A Study of Advanced Visualization Literacy

Elif Emel Fırat

20254027
School of Computer Science

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

2022

This thesis is dedicated to my dear family.

Abstract

Visualization literacy is essential with the widespread advent of visualization techniques to convey complex data. Two noteworthy facets of literacy are user understanding and the discovery of visual patterns with the help of graphical representations. This thesis focuses on studying and advancing visualization literacy and aims to assist understanding and interpretation of visual designs by developing interactive visualization techniques. Our primary objective is to study treemap and parallel coordinates graphic designs which are commonly used to display data. However, they are not easily understood visual representations.

The research literature provides valuable guidance and opportunities for further studies. We begin by completing a comprehensive survey of literature in interactive visualization literacy, where we identify the previous related research and address past and future trends.

One goal is to identify and address barriers to treemap literacy, a popular visual design to display hierarchical data, with the intend of improving novices users' treemap visualization literacy skills. We examine the results of two years of an information visualization assignment in order to investigate the barriers to understanding and creating treemaps. Also, we create a treemap visualization literacy test and propose a pedagogical tool that facilitates both teaching and learning of treemaps to advance treemap visualization literacy.

We also identify and explore barriers to comprehending parallel coordinates plots (PCPs), one of the advanced graphical representations for displaying multivariate and high-dimensional data. We analyze the obstacles to PCP literacy, design a PCP test, and introduce interactive educational software that assists the teaching and learning of PCPs by offering a more active learning experience.

Finally, the parallel coordinates literacy study inspired us to enhance dense parallel coordinates plots. We introduce techniques that facilitate understanding and interpretation of this complex visual design and present data in dense areas, since increasing data size and complexity may make it challenging to decipher and uncover trends and outliers in a confined space. This project is inspired by the collaborative project RAMP VIS [1].

For reproducibility, both treemap and parallel coordinates literacy tests are provided in Appendix A and B. In addition, the link to software demonstration videos is available in a table presented in the Introduction. More information on the data from the collaborative project presented in Chapter 3 is available in Appendix C. Finally, we offer a supplementary literature review focusing on inclusivity and diversity in data visualization education in Appendix D.

DECLARATION

This work has not been previously accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signature: ....

Date: 8^{th} of June 2022

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signature: ....

Date: 8^{th} of June 2022

I hereby give my consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organizations.

Signature: ....

Date: 8^{th} of June 2022

Contents

Acknowledgements

First and foremost, I would like to express my deep and sincere gratitude to my supervisor Dr Robert S Laramee, for his support and guidance throughout my PhD and my Master's projects before that. His vision, sincerity, and motivation have inspired me. We had many interesting discussions over this period on various topics. He has taught me the methodology to carry out the research and provided many opportunities. Also, I would like to thank him for his friendship, empathy, and great sense of humor.

I would also like to thank the co-authors of my papers: Dr Alena Denisova for her hard work and support; Dr Alark Joshi for his willingness to help and encouragement; Dr Max Wilson for his unsurpassed experience and knowledge; and Dr Ben Swallow for his expertise and excellent feedback. Special thanks to Dr Rita Borgo, Dr Cagatay Turkay, and other RAMP VIS team members for sharing their experiences with me.

I gratefully acknowledge funding from my sponsor, the Ministry of Education of the Turkish Republic. Without their financial and administrative support, it would never have been possible to commence the MSc and PhD degrees.

I am incredibly grateful to my family for their love, caring, supports, and sacrifices to pursue my goals. Special thanks to my friends: Dr Michele Mesiti, Dr Dylan Rees, Dr Liam McNabb, Dr Mohammad Alharbi, and Dr Mehdi El Krari for their infinite patience, assistance, and support during my degree. Finally, I would like to thank my friends in Nottingham and Swansea for their support and the needed distractions.

List of Publications

- Elif E Firat and Robert S Laramee, ***Towards a Survey of Interactive Visualization for Education***, The Computer Graphics and Visual Computing (CGVC) Conference 2018, pages 91-101, 12-14 September 2018, Swansea, UK, DOI: 10.2312/cgvc.20181211
- Elif E Firat and Robert S Laramee, ***Inclusivity for visualization education: A brief history, investigation and guidelines***, Dialogue with the Creative Economy (Dialogo com a Economia Criativa), Special Issue on Visualization, Volume 4, Number 12, pages 146-160, December 2019, DOI: 10.22398/2525-2828.412146-160 [Appendix D]
- Elif E Firat, Alena Denisova, and Robert S Laramee, ***Treemap Literacy: Classroom-based Investigation***, Eurographics Education Papers, Eurographics & Eurovis 2020 (EGEV 2020), 25-29 May 2020, Norrkoping, Sweden, DOI: 10.2312/eged.20201032
- Alexandra Diehl, Elif E Firat, Thomas Torsney-Weir, Alfie Abdul-Rahman, Benjamin Bach, Robert S Laramee, Renato Pajarola, and Min Chen, ***VisGuided: A Community-driven Approach for Education in Visualization*** Eurographics Education Papers, Eurographics 2021 (EG 2021), 3-7 May 2021, Vienna, Austria, DOI: 10.2312/eged.20211003
- Xiaoxiao Liu, Mohammad Alharbi, Joe Best, Jian Chen, Alexandra Diehl, Elif E Firat, Dylan Rees, Qiru Wang, and Robert S Laramee, ***Visualization Resources: A Starting Point*** The 25th International Conference on Information Visualization, (IV) 2021, 5-16 July 2021, Sydney, Australia, DOI: 10.1109/IV53921.2021.00034
- Alark Joshi, Katy Börner, Robert S Laramee, Lane Harrison, Elif E Firat, and Bum Chul Kwon, ***Visualization Literacy for General Audiences-Can We Make A Difference?***(2021) Panel at IEEE VIS 2021, Virtual, 24-29 October 2021
- Elif E Firat, Alark Joshi, and Robert S Laramee, ***Interactive Visualization Literacy: The State-of-the-Art***, Information Visualization Journal, forthcoming, 2022, DOI: 10.1177/14738716221081831 [Chapter 2]
- Elif E Firat, Alark Joshi, and Robert S Laramee, ***VisLiteE: Visualization Literacy and Evaluation***, IEEE Computer Graphics and Applications (IEEE CG&A), forthcoming, 2022
- Ilena Peng, Elif E Firat, Alark Joshi, and Robert S Laramee, ***Evaluating the Impact of Mastery Learning on Parallel Coordinates Literacy***, Eurographics Education Papers, Eurographics, forthcoming, 2022
- Elif E Firat, Alena Denisova, Max L Wilson, and Robert S Laramee, ***P-Lite: A Study of Parallel Coordinate Plot Literacy***, Visual Informatics Journal, forthcoming, 2022 [Chapter 4]
- Michael Dunne, Hossein Mohammadi, Peter Challenor, Rita Borgo, Thibaud-Porphyre, Elif E Firat, Cagatay Turkay, Thomas Torsney-Weir, Ian Vernon, Richard Reeve, Hui Fang, and Ben Swallow, ***Uncertainty Quantification: A Tutorial on a Stochastic Epidemic Model***, Epidemics Journal, forthcoming, 2022
- Alfie Abdul-Rahman, Daniel Archambault, Benjamin Bach, Rita Borgo, Min Chen, Jessica Enright, Hui Fang, Elif E Firat, Euan Freeman, Tuna Gonen, Claire Harris, Radu Jianu, Nigel W. John, Saiful Khan, Andrew Lahiff, Robert S. Laramee, Louise Matthews, Sibylle Mohr, Phong H. Nguyen, Alma A. M. Rahat, Richard Reeve, Panagiotis D. Ritsos, Jonathan C. Roberts, Aidan Slingsby, Ben Swallow, Thomas Torsney-Weir, Cagatay Turkay, Robert Turner, Franck P. Vidal, Qiru Wang, Jo Wood, Kai Xu, ***Visualization for***

Epidemiological Modelling: Challenges, Solutions, Reflections & Recommendations, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, The Royal Society, 2022 DOI: 10.1098/rsta.2021.0299

Under Review

- Elif E Firat, Ben Swallow, and Robert S Laramee, ***DPCP Vis: Techniques for Dense Parallel Coordinate Plots***, Under Review [Chapter 5]
- Xiaoxiao Liu, Mohammad Alharbi, Joe Best, Jian Chen, Alexandra Diehl, Elif E Firat, Dylan Rees, Qiru Wang, and Robert S Laramee, ***Open Visualization Resources: A Comprehensive Overview***, Under Review

Contents

List of Tables

List of Figures

Chapter 1

Introduction

"The greatest value of a picture is when it forces us to notice what we never expected to see."

—John Tukey, Mathematician (1915-2000)

Contents

1.1 Data Visualization

Data visualization is defined as "*the use of computer-supported, interactive, visual representations of data to amplify cognition*" by Card [27]. Converting data into visual components makes data visualization a powerful tool for exploring and making sense of the data. The graphical representation of data assists in identifying the patterns, trends, outliers in behavior and reveal new perspectives. Data visualiza-

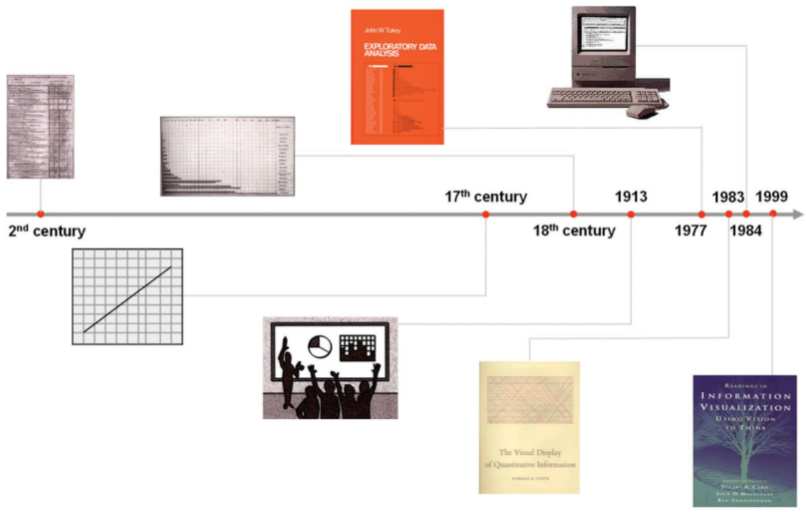


Figure 1.1: *The timeline of data visualization history. Image courtesy of Few and Edge [2]*

tions also can save time by displaying data in an overview image, enabling users to see the entire data set in a single image and quickly understand the patterns [28]. The use of visual designs has a long history and still evolves rapidly (see Figure 1.1). The first table that was recorded was made in Egypt in the 2nd century to organize astronomical data as a navigational tool [2]. During the Portuguese expansion, the kingdom's wealth and boundaries were shown on the Cantino map in the sixteenth-century [29]. William Playfair invented several forms of diagrams to display statistical information later in the eighteenth century using those representational approaches [30]. The first infographics emerged in journals and magazines in the 1970s to condense information [29]. Edward Tufte's book *The Visual Display of Quantitative Information*, released in 1983, demonstrated that there are many effective techniques of visualizing data [2]. Moreover, the survey of information visualization books [31] presents an overview to identify the most relevant books in this field.

Card *et al.* [3] defined the visualization pipeline (see Figure 1.2), which outlines how visualization is created from raw data. In the visual mapping process, focus data characteristics are connected to visual structures after being derived in the data transformation step. This stage fills in the blanks between abstract data and human-perceivable forms. As a result, visual mapping is an important phase to assure the expressiveness of the final representation.

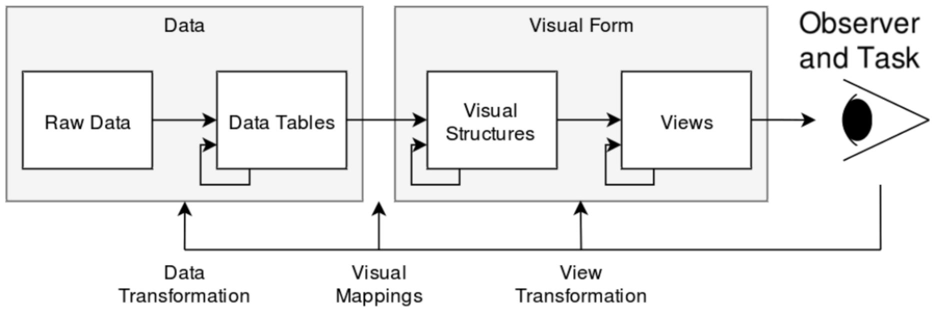


Figure 1.2: *Information Visualization pipeline model introduced by Card et al. [3]. Image courtesy of McNabb et al. [4]*

1.2 Information Visualization

Information visualization and scientific visualization are two main sub-fields of data visualization. Scientific and information visualization focuses on translating data into a visual form used to gain insight and knowledge. Gershon and Eick [32] propose a definition of information visualization as, "*a process of transforming data and information that are not inherently spatial into a visual form, allowing the user to observe and understand the information. This contrasts with scientific visualization, which frequently focuses on spatial data generated by scientific processes.*" The standard methods used in information visualization are line charts, bar charts, pie charts, scatter plots, treemaps or parallel coordinates. To effectively display information using visualization techniques, a set of rules are followed based on visual perception, and cognition [2]. By increasing the number of information visualization applications and software to analyze data, Bikakis [33] identifies aspects that modern visualization should be capable of handling effectively, such as real-time interaction and visual scalability.

1.3 Data Literacy

The phrase data literacy refers to a collection of abilities that revolve around the use of data in everyday thinking and reasoning to solve real-world challanges [34]. For example, Vahey [35] defines data literacy as "*the ability to formulate and answer questions using data as part of evidence-based thinking; use appropriate data, tools, and representations to support this thinking; interpret information from data, and use data to solve real problems and communicate their solutions.*" As a result, data literacy is a crucial ability to extract meaning from data and make rational decisions that are also closely linked and required to interpret visual designs.

1.4 Visualization Literacy

The focus of this thesis is visualization literacy. The simple definition of visualization literacy is the ability to read, interpret, and understand the information presented in graphical designs [36]. Visualization literacy has also been defined as "*The ability to make meaning from and interpret patterns, trends, and correlations in visual representations of data*" [5]; "*the ability to confidently use a given data visualization to translate questions specified in the data domain into visual queries in the visual domain, as well as interpreting visual patterns in the visual domain as properties in the data domain*" [37]; "*the ability and skill to read and interpret visually represented data in and to extract information from data visualizations*" [22]. Moreover, Börner *et al.* [20] define the ability of reading and construction of information visualizations as essential as the ability to read and write text. The neighboring concepts of static graph literacy, sensemaking, and visualization to support general education are not our main focus and are explained further in the Subsection 2.1.1.

Graphics technology demonstrates complex and diverse concepts and has become an integral part of many fields. However, big datasets and complex relationships between data dimensions can make analysis and interpretation difficult. Being able to evaluate, interpret, and analyze the meanings of data and the links between data dimensions presented in visual designs, visualization literacy is a fundamental ability for everyone from young to old. Although interactive visualization methods play a crucial role in simplifying and conveying a meaningful understanding of complex systems, the frequent use of visual images for exploratory analysis of data requires improving visualization literacy skills for novice users.

Several studies focus on visualization literacy. Boy *et al.* [37] developed a method for visualization literacy evaluation using the Item Response Theory (IRT) [38] by creating prospective test items that measure a user's visualization literacy level. The aim was to create an efficient and reliable test using line graphs, bar charts, and scatterplots for identifying users with lower visualization literacy. Similarly, an evaluation tool, created by Maltese *et al.* [39], aimed at investigating the ability of groups with varying levels of experience in STEM fields to read and interpret graphical representations. Ruchikachorn and Mueller [15] present a learning-by-analogy method that illustrates an unfamiliar visual design by displaying a step-by-step transformation from another design. The transformation concept promotes comprehension of the uncommon visual representation by interacting with the transitions. They focus on understandings of parallel coordinates, the hyperbox, spiral chart, and treemap.

Alper *et al.* [8] investigate visualization literacy teaching methods for elementary school children and present an online platform C'est La Vis, that enables students to create and interact with data visualizations and is used by instructors in the classroom for teaching the visualization by creating exercises for children. Moreover, Chevalier *et al.*[40] present an evidence-based discussion of visualization literacy, suggestions for improving it in early education, and future research directions for visualization literacy. Most recently, studying the impact of cognitive characteristics to advance users' visualization literacy has become essential. Thus, Lee *et al.* [17] concentrated on testing the correlation between visualization and cognitive features, such as cognitive ability, cognitive motivation, and cognitive style. Börner *et al.* [20] propose a data visualization literacy framework (DVL-FW) to guide the visualization literacy teaching and assessment. The study provides a set of guidelines and an evaluation that can be utilized to measure and advance visualization literacy.

Further work for enhancing the teaching and learning experience are presented by Kwon and Lee [16]. The work focuses on parallel coordinates, an efficient method to display multidimensional data, to study the impacts of multimedia learning environments for teaching data visualization to non-expert users. The inspiration behind this research is to examine the active learning theory.

1.5 Contributions

The contributions of this thesis are as follows:

Literature Survey: The research literature on visualization literacy provides valuable guidance and essential opportunities for further studies in this field. We present a state-of-the-art (STAR) in Chapter 2, and the STAR examines and classifies prior research on visualization literacy that analyzes how well users understand novel data representations [41]. We categorize existing relevant research into unique subject groups that facilitate and inform comparisons of related literature and provide an overview. Additionally, the STAR/classification also provides an overview of the various evaluation techniques used in this field of research due to their challenging nature. Our novel classification enables researchers to find both mature and unexplored directions that may lead to future work. The review serves as a valuable resource for both beginners and experienced researchers interested in the topic of visualization literacy.

Advancing Treemap Literacy: This research aims to identify and address barriers to treemap literacy to improve a non-expert user's treemap visualization literacy skills. We provide the study in Chapter 3, based on a conference publication [42]. First, we present the results of two years of an information visualization assignment, which are used to identify the barriers to and challenges of understanding and creating treemaps. From this, we develop a treemap visualization literacy test. Then, we propose a pedagogical tool that facilitates both teaching and learning of treemaps and advances treemap visualization literacy. To investigate the efficiency of this educational software, we then conduct a classroom-based study with 25 participants. Finally, we identify the properties of treemaps that can hinder literacy and cognition based on the treemap visualization literacy test results.

Advancing Parallel Coordinates Plots (PCPs) Literacy:
We identify and investigate barriers to comprehending PCPs and present a user study in Chapter 4. We analyze the barriers to reading, understanding, and creating PCPs. We develop a parallel coordinates literacy test with diverse images generated using popular PCP software tools. The test uncovers evidence of barriers to PCP literacy and evaluates the user's literacy skills. We introduce an interactive educational tool that assists the teaching and learning of parallel coordinates by offering a more active learning experience. We aim to advance a novice user's parallel coordinates literacy skills using this pedagogical tool. Based on the hypothesis that an interactive tool that links traditional Cartesian Coordinates with PCPs interactively will enhance PCP literacy further than static slides, we compare the learning experience using traditional slides with our novel software tool and investigate the efficiency of the educational software with an online, crowdsourced user-study. We analyze the features of PCPs that can obstruct literacy and reasoning.

Understanding Data in Dense PCPs: We introduce novel visual designs inspired by the RAMP VIS project [1] and the results of our PCP literacy study introduced in Chapter 4. A dense PCP image resulting from overlapping edges may cause patterns to be covered. We develop techniques to explore the relationship between data dimensions to uncover trends in the data in Chapter 5. We present correlation glyphs in the PCP view to reveal the strength of the correlation between adjacent axis pairs and an interactive glyph lens to uncover links between data variables by investigating the edge intersections. We also present a subtraction operator to identify differences between two similar multivariate data sets and relationship-guided dimensionality reduction based on collapsing of axis pairs. We finally discuss a case study of our techniques on ensemble data and provide feedback from domain experts in epidemiology.

1.6 Challenges

Evaluation of a user's visualization literacy skills is a major challenge. We consider three main challenges with respect to user-studies and visual design choices.

Advanced Visual Designs: Treemap and PCP designs, as they are advanced visual representations, require an understanding of more design components than simple visual designs that make it difficult for users to interpret. These characteristics can be identified as hierarchical data, size, and number of rectangles for treemaps in Chapter 3; high dimensionality and overplotting of large data sets, understanding the links between data dimensions for the PCP design in Chapters 4 and 5.

Literacy Tests: One of the challenges we encountered was the developing visualization literacy tests. In order to create a literacy test, finding or creating appropriate images was a difficult process. In addition to that, isolating the different factors and barriers to treemap and PCP literacy is very difficult. Contrary to our initial hypothesis that literacy test questions might focus on understanding only one aspect of a given visual design, literacy test questions often require a user to comprehend at least three features of the design simultaneously. We address these challenges in Chapters 3 and 4.

User-study Participants: In order to evaluate the effectiveness of the educational tools we developed, we recruited participants. Finding appropriate participants during summer time and lockdown for a user study designed for a classroom environment was a great challenge. Due to the fact that the classroom-based user-study introduced in Chapter 3 was conducted in the summer period, it was not easy to reach an appropriate and larger number of participants. For the user-study presented in Chapter 4, the Covid-19 lockdown pushed us in different directions, such as changing the study design and recruiting participants for engagement with a crowd-sourcing platform.

1.7 Thesis Overview

The rest of this thesis is structured as follows. Chapter 2 presents a survey of interactive visualization literacy literature. The chapter provides a novel classification that assists the reader in discovering topics of interest and offers a summary of previous related literature. Chapter 3 identifies barriers to treemap literacy, presents a treemap literacy test, an educational treemap tool, and the results from a classroom-based user study. Chapter 4 investigates barriers to understanding PCPs, provides a PCP literacy test, and introduces an educational PCP tool to advance PCP literacy. The results of a crowdsourcing experiment are also presented in this chapter. The techniques aimed at investigating relationships between data dimensions to reveal trends in PCPs are presented in Chapter 5. This chapter features novel interaction methods featuring an overview of the connection between data dimensions and introducing a subtraction operator and dimensionality reduction. Conclusions are drawn in Chapter 6, and future work is presented.

Moreover, the full treemap and parallel coordinates plot literacy tests introduced in Chapters 3 and 4 are provided Appendices A and B respectively. More information on the Covid-19 data presented in Chapter 3 is available in Appendix C. Finally, we offer a supplementary literature review focusing on inclusivity and diversity in data visualization education [43] in Appendix D.

1.8 Video Demonstrations

Each chapter contains a supplementary video that highlights our contributions and demonstrates the software tools we created. We recommend watching these videos for a general sense of the contributions presented in chapters. The video demos and video links are included in Table 1.1. In addition, video links of presentations in conferences and a panel are provided in Table 1.2.

Chapter	Title	Video URL
Chapter 3	Treemap Literacy: A Classroom-based Investigation	https://vimeo.com/660529352
Chapter 4	A Study of Parallel Coordinate Plot Literacy	https://vimeo.com/456884883
Chapter 5	Techniques for Dense Parallel Coordinate Plots	https://vimeo.com/652208042

Table 1.1: The table displays video demonstrations of our studies presented in following chapters. Chapter 3 and 4 have additional supplementary videos in each related chapter.

Chapter	Title	Video URL
CGVC	Towards a Survey of Interactive Visualization for Education	https://bit.ly/3EAmHSM
Eurographics	Treemap Literacy: A Classroom-based Investigation	https://bit.ly/32DPjh1
IEEE Vis Panel	Visualization Literacy for General Audiences: Can We Make a Difference?	https://bit.ly/3ezfTdA

Table 1.2: The table shows conference and panel presentation videos of from the events that were completed during the PhD.

1.9 PhD Timeline: 2018–2021

We provide a timeline of this PhD project since the PhD was completed under unusual and historical circumstances. I started my PhD study in January 2018 at Swansea University, displayed in Figure 1.3, and have presented my work at various events. I presented a poster at the first College of Science (CoS) Postgraduates (PGR) Conference in 2018. I presented a survey paper on visualization education [44] at the Computer Graphics and Visual Computing (CGVC) Conference in September 2018. In addition, a treemap tool was developed for the treemap literacy study and I presented the tool in two meetings with Professor Ben Shneiderman (University of Maryland) and Dr. Max Wilson (University of Nottingham). I also introduced the treemap tool at the second CoS PGR Conference in 2019. A paper on Inclusivity for Visualization Education [43] was published in Journal Dialogue with Creative Economy in 2019 (see Appendix D). In March 2020, I transferred to the University of Nottingham with my supervisor. The treemap literacy study [42] was presented

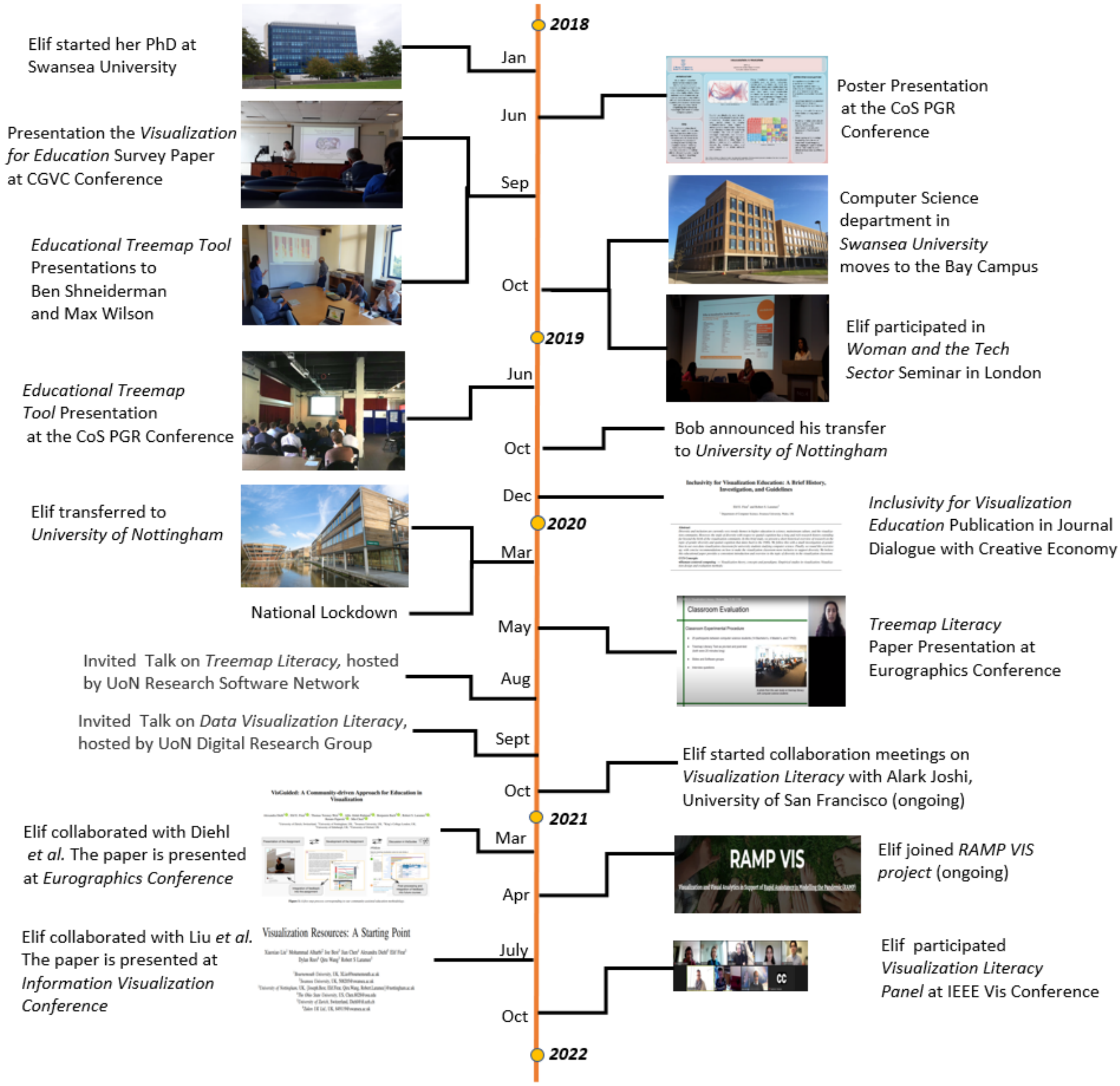


Figure 1.3: *Timeline of the most key events, presentations, publications and contribution of the collaborative work with different academics and research groups throughout the PhD (2018-2021).*

at Eurographics Conference in May 2020. I was invited to two talks hosted by the University of Nottingham Software Network and Digital Research groups, and I delivered presentations on treemap and data visualization literacy. Moreover, I have started to collaborate with Alark Joshi (University of San Francisco) and contribute to papers in the visualization literacy field in 2020. Moreover, two studies were published as a result of collaborations with Diehl *et al.* [45] and Liu *et al.* [46] in 2021. Also, I was on the Visualization Literacy Panel [47] at IEEE Vis Conference in October 2021. Finally, I joined the RAMP VIS project and started to attend regular meetings in 2021, and wrote a collaborative paper based on this project [48].

Chapter 2

Interactive Visualization Literacy: The State-of-the-Art

"We are drowning in information but starved for knowledge."

—John Naisbitt, Author (1929-2021)

Contents

We started our research by reviewing the literature on visualization literacy. As a result, we discovered several articles published in this area. However, we see that no up to date study has been published that summarizes and analyzes the literature on visualization literacy. Therefore, this chapter presents a state-of-the-art literature review on visualization literacy that provides valuable meta-data and guides researchers interested in the field. This chapter is the result of the STAR [41].

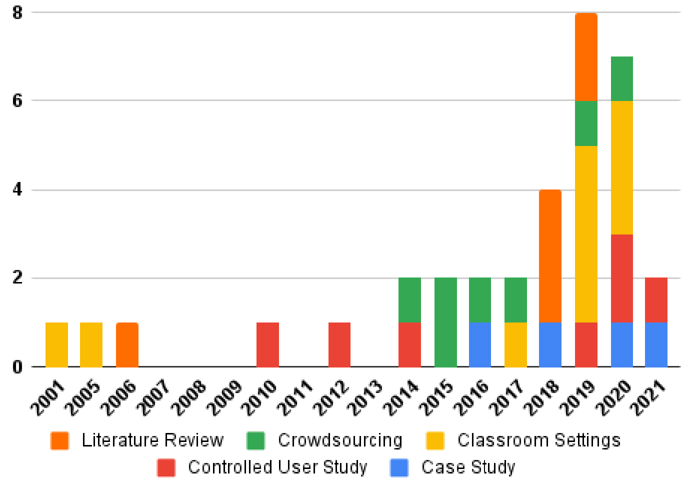


Figure 2.1: *Number of papers by publication year and evaluation method used. There are 34 papers in total. We stopped searching for literature in the middle 2021.*

2.1 Introduction and Motivation

Visualization literacy is an essential skill required for comprehension and interpretation of complex imagery conveyed by interactive visual designs. Developing visualization literacy is essential to support cognition and evolve towards a more informed society [40]. Gaining a deeper understanding of the visualization literacy of a cohort of participants or domain experts has become a prominent theme in the information visualization community. Visualization literacy was described as an essential skill in the IEEE VIS 2019 keynote talk by Börner [49]. Few studies were published in the previous 20 years; however, there was an increase in the last 7 years in this field as shown by the graph in Figure 2.1. If we look at different categories, there is no obvious trend yet it due to immaturity in the field. In recent years, more studies feature classroom-based evaluation and literature reviews.

The Merriam-Webster dictionary defines the term literacy as "the ability to read and write" [50]. Literacy is described as the ability to comprehend and use something with an emphasis on the consumption aspect when the term is combined with other subjects like information literacy, health literacy, etc. More specifically, visualization literacy is defined by Boy *et al.* as "a concept generally understood as the ability to confidently create and interpret visual representations of data [37]". Börner *et al.* explain, "the ability to make meaning from and interpret patterns, trends, and correlations in visual representations of data" [5], while Lee *et al.* refer to it as "the ability and skill to read and interpret visually represented data in and to extract information from data visualizations" [22]. There are also related concepts such as visual literacy which is defined by Bristor and Drake as the, "ability to understand, interpret, and evaluate visual messages" [51]. Ametller and Pintó state that visual literacy "encompasses the ability to read (understand or make sense of) as well as write (draw) visual representations" [52] while Bradent and Hortinf identify it as "the ability to think, learn, and express oneself in terms of images" [53].

In this chapter, we present a literature review of visualization literacy to inform both mature and unsolved problems and to convey trends emerging from visualization literacy to readers who are interested in this topic as a research direction. The study also provides an overview of the evaluation methods used in visualization literacy studies. To investigate the state-of-the-art systems implemented for advancing literacy skills, we survey and classify a variety of literacy research. The contributions of this state-of-the-art report (STAR) are as follows:

- The first survey of its kind on the topic of visualization literacy with a special focus on evaluation with a total of 34 main papers with an additional 45 related publications.
- A novel literature classification of research papers in this area
- Beneficial meta-analysis to facilitate comparison of the literature
- Indicators in the field of both mature themes and unsolved problems

We collect literature referenced in this survey in an online resource using an interactive literature browser called SurVis [54]. This can be found at the following URL: http://www.cs.nott.ac.uk/~psxef1/index.html

The rest of the survey is organized as follows. We first present an overview of the related work that contains previous relevant papers that examine visualization literacy. The subsequent section provides a review of visualization subjects and technologies used to enhance users' ability to understand and interpreting visual representations in different research fields. We later present a discussion of future work and open directions for research in this field.

2.1.1 STAR Scope

In this state-of-the-art report, we provide an overview of visualization papers that *examine/test/study* users' visualization literacy skills and improve the literacy skills of understanding and creating advanced visual designs. Studies that concentrate on data visualization literacy using interactive visualization techniques are within the scope of this survey. The STAR includes papers to investigate the ability of reading, understanding, interpreting, and constructing visual designs. The main criterion is to examine how the work advances user's basic comprehension and interpreting visual representations of data.

The research topics and papers presented here introduce methods or software that include advanced and interactive graphical representations developed and used for improving visualization literacy skills. A major challenge is to evaluate the effectiveness of the target methodologies and technologies for increasing a user's understanding with the support of interactive visualization systems. Evaluating the effectiveness of an interactive visualization technique to advance visualization literacy is a non-trivial endeavor. As such this survey pays particular attention to the type of evaluation used when examining the literature.

2.1.2 Out of Scope

Our criterion for including research in the STAR is that visualization literacy is the focus. Papers concentrated on the neighboring concepts of static graph literacy, sensemaking, visualization to support general education, or general user-studies in visualizations are out of scope. Including these related topics would make the survey unmanageable in size and would detract from its primary focus.

1. Static Graph Literacy There is considerable research in the field of vision psychology that studies static graph literacy. For example, Simkin and Hastie [55] collect evidence that observers make assumptions about what types of information is contained in graphs. Another study by Pinker [56] offers a theory that is used to estimate what makes a person better or worse at reading graphs and what makes a graph better for transferring given information to a reader. Shah and Hoeffner [57] present principles for displaying graphs to students and determine the implications of graph understanding research on the pedagogy of graphical literacy skills. Similarly, the purpose of the research by Moore-Russo *et al.* [58] is to determine how a group of mathematics teachers reasoned about spatial tasks by looking at different components of spatial literacy. Since the focus of this survey is interactive visualization, we do not include static graph literacy studies from vision psychology.

2. Sensemaking Research on sensemaking is considered out of the scope of this survey. Pirolli and Card [59] define sensemaking as an ongoing process that translates raw data into the reasoning of information progressively. Another definition of sensemaking is "the process of collecting, representing and organizing complex information sets based on a particular problem in such a way that can help us understand the problem better and make sense of it more effectively" [60]. Lee *et al.* define the term sensemaking as, "conscious efforts to achieve understanding of how to interpret visual objects and underlying content in an information visualization" [61]. Sensemaking *focuses on comprehension of the underlying data* rather than understanding a given visual design. In our survey the focus is on basic understanding of a given *visual design* rather than the underlying data. If the focus of a research paper is data-centric, then it is out of scope.

3. Visualization to Support General Education Our work is focused on visualization literacy for general visualization. We recognize that visual designs play a key role in simplifying and conveying meaningful understanding of complex systems. Visualization tools can assist the educational process and enhance cognition for all types of users. Schwab *et al.* [62] provide a web-based environment called booc.io that enables users to study specific concepts and facilitates creating hierarchical structures. The tool enables linear and non-linear presentation of content such as lecture slides, book chapters, and videos.

Another study by Silva *et al.* [63] provides experiences using VisTrails as an environment to teach scientific visualization. VisTrails is an open-source tool designed to assist research on computational tasks such as data analysis and visualization.

The work presented by Contero *et al.* [64] aims to improve engineering students' visualization skills using a web-based graphics application and a sketch-based modeling system. Websites are used in the course which enable students to implement 3D graphical content offering richer features to improve students' visualization skills. The focus here is on improving education, not visualization literacy. Visualization

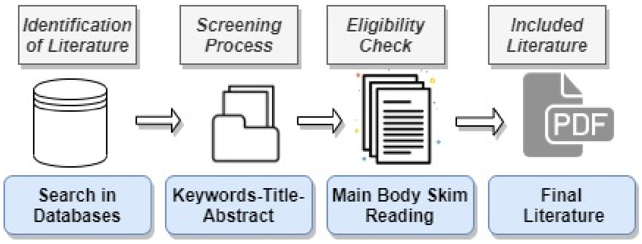


Figure 2.2: *Sampling process of the literature*

used to support education is out of scope because it assumes the user understands the visual designs. Firat and Laramee [44] review research that concentrates on teaching or learning materials in classrooms or distance learning systems using advanced special visualization techniques.

4. User-Studies in Visualization General user-studies in visualization are generally not included in this survey because they focus on a specific visual design optimization. McNabb and Laramee [65] present an overview that includes research on user studies. Fuchs *et al.* [66] survey a number of papers on glyph design to inform understanding of trade-offs in the glyph design space. In general, user studies on visualization assume that users already have a basic understanding of a visual design. Visualization literacy does not necessarily make that assumption. Studies of literacy do not presume an a priori level of comprehension for a given visual representation. They test basic visualization literacy.

2.1.3 Literature Search Methodology

Our methodology uses a systematic search of the existent literature in the field of visualization literature with an emphasis on papers published at the annual IEEE VIS Conference [67]. We specifically focused on the keywords “Visualization”, “Literacy”, “Teaching”, “Learning”, “Understanding”, “Interpretation”, and “Construction” in our search for literature. We also included the following resources for our literature search:

- IEEE Xplore [68]
- Google Scholar [69]
- Vispubdata [70]
- Eurographics education papers [71]

- Eurographics Digital Library [72]
- ACM Digital Library [73]
- IEEE Pacific Visualization Conference [74]

To further expand our collection of resources, we examine all the papers that were cited by the papers in our collection. The related work section of each paper was also examined for additional sources of visualization literacy research. A Survey of Surveys (SoS) [65] indicated that there was an absence of surveys on visualization literacy. A survey of interactive visualization for education [44] does not include any study on visualization literacy either. This is also relatively little material provided on this topic in information visualization books [31]. We checked all the literature cited in the related work section of our previous study on treemap literacy [42]. As a result of our search for the keyword "Visualization Literacy" on Google Scholars, we found approximately 90,000 matches. But after reviewing the 70 articles in the first 7 pages, the studies began to lose relevance. We checked the references of each and used Google Scholar's *"cited by"* feature to discover more research. In the next step, after a quick review of the article and checking the eligibility of the articles, the number of studies in our collection was finally reduced to 34. The figure 2.2 shows the general sampling procedure.

2.1.4 Literature Overview

To categorize the papers and projects we examined, we developed a novel classification. Table 2.3 summarizes how each research paper is classified. We carefully examine the evaluation methods in each paper and further categorize the evaluation method used in the paper as well as providing the number of participants involved in the evaluation. For evaluation methods, we identify five categories: **in the wild**, **controlled user study**, **classroom-based evaluation**, **crowd-sourced evaluation**, and a **meta-review** of related literature. The categories are presented in ascending order according to the approximate number of participants involved in the evaluation process. The category called "review of literature" includes papers based-on literature surveys rather than providing an explicit evaluation. Using this approach above, we define a matrix for the literature classification (the columns in Table 2.3).

- **In the Wild:** This evaluation method includes observing and recording a group of participants in a public setting and how this changes over the time in an uncontrolled environment [75]. The goal is defined by Roger and Marshall [76] as "*understanding how technology is and can be used in the everyday/real world, in order to gain new insights about: how to engage people/communities in various activities, how people's lives are impacted by a specific technology, and what people do when encountering a new technology in a given setting.*" It is one of the preferred assessment methods that involves a use-case of how the given software is used in a public environment. Research presented by Börner *et al.* [5] incorporates this evaluation technique.

- **Controlled User Study:** A controlled user study is an experiment conducted in a controlled laboratory environment. Individual participants are asked to use new interactive and visual designs and perform specific tasks. Task performance time and correctness are measured and evaluated. Grammel *et al.* [77] and Kodagoda *et al.* [78] prefer this method.

- **Classroom:** Researchers prepare pre- and post-experiment tests and examine a visual designs' effectiveness in a classroom environment based on a group of students. Task performance is evaluated on a cohort level. Pre- and post-experiment tests in a classroom evaluation environment are the most popular across all categories. Bishop *et al.* [11] and Firat *et al.* [42] incorporate a classroom evaluation.

- **Crowdsourced evaluation:** This method includes studies that are conducted and evaluated online. Researchers collect feedback from a wide geographically-distributed pool of participants in order to collect the largest amount of participant data possible. Crowdsourcing using Amazon's Mechanical Turk offers a large number of experimental participants in a very short time at reasonable costs for obtaining participant data. Boy *et al.* [37] use Amazon's Mechanical Turk to assess visual designs developed for education.

- **Literature Review:** This category is created to identify papers which do not provide any explicit evaluation technique. The studies by Lee *et al.* [17] present literature reviews of visual systems. This is a kind of meta-evaluation.

Evaluation Method & Characteristics	***Physical Distance***	***Proximity of the Observations***	***Control over the Environment***	***Number of Participants***
In the Wild	Medium	Medium	Medium	30–400
Controlled User Study	Close	Close	High	10–180
Classroom-based User Study	Close	Medium	Medium	10–50
Crowdsourcing Study	Far	Distant	Less	30+
Literature Review	N/A	N/A	N/A	N/A

Table 2.1: A summary of the evaluation methods and different characteristics of the classification categories.

For each category we describe, there is a physical distance involved between the participants and the researcher. For example, classroom-based and controlled user studies involve very close distances meaning that experiments are conducted in the same room. At the same time, crowdsourcing evaluation involves participation across the globe. Another characteristic is the level-of-detail of the observations that can be recorded based on the distance between experimental participants and observers. The level of observational detail for each participant differs for each type of evaluation. For instance, observations are made with the studies in the wild by paying attention to an uncontrolled group of individuals. User-studies support the highest level of detail for making observations. Usually measuring every individual task sometimes with supplementary video. The focus of the observation is a group in a classroom style evaluation while it is a distant larger group of people in a crowdsourcing study. We also have different levels of control over the environment. We have strict control over the environment for user-studies. There is a higher level of control over the environment with a lab-based user-study than a classroom-based study. The number of participants also changes depending on your evaluation method. It is usually around 10–50 people in a classroom-based study, while it's more the crowdsourcing study e.g., 30–200. These characteristics are summarized in the Table 2.1. Table 2.3 presents that classroom-based evaluation is the most popular followed by crowdsourced evaluation. Also, some papers did not provide the number of participants involved in their evaluation. This is indicated with a 'N/A' in the table.

2.2 Background

In this section we describe related surveys that systematically review papers with visualization user studies. A survey provided by Fuchs *et al.* [66] reviews 64 research papers with quantitative controlled studies focused on data glyphs to help researchers and practitioners gain understanding, to find the most relevant papers, and obtain an overview of the use, design, and future research directions involving glyphs.

Johansson and Forsell [79] provide a comprehensive literature review that examines user-centric assessments and explores usability challenges with parallel coordinates. They present 23 papers in four categories: analysis of axis configurations, comparison of clutter reduction approaches, practical application of different parallel coordinates, and comparison of parallel coordinates with other analytical techniques. The survey identifies challenges within the field and provides guidelines for possible future studies.

ACRL Knowledge Literacy Standard [80] requires students to assess and integrate sources into their knowledge base. There are sufficient studies, in both the evaluation and critique of data visualization resources, supported by considering these elements separately. *Evaluation* corresponds to the basic questions asked in order to determine the quality, accuracy and reliability of a particular visual design. *Critique* is an analysis raised to the next level and seeks to answer the question of whether, for a particular application, a particular data visualization is among the best in its field.

Firat and Laramee [43] present a historical overview of studies on gender diversity and spatial cognition and share gender bias research findings in data visualization classrooms for university students studying computer science. The paper offers concise recommendations on how to make the visualization classroom more inclusive in order to encourage diversity.

2.3 The State-of-the-Art on Interactive Visualization Literacy

This section presents a collection of important re-occurring themes related to visualization literacy and associated research papers. Each research paper is summarized

Table	*Description*
Table 2.1	A summary of the evaluation methods and different characteristics of the classification categories.
Table 2.3	An overview of the visualization literacy literature comparing classification categories, target participants, and number of participants involved in the evaluation.
Table 2.4	A summary table of literature that comparing evaluation controlled user study experiments, the age range,and the average age of the participants involved in the study and duration of the study
Table 2.5	A table indicating the data themes used in the literature.
Table 2.6	A classification table of studies that use a classroom evaluation approach. Literature is classified as using a united or divided class. Education level of participants is provided.
Table 2.8	A summary table of literature that presents experiments carried out utilizing crowdsourcing platforms.
Table 2.7	An overview of the literature and the visual designs evaluated. The table displays the type of visual designs tested by column.
Table 2.9	An overview table summarizing the contributions of the literature for research purpose on visualization literacy.
Table 2.10	A summary table of literature that compares the number of references are provided in literature reviews on visualization literacy.
Table 2.11	A table of future research directions discussed for each paper. The directions displayed represent common research areas that reoccur in the literature.

Table 2.2: A summary of meta-analysis provided in this literature review.

in a systematic way to guide the literature review [81]. Each paper is placed in its respective category (in the wild, controlled study, classroom study, crowdsourced evaluation, or literature reviews) to facilitate comparison.

Summary of Meta-Analysis: A summary of meta-analysis provided in this study is presented in Table 2.2. In Table 2.3 we compare the evaluation methods of the different papers while presenting the literacy skills tested and the target participants. Table 2.4 presents literature with controlled user study experiments and presents the age range and the average age of the participants involved in the study as well as the duration of the study. We look at data themes used in the literature and give an overview in Table 2.5. The literature involving classroom-based user study settings is compared in Table 2.6. The table summarizes literature as using a united or divided classroom for the study and the education level of the participants. Table 2.8 indicates a summary of literature that presents experiments carried out utilizing crowdsourcing platforms with the choice of online platforms. Table 2.7 provides an overview of the literature and the visual designs evaluated. Table 2.9 summarizes the contributions in the papers that are classified based on common themes. The literature review papers are compared in Table 2.10 with the number of references. Finally, Table 2.11 presents future research directions discussed in each paper.

Literature & Categories	***In the Wild***	***Controlled User Study***	***Classroom-based User Study***	***Crowdsourcing Study***	***Literature Review***	***Target Participants***	***Number of Participants***
Baker *et al.* [82]			***R+W***			***A***	52
Delmas *et al.* [83]			***R***			***A***	1464
Schönborn *et al.* [18]					***R***	***A+C***	N/A
Grammel *et al.* [77]		***R+W***				***A***	4, 9
Kodagoda *et al.* [78]		***R***				***A***	24
Boy *et al.* [37]				***R***		***A***	34, 37, 34, 39
Huron *et al.* [84]		***R+W***				***A***	12
Ruchikarhorn and Mueller [15]				***R***		***A***	22, 11 ,11
Maltese *et al.* [39]				***R***		***A***	202
Kwon and Lee [16]				***R+W***		***A***	120
Börner *et al.* [5]	***R***					***A+C***	273
Alper *et al.* [8]			***R+W***			***C***	6, 15
Lee *et al.* [22]				***R***		***A***	65, 191
Wojton *et al.* [85]	***R***					***A***	388
Chevalier *et al.* [40]					***R+W***	***C***	N/A
Zoss *et al.* [19]					***R+W***	***A+C***	N/A
Mansoor and Harrison [86]					***R***	***A+C***	N/A
Börner *et al.* [20]					***R+W***	***A+C***	N/A
Stoiber *et al.* [21]					***R***	***A+C***	N/A
Lee *et al.* [17]				***R***		***A***	178
Fuchs *et al.* [9]			***R***			***A***	28
Gäbler *et al.* [10]			***R***			***C***	23
Bishop *et al.* [11]			***R***			***C***	24
Lallé *et al.* [6]		***R***				***A***	119
Krekhov *et al.* [12]			***R+W***			***A***	11
Firat *et al.* [42]			***R***			***A***	25
Wang *et al.* [13]			***R+W***			***A***	11
Rodrigues *et al.* [14]			***R***			***A***	22
Huynh *et al.* [7]		***R***				***C***	33
D'Ignazio and Bhargava [87]	***R+W***					***A+C***	N/A
Donohoe and Costello [88]				***R***		***A***	32
Barral *et al.* [89]		***R***				***A***	56, 119
Barral *et al.* [90]		***R***				***A***	56, 119
Peppler *et al.* [91]	***R+W***					***A+C***	~33
Total: 34	4	7	10	7	6		

Table 2.3: An overview of the visualization literacy literature with classification categories, target participants, and number of participants involved in the evaluation. The evaluation technique that each research paper uses is categorized into: in the wild, controlled user study, classroom setting, crowdsourcing, and literature review. '***R***' indicates that reading and understanding are tested whereas '***W***' indicates where the ability to construct (write) a visual design is evaluated. The target participants in the studies are identified as '***A***' adult and '***C***' children.

2.3.1 Visualization Literacy In the Wild

This subsection introduces papers in which a study is conducted in the wild in order to demonstrate the idea presented in the research. Study participants in this category are members of the public. The exact number of participants is not controlled, neither is the selection process of participants. Each study provides a use-case scenario for the given software and testing it in an uncontrolled public environment. This evaluation method is one of the methods used to evaluate visualization systems. Börner *et al.* [5] study the familiarity of young and adult museum visitors with a selection of visual designs. A study is conducted in three US science museums, considered informal learning environments. Börner *et al.* [5] chose 20 visualizations from textbooks and widely used online visualization libraries such as the D3.js library [92]. These visual designs consist of two charts, five maps, eight graphs,and five network layouts (see Figure 2.3). Five of the 20 visual designs were displayed to visitrs of the science museums. Museum visitors are asked to state their familiarity with the visual designs and to identify the name of the design.

Some 127 youths aged between 8-12 years old and 143 adults participate in a pre-test experiment. Visitors with a known perceived gender comprise 110 youth and 117 adults. Before exploring the set of five visualizations, participants were asked to report their interest in science, math, and art on a scale of 1-10. During the test, visitors are asked questions about data and data presentations. During a post-test, a total of 53 subjects sorted the five visual designs in order from easiest to most difficult to read. The results indicate strong experimental evidence that a very high proportion of the studied population, both adult and youth cannot name or interpret visual representations beyond very basic charts. They show low performance on the main aspects of data visualization literacy. The results indicate charts are easiest to read, followed by maps, and then graphs. Network layouts were identified as the most difficult to read.

A study by Wojton *et al.* [85] proposes a Simplicity-Familiarity Matrix, a study-driven model for integrating advanced data visualizations into an exhibition design to ensure all museum visitors can understand the visualizations and participate. This model derives from a data literacy study. The method of creating a data visualization was used to examine those aspects of data visualization are simpler or difficult for

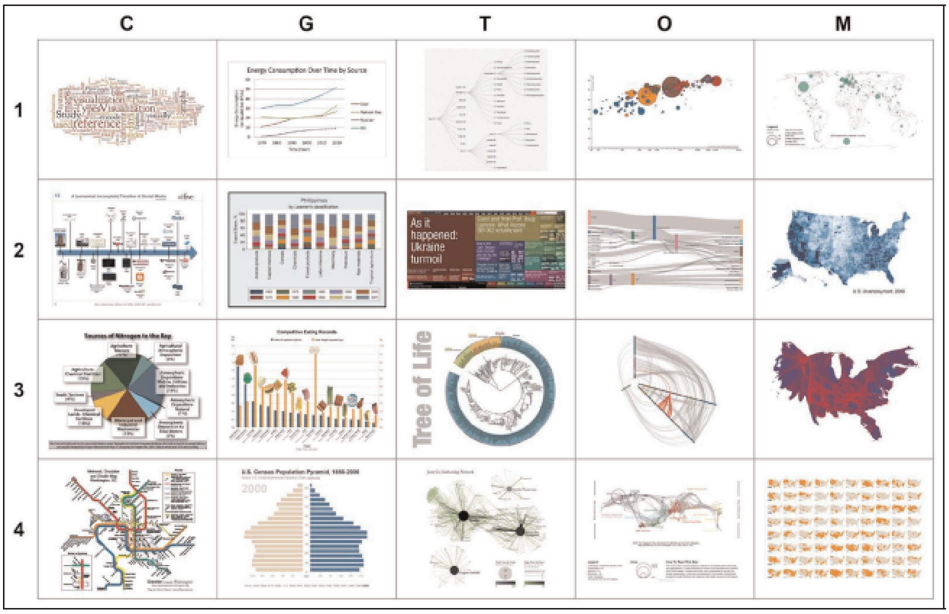


Figure 2.3: *Four sets of five visualizations each row represents one set. All four rows make up the complete set of all 20 visualizations used in the study. Image courtesy of Börner et al. [5]*

visitors to comprehend quickly and correctly. The study was performed in four museums and one aquarium to collected data that was driven by the question, *"How do people engage with/ understand reference systems?"*. A total of 250 adults and 138 youth participate. Four visualization booklets consist of a base and layers required to create a visualization. During the design construction, the participants were shown the base and asked to identify what it was communicating. The finding addresses key concerns and problems related to data visualization across two spectrums: simplicity and familiarity. The results also indicate that that visual designs with more than average complexity are difficult to comprehend. The use of the Simplicity-Familiarity Matrix can be used to increase visitors understanding.

D'Ignazio and Bhargava [87] propose teaching methods focused on feminist theory, procedure, and design to address inequalities. Via three case studies, they explore what feminism can offer visualization literacy, with the intention of improving self-efficacy for women and less-represented groups. They demonstrate creative ways to to work with data and develop learning spaces.

The first case study concentrates on the process of collaborating with community groups, Groundwork Somerville, and local youth design to paint a data-driven story as a community mural. This example of a 'data mural' documents an action-oriented, community-based project that builds data literacy. In the second case, an exercise

Literature	***Age Range & Average Age***	***Duration***
Grammel *et al.* [77]	20-24, 21	45 min
Kodagoda *et al.* [78]	35-50, 39	–
Huron *et al.* [84]	22-43, 28	70 min
Lallé *et al.* [6]	18-69, 26	90 min
Huynh *et al.* [7]	11-13, –	60 min
Barral *et al.* [89]	19-69, 26	115 min
Barral *et al.* [90]	19-69, 26	115 min

Table 2.4: The summary table of literature that introduces evaluation controlled user study experiments, the age range, and the average age of the participants involved in the study and duration of the study.

called 'Asking Questions' from the DataBasic.io suite of resources and activities is described. DataBasic.io includes web-based tools for beginners that incorporate principles for working with data varying from quantitative text analysis to network analysis. For the purposes of this case they focus on the WTFcsv tool and its accompanying learning activity 'Asking questions'. The third case study is an activity designed to allow people to exercise the ability to argue with data in order to persuade people to take action, called 'ConvinceMe'.
Moreover, they provide six conceptual principles for feminist data visualization, drawing from work in feminist science and technology studies, feminist human-computer interaction, feminist digital humanities, and critical cartography & GIS.
The study by Peppler *et al.* [91] investigates what design features can assist data visualization literacy in science museums. They conduct a study and collect video data from 11 visitor groups that participated in visualization reading and writing in a science museum exhibition. Visitors are encouraged to interact with and interpret data, which consists of personal information records. Furthermore, tasks related to the data visualization framework are displayed around a screen. Participants are shown images and given tasks like finding themselves in the data, comparing with others, or changing their data. Responses collected from participants are analyzed thematically. Results indicate how the exhibit's design features resulted in the identification of single data records, data patterns, incorrect measurements, and distribution rate. The findings address design practices to aid data visualization literacy in museums.

2.3.2 Visualization Literacy and Controlled User Study-Based Evaluation

A controlled user study is an investigation carried out in a controlled laboratory environment. Participants are required to undertake given tasks interacting with visual interfaces. The success rate and completion times for each individual task are recorded. Generally, the experiment is performed one participant at a time. Grammel *et al.* [77] and Kodagoda *et al.* [78] chose a controlled user study method for the evaluation. Table 2.4 presents literature that incorporates controlled user study experiments for evaluation and provides participants' age range and the average age. We can see from the table that a wide age range of participants are involved in the user studies and the duration of the studies which averages 60-70 minutes.

Grammel *et al.* [77] investigate the barriers novices face in interpreting and expressing visual designs when developing tools that enable users to create good graphical representations. The study examines how novices create a visual representation. They present abstract models to identify obstacles to understanding data and tool implications from their findings to reduce uncovered obstacles. In a series of a pilot studies, four participants were asked to define the visualizations they want to see while a mediator generates designs using Tableau and shows the participants the resulting visual layouts. After a few pilot studies, nine students from the business school participate in an experiment. Students are observed for 45 minutes while constructing visual layouts. They are requested to analyze their visual designs verbally. The study ends with an interview questionnaire to explore the resulting visual layouts and problems encountered while students were constructing visual designs. This study reports three visual mapping process barriers: i) selecting the parameters to map to the visual variables, ii) selecting the visual marks to be used, and iii) decoding and interpreting the visual result.

The research offers implications for tool design based on empirical evidence. Some suggestions for developing tools include providing search facilities for retrieving data, offering visual design suggestions, supporting iterative specification and providing explanations, and support for learning of the design.

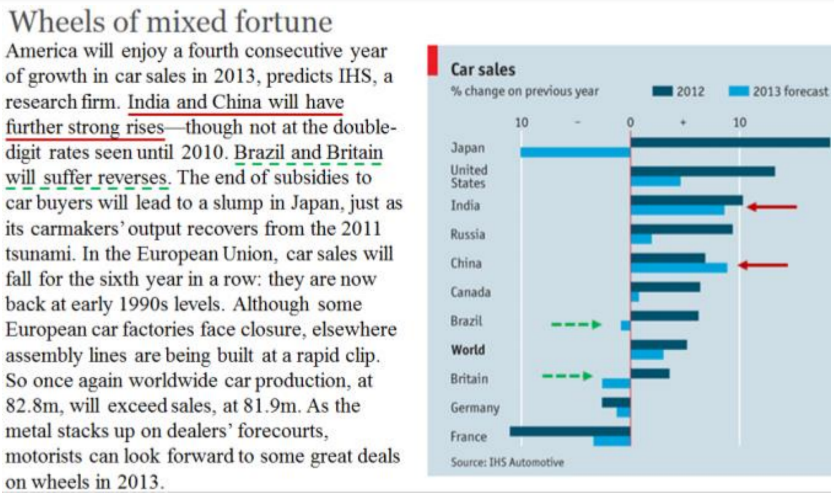
A study by Kodagoda *et al.* [78] describes the challenges low literacy (LL) users face while searching for information online. They derive a set of design principles for visual interfaces suitable for LL users. This research identifies two difficulties: understanding LL users in a way that facilitates new designs and understanding the problem that needs addressing. Based on the differences in LL users' reading skills and perceived mental models, recommendations on designing a user interface for LL users are suggested.

Invisque (INteractive VIsual Search and Query Environment) [93] was developed for creating queries and searching for information in an interactive and visual approach. Invisque focuses on a set of design principles advantageous for LL users. Invesque decreases the cognitive load of word-for-word reading by providing information in boxes on white space and showing the amount of data visible that can be modified through the use of a slider.

The effectiveness of Invisque is evaluated by comparing LL users' performance with the performance of HL (High Literacy) users by using a traditional website for search and retrieval tasks. Some 12 HL and 12 LL participants from a local Citizens Advice Bureau (CAB) with computer and internet literacy are recruited for the study (12 female, 12 male) with a mean age of 39. The study is conducted in a lab. Each participant performs six tasks in total, three with Invisque and three with the Adviceguide website which starts from a general overview and then requires a deep search to access more specific topics. For each interface, participants perform one easy, one medium and one difficult task. Cognitive Task Analysis (CTA) techniques are used to understand the users' decision process during their tasks. Techniques such as think-aloud, user observation, semi-structured interviews, and questionnaires focusing on the systems were used as methods for data collection. Results indicate that Invisque enhances LL users' performance and changed their behavior strategies.

Huron *et al.* [84] explore how users build their visualizations and what kinds of visualizations they create. They introduce a visual mapping model to explain how users utilize tokens to form a visual arrangement that conveys their data as well as providing implications for designing tools.

The study's goals are to understand more about the visual mapping process, determine what makes the process easy or difficult for users, and investigate the suitability of constructive authoring of infovis as a method to construct images. Some 12 participants are assigned three tasks (create, update and annotate a visualization)


Wheels of mixed fortune

America will enjoy a fourth consecutive year of growth in car sales in 2013, predicts IHS, a research firm. India and China will have further strong rises—though not at the double-digit rates seen until 2010. Brazil and Britain will suffer reverses. The end of subsidies to car buyers will lead to a slump in Japan, just as its carmakers' output recovers from the 2011 tsunami. In the European Union, car sales will fall for the sixth year in a row: they are now back at early 1990s levels. Although some European car factories face closure, elsewhere assembly lines are being built at a rapid clip. So once again worldwide car production, at 82.8m, will exceed sales, at 81.9m. As the metal stacks up on dealers' forecourts, motorists can look forward to some great deals on wheels in 2013.

Figure 2.4: *An example of an MSNV document with multiple references, with the first two underlined for easier identification. Image courtesy of Lallé et al. [6]*

based on a given financial scenario to represent using tokens. The video of the whole user study process and the photos of visual designs are captured. Also, participants are interviewed on how they created designs to collect more information about construction process. By examining the collected data, the visual mapping process was analyzed as three activities: construction, computation, and storytelling. They provide details of the logical tasks and actions of visual mapping (e.g. build data, build and combine, construct etc.).

Lallé *et al.* [6] investigate gaze-built adaptations as a way of promoting the production of visualizations in narrative text, known as the Magazine-Style Narrative Visualization (MSNV) and focus on the MSNVs with bar charts and one of the most widely used visualizations found in MSNV documents: newspapers, blogs and textbooks (see Figure 2.4). They also explore the possible value of long-term user characteristics in order to further customize the delivery of gaze-driven adaptation in MSNVs.

In order to assess the gaze-driven adaptive highlights for MNSVs, they compare the output of the group of users who read MNSV (see Figure 2.4) with the highlights of the intervention (adaptive group) and the control group that reads the same MSNV without highlights. They used a group of 14 bar chart-based MSNVs, extracted by Toker *et al.* [94] from the existing 40 MSNV datasets, e.g. Pew Research, The Guardian or The Economist. In total, 119 individuals were recruited through advertisements on campus and on the Craigslist (63 for adaptive study, 56 for control

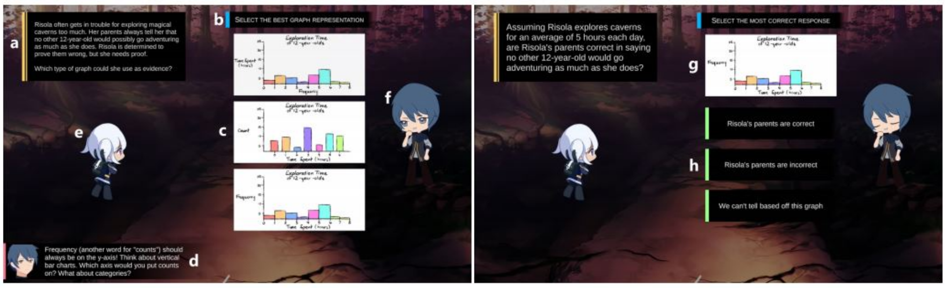


Figure 2.5: *Activity View. This view consists of a question display box (a), instructions (b), choices to select from (c), a feedback box (d), and character sprites (e, f). Follow-up "interpret the chart" questions include a picture of the previously-chosen chart (g), and choices of written answers (h). Image courtesy of Huynh et al. [7]*

study). The 90 minute session begins with participants receiving an eye-tracker calibration. Next, participants are provided the assignment of reading the MSNV document on the computer screen and signaling when they are finished by pressing 'next'. Participants see a monitor with a collection of questions that reveal their view of the document and test their understanding.

Five user characteristics are specified in order to test participants. The first characteristic, *visual literacy* relates to how well users can process visualization. The other three *verbal WM, reading abilities, and verbal IQ* relate to the ability to process textual elements. The last one, *for need cognition* is a personality trait that determines how much users like cognitive activities. The results indicate that the interventions enhanced the performance of users with low visualization literacy, while the interventions did not affect high literacy users. Barral *et al.* [89] expand this research by identifying the particular visual behaviors that adaptive guidance produces in users, based on their level of visualization literacy. Barral *et al.* [90] extend their previous user study [6] by examining the speed with which users process newly triggered intervention in order to understand how effective interventions are to direct the attention of the user in the visualization and how this adaptive mechanism influences the users according to their visual literacy levels.

Huynh *et al.* [7] introduce a novel story-based role-playing game to facilitate visualization literacy education for children aged between 11-13 that features a series of exercises, overarching narratives, and mechanics to reinforce narrative elements. The game contains educational activities that focus on pie charts and histograms presented as multiple-choice questions, consisting of a chart question followed by an interpretation. The game consists of three key views: the Action View showing

Literature	*Data Sources*
Baker *et al.* [82]	Global properties, Brand Price and Quality
Delmas *et al.* [83]	Baseball, Food Cost, Duration of Sleep, etc.
Schönborn *et al.* [18]	Biochemistry Education
Grammel *et al.* [77]	Sales Data
Kodagoda *et al.* [78]	Social Service Website
Boy *et al.* [37]	Monthly Unemployment Rates
Huron *et al.* [84]	Fictional Financial Data
Ruchikarhorn and Mueller [15]	Energy, Time
Maltese *et al.* [39]	Average Temperature, Greenhouse gases, Tornado Events, etc.
Kwon and Lee [16]	Features of Cars
Börner *et al.* [5]	Sources of Nitrogen, Energy Consumption, US Unemployment, US Population, etc.
Alper *et al.* [8]	Flowers, Animals, Ingredients, etc.
Lee *et al.* [22]	Oil Price, Internet Speed, Cost of Food, etc.
Wojton *et al.* [85]	Health and Wealth of Nations, International Airports, Competitive Eating Records, etc.
Lee *et al.* [17]	Oil Price, Internet Speed, Cost of Food, etc.
Fuchs *et al.* [9]	Shape , Points, Clustering Algorithms
Gäbler *et al.* [10]	Fictional Data, Fictional Characters
Bishop *et al.* [11]	Fictional Data
Lallé *et al.* [6]	Car Sales, Syrian Refugees, World Economy
Krekhov *et al.* [12]	Music, Eating, Screens, Water, etc.
Firat *et al.* [42]	Market, Earthquakes, Investment Funds, Medals, Health etc.
Wang *et al.* [13]	Olympics, Iris, Cliques, Clusters, or Bridges
Rodrigues *et al.* [14]	Mock Data
Huynh *et al.* [7]	Constellations, Fictional Data, Fictional Characters
D'Ignazio and Bhargava [87]	Data-driven Story
Donohoe and Costello [88]	Oil Price, Internet speed, Cost of Food, etc.
Barral *et al.* [89]	Car Sales, Syrian Refugees, World Economy
Barral *et al.* [90]	Car Sales, Syrian Refugees, World Economy
Peppler *et al.* [91]	Personal data

Table 2.5: A table indicate the data themes used in the literature.

the puzzles, the Dialog View and the Exploration View to support the narrative component of the game (see Figure 2.5). In order to provide a narrative component to the tasks, players are engaged in a game world where they can discover and find characters to communicate with them by dialogues.

To evaluate the effect of narrative elements in games based on visualization literacy, a study is designed to evaluate two game variations: one with and one without narrative elements. A total of 33 children aged 11–13 participate in the study. Participants are tested in independent sessions conducted in a lab with only the participant and investigator. The study is designed with four phases: pre-test, play time, post-test and interview. The participants were asked about their experience with pie charts and histograms in the pre-test phase. In play time, participants were all provided with the same activities and randomly allocated to either the non-narrative (i.e. activities only) presented in sequence without additional context) or with-narrative (i.e., activities, exploration and dialogue) context. In the post-test, participants are provided a paper-based test to evaluate changes in their understanding of the subject. Then, participants are asked to express their thoughts on the different aspects of the game in the interview phase. The findings indicate that the

Literature	***United Classroom***	***Divided Classroom***	***Education Level***
Baker *et al.* [82]	✓		High School
Delmas *et al.* [83]		✓	High S., Higher E.
Alper *et al.* [8]	✓		Primary School
Fuchs *et al.* [9]	✓		Higher Education
Gäbler *et al.* [10]	✓		High School
Bishop *et al.* [11]	✓		Primary School
Krekhov *et al.* [12]	✓		Higher Education
Firat *et al.* [42]		✓	Higher Education
Wang *et al.* [13]	✓		Higher Education
Rodrigues *et al.* [14]	✓		Higher Education
Total: 10	8	2	

Table 2.6: A classification table of studies that use a classroom evaluation approach. Literature is classified as using a united (the entire classroom of students) or divided (the classroom is divided in half: a control group and an experimental group). Participants' education level (primary school, high school, or higher education) is provided in the table.

narrative elements provide a substantial positive effect on children's interaction and enjoyment although it requires players to spend much more time engaging with elements. Interviews reveal that children in the narrative condition setting are usually satisfied with the story and related interactions.

Data Sources: Table 2.5 provides an overview of the data sets that are displayed and used in the literacy evaluation in the literature. The data sources span a very wide breadth of different subjects and categories and do not show convergence on any particular subjects. While some fictional data is chosen for a few studies, most of the selected data sets are non-fictional based on convenience that can be easily accessible online. The table does not indicate any special data source theme that the researchers studied in the visualization literacy field. For example, Huynh *et al.* [7] introduces role-playing games which include asking questions using fictional data. Börner *et al.* [5] present test questions shows data on energy consumption and US population to assess museum visitors' literacy level while the study by Fuchs *et al.* [66] is an example to a special case in which the data set used to increase users' understanding of clustering algorithms.

2.3.3 Classroom-Based Evaluation

In a classroom setting, researchers design tests for pre- and post-experiments and investigate the visualization literacy skills of users based on participants' answers to questions. In this category, a cohort of participants carry out an experiment as a group simultaneously, usually in a classroom. Preparing questionnaires to ask in pre- and post-experiments in a classroom environment is the most popular evaluation method among all categories. Papers by Alper *et al.* [8] and Fuchs *et al.* [9] present examples of a classroom-based assessment.

Table 2.6 displays a summary of studies that use a classroom evaluation approach. Evaluation categories are further sub-divided according to the classroom evaluation method. In some cases, the entire class experiences the same education: pre-test, a new educational technology, and a post-test. We call this a united evaluation. In other evaluations, the classes are split in half. The whole class takes the same pre- and post-tests. However one half of the class is taught the traditional way, while the other half uses new visualization technology. We call this a divided classroom evaluation. Table 2.6 indicates that researches mainly prefer the united classroom approach for the experimental setting. We also provide the education level of the participants (primary school, high school, or higher education) involved in the study.

Baker *et al.* [82] investigate middle school students' reasoning abilities with three graphical representations: histograms, scatterplots, and stem-and-leaf plots. They run an experiment to see how novices perform when it comes to interpreting, generating, and selecting visual representations. In the study, 52 students from grades 8 and 9 completed 3-4 exercises in which they were asked to draw a histogram, scatterplot, or stem-and-leaf plot, in response to a set of interpretation questions for each visual design or to select the most appropriate representation for a particular question.

Overall, students performed moderately well on graph interpretation, with an average of 56% correct answers. Student performance on graph selection and generation is quite poor. In graph selection, students performed no better than 25% correctness. They were also unsuccessful in producing histograms and scatterplots. The performance of the 15 students to generate stem-and-leaf plots was relatively poor, with only 20% of them scoring completely correct. This result, however, was significantly better than the performance of the 12 students who generate histograms and the 12 students who generate scatterplots (0%).

Delmas *et al.* [83] define graph comprehension as the ability to translate a graph or a table and being able to interpret connections or major elements in a graph. The focus is to evaluate learning results in a first statistics course through the As-

Figure 2.6: *Deployment in grade 2 showing the setup in the classroom, discussions between students and written activity. Image courtesy of Alper et al.* [8]

sessment Resource Tools for Improving Statistical Thinking (ARTIST) [95] project over three years. The project is designed to address the challenge of evaluation difficulty in statistics education. In addition, the project team develops an overall Comprehensive Assessment of Outcomes in Statistics (CAOS) that includes a group of multiple-choice items to assess student's comprehension and reasoning on the topic of variability when interpreting distributions, and making comparisons.
The evaluation data is collected by testing high school and college students. A group of 909 students take the CAOS test (97 students from high schools and 812 students from universities), and 555 students take one or more of the ARTIST topic scales (205 students from 4 high schools and 350 students from universities). All questions are multiple choice. Results indicate that students do not have difficulty understanding simple histograms and matching different graphs of the same data, as long as they have clear features to guide them. When students are asked to coordinate more information, the matching is more challenging. Students perform well when matching graphs to the definition of variables. Participants also show difficulty in many aspects of reasoning about images of distributions. They mainly have difficulty reading the data when the bars contain intervals of values rather than single values of a variant.

Alper *et al.* [8] investigate visualization literacy teaching methods for elementary school children and present an online platform C'est La Vis, that enables students to create and interact with visual data representations. It is used by instructors in the classroom by creating exercises for children (see Figure 2.6). Alper *et al.* [8] provide the results of an investigation of visualization types taught in grades K-4, in a formative study. They analyze visuals designs included in elementary textbooks and study textbooks that follow the US common core standards. These include five math eTextBooks from the Go Maths collection, six French by Éditions Hatier

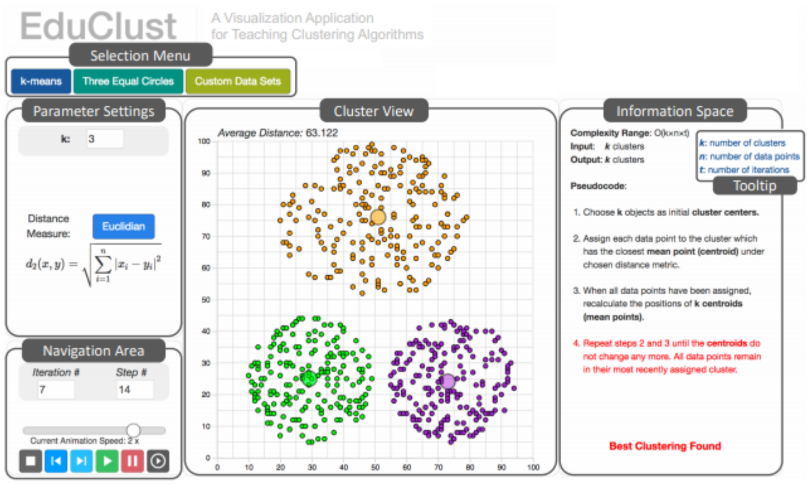


Figure 2.7: *The five main components of EduClust visualization application. Image courtesy of Fuchs et al. [9]*

and eight Turkish elementary math textbooks provided by the Turkish Ministry of Education [8].

Students interacting with the tool are evaluated in a field study that aims to understand their interest and understanding of the exercise and to collect feedback from the teachers on how the tool enhances current teaching in the classroom. Some 15 students, split into small groups, from two classrooms (grades K and 2) have their activities observed. An observer takes notes during the sessions with C'est La Vis, occasionally asking or answering questions from students. The main goals are to understand touch interactivity, verbal activity and class dynamics. Observers reported 13 students interacting with the app as playing a game rather than learning. A selection of 6 students also verbalize visualization literacy concepts (how to read an axis), and they are generally willing to use the app. Also, 16 teachers are surveyed to identify educational strategies for teaching simple visual designs. As a result, a set of design goals are provided to enhance visualization literacy in early grades.

Fuchs *et al.* [9] develop EduClust, an online application that assists both learning and teaching of clustering algorithms. This application combines visual designs, interaction, and intermediate clustering results to facilitate the comprehension and teaching of clustering algorithms. A web-based tool is developed that enhances the teaching and learning of clustering algorithms for both lecturers and students (see Figure 2.7). The tool facilitates rapid exploration for quick understanding of clustering algorithms with interactive data sets. EduClust is evaluated in a classroom environment where participant feedback is collected.

Figure 2.8: *An example of a quiz asking about which of the three pie charts corresponds to the bar chart. Image courtesy of Gäbler et al. [10]*

Students are shown different clustering algorithms in the classroom to answer a questionnaire based on their experience of the application. The results of the feedback indicate that 50% of the class strongly agreed that EduClust helps them comprehend the clustering algorithms. Also, 47% of students strongly agree that they would benefit from the tool for exam preparation. Moreover, students are asked to share their thoughts on the current implementation of the different algorithms and whether it would advance their comprehension of the algorithmic behavior. Some 22 students (77%) agreed or strongly agreed with this statement while five students (18%) were neutral in their decision and one student (4%) disagreed.

Gäbler *et al.* [10] developed Diagram Safari, an interactive mobile game for teaching diagrams and charts to children aged 9-11. The game is about learning how to construct bar charts, how to read and interpret them and how to match them to pie charts. In the game, the player navigates a ball across a bar chart by adjusting the bar height. The game includes numerous challenges in the form of quizzes, interactions between drag and drop, and it is designed in a visually appealing format for children (see Figure 2.8).

The game is tested with 23 children in the fourth grade of primary school to obtain initial input from the target group evaluating the game design and playability. First, the children complete a questionnaire about their diagram literacy. Then, four tasks are given that require the bar chart to be assigned to the appropriate pie chart (see Figure 2.8). This is used to assess the ability of bar charts to correlate with pie charts. The kids play the game for about 15 minutes. In the last step, they are asked to complete a second questionnaire that examines their perspectives on the

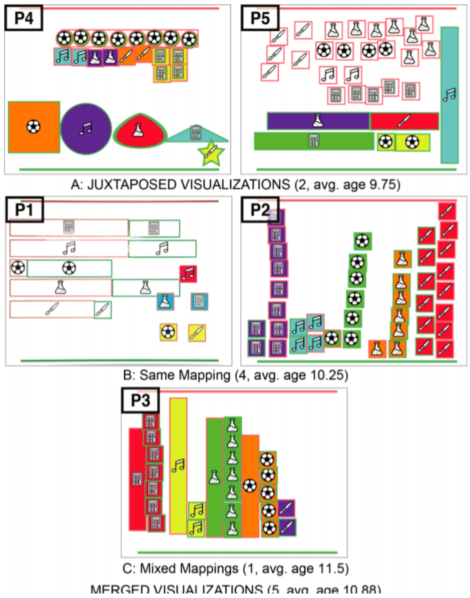


Figure 2.9: *Linked visual representations created by pairs in Task 3. Image courtesy of Bishop et al. [11]*

game. This questionnaire includes "Do you like the game", "Would you like to play the game at home?". The result indicates that children assigned on average 3.76 of 4 (%94) bar charts correctly to the corresponding pie chart in the first step. After playing the game, they assigned 5.76 out of 6 (%96) bar charts correctly. The study also indicates the children's interest in the subject of diagrams and that they are motivated to continue playing.

Bishop *et al.* [11] investigate the visual reasoning processes of children while they participate in free-form visualization (not regular visual designs) practices with Construct-A-Vis, a tablet-based, free-form visualization application. The tool is designed to (1) help the development of free-form visualization, (2) make data-to-visual mappings clear, (3) offer visual input for scaffolding, (4) provide functionality for various ability levels and (5) facilitate shared activities. A qualitative study including three tasks increasing in complexity is conducted to test the Construct-A-Vis tool with 24 elementary school students aged 5-12. The students are asked to visualize a fictional data set about school subjects (maths, music, sports, science, arts, represented as icons) using the tool. The study focuses on examining whether children construct meaningful visualizations with Construct-A-Vis without formal instructions or models, the types of processes that children adopt while making individual and group visualizations, and how the tool encourages active participation in children's data visualization.

In Task 1, children are provided icons, color, and the shape for their visual designs. The task is designed to promote specific mappings between data points and symbols by providing icons that correlate directly to school subjects. Task 2 requires children to use size and color for their representations. The aim is to lead children toward abstract mappings using size for coding values. In Task 3, the children are provided with an additional shared tablet and told to jointly create a single representation showing an overview of their combined data sets (see Figure 2.9). The purpose is to support the communication to children. The results indicate that children are engaged in the visualization process and that processes lead to effective discussions and behaviors.

The paper by Krekhov *et al.* [12] seeks for an opportunity to develop computer graphics and visualization courses in a way that would allow students to create visualizations that are easy to understand, engage the students, and memorable. The design of the course is especially inspired by the book Dear Data [96], in which visualization was generated by the composing of visualization information and creativity. The purpose of the research is to enable participants to focus on design thinking and hands-on exploration of the visualization without being compelled to proceed in a linear manner that is often prescribed by existing tools.

The paper presents a 12-week teaching experiment and designs a course curriculum. The purpose of the course how to transform various datasets into useful and engaging visualizations for a wider audience. The course is divided into the following components: understanding data, visualizing data, and design thinking. For each session, students are assigned the task of creating an appealing, detailed visualization based on the data they are given. Topics include such as water, music, eating (see Figure 2.10). After the submission of the design, students present their outcomes at weekly meetings. They learn more from feedback received from the lecturers and from the other outcomes.

Some 11 students participated in the study who identified themselves as being beginners in the field of visualization. During the first 6 weeks, subjects worked on their own, while in the second half, they are divided into groups of 2-4 participants and asked to construct a single collective visualization. During the course, participants were surveyed and an online questionnaire was presented at the end of the

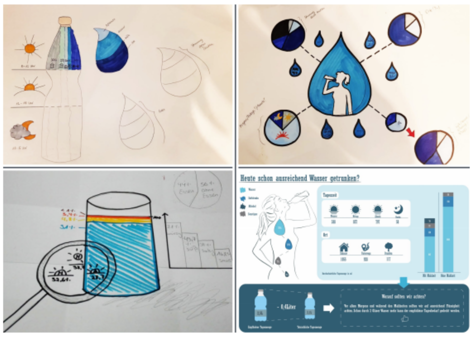


Figure 2.10: *An example design of the student's early and intermediate prototypes. Image courtesy of Krekhov et al. [12]*

experiment. The qualitative results reveal that the concept of thinking motivates novices to experiment with a wider variety of visualization methods. Ten students state that they formulated the main message or purpose of their visualization during or before data collection and mention that the most important factors to a good visualization are the appeal, metaphor, and comprehension.

Research by Firat *et al.* [42] identifies the barriers and challenges of understanding and creating treemaps by examining the results of two years of an information visualization assignment. In order to assess the barriers, a treemap visualization literacy test is developed. In addition, a pedagogical tool that advances treemap visualization literacy is introduced. Firat *et al.* [42] conducted a classroom-based experiment with 25 computer science students from undergraduate and postgraduate levels to evaluate the participants' treemap literacy and the effectiveness of the treemap tool. Participants are assigned into one of two groups and both groups answer the pre-test treemap questionnaire. In the first session, one group experiences traditional treemap teaching using slides. Another group is introduced to the interactive treemap tool. On completion of teaching, all participants answer the post-tutorial and interview questionnaires. The results show that students who attended the slide session answered on average 79% of the post-intervention test questions correctly, while the students who attended the software session answered 89% correctly. Also, participants' feedback indicates that the treemap software offered them an effective learning experience.

Wang *et al.* [13] present the notion of 'cheat sheets' to support data visualization education. Cheat sheets are a collection of graphic descriptions and text annotations, like data comics. Cheat sheets enable a broad audience to understand the data

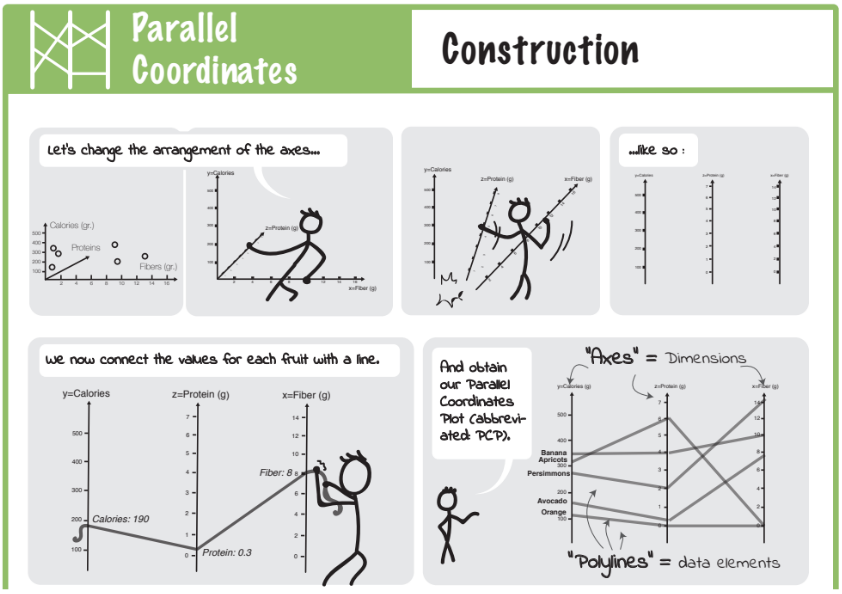
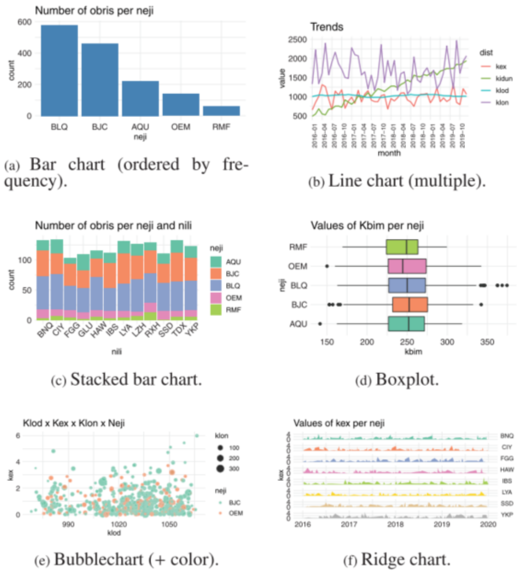


Figure 2.11: *Part of Construction for PCP, showing its "creation" from three-dimensional scatterplot in a comic-strip. Image courtesy of Wang et al. [13]*

visualization techniques and support two scenarios (i) first-time learning assisted by slides, posters, books, or activities; and (ii) as testing material during the actual data discovery process.

Six forms of cheat sheets are introduced: Anatomy, Construction (see Figure 2.11), Well-Known Relatives, Visual Patterns, Pitfalls, and False-Friends. Cheat sheets are a combination of six forms and describe individual aspects of visualization techniques. Cheat sheets types are organized for a specific presentation purpose: by type and by technique.

Wang *et al.* [13] conduct a user study with participants from a local university. Answers from 11 participants who reported that they are novice with visualization are analyzed. For each of the three techniques (boxplots, PCP, matrix), a cheat sheet is produced. Participants are provided a visualization example of a given technique and asked to respond to a small quiz to evaluate their understanding. During the quiz, participants receive printouts of anatomy, visual patterns, and pitfalls. Next, participants are asked to think a loud and write down the content that was unclear. Finally, participants receive a questionnaire asking them to rate comprehensibility, aesthetics, usefulness, etc. The results indicate that novice participants liked and considered the cheat sheets useful for improving comprehension of complex visualizations and that the development of cheat sheets facilities understanding of novel techniques.



(a) Bar chart (ordered by frequency).
(b) Line chart (multiple).
(c) Stacked bar chart.
(d) Boxplot.
(e) Bubblechart (+ color).
(f) Ridge chart.

Figure 2.12: *A sample of the visual designs used in the study presented in the paper by [14]. Image courtesy of Rodrigues et al. [14]*

Rodrigues *et al.* [14] analyze individuals' initial questions to ask when they first experience a visualization. In this way, they understand the common mistakes individuals make when asking data related questions in an attempt to interpret the data. They designed a study to gauge the quality of data-related questions generated by individuals with low visualization literacy skills when they are shown different types of visual design. A group of 22 participants from graduate and undergraduate studies who self-assessed their prior (no or limited) experience of data visualizations are involved in the study. The research is performed through an online questionnaire, which displays visualizations in random order. Each participant is asked to generate up to five questions about the underlying data that could be answered by analyzing the images. For each question, they are asked to indicate the amount of effort needed to produce the question.

For the study, 20 visualizations (bar chart, heatmap, chord, Sankey, network, histogram, scatterplots, etc.) are created with a dataset made of variables given meaningless names (e.g. klon, neji) (see Figure 2.12). The questionnaire collected a total of 1058 responses. The responses are examined by three researchers to created standardized versions of the questions as an attempt to reduce the number of unique questions. The 8000 clear questions are classified as 'OK' and 250 problematic questions are classified as 'problem'. The clear questions are reviewed to describe the different types of questions people can answer through each visualization. The prob-

Literature & Visual designs	*Bar Chart*	*Scatterplots*	*Line Chart*	*Pie Chart*	*Histogram*	*Choropleth Map*	*Stack Chart*	*Treemap*	*Illustration*	*Parallel Coord.*	*Network*	*Pictograph*	*Matrix*	*Boxplots*	*Spiral Chart*	*Buble Chart*	*Chord Diagram*	*Sankey Diagram*	*Stem and Leaf*	*GUI*
Baker *et al.* [82]																				
Delmas *et al.* [83]																				
Grammel *et al.* [77]																				
Kodagoda *et al.* [78]																				
Boy *et al.* [37]																				
Huron *et al.* [84]																				
Ruchikarhorn and Mueller [15]																				
Maltese *et al.* [39]																				
Kwon and Lee [16]																				
Börner *et al.* [5]																				
Alper *et al.* [8]																				
Lee *et al.* [22]																				
Wojton *et al.* [85]																				
Fuchs *et al.* [9]																				
Lee *et al.* [17]																				
Gäbler *et al.* [10]																				
Bishop *et al.* [11]																				
Lallé *et al.* [6]																				
Krekhov *et al.* [12]																				
Firat *et al.* [42]																				
Wang *et al.* [13]																				
Rodrigues *et al.* [14]																				
Huynh *et al.* [7]																				
D'Ignazio and Bhargava [87]																				
Donohoe and Costello [88]																				
Barral *et al.* [89]																				
Barral *et al.* [90]																				
Peppler *et al.* [91]																				
Total	15	13	9	8	8	8	7	5	3	3	3	3	2	2	1	1	1	1	1	1

Table 2.7: An overview of the literature and the visual designs evaluated. The table displays the type of visual designs tested by column and chronically sorted literature on the y axis. Each individual paper is colored according to evaluation techniques used: blue: In the wild, pink: Controlled User-study, orange: Classroom-based, green: Crowdsourcing. Literature review papers are left out of this table.

lematic questions are examined and classified to understand which kinds of problems in the questions occurred more often for each type of image. The findings of the study can be an important source for teaching visual designs as they reveal and identify common errors that individuals make when thinking about visually presented data.

Table 2.7 provides an overview of which visual designs are included in the evaluation in previous visualization literacy studies, as well as expressing the evaluation methods in the studies using color according to our main classification (see Table 2.3). The table indicates that the most evaluated designs in literacy are bar charts and scatterplots. In contrast, images such as bubble charts, spiral charts, sankey diagrams, and chord diagrams have only been evaluated in one study.

2.3.4 Visualization Literacy and Crowdsourced-Based Evaluation

Some studies prefer to conduct experiments using an online platform to recruit a large number of participants from a geographically diverse pool of participants. Crowdsourcing using Amazon's Mechanical Turk (MTurk) offers access to a great number of participants at affordable prices for collecting data in a relatively short

Literature & Crowdsourcing Platforms	***Amazon Mechanical Turk***	***Other***
Boy *et al.* [37]	✓	
Ruchikarhorn and Mueller [15]	✓	
Maltese *et al.* [39]		Online Test
Kwon and Lee [16]	✓	
Lee *et al.* [22]	✓	
Lee *et al.* [17]	✓	
Donohoe and Costello [88]		Online Test
Total: 7	5	2

Table 2.8: A summary table of literature that presents experiments carried out utilizing crowdsourcing platforms. The table indicates studies which experiments use the most popular platform, Amazon Mechanical Turk, or an online test designed by authors for gathering responses.

period of time. For example, Kwon and Lee [16] and Boy *et al.* [37] chose to engage participants and designed online experiments using MTurk. Table 2.8 summarizes visualization literacy literature that designs experiments carried out utilizing crowdsourcing platforms or create online tests for sharing with crowd. The studies are grouped according to the type of platform used for collecting participants' responses. We can see that Amazon Mechanical Turk is a popular platform for crowdsourced studies.

Boy *et al.* [37] aim to develop a method for visualization literacy evaluation. They use the Item Response Theory (IRT) [38] to separate the impacts of item difficulty and examinee ability. The main purpose is to create fast, efficient, and reliable tests that researchers can use to identify test takers with lower visualization literacy ability. The tests are developed based on a 3-part structure. These are a stimulus, a task, and a question. The stimuli are the selected visual designs being studied. Tasks are defined based on the visual operations and mental projections that a participant performs when answering a given question. Tasks and questions are linked. This distinction is emphasized because different orientations of a question could influence participants' performance.

They focus on tasks that require only basic intelligence, such as identifying minimum, maximum, variation, intersection, average, and comparison. They test the user's ability to find these characteristics on line graphs using Amazon Mechanical Turk with 40 participants. They also perform experiments using bar charts and scatterplots. The results indicate that IRT modelling is useful for differentiating and assessing visualization literacy, especially for examinees with lower ability.

The aim of the research paper by Maltese *et al.* [39] is to examine differences in data visualization ability along a spectrum of expertise from novice undergraduate students to STEM practitioners to gain a better understanding of how users interpret graphical representations of data. The study reports on the design of the data visualization and evaluation results. In order to collect data, participants respond to questions while viewing given graphs and tables to test their ability to read and interpret them. Task performance data is collected from teaching staff and doctoral students with a range of science expertise in science education. Some 19 of 20 core test items were visualizations from widely published textbooks, government websites, or published reports.

Maltese *et al.* [39] conducted an analysis to better understand the psychometric features of the items (internal consistency for dichotomous items, item difficulty, item discrimination) used in the study evaluation. Some 202 participants, mainly university and college graduates (68%) and graduate students (9%) participate the study and report the average number of STEM classes that they completed. Their scores from an online assessment of the 20 test items range between 6-18 correct answers. A reasonable correlation was found between the number of STEM coursework participants completed and their performance, but overall this relationship is not strongly positive. The findings indicated that even participants that completed advanced science and mathematics coursework found it difficult to interpret basic data representations.

Ruchikachorn and Mueller [15] present a learning-by-analogy technique that explains an unfamiliar visualization method by showing a step-by-step conversion between two visual designs. The research shows the concept using four visualization pair examples such as a data table and parallel coordinates, a scatterplot matrix and hyper box, a linear chart and spiral chart (see Figure 2.13), and a hierarchical pie chart and Treemap. The participants understand the uncommon visual designs more quickly after they interact with the transitions.

In the first stage of evaluation, a short task and questionnaire are prepared to test 22 participants via Amazon Mechanical Turk. The pair of linear and spiral charts are chosen. Eleven participants are shown the spiral chart and 11 are shown the linear chart. Results indicate that only half of the spiral chart answers are correct, while all answers on the linear chart are correct.

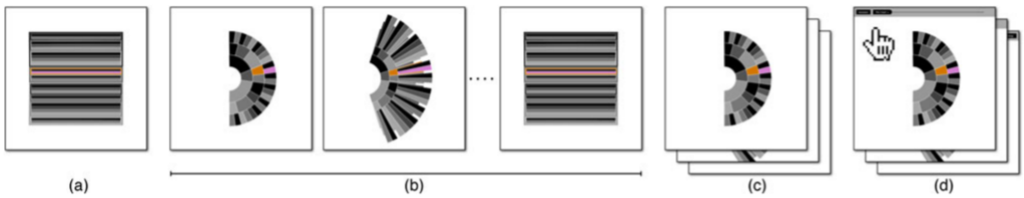


Figure 2.13: *A pie chart–treemap pair example: (a) target visualization, (b) series of the intermediate images from the source visualization, (c) animation, and (d) interactive visualization. Image courtesy of Ruchikachorn and Mueller [15]*

For the first main study, six male and five female participants are asked about their background and are given the source visualization descriptions. Four visualizations and morphings are displayed and ordered as (a) the target visualization, (b) a series of the intermediate from the source visualization, (c) an animation, and (d) an interactive visualization (see Figure 2.13). This order is chosen to demonstrate the interaction cost from the smallest to highest. The order also indicates how much help participants need to comprehend and learn the unfamiliar visual designs. All participants describe the morphing as an effective tutoring way to understand new visual designs. A further 11 participants are chosen to test if they can read the underlying data and view the trend on the target visualization. Participants were shown the target visualizations with different data sets before being asked open-ended questions about their comprehension of the data. Results indicate 7 out of 10 participants could read and provide observations from the target visualizations. The other participants had already seen similar visual representations prior to the experiment.

Kwon and Lee [16] focus on Parallel Coordinates, an efficient method to display multidimensional data, to study the impacts of multimedia learning environments for teaching data visualization to non-expert users. The inspiration behind this research is to examine active learning theory. Four experimental conditions are created: baseline, interactive, static, and video. The baseline condition contains a single-page description of how data is presented in parallel coordinates. In the interactive condition, the user can draw parallel coordinates by entering values and creating edges. The static condition displays instructions with screenshots taken from the interactive condition without providing feedback. The video provides screen activities of a walk-through of the activities in the Interactive mode, so it includes the same feedback. The other three conditions provide a description and a tutorial

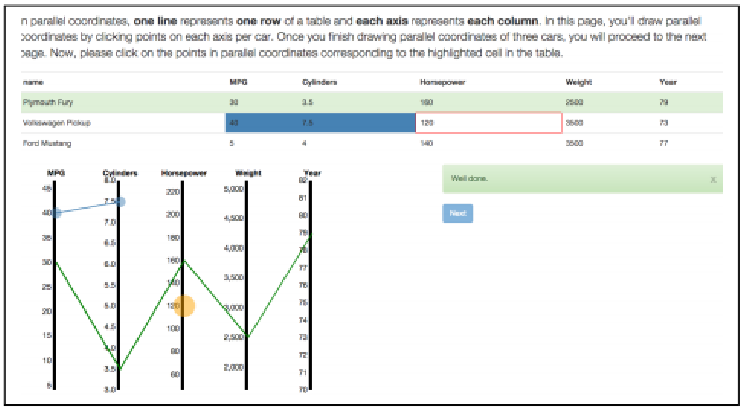


Figure 2.14: *The Build tutorial page: as people click on points in parallel coordinates, lines are drawn connecting them. Image courtesy of Kwon and Lee [16]*

using different media (see Figure 2.14).

An experiment was conducted on Amazon Mechanical Turk with 75 male and 45 female participants (30 people per condition). After the tutorial session, participants are asked 18 questions based-on tasks such as mapping between data points and visual elements, data distribution, comparison and similarities. They are also given 6 interview questions related to the tutorial. Results indicate that participants with the interactive condition perform better than the static and baseline conditions, and stated that they had a better experience than those with the static condition.

Lee *et al.* [22] develop a test to assess ordinary users' visualization literacy skills, especially users who are not experts in data visualization. Three different sources are examined: K-12 curriculum, data visualization authoring tools, and news articles in order to determine the content of the test. They organize a pilot study before generating the test items to analyze the usage of vocabulary and phrases when test takers read and interpreted the data visualizations, which may influence test participants' performance. After developing a group of test items, domain experts review them to ensure the test contains appropriate contents and tasks. A total of 191 participants (MTurk) consisting of 105 females and 86 males with an age range of 19-72 take the visualization literacy test. The test includes 54 test questions composed of 34 four-option multiple choices, 3 three-option multiple-choice, and 17 true-false questions. Based on the results, all the items are reviewed in order to eliminate inappropriate items and finalize test items for the Visualization Literacy Assessment Test (VLAT). A final experiment is preformed with finalized VLAT test item choices. A total of 37 people (MTurk) 14 females and 23 males in the age range of 22-58 participated in the

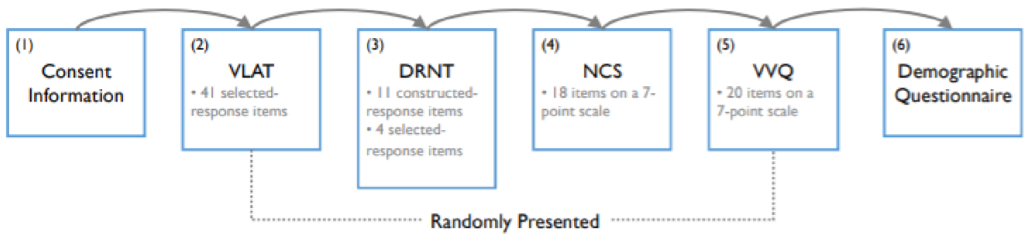


Figure 2.15: *The experiment procedure that consists of six stages. Stages 2, 3, 4, and 5 were randomly presented to the participants. Image courtesy of Lee et al. [17]*

study. The experiment is designed to measure visualization literacy and the ability to learn an unfamiliar visualization. Participants completed 53 questions and were redirected to a Parallel Coordinates Plot (PCP) test with an online learning tutorial developed by Kwon and Lee [16]. After the tutorial material, participants are asked to answer 13 test items related to PCPs. The result shows that visualization literacy is positively linked with the users' ability to learn an unfamiliar visualization.

The research by Lee *et al.* [17] aims to find the connection between visual literacy and the following three cognitive characteristics: numeracy as cognitive ability, cognitive motivation and cognitive style. An experiment with 178 participants using MTurk is conducted. Participants are evaluated against four categories: visualization literacy, numeracy, need for cognition, and visualizer/verbalizer. Participants are asked to complete four assessments: a Visualization Literacy Assessment Test (VLAT) [22], Decision Research Numeracy Test (DRNT) [97], Need for Cognition Scale (NCS) [98], and Verbalizer-Visualizer Questionnaire (VVQ) [99] (see Figure 2.15).

For example, VLAT asks the participants to choose, within a time limit, the best response for each item. The DRNT asks participants to respond quickly and accurately rather than with a time limit. The NCS and the VVQ ask participants to show in what manner each object represents its cognitive features on a 7-point scale ranging from a strong disagree to a strong agree. At the end of the experiment, the participants are required to complete a demographic questionnaire. Results indicated that an individual's numeracy and need for cognition are positively correlated to individual's visualization literacy. However, the visualizer-verbalizer cognitive style did not indicate a correlation with visualization literacy.

Research by Donohoe and Costello [88] evaluates participant's perceived utility and

Literature & Contributions	*VL Test*	*Vis Tool*	*Other*
Baker *et al.* [82]			Results of novices' performance on VL
Delmas *et al.* [83]	✓		Results of assesment of reasoning visual designs
Grammel *et al.* [77]			A model to identify barriers, Implications for tool design
Kodagoda *et al.* [78]			Identification of barriers, Design principles for novices
Boy *et al.* [37]	✓		Assessment of examinee's VL level, Definition of VL
Huron *et al.* [84]			Visual mapping model, Implications for tool design
Ruchikarhorn and Mueller [15]			Introducing a familiarity concept
Maltese *et al.* [39]	✓		Evaluation result of an assessment test
Kwon and Lee [16]		✓	Testing result of the active learning theory
Börner *et al.* [5]	✓		Testing result of familiarity with visual designs
Alper *et al.* [8]		✓	Identification of barriers to VL and design principles
Lee *et al.* [22]	✓		Evidence for the validity of the test
Wojton *et al.* [85]			Introducing the simplicity-familiarity matrix
Lee *et al.* [17]			Relationship between VL and cognitive characteristics
Fuchs *et al.* [9]		✓	Result of use-case scenario
Gäbler *et al.* [10]		✓	Result of evaluation and playtesting a game
Bishop *et al.* [11]		✓	Insights into the design of free-form visualization
Lallé *et al.* [6]			Investigation of gaze-driven adaptation in narrative visualization
Krekhov *et al.* [12]			Insights into the thinking process and the visualization pipeline
Firat *et al.* [42]	✓	✓	Identifying barriers to treemap literacy, Result of tool evaluation
Wang *et al.* [13]			Introducing visualization teaching concept
Rodrigues *et al.* [14]	✓		Results of novices' performance on assessment test
Huynh *et al.* [7]		✓	Evaluation of narrative elements in a game
Donohoe and Costello [88]	✓		Evaluation of users' VL level
Barral *et al.* [89]			Investigation of gaze-driven adaptation in narrative visualization
Barral *et al.* [90]			Investigation of gaze-driven adaptation in narrative visualization
Peppler *et al.* [91]		✓	Exploration design aspects to support VL

Table 2.9: An overview table summarizing the contributions of the literature for research purpose on visualization literacy (VL). Contributions in the papers are classified based on common themes: 1) creating a VL test to evaluate user's VL level, 2) developing a visualization software or game to support VL 3) or other. The rest of contributions are briefly explained in the *other* column. The number of citations of each paper is also shown. Literature reviews papers are left out of this table.

confidence in understanding visual designs by modifying current research tools used in other studies.

A questionnaire is designed that consists of two questions on perceived usefulness [100] and two modified skills questions to test participant's perception of their peer's literacy level and evaluate their perceived skill. These questions are followed by 24 multi-choice test items from VLAT [22] covering six data visualizations based on eight tasks. The study is sent to 157 prospective participants and responses are returned by 32 participants (20.4%). The results reveal that visual designs are useful, but the goal of some data visualizations is not always understood. Findings also indicate that participants consider their data visualization literacy to be higher than their peers' assumption. In contrast to their high confidence, their literacy level was sometimes lower.

Table 2.9 summarizes the contributions provided in the literature and shows common research direction in the field. The main themes in visualization literacy literature are grouped: 1) tests that are created to assess users' visualization literacy level, 2) developed tools or games aimed at advancing user's visualization literacy level or support learning visual designs, 3) other. While the studies generally focus on examining the users' visualization literacy skills using a test and assessing the test results, on the other hand, the impact of the developed tools on the users is evaluated. The rest of the contributions are provided in the *other* column. The novelty in the literature includes the effects of tool designs, the results of the evaluation of users' visualization skills, and the identification of barriers to visualization literacy, etc. Kwon and Lee [16] introduce a parallel coordinates tool and a tool demo was used in an experiment. The paper discusses the test results of active learning theory.

2.3.5 Literature Reviews on Visualization Literacy

This category is intended to collect literature and does not feature any specific assessment methodology. The literature is summarized in the (see Table 2.10) with number of references provided. This provides a type of meta-assessment. A study by Chevalier *et al.* [40] is an example of a literature review.

Schönborn *et al.* [18] describe the value of visualization in biochemistry education and support the idea of teaching visual literacy and skills using visualization tools as key components of all education programs in biochemistry. A selection of 10 guidelines are introduced to encourage visualization and visual literacy in biochemistry education. At the molecular organizational level, students may need to translate a more practical electron micrograph of the binding complex from various representations of antibodyantigen binding ranging from a molecular representation to a stylized two-dimensional diagram or computer image (see Figure 2.16). This implies, among other things, that students may need to make sense of an abstract representation of a molecular phenomenon alongside stylised and concrete representations of the same phenomenon, something that students find very challenging. Therefore, students are required not only to translate between the macro, micro, and

Literature	***Number of References***
Schönborn *et al.* [18]	47
Chevalier *et al.* [40]	17
Zoss *et al.* [19]	37
Mansoor and Harrison [86]	21
Börner *et al.* [20]	72
Stoiber *et al.* [21]	55
Total: 6	

Table 2.10: The summary table of literature that introduces the number of references are provided in literature reviews on visualization literacy.

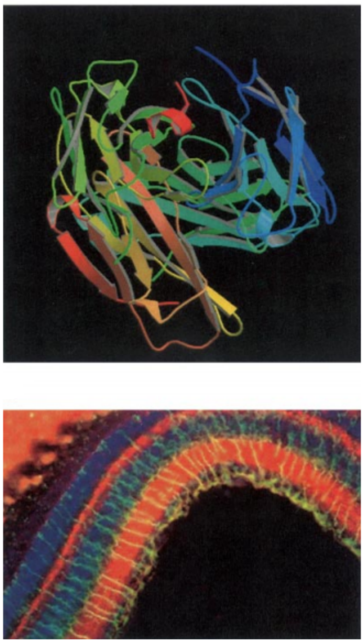


Figure 2.16: *Multiple ERs of antibody-antigen binding on a continuum from abstract to stylized (top) and to realistic (bottom). Image courtesy of Schönborn et al. [18]*

molecular levels of organization, but also between external representations (ERs) presenting phenomena at each level of abstraction, which becomes extremely cognitively challenging for students in combination. However, it would not be possible to interpret ERs without visualization and the associated processes and abilities of human imagination, studying, teaching, and analysis of molecular structure.

Another study by Chevalier *et al.* [40] presents an evidence-based discussion of visualization literacy and how it can be improved in early education and provides future research directions on visualization literacy. Chevalier *et al.* [40] investigate how children study visual designs and how their visualization literacy skills are improved at elementary school. They collected data from teachers on how much they use visual materials in class. Moreover, C'est la Vis, a tool that supports teaching and learning of pictographs and bar charts, is used to acquire data about child learning activities and interaction with the tool.

Three thought-provoking learning paradoxes arising from empirical information collected and observations in the field are described.

- Visuals are omnipresent in grades K–4.
- Teachers believe visual images are intuitive.
- Elementary students learn to read and create visual designs in early grades.

Figure 2.17: *Sample network visualizations, using a circular layout algorithm (a), a geographic layout (b), and a science map (c). Image courtesy of Zoss et al. [19]*

Moreover, three specific insights are derived and help inform the design of future visualization teaching materials for early education.

- Technology could curtail learning: Children interact with technology, especially when it features visuals and animations. Children may focus on solving exercises rather than concentrate on learning on underlying concepts.
- Technology could curtail social interactions: Teachers believe social interactions and verbal formulation, or reformulation, of knowledge acquired is necessary to the learning process.
- Technology can be too helpful, preventing beginners from practicing other abilities they need to obtain, and the advantages were sometimes skeptical.

Zoss *et al.* [19] define network visualization literacy (NVL) as the ability to read, interpret, and visualize different types of networks. In this paper, they provide on a series of topics that attempt to develop a more objective understanding for NVL including how to evaluate NVL, the role of NVL in teaching and learning, and suggestions based on understanding of the effective ways to enhance NVL. Challenges to interpret visualizations arise due to a lack of clarification about the limitations of network visualizations in the understanding of very complicated structures and the characteristics network components (see Figure 2.17). Zoss *et al.* [19] study some three aspects of NVL: *Representational Literacy, Metaphoric Literacy, and Topological Literacy.*

Zoss *et al.* [19] state that research is mainly based on experimental studies of the understanding of network visualization, restricted to particular tasks by design. For a better understanding of network visualization literacy, the visualization community must also take into account both how individuals interpret network images in

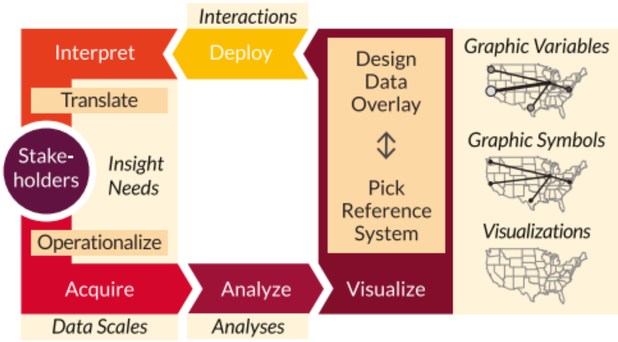


Figure 2.18: *Process of data visualization construction and interpretation with major steps. Image courtesy of Borner et al. [20]*

their everyday lives and how they acquire the skills required to create their own network representations. Thus, a mix of formal and informal education is necessary to enable more users to read and visualize network data. The paper presents three current approaches: Connections: The Nature of Networks (a public science museum exhibition at the New York Hall of Science), NetSci High (a research program for high school students), and the Information Visualization MOOC course at Indiana University. The paper also provides recommendations for improving network visualization literacy based on the review of relevant research and experiences with teaching and learning with network visualization.

Mansoor and Harrison [86] provide a case for combining parallel threads of data visualization literacy and visualization bias. The study address research in cognitive biases which claims that cognitive ability and experience can have an effect in how responsive a person is to a particular type of bias [101]. Mansoor and Harrison [86] review previous work on visualization biases to demonstrate how visualization literacy and biases may relate. For example, they cover research on attraction bias and availability bias by Dimara *et al.* [[102], [103]] and address how data literacy interventions potentially affect their analyses and resulting discussions. The paper also includes studies proposing the use of visualizations to mitigate bias, such as Dragicevic *et al.* [104], and demonstrates how results in visualization literacy [22] can facilitate their efficacy. These examples indicate that, as data visualization research continues to identify biases that occur in visualizations, the influence of individuals' abilities can be an significant factor for analysis and design.

Research by Borner *et al.* [20] provides a framework for data visualization literacy

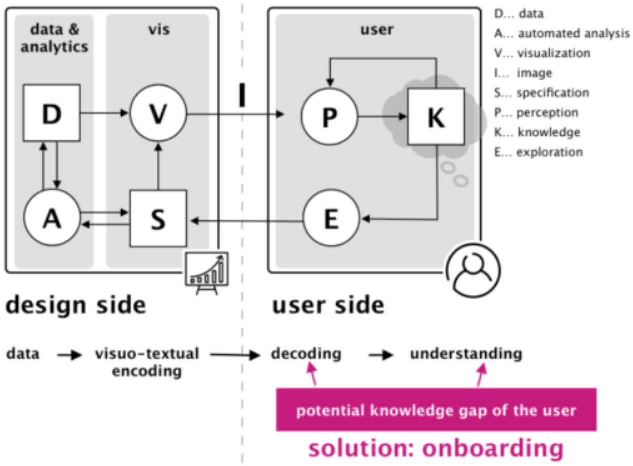


Figure 2.19: *Visualization onboarding in visual analytic system. Image courtesy of Stoiber et al. [21]*

(DVL-FW) that has been specifically developed to describe, teach and analyse DVL. The DVL-FW facilitates both reading and construction of visual designs, a pairing similar to that of both reading and writing in textual literacy as well as comprehension and application in computational literacy. Although DVL requires textual, mathematical and visual literacy skills, DVL-FW relies on key DVL concepts and procedural knowledge. The main process for the construction and interpretation of data visualization is defined and the process is interconnected with the typology of DVL-FW that contains 7 core types. These were derived from a broad literature review and collected from feedback gained from projects in the Information Visualization massive open online course [105]. Borner *et al.* [20] address the important role of stakeholders and describe the five steps (acquire, analyze, visualize, deploy, and interpret) (see Figure 2.18) of the system and their relationship to typology.

Borner *et al.* [20] introduce selected activities that support learning and evaluation of data visualization literacy (DVL) such as assessment of graphic symbols/variables knowledge, naming and classifying of visualizations, assessing students' ability to interpret visualizations and assessing students' ability to create and evaluate visualizations with practical assignments. Additional theoretical lectures and practical exercises are accessible online via the IVMOOC [105]. DVL-FW typology, activities and evaluations outlined in the paper have been applied in the Information Visualization course at Indiana University, supplying initial evidence that the framework can be used to instruct and test DVL. Information on student outcomes, success and feedback have directed the improvement of DVL-FW.

Another study by Stoiber *et al.* [21] introduces the design space of visualization onboarding and structures it along with the Five W's and How tool (see Figure 2.19). The process of assisting users in reading, interpreting, and extracting information from visual representations of data is defined as visualization onboarding (WHAT). This supports observers in dealing with large and complex information structures, to make visualizations more comprehensible (WHY). Another aspect is to determine the knowledge gap that the user has. The users' prior knowledge considered for developing onboarding concepts is presented such as domain knowledge, data knowledge, knowledge of visual encoding & interaction concepts, and analytical knowledge (WHO). Other relevant aspects of how visualization onboarding is implemented are defined as onboarding type, context sensitivity, and interaction (HOW). The visualization onboarding system can be integrated internally in the visualization or external source (WHERE). Visualization onboarding concepts can be connected before or during the use of visualization tools (WHEN).

2.4 Future Work

We examined each paper to identify common research areas that are discussed in each individual paper presented in Table 2.3 and summarize the common future research directions in Table 2.11. The summary facilities identifying a number of potential research areas in the scope of visualization literacy.

Further Evaluation: The most common future research goal identified in eight papers is to continue the investigation with new experimental settings including different parameters or materials with the aim of understanding barriers to visualization literacy. For instance, Firat *et al.* [42] recommend conducting more studies with a large and diverse group from non-computer science fields to reinforce the efficacy and to study the impact of participants' familiarity with the design. Similarly, to reproduce the findings, Rodrigues *et al.* [14] intend to perform a similar study with more participants while Gäbler *et al.* [10] focus on evaluating the effectiveness of game's education in terms of improving visualization literacy.

New Visual Designs: Much of the research uses specific visual methods (see Table 2.7) and targets incorporating various visual representations for further investigation on advancing visualization literacy. Boy *et al.* [37] aim to assess the suitability of

Future Works & Literature	Grammel *et al.*[77] [2010]	Kodagoda *et al.* [78] [2012]	Boy *et al.* [37] [2014]	Ruchikarhorn and Mueller [15] [2015]	Maltese *et al.* [39] [2015]	Kwon and Lee [16] [2016]	Börner *et al.* [5][2016]	Alper *et al.* [8] [2017]	Lee *et al.* [22] [2017]	Chevalier *et al.* [40]	Mansoor and Harrison [86] [2018]	Lee *et al.* [17] [2019]	Fuchs *et al.* [9] [2019]	Gäbler *et al.* [10] [2019]	Lallé *et al.* [6] [2019]	Krekhov *et al.* [12]	Firat *et al.* [42] [2020]	Wang *et al.* [13] [2020]	Rodrigues *et al.* [14]	D'Ignazio and Bhargava [87] [2020]	Donohoe and Costello [88] [2020]
Further Evaluation	X		X	X	X		X							X			X	X	X	X	
New Visual Designs			X				X	X					X		X			X			X
Improving Software	X	X											X	X			X				
Improving Literacy Test			X						X								X		X		X
Larger/New Target Group					X			X									X		X		
Studying Cognitive Impact										X	X	X									X
New/Larger Datasets						X											X				

Table 2.11: The table shows future research directions discussed for each paper. The directions displayed represent common research areas that reoccur in the literature and sorted according to occurrence frequency.

the approach studied for other forms of representation (e. g., parallel coordinates, node-link diagrams, star plots, etc.). Similarly, Lallé *et al.* [6] would like to study gaze-driven adaptation in MSNVs with different visual designs beyond bar charts. Wang *et al.* introduce *cheat sheets* and suggest including more examples of patterns and illustrations that will be combined with each cheat sheet as a part of an extension of their work.

Improving Software: Another common future work direction is developing the visualization tools introduced further by including new features to the applications to support visualization literacy. Fuchs *et al.* [9] suggest advancing the software into a framework for enabling users to upload new algorithms and visual designs. Kodagoda *et al.* [78] recommend investigating the design principles introduced, as the system may be appropriate for users with low literacy, but it may influence the performance of users' with high literacy in some cases. The future goals of Firat *et al.* [42] are to enable users to upload new hierarchical datasets and displaying new treemap layout algorithms in their pedagogical treemap tool.

Larger/New Target Group: In order to gain a better understanding of the visual literacy skills of individuals from various ages and backgrounds as well as achieving more reliable results, some papers suggest conducting experiments with larger or different target groups. Alper *et al.* [8] intend to explore how the approach presented in their paper can be adopted to instruct adult populations rather than children. The future focus of Maltese *et al.* [106] is to collect data from a broad sample of the population in order to increase understanding of the development of visualization literacy skills.

Studying Cognitive Impact: Individuals' cognitive differences affect visualization literacy skills. Chevalier *et al.* [40] recommend exposing examples of perceptual and cognitive biases that influence interpretation to raise awareness in this area. Similarly, Mansoor and Harrison [86] indicate that studying the relation between cognitive bias and visual literacy can be useful in the understanding of visual design. Lee *et al.* [17] suggest examining the impact of other cognitive characteristics on visualization literacy including personality, level of experience, and demographic information.

New/Larger Datasets: The type and size of a dataset plays an important role in individuals' comprehension of visual designs. For example, the size of the data is a barrier to treemap literacy [42]. Instead of using simple datasets, Kwon and Lee [16] suggest using real-world datasets that can uncover hidden insight that would require more particular expertise than learning the simple principles of parallel coordinates. The future work of Firat *et al.* [42] includes a wider variety of datasets in their literacy test.

2.5 Chapter Summary

In this chapter, we present a comprehensive review of the literature in visualization literacy, classifying the literature into five groups. We summarize each publication to provide insight into the study and presented guidelines for improving literacy skills. In addition, we provide tables that include extensive meta-data that summarize the characteristics in the literature and facilitate comparisons. We also reviewed and discussed common future directions in the paper collection. This complete review provides an essential and unique starting point for beginners and experienced researchers in the field.

Chapter 3

Treemap Literacy: A Classroom-Based Investigation

"Tell me and I forget. Teach me and I remember. Involve me and I learn."

—Benjamin Franklin, Politician (1706-1790)

Contents

Advanced data visualization techniques are often significantly challenging to comprehend and interpret. Based on our literature review on visualization literacy, we can see that one of the less studied visual designs is treemaps. Therefore, this chapter focuses on the treemap, one of the advanced visual designs for displaying hierarchical data. We aim to explore and examine the barriers to treemap literacy. We also introduce a treemap literacy test to evaluate users' literacy ability while attempting to improve understanding of the design with novel educational software. Finally, we utilize the literacy test and educational tool for a user-study and report its results. This chapter is based on a paper published at Eurographics [42].

3.1 Introduction

Visualization is becoming a fundamental component of education. The use of visual design in pedagogy has a long history and is still evolving rapidly. Enhancing the educational process by enabling a better understanding of a given subject with graphical representations and promoting visualization literacy skills are important challenges. Visualization literacy is recognized as an important direction of research, indicative in workshops at EuroVis 2014 "*Towards Visualization Literacy*" [107] and at IEEE VIS 2014 "*Towards an Open Visualization Literacy Testing Platform*" [108]. It is also widely studied in the visualization community, e.g. [5, 22, 37]. For purpose of this study, we define treemap literacy as the ability to construct and interpret treemaps.

Treemaps are an efficient way to represent hierarchical data and they require a special layout algorithm. But displaying large hierarchical data sets increases the complexity of the treemap, causing difficulty in treemap comprehension. Poor design parameter choices for a treemap can cause ambiguity and pose challenges in exploring the information represented in the treemap [109]. An investigation into the barriers of interpreting and designing useful treemaps is essential to enhance their effectiveness and intelligibility. Hence, the focus of this study is to identify these barriers to enable a complete literacy of treemaps.

This study is the first one of its kind focusing on treemaps. While the challenges posed by treemaps are not exclusive to this type of visualization, treemaps do have unique properties such as representing hierarchical data and requiring a special layout algorithm. We propose a novel treemap literacy test to assess the barriers to treemap literacy and advance a user's treemap literacy skills by designing an effective pedagogical tool that enables novices to improve their skills of reading, comprehending, interpreting, and creating treemaps. The tool attempts to transform the passive learning experience to an active learning process. Moreover, the educational software supports the analysis of hierarchical data and facilitates correct observations of that which it represents. The research prototype tool demonstrates the correspondence between the traditional tree structure and a treemap design simultaneously.

In order to investigate the potential impact, the result of an experiment conducted in a classroom environment with participants from a computer science department is reported. This study presents the results of the treemap evaluation using the educational tool in an attempt to improve understanding of users' visualization literacy abilities. The main contributions of this study are as follows:

1. Identifying and investigating the barriers to treemap literacy;
2. Introducing a treemap visualization literacy test and conducting a classroom-based user study to evaluate the impact of an interactive tool for the comprehension of treemaps;
3. Developing a novel pedagogical application that facilitates both teaching and the learning of treemaps, advancing treemap visualization literacy.

Supplementary Material	**URL**
Treemap Literacy Presentation	https://bit.ly/3r2QSyn
Treemap Software Demo	https://bit.ly/3ADdtEX
Pre-Intervention Test	http://bit.ly/2kHChcn
Post-Intervention Test	http://bit.ly/2kfsdqQ

Table 3.1: Supplementary materials with URLs.

Table 3.1 provides a summary of supplementary material for this literacy study. The supplementary material makes the study fully reproducible.

3.2 Background

Several studies focus on comprehension and interpretation of visual designs and assess users' understanding of visual representations. We start by examining the related literature on visualization literacy through the Survey of Surveys (SoS) on information visualization [4] and a survey of information visualization books [110]. A survey of interactive visualization for education [44] does not include any study on visualization literacy.

The survey by Scheibel *et al.* [111] presents an overview of treemap layout algorithms and describes the more effective use of treemap visualization techniques. The study presents an extensive classification of treemap layouts and presents the difficulty in understanding the characteristics of different algorithms. A study by Stasko *et al.* [112] compares treemap and sunburst charts and presents a study of two space-filling information visualization techniques for depicting file hierarchies. The sunburst technique assisted task performance more frequently, both in correctness and in time, especially for larger file hierarchies. The detailed depiction of the structure appeared to be a major contributor to this advantage, as participants in the study preferred the Sunburst chart overall. This outcome reflects the difficulties in comprehending data displaying techniques and hierarchies.

The goals of the research by Long *et al.* [113] are to explain the challenge that node-link diagrams encounter and to determine which types of treemaps are more helpful in understanding tree structure representation. The study focuses on understanding various types of treemaps and assists users in perceiving various parts of information with a hierarchical structure. Similarly, Müller *et al.* [114] focus on hierarchies and undertake user research that examines three of the most often used hierarchical data visualizations: node-link, treemap, and icicle plot. These three visualization techniques were tested with four tasks that are common for these types of visual designs. According to the statistical analysis results, participants' performed the worst with the treemap tasks. Because treemaps do not explicitly follow a structural order with multiple hierarchy levels, participants must continually reorient and recall which components have already been processed. They assume that this disparity contributes to the participants' poor performance with treemaps.

Tu and Shen [115] present a new treemap design algorithm to minimize abrupt changes in layout and establish clear visual patterns, and build a contrast treemap to compare attributes in one treemap from two snapshots of hierarchical data. An experiment to test the new layout and a user study to compare the data and examine the changes were conducted. Moreover, Tu and Shen [116] introduce Balloon Focus, a seamless technique for treemaps in multi-focus+context. A user study was conducted with 12 participants who were asked to perform a variety of tasks as well as a case study on the use of the system to convey NBA statistics.

Ziemkiewicz and Kosara [117] investigated how the structure of a visualization affects how we interpret it. They evaluated the effects of a visual metaphor and a verbal metaphor on understanding of tree visualizations by measuring the participants' data comprehension questions on either a treemap or a node-link diagram. Another work by Woodburn *et al.* [118] compared three common visualizations for hierarchical quantitative data, treemaps, icicle plots and sundown charts with a controlled user study with 12 participants. The study looked at performance task accuracy of the visualizations and the participant's visual designs preferences.

3.3 The Challenges of Interpreting Treemaps

Treemaps are a good solution for presenting large hierarchical data sets. The available screen space is divided into rectangles that are scaled, placed, and color-mapped to the variables in the data [119]. They do, however, present certain challenges to some. One aim of our study is to test the hypotheses related to the existence of these barriers. We examined the twelve most popular visual designs presented in Figure 3.1 and observed differences between treemap and other designs. These observations relate to the requirements of creating successful treemaps. The observations reflect differences between treemaps and the other visual designs and show how treemaps can be more complex. We show the connection between these observations and our hypotheses in parentheses. The design complexity results in barriers to comprehension and interpretation of treemap visualization. In addition, we investigated the barriers based on the review of related literature ([112, 113, 114, 120, 121]) and examined coursework submissions from Data Visualization modules for two years and we noticed that these errors came up repetitively in the submissions (see Section 3.4). This is how we identified the five hypotheses to treemap literacy barriers: *H-Hierarchy, H-Layout, H-Size, H-Labels, and H-Legend.*

- **H-Hierarchy** One of the barriers to treemap literacy is likely based on the fact that treemaps convey hierarchical data. A treemap displays the relationship between hierarchically structured data attributes. Identifying the multiple levels of the hierarchy can be a challenge to treemap comprehension. The paper by Stasko [112] reflects the difficulties in comprehending data displaying techniques and data hierarchies.

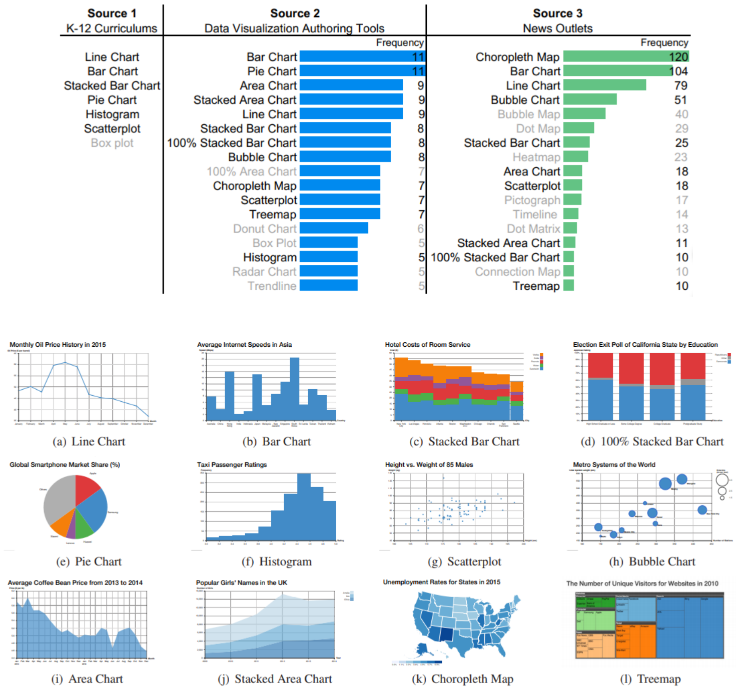


Figure 3.1: *(a, top) Data visualization types surveyed from three sources: K-12 curricula, data visualization authoring tools, and popular news outlets. (b, bottom) The 12 visual designs that compose the VLAT. Images courtesy of Lee et al. [22]*

- **H-Layout** The layout algorithms build useful treemaps by controlling the placement and aspect ratios of the rectangles that compose a treemap. Algorithms aim to increase the visibility of small items in a single image. However, the complexity of the layout algorithm, failing to maintain the order of the data, and layouts that are difficult to visually explore [112, 122] may lead to challenges in comprehension.

- **H-Size** The larger the data set size, the more difficult a treemap image would be to understand, because a larger number of rectangles results in higher visual complexity. Rosenbaum and Hamman [120] states that visualization of large dataset shares a common flaw with many other visualizations: data details are difficult to understand.

- **H-Labels** Node labels enable users to identify which variable a given treemap rectangle corresponds to. Absence of labels or limited display of the labels shown in a designated screen space can cause difficulty in understanding and interpretation.

- **H-Legend** A color legend situated near the treemap can be used to represent value ranges visually. The absence of a color scale can lead to barriers in treemap interpretation.

The last two hypotheses relate to the simple absence or presence whereas the rest are more algorithmic in nature.

In this chapter, we study these barriers preventing users from interpreting and comprehending treemaps correctly. Afterwards, we attempt to improve literacy skills in understanding of treemaps.

3.3.1 Comparisons with Most Common Visual Designs

Our work is inspired by the Visualization Literacy Assessment Test (VLAT) developed by Lee *et al.* [22]. VLAT identifies three major sources to search and determine the most popular visualizations to incorporate in their test [22]. Figure 3.1 (a) compiles the most frequently used visual designs from three different sources: the K-12 educational programs (core state standards for mathematics) [123, 124, 125, 126], data visualization authoring tools (**Google Chart Tools**, **D3.js**, and news articles (The New York Times, The Guardian, and The Washington Post).

They identify data visualization designs included in the curriculum and the designs most often used in authoring tools and popular news outlets. Some of the visual designs covered by educational programs, however, are not as popular with authoring tools and news articles. Figure 3.1 (a) indicates that the Choropleth Map is the most frequently used visual design in news articles although it was not included in the K-12 curriculum. Conversely, pie charts and histograms are used in the educational program and supported by tools, but they were not the most frequently used visualization types in news articles. Figure 3.1 (b) illustrates the 12 data visualizations chosen for VLAT, selected from the most popular visualization types used in news articles e.g. Treemaps, Choropleth Maps, Scatterplots.

It is evident from Figure 3.1 (b) that the treemap design features several characteristics that distinguish it from the other most popular visual designs shown:

- A treemap is not based on a simple Cartesian (nor geo-spatial) coordinate system. (**H-Layout**)

- It utilizes a layout algorithm, as opposed to a simple lookup table in order to guide the placement of geometric primitives such as rectangles, labels, and edges. (**H-Layout**)
- It is the only visual design out of the 12 most popular that incorporates hierarchical data. (**H-Hierarchical**)
- The treemap requires a more sophisticated label placement algorithm than the other visual designs. (**H-Labels**)
- In Figure 3.1 (b), particularly, the treemap is the only example that does not feature a color legend (where necessary). (**H-Legend**)
- The treemap does not feature labelled and numbered axes like the other visual layouts. (**H-Labels**)
- The treemap is the visual design that can be used to display the most individual data samples with the exception perhaps of scatterplots. (**H-Size**)

3.4 Treemap Literacy Assessment

The data visualization module at our university has been taught to final-year undergraduate and master's level students since 2006. The course consists of two-hour lecture and one-hour labs run weekly during one semester. As the construction of a visual design is a way of assessing visualization literacy suggested by Borner [5], we explored how effective students were at creating a treemap by looking at the historical results of the information visualization assignment in 2018 and 2019. Thus, we sought to assess the students' strengths and weaknesses in generating the treemap images, as well as their level of comprehension and interpretation.

Based on our hypotheses and the work of Lee *et al.* [22], we derived criteria that enabled the assessment of treemap literacy. The criterion consisted of questions examining treemap features that were correctly interpreted by the user, including the hierarchy, internal nodes, leaf nodes, labels and legends, and color mapping. The main purposes of treemap questions are to provide students with a list of requirements for successful treemap creation, as well as to evaluate their treemap submissions later. These questions are also a resource for researchers who need to assess users' treemap literacy skills (See Figure 3.2 and Figure 3.3).

The results of the treemap literacy assessment indicate how many treemap features the students correctly incorporate and interpret while creating an appropriate image. The list of literacy test questions was provided to students for first assignment in 2019.

1. **Image:** the treemap image you are describing
2. **Name of Tool:** The tool that was used to generate the treemap
3. **Country:** Name of country(s) data shown
4. **Disease:** Name of disease(s) shown
5. **Year:** The year(s) or time-span of data shown
6. **Data Preparation:** A helpful description of how you prepared the data
7. **Color:** What is color mapped to?
8. **Hierarchy:** What is the data hierarchy contained in the treemap?
9. What leaf node size is mapped to?
10. How are the leaf nodes laid out or positioned?
11. What are internal nodes mapped to?
12. What is internal node size mapped to?
13. Which treemap node layout algorithm is used?

The treemap literacy criterion was applied retroactively to evaluate treemaps submitted by students as part of an information visualization assignment. In 2018, 83 computer science students enrolled in the data visualization module. For the information visualization coursework in 2018, students were required to submit five visual designs to study the Public Health Data of England using existing visualization software (see Figure 3.4 and 3.5). Public Health Data of England [127] is a geographically hierarchical data with England divided into a hierarchy of areas and diagnosis [128]. Students were asked to create and explain at least five unique visual designs using existing data visualization tools. Although there was no explicit requirement to generate a treemap, 68 students in the class attempted to create a treemap as a part of the assignment. Only 38 out of the 68 students specified how they prepared the data before producing the treemap. Usually, this involves formatting to create a hierarchy. This result indicates that the data pre-processing required for a treemap can lead to barriers in treemap generation (**H-Hierarchy**). Some 16 (out of 68) treemaps did not feature a critical feature of a treemap, namely a data hierarchy (**H-Hierarchy**), even when they were explicitly informed about this challenge. 65 of the 68 students defined what color was mapped to, and the color-map was described correctly by 58 (out of 68) of them (**H-Color**).

We examined the students' ability to explain the internal and leaf nodes displayed on the treemap, concluding that students struggled more to describe the internal nodes. 57 out of the 68 students who provided treemaps defined the lowest level (leaf node) rectangles and the leaf node size correctly. Only 42 students were able to describe what the internal node rectangles represent accurately, and only 35 students explain what the size of internal rectangles represents (**H-Hierarchy**). Again, this provides evidence that the hierarchical aspect of treemaps can be a challenging concept for some. Only 42 students provided labels or a legend, in spite of the software being used in class allowing for the creating of a legend (**H-Labels**, **H-Legend**). A correct interpretation of the treemap and unique observations were provided by only 44 students.

We examined the information visualization coursework of the 2019 class using the treemap literacy assessment described previously (see Figure 3.3). As a modification to the previous year's assignment, we asked students to go into greater depth and create a treemap image from the Project Tycho data [129] in addition to generating five images in the first part of the coursework. We provided students with 18 explicit questions that assess treemap literacy related to the color mapping, data hierarchy, internal nodes, leaf nodes, labels and legends, software choice, and the treemap layout algorithm.

Some 66 students attempted the coursework, and only two of them did not provide a treemap example. While 64 students in the class mentioned the software tool

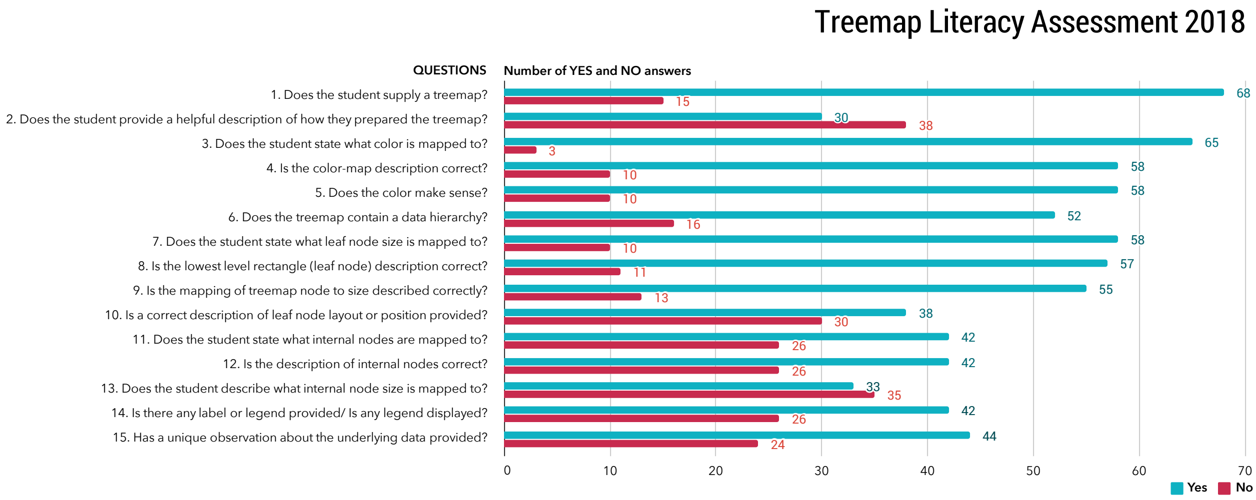


Figure 3.2: *A treemap literacy assessment test results from the information visualization assignment in 2018. Questions 2, 6, 11, 12, and 13 indicate difficulties with the hierarchical aspect of treemaps.*

used to create a treemap, only 42 students (of 64) supplied a detailed description of the treemap example. Figure 3.3 demonstrates that colors used in the treemap were identified appropriately by 59 out of 64 students (**H-Color**). However, 19 out of the 68 treemaps did not contain a data hierarchy, and 20 of them were not defined correctly (**H-Hierarchy**), pointing towards the challenging nature of the hierarchical aspect of treemaps. Some 49 out of the 64 students were able to correctly identify what the lowest level rectangle size was mapped to, but only 35 of leaf node descriptions were accurate. Leaf node layout or position was described incorrectly by 31 students. Similar to the 2018 test results, identification of an internal node was a challenge for students in comparison with identification of the leaf nodes.

All students attempted to define what the internal node size represents, but only 35 out of the 64 students did so accurately (**H-Hierarchy**). In contrast to 2018, all treemaps had a label and a legend. Only 25 students provided unique treemap observations and 29 correctly identified the layout algorithm used (**H-Layout**), indicating that the layout algorithm is a barrier to treemap literacy. Overall, considering that the students taking the course are all in their later stages of the computer science degrees, the error rates and the interpretation of treemaps and topics related specifically to H-Hierarchy and H-Layout can be considered somewhat high.

Additionally, we investigated the software used, the treemap layout algorithm (see Figure 3.4 and Figure 3.5), the students' observations about the data from looking

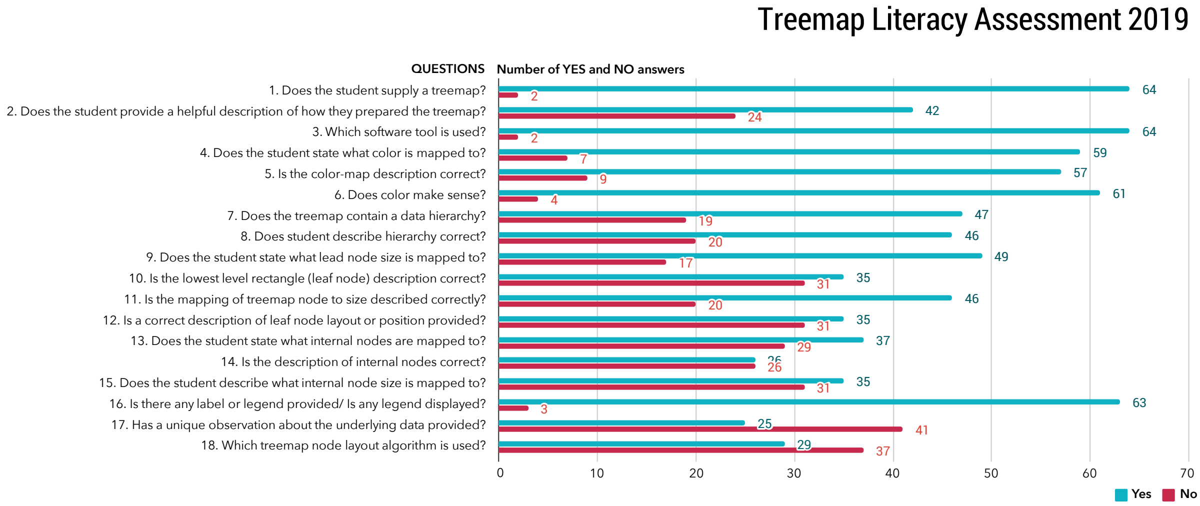


Figure 3.3: *A treemap literacy assessment test results from the information visualization assignment in 2019.*

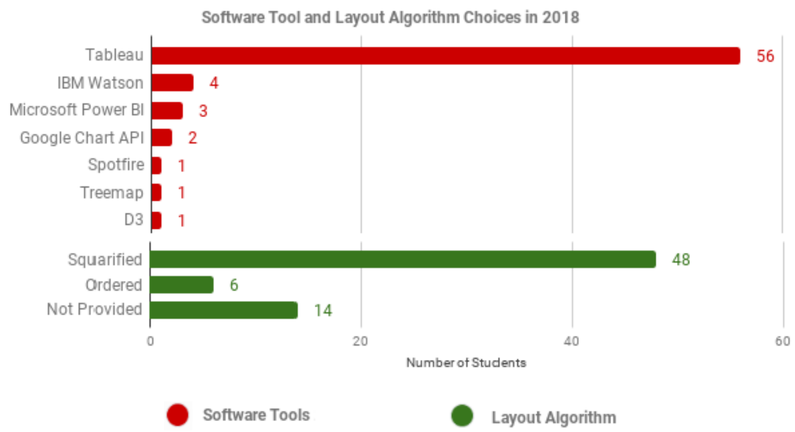


Figure 3.4: *The results of software tools used and treemap layout algorithm in 2018.*

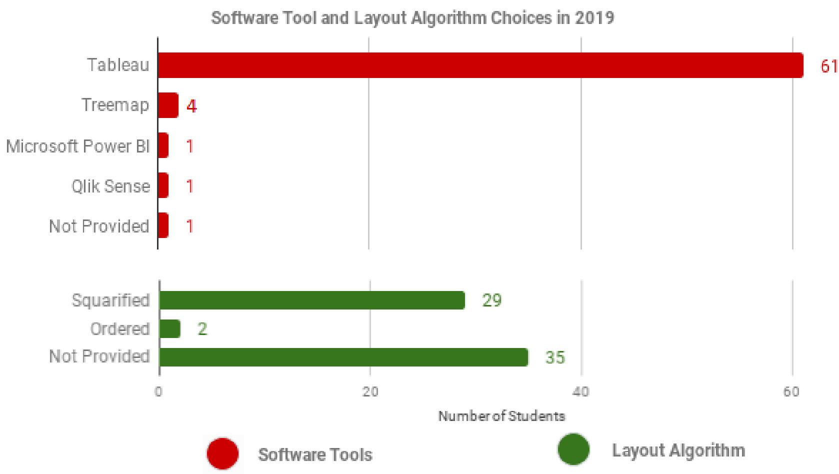


Figure 3.5: *The results of software tools used and treemap layout algorithm in 2019.*

at the treemap image, and how students prepare the data to generate the treemap in both years respectively. Students created treemaps using Tableau [130], IBM Watson [131] etc. Students used a squarified or ordered treemap algorithm. Students' unique treemap observations were solicited to assess students' abilities at the interpreting the treemaps they produced. We also assessed the accuracy of the students' answers regarding their treemap submission. Our treemap literacy evaluation result provides insights into students' treemap literacy skills and enhances our understanding of the barriers to treemap construction. This evidence guides the development process of our educational treemap application.

The investigation of students' software choice showed that even though some students did not provide any software identification, 56 students (82%) preferred to use **Tableau** [130], one of the popular commercial tools to create visualizations. However, caution must be used when interpreting this because we provide an explicit Tableau tutorial for the class. The other tool choices were **IBM Watson** [131], **Microsoft Power BI**[132] etc. Forty eight students (71%) used a squarified treemap algorithm to obtain a treemap (see Figure 3.4). The overall result of software and treemap layout algorithm choice of the class in 2019 is displayed in the Figure 3.5.

3.5 Treemap Visualization Literacy Test

We developed a treemap visualization literacy test to measure a user's treemap literacy skills and identify the barriers to comprehension of a treemap. In-class investigation is based on treemap construction whereas the treemap visualization literacy test focuses more on treemap interpretation. We searched for appropriate treemap examples with diverse treemap visual designs to test the comprehension of users with varying levels of treemap literacy and enable them to attempt a range of questions related to treemaps. We first selected examples with correct hierarchical data structure and eliminated examples without labels or a legend – these provide clues as to what internal and leaf nodes are mapped to and how color is used on the treemap design. Finally, we ensured that the treemaps are of high quality, excluding low resolution images (see Figure 3.6).

We surveyed four sources to find a number of treemaps chosen from each source as follows: The Visualization Literacy Assessment Test by Lee *et al.* [22] (1 image), The

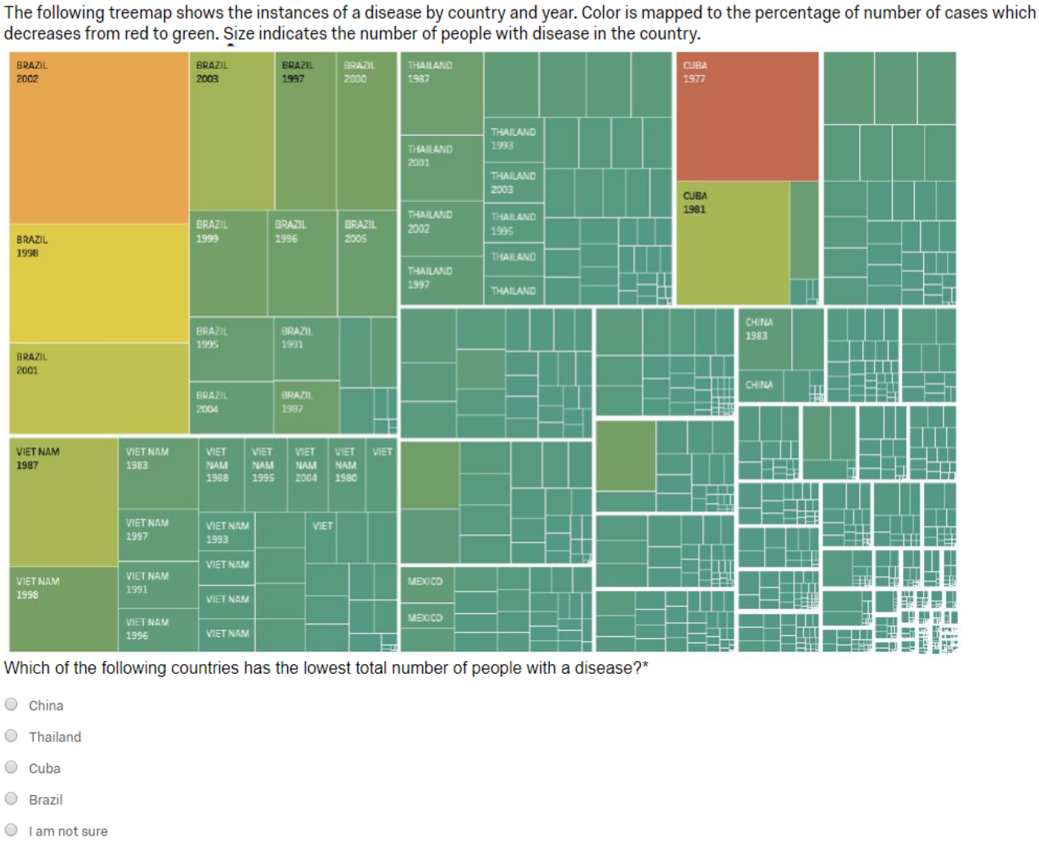


Figure 3.6: *An example question from the Treemap Visualization Literacy Test.*

Book of Trees: Visualizing Branches of Knowledge by Manuel Lima [133] (3 images), Google keyword search for "Treemap" (3 images), and students' treemap examples submitted for the information visualization coursework for the Data Visualization course (6 images). We used Google search engine that provides high resolution, and interpretable treemap images with a correct hierarchy. All treemaps were static for consistency.

Once the treemap examples were selected, we prepared five questions for each treemap image that test the comprehension of different aspects of a treemap. The test was prepared to explore the user's ability to make sense of the treemap by asking them a variety of questions. Answering treemap literacy test questions requires the evaluation of multiple factors. Therefore, the questions were coded to identify how users perform in interpreting the data hierarchy, internal and leaf nodes, labels color mapping, a range of data sizes, a legend and layout algorithm (see Hypotheses) using treemap literacy skills. Each question in the literacy test required understanding of at least two treemap features (see Figure 3.24). The full list of questions is available at **pre-intervention test** [134] and **post-intervention test** [135].)

3.6 A Pedagogical Treemap Tool

In order to improve treemap literacy, we developed an instructional software tool for classroom use. The treemap application facilitates understanding of a hierarchical data structure and supports accurate observations by displaying the data correspondence between a traditional tree structure and a treemap layout simultaneously (Figure 3.7). Our treemap tool can be used on different kinds of data, which can be set up by the user. The educational tool was designed following the Sedlmair [136] design methodology, and the design process did not explicitly include the end-user's needs. However, the software was designed based on the assumption that most users understand the traditional tree hierarchy. Therefore, we introduced a tree view to depict the data hierarchy and the correspondence with the treemap view, and coordinated the two views using animation. The key design decisions that were made were to support interaction with tree view and treemap and they were linked together and synchronized. So users can clearly see the correspondence between the traditional tree and treemap views. We considered several different layout algorithms however, they were complicated to implement. We followed the traditional treemap structure, slice and dice [137] which was the most fitting layout algorithm to explain the treemap concept. The software is developed using C++ and the Qt framework [138].

Tree Features: The tree view (Figure 3.7 on the left) enables users to analyze the hierarchical data structure and control the treemap.

- The user can hover the mouse over any rectangle in the tree view. The rectangle is highlighted, increases in size and displays a tooltip with the underlying data values.
- The user can click on any internal node, and the user selected node dynamically displays its child nodes.
- If an internal node displays its children and the user clicks on it again, the child nodes collapses into a representative rectangle.
- In addition, the tree view displays labels identifying the levels of the hierarchy.

Treemap Features: The treemap (Figure 3.7 on the right) demonstrates the hierarchical data structure with a layout algorithm.

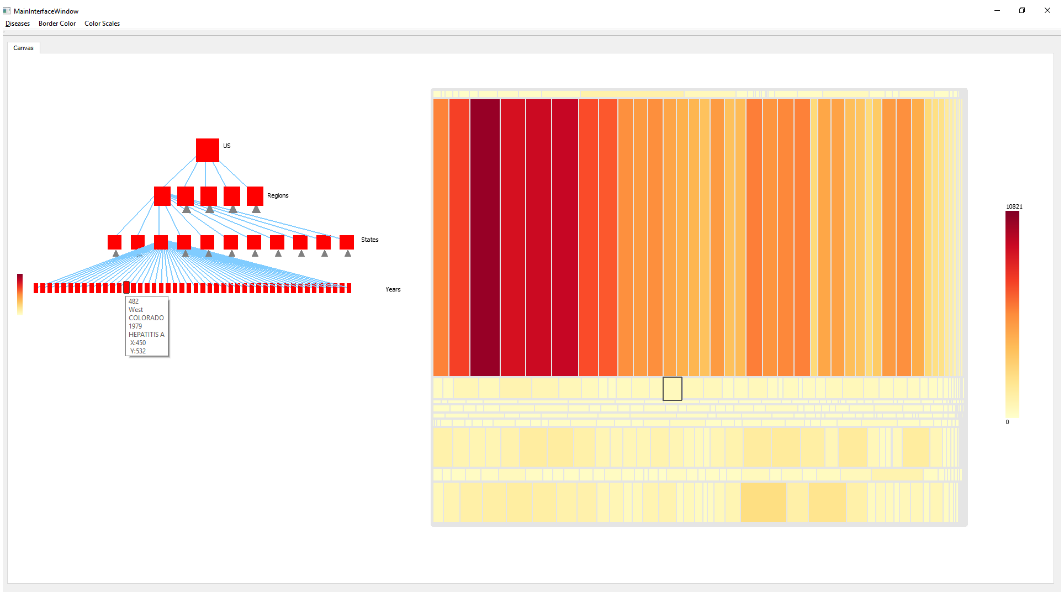


Figure 3.7: *Instructional treemap tool interface with traditional tree structure (left) and linked treemap visualization (right).*

- The treemap displays the equivalent of the tree view using a treemap layout algorithm, in this case, slice and dice [137].
- The treemap view has a user-modifiable color-map where color is mapped (redundantly) to the size of each leaf node rectangle.
- The user can hover the mouse over any rectangle, and a tooltip shows the underlying data values. The rectangle is also highlighted/enlarged.

Coordinated and Linked Treemap and Tree Features: The traditional tree and treemap views are linked and synchronized. Interacting with either one causes updates to other. Interactive control of drawing treemap and tree view allows users to determine the properties of the data hierarchy and provides real-time feedback.

- The treemap view is updated whenever the user clicks on a node on the tree view.
- The treemap view reflects the number of internal and leaf nodes shown in the tree view.
- If the user hovers the mouse cursor on any node or rectangle in the tree view the corresponding node or rectangle is highlighted/enlarged in the treemap view, and vice-versa.

Menu and User Options: The menu options offer more features to the user. The 'disease' menu option lets the user choose between a selection of diseases to visualize from Project Tycho [129]. The list of diseases includes: Hepatitis A, Measles, Rubella, Mumps, Polio, Pertussis, and Smallpox.

Color Selection and Color Legend: Six color scales are provided in order to explore different mappings. We utilize a color library [139] with the assistance of Colowbrewer [140], an online source for selecting color scales.

- The user can choose any color scale among the given color-map options.
- The color legend is updated based on user choice of color scale options.
- The color legend shows maximum and minimum values of the smallest level of the current treemap view and represents color distribution on the treemap according to the current range of data values.
- Maximum and minimum values are updated when the user chooses a disease, region, state and year for the treemap using tree view.

In our classroom-based experiment, we used Project Tycho – a large-scale data of the US records disease incidence frequency data between the years 1888-2014, recorded weekly. The dataset, provided by the Public Health Dynamics Laboratory at the University of Pittsburgh Graduate School of Public Health, provides a record of the number of cases or deaths due to a given disease in a specific location over a time duration e.g. 5 people diagnosed with Hepatitis A in Alabama in week 33, 1966. For our study, we selected a group of diseases recorded based on the states (some of them contain specific cities). In order to create a hierarchy, we grouped states for each disease according to five regions in the US (West, Southwest, Midwest, Southeast, and Northeast [141]) as a level in the hierarchy.

3.7 Classroom Evaluation

We designed a classroom-based user study to evaluate the participants' treemap literacy and the effectiveness of the pedagogical treemap software. We provided two tests, a **pre-test** [134] and **post-test** [135], which featured 30 and 27 questions, respectively. Both tests contain a collection of treemaps, multiple choice questions, and a description of each treemap (see Figure 3.6). Both the treemap designs and the data sets used in this study varied in their complexity. For each correct answer,

Figure 3.8: *A photo from the user study on treemap literacy with computer science students.*

students were given 1 point in both tests. After the pre- and post-intervention tests, 12 open-ended interview question (see Section 3.7.5) were given to participants to collect feedback. These tests were administered using Qualtrics [142], an online survey tool for collecting data.

3.7.1 Experimental Classroom Procedure

The experiment was run in a classroom environment. Some 25 computer science students (2 female) were recruited to participate in the study. Participants were students at different degree levels (14 Bachelor's, 4 Master's, and 7 PhD). The age of participants ranged from 18 to 38. Only 4 students had a data visualization background from various taught classes. Participants were randomly assigned to one of the two groups: a presentation SLIDES group and a SOFTWARE group. The participants in the SOFTWARE group were provided with a treemap software demonstration and given time to interact with the educational treemap tool (see Figure 3.8). The SLIDES group was shown only traditional treemap slides, used for teaching treemap concepts. Each participant was provided with an Amazon voucher upon the study's completion.

We described the procedure of our study and asked for the students' consent to participate. Upon their agreement, we provided all participants with the pre-intervention test treemap questionnaire, which took approximately 20 minutes to complete. After the completion of the pre-intervention test, we randomly sampled half of the participants to be allocated to the SOFTWARE demonstration group (every other participant).

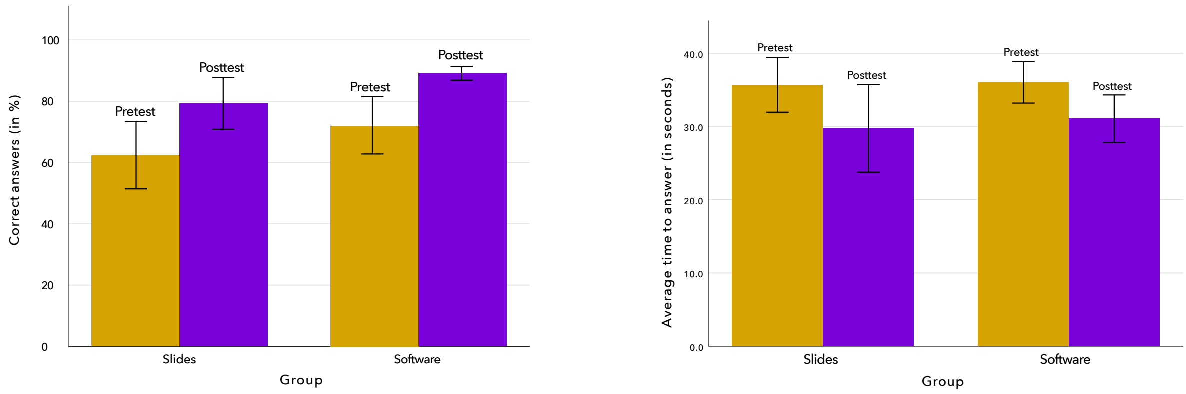


Figure 3.9: *(Left) The percentage of correctly answered questions in the pre- and post-intervention tests for* SOFTWARE *demonstration and* SLIDES *groups. Error Bars (95% CI). (Right) The average time participants in* SOFTWARE *demonstration and* SLIDES *groups spent answering questions in the pre- and post-intervention tests. Error Bars (95% CI).*

Both sessions, SLIDES and SOFTWARE, were delivered by the same member of academic staff to eliminate the possibility of a delivery confound. To facilitate this, half of the students had the traditional slides delivered to them, while the other half (the SOFTWARE group) waited in a different room. Once the slides session finished, the SOFTWARE group switched rooms with the SLIDES group. Both the SOFTWARE and the SLIDES sessions were 20 minutes long.

The SOFTWARE group were introduced to the pedagogical treemap application. They were then asked five questions related to the Project Tycho data set provided. The students answered the questions verbally by exploring the dataset for answers using the features of treemap software.

The SLIDES group returned to the classroom once the software session was over. Both groups were then given the post-intervention test questionnaire. Upon its completion, all participants answered 12 interview questions, referring to their background, the test questions, and the treemap software.

3.7.2 Quantitative Results of Test Data

The data we collected was normally distributed, as indicated by the Shapiro-Wilk test for both pre- and post-intervention test groups. Hence, we used one-way ANOVA for our data analysis (significance level at $\alpha = 0.05$).

The pre-intervention test results did not differ significantly between the two groups: $F(1, 24) = 1.841$, $p = 0.188$, $\eta_p^2 = 0.074$. Those students who then attended the SLIDES session answered on average 62% of the pre-intervention test questions (SD = 19%), and the students who then took part in the SOFTWARE demo answered

	PRE-INTERVENTION TEST					POST-INTERVENTION TEST				
	Slides	Software	$F(1,24)$	p	η_p^2	Slides	Software	$F(1,24)$	p	η_p^2
C: Color	58.7 ± 22.8	74.2 ± 16.4	3.763	0.065	0.141	72.6 ± 20.1	88.0 ± 10.0	5.662	0.026*	0.198
H: Hierarchy	65.6 ± 20.8	71.6 ± 18.2	0.577	0.455	0.024	77.7 ± 17.5	87.3 ± 5.2	3.298	0.082	0.125
LN: Leaf node	59.2 ± 21.1	68.3 ± 15.3	1.504	0.233	0.061	81.4 ± 12.5	89.5 ± 5.0	4.398	0.047*	0.161
LB: Label	65.7 ± 19.0	77.3 ± 15.5	2.742	0.111	0.107	79.8 ± 15.3	89.2 ± 4.9	4.161	0.053	0.153
LA: Layout algorithm	58.0 ± 20.3	68.6 ± 17.2	1.973	0.173	0.079	75.5 ± 17.6	87.0 ± 5.0	4.770	0.039*	0.172

Table 3.2: The results of pre- and post-intervention tests for the Slides and Software groups ($M \pm SD$ in percentages), based on the categories of questions. Significant results are shown as follows: $*p < 0.05$.

72% of these questions (SD = 16%). However, the results of the post-intervention test differed significantly between the two groups: $F(1, 24) = 5.074$, $p = 0.034$, $\eta_p^2 = 0.181$. Those who attended the slides session answered on average 79% of the post-intervention test questions (SD = 15%), which was significantly lower than the results of the students who interacted with the software – they answered 89% of the questions correctly (SD = 4%).

The slides group have seen a 17% improvement in their results from pre-intervention test to post-intervention test (SD = 18%) and the software group have improved their results on average by 17% (SD = 17%) (Figure 3.9 on the left). There was no significant difference between the two groups with regards to their treemap literacy improvement. Participants in both groups answered the post-intervention test questions faster than the pre-intervention test ones: $F(1, 24) = 23.222$, $p < 0.001$, $\eta_p^2 = 0.492$ (Figure 3.9 in the right). There was no interaction effect based on the manipulation.

Nonetheless, as we hypothesized that the software would provide additional support to the students in overcoming different barriers to understanding treemaps, we have also looked at the students' performance in the different parts of the test aimed at measuring one's comprehension of different attributes of a treemap. We did so by looking at the question classification based on the treemap features that could influence the participants' answers in the test.

To investigate where participants struggle the most and evaluate their visualization literacy skills, we developed a variety of treemap test questions, considering treemap features such as: Color legend, Hierarchy and Internal node comprehension, Leaf node, Labels, and Layout Algorithm. There was no significant difference between the two groups of participants taking the pre-intervention test in any of the five categories (Table 3.2).

In the post-intervention test, there was no difference in the results obtained by participants in both groups for the questions about neither Hierarchy and Internal nodes nor Labels. However, participants who interacted with the software performed much better in the questions related to the Color legend, Leaf nodes, and Layout Algorithms (Table 3.2).

The number of correct answers in the pre-intervention test for all participants correlated negatively with the number of rectangles on a treemap ($r = -0.520$, $p = 0.003$). The higher the number of leaf nodes the more difficult the label placement, and hence, they can be more difficult to interpret. However, there was no correlation between the number of rectangles and the correct answers of participants in the post-intervention test ($r = -0.084$, $p = 0.677$). Similarly, there was no correlation between the number of rectangles on a treemap and the amount of time participants spent answering each question: $r = 0.207$, $p = 0.123$.

3.7.3 Qualitative Analysis of Interview Data

Thematic analysis was jointly conducted by the authors using the feedback gathered from the interview session. For this, we followed the procedure by Braun and Clark [143]. We first familiarized ourselves with the answers to understand the participants' experiences of the treemap literacy test and the treemap instructional sessions. We used a deductive approach to establish the themes based on the barriers of the treemap literacy identified in the hypotheses. However, we did not limit the analysis to the barriers alone and looked for further insights into the challenges of treemap literacy through the feedback.

We identified five themes in our analysis of the qualitative data: *Hierarchy, Labels, Colors, Layout, and Size.* These themes indicate that our analysis had strong links to the initially identified barriers.

Hierarchy In contrast to the quantitative data, qualitative feedback gathered from the students in both groups did not highlight many problems regarding the exploration of the hierarchical relationship between data features on treemap designs. Only one student mentioned this category in their feedback: *"I had a hard time recognizing the levels of hierarchy in the images."* (P10, SLIDES).

Labels Finding the right answer to the questions was possible through information provided by labels. The visibility of labels on treemaps played a major role in participants' performance during the test. Feedback from five students indicated that too much or too little information about data was a challenge. This challenge might have been a byproduct of the cluttered visual design showing a large dataset: *"Visual information is easier to process, so a lot of questions were simple."* (P24, SOFTWARE) and *"[In] the question with commodities, labels where not visible and too many of them."* (P11, SLIDES).

Colors Questions requiring interpretation of the treemap color-mapping were indicated as difficult by four students. This was despite the inclusion of a color legend or color description: *"Treemap visualizations and attached explanation of mapping color/size were clear."* (P19, SOFTWARE), but *"It [was] not clear how to compare [the nodes] and sometimes too many colors."* (P11, SLIDES).

Layout Finding a specific data point among the treemap rectangles could be enhanced through the understanding of the layout algorithm. Feedback showed that only three participants explicitly struggled with this aspect, e.g.: *"Some of the categories or specific data was hard to find."* (P06, SOFTWARE). However, *"knowledge of the domain represented seems to be very useful to answer quickly, as you know where to look."* (P01, SOFTWARE)

Size Qualitative evidence showed that both groups found data size to be an issue regardless of which group they were in. Feedback from 10 students mentioned the size of the data as a major barrier to being able to correctly interpret the treemap, e.g. *"The treemap contains many boxes that are hard to see."*(P15, SOFTWARE) and *"The more data being represented translated in more convoluted/dense treemaps which made certain things hard to spot."* (P17, SLIDES).

We also analyzed the feedback that referred specifically to our pedagogical application, which was obtained from the students who had interactive practice with the software. We coded the feedback based on the features of the software that were perceived as having a positive effect on the student experience and the feedback referring to the features that could be improved in the future software development iterations. Two most prominent themes emerged: *Hierarchy* and *Interaction and Active Learning*.

Hierarchy Responses collected showed that most participants in the SOFTWARE group found the ability to freely interact with and explore the hierarchical data structure particularly helpful. Five students commented on the difficulty of interacting with the hierarchy, e.g. *"It breaks the tree down so you can only view what you want to see."* (P07) and *"I can see the relationship and the categories of the different data."* (P15).

Interaction and Active Learning Students who participated in the interactive software session predominantly responded positively to their active learning experience, e.g. *"The visual feedback when hovering and the pop up were helpful"* (P01), *"Hands-on approach was effective"* (P19) and *"The tree next to the treemap allow[ed] me to view the path. The boxes in the treemap where also highlighted when you hovered over them in the tree"* (P07).

3.7.4 Supplementary Analysis

After completing the quantitative and qualitative analyzes, we examined the pre-and post-intervention test results of the participants in both groups more closely. Figures 3.10 and 3.11 indicate number of questions answered correctly by each participant in SLIDES and SOFTWARE groups. This review enables us to view the performance of each participant in the pre-intervention test with the post-intervention test.

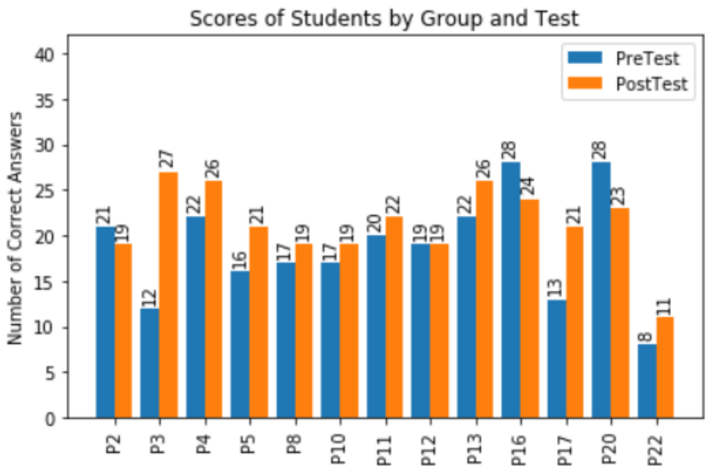


Figure 3.10: *The score of pre- and post-test students for the* SLIDES *group.*

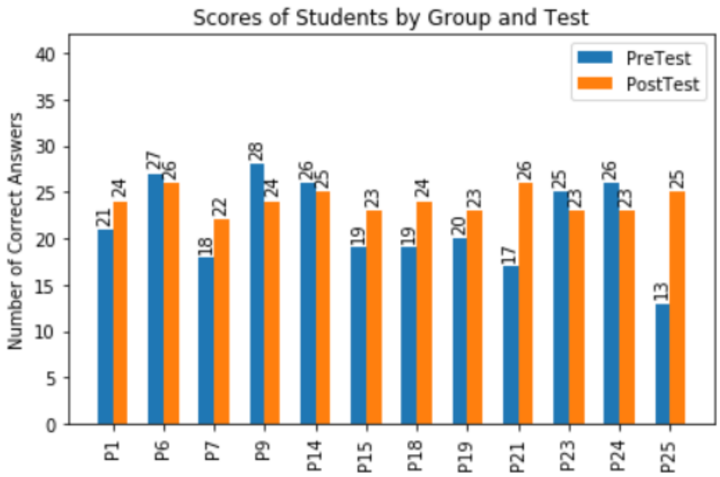


Figure 3.11: *The score of pre- and post-test students for the* SOFTWARE *demonstration group.*

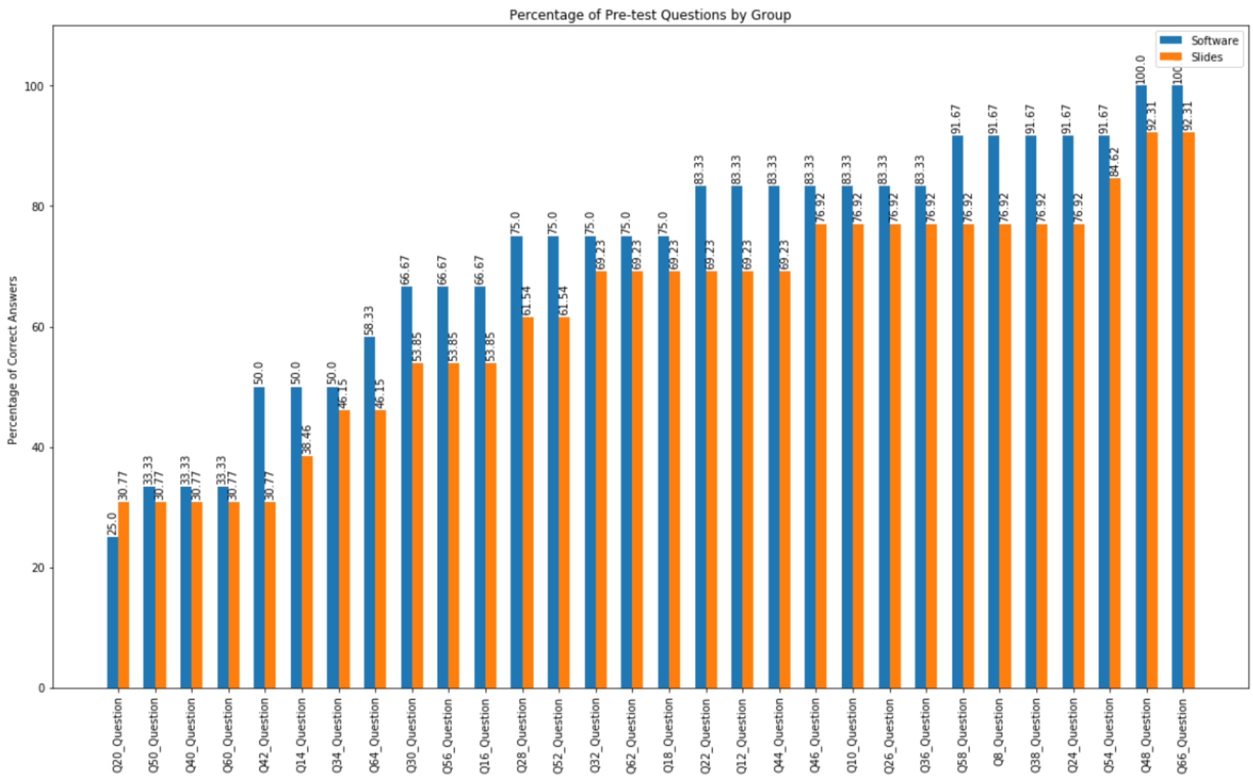


Figure 3.12: *The percentage of correct answers on pre-intervention test questions for the* SOFTWARE *demonstration and* SLIDES *groups.*

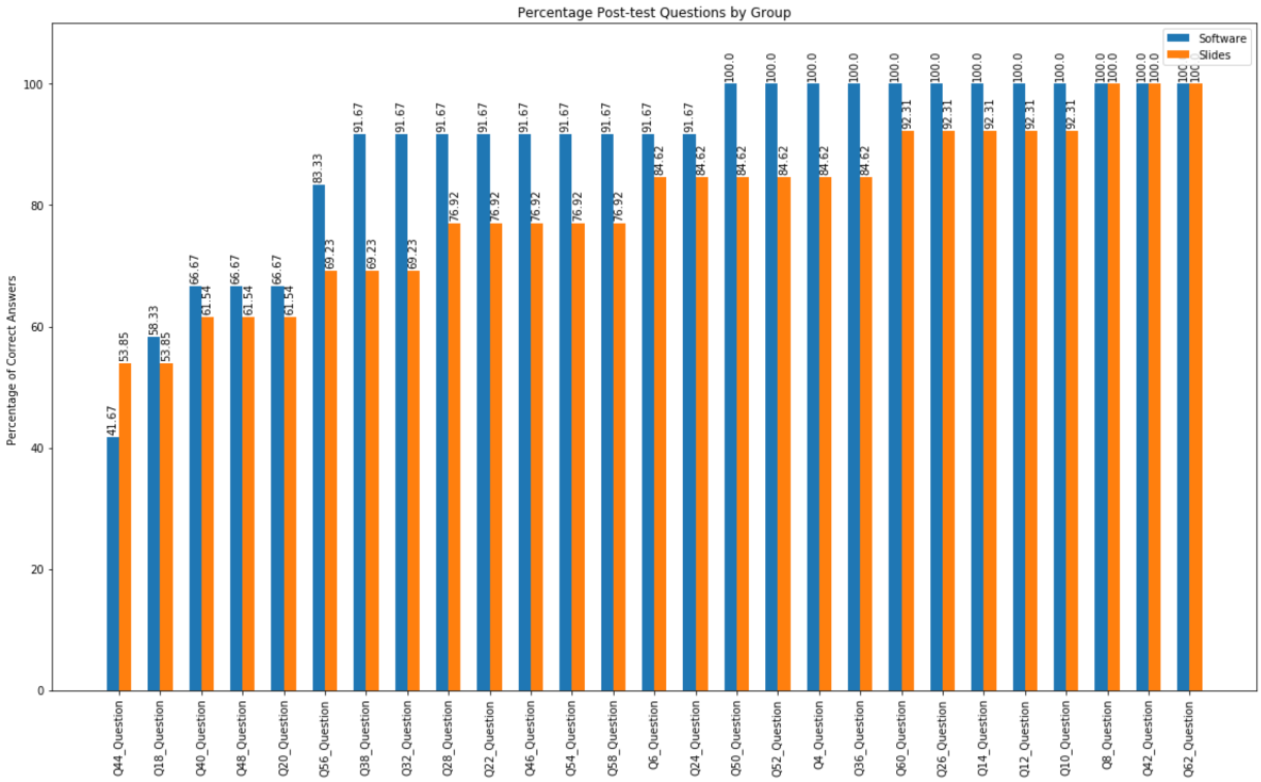


Figure 3.13: *The percentage of correct answers on post-intervention test questions are shown for the* SOFTWARE *demonstration and* SLIDES *groups.*

In addition, we created two graphs showing the percentages of participants in the slide and software group who answered the questions correctly in both the pre-intervention test and the post-intervention test, shown in Figures 3.12 and 3.13. We ranked the percentages of answering the questions in both groups from smallest to greatest. The result shows us the hardest and easiest questions in both tests. The graphs show the easiest question for the pre-test is Q66 on page 173 while the most difficult is Q20 on page 171. For the post-test, we see that the easiest question is Q62 on page 183, while the most difficult is Q44 on page 187. Note that the pre-test question numbers are mapped to the blue color, while the post-test ones are in red. Based on the hypothesis that one of the factors affecting the participants' performance might be the number of rectangles in the treemap, we created graphs that view the relationship between the number of rectangles in the treemap and the number of questions that participants answered correctly in the pre-intervention and post-intervention tests (see Figures 3.14 and 3.15). The results generally do not show a strong correlation between the number of rectangles in the treemap in terms of correctly answering the questions.

In addition, we created two additional graphs for both intervention tests to see the relationship between the number of rectangles in the treemap and the time spent to answer the question. We see that more time is spent answering questions with treemaps that are dense in terms of the number of rectangles. The findings do not indicate a direct relationship between the number of rectangles and the time spent to answer the question correctly (see Figures 3.16 and 3.17).

Please see Appendix A for the pre-intervention and post-intervention test questions.

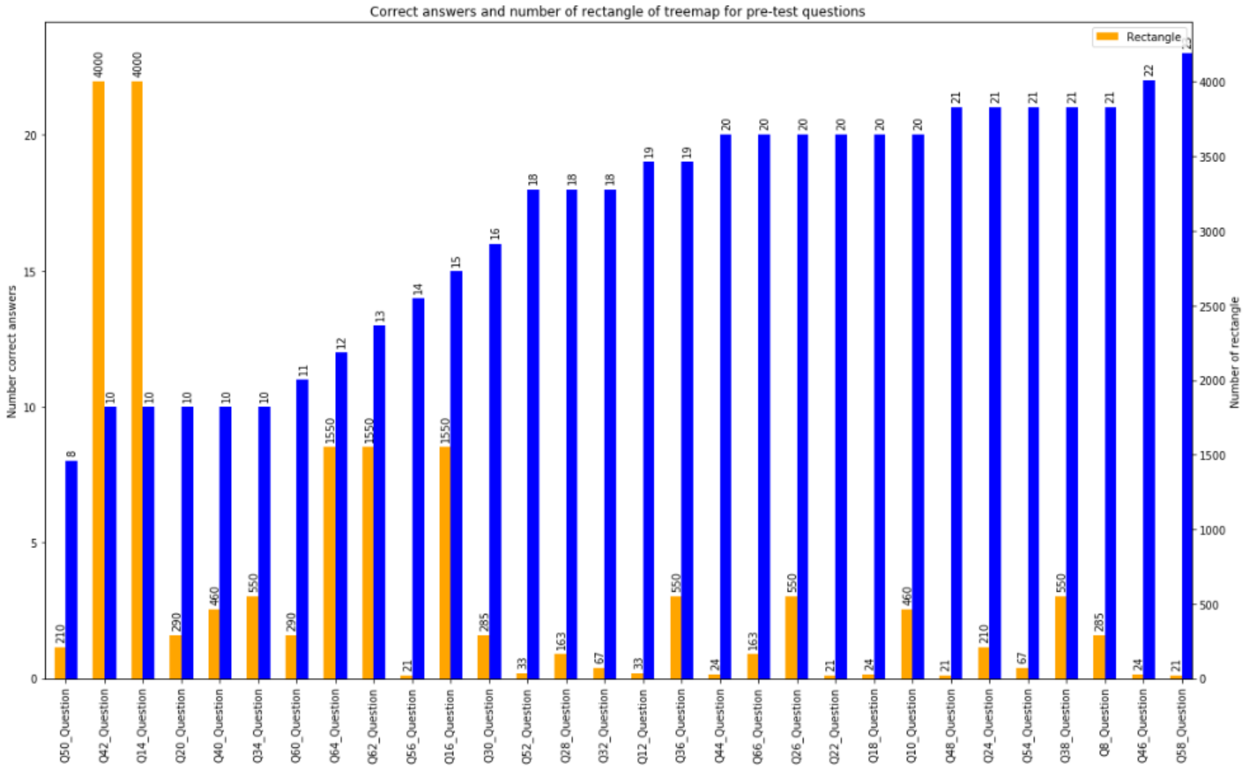


Figure 3.14: *The number of rectangles on a treemap versus the number of correct answers on the pre-intervention test.*

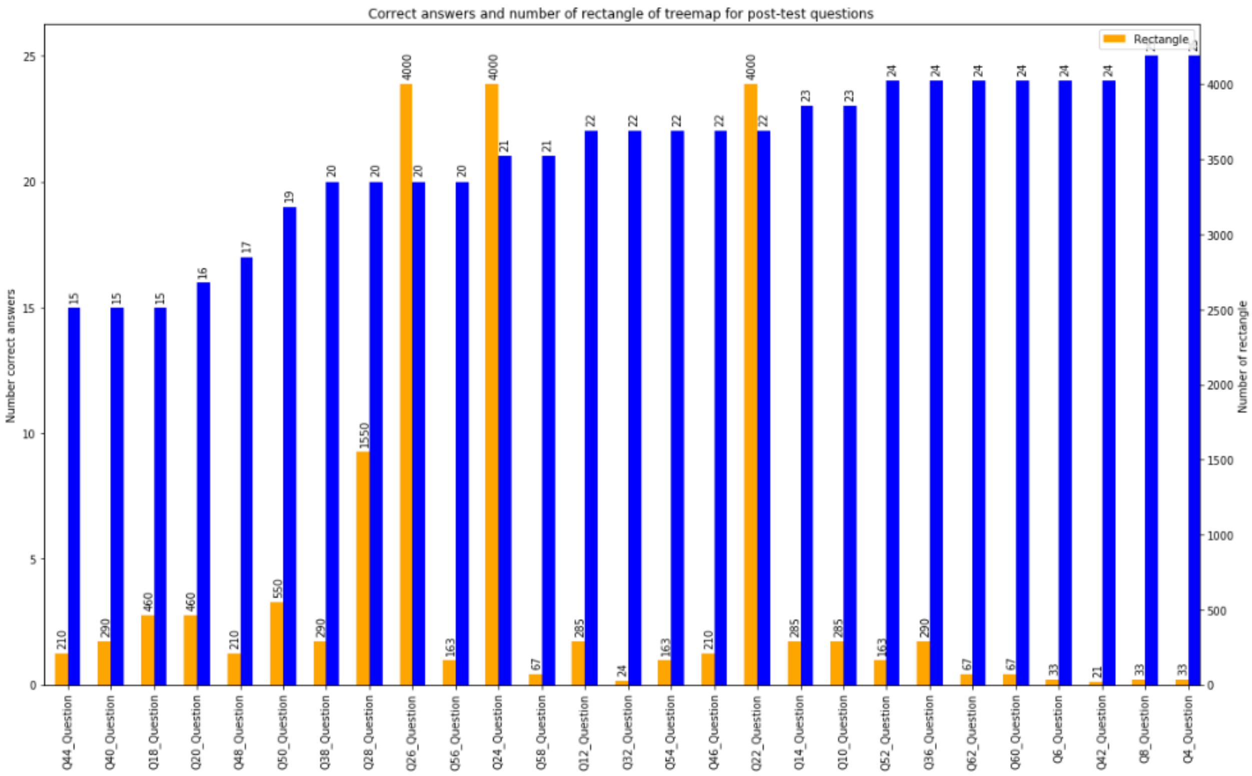


Figure 3.15: *The number of rectangles on a treemap versus the number of correct answers on the post-intervention test.*

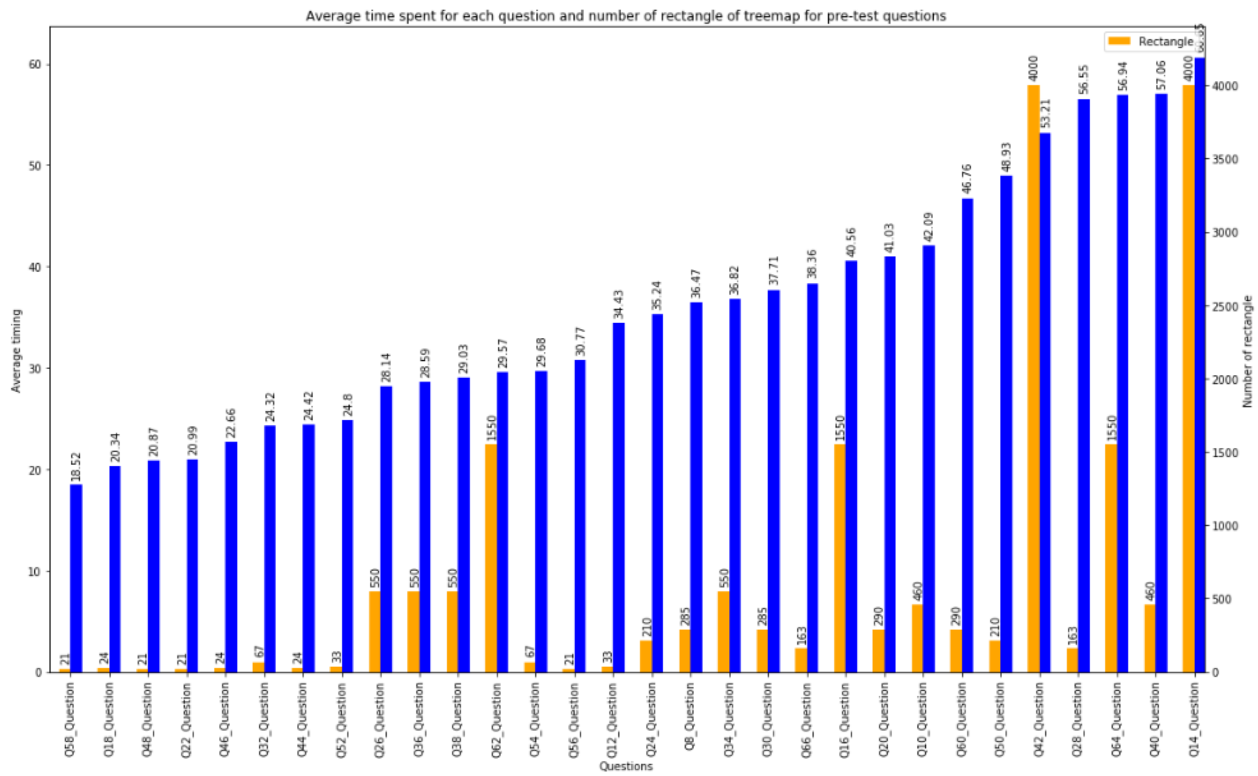


Figure 3.16: *The number of rectangles on a treemap versus the average time spent on each question on the pre-intervention test.*

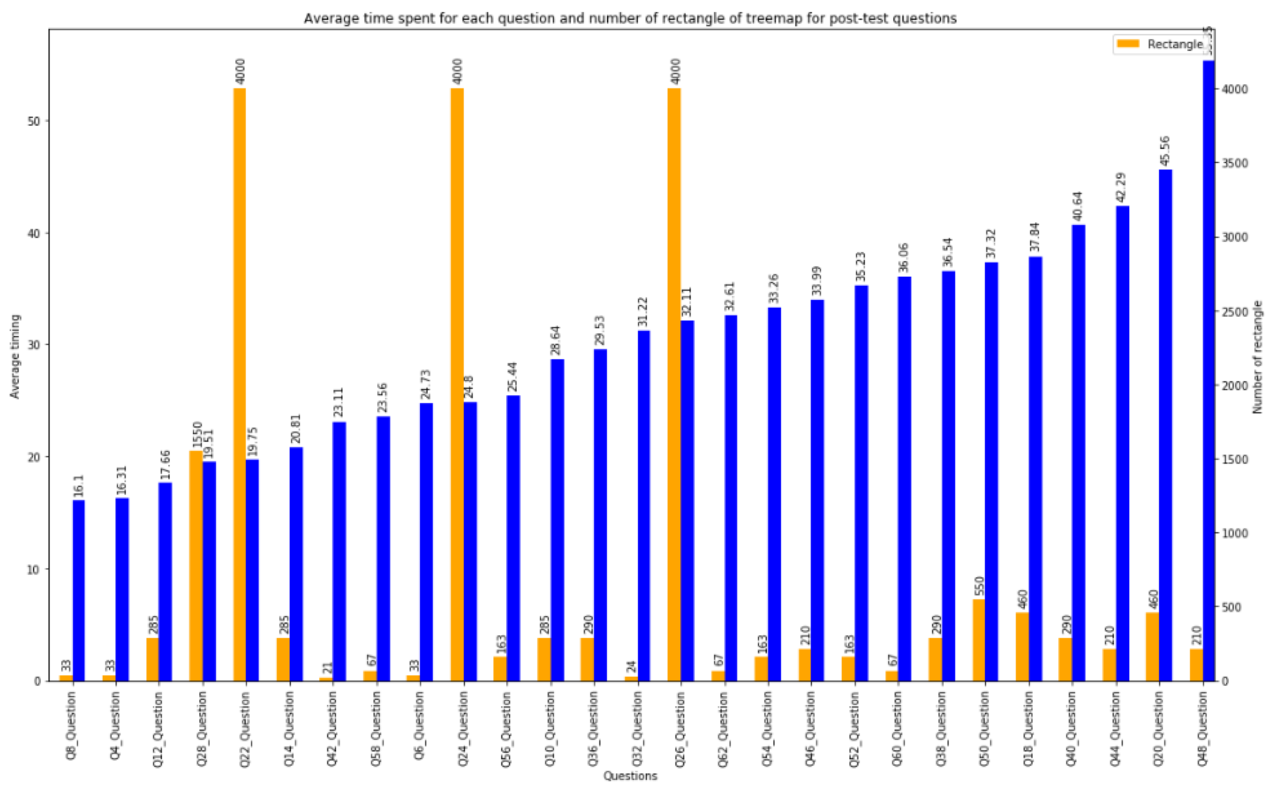


Figure 3.17: *The number of rectangles on a treemap versus the average time spent on each question on the post-intervention test.*

3.7.5 Post-Interview Questions and Analysis of Participants' Responses

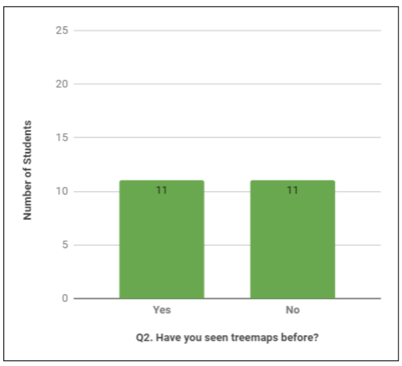
We prepared a post-interview 12 questions given to the participants after completion of post-intervention test, in which we can evaluate the participants' familiarity with data visualization and treemaps, the difficulty of the test questions, and their opinions on the effectiveness of the treemap software. The questions are as follows:

1. Have you seen treemaps before? If yes, where?*
2. Do you have a background in Data Visualization? If so, what is it?*
3. How difficult did you find the test questions?*
   1-Not at all, 7-Very much
4. Please expand why you felt the test questions difficult or easy.*
5. Did you struggle to answer any questions? If yes, what in particular did you struggled with?*
6. How helpful was the Treemap software and software demonstration? (Not applicable for the participants who took slide demonstration)
   1-Not at all, 7-Very much
7. Why (or why not) do you think it was helpful? (Not applicable for the participants who took slide demonstration)
8. Was the Treemap software effective enough to visualize the data hierarchy? (Not applicable for the participants who took slide demonstration)
   1-Not at all, 7-Very much
9. Why (or why not) do you think it was effective enough? (Not applicable for the participants who took slide demonstration)
10. Do you think you perform better on the test after the Treemap software demonstration? (Not applicable for the participants who took slide demonstration)
11. Do you recommend any improvements to the software? (Not applicable for the participants who took slide demonstration)

12. What is your level of English proficiency?*

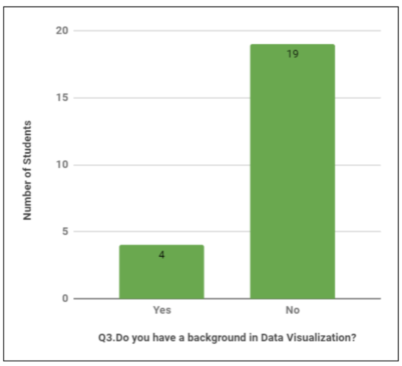
    - Fluent/Native or Bilingual
    - Full Professional
    - Professional Working
    - Limited Working
    - Elementary

We analyzed the answers collected from the post-interview. While half of the participants stated that they had seen the treemap before, only 4 of them had a background in data visualization (see Figures 3.18 and 3.19). Figure 3.20 shows the number of participants and the degree of how difficult they find the test questions. The average was found to be 3.7 out of 7 while the Figure 3.21 displays the degree of how helpful they find the treemap software and software demonstration with an average of 4.8 out of 7. The participants gave an average of 5.7 out of 7 when asked how effective the treemap software was in showing the hierarchy in the data (see Figure 3.22). Finally, the participants also have language proficiency in English and they are fluent in English, except 2 participants (see Figure 3.23). (Post-interview questions are ranked based on this scale (1-Not at all, 7-Very much))



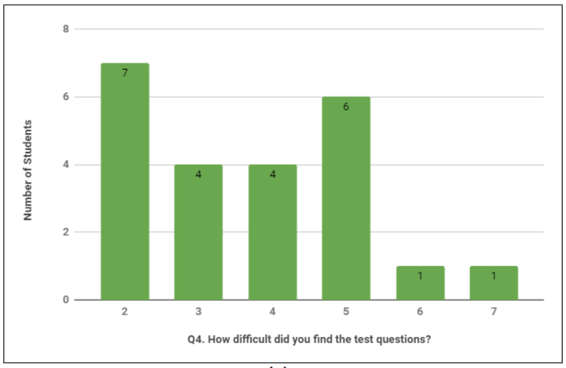
(a)

Figure 3.18: *Yes or no response of participants on whether they have seen a treemap before*



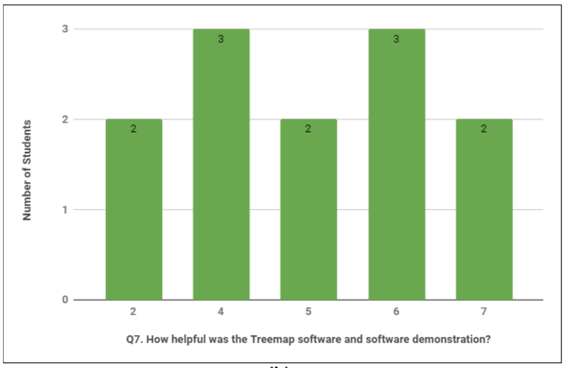
(b)

Figure 3.19: *Yes or no response of participants whether they have data visualization background*



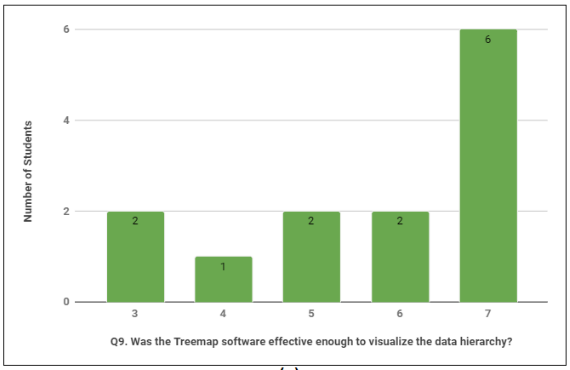
(a)

Figure 3.20: *Number of participants and how difficult they find the treemap test questions*



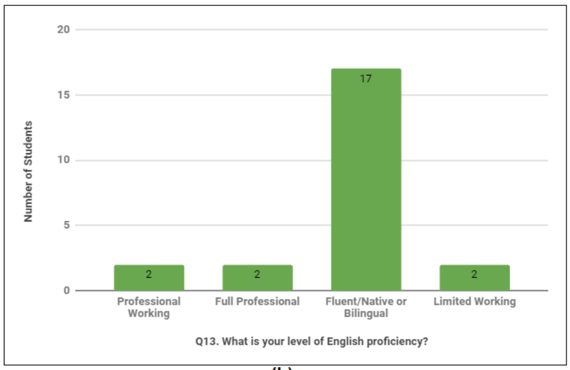
(b)

Figure 3.21: *Number of participants to degree how helpful they find the treemap software*



(a)

Figure 3.22: *Number of participants to degree how effective they find the treemap software*



(b)

Figure 3.23: *Answers of participant's level English proficiency*

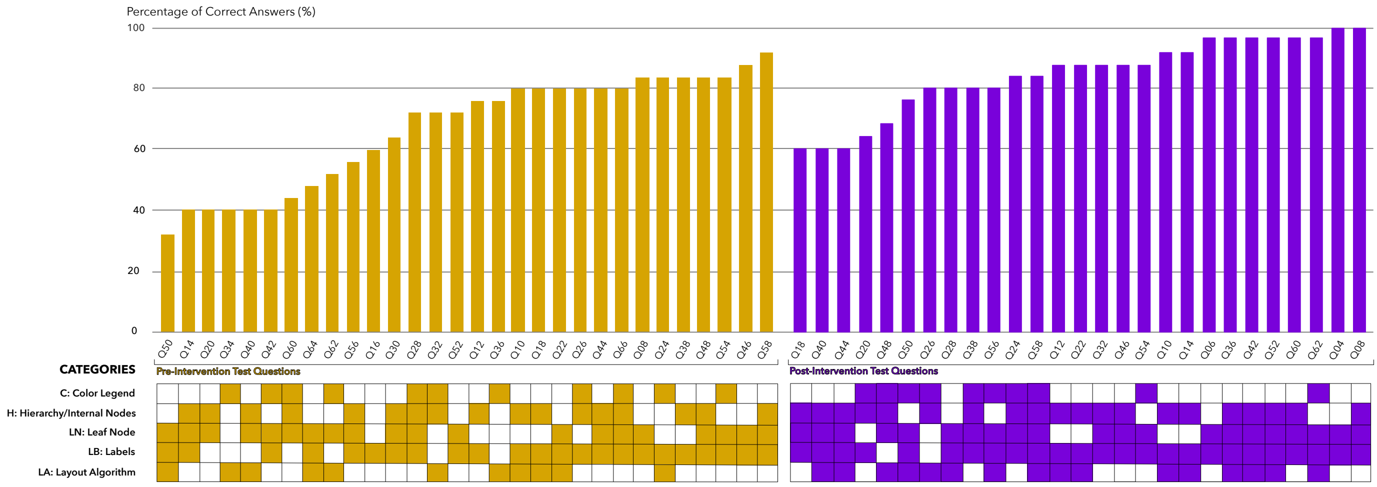


Figure 3.24: *The percentage of correct answers and the classification of questions in the pre- and post-intervention tests.*

3.8 Discussion and Limitations

We coded every question with respect to the treemap aspects that are necessary to understand in order the answer questions correctly. The classification for the pre- and post-intervention questions (see Figure 3.24) that are ranked from the easiest to the most difficult based on the number of correctly answered questions. The corresponding aspects of each question are annotated below it. Contrary to our initial belief that questions might focus on only one aspect, for example, hierarchy, questions require a user to understand three or four features of the treemap simultaneously. This finding indicates that perception of multiple aspects of a treemap is required for its complete understanding and is a barrier to treemap literacy. We also noticed that the most difficult questions were characterized by dense rectangles with only partial labels. Of course, the more dense the rectangle, the more difficult it is (or impossible) to place labels.

Despite the relatively small sample size, the open-ended questions allowed us to gather sufficient data to interrogate and posit reasons for the students' positive or negative experiences with the software, as per our intentions for this study. The majority of the participants who had the opportunity to interact with the software provided positive feedback regarding their experiences.

Nonetheless, a self-selection bias, as well as availability bias, might have played a role in shaping our findings, as the participant recruitment happened over the summer period. Finally, some of the students taking part in the study had some background in data visualization, which could have impacted their ability to navigate treemaps. Despite the split of students with this background between the two groups being equal, we hope to investigate our hypotheses further using audiences from broader backgrounds in our future studies.

3.9 Chapter Summary

We presented a study that explores potential barriers to treemap interpretation and comprehension. In addition, a novel treemap literacy test is presented, which comprises a range of treemap designs and treemap questions classified based on treemap attributes. This work provides a better understanding of the barriers to complete understanding of a treemap and a way for advancing treemap literacy.

Furthermore, we created an interactive educational treemap tool to aid in the training and understanding of a treemap design that enhances for the study of a hierarchical data structure. The results of the user-study indicate that students who interacted with the software outperformed students who learned through slides alone. Also, participants' feedback shows that pedagogical treemap software offers an effective learning experience through more straightforward and faster access to treemap characteristics.

Chapter 4

P-Lite: A Study of Parallel Coordinate Plot Literacy

"If you think in terms of a year, plant a seed; if in terms of ten years, plant trees; if in terms of 100 years, teach the people."

—*Confucius, Philosopher, (551-479 BC)*

Contents

Following the initial study on improving advanced visualization literacy skills of users by focusing on the treemap design, in this chapter, we concentrate on another advanced visual design, parallel coordinate plots used to display multidimensional data. This chapter investigates barriers to parallel coordinate plot literacy and provides a novel literacy test to evaluate users' parallel coordinates literacy skills.

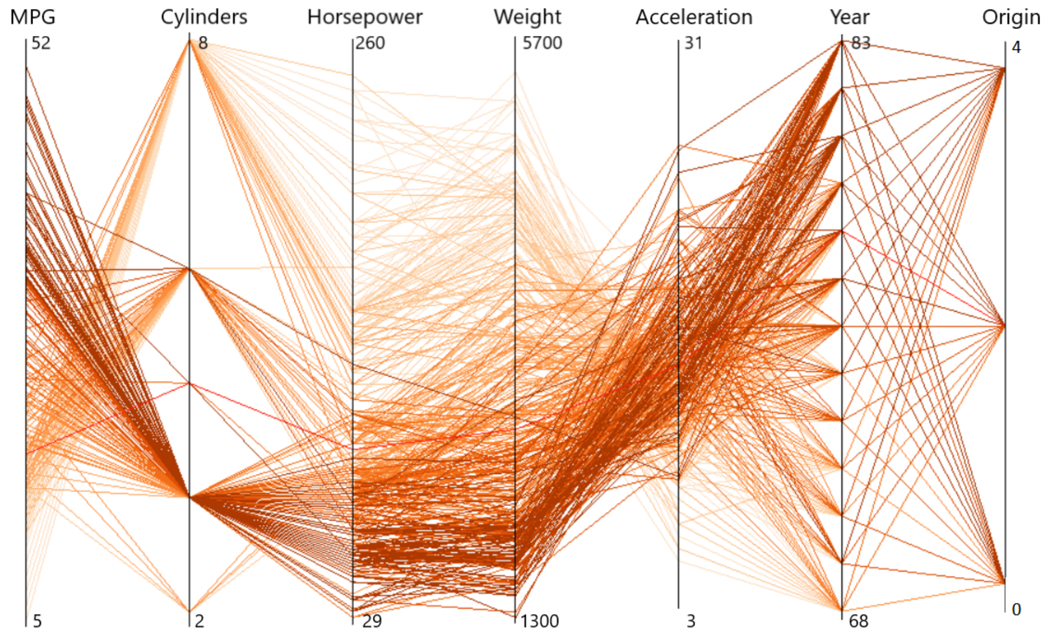


Figure 4.1: *An example parallel coordinate plots of car dataset with 7 attributes. The image was created using Xmdv [23].*

In addition, we developed pedagogical software to increase the literacy level of users and report the results of our user-study. This chapter is based on a technical report [144].

4.1 Introduction

Parallel coordinates plots (PCPs) provide a graphical representation of multidimensional relationships through the use of parallel axes (see Figure 4.1 and 4.2). This design can display high-dimensional data with up to 10-15 dimensions in practice, as each axis is visually separated [145]. Each polyline represents a data record that intersects the parallel axes at given points that indicate the value of individual dimensions. In comparison to Cartesian Coordinate Plot (CCP), for example, PCPs display this multidimensional data in a plane that offers additional advantages (see Figure 4.2). The process of plotting data is different in the CCP and PCP spaces. For the purposes of this chapter, we define PCP literacy as the ability to correctly read, interpret, and construct PCPs. PCP literacy is essential for any user who is interested in understanding multidimensional data, as this is what separates PCPs from other more common visual designs. PCPs, however, have a reputation of being difficult to comprehend, called an expert-only visual design, especially if the

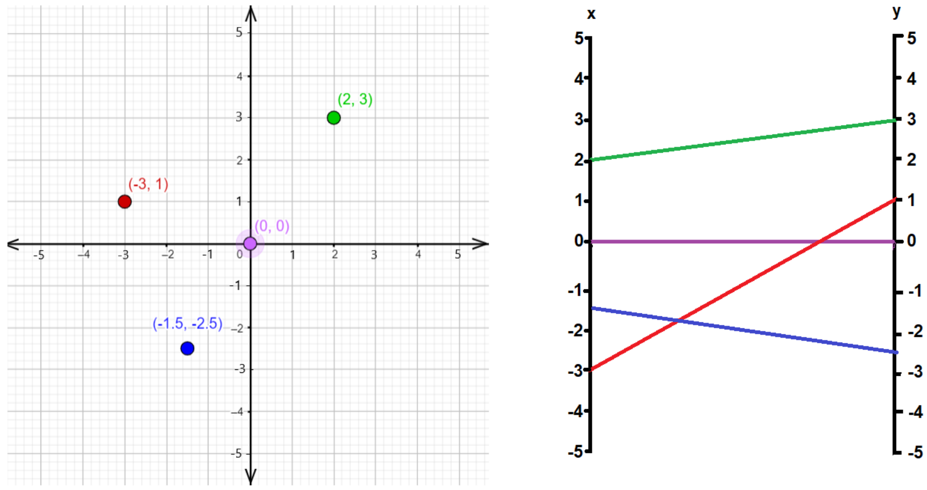


Figure 4.2: *An example of Cartesian Coordinate Plot and Parallel Coordinate Plots with a 2D point data.*

implementation lacks essential features e.g. interaction [146]. To improve the comprehension of PCPs, an assessment of barriers to the understanding and creation of PCPs is required. The aim of this study, therefore, is to identify cognitive obstacles to the successful interpretation of PCPs, and use them to design learning and evaluation of PCP literacy.

We introduce a novel literacy test to investigate the barriers to PCP literacy. In the test, we include datasets and images generated using popular, off-the-shelf PCP tools. Based on our experience of teaching PCPs in the classroom, we develop an interactive pedagogical tool that advances PCP literacy skills by enabling novices to enhance their comprehension, interpretation, and construction of PCPs. As well as supporting literacy skills of users, the tool empowers the effective transformation of data into knowledge and can be used to support an active learning experience in the classroom. Our fundamental hypothesis is that a software tool that interactively links CCP and PCPs will advance PCP literacy more than static slides alone. We assess the learning experience using traditional slides versus our novel software tool and investigate the efficiency of the educational software on PCP literacy with an online, crowdsourced user-study. The main contributions of this chapter are:

1. Identifying and investigating the barriers to PCP literacy;

2. Developing a novel educational tool that facilitates both the teaching and learning of parallel coordinates, advancing parallel coordinates literacy;

3. Developing a novel PCP literacy test [24] and conducting a user-study to assess the impact of our pedagogical software on the comprehension of parallel coordinates.

Supplementary Material	**URL**
Educational PCP Software [25]	https://bit.ly/3ddTMJl
Instructions for the Experiment [147]	https://bit.ly/36UobZH
Slides Video Tutorial [148]	https://bit.ly/36MSRvU
Software Video Tutorial [149]	https://bit.ly/3ix3ZSZ
PCP Literacy Test [24]	https://bit.ly/3xOUEMr

Table 4.1: The table summarizes the supplementary materials with URLs.

Table 4.1 provides a summary of supplementary material for this literacy study. The supplementary material makes the study fully reproducible.

The rest of the chapter is organized as follows: Section 2 introduces related work that includes related literature on visualization literacy and PCPs with a user-study evaluation. Section 3 identifies and presents some of the challenges of interpreting parallel coordinates. The development of a parallel coordinates literacy test is presented in Section 4 and the educational literacy tool is introduced in Section 5. Section 6 explains how the user-study was conducted and Section 7 analyses the results. Discussion is introduced in Section 8. Finally, a conclusion is presented in Section 9.

4.2 Background

There are several previous projects on visualization literacy that study the influence of visual designs on a user's understanding and advance comprehension of visual interfaces. Some studies investigate the user's visualization literacy skills by providing reviews of user-study results. We searched the Survey of Surveys (SoS) on information visualization [4], a survey of interactive visualization for education [44], and a study of information visualization books [110] for related literature on visualization literacy. In this section, we present a collection of parallel coordinate papers that include user-studies with PCPs. While visualization literacy papers focus on investigation and improving novice user's literacy skills, papers with user-studies evaluate PCPs to inform and evaluate the usability of the design. Differences between previous studies and the work presented here are summarized in Table 4.2.

Inselberg has written extensively on parallel coordinates and provides an in-depth guide [150]. We found three survey papers using the SoS [4] that concentrate on parallel coordinates. A survey by Dasgupta *et al.* [151] seeks to identify various sources of uncertainty in screen space and link them to different uncertainty effects on the user. They review the research on parallel coordinates and use a taxonomy to classify different techniques to reduce ambiguity [151]. Another survey provided by Heinrich and Weiskopf surveys the parallel coordinates literature and develops a categorization, aiming to guide research into new topic-related directions [152]. Johansson and Forsell present a detailed literature review that focuses on user-centered evaluation and analyze the usability of parallel coordinates. The goal is understanding how people use PCPs to identify barriers to PCP literacy as well as providing a set of guidelines for future studies [79].

Yang *et al.* [153] propose a system for interactive hierarchical displays (IHDs) to tackle the clutter challenge associated with hierarchical multivariate visualization (HPC) techniques in the study of large datasets. For the evaluation of IHDs, 20 users are asked to identify patterns using two separate forms of parallel coordinate plots: flat parallel coordinates (FPC) and HPC. Patterns have been broadly described as either clusters or outliers. Participants found 8-9 out of the 25 patterns that were used. Subjects using HPC are more successful than subjects using FPC in finding trends in large datasets. This work assesses whether the proposed framework of IHDs provides users with effective help in exploring large datasets. Their study assumes that participants already have a basic knowledge of PCPs whereas ours does not.

Siirtola [154] compares two functionally different visual designs: PCPs and the Reorderable Matrix (RM), and examines how the two visual designs can be combined. An experiment is conducted where 20 participants performed tasks with an application featuring RM and PCP views of the same data, with and without linking. Results indicate that while the view linking originally slows the performance of user tasks, it accelerates learning and is well received by users. Similarly, our study indicates links between CCPs and PCPs rather than RM and PCPs.

Siirtola and Räihä [146] discuss the methods of interaction found in PCP browsers and review them in accordance with existing user interface guidelines. An empirical study is performed in which the usability of PCPs is tested. They indicate the rich interaction opportunities of parallel coordinates and illustrate the value of interactivity for the method. We also incorporate an interaction to display the correspondence between CCP and PCP designs that advances user understanding.

Lind *et al.* [155] provide a new layout of axes for multiple PCP displays to assist users. The display of many-to-many relational parallel coordinates is designed to enhance the visual exploration of relations between variables and can be used to explore objects or object groups. Our work differs by presenting a standard PCP layout rather than a new, enhanced design.

The goal of Claessen and Wijk [156] is to allow users to freely identify and position flexible, linked coordinate axes and specify novel visual layouts by linking these axes with a flexible layout. This enables users to compare scatterplots, PCPs, and radar charts and also create a highly customized layout. The method is tested with 10 users who considered the idea easy to grasp and highly encouraging. This work presents new PCP features e.g. axes layouts that can potentially provide future work for our study. Their study assumes that participants already have a basic knowledge of PCPs whereas ours does not.

	Research Paper	User-study Theme	Visual Design	Image Generation	Evaluation Method	Evaluation Tool	Participant	Task \ Question
Visualization Literacy	Boy *et al.* [37]	Calibration of visualization literacy test	BC, LC, SP	Manual	Crowdsourcing	Amazon MTurk [157]	43	48T
	Ruchikachorn and Mueller [15]	Testing new framework that links unfamiliar visualizations to familiar ones	LC, PC, S, SP, PCP, T	Processing[158]	Crowdsourcing	Amazon MTurk [157]	22	12Q
	Börner *et al.* [5]	Determining the familiarity of users with different visual designs	BC, PC	Manual	In-field	Forms	273	100Q
	Alper *et al.* [8]	Testing users' interests and understanding of their activity	BC, P	C'est La Vis (C)	Classroom	Tablets	21	7T
	Lee *et al.* [22]	Evaluation of content validity and test reliability	BC, CM, H, LN, PC, SC, SP, T	Manual	Crowdsourcing	Amazon MTurk [157]	297	53Q
	Firat *et al.* [42] [CH3]	Evaluation of the treemap test and educational treemap tool	T	The Book of trees[133], students' designs, Google search	Classroom	Desktop Computers	25	57Q
User Studies with PCP	Siirtola [154]	Testing the effect of linking RM and PCP	RM, PCP	RM–PCP Browser (C)	Controlled User-study	Desktop Computers	20	20T
	Yang *et al.* [153]	Assessment of the IHD framework	SC, SG, SP, PCP	IHD Framework (C)	Controlled User-study	Desktop Computers	20	1T
	Siirtola and Räihä [146]	Evaluation of the PCP vs SQL	PCP	PCP Explorer (C)	Controlled User-study	Desktop Computers	16	16Q
	Lind *et al.* [155]	Evaluation of the new PCP axis layout	PCP	Many-to-Many Layout (C)	Controlled User-study	Desktop Computers	12	21Q
	Claessen and Wijk [156]	Evaluation of the new PCP prototype usability	PCP	FlinaView (C)	Controlled User-study	Desktop Computers	10	7Q
	Rosenbaum *et al.* [159]	Testing effectiveness of the new style in pattern detection	PCP	Progressive PCP (C)	Crowdsourcing	Amazon MTurk [157]	43	20Q
	Palmas *et al.* [160]	Evaluation of the new vs classic PCP	PCP	Edge-bundling Layout (C)	Crowdsourcing	University email	137	2T
	Kanjanabose *et al.* [161]	Comparing SPs and PCPs	PCP, SP	Manual	Controlled User-study	Desktop Computers	42	4T
	Kwon and Lee [16]	Investigating the efficacy of the online learning environments	PCP	Manual	Crowdsourcing	Amazon MTurk [157]	120	18Q
	Our Work	**Evaluation of PCP literacy barriers and test development**	**PCP**	**High-D** [162], **Mondrian** [163], **Quadrigam** [164], **PCP Tool**, **Xmdv** [23], **XDat** [165]	**Crowdsourcing**	**Amazon MTurk** [157], **Qualtrics** [142]	**60**	**28Q**

Table 4.2: An overview of the related literature of visualization literacy and PCPs with user-studies. The columns are: user-study themes, visual designs tested, image generation tools and methods, evaluation techniques, the tools used for evaluation, the number of participants included in the user-studies and number of tasks **(T)** or questions **(Q)** asked are provided. The evaluation technique that each research paper uses is categorized into: controlled user-study, classroom setting, and crowdsourcing. Abbreviations used for visual designs include **BC:** Bar Chart, **CM:** Choropleth Map, **H:** Histogram, **LC:** Line Chart, **PC:** Pie Chart, **RM** Reorderable Matrix, **PCP:** Parallel Coordinates, **P:** Pictographs, **S:** Spiral Chart, **SC:** Stack Chart, **SP:** Scatterplot, **T:** Treemap, **IHD:** Interactive Hierarchical Display. **(C)** indicates that the paper introduces a customized tool for image generation.

Rosenbaum *et al.* [159] implement positive PCP to overcome data collection and design challenges due to processing large volumes of data. A systematic study was performed with 43 participants to compare the usefulness of progressive PCP with regular PCP. The participants were asked to perform a variety of exercises, such as recognizing patterns in instances and looking for similarities in various stages in refinement. The findings show there was no major difference between the two methods in terms of accuracy. However, progressive PCP were slightly quicker for pattern detection and, on average, just 37% percent of the data was required to identify the patterns. Their study requires that participants start with a basic PCP literacy level whereas ours does not.

Palmas *et al.* [160] introduce an edge-bundling technique using density-based clustering for each dimension. It enables the clustered lines to be rendered using polygons, significantly reducing the rendering time. They also develop attribute relations with this technique to promote multidimensional clustering. A web-questionnaire with two tasks is given to compare the classic PCP against the new visual design in a user-study. The link to the questionnaire is sent to computer science students and researchers at a local university and in total 137 respondents are analyzed. Their user study assumes basic PCP literacy before participation, whereas our does not.

Kanjanabose *et al.* [161] conduct an experiment including 42 participants to compare scatterplots (SP) and PCP and measure user performance in terms of accuracy and response time using four specific tasks. Three levels of task difficulty are given to users on three representations as a data table, SP and PCP. Similarly, we display the connection between a CCP and PCP, and gauge participants' level of PCP literacy.

Kwon and Lee [16] focus on parallel coordinates to study the impacts of multimedia learning environments for teaching data visualization to non-expert users by examining the effects of active learning theory. To research the efficiency for data visualization education, 18 questions are given to 120 participants in an experiment based on tasks such as mapping between data points and visual elements, data distribution, comparison and similarities. In this research, focus on parallel coordinates to study the impacts of multimedia learning environments for teaching data visualization to non-expert users by examining the effects of active learning theory. The video tutorial and the static tutorial approaches are similar to ours, but this study does not include viewing correlation with the standard method (CCP) to accelerate learning, identify barriers to the comprehension of a PCP, and develop a literacy assessment test.are similar to ours, but the study does not include viewing correlation with the standard method (CCP) to accelerate learning, identify barriers to the comprehension of a PCP, and develop a literacy assessment test.

The difference between research on literacy and research involving user-studies is that, in general, user-studies on PCPs assume a basic prior knowledge of the visual design whereas literacy studies assume no prior knowledge. This is because user studies typically focus on a specific PCP design optimization over the standard layout. Our work assumes no prior knowledge and develops a tutorial to promote literacy. Table 4.2 summarizes how our work compares to previous related work. This study is the first one of its kind focusing on PCPs. Unlike previous work, it compares and links CCPs with PCPs. We also compare the visual design of PCPs with the most popular visual designs to inform and identify the specific barriers to PCP literacy. We design a novel pedagogical tool and conduct a crowdsourced user-study to gather evidence and study barriers to PCP literacy. The focus of our research is specifically on PCP literacy as we provide analysis and guidance to address the barriers to reading, understanding, and interpreting PCPs. Together, our work is a unique combination of barrier identification, PCP tool evaluation, PCP literacy test development, and a user study on PCP literacy.

4.3 The Challenges of Interpreting PCPs

PCPs provide a visual solution to study the properties of multivariate and high dimensional data. A PCP focuses on a continuous vertical dimension when positioning points along axes compared to a Cartesian Coordinate Plot (CCP) which follows both a continuous vertical and horizontal plotting process.

We identify at least seven barriers to PCP literacy based on a review of previous literature (Section 4.2) and survey papers on this topic [79, 151, 152]. This process is supported by our experienced teaching PCPs in the classroom with traditional slides and an assessment of coursework submissions in the Data Visualization module. We systematically reviewed coursework submissions from previous years and identified patterns of errors that are often associated with PCP comprehension and construction, e.g. understanding high dimensional data, overplotting, and relationships between data attributes [152]. This is how we identified hypotheses to PCP literacy barriers: *Space, Multivariate, Correlation, Distribution, and Order*

- **(S) Space** PCPs use an alternative layout of space when compared to CCPs and other popular designs. The most popular visual designs are based on two orthogonal axes whereas PCPs are based on repeated (typically 2-10) parallel axes [22]. Unfamiliarity with this use of axes can create a barrier when interpreting PCPs.
- **(M) Multivariate** An obstacle to parallel coordinates comprehension is the requirement understanding of multivariate ($n \geq 3$) or high-dimensional ($n \geq 5$) data attributes and their relationships [166].
- **(C) Correlation** Identifying the correlation and relationships between data dimensions, as well as knowing how to interpret the slope of the edges, is one of the obstacles to parallel coordinates interpretation. The slope of edges between axes can convey a correlation. This barrier also requires understanding the statistical terminology alone, i.e. correlation, outside the context of PCPs, which make it more complex.
- **(D) Distribution** One of the barriers to interpretation is based on the spread of edges over screen space. The more uneven the distribution of edges is, the more difficult it may be to follow polylines as they cross and obstruct one another. Overplotting can result in higher visual complexity and occlusion [152].
- **(O) Order** Parallel coordinates rely on an axis layout order that specifies placement of axes in screen space. The location of the axes may create an obstacle to understanding the relationships between data dimensions that are not adjacent neighbors in a PCP [152].

We also identify two general visualization barriers with respect to labels and legends. However, we chose not to investigate these general concerns further, because lack of labels and legends means that the PCP design is incomplete. For example, it is difficult to identify data axes with no labels. Although we do not investigate them further, for reference they are:

- **Labels** Labels and minimum-maximum values on the axes facilitate user understanding. Missing axis labels can prevent PCP literacy.

- **Legend** When a color mapping is used in a PCP, a color legend presents the range of values of a data attribute. The absence of a color legend can obstruct understanding of a PCP.

4.3.1 Comparisons with Most Common Visual Designs

Lee *et al.* [22] investigated the 12 most popular visual designs included in the education curriculum, and the most frequently used visual designs in news articles. From *their* Figure 3.1 in Chapter 3, we observe that parallel coordinates plot is *not* among the most popular graphical representations because a PCP is a visual design which is difficult to comprehend. Considering the identified barriers, PCPs have a number of characteristics that differentiate them from the other most common visual designs Lee *et al.* describe:

- It uses a unique space and axis arrangement to plot points along vertical axes that represent data variables. If we look at Figure 3.1 carefully, all other designs use a Cartesian coordinate space except for pie charts and treemaps. Axes are generally orthogonal whereas in PCPs they are in parallel and there are often several of them. (**Space**)

- Parallel coordinates design is used to display multivariate data. Figure 3.1 contains only one multivariate visualization (the stacked bar chart). We believe this is one of the barriers to a comprehension of a PCP. (**Multivariate**)

- The PCP is the only design that focuses on identifying direct relationships between multiple ($n \geq 3$) attributes. (**Correlation**)

- The PCP presents a high number of data records using polylines that cross parallel axes. (**Distribution**)

- The PCP is the only visual design that requires decisions with respect to the order of variates other than the treemap. (**Order**)

- Each data variate is associated with a specific category which may have different types of a attributes. (**Labels**)

Features & Tools	Color Mapped Polylines	Customizable Color map	C. Polyline Color Choice	C. Background Color	Axis Labels	C. Axis Labels	Min-Max Values	C. Min-Max Values	Read Text Data	Removable Axes
GGobi [167]										
High-D [162]										
Mondrian [163]										
Quadrigam [164]										
Sliver [168]										
PowerBI [169]										
Spotfire [170]										
SPSS [171]										
XDat [165]										
Xmdv [23]										

Table 4.3: Tools used to create PCP images for the literacy test. The table shows name the tool and some important features supported such as color mapped polylines, customizable color map, customizable polyline color choice, customizable background color, axis labels, customizable axis labels, min-max values, customizable min-max values, ability to read text data, and removable axes. A green cell indicates support for the corresponding feature.

It is not possible to guarantee that these are all barriers with respect to interpret PCPs. There may also be other barriers and investigating further barriers is a very good opportunity for future work.

4.4 Developing a PCP Literacy Test

We propose a PCP literacy test to assess an individual's literacy skills with PCP images [24]. In order to develop an effective literacy test, we surveyed both related work and parallel coordinates tools for appropriate software and related data while considering the various software features. Our primary hypothesis is that the students taught with the interactive tool demonstration will perform better on the literacy test than those taught with static slides.

4.4.1 Identifying PCP Tools

We reviewed visualization tools that create PCPs using online software collections provided by Keshif [172] and Kirk [173] (see Table 4.3). In addition, we searched for the most frequently used visualization tools to generate PCPs using Google. We also examined students' submissions of an information visualization assignment for two years in the Data Visualization course we taught at Swansea University. Students created parallel coordinates examples with a description of which tools or technologies were used and the datasets used to create their images (see Figure 4.3). We incorporate teaching experience to inform and identify barriers to PCP literacy.

Figure 4.3: *Parallel coordinates tool selection from computer science students' on information visualization assignments in 2018 and 2019.*

Assessing these assignments, we saw some of the same errors repeatedly such as limited understanding of correlation between data variables and multivariate data. We assessed and compared each tool systematically based on the quality of the rendering, the use of color, the presence and legibility of axis labels, and clearly marked axis scales. By quality of rendering, we mean rendering without too much line aliasing. We do not have a specific color that we believe that is required, but generally darker colors on a light background with a good level of contrast. We looked for good quality color mapping too. Moreover, axis labels are necessary in order to interpret what each axis is for in a particular experiments. After testing PCP tools from a variety of sources, we identified a collection of tools to create quality PCPs for the literacy test as follows: **Ggobi** [167], **High-D** [162], **Mondrian** [163], **Quadrigam** [164], **Microsoft PowerBI** [169], **Sliver** [168], **Spotfire** [170], **IBM SPSS Statistics** [171], **XDat** [165], and **Xmdv** [23]. These tools provide quality parallel coordinate plots with clear designs, coloring, and labeling options (see Table 4.3 and Figure 4.4).

Table 4.3 indicates important features of the tested PCP tools. The color-mapped polylines column identifies the ability to apply a color map to a set of polylines while the customizable color map column indicates the option to change the color map. Customizable polyline color is important for changing the color of individually selected polylines. Customizable backgrounds indicate the option to update the background color. Axis labels refer to visible labels at the top or bottom of each axis

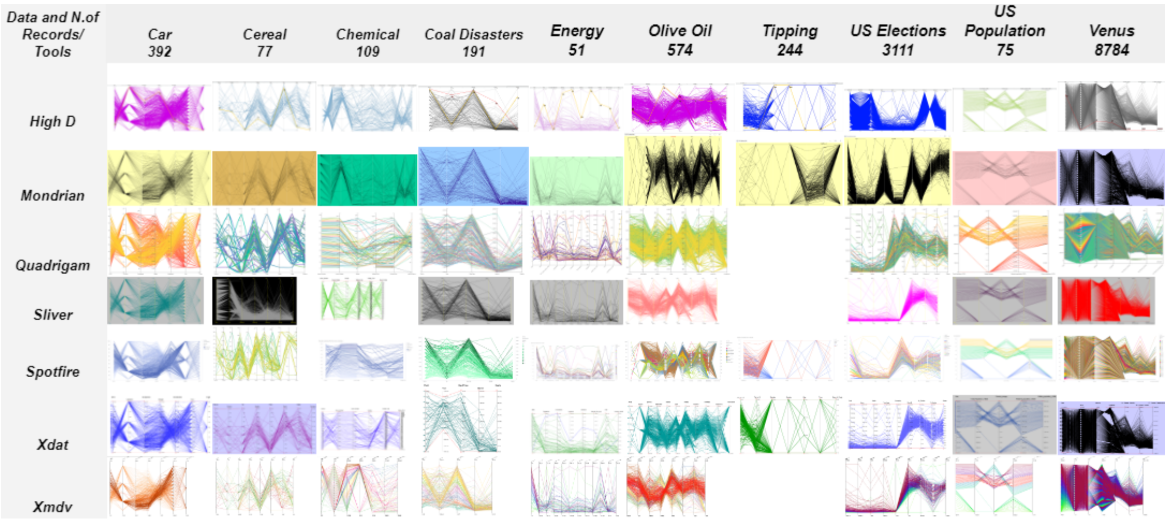


Figure 4.4: *Parallel coordinates matrix used to develop the PCP literacy test. The matrix indicates a name of dataset and tools used to create PCP images. Since Tipping dataset has many text attributes, PCPs images were not created using tools that cannot read text data (see Table 4.3).*

and some tools facilitate modification of axis labels through the software (customizable axis labels). This feature is vital for creating legible PCP images for end users. We use customizable polyline color to create images of polylines with special colors that can be tracked across the display and hence multiple dimensions. We modified some of axes labels to make them more legible. For example, if the axis labels are too long or small, we shortened the labels or increased the font size. Min-max values indicates the software displays min and max attribute values and customizable min-max values implies that user can change these values using the software. As a consequence of changing min-max values, the axes should be re-scaled. We modified a number of min-max values to simplify some of the PCP images including removing a number of unnecessary decimal places. Some tools can read only numerical values. This is important because some datasets features text data. Tools that can read text data are indicated under the *Read Text Data* column. The remove axes feature can enable making some axes visible or invisible on the image. We incorporated this feature to lower the dimensionality of some the datasets thus simplifying some of the images, for example, we aimed keeping *the number dimensions* ≤ 10.

Dataset Name	Number of Records	Number of Dimensions	Name of Dimensions	Data Source	Description
Car	392	7	miles per gallon, cylinders, horsepower, weight, acceleration, year and origin	**Xmdv** [174]	Compares cars produced between 1968 and 1983.
Cereal	77	11	calories, protein, fat, sodium, fiber, carbohydrates, sugar, potassium, vitamins and minerals, display shelf, weight, and number of cups for each serving	**Xmdv** [174]	Provides information on nutritional properties in specific cereal products.
Chemical Elements	109	33	name, period group, chemical symbol, atomic mass, year of discovery, density, melting point, boiling point, enthalpy of fusion, enthalpy of vaporisation, molar entropy, enthalpy of atomisation, etc.	**High-D** [175]	Records characteristic attributes of chemical elements in the periodic table of elements.
Coal Disaster	191	5	months, years, day of year, interval, and deaths	**Xmdv** [174]	Records a number of coal-mining accidents between March 15, 1851 and March 22, 1962.
Energy	51	12	state, total energy consumption, per capita energy consumption, residential sector, commercial sector, industrial sector, transportation sector, petroleum, natural gas, coal power, hydroelectric power, and nuclear electric power	**Xmdv** [174]	Records the energy consumption in US states in terms of energy types and sectors that use the energy.
Olive oil	574	11	area, region, palmitic, palmitoleic, stearic, oleic, linoleic, linolenic, arachidic, eicosenoic, test/training	**Mondrian** [163]	Describes eight chemical measurements on different olive oil samples produced in different Italian regions.
Tipping	244	7	total bills, tip collected, gender, smoker, day, time (day/night), size of team worked	**Mondrian** [163]	The data was collected to show tipping behavior in a restaurant located in a shopping mall.
US Election	3111	53	name, state name, Bush, Kerry, Nader, total, male, female, obese, unemployed, rent, pcturban, urbrural, etc.	**Mondrian** [163]	Records information on the 2004 US presidential election that includes characteristics of on voters' demographics with state name for three candidates.
US Population	75	5	year, total population, percent change, resident population, and civilian population	**Xmdv** [174]	Provides US Population census data from 1900-2006 in thousands.
Venus	8784	7	date, hour, latitude, longitude, PV Plasma, Velocity (km/sec), PD Plasma Density (no/cc) and PT Plasma Temperature (K).	**Xmdv** [174]	The Venus atmospheric data is time-oriented and collected from a NASA mission.

Table 4.4: The original datasets used to develop the PCP literacy test (before modification). The table indicates a name of dataset, number of records, number of dimensions, name of dimensions, data source, and description. Column 5 provides the URL of each data set via citation.

4.4.2 Exploring Datasets

Creating appropriate PCPs for the literacy test is important to assess users' literacy levels. Therefore, we carefully sought to find appropriate and interpretable multivariate datasets through the same websites of tools (e.g. Xmdv datasets [174]) that we selected for creating PCP images. In order to select appropriate datasets for the literacy test, we constrained them using the following criteria:

- Multi-dimensional including 4 or more dimensions,
- A minimum size of 5 records and a maximum of 10,000,
- Studying a theme or behavior that is not too common,
- Clearly documented data dimensions.

The datasets search is intended to create an adequate variety of images and enable test takers with various literacy skill levels to attempt a range of questions. We found 10 datasets identified in Table 4.4. We noticed many previous related papers only test one or two data sets. We wanted a wider variety than many of the previous related work. We also created images with a balance of clusters, correlations, patterns, and outliers across images. The axes are generally ordered in the same order as the original data.

4.4.3 PCP Image Generation Process

Together, we used these datasets and tools to create quality PCP images for the literacy test questions. We imported the datasets using each tool, removed some redundant dimensions, updated some polylines, and background colors for higher quality images. We encountered some difficulties generating PCPs due to some limitations of tools. Not every tool provides the same features such as displaying visible and large axes labels, a color map, showing min-max values, and reading non-numerical values (see Table 4.3). These barriers inspired us to try each combination of dataset with different tools to determine the most appropriate images for the test (see Figure 4.4). Only some images in Figure 4.4 were selected – we tried to choose only those which were of high quality.

Our image selection criteria was also based on using anti-aliasing and color mapping for polylines. As color is useful for a range of purposes, such as aesthetics, color mapping, and visibility options, we used color to experiment with these parameters. For example, color mapping can be mapped to a data variable for one dimension. Color is also used to modify the background in order to display a contrasting polyline color. We followed some rules such as choosing visually appealing colors although this is subjective. We tested for color blindness and ensured that colors of the polylines and background did not interfere with each other e.g. yellow polylines on a white background. We left out the images that have a distracting background or background color that interferes with the polyline color. Moreover, we wanted to make sure that clusters, patterns, outliers, and correlations were balanced across images. For the images that ended up in the test, refer to Table 4.1.

4.4.4 Deriving a PCP Literacy Test

The test was created to examine the understanding of different aspects of parallel coordinates design and the methods of visual exploration and analysis. For the literacy test, a matrix of parallel coordinates images for each dataset and tool (see Figure 4.4) were used to inform the most appropriate examples which provide a good color choice and display data attributes correctly in an image. From the collection, we chose the highest quality images based on the properties in Table 4.3. Once the PCP image selection was completed for the test, we prepared questions with multiple-choice answers for each PCP image. Test questions were derived from previous user-studies [16, 146, 156, 161], involving parallel coordinates and targets the barriers identified in Section 4.3. We chose to derive some test questions from these studies since they have been subject to a refereeing process and approved by reviewers.

PCP Question Classification: We categorized each question with respect to the barriers we identified in Section 4.3. An example is shown in Figure 4.5.

- **(S) Space** requires the user to search the PCP for specific data points.
- **(M) Multi-variate** requires an understanding of two or more variables.
- **(C) Correlation** identifies questions that target understanding of the rela-

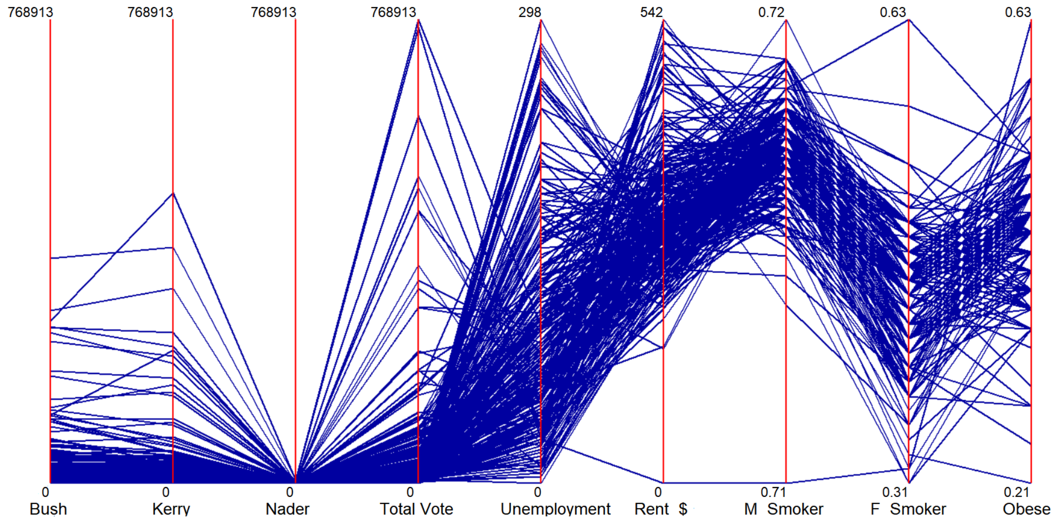


Figure 4.5: *An example PCP literacy test question. Which variable has an indirect correlation with the unemployment rate? Options: A) Bush, B) Kerry, C) Nader, D) Not sure, E) None of the above. The full set of questions can be found at [24].*

tionship between two attributes; direct, indirect, or no relationship.

- **(D) Distribution** indicates those questions that require an understanding of polyline distribution; characterized by how edges are distributed along an axis.
- **(O) Order** identifies questions that are influenced by axis order. (non-neighboring axes)

As a result, we compiled a total of 28 questions for the PCP literacy test. Each question in the test requires understanding at least two challenges to answer. For example, Figure 4.5 displays information from the 2004 US presidential election that includes records of voters' demographics including the names of three political candidates. The data variables are presented as state name, candidates Bush, Kerry, and Nader the total number of votes, rent in US Dollars, unemployment, and obesity rates. The question is *"Which variable has an indirect correlation with the unemployment rate?"* In order to answer this question, the user is required to understand **Space**, **Multivariate**, and **Distribution** aspects. Figures 4.9 and 4.10 record the classification of the each test question in terms of PCP literacy barriers and indicate which aspects of PCPs a user is required to understand in order to answer it. Each image in the literacy test is accompanied by a description of each data dimension. We believe if we do not describe the data dimensions, it would be very difficult to answer the questions without knowing what the dataset and data columns are.

4.5 Developing an Educational PCP Literacy Tool

In addition to a PCP literacy test, we developed a novel, interactive pedagogical tool to enhance a non-expert user's PCP literacy skills. Testing many different existing PCP tools in conjunction with developing the PCP literacy test was very informative when developing our own pedagogical software. Our tool is intended to facilitate the interpretation and exploration of multivariate data as well as enable users to create and interpret PCPs interactively. The software features correspondence between CCP and PCP views and links the points in Cartesian space with polylines in parallel coordinates space (see *PCP Literacy Tool Demonstration Video* [25]). The pedagogical tool is developed as a desktop application and the design is inspired by Alfred Inselberg's software [150] that draws the correspondence between CCP and PCPs. This is the way that Inselberg himself taught PCPs after many years of using slides. We added new functionality to his software; in particular, we increased the number of dimensions, included glyphs, labels, and color mapping. Developing PCP software based on this approach is a novel aspect of our work. The pedagogical tool was created using the Sedlmair [136] design approach, and the design process did not explicitly contain the end-user requirements. The software was built around the idea that most users comprehend CCP with point data. As a result, we introduced the CCP view, which allows users to draw points and see how they correspond with the edges on the PCP view. We also provide the reverse functionality (drawing edges on PCP to draw a point on the CPP) to enable user to have better understanding. The software was created with the C++ language and the Qt framework [138].

Cartesian Coordinate Plot Features (CCP): Seeing the correspondence between the traditional CCP and the new PCP can help users understand the PCP. The interface (Figure 4.6, left) shows a CCP where users can specify points and interactively position them in a traditional 2D coordinate system by clicking on them. Moreover, a user can right-click on each point to resize, change the shape, return to the default shape/location, and remove the point. The color variable is mapped to the x-position of each point by default. Any update to a point will update the x-position and color accordingly.

Parallel Coordinates Plot (PCP) Features: The tool (Figure 4.6, right) presents a PCP with up to five attributes that represent the dimensions such as x-position

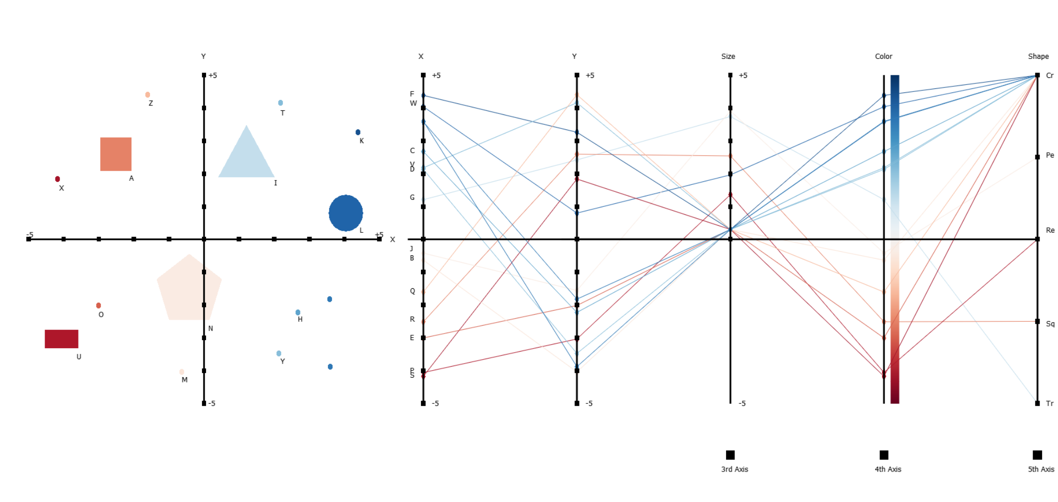


Figure 4.6: *Pedagogical tool interface with Cartesian coordinate space (left) and the corresponding parallel coordinates plot (right). Our PCP Literacy Tool Demonstration Video can be found at [25].*

and y-position in Cartesian space, the size of each point, color, and shape. Each component (edge or point) of the PCP corresponds to a component (point or edge) in the CCP. The size, color, and shape axes can be toggled interactively to observe the direct connection between dimensions in the CCP and PCP. For each point rendered in the CCP, an edge is created in parallel coordinates. The user can hover the mouse over any point or edge. This highlights and displays the correspondence between the CCP and PCP views. Interactive control of either point or edge updates the other in both spaces. Moreover, users can create edges between x and y axes in the PCP that result in the corresponding points in Cartesian space with default size, color and shape settings. This feature facilitates the creation of points by drawing edges and increases the user's comprehension of the relationship between points and edges.

Menu and User Options: The menu options provide more features to the user. The file menu can save the points and corresponding polylines on the screen or load from a previous file as well as deleting all of the current points. These features were intended for a classroom-based experiment which was not possible. However, these features are useful for testing purposes and creating tutorial videos.

Labels: In order to convey the link between the points and polylines, each point and polyline can be labelled with a letter. This feature facilitates understanding by matching a point in the CCP view and corresponding polyline in the PCP with the assistance of the corresponding labels. Also, the user can toggle display labels and change the label font size.

4.6 Experimental Design and Crowdsource User-study

The purpose of this experiment is to evaluate the PCP literacy skills of participants with our novel test by concentrating on the identified barriers, as well as seeking evidence of the educational PCP software efficacy compared to traditional slides in advancing novices' PCP literacy skills.

4.6.1 Experimental Setting

Our original intention was to conduct an in-class user-study similar to that of Firat *et al.* [42] to evaluate our pedagogical software tool. Hence, we believe our software helps both students and the general population. However, enabling the interaction of participants with the PCP software tool during an in-person lecture was no longer feasible. As the researchers were not teaching a visualization class at the time of the pandemic, it wasn't possible to conduct this study with the students at the university where the researchers are based. Thus, we chose crowdsourcing for an empirical evaluation of PCP images and software during this challenging period and we designed a crowdsourced user-study as opposed to a classroom-based one. This choice is inspired by the many crowdsourced experiments performed in the visualization literature e.g. Borgo *et al.* [176]. Crowdsourcing enables us to reach users from a large and diverse pool of backgrounds and to evaluate both varying levels of PCP literacy and the efficacy of the educational PCP tool. We designed the experiment incorporating the checklist provided by Borgo *et. al* [176] for completeness and to minimize error. We split our PCP literacy test for the study into two parts: pre-tutorial and post-tutorial. Both tests contain 14 randomly assigned literacy questions from a larger collection. Each question consists of a PCP image, a description of each image and a multiple-choice question with a single correct answer. Each question is accompanied by a description of each data dimension. We believe if we do not describe the data dimensions, it would be very difficult to answer the questions without knowing what the dataset and data columns are. However, designing a PCP literacy test that requires no data descriptions may be a good future work even though it is challenging. We also used additional PCP questions as simple screening tests. Each PCP design, dataset, and question provided in the test varies in complexity. The tests [24] were administered using Qualtrics [142], an online survey tool for collecting data.

4.6.2 Condition Groups

The user-study participants were divided into two experimental conditions: SLIDES and SOFTWARE. We investigated the impact of our educational PCP software that shows the correspondence between a CCP and PCP, based on improving the PCP literacy of users by comparing the novel software to the traditional PCP slides. For the purpose of this study, we prepared a video tutorial for the SLIDES condition that explains the features of PCP with traditional lecture-style slides only [148] because this is the most commonly used presentation material in the classroom. Time constraints prevent the use of interaction with the dozens of visual designs taught in a visualization class. For the SOFTWARE condition, we prepared a video tutorial that briefly introduces PCPs with slides but also displays the interaction featured in the pedagogical tool demo by showing each of the features described in Section 4.5 [149].

A hands-on approach to interacting with the educational tool is not part of our assessment because it introduces a new variable into the experiment (e.g. hardware). This would make the experiment more complicated and could confound the results. For example, the interaction is different depending on the input device being used (mouse, touchpad, trackball, touchscreen, joystick, etc.). In a crowdsourced experiment, we do not have as much control over the hardware input-device used. Moreover, users often encounter static PCPs rather than tools with which interaction is possible. This is especially true in the classroom and in presentations in general. This also indicates the importance of studying every challenge with respect to PCPs that may enable or prevent users to advance PCP literacy. In some cases the investigation provided by interaction eliminates some barriers to PCP comprehension. Given the importance of interaction, studying this topic is a good direction for future work.

The SLIDES [148] and SOFTWARE [149] tutorial videos are approximately 10 minutes long. Overall, it took approximately 40 minutes for each participant to complete the study. We also believe that length of the tutorial video and the test impact the number of valid responses collected in the user-study. To mitigate for the bias between two groups in terms of the content taught, we delivered the same essential information about PCPs in both tutorial videos e.g. example PCP designs, relationships between axes (direct, indirect, or none). This means that with respect to the topics presented, the slides were the same for both groups. The essential difference between the SLIDES and SOFTWARE groups was the software demonstration video, which therefore comes down to the animation used in the software that interactively displays the links between a CCP and a PCP. Thus, we rely on participants previous knowledge of CCPs in the software. Both groups answer the same set of PCP literacy test questions, e.g. the independent variable.

4.6.3 Participants Screening

We recruited participants through Amazon Mechanical Turk (MTurk) and established a strict screening process to improve data quality. Participants were required to have a total of 1,000 or more approved HITs and a 95% or greater HIT approval rate on MTurk. Before the experiment started, we ruled out participants who self-reported that they were using a mobile phone since the PCP images require significant screen space due to their high level of detail. We also screened out participants who did not recognize the numbers in color blind test images as our PCP images do not use color-blind safe colors. We also identified participants that attempted to take part in the study more than once. We used longitude and latitude data and looked at time stamp information to screen these participants. The responses obtained from those participants were removed from the data to analyze.

4.6.4 Pilot Study-1

Our experiment was initially started with a pilot study to test the experimental design and procedures. A total of 15 participants were recruited regardless of their background and all of them completed the tests. Although participants are informed on the importance of watching the video tutorial in the instruction video, our screening identified that only 3 participants spent the time necessary (or close) to watch the entire tutorial according the data provided by Qualtrics. This initial trial informed us that we need to ask questions specifically about the video tutorials to encourage participants to watch the tutorial video until the end and to pay attention to it.

4.6.5 Pilot Study-2

Based on the experience obtained from the first pilot study, we prepared 3 questions on the tutorial video with multiple correct answers (e.g., *Which topics were mentioned in the video tutorial?, What was the example data about on the last slide of the video tutorial?*, *Which one of the following was a data attribute in the last example?*) to assess their engagement with the presented material. We followed the same procedure in the second pilot study with 10 participants. As a result, even if some participants watched the tutorial video in its entirety, all but 2 participants selected random answers to the video tutorial on all 3 screening questions. Consequently, we decided to carefully screen for attentive participants to gather quality data for the main study.

4.6.6 Experimental Procedure

We published the link of the survey on the MTurk website [157] and asked for Turkers' consent to participate in the study. With their agreement, participants were provided with an instruction video that describes the experimental procedure, followed by some demographics questions. After that, they began the pre-tutorial test which consists of 14 randomly selected PCP images and multiple-choice questions with the goal of determining the current level of a participant's PCP literacy. Upon the completion of a pre-tutorial test, one of the video tutorials was randomly assigned to participants (SLIDES or SOFTWARE). The participants were instructed to watch the video tutorial carefully until the end.

After the video tutorial, participants answered questions about the tutorial and took the post-video tutorial test. The correct answers to pre- and post-video tutorial tests were counted and participants were awarded 1 point in both tests for each correct answer and 0 points for any incorrect answers. To evaluate a participant's PCP literacy using this test, we then calculated the percentages of correctly answered questions by adding all points up and dividing the result by the overall number of questions. These percentages were then used as the data samples for our analysis.

After both tests, 7 open-ended questions were given to collect feedback about the experiment and measure their confidence in the material. Additionally, 4 questions are asked specifically about the educational tool to the participants included in the SOFTWARE condition. After giving the questionnaire for feedback collection, participants were provided with a unique completion code by Qualtrics that is submitted to the survey page on the MTurk website to complete the study.

Each part of the PCP literacy test (pre- and post-tutorial video) begins with two multiple-choice PCP questions as screening checks before randomly selected PCP literacy questions (e.g., *How many parallel axes are there in the image?*, *What is the dataset about?*). Three questions about the tutorial video follow the completion of the video tutorial session. Correct responses to these simple questions about the videos, and the recorded time spent on playing the tutorial video, helped us screen for attentive user-study participants. Inattentive participants were excluded from the data.

4.6.7 Data Collection and Filtering

We continued with the recruitment process until we obtained 60 valid responses from the same number of participants for both experimental conditions in order to provide a balanced comparison and effective quantitative analysis. Initially, we identified a total of 202 attempts to participate in the experiment. After filtering the participants based on color vision deficiency, using mobile phones, or those who tried to participate more than once from the same location, 170 participants remained. In the next step, we looked at the time participants watched the video tutorial, the answers given by the participants to the screening questions, and the simple questions about the video tutorial. At the end of filtering process, we collected responses from a total of 60 participants which is a comparable sample size to previous crowdsourced studies [176].

The number of participants involved in each crowdsourced user-study in Table 4.2 shows that three studies use fewer than 60 participants and three others use more than 60 participants. This puts our 60 participants at the median of the related literature. We also believe that the duration of the study has an impact on collecting data provided by attentive participants (approximately 40 minutes). Thus, we screened out a total of 110 participants (65%) from the experiment after about a

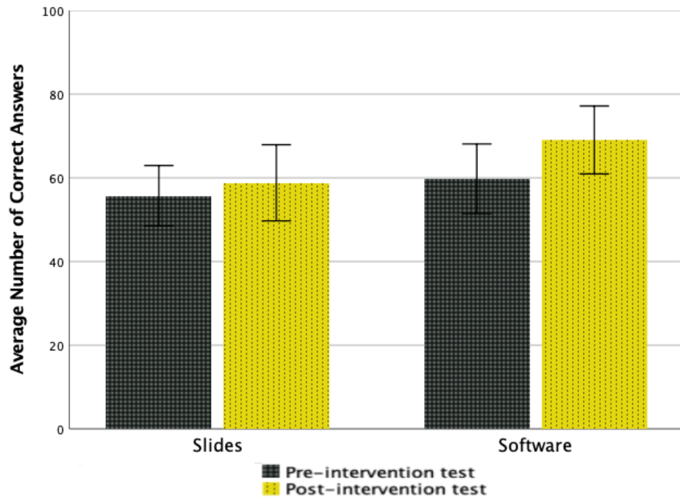


Figure 4.7: *The percentage of correctly answered questions in the pre- and post-intervention tests for* SOFTWARE *demonstration and* SLIDES *groups. Error Bars (95% CI).*

week spent for collecting data.

Among the 60 participants (30 participants per condition SLIDES and SOFTWARE), the majority of the participants reported that they were from the United States (29 US, 15 India, 10 Europe, 3 Brazil, 1 Bangladesh, 1 Indonesia, 1 Pakistan). Participants (36 male and 24 female) had different degree levels: 28 Bachelor's, 17 Master's, 8 Ph.D., and 7 High School. The age of participants ranged from 18 to 60 with an average age of 33.4 years.

4.7 Results

Our primary hypothesis was that the students taught with the interactive tool demonstration would perform better on the literacy test than those taught with static slides.

4.7.1 Quantitative Results of Test Data

The percentages calculated for each participant were normally distributed, as indicated by the Shapiro-Wilk test for both pre- and post-intervention test groups. Hence, we used one-way ANOVA for our data analysis (significance level at $\alpha = 0.05$). The pre-intervention test results did not differ significantly between the two groups: $F(1,58) = 0.564$, $p = 0.456$, $\eta_p^2 = 0.010$. Those participants who then watched the SLIDES video correctly answered on average 55.71% of the pre-intervention test questions (SD = 19.35%), and the participants who were in the SOFTWARE condition

	SLIDES					SOFTWARE				
	Pre-Intervention	Post-Intervention	$F(1,29)$	p	η_p^2	Pre-Intervention	Post-Intervention	$F(1,29)$	p	η_p^2
S: Space	56.33 ± 25.26	59.57 ± 26.34	0.392	0.536	0.013	64.47 ± 27.73	72.70 ± 22.47	3.724	0.063	0.114
M: Multi-variate	52.90 ± 19.30	54.77 ± 27.72	0.122	0.729	0.004	55.93 ± 22.07	68.30 ± 24.00	10.153	0.003*	0.259
C: Correlation	47.60 ± 30.70	56.43 ± 31.88	1.018	0.321	0.034	46.00 ± 30.40	56.93 ± 29.71	3.602	0.068	0.110
D: Distribution	60.07 ± 31.34	63.97 ± 19.64	0.346	0.561	0.012	65.37 ± 30.32	72.90 ± 20.88	2.018	0.166	0.065
O: Order	42.50 ± 34.45	54.97 ± 33.37	1.778	0.193	0.058	41.03 ± 33.50	54.03 ± 32.06	3.862	0.059	0.118

Table 4.5: The results of pre- and post-intervention tests for the SLIDES and SOFTWARE groups ($M \pm SD$ in percentages), based on the categories of questions. Significant results are shown as follows: $*p < 0.05$.

answered 59.76% of these questions (SD = 22.31%).

As for the post-intervention test results, those who were in the SLIDES condition on average responded correctly to 58.81% of the questions (SD = 24.36%), while the subjects who watched the SOFTWARE video answered 69.05% of the questions correctly (SD = 21.74%). This difference, however, was not significant: $F(1,58) = 2.950$, $p > 0.05$.

The SLIDES group have seen a 3% increase in the percentage of correctly answered questions from pre-intervention test to post-intervention test. This increase was not significant: F(1,29) = 1.796, p = 0.191, $\eta_p^2 = 0.058$. On the other hand, the SOFTWARE group performed significantly better in the post-intervention test than in the pre-intervention test, having improved their results on average by 9%: F(1,29) = 8.092, p = 0.008, $\eta_p^2 = 0.218$ (Figure 4.7).

Further to these findings, we hypothesized that the software tutorial video would provide additional support in overcoming different barriers to understanding PCPs. Thus, we have also looked at the participants' performance in the different parts of the test aimed at measuring one's comprehension of different attributes of parallel coordinates. We did so by looking at the question classification based on the parallel coordinates features that could influence the participants' answers in the test.

In the post-intervention test, there was no difference in the results obtained by participants in both groups for the questions about Correlation (**C**), Distribution (**D**), and Order (**O**). However, participants who were in the SOFTWARE condition performed much better than participants in the SLIDES condition when answering the questions related to the Space (**S**) (F(1,58) = 4.316, p = 0.042, $\eta_p^2 = 0.069$) and Multi-variate (**M**) (F(1,58) = 4.087, p = 0.048, $\eta_p^2 = 0.066$) categories.

Within condition, comparing pre- and post-intervention, participants in the SLIDES group did not improve their performance significantly in any specific category of the questions. Within the SOFTWARE condition, our participants significantly increased the number of correct responses for the Multi-variate (**M**) category of questions (Table 4.5), but no significant improvement was observed in the other four question categories.

4.7.2 Qualitative Analysis of Feedback Data

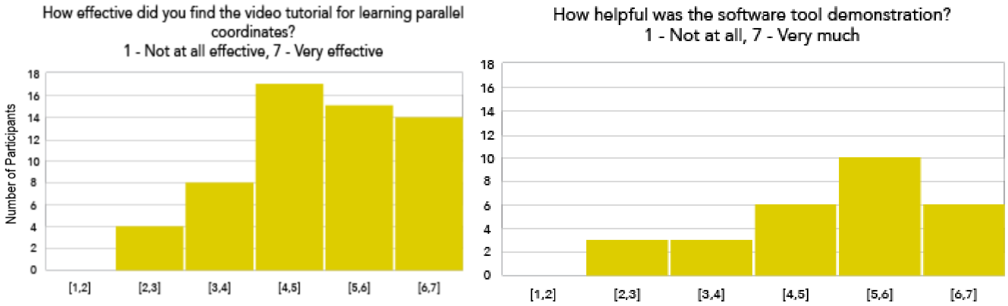
To analyze the feedback collected from our participants at the end of the survey, we followed a lightweight approach related to the initial inductive thematic analysis steps as described by Braun and Clarke [143]. We divide the themes based upon whether they come from questions about the PCP literacy test, or from questions about our PCP literacy tool.

PCP Literacy Test

We identified three themes about the PCP literacy test questions: *Impact of the tutorial video*, *Distribution*, and *Correlation.*

***Impact of the tutorial video*:** Participants were asked to reflect on their experience of answering post-intervention test questions. The feedback from some participants in both conditions indicates a high difficulty level of the test questions although they watched the tutorial video and answered randomly chosen questions. Some 14 participants stated questions were generally difficult: *"All questions are very difficult to find the answers."* (P1) and *"More difficult but not much"* (P15). This result reveals that although the test questions have varying difficulties, the PCP literacy skills of these participants are weak. However, 16 participants (27%) stated that the post-tutorial test was easy or became easier after watching the video tutorial: *"After the tutorial, it is easy to identify what is meant by the image, the only normal difficulty of evaluating a problem"* (P42). It is clear that most of participants found the video tutorial very helpful in Figure 4.8. Participants also indicated an increase in confidence after watching the tutorial video.

***Distribution (D)*:** A high number of data records using polylines that cross parallel axes result in overplotted PCP designs – a barrier to PCP literacy. This challenge may be the result of not understanding crowded PCP images. This can play a significant role in participants' performance during the test. This difficulty is expressed by 5 participants (8%): *"There were complex plots with intermeshing lines being hard to follow the destination"* (P33) and *"[The questions were] very difficult because the*



(a) (b)

Figure 4.8: *Participants' answers to feedback questionnaire about video tutorials.*

data showed was very messy" (P48).

***Correlation (C)*:** Observing values and axes labels to understand the relationship between two neighboring data attributes (direct, indirect, or none) is required to answer correlation questions correctly. Consistent with the quantitative data analysis, the qualitative feedback collected from participants in both groups did not show statistical improvement in every category (except the Multi-variate and Space categories in the SOFTWARE group). According to the feedback data, some 5 participants (8%) stated that they had difficulty understanding the correlation topic and answering the questions related to it: *"Before and after [the tutorial video], I still had the most difficulty with direct and indirect correlation."* (P23), *"I was torn between the direct and indirect."* (P27), and *"I had a hard time remembering the difference between direct and indirect."* (P47). This study gathers both qualitative and quantitative evidence identifying correlation as a barrier to PCP literacy. We hypothesize that this is due to the mathematics background required to understand this topic.

Educational PCP Software – We also evaluated the feedback directly relevant to our pedagogical application. The feedback was collected from participants who were provided with the tutorial video of the educational PCP software. Most of the participants (73%) reported that the educational PCP tool was effective and helpful for a better understanding of the relationships between data attributes. We coded the feedback based on the features of the software tool that were recognized as having a positive learning effect on the participants. We identified two aspects: *Multi-variate*, and *Correspondence and Animation.*

***Multi-variate (M)*:** Some six participants (20%) identified that the PCP tool was helpful to understand the PCP concept and convey the mapping of multivariate data which is one of the main challenges of high dimensional data. Participants stated: *"Connecting multiple data [attributes] in a single graph is better to get information easily."* (P41) and *"It [the tool] helps in understanding data with more than 2 variables."* (P43), and *"It [the tool] provides a direct visual demonstration of the mapping between axes."* (P50).

***Correspondence and Animation*:** The educational tool demonstrates the correspondence between views of the CCP and the PCP using animation. Some five participants (17%) stated how the feature enabled them to understand the data through interaction between plots: *"It [the tool] provides to represent the data on both charts at the same time that explained everything clearly."* (P33) and *"It gave a good real-time illustration of the relation between the two coordinate types, and it [the tool] gave a clear way of displaying the data, with a more easily identifiable scale."* (P38), and *"Being able to see the changes from parallel coordinates chart to Cartesian chart made for better understanding of the relationships between data attributes."* (P56)

The Summary of Evidence– The results of quantitative analysis indicate that the software group has experienced a 9% increase in the percentage of correctly answered questions from the pre-intervention test to the post-intervention test while the slides group had improved their results on average by only 3%. According to the feedback collected from our participants at the end of the survey, most of the participants (73%) in the group who watched the software video reported that the educational PCP tool was effective and helpful for a better understanding of the relationships between data attributes. They rated the software tool demonstration on average of 5-6 out of 7 (1-Not at all, 7-Very much) in terms of being helpful (see Figure 4.8b) and responded to our question on the efficiency of the software tool positively e.g. *"Because it [the tool] allows learners to see how the graph works in action."* (P123) and *"It [the tool] gave a thorough explanation. And so, I felt like I could move my way around more effectively."* (P105). These findings provide evidence that the software tool facilitates PCP literacy and supports our hypothesis.

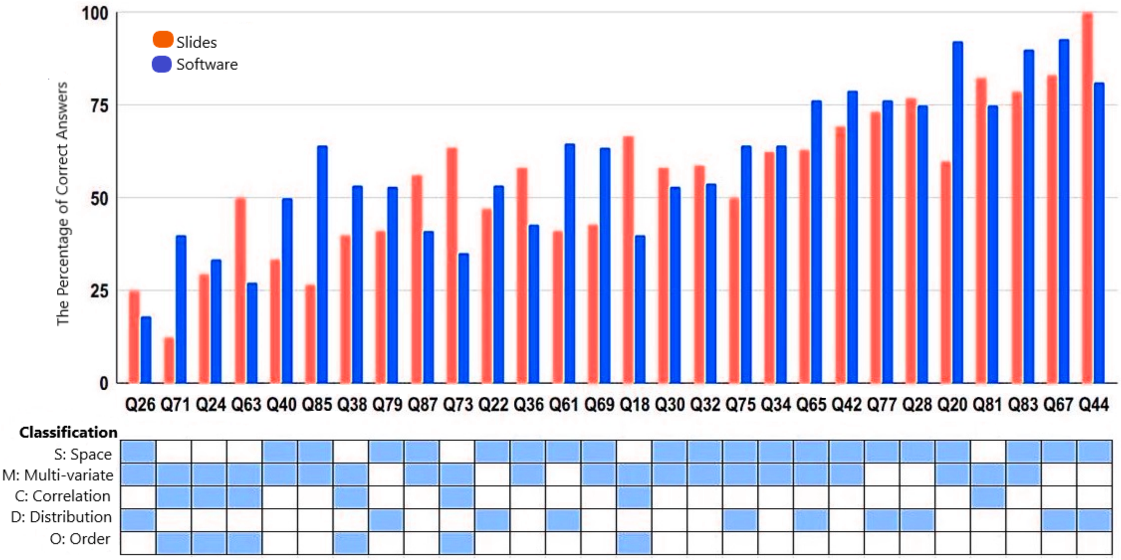


Figure 4.9: *The percentage of correct answers on the pre-tutorial test by groups and the classification of questions. The questions are ranked from the most difficult to the easiest.*

4.8 Discussion

4.8.1 Designing Questions for the PCP Literacy Test

Figures 4.9 and 4.10 couples the classification of PCP literacy test questions with percentage of correct answers given by the participants in both pre- and post-tutorial tests. Questions are ranked from the easiest to the most difficult based on the percentage of correctly answered questions for both groups. The corresponding testing categories of each question are given below (see Appendix B for example questions). Originally, we thought we could design test questions that isolate the different cognitive processes required to interpret the PCP images. However, this did not turn out to be true. Observers usually have to understand two or three cognitive properties of PCPs in order to successfully answer PCP questions. We hypothesize that this confounding effect is itself a PCP literacy barrier. In addition, the results indicate that some participants may not be very familiar with the new terminology. Understanding the terms may also be a barrier to PCP literacy (e.g. Correlation). One of the participants stated that *"There were a lot of terminologies to remember, and so that increased difficulty."* (P44). This indicates reinforcing the prerequisite terminology is an essential step in advancing a non-expert users' PCP literacy level.

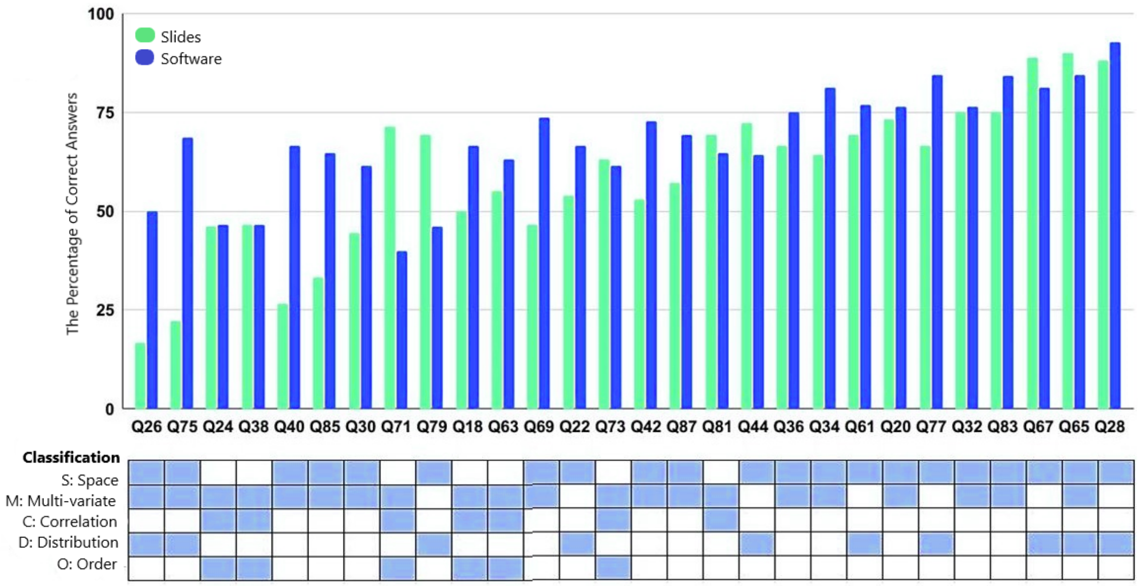


Figure 4.10: *The percentage of correct answers on the post-tutorial test by groups and the classification of questions. The questions are ranked from the most difficult to the easiest (different to pre-tutorial).*

4.8.2 The effectiveness of the PCP Literacy Tool

The pedagogical PCP tool shows the correspondence between CCP and PCP views to improve PCP understanding. The two views are connected and synchronized so that the mapping of multivariate data to geometric primitives is clearly illustrated. To lessen the impact of barriers to PCP literacy, the PCP tool facilities toggling axes in the display between different data attributes **(M, C, O)**. Points in the CCP can be added, deleted, or moved to demonstrate data distribution **(D, S)**. The experimental results and Figure 4.8b reveal the positive effect of the educational tool video tutorial on PCP literacy, based on participants' experience with the PCP tool demonstration. Although participants did not explicitly address all individual barriers to understanding PCPs in their feedback, overall feedback on the tool was very positive as indicated in Figures 4.8 (a) and 4.8 (b).

Answering the literacy test questions requires perceiving more than one property of PCP. To help others who are not significantly influenced by the literacy tool, we can add new features such as loading new datasets and axis reordering axes as future work. A user-study dedicated to interaction in a controlled lab (an experiment which is currently not possible) might also be insightful.

4.9 Chapter Summary

We examined barriers to understanding PCPs. We introduced a novel literacy test that includes tools and datasets to assess non-expert users' literacy skills with a variety of PCP images. We also compiled a list of the most often used visualization tools and multivariate datasets for generating PCPs for our literacy test. To aid with the comprehension and investigation of multivariate data, we created an educational PCP tool and used the tool for the evaluation of users' PCP literacy skills. The results showed that students who trained with the pedagogical practice training video scored better on the post-tutorial literacy test. Participants stated that depicting the relationship between the PCP and CCP of view is a convenient method to improve PCP literacy and facilitate understanding of multidimensional data.

Chapter 5

DPCP Vis: Techniques for Dense Parallel Coordinate Plots

"To acquire knowledge, one must study; but to acquire wisdom, one must observe."

—Marilyn vos Savant, Writer (1946-)

Contents

In Chapter 4, we gain a deeper understanding of barriers to PCP literacy. Informed by our study in Chapter 4, this chapter addresses the challenges of understanding large amounts of data due to overplotting and the relationship between data dimensions on the PCPs. To help overcome these challenges we introduce interaction techniques to facilitate comprehension of data and data dimensions. The RAMP VIS project [1] inspires the study, and the results of our user-study provided in Chapter 4 reveal that identifying correlation can be a barrier to PCP literacy. Finally, we discuss our novel techniques and provide feedback from domain experts. This chapter is based on a technical report [48].

5.1 Introduction and Motivation

The Parallel Coordinate Plot (PCP), introduced by Inselberg [177], is a visual design showing multidimensional relations using parallel axes. PCPs facilitate data exploration and understanding relationships for multivariate data. One of the well-known challenges with PCPs is associated with overplotting. Rendering thousands of polylines causes overlapping edges that may obscure the underlying patterns in the image, especially in high data density areas [178]. In these cases, interaction can be crucial in exploring the data and minimizing ambiguity. However, processing and analyzing overplotted data requires new approaches to support understanding. We propose novel visual feature and interaction methods to address challenges in PCPs that occur as a result of overlapping line segments. We introduce interactive glyph lenses that enable users to explore an overplotted area using a dynamic lens that hovers over the PCP based on mouse location. This interaction summarizes edges that intersect with the lens represented by arrow glyphs showing the average slope of a dense collection of edges. To convey relationships between dimensions, we display arrow glyphs placed below each adjacent pair of axes that indicate the correlation. We introduce a dimension reduction technique that enables users to evaluate a PCP by looking at the correlation value between neighboring axes and collapsing axis pairs that do not add information to the display. We also present a user option we call a subtraction operator, Δ, that displays the difference between two multi-dimensional data sets for quick comparison. The Δ operator addresses the unsolved problem of visually comparing multivariate ensemble data. In this chapter, we specifically concentrate on interaction techniques for dense PCPs. The main contributions of this study are as follows:

- The introduction of interactive correlation glyphs for adjacent axis pairs
- Novel dynamic glyph lenses to support data analysis and comprehension
- A subtraction operator, Δ, to indicate differences between two multi-dimensional data sets
- Relationship-guided dimensionality reduction based on collapsing of axis pairs to reduce redundancy

We evaluate our methods with a case study based on the simulation of Covid-19 contagion behavior together with a modelling expert in this area. Visual comparison of ensemble data is considered an unsolved problem [179].

The rest of the chapter is organized as follows: In Section 2, we review the previous work on reducing the impact of clutter in PCPs. In Section 3, we demonstrate interaction design including correlation glyphs, dynamic and static lenses, and the Δ operator. In Section 4, we discuss the performance of our visualizations and provide feedback from domain experts. Section 5 wraps up with conclusions and future work.

5.2 Background

Displaying a large multivariate data set in a 2D space has always been a challenge for data exploration due to over-plotting and clutter. We start by reviewing related surveys and focus on literature for the discovery of the information in dense and cluttered areas in PCPs.

Surveys: Dasgupta *et al.* [151] investigate different types of ambiguity in the PCP images and introduce a taxonomy for classifying them to reduce uncertainty. By creating a taxonomy, they aim to detect distinct sources of uncertainty in the design and link them to different impacts of uncertainty for the user. Similarly, Heinrich and Weiskopf [152] propose a taxonomy and assessment of strategies for modeling, visualizing, analyzing, and interacting with PCPs, as well as a classification of common tasks for investigation. Johansson and Forsell [79] summarize and categorize studies on evaluating PCPs. A thorough examination of previous research presents user-centered evaluations to report on the human-centered aspects of PCPs.

In this section, we focus primarily on previous work on PCPs that address visual clutter and ambiguity. We briefly introduce solutions to analyze large data on PCPs. In general, the methods for reducing the impact of clutter on dense displays can be categorized as frequency-based, using interaction and brushing, clustering, and edge-processing.

Frequency-based: Artero *et al.* [180] present a method for creating frequency and density plots from PCPs. The new plots enable interactive data exploration of large and high-dimensional data, enabling users to remove noise and highlight data-rich areas. Work by Geng *et al.* [178] proposes angular histograms and attribute curves that enable users to investigate clustering and linear correlations in large data sets to address over-plotting and clutter in PCPs. The state-of-the-art reported by Henrick and Weiskopf [152] has a particular subsection on frequency-based techniques that address aggregating edges together as an approach to overplotting and provides numerous methods for aggregating the data [159, 181, 182, 183, 184]. Our work incorporates a frequency-based approach that counts the number of edge intersections with an interactive lens.

Interaction and Brushing: Blass *et al.* [185] present quantization and compression techniques for data pre-processing, as well as joint density distributions for adjacent variables enabling efficient GPU-based rendering of PCPs. In addition, they propose faster brushing methods for interactive data selection in several linked views. Raidou *et al.* [186] introduce a novel technique, Orientation-enhanced PCPs, to improve the view by visually enhancing segments of each PCP line emphasizing slope when there are several overlapping edges or when outliers and structures are obscured by noise. A novel effective selection method, the Orientation-enhanced Brushing (O-Brushing) is also presented that eliminates unnecessary user interaction. Another brushing method to enhance dense PCPs by Roberts *et al.* [187] introduces higher-order, smart data-driven brushing, and sketch-based brushing. The sketch-based brush is generated by connecting mouse clicks across the PCP on each axis at the chosen brush-axis intersection. Smart brushing assists the user during interaction by revealing patterns at run time. Some of our methods are based on interaction, however, none involve traditional brushing on PCPs.

Clustering: Data clustering is one method for reducing clutter in a PCP. Fua *et al.* [181] use hierarchical clustering to create a multiresolution representation of the data, and a variation on the PCP to express aggregated information for the clusters that facilitates navigation and filtering to explore the patterns and trends in the data. Ellis and Dix [188] propose several approaches for measuring occlusion by interactively adjusting the level of sampling. They explore three algorithms (raster, line, random) to measure the degree of occlusion. When compared to other algorithms, the raster algorithms result in higher accuracy.

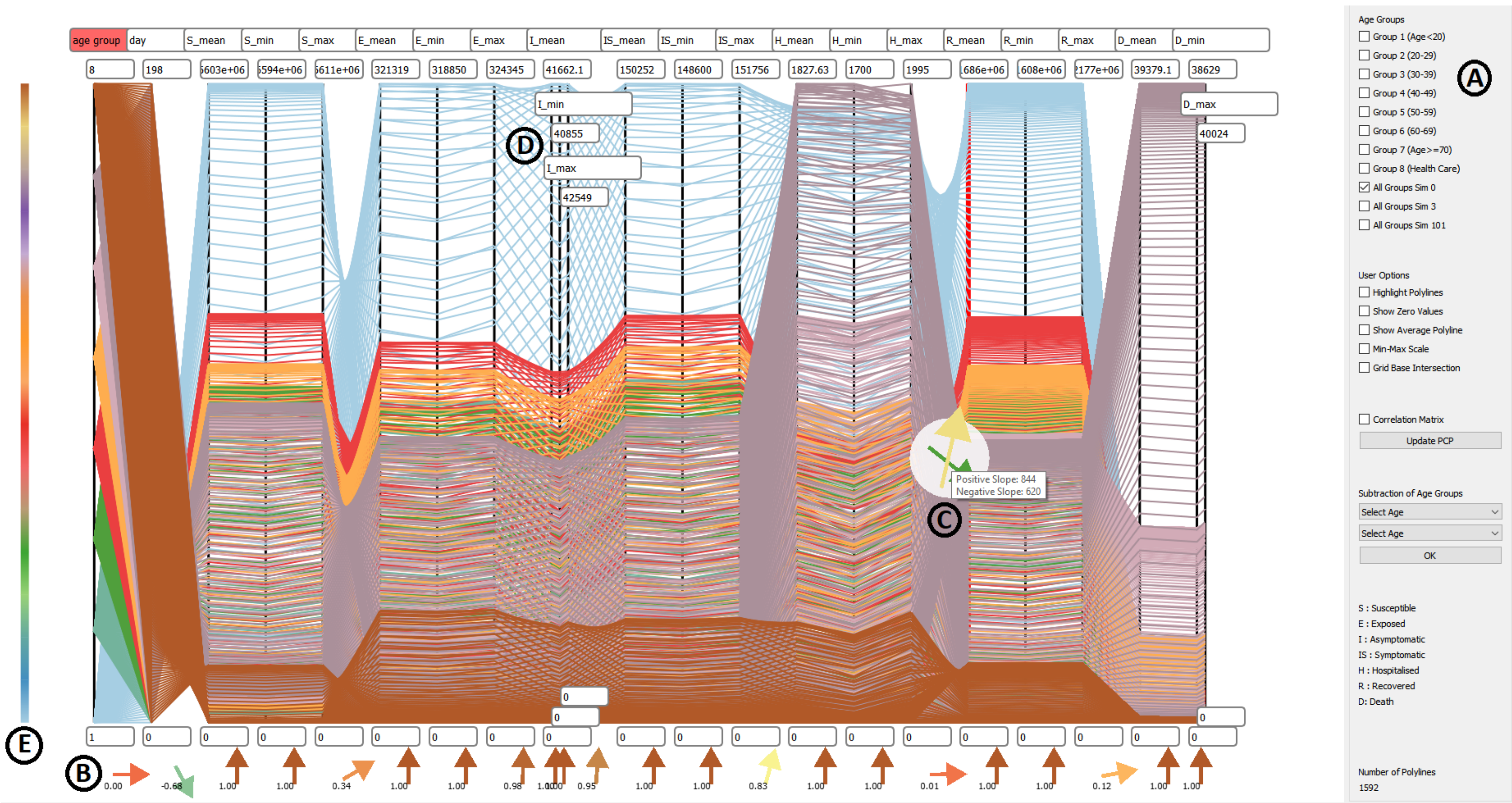


Figure 5.1: *Overview of the PCP software tool. (A) The image displays user options, (B) the data with correlation glyphs under each axis pair, (C) interactive feedback in a dense area with an arrow glyphs lens, (D) collapsed axes pairs with stacked labels, and (E) a color legend. The PCP displays the predictions that the number of recovery in those under the age of 20 (Group 1) and the number of deaths in patients over the age of 70 (Group 7) will be higher than in other age groups. It also shows that mortality is lower for health care workers.*

In addition to hierarchical clustering and calculating polyline occlusion techniques, Johansson *et al.* [189] use transfer functions to display different characteristics of clusters and transform each K-means-derived cluster to high-precision structural texture that, applied to a colored polygon, creates the cluster's final visual appearance. Our methods do not use explicit clustering. However, the lens we introduce summarize the edges that pass through them depicting average slope.

Edge-Processing: McDonnell and Mueller [190] introduce a technique that shows each data point as a poly curve to facilitate edge bundling and declutter the display. Palmas *et al.* [160] present an edge-bundling technique that applies density-based clustering for each dimension. It represents the clustered lines as polygons, which reduces rendering time. They also use this strategy to enhance multidimensional clustering by developing attribute connections. Divino *et al.* [191] describe an edge bundling strategy used in PCPs to expose cluster information directly from the overview. The edge-bundling survey by Lhuillier *et al.* [192] presents a data-based taxonomy for classifying bundling methods and introduces a framework to describe the steps of bundling algorithms. The survey provides a subsection on PCPs and describes edge bundling papers that apply edge bundling for reducing the clutter and increasing readability [184, 190, 193, 194]. Our dynamic lens could be considered as a kind of edge processing technique.

In contrast to previous work, the techniques we describe generally focus on the space between axis pairs rather than on axes themselves. Most previous literature focuses on either the parallel axes or the polyline edges. We focus on supporting cognition of relationships between axis pairs in the context of dense PCPs. We introduce novel techniques to facilitate data analysis guided by correlation glyphs between neighboring axis pairs, showing the differences between data sets using a subtraction operator, and enabling the user to reduce dense areas and dimensions by collapsing axis pairs.

5.3 Visualization Design

This work is completed in part in partnership with the RAMP VIS [1] team, who are assisting the modelling scientists and epidemiologists in the Scottish COVID-19 Response Consortium (SCRC) [195]. They provide the ensemble data (see Section

5.3.1), which is a significant motivation of the techniques we present here. In addition, we have brainstorming sessions with visualization experts, modellers, and statisticians. These conversations aid in the comprehension of the data simulations, the exploration of the most significant input parameters. We also identify the requirements of modelling scientists and epidemiologists as well as getting the visualization experts' feedback on the software. We follow the design study methodology introduced by Sedlmair [136] to evaluate a specific real-world problem encountered by domain experts and design a visualization system to aid in problem-solving. In addition to that, we adopt the nested four-level model for visualization design and evaluation by Munzner [196]. First, we describe the problems and data before mapping them into abstract processes and data types. We then created the visual encoding and interface to enable those processes. Finally, we develope an algorithm to perform that design efficiently. The software was written in C++ using the Qt framework [138].

In order to convey the strength of the correlation between axis pairs, correlation glyphs for each adjacent pair (Section 5.3.2) are presented in the PCP view. This provides users with a summary perspective of the multivariate relationships and an improved understanding of the link between axis pairs, which may not be visible by glancing at a dense set of edges. One of our techniques for dense displays is based on detecting the intersection of the edges with a glyph lens. The lens offers interactive feedback to the user as a function of the current mouse position that specifies center of the lens in the PCP (Section 5.3.3) in dense areas where the relationship between the axes may be difficult to interpret. The Δ operator (Section 5.3.4) is one of the techniques developed in order to understand the difference between two comparable data sets. Also, axis pairs can be collapsed (Section 5.3.5) through a selection that enables users to view a reduced set of axes, motivated by redundant information.

Figure 5.1 shows an overview of the PCP tool we developed that allows a user to view different data sets via the user interface on the right of the screen (A). To demonstrate the relationship between each adjacent axis pair, correlation arrow glyphs are positioned under the PCP view (B). The figure also shows an example of a dynamic edge glyph lens (C) and some collapsed axes with stacked labels (D). The color scale on the left (E) is initially mapped to the edges on the first axis. This can be updated by selecting another axis. One of the user options offered by the tool is to display data labels and points where an edge crosses the axes by hovering the mouse over the edges and highlighting them. In addition, features such as rendering the average edge by taking the average of all edges and showing the zero point on the axes are also supported.

5.3.1 Ensemble Data from a Covid-19 Simulation

The ensemble data we study is a major motivation for the techniques we develop here. RAMP VIS [1] is a VIS volunteer group that responded to a call by the Scottish COVID-19 Response Consortium (SCRC) to support modeling scientists and epidemiologists [195]. The primary objective is to build a stronger and improved understanding of possible strategies to deal with the Covid-19 outbreak in the United Kingdom. We study the ensemble data set provided by the modelers by processing the large amount of simulation data given to the RAMP VIS group in our study. The data includes hundreds of time series for different regions of Scotland and different indicators (e.g., test, case, hospitalized, and fatality) and different age groups. The ensemble data is aggregated based on eight age groups and contains 23 parameters (see Figure 5.1). Each age group exemplifies an age interval (e.g. *Group 1* $\rightarrow$ [age$\leq$20], *Group 2* $\rightarrow$ [20-29], ...,*Group 7* $\rightarrow$ [70$\leq$age], and *Group 8* $\rightarrow$ Healthcare Workers). The data contains the total numbers of susceptible, exposed, asymptomatic, symptomatic, hospitalized, recovered, deceased patients with a minimum, maximum, and mean values. Each age group is recorded on daily basis for 198 days. Each row in the data set represents a record of one day. See the Supplementary Material for a more detailed description of the ensemble data.

By investigating the ensemble data in our novel PCP software, we aim to assist users in exploring models such that users can interactively compare outcomes across age

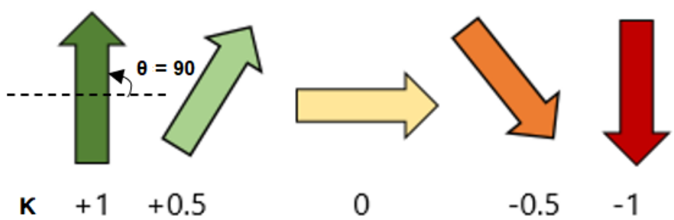


Figure 5.2: *The figure shows the glyphs that represent the correlation coefficient value between adjacent axis pairs displayed in the $\theta \in$ [-90, +90] range.*

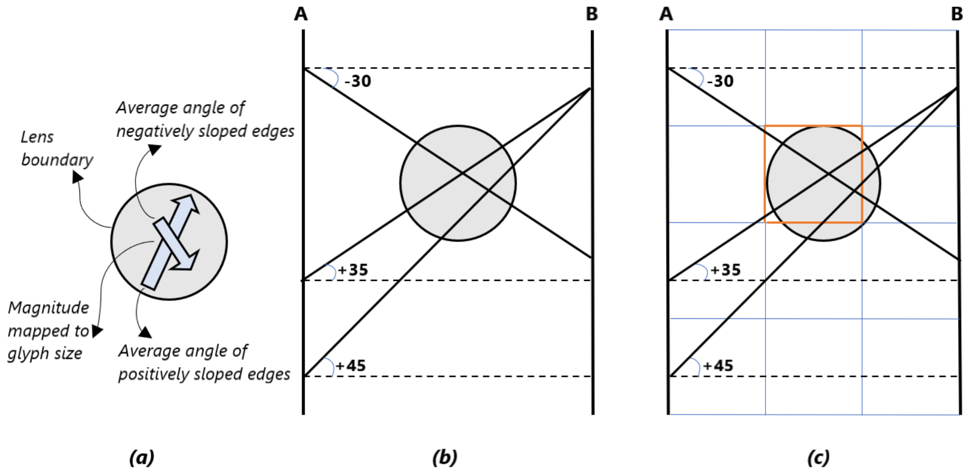


Figure 5.3: *An overview of (a) the glyph lens, (b) edge intersection summary with the dynamic edge glyph lens, and (c) edge-grid intersection with the grid-based edge glyph lens. This figure shows two attributes in the PCP and three line edges that connect A and B. After the detection of the intersecting edges for both, arrows are shown as in the lens (a) representing the edges. Since there are two positively sloped and one negatively sloped edges showing the relationship between A and B, the arrow representing the positive slope is longer than the other as it indicates two edges.*

groups, identify differences between simulation parameters, and observe patterns as well as reveal outliers and features in the data.

5.3.2 Axis Correlation Glyph

The correlation coefficient is beneficial to identify relationships between the two variates. For some PCP examples, overlapped edges may create clutter and users may have difficulty viewing patterns between axes. Results of a previous user-study on PCP understanding reveal that identifying correlation can be a barrier to PCP literacy [144]. Deriving the slope of the edges and interpreting the links between data variables by looking at the PCP image can be challenging. Therefore, we introduce arrow glyphs for each pair of axes to present correlation values explicitly (see Figure 5.1, (B)). The appropriate design of glyphs is critical for usability and successful visual communication. Relevant visual channels should be carefully selected and integrated for an effective glyph design [197]. The study by Fuchs *et al.* [66] methodically gathers and categorizes the literature on data glyphs, describing their designs, questions, data, and tasks. The arrow glyph is included in the "One-to-One Mapping" category. Borgo *et al.* [198] describe that glyph design can use a variety of visual channels, including shape, color, texture, size, and orientation. Our glyph design reveals the relationship between axes-pairs by presenting an arrow shape, using a peer-reviewed color library [139] and direction of the slope for the correlation value. In addition, the color was consistent with the polylines and the color scheme used in the PCP has also been adapted to the correlation glyphs based on κ.

Design Justification: The axis correlation glyph offers users a convenient way to interpret the relationship between two dimensions by glancing at the correlation glyphs. For dense PCPs, it may be difficult to determine relationships between data dimensions by observing the slope of the edges. We use an arrow glyph that conveys correlation value using slope information. The arrow glyph reveals the trend between dimensions using both the slope and direction. There are several other options possible here. Both bar charts and pie charts can encode the same information such as a number of intersecting edges and average slope. However, we wanted to map slope of edges to a glyph with slope intuitively built in. Arrow glyphs already have these characteristics naturally build in whereas other charts and glyphs generally do not.

The correlation values, κ, are calculated using Pearson's Correlation Coefficient [199] for each axis pair. The arrow glyph represents each pairwise coefficient value. The individual distributions of the two related axis pairs are shown in the range $\kappa \in$ [-1, +1] and the arrow glyphs represent the range $\theta \in$ [-90, +90] and correspond to the correlation values, κ, indicating negative and positive relationships respectively (see Figure 5.2).

5.3.3 Dynamic Edge Glyph Lens

The underlying structure in the data is not always obvious in PCPs. The dense PCP resulting from overlapping of the edges may cause information to be covered. This may make it difficult for the user to interpret the existing correlation and observe patterns. Thus, we introduce a glyph lens designed to reveal information that may be obscured by edge overplotting. Observing the dynamic glyph by hovering the lens over the edges offers the user a summary of the edges and of the average slope, θ_{AVG}, of the edges represented by arrows.

Design Justification: This is a special type of lens that focuses on the space between the axes as opposed to the axes themselves. Frequency-based approaches previously presented in the related work focus primarily on axes instead of relationships

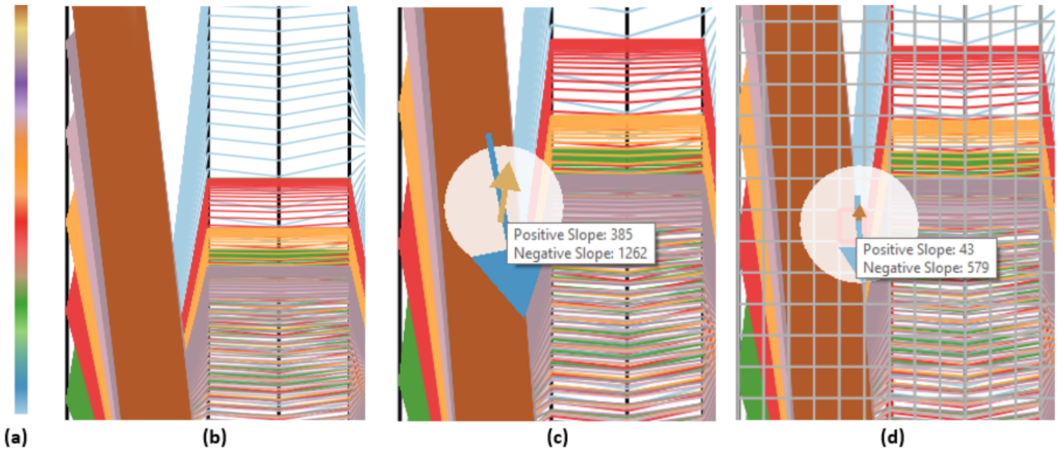


Figure 5.4: *An overview of (a) a color legend, (b) a dense area in the PCP, and (c) summary of edges in the same area with dynamic edge glyph lens, and (d) grid-based edge glyph lens (see section 5.3.3). The numbers indicate the number of edge intersections with the lens.*

between axes. Our dynamic edge glyph lens solution offers a user interaction-based feature integrated into the PCP to uncover the trends between axes and improve the interpretation of the data (see Figure 5.3). We chose the same arrow glyphs as in Figure 5.2 because they intuitively encode slope and thus the correlation between axes. Other charts and glyphs can encode this same information but not intuitively because the slope is not the predominant characteristic of most charts and glyphs, e.g., pie charts, bar charts, etc.

To address the overlap problem, we focused on the intersection of the edges with the lens, starting from the left axis and ending on the right axis (in any pair). The dynamic edge glyph shows the number of edges that intersect with the lens and average slope, θ_{AVG}, of each intersecting edge (see Figure 5.1, (C)). After calculating θ_{AVG}, the edges intersecting the lens are grouped according to whether the edge has a positive or negative slope. The two groups are represented by two arrows placed in the lens glyph (see Figure 5.3a). The upward arrow in the glyph lens represents the average positive, $\theta_{\text{AVG}}+$, and the other represents the average negatively sloped edges, θ_{AVG}-. The resulting arrows are designed similar to the correlation glyph arrow (Section 5.3.2). They display the angle, $\theta \in$ [-90, +90] by calculating the average angles of inclination $\theta_{\text{AVG}}+$, θ_{AVG}- (see Figure 5.3b). The magnitude of the arrows is also scaled by the number edges (with positive and negative slope) that intersect with the lens. The color of the arrows is mapped to the color legend provided. This interactive feature facilitates uncovering hidden correlation information between data axes by hovering the lens and observing the trends in the data.

See Figure 5.4. An accompanying detailed demonstration video [200] introduces our previously mentioned techniques.

Grid-Based Edge Glyph Lens: One limitation we encounter with a dynamic lens is run-time edge detection, which may slow down when there are too many edges. With very large data sets, the performance of detecting edge intersections starts to degrade. One way to address this challenge is to pre-compute a summary of edge intersections in a static grid and then display the meta-data, rather than trying to calculate the edge intersections at run-time. In order to pre-compute edge intersections, we divide the space between two neighboring axes into nxm squares, [3x30] square cells by default. As an example, we use a resolution of [66x30] cells in our display space for 22 axis pairs. To calculate the edge and grid cell intersection, we adopt the technique presented by Ericson [201].

For each grid cell, we pre-computed and store the edge intersections. Thus, we identify the number of positive and negative edges and their average slope, $\theta_{\text{AVG}}+$, θ_{AVG}-, in each cell. In the pre-computed grid view, the summary information for the edges that overlap with a given grid cell and containing the center of the lens is used (see Figures 5.3c and 5.4d).

5.3.4 Multivariate Subtraction Operator, Δ

Plotting two data sets on the same PCP or two adjacent PCPs is a common approach for comparison. However, both of these can lead to challenges with large data sets as both may be dense to start with. We introduce a multivariate subtraction operator, Δ, that we can apply to compare two similar data sets on the same PCP.

Design Justification: In our case, we have ensemble data from a Covid-19 simulation, thus, the simulation configurations are directly comparable. The Covid-19 simulation data is major inspiration for our features because the modelers are very interested in comparing different simulation configurations. The Δ operator reveals the differences between similar data sets e.g., the case of ensemble data. The variation between data attributes such as hospitalization or recovery numbers can be interpreted quickly. Plotting the difference S_Δ between two simulations, S_1 and S_2, in the same space as S_1 and S_2 themselves is simple, fast, and intuitive.

In order to perform the multivariate subtraction, the attributes of the data sets

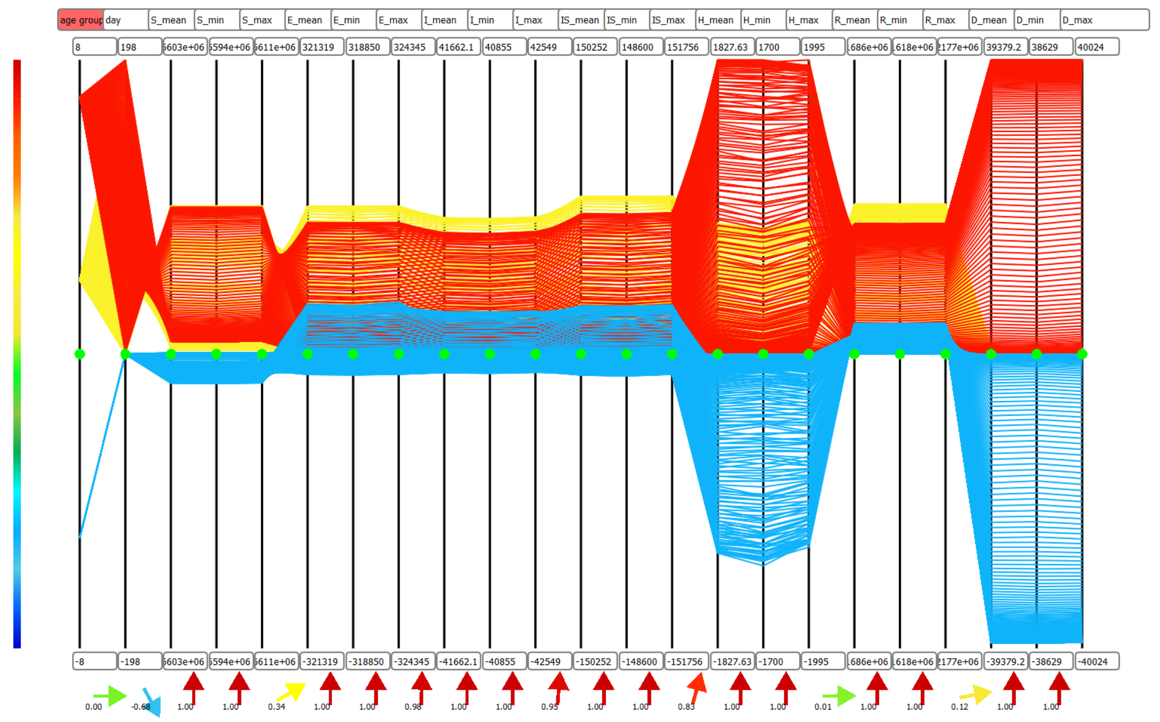


Figure 5.5: *Multivariate subtraction performed on the Group 1 ([d_{age} ≤20]) and Group 7 ([70≤ d_{age}]) in yellow and red respectively. The difference, Δ, is shown in the PCP with blue polylines. Using Δ, multivariate differences between age groups become obvious with respect to hospitalizations, h, and mortality, d. Green points on each axis address zero values on the axis.*

are the same and in the same order, such as the Covid-19 simulation [195] we use. The edges of the difference obtained after the subtraction can also be rendered and shown in the PCP (see Figure 5.5). As a result of plotting the difference data, S_Δ, labels for minimum, d(min), and maximum, d(max), values are updated.

The subtraction operator, Δ, is implemented to highlight changes in simulation output parameters for different input configurations that may or may not be similar. We perform subtraction on two configurations selected through the user interface (see Figure 4.6 (A)). The second selected, S_2, is subtracted from the first, S_1. This operation is applied by subtracting the corresponding values in the same dimensions. Given a simulation, S, with dimensions $S(d_0, d_1, ..., d_n)$ the subtraction operator computes the difference, Δx, between data values, x, that correspond to one another e.g.,

$$S_\Delta = S_1(d_n(x_m)) - S_2(d_n(x_m))$$

Where d_n is a given dimension and m is a given data index. With the selection of S_1 and S_2, the maximum value of the axis, d(max), is derived as the maximum value, d(min), of both $S_1(d_n)$ and $S_2(d_n)$, and the minimum value is set as $-1 \times d(max)$.

The S_Δ obtained as a result of subtraction is plotted on the PCP. Positive or negative differences can be seen within the updated d(min) and d(max).

Figure 5.5 displays the output of the subtraction operator, Δ, applied to *Group 1* ([$\leq$20]) and *Group 7* ([70$\leq$ d_{age}]) provided in the Covid-19 simulation [195]. The calculation is performed by subtracting *Group 7* from *Group 1* plotted with polylines in yellow and green respectively. We can see an example of this by looking at the age group dimension. By subtracting the values of d_{age}, the result is -6 *(1-7 = -6)*. The edges representing the difference between two data sets are plotted within age groups $\in$ [-8, +8], shown in red. Green points on each axis indicate zero values for each dimension and enable viewing the negative differences. The result is shown in Figure 5.5. The number of hospitalizations, h, and deaths, d, in patients over 70 years of age is much greater than in patients under 20 years of age.

5.3.5 Dimensionality Reduction by Collapsing Axis Pairs

The purpose of using parallel coordinates is to expose particular features in the multivariate data. However, the essential information sometimes may not be obvious due to overlapping edges and a high number of dimensions plotted in the PCP. The images vary depending on the order of axes. In order to display the relationship between dimensions, we use glyphs showing the tendency between each axis pair and the corresponding correlation, κ, (see Section 5.3.2). By using on these correlation glyphs, the user may exploit relationship-guided dimensionality reduction via collapsing of axis pairs.

Design Justification: The high-dimensional ensemble data is based on eight age groups and contains 23 parameters with minimum, maximum and mean values of each indicator. The data includes repetitive information. We introduce this user option that gives a different perspective on the data dimensions by removing some of the redundant elements that do not add new information to the PCP. The objective of collapsed axis pairs is to decrease the number of dimensions and depict a less complex PCP view e.g., especially for values of $\kappa = 1$. This feature enables the user to explore and display the relationship between dimensions, d, that they choose to emphasize and with less redundant information (see Figure 4.6 (D)).

The user option provides a new view of the data dimensions by reducing some of the

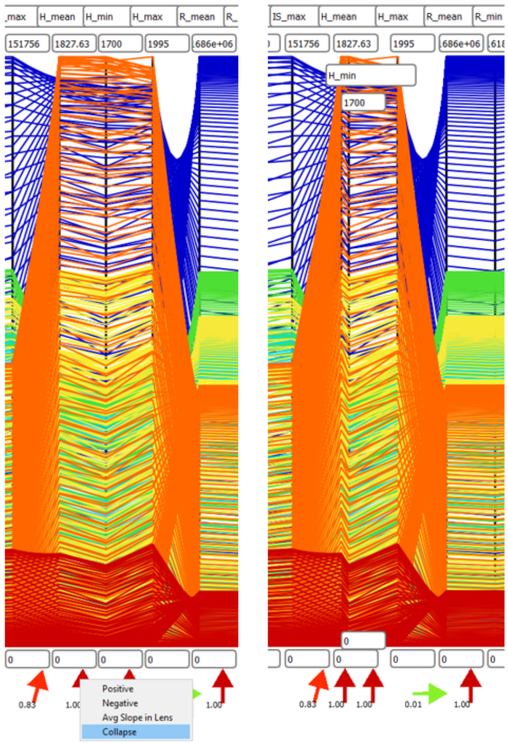


Figure 5.6: *The collapsing of the h_{mean} and h_{min} axis pair by right-clicking on the correlation glyph showing the relationship between them. The labels of h_{min} are stacked to indicate the collapsing process.*

redundant dimensions that do not present a particularly notable pattern in the PCP. Collapsing axes can be guided by observing correlation glyphs. For example, pairwise axes with a correlation κ of unity may be collapsed without loss of information. The process is performed by right-clicking on a correlation glyph for a given axis pair and reducing the space between them by translating the right axis closer to the left axis. In the new layout, the axis name and maximum value labels of the right axis are stacked under the left labels of the pair axis while the minimum label is placed on the top of other minimum value labels. The collapsing procedure can be undone by right-clicking on the same correlation glyph to obtain the previous PCP view. Figure 5.6 demonstrates an example of axis pair collapsing between $\mathrm{h}_{\mathrm{mean}}$ and $\mathrm{h}_{\mathrm{min}}$ (h:*Hospitalization*). Selected collapsed axis groups are data variables with, $\kappa = +1$, in other words, showing a direct relationship. As a result of the collapsing of an axis pair, two dimensions are positioned side-by-side and axis labels stacked on top of each other are displayed. Figure 5.1 (D) shows an example where 3 dimensions are juxtaposed after collapsing two-axis pairs. With the dimensionality reduction feature, redundant and repetitive information that makes it more challenging to reveal patterns in the data can be excluded.

Additional Features: In addition to the previous features we introduced, the software includes features that are helpful in exploring the ensemble simulation data. We provide a feature that allows the min and max labels to be updated such that the axis data in a given range can be scaled. We enable the user to change the given default static grid resolution by entering a new value for the x and y axis of the PCP through the user interface. We offer six different color scales for color mapping in the PCP using a color library by Roberts *et al.* [139]. We also introduce the features of drawing the average polyline using the average of the edges, or rendering the positive and negative sloped edges by right-clicking on any area of the PCP, using focus+context. Finally, we developed a κ matrix to understand the relationship between each data dimension combination. In the matrix, the user can select one of the dimensions and sort the correlation values from smallest to largest.

5.4 Evaluation

We provide three use cases to evaluate our techniques and provide a demo video for these use cases. We demonstrated the software to the domain expert and reported feedback collect from the expert in this section. Our primary goals are to evaluate the usefulness of our approaches in comparing multidimensional data and determining the most effective parameters, and to make sure that they enable users to make some new observations.

5.4.1 Case-Study

In this section, three use cases demonstrate the effectiveness of our techniques in understanding underlying trends in the Covid-19 ensemble data.

Use Case 1: Multivariate Comparison of Age Groups To explore the multivariate differences between age groups, we used the Δ operator between two age groups in the first simulation configuration presented in Figure 5.1. For example, we render the relationship between the simulation results under age 20 (Group 1) and above age 70 (Group 7) (see Figure 5.5) by applying the Δ operator to these age groups. We observe that the hospitalization and mortality numbers are much higher compared to Group 1.

Use Case 2: Comparing Input Parameter Values, $\mathbf{p_{inf}}$ Probability of infection, p_{inf}, is one of the most interesting input parameters of the simulation according to the simulation domain experts. We selected the two simulations with the minimum and maximum, p_{inf} (min) and p_{inf} (max), for input parameter values. Then we utilized the Δ operator to compare the outcomes for these two simulations to investigate how influential the p_{inf} parameter is and understand how input parameter values influence the output. To compare two simulations, we sorted simulations by the p_{inf} value and included all age groups in Simulation 3 with the lowest p_{inf} value and Simulation 101 with the highest p_{inf}. We then used Δ operator to render the difference between these simulations. As a result of Δ, Simulation 101 shows a very clear difference for all output parameters compared to Simulation 3 (see Figure 5.7). The Δ operator indicates that p_{inf} is a very influential input parameter.

Use Case 3: κ-guided Dimensionality Reduction We examine the PCP in Figure 5.1 and the correlation glyphs under each axis pair. We observe that there is always a direct relationship between the mean, min and max values of each parameter in the output. We used this observation to reduce the redundant dimensions and produce a new image with the redundant axes removed. The dimensionality reduction technique we utilize by collapsing axis pairs results in an image that reduces the number of dimensions by almost 50% in the PCP (see Figure 5.8). Note that the pairwise glyphs are also preserved and remind the user of the redundancy.

5.4.2 Domain Expert Feedback

This work is partially carried out in collaboration with RAMP VIS [1] team, who support the modelling scientists and epidemiologists in the Scottish COVID-19 Response Consortium (SCRC) [195] (see Subsection 5.3.1). We had three meeting sessions, including visualization experts, modellers, and statisticians. The brainstorming sessions facilitated understanding of the data simulations and exploring the most influential input parameters. We organized a feedback session and interviewed Dr Ben Swallow, with a PhD in Statistics and working in the School of Mathematics & Statistics, University of Glasgow. He has been working in statistical simulation and estimation for seven years and has spent approximately four years on epidemiological studies. Some of his work focuses on Bayesian parameter inference and model selection and methods for zero-inflated data. Our interview questions

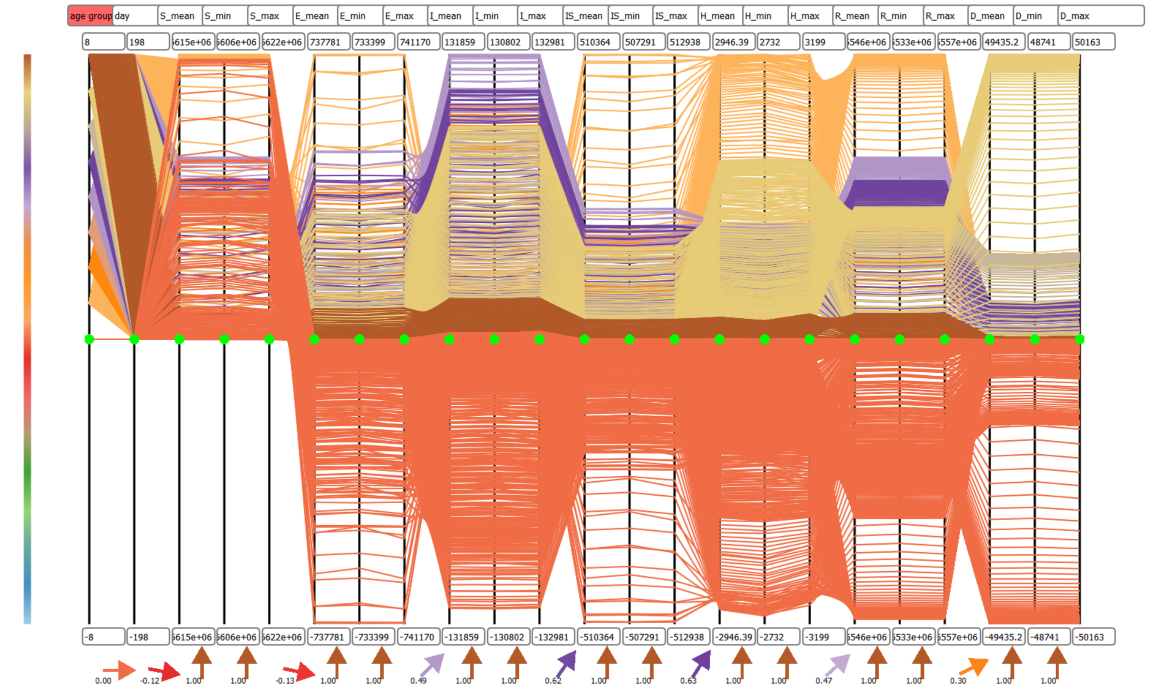


Figure 5.7: *This figure displays the subtraction operator applied to Simulation 3 with lowest p_{inf} and Simulation 101 with highest p_{inf}.*

were adopted from Hogan *et al.* [202].

Correlation Glyph: We demonstrated the correlation glyphs, and he reported: *"It's a really good way of guiding the dimension reduction when you have so much information. Users are trying to find a way of deciding how to reduce it down and extract information. It's pretty cool."*

Dynamic and Grid-based Edge Glyph Lenses: When we presented the both glyph lenses to watch the behavior of the glyphs and discover areas with a lot of variation, he stated that the feature is useful and added; *"I think it's just another way of looking at the kind of sensitivity to that particular parameter and in what direction it's going. I particularly know the type of people that would likely use this. I think you can get this through more hardcore mathematical sensitivity analysis, but I think getting an idea of a sensitivity across regions of parameters and different parameters will be very welcome. It would be huge benefit of having this type of software. Yes, I really like that."*

Dimensionality Reduction: We mentioned that there are a lot of redundant dimensions in the data and to the expert. He agreed on this and reported: *"Yes, that's what we found from the mathematical analysis as well. It was p_{inf} and P_s that we really the only two parameters that had any impact at all. It seems that that's*

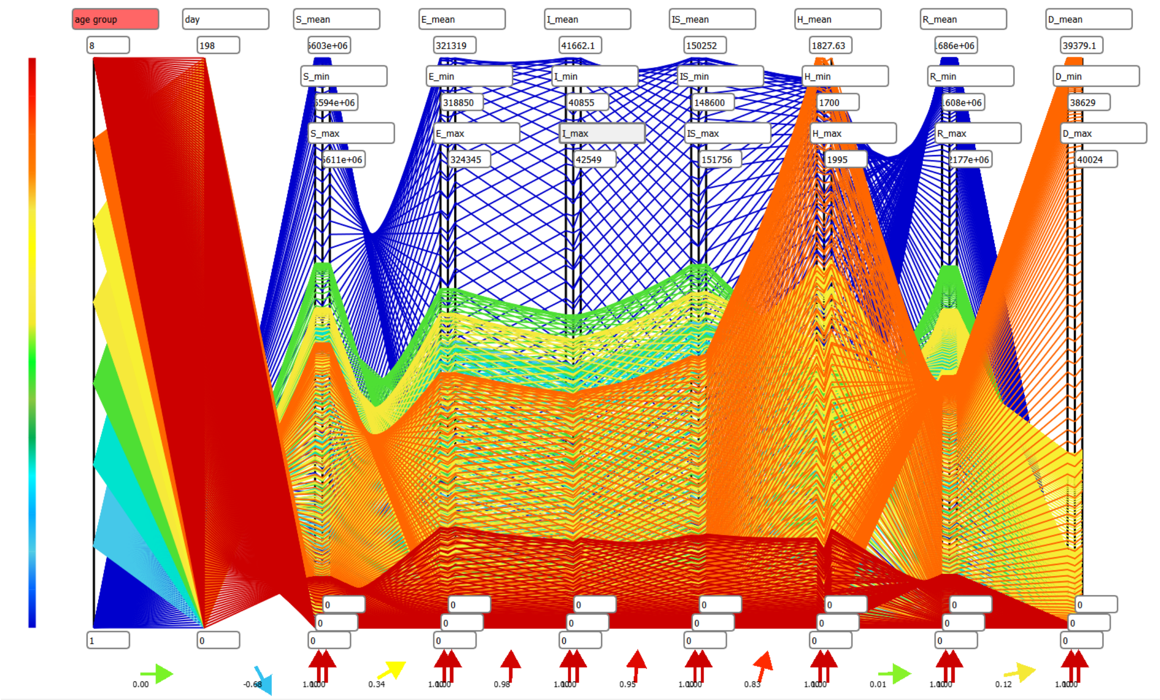


Figure 5.8: *This figure demonstrates dimensionality reduction applied on axis pairs with κ = 1.*

what's being visualized here and in a much more clear manner."

Use Case 1: Comparing Age Groups When we first demonstrated the subtraction operator, Δ, to the expert, he liked the concept of presenting the difference between two multidimensional data sets visually to compare them and stated: *"I think it's highlighting differences. The differences are going to be specific to particular groups or compartments of the model. So I like being able to observe that. From a policy point of view, you think "if I change this parameter, what's it going to change?" If it has a negative impact on say younger people, in terms of the number of cases, but maybe it reduces the deaths in another age category, then that's going to be useful from a policy perspective rather than just saying, "well, we've just looked at the combined groups". There's going to be more cases in group two. You know group two is going to be less impacted by Covid-19 in general. And knowing how it's affecting things in a more detailed and visual way, I think, is really useful."*

Use Case 2: Comparing Input Parameter Values, p_{inf} The p_{inf} input parameter has a significant influence on the outcome, and the difference between simulations verifies that. We demonstrated this in Figure 5.7 and asked the expert if he finds this helpful. He commented: *"I would like confirming what we have done already [formal mathematical sensitivity analysis], or if we used the software first and looking at what we think might be the most important parameters. You know most of the* *model developers will have an idea of which is the most important parameters are. Visualizing that is very useful for confirmation."*

We also asked the expert how he figured out the most influential parameters without the software and how long it takes. He reported: *"We normally have to do a full mathematical sensitivity analysis of the model. You could look at things like histograms of the output, so they would tend to be either visualized viewpoints, but probably nowhere near as sophisticated as this. Or kind of formal mathematical way you look at things like the derivative, i.e., changes in the output as a function of the different parameters. But that's a lot more complex and time-consuming than this. The process really depends on the complexity of the model and number simulations you have to do, but it would take probably at least a couple of hours to run mathematical analysis. Because you generally use a Monte Carlo approach across lots of simulations as you are plotting here. But again, there are lots of different questions that you could ask using this the PCP software in terms of the sensitivity across time, different age groups, and different classes. You would have to do it on the separate simulation or sensitivity analysis for each of those different configurations, whereas here you have the option of interrogating them all in one go or very quickly switching between the different questions that you might want to ask of them to the model."*

Use Case 3: κ-guided Dimensionality Reduction Dimensionality reduction and axis ordering are still considered unsolved problems. We demonstrated our κ-guided dimensionality reduction features by collapsing axis pairs (see Figure 5.8). We asked the expert if the feature let him see anything that he might previously have not been able to see or make some new observations or hypotheses. He reported: *"One of the common aspects of these types of models is over parameterization. When you try and estimate the parameters, if the model is not sensitive at all to the input parameters, then no matter how much you try and make any inference, it is not going to be useful at all. So from that perspective, I think this feature is useful. The standard approach to deal with overparameterization is that if you have got parameter redundancies to make some model reduction - that's quite complex to do without a good understanding of the model and where it is lack of sensitivity arises. So I think it would be really helpful in deciding how to think about either combining outputs into* *a single one. For example, if there was age differentiation or there was no impact on the parameters on age, then I think you would see that here and in terms of looking at the different compartments, I think that is really useful. Parameter redundancy is generally quite a useful way of guiding model reduction and that would be very helpful there."*

We asked the expert if the feature might increase confidence in terms of the correctness or accuracy of the simulations. He stated: *"Yes, I'm sure. If you are seeing some of the maximum numbers, if you knew, for example, that hospitalizations never got above a particular point but your model is consistently estimating numbers of hospitalizations to be in the hundreds of thousands, and you know that's not realistic, you would probably have some lack of confidence in that model. I think that could be something else that this helps with. In terms of focusing where you perhaps want to do data collection as well, if you know there's a lot of sensitivity. It seems like hospitalization in this model are a very sensitive, very valuable output. Then you might try and focus your data collection on that when you want to make inferences and try and estimate these parameters. That would probably be a good way of guiding that decision as well."*

5.5 Limitations of the Tool and Future Improvements

We introduced techniques for investigating and understanding dense PCPs, including sorting the correlation values of axis pairs in ascending order and updating the PCP view accordingly. However, with the traditional PCP axis ordering, plotting axis pairs according to pairwise order is not feasible. Another limitation we encounter with a dynamic lens is run-time edge detection, which may slow down when there are too many edges. In addition, with large data sets, the performance of detecting edge intersections starts to degrade. A future improvement could be exploring techniques to speed up the process. Scalability is one of the limitations, i.e., how to arrange axis labels when ten or more pairs of neighboring axes are collapsed. The current version of the software does not allow viewing multiple lenses, adjusting lens size, or additional operators, such as addition, multiplication, and division of simulation data sets. These limitations can be promising for future endeavors.

5.6 Chapter Summary

This chapter introduces interactive glyph lenses, which enable users to investigate an overplotted region with a dynamic lens that hovers over the PCP based on mouse location. We provide an arrow glyph beneath each neighboring axis pair to highlight the relationship between dimensions. In addition, we present a dimension reduction approach that allows users to simplify a PCP based on the κ correlation value between neighboring axes and collapsing axis pairings that do not contribute information to the display. We also give a user option, Δ, which displays the difference between two multidimensional data sets for comparison. Finally, in collaboration with a modelling specialist, we assess our approaches using a case study based on a simulation of Covid-19.

Chapter 6

Conclusion

Contents

6.1 Main Contributions

This thesis studies methods for enhancing beginners' visualization literacy skills as well as barriers to visualization literacy. This section includes a discussion of how each chapter contributes to the objectives, followed by more general conclusions, limitations, and suggestions for further research.

Chapter 2: Interactive Visualization Literacy: The State-of-the-Art This chapter contributes a literature review of visualization literacy papers published from 2001-to the present. We provide a novel classification of literacy research that enables readers to explore published literature. This classification emphasizes the evaluation method chosen to test individuals' visualization literacy skills, presents guidelines for improving literacy skills, and indicates factors that affect individuals' understanding of various visual designs. This STAR also offers overview tables guided by the evaluation method-based classification. The tables present meta-data that facilitate literature comparisons including visual designs, the number of participants involved in the study, target groups (e.g., age), chosen study platforms more. The survey offers valuable information identifying experimental settings required to assess individuals in uncovering problems in the area as well as having a more complete understanding of advancing visualization literacy skills. Moreover, we share an overview of future work from the literature that enables readers to identify areas of open research subjects in this scope. We believe our survey is beneficial for both new or experienced researchers interested in visualization literacy.

Chapter 3: Treemap Literacy: A Classroom-based Investigation In this chapter, we present a study that investigates possible barriers to the interpretation and comprehension of treemaps. A novel treemap literacy test is introduced that includes a variety of treemap designs and treemap questions with a test question classification based on treemap features. This chapter offers researchers a better understanding of barriers to a complete comprehension of a treemap and a method to advance treemap literacy.

Moreover, we develop an interactive pedagogical treemap application for training and cognition of a treemap design that supports the exploration of a hierarchical data structure. The educational treemap software transforms a passive study into active practice in classrooms and can be used as a replacement for traditional treemap teaching approaches. Results of the user study indicate that the students who interacted with the software outperformed students who only learned through slides before taking a treemap literacy test. Furthermore, participants' feedback signifies that the pedagogical treemap software offers an effective learning experience through easier and quicker understanding of treemap properties.

Chapter 4: P-Lite: A Study of Parallel Coordinate Plot Literacy In this chapter, we investigate barriers to PCP literacy based on the research literature and our teaching experience in data visualization classes. We provide insight into the barriers to a complete understanding of the concept of the PCP. A novel literacy test is presented to evaluate non-expert users' literacy skills with a range of PCP images using a collection of tools and datasets. We also collated the most frequently used visualization tools and multivariate datasets used to generate PCPs for our literacy test.

Furthermore, we developed an educational PCP tool to facilitate the interpretation and exploration of multivariate data as well as enabling users to learn how to create and interpret PCPs interactively. The software features correspondence between CCP and PCPs views and connects the points in Cartesian space with polylines in parallel coordinates space. Results of the user study reveal that participants taught with the pedagogical application tutorial video performed better on the PCP literacy test. Participants indicate that the illustration of the relation between the two views offers a simple way of improving PCP literacy as well as understanding data with more than two dimensions.

Classroom vs. Crowdsourcing We conducted two studies to evaluate educational Treemap and PCP tools. Both approaches have pros and cons. The crowdsourcing experiment could recruit people from a large pool with diverse backgrounds and collect data in a shorter time. However, there is no guarantee of collecting data from attentive participants. For a classroom experiment, it is feasible to recruit people with similar backgrounds and conduct an investigation in uniform conditions, even though setting up the experiment takes time. There is also more control over the experiment to collect credible results.

Chapter 5: DPCP Vis: Techniques for Dense Parallel Coordinate Plots This chapter presents interactive glyph lenses, which enable users to explore an over-plotted PCP with a dynamic lens that hovers over the PCP based on mouse position. This interaction summarizes the edges that overlap with the lens, represented by arrow glyphs that show the average slope, θ_{AVG}, of a dense collection of edges. We display an arrow glyph below each adjacent axis pair that indicates the correlation between dimensions. We present a dimension reduction technique that allows users to simplify a PCP based on the correlation value, κ, between adjacent axes and collapsing axis pairs that do not add information to the display. We also provide a user option we call a subtraction operator, Δ, which displays the difference between two multidimensional data sets for comparison. We evaluate our techniques with a case study based on a simulation of Covid-19 in collaboration with a modeling expert.

6.2 Future Work

Each idea discussed in this thesis has the possibility for growth. Potential directions for further improvement are presented below.

Interactive Visualization Literacy: The State-of-the-Art

With regards to our literature review, we have two unique directions as suggestions for further improvement in the visualization literacy field in addition to the most frequent future work directions presented in our literature review in Chapter 2.

Visibility: We note that visualization literacy is not a very visible sub-field yet. Even though data visualization is growing in prominence, the significance of visualization literacy does not yet stand out in research communities. The amount of literature we presented in the survey also supports this idea. Gaining visibility and momentum is necessary in order to improve literacy skills which enable effective use of visualization in various research areas.

Standards: Some basic subjects have a standard assessment test e.g. mathematics, languages, and analytic reasoning. Although some studies [22, 37] have taken the first steps in this direction by providing visualization literacy tests, we suggest developing a series of a standardized assessment tests for visualization literacy that can vary according to the complexity of visual designs and data sets for students with different backgrounds.

Treemap Literacy: A Classroom-based Investigation

We introduced the treemap literacy test that enables evaluation of users treemap literacy skills in Chapter 3. For an even more reliable test for further research, improving the literacy test with a wider variety of data and treemap visualization designs is recommended. More studies are recommended with a more diverse group and more participants to reinforce the efficacy of the educational treemap tool. Further research with participants from non-computer science backgrounds for investigating the influence of users' familiarity with treemaps on the study result would be interesting. Also, analyzing the experimental results from the varied background of the participants can be a helpful next step to understand treemap visualization literacy skills.

Improvements to the pedagogical software have been identified for the treemap view, nesting the top level of the data hierarchy and providing labels for each rectangle directly instead of requiring mouse-over interaction are potential further attributes that might be possible. In addition, enabling users to display large datasets with a different layout algorithms and a greater number of data hierarchies, interacting with the treemap for additional nesting exploration, and keyboard control are future endeavors.

P-Lite: A Study of Parallel Coordinate Plot Literacy

Distance learning, a growing trend, offers a solution to access lecture material online as well as a number of facing disadvantages such as difficulty staying engaged or receiving immediate feedback. Our crowdsourcing user study in Chapter 4 features similarities to distance learning settings in which a video tutorial is provided to instruct the PCP design. Our findings reveal only 35% of the participants (60 out of 170) spent the time necessary (or close) to watch the entire tutorial. This indicates that keeping students engaged while watching a tutorial video online is a non-trivial challenge. We were not anticipating this barrier. For future research, we can conduct a similar experiment in a classroom setting and offer participants interaction with the educational PCP software instead of a video tutorial. Thus, we can examine the results to investigate improvement between crowdsourcing and classroom-based user-study settings.

Moreover, further studies are recommended to reinforce the effectiveness of the educational tool where interaction is necessary. Also, we intend to teach PCP design to novices by presenting the correspondence between a CCP and PCP, including advanced PCP features (e.g. axis flipping, axes reordering, removing axes) in the pedagogical PCP tool. Improving the PCP literacy test with a wider variety of datasets and asking users to provide their observations given a PCP image using a think-aloud protocol would be interesting for further study.

DPCP Vis: Techniques for Dense Parallel Coordinate Plots

In Chapter 5, we presented techniques for exploring and understanding dense PCPs that could involve sorting axis pairs correlation value in ascending order and updating the PCP view accordingly. However, plotting axis pairs according to pairwise, κ, is not feasible with the traditional PCP axis ordering. Therefore, introducing a new axis sorting approach to convey the axes' relationships can be a future endeavor. Another limitation is scalability, i.e., how to arrange axis labels when 10 or more pairs of neighboring axes are collapsed. Future directions can involve, including multiple lenses, adjustable lens size, and additional operators, such as addition, multiplication and division of simulation data sets.

Bibliography

[1] "RAMP VIS: Visualization and visual analytics in support of rapid assistance in modelling the pandemic (ramp)." 2020. [Online]. Available: https://sites.google.com/view/rampvis

[2] S. Few and P. Edge, "Data visualization: past, present, and future," *IBM Cognos Innovation Center*, 2007.

[3] S. T. Kard, J. D. Mackinlay, and B. Scheiderman, *Readings in Information Visualization, using vision to think*. San Francisco: Morgan Kaufmann, 1999.

[4] L. McNabb and R. S. Laramee, "Survey of surveys (SoS)-mapping the landscape of survey papers in information visualization," in *Computer Graphics Forum*, vol. 36, no. 3. Wiley Online Library, 2017, pp. 589–617. [Online]. Available: https://doi.org/10.1111/cgf.13212

[5] K. Börner, A. Maltese, R. N. Balliet, and J. Heimlich, "Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors," *Information Visualization*, vol. 15, no. 3, pp. 198–213, 2016. [Online]. Available: https://doi.org/10.1177/1473871615594652

[6] S. Lallé, D. Toker, and C. Conati, "Gaze-driven adaptive interventions for magazine-style narrative visualizations," *IEEE Transactions on Visualization and Computer Graphics*, 2019.

[7] E. Huynh, A. Nyhout, P. Ganea, and F. Chevalier, "Designing narrative-focused role-playing games for visualization literacy in young children," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 924–934, 2020.

[8] B. Alper, N. H. Riche, F. Chevalier, J. Boy, and M. Sezgin, "Visualization literacy at elementary school," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 5485–5497. [Online]. Available: https://doi.org/10.1145/3025453.3025877

[9] J. Fuchs, P. Isenberg, A. Bezerianos, M. Miller, and D. Keim, "Educlust-a visualization application for teaching clustering algorithms." Eurographics–40th Annual Conference of the European Association for Computer Graphics, 2019, pp. 1–8. [Online]. Available: http://dx.doi.org/10.2312/eged.20191023

[10] J. Gäbler, C. Winkler, N. Lengyel, W. Aigner, C. Stoiber, G. Wallner, and S. Kriglstein, "Diagram safari: a visualization literacy game for young children," in *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*, 2019, pp. 389–396.

[11] F. Bishop, J. Zagermann, U. Pfeil, G. Sanderson, H. Reiterer, and U. Hinrichs, "Construct-a-vis: exploring the free-form visualization processes of children,"

IEEE Transactions on Visualization and Computer Graphics, vol. 26, no. 1, pp. 451–460, 2019.

[12] A. Krekhov, M. Michalski, and J. Krüger, "Integrating Visualization Literacy into Computer Graphics Education Using the Example of Dear Data." The Eurographics Association, 2019.

[13] Z. Wang, L. Sundin, D. Murray-Rust, and B. Bach, "Cheat sheets for data visualization techniques," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.

[14] A. M. B. Rodrigues, G. D. J. Barbosa, H. C. V. Lopes, and S. D. J. Barbosa, "What questions reveal about novices' attempts to make sense of data visualizations: patterns and misconceptions," *Computers & Graphics*, 2020.

[15] P. Ruchikachorn and K. Mueller, "Learning visualizations by analogy: Promoting visual literacy through visualization morphing," *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 9, pp. 1028–1044, 2015. [Online]. Available: https://doi.org/10.1109/TVCG.2015.2413786

[16] B. C. Kwon and B. Lee, "A comparative evaluation on online learning approaches using parallel coordinate visualization," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 993–997. [Online]. Available: http://dx.doi.org/10.1145/2858036.2858101

[17] S. Lee, B. C. Kwon, J. Yang, B. C. Lee, and S.-H. Kim, "The correlation between users' cognitive characteristics and visualization literacy," *Applied Sciences*, vol. 9, no. 3, p. 488, 2019. [Online]. Available: http://dx.doi.org/10.3390/app9030488

[18] K. J. Schönborn and T. R. Anderson, "The importance of visual literacy in the education of biochemists," *Biochemistry and molecular biology education*, vol. 34, no. 2, pp. 94–102, 2006.

[19] A. Zoss, A. Maltese, S. M. Uzzo, and K. Börner, "Network visualization literacy: Novel approaches to measurement and instruction," in *Network Science In Education*. Springer, 2018, pp. 169–187.

[20] K. Börner, A. Bueckle, and M. Ginda, "Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments," *Proceedings of the National Academy of Sciences*, vol. 116, no. 6, pp. 1857–1864, 2019. [Online]. Available: https://doi.org/10.1073/pnas.1807180116

[21] C. Stoiber, F. Grassinger, M. Pohl, H. Stitz, M. Streit, and W. Aigner, "Visualization onboarding: Learning how to read and use visualizations," *IEEE Workshop on Visualization for Communication*, 2019.

[22] S. Lee, S.-H. Kim, and B. C. Kwon, "Vlat: Development of a visualization literacy assessment test," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 551–560, 2017. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2016.2598920

[23] "Xmdv Tool," http://davis.wpi.edu/xmdv/, Last accessed: 2020.

[24] "Parallel Coordinates Literacy Test," https://swanseasom.au1.qualtrics.com/jfe/form/SV_2gxKduwQjtAQIuN, Last accessed: 2020.

[25] "A Tool to Support Parallel Coordinates Literacy Video," https://vimeo.com/456884883, Last accessed: 2020.

[26] M. C. Velez, D. Silver, and M. Tremaine, "Understanding visualization through spatial ability differences," pp. 511–518, 2005.

[27] M. Card, *Readings in information visualization: using vision to think*. Morgan Kaufmann, 1999.

[28] R. Krum, *Cool infographics: Effective communication with data visualization and design*. John Wiley & Sons, 2013.

[29] M. Aparicio and C. J. Costa, "Data visualization," *Communication design quarterly review*, vol. 3, no. 1, pp. 7–11, 2015.

[30] W. Playfair, *Playfair's commercial and political atlas and statistical breviary*. Cambridge University Press, 2005.

[31] D. Rees and R. Laramee, "A survey of information visualization books," in *Computer Graphics Forum*. Wiley Online Library, 2018, pp. 610–646.

[32] N. Gershon and S. Eick, ""Foreword", Proc. IEEE Symp." *Information Visualization, IEEE CS Press*, 1995.

[33] N. Bikakis, "Big data visualization tools," *Encyclopedia of Big Data Technologies, Springer*, 2018.

[34] A. Wolff, D. Gooch, J. J. C. Montaner, U. Rashid, and G. Kortuem, "Creating an understanding of data literacy for a data-driven society," *The Journal of Community Informatics*, vol. 12, no. 3, 2016.

[35] P. Vahey, L. Yarnall, C. Patton, D. Zalles, and K. Swan, "Mathematizing middle school: Results from a cross-disciplinary study of data literacy," in *Annual Meeting of the American Educational Research Association, San Francisco, CA*, 2006.

[36] R. E. Wileman, *Visual communicating*. Educational Technology, 1993.

[37] J. Boy, R. A. Rensink, E. Bertini, and J.-D. Fekete, "A principled way of assessing visualization literacy," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1963–1972, 2014. [Online]. Available: https://doi.org/10.1109/TVCG.2014.2346984

[38] M. D. Reckase, *Multidimensional item response theory models*. Springer, 2009.

[39] A. V. Maltese, J. A. Harsh, and D. Svetina, "Data visualization literacy: Investigating data interpretation along the novice–expert continuum," *Journal of College Science Teaching*, vol. 45, no. 1, pp. 84–90, 2015. [Online]. Available: https://doi.org/10.2505/4/jcst15_045_01_84

[40] F. Chevalier, N. H. Riche, B. Alper, C. Plaisant, J. Boy, and N. Elmqvist, "Observations and reflections on visualization literacy in elementary school," *IEEE computer graphics and applications*, vol. 38, no. 3, pp. 21–29, 2018. [Online]. Available: http://dx.doi.org/10.1109/MCG.2018.032421650

[41] E. E. Firat, A. Joshi, and R. S. Laramee, "Interactive visualization literacy: The state-of-the-art," *Technical Report, School of Computer Science, University of Nottingham*, 2021.

[42] E. E. Firat, A. Denisova, and R. S. Laramee, "Treemap Literacy: A Classroom-Based Investigation," in *Eurographics 2020 - Education Papers*, 2020, pp. 29–38. [Online]. Available: http://dx.doi.org/10.2312/eged.20201032

[43] E. E. Firat and R. S. Laramee, "Inclusivity for visualization education: a brief history, investigation, and guidelines," *Diálogo com a Economia Criativa*, vol. 4, no. 12, pp. 146–160, 2019.

[44] E. E. Fırat and R. S. Laramee, "Towards a survey of interactive visualization for education," *EG UK Computer Graphics & Visual Computing*, pp. 91–101, 2018. [Online]. Available: https://doi.org/10.2312/cgvc.20181211

[45] A. Diehl, E. E. Firat, T. Torsney-Weir, A. Abdul-Rahman, B. Bach, R. S. Laramee, R. Pajarola, and M. Chen, "Visguided: A community-driven approach for education in visualization," in *Eurographics 2021 (EG 2021)*, 2021.

[46] X. Liu, M. Alharbi, J. Best, J. Chen, A. Diehl, E. E. Firat, D. Rees, Q. Wang, and R. S. Laramee, "Visualization resources: A starting point," in *2021 25th International Conference Information Visualisation (IV)*. IEEE, 2021, pp. 160–169.

[47] A. Joshi, K. Börner, R. S. Laramee, L. Harrison, E. E. Firat, and B. C. Kwon, "Visualization literacy for general audiences-can we make a difference?" *IEEE Vis*, 2021.

[48] E. E. Firat, B. Swollow, and R. S. Laramee, "Techniques for dense parallel coordinate plots," *Technical Report, School of Computer Science, University of Nottingham*, 2021.

[49] K. Börner, "Data visualization literacy," *Keynote Talk, IEEE Transactions on Visualization and Computer Graphics*, 2019.

[50] "Merriam-Webster Dictionary," https://www.merriam-webster.com/, 2019, last Accessed: May 2019.

[51] V. J. Bristor and S. V. Drake, "Linking the language arts and content areas through visual technology," *THE Journal (Technological Horizons In Education)*, vol. 22, no. 2, p. 74, 1994.

[52] J. Ametller and R. Pintó, "Students' reading of innovative images of energy at secondary school level," *International Journal of Science Education*, vol. 24, no. 3, pp. 285–312, 2002.

[53] R. A. Bradent and J. A. Hortinf, "Identifying the theoretical foundations of visual literacy," *Journal of Visual Verbal Languaging*, vol. 2, no. 2, pp. 37–42, 1982.

[54] F. Beck, S. Koch, and D. Weiskopf, "Visual analysis and dissemination of scientific literature collections with survis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 180–189, 2015.

[55] D. Simkin and R. Hastie, "An information-processing analysis of graph perception," *Journal of the American Statistical Association*, vol. 82, no. 398, pp. 454–465, 1987.

[56] S. Pinker, "A theory of graph comprehension," *Artificial intelligence and the future of testing*, pp. 73–126, 1990.

[57] P. Shah and J. Hoeffner, "Review of graph comprehension research: Implications for instruction," *Educational Psychology Review*, vol. 14, no. 1, pp. 47–69, 2002. [Online]. Available: http://dx.doi.org/10.1023/A:1013180410169

[58] D. Moore-Russo, J. M. Viglietti, M. M. Chiu, and S. M. Bateman, "Teachers' spatial literacy as visualization, reasoning, and communication," *Teaching and Teacher Education*, vol. 29, pp. 97–109, 2013.

[59] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," in *Proceedings of international conference on intelligence analysis*, vol. 5. McLean, VA, USA, 2005, pp. 2–4.

[60] D. M. Russell, R. Jeffries, and L. Irani, "Sensemaking for the rest of us," in *CHI workshop on Sensemaking*, 2008, pp. 1226–1230.

[61] S. Lee, S.-H. Kim, Y.-H. Hung, H. Lam, Y.-a. Kang, and J. S. Yi, "How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 499–508, 2015.

[62] M. Schwab, H. Strobelt, J. Tompkin, C. Fredericks, C. Huff, D. Higgins, A. Strezhnev, M. Komisarchik, G. King, and H. Pfister, "booc. io: An education system with hierarchical concept maps and dynamic non-linear learning plans," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 571–580, 2017.

[63] C. T. Silva, E. Anderson, E. Santos, and J. Freire, "Using vistrails and provenance for teaching scientific visualization," in *Computer Graphics Forum*, vol. 30, 1, 2011, pp. 75–84.

[64] M. Contero, F. Naya, P. Company, J. L. Saorin, and J. Conesa, "Improving visualization skills in engineering education," *IEEE Computer Graphics and Applications*, vol. 25, no. 5, pp. 24–31, 2005.

[65] L. McNabb and R. S. Laramee, "Survey of surveys (SoS)-mapping the landscape of survey papers in information visualization," in *Computer Graphics Forum*, vol. 36, 3. Wiley Online Library, 2017, pp. 589–617.

[66] J. Fuchs, P. Isenberg, A. Bezerianos, and D. Keim, "A systematic review of experimental studies on data glyphs," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 7, pp. 1863–1879, 2016.

[67] "IEEE Vis," http://ieeevis.org/year/2018/welcome, 2019, last Accessed: August, 2019.

[68] "IEEE Xplore," http://ieeexplore.ieee.org/Xplore/home.jsp, 2019, last Accessed: August, 2019.

[69] "Google Scholar," https://scholar.google.co.uk/, 2019, last Accessed: August, 2019.

[70] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. D. Stolper, M. M. Sedlmair, J. Chen, T. Möller, and J. Stasko, "vispubdata.org: A metadata collection about IEEE visualization (VIS) publications," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, 2019.

[71] "Eurographics Educational Activities," https://www.eg.org/wp/educational-activities/, 2019, last Accessed: August, 2019.

[72] "Eurographics Digital Library," https://diglib.eg.org/, 2019, last Accessed: August, 2019.

[73] "ACM digital library," https://dl.acm.org/, 2019, last Accessed: August 2019.

[74] "IEEE pacific visualization conference," https://www.pvis.org, 2019, last Accessed: August, 2019.

[75] A. Chamberlain, A. Crabtree, T. Rodden, M. Jones, and Y. Rogers, "Research in the wild: Understanding 'in the wild'approaches to design and development," in *Proceedings of the Designing Interactive Systems Conference*, 2012, pp. 795–796.

[76] Y. Rogers and P. Marshall, "Research in the wild," *Synthesis Lectures on Human-Centered Informatics*, vol. 10, no. 3, pp. i–97, 2017.

[77] L. Grammel, M. Tory, and M.-A. Storey, "How information visualization novices construct visualizations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 943–952, 2010.

[78] N. Kodagoda, B. Wong, C. Rooney, and N. Khan, "Interactive visualization for low literacy users: from lessons learnt to design," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1159–1168.

[79] J. Johansson and C. Forsell, "Evaluation of parallel coordinates: Overview, categorization and guidelines for future research," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 579–588, 2015.

[80] "Visual literacy competency standards for higher education," 2011, Last accessed: December 2019, http://www.ala.org/acrl/standards/visualliteracy.

[81] R. S. Laramee, "How to read a visualization research paper: Extracting the essentials," *IEEE Computer Graphics and Applications*, vol. 31, no. 3, pp. 78–82, 2011.

[82] R. S. Baker, A. T. Corbett, and K. R. Koedinger, "Toward a model of learning data representations," in *Proceedings of the 23rd annual conference of the Cognitive Science Society*. Mahwah NJ, 2001, pp. 45–50.

[83] R. Delmas, J. Garfield, and A. Ooms, "Using assessment items to study students' difficulty reading and interpreting graphical representations of distributions," in *Fourth Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, 2005, pp. –.

[84] S. Huron, Y. Jansen, and S. Carpendale, "Constructing visual representations: Investigating the use of tangible tokens," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2102–2111, 2014.

[85] M. A. Wojton, D. Hayde, J. E. Heimlich, and K. Börner, "Begin at the beginning: A constructionist model for interpreting data visualizations," *Curator: The Museum Journal*, vol. 61, no. 4, pp. 559–574, 2018.

[86] H. Mansoor and L. Harrison, "Data visualization literacy and visualization biases: Cases for merging parallel threads," in *Cognitive Biases in Visualizations*. Springer, 2018, pp. 87–96.

[87] C. D'Ignazio and R. Bhargava, "Data visualization literacy: A feminist starting point," *Data Visualization in society*, p. 207, 2020.

[88] D. Donohoe and E. Costello, "Data visualisation literacy in higher education: An exploratory study of understanding of a learning dashboard tool," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 17, pp. 115–126, 2020.

[89] O. Barral, S. Lallé, and C. Conati, "Understanding the effectiveness of adaptive guidance for narrative visualization: a gaze-based analysis," in *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 2020, pp. 1–9.

[90] O. Barral, S. LallÉ, A. Iranpour, and C. Conati, "Effect of adaptive guidance and visualization literacy on gaze attentive behaviors and sequential patterns on magazine-style narrative visualizations," *ACM Transactions on Interactive Intelligent Systems (TiiS)*, vol. 11, no. 3-4, pp. 1–46, 2021.

[91] K. Peppler, A. Keune, and A. Han, "Cultivating data visualization literacy in museums," *Information and Learning Sciences*, 2021.

[92] "D3.js: Data-Driven Documents," http://d3js.org, last accessed: May 2019.

[93] A. Samuel, G. Shankar, A. Sivakumar, G. Thulasidoss, R. Kumar, B. Balachandran, K. Palani, and S. Angadi, "Electronic discovery systems and workflow management method," 2014, uS Patent App. 14/211,413.

[94] D. Toker, C. Conati, and G. Carenini, "Gaze analysis of user characteristics in magazine style narrative visualizations," *User Modeling and User-Adapted Interaction*, vol. 29, no. 5, pp. 977–1011, 2019.

[95] J. Garfield, R. Delmas, and B. Chance, "The assessment resource tools for improving statistical thinking (ARTIST) project. nsf ccli grant asa-0206571," 2002.

[96] G. Lupi and S. Posavec, *Dear data*. Chronicle books, 2016.

[97] E. Peters, N. Dieckmann, A. Dixon, J. H. Hibbard, and C. Mertz, "Less is more in presenting quality information to consumers," *Medical Care Research and Review*, vol. 64, no. 2, pp. 169–190, 2007.

[98] J. T. Cacioppo, R. E. Petty, and C. Feng Kao, "The efficient assessment of need for cognition," *Journal of personality assessment*, vol. 48, no. 3, pp. 306–307, 1984.

[99] J. R. Kirby, P. J. Moore, and N. J. Schofield, "Verbal and visual learning styles," *Contemporary educational psychology*, vol. 13, no. 2, pp. 169–184, 1988.

[100] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models," *Management science*, vol. 35, no. 8, pp. 982–1003, 1989.

[101] K. E. Stanovich and R. F. West, "On the relative independence of thinking biases and cognitive ability." *Journal of personality and social psychology*, vol. 94, no. 4, p. 672, 2008.

[102] E. Dimara, P. Dragicevic, and A. Bezerianos, "Accounting for availability biases in information visualization," *arXiv preprint arXiv:1610.02857*, 2016.

[103] E. Dimara, A. Bezerianos, and P. Dragicevic, "The attraction effect in information visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 471–480, 2016.

[104] P. Dragicevic and Y. Jansen, "Visualization-mediated alleviation of the planning fallacy," *IEEE VIS, DECISIVe : Workshop on Dealing with Cognitive Biases in Visualisations*, 2014.

[105] K. Börner and D. E. Polley, *Visual insights: A practical guide to making sense of data*. MIT Press, 2014.

[106] A. V. Maltese, J. Harsh, and D. Svetina, "Interpretation of graphical representations along the novice –expert continuum," *J Coll Sci Teach*, 2011. [Online]. Available: https://doi.org/10.2307/749671

[107] M. Romero, M. Velez, G. McInemy, D. Silver, and M. Chen, "Towards visualization literacy," *EuroVis Workshop*, 2014. [Online]. Available: https://www.kth.se/profile/marior/page/eurovis-2014-workshop-towards-visualiza

[108] S.-H. Kim, J. Boy, S. Lee, J. S. Yi, and N. Elmqvist, "Towards an open visualization literacy testing platform," *IEEEVIS 2014 Workshop*, 2014. [Online]. Available: http://visualizationliteracy.org

[109] N. Kong, J. Heer, and M. Agrawala, "Perceptual guidelines for creating rectangular treemaps," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 990–998, 2010. [Online]. Available: http://dx.doi.org/10.1109/tvcg.2010.186

[110] D. Rees and R. S. Laramee, "A survey of information visualization books," *Computer Graphics Forum*, vol. 38, no. 1, pp. 610–646, 2019. [Online]. Available: https://doi.org/10.1111/cgf.13595

[111] W. Scheibel, D. Limberger, and J. Döllner, "Survey of treemap layout algorithms," in *Proceedings of the 13th International Symposium on Visual Information Communication and Interaction*, 2020, pp. 1–9.

[112] J. Stasko, R. Catrambone, M. Guzdial, and K. McDonald, "An evaluation of space-filling information visualizations for depicting hierarchical structures," *International journal of human-computer studies*, vol. 53, no. 5, pp. 663–694, 2000.

[113] L. K. Long, L. C. Hui, G. Y. Fook, and W. M. N. W. Zainon, "A study on the effectiveness of tree-maps as tree visualization techniques," *Procedia Computer Science*, vol. 124, pp. 108–115, 2017.

[114] N. H. Müller, B. Liebold, D. Pietschmann, P. Ohler, and P. Rosenthal, "Hierarchy visualization designs and their impact on perception and problem solving strategies," in *Proceedings of the International Conference on Advances in Computer-Human Interactions*, 2017, pp. 93–101.

[115] Y. Tu and H.-W. Shen, "Visualizing changes of hierarchical data using treemaps," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1286–1293, 2007. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2007.70529

[116] ——, "Balloon focus: a seamless multi-focus+context method for treemaps," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1157–1164, 2008. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2008.114

[117] C. Ziemkiewicz and R. Kosara, "The shaping of information by visual metaphors," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1269–1276, 2008. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2008.171

[118] L. Woodburn, Y. Yang, and K. Marriott, "Interactive visualisation of hierarchical quantitative data: an evaluation," *arXiv preprint arXiv:1908.01277*, 2019. [Online]. Available: http://dx.doi.org/10.1109/VISUAL.2019.8933545

[119] R. C. Roberts, C. Tong, R. S. Laramee, G. A. Smith, P. Brookes, and T. D'Cruze, "Interactive analytical treemaps for visualisation of call centre data," in *Eurographics Italian Chapter Conference*, 2016, pp. 109–117. [Online]. Available: http://dx.doi.org/10.2312/stag.20161370

[120] R. Rosenbaum and B. Hamann, "Progressive presentation of large hierarchies using treemaps," in *International Symposium on Visual Computing.* Springer, 2009, pp. 71–80.

[121] A. Macquisten, A. M. Smith, and S. J. Fernstad, "Evaluation of hierarchical visualization for large and small hierarchies," in *2020 24th International Conference Information Visualisation (IV).* IEEE, 2020, pp. 166–173.

[122] B. B. Bederson, B. Shneiderman, and M. Wattenberg, "Ordered and quantum treemaps: Making effective use of 2d space to display hierarchies," *ACM Transactions on Graphics (TOG)*, vol. 21, no. 4, pp. 833–854, 2002. [Online]. Available: https://doi.org/10.1016/B978-155860915-0/50033-0

[123] "New York State P-12 Common Core Learning Standards for Mathematics," https://www.engageny.org/resource/new-york-state-p-12-common-core-learning/-standards-for-mathematics), New York State Education Department, 2013.

[124] C. D. of Education, *The California Common Core State Standards: Mathematics*, 2013. [Online]. Available: https://www.csai-online.org/resource/194

[125] "The 2014 Indiana Academic Standards for Mathematics," https://www.doe.in.gov/standards/mathematics), Indiana Department of Education, 2014.

[126] "National Governors Association and Others, Washington, DC," http://www.corestandards.org/about-the-standards/development-process/), Common Core State Standards, 2010.

[127] "Public health england," https://www.gov.uk/government/organisations/public-health-england, last accessed: 2019.

[128] C. Tong, L. McNabb, R. S. Laramee, J. Lyons, A. Walters, D. Berridge, and D. Thayer, "Time-oriented cartographic treemap for visualization of publice health care data," in *Proceedings of the Conference on Computer Graphics and Visual Computing (CGVC)*, vol. 1, 2017. [Online]. Available: http://dx.doi.org/10.2312/cgvc.20171273

[129] W. G. van Panhuis, A. Cross, and D. S. Burke, "Project tycho 2.0: a repository to improve the integration and reuse of data for global population health," *Journal of the American Medical Informatics Association*, vol. 25, no. 12, pp. 1608–1617, 2018. [Online]. Available: https://doi.org/10.1093/jamia/ocy123

[130] "Tableau desktop," https://www.tableau.com/products/desktop, last accessed: 2019.

[131] "IBM watson," https://www.ibm.com/watson, last accessed: 2019.

[132] "Microsoft PowerBI Desktop," https://powerbi.microsoft.com/en-us/desktop/, 2003, last accessed: 2019.

[133] M. Lima, *The book of trees: Visualizing branches of knowledge.* Princeton Architectural Press, 2014.

[134] "Treemap Literacy Pre-Intervention Test," 2019, http://bit.ly/2kHChcn.

[135] "Treemap Literacy Post-Intervention Test," 2019, http://bit.ly/2kfsdqQ.

[136] M. Sedlmair, M. Meyer, and T. Munzner, "Design study methodology: Reflections from the trenches and the stacks," *IEEE transactions on visualization and computer graphics*, vol. 18, no. 12, pp. 2431–2440, 2012.

[137] B. Shneiderman and M. Wattenberg, "Ordered treemap layouts," in *IEEE Symposium on Information Visualization, 2001. INFOVIS 2001.* IEEE, 2001, pp. 73–78. [Online]. Available: https://doi.org/10.1109/INFVIS.2001.963283

[138] "Qt cross-platform software development tool," https://www.qt.io/, Last accessed: 2022.

[139] R. C. Roberts, L. McNabb, N. AlHarbi, and R. S. Laramee, "Spectrum: a c++ header library for colour map management," in *Proceedings of the Conference on Computer Graphics & Visual Computing.* Eurographics Association, 2018, pp. 135–141. [Online]. Available: https://doi.org/10.2312/cgvc.20181218

[140] M. Harrower and C. A. Brewer, "Colorbrewer. org: an online tool for selecting colour schemes for maps," *The Cartographic Journal*, vol. 40, no. 1, pp. 27–37, 2003. [Online]. Available: http://dx.doi.org/10.1179/000870403235002042

[141] "United states regions," 2021, https://www.nationalgeographic.org/maps/united-states-regions/.

[142] S. M. Smith, R. Smith, J. Smith, and S. Orgill, "Qualtrics," 2002, last accessed: July 2019, https://www.qualtrics.com.

[143] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006. [Online]. Available: https://doi.org/10.1191/1478088706qp063oa

[144] E. E. Firat, A. Denisova, M. L. Wilson, and R. S. Laramee, "A study of parallel coordinate plot literacy," *Technical Report, School of Computer Science, University of Nottingham*, 2021.

[145] R. Kosara, F. Bendix, and H. Hauser, “Parallel sets: Interactive exploration and visual analysis of categorical data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 4, pp. 558–568, 2006.

[146] H. Siirtola and K.-J. Räihä, “Interacting with parallel coordinates,” *Interacting with Computers*, vol. 18, no. 6, pp. 1278–1309, 2006.

[147] “Instructions for the Experiment Video,” https://www.youtube.com/watch?v=ReNKoU9AI5U, Last accessed: 2020.

[148] “Introduction to Parallel Coordinates: A Tutorial (Using Slides),” https://www.youtube.com/watch?v=_G7eJHtgH5Y, Last accessed: 2020.

[149] “Introduction to Parallel Coordinates: A Tutorial (Using Software),” https://www.youtube.com/watch?v=JUfTNNlv97A, Last accessed: 2020.

[150] A. Inselberg, *Parallel coordinates*. Springer, 2009.

[151] A. Dasgupta, M. Chen, and R. Kosara, “Conceptualizing visual uncertainty in parallel coordinates,” in *Computer Graphics Forum*, vol. 31, 3pt2. Wiley Online Library, 2012, pp. 1015–1024.

[152] J. Heinrich and D. Weiskopf, “State of the art of parallel coordinates,” in *Eurographics State-of-Art reports (STARs)*. Eurographics Association, 2013, pp. 95–116.

[153] J. Yang, M. O. Ward, and E. A. Rundensteiner, “Interactive hierarchical displays: a general framework for visualization and exploration of large multivariate data sets,” *Computers & Graphics*, vol. 27, no. 2, pp. 265–283, 2003.

[154] H. Siirtola, “Combining parallel coordinates with the reorderable matrix,” in *Proceedings International Conference on Coordinated and Multiple Views in Exploratory Visualization-CMV 2003-*. IEEE, 2003, pp. 63–74.

[155] M. Lind, J. Johansson, and M. Cooper, “Many-to-many relational parallel coordinates displays,” in *2009 13th International Conference Information Visualisation*. IEEE, 2009, pp. 25–31.

[156] J. H. Claessen and J. J. Van Wijk, “Flexible linked axes for multivariate data visualization,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2310–2316, 2011.

[157] “Amazon Mechanical Turk,” https://requester.mturk.com/, Last accessed: 2020.

[158] C. Reas and B. Fry, “Processing: a learning environment for creating interactive web graphics,” in *ACM SIGGRAPH 2003 Web Graphics*, 2003, pp. 1–1.

[159] R. Rosenbaum, J. Zhi, and B. Hamann, “Progressive parallel coordinates,” in *2012 IEEE Pacific Visualization Symposium*. IEEE, 2012, pp. 25–32.

[160] G. Palmas, M. Bachynskyi, A. Oulasvirta, H. P. Seidel, and T. Weinkauf, “An edge-bundling layout for interactive parallel coordinates,” in *2014 IEEE Pacific Visualization Symposium*. IEEE, 2014, pp. 57–64.

[161] R. Kanjanabose, A. Abdul-Rahman, and M. Chen, “A multi-task comparative study on scatter plots and parallel coordinates plots,” in *Computer Graphics Forum*, vol. 34, 3. Wiley Online Library, 2015, pp. 261–270.

[162] "High-D Tool," https://www.high-d.com/, Last accessed: 2021.

[163] "Mondrian Tool," http://www.theusrus.de/Mondrian/, Last accessed: 2020.

[164] "Quadrigram Tool," https://www.quadrigram.com/, Last accessed: 2020.

[165] "Xdat Tool: A free parallel coordinates software," https://www.xdat.org/, Last accessed: 2020.

[166] M. O. Ward, G. Grinstein, and D. Keim, *Interactive data visualization: foundations, techniques, and applications.* CRC Press, 2010.

[167] "GGobi Data Visualization System," http://ggobi.org/, 2021, Last accessed:.

[168] "Sliver Tool: Multivariate Data Visualization Software," http://www.sliversoftware.com/features.htm, Last accessed: 2020.

[169] "Microsoft PowerBI," https://powerbi.microsoft.com/en-us/downloads/, Last accessed: 2020.

[170] "Spotfire Tool," https://www.tibco.com/products/tibco-spotfire, Last accessed: 2020.

[171] "IBM SPSS Statistics," https://www.ibm.com/support/pages/downloading-ibm-spss-statistics-26, Last accessed: 2020.

[172] M. A. Yalçın, N. Elmqvist, and B. B. Bederson, "Keshif: Rapid and expressive tabular data exploration for novices," *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 8, pp. 2339–2352, 2017.

[173] A. Kirk, "Visualizing Data," https://www.visualisingdata.com/resources/, Last accessed: 2020.

[174] "Xmdv Datasets," http://davis.wpi.edu/xmdv/datasets.shtml, Last accessed: 2020.

[175] "High-D Datasets," https://www.high-d.com/datasets/, Last accessed: 2020.

[176] R. Borgo, L. Micallef, B. Bach, F. McGee, and B. Lee, "Information visualization evaluation using crowdsourcing," in *Computer Graphics Forum*, vol. 37, 3. Wiley Online Library, 2018, pp. 573–595.

[177] A. Inselberg and B. Dimsdale, "Parallel coordinates: a tool for visualizing multi-dimensional geometry," in *Proceedings of the First IEEE Conference on Visualization: Visualization90.* IEEE, 1990, pp. 361–378.

[178] Z. Geng, Z. Peng, R. S. Laramee, J. C. Roberts, and R. Walker, "Angular histograms: Frequency-based visualizations for large, high dimensional data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2572–2580, 2011.

[179] J. Wang, S. Hazarika, C. Li, and H.-W. Shen, "Visualization and visual analysis of ensemble data: A survey," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 9, pp. 2853–2872, 2018.

[180] A. O. Artero, M. C. F. de Oliveira, and H. Levkowitz, "Uncovering clusters in crowded parallel coordinates visualizations," in *IEEE Symposium on Information Visualization.* IEEE, 2004, pp. 81–88.

[181] Y.-H. Fua, M. O. Ward, and E. A. Rundensteiner, *Hierarchical parallel coordinates for exploration of large datasets.* IEEE, 1999.

[182] H. Siirtola, "Direct manipulation of parallel coordinates," in *2000 IEEE Conference on Information Visualization. An International Conference on Computer Visualization and Graphics.* IEEE, 2000, pp. 373–378.

[183] G. Andrienko and N. Andrienko, "Parallel coordinates for exploring properties of subsets," in *Proceedings. Second International Conference on Coordinated and Multiple Views in Exploratory Visualization, 2004.* IEEE, 2004, pp. 93–104.

[184] J. Heinrich, Y. Luo, A. E. Kirkpatrick, H. Zhang, and D. Weiskopf, "Evaluation of a bundling technique for parallel coordinates," *CoRR*, vol. abs/1109.6073, 2011.

[185] J. Blaas, C. Botha, and F. Post, "Extensions of parallel coordinates for interactive exploration of large multi-timepoint data sets," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1436–1451, 2008.

[186] R. G. Raidou, M. Eisemann, M. Breeuwer, E. Eisemann, and A. Vilanova, "Orientation-enhanced parallel coordinate plots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 589–598, 2015.

[187] R. C. Roberts, R. S. Laramee, G. A. Smith, P. Brookes, and T. D'Cruze, "Smart brushing for parallel coordinates," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 3, pp. 1575–1590, 2018.

[188] G. Ellis and A. Dix, "Enabling automatic clutter reduction in parallel coordinate plots," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 717–724, 2006.

[189] J. Johansson, P. Ljung, M. Jern, and M. Cooper, "Revealing structure in visualizations of dense 2d and 3d parallel coordinates," *Information Visualization*, vol. 5, no. 2, pp. 125–136, 2006.

[190] K. T. McDonnell and K. Mueller, "Illustrative parallel coordinates," in *Computer Graphics Forum*, vol. 27, no. 3. Wiley Online Library, 2008, pp. 1031–1038.

[191] R. S. Divino, G. Carlos, B. S. Meiguins *et al.*, "A visual representation of clusters characteristics using edge bundling for parallel coordinates," in *2017 21st International Conference Information Visualisation (IV).* IEEE, 2017, pp. 90–95.

[192] A. Lhuillier, C. Hurter, and A. Telea, "State of the art in edge and trail bundling techniques," in *Computer Graphics Forum*, vol. 36, no. 3. Wiley Online Library, 2017, pp. 619–645.

[193] H. Zhou, X. Yuan, H. Qu, W. Cui, and B. Chen, "Visual clustering in parallel coordinates," in *Computer Graphics Forum*, vol. 27, no. 3. Wiley Online Library, 2008, pp. 1047–1054.

[194] G. Palmas and T. Weinkauf, "Space bundling for continuous parallel coordinates," in *Proceedings of the Eurographics/IEEE VGTC Conference on Visualization: Short Papers*, 2016, pp. 61–65.

[195] "The Scottish COVID-19 Response Consortium: Home page." [Online]. Available: https://www.gla.ac.uk/research/az/scrc/

[196] T. Munzner, "A nested model for visualization design and validation," *IEEE transactions on visualization and computer graphics*, vol. 15, no. 6, pp. 921–928, 2009.

[197] K. Koc, A. S. McGough, and S. Johansson Fernstad, "Peaglyph: Glyph design for investigation of balanced data structures," *Information Visualization*, vol. 21, no. 1, pp. 74–92, 2022.

[198] R. Borgo, J. Kehrer, D. H. Chung, E. Maguire, R. S. Laramee, H. Hauser, M. Ward, and M. Chen, "Glyph-based visualization: Foundations, design guidelines, techniques and applications." in *Eurographics (State of the Art Reports)*, 2013, pp. 39–63.

[199] K. A. Bollen and K. H. Barb, "Pearson's r and coarsely categorized measures," *American Sociological Review*, pp. 232–239, 1981.

[200] "Parallel coordinates plot software." [Online]. Available: https://vimeo.com/652208042

[201] C. Ericson, *Real-Time Collision Detection.* CRC Press, Inc., 2014.

[202] T. Hogan, U. Hinrichs, and E. Hornecker, "The elicitation interview technique: Capturing people's experiences of data representations," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 12, pp. 2579–2593, 2015.

[203] "Covid-19 EERAModel." [Online]. Available: https://github.com/ScottishCovidResponse/Covid19_EERAModel

[204] A. Pease and B. Pease, *Why Men Don't Listen & Women Can't Read Maps.* Pease International Pty Ltd, 2003.

[205] J. O'Connor, "Structural visualization," *Boston, Human Engineering Laboratory*, 1943.

[206] J. P. Guilford, "The guilford-zimmerman aptitude survey." *Personnel & Guidance Journal*, 1956.

[207] R. E. Stafford, "Sex differences in spatial visualization as evidence of sex-linked inheritance," *Perceptual and motor skills*, vol. 13, no. 3, pp. 428–428, 1961.

[208] D. Witelson, "Sex and the single hemisphere: Specialization of the right hemisphere for spatial processing," *Science*, vol. 193, no. 4251, pp. 425–427, 1976.

[209] D. P. Waber, "Sex differences in cognition: a function of maturation rate?" *Science, 192,*, pp. pp.425–427, 1976.

[210] B. Sanders, M. P. Soares, and J. M. D'Aquila, "The sex difference on one test of spatial visualization: A nontrivial difference," *Child development*, pp. 1106–1110, 1982.

[211] P. P. Gilmartin and J. C. Patton, "Comparing the sexes on spatial abilities: map-use skills," *Annals of the Association of American Geographers*, vol. 74, no. 4, pp. 605–619, 1984.

[212] P. P. Gilmartin, "Maps, mental imagery, and gender in the recall of geographical information," *The American Cartographer*, vol. 13, no. 4, pp. 335–344, 1986.

[213] C. P. Benbow and J. C. Stanley, "Sex differences in mathematical reasoning ability: More facts," *Science*, vol. 222, no. 4627, pp. 1029–1031, 1983.

[214] W. W. Beatty and A. I. Tröster, "Gender differences in geographical knowledge," *Sex Roles*, vol. 16, no. 11-12, pp. 565–590, 1987.

[215] K.-t. Chang and J. R. Antes, "Sex and cultural differences in map reading," *The American Cartographer*, vol. 14, no. 1, pp. 29–42, 1987.

[216] D. F. Halpern, "The disappearance of cognitive gender differences: What you see depends on where you look." 1989.

[217] D. Goldstein, D. Haldane, and C. Mitchell, "Sex differences in visual-spatial ability: The role of performance factors," *Memory & Cognition*, vol. 18, no. 5, pp. 546–550, 1990.

[218] M. Tory, "Mental registration of 2d and 3d visualizations (an empirical study)," pp. 371–378, 2003.

[219] K. Rehm, K. Lakshminaryan, S. Frutiger, K. A. Schaper, S. C. Strother, J. R. Anderson, D. A. Rottenberg *et al.*, "A symbolic environment for visualizing activated foci in functional neuroimaging datasets," *Medical Image Analysis*, vol. 2, no. 3, pp. 215–226, 1998.

[220] D. F. Halpern, "Sex differences in cognitive abilities," 2000.

[221] D. Kimura, *Sex and cognition*. MIT press, 2000.

[222] A. A. Rizzo, J. G. Buckwalter, J. S. McGee, T. Bowerly, C. v. d. Zaag, U. Neumann, M. Thiebaux, L. Kim, J. Pair, and C. Chua, "Virtual environments for assessing and rehabilitating cognitive/functional performance a review of projects at the usc integrated media systems center," *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 4, pp. 359–374, 2001.

[223] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," *Science*, vol. 171, no. 3972, pp. 701–703, 1971.

[224] M. Czerwinski, D. S. Tan, and G. G. Robertson, "Women take a wider view," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2002, pp. 195–202.

[225] Paraview, https://www.paraview.org/, last accessed: June 2019.

[226] Inviwo, http://www.inviwo.org/, last accessed: June 2019.

[227] W. Montgomery, "The inclusive learning and teaching handbook,," *Teaching Exceptional Children 33.4*, 2001.

[228] E. Rodrigue, "The inclusive learning and teaching handbook," *University of Sheffield, South YorkShire, United Kingdom*, 2010.

[229] J. C. Roberts, C. Headleand, and P. D. Ritsos, "Sketching designs using the five design-sheet methodology," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 1, pp. 419–428, 2015.

Appendix A

Treemap Literacy Test

A.0.1 Pre-Intervention Test

1. Consent Form (Q1)
2. Please write first 5 digits of your student number. (Q2)
3. What is your current education level? (Q3)

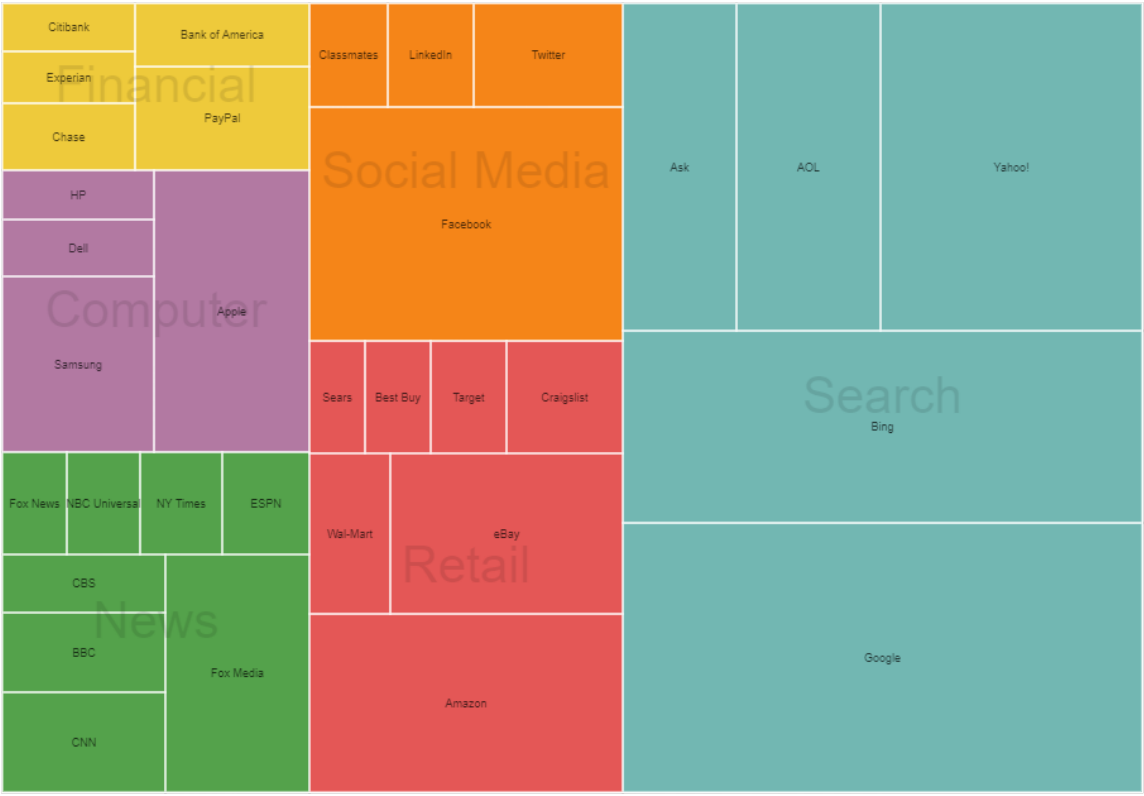
   A. Bachelor's B. Master's C. Ph.D. D. Other
4. What is your age? (Q4)

   A. 18-22 B. 23-27 C. 28-32 D. 33-37 E. 38-Above
5. What is your gender? (Q5)

   A. Female B. Male C. Prefer not to say
6. Are you color blind? (Q6)

   A. Yes B. No

The treemap shows the number of unique visitors for a range of categorized websites in 2010. The size of each node represents the number of unique visitors, and the color symbolize the different categories.(7, 8)



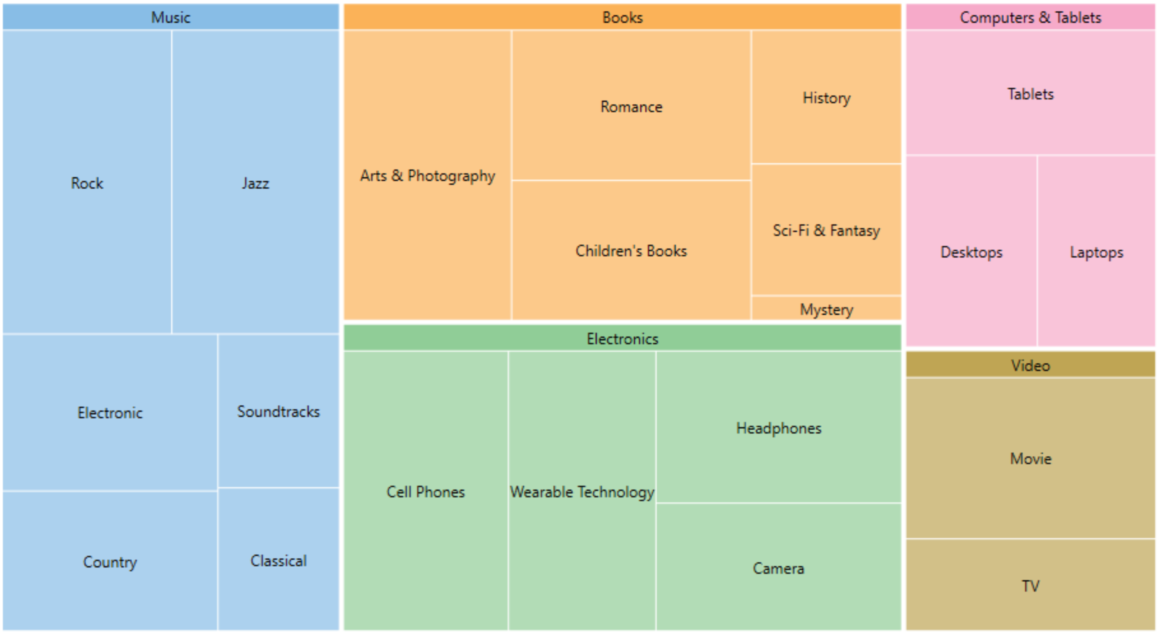
7. Which of the following website categories received more visits than the News category? (Q12)

   A. Computer B. Finance C. Social Media D. Retail E. Not Sure

8. The Apple website received more visits than Facebook. (Q52)

   A. True B. False C. Not Sure

The following treemap displays the distribution of money spent on entertainment. Color is mapped to entertainment types and size is mapped to money spent on each purchase. (9, 10, 11, 12)



9. Which entertainment industry has the least amount of money spent on it? (Q22)

   A. Music B. Books C. Video D. Electronics E. Not Sure

10. People spend more money on jazz music in comparison to history books. (Q48)

    A. True B. False C. Not Sure

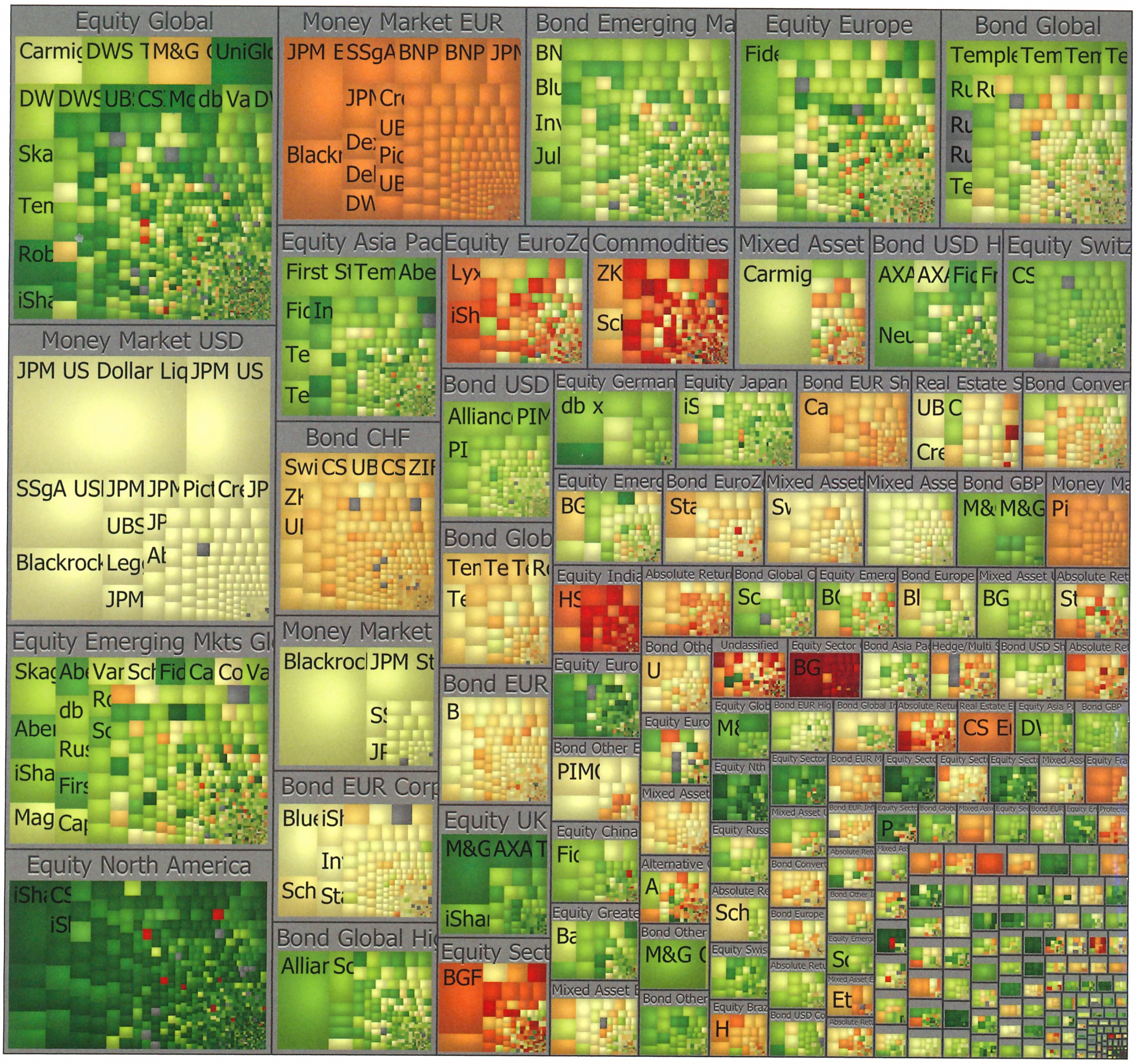
11. Which entertainment category has similar amount of money spent on all items in this category? (Q56)

    A. Books B. Music C. Computer & Tablets D. Video E. Not Sure

12. To which entertainment category does headphones belong to? (Q58)

    A. Electronics B. Music C. Computer & Tablets D. Video E. Not Sure

The following treemap shows investment funds registered in Switzerland. The size of each rectangle is proportional to total assets under each fund's management. Colors represent the funds' six-month performance with red showing losses and green profit. (13, 14)



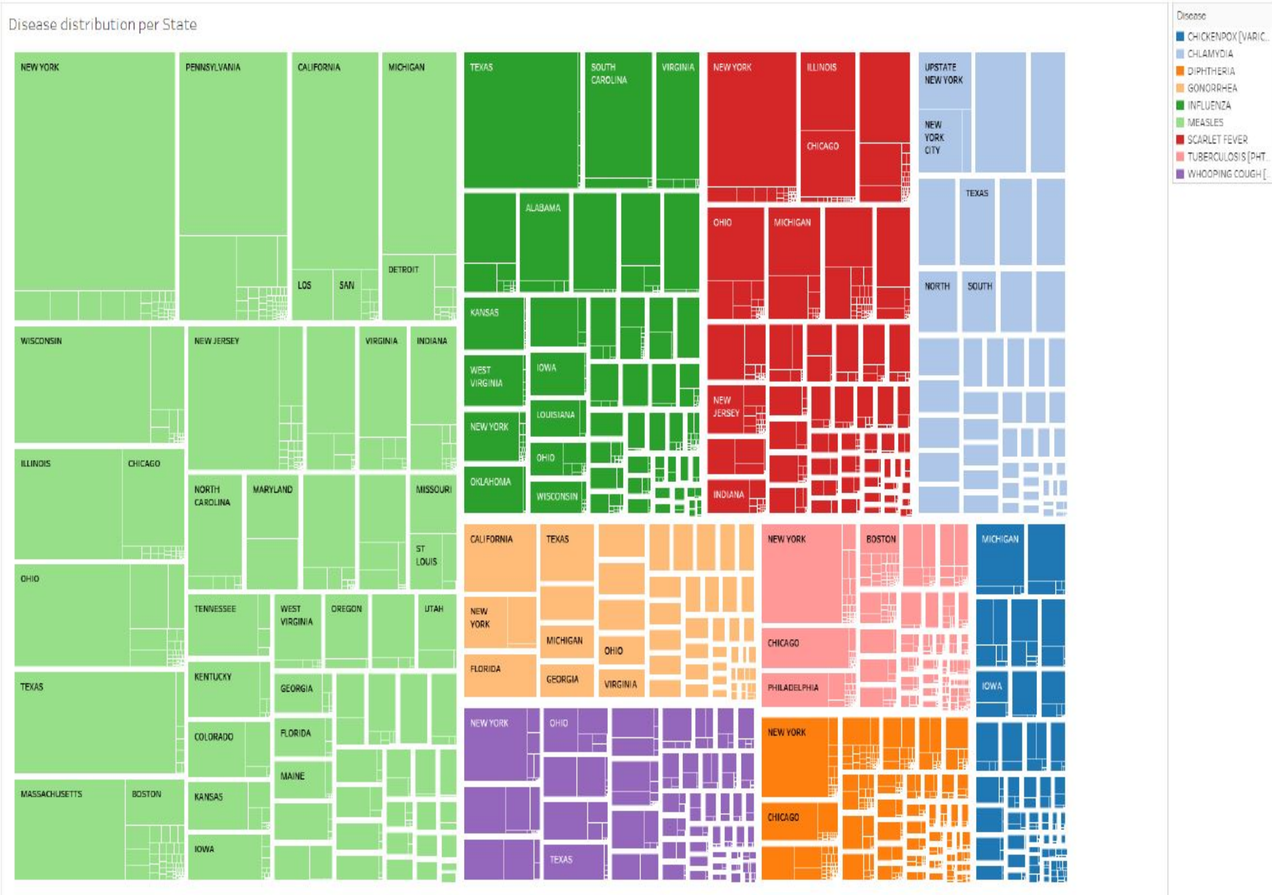
13. Which two markets do not have similar total assets? (Q14)

A. Equity Europe-Money Market USD B. Bond CHF-Equity Global
C. Equity Europe –Money Market EUR D. Money Market-Commodities
E. Not Sure

14. Which of the following funds have the largest assets despite showing a loss in the past six months? (Q42)

A. Money Market EUR B. Equity Asia Pacific C. Bond Global
D. Equity Europe E. Not Sure

The following treemap shows the distribution of states and their cities according to a total number of cases for different historic diseases. (15, 16, 17, 18)



15. Which state has the greatest number of cases of Measles? (Q26)

    A. Texas B. Ohio C. California D. New York E. Not Sure

16. How many diseases have a lower number of case than Chlamydia? (Q34)

    A. 6 B. 7 C. 5 D. 4 E. Not Sure

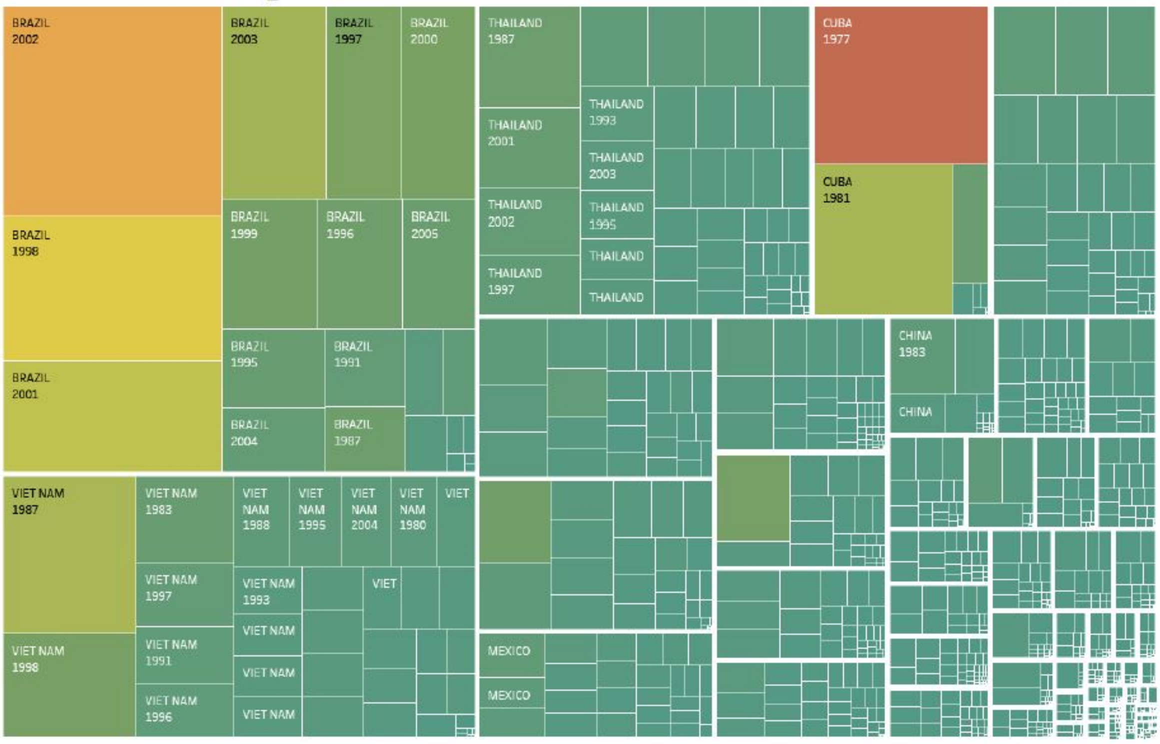
17. Which of the following disease had a higher number of cases than Scarlet Fever? (Q36)

    A. Diphtheria B. Measles C. Chlamydia D. Chickenpox E. Not Sure

18. More Diphtheria cases are recorded in New York than Chicago. (Q38)

    A. True B. False C. Not Sure

The following treemap shows the instances of a disease by country and year. Color is mapped to the percentage of number of cases which decreases from red to green. Size indicates the number of people with disease in the country. (19, 20, 21)



19. Which of the following countries has the lowest total number of people with a disease? (Q16)

    A. China B. Thailand C. Cuba D. Brazil E. Not Sure

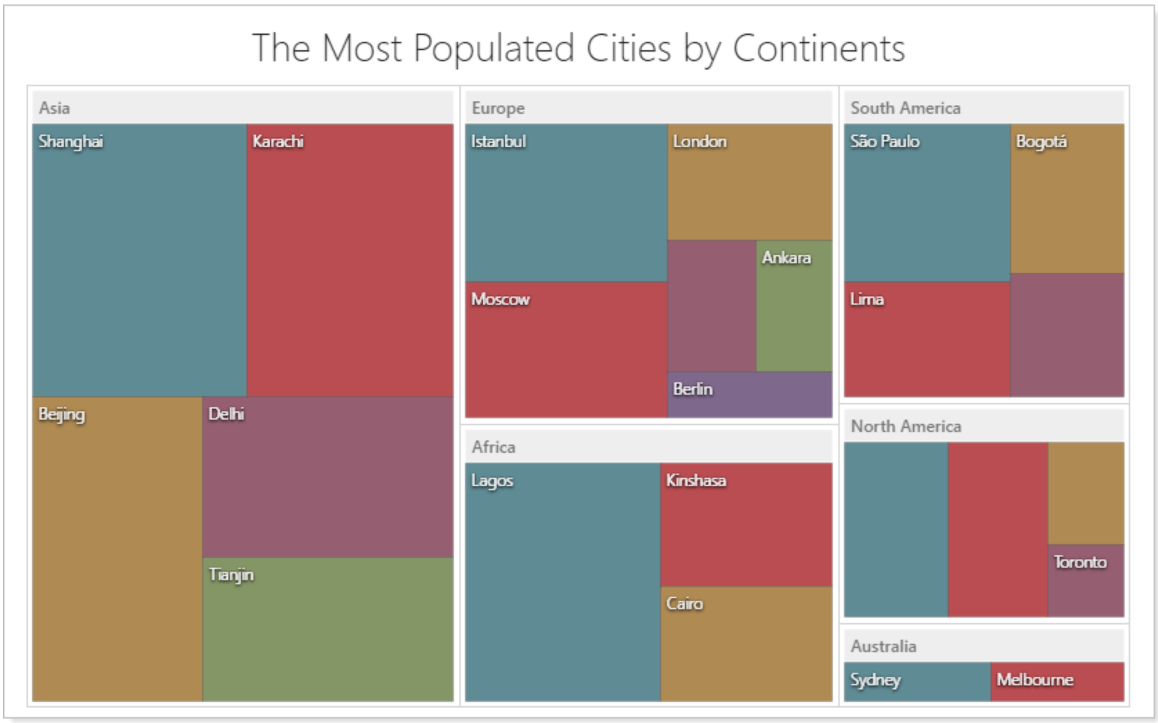
20. In which year and country is the highest percentage of disease cases? (Q62)

    A. Cuba-1981 B. Vietnam-1987 C. Cuba-1977 D. Brazil-2002
    E. Not Sure

21. Which year has the smallest percentage of cases in Brazil? (Q64)

    A. 2003 B. 2001 C. 1998 D. 1997 E. Not Sure

The following treemap displays the most populated cities by continents. The population is indicated by node size. (22, 23, 24)



22. The population of Asia is nearly equal to the total population of both Europe and Africa combined. (Q18)

    A. True B. False C. Not Sure

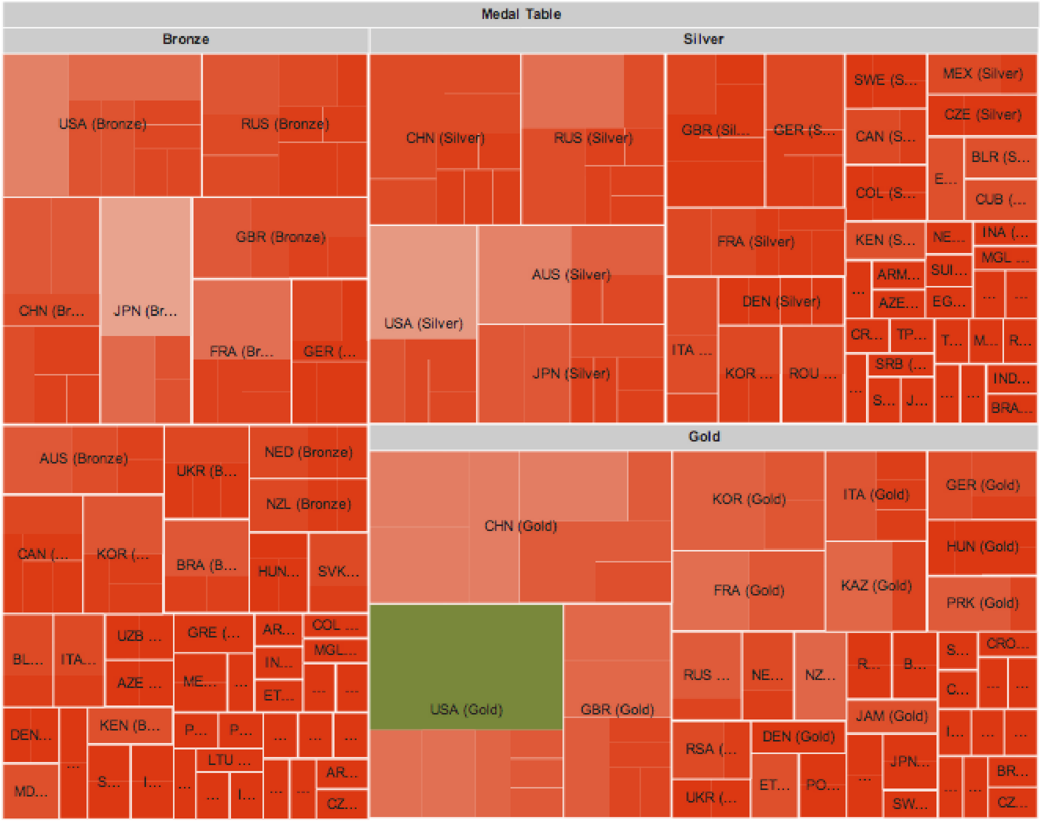
23. Karachi belongs to the most crowded continent. (Q44)

    A. True B. False C. Not Sure

24. Lagos is less populated than Kinshasa and Cairo. (Q46)

    A. True B. False C. Not Sure

This following figure shows a medal standings table for three different medals sorted by countries and sports branch. Size and color are mapped to the number of medals. The number of medals decreases from green to light red. (25, 26)



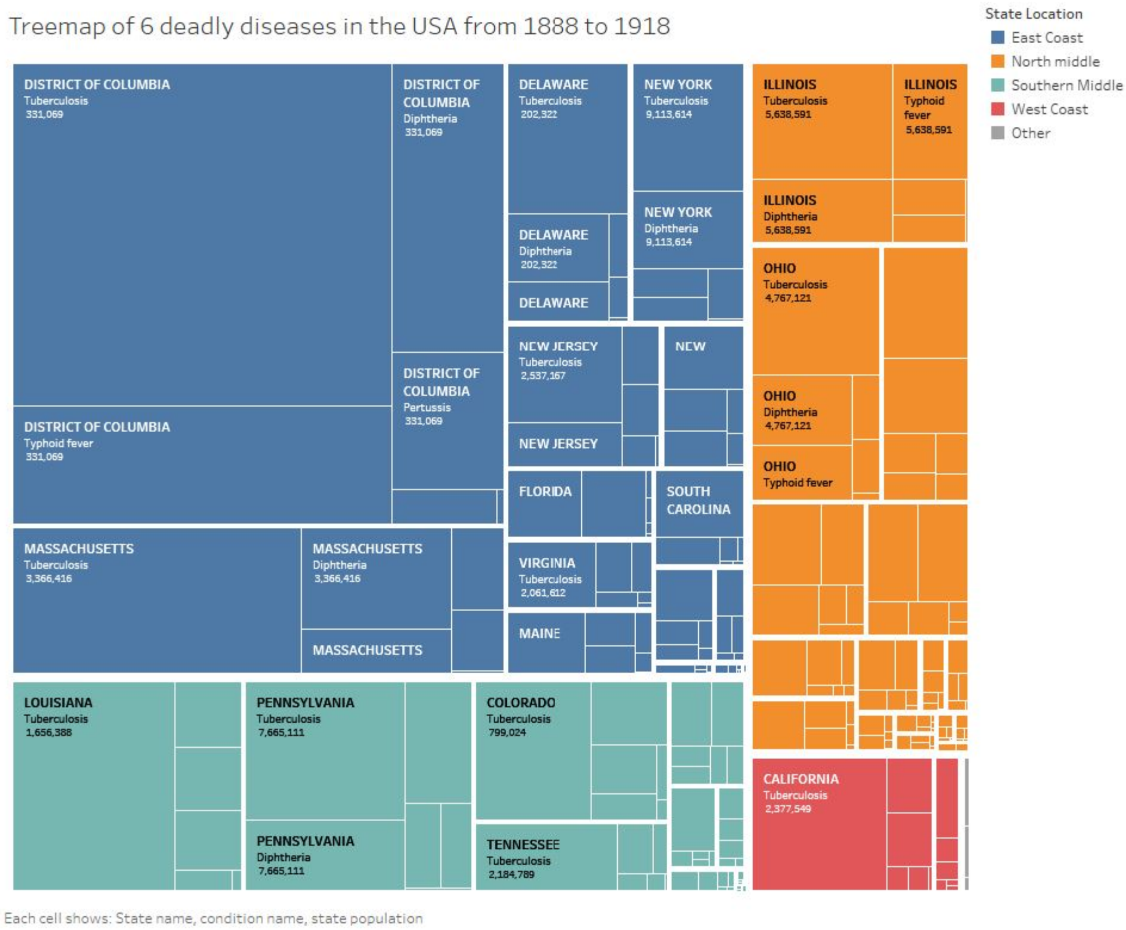
25. Does the number of countries with a silver medal equal the number with a gold? (Q20)

A. True B. False C. Not Sure

26. Which country has the highest number of gold medals awarded for a single event? (Q60)

A. GBR B. CHN C. USA D. AUS E. Not Sure

This treemap demonstrates historic deadly diseases in the USA. Size is mapped to the number of deaths. (27, 28)



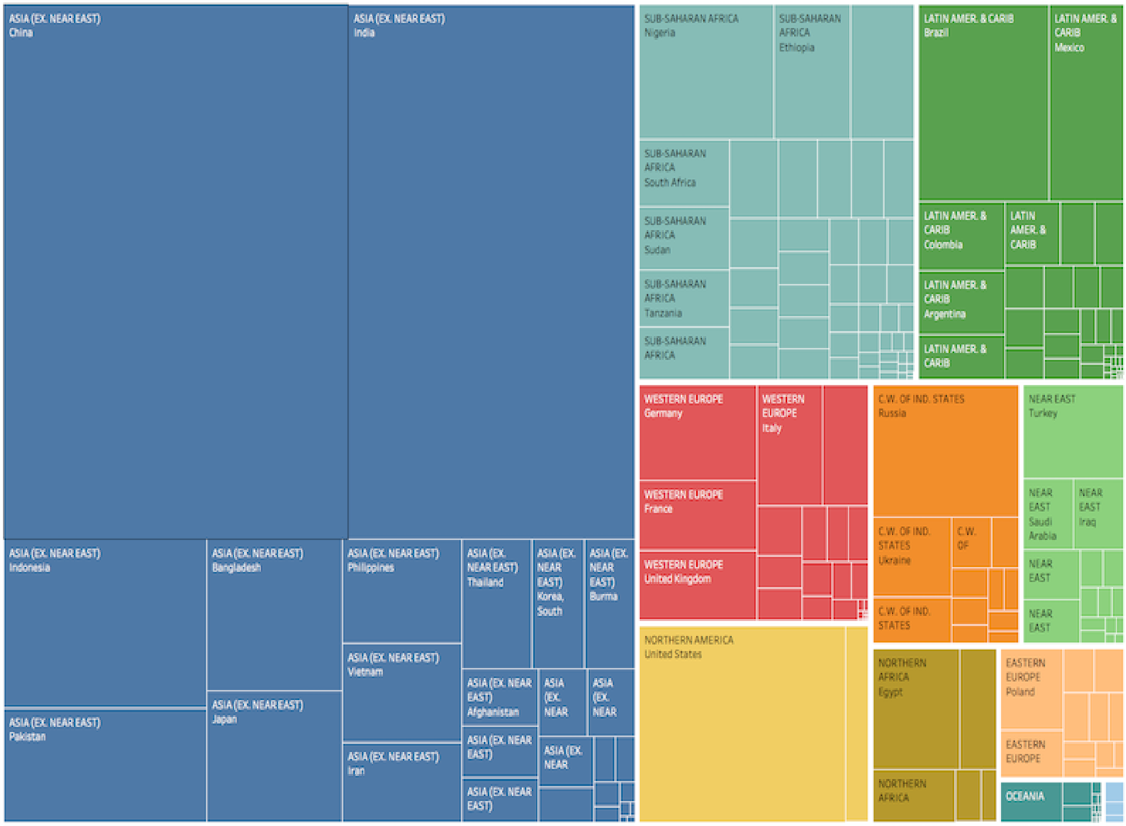
27. Which region had the highest number of deaths? (Q24)

A. North Middle B. West Coast C. Southern Middle D. East Coast
E. Not Sure

28. Which of the following disease causes the fewest deaths in the USA? (Q50)

A. Tuberculosis B. Diphtheria C. Typhoid fever D. Pertussis
E. Not Sure

The treemap displays the population of countries which are grouped by regions. Color indicates the regions and node size indicates the population. (29, 30)



29. Which of the following regions has a higher population than Western Europe? (Q28)

    A. Northern America B. Near East C. Latin America D. Eastern Europe E. Not Sure

30. To which region does Brazil belong, it being the largest population in that region? (Q66)

    A. Northern America-Near East B. Sub Saharan Africa-Asia C. Near East D. Latin America E. Not Sure

The following treemap displays the map of the market. The area of a rectangle corresponds to the market capitalization of the company, and the color indicates you how the stock price has changed since the previous market close with an increase in green and decrease in red. (31, 32)

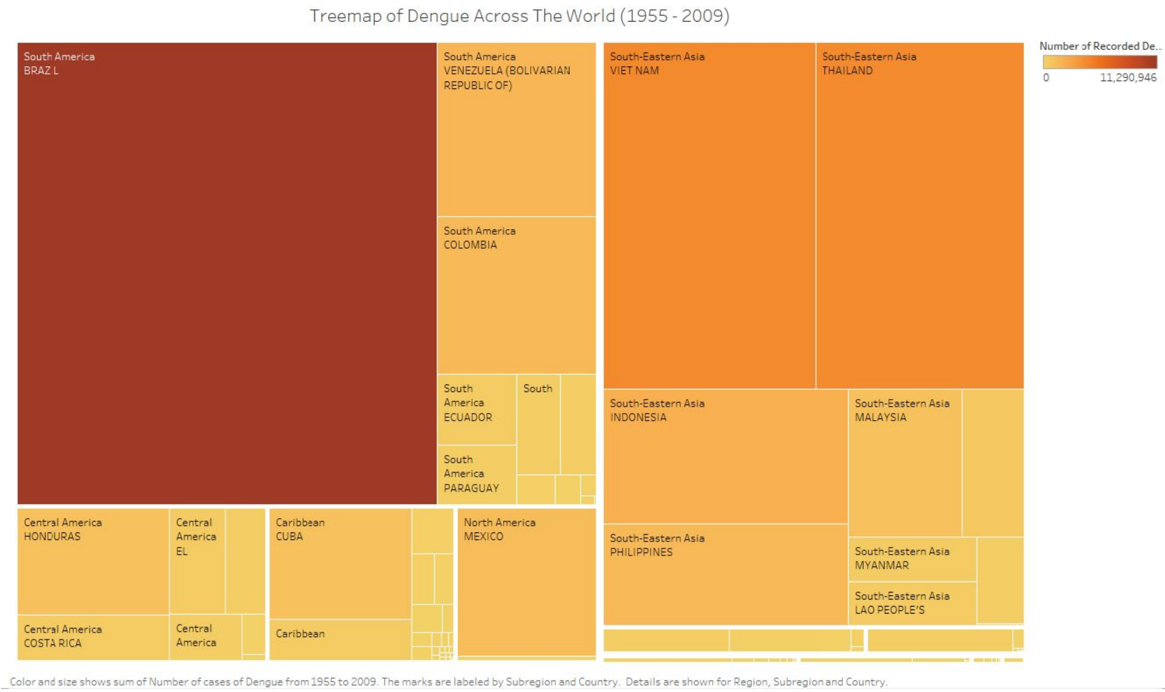
31. Which market had the largest volume of trading? (Q8)

    A. Communication B. Health Care C. Financial D. Transport
    E. Not Sure

32. Which two financial sectors have similar market capitalization? (Q30)

    A. Consumer Cyclicals and Technology B. Transport and Financial
    C. Energy and Capital Goods D. Health Care and Utilities E. Not Sure

The treemap displays the number of dengue cases observed in countries which are group by region. Size and color are mapped to the number people with dengue. (33, 34)



33. Which region had the highest incidence of dengue? (Q32)

    A. Caribbean B. Central America C. North America
    D. South America E. Not Sure

34. Vietnam and Thailand had a similar incidence of dengue. (Q54)

    A. True B. False C. Not Sure

Overview of the world's deadliest earthquakes since 1900. The magnitude of an earthquake is indicated by the color, with red indicating the highest magnitude. Node size represents the number of deaths caused by each earthquake. (35, 36)

35. Which two countries have a similar number of deaths? (Q10)

A. China-Mexico B. Taiwan-India C. Greece-Iran
D. Japan-United States E. Not Sure

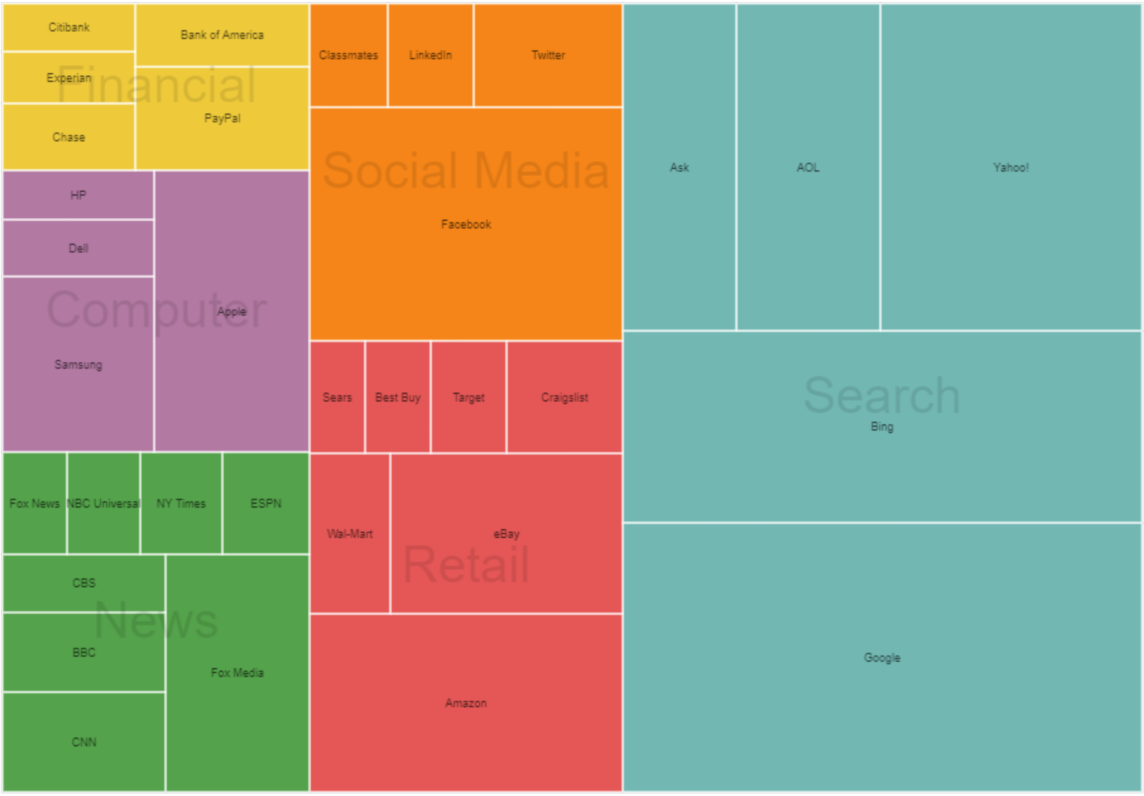
36. Which of the following earthquakes occurred in the country with the largest number of deaths? (Q40)

A. Peru: ICA, Nazca B. China: Xinjiang, Turkestan
C. Ecuador: Off Coast D. Japan-Honshu E. Not Sure

A.0.2 Post-Intervention Test

1. Please write first 5 digits of your student number. (Q1)
2. Please enter the Group ID (**S** Slides only, **SD** Software Demo only) (Q2)

   The Treemap shows the number of unique visitors for a range of categorized websites in 2010. The size of each node represents the number of unique visitors, and the color symbolize the different categories. (3, 4)



3. Which website had the most unique visitors in 2010? (Q4)

   A. Facebook B. Amazon C. Bing D. Google E. Not Sure

4. The number of unique visitors to Amazon was more than that of Yahoo in 2010. (Q6)

   A. True B. False C. Not Sure

5. Samsung is in the Financial category. (Q8)

   A. True B. False C. Not Sure

Overview of the world's deadliest earthquakes since 1900- The magnitude of an earthquake is indicated by the color, with red indicating the highest magnitude. Node size represents the number of deaths caused by each earthquake. (6, 7, 8)

6. Which country has the most deaths from earthquakes? (Q16)

   A. Indonesia B. Russia C. Chile D. Japan E. Not Sure

7. Which of the following earthquakes had the largest magnitude? (Q18)

   A. Alaska B. Chile: Atacama C. Indonesia: Sumatra
   D. Japan: Honshu E. Not Sure

8. Which of the following earthquakes had the largest magnitude? (Q20)

   A. Russia: Kuril Island B. China: Xinjiang province C. Chile: Chillan D. Mexico: Guerrero E. Not Sure

The following treemap displays the map of the market. The area of a rectangle corresponds to the market capitalization of the company, and the color indicates you how the stock price has changed since the previous market close with an increase in green and decrease in red. (9, 10, 11)

9. The market capitalization for the Transport sector is greater than that of Basic Materials. (Q10)

   A. True B. False C. Not Sure

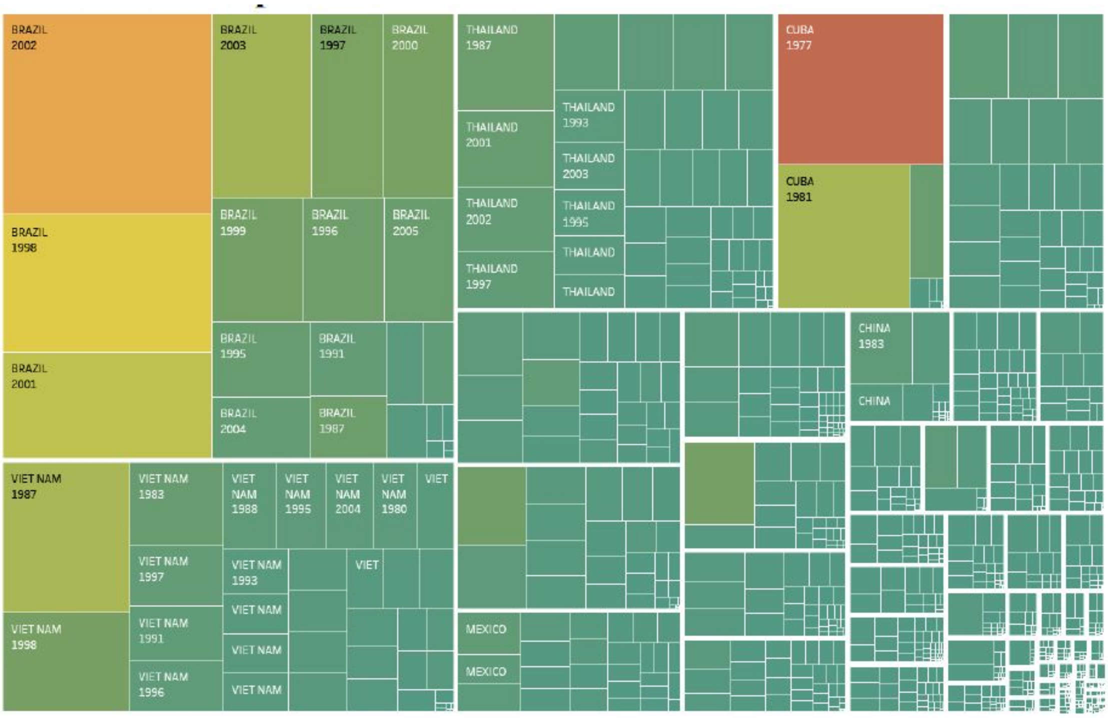
10. The Technology sector has a greater market capitalization than the Communication sector. (Q12)

    A. True B. False C. Not Sure

11. Which of the following sectors have a larger market capitalization than the Energy sector? (Q14)

    A. Utilities B. Basic Materials C. Transport D. Financial E. Not Sure

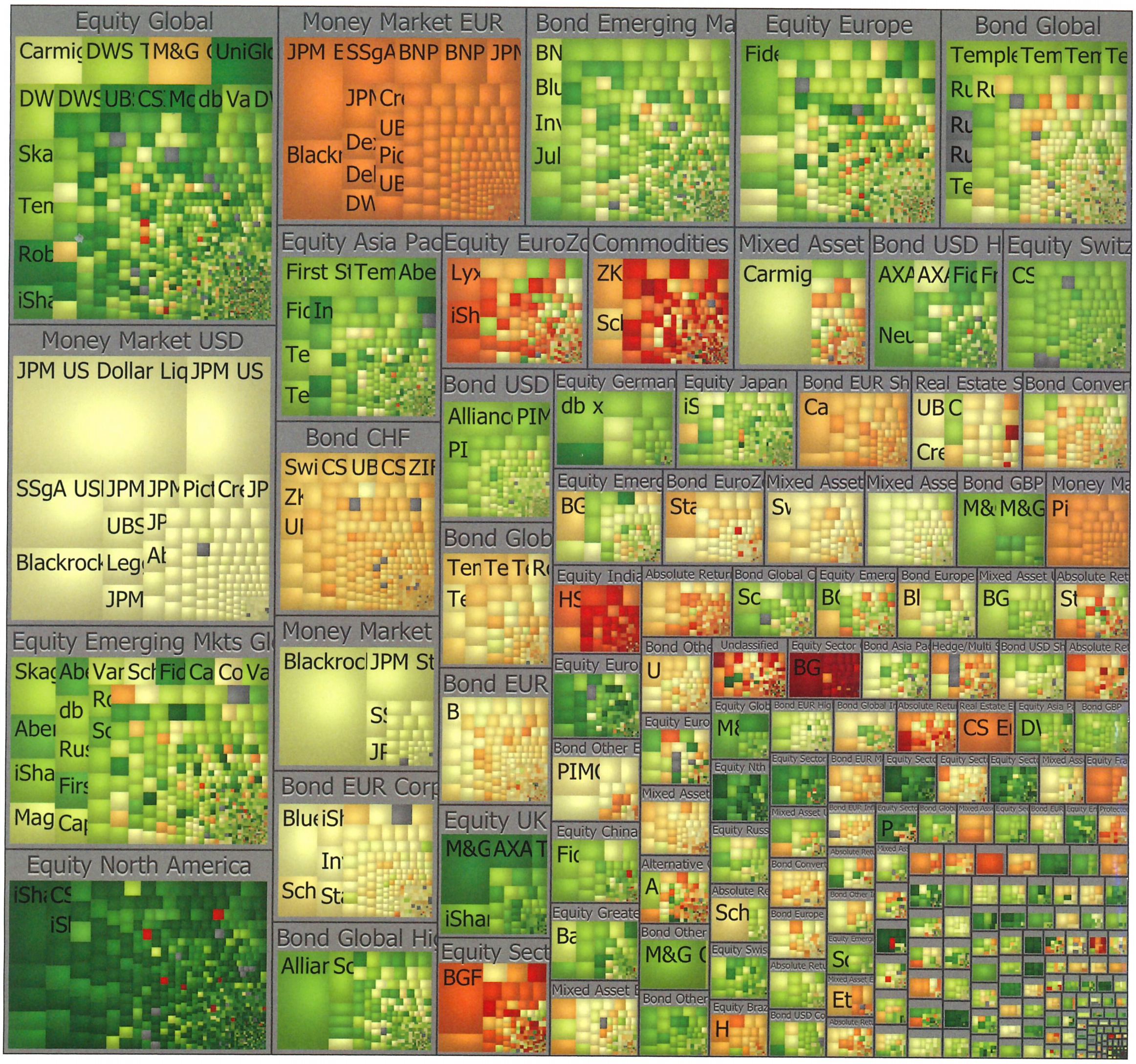
The following treemap shows the instances of a disease by country and year. Color is mapped to the percentage of number of cases which decreases from red to green. Size indicates the number of people with disease in the country. (12, 13)



12. Which country has the highest number of cases across all years? (Q28)
    A. Cuba B. Vietnam C. Thailand D. Brazil E. Not Sure

13. Which country has the year with the highest percentage of cases? (Q30)
    A. Vietnam B. Thailand C. Cuba D. Brazil E. Not Sure

The following treemap shows investment funds registered in Switzerland. The size of each rectangle is proportional to total assets under each fund's management. Colors represent the funds' six-month performance with red showing losses and green profit. (14, 15, 16)



14. Which market has the largest investment funds? (Q22)

    A. Money Market EUR B. Equity North America C. Money Market D. Equity Global E. Not Sure

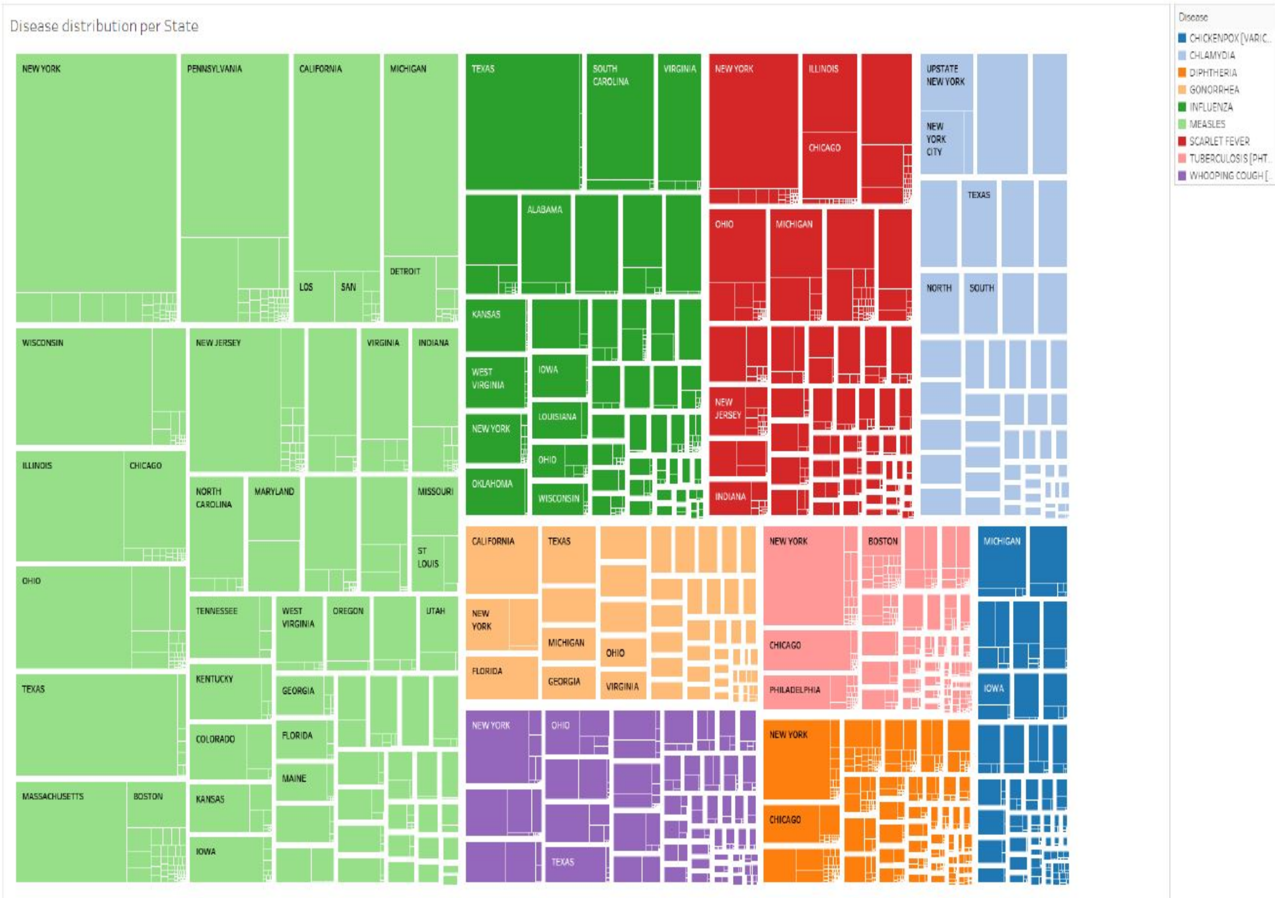
15. Which market's total assets has the largest profit? (Q24)

    A. Equity North America B. Money Market EUR C. Commodities D. Mixed Assets E. Not Sure

16. Which of the following funds have the largest assets despite showing a loss in the past six months? (Q26)

    A. Money Market EUR B. Equity Asia Pacific C. Bond Global D. Equity Europe E. Not Sure

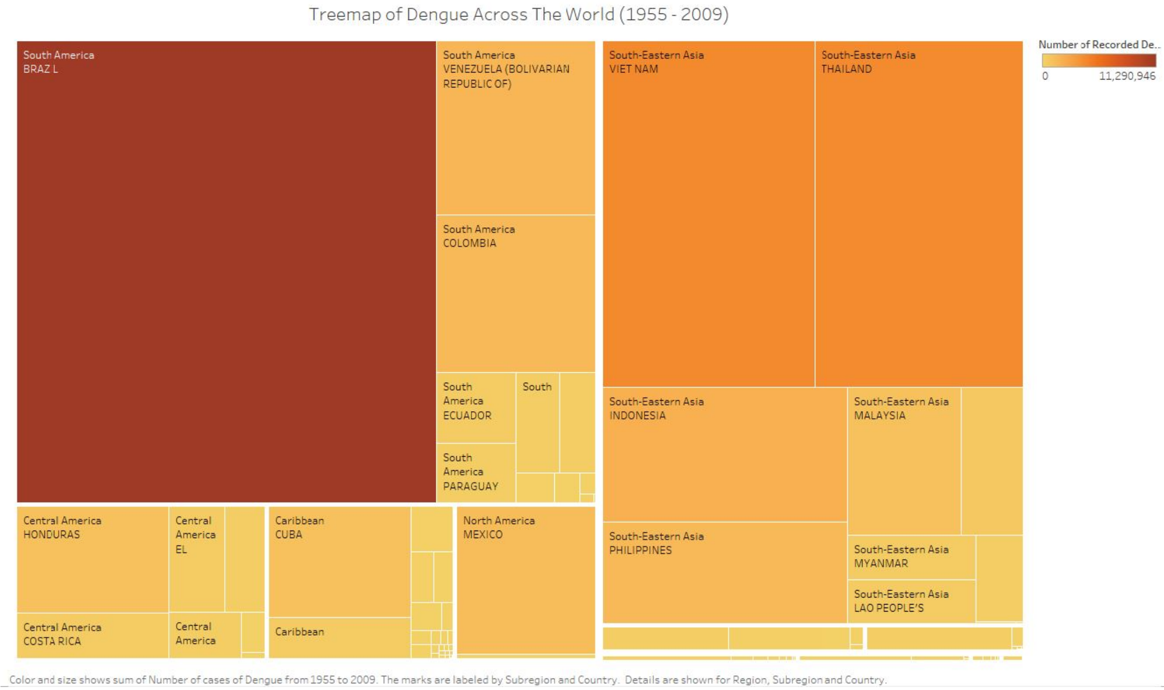
The following treemap shows the distribution of states and their cities according to a total number of cases for different historic diseases. (17)



17. Which of the following diseases does New York not have the highest number of cases? (Q50)

A. Scarlet Fever B. Influenza C. Whooping Couch
D. Tuberculous E. Not Sure

The treemap displays the number of dengue cases observed in countries which are group by region. Size and color are mapped to the number people with dengue. (18, 19, 20)



18. Which country had the second highest number of dengue cases in South America? (Q58)

    A. Brazil B. Ecuador C. Paraguay D. Venezuela E. Not Sure

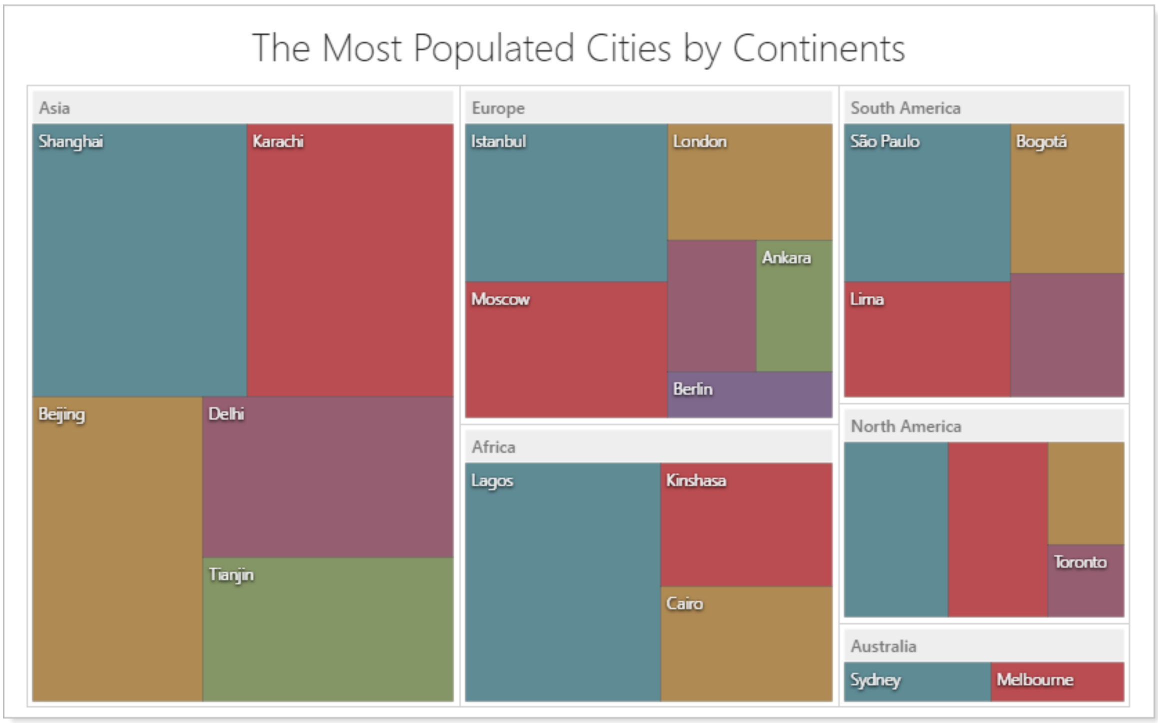
19. To which region does Cuba belong, having the highest rate of Dengue in that region? (Q60)

    A. Central America B. North America C. South America
    D. Caribbean E. Not Sure

20. The total number of people with Dengue in Central and North America is more significant than people with Dengue in Brazil. (Q62)

    A. True B. False C. Not Sure

The following treemap displays the most populated cities by continents. The population is indicated by node size. (21, 22)



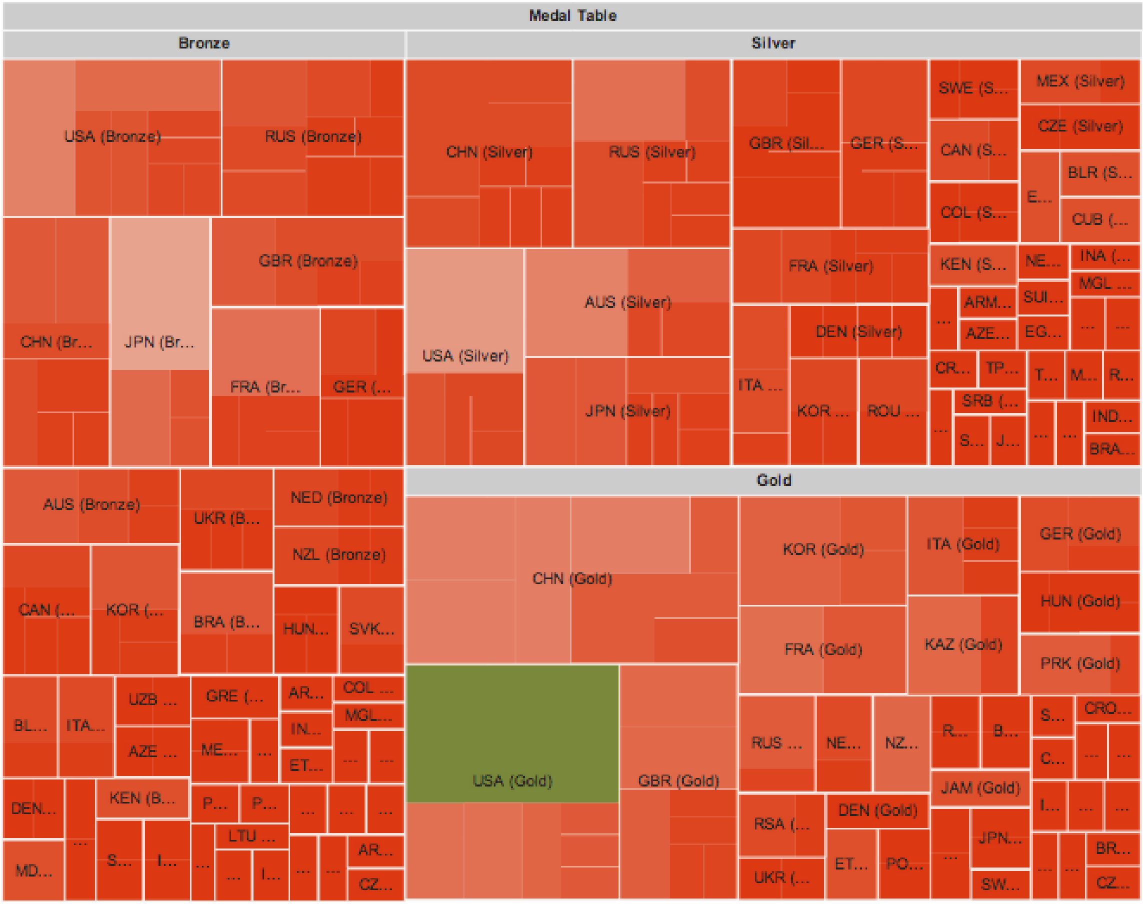
21. Which of the following cities is located within a continent that is neither the least populated nor the most populated? (Q32)

    A. Sydney B. Shanghai C. Delhi D. Istanbul E. Not Sure

22. Which of the following continents has not a higher population than the continent containing Toronto? (Q34)

    A. Africa B. South America C. Europe D. Australia E. Not Sure

This following figure shows a medal standings table for three different medals sorted by countries and sports branch. Size and color are mapped to the number of medals. The number of medals decreases from green to light red. (23, 24, 25)



23. Which of the following countries has more bronze medals than FRA? (Q36)

    A. KOR B. RUS C. JPN D. CAN E. Not Sure

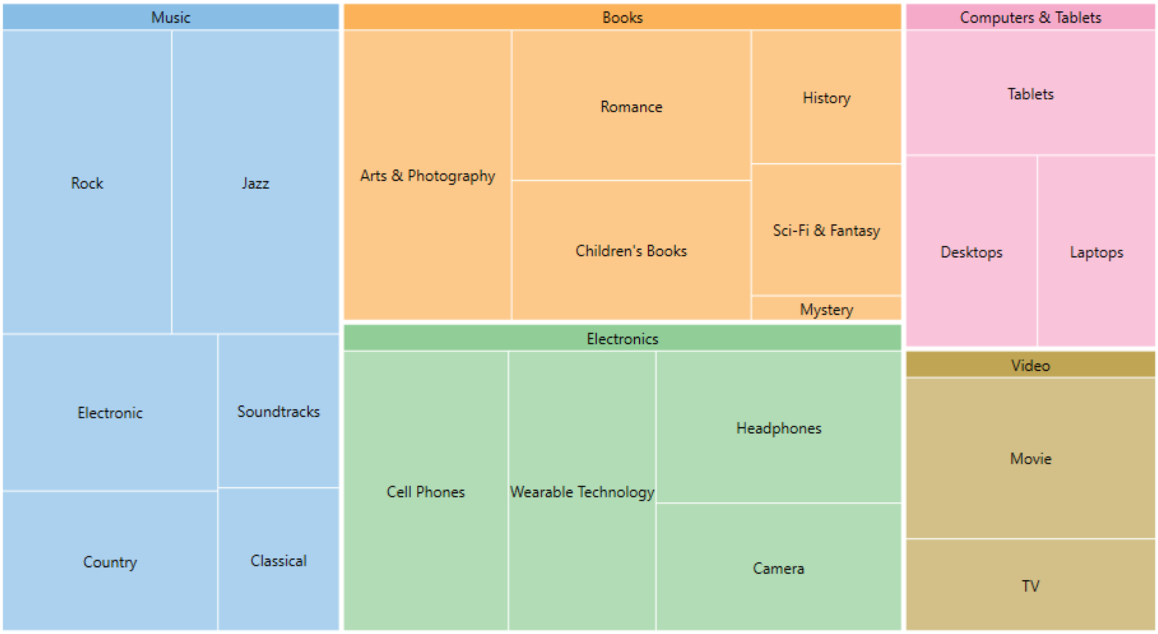
24. In how many sport categories did Russia win silver medals? (Q38)

    A. 9 B. 8 C. 7 D. 2 E. Not Sure

25. Which of the following nation has more gold medals than silver and bronze medals? (Q40)

    A. KOR B. UKR C. CHN D. CAN E.Not Sure

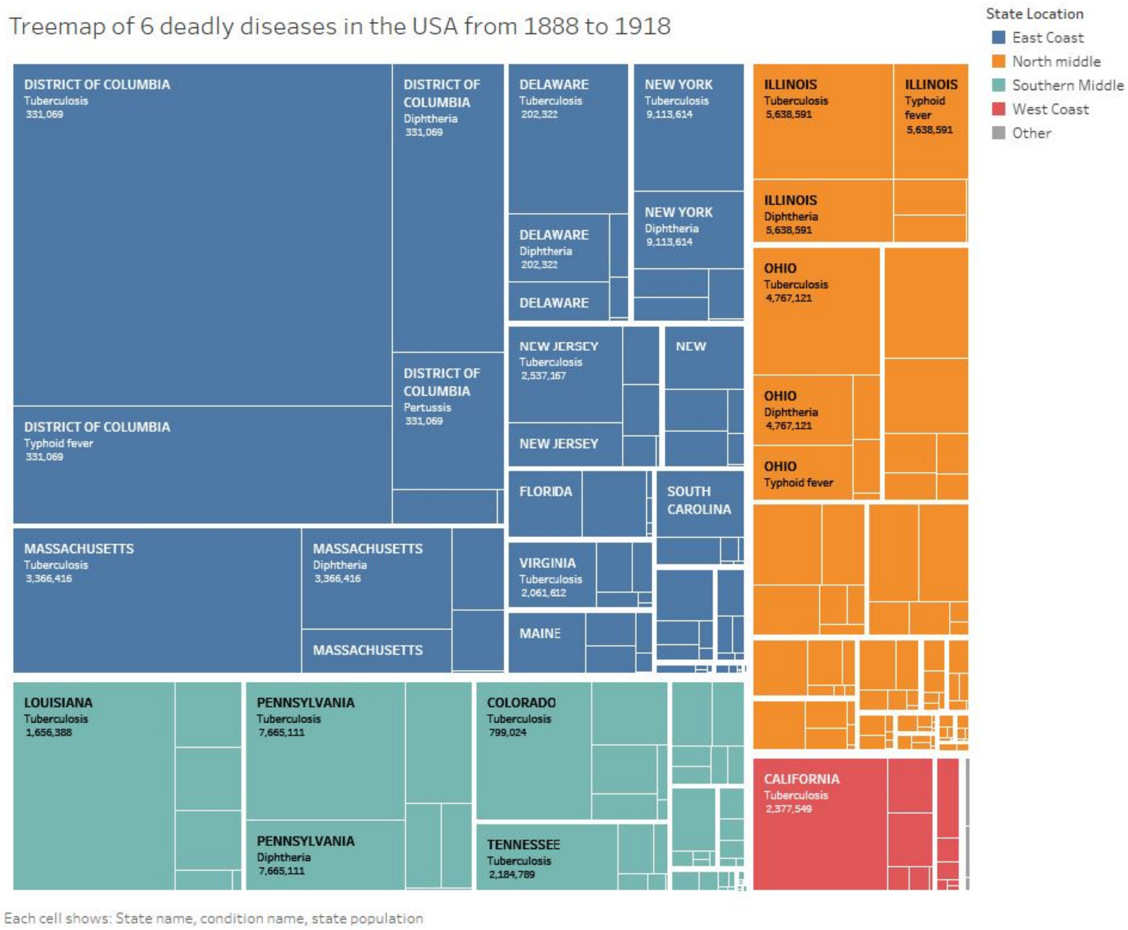
The following treemap displays the distribution of money spent on entertainment. Color is mapped to entertainment types and size is mapped to money spent on each purchase. (26)



26. On which of the following entertainment item was less money spent on than TV? (Q42)

    A. Rock albums B. Mystery books C. Tablets D. Cameras E. Not Sure

This treemap demonstrates historic deadly diseases in the USA. Size is mapped to the number of deaths. (27, 28, 29)



27. Which state has more deaths from Tuberculosis than deaths in Louisiana? (Q44)

    A. Pennsylvania B. Colorado C. California
    D. District of Colombia E. Not Sure

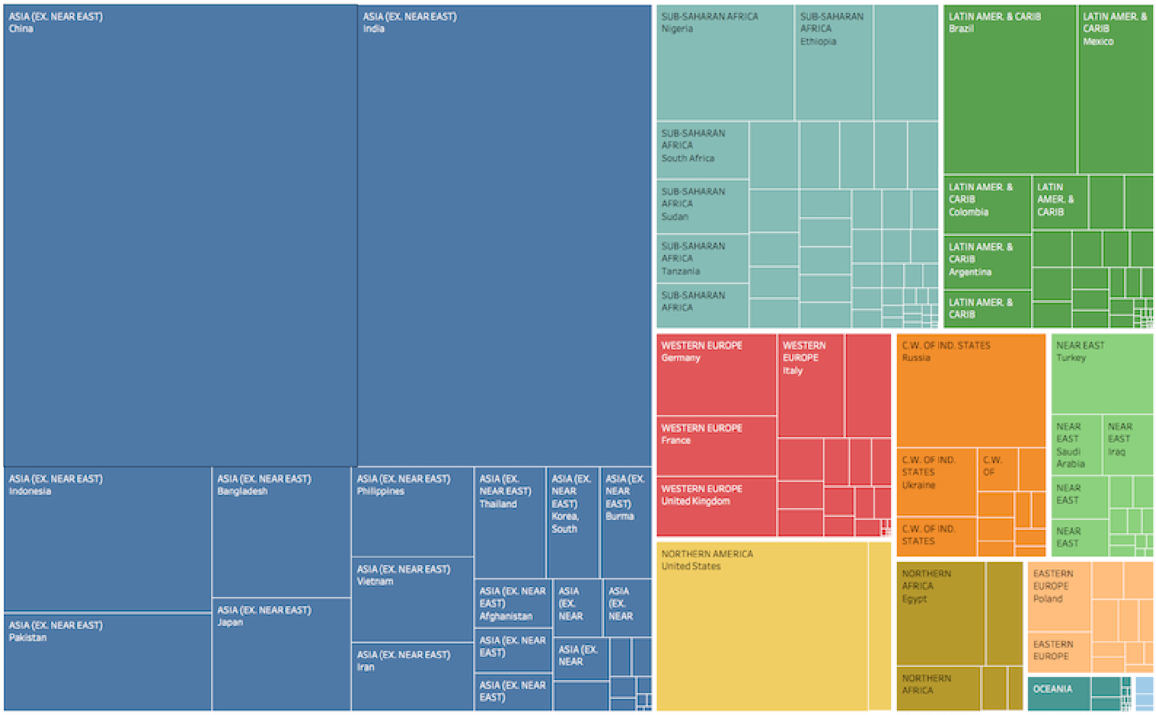
28. Deaths caused by Typhoid fever is more than deaths caused by Diphtheria in Ohio. (Q46)

    A. True B. False C. Not Sure

29. In which of the following east coast states did more people die from Tuberculosis than died in California? (Q48)

    A. New Jersey B. Florida C. New York D. Massachusetts
    E. Not Sure

The treemap displays the population of countries which are grouped by regions.Color indicates the regions and node size indicates the population. (30, 31, 32)



30. Which of the following Asian countries has a larger population than Japan? (Q52)

    A. Thailand B. Pakistan C.Vietnam D. Philippines E. Not Sure

31. The population of the United States is smaller than the population of Russia. (Q54)

    A. True B. False C. Not Sure

32. The number of populations in Asia (Ex. Near East) is nearly equal to the population of the rest of the regions combined. (Q56)

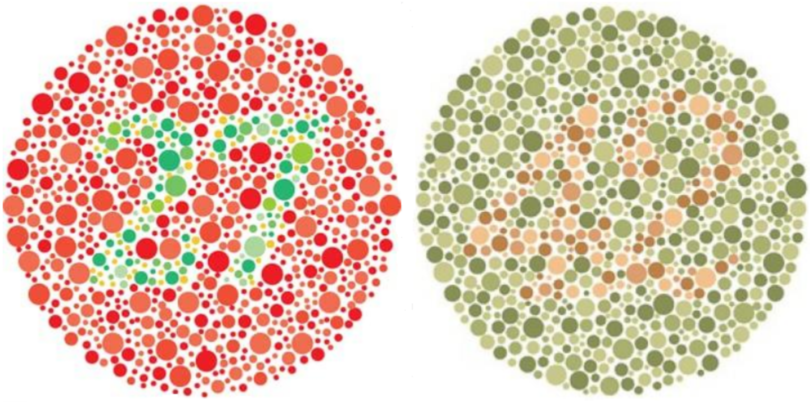
    A. True B. False C. Not Sure

Appendix B

Parallel Coordinates Literacy Test

1. Consent Form (Q1)

2. Thank you for accepting the HIT. Please watch the video for the instructions. (Q2)



3. Which two numbers are shown in the images above? (Q3)

   A. 87-48 B. 27-47 C. 27-42 D. 87-22

4. Which device are you using? (Q4)

   A. Desktop B. Laptop C. Tablet D. Mobile Phone

5. Please write your Amazon Mechanical Turk Worker ID. (Q5)

6. Please check the time right now. Is the current minute an even (e.g. 11.02) or an odd (e.g. 11.01) number? (Q6)

   A. Even B. Odd

7. What is your age? (Q7)

   A. 18-22 B. 23-27 C. 28-35 D. 36-45 E. 46-above

8. What is your gender? (Q8)

   A. Female B. Male C. Prefer not to say

9. Where are you from? (Q9)

   A. China B. Europe C. India D. USA E. Other (Space to type)

10. What is your current education level? (Q10)

    A. High School B. Bachelor's C. Master's D. Ph.D.

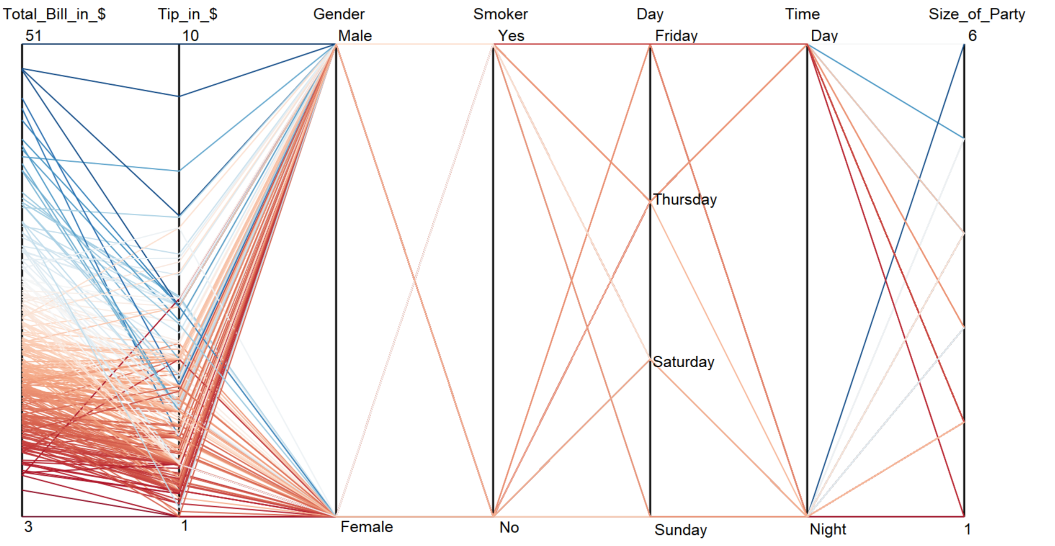
11. What is your English proficiency? (Q11)

    A. Fluent/Native B. Full Professional C. Professional Working
    D. Limited Working E. Elementary

12. Have you seen parallel coordinates before? If yes, where? (Q12)
    (Space to type)

13. Do you have a background in Data Visualization? If so, what is it? (Q13)
    (Space to type)

The image shows tipping behavior in a restaurant located within a shopping mall. The data attributes presented by polylines are total bills amount, tip amount, day and time of tip, the gender, the smoking habit of a worker, and the size of the dining party. (1, 2)



1. How many attributes does the data set have? (Q14)

   A. 5 B. 6 C. 7 D. 8 E. Not sure F. None of the above

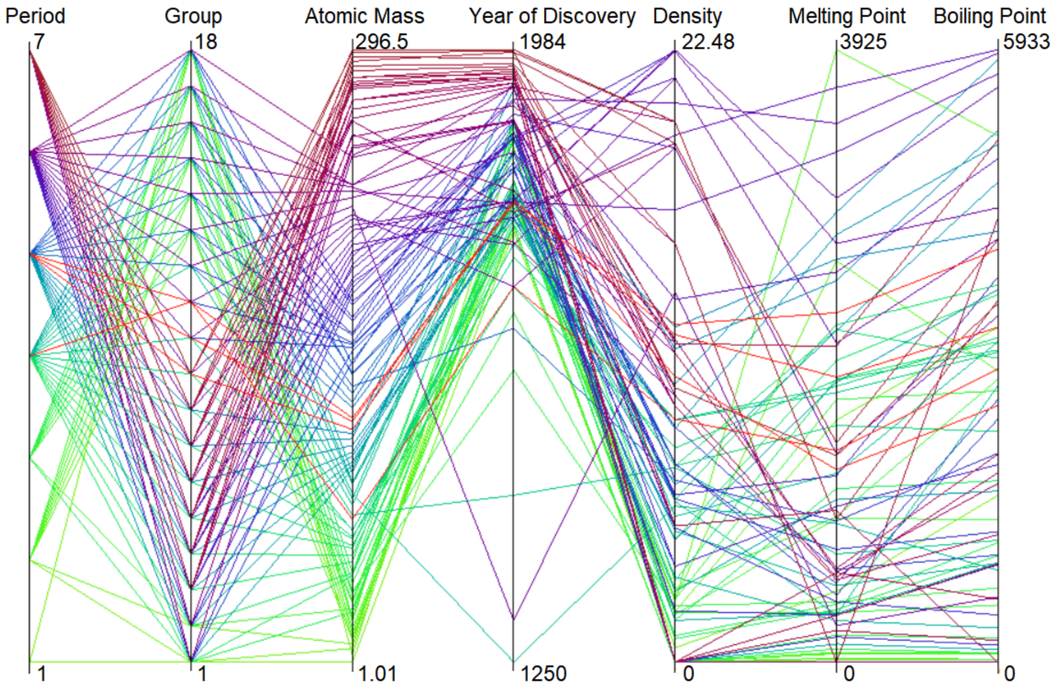
2. On which day of a week was the tipping information not recorded? (Q22)

   A. Friday B. Saturday C. Sunday D. Monday E. Not sure
   F. None of the above

3. What kind of relationship is there between the total bill and tip amounts? (Q78)

   A. Direct B. Indirect C. Not sure D. None of the above

This image displays the distribution of attributes for the periodic table elements. Each polyline represents one element. Each axis is mapped to an element attribute. The attributes of each element that are mapped to axes are the period number, group number, atomic mass, year of discovery, density, melting point, the boiling point. (4, 5)



4. How many parallel axes are there in the image? (Q16)

   A. 7 B. 8 C. 9 D. 10 E. Not sure F. None of the above

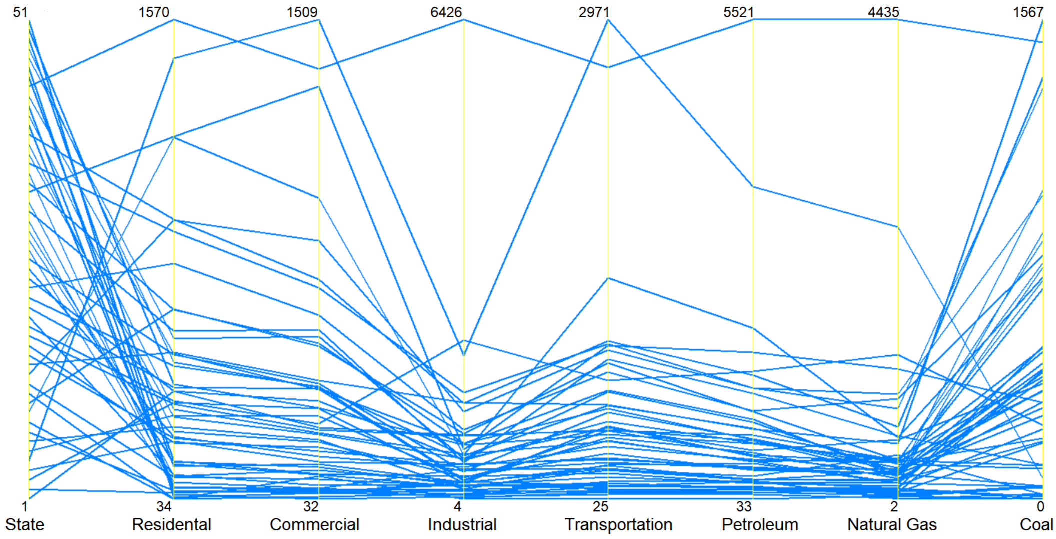
5. What is the approximate maximum value of the boiling point for all elements? (Q42)

   A. 1050 B. 3050 C. 4600 D. 5900 E. Not sure F. None of the above

6. Elements with the greatest atomic mass were found mostly in the 1980s. (Q65)

   A. True B. False C. Not sure

This parallel coordinate plot displays the energy consumption of US states in terms of energy types and sectors that use energy. The attributes of the energy consumption and generation sources presented on the parallel coordinates plot by the state are residential, industrial, transportation, petroleum, natural gas, and coal. (7, 8)



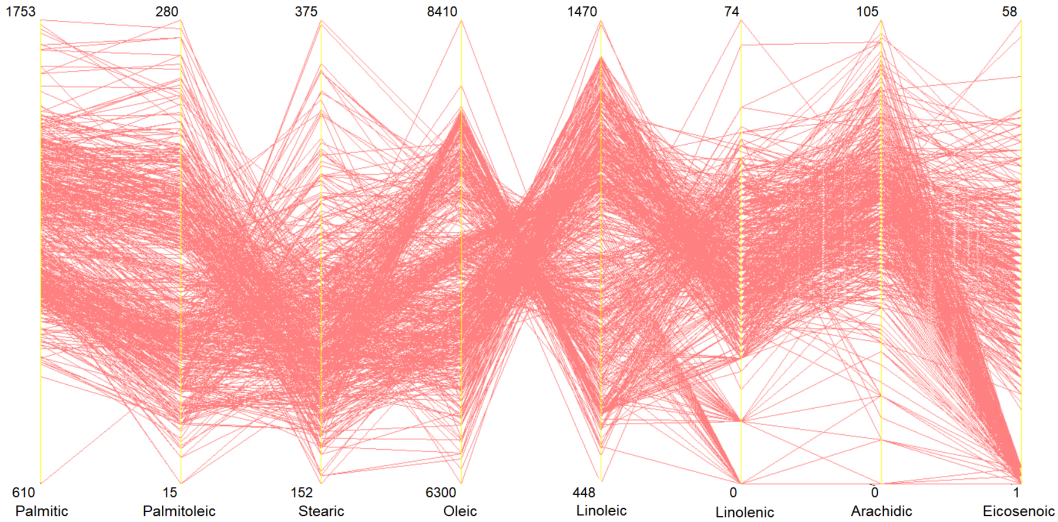
7. What kind of correlation is there between the residential and commercial attributes? (Q18)

   A. Direct B. Indirect C. None D. Not sure

8. What is the commercial energy value for the state which has a 2971 transportation consumption value? (Q69)

   A. 1509 B. 1570 C. 1567 D. 5521 E. Not sure F. None of the above

The image shows chemical measurements from different olive oil samples produced across varying Italian regions. Each olive oil is represented by a polyline and represents measures of the following attributes: palmitic, palmitoleic, stearic, oleic, linoleic, linolenic, arachidic, and eicosenoic. (9, 10)



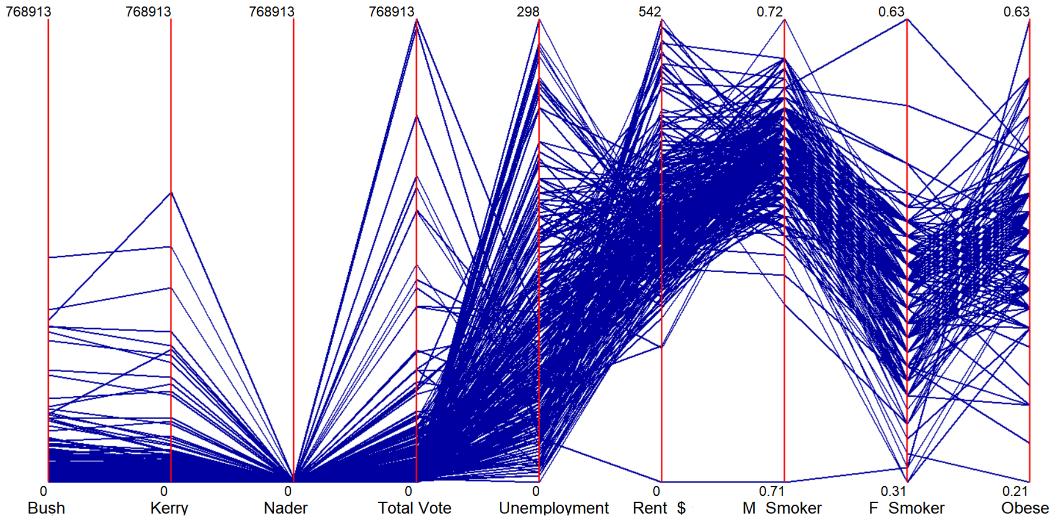
9. What was the maximum arachicid value measured? (Q20)

   A. 105 B. 280 C. 375 D. 1470 E. Not sure F. None of the above

10. Which chemical measurement has an indirect correlation with the attribute of linoleic? (Q71)

    A. Oleic B. Stearic C. Palmitoleic D. Eicosenoic E. Not sure
    F. None of the above

The image shows information from the 2004 US presidential election that includes records of voters' demographics including the names of three political candidates. The data variables are presented as the state name, candidates Bush, Kerry, and Nader, the total number of votes, rent in US Dollars, unemployment, female and male smokers, unemployed, and obesity rates. (11, 12)



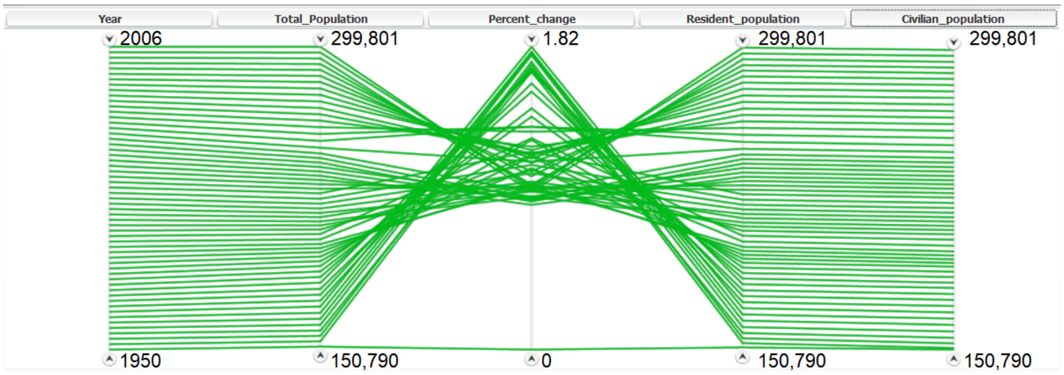
11. Which variable has an indirect correlation with the unemployment rate? (Q24)

    A. Nader B. Rent C. M_Smoker D. Obese E. Not sure F. None of the above

12. Which candidate has the maximum number of votes for an individual state? (Q75)

    A. Bush B. Kerry C. Nader D. Not sure E. None of the above

This image provides information about the US Population from 1950 to 2006 in the thousands. The US Population data variables are plotted on axes such as year, total population, the percent change in the population, resident population, and the civilian population. (13, 14)



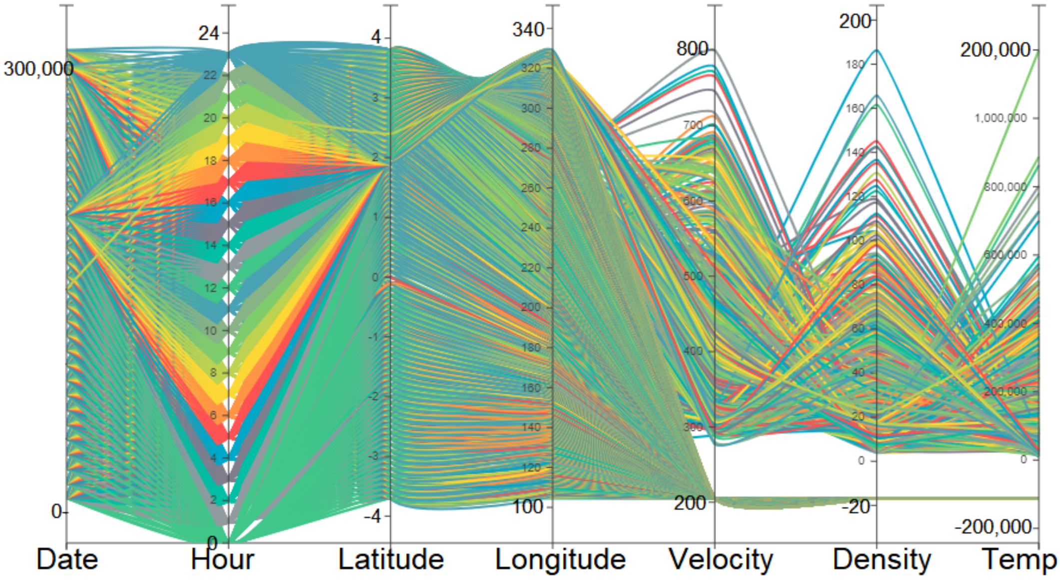
13. The percentage of change started to increase after a certain year. (Q26)

    A. True B. False C Not sure

14. Approximately, what is the maximum percent change over the years? (Q77)

    A. 1.5 B. 1.8 C. 1.9 D. 2.9 E. Not sure F. None of the above

The image plots the characteristics of time-oriented attributes of the planet Venus. The attributes of Venus are presented as date, hour, latitude, longitude, plasma velocity, plasma density, and plasma temperature. (15, 16, 17)



15. What is the approximate maximum value for velocity? (Q28)

    A. 500 B. 600 C. 700 D. 800 E. Not sure F. None of the above

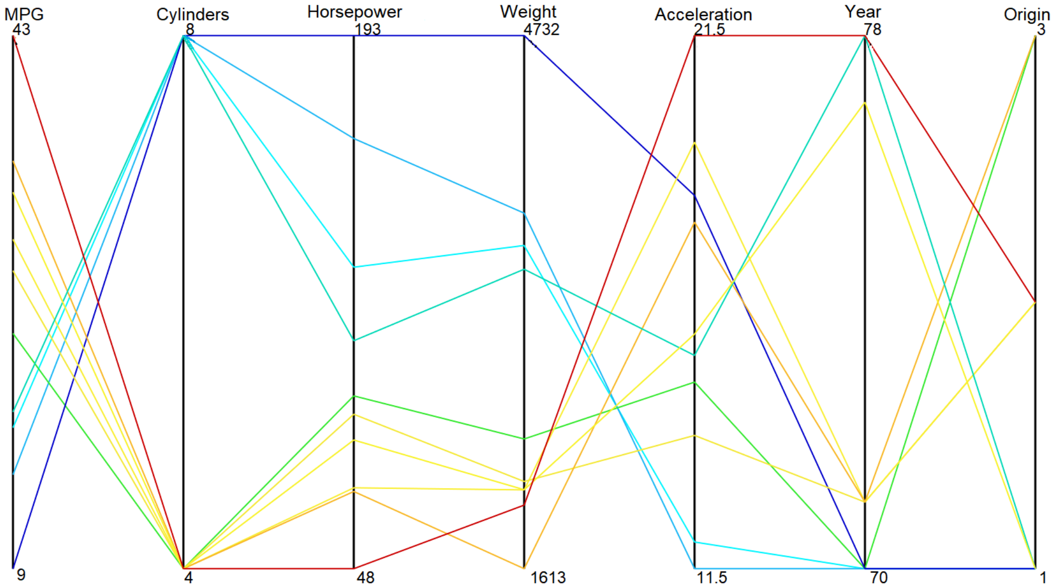
16. How many attributes does the data set have? (Q57)

    A. 4 B. 5 C. 6 D. 7 E. Not sure F. None of the above

17. Which attribute has values that are the most proportionally and widely spread? (Q79)

    A. Hour B. Longitude C. Latitude D. Density E. Not sure F. None of the above

The parallel coordinate plot shows 7 variables for a collection of cars. Each polyline represents a car with attributes as follows: miles per gallon (MPG), cylinders, horsepower, weight, acceleration, production year, and origin. The image displays 10 cars worth of data and each color represents a car type. (18, 19)



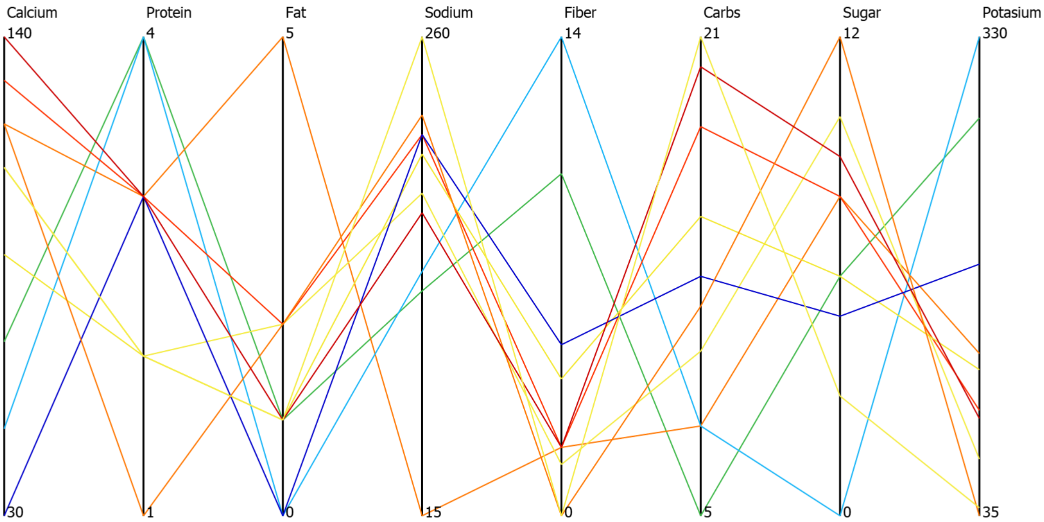
18. Approximately, what are the horsepower value and production year of the car which has 9 MPG value? (Q30)

    A. 4-54 B. 48-54 C. 193-70 D. 1613-70 E. Not sure F. None of the above

19. Which color represents a car with a direct correlation between year and acceleration? (Q81)

    A. Purple B. Red C. Green D. Orange E. Not sure F. None of the above

The image shows the nutritional values of specific cereal products. Each cereal is represented by a polyline and a color on the parallel coordinates plot. The nutritional values are calcium, protein, fat, sodium, fiber, carbs, sugar, and potassium. The image displays 10 kinds of cereal worth of data. (20, 21)



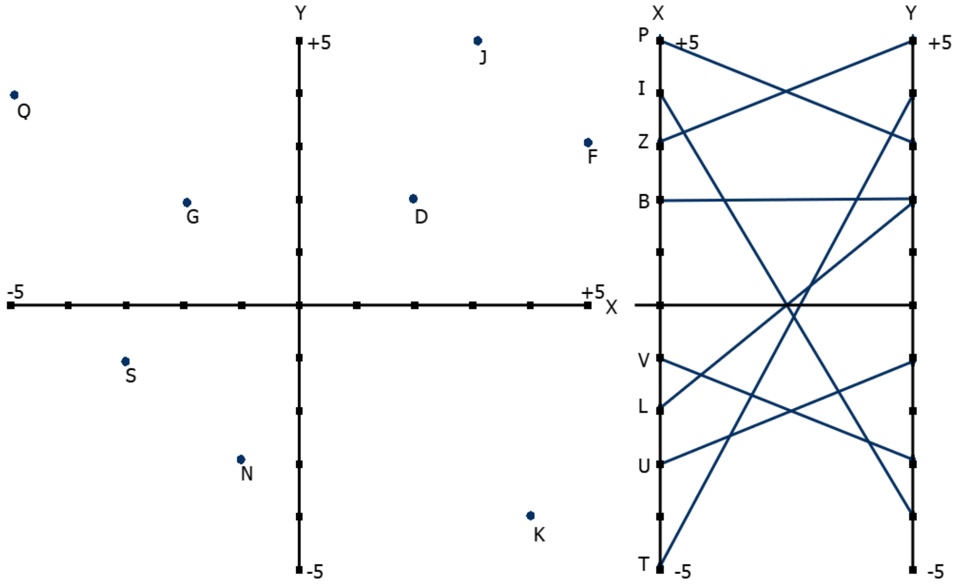
20. Which cereal mapped to color has the maximum value 14? (Q32)

    A. Green B. Red C. Purple D. Blue E. Not sure F. None of the above

21. Which cereal type represented with color has the maximum values of both fat and sugar attributes? (Q83)

    A. Red B. Orange C. Pink D. Blue E. Not sure F. None of the above

This image shows the correspondence between Cartesian Coordinates (left) and Parallel Coordinates (right). Each point and the edge is drawn within the range [-5, +5]. Each point and polyline are labeled with unique letters intended to be used to answer the questions. (22, 23)



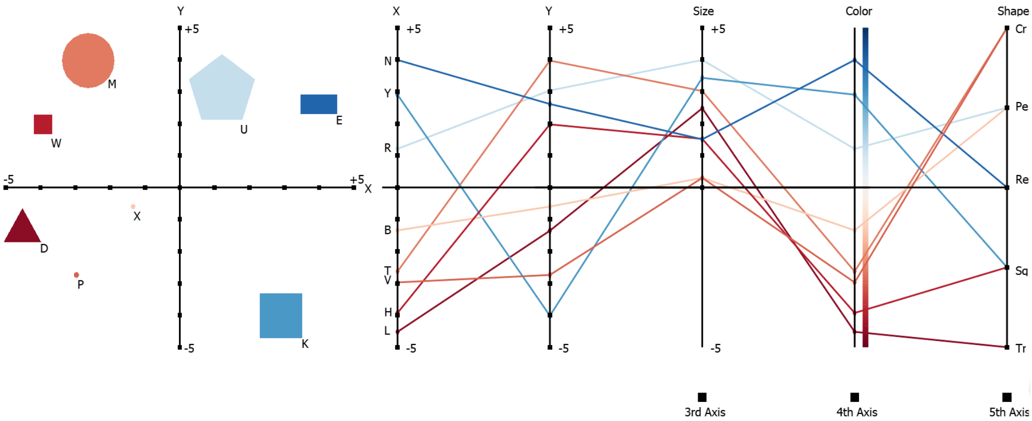
22. Which edge has negative X and Y values? (Q34)

    A. T B. I C. L D. V E. Not sure F. None of the above

23. Using the labels, what is the corresponding point of edge T? (Q85)

    A. S B. Q C. G D. N E. Not sure F. None of the above

This image shows the correspondence between Cartesian Coordinates (left) and Parallel Coordinates (right). Each shape on the Cartesian Coordinates plot is represented by a polyline in the Parallel Coordinates plot. The data variables are X position, Y position, Size, Color, and Shape. Each shape and polyline are labeled with unique letters intended to be used to answer the questions. (24, 25)



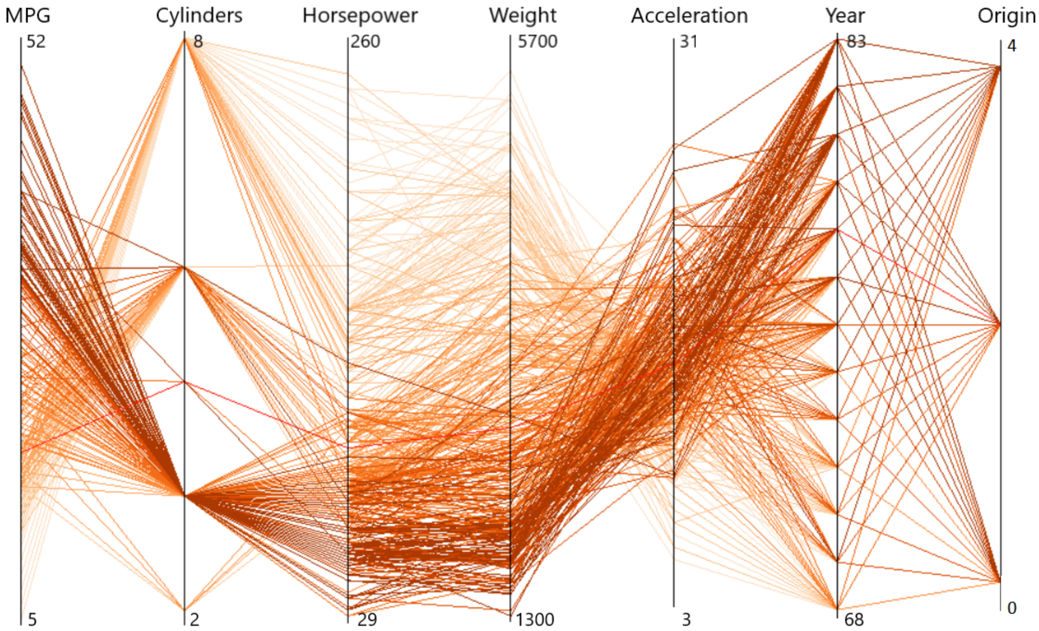
24. What is the corresponding polyline of shape D? (Q36)

    A. N B. H C. L D. B E. Not sure F. None of the above

25. Which shape does the polyline labeled with R represent? (Q87)

    A. Triangle B. Square C. Rectangle D. Pentagon E. Not sure
    F. None of the above

The parallel coordinate plot shows 7 variables for a collection of cars. Each polyline represents a property of a car with attributes as follows: miles per gallon (MPG), cylinders, horsepower, weight, acceleration, production year, and origin. (26, 27)



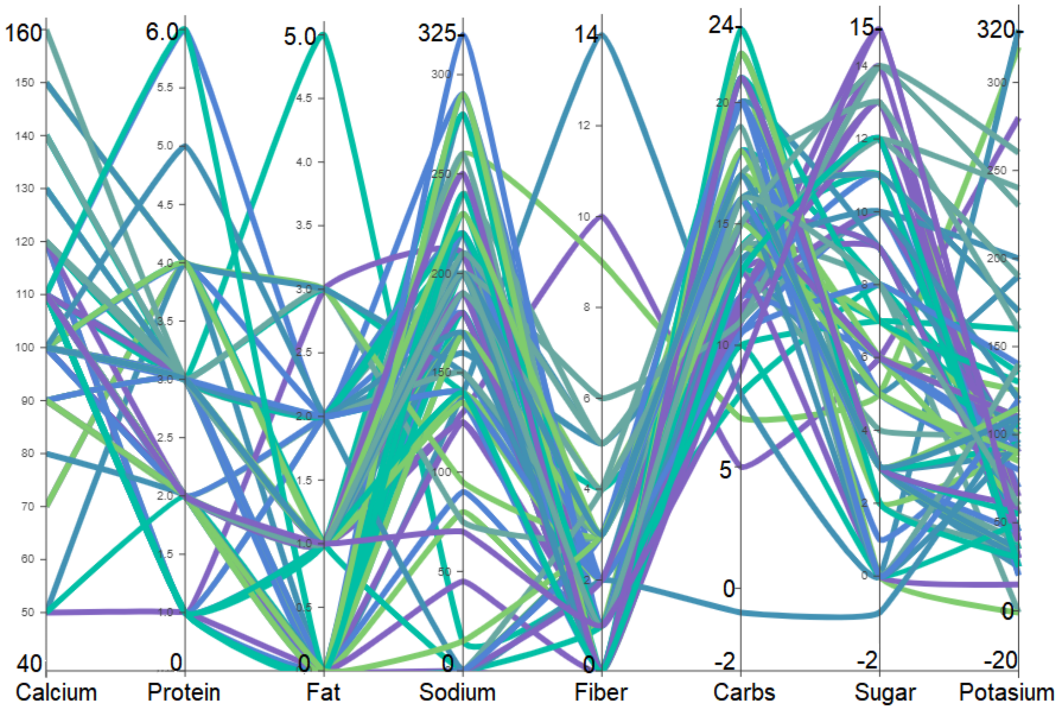
26. Which attribute has an indirect correlation with the car weight? (Q38)

A. Horsepower B. Weight C. Acceleration D. Origin E. Not sure F. None of the above

27. How many distinct car cylinder values are there? (Q61)

A. 3 B. 5 C. 7 D. 9 E. Not sure F. None of the above

The image shows the nutritional values of specific cereal products. Each cereal is represented by a polyline on the parallel coordinate plot. The nutritional values are protein, calcium, fat, sodium, fiber, carbs, sugar, and potassium. (28, 29)



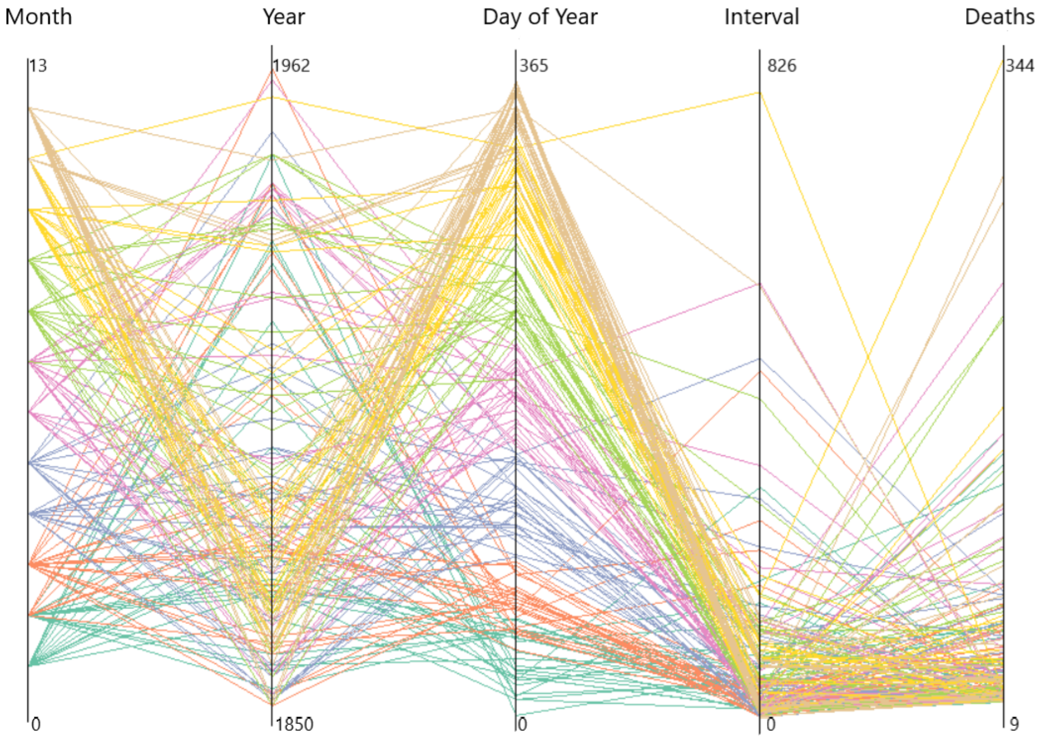
28. For which nutrient do cereals contain the lowest maximum value? (Q40)

    A. Fat B. Sodium C. Carbs D. Potassium E. Not sure
    F. None of the above

29. There is a direct relationship between sodium and fiber in the data. (Q63)

    A. True B. False C. Not sure

The image shows information about coal-mining disasters for over a century. Each polyline represents one coal disaster with information such as a month of the accident, year of the accident, day of the year, intervals between coal mining disasters, and the number of deaths. (30, 31)



30. There were generally more disasters recorded in the earlier years. (Q44)

A. True B. False C. Not sure

31. Approximately, what is the range of values for the year attribute? (Q67)

A. 1830-1952 B. 1850-1962 C. 1910-1955 D. 1810-1975 E. Not sure
F. None of the above

The main purpose of this study is to understand the impact of video in learning parallel coordinates. For this reason, you need to watch the video carefully to the end. After watching the video, you will be asked questions about the video tutorial. If you cannot answer these questions correctly, your answers will not be included in the experiment.

32. Slides Video Tutorial (Q46)

33. Software Video Tutorial (Q47)

    Questions for Attention Check

34. Which topics were mentioned in the video tutorial? (Q49)

    A. Cartesian Coordinates B. Statistical Results C. Bar Charts
    D. Profit from Sales E. Data Attribute Relationship
    F. Colors of Polylines

35. What was the example data about on the last slides of the video tutorial? (Q53)

    A. Diseases B. Flights C. Trading D. Cars E. Not sure
    F. None of the above

36. Which one of the following was a data attribute in the last example? (Q55)

    A. Cylinders B. Price C. Year D. Airbags E. Horsepower
    F.Camera System E.Weight

Appendix C

Covid-19 Simulation Data

The simulation model used is from the Epidemiology, Economics and Risk Assessment (EERA) model [203]. The model incorporates an inference process to estimate the range of parameters of interest and the ranges of parameters to extract parameter configurations. In this case, there are 160 parameter configurations. For each configuration there are multiple simulation runs. In this case, 1000 runs result in different predictions.
The model takes the same set of input parameters, called simulation configurations that yield different output results for each run. The model aims to provide the range of output possibilities for each possible prediction. For each output result, minimum, maximum and mean values of output parameters are provided.
For the model, there is a long list of parameters, some are inferred, some are estimated a priori, and some are fixed across runs. Here are the critical input parameters:

- **nsse_cases:** Normalised sum of square error for the number of cases
- **nsse_deaths:** Normalised sum of square error for the number of deaths
- **p_inf:** Probability of infection
- **p_hcw:** Probability of infection (Healthcare worker)
- **c_hcw:** Mean number of healthcare worker contacts per day
- **d:** Proportion of population observing social distancing
- **q:** Proportion of normal contact made by people self-isolating
- **p_s:** Age-dependent probability of developing symptoms
- **rrd:** Risk of death if not hospitalized
- **lambda:** Background transmission rate

For each age group (8 age groups) there are;

- 200 days of predicted time-series of each output data dimension in the model
- 16 distinct output data dimensions (see the list below)

The model generates a number of output files for each run. In total, 160 (parameter configurations) $\times$ 16 (data dimensions) $\times$ 1000 (runs) $\times$ 8 (age groups) = 20,480,000 time series of 200 days each. The data we display by default is the first configuration. The output simulation parameters are as follows:

- **Age Group:** Age groups ID are used in the model.
- **Day:** The day for the record
- **S:** Number of susceptible individuals (not infected)
- **E:** Number of infected individuals but not yet infectious (exposed)
- **I:** Number of infected and infectious asymptomatic individuals
- **IS:** Number of infected and infectious symptomatic individuals
- **H:** Number of infected individuals that are hospitalized
- **R:** Number of infected individuals that are recovered from infection
- **D:** Number of dead individuals due to disease

The age groups ID as used in the model are here:

- **Group 1:** Under 20
- **Group 2:** 20-29
- **Group 3:** 30-39
- **Group 4:** 40-49
- **Group 5:** 50-59
- **Group 6:** 60-69
- **Group 7:** 70+
- **Group 8:** Health Care Workers

Appendix D

Inclusivity for Visualization Education: A Brief History, Investigation, and Guidelines

D.1 Introduction

Diversity and inclusion have become more and more important in several fields. The main goal of this study is to address the topic of diversity and inclusion in the visualization community. In particular, we aim to give an overview of studies on diversity concerning spatial cognition and improve issues related to this subject. To get a better understanding of the impact of possible bias in visualization education, we investigate gender bias in data visualization class and evaluate the computer science students' scores. Additionally, we provide recommendations on how to make the visualization classroom more inclusive, supporting diversity and inclusion. The contributions of this paper are:

- A compilation of the historical study of gender bias in spatial cognition.
- The reporting of student scores from a data visualization class, focusing on gender bias.
- To provide a list of recommendations for visualization teachers, helping cater to a more diverse and inclusive classroom.

We start with a brief history of research on spatial cognition diversity with respect to gender. This is followed by a small investigation of evidence of gender bias in a data visualization course. This is followed by recommendations for inclusion in a visualization classroom.

D.2 A Short History of Diverse Cognition

The study of spatial cognition and perception with respect to gender has a long history dating back to the 1940s. Spatial ability is the capacity of imagining the shape of an object, its dimensions, co-ordinates, aspect ratio, movement, and geography. Picturing an object being rotated in space, turning around an obstacle and seeing things from a three-dimensional view can be included in the description of spatial ability [204].

1940s: In 1943, O'Connor uses a performance test called the "Wiggly Block" and discovers that almost 25% of females outperform the average of males in spatial

ability [205]. The earliest research paper reference we found on the topic dates back to 1943.
1950s: Guilford [206] designs seven tests including spatial visualization, spatial orientation, and perceptual speed to assess a diverse tendency factor or primary mental ability. Results indicate that scores for females are significantly lower than scores of males in spatial visualization and spatial orientation. In contrast, females perform better than men in a perceptual speed test however the findings are not conclusive in that participants represent their gender.
1960s: In one study from 1960s, 104 males and females are given the Identical Blocks Test, as a standard test of spatial visualization, developed by Stafford in 1961. Results suggest that a gender-linked recessive gene may affect spatial abilities. Also, test scores demonstrate that females' scores are significantly lower than males [207].
1970s: The right hemisphere plays a key role in spatial, holistic cognitive processing, and handles visual and tactile spatial processing skills. An experiment with 200 boys and girls between 6 and 13 years of age demonstrates that the right hemisphere of boys is powerful in processing non-linguistic spatial information by about age 6 [208]. Conversely, the right hemisphere in girls is not spatially developed even by the age of 13. These outcomes indicate that boys have greater hemisphere specialization and there is a gender dimorphism in the neural organization related to spatial cognition. The outperformance of males to females on many spatial tests is related to neural dimorphism. Spatial ability is connected to gender chromosomes and testosterone. Genetic and hormonal agents are the factors of neural dimorphism for the two gender [208].
Research by Waber [209] indicates that gender difference in adolescence has benefits for males with respect to spatial ability. He reveals that early maturers score lower on spatial ability tests than late maturers. Since puberty occurs earlier in females than in males, adolescent females are expected to have poorer performance on the spatial test.
1980s: Sanders et al. [210] indicate that male performance is higher on a task that required subjects to mentally rotate three-dimensional arrays of cubes. Gilmartin and Patton [211] test college students and undergraduates to study the student skills with cartographic illustrations conveying geographic information. The gender based differences are observed in the younger age groups, where males outperform females and map use scores for females and boys are almost similar between college students. Research by Gilmartin [212] investigates the effect of mental imagery on recall of spatial information and whether there is gender-based differences in ability to employ such a visualization technique. Subjects in three different groups are given maps with geographic text, illustrated with text and mental images of the text to read. Outcomes of the research indicate that gender has a significant impact on recall of spatial relationships where men score higher than women on reading the text with maps.
Dr Camilla Benbow, a psychology professor at Iowa State University, scanned the brains of more than a million males and females to examine their spatial ability and reports the distinction between gender is visible by the age of four. She finds that while females are successful at perceiving two dimensions in the brain, males have a better ability to perceive the third dimension [204].
Dr Benbow and Dr Stanley [213] test a set of talented children and discover that males outperform females at spatial mathematics by 13 to 1. Males can build a block building from 2D plans faster and easier than females. Males can predict angles precisely.
Beatty and Tröster [214] perform a survey of 1800 undergraduate students indicating

that males could locate targets on a US map more accurately than females. A group of 202 male and female college students are tested by Chang and Andes [215] to investigate gender differences when reading reference and topographic maps. Results indicate that males outperform females in reading reference and topographic maps. The gender difference in mental rotation ability was, in 1989, the largest cognitive gender difference documented in the literature [216].

D.3 Diverse Spatial Cognition in Popular Literature

We highlight some findings from a popular book [204] on the subject of spatial cognition with respect to gender.
1990s: A perspective on gender differences in spatial ability is published by Goldstein [217]. He supports that males outperform females on spatial ability tests. However, his study also indicates that females' spatial skills may be masked due to lower confidence in their ability. If the confidence level of females could be lifted through experiments, their performance on a timed test might be identical to that of males. Males outscore females in spatial ability by a ratio of 4.1 on three-dimensional video tests. Spatial ability also enables a man to rotate a map in his mind and understand directions. The spatial field in his brain can store this information for future events. Research indicates that a man's brain can measure speed and distance to understand when to change direction [204].
Reading maps and understanding current location are related to spatial ability. Brain scans demonstrate male's most powerful ability spatial ability is situated mainly in the right brain of men. Spatial ability is located in both women's left and right brain hemispheres. However, most women have limited spatial ability. Only 10% of them have spatial abilities as dynamic as men according to Pease *et al.* [204].
The anatomy of Albert Einstein's brain was examined and compared to the preserved brains of 35 men and 56 women with average intelligence by Dr Sandra Witelson with other scientists at McMaster University [208]. It was found that the spatial area of Einstein's brain, connected to his mathematical skills, is 15% larger on both sides than in average men and women.
In males, the right brain improves faster than the left brain and develops more connections with-in itself and fewer connections with the left brain. In females, both sides of brain improve at a equal speed which provides females with a more diverse range of skills. They also have more connections between left and right through a thicker corpus callosum which results in a tendency of females to be more ambidextrous than males. Many more women have difficulty identifying their left hand from their right.
Pease *et al.* [204] state, "Testosterone hormones inhibit the left-brain growth in boys as a trade-off for greater right-side development, giving them a better spatial ability for hunting. Studies of children between the ages of five and eighteen show that boys outstrip girls in their ability to move a beam of light to hit a target, reproduce a pattern by walking it out on the floor, assemble a range of three-dimensional objects and solve problems requiring mathematical reasoning. All these skills are located mainly in the right brain of at least 80% of men and boys".
Dr D. Wechsler develops IQ tests to remove sexual bias against men or women during spatial tests. People from cultures of primitive races to developed city residents are examined. Findings indicate that although women have slightly smaller brains, women exceed men in intelligence being around 3% smarter than men. When participants are asked to solve maze puzzles, men outperform women, scoring 92% without factoring in culture [204].

The British Cartography Society declares that half of its members are female, and maps are designed and edited by women as well. British cartographer Alan Collinson states, "Map design is a two-dimensional task in which women are equally as capable as men". Alan Collinson adds that the most women face challenges when reading and navigating with the map because 3D perspective is required to navigate a route. Tourist maps have a 3D perspective. Thus trees, mountains, and other landmarks are featured on such maps which support women to perform better when reading the maps. His study also indicates that men have the ability to cognitively turn a 2D map into a 3D view in their mind, but most women don't seem to be able to do this [204].

Woman develop the spatial ability on both sides of the brain thus this may interfere with speech function. If a woman is given a street directory, she will stop talking before she turns the guide around. Whereas men will maintain speech. However, he will turn the radio off because he can-not manage to auditory processing while he is engaging his map reading abilities [204].

Practice and recurrence of a task increase brain connections. As an example, brain mass of a retired person who has nothing to do decreases over time. Whereas active intellectual interests protect brain mass and even improve it [204].

D.4 Diverse Spatial Cognition in Visualization Literature

The impact of cognitive abilities on the understanding of visual designs and what features of the visual designs influence users' understanding are considered unsolved problems. Velez *et al.* [26] aim to study spatial ability in a varied subject group to examine subjects' perceptions of complex visualizations. Velez *et al.* [26] contribute a basic visualization test experiment to assess comprehension and difficulty when visualizing spatial ability and report several outcomes from the experiment.

D.4.1 Challenges in Visualization Understanding

Classic 2D visualizations feature 2D slices and 3D volumetric orthographic projections. Research indicates that projection and slice-based visualizations are not ideal for shape understanding [218] and comprehension of 3D space layout [219]. Therefore, combining 2D and 3D methods, cross sections or orthogonal projections unified with 3D position references are suggested.

Spatial ability is interrelated to skills involving the retrieval, retention and conversion of visual information from a spatial source [220]. Researchers sub-divide the concept of spatial ability into more specific factors to help understanding. Six spatial factors are described by Kimura [221] and can be identified by experimental assessment. These are: spatial orientation, spatial location memory, targeting, spatial visualization, disembedding and spatial perception.

D.4.2 Visualization Comprehension

Rizzo *et al.* [222] studies subjects that have difficulty with the Shepherd and Metzler [223] mental rotation test. Some users can develop their mental rotation ability through training in a VR environment. Gender differences in 3D virtual environment navigation is studied by Czerwinski [224], [218]. Their results indicate that larger displays and a broad field of view enhanced female performance in navigation tasks and was comparable to male performance.

An experiment is designed to assess a person's ability to understand visual designs

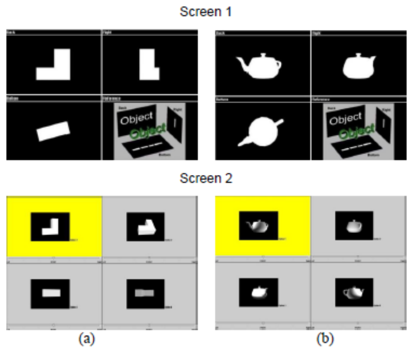


Figure D.1: *Two object examples used in a visualization test. The figures in Screen 1 show the orthogonal projections and Screen 2 shows four possible answers. The correct answer appears highlighted [26]. Image courtesy of Velez et al. [26]*

and define how this ability is related to spatial skills. Velez *et al.* [26] describe an experimental method that focuses on fundamental visualization tasks by designing a simple visualization test that requires participants to form a mental picture of a 3D object based on its 2D projections. The test design is simple enough for inexperienced users and resembles standard spatial ability tests. The goal is to understand what makes a 3D visualization challenging based on the features of the objects and their visualization presentation.
A total of 56 students, half of them female, aged between 18-31 studying or graduated from US University participate in the study. Each experimental session takes approximately two hours. In the first hour, participants are given five paper-based cognitive factor tests. After the paper tests, computer-based visualization tests are administered. Subjects are seated in front of desktop computers on which the orthogonal projection test is administered and are expected to answer 38 questions. Figure D.1 provides two examples used in the visualization test.

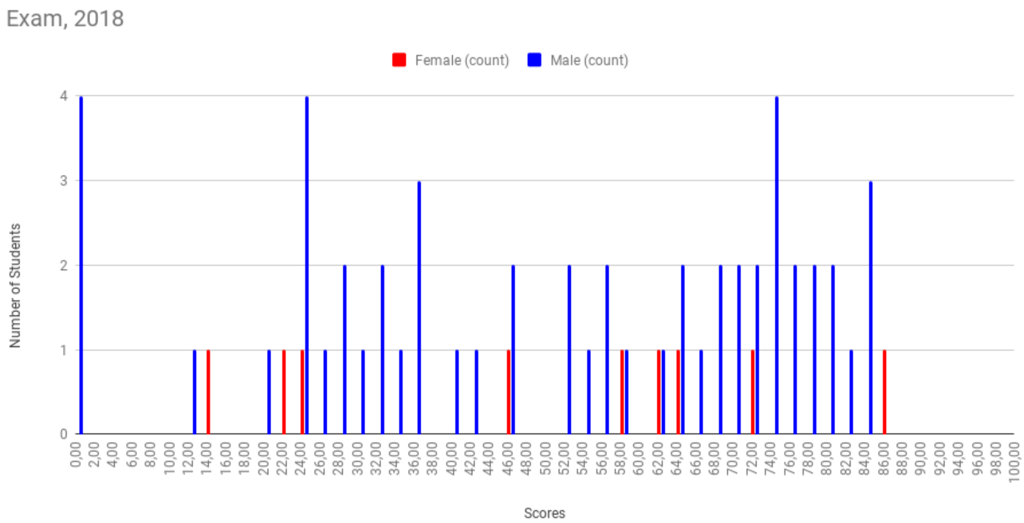


Figure D.2: *Histogram of the Data Visualization Exam scores in 2018 includes the average score for males (50) and females (50). Gender is indicated.*

Data from the experiment is analyzed based on gender for visualization ability for each of the spatial skill tests. The result indicates that males perform distinctly better on the visualization test than females. Velez *et al.* [26] determine that geometric objects, the number of original and hidden surfaces, edges and vertices relate to accuracy, and that low spatial ability participants are able to understand

only basic geometrical objects such as cubes and cones.
Velez *et al.* [26] summarize some results from the study.

- Spatial ability diversity in the population is quite large. They argue that they examine a large enough subject pool.
- Understanding a visualization is related to spatial ability. The level of spatial skill can help to explore the reasons behind comprehension problem by comparing distinct comprehension mistakes. Spatial ability can also be used to classify the population so that visualizations can be customized for different spatial ability groups.
- Time is not related to visualization accuracy. Time to understand projections does not affect accuracy in a visualization test. Using time to assess the properties impact on understanding of visualization is not recommended.
- The number of geometric properties influences visualization accuracy and examination time. This result indicates that if an animation is aimed for the general population the speed of animations for complex objects presented can be a strong influence.
- The hidden geometric properties in the visual representation of object influence the accuracy of visualization. Hidden objects in the visualization make the image difficult to understand for a user with low spatial ability. Rotating the object will enhance understanding for a user.
- Small rotation is difficult to detect. To identify small changes of rotation with small angles when comparing two objects is a difficult task.

D.5 Investigating Evidence of Gender Bias in a Data Visualization Class

Surprised and interested in the gender-based diversity in spatial cognition research literature, we decided to look for evidence of gender bias in the data visualization module at Swansea University, Wales. The Data Visualization module has been taught to third-year undergraduate and master level students since 2006. The course includes two lecture hours and a one-hour lab session each week during the semester. Many students enrolled in this course are from overseas countries, and most of the class consists of male students, about 85% of the class, as do all computer science classes at Swansea University.
The assessment consists of one exam at the end of the semester, and two assignments, one of which is focused on information visualization and the other on scientific visualization. Students are provided with a data set and asked to create and explain at least five unique visual designs using existing data visualization tools for the first assignment. For the next assignment, students must modify source code given by the lecturer and use this code to produce volume visualizations with the help of existing volume renderers and describe how they obtain the visual representations. In addition to the programming aspect, they also use a volume data set provided to them to use volume visualization software such as ParaView [225] and Inviwo [226]. Both assignments are marked by focusing on students' visual designs and the description of each image.
We aim to investigate for evidence of gender bias and its impact on students in our class by comparing scores for both genders. We produce six histograms for this investigation. Three histograms consist of the 2018 exam and coursework assessment

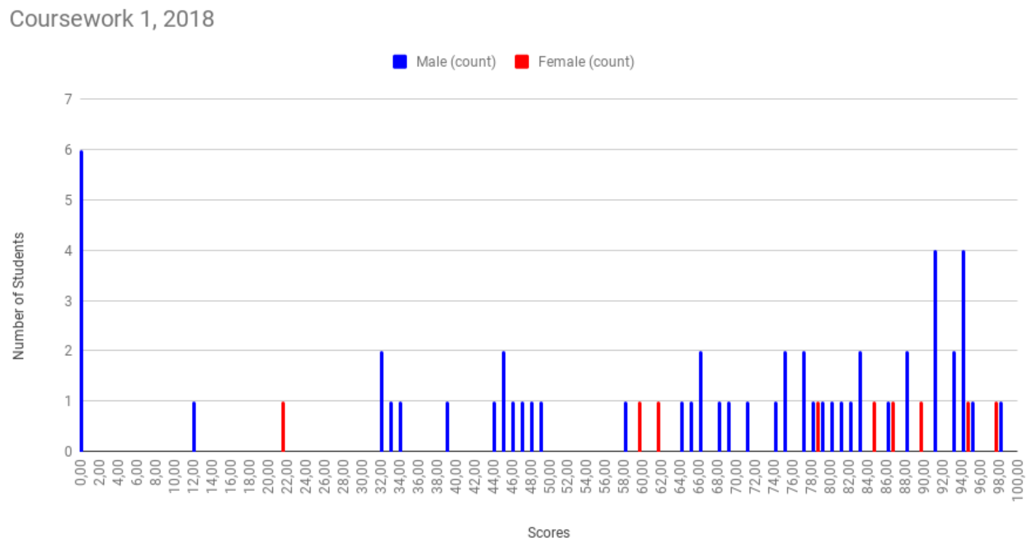


Figure D.3: *Histogram of the Data Visualization Coursework 1 in 2018 including the average score for males (64) and females (74). Gender is indicated.*

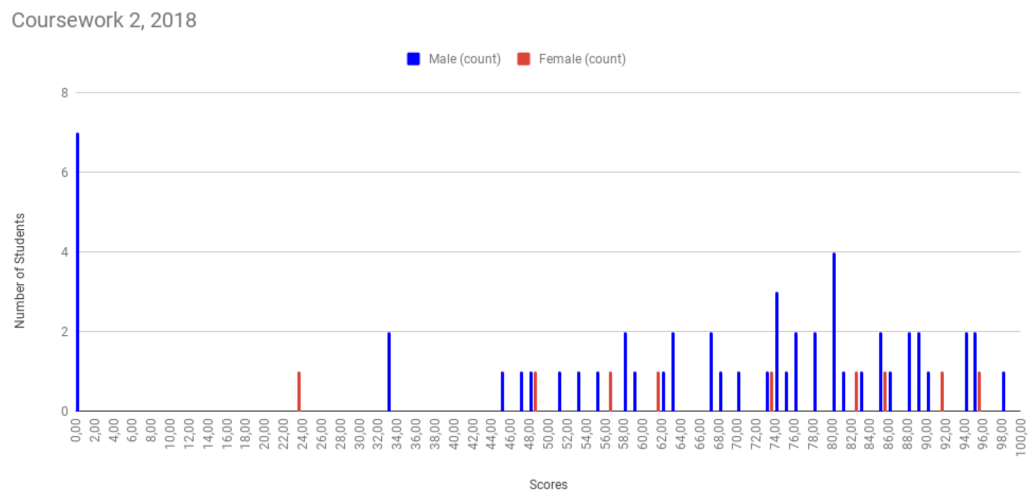


Figure D.4: *Histogram of the Data Visualization Coursework 2 in 2018 includes the average score for males (63) and females (68). Gender is indicated.*

results. The other three histograms represent five years' worth of exam and assignment assessment results from 2013-2017.

2018 Exam Results

We classify students according to gender. The histogram for 2018 exam results (See Figure D.2) displays the minimum and maximum exam scores as 0 and 86 respectively. Only seven students, one of them female, score higher than 80 (In the UK, scores above 70 are considered excellent.).
Similar to higher scores, seven student scores are lower than 20 with four zeros by male students. Much to our surprise, our initial investigation indicates that exam results appear to be evenly distributed between male and female grades from 0 to 86. However, we do notice that a binomial distribution with a dip in the middle with a score of 50. Upon further investigation we noticed a different kind of bias based on language (Asian vs English). This cannot be seen in the histogram but was observed first hand by the teacher.

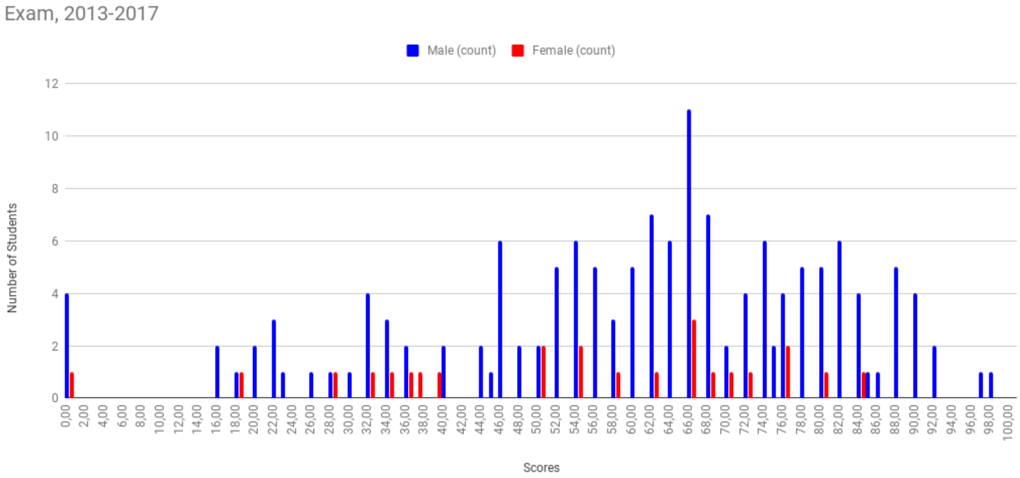


Figure D.5: *Collective histogram of the Data Visualization Exam from 2013-2017 includes the average score for males (60) and females (53). Gender is indicated by color.*

2018 Information Visualization Assignment

The next two histograms reflect the distribution of the first and second assignment grades. For the information visualization assignment, the females and males perform similarly, and their scores are very close include the average score for males and females. One female and one male student have a score of 98 while six male students receive 0 on the first assignment. Zero indicates that a student does not submit any assignments (See Figures D.3, D.4).

2018 Scientific Visualization Coursework

Male and female student score better results than on the first assignment. Almost half of the class scores higher than 80 for both genders with some at almost 100% while there are seven 0 scores by male students. None of the graphs for 2018 indicates any gender-related differences.

2013-2017 Exam Results

Furthermore, we create three more histograms over a five-year period for the exam and two coursework for the years 2013-2017 in order to find evidence of gender bias. The histogram for five years' worth of exam results indicates that the distribution of scores is very similar for both genders. Similar to the histograms for 2018, not many female students enrol in our module over these five years. There are 24 females and 148 males enrolled in the Data Visualization module over the 5 years. However, exam results of female and male students show a similar tendency, and they have a similar distribution of exams scores from 2013 to 2017. We observe a peak at a score of 66 with eleven students, and five students receive 0 while only two students receive 98 in five years (See Figure D.5).

2013-2017 Coursework Results

In addition, we have two histograms for coursework covering information and scientific visualization over five years (See Figures D.6, D.7). The histogram representing the first assignment demonstrates remarkable results for 19 students who score 0 and five students scoring 100. Another outcome shows a score of 81 with seven male and one female student. The distribution of male and female student grades in the histogram is very similar to the distribution of the first course-work in 2018. A large subset of the student population performs well since their scores are mainly higher than 50. We also analyse the results of the second assignment over the same period. Compared to the exam and first coursework over a five-year period, the second coursework result is at a higher level, and students receive full marks. Nine males and two females receive the full mark on the scientific visualization assignment while

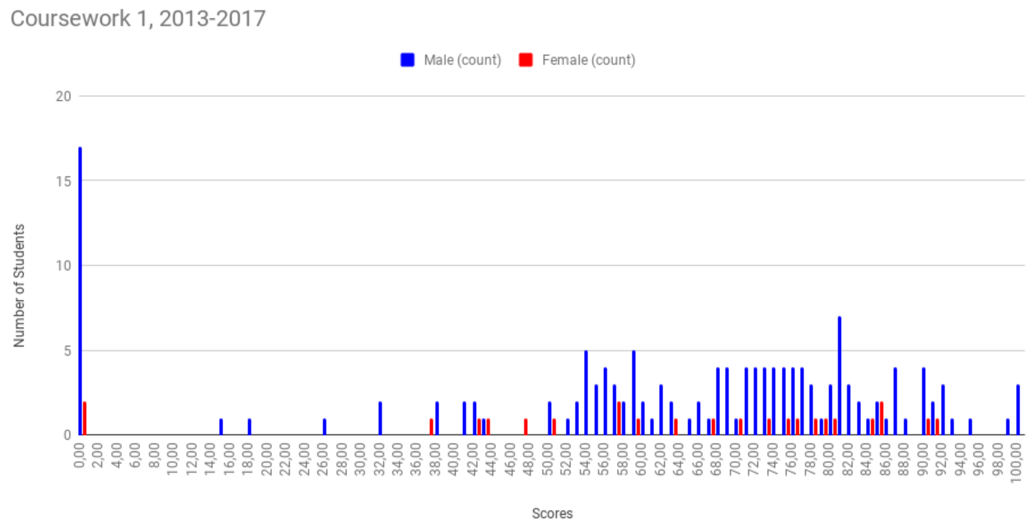


Figure D.6: *Collective histogram of the Data Visualization Coursework 1 scores from 2013-2017 include the average score for males (61) and females (62). Gender is mapped to color.*

seven male and one female receive 98. Eight male students receive a score of 0. We also observed on the graph that gender bias is not reflected in the grade distribution. This investigation indicates that scores are generally evenly distributed between male and female students and gender bias is not evident on exams or assignments in histogram graphs for 2018 or over a five-year period from 2017-2013. Although we did not notice any gender bias, we notice cultural bias in our class. From our perspective, overseas students, especially Chinese students, struggle with the understanding the language. In the next sub-section, we provide some recommendations for inclusive teaching.

D.6 Recommendations for More Inclusive Teaching in Data Visualization

Our Data Visualization course includes two assignments and students are provided data sets to examine in order to produce images. However, some students have difficulties analysing the data set and cannot generate sensible visualizations. One suggestion could be that students should have a chance to select their own data set to analyze and visualize. Students will likely choose data sets that they are interested in and thus perhaps engage more with assignments. Another challenge in the class is language. A language barrier may be present in the class, and many Asian students are limited in their understanding of the lecturer including lecture notes, assignments and exam questions due to a language barrier. This is mainly observed with master level students rather than third-year undergraduate students. The third-year students have time to improve their language skills until their final year at university. A possible recommendation to address this challenge is higher university entry English language requirements. Another approach to address this challenge may be to divide the class based on language skills for the Data Visualization module. The lecturer may adopt a different teaching approach considering students' language diversity. In addition, the lecturer can record the lectures and upload them to an easily accessible web environment such as YouTube. This method enables students to watch the lecture videos multiple times and support them to compensate for not understanding part of lectures. Also, YouTube facilitates subtitles which enable

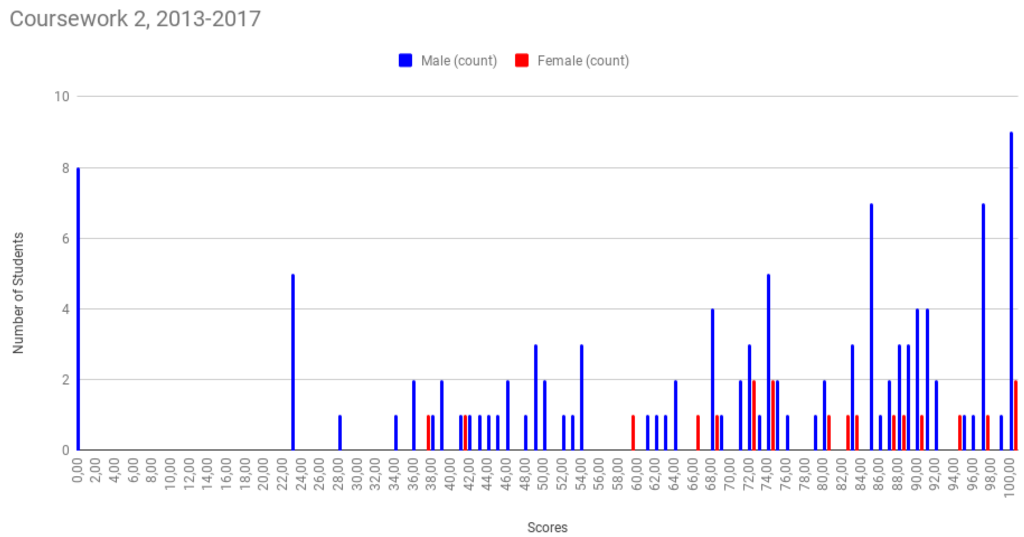


Figure D.7: *Collective histogram of the Data Visualization Coursework 2 from 2013-2017 includes the average score for males (67) and females (77). Gender is mapped to color.*

students to follow and understand lecturers.
Our research is concerned with making classrooms more inclusive due to a variety of cultures, student backgrounds and student learning types. Inclusive teaching has many potential benefits such as being collaborative, engaging, and supports the understanding that considers a diverse student body.
Montgomery [227] describes culturally responsive classrooms that consider culturally diverse students. Students need to engage with the subject topic and the tasks that are given them. Instructional approaches and individual teaching attitudes can encourage all students to get involved in learning activities that will lead to improved academic success. Another point described by Montgomery [227] is that the improvement of instructional programs that avoid failure and increase opportunities for achievement should be the goal of every lecturer. Furthermore, Rodriguez-Falcon *et al.* [228] at Sheffield University provide recommendations to adopt a lecturer's teaching approach to meet the needs of a diverse community with inclusive teaching. These suggestions are:

- Lecturers can use clear language and not speak very quickly.
- Handouts and presentations are written and organized clearly. This means a combination of correct color and font size, clear graphs and images. All course material is accessible online.
- The lecturer gives the impression that they are available to answer students' questions and approach them positively for personal engagement.
- Instructors can explain the processes of assessment and feedback. They do not assume students already know the evaluation structure.
- The lecturer chooses common visualization examples for all students, especially students who have different cultural backgrounds who can be familiar with the example.
- Instructors can break up visualization lectures to ask questions or include short 'partner-work' sessions.

Other approaches to inclusive teaching for visualization teachers are to support critical and analytic thinking and considering conducting a peer mentoring scheme for assistance for all students. As a result of this, a diverse community can communicate and assist the learning process for each other in the class environment. Moreover, students should be encouraged to feel comfortable in the classroom and participate in the lecture with their ideas, thoughts and questions. More specific to visualization:

- Those visualization modules that require a visualization project may consider allowing students to propose their own project as an alternative to a prescribed project. Another option is to allow students to choose between two or more options when selecting visualization assignments or visualization projects. Providing options may support a more diverse student background.

- Data visualization assignments can enable students to generate and collect their own data to visualize rather than using a given data set which students are not familiar to or struggle to analyze.

- We also recommend encouraging the use of diverse hardware including a range of display devices of varying size.

- Another inclusive approach to data visualization class includes the use of low-tech methods such as hand-sketched visual designs life those described by Roberts *et al.* [229].

- And finally, more inclusive teaching can be supported by social media. Social media groups can facilitate convenient and frequent communication between students in a diverse classroom.