Visualizing Ambiguity: A Grammar of Graphics Approach to Resolving Numerical Ties in Parallel Coordinate Plots

Denise Bradford

2026-01-01

Table of contents

1 Introduction

This proposal outlines a systematic approach to visually distinguish tied numerical values in multidimensional datasets by employing parallel coordinate plots (PCPs). Parallel coordinates, first popularized by Alfred Inselberg, are a powerful technique for investigating patterns across multiple attributes simultaneously (Inselberg 2009). However, when datasets contain exact numerical ties, the resulting overlapping lines in PCPs can remove useful information.

To address this, we propose a uniform method for collected spacing to tied values. Pending approval, this feature will be integrated into `ggpcp` to help users resolve numerical ties more clearly. This tool will use even spacing to fix overlapping lines, ensuring parallel coordinate plots remain readable. My goal is to streamline the workflow for anyone who needs to track specific data points while still analyzing aggregate trends

Our approach is inspired by recent work on generalized parallel coordinate plots (GPCPs), an extension of PCPs that supports categorical variables (VanderPlas et al. 2023). The `ggpcp` package for R implements these GPCPs using a grammar of graphics framework, which seamlessly incorporates both continuous and categorical variables in a single parallel coordinate plot. One of the key contributions of that work is a robust tie-breaking mechanism for categorical variables, implemented through the `pcp_arrange()` function with methods hierarchical sorting. This ensures that individual observations can be traced across multiple dimensions, even when categories induce identical or "tied" values.

By adding a similar numerical tie-breaking technique for continuous data, we further enhance GPCPs' capacity to handle the visualization of real-world datasets exhibiting ties in numerical and categorical data.

2 Background and Motivation

2.1 Parallel Coordinate Plots

Parallel coordinate plots assign each dimension of an n-dimensional dataset to a vertical axis arranged in parallel (Wegman 1990). Each observation is drawn as a polyline connecting its values on these axes, providing a visual representation that can illuminate underlying data structures.

2.2 Numerical Ties and Visual Overlap

When multiple observations share the same value in a given dimension, their polylines perfectly overlap, creating "visual collisions." This masks information about distribution, density, or potential outliers. The treatment of ties is an aspect not generally addressed in the original parallel coordinate plots of Inselberg (1985) and Wegman (1990). However, the `ggpcp` package implementation has demonstrated that careful tie-handling is essential for both continuous and categorical variables.

The `ggpcp` package separates data management from visual rendering into by implementing multiple stages of data transformation: selection (implicit reshaping), scaling and sorting, a step important because of the ggpcp's treatment of categorical variables (VanderPlas et al. 2023).

2.2.1 The Challenge of Visualizing Mixed-Type Data

Parallel coordinate plots (PCPs) have been established as valuable tools for exploratory data analysis of high-dimensional numerical data since their introduction (Inselberg 1985; Wegman 1990). However, the use of PCPs is fundamentally limited when working with mixed categorical-continuous data. As VanderPlas et al. (2023) note in their introduction to generalized parallel coordinate plots (GPCPs), existing solutions for categorical values become insufficient when attempting to maintain visual continuity across both data types.

The treatment of ties, multiple observations sharing the same value, represents a critical design decision that affects both perceptual effectiveness and analytical utility. As noted earlier, the treatment of ties—where multiple observations share the same value—is a critical design decision that shapes both the perceptual effectiveness and the analytical utility of a visualization. This observation extends naturally from categorical to numerical variables.

The ability to follow individual observations is central to the analytical power of PCPs, enabling users to identify patterns, outliers, and relationships that span multiple variables simultaneously. The analytical effectiveness of parallel coordinate plots (PCPs) depends entirely on the Gestalt principle of continuity, which posits that the human visual system inherently follows lines that maintain a consistent trajectory. PCPs use object-based visual attention to show a multidimensional data point as one long line. This allows researchers to preserve the whole profile of an observation in working memory while it "threads" along several axes. This ability to stick to one path makes it easier to see complex multivariate patterns and outliers that would be hidden in summary statistics.

But as VanderPlas et al. (2023) points out, the initial frameworks made by Inselberg (1985) and Wegman (1990) didn't adequately deal with the "treatment of ties." This often leads to confusing crossings that go against the idea of common fate. When lines perfectly overlap at a tie, the data point loses its "objectivity," which forces the analyst to transition between a general and a detailed perspective. This study seeks to rectify this issue by ensuring uniform spacing of the lines.

2.2.2 Existing Solutions for Categorical Ties in ggpcp

The `ggpcp` package currently addresses categorical ties through sophisticated tie-breaking algorithms, . The package implements hierarchical sorting through the `pcp_arrange(data, method, space)` function, which orders observations based on a hierarchical application of variable values. The `method` parameter determines the sequence in which variables are considered when resolving ties in the arrangement. The `space` parameter specifies the proportion of the y-axis allocated for spacing between levels of categorical variables (Figure 1).

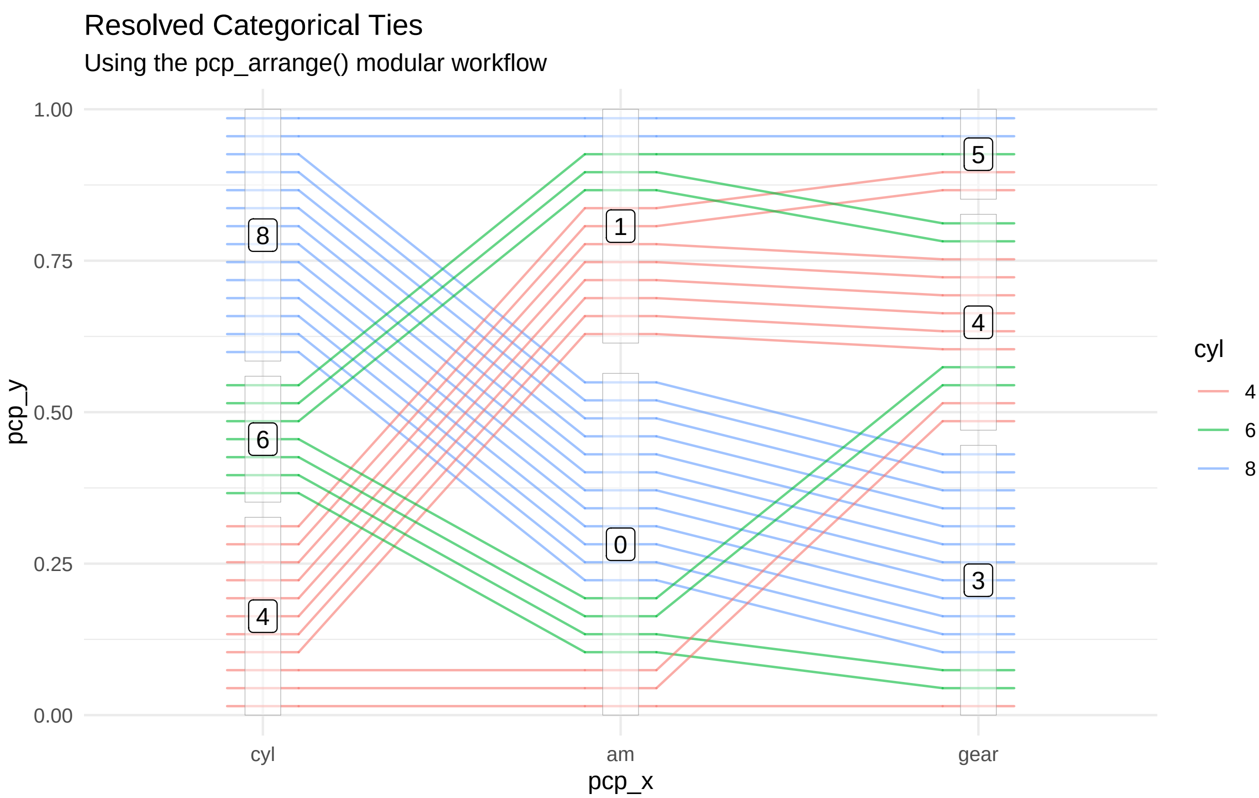

Figure 1

The ability to follow individual observations is central to the analytical power of PCPs, enabling users to identify patterns, outliers, and relationships that span multiple variables simultaneously. This hierarchical sorting approach serves as "external cognition," the additional computational processing reduces the cognitive load required to untangle overlapping lines in the parallel coordinate plot. The categorical tie-breaking creates equispaced points along the axis that reduces line crossings and allows users to follow individual observations from left to right through the plot even for categorical variables.

- Step 1: Identify the Parallel Axes

Begin by identifying each vertical axis in the plot. Each axis represents one variable from the dataset. The axes are typically arranged from left to right, and the order may be determined by the data analyst to highlight specific relationships or minimize visual clutter.

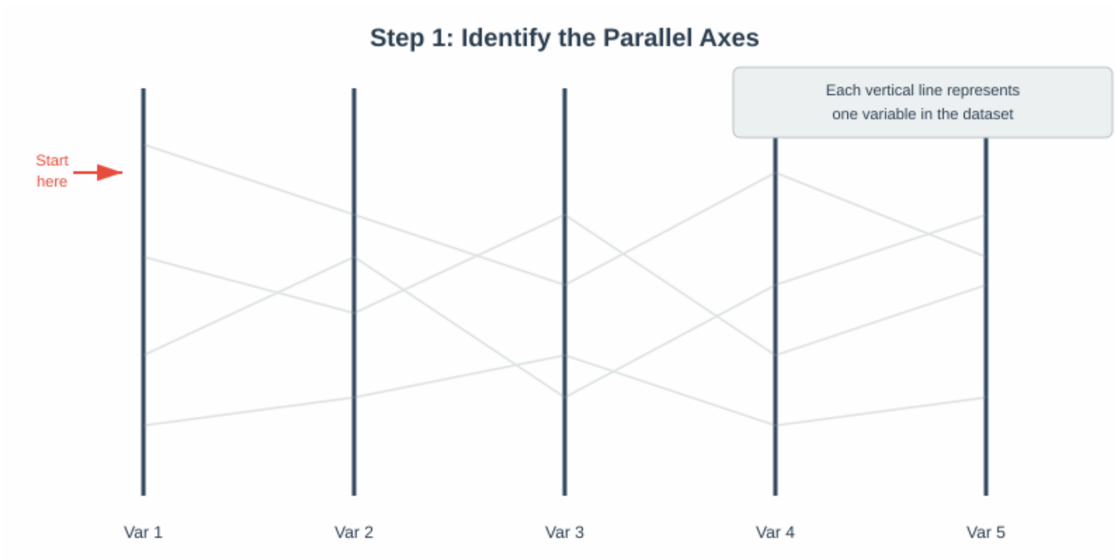

Figure 2: The parallel axes form the structural framework of the visualization.

- Step 2: Locate the Starting Point

Find the observation of interest on the leftmost axis. The vertical position indicates the scaled value of that observation for the first variable. If you are examining a highlighted or color-coded observation, look for its distinctive marker at this starting position.

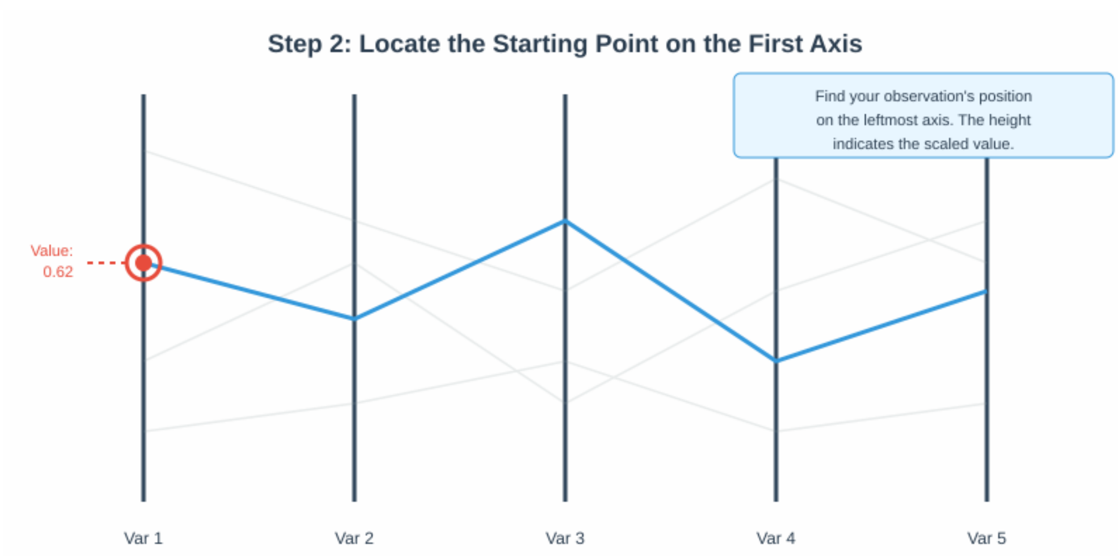

Figure 3: The starting point is identified on the first axis with its corresponding value.

- Step 3: Follow the Line Segment

Trace the line segment from the starting point to its intersection with the next axis. The human visual system naturally follows smooth, continuous paths due to the Gestalt principle of good continuation. This principle allows viewers to perceive connected lines as unified objects, making it easier to track observations across multiple variables.

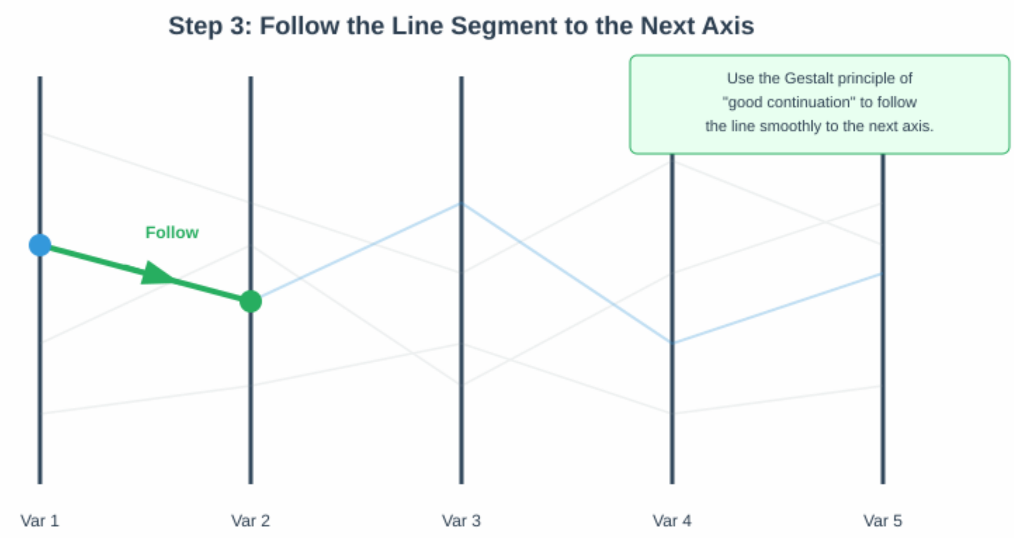

Figure 4: Following the line segment uses the Gestalt principle of good continuation.

- Step 4: Read Values at Intersections

At each axis intersection, the vertical position of the line indicates the observation's value for that variable. Read these values to understand how the observation changes across different dimensions of the data. The slope of line segments between axes provides information about the relationship between consecutive variables for that specific observation.

- Step 5: Continue Across All Axes

Repeat the tracing process for each consecutive pair of axes until reaching the rightmost axis. By following the complete path, you obtain a comprehensive view of how that particular observation behaves across all measured variables. The vertical positions along each axis represent scaled values that can be interpreted as quantiles when the data are appropriately transformed. This enables identification of unique characteristics, cluster membership, or outlier status.

2.2.3 Current State: ggpcp's Categorical Tie-Breaking

The ggpcp package implements a sophisticated tie-breaking algorithm for categorical variables that maintains individual observation traceability. The approach spaces observations evenly within categorical levels:

> "All observations are spaced out evenly. This results in a natural visualization of the marginal frequencies along each axis (additionally enhanced by the light gray boxes

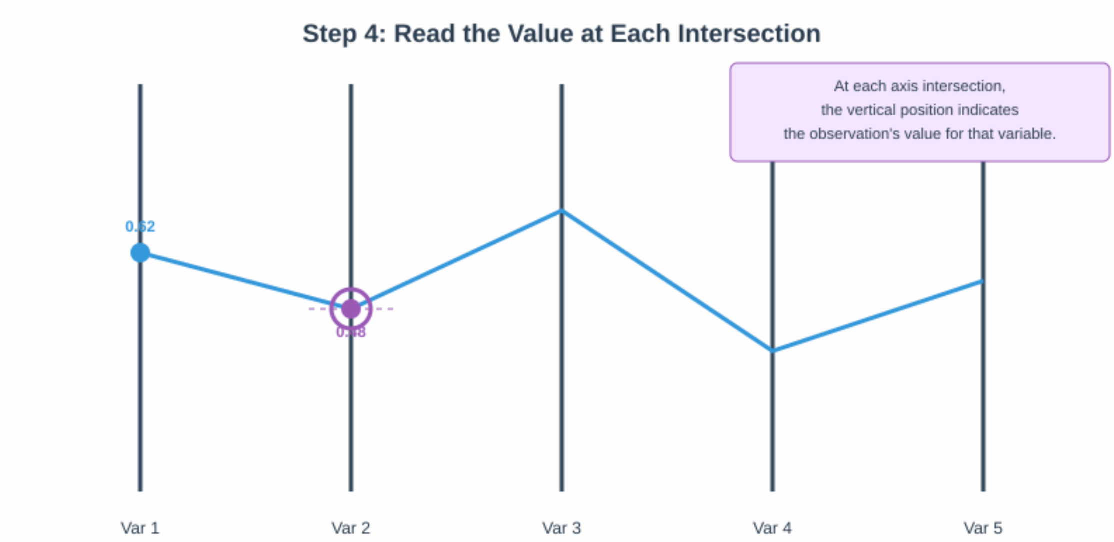

Figure 5: Values are read as approximate scaled value relative to the total range.

Figure 6: The complete traced path reveals the observation's values across all variables.

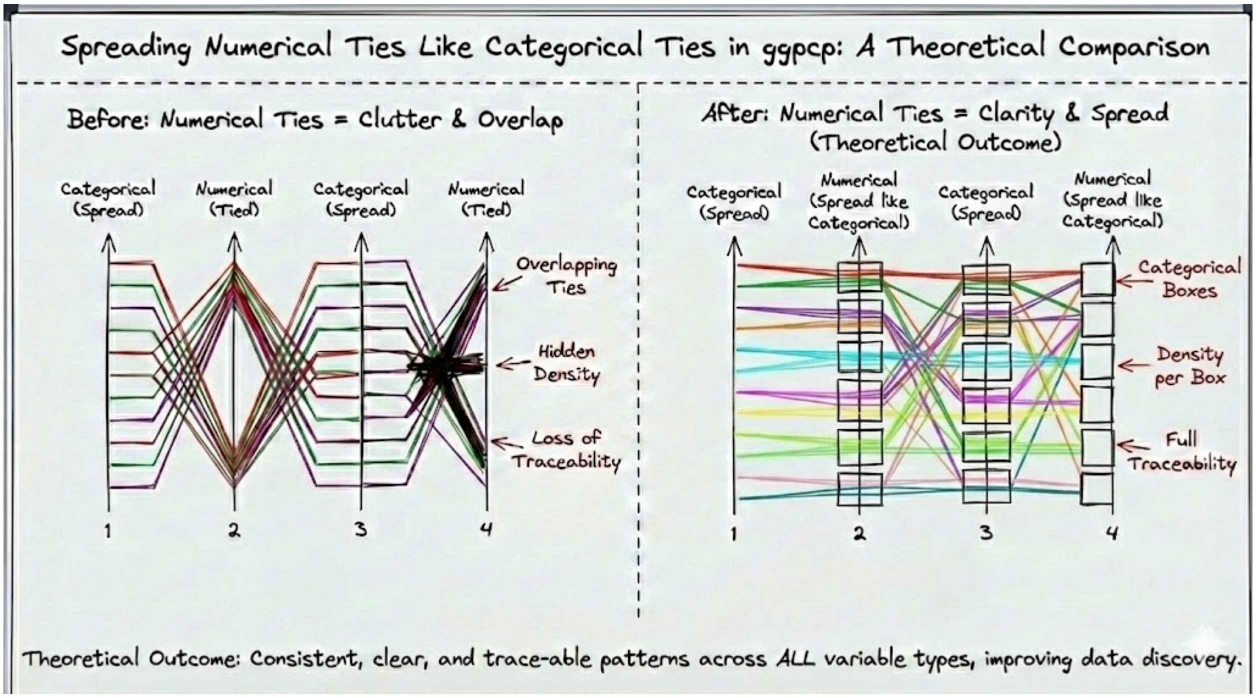

Figure 7: Comparison of standard parallel coordinate plot representation and the Even Tie Spreading Algorithm for resolving numerical ties.

> grouping observations in the same category) that is not as prominent in the previous three panels. The ordering of the observations within the level is such that a minimal number of line crossings occurs between the axes." (p. 11)

The algorithm achieves this through hierarchical sorting implemented in `pcp_arrange(data, method, space)`, where the `space` parameter specifies the proportion of the y-axis used for spacing between categorical levels. This optimization can be formalized as:

$$d_i = \frac{S_i - S_i^- - S_i^+}{n_i - 1}$$

where:

- S_i is the total space allocated to category i
- S_i^- is the spacing below category i
- S_i^+ is the spacing above category i
- n_i is the number of observations in category i
- d_i is the optimal spacing distance between consecutive observations

2.2.4 Alternative Approach: Hammock Plots

Hammock plots, introduced by Schonlau Schonlau and Yang (2024), take a fundamentally different approach to handling both categorical and numerical variables. Rather than using individual lines, hammock plots employ two-dimensional boxes to connect adjacent axes, with box width proportional to the number of observations.

As Schonlau and Yang (2024) describes:

> "Like a parallel coordinate plot, the axes are aligned parallel to one another. Categories of adjacent variables are connected by boxes. (The boxes shown are parallelograms; I use the word boxes for simplicity). The width of boxes is proportional to the number of observations." (p. 3)

For numerical variables specifically, hammock plots maintain constant-width boxes throughout the visualization. Schonlau and Yang (2024) explains the spatial constraint this imposes:

> "When treating this variable as numerical, the range from 0 to 20 leaves 1/21th of the space for each unit length. Consequently, the widths of the boxes have to be more frugal." (p. 21)

This creates a fundamental trade-off: hammock plots explicitly encode frequency through box width but sacrifice individual observation traceability. As Schonlau and Yang (2024) notes, "For small data sets, GPCP plots beautifully show all individual observations whereas hammock plots require highlighting to feature individual observations" (p. 19).

2.2.5 Visual Comparison:

A key visual difference emerges when connecting categorical to numerical variables. VanderPlas et al. (2023) observe:

> "When many observations have the same value for a categorical and an adjacent numerical variable, the corresponding area looks like a triangle… Notice the lines/boxes between the variables hospitalizations and comorbidities in the GPCP (Figure 13) and hammock plots (Figure 2). Most of the observations are in the boxes leading from hospitalizations=0 to either comorbidities=0 or comorbidities=1. This is far more obvious in the hammock plot than in the GPCP plot." (p. 19)

When ggpcp connects categorical and numerical variables, the tie-breaking algorithms keep the triangular patterns clear (Figure 8). The `ggpcp` package makes sure that concentration patterns stay visible by keeping the vertical spacing consistent through the space parameter and the systematic ordering through the method parameter. This doesn't make it harder to understand individual observation paths. Box selections and triangular density representations have very different uses. Box selections separate discrete data subsets for filtering and brushing interactions, while triangles show a continuous aggregate view of data relationships and density between dimensions. Users do not have to choose between showing overall patterns and being able to follow specific cases through the plot.

This observation indicates that each method possesses perceptual advantages in varying contexts, necessitating a thorough comparative analysis based on perceptual science. To understand our tie-resolution methods, it is important to know about the perceptual and practical problems that come with parallel coordinate plots. This section combines real-world evidence from the visualization literature that directly affects our design choices and evaluation plan.

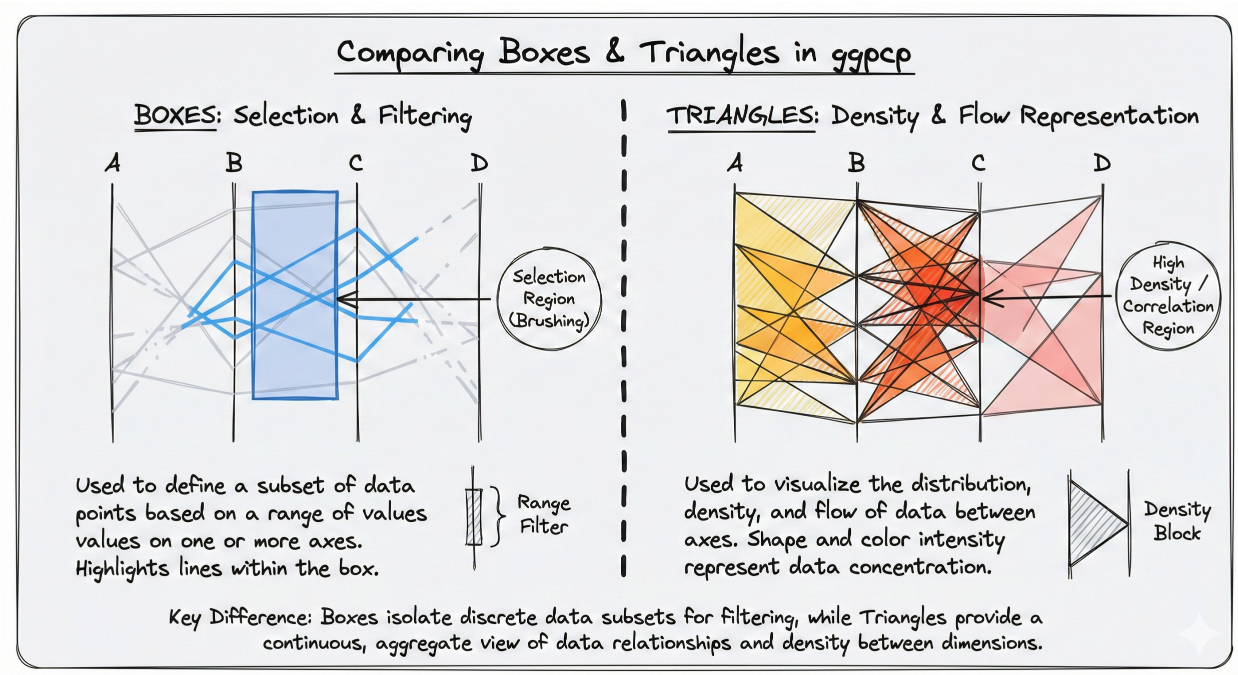

Figure 8: The visual proof of the perceptual trade-offs they describe regarding how data transitions between axes.

3 Handling Numerical Ties

A significant challenge arises when multiple observations share identical values on an axis. In traditional PCPs, these observations converge to a single point, creating overlapping lines that make individual tracking impossible. The ggpcp package addresses this through the `tie_spread` algorithm.

3.1 The Problem: Overlapping Lines

When multiple observations have the same value for a variable, their lines converge to a single point on that axis. This convergence creates visual clutter and breaks the continuity needed for individual observation tracking.

3.2 The Solution: Tie Spreading

The `tie_spread` algorithm splits up tied observations by evenly spreading them out over a small range around their original value. This range is set by default to balance the need for effective tie-breaking with spatial perception. This makes sure that separated observations stay visually grouped near their shared value while still being different enough to trace individual paths through the plot.

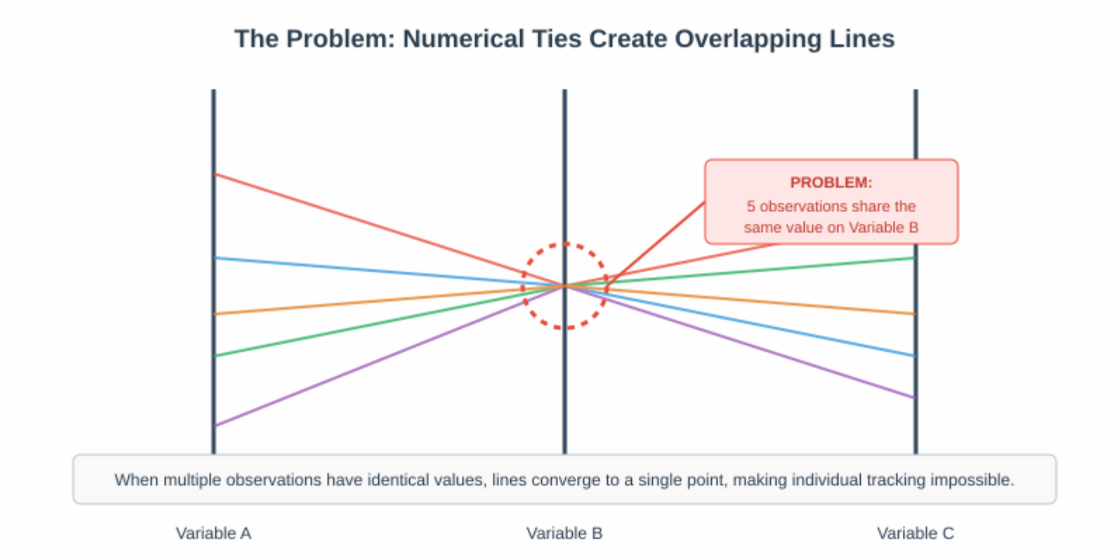

Figure 9: Numerical ties cause lines to converge at a single point.

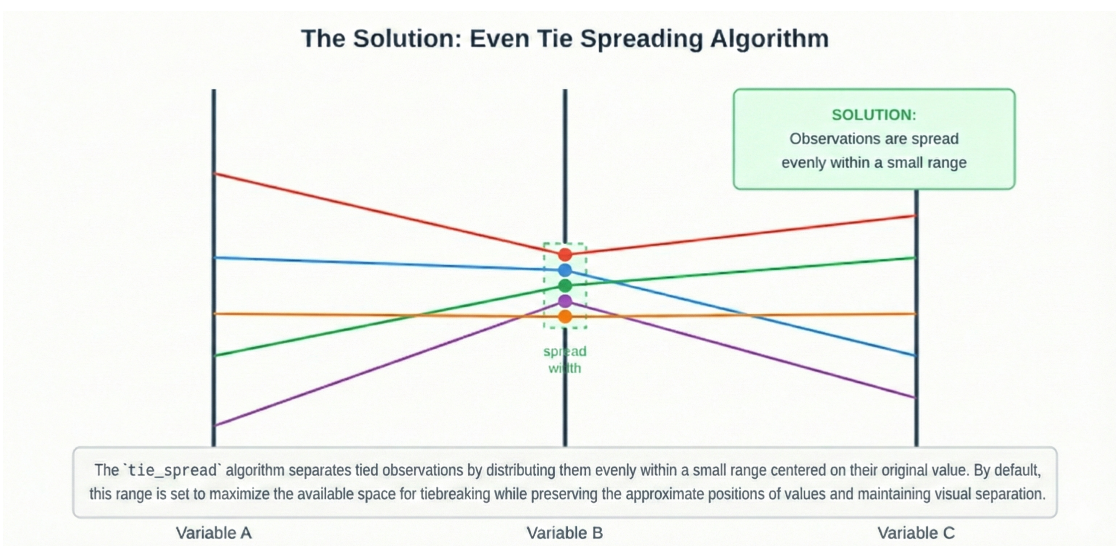

Figure 10: Tie spreading maintains visual separation while preserving approximate positions.

3.3 Hierarchical Sorting for Minimal Crossings

Beyond simply spreading tied values, the ggpcp package implements hierarchical sorting to minimize line crossings as in categorical tie breaking.

By ordering observations within a tie group based on their values on adjacent axes, the package leverages the Gestalt principle of common fate. This strategy ensures that observations with similar trajectories are positioned in close proximity, creating cohesive visual bands. These bands facilitate the perception of distinct groups as they move together through high-dimensional space, reducing visual noise and highlighting underlying patterns.

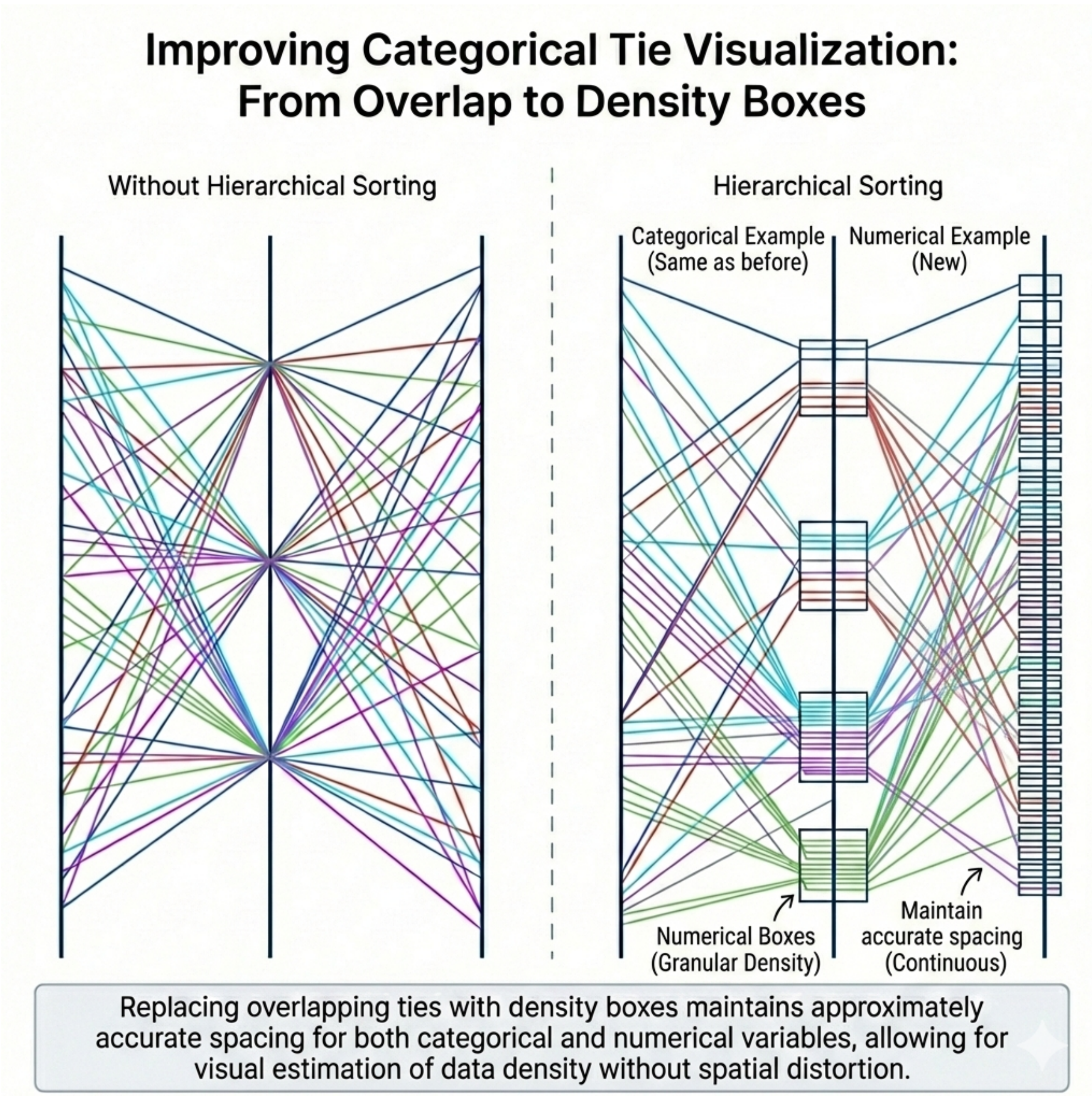

Figure 11: Hierarchical sorting reduces line crossings, making patterns easier to perceive.

3.4 Theoretical Framework: Perception-Driven Design

Effective data visualizations must be grounded in how people actually see and process information. This theoretical framework draws on research from cognitive psychology and visual perception to explain how viewers interpret visual displays. The framework covers four key areas that trace the path from seeing a visualization to understanding its meaning. **Stage 1** explores how Gestalt principles help viewers group visual elements into meaningful patterns. **Stage 2** examines how visual working memory and attention allow people to hold and manipulate information while analyzing a visualization. **Preattentive processing** describes how certain visual features can be detected quickly and automatically, enabling rapid visual search before focused attention is required. Finally, **external cognition and computational offloading** addresses how visualizations act as external aids that extend our mental capabilities by shifting some cognitive work from our minds to the visual display itself. These four components provide a foundation for understanding how design choices affect visualization effectiveness.

3.4.1 Stage 1: Pattern Perception and Gestalt Principles

The first stage involves active pattern perception processes that segment the visual scene into coherent regions. The Gestalt laws of perceptual organization, articulated by German psychologists in 1912, remain fundamental principles for visualization design:

- **Proximity:** Objects near each other are perceived as belonging together
- **Similarity:** Objects sharing visual properties (color, shape, size) are grouped perceptually
- **Connectedness:** Objects connected by lines are seen as related

- **Continuity:** The visual system prefers interpretations with smooth, continuous paths
- **Symmetry:** Symmetric patterns are more easily perceived and remembered
- **Closure:** Incomplete shapes are perceptually completed
- **Common Fate:** Objects moving together are perceived as a group
- **Relative Size:** Smaller elements are perceived as figures against larger backgrounds

VanderPlas et al. (2023) explicitly invoke these principles in their design rationale for ggpcp:

> "By reducing the number of line crossings at non-axis points simplifies the plot, reducing the overall cognitive load required to 'untangle' (literally and metaphorically) the individual observations and leveraging the Gestalt principle of common fate." (p. 4)

The principle of **good continuation** is particularly relevant to parallel coordinate plots: can viewers smoothly follow individual lines through numerical ties, or do visual discontinuities disrupt the perceptual flow?

3.4.2 Stage 2: Visual Working Memory and Attention

The seconnd stage involves conscious attention and visual working memory, which holds only a limited number of objects (typically 3-4) for active processing.

This three-stage model has direct implications for evaluating tie-breaking strategies:

Equispaced Lines:

- Leverage preattentive orientation and position
- Support continuous line following (good continuation)
- Implicitly encode frequency through density (requires stage 1 processing)

Constant-Width Boxes:

- Leverage preattentive size and area
- Explicitly encode frequency (reduced cognitive load)
- May disrupt line continuity (challenges good continuation)

3.5 Preattentive Processing and Visual Search

A key distinction between the two approaches concerns the nature of visual search required for different analytical tasks. Ware distinguishes between:

- **Parallel Search (Preattentive):** Target pops out immediately, search time independent of number of distractors
- **Serial Search (Attentive):** Search time increases linearly with number of items, requiring conscious attention

For tracing individual observations through equispaced lines, the task may benefit from preattentive processing of orientation and position. Each observation maintains a consistent visual identity as a continuous line with specific orientation. In contrast, tracing through constant-width boxes requires conscious attention to reconstruct the path, as the area-based encoding does not support preattentive line following.

However, for frequency estimation tasks, constant-width boxes may have an advantage. Equispaced lines require inferring frequency from density, which is a second-order visual property requiring pattern perception (stage 1).

3.5.1 Gestalt Principles and Line Continuity

The principle of **good continuation** states that the human visual system preferentially perceives smooth, continuous contours over interpretations requiring abrupt changes in direction. This principle directly impacts the effectiveness of parallel coordinate plots for tracing individual observations.

Equispaced lines maintain continuity by representing each observation as a continuous polyline extending from the leftmost to the rightmost axis. When observations share identical numerical values (ties), the equispacing mechanism distributes lines within each category so that they remain visually distinct without introducing discontinuities. The ggpcp framework implements this through hierarchical sorting, which minimizes line crossings at axis intersections and leverages both the Gestalt principle of good continuation and the principle of common fate, wherein lines with similar values across multiple axes appear to move together through the display.

Constant-width boxes represent observations as aggregated area segments rather than individual lines. This encoding communicates aggregate quantities effectively and provides an immediate visual representation of bivariate contingency relationships. However, tracing a single observation through a box requires viewers to mentally reconstruct the implied path, engaging conscious attention and working memory. In this sense, the two approaches serve distinct analytical purposes: equispaced lines support individual-level tracing, while boxes support aggregate-level frequency comparison.

This distinction becomes increasingly relevant as dataset size grows. With many observations, equispaced lines may coalesce into ribbon-like bands, yet the continuity principle still applies—viewers can perceive flow patterns even when individual lines are indistinguishable. Boxes, by contrast, maintain their aggregate interpretation but are not designed to preserve individual-level information. Rather than viewing these as competing representations, they can be understood as complementary: the line-based approach is optimal when the analytical task involves following specific cases or detecting outliers, whereas the box-based approach is optimal when the task involves comparing category frequencies or understanding distributional patterns.

3.6 Implementation and Evaluation Plan

3.6.1 Phase 1: Algorithm Development (Weeks 1-4)

Deliverable: Functional R package extension with comprehensive documentation

Tasks:

1. Extend `pcp_arrange()` to detect and handle numerical ties
2. Implement `optimize_spacing_numerical()` or `tie_spacing()` function with perceptually-motivated parameters
3. Handle edge cases while preserving perceptual properties:
   - Single unique value (centered position)
   - Extreme skew (adaptive buffer sizing)
   - Missing values (explicit separation region)
4. Integration testing with existing ggpcp workflow

Code Structure (Simplified):

3.7 Phase 2: Perceptual Validation Studies (Weeks 5–7)

Study 1: Frequency Perception

- **Duration:** Weeks 5–7
- **Participants:** n = 60 (30 per condition)
- **Design:** Between-subjects

Procedure:

1. Magnitude estimation (12 trials): “What percentage have value X?”
2. Ordinal comparison (12 trials): “Which value is more frequent?”

3. Ratio judgment (12 trials): "A is how many times B?"
4. Confidence ratings (7-point scale) for each response

Analysis:

- Absolute percentage error for magnitude estimation
- Accuracy rates for ordinal and ratio tasks
- Bias analysis: systematic over/under-estimation
- Confidence calibration: accuracy vs. subjective confidence

3.8 Phase 3: Comparative Benchmarking (Weeks 6–10)

Computational Metrics:

1. Rendering performance (time, memory, scalability n = 2 to 100)
2. Visual quality metrics (Dennig et al. 2021): line crossing count, ribbon overlap, visual clutter index
3. Perceptual quality estimates: modeled eye movements, predicted visual search time

Case Studies:

- Palmer Penguins: Mixed categorical-numerical data with known correlation structure
- Iris Dataset: Multiple numerical variables with natural ties
- Asthma Data (Schonlau and Yang 2024): Direct comparison with published hammock plot

3.9 Phase 4: Integration and Dissemination (Weeks 8–12)

Deliverables:

1. **R Package Update (ggpcp Pull Request):**
    - CRAN submission with full documentation
    - Vignette: "Handling Numerical Ties in Parallel Coordinate Plots"
    - Unit tests achieving $> 95\%$ coverage
2. **Academic Paper:**
    - Target: IEEE Transactions on Visualization and Computer Graphics
    - Submission deadline: March 2026 (IEEE VIS)
3. **Supplementary Materials:**
    - Open Science Framework repository
    - All experimental stimuli and data
    - Reproducibility package

4 Timeline and Milestones

Table 1: Research Timeline (12 Weeks)

Week	Milestone	Deliverable	Perceptual Focus
1–2	Literature synthesis	Annotated bibliography with Gestalt framework integration	Gestalt framework integration
3–4	Core algorithm development	Working R code with perceptual properties verified	Perceptual properties implemented
4	IRB submission	Approved protocol for human subjects research	Human subjects clearance
5–7	Study 1: Frequency perception	Raw data and statistical analysis	Magnitude estimation validation
8–10	Benchmarking	Performance report with computational metrics	Computational metrics
10–12	Integration	Draft manuscript and R package release	Synthesis and dissemination

5 Expected Contributions

5.1 Theoretical Contributions

Unified Tie-Breaking Theory: Formalization of equispaced optimization for both categorical and numerical variables, with key perceptual properties (separability, continuity, space efficiency)

5.2 Practical Contributions

1. **Production Software:** Open-source R package extension immediately usable by practicing data scientists and researchers
2. **Evidence-Based Guidelines:** Design recommendations grounded in empirical evidence and perceptual theory:
   - When to use equispaced lines vs. constant-width boxes
   - How to set spacing parameters for optimal perceptual clarity
   - Task-specific visualization selection criteria
3. **Benchmark Suite:** Reusable experimental framework and stimuli for evaluating future parallel coordinate plot innovations

5.3 Methodological Contributions

1. **Theory-Driven Evaluation:** Demonstrates how established perceptual theory (Gestalt principles) can generate specific, testable hypothesis
2. **Multi-Method Approach:** Integrates behavioral experiments with computational benchmarking to triangulate findings

Table 3: Application Domains for Research Impact

Domain	Application	Data Characteristics
Healthcare	Patient trajectory visualization	Categorical diagnoses + numerical measurements
Manufacturing	Quality control monitoring	Defect categories + continuous sensor readings
Social Science	Survey analysis	Demographic categories + Likert scales
Finance	Portfolio analysis	Categorical sectors + numerical performance metrics
Education	Learning analytics	Course completion + assessment scores

3. **Reproducible Research:** Complete open science package enabling replication and extension by other researchers

6 Broader Impact

This research addresses visualization challenges across numerous domains where mixed categorical-numerical data is common:

By providing both theoretical understanding and practical tools, this work enables more effective exploratory data analysis across these domains, ultimately supporting better data-driven decision making.

7 Limitations and Future Directions

7.1 Study Limitations

- **Sample Characteristics:** University student population may not generalize to domain experts
- **Experimental Constraints:** Laboratory tasks may lack ecological validity
- **Scope:** Limited to static visualizations; interactive features not evaluated

7.2 Future Extensions

- Longitudinal study with domain experts in real analytical workflows
- Investigation of interaction effects between individual differences and visualization method
- Hybrid approaches: smooth interpolation between lines and boxes based on dataset properties
- Integration with animation and interactive highlighting techniques

8 Conclusion

This thorough examination proposal tackles a key issue in information visualization by basing design choices on how people see things. The suggested equispaced line method builds on ggpcp's well-known categorical tie-breaking algorithm for numerical variables. This creates a single framework that keeps the visual flow going, uses preattentive processing, and follows Gestalt principles of good continuation and common fate. Equispaced lines provide distinct perceptual trade-offs relative to hammock plots' constant-width boxes: they enhance individual observation traceability and facilitate smoother visual flow, possibly at the expense of explicit frequency encoding.

This research advances theory by illustrating how Gestalt principles and perceptual processing models formulate testable hypotheses regarding visualization effectiveness; it enhances practice by providing open-source software and evidence-based design guidelines; and it enriches methodology by showcasing theory-driven, multi-method evaluation. To be effective, visualization must be based on how people perceive things, especially how they use Gestalt principles of proximity, similarity, continuity, and common fate to organize and make sense of visual features. This study rigorously examines the principle by inquiring not merely "which visualization appears superior?" but rather "which visualization more effectively corresponds with perceptual organization principles for particular analytical tasks?"

The answer, like most questions about visualization, is complicated and depends on the situation. But now it will be based on real-world data and perceptual theory.

9 Appendix

9.1 Visual Clutter and Information Density (To investigate)

Visual clutter constitutes a fundamental limitation on visualization effectiveness. Clutter occurs when too many visual elements compete for attention, overwhelming the viewer's capacity to extract meaningful patterns from the display.

The two approaches manage clutter through fundamentally different mechanisms. For equispaced lines, clutter increases with the number of observations, but hierarchical sorting minimizes line crossings and thereby reduces visual complexity. The approach exhibits graceful degradation: as density increases, individual lines transition perceptually into ribbons and eventually into filled areas, yet individual observations remain theoretically accessible for highlighting or interactive selection. For constant-width boxes, clutter depends primarily on the number of unique value combinations rather than on raw observation counts. Aggregation inherently reduces clutter, though at the cost of individual-level accessibility. The resulting display provides a clear representation of bivariate contingency relationships.

Dennig et al. (2021) formalize clutter metrics for parallel sets visualizations, measuring ribbon overlap, crossing angles, and ribbon width variance. These metrics can be adapted to compare the two approaches quantitatively and to guide the selection of dimension and category orderings that minimize visual complexity.

References

Dennig, Frederik L., Maximilian T. Fischer, Michael Blumenschein, Johannes Fuchs, Daniel A. Keim, and Evanthia Dimara. 2021. "ParSetgnostics: Quality Metrics for Parallel Sets." *Computer Graphics Forum* 40 (3): 375–86. https://doi.org/10.1111/cgf.14314.

Inselberg, Alfred. 1985. "The Plane with Parallel Coordinates." *The Visual Computer* 1 (2): 69–91. https://doi.org/10.1007/BF01898350.

———. 2009. "Parallel Coordinates: Visual Multidimensional Geometry and Its Applications." *Springer Science & Business Media*.

Schonlau, Matthias. 2003. "The Hammock Plot: Visualizing Mixed Categorical and Numerical Data." In *Proceedings of the American Statistical Association*.

Schonlau, Matthias, and Rosie Yuyan Yang. 2024. "Hammock Plots: Visualizing Categorical Data Beyond Parallel Coordinates." *Journal of Computational and Graphical Statistics*.

VanderPlas, Susan, Yawei Ge, Antony Unwin, and Heike Hofmann. 2023. "Penguins Go Parallel: A Grammar of Graphics Framework for Generalized Parallel Coordinate Plots." *Journal of Computational and Graphical Statistics* 32 (4): 1405–20. https://doi.org/10.1080/10618600.2023.2181762.

Wegman, Edward J. 1990. "Hyperdimensional Data Analysis Using Parallel Coordinates." *Journal of the American Statistical Association* 85: 664–75.