DATA SCIENCE, DASHBOARDS, AND THE WAY IT WORKS WITH STATISTICS

by

Denise Renee Bradford

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Statistics

Under the Supervision of Susan R. VanderPlas, Ph.D

Lincoln, Nebraska

Month, Year

DATA SCIENCE, DASHBOARDS, AND THE WAY IT WORKS WITH STATISTICS

Denise Renee Bradford, Ph.D.

University of Nebraska, Year

Adviser: Susan R. VanderPlas, Ph.D

Here is my abstract. *(350 word limit)*

COPYRIGHT

© Year, Denise Renee Bradford

DEDICATION

Dedicated to...

ACKNOWLEDGMENTS

Thank you to all my people!

Table of Contents

List of Figures **vii**

List of Tables **viii**

1 General Introduction **1**

1.1 Exploratory Data Analysis (EDA) 1

1.2 Open-Source 5

1.3 Origins of Open Source 7

1.4 Open Access Data Repositories 7

2 Chapter Paper on Rural Shrink Smart Manuscript submitted to Journal of Data Science Special Issue **14**

2.1 Abstract 14

2.2 Introduction 14

2.3 Data Description 16

2.4 Dashboard Design Considerations 18

2.5 Guiding Design Principles 22

2.6 Dashboard Design Process 23

2.7 Discussion 31

2.8 Future Work 32

2.9 Conclusions 33

3 Dashboard Poetry **35**

3.1 Introduction 35

3.2 Dashbaord Construction 45

3.3 Cognitive Principles 49

3.4 Ensemble Perception 59

4 Does Color Help?: Multidimensional Ensembles in Bland-Altman Plots: An Exploration **66**

4.1 Introduction: 66

4.2 Methods: 73

4.3 Discussion: 75

Conclusion **76**

A The First Appendix **77**

B The Second Appendix, for Fun **79**

Colophon **80**

References **85**

List of Figures

2.1 Diagram of considerations for our dashboard design process. . . . 19
2.2 Initial dashboard design sketch (top) and implementation (bottom). 26
2.3 A second iteration of the sketched design (top) and the implementation (bottom). 28
3.1 Grammar of Graphics Diagram of Wickham and Wilkinson's work 43
3.2 Grammar of Graphics Diagram of Wickham and Wilkinson's work 44
3.3 An example of how the retina signals the visual cortex 50
3.4 Working Memory Model created by Baddeley (left) and Information Processing Model created by Atkinson and Shiffrin (right) . . 58
3.5 Working Memory Model created by Baddeley (left) and Information Processing Model created by Atkinson and Shiffrin (right) . . 58

List of Tables

Chapter 1

General Introduction

1.1 Exploratory Data Analysis (EDA)

John Tukey was the first to organize the collection and methods associated with philosophy into Exploratory Data Analysis (EDA). Previous research by Tukey focused on graphics as a tool for exploratory analysis. In "Exploratory Data Analysis," Tukey wrote that graphics and charts often display data with more enhanced understanding than a table, (Tukey & Wilk, 1966). Tukey outlines detailed the types of different graphics and in which situations to utilize these graphics. He was a strong advocate for the importance of EDA as a crucial first step in the data analysis process and emphasized the need for visualization and interactive techniques to understand patterns and relationships in data.

Tukey's Principles of EDA have become a cornerstone in the field of statistics and have been adopted by data professionals in various industries. Tukey's principles have enabled data professionals to understand complex data sets better and make more informed decisions by emphasizing the importance of visual exploration, data characterization, and model critique. In this way, Tukey's

Principles have revolutionized our data analysis approach and become the foundational framework for EDA.

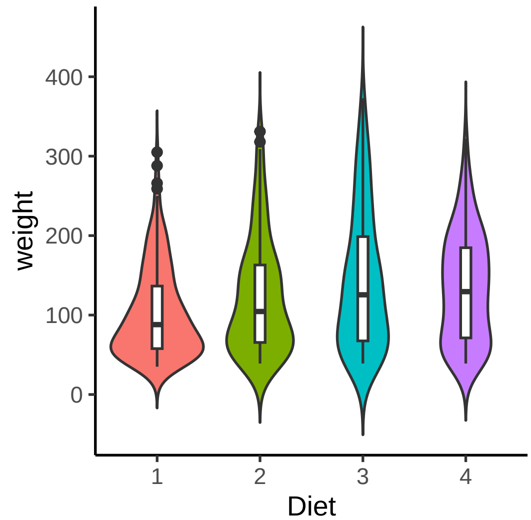
Tukey's Principles in EDA:

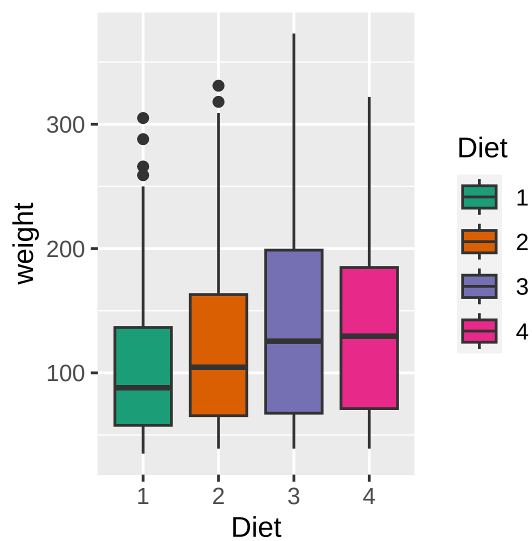
1. Graphical exploration, looking for patterns or displaying fit, the method demonstrates things about data that a single numeric metric does not understand. This has been useful in graphing the data before you develop summary statistics.

2. Describing the general patterns of the data. This step should be insensitive to outliers. In general, think about the types of resistant measures (i.e., median or mean). This step is making sure to determine data patterns.

3. The natural scale/state that the data are at their best. This will be the step at which the scale of data can be helpful for analysis. The reexpressing data to a new scale by taking the square root or logarithmic scale.

4. The mostly known parts of EDA but is done in the way of accessing fit of the data. This is taught in every statistics 101 class. The growth of machine learning and prediction methods have now used residuals more in the toolbox to assessing the best prediction models.





Data visualizations are an integral part of the EDA process, enabling analysts to discern patterns and relationships in the data that would otherwise be difficult to discern from tabular data alone. Through data visualization, analysts can quickly identify trends, outliers, and other patterns that may be missed through numerical analysis alone. Moreover, visualizations facilitate the communication of findings to non-technical stakeholders, allowing them to comprehend complex data sets more efficiently. Through visualizations, analysts can also identify potential issues or biases in the data, resulting in better decisions and models. Thus, visualizations play a crucial role in the EDA process by enabling analysts to more effectively explore, comprehend, and communicate data-derived insights. During the initial EDA stage, an analyst may find that a variable or a covariate is directly related to the dependent variable when looking at a correlation heatmap or a scatterplot. The basic understanding can be formalized to visualize the discovery process.

The field of graphical communication, which is directly related to EDA, semiology, and their use in touch, has been a valuable tool and extension of the EDA thoughts that Tukey expressed. One of the fundamental principles of semiology is the relationship between signifier and signified, in which a visual element (the signifier) represents a particular meaning or concept (the signified), (**barthes1972?**). Another essential concept in semiology is using syntax and semantics to convey meaning in graphic communication effectively. This includes both the syntax and semantics of a graphic's visual elements, (Monmonier, 1985).

Using color to represent data on maps is an example of successful graphical communication utilizing semiology. By using different colors to represent different data points, viewers can comprehend patterns and relationships in the data quickly and easily. Jacques Bertin writes in "Semiology of Graphics" that color can be used to "emphasize a point, distinguish one category from another, or establish a relationship between two points", (Monmonier, 1985). In addition, Bertin explains that the use of color can help overcome language barriers, making it easier for the audience to comprehend the presented information.

The application of semiology in graphical communication is not devoid of obstacles. One difficulty is the possibility of misinterpretation, in which viewers may assign a different meaning to a visual element than was intended, (Monmonier, 1985). Another concern is the possibility of cultural differences in interpretation, in which a visual element may have a different meaning in one culture versus another, (Norman, 2013).

Exploratory Data Analysis (EDA) analyzes and summarizes a dataset to discover patterns, trends, and insights. It is a crucial step in the data analysis process and is often used to identify which variables are essential, what the data looks like, and what the underlying structure of the data is. EDA is typically done using various techniques, such as visualizations, statistical summaries, and data transformations.

1.2 Open-Source

Open-source data, deeply rooted in the foundational principles of shared knowledge that has driven human progress for millennia, has evolved significantly over the decades. Originating from the frustrations of figures like Richard Stallman in the 1980s, who championed the free software movement and paved the way for the creation of transformative licenses such as the GNU General Public License, the concept has grown to encompass more than just software. (Stallman, 2002) With organizations like the Open Source Initiative standardizing open-source practices and monumental projects like Linux showcasing its potential, open-source principles expanded by the late 1990s to include data. This movement towards openness has been marked by a commitment to transparency, collaboration, and unrestricted access, revolutionizing how we perceive and interact with data in the modern era.

The principles of open-source data revolve around the concepts of free access, transparency, collaboration, and redistribution, fostering a community-driven approach to data sharing and utilization. At its core, open source is about making the source content (usually software code) freely available. This transparency allows anyone to review, inspect, and understand the source. The freedom to use, modify, distribute, and study the source content without restrictions is a cornerstone of open source. This is articulated in various open source licenses, like the GNU General Public License (GPL), which emphasizes the rights of end users. Open source projects thrive on collaborative efforts. Diverse groups of people from around the world contribute, enhancing the project's robustness and creativity. Tools like Git and platforms like GitHub have made this collaborative approach more streamlined. Open source projects often foster strong communities. These communities not only contribute code but also offer support, documentation, and strategies for the project's future. These communities' sense of belonging and shared purpose is a driving force behind many successful open-source projects. Contributions to open source projects are often evaluated based on merit. The best ideas or implementations, regardless of their source, are adopted, promoting a culture of excellence. The principle of redistribution ensures that modified versions of open source content remain open. This ensures a perpetual cycle of community-driven improvement and access. Open source principles prioritize the end user's interests. This user-centric approach often leads to software or content that's more aligned with what users genuinely need and want.

1.3 Origins of Open Source

The concept of open source originated long before computers, where shared knowledge formed the basis of human progress. With the advent of computers, Richard Stallman, frustrated with a printer that didn't have available source code, initiated the free software movement in the 1980s and established the GNU Project and the Free Software Foundation (Stallman, 2002). Linux, an open-source operating system kernel initiated by Linus Torvalds in 1991, became one of the most popular examples of open-source success. The Apache HTTP Server, released in 1995, became another success story, powering a large fraction of the internet (Weber, 2004). Open-source principles began expanding from software to data in the late 1990s and early 2000s. The idea was to share data sets for public use without restrictions. Projects like the Human Genome Project advocated for open data to advance science and medicine (Sulston & Ferry, 2002).

1.4 Open Access Data Repositories

In the digital age, open access data repositories have become crucial platforms, transforming how data are shared, accessed, and stored among academic and research communities. These repositories are online spaces created specifically to store datasets and make them available to everyone. In line with the principles of open science, which promote knowledge sharing, transparency, and reproducibility in research endeavors, their main objective is to democratize access to data.

These repositories serve many different purposes at their core. They support openness in the scientific method first and foremost. It enables other researchers and the general public to examine, confirm, and replicate research findings by providing unrestricted access to datasets. Collaboration is also encouraged by this open model. Datasets are accessible to researchers in a variety of fields and locations, providing opportunities to expand upon, combine, or compare various sets of data. Eliminating research duplication is a significant additional benefit. Access to data from earlier studies reduces the need to repeatedly collect similar data, which saves time, effort, and resources. Additionally, these repositories are essential for the long-term preservation of data, guaranteeing that datasets are accessible for future scientific research and for posterity.

A closer look at the features of these repositories reveals several standard components. Metadata is paramount; this descriptive information elucidates the context, origin, methodology, and other pertinent details of the datasets, ensuring they are comprehensible to those accessing them. Many repositories also emphasize the citability of datasets. By assigning a Digital Object Identifier (DOI), datasets become easily referable in academic and research contexts. Licensing is another cornerstone. To clarify usage rights and conditions, datasets are often accompanied by explicit licenses, with Creative Commons licenses being particularly prevalent. These licenses can stipulate various conditions, with attribution being a common requirement. Furthermore, the user experience is enhanced through tools that facilitate easy searching, accessing, and downloading of datasets.

However, like all systems, open access data repositories face challenges. Ensuring data privacy is a significant concern, especially in datasets derived from human subjects, such as in health or social research. Standardization is another hurdle. Given the diverse range of researchers and datasets being uploaded, achieving a standardized format for data and metadata is daunting. And not to be overlooked is the challenge of sustainability. Maintaining a sophisticated digital platform requires both financial resources and technical expertise. Ensuring the longevity of such repositories, especially in a rapidly evolving digital landscape, remains a concern.

Several prominent examples underscore the significance and diversity of open access data repositories. Zenodo, developed under the European OpenAIRE program, serves as a multifaceted platform catering to researchers across disciplines. Dryad specializes in datasets linked to scientific publications, predominantly in life sciences and biomedicine. Figshare offers a broader spectrum, allowing researchers to deposit a range of research outputs, including datasets, making them available to the public.

Historically, the rise of these repositories can be contextualized within the broader open science movement. This movement, gathering momentum over recent decades, has been advocating for a more transparent, accessible, and shared research process. As technological advancements surged, leading to an exponential growth in digital data, the emphasis on preserving and sharing this data treasure trove intensified.

Today, the significance of open access data repositories is further underscored by policies and guidelines from influential bodies. A growing number of funding agencies and academic journals are either mandating or recommending that data underpinning research findings be housed in open access repositories. This trend not only ensures that research outputs, especially data, remain accessible but also reflects a broader societal push towards transparency, especially when research is funded by public coffers.

R has become a prominent and adaptable programming language within the domain of data analysis and exploratory data analysis (EDA). A number of crucial R packages have been created to streamline the processes of data manipulation, visualization, and statistical analysis. One of the prominent programs in this category is ggplot2, which was developed by Hadley Wickham. This package is widely recognized for its versatility and aesthetic appeal in generating a diverse range of visually appealing and informative data visualizations (**wickham2016?**). Wickham et al. have developed dplyr and tidyr, which offer convenient functionalities for filtering, summarizing, and reshaping data frames, hence enhancing the efficiency of data manipulation activities (**wickham2021?**). The data.table package, developed by Matt Dowle, is widely recognized for its remarkable speed and efficiency in handling huge datasets, particularly in the context of performance-oriented data manipulation tasks (**dowle2021?**). Although not classified as a R package, it is noteworthy to include Wes McKinney's Python library, pandas, due to its significant influence in the field of data manipulation and analysis in the Python programming language (**mcKinney2010?**). Within the field of psychology and psychometrics, the psych package developed by William Revelle serves as an extensive tool for doing factor analysis and reliability analysis (**revelle202?**).

Furthermore, there are other software packages available that may be utilized by researchers and analysts to enhance their data management and analysis processes. For instance, the naniar package, developed by (**tierney2020?**), is specifically designed for handling missing data. Another useful package is summarytools, created by (**comtois2021?**), which enables the generation of descriptive statistics. Additionally, the corrplot package, developed by Wei and Simko (Wei et al., 2017), provides significant tools for displaying correlation matrices. These packages provide researchers and analysts a range of valuable resources to aid in their work. The FactoMineR package, developed by Husson, Lê, and Pagès (**husson2020?**), offers crucial techniques for conducting multivariate data analysis, including principal component analysis and clustering. These software programs, created by renowned individuals in the discipline, jointly provide data analysts and scientists with the necessary tools to efficiently investigate and analyze data.

Here are some popular R packages for exploratory data analysis along with a brief summary:

1. **ggplot2**:
    - Summary: ggplot2 is a powerful data visualization package for creating a wide range of high-quality plots and charts. It follows the Grammar of Graphics framework, making it highly customizable and flexible.
    - Author: Hadley Wickham
2. **dplyr**:
    - Summary: dplyr is a package for data manipulation. It provides a set of easy-to-use functions that facilitate common data manipulation tasks such as filtering, selecting, grouping, and summarizing data.
    - Author: Hadley Wickham

3. **tidyr**:
   - Summary: tidyr is another package by Hadley Wickham that is used for tidying and reshaping data. It helps in converting data from wide to long format and vice versa, making it suitable for analysis and visualization.
4. **data.table**:
   - Summary: data.table is an efficient package for data manipulation and analysis. It is particularly well-suited for large datasets and complex operations.
5. **pandas** (not an R package, but widely used in Python for EDA):
   - Summary: pandas is a Python library for data manipulation and analysis. It provides data structures like DataFrames and Series, along with a wide range of functions for data cleaning and analysis.
6. **psych**:
   - Summary: The psych package in R provides various functions for conducting psychometric and psychological research. It includes tools for factor analysis, reliability analysis, and data visualization.
7. **naniar**:
   - Summary: naniar is an R package that helps with missing data exploration and visualization. It provides functions for identifying, visualizing, and handling missing data in datasets.
8. **summarytools**:

- Summary: summarytools is a package for creating beautiful and informative summary tables and charts for descriptive statistics. It can be helpful for summarizing and exploring data.

9. **corrplot**:

   - Summary: corrplot is a package for visualizing correlation matrices. It provides various plot types and customization options for exploring relationships between variables.

10. **FactoMineR**:

    - Summary: FactoMineR is a package for multivariate data analysis, including methods for principal component analysis (PCA), factor analysis, and clustering. It is useful for exploring complex datasets.

Chapter 2

Chapter Paper on Rural Shrink Smart Manuscript submitted to Journal of Data Science Special Issue

2.1 Abstract

Many small and rural places are shrinking. Interactive dashboards are the most common use cases for data visualization and context for exploratory data tools. In our paper, we will explore the specific scope of how dashboards are used in small and rural area to empower novice analysts to make data-driven decisions. Our framework will suggest a number of research directions to better support small and rural places from shrinking using an interactive dashboard design, implementation and use for the every day analyst.

2.2 Introduction

As the amount of data has increased in nearly every facet of life, the need to make sense of that data in an approachable, accessible form has become ever more important. As a result, many companies and organizations use interactive dashboards to present these data in a more useful and visually appealing form (Sarikaya, Correll, Bartram, Tory, & Fisher, 2019).

In many cases, dashboards support viewers' information processing, helping to make sense of complex data, navigate through a dataset, and supporting decision making based on the data.

Dashboards are often used, as with the car display of the same name, to provide summary information about many separate attributes of a common entity. One glance at a car's dashboard will tell you the speed, RPM, engine temperature, amount of gas in the tank; more importantly, however, the goal is not for the user to remember all of these characteristics, but to assess whether any of these quantities is outside of the expected range. Similarly, interactive dashboards for data are often used to display many different attributes and performance metrics which are of importance for stakeholders.

In this paper, we discuss the process of designing a dashboard to present publicly available government data to stakeholders in small Iowa towns to facilitate decision making and objective comparison with other similarly-situated towns.

Some communities continue to thrive as they lose population because they adapt, maintaining quality of life and community services for residents while investing in the future. This process, *smart shrinkage*, is important for rural areas who have experienced shrinking populations for decades. As small rural towns do not have access to data scientists or even the ability to easily leverage data collected locally to support decisions, our research team will provide communities with data about services in small town Iowa in order to assist with developing strategies to improve quality of life for their residents amid shrinking populations (Rural Shrink Smart Team, 2022). We hope to allow towns to explore their own data and compare to other similar towns, centering decision-making on data in the context of small-town Iowa life.

2.3 Data Description

The Smart and Connected Community (SCC) dashboard data are primarily assembled from `data.iowa.gov` (State of Iowa, 2020), with some additional datasets assembled from federal and private sources. Most of these data sets are collected at a town/city or county spatial resolution, requiring us to carefully join data to ensure that these differences are respected while collating relevant information at the city level. In addition to the more commonly available statistics derived from e.g. the census and American Community Survey, `data.iowa.gov` contains several unique data sets, including local liquor sales, school building locations, town budgets and expenditures, hospital beds, Medicaid reimbursements, and other details that may provide information about local quality of life.

Data available on Iowa's data portal were augmented in some cases with higher-quality data sets in cases where the Iowa data were out of date or insufficiently accurate. Data collected from ELSI (National Center for Education Statistics, 2020) from `https://nces.ed.gov` were used to show the distance to any private or public school. The National Center for Education Statistics (NCES) is the primary federal entity for collecting and analyzing data related to education (Zarecor, Peters, & Hamideh, 2021).

Data collected from the Index of Relative Rurality (IRR) (USDA - ERS, 2020a) were used in the SCC dashboard to help classify the towns. The Index of Relative Rurality (IRR) is a continuous, threshold-free, and unit-free measure of rurality. It is an alternative to the traditional discrete threshold-based classifications.The IRR ranges between 0 (low level of rurality, i.e., urban) and 1 (most rural). Four steps are involved in its design:

1. Identifying the dimensions of rurality: population size, density, remoteness, and built-up area.

2. Selecting measurable variables to adequately represent each dimension:

   - Size: logarithm of population size
   - Density: logarithm of population density.
   - Remoteness: network distance.
   - Built-up area: urban area (as defined by the US Census Bureau) as a percentage of total land area.

3. Re-scaling the variables onto bounded scales that range from 0 to 1.

4. Selecting a link function: unweighted average of the four re-scaled variable.

Data collected from Rural Urban Commuting Area Codes (USDA - ERS, 2020b) were used to help identify towns with commuting behaviors in our rural areas. The rural-urban commuting area (RUCA) codes classify U.S. census tracts using measures of population density, urbanization, and daily commuting. This data is on a zip code-level that will help identify those communities that commute to more urban areas. The most recent RUCA codes are based on data from the 2010 decennial census and the 2006-10 American Community Survey. The classification contains two levels. Whole numbers (1-10) delineate metropolitan, micropolitan, small town, and rural commuting areas based on the size and direction of the primary (largest) commuting flows.

One of the interesting features of this assembled data set is that missing data can be missing for multiple reasons: not all state data is complete, but data about certain services may also be missing because towns do not offer that service. Thus, in addition to the usual challenges of working with real-world data that is “messy” in a variety of ways, we also have to contend with missing data that is missing due to the size of the community or the lack of services. This makes both visualization and statistical analysis more complicated (and more interesting).

2.4 Dashboard Design Considerations

One problem we identified early in the process of assessing smart-shrinkage strategies in small towns is that these towns do not have the resources to make data-driven decisions. Typically, small towns in Iowa are managed by at most a few part-time employees or volunteers. In some cases, essential management functions of the town are paid, but the municipalities we are interested in do not have sufficient funding to hire professionals to gather and analyze data.

As part of a wider project investigating the strategies towns use to maintain quality of life amid shrinking population, our research team provides communities with data about their own town, but also comparable towns across the state which may have a different approach to city services. In combination with other engagement strategies that are more qualitative, we hope to use this interactive dashboard approach to assist small Iowa cities with generalizing and developing strategies to improve or maintain quality of life amid shrinking populations.

One factor at the forefront of our visualization design is the importance of

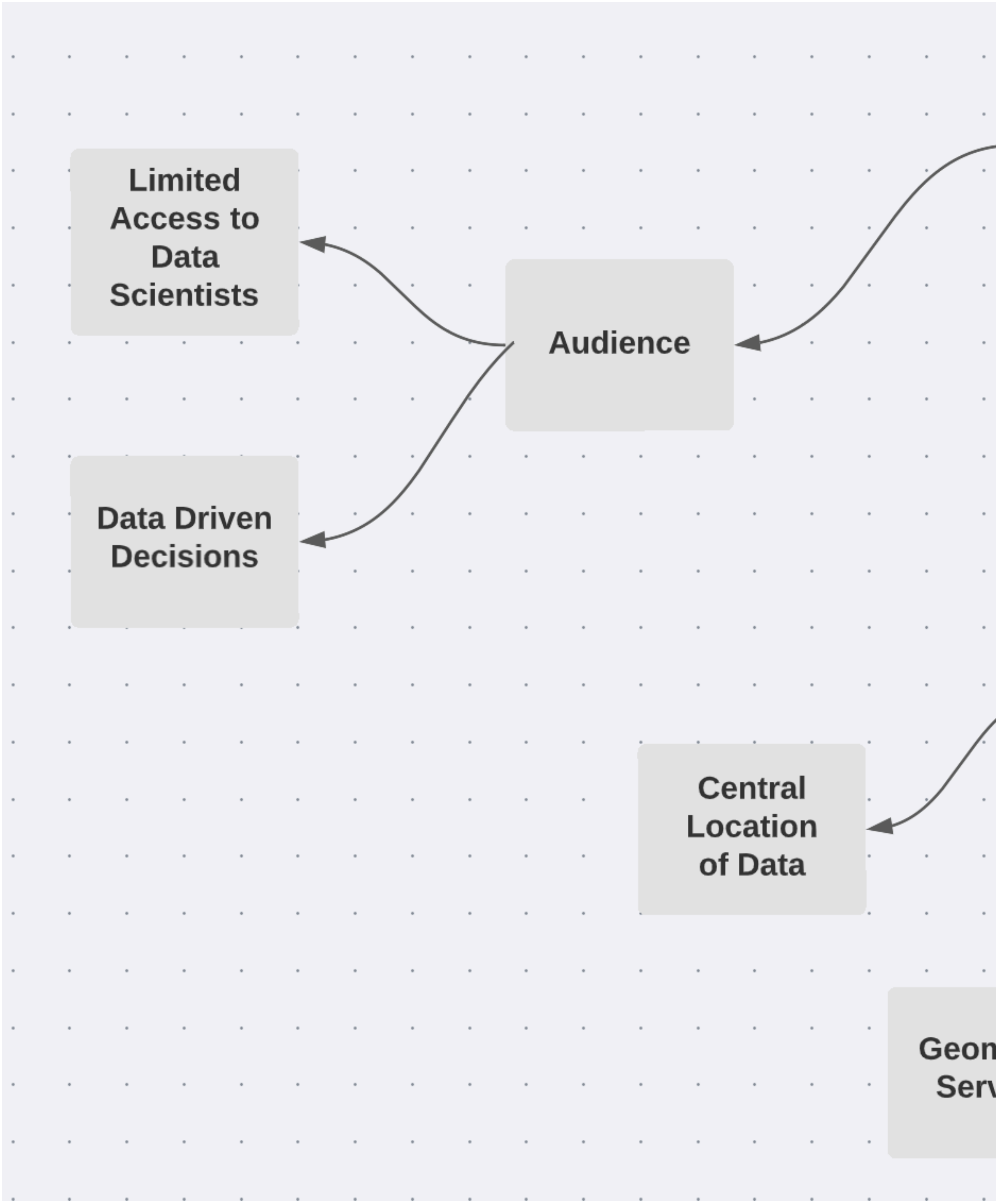


Figure 2.1: Diagram of considerations for our dashboard design process.

reducing the cognitive demands on viewers: we have assembled an incredible amount of data, and it is easy for even statisticians who deal with much larger datasets to get lost in the details of this data. At the same time, we want to invite viewers to engage with the data - to imagine, to draw comparisons, to generalize across towns, and to integrate outside information into the conclusions drawn based on the data we present. This invitation to engage with the data is similar to the approach advocated in Guided Discovery Learning, a framework leverages hints, feedback, and other helpful information to guide users in interactive exploration (DeDonno, 2016).

We expect that users will be interested in “sets” of variables from the wider dataset, which we assembled based on quality of life factors in the Iowa Small Town Poll (Peters, 2019). For instance, users might be interested in medical and social services available to residents, such as a local primary care clinic, nursing homes which are within driving distance, and the distance to the nearest emergency room; these factors might be explored separately from variables describing the services provided directly by city government, such as parks and recreation expenditures, snow removal services, and the distance to the closest fire station.

As a consequence of this massively multivariate structure, we very quickly focused on the use of parallel coordinate plots; other alternatives, such as tours (Wickham, Cook, Hofmann, & Buja, 2011), require much more sustained attention to interactive plots as well as a deeper understanding of projections in multidimensional space which we cannot assume our users will have. Introduced in the 1880s (d’Ocagne, 1885), parallel coordinate or parallel set plots feature a series of vertical axes representing different variables arranged horizontally, with lines connecting each observation. When representing categorical data, parallel set plots may show "blocks" of data instead of individual lines, and are useful for representing conditional relationships between adjacent variables (Bendix, Kosara, & Hauser, 2005); modifications of this design, such as common-angle plots (Hofmann & Vendettuoli, 2013), address the issues which arise due to line-width illusions VanderPlas & Hofmann (2015a). Parallel coordinate plots have been generalized to allow for continuous data and additional summaries beyond individual data points, such as densities (Heinrich & Weiskopf, 2009). In this paper, we use the `ggpcp` package, which leverages the grammar-of-graphics framework introduced in Wickham (2016), allowing us to use not only parallel coordinate plots, but also to overlay other statistical summaries, such as boxplots or violin plots, to provide additional context about the marginal distributions of each variable in addition to allowing for exploration of the multivariate space.

We also anticipate that users will be interested in comparing their town to other, similar towns. We will discuss the different ways that this comparison strategy was implemented in each dashboard in the next section, which describes the evolution of the dashboard over time and accounting for feedback from users and other researchers on the wider project.

One final component of this project is that our dashboard is part of a wider effort to work with towns to understand the different strategies used to maintain resident quality of life amid shrinking populations. Thus, while the town leaders are our primary audience, we also are creating this applet for use in parallel with a team of other researchers: sociologists, economists, city planning specialists, and artists. These researchers opinions and feedback about the dashboard are also useful and important, as they regularly work with town leaders in different capacities and have an understanding of what factors are most important to them and what types of questions these leaders may have when faced with data and unfamiliar statistical visualizations.

Throughout the design process, we will assess our visualizations to determine which strategies for user interface and interactive graphics design are most useful to empower town leaders to make discoveries in publicly available data assembled with a focus on items that impact rural quality of life.

2.5 Guiding Design Principles

Research on dashboard creation and interactive visualization tends to be very task-specific and hard to apply to more generalized settings. That is, it is relatively easy to create a dashboard that works for a particular task, but it is hard to generalize from that process what will work for the next dashboard. With this in mind, we set out to clearly document our intentions at each stage of the design and evaluation process, with the goal of gathering some useful information about general dashboard design from the process of creating this specific dashboard.

Thus, our initial set of dashboard design principles is as follows:

- The town leaders are the focus audience; thus, the town itself should be the central focus of the app.
- We should facilitate comparisons with other towns in order to allow the user to explore other potential solutions to offering services that enhance resident quality of life.

- We will present the user with peer comparisons in order to widen the scope of exploration beyond the initial set of obvious peers in the local region.

- We will implement feedback mechanisms that allow us to provide more detailed data and respond to feature requests to improve the dashboard design over time.

As with many dashboards, this project is under continuous development; while it makes for an unsatisfactory conclusion, we do not have a "final" dashboard design because the application will continue to evolve. However, we have some useful insights into the process of creating an application designed to invite users to explore a large and complex dataset that we believe to be a useful contribution to work in this area.

2.6 Dashboard Design Process

2.6.1 Dashboard Components

In this section, we discuss the philosophy behind the basic "building blocks" of the dashboard. This philosophy is present in all of the iterations of the dashboard that we present in this discussion, and we will evaluate the overall philosophy's effectiveness in the conclusion.

The large set of publicly available data (primarily from `data.iowa.gov`) we have assembled is useful, but we must be careful with how we present this data because it would be easy to overwhelm the user with small details that mask the bigger picture. We select a small subset of towns (out of the 999 towns in Iowa) and a small subset of variables of interest to start with, and then allow the user to increase the complexity of the display in accordance with their interest. This avoids some of the pitfalls of dashboard design that can easily lead to user overload (Few, 2006a).

Our primary objective is to provide users with a town-centric approach: their town is at the center of our application, and comparisons to other, similar towns are secondary. As a result, the next component of the dashboard is intended to provide a brief overview of the information we have about a specific town of interest. This design is based on research into visualization sensemaking (Lee et al., 2016), in that we allow users to explore outward from the familiar to the unknown. The map visuals were built using Open Source Routing Machine (OSRM) route functions (Luxen & Vetter, 2011) in R (R Core Team, 2022) to amplify the accuracy of the distances from necessary services in town-centric point. OSRM allows for finding the "As the Crow Flies" distance and time on the road for our vital services map, since OSRM technology is similar to Google maps.

When faced with the next component, a parallel coordinate plot (PCP), a novice user will be able to determine two basic components: Visual Object (textual objects and non-textual objects) and Frame (frame of content and frame of visual encoding).

Taken together, the app is a single page; the initial "solid ground" which the user explores from consists of maps showing the route from the center of town to necessary services, including the fire department, schools, post offices, and hospitals. In version 2, as shown in Figure 2.3, the map portion is condensed, and more space is given to value boxes that show vital statistics about the town's QoL and financial metrics. This relatively straightforward display is followed by a parallel coordinate plot that allows the user to see similar towns along dimensions such as economic indicators or population size.

2.6.2 Initial Draft

The initial design sketch and implementation are shown in Figure 2.2.

Users' towns are at the center of our application, and comparisons to other, similar towns are secondary. As it can be extremely difficult to predict which towns are optimal for comparison purposes (similar may involve population, region, economic indicators, sports rivalries, and any number of other variables), we allow users to modify a set of suggested comparison towns to indicate other towns of interest.

We implemented some suggested town comparisons using unsupervised clustering methods to help our towns make decisions that are informed in comparison to similar towns, for budget size, population size and location. We initially focused on determining the next five to ten similar towns, based on distances to services. This feature became an important diagnostic for our data quality, as it became clear that towns which were grouped with big cities but which did not have a large population were so grouped because of missing data. Unfortunately, this clustering feature was not as useful to the application users, as they came to the dashboard with a pre-existing set of towns to compare to; our suggested comparisons were in the way.

The initial dashboard design featured several responsive maps showing the distance to the nearest hospital, fire department, post office, and school. These maps were ineffective for several reasons:

- Town residents already know this information (though it was useful for

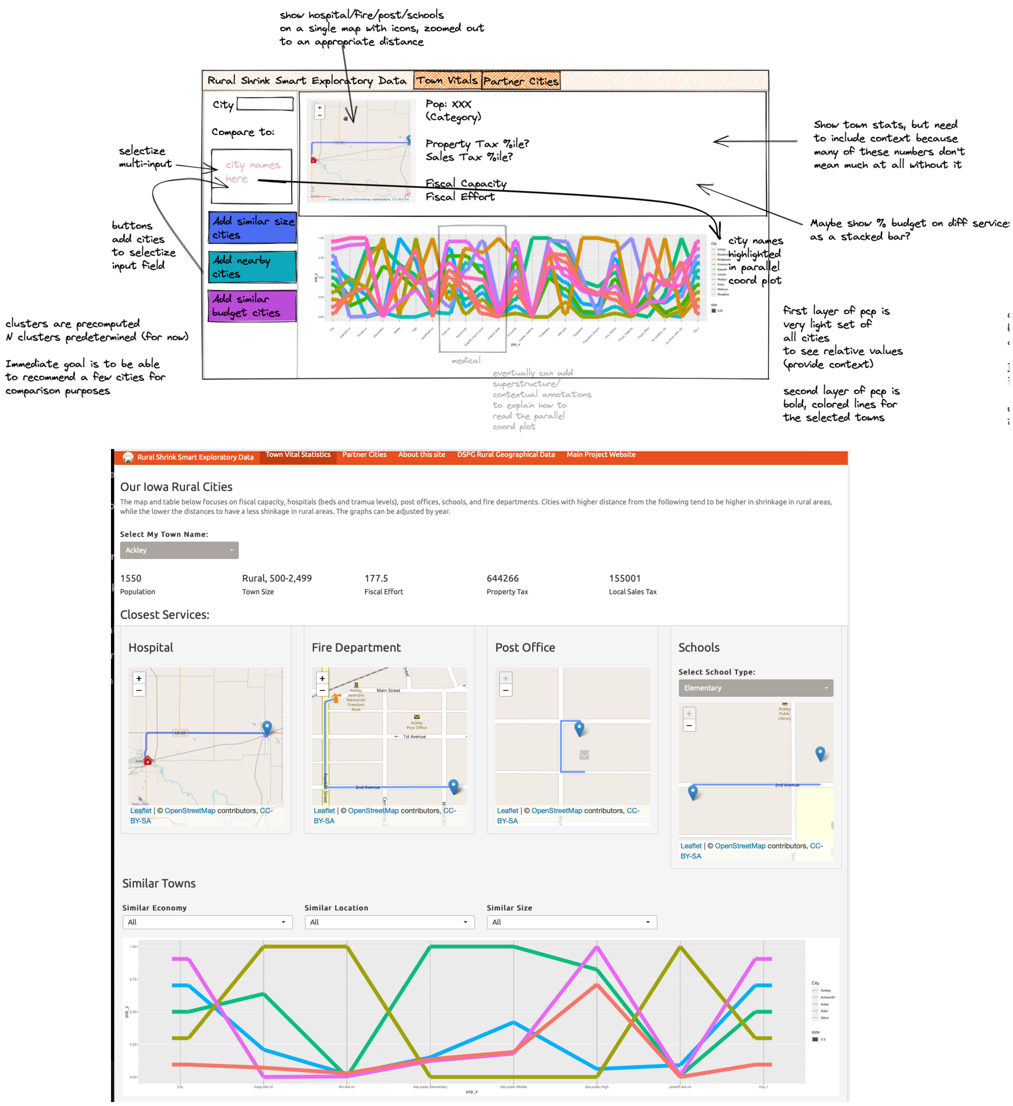


Figure 2.2: Initial dashboard design sketch (top) and implementation (bottom).

us as the dashboard designers, because we aren't nearly as familiar with the 900+ small towns in Iowa)

- We computed distance from services relative to the center of town - coordinates provided in the data from data.iowa.gov. Generally speaking, the post office is at the center of town and the fire department is usually very close to the center of town; these two maps were useless. The school and hospital maps were less useless, but still did not provide particularly useful information to people already familiar with the town.
- It became clear that it might be more useful to show the comparison towns on a map (relative to the town of interest) so that users could compare geographical ratings for unfamiliar data to familiar data.

In addition, we received feedback on the parallel coordinate plot at the bottom of the app which was surprising: the viewers (in this case, other researchers on the team) were not as intimidated by the parallel coordinate plot as we had expected. They did need some explanation of how to read the plot, and these hints need to be included in the dashboard, but they grasped the fundamental idea of the plot very quickly.

Our conclusion, based on this initial dashboard draft, was that we needed to restructure the application. Our attempt to show familiar information first to "build up" to the more unfamiliar structure of a parallel coordinate plot was not effective; there was too much clutter and not enough new information to draw users in.

Figure 2.3: A second iteration of the sketched design (top) and the implementation (bottom).

2.6.3 Redesign

In the initial design, we included a map for each vital service, this initially created a lag for the users' experience. As a result, we cached map directions from OSRM for each service in our database, which drastically reduced the response time for the user. Our initial design did not naturally focus the user's eye on the most important parts of the dashboard; the redesign allowed for a cleaner flow from the top to the bottom.

In addition to the timing due to the map loading slowly, we added the vital statistics at the county level to allow for a more robust understanding of the town and it's surroundings. The rurality index provided a better classification and the USDA sources allowed for the town to understand the impact of the closest major city due to commuting for work and shopping at larger stores not available within the town.

We also modified the parallel coordinate plots in several ways:

- Our x-axis had a large number of variables that we as researchers believed to be the most strongly associated with quality of life. However, there were still too many variables for users to successfully parse. We reduced the number of variables, focusing on variables that had the highest data quality, and we grouped these variables by quality of life factor (Peters, 2019).
- Originally, parallel coordinate bands were scaled based on the selected comparison towns. This had the effect of truncating the range of variables and over-emphasizing differences between selected towns relative to the overall range of each variable over all towns in our data set. We

chose to show all towns in the data set in a very light α grey color to provide some information about the overall range of each variable. Unfortunately, even with the low-α value, this increased the visual complexity of the plot and confused users. Future iterations will likely make use of another aesthetic, such as boxplots or violin plots, to show the range of values for all towns, and then use lines only for towns that are selected by the user. This should strike a balance between visual complexity and representing the data accurately.

- We noticed that users did not make use of our suggested comparison towns, and so we removed that option in favor of allowing users to enter their own comparison towns directly. Users already had pre-determined towns they wanted to compare to, and our suggestions were just in the way.

While not all of these modifications were well received in our second round of user testing, the changes did incrementally move the dashboard display towards our goal of allowing users to explore the data and engage with it. We continued to be surprised with how well users reacted to the parallel coordinate plots, which we initially thought might be too abstract for users unfamiliar with multivariate data displays, but the ability to compare towns across multiple dimensions and examine the similarities and differences between their approaches to different services seemed to be intuitive for users once they understood that each vertical axis was a different variable.

2.7 Discussion

Our dashboard design philosophy worked primarily to promote a town-centric approach application with comparisons to other similar towns being secondary. This approach created a way for the user to see their town information at the top of the page and to explore the PCP after reviewing their own town's essential statistics. The PCP in the lower part of the dashboard allowed for the user to see the plot and adjust to the fact that they could add more towns to the plot, providing an opportunity to explore the wider dataset from a base of familiar knowledge.

While we initially framed the design around guided discovery learning, the approach did not seem to suffice for our user base; instead, we found that users were more drawn to the unfamiliar from the start. We will likely leverage this in future iterations by using visual forms such as flower plots to draw the users in; even though these plots are not ideal for numerical display of data, the visual novelty and aesthetic appeal will provide some motivation to continue exploring and thinking about the data.

One factor that we have briefly considered and have seen hints of in our user feedback is that towns may not want to be compared negatively with other towns. While users have very definite ideas about which towns they would like to compare to, we can always mask the town names and move back to comparisons based on town size and other factors (for instance, whether or not a town is the county seat is a factor that is important outside of population). Using this approach, we would label each town as "Town 1", "Town 2", and so on, which would eliminate some of the fears about negative comparisons, but would also remove some of the novelty of the data dashboard for our users and would prevent users from drawing on their own outside knowledge about each of the comparison towns.

We also recognize that we need to leverage the expertise of others in our research team: we are working with artists, researchers in architecture, economists, and sociologists; these researchers provide outside knowledge that we do not have and may be able to help us create insightful use-cases to showcase the app and teach towns how to use it. We can also leverage the app to connect users with our research team, providing additional value to those who use the applet and facilitating development of strategies for maintaining quality of life amid shrinking populations.

2.8 Future Work

One avenue we will explore in future iterations of the dashboard is to incorporate other dashboards generated by different groups within this project. This will create a wider field of information to explore: for instance, some of the additional work will focus on the 99 towns featured in the Iowa Small Town Poll; this will allow us to showcase survey-based measures of quality of life alongside the more objective measurements assembled in the dataset discussed in this paper. While at least one tab of this omni-dashboard will still focus on wider EDA and discovery, we hope to incorporate other information as well to provide a more well-rounded data display encompassing most of the facets of this complex project.

We are also mindful of a distinction between "eye candy" and purpose-driven data visualization. While we have typically focused on the latter, there is certainly a place in our dashboard for the former as well. "Eye candy" visualization is intended to draw the viewer in and motivate them to explore; while these visualizations may not be particularly effective at communicating quantitative information, if they motivate the user to engage with the rest of the dashboard, they still serve a purpose. It is with this mindset that we intend to explore the use of flower plots - the artistic opportunities combined with the display of quantitative information (even in a form that isn't optimal for quantitative comparisons) may be useful to engage viewers before transitioning to more useful data visualizations intended to provide accurate quantitative comparisons.

EDA can be a difficult for a variety of groups of people, novice users and experienced researchers. One of the more difficult components of this project has been clearly articulating the purposes of EDA to a diverse group of researchers unfamiliar with the concept. One of the most useful parts of this dashboard iteration process has been as an aid to data discovery: that is, the dashboard motivated us to find additional data sources and incorporate them into the project. Having conversations with other researchers about the EDA process helped to facilitate these conversations, as each discussion seemed to uncover additional data sources that someone remembered after looking at the dashboard. While this facet of the dashboard process may be difficult to study formally, it would be an interesting avenue for investigation.

2.9 Conclusions

In this paper, we have documented the process of designing a dashboard for exploration and visualization of a large and complex data set assembled from many different sources. Our primary audience was leaders of small towns in Iowa, with a secondary audience of researchers in fields other than statistics collaborating on this project with us. Through the process of revising our dashboard, we found that the idea of guided discovery learning as implemented in our first version did not work as well as we had anticipated. It was more important to focus on allowing users to explore their questions about the dataset by facilitating user-driven comparisons and exploration, rather than attempting to anticipate user desires by providing comparison towns. In addition, we found that it would be more effective to draw users in with novel visual displays, as these seemed to attract more interest than providing known facts and an opportunity to explore outwards from an initial area of familiarity.

While it is hard to apply the findings from one fairly specific visualization project more widely, there is a lack of resources in this area that provide both design philosophies and actual analysis of user feedback in a qualitative sense. We have attempted to address this dearth of information by providing the design strategies, user feedback, and our planned and executed modifications, in the hopes that others facing the daunting challenge of designing a dashboard for EDA may learn something from our experiences.

Chapter 3

Dashboard Poetry

3.1 Introduction

Statisticians use graphs in almost every stage of their work. They create charts to summarize and explore new data and identify potential problems and opportunities (Tukey & Wilk, 1966). Models are fit based on relationships between variables which are often identified visually (Hullman & Gelman, 2021). We identify problems with those models based on residual plots and other visual diagnostics. When our modeling work has been completed, we present our results to interested parties using visual displays, because non-statisticians often find it easier to understand data and models through an intuitive visual medium rather than through the mathematical formulae which underlie the statistical work.

When we create visualizations for public consumption, we have to consider both perceptual factors and the target audience's domain knowledge. In addition, not all visual displays have equal perceptual value Aspillaga (1996). The best graphics are designed to account for both the dataset and the intended audience's features. Some design constraints stem from limitations of the human perceptual system and are common to most potential consumers of the visualization. For example, the sine illusion affects anyone with binocular depth perception, and color recommendations are built around the specific characteristics of the human retina (VanderPlas & Hofmann, 2015b). Other design constraints are due to the audience's experience level and if they are used to working with data and understand specialized techniques (e.g., enough familiarity with principal component analysis such that a plot of factor loadings might be useful). Given the wide range of uses for graphs and visual data displays in statistical modeling, it is unsurprising that some graphs are more useful for specific applications, such as exploratory analysis, and are unsuitable for other applications, such as presenting to an outside group.

Most research in statistical graphics has been done on static graphics; usually, research also strips away all but the most essential contextual information, sacrificing external validity for statistical control. As a result, it can be hard to generalize this research to practical applications, where the contextual information surrounding the data is critical and the chart does not just exist in a vacuum.

In the real world, however, conventions and familiarity often win out over best practice validated by perceptual experiments. In "The Commercial and Political Atlas," Playfair used various types of graphical representations to illustrate economic and political data. He included a chart that he called a "circle chart" or "pie chart" to display the distribution of imports and exports of England in 1781 (Playfair, 2005). Scholars frequently choose to utilize pie charts as a means of presenting data, acknowledging the widespread familiarity of this particular chart format among readers (Edward R. Tufte, 2001). Discrete categories can be effectively highlighted by their utilization. Moreover, researchers have the ability to enhance the accessibility of their study findings to a broader range of individuals. However, there are certain limitations that should be acknowledged. Pie charts may not be appropriate for visualizing datasets that have a large number of categories or require precise comparisons. This is because it can be difficult to precisely perceive slight differences in the sizes of the pie chart's wedges. In addition, the presence of inaccurate labeling and scaling has the potential to result in misinterpretation. Hence, it is imperative for researchers to exercise prudence while utilizing pie charts as a visualization tool for categorical data, in order to guarantee precise and significant portrayal.

Dashboards have seen a significant increase in day-to-day usage as a potent data visualization and decision-making tool. The John Hopkins COVID-19 Dashboard is a reputable online platform that offers current and comprehensive data and statistics pertaining to the worldwide COVID-19 epidemic. The tool mentioned earlier functions as a highly helpful instrument in monitoring the dissemination of the virus, providing up-to-date information on the prevalence of cases, fatalities, and immunization rates within various nations and geographical areas. This dashboard is widely utilized by researchers, politicians, and the general public as a reliable and authoritative resource for continuously monitoring the ongoing effects of the epidemic ("John hopkins university covid-19 dashboard," n.d.). The proliferation of dashboards can be attributed to several factors, including the growing availability of data from various sources and the increasing need for organizations to extract actionable insights. Steven Few found that the widespread use of dashboards is attributable to their capacity to present key performance indicators and relevant metrics in a visually appealing and easily digestible format, (Few, 2006b). Moreover, technological advancements and the development of user-friendly dashboard platforms have facilitated the creation and effective utilization of dashboards by individuals from diverse fields. Dashboards have revolutionized data analysis and presentation, allowing users to gain valuable insights and make data-driven decisions more effectively.

Each chart on the dashboard contributes to the overall comprehension of the situation, similar to how each sentence in a paragraph contributes to the larger concept. A chart is similar to a sentence in that it presents a straightforward piece of information and visually represents data to facilitate comprehension. For example, a chart could be a bar graph depicting sales over a year, a pie chart illustrating the percentage distribution of a budget, or a scatter plot illustrating the correlation between two variables. A dashboard may combine multiple graphs, tables, and metrics to provide an all-encompassing view of a company's performance, a project's development, or market trends.

However, a counterexample to this analogy could be a poorly designed dashboard presenting overwhelming information without clear organization or hierarchy. In such a case, the charts may compete for attention and confuse the reader, similar to how a paragraph with too many disjointed sentences can lead to confusion and a lack of coherence. This counterexample highlights the importance of thoughtful design and effective communication in creating an informative and comprehensible dashboard. The research objective of this study is to investigate how changes in real-time data displayed on a dashboard affect ensemble perception and the user's ability to make accurate and rapid decisions based on summary statistics in a dynamic environment.

John Tukey was the first to organize the collection and methods associated with philosophy into exploratory data analysis (EDA). Tukey used graphics as a tool for exploratory analysis. In "Exploratory Data Analysis," Tukey wrote that graphics and charts often display data with more enhanced understanding than a table (Tukey & Wilk, 1966). Tukey outlines in detail the types of different graphics and in which situations to utilize them. He was a strong advocate for the importance of EDA as a crucial first step in the data analysis process and emphasized the need for visualization and interactive techniques to understand patterns and relationships in data.

John Tukey's exploratory data analysis (EDA) principles revolutionized the way we approach data exploration and visualization, emphasizing the importance of understanding data's underlying structure and patterns before diving into formal statistical analysis. Tukey's Principles in EDA:

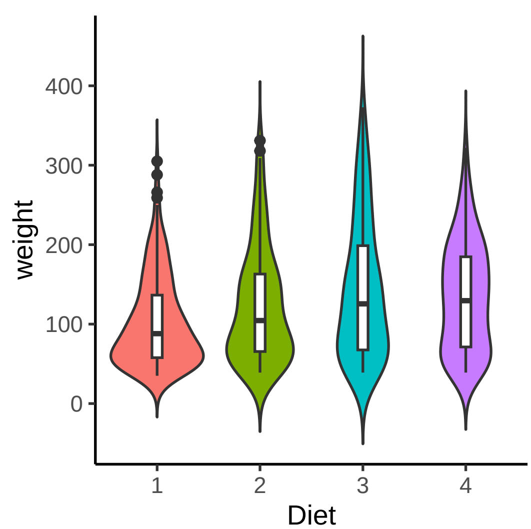
1. Through graphic exploration (looking for patterns or displaying fit), the method demonstrates things about data that a single numeric metric does not understand. This has been useful in graphing the data before you develop summary statistics.

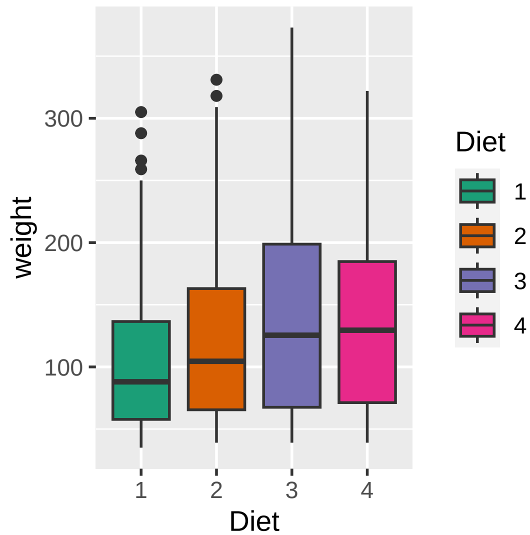
2. Describing the general patterns of the data This step should be insensitive to outliers. In general, think about the types of resistant measures (i.e., median or mean). This step makes sure to determine data patterns.

3. The natural scale or state in which the data are at their best. This will be the step at which the scale of the data can be helpful for analysis.

4. The most known part of EDA is done by assessing the fit of the data. This is taught in every statistics 101 class. With the growth of machine learning and prediction methods, residuals are now more widely used in the toolbox for assessing the best prediction models.

Tukey's Principles of EDA have become a cornerstone in the field of statistics and have been adopted by data professionals in various industries. Tukey's principles have enabled data professionals to understand complex data sets better and make more informed decisions by emphasizing the importance of visual exploration, data characterization, and model critique. In this way, Tukey's Principles have revolutionized our data analysis approach and become the foundational framework for EDA.





Data visualizations are an integral part of the EDA process, enabling analysts to discern patterns and relationships in the data that would otherwise be difficult to discern from tabular data alone. Through data visualization, analysts can quickly identify trends, outliers, and other patterns that may be missed through numerical analysis alone. Visualizations facilitate the communication of findings to non-technical stakeholders, allowing them to comprehend complex data sets more efficiently. Also, analysts can use visualizations to identify potential issues or biases in the data, resulting in better decisions and models.

Using color to represent data on maps is an example of successful graphical communication utilizing semiology. By using different colors to represent different data points, viewers can comprehend patterns and relationships in the data quickly and easily. Jacques Bertin writes in "Semiology of Graphics" that color can be used to "emphasize a point, distinguish one category from another, or establish a relationship between two points", (Monmonier, 1985). In addition, Bertin explains that the use of color can help overcome language barriers, making it easier for the audience to comprehend the presented information.

By utilizing visual elements such as bars and lines to represent data, graphs can make complex information more understandable to viewers. For instance, a line graph can be used to illustrate the change in the value of a stock over time, making it easier for investors to identify trends and patterns. Leland Wilkinson writes in his book “The Grammar of Graphics” that “graphical methods are not only superior to other forms of communication but also superior to numerical or verbal methods for certain types of data and reasoning” (Wilkinson, 2012).

It proposes that any statistical graphic can be broken down into a set of essential components, or “grammar,” that can be combined in different ways to create a wide range of visualizations, following a layered approach to describe and construct visualizations or graphics in a structured manner.

The components of the grammar of graphics include:

- Data: The raw data being visualized represents a set of observations or values.
- Aesthetic Mappings: The mapping of data variables to visual properties such as position, color, shape, and size.
- Scales: The mapping of data values to visual values, such as mapping a numerical value to a bar height.
- Geometries: The basic shapes representing the data, such as points, lines, bars, and histograms.

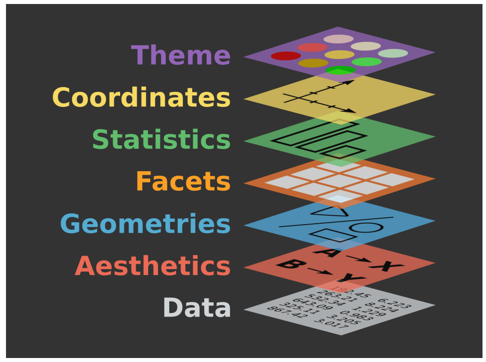


Figure 3.1: Grammar of Graphics Diagram of Wickham and Wilkinson's work

- Facets: The plot division into multiple subplots, each representing a different subset of the data.

For example, a bar chart can be created by mapping a categorical variable to the x-axis, mapping a numerical variable to bar heights, and using rectangular bars as the geometry. Moreover, mapping two numerical variables can create a scatter plot to the x and y positions and use points as the geometry. Finally, the "Grammar of Graphics" provides a systematic way of thinking about visualizations, making it easier to choose the appropriate visual representation for a given dataset.

A dashboard is a visual display of the essential information needed to achieve one or more objectives, consolidated and arranged on a single screen so the data can be monitored at a glance (Few, 2006b). Dashboard design creates visually informative and interactive interfaces that present data and key performance indicators (KPIs) in a consolidated and simple-to-understand format. The objective of KPIs are to provide users with insights and enable them to make intelligent decisions based on the presented data.
As organizations increasingly rely on data-driven decision-making, well-designed dashboards become pivotal.

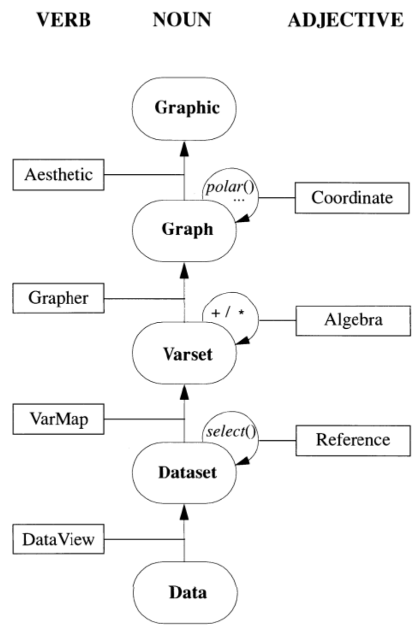


Figure 3.2: Grammar of Graphics Diagram of Wickham and Wilkinson's work

The literature on dashboard design provides a comprehensive roadmap for understanding and implementing effective dashboards, focusing on critical applications such as evaluation criteria in healthcare, learning dashboards in educational settings, design patterns and trade-offs, academic literature reviews, and practical tips for implementation. Each of these applications offers unique perspectives and actionable insights. If the data used in the design and implementation of dashboards is flawed or incomplete, it can lead to misleading insights and ineffective decision-making. Additionally, different user groups' specific needs and preferences may not always align with the applications provided.

A systematic literature review by (Schwendimann et al., 2016) discusses the state-of-the-art in learning dashboards. The paper identifies critical design features, dividing them into functional and visual features.

This study is particularly useful for educational institutions implementing learning analytics dashboards. Conducted as a systematic literature review, the study categorizes critical design features into functional and visual aspects, providing a comprehensive understanding of what makes a learning dashboard effective. By identifying key design elements and their impact on student performance, the paper is a foundational resource for educators and administrators looking to leverage dashboards to enhance educational outcomes.

3.2 Dashbaord Construction

Given that the intended audience has limitations, there are design constraints around the data, and the audience has the ability to successfully use the graphical displays of the data, what can we take from this body of research that applies to more complicated sets of graphics? How do we maintain user attention, create a desire to explore, and accurately communicate the data through the medium of an interactive data dashboard? Solutions to these questions can start with a dashboard.

A dashboard is a visual display of the essential information needed to achieve one or more objectives, consolidated and arranged on a single screen so the data can be monitored at a glance (Few, 2006b).

Dashboards can present various statistical data, such as financial performance, website traffic, or customer engagement metrics. They allow users to quickly and easily understand complex data sets by using visual elements such as charts, graphs, and tables to display the information. Additionally, statistics can be used to analyze data presented on a dashboard, providing insights into trends and patterns that can inform decision-making.

While a dashboard can be handy, it may be worth mentioning that a poorly designed dashboard will not be used. A dashboard should be concise, clear, and intuitive when displaying components in combination with a customized list of user requirements.

Much of the work done in statistical research and dashboard design involves collaboration with other researchers and users. While this may be the best for the growth of the discipline, one will find that working with collaborators with non-STEM backgrounds Dashboards can help understand and support many data types for essential business objectives. There are many different ways to label and utilize dashboards of different kinds.

Combining two compelling graphics does not necessarily result in a successful visualization. In certain instances, suboptimal combinations can result in confusion, misinterpretation, and the failure to convey the intended message. Combining two charts with distinct scales or units is an example of suboptimal graphic design, which can result in misinterpretation and flawed comparisons.
For example, if a bar chart displaying the number of sales is combined with a line chart showing revenue, meaningful comparisons between the two metrics can be challenging. According to a study conducted by Cleveland and McGill, people frequently make inaccurate judgments when comparing graphs with different scales (Cleveland & McGill, 1984).

In addition, combining two difficult-to-compare graphics with redundant visual cues or unnecessary embellishments such as colors, 3D effects, or patterns can increase cognitive load and reduce the dashboard's effectiveness. Although adding extra elements to a chart or graph may be tempting, doing so can detract from the primary message and make it more difficult for the audience to focus on the essential information. Tufte discovered that adding unnecessary visual cues to a graph decreases its effectiveness because viewers are more likely to focus on the embellishments rather than the data (Edward R. Tufte, 1985). Most importantly, dashboards should leverage people's visual and cognitive capabilities.

Static Visualization is commonly used in the communication phase of data science workflows, and data scientists sometimes use them as part of the analysis. John Tukey's EDA methods are currently known and well-vetted in the field. However, Satyanarayan et al. addressed this by introducing a high-level grammar of graphics called "Vega-Lite," which presents a set of standardized linguistic rules for producing interactive information visualizations using a concise JSON format for data to be represented by the grammar (Satyanarayan, Moritz, Wongsuphasawat, & Heer, 2016). Vega-Lite has been directly implemented in R via the `ggvis` package using the same - albeit slightly lower-level.

Understanding cognitive load is crucial for designing compelling data visualizations, as it influences how users perceive, process, and remember the data presented in the visualization. When designing visuals, it is essential to consider the cognitive load they may place on the viewer. Cognitive load is the amount of mental effort required to process information, and minimizing it can enhance a graphic's effectiveness. In addition, displaying as much raw data as possible while minimizing cognitive load can improve the graphic's clarity and precision. Here are some general guidelines for making better graphics with works from Few (**few2012?**), Tufte (**tufte1983?**), and Cairo (Cairo, 2016):

1. Keep it simple - Avoid overwhelming the viewer with too much information at once by employing a clear and concise design with minimal

distractions.

2. Use visual hierarchy - Utilize size, color, contrast, and placement to highlight important information and direct the viewer's focus.

3. Choose appropriate charts - Choose the chart type that best illustrates the data and facilitates comprehension.

4. Label clearly - Use labels that are clear and concise for axes, legends, and other essential information to avoid confusion.

5. Use data-to-ink ratio - Focus on the data by minimizing the amount of non-data ink, such as decorative elements or excessive grid lines.

6. Avoid distortion - Use appropriate scaling and avoid distortions to ensure that the graphics accurately represent the data.

7. Provide context - Add context to assist the viewer in comprehending the significance of the data and its relevance to the topic.

These principles are based on cognitive psychology and understanding how the human brain processes visual information. Cowan suggested that the average person can only hold two to six pieces of information at a time (Cowan, 2001). By applying these principles to dashboard design, designers can create visual arrangements that make it easier for viewers to understand the relationships between data elements. For example, proximity can be used to group related elements together, while symmetry can be used to create balance and harmony in the overall layout of the dashboard. At its most basic, the entire form is perceived (or emerges from our visual pathways) as opposed to its component parts. For instance, if a scatter plot and a bar chart are combined, the resulting visualization may be difficult to interpret due to the two graphics types' incompatibility. Hullman et al. (Hullman, Adar, & Shah, 2011) discovered that viewers had trouble understanding a visualization that included a scatter plot and a line chart, which brought attention to this issue. Moving forward, designers must navigate the delicate balance of complexity and comprehension to ensure that dashboards serve as a potent tool for conveying information succinctly and effectively, enhancing the viewer's ability to grasp and analyze the presented data seamlessly.

3.3 Cognitive Principles

Perception is a biological process involving sensory systems and neural mechanisms. The retina, a multi-layered tissue, lines the back of the eye and converts photons into electrical impulses that travel along the optic nerve to the brain. This process is crucial for understanding perception. Sensory modality, such as sight, hearing, and touch, has specialized receptors that convert physical stimuli into electrical signals that the brain can interpret. For instance, light enters the eye and activates photoreceptor cells in the retina, (Hubel & Wiesel, 2004).

Human perception is an essential component of data visualization that can significantly enhance both the content and quantity of displayed information (Ware, 2012). Perception refers to the organization, interpretation, and conscious experience of sensory data. Perception is also defined as "the process of recognizing (being aware of), organizing (gathering and storing), and interpreting (binding to knowledge) sensory information" (Ward, Grinstein, & Keim, 2010). Ward et al. explain the notion of perception as follows: "The

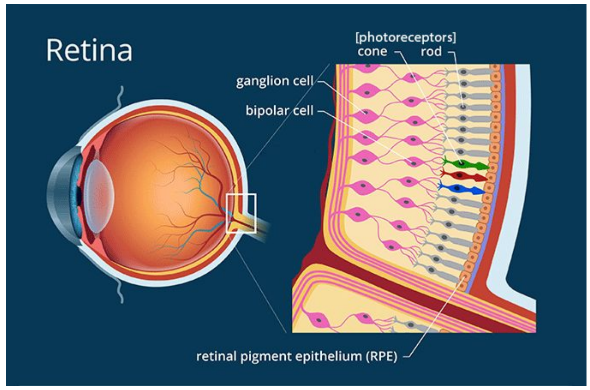


Figure 3.3: An example of how the retina signals the visual cortex

brain makes assumptions about the world to overcome the inherent ambiguity in all sensory data and in response to the task at hand."

The principles of eye-tracking involve the investigation of eye movements and fixations during visual perception. Eye-tracking technology permits researchers to monitor and record where individuals look and how their gaze traverses a visual scene. This data can be utilized to analyze patterns of attention, gaze behavior, and the sequence of fixations. The principles of eye-tracking provide valuable information regarding how individuals allocate their attention, which elements attract their gaze, and how they visually explore and process information.

Gestalt principles, on the other hand, examine how humans perceive and organize visual elements into meaningful patterns and wholes. These principles originated in the field of Gestalt psychology, which emphasized that perception is influenced by the arrangement and grouping of its constituent parts. The Gestalt principles of proximity, similarity, closure, and continuity describe how our brains organize visual stimuli to form coherent and meaningful perceptions.

Perceptual grouping is a fundamental process in visual perception that involves organizing individual graphical elements into coherent perceptual units based on their inherent properties and spatial relationships. It helps us make sense of the complex visual world by grouping elements that belong to the same object or structure and separating elements that belong to different entities. Gestalt psychologists have extensively studied the concept of perceptual grouping, proposing principles such as proximity, similarity, closure, and continuity as grouping's underlying mechanisms. These principles govern our perception of objects, edges, contours, and patterns, enabling us to perceive organized and meaningful visual information (Wertheimer, 1938), (Wagemans et al., 2012) and (Palmer, 2002).

The Gestalt principles play a crucial role in directing eye movements and cognitive processes involved in scanning scenes. The integration of scene scanning with Gestalt grouping concepts synergistically contributes to the facilitation of visual scene perception and comprehension. In the course of scene scanning, individuals frequently employ Gestalt grouping principles, such as proximity and similarity, in an unconscious manner to arrange diverse visual elements into cohesive groups. This process facilitates the rapid and efficient interpretation of intricate scenes (WERTHEIMER, 1923) and (Wertheimer, 1938). They facilitate the efficient assimilation and processing of visual information by emphasizing specific clusters or patterns in the visual field.

Scene scanning is the cognitive process of visually examining a given scene in order to acquire relevant information pertaining to the surrounding environment. The aforementioned procedure encompasses the utilization of both ocular motions and cognitive mechanisms to comprehend and analyze the visual data inherent in a given scenario. The subject matter is frequently examined within the framework of disciplines such as psychology, neuroscience, and computer vision.

The process of exploring visual scenes involves an intricate interaction of cognitive processes that influence an individual's perception and interaction with their surroundings. The saliency-driven focus is a key mechanism that guides scene scanning. This process involves both overt and covert shifts of visual attention, which are directed towards specific characteristics in the environment. Itti and Koch have extensively studied and explained this phenomena. Furthermore, the cognitive processes involved in perceiving scenes frequently involve a higher-level comprehension that entails a synergistic equilibrium between bottom-up processes driven by stimuli and top-down strategies driven by information (Itti & Koch, 2000). This theoretical perspective is supported by the work of Henderson and Hollingworth (Henderson & Hollingworth, 1999). The aforementioned phenomena are applicable to various cognitive processes such as reading, visual searches, and scene perception, as demonstrated by (Rayner, 2009). In this study, it was found that eye movements during these activities are notably impacted by individual cognitive processes, including linguistic comprehension and visual interpretation. Moreover, the interpretation and allocation of attention during scene scanning are influenced by the context in which visual items are encountered, as emphasized by (Bar, 2004). The significance of global features in the search for objects in real-world settings is crucial for guiding eye movements and attention within a contextual framework, (**torrablba2006?**).

By examining the complexities of scene scanning, it becomes evident that attention plays a crucial role in directing the eyes systematically as they analyze and comprehend the visual narratives present in each gaze. This process contributes to a complex cognitive framework that integrates perception and interpretation through an interactive and dynamic relationship.

Starting with Wertheimer's experiments on the perception of motion, he examined the phenomenon of apparent motion, which refers to the perception of motion when successive stationary stimuli are presented. Wertheimer investigated the principles underlying the perception of motion and provided insights into the Gestalt principles of visual perception by conducting a series of perceptual experiments.

(Wertheimer, 1912) investigated the phi motion, a type of motion in which the rapid presentation of two stationary stimuli creates the perception of movement. By manipulating the temporal and spatial arrangement of stimuli systematically, Wertheimer was able to identify critical factors that influenced the perception of motion. His experiments revealed that motion perception results from the interaction between sensory input and perceptual organization processes.

The work of Wertheimer emphasized the significance of perceptual grouping principles, such as proximity and similarity, in the perception of motion. He hypothesized that the visual system tends to group stimuli that are close in space or have similar properties, resulting in the perception of a continuous and smooth motion between them. The results of Wertheimer's experiments on the perception of motion were instrumental in developing Gestalt psychology, which emphasized the significance of holistic and organized perceptual experiences. Wertheimer's work influenced our understanding of how the brain processes visual information to create motion perception by laying the groundwork for subsequent studies on motion perception.

According to (Goldstein & Cacciamani, 2021), preattentive processing automatically extracts and analyzes basic features such as color, shape, orientation, and motion. These features are processed in parallel across the visual field, allowing for the rapid detection and identification of salient environmental stimuli. Preattentive processing occurs effortlessly and outside conscious awareness, laying the foundation for subsequent attentional selection and more elaborate perceptual processing.

The theories of Goldstein emphasize the significance of preattentive processes in various perceptual domains. For instance, he discusses the preattentive analysis of visual features such as color and orientation, which contribute to the visual perception of objects, scenes, and graphic patterns. In the auditory domain, preattentive processes automatically extract basic acoustic features like pitch and loudness. This makes it easier to find the source of sounds and tell them apart.

By examining preattentive processing, Goldstein's theories provide a framework for comprehending the initial stages of sensory processing and the automatic extraction of fundamental perceptual features. These theories have significant implications for understanding how automatic and controlled processes shape perception.

Perception and attention are crucial cognitive processes that allow users to interpret and make sense of data visualizations. Perception refers to the manner in which we interpret and organize sensory information from our environment, whereas attention refers to the capacity to selectively focus on particular aspects of this information (McCallum, n.d.). Expectations of perception and attention are important in data visualization interactions, however expertise is the in-depth knowledge and skills that come from having a lot of experience and learning over a long period of time.

In addition to perception and focus, domain-specific knowledge is essential for understanding and interacting with data visualizations. Expertise in a particular field can enable individuals to better interpret and comprehend the significance of the presented data, as well as identify potential biases or errors in the visualization. Thus, the ability to perceive and interact with data visualizations requires a combination of perceptual and attentional processes, as well as domain-specific knowledge, to interpret and comprehend the presented information. This suggests that data visualization involves the misuse of human visual perception in the visual presentation of data. Assigning meaning to visualization is not a statistical or computational step but a cognitive one. Each step in the data analysis process is part of a more extensive mental process of constructing meaning with important cognitive-based concepts.

Short-term memory (STM), which is often referred to as working memory, represents a cognitive stage characterized by the brief storage and processing of information, requiring significant cognitive resources for memory preservation. In the publication titled "Working memory: Theories, models, and controversies" authored by Alan Baddeley, it is asserted that short-term memory (STM) is a system with restricted capacity that is susceptible to both interference and decay (A. Baddeley, 2012). Selective attention plays a crucial role in the preservation of short-term memory (STM) since it enables individuals to effectively filter out extraneous information and focus on pertinent stimuli (Cowan, 2001). According to (Alvarez & Cavanagh, 2004), the utilization of visual aids, such as charts and diagrams, has the potential to enhance short-term memory by facilitating more efficient encoding and retention of information. The utilization of visual aids, such as charts, has the potential to increase our short-term memory. Moreover, annotations can also serve to facilitate short-term memory. According to (Alvarez & Cavanagh, 2004), the act of incorporating annotations, such as notes or highlights, to the information we aim to retain can enhance our ability to recall the information at a later time.

According to the Feature Integration Theory (FIT), STM is composed of two stages: pre-attentive processing and focused attention (A. Treisman, 1998). Parallel and independently, the brain processes the physical characteristics of an object, such as its color, shape, and orientation, during pre-attentive processing. However, focused attention is required to bind these features into a coherent object representation in STM. STM can be improved through various strategies, such as rehearsal, chunking, and elaboration (Oberauer, 2009). For example, by repeating a phone number several times or breaking it down into chunks of two or three digits, we can increase the likelihood of it being stored in STM. Similarly, by elaborating on the information we want to remember, such as creating mental associations or visual images, we can enhance its retention in STM (Bui & Myerson, 2014).

STM is a dynamic and malleable cognitive system that is crucial to our daily lives. Understanding the mechanisms underlying STM and how to improve it can have significant implications for learning, memory, and the treatment of memory disorders. By analyzing the relationship between attention and working memory, we can gain insight into how we construct meaning from the information in our environment.

Gestalt psychology indicates that humans actively construct meaning by organizing information into patterns and wholes (Wertheimer, 1938). Both top-down and bottom-up processing are involved in the process of meaning construction. Bottom-up processing entails analyzing sensory data from the environment and constructing perceptions based on this data. Top-down processing reflects the influence of prior knowledge, expectations, and context on the perception and interpretation of incoming sensory data.

Together, top-down and bottom-up processing facilitate the encoding and retrieval of information from STM. Selective attention, which is the ability to focus on relevant information while ignoring irrelevant information, is an example of top-down processing that aids in the encoding and retrieval of information in STM (Cowan, 2010).
According to FIT, perceiving objects involves both the bottom-up analysis of individual features and the top-down processing of higher-level features in order to form a complete perception (A. M. Treisman & Gelade, 1980).

The Gestalt principles of perception address how humans construct meaning from sensory data through both bottom-up and top-down processing. Both types of processing are involved in encoding and retrieving information, which has significant implications for understanding the mechanisms of STM.

Baddeley expanded our understanding of working memory by emphasizing its active processing nature, expanding upon the model of Atkinson and Shiffrin's Information Model developed in the 1968, which emphasizes the process of encoding, storage, and retrieval. Unless actively practiced, short-term memory has a limited capacity and a short duration of retention. If infor-

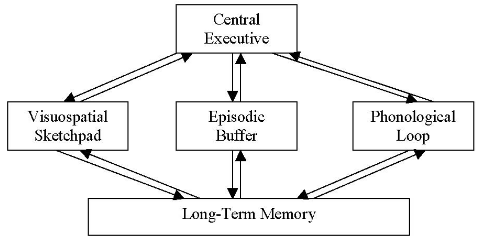


Figure 3.4: Working Memory Model created by Baddeley (left) and Information Processing Model created by Atkinson and Shiffrin (right)

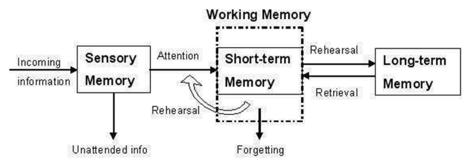


Figure 3.5: Working Memory Model created by Baddeley (left) and Information Processing Model created by Atkinson and Shiffrin (right)

mation is deemed significant or sufficiently rehearsed, it can be encoded and transferred to long-term memory, which has an almost unlimited capacity and long-term storage. The influential model developed by Baddeley, known as the working memory model, proposed a more complex structure with multiple components, (Baddeley Alan D., 1976).

The Baddeley Memory Model is an updated and influential model of working memory. It includes the phonological loop (maintenance of verbal information), the visuospatial sketchpad (maintenance of visual and spatial information), the central executive (attentional control), and the episodic buffer (integrated storage). Together, these components facilitate the active processing and temporary storage of data in working memory.

Visual and spatial information is processed and temporarily stored in the visuospatial sketchpad when individuals view statistical graphics. The central executive component facilitates the interpretation and analysis of the presented data by directing attention to pertinent aspects of the graphic. Utilizing working memory resources effectively can aid in comprehending and remembering the statistical information conveyed by the graphics.

Together, Baddeley's model of working memory provided a comprehensive framework for studying memory and cognition. They contributed to the understanding of how information is processed, encoded, stored, and retrieved in human memory systems, laying the groundwork for subsequent research and theories in cognitive psychology.

3.4 Ensemble Perception

Ensemble perception is the cognitive ability to quickly derive summary statistics from sets of similar items (Chong & Treisman, 2003). This component of visual cognition enables the visual system to summarize and describe collections of comparable objects or features effectively. This capability is often activated at a glance and is crucial for making sense of complex visual scenes.

David Whitney's fundamental review outlines the core principles of ensemble perception, emphasizing the extraction of summary statistical information from groups of similar objects (Whitney, Haberman, & Sweeny, 2014) and (Dakin & Watt, 1997). Ensemble perception pertains to the capacity of the human visual system to efficiently and swiftly extract statistical data from a collection of items or objects, as opposed to separately processing each item. The application of ensemble perception can be observed through the utilization of Gestalt principles of visual perception, including proximity (referring to the nearness of objects), similarity (pertaining to similar appearance), and closure (denoting the completion of shapes). These concepts facilitate the perception of patterns and groups in visual scenes without necessitating a detailed examination of each individual piece.

Huberman et al. further discussed that individual differences exist in ensemble perception capabilities, indicating the presence of multiple, independent levels of ensemble representation (J. Haberman, Brady, & Alvarez, 2015). Recent studies have expanded on these principles.
Khayat et al.'s work explores how ensemble perception can create a unified perception from groups of similar objects and also delves into the implicit perception and memory of set statistics (Khayat, Ahissar, & Hochstein, 2023) and (Khayat, Fusi, & Hochstein, 2021). Examples of perception and memory of set statistics in visual information can best be described when you look at a field of flowers with various colors. Your perception quickly summarizes the set statistics of color, such as the overall color distribution, or in a crowd of people, your perception helps you group individuals into categories based on shared characteristics like clothing color, height, or age. In brief, the cognitive processes of perceiving and remembering statistical information in visual stimuli are crucial for efficiently comprehending and retaining the statistical characteristics of visual scenes and objects. These mechanisms play a significant role in enhancing our comprehension and analysis of the visual environment.

The study "Perceptual History Biases in Serial Ensemble Representation" by Khayat et al. focuses on ensemble perception, explicitly examining how past visual experiences influence the perception of current visual ensembles.
The study investigates the serial dependence of ensemble perception when each ensemble set is presented simultaneously but spatially distributed over

the screen.

This suggests that the objects and our prior experiences with similar ensembles impact how we perceive groups of similar objects (Khayat et al., 2023).

This adds a layer of complexity to the understanding of ensemble perception, which is generally considered the visual system's ability to summarize groups of similar objects into a unified perception efficiently.

The study "Perceiving ensemble statistics of novel image sets" by Khayat et al. focuses on how the human visual system perceives summary statistics of sets of stimulus elements.

The study is particularly interested in how we perceive novel image sets and hypothesizes that our capacity to summarize statistical data from these sets affects how well we can comprehend and interpret new visual information (Khayat et al., 2021). This research contributes to the broader field of ensemble perception, which explores how the visual system can efficiently represent groups of similar objects as a unified perception.

The study implies that not only can the visual system quickly grasp the "gist" or essence of familiar visual ensembles, but it can also do so for novel sets of images.

This ability to quickly summarize statistical information from new visual stimuli could be a fundamental feature of human perception (Khayat & Hochstein, 2018).

Other research has investigated the role of ensemble perception in both high- and low-level visual information, such as emotion and brightness, and how it can even operate when scene information is incomplete (Chakrabarty & Wada, 2020) and (J. M. Haberman & Ulrich, 2019).

Furthermore, ensemble perception is not just a specialized function but a pervasive aspect of visual perception.

It has been discussed holistically to engage a general audience and has been shown to condense redundant information into summary statistical representations (Corbett, Utochkin, & Hochstein, 2023) and (Whitney & Manassi, 2022).

Stable ensemble representations have been found to facilitate visual search, even when they are not predictive of a target location (Utochkin, Choi, & Chong, 2023). The study focuses on a coding model that emphasizes the crucial role of the "pooling layer" in ensemble perception.

3.4.1 Ensemble Visualization

Ensemble visualization refers to the graphical representation of ensemble data, which typically consists of multiple related datasets. It aims to provide a comprehensive view that allows for rapidly extracting visual statistics and insights about distributed information.

Ensemble visualization techniques can vary, including Ensemble Surface Slicing (ESS) to visualize regions of similarity and difference among surfaces and statistical visualization techniques to identify areas of interest quickly (Alabi et al., 2012) and (Potter et al., 2009).

The aim is to provide a comprehensive view that allows for the rapid extraction of visual statistics and insights about distributed information.

By combining various statistical visualization techniques, Potter presents the Ensemble-Vis framework in his paper as a way for scientists to identify areas of interest quickly, (Potter et al., 2009). The paper's thorough methodology, which combines various statistical visualization techniques to build a more helpful ensemble visualization framework, makes it noteworthy. This framework has been referenced by many other studies, leading to the development of new visualization methods.

By exploring the effects of ensemble and summary displays on data interpretation, Padilla's research adds valuable insights to the field, (Padilla, Ruginski, & Creem-Regehr, 2017). Padilla investigates how ensemble displays, which plot multiple data points on the same Cartesian coordinate plane, affect the viewer's understanding and interpretation of the data. The paper contributes to the theory by examining the cognitive aspects of ensemble visualization. However, a counterexample to the effectiveness of ensemble displays can be observed when the data points are highly overlapping and densely packed, leading to visual clutter and difficulty distinguishing individual data points. In such cases, ensemble displays may hinder the viewer's understanding and interpretation of the data rather than enhance it. This highlights the importance of considering the specific characteristics of the presented data when deciding whether to use ensemble displays. Additionally, summary displays, which provide an overview of the data, can be a helpful alternative in situations where ensemble displays may be less effective. Summary displays condense the data into crucial statistics or visual representations, allowing for more straightforward interpretation and analysis, especially when dealing with complex or dense datasets. Further research is needed to explore the optimal conditions for ensemble and summary displays to maximize their effectiveness in data interpretation.

3.4.2 Multidimensional Ensembles

Initial research on ensemble perception primarily focused on one-dimensional ensembles, where summary statistics are extracted from a single feature or dimension (J. Haberman et al., 2015). For example, in a study on facial expression perception, participants were presented with an ensemble of faces displaying different emotions. In this case, the pooling layer would analyze the overall emotional expression of the ensemble, summarizing the various individual facial expressions into one general emotion perception. This model allows researchers to understand how humans perceive and interpret complex emotional expressions more systematically. However, as Maule and Franklin note, real-world scenes often consist of complex, multidimensional attributes, and research has gradually shifted towards understanding how the human visual system processes these more intricate ensembles (Maule & Franklin, 2015). Research on multidimensional ensembles has explored how people simultaneously perceive summary statistics across multiple attributes, such as size and color.

Darkin and Watt were among the first to explore how orientation statistics are computed from visual textures, extending the concept of ensemble perception into a multidimensional setting, (Dakin & Watt, 1997). Haberman et al. expanded this research by showing that individual differences exist in ensemble perception capabilities, suggesting multiple, independent levels of ensemble representation exist, (J. Haberman et al., 2015). The existing literature on multidimensional ensembles in visual information covers various topics, from neuroscience to data visualization.

For instance, studies have explored the role of neuronal ensembles in control-

ling visually guided behavior and their influence on visual working memory (Carrillo-Reid, Han, Yang, Akrouh, & Yuste, 2019). Research has also delved into the use of aggregated plots for multidimensional visual analysis, although these don't explicitly mention ensembles (Fofonov & Linsen, 2018).
Fast ensemble representations have been investigated to understand high-level perceptual impressions based on visual information (Leib, Kosovicheva, & Whitney, 2016).

Additionally, the aesthetic complexity of visual information has been quantified using information theory, offering a potential framework for ensemble-based representations (Karjus, Solà, Ohm, Ahnert, & Schich, 2023). But a detailed example of why you shouldn't use aggregated plots for multidimensional visual analysis is when the individual data points in the ensemble show significant differences. Data aggregation may obscure important patterns in such scenarios and lead to misleading interpretations. Further, ensembles may fail to capture the fine-grained details and nuances present in the individual plots, compromising the overall accuracy and precision of the analysis. The long-term stability of neuronal ensembles in the visual cortex has been studied, shedding light on their potential role in visual perception (Pérez-Ortega, Alejandre-García, & Yuste, 2021). Ensemble visualization techniques, particularly in computer simulations, have also been reviewed (Afzal et al., 2019). Lastly, the structure of neural networks has been shown to affect working memory, which could have implications for visual ensembles (Leavitt, Pieper, Sachs, & Martinez-Trujillo, 2017).

Chapter 4

Does Color Help?: Multidimensional Ensembles in Bland-Altman Plots: An Exploration

4.1 Introduction:

"How can multidimensional ensembles be effectively incorporated into Bland-Altman plots to enhance the accuracy and comprehensibility of assessing agreement between multiple measurement techniques?" This research question explores the theoretical and practical implications of incorporating ensemble perception and multidimensional data into Bland-Altman plots, traditionally used for comparing two measurement methods. By doing so, the research seeks to advance the utility and interpretability of these plots in complex, real-world scenarios where multiple dimensions are often at play. The interpretation of Bland-Altman plots is conventionally one-dimensional, focusing primarily on the mean difference and limits of agreement (Bland & Altman, 1986). However, in real-world applications, the measurements under consideration often contain multiple dimensions that could contribute to interpreting the agreement or disagreement between two techniques (Chong & Treisman, 2003). A detailed counterexample of the utility and interpretability of Bland-Altman plots in complex, real-world scenarios can be seen where two techniques are being compared for measuring blood pressure. The Bland-Altman plot may show good agreement between the mean difference and the limits of agreement, suggesting high concordance. However, when considering additional dimensions such as accuracy and precision in different subgroups (e.g., age, gender), it could reveal significant discrepancies and limitations in the interpretation of Ensemble coding, a perceptual mechanism that provides a statistical summary of a visual scene, offers a promising solution by facilitating the rapid extraction of variability information (Alvarez, 2011). The theory of ensemble perception offers a framework for understanding how these multiple dimensions could be processed simultaneously (J. Haberman et al., 2015) and (Whitney et al., 2014).

4.1.1 Ensemble Perception

Ensemble perception is the cognitive ability to quickly derive summary statistics from sets of similar items (Chong & Treisman, 2003). This component of visual cognition enables the visual system to summarize and describe collections of comparable objects or features effectively. This capability is often activated at a glance and is crucial for making sense of complex visual scenes.

David Whitney's fundamental review outlines the core principles of ensemble perception, emphasizing the extraction of summary statistical information from groups of similar objects (Whitney et al., 2014) and (Dakin & Watt, 1997).

(J. Haberman et al., 2015) further discussed that individual differences exist in ensemble perception capabilities, indicating the presence of multiple, independent levels of ensemble representation.

Recent studies have expanded on these principles.

Khayat et al.'s work explores how ensemble perception can create a unified perception from groups of similar objects and also delves into the implicit perception and memory of set statistics (Khayat et al., 2023) and (Khayat et al., 2021).

The study "Perceptual History Biases in Serial Ensemble Representation" by Khayat et al. focuses on ensemble perception, explicitly examining how past visual experiences influence the perception of current visual ensembles.

The study investigates the serial dependence of ensemble perception when each ensemble set is presented simultaneously but spatially distributed over the screen.

This suggests that the objects and our prior experiences with similar ensembles impact how we perceive groups of similar objects (Khayat et al., 2023). This adds a layer of complexity to the understanding of ensemble perception, which is generally considered the visual system's ability to summarize groups of similar objects into a unified perception efficiently.

The study "Perceiving ensemble statistics of novel image sets" by Khayat et al. focuses on how the human visual system perceives summary statistics of sets of stimulus elements.

The study is particularly interested in how we perceive novel image sets and hypothesizes that our capacity to summarize statistical data from these sets affects how well we can comprehend and interpret new visual information (Khayat et al., 2021). This research contributes to the broader field of ensemble perception, which explores how the visual system can efficiently represent groups of similar objects as a unified perception.

The study implies that not only can the visual system quickly grasp the "gist" or essence of familiar visual ensembles, but it can also do so for novel sets of images.

This ability to quickly summarize statistical information from new visual stimuli could be a fundamental feature of human perception (Khayat & Hochstein, 2018).

Other research has investigated the role of ensemble perception in both high- and low-level visual information, such as emotion and brightness, and how it can even operate when scene information is incomplete (Chakrabarty & Wada, 2020) and (J. M. Haberman & Ulrich, 2019).

Furthermore, ensemble perception is not just a specialized function but a pervasive aspect of visual perception.

It has been discussed holistically to engage a general audience and has been shown to condense redundant information into summary statistical representations (Corbett et al., 2023) and (Whitney & Manassi, 2022).

Lastly, stable ensemble representations have been found to facilitate visual search, even when they are not predictive of a target location (Utochkin et al., 2023). The study focuses on a coding model that emphasizes the crucial role of the "pooling layer" in ensemble perception.

Ensemble perception refers to the ability of the visual system to summarize information from a group of similar objects. The "pooling layer" in the model likely serves as a computational mechanism for aggregating or summarizing this information, potentially providing insights into how the brain processes complex visual scenes. The study aims to provide a more structured understanding of ensemble perception by introducing a model highlighting the importance of a specific computational layer, known as the "pooling layer," in summarizing visual information.

4.1.2 Multidimensional Ensembles

Initial research on ensemble perception primarily focused on one-dimensional ensembles, where summary statistics are extracted from a single feature or dimension (J. Haberman et al., 2015). For example, in a study on facial expression perception, participants were presented with an ensemble of faces displaying different emotions. In this case, the pooling layer would analyze the overall emotional expression of the ensemble, summarizing the various individual facial expressions into one general emotion perception. This model allows researchers to understand how humans perceive and interpret complex emotional expressions more systematically. However, as (Maule & Franklin, 2015) notes, real-world scenes often consist of complex, multidimensional attributes, and research has gradually shifted towards understanding how the human visual system processes these more intricate ensembles. Research on multidimensional ensembles has explored how people simultaneously perceive summary statistics across multiple attributes, such as size and color.

(Dakin & Watt, 1997) were among the first to explore how orientation statistics are computed from visual textures, extending the concept of ensemble perception into a multidimensional setting. (J. Haberman et al., 2015) expanded this research by showing that individual differences exist in ensemble perception capabilities, suggesting multiple, independent levels of ensemble representation exist. The existing literature on multidimensional ensembles in visual information covers various topics, from neuroscience to data visualization.

For instance, studies have explored the role of neuronal ensembles in controlling visually guided behavior and their influence on visual working memory (Carrillo-Reid et al., 2019). Research has also delved into the use of aggregated plots for multidimensional visual analysis, although these don't explicitly mention ensembles (Fofonov & Linsen, 2018).

Fast ensemble representations have been investigated to understand high-level perceptual impressions based on visual information (Leib et al., 2016).

Additionally, the aesthetic complexity of visual information has been quantified using information theory, offering a potential framework for ensemble-based representations (Karjus et al., 2023). But a detailed example of why you shouldn't use aggregated plots for multidimensional visual analysis is when the individual data points in the ensemble show significant differences. Data aggregation may obscure important patterns in such scenarios and lead to misleading interpretations. Further, ensembles may fail to capture the fine-grained details and nuances present in the individual plots, compromising the overall accuracy and precision of the analysis. The long-term stability of neuronal ensembles in the visual cortex has been studied, shedding light on their potential role in visual perception (Pérez-Ortega et al., 2021). Ensemble visualization techniques, particularly in computer simulations, have also been reviewed (Afzal et al., 2019). Lastly, the structure of neural networks has been shown to affect working memory, which could have implications for visual ensembles (Leavitt et al., 2017).

4.1.3 Bland-Altman Plots

The study of multidimensional ensembles has seen applications in the field of data visualization. (Szafir, 2017) showed that understanding color differences could improve the design of visualizations that require the viewer to integrate multiple pieces of information. This work suggested that effective visualization tools could be designed by leveraging the human ability to process multidimensional ensembles rapidly.

Bland-Altman Plots are widely used in clinical research for assessing the agreement between two measurement techniques by plotting the differences against the means (Bland & Altman, 1986). They offer a straightforward representation of data, making them a popular choice for visualizing measurement bias. However, the plots are conventionally one-dimensional, primarily focusing on mean differences and limits of agreement.

4.1.4 The intersection of Multidimensional Ensembles and Bland-Altman Plots

The application of multidimensional ensembles in Bland-Altman Plots could potentially offer a more nuanced understanding of data. For example, integrating color coding to indicate standard deviation and shape variations to indicate skewness could offer additional layers of information in a single plot (Szafir, 2017). Such an approach could leverage our innate abilities in ensemble perception to offer a more comprehensive assessment of agreement between multiple sets of measurement techniques (Szafir, Haroz, Gleicher, & Franconeri, 2016).

4.1.5 Gaps in Research

While the fields of ensemble perception and data visualization have seen significant growth, there is a lack of research focusing on the application of ensemble perception, particularly multidimensional ensembles, in Bland-Altman Plots. This gap points to the need for empirical studies designed to validate theoretical frameworks and to assess the practical utility of incorporating multidimensional ensembles into Bland-Altman Plots.

4.2 Methods:

Ensemble perception involves the rapid and often unconscious extraction of summary statistics from a set of similar items (Chong & Treisman, 2003), (Whitney et al., 2014). In the context of multidimensional ensembles, this would refer to the simultaneous extraction of various features such as mean, variance, and other statistical attributes across multiple dimensions (Maule & Franklin, 2015), (J. Haberman et al., 2015). Understanding this concept will provide a unique way to interpret Bland-Altman plots that contain data from multiple dimensions.

Traditionally, Bland-Altman plots present the difference between two sets of measurements against their mean, providing a graphical representation of agreement or bias. However, this one-dimensional representation might not capture the full complexity of real-world data, where measurements can often be multidimensional (Dakin & Watt, 1997). Given the human brain's ability to rapidly process multidimensional ensemble statistics (J. Haberman et al., 2015), incorporating this aspect into Bland-Altman plots might yield a more nuanced interpretation (Bauer, 2009), (Szafir et al., 2016).

One way to incorporate multidimensionality into Bland-Altman plots is by adding layers that represent different statistical attributes. For example, varying shades of color could indicate the standard deviation within each plotted point, and shape variations could indicate another dimension such as skewness or kurtosis (Szafir, 2017). This model would require human observers to simultaneously extract multiple summary statistics, leveraging the brain's capabilities in ensemble perception (Chong & Treisman, 2003).

In this proposed study, we utilize color coding to represent varying levels of variability in Bland-Altman plots. To ensure the universal interpretability of the plots (Ware, 2012), the color palette will be selected with care, taking into account potential issues such as color blindness.

For this study, two Bland-Altman plots will be generated. A traditional plot without color-coding and a plot using the color-coding technique we propose. Both plots will be presented to a group of participants that includes both experts and non-experts in data interpretation and statistics. The participants will be required to interpret the plots and complete a questionnaire to assess their comprehension and speed of interpretation.

For a layer of interactivity in our study, we will generate interactive Bland-Altman plots incorporating ensemble coding of variability through color gradations. Interactivity will be implemented via D3, providing detailed information about each data point when hovered over, as well as zoom features allowing users to zero in on areas of interest.

4.2.1 Design of the User Studies

User studies will be designed to assess the effectiveness of the interactive Bland-Altman plots in conveying information about the variability of data points. The studies will consist of two parts:

Task-based Evaluation: Participants will be given a set of tasks to complete using both the interactive color-coded Bland-Altman plot and a traditional static Bland-Altman plot. Tasks will involve identifying specific data points, interpreting data variability, and answering questions about the overall data trend. Metrics such as task completion time, success rate, and error rate will be recorded (Rubin & Chisnell, 2008).

Subjective Evaluation: After completing the tasks, participants will be asked to fill out a questionnaire assessing their user experience. The questionnaire will include items related to the perceived ease of use, satisfaction, and preference between the traditional and interactive color-coded Bland-Altman plots.

4.3 Discussion:

We anticipate that the application of ensemble coding for variability will aid in the comprehension of Bland-Altman plots. Based on ensemble coding principles, the color-coded plot should enable faster and more accurate interpretation of data variability (J. Haberman & Whitney, 2012). This method has the potential to enhance the interpretability and utility of these graphs, making them accessible to a broader audience and facilitating more efficient data communication.

Conclusion

If we don't want Conclusion to have a chapter number next to it, we can add the `{-}` attribute.

More info

And here's some other random info: the first paragraph after a chapter title or section head *shouldn't be* indented, because indents are to tell the reader that you're starting a new paragraph. Since that's obvious after a chapter or section title, proper typesetting doesn't add an indent there.

Appendix A

The First Appendix

This first appendix includes all of the R chunks of code that were hidden throughout the document (using the `include = FALSE` chunk tag) to help with readibility and/or setup.

In the main Rmd file

```
library(knitr)
library(palmerpenguins)
library(tidyverse)
library(nycflights13)
data(flights)

library(ggpcp)
library(ggplot2)
library(dplyr)
data(nasa)

library(scales)
library(datasets)
data("ChickWeight")
library(formatR)
```

In Chapter 4:

Appendix B

The Second Appendix, for Fun

Colophon

This document is set in EB Garamond, Source Code Pro and Lato. The body text is set at 11pt with *lmr*.

It was written in R Markdown and LaTeX, and rendered into PDF using huskydown and bookdown.

This document was typeset using the XeTeX typesetting system, and the University of Washington Thesis class class created by Jim Fox. Under the hood, the University of Washington Thesis LaTeX template is used to ensure that documents conform precisely to submission standards. Other elements of the document formatting source code have been taken from the Latex, Knitr, and RMarkdown templates for UC Berkeley's graduate thesis, and Dissertate: a LaTeX dissertation template to support the production and typesetting of a PhD dissertation at Harvard, Princeton, and NYU

The source files for this thesis, along with all the data files, have been organised into an R package, xxx, which is available at `https://github.com/xxx/xxx`. A hard copy of the thesis can be found in the University of Washington library.

This version of the thesis was generated on 2023-10-14 21:06:47. The repository is currently at this commit:

The computational environment that was used to generate this version is as follows:

```
## - Session info -------------------------------------------
##  setting  value
##  version  R version 4.2.2 (2022-10-31)
##  os       macOS Big Sur ... 10.16
##  system   x86_64, darwin17.0
##  ui       X11
##  language (EN)
##  collate  en_US.UTF-8
##  ctype    en_US.UTF-8
##  tz       America/Detroit
##  date     2023-10-14
##  pandoc   2.19.2 @ /Applications/RStudio.app/Contents/Resources/app/quarto/bin/tools/ (
##
## - Packages -----------------------------------------------
##  package       * version date (UTC) lib source
##  assertthat      0.2.1   2019-03-21 [1] CRAN (R 4.2.0)
##  bookdown        0.33    2023-03-06 [1] CRAN (R 4.2.0)
##  cachem          1.0.7   2023-02-24 [1] CRAN (R 4.2.0)
##  callr           3.7.3   2022-11-02 [1] CRAN (R 4.2.0)
##  cli             3.6.1   2023-03-23 [1] CRAN (R 4.2.0)
##  colorspace      2.1-0   2023-01-23 [1] CRAN (R 4.2.0)
##  crayon          1.5.2   2022-09-29 [1] CRAN (R 4.2.0)
##  devtools        2.4.5   2022-10-11 [1] CRAN (R 4.2.0)
##  digest          0.6.31  2022-12-11 [1] CRAN (R 4.2.0)
##  dplyr         * 1.1.2   2023-04-20 [1] CRAN (R 4.2.0)
##  ellipsis        0.3.2   2021-04-29 [1] CRAN (R 4.2.0)
##  evaluate        0.21    2023-05-05 [1] CRAN (R 4.2.0)
##  fansi           1.0.4   2023-01-22 [1] CRAN (R 4.2.0)
```

```
## farver          2.1.1   2022-07-06 [1] CRAN (R 4.2.0)
## fastmap         1.1.1   2023-02-24 [1] CRAN (R 4.2.0)
## forcats       * 1.0.0   2023-01-29 [1] CRAN (R 4.2.0)
## formatR       * 1.14    2023-01-17 [1] CRAN (R 4.2.0)
## fs              1.6.2   2023-04-25 [1] CRAN (R 4.2.0)
## generics        0.1.3   2022-07-05 [1] CRAN (R 4.2.0)
## ggpcp         * 0.2.0   2022-11-28 [1] CRAN (R 4.2.0)
## ggplot2       * 3.4.2   2023-04-03 [1] CRAN (R 4.2.0)
## glue            1.6.2   2022-02-24 [1] CRAN (R 4.2.0)
## gtable          0.3.3   2023-03-21 [1] CRAN (R 4.2.0)
## hms             1.1.3   2023-03-21 [1] CRAN (R 4.2.0)
## htmltools       0.5.4   2022-12-07 [1] CRAN (R 4.2.0)
## htmlwidgets     1.6.2   2023-03-17 [1] CRAN (R 4.2.0)
## httpuv          1.6.9   2023-02-14 [1] CRAN (R 4.2.0)
## knitr         * 1.42    2023-01-25 [1] CRAN (R 4.2.0)
## labeling        0.4.2   2020-10-20 [1] CRAN (R 4.2.0)
## later           1.3.0   2021-08-18 [1] CRAN (R 4.2.0)
## lifecycle       1.0.3   2022-10-07 [1] CRAN (R 4.2.0)
## lubridate     * 1.9.2   2023-02-10 [1] CRAN (R 4.2.0)
## magrittr        2.0.3   2022-03-30 [1] CRAN (R 4.2.0)
## memoise         2.0.1   2021-11-26 [1] CRAN (R 4.2.0)
## mime            0.12    2021-09-28 [1] CRAN (R 4.2.0)
## miniUI          0.1.1.1 2018-05-18 [1] CRAN (R 4.2.0)
## munsell         0.5.0   2018-06-12 [1] CRAN (R 4.2.0)
## nycflights13  * 1.0.2   2021-04-12 [1] CRAN (R 4.2.0)
## palmerpenguins * 0.1.1  2022-08-15 [1] CRAN (R 4.2.0)
## pillar          1.9.0   2023-03-22 [1] CRAN (R 4.2.2)
## pkgbuild        1.4.0   2022-11-27 [1] CRAN (R 4.2.0)
## pkgconfig       2.0.3   2019-09-22 [1] CRAN (R 4.2.0)
## pkgload         1.3.2   2022-11-16 [1] CRAN (R 4.2.0)
## prettyunits     1.1.1   2020-01-24 [1] CRAN (R 4.2.0)
```

```
## processx        3.8.1   2023-04-18 [1] CRAN (R 4.2.0)
## profvis         0.3.7   2020-11-02 [1] CRAN (R 4.2.0)
## promises        1.2.0.1 2021-02-11 [1] CRAN (R 4.2.0)
## ps              1.7.5   2023-04-18 [1] CRAN (R 4.2.0)
## purrr         * 1.0.1   2023-01-10 [1] CRAN (R 4.2.0)
## R6              2.5.1   2021-08-19 [1] CRAN (R 4.2.0)
## RColorBrewer    1.1-3   2022-04-03 [1] CRAN (R 4.2.0)
## Rcpp            1.0.10  2023-01-22 [1] CRAN (R 4.2.0)
## readr         * 2.1.4   2023-02-10 [1] CRAN (R 4.2.0)
## remotes         2.4.2   2021-11-30 [1] CRAN (R 4.2.0)
## rlang           1.1.1   2023-04-28 [1] CRAN (R 4.2.0)
## rmarkdown       2.20    2023-01-19 [1] CRAN (R 4.2.2)
## rstudioapi      0.14    2022-08-22 [1] CRAN (R 4.2.0)
## scales        * 1.2.1   2022-08-20 [1] CRAN (R 4.2.0)
## sessioninfo     1.2.2   2021-12-06 [1] CRAN (R 4.2.0)
## shiny           1.7.4   2022-12-15 [1] CRAN (R 4.2.0)
## stringi         1.7.12  2023-01-11 [1] CRAN (R 4.2.0)
## stringr       * 1.5.0   2022-12-02 [1] CRAN (R 4.2.0)
## tibble        * 3.2.1   2023-03-20 [1] CRAN (R 4.2.0)
## tidyr         * 1.3.0   2023-01-24 [1] CRAN (R 4.2.0)
## tidyselect      1.2.0   2022-10-10 [1] CRAN (R 4.2.0)
## tidyverse     * 2.0.0   2023-02-22 [1] CRAN (R 4.2.0)
## timechange      0.2.0   2023-01-11 [1] CRAN (R 4.2.0)
## tzdb            0.3.0   2022-03-28 [1] CRAN (R 4.2.0)
## urlchecker      1.0.1   2021-11-30 [1] CRAN (R 4.2.0)
## usethis         2.1.6   2022-05-25 [1] CRAN (R 4.2.0)
## utf8            1.2.3   2023-01-31 [1] CRAN (R 4.2.0)
## vctrs           0.6.2   2023-04-19 [1] CRAN (R 4.2.0)
## withr           2.5.0   2022-03-03 [1] CRAN (R 4.2.0)
## xfun            0.37    2023-01-31 [1] CRAN (R 4.2.0)
## xtable          1.8-4   2019-04-21 [1] CRAN (R 4.2.0)
```

```
##  yaml              2.3.7   2023-01-23 [1] CRAN (R 4.2.0)
##
##  [1] /Library/Frameworks/R.framework/Versions/4.2/Resources/library
##
## ----------------------------------------------------------
```

References

Afzal, S., Hittawe, M. M., Ghani, S., Jamil, T., Knio, O., Hadwiger, M., & Hoteit, I. (2019). The state of the art in visual analysis approaches for ocean and atmospheric datasets. In *Computer graphics forum* (Vol. 38, pp. 881–907). Wiley Online Library.

Alabi, O. S., Wu, X., Harter, J. M., Phadke, M., Pinto, L., Petersen, H., et al.others. (2012). Comparative visualization of ensembles using ensemble surface slicing. In *Visualization and data analysis 2012* (Vol. 8294, pp. 318–329). SPIE.

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131.

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*(2), 106–111.

Aspillaga, M. (1996). Perceptual foundations in the design of visual displays. *Computers in Human Behavior*, *12*(4), 587–600.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, Alan D. (1976). *The psychology of memory / alan d. baddeley* (pp. xvii, 430 p. :). Book, Basic Books New York.

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, *5*(8), 617–629.

Bauer, B. (2009). Does stevens's power law for brightness extend to perceptual brightness averaging? *The Psychological Record*, *59*, 171–185.

Bendix, F., Kosara, R., & Hauser, H. (2005). Parallel Sets: Visual Analysis of Categorical Data. In *2005 IEEE symposium on information visualization (INFOVIS'05)* (pp. 133–140). IEEE. http://doi.org/10.1109/INFOVIS.2005.27

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310.

Bui, D. C., & Myerson, J. (2014). The role of working memory abilities in lecture note-taking. *Learning and Individual Differences*, *33*, 12–22.

Cairo, A. (2016). *The truthful art: Data, charts, and maps for communication*. New Riders.

Carrillo-Reid, L., Han, S., Yang, W., Akrouh, A., & Yuste, R. (2019). Controlling visually guided behavior by holographic recalling of cortical ensembles. *Cell*, *178*(2), 447–457.

Chakrabarty, M., & Wada, M. (2020). Perceptual effects of fast and automatic visual ensemble statistics from faces in individuals with typical development and autism spectrum conditions. *Scientific Reports*, *10*(1), 2169.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404.

Cleveland, W. S., & McGill, R. (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*(387), 531–554.

Corbett, J. E., Utochkin, I., & Hochstein, S. (2023). *The pervasiveness of ensemble perception: Not just your average review.* Cambridge University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114.

Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, *19*(1), 51–57.

d'Ocagne, M. (1885). Coordonnées parallèles et axiales : Méthode de transformation géométrique et procédé nouveau de calcul graphique déduits de la considération des coordonnées parallèles. *Gauthier-Villars*, 112. Retrieved from https://archive.org/details/coordonnesparal00ocaggoog/page/n10

Dakin, S. C., & Watt, R. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*(22), 3181–3192.

Day, R. H., & Stecher, E. J. (1991). Sine of an illusion. *Perception*, *20*, 49–55.

DeDonno, M. A. (2016). The influence of IQ on pure discovery and guided discovery learning of a complex real-world task. *Learning and Individual Differences*, *49*, 11–16.

Few, S. (2006b). *Information dashboard design: The effective visual communication of data.* Newton, MA: O'Reilly Media, Inc.

Few, S. (2006a). *Information dashboard design: The effective visual communication of data.* Newton, MA: O'Reilly Media, Inc.

Fofonov, A., & Linsen, L. (2018). A top-down interactive visual analysis approach for physical simulation ensembles at different aggregation levels. *Information*, *9*(7), 163.

Goldstein, E. B., & Cacciamani, L. (2021). *Sensation and perception, 11th edition.* Cengage Learning.

Haberman, J. M., & Ulrich, L. (2019). Precise ensemble face representation given incomplete visual input. *I-Perception*, *10*(1), 2041669518819014.

Haberman, J., Brady, T. F., & Alvarez, G. A. (2015). Individual differences in ensemble perception reveal multiple, independent levels of ensemble representation. *Journal of Experimental Psychology: General*, *144*(2), 432.

Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. *From Perception to Consciousness: Searching with Anne Treisman*, *33*.

Heinrich, J., & Weiskopf, D. (2009). Continuous Parallel Coordinates. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 1531–1538. http://doi.org/10.1109/TVCG.2009.131

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*(1), 243–271.

Hofmann, H., & Vendettuoli, M. (2013). Common Angle Plots as Perception-True Visualizations of Categorical Associations. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2297–2305. http://doi.org/10.1109/TVCG.2013.140

Hubel, D. H., & Wiesel, T. N. (2004). *Brain and visual perception: The story of a 25-year collaboration*. Oxford University Press.

Hullman, J., Adar, E., & Shah, P. (2011). Benefitting infovis with visual difficulties. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2213–2222.

Hullman, J., & Gelman, A. (2021). Designing for Interactive Exploratory Data Analysis Requires Theories of Graphical Inference. *Harvard Data Science Review*, *3*(3).

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10-12), 1489–1506.

John hopkins university covid-19 dashboard. (n.d.). Retrieved from `https://coronavirus.jhu.edu/map.html`

Karjus, A., Solà, M. C., Ohm, T., Ahnert, S. E., & Schich, M. (2023). Compression ensembles quantify aesthetic complexity and the evolution of visual art. *EPJ Data Science*, *12*(1), 21.

Khayat, N., Ahissar, M., & Hochstein, S. (2023). Perceptual history biases in serial ensemble representation. *Journal of Vision*, *23*(3), 7–7.

Khayat, N., Fusi, S., & Hochstein, S. (2021). Perceiving ensemble statistics of novel image sets. *Attention, Perception, & Psychophysics*, *83*, 1312–1328.

Khayat, N., & Hochstein, S. (2018). Perceiving set mean and range: Automaticity and precision. *Journal of Vision*, *18*(9), 23–23.

Leavitt, M. L., Pieper, F., Sachs, A. J., & Martinez-Trujillo, J. C. (2017). Correlated variability modifies working memory fidelity in primate prefrontal neuronal ensembles. *Proceedings of the National Academy of Sciences*, *114*(12), E2494–E2503.

Lee, S., Kim, S.-H., Hung, Y.-H., Lam, H., Kang, Y., & Yi, J. S. (2016). How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking. *IEEE*, *22*, 499–508.

Leib, A. Y., Kosovicheva, A., & Whitney, D. (2016). Fast ensemble representations for abstract visual impressions. *Nature Communications*, *7*(1), 13186.

Luxen, D., & Vetter, C. (2011). Real-time routing with OpenStreetMap data. In *Proceedings of the 19th ACM SIGSPATIAL international con-*

ference on advances in geographic information systems (pp. 513–516). New York, NY, USA: ACM. http://doi.org/10.1145/2093973.2094062

Maule, J., & Franklin, A. (2015). Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision*, *15*(4), 6–6.

McCallum, W. C. (n.d.). Attention. https://www.britannica.com/science/attention.

Monmonier, M. (1985). *Annals of the Association of American Geographers*, *75*(4), 605–609. Retrieved from http://www.jstor.org/stable/2563117

National Center for Education Statistics. (2020). National center for education statistics. https://nces.ed.gov/.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. Basic books.

Oberauer, K. (2009). Design for a working memory. *Psychology of Learning and Motivation*, *51*, 45–100.

Padilla, L. M., Ruginski, I. T., & Creem-Regehr, S. H. (2017). Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. *Cognitive Research: Principles and Implications*, *2*, 1–16.

Palmer, S. E. (2002). Perceptual organization in vision. *Stevens' Handbook of Experimental Psychology*.

Pérez-Ortega, J., Alejandre-García, T., & Yuste, R. (2021). Long-term stability of cortical ensembles. *Elife*, *10*, e64449.

Peters, D. J. (2019). Community Resiliency in Declining Small Towns: Impact of Population Loss on Quality of Life over 20 Years. *Rural Sociology*, *84*(4), 635–668. http://doi.org/10.1111/ruso.12261

Playfair, W. (2005). *Playfair's commercial and political atlas and statistical breviary*. Cambridge University Press.

Potter, K., Wilson, A., Bremer, P.-T., Williams, D., Doutriaux, C., Pascucci, V., & Johnson, C. R. (2009). Ensemble-vis: A framework for the statistical visualization of ensemble data. In *2009 IEEE international conference on data mining workshops* (pp. 233–240). IEEE.

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from `https://www.R-project.org/`

Rayner, K. (2009). The 35th sir frederick bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Rubin, J., & Chisnell, D. (2008). *Handbook of usability testing: How to plan, design, and conduct effective tests*. John Wiley & Sons.

Rural Shrink Smart Team. (2022). Rural shrink smart. `https://ruralshrinksmart.org/`.

Sarikaya, A., Correll, M., Bartram, L., Tory, M., & Fisher, D. (2019). What do we talk about when we talk about dashboards? *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 682–692. http://doi.org/10.1109/TVCG.2018.2864903

Satyanarayan, A., Moritz, D., Wongsuphasawat, K., & Heer, J. (2016). Vega-lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 341–350.

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., … Dillenbourg, P. (2016). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies*, *10*(1), 30–41.

Stallman, R. (2002). *Free software, free society: Selected essays of richard m. stallman*. Lulu. com.

State of Iowa. (2020). Iowa data portal. https://data.iowa.gov.

Sulston, J., & Ferry, G. (2002). *The common thread: A story of science, politics, ethics, and the human genome*. Random House.

Szafir, D. A. (2017). Modeling color difference for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 392–401.

Szafir, D. A., Haroz, S., Gleicher, M., & Franconeri, S. (2016). Four types of ensemble coding in data visualizations. *Journal of Vision*, *16*(5), 11–11.

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *353*(1373), 1295–1306.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Tufte, Edward R. (1985). The visual display of quantitative information. *The Journal for Healthcare Quality (JHQ)*, *7*(3), 15.

Tufte, Edward R. (2001). *The visual display of quantitative information (2nd edition)*. USA: Graphics Press.

Tukey, J. W., & Wilk, M. B. (1966). Data analysis and statistics: An expository overview. In *Proceedings of the november 7-10, 1966, fall joint computer conference* (pp. 695–709).

Unwin, A., Volinsky, C., & Winkler, S. (2003). Parallel coordinates for exploratory modelling analysis. *Computational Statistics & Data Analysis*, *43*, 553–564. http://doi.org/10.1016/S0167-9473(02)00292-X

USDA - ERS. (2020a). Rural classifications. `https://www.ers.usda.gov/topics/rural-economy-population/rural-classifications/`.

USDA - ERS. (2020b). Rural-urban commuting area codes. `https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/`.

Utochkin, I. S., Choi, J., & Chong, S. C. (2023). A population response model of ensemble perception. *Psychological Review*.

VanderPlas, S., & Hofmann, H. (2015a). Signs of the sine illusion–why we need to care. *Journal of Computational and Graphical Statistics*, *24*(4), 1170–1190. http://doi.org/10.1080/10618600.2014.951547

VanderPlas, S., & Hofmann, H. (2015b). Signs of the sine illusion—why we need to care. *Journal of Computational and Graphical Statistics*, *24*(4), 1170–1190.

Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., Van der Helm, P. A., & Van Leeuwen, C. (2012). A century of gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, *138*(6), 1218.

Ward, M., Grinstein, G., & Keim, D. (2010). *Interactive data visualization: Foundations, techniques, and applications* (p. 81). http://doi.org/10.1201/b10683

Ware, C. (2012). *Information visualization: Perception for design*. Morgan Kaufmann.

Weber, S. (2004). *The success of open source*. Harvard University Press.

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y., Zemla, J., et al. (2017). Package "corrplot." *Statistician*, *56*(316), e24.

Wertheimer, M. (1912). Experimentelle studien uber das sehen von bewegung (translated extract reprinted as "experimental studies on the seeing of motion"). *Zeitschrift Fur Psychologie*, *61*, 161–165.

WERTHEIMER, M. (1923). Laws of organization in perceptual forms. First published as untersuchungen zur lehre von der gestalt II. *Psychologische Forschung*, *4*, 301–350.

Wertheimer, M. (1938). Laws of organization in perceptual forms.

Whitney, D., Haberman, J., & Sweeny, T. D. (2014). 49 from textures to crowds: Multiple levels of summary statistical perception. *The New Visual Neurosciences*, 695–710.

Whitney, D., & Manassi, M. (2022). Ensemble perception: Stacking the hay to find the needle. *Current Biology*, *32*(22), R1264–R1266.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). Springer-Verlag New York. Retrieved from `https://ggplot2.tidyverse.org`

Wickham, H., Cook, D., Hofmann, H., & Buja, A. (2011). tourr: An R Package for Exploring Multivariate Data with Projections. *Journal of Statistical Software, Articles*, *40*(2), 1–18. http://doi.org/10.18637/jss.v040.i02

Wilkinson, L. (2012). *The grammar of graphics*. Springer.

Zarecor, K. E., Peters, D. J., & Hamideh, S. (2021). Rural smart shrinkage and perceptions of quality of life in the american midwest. In *Handbook of quality of life and sustainability* (pp. 395–415). Springer.