Visualization is based on exploiting the human visual system as a means of communication. Discuss about the above statement.

Exploring data visualization: how it works and its applications

Introduction

While data visualization might sound very much like a recent tech buzzword, it has a deep historical root in human history that dates back to the days as early as before the 17th century. Along with the progress of statistical theory, technological and graphical innovations, statistical visualization had evolved significantly and started to be adopted in many fields. However, it was only from the 1960s onwards, the popularization of computers had resulted in the wide adoption of computer-generated statistical graphics that has revolutionized the field of data visualization. [1] As it has become so convenient and efficient to produce computer generated data visualization, the field of data visualization bloomed.

Moreover, with the arrival of the age of data explosion, the field of data visualization develops at an unbelievable pace in both theory and practice. In this essay that discusses about the topic of data visualization, there are two main topics that I hope to delve into:

1. How does data visualization exploit the human visual system to augment our cognitive capabilities?

2. Types of data visualizations and their respective usage:
    - exploratory data visualization
    - presentation data visualization
    - interactive data visualization

How does data visualization exploit the human visual system to augment our cognitive capabilities?

Data Visualization and external cognition

“The power of the unaided mind is highly overrated. Without external aids, memory, thought, and reasoning are all constrained. But human intelligence is highly flexible and adaptable, superb at inventing procedures and objects that overcome its own limits. The real powers come from devising external aids: it is things that make us smart” (Norman, 1993, p.43) [2]

It seems that the core of human intelligence lies in our ability to devise external aids that augment our cognition: drawings, diagrams, symbols, languages, computers… Each and every single tool develops to serve its unique purposes that allow humans to push our cognitive boundaries beyond the limitations of our internal cognition significantly, which becomes an inseparable part of human brain power of today, known as external cognition.

External cognition is concerned with the interaction of cognitive representations and processes across the external/internal boundary in order to support thinking. (Kard and Mackinlay, 1999) [3]. In the context of data visualization, what we attempt to do is the abstraction of insights from a collection of observations or data in the form of visual representation which are usually descriptive, concentrating on 'raw' data and simple summaries.

The reason that data visualization is necessitated is due to the fact that our human brains are very incapable of processing long, mundane, oftentimes numerical data. However, humans are very good at detecting visual patterns quickly as one-third of the brain cortex neurons is devoted to **visual perception and pattern recognition**. [4] It is thus no wonder that we can make use of data visualization to exploit this neurological machinery to comprehend extremely complex data by the standard of our internal cognition. By transforming numbers into shapes and colors, the symbolic into the geometric, the data is translated into a form that is much more accessible and comprehensible to our very visual oriented brain.

Computer-generated data visualizations
Despite its importance and long-established history, data visualization remained very limited in its usage before the relatively recent development in computer-generated graphics as it was a tremendous pain in the neck to do so manually. With the advancement of computation power and increasing convenience in generating data visualization, it becomes increasingly convenient for us to visualize data that leads to the bloom of data visualization.

Types of data visualizations and their specific applications:

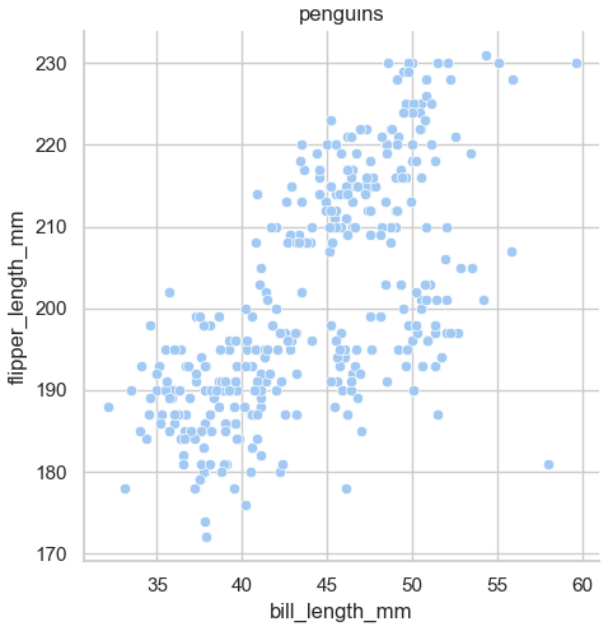
While data visualization appears very impressive on its theoretical level, how is it usually being used in our daily practice? By aggregating, modeling and translating the data into a data graphic, it reveals data features that we might not notice otherwise: unusual distributions of data, local patterns, clusterings, trends, outliers, and so on. Thanks to our visual perception and pattern recognition capabilities, our brain can make sense of complex data very intuitively and quickly through visualization, allowing us to dig into deeper understanding of the data and its insights.

Exploratory data visualization

One of the most common examples is its use in exploratory analysis, e.g.:

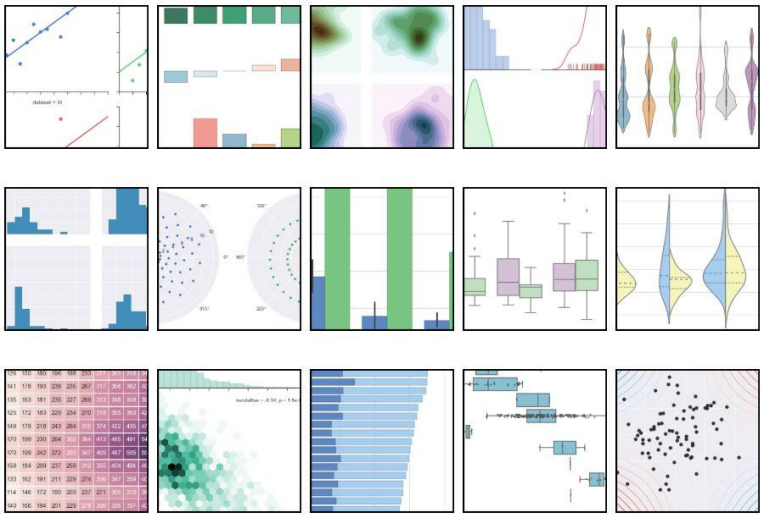
	species	island	bill_length_mm	bill_depth_mm	flipper_length_mm	body_mass_g	sex
0	Adelie	Torgersen	39.1	18.7	181.0	3750.0	Male
1	Adelie	Torgersen	39.5	17.4	186.0	3800.0	Female
2	Adelie	Torgersen	40.3	18.0	195.0	3250.0	Female
3	Adelie	Torgersen	NaN	NaN	NaN	NaN	NaN
4	Adelie	Torgersen	36.7	19.3	193.0	3450.0	Female

Above is the very commonly used “Penguins” dataframe head consisting of 5 of 344 observations in the dataset. While it is a very detailed, structured documentation, there is no way we can make sense of it in our brain.



But through visualization, we are able to abstract and make sense of certain attributes of the dataset at a glance. In this example, we plot the penguins’ flipper length against its bill length and get an understanding of their distribution and patterns. By visualizing all the attributes that we wish to study, our brain can create a comprehensible representation of the dataset by overlaying these data graphs on one another. As our brain is terrific in processing visual information and recognizing data, it is incredibly useful for us to understand the data, explore trends and patterns and generate interesting insights.

Figure below shows some examples of very commonly used exploratory visualization, e.g. bar chart, scatterplot, line graph, pie chart, Venn diagram…



(Source: Exploratory Data Analysis with Python - DataScienceCentral.com)

In exploratory data visualization, the main audience that we should really communicate to is ourselves. Instead of a beautiful graphic, we need many, many imperfect graphics that are solely for one purpose: provide alternative views and additional information for ourselves. The more the visualizations we have, the more accurately our brain can construct its understanding on the data as it has more cuts into the data that allows for a more comprehensive structure of understanding to be built.

This approach is only made available through the advancement of computed visualization technologies as it has made the slow process of drawing and redrawing data graphics become incredibly fast and flexible. Nonetheless, the progress is still far from its endgame. As I'm writing now, there is tons of research and development being done to improve our understanding on the theory of graphics, build new tools for more complex visualizations, optimize existing functions with higher standards and flexibility, etc.

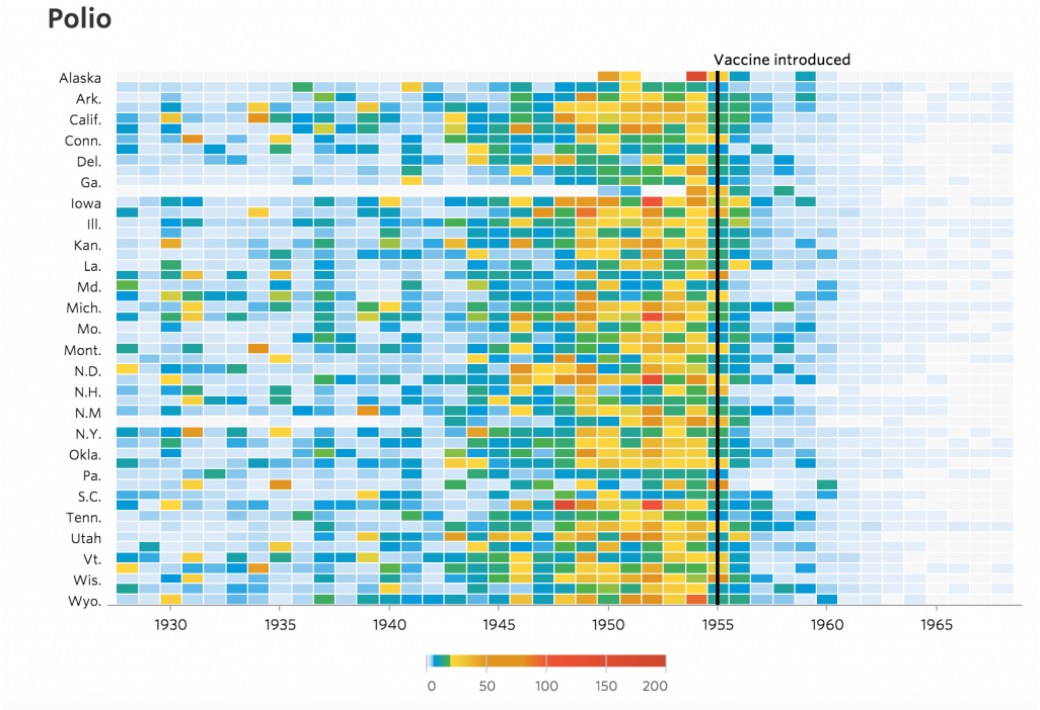
At its current stage, exploratory data visualization is very much used for "data cleaning, exploring data structure, detecting outliers and unusual groups, identifying trends and clusters, spotting local patterns, evaluating modeling output, and presenting results." [6]

Presentation data visualization

Meanwhile, in the realm of presentation data visualization, the graphic produced has a very different purpose from that of exploratory visualizations: to convey known information to the crowd with the greatest possible effectiveness.

Thus, it is a must for the graphic to be well-designed to attract the audience's attention and relay meaningful information concisely. There is a good old saying in the world of journalism, "a picture is worth a thousand words." It is a good benchmark to set for presentation data visualization too. There is no limit to the type of design needed for presentation data graphics.

Much like good story-telling, there is only one rule:engage your audience and get your message across. For instance:



(source: [Battling Infectious Diseases in the 20th Century: The Impact of Vaccines - WSJ.com])

Data graphic above conveys the impact of vaccines in the reduction of Polio very effectively. The audiences are able to capture the significant reduction of Polio cases after introduction of vaccination at a glance.

Interactive data visualization

With our increased ability in data mining, we are coming to deal with more and more complex datasets to derive insights. If these are represented by static visualizations, it would result in an impractical amount of figure to fully represent the complexity of the dataset, that will make navigation between these figures a disastrous experience. For instance, in their work of urban planning, the planners will have to open up maps of different levels: state, district and street levels and navigate through them back and forth, which is terribly inconvenient.

Nevertheless, the interface of this process can be improved indefinitely with the introduction of interactive data visualization. Its application helps by “organizing and visually displaying data in an easy to use interface, seamlessly providing navigation, selection, and representation of data.” [6] Moreover, compared with static information visualization technologies, interactive data visualization provides another competitive edge as it allows users to control which data to display, how to represent the data, or both.

One of the most readily available examples of interactive data visualization that we use frequently is google map. On google map, we are able to access to its geolocation data with

incredible flexibility in control, e.g. the location or coordinates (by identifying the location), the scale of map (by zooming in and out), the type of view (plan view, street view or 3D-view). It has been so convenient for us users that it has replaced most of the physical maps in the world, which shows how big a potential interactive data visualization may be.

Conclusion

In conclusion, data visualization has played an essential role in augmenting our cognition, allowing us to understand very complex data (by the standard of our internal cognition) efficiently. It is deeply rooted in our neurological development, as our brain has evolved to be terrific at visual perception and pattern recognition. There, we exploit this feature of our brain and develop strategies and techniques to allow us to surpass our internal cognition and get a grasp on our data more efficiently. In practice, there are 3 types of data visualization being deployed: exploratory data visualization, presentation data visualization and interactive data visualization. Each of which can be applied to their respective unique situations and help humans to learn better, understand more and achieve more by exploiting our versatile visual system.

References:

1. Friendly, M. (2008). A brief history of data visualization. In *Handbook of data visualization* (pp. 15-56). Springer, Berlin, Heidelberg.Available from: https://www.researchgate.net/publication/226400313_A_Brief_History_of_Data_Visualization [accessed Apr 22 2022].
2. Norman, D. (1994). Visual representations In Things That Make Us Smart: Defending Human Attributes in the Age of the Machine. Retrieved from: Visualization: A cognition amplifier (wiley.com)
3. Card, M. (1999). *Readings in information visualization: using vision to think*. Morgan Kaufmann. Available from: Readings in Information Visualization, Using vision to think (researchgate.net)
4. Valle, M. (2013). Visualization: A cognition amplifier. *International journal of quantum chemistry*, *113*(17), 2040-2052. Available from: Visualization: A cognition amplifier (wiley.com)
5. Unwin, A., (Feb 01, 2020), Why is Data Visualization Important? What is Important in Data Visualization?, *Harvard Data Science Review. Retrieved from URL:* Why is Data Visualization Important? What is Important in Data Visualization? · Issue 2.1, Winter 2020 (mit.edu)
6. Janvrin, D. J., Raschke, R. L., & Dilla, W. N. (2014). Making sense of complex data using interactive data visualization. *Journal of Accounting Education*, *32*(4), 31-48. Available from: Making sense of complex data using interactive data visualization - ScienceDirect