Consider New York City’s elaborate subway system. The New York City subway map organizes the complexity of this transit system to help passengers navigate from one location to another. The visualization removes unnecessary geographic information and adds information related to train schedules and transfers. The system is still difficult to traverse, but the map’s visual clarity makes it manageable. Maps are an early form of visualization—maps of the stars were created before recorded history—but they are only one among a myriad of techniques available to designers. Visualization also helps communicate abstract information and complex processes.
DATA INTO FORM

People have a remarkable ability to understand data when it's presented as an image. As researcher Stuart K. Card says, "To understand something is called 'seeing' it. We try to make our ideas 'clear,' to bring them into focus, to 'arrange' our thoughts." Like written words, visual language is composed to construct meaning. Our brains are wired to make sense of visual images. In contrast, it can take years of education to develop the ability to read even the simplest article in a newspaper.

The fundamentals of visual understanding, originally pursued by Gestalt psychologists in the early twentieth century, are now researched at a deeper level within the field of cognitive psychology. The findings of this research have been communicated within the visual arts by educators including Gyorgy Kepes, Donal A. Don, and Rudolf Arnheim, as well as through the work of visualization pioneers such as William Playfair, John Tukey, and Jacques Bertin. Data presentation techniques that combine our innate knowledge with learned skills make data easier to understand. In The Visual Display of Quantitative Information, Edward Tufte presents a data set and representation that supports this claim. Compare the tabular data to the scatterplot representation to see how the patterns become immediately clear when presented in the second format.

In his book Semiology of Graphics: Diagrams, Networks, Maps, Bertin presents another clear example of the communicative power of visual representation.

The maps of France on the left and right both present the same socio-geographic data, divided by canton (a French territorial subdivision). The representation on the right replaces each number with a circle sized to correspond to the numerical value. We can spend time analyzing the left map to see where there are concentrations of larger numbers, but on the right map we instantly comprehend the increased density in the upper left.

In the same book, Bertin introduces a series of variables that can be used to visually distinguish data elements: size, value, texture, color, orientation, and shape. For example, a bar chart distinguishes data through the height of each bar, and different train routes on a transit map are typically distinguished with color. For visualizations using only one variable, each element can be used in isolation. For multivariate visualizations (containing more than one variable) elements are combined. When applying form to data, there are always questions about goodness of fit, meaning how well the representation fits the data. Visualizations can mislead as well as enlighten. As Tufte warns in Visual Explorations: Images and Quantities, Evidence and Narrative, "There are right ways and wrong ways to show data; there are displays that reveal the truth and displays that do not." In Bertin's maps of France, the goodness of fit of the
representations reveals information hidden within the data. Because each piece of data derives from a particular canton, associating that data with its location on a map allows us to see regional patterns. Presenting the data in a table that is organized alphabetically would not reveal this pattern. By applying the same visualization technique to a different map—a map of Europe, for example—would also not work as well. The Berlin map works because each canton is roughly the same size, but the size differences among European countries would dilute the visual patterns needed for interpretation. In this case, the data is tightly linked to a source (geography), but in other instances data can be more abstract, such as when revealing patterns in language. The visualization techniques that follow in the next section present other options for revealing patterns in data.

There are hundreds of distinct visualization techniques that can be organized into categories, including tables, charts, diagrams, graphs, and maps. When creating a new visualization, one technique is selected instead of another based on the organization of the data and what the visualization is meant to convey. Data representations that commonly appear in newspapers, such as bar charts, pie charts, and line graphs, were all developed before people relied on software; in fact, most commonly used data representation techniques are only useful for representing simple data (1- and 2-D data sets). These techniques are automatic within frequently used software tools such as Microsoft Excel, Adobe Illustrator, and related programs. Visualizing information, once a specialized activity, is becoming a part of mass culture.

Writing new software is one approach to move beyond common data representations. New visualization techniques emerge as researchers and designers write software to fulfill their growing needs. The treemap technique is a good example to demonstrate the origins and evolution of a new visualization. It also shows how techniques often arise within a research group and are visually refined as designers use them in diverse contexts.

A treemap is a visualization that utilizes nested rectangles to show the relationships between one or more data elements. They are effective because they allow for easy 2-D size comparisons. The story of the first treemaps, from their origin to the present, is documented by the technique's originator, Ben Shneiderman, at the University of Maryland. The first treemaps were developed in 1990 as a way to show the memory usage on a computer hard drive. After subsequent applications and further development, the public at-large was introduced to treemaps by Map of the Market, an Internet application created by SmartMoney.com in 1998. This application introduced the innovation of making the tiles close to square, rather than using the thin tiles of previous treemaps, to increase legibility. The circular treemap technique explored by interface designer Kai Wetzel in 2003 pushed the form of treemaps even further. Wetzel worked on this representation as one of many ideas for a Linux operating system interface. He recognized that the approach wastes space and the algorithm is slower, but the aspect ratio of each node is the same. In 2004, a Newsmap application by Marcos Weskamp applied treemaps to the headlines of news articles collected from the Google News aggregator. The treemap representation makes it easy to see how many articles are published within each news category. For example, the visualization makes clear that, in England, the highest volume of published articles is world news rather than national stories, while in Italy, the reverse is true. By 2007, the refinement of these and other initiatives, the treemap technique had replaced the New York Times with the expectation that a general audience can understand it.

The era of modern data analysis began with the 1890 U.S. Census. The Census Bureau
Some early forms of data navigation might have included flipping through clay tablets, moving through a room painted with hieroglyphs, and rolling and unrolling a scroll. Early books improved upon scrolls because they allowed the reader to move quickly between sections in the text and could be smaller and therefore easier to carry. Book conventions such as the index, page numbers, and table of contents developed slowly. Despite thousands of years of refinement and the widespread proliferation of the internet, we’re still scrolling and viewing data on pages. The unique tool for looking at and navigating pages on the web is the hyperlink, a link from one page to another. Ted Nelson coined the phrase hypertext in the 1960s to describe this concept. Since that time, designers and researchers have pushed forward this and other innovative navigation concepts by writing software.

As an example, imagine the data inside a thesaurus. There’s a list of thousands of words in addition to all of the relations from each word to others. To explore this data, you look up one word, which you may then follow to another, and so on. Even the small amount of time needed to hunt for the next word can break the flow. The Visual Thesaurus software, written by Thinkmap, makes navigating language relations a more fluid experience. The software shows a network of words related to the currently selected word. Clicking on one of the outgoing words makes it the center, and new words appear that relate to it, while the former relations disappear. The interface allows the user to see the context around the current selection, but avoids overwhelming the senses with additional layers of non-relevant information.

Spatial navigation is an emerging technique for exploring data, but it has roots that are thousands of years old. The ancient memorization technique of loci, sometimes called a Memory Palace, places information inside imagined rooms within the mind to enhance recall by associating data with a mentally navigable space.

The sci-fi novels of William Gibson introduced intriguing concepts for spatial data navigation. In Neuromancer, published in 1984, he wrote about “rich fields of data” and described a vision of cyberspace: “A graphic representation of data abstracted from the banks of every computer in the human system.” Although Gibson’s world was fictional, a related real-world concept was developed by designer Lisa Strausfeld in 1995. She describes her software, Financial Viewpoints, as follows:

Imagine yourself without size or weight. You are in a zero-gravity space and you see an object in the distance. As you fly towards it, you are able to recognize the object as a financial portfolio. From this distance the form of the object conveys that the portfolio is doing well. You move closer. As you near the object, you pass through an atmosphere of information about net assets and overall return statistics. You continue moving closer. Suddenly you stop and look around. The financial portfolio is no longer an object, but that you now inhabit. Information surrounds you. At that time, Strausfeld was a research assistant in the Visible Language Workshop (VLT) at the MIT Media Lab. The research group was directed by Muriel Cooper, who set out to discover what graphic design could mean in the new era of communications, through the use of software applications. David Smail, another researcher in the group, used software to present large bodies of text within a single navigable environment. His Virtual Shakespeare presents the entire works of William Shakespeare within one continuously navigable space. From the long view, only the names of individual plays, such as Hamlet and Henry V, are visible, but as you zoom closer, the acts come into view as rectangular textures, and finally it’s possible to read the dialog and stage directions.
A time-series visualization shows data collected over a long period within a single image. It compresses many moments into a single frame. A time-series image can be a single, static image or it can be an animated image that combines data through motion. By using time as the ordering principle, changes become clearer.
the plane is and where it was, therefore implying its destination. The given insight into the nature of the inevitable high-ware far above the ground. Population therapist are visible through the densities of patterns.
As social, political, and technical networks become denser and more complex, widespread interest in visualization is growing. Provocative visualizations help us to better understand the sometimes invisible relationships that affect our world. Network diagrams frequently include two types of elements: nodes and connections. A node is an individual element (a person, country, or computer) and connections show relationships between the nodes. Visualizations help us to see different types of networks: centralized (star), decentralized (hybrid), and distributed (grid or mesh). These different organizations were elegantly diagrammed in 1962 by Paul Baran, one of the conceptual architects of the Internet.
Most maps show many layers of information within a single surface. For example, a single map might show the locations of roads, landmarks, topography, and political borders. Because the sophisticated language and representation of maps are so familiar, they provide a good foundation for additional layers of information. Adding changes in time and geometric distortion are effective ways to push the conventions further.
Images were created to think about mathematics long before computers were invented to calculate and visualize. For instance, Euclid (circa 300 BCE) constructed diagrams to show relationships between geometric elements and physical models. A paper model of a Möbius strip and a glass model of a Klein bottle can make these intriguing surfaces approachable to a wider audience, beyond those who understand the equations behind them. Before the era of the personal computer, the Mathematica exhibition, created by the Eames Office in 1961, presented diagrams, objects, and machines to demystify basic mathematical principles. Mathematica is also the name of a powerful program used within the sciences for calculations and visualizations. This program, along with other software development projects, cleared the path to new categories of mathematics visualization, including the popular fractal images of the Mandelbrot set. Mathematicians, artists, and architects are actively mining the structures of numbers and equations to produce visual images for pleasure and insight.