EFFECTIVE DATA VISUALISATION

By Shesh Murali

Introduction

Data visualisation is the graphical representation of data and information, it acts as a body of various methods for the exploration of quantitative datasets (Graham, 2017). By delving through data analysis, data visualisation tools can be applied to analyse and interpret various aspects of culture, behaviour and society. Gaining insights, identifying patterns and communicating findings allows one to highlight the significance and impact of the effect of research. By simplifying complex datasets into understandable and successful data visualisers helps find solutions to real world problems (Midway, 2020). There is a great importance in conveying complex information so that clarity, understanding and engagement is kept. These specific aspects, reduces the likelihood of misunderstanding so that the audience comprehends messages accurately, hence leading to better retention time. Retention time is dependent on the interactive elements reflected in data visualisers, by introducing features and other elements can engage viewers and improve comprehension (Bobek & Tversky, 2016). Therefore, directly impacting user experience and engagement, ultimately contributing to the efficiency of the research. In the field of data science leveraging visuals effectively, they can provide value to organisations. Data visualisations are categorised into several types based on the nature of how the data is represented and the goals achieved from visualisation; Static and interactive Statistical visualisations, Geospatial Visualisation, Hierarchical Visualisation and Network Visualisations (Fry, 2008). However, this study will delve into the importance of static and interactive statistical visualisations, understand its position when solving problems and discussion, the nature of data and collection methods, a case study into effective data visualisation through static and interactive statistical visualisations and finally an evaluation of good and bad design practices. Ultimately, by systematically exploring various aspects of effective data visualisation, perpetuates clear data analysis and communication, therefore further bolstering the practice of data science

Problem Identification and Discussion

Data visualisation, like said before, is a powerful tool for understanding and communicating data insights, however it comes with several sets of challenges and pitfalls. These problems arise mainly due to the interpretation and presentation of the data. Misleading visualisations can occur due to the irregular scaling. The non uniform and non linear scaling of data variables, truncating axes and selective data inclusion skew the perception of data. These practices omit crucial information ,exacerbate certain data points or even compress certain data at one end of the spectrum (Mößner, 2014). Hence, creating a false sense of significance or trends within the data, undermining the integrity and reliability of the visualisation. While on the other hand, overly complex visualisation can just as hinder users. Too many graphical elements can overwhelm users which makes it difficult to discern the intended patterns and insights. This creates cognitive overload, which in turn creates difficulties in processing information, identifying patterns and making connections, hindering the effectiveness of data visualisation to be used as a tool (The Psychology of Data Visualization: How Humans Perceive and Interpret Visual Data, n.d). Therefore, to combat these problems, there should be a great importance in choosing the appropriate visualisation technique based on the data and who the target audience is. The preferences and characteristics of the target audience differ with their knowledge on the subject matter, technical expertise and cognitive abilities influence their ability to interpret data visualisation (Few, 2004). For instance, if data is presented to the general public it might need to be accommodated by utilising simpler visuals such as clear labels and minimal complexity to facilitate comprehension. Contrastingly if the audience were other analysts and data scientists detailed annotation and interactive features that help explore the intricacies of the data. By doing so data visualisation plays a vital role across various disciplines by contributing to the advancement of how information is interpreted more efficiently (Data Visualisation – a Scientific Treatment, n.d.). Moreover, it enables researchers to communicate their results to a broader audience, such as policymakers and the public, hence bolstering transparency and engagement for the research. This enhances the relevance and the integrity within the research community contributing the relevance of research, to further aid the public in societal developments and policy making.

Nature of Data and Collection Methods

Static and interactive statistical visualisations are the graphical representation utilised to display data, these include charts, graphs, maps and infographics (Tableau, 2018). The data is typically in a fixed format with predetermined axes, visual elements and scales, curated and simplified to highlight specific trends and patterns. These static aspects of visualisation often convey key insights that are quick glances into the context of the research or report. While the dataset for the interactive visualisation may be more comprehensive, as it encompasses a broader range of variables and observations. Users more often interact with the visuals by altering parameters, filtering and drilling into specific subsets, thus enabling deeper analysis and insights. Static visuals enable audiences to make simple informed decisions based on the information presented, to counteract the lack of interactivity interactive visuals are paired with this for data exploration and hypothesis testing. Ultimately both foster empowerment, engaging users with the data are more personalised and interactive. The method of data collection for static and interactive statistical visualisation can depend on the nature of the data and dependent on the type of data. For instance, sensitive data may require static visualisations to maintain privacy and security, while interactive visualisations may be more suitable for exploring non-sensitive data to understand and uncover insights and patterns (Static vs Interactive Visualizations – How to Choose, n.d.). Through means of surveys, publicly available datasets or even observational studies collect the data for analysis to cover a wide range of topics and domains (Young et al., 2011). Additionally, real-time data streams like live databases, sensors and IoT devices collect data from physical environments such as temperature or motion and are transmitted visually for analysis and live databases provide up to date information from online sources. All enable researchers to interactively explore data trends and insights.

Case Study: Effective Data visualisation Part I

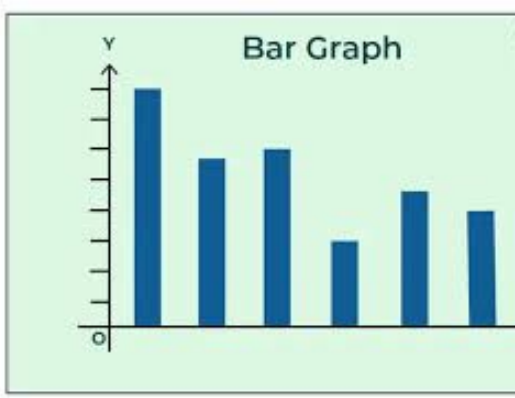
A case study is required to delve into the notable examples of data visualisation to exemplify the principles. The first is an example of an interactive statistical visualisation is a dynamic dashboard. A dashboard that visualises sales data for a retail company. Users interactively use it to compare sales performances over regions and categories, explore sales trends over time and also drill down into specific subsets of the data, an example of that could be a sale by customer segment

Possible interactive features could be

- **FIltering**: Users can apply filters to focus on specific product categories, time periods or geographical regions of interests
- **Interactive charts**: Users can click on elements within charts to filter or highlight related data
- **Dynamic cross-filtering**: Selecting data in one chart automatically updates related charts, thus allowing for seamless exploration of datasets
- **Data aggregation**: Users can aggregate data at different levels of time, like daily, monthly or weekly

This visualisation encourages active engagement from the audience. allowing to explore the data and interaction with different elements with the visual. Hands on approach promotes customization and a deeper understanding, empowering each audience member to extract the relevant insights. Additionally, dynamic feedback facilities hypothesis testing through instant feedback and analysis. Overall, the interactive leveraged user engagement and bolstered these techniques to effectively communicate insights and facilitate interaction with the data and audience

Case Study Part II



An example of static statistical visualisation would be a bar chart representing the disruption of sales separated by product category for a specific time period. This visualisation, each products category is represented by a bar, with the y axis or height of the bar corresponding with the total sales generated by the customers

Possible features include

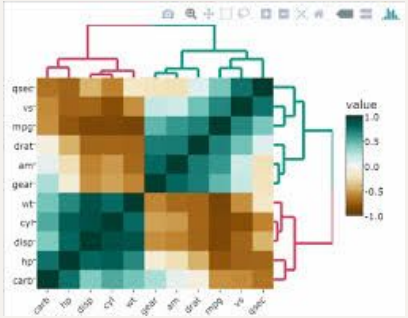
- **Fixed layout**: The layout is predetermined and provide clarity making them easier to understand
- **Fixed data**: The data displayed is static and does not update hence the can be easily shared and distributed through pdfs or images making them easily accessible
- **Labels and annotations**: Chart can include labels for product categories and values indicating the sales for each category, thus staying consistent and maintains their appearance and layout

Status visualsions encourage clear presentation where the things being presented are straightforward and easily understood, allowing the audience to quickly grasp. Additionally, with a focused message it communicates the important findings concisely. Consistency and accessibility helps reach the data to a larger audience through giftable images and different viewing platforms and reduces the risk for misinterpretations. Ultimately, while it may be less dynamic the simplicity is a valuable tool. But, both these types of graphs paired together can enrich data exploration and communication providing users with complementary insights that enhance the overall research.

Evaluation of Good and Bad Design Practices Part I

For the purpose of further understanding, a comparative analysis has to be conducted with a bad and food design practice with interactive and static statistical visualisations

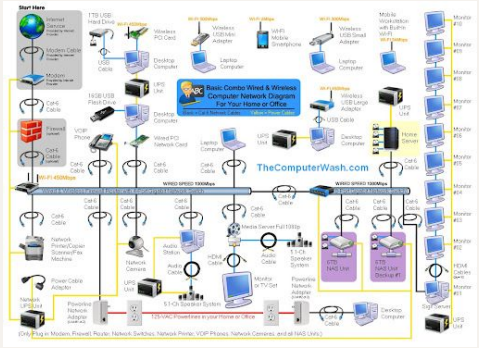
Good Interactive Statistical Visualisation



Interactive Heatmap

- **Design**: The heat map efficiently uses colour gradient to represent varying levels of metrics across different categories or time periods
- **Interactivity**: Users can hover over cells to see specific values and click legend time to increase visibility
- **Usability**: Intuitive controls and clear labelling allowing for exploration and identifying trends
- **Performance**: The visualisation is responsive and loads quickly

Bad Interactive Statistical Visualisation

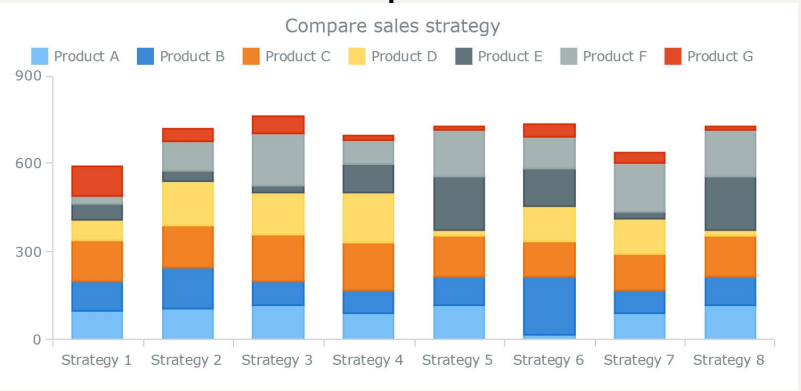


Overly Complex Network Diagram

- **Design**: Network Diagram too cluttered with many nodes and edges hard to discern meaningful patterns
- **Interactivity**: Users can zoom and navigate however the lack of filtering overwhelms users though
- **Usability**: Users struggle to interpret the visualisation due to increased complexity
- **Performance**: The visualisation is not that responsive due to large datasets.

Evaluation of Good and Bad Design Practices Part II

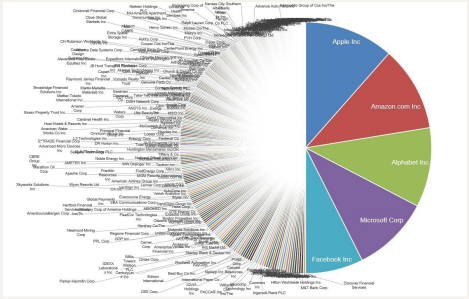
Good Static Statistical Graph



Well Designed bar chart

- **Design**: The bar chart uses appropriate visual cues, contrasting colours and clear labels
- **Clarity**: The data is presented in a straightforward manner allowing users to grasp information easily
- **Accuracy**: Accurately represent the underlying data and avoids misleading data

Bad Static Statistical Graph



Misleading Pie chart

- **Design**: Uses to many slices and similar colours leading to confusion
- **Clarity**: The visualisation lack clarity hard for interpretation between categories
- **Accuracy**: Pie chart distors the data with unequal slices leading to misinterpretation

In summary simplicity, clarity and relevance are paramount for the visualisation whether they are static or interactive. Effective communication provides input for users to explore the data. Contrastingly, bad designs suffer from issues like complexity, lack of clarity and poor accuracy, hindering users ability to understand and extract meaning.

Conclusion

In conclusion, data visualisation is a powerful tool for both communicating and understanding insights derived from complex datasets. Through exploration of quantitative data visualisation techniques, it enables researchers and the audience to gain insights, identify patterns and effectively communicate findings. However, the effectiveness of data visualisation adheres to the principles of simplicity, clarity and relevances.

Contrastingly, misleading visualisations and complex designs impede understanding and disrupt decision making. Hence, it is crucial to choose appropriate visualisation techniques and models based on the nature of the data and characteristics. Thoughtful design considerations, like intuitive interactivity, accurate representation of data and clear labelling. These are all essential for facilitating comprehensions

Moreover,understanding what status or interactive statistical visualisation and the distinction between them underscores the importance of providing users with a tailor experience catering to the specific needs and preferences. Status visualisation offers quick insights and facilitates quick simple decision making while interactive visualisation empowers users to explore data more deeply. Therefore fostering understanding of complex datasets.

Ultimately, in the realm of data science, effective data visualisation plays a pivotal role in driving informed decision making and creating values for organisations. By leveraging visuals effectively, Data Scientists can communicate insights efficiently and therefore contribute the advancement of various fields and facilitates societal development and policy making

References

Bobek, E., & Tversky, B. (2016). Creating visual explanations improves learning. *Cognitive Research: Principles and Implications*, *1*(1). https://doi.org/10.1186/s41235-016-0031-6

Data Visualisation – A Scientific Treatment. (n.d.). Www.linkedin.com. Retrieved April 17, 2024, from https://www.linkedin.com/pulse/20141128184842-3170720-data-visualisation-a-scientific-treatmen

Few, S. (2004). *Eenie, Meenie, Minie, Moe: Selecting the Right Graph for Your Message*. https://www.perceptualedge.com/articles/ie/the_right_graph.pdf

Fry, B. (2008). Visualizing Data. In *Google Books*. "O'Reilly Media, Inc."
https://books.google.com.au/books?hl=en&lr=&id=RRswXg4pJhcC&oi=fnd&pg=PR3&dq=Visualizing+Data:+Exploring+and+Explaining+Data+with+the+Processing+Environment%22+by+Ben+Fry+-+T&ots=7b4k_xjXJ8&sig=HjH51dEvjd815Rr2YJrtZJWPEng#v=onepage&q&f=false

Graham, E. (2017). Introduction: Data Visualisation and the Humanities. *English Studies*, *98*(5), 449–458. https://doi.org/10.1080/0013838x.2017.1332021

Midway, S. R. (2020). Principles of Effective Data Visualization. *Patterns*, *1*(9), 100141. sciencedirect. https://doi.org/10.1016/j.patter.2020.100141

Mößner, N. (2014). Visual Information and Scientific Understanding. *Axiomathes*, *25*(2), 167–179. https://doi.org/10.1007/s10516-014-9246-7

Static vs Interactive Visualizations – How to Choose. (n.d.). Www.linkedin.com. Retrieved April 17, 2024, from
https://www.linkedin.com/pulse/static-vs-interactive-visualizations-how-choose-ali-allage?trk=mp-reader-card

Tableau. (2018). *Data visualization beginner's guide: a definition, examples, and learning resources*. Tableau Software; Tableau. https://www.tableau.com/learn/articles/data-visualization

The Psychology of Data Visualization: How Humans Perceive and Interpret Visual Data. (n.d.). Www.linkedin.com. https://www.linkedin.com/pulse/psychology-data-visualization-how-humans-perceive-interpret-wagh

Young, F. W., Valero-Mora, P. M., & Friendly, M. (2011). *Visual Statistics : Seeing Data with Dynamic Interactive Graphics*. Wiley.