**Chapter**

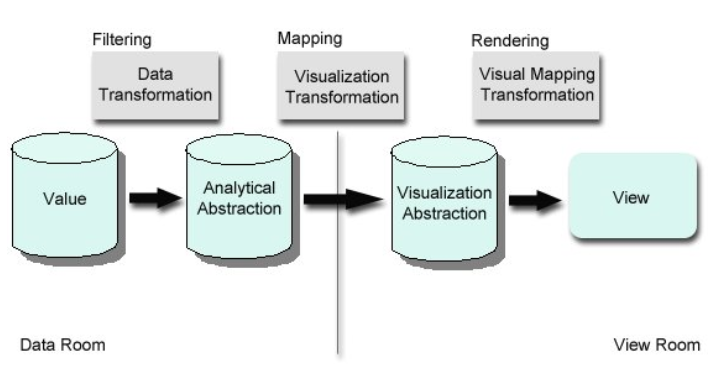
**Data Visualization**

* **Introduction**
* The ways we structure and visualize information are changing rapidly and getting more complex with each passing day. Thanks to the rise of social media, the ubiquity of mobile devices, and service digitization, data is available on any human activity that utilizes technology. The generated information is hugely valuable and makes it possible to analyse trends and patterns, and to use big data to draw connections between events. Thus, data visualization can be an effective mechanism for presenting the end user with understandable information in real time.
* Over the past three decades, computer graphics methods have fuelled a growing understanding of physical phenomena, which in turn have helped scientists and engineers substantially improve the quality of life.
* Today the relatively new field of 3-D scientific visualization has made a major impact on the display of behaviour. Originally developed to address large-scale visualization needs such as medical imaging, real-time data reduction, and satellite imagery, the techniques of scientific visualization are now being applied to numerical analysis and simulation to understand complex volumetric, multidimensional or time-dependent behaviour.
* Computer Visualization represents a single, unified collection of computer graphics techniques for the scientific visualization of behaviour. It is a reference work for professionals who use computing techniques to visualize data, and for the academic and software communities who support them.
* Visualization is defined in the dictionary as “a mental image.” In the fields of computer graphics and engineering design the term has a much more specific meaning: The technical specialty of visualization concerns itself with the display of behaviour, and, particularly, with making complex states of behavior comprehensible to the human eye.



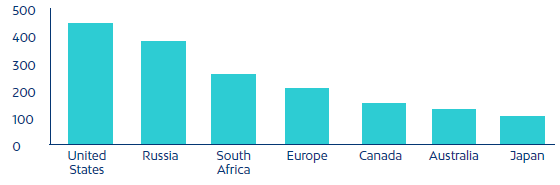
*Figure 1: Visualization system*

* 1. **What is data visualization?**
* Data visualization is the process of acquiring, interpreting and comparing data in order to clearly communicate complex ideas, thereby facilitating the identification and analysis of meaningful patterns.
* Data visualization can be essential to strategic communication: it helps us interpret available data; detect patterns, trends, and anomalies; make decisions; and analyze inherent processes. All told, it can have a powerful impact on the business world.
* **The data visualization process**
* **Filtering & processing**. Refining and cleaning data to convert it into information through analysis, interpretation, contextualization, comparison, and research.
* **Translation & visual representation**. Shaping the visual representation by defining graphic resources, language, context, and the tone of the representation, all of which are adapted for the recipient.
* **Perception & interpretation**. Finally, the visualization becomes effective when it has a perceptive impact on the construction of knowledge.



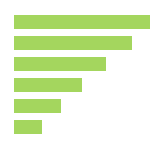
*Figure 2: Data visualization process*

* ***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.
  1. **Data visualization chiefly helps in 3 key aspects of reports and statements:**
* **1) Explaining**
* Visuals aim to lead the viewer down a path in order to describe situations, answer questions, support decisions, communicate information, or solve specific problems. When you attempt to explain something through data visualization, you start with a question, which interacts with the data set in such a way that enables viewers to make a decision and, subsequently, answer the question. For example: This graphic below could clearly explain the country with the greatest demand for a certain product compared globally, in a concrete month.



*Figure 3: Explaining visualization example*

* **2) Exploring**
* Some visuals are designed to lend a data set spatial dimensions, or to offer numerous subsets of data in order to raise questions, find answers, and discover opportunities. When the goal of a visual is to explore, the viewers start by familiarizing themselves with the dataset, then identifying an area of interest, asking questions, exploring, and finding several solutions or answers.
* **3) Analysing**
* Other visuals prompt viewers to inspect, distill, and transform the most significant information in a data set so that they can discover something new or predict upcoming situations.
* **Data relationships**
* Data relationships can be simple, like the progress of a single metric over time (such as visits to a blog over the course of 30 days or the number of users on a social network), or they can be complex, precisely comparing relationships, revealing structure, and extracting patterns from data. ***There are seven data relationships to consider:***
* **Ranking**: A visualization that relates two or more values with respect to a relative magnitude. For example: a company’s most sold products.



*Figure 4: ranking example*

* **Deviation**: Examines how each data point relates to the others and, particularly, to what point its value differs from the average. For example: the line of deviation for tickets to an amusement park sold on a rainy versus a normal day.



*Figure 5:deviation charts*

* **Series over time**: Here we can trace the changes in the values of a constant metric over the course of time. For example: monthly sales of a product over the course of two years.



*Figure 6: Series over time chart*

* **Distribution**: Visualization that shows the distribution of data spatially, often around a central value. For example: the heights of players on a basketball team.



*Figure 7: Distribution chart*

* **Nominal comparisons**: Visualizations that compare quantitative values from different subcategories. For example: product prices in various supermarkets



*Figure 8: Nominal comparison chart*

* **Correlation**: Data with two or more variables that can demonstrate a positive or negative correlation with one another. For example: salaries based on level of education



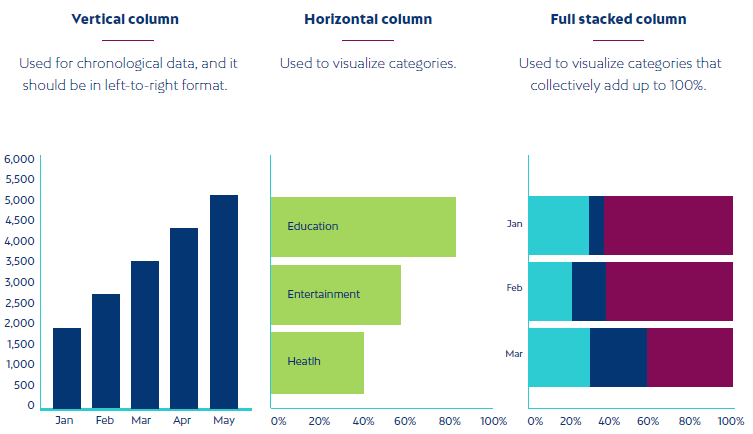
*Figure 9: correlation chart*

* **Partial and total relationships**: Show a subset of data as compared with a larger total. For example: the percentage of clients that buy specific products.



*Figure 10: Partial and total relationship chart*

* 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.
* **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.

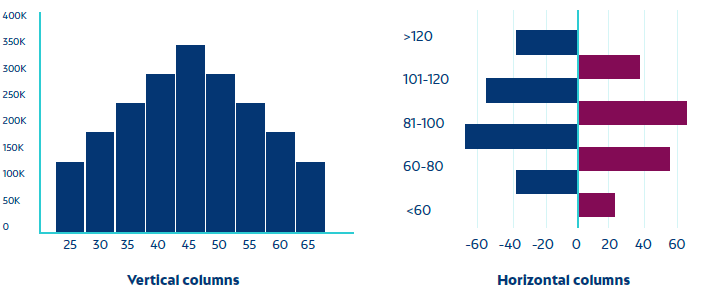


*Figure 11: Bar chart*

* **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

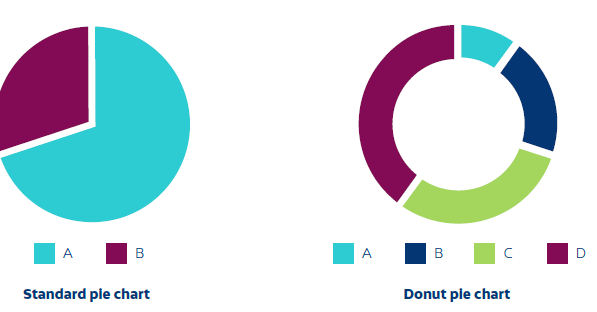


*Figure 12: Histogram chart*

* **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:

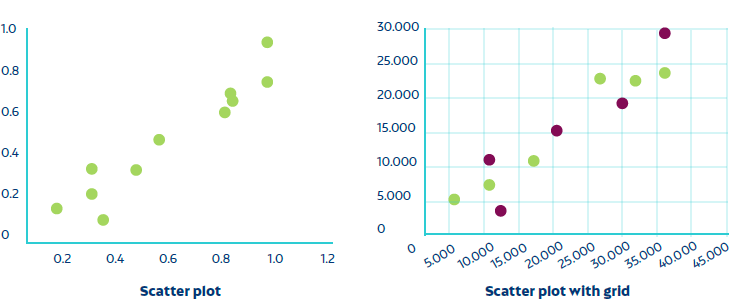
• ***Standard***: Used to exhibit relationships between parts.

• ***Donut***: A stylistic variation that facilitates the inclusion of a total value or a design element in the center.



*Figure 13: Pie chart*

* **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.

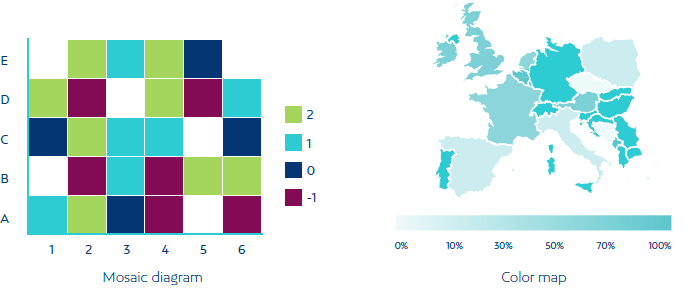


*Figure 14: Scatter plot chart*

* **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:

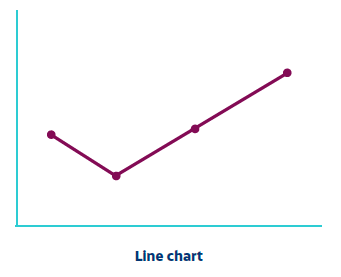
• Mosaic diagram

• Color map



*Figure 15: Heat maps chart*

* **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.



*Figure 16:Line chart*

* **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.



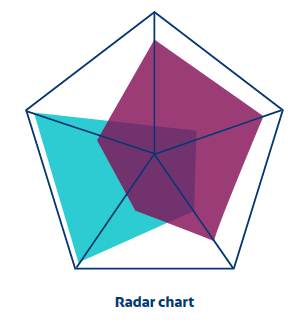
*Figure 17: Bubble plot chart*

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



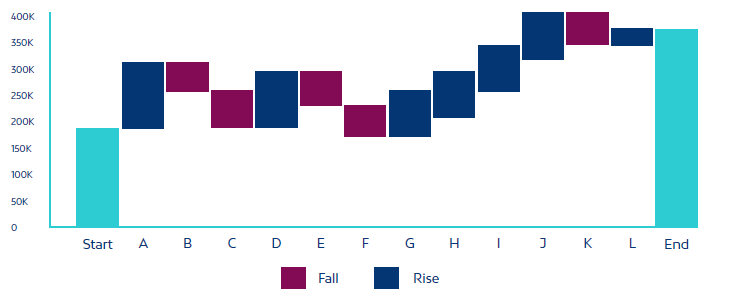
*Figure 18: Bubble map chart*

* **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.



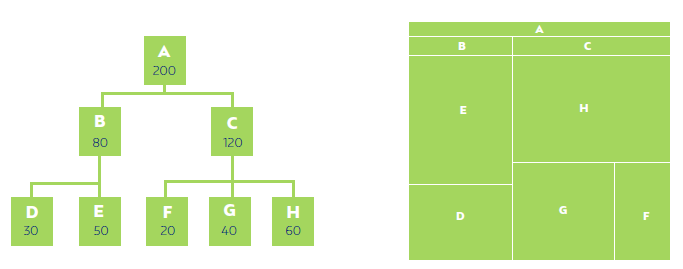
*Figure 19:Radar chart*

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



*Figure 20: Waterfall chart*

* **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.



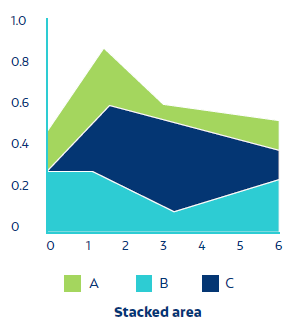
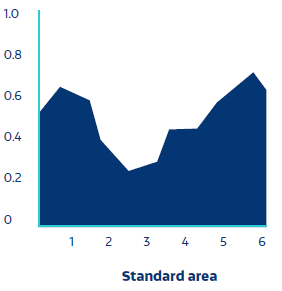
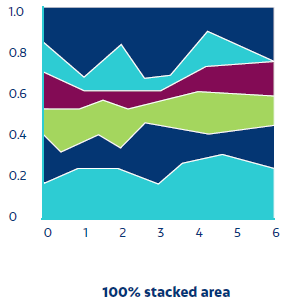
*Figure 21: Tree map*

* **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.

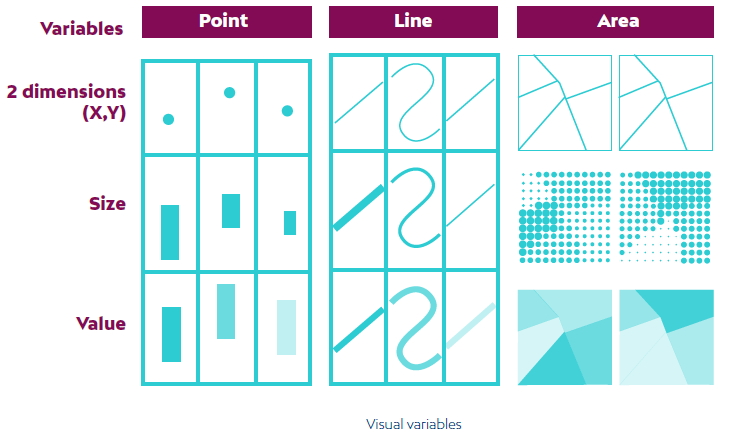
*Figure 22:Area chart*

* **Visual variables and their semantics**
* Visual variables are the building blocks of visual representation. They conform to an order and spatial context in order to convey a quantitative message. These resources can be used to categorize meaningful properties and amplify the message being represented. Let’s take a look at their semantics:

***Point***: Has no dimensions and indicates a place.

***Line***: Has one dimension and indicates length and direction.

***Plane***: Has two dimensions and indicates space and scale.



*Figure 23: Visual variables and their semantics*

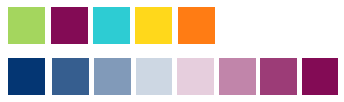
* **Using consistent and attractive color schemes**
* Color is one of the most powerful resources for data visualization, and it is essential if we are going to understand information properly.
* Color can be used to categorize elements, quantify or represent values, and communicate cultural attributes associated with a specific color.
* It dominates our perception and, in order to analyze it, we must first understand its three dimensions.
* **Hue**: this is what we normally imagine when we picture colors. There is no order to colors; they can only be distinguished by their characteristics (blue, red, yellow, etc.).
* **Brightness**: the color’s luminosity. This is a relative measure that describes the amount of light reflected by one object with respect to another. Brightness is measured on a scale, and we can talk about brighter and darker values of a single hue.
* **Saturation**: this refers to the intensity of a given color’s hue. It varies based on brightness. Darker colors are less saturated, and the less saturated a color is, the closer it gets to gray. In other words, it gets closer to a neutral (hueless) color.
* **three different kinds of color schemes**
* **Monochromatic sequential palettes or their analogue**
* These palettes are great for ordering numeric data that progresses from small to large. It is best to use brighter color gradients for low values and darker ones for higher values.



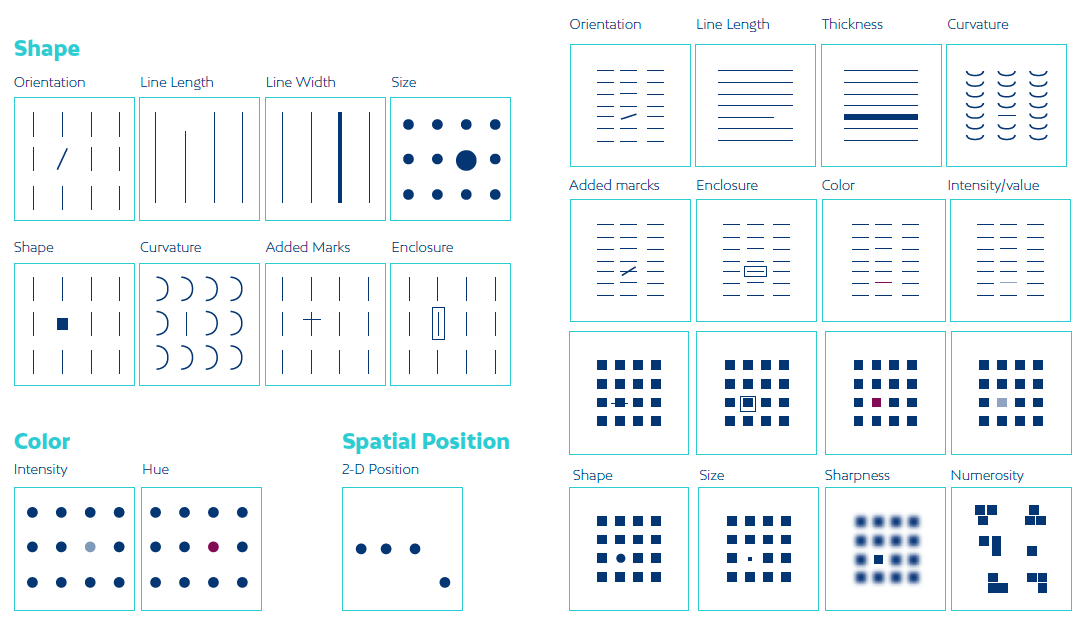
* Thus, brightness levels can be used as a visible, coherent aspect of a graphic scheme. Sequential color schemes make it possible to create a smooth, low-contrast design. This color scheme is better for an image than for data visualization
* **Diverging palettes**
* These are more suitable for ordering categorical data, and they are more effective when the categorical division is in the middle of the sequence. The change in brightness highlights a critical value in the data, such as the mean or median, or a zero. Colors become darker to represent differences in both directions, based on this meaningful value in the middle of the data.



* **Qualitative palettes**
* These are better for representing ordinal or categorical data to create primary visual differences between categories. Most qualitative schemes are based on differences in hue, with differences in brightness between the colors.

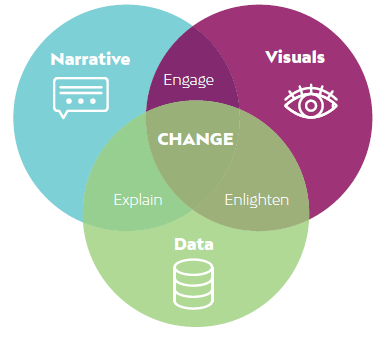


* Finally, don’t forget to use palettes that are comprehensible to people who can’t see color. Color blindness is a disability or limited ability that makes it difficult to distinguish certain pairs of colors, such as blue and yellow, or red and green. One strategy for avoiding this problem is to adapt designs that use more than just hue to codify information; create schemes that slightly vary another channel, such as brightness or saturation.
* **Prioritize patterns in your visualizations:**
* 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.
* The visual below lists preattentive attributes that represent aspects of lines and planes when visualizing and analyzing graphic representation: shape, color, and spatial position.



* **Data storytelling**
* We all love good stories, and data is one of the best tools for telling them. Millions of pieces of data are generated every day. They could be converted into great stories, but instead they are left unused. It’s time to change all that. It’s time to start telling stories that draw their power from data. So-called “data storytelling” is nothing more than placing a structured focus on the way we use data to communicate insights. It relies on three key elements:

narrative, visualization, and data.



*Figure 24: Data storytelling*

* **What do we get when we combine these elements?**
* **Data + Narrative**

Data can be insights; they are drawn from study and analysis. Their nature can propose the narrative context.

* **Visualization + Data**

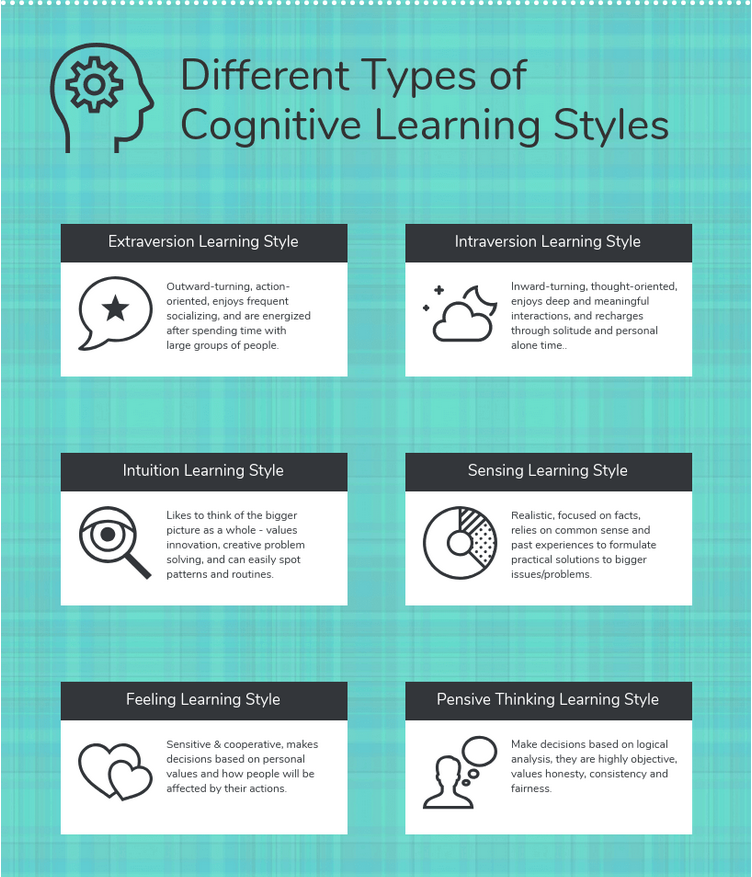
Visualization shines a light on our data by enabling us to rapidly process large volumes of data in a visual system.

As more data series are represented, we rely less on the verbal and more on the visual. Thus, we can enlighten our audience with insights that they may not have otherwise seen

* **Narrative + Visualization**

The story must motivate. It must have a plot, highs and lows, and an arc of emotional connection in order to draw in and entertain our audience.

* Data + Visualization + Narration = Successfully using our data to tell a story, wield influence, and effect the desired change.
  1. **Infographics**
* infographics (a clipped compound of "information" and "graphics") are graphic visual representations of information, [data](https://en.wikipedia.org/wiki/Data), or [knowledge](https://en.wikipedia.org/wiki/Knowledge) intended to present information quickly and clearly.
* An infographic is a collection of imagery, charts, and minimal text that gives an easy-to-understand overview of a topic.
* They can improve cognition by utilizing graphics to enhance the human visual system's ability to see patterns and trends. Similar pursuits are information visualization, data visualization, [statistical graphics](https://en.wikipedia.org/wiki/Statistical_graphics), information design, or [information architecture](https://en.wikipedia.org/wiki/Information_architecture).
* An infographic is a visual representation of a large collection of information. It tells a more comprehensive story than a single data visualization and often offers subjectivity. While an infographic can cover a data-heavy topic, it can just as easily explore a highly subjective one, such as the best beach vacations or most effective cleaning methods.
* Infographics are great for making complex information easy to digest. They can be helpful anytime you want:
* Provide a quick overview of a topic
* Explain a complex process
* Display research findings or survey data
* Summarize a long blog post or report
* Compare and contrast multiple options
* Raise awareness about an issue or cause
* An infographic can contain one or more data visualizations. These are often accompanied by short blurbs, quotes, or other pieces of text that elaborate on the topic or tie a collection of visualizations together.
* The primary purpose of an infographic is to explore a complex and comprehensive topic. While the infographic will usually offer several pieces of quantifiable and statistical data, it doesn’t always carry the viewer through to a single conclusion. Infographics often educate viewers on a topic so they can ultimately come to their own final conclusion about it.
* ***The three parts of all infographics are the visual, the content, and the knowledge.***
* The **visual** consists of colors and graphics. There are two different types of graphics – theme, and reference. Theme graphics are included in all infographics and represent the underlying visual representation of the data. Reference graphics are generally icons that can be used to point to certain ***data***, although they are not always found in infographics. Statistics and facts usually serve as the content for infographics and can be obtained from any number of sources, including census data and news reports. One of the most important aspects of infographics is that they contain some sort of insight into the data that they are presenting – this is the ***knowledge***.
* **Examples** 
  1. Small businesses and entrepreneurs use infographics to reach new audiences and increase brand awareness



* 1. Governments use infographics to share statistics and census data:



* 1. Educators and trainers use infographics to make content more memorable for students and employees:



* **What are the different types of infographics?**

1. Statistical infographics
2. Informational infographics
3. Timeline infographics
4. Process infographics
5. Geographic infographics
6. Comparison infographics
7. Hierarchical infographics
8. List infographics
9. Resume infographics

* **Informational infographics**: for example, are typically more text-heavy than the other types of infographics. Like this example from Course Hero below, they work best as summative, standalone pieces that provide a high-level explanation of a topic.
* **Statistical infographics**, on the other hand, are more focused on numbers, charts, and data If you want to visualize survey results, present data from multiple sources, or backup an argument with relevant data, then a statistical infographic is the best infographic to do that.
* [**Timeline infographics**](https://venngage.com/features/timeline-infographics) are the best type of infographic for visualizing the history of something, to highlight important dates, or to give an overview of events (for example, a project timeline).
* **Process infographic templates**: While a timeline infographic will highlight points in time, [a process infographic](https://venngage.com/features/process-infographics) is the best infographic for providing a summary or overview of the steps in a process.
* Process infographics will allow you to simplify and clarify each step. Most process infographics follow a straightforward top-to-bottom or left-to-right flow. Numbering the steps will make your process easy to follow.
* **Geographic infographic templates** Do you want to visualize location-based data, demographic data or large quantities of data? In those cases, a geographic infographic is your best bet. Geographic infographics use map charts as the focus visual. Different types of map charts work better for different types of data.
* **Comparison infographic templates:** A lot of people have trouble picking between multiple options. Whether you want to compare options in an unbiased way, or you want to make one option seem better, a [comparison infographic](https://venngage.com/features/comparison-infographic) if the best infographic to do that.
* **Hierarchical infographic templates**: A hierarchical infographic can organize information from greatest to least
* **List infographic templates**: If you want to share a collection of tips, or a list of resources, or a list of examples, then why not create…a list infographic!
* **Infographic resume templates:** With the current job market being so saturated, job seekers need to find creative ways to set themselves apart. That’s why [infographic resumes](https://venngage.com/features/resume-maker) have gained a lot of popularity in recent years. Infographic resumes won’t be able to entirely replace a traditional resume in most cases. But they’re a great visual document to bring to an interview, to publish on your portfolio site, or to include in an email application.

Fonts and Typography[[1]](#footnote-0)

**What is typography?**

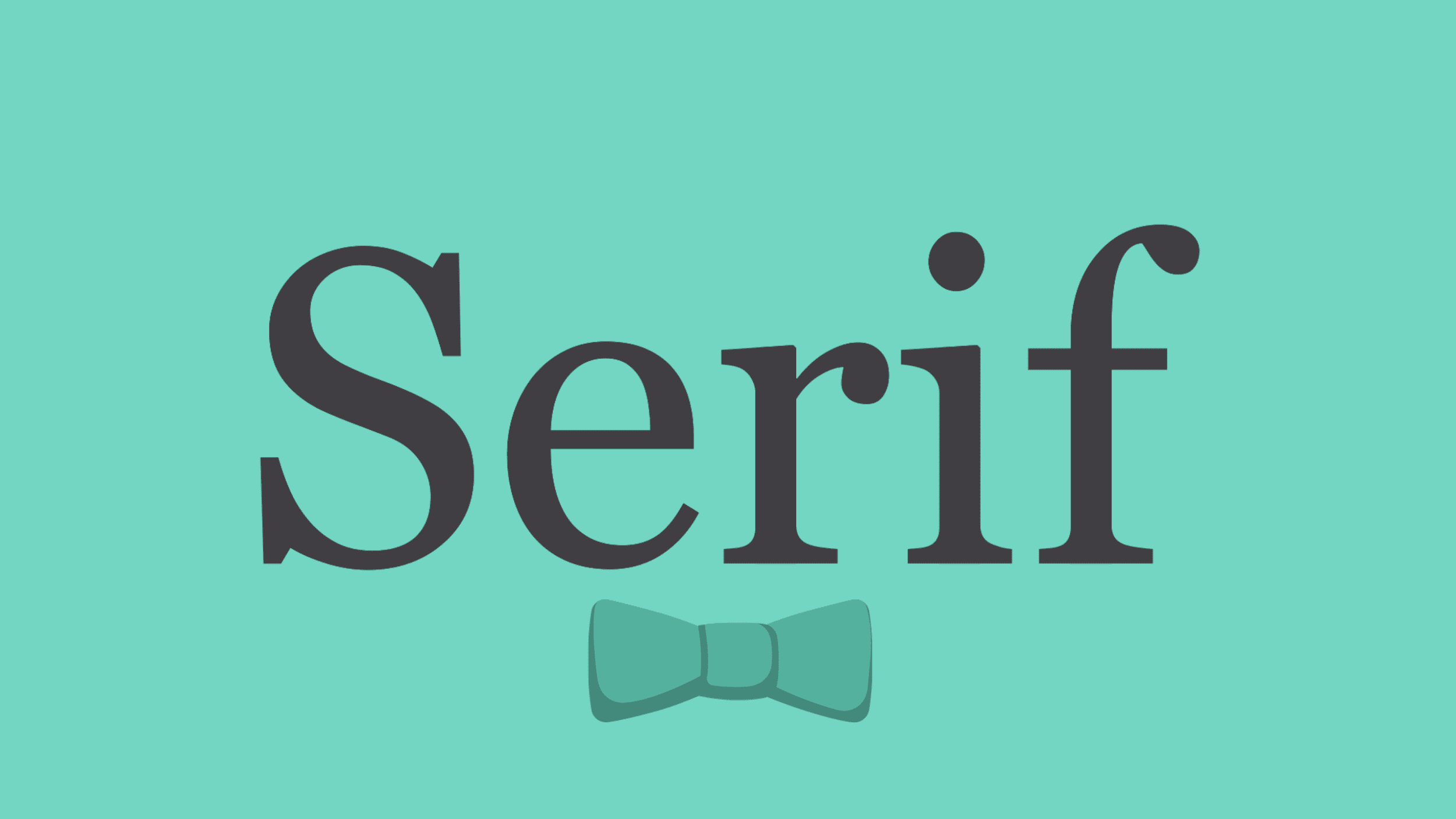
Typography is everywhere we look. It's in the books we read, on the websites we visit, even in everyday life—on street signs, bumper stickers, and product packaging.



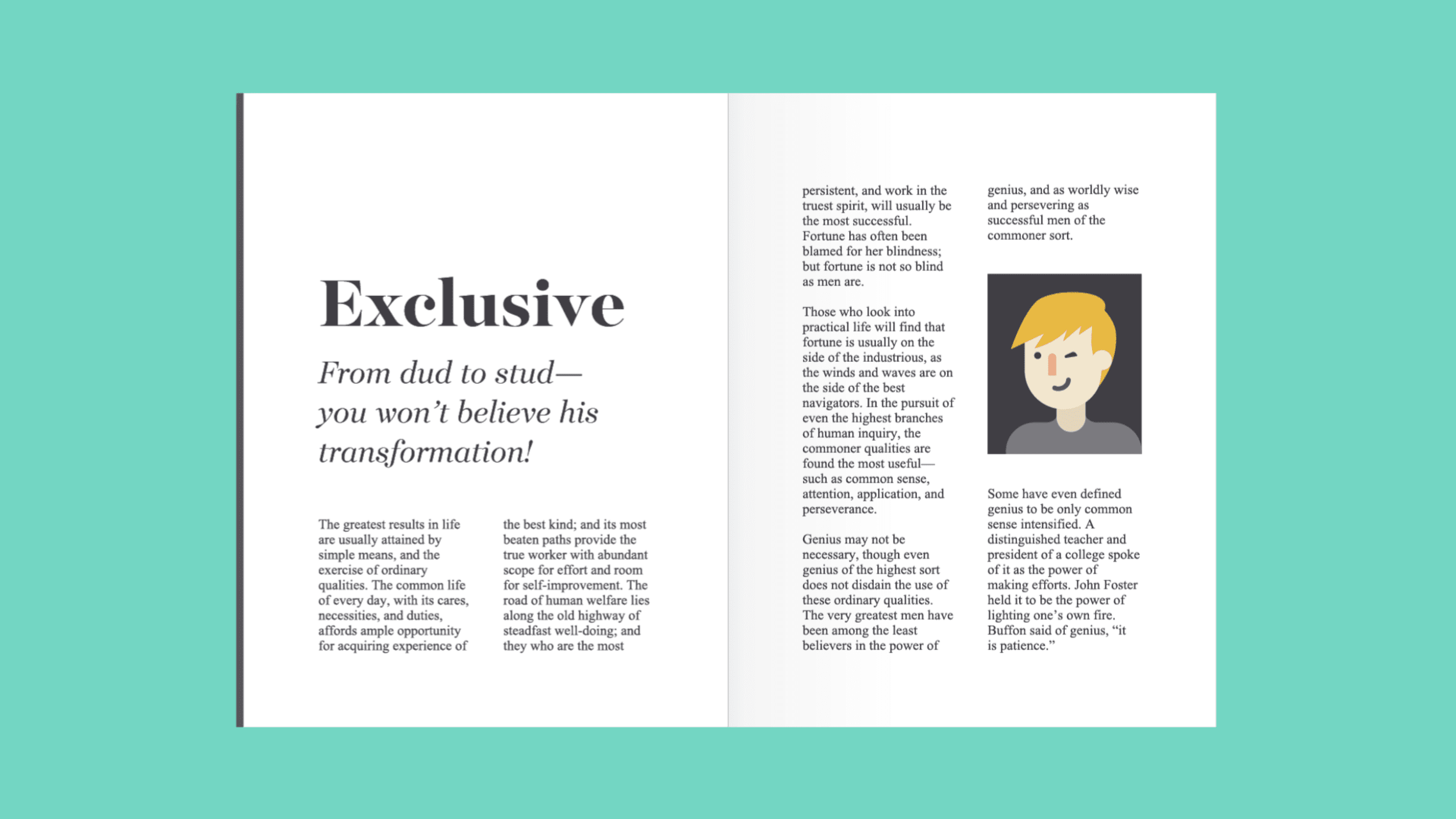
But what exactly is typography? Simply put, **typography is the style or appearance of text**. It can also refer to the art of working with text—something you probably do all the time if you create documents or other projects for work, school, or yourself.

**Example: Serif fonts**

Serif fonts have little strokes called **serifs** attached to the main part of the letter.



Because of their classic look, they're a good choice for more **traditional projects**. They're also common in print publications, like magazines and newspapers.



1. <https://edu.gcfglobal.org/en/beginning-graphic-design/typography/1/> [↑](#footnote-ref-0)