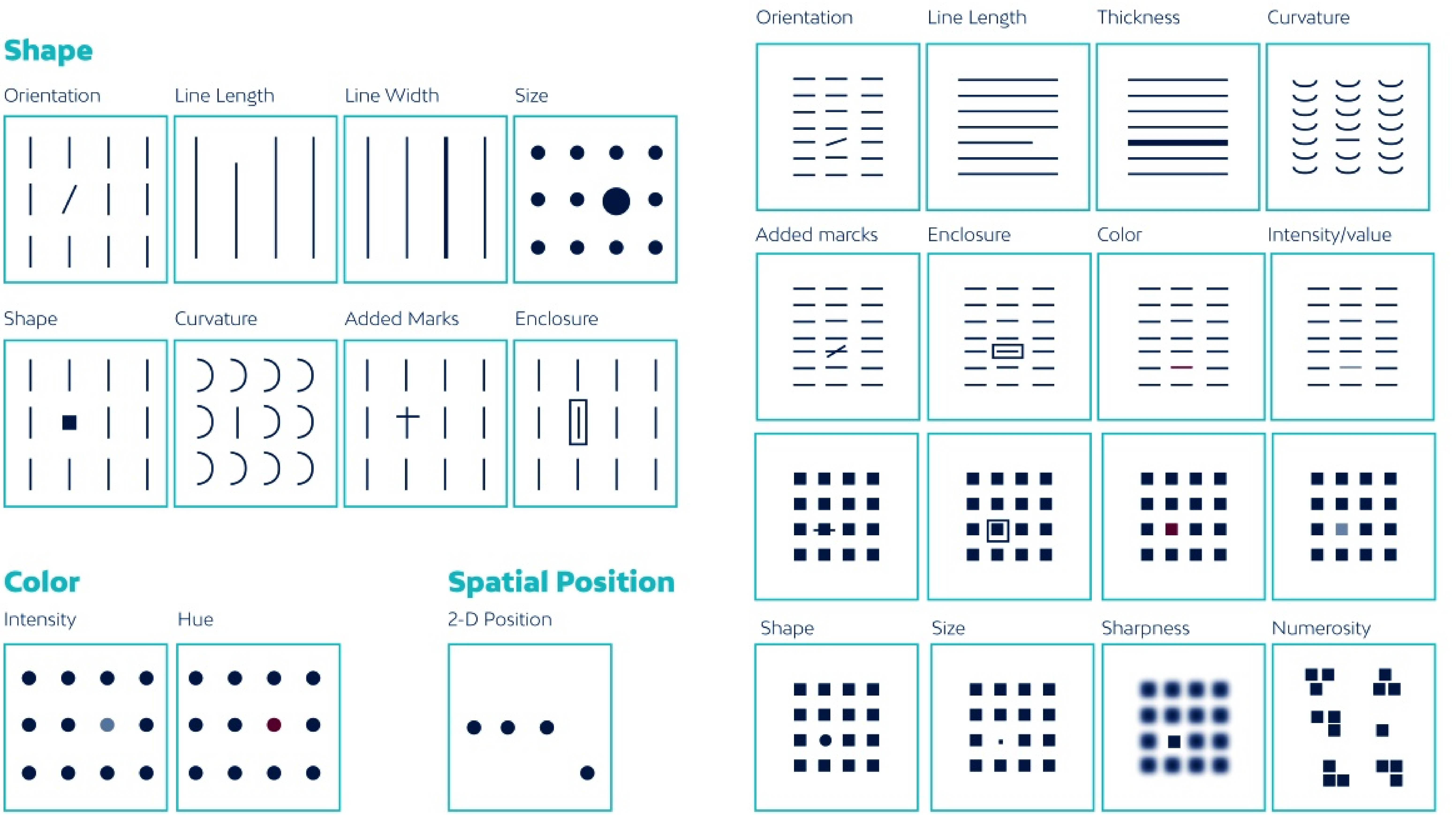
← **Top down process reinforces relevant information**



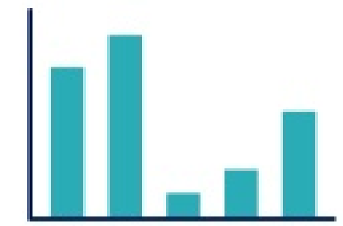
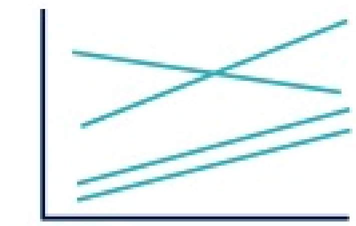
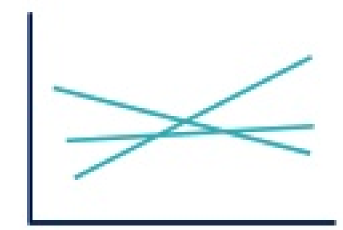
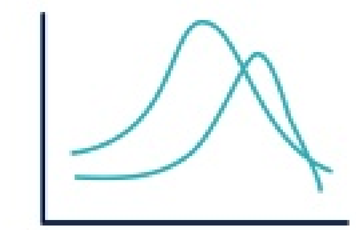
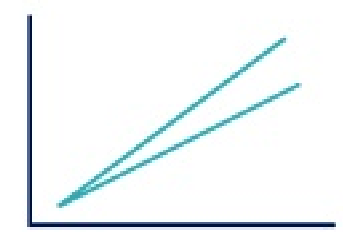
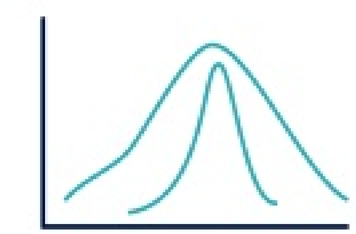
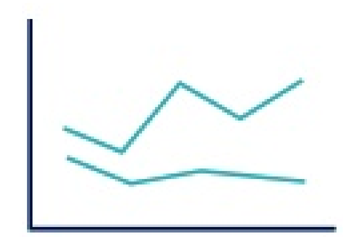
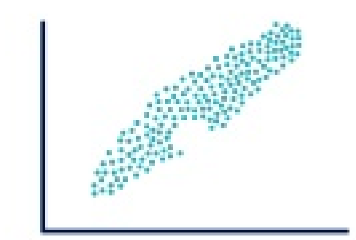
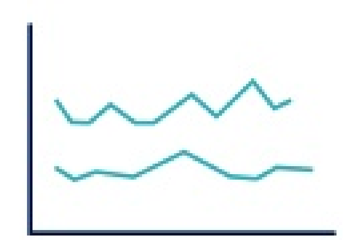
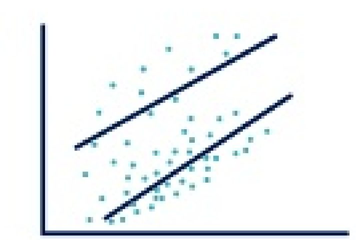
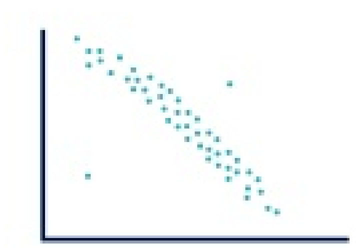
Preattentive attributes enhance object perception and cognition processes, leveraging our mind's visual capacities. Good data visualizations deliberately make use of these attributes because they boost the mind's discovery and recognition of patterns such as lines, planes, colors, movements, and spatial positioning.[9]

9 Dondis, D.A. (2015). La sintaxis de la imagen: introducción al alfabeto visual. Editorial Gustavo Gili: Barcelona
Meirelles, I. (2014). La información en el diseño. Barcelona: Parramón.

The visual below lists preattentive attributes that represent aspects of lines and planes when visualizing and analyzing graphic representation: shape, color, and spatial position.



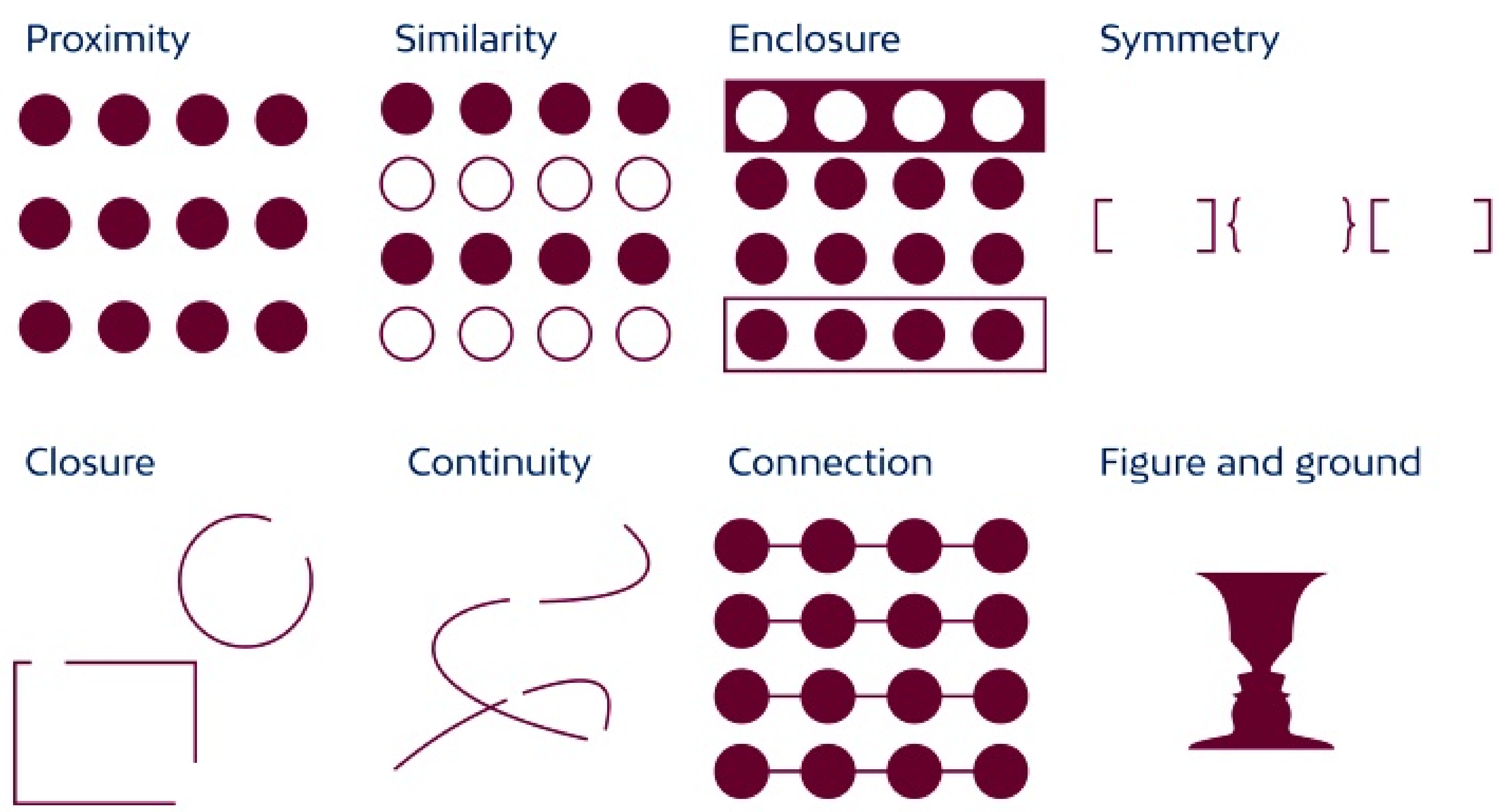
Detecting patterns is fundamental to structuring and organizing visual information. When we create visuals, we often want to highlight certain patterns over others. Preattentive attributes are the alphabet of visual language; analytic patterns are the words that we write by using them. When we see a good visualization, we immediately detect the preattentive attributes and recognize analytic patterns in the visualization. The following table summarizes a few basic analytic patterns:

Pattern	Example	Pattern	Example
High, low and in between		Non-intersecting and intersecting	
Going up, going down and remaining flat		Symmetrical and skewed	
Steep and gradual		Wide and narrow	
Steady and fluctuating		Clusters and gaps	
Random and repeating		Tightly and loosely distributed	
Straight and curved		Normal and abnormal	

Analytic patterns

We have seen how preattentive attributes and patterns make it possible to process and analyze visual information; they also enable us to improve pattern discovery and perceptive inferences and provide processes for solving visualization problems.

Gestalt's principles are the principles that enable us to understand the requirements posed by certain problems so that we see everything as an integral, coherent whole. It involves proximity, similarity, shared destiny, "pragnanz" or pithiness, closure, simplicity, familiarity, and discernment between figure and ground.



Gestalt's principles

According to Dondis (2015), Gestalt's principles help describe the way we organize and merge elements in our minds. They quiet the noise of the graphics so that we relate, combine, and analyze them. These principles come into play whenever we analyze any sort of visualization. Only position and length can be used to accurately perceive quantitative data. The other attributes are useful for perceiving other sorts of data, such as categorical and relational data.

We'll close this section with one piece of practical advice on how to effectively visualize data. Colin Ware in *The Visual Thinking: for Graphic Design* (2008) summarizes the importance of always being mindful of preattentive attributes and patterns when designing a visualization:

"Good design optimizes the visual thinking process. The choice of patterns and symbols is important so that visual queries can be efficiently processed by the intended viewer. This means choosing words and patterns each to their best advantage."