Data Visualization

Idea In Short

Data Visualization is the most elegant way to bridge the gap between data scientists and end-users. Lately, I have been reading my some papers on good visualization techniques and here is a list of must-reads I feel that every data scientist should review.

Data Visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, you can take the concept a step further by using technology to drill down into charts and graphs for more detail, interactively changing what data you see and how it's processed.

History & today

The concept of using pictures to understand data has been around for centuries, from maps and graphs in the 17th century to the invention of the pie chart in the early 1800s. It's technology, however, that truly lit the fire under Data Visualization. Computers made it possible to process large amounts of data at lightning-fast speeds. Today, Data Visualization has become a rapidly evolving blend of science and art that is certain to change the corporate landscape over the next few years.

Importance Of Data Visualization

Data Visualization is important because it allows trends and patterns to be more easily seen. With the rise of Big Data, we need to be able to interpret increasingly larger amounts of data. Machine Learning makes it easier to analyse data. Subsequently, Data Visualization is important, not only for data scientists, but also for professionals in any career, including those that work in professional services. Whether you work in finance, marketing, tech, design, or anything else, you need to visualize data. According to Aberdeen Group1, managers who utilize data visualizations are 28% more likely to find relevant information

compared with those who use managed dashboards and reporting tools. They also discovered that 48% of those who use data visualizations are able to find the information they want without the help of tech support. In short, visualizing data provides key stakeholders with the ability to easily connect the dots among multi-dimensional data sets. According to the HBR article2 - Visualizations That Really Work:

> Complex systems—business process workflows, for example, or the way customers move through a store—are hard to understand, much less fix, if you can't first see them.

In general, a visual summary of information makes it easier to identify patterns and trends than looking through thousands of rows on a spreadsheet. It's the way the human brain works! Since the purpose of data analysis is to gain insights, data is much more valuable when it is visualized. Even if a data analyst can pull insights from data, it will be more difficult to communicate the meaning without visualization. Charts and graphs make communicating data findings easier even if you can identify the patterns without them.

Engagement

Today, in the information age, we are all overwhelmed with information from various channels & media. The content to assimmilate is never-ending. Hence, the more we can be enticed & engaged with content in a way that allows us to learn & understand at a glance, the more likely we are able to make decisions about what information is worth our time for a deeper dive. Thus, Data Visualization is about illuminating & providing actionable knowledge, so we can spend more time doing analysis that's worth doing. In other words, accurate information used in data visualization tools fuel better decision-making in an organization. Furthermore, data consolidation is another advantage of data visualization. Instead of going through multiple pages of texts, data visualization consolidates all the data and presents in the form of one or more visuals. This can help you save a lot of precious time that can be used to put this data to good use.

Expert opinions

With time, we faced a new challenge – surviving what Richard Saul Wurman called the tsunami of bits of information that was cresting the horizon. He also introduced the concept

of the black hole between data and knowledge3. The more knowledge we discover, the wider the gap between what we know & what we think we should know. Alberto Cairo, a Spanish information designer, professor, and author 4uses the term wisdom as the final point of the transformation of reality into knowledge:

> We reach wisdom when we achieve a deep understanding of acquired knowledge, when we not only get it," but when new information blends with prior experience so completely that it makes us better at knowing what to do in other situations, even if they are only loosely related to the information from which our original knowledge came

The concept of the citizen data scientist is on the rise5. In fact, skill sets are changing to accommodate a data-driven world. In any case, it is increasingly valuable for professionals to be able to leverage data to make decisions & use visuals to tell stories of when data informs the who, what, when, where, and how. Traditionally, education drew a distinct line between creative storytelling & technical analysis. However, the modern business world emphasizes skills that cross between the two. That is to say, data visualization sits right in the middle of analysis & visual storytelling. Furthermore, Data Visualization tools, such as Tableau, Microsoft Power BI, Qlik, etc. have played an important part in democratizing data & analytics. To that end, the adoption of these tools across organizations is making data-driven insights available to employees throughout an organization.

Science behind Data Visualization

To understand the science behind data visualization, we must first discuss how humans gather and process information. In collaboration with Amos Tversky, Daniel Kahneman did extensive research on how we form thoughts, and concluded that we use one of two methods:

System I:

Describes thought-processing that is fast, automatic, and unconscious. We use this method quite frequently in our everyday lives and can accomplish the following:

- Read text on a sign

- Determine where the source of a sound is
- Solve 1+1
- Recognize the difference between colors
- Ride a bike

System II

Describes a slow, logical, infrequent, and calculating thought and includes:

- Distinguish the difference in meaning behind multiple signs side-by-side
- Recite your phone number
- Understand complex social cues
- Solve 23x21

Through these two systems, Kahneman explains why humans struggle to think in terms of statistics. For instance, System I thinking is based on heuristics and biases to handle the volume of stimuli we encounter daily. Therefore, we should seek to ensure that data is presented in a way that correctly communicates to our System I thought process. As a result, our System II thought process can accurately analyze the data. Furthermore, our unconscious System I has the ability to process about 11 million pieces of information/second. However, our conscious System II can process only 40 pieces of information/second. In any case, our sub-conscious system processes more information through vision. Hence, Data Visualization is a perfect solution to communicate patterns and insights from data sets.

Recommended Publications

1. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. William S. Cleveland; Robert McGill
2. Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays. Christopher Ahlberg and Ben Shneiderman
3. High-Speed Visual Estimation Using Preattentive Processing. C. G. Healey, K. S. Booth and J. T. Enns
4. The Structure of the Information Visualization Design Space. Stuart K. Card and Jock Mackinlay
5. Automating the Design of Graphical Presentations of Relational Information.

Jock Mackinlay
6. How NOT to Lie with Visualization. Bernice E. Rogowitz, Lloyd A. Treinish
7. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. Ben Shneiderman

Recommended Books

With Data Visualization growing in interest & popularity, the number of resources available is huge. As time management is a vital skill for a consultant, I did the research for you. To save you some precious time, I went through a dozen serious websites listing the best books that I feel should help you get started. In short, here are the 10 most recommended books.

1. Storytelling with Data: A Data Visualization Guide for Business Professionals
2. The Truthful Art: Data, Charts, and Maps for Communication
3. Show Me the Numbers: Designing Tables and Graphs to Enlighten
4. Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations
5. Data Points: Visualization That Means Something
6. The Wall Street Journal Guide to Information Graphics: The Dos and Don'ts of Presenting Data, Facts, and Figures
7. Stories That Stick: How Storytelling Can Captivate Customers, Influence Audiences, and Transform Your Business
8. Presenting Data Effectively: Communicating Your Findings for Maximum Impact
9. Visualize This: The FlowingData Guide to Design, Visualization, and Statistics
10. Now You See It: Simple Visualization Techniques for Quantitative Analysis

Recommended Podcasts

If you're a consultant on the move, you may prefer to make the best use of your dead time & learn through podcasts. Please search for these podcasts in your favorite podcast application.

- Data Stories
- storytelling with data podcast
- The Present Beyond Measure Show: Data Visualization, Storytelling & Presentation for Digital Marketers

- Data Viz Today
- Data Skeptic

Summary

Recommended publications include works by Cleveland, Shneiderman, and Mackinlay covering graphical perception, preattentive processing, and visualization design space, providing a solid academic foundation for practitioners.