Data visualization is the representation of data through use of common graphics, such as charts, plots, infographics, and even animations. These visual displays of information communicate complex data relationships and data-driven insights in a way that is easy to understand.

Data visualization can be utilized for a variety of purposes, and it's important to note that is not only reserved for use by data teams. Management also leverages it to convey organizational structure and hierarchy while data analysts and data scientists use it to discover and explain patterns and trends. Harvard Business Review (link resides outside IBM) categorizes data visualization into four key purposes: idea generation, idea illustration, visual discovery, and everyday dataviz. We'll delve deeper into these below:

Idea generation

Data visualization is commonly used to spur idea generation across teams. They are frequently leveraged during brainstorming or Design Thinking sessions at the start of a project by supporting the collection of different perspectives and highlighting the common concerns of the collective. While these visualizations are usually unpolished and unrefined, they help set the foundation within the project to ensure that the team is aligned on the problem that they're looking to address for key stakeholders.

Idea illustration

Data visualization for idea illustration assists in conveying an idea, such as a tactic or process. It is commonly used in learning settings, such as tutorials, certification courses, centers of excellence, but it can also be used to represent organization structures or processes, facilitating communication between the right individuals for specific tasks. Project managers frequently use Gantt charts and waterfall charts to illustrate workflows. Data modeling also uses abstraction to represent and better understand data flow within an enterprise's information system, making it easier for developers, business analysts, data architects, and others to understand the relationships in a database or data warehouse.

Data visualization

Data visualization is a critical step in the data science process, helping teams and individuals convey data more effectively to colleagues and decision makers. Teams that manage reporting systems typically leverage defined template views to monitor performance. However, data visualization isn't limited to performance dashboards. For example, while text mining an analyst may use a word cloud to to capture key concepts, trends, and hidden relationships within this unstructured data. Alternatively, they may utilize a graph structure to illustrate relationships between entities in a knowledge graph. There are a number of ways to represent different types of data, and it's important to remember that it is a skillset that should extend beyond your core analytics team.

Types of data visualizations

The earliest form of data visualization can be traced back the Egyptians in the pre-17th century, largely used to assist in navigation. As time progressed, people leveraged data visualizations for broader applications, such as in economic, social, health disciplines. Perhaps most notably, Edward Tufte published The Visual Display of Quantitative Information (link resides outside IBM), which illustrated that individuals could utilize data visualization to present data in a more effective manner. His book continues to stand the test of time, especially as companies turn to dashboards to report their performance metrics in real-time. Dashboards are effective data visualization tools for tracking and visualizing data from multiple data sources, providing visibility into the effects of specific behaviors by a team or an adjacent one on performance.

- **Tables**: This consists of rows and columns used to compare variables. Tables can show a great deal of information in a structured way, but they can also overwhelm users that are simply looking for high-level trends.

- **Pie charts and stacked bar charts:** These graphs are divided into sections that represent parts of a whole. They provide a simple way to organize data and compare the size of each component to one other.

- **Line charts and area charts:** These visuals show change in one or more quantities by plotting a series of data points over time and are frequently used within predictive analytics. Line graphs utilize lines to demonstrate these changes while area charts connect data points with line segments, stacking variables on top of one another and using color to distinguish between variables.

- **Scatter plots**: These visuals are beneficial in reveling the relationship between two variables, and they are commonly used within regression data analysis. However, these can sometimes be confused with bubble charts, which are used to visualize three variables via the x-axis, the y-axis, and the size of the bubble.

- **Heat maps:** These graphical representation displays are helpful in visualizing behavioral data by location. This can be a location on a map, or even a webpage.

- **Tree maps**, which display hierarchical data as a set of nested shapes, typically rectangles. Treemaps are great for comparing the proportions between categories via their area size.

Data visualization best practices

With so many data visualization tools readily available, there has also been a rise in ineffective information visualization. Visual communication should be simple and deliberate to ensure that your data visualization helps your target audience arrive at your intended insight or conclusion. The following best practices can help ensure your data visualization is useful and clear:

Set the context: It's important to provide general background information to ground the audience around why this particular data point is important. For example, if e-mail open rates were underperforming, we may want to illustrate how a company's open rate compares to the overall industry, demonstrating that the company has a problem within this marketing channel. To drive an action, the audience needs to understand how current performance compares to something tangible, like a goal, benchmark, or other key performance indicators (KPIs).

Data visualization and IBM

While there are various data visualization tools on the market, Cognos Analytics is IBM's business intelligence and data visualization tool. Cognos Analytics's self-service platform integrates cognitive computing technology, including artificial intelligence and machine learning, to make it easy for organizations to visualize data, share new insights, and encourage data-driven decision-making. To learn how to visualize your data with Cognos and other every day tools, like Excel, please sign-up for our "Data Visualization and Dashboards with Excel and Cognos" course on Coursera (link resides outside IBM).

IBM Planning Analytics is IBM's artificial intelligence-infused integrated planning solution that automates planning, forecasting, and budgeting. By accelerating processes and obtaining more reliable results, Planning Analytics powers more intelligent workflows that drive greater accuracy and efficiency. IBM Planning Analytics is built on IBM's powerful calculation engine TM1, that allows businesses to harness data to inform the best possible business decisions.

IBM Watson Studio provides the environment and tools to help businesses solve problems by collaboratively working with data. Businesses can choose the tools they need to analyze, visualize, cleanse, and shape data and to create and train machine learning models.

Benefits of data visualization

Get accurate big data visualization without IT involvement, without a visualization specialist and the wait.