Curtain Graphs: Using a Floating Baseline for Comparison in a Two-dimensional Graphical Space

Kassandra Raymond and Andrew Hamilton-Wright
School of Computer Science, University of Guelph, Canada

Keywords: Relative and Absolute Measurement, Multi-dimensional Data, Floating Baseline, Inter-series Comparison, Time-series, Data Analysis.

Abstract: We present a novel visualization tool designed to provide support for the analysis of data sets focused around deviation from a baseline and including data from multiple series. The incorporation of a floating baseline makes the curtain graph distinct from waterfall plots and bar charts. Each data series therefore has a visual anchor that assists in interpretability, gives focus, and provides a means for easily broadening analysis across all presented series. The use of this tool in real-world examples based on relative and absolute comparisons is discussed.

1 INTRODUCTION AND RELATED WORK

This paper presents a novel tool for inter- and within-series comparisons where a defining feature of the series is its relationship to a particular anchor point, representing an initial condition or exemplar. This static graph can be used to visualize comparisons between multi-dimensional numerical and categorical data.

1.1 Background

The complexity of data analysis is increasing, both in terms of the complexity of the data sets themselves, and in the insights expected to be derived from them. The relationship between multiple variables of different types frequently captures the critical insight allowing us to understand a multi-dimensional data problem. This dimensionality, coupled with data of different types, such as discrete, continuous and categorical, can create a complicated set of data which can be difficult to visualize. Moreover, we are often limited to static plots when creating reports, books or articles, causing it to be increasingly difficult to display multi-dimensional data.

The field of information visualization has a long history of both broadly applicable and context specific visualization techniques. Great ideas in the field build on the work of pioneers such a Bertin (2001) and we see books regularly appearing that provide excellent advice in how to proceed both in general and in specific cases (Shneiderman, 1996; Tufte, 2001, 1991, 1997, 2006; Cairo, 2012, 2016; Evergreen, 2020).

A central focus within this long history is the identification of tools specific to a certain decision making paradigm. Inside of a decision making context, identification of a particular narrative (Klanten et al., 2011) or decision point (Agrawala et al., 2011) allows construction of a context for the specific data interpretation needs (Cairo, 2016).

Evaluation of visualization techniques has proceeded by applying various metrics to elucidate the understandability, utility and efficacy of the tool (Zen, 2013; Agrawala et al., 2011; Daru, 2001), including metrics such as the data-ink ratio (Tufte, 2001; Inbar et al., 2007) and identification of patterns driving our understand such as 'small multiples' (Tufte, 1991, pp. 67–80). This analysis has resulted in toolkits (Heer et al., 2005; Klimov et al., 2010) and advice (Shneiderman, 1996; Evergreen, 2020; Hicks, 2009; Metoyer et al., 2012) based around selecting exactly the correct visualization for a specific purpose, resulting in tailored visualizations for specific applications (de Almeida and Roselli, 2017; Klimov et al., 2010; Klanten et al., 2011).

In this paper we target a specific instance of data set interpretation, focusing on a multi-dimensional data set whose interpretation is based on the context of changes relative to a particular exemplar point—either an initial sample forming a point in time, or a particular datum from which all of the other points derive their meaning in some other form.

Raymond, K. and Hamilton-Wright, A.
Curtain Graphs: Using a Floating Baseline for Comparison in a Two-dimensional Graphical Space.
DOI: 10.5220/0008874001250132
In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020) - Volume 3: IVAPP*, pages 125-132
ISBN: 978-989-758-402-2; ISSN: 2184-4321

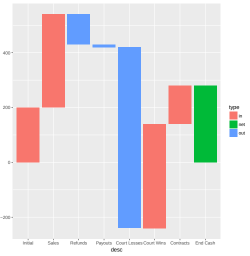

Figure 1: Waterfall plot showing consecutive changes in a series, from Few (2006).

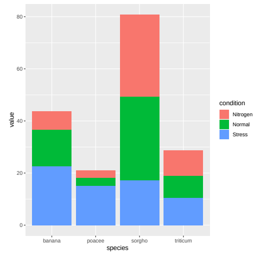

Figure 2: Stacked barchart showing series data in summation, from Holtz (2019).

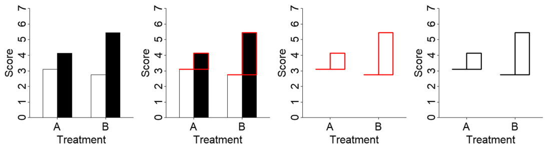

Figure 3: Hat graphs (rightmost two plots), as related to barcharts of the same data; reproduced from Witt (2019).

A related visualization technique is the 'waterfall chart' (Metoyer et al., 2012; Few, 2006; Evergreen, 2020), which is constructed as a sequence of bars with a start and end showing change at an interval described on the *x* axis relative to the previous bar, as shown in Figure 1, which is reproduced from the data discussed in Few (2006).

In a waterfall chart, there are three assumptions that are different from those of our needs:

- the initial value from which the first change starts is assumed to be zero;
- the changes as one works through the series are to be compared primarily only to their neighbours; and
- there is only one series in the chart.

Another related visualization is the common stacked bar chart, as discussed in Talbot et al. (2014) and shown in Figure 2. In this presentation, data from several series over a common baseline are combined into stacks. however the readability within a given series is compromised in support of an overall understanding of the extent of the sum of the series. The difficulty of visually separating a single series from the stack is evident.

Other work improves on the data-ink ratio of the bar char, providing the hat graph, shown in Figure 3, reproduced from the paper developing this technique, (Witt, 2019). In the hat graph, only the tops of the bars of a bar chart, providing a floating appearance of the relevant data, and using as visual anchor the top of the first bar. This reduces the amount of int required, to present the same data as the analogous bar chart, as shown in Figure 3.

1.2 Identifying the Need

We wish to extend the use of the waterfall plot for the specific context in which we have multiple series of data, and where the initial point is the focus of comparison across all points on the *x* axis. We break the association with a single baseline, allowing various types of comparisons across and among the data series within the plot, unifying the whole within a common *y* axis.

The next section of this paper will outline our proposed solution to this problem, and outline the defining features of our idea. Following this we will evaluate our tool on several interesting data sets drawn from real world problems. Finally, we present a discussion summarizing the data trends, patterns and comparisons that can be expressed by the curtain graph.

2 PROPOSED VISUALIZATION

The motivation for this visualization comes from the desire to represent three-dimensions within a clear two-dimensional set of axes. The measures along the *x* dimension form a series relative to a fixed starting location, but where this starting location is not necessarily the same for other series within the same *x* variable. The *y* axis provides a common system for interpretation, and a set of data series arranged horizontally captures a variable in the *z* axis.

Thus, we present the curtain graph (Figure 4). The curtain graph is a novel, comprehensible method of visualizing high dimensional, multivariate data. This plot allows correlations of various attributes of a data set involving relative and absolute value comparisons to be easily identified on a single plot.

2.1 The Curtain Graph

The fundamental property of the curtain graph is the inclusion, within a single *y* axis, of a series of barchart type representations, allowing the visualization to be especially useful when comparing data with multiple treatment types. Figure 4 shows a detailed example of the curtain graph.

In Figure 4, each of the bar chart representations, which we will now refer to as 'subplots,' are placed vertically in a position on the curtain graph which acts as an anchor for that subplot. This anchor represents a common baseline, from which each bar from each subplot extends. A single subplot is identified in Figure 4 within the single line yellow box.

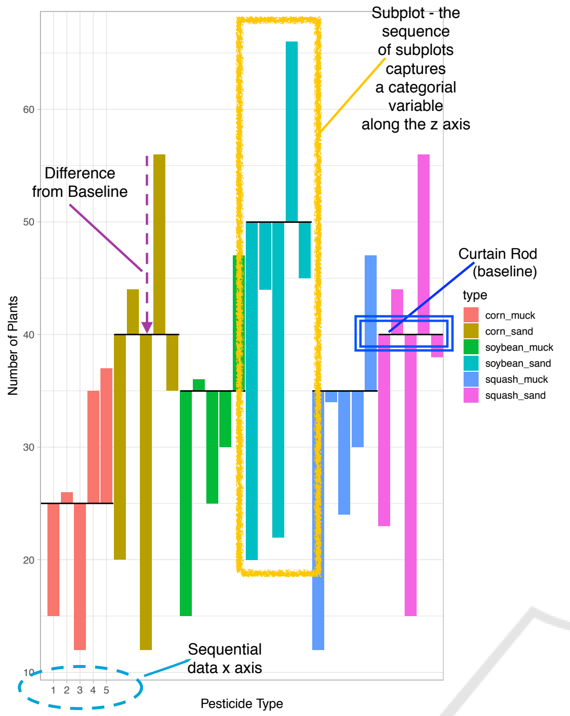

Figure 4: The anatomy of a curtain graph; colour/line variations used to show the graphical terminology used to refer unique areas on the plot. This figure was rendered from R using `ggplot2` and modifications to the `geom_rect` functionality.

We will refer to the this floating baseline as the 'curtain rod' as it provides the anchor for the strips forming the 'curtain' to ascend or descend. This floating axis rod is shown in Figure 4 enclosed by a blue double box near the right hand side of the figure.

The subplot consisting of the floating curtain rod axis and the attached bars can be placed in a sequence representing some categorical variable. Figure 4 presents the data obtained when measuring the response variable "Number of Plants" obtained when applying one of six (numbered) pesticide preparations to plants growing in various 'muck' and 'sand' soils. For each soil type there is a base expected yield, and the pesticide performance is measured against this base yield.

Such a sequence is shown in Figure 4 within the dotted ellipse identifying the individual bars making up the curtain running along the rod presented within the leftmost subplot of the figure. This sequence of the bars allows a second variable within an inner x axis, providing side-by-side comparison of some variable for whom the values form the basis of the most sensitive consideration in the visualization design (as opposed to the set of subplots forming the outer, or z axis, following the advice from Talbot et al. (2014). In this case, this would be pesticide type.

By placing the bars of the subplots all together at the same y coordinate, the absolute and relative position of each of the bar from the curtain rod of the bars respective subplot can be identified. Each subplot contains the data for a specific instance of a variable of interest for which we have performed repeated measures. The same bar within each curtain then pertains to a measure of the same variable within a different data series.

As noted in Cleveland and McGill (1984), position along a given access is an easily apprehended measure, and in particular more easily understood than data represented purely by length. In the extensive experimentation provided by Talbot et al. (2014), these conclusions are extended by statements about relative bar length, separation and position that support our placement within disparate series.

Because the curtain rod acts as the anchor for comparison of the variables within each subplot, the units of measure forming the y axis must be conductive to measurement within this number line. Each bar of the curtain provides an easily approachable measure of the degree of deviation from the baseline, while the absolute position of the ends of these bars provides an ability to compare the data across the full graphic, in absolute terms. This allows both simple absolute comparisons of values as typically found on a standard bar chart, or in a waterfall plot, as well as relative comparisons within the series attached to each floating curtain rod baseline.

Note the distinction from the hat graph described by Witt (2019); while both graph types display an apparently floating sequence, in the hat graph, this sequence is driven by the need to display only the tops of bars arising in the same direction from a baseline at an axis. Here, the baseline itself is floating, and the potential variability in sign requires us to keep the bars linking the data with each floating baseline.

While the properties commonly held on any number-line axis measurement (such as linear, logarithmic *etc.*) must be considered for the y axis, the x axis can be comprised of other variable types such as ordinal or categorical. The x axis of the curtain graph is comprised of the axes on the subplots. The number of labels on each x axis will correspond to the number of bars extending from each curtain rod. For example, with three subplots, each containing six bars, the x axis would consist of three sequences of multiply measured variables with six labels.

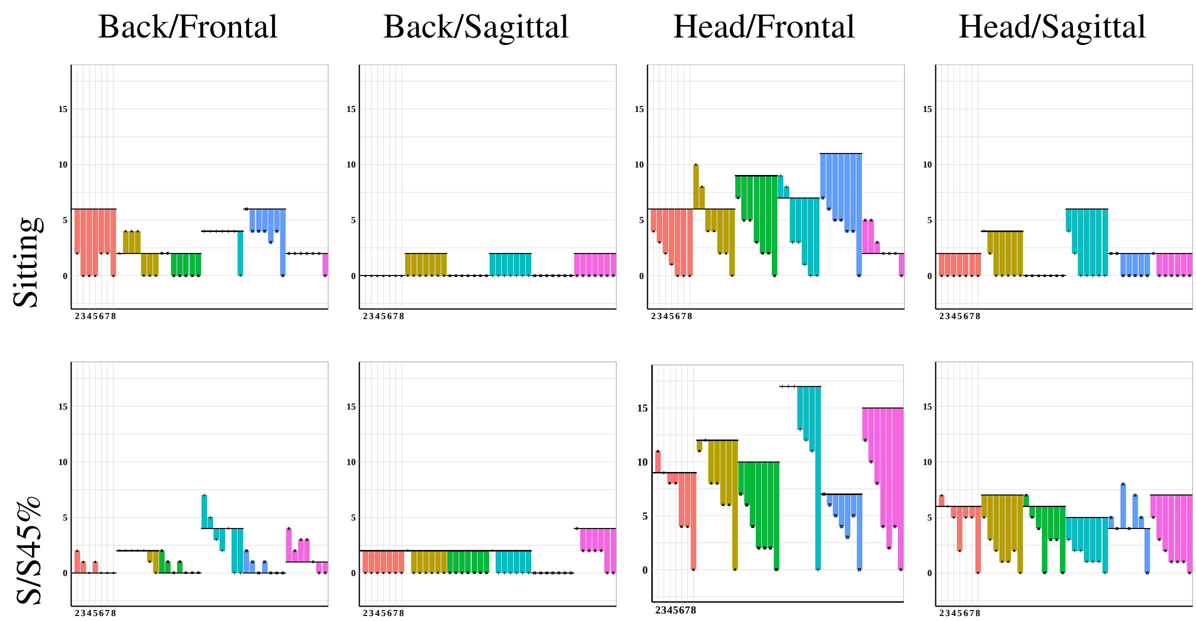

Figure 5: Four subplots showing the comparison of five dimensions of data using the curtain graph. Subplot curtain identifiers: Back Avoid. Back Risk. Fatigue Avoid. Fatigue Risk. Neck Avoid. Neck Risk.

3 REAL-WORLD APPLICATIONS

The applications of the curtain graph are very robust; it can be used in many fields, with different types and dimensionalities of data. To understand the type of data that can be expressed using the curtain graph, we outline three real-world examples of the visualization.

3.1 Example: Posture Data

The data analysis problem which spurred our interest in this type of visualization is a problem in postural data analysis. (This data and the soil type data available at the authors' website `https://qemg.uoguelph.ca/data/`.) This data analysis is characterized by evaluating observations regarding perception ('perception') across multiple channels simultaneously ('channel'), as well as across multiple experimental setups ('modality'). For each of these, we needed a visualization to explore the relationship between the effects of progressive degrees of application of a new filter ('filtering') and the response variable, which in this case was the number of patterns obtained under the filtering strategy ('number of patterns').

The dataset: Researchers are studying sedentary behaviour in office workers and are concerned with understanding gestures that are associated with the risk and avoidance of back pain, neck pain and fatigue in four different bodily modalities at three different workstations. The results of this work were presented at the July 2019 meeting of the Canadian Association of Ergonomists (Raymond et al., 2019).

The data consists of angular data obtained from the head and neck measured as an incline from vertical. This data was obtained for a the set of perceptions mentioned above (risk and avoidance of fatigue and of pain at the neck, and of pain at the back), and under experimental modalities of controlled sitting and standing alternating within a 20 minute cycle. Example data shown here includes at 45% standing (S/S45%) and sitting, for several channels. Only the channels 'Head/Sagittal' and 'Back/Frontal' are reproduced here, due to space constraints.

This results in a numerical y axis with a numerical x axis along each curtain rod and a categorical z between the series of bars (Figure 5).

In Figure 5 the curtain graphs are arranged in a grid table similar to a matrix. Note that the axes, and colour/pattern in each curtain graph is consistent with the other curtain graphs. The column and rows of the matrix are based on the planes of the body and each workstation, respectively.

Representation of this five dimensional problem gave birth to this visualization strategy through the observation that the critical comparison was the iterative application of the 'filtering' factor. As the filtering is occurring relative to the state of a baseline data set, new values for filtering are understood only in the context of observations given in relation to the initial or base point for that series — we can see any changes as a refinement of an initial setup. Applying successively greater amounts of filtering yields a series of points in a logical progression. This requires a local display space in order to be coherently understood relative to the base position, but which also require comparison with our other variables.

As shown in Figure 5, the curtain plot displays this data in order of comparison precedence, moving up from the number of patterns obtained through successive filtering through data channel and experimental modality. Specifically:

1. In a typical multi-plot grid, the two least critical variables lay out a grid of inner plots, indexed by channel and experimental modality. This reduces the remaining visualization dimensionality needs to three.
2. Within each plot, we can then associate the number of patterns on the y axis with both perception, shown as the group of measures within a given series placed along a specific curtain rod, and the filtering values arranged along the rod.

The filtering setting is the experimental variable that we wish to fall under the closest scrutiny. By setting these within the adjacent bar sequence arranged along each 'curtain rod' we easily see the effects of increasing the filtering control, moving from no filtering at the leftmost end of the rod through to a filter setting of 8 units at the rightmost end of each series. The curtain rod bars themselves float in space at the baseline position provided by no filtering, which allows easy and direct comparison of any of the number of patterns obtained by different filtering settings within a series to the number obtained at baseline.

For example, in the lower leftmost plot in Figure 5 (S/S 45%, Back/Frontal) the sequences for 'back avoid' and 'back risk' are easily recognized as having little change relative to their baseline position, and also immediately indicate that the direction of the changes that are seen are in opposite sign to one another, back avoid projecting upwards from the rod, while back risk descends. The relative position of these two rods in this plot show that while the range is the same, the opposite change is shown. Comparing these two series to the other series in the plots, we can easily locate those where significant change is apparent, as well as the overall trend of the change.

In the display for both experimental modalities of the Head/Frontal data large visible trends of decreasing measures are shown. The cascade of bars down from the curtain rod gives us the name for the tool. Again, several insights are easily captured here that would not be available without the relative positioning of the curtain rod chart:

- the similarity of structure across all the subplots;
- the right shifting along the rod of 'back risk' and 'fatigue risk' within S/S45%; and
- the similar trend but different initial sign in 'back avoid'.

Turning to the Back/Sagittal column, we easily see here the consistency of the changes noted, as well as the distinct representation of an unfilled baseline curtain rod. While in this data all of the empty rods happen to be at the $y = 0$ axis, it is easy to imagine plots in which they might be scattered.

Overall, this plot type employs the type of immediate insight described as 'small multiples' by Tufte (1991, pp. 67–80) which allows immediate recognition of visually accessible patterns within the display. As the purpose here is to facilitate recognition of both overall trends as well as patterns of association of the 'number of patterns' variable across the many plots, the concept of the small multiple is particularly important here. For example, the power of the plot is shown in particularly striking fashion by the way that 'neck avoid' on the lower right plot leaps from the page (S/S 45%, Head/Sagittal), showing how unusual cases are easily recognized relative to the floating baseline.

Movement along the x axis can be identified as progression when labels are numerically and temporally sequential in Figure 5. Here, the progression in the x axis corresponds to the increase in the length of the minimum persistent duration filter. Further, when the curtain graphs are grouped together in a matrix, the patterns of different combinations of data can be easily identified.

3.2 Example: Crop Data

The curtain graph can also be used to visualize a data set consisting of multiple categorical factors as well as numerical data. Several modifications to colour/pattern and symbols are shown to emphasize specific trends in the data set.

The dataset: Researchers are interested in studying the survival rate of three crops, in two different soil types after the application of five pesticides compared to when there are no pesticides applied. The crops are soybean, corn and wheat and the soil types are sandy and muck soil (sapric organic soils). The pesticides used are simply labelled one to five, and zero (for no pesticide applied).

Figure 6 displays the data on the curtain graph. Colours are used to distinguish the subplots which represent the different soil/crop types. Each bar within a subplot shows the plant survival rate of the crop it represents in comparison to the baseline. Here,

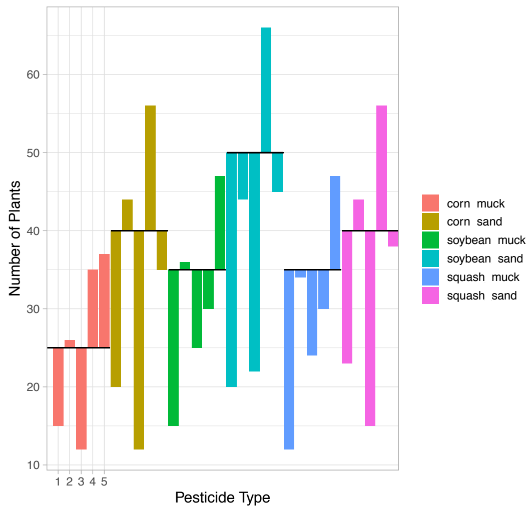

Figure 6: A curtain plot showing the data from the "crop" data set.

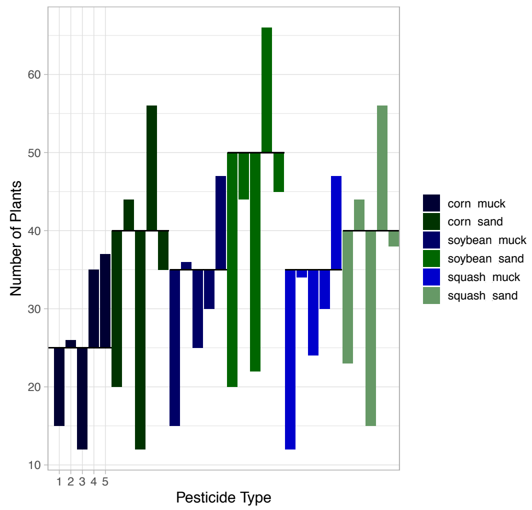

Figure 7: A curtain plot showing the data from the "crop" data set, with updated colour scheme to emphasize soil type.

pesticide level zero (no pesticide applied), acts as the curtain rod (baseline), which is the solid black line on each subplot. The extension of the bars from the curtain rod therefore represents the absolute and relative difference in survival rate from when there is no pesticide applied.

Note that while here, 'no pesticide applied', acts as the curtain rod, if the researcher was interested in the difference in survival rate from a gold standard pesticide, they could simply use that pesticide as the curtain rod anchor point.

The curtain graph can be modified using colour, pattern or symbols to visually emphasize different trends or patterns in the data. For example, the variation in pattern between soil type can be explored in this manner. From Figure 6, the viewer can conclude that for this experience, the baseline survival rate of all crops was higher when plants were grown in the soil type identified as muck soil, compared to sandy. To easily distinguish this trend, the curtain plot could be refined using shades of the same colour or varieties of the same pattern, to show which that these crops were all grown in the same soil type (Figure 7).

3.3 Example: Non-linear Axes

The curtain graph is robust in its ability to accommodate non-linearity in the y axis, in spite of the need to support both relative and absolute data comparisons. For example, the curtain graph can be used with a logarithmic axis.

The dataset: This UNICEF (2019) dataset (available from UNICEF at `https://mics.unicef.org`) consists of the estimated number of deaths by diarrhoeal disease in children under five in the year 2000

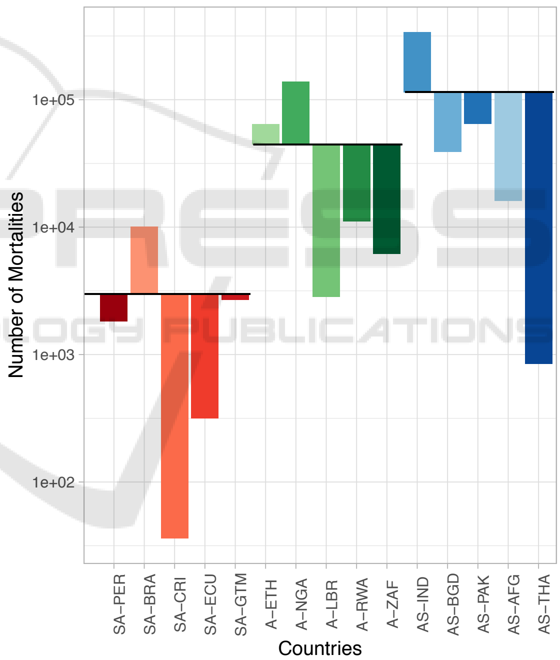

Figure 8: A curtain plot showing the data from the 'UNICEF' data set. Sublots with data beginning 'SA-' are South America, 'A-' are Africa, and 'AS-' are Asia. Countries are rendered in shades of red, green and blue respectively.

by country. The countries used were organized into geographic region. South America is represented by Peru (SA-PER), Brazil (SA-BRA), Costa Rica (SA-CRI), Ecuador (SA-ECU) and Guatemala (SA-GTM). Africa is represented by Ethiopia (A-ETH), Nigeria (A-NGA), Liberia (A-LBR), Rwanda (A-RWA) and South Africa (A-SA) and Asia consists of data from India (AS-IND), Bangladesh (AS-BGD), Pakistan (AS-PAK), Afghanistan (AS-AFG) and Thailand (AS-THA).

Figure 8 shows the described UNICEF dataset using a $\log_{10} y$ axis. Here, each subplot represents a geographic region, while the bars each represent a different country within that region. Visually, each geographic region is a different colour, with the bars consisting of different hues of this colour. The coordinates of the curtain rod were calculated by taking the mean mortality rate of children under five due to diarrhea in the region.

We chose the mean as the anchor point for the curtain rod baseline to provide an example of a different type of analysis, where variation around a central theme, rather than a series based on initial conditions, is the focus of the visualization.

As there is still a common space defined by the y axis, and there is a meaning to the anchor point and the extent of the curtain bars, all of our previous observations on interpretability still hold in this scenario.

4 DATA PROPERTIES

The fundamental property of this graph is the inclusion, within a single y axis, of a series of bar chart type representations, where the vertical position of the baseline of each of the individual sub-charts is placed at a position representing the anchor of that data series. By placing these together within the same set of coordinate systems, both the absolute and relative position of each of the bars within the various sub-charts correctly represents the data changes both across and between the data series.

The units of the measure forming the y axis must therefore be conducive to measurement within this number line, so this measure must obey the properties of a metric space — in other words the properties of a distance measurement must hold:

- points falling at the same position must indicate that there is no difference in the measurement (there must be only one measured value that falls at each point);
- the distance between any two points mapped to the y axis must be the same regardless whether measured in the positive or negative direction; and
- the triangle inequality must hold, such that the sum of two consecutive distances on this line must represent a difference at least as large as that between the two end points.

These properties commonly hold on any number line based axis measurement, such as linear, logarithmic, *etc*. The other axes within the visualization are free to be other non-metric types, such as ordinal or categorical.

The motivation for this visualization comes from the desire to represent three dimensional data within a clear two-dimensional plotting space. We are interested in the case where the measures along our x dimension form a series relative to a fixed starting location, but where this starting location is not necessarily the same for the other series within the x variable. The existing waterfall chart also presents a series of values relating to a fixed starting point, but in that case, the starting point must be $y = 0$. In addition, we combine multiple series with different starting points in the same plot.

5 CONCLUSIONS AND FUTURE WORK

The visualization tool presented allows for the visualization of high-dimensional multivariate data in one conceptual two-dimensional plot. We are deliberate with the layout of the curtain graph such that it allows for comparisons between and within the interrelated data groupings of a given data set. The curtain graph is flexible in its ability to accommodate data of different types; while the y axis must be numerical, while the x and z axes can be categorical, ordinal or numerical.

In this paper, we outline three different real-world examples from different disciplines and data sets, providing the reader examples of the benefits of using the curtain graph and its ability to visualize information in a fashion that was not previously as coherent. Further, we provide the properties of data that must be true for an individual to use the curtain graph.

Future work will consist of creating a software package for the creation of the curtain graph. We currently use modifications to the 'geom-rect' functionality in R's `ggplot2` package (Wickham, 2016) to create the graphs. Ideally, this would allow a user to simply format their data in a data frame and curate the graph using a function call. This will allow for use of the visualization from a broader audience.

ACKNOWLEDGEMENTS

We would like to thank Nancy Black of the University of Moncton, New Brunswick, for providing us with the data that inspired the early workings of our visualization. Further, thank you to Jack Johnson for motivating the use of the term 'subplots' in our work.

REFERENCES

Agrawala, M., Li, W., and Berthouzoz, F. (2011). Design principles for visual communication. *Communications of the ACM*, 54(4):60–69.

Bertin, J. (2001). Matrix theory of graphics. *Information Design Journal*, 10(1):5–19.

Cairo, A. (2012). *The Functional Art*. Pearson Education.

Cairo, A. (2016). *The Truthful Art*. New Riders.

Cleveland, W. S. and 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.

Daru, M. (2001). Jacques Bertin and the graphic essence of data. *Information Design Journal*, 10(1):20–25.

de Almeida, A. T. and Roselli, L. R. P. (2017). Visualization for decision support in fitradeoff method: Exploring its evaluation with cognitive neuroscience. In *Proceedings of the 2017 International Conference on Decision Support System Technology*.

Evergreen, S. D. (2020). *Effective Data Visualization: the Right Chart for the Right Data*. Sage Publications Ltd., second edition.

Few, S. (2006). Graph designs for reviewing transactions and the changing balance. Technical report, Perceptual Edge.

Heer, J., Card, S. K., and Landay, J. A. (2005). prefuse: a toolkit for interactive information visualization. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI'05)*.

Hicks, M. (2009). *Trends in Interactive Visualization*. Advanced Information and Knowledge Processing. Springer.

Holtz, Y. (2019). Grouped, stacked and percent stacked barplot in ggplot2. Technical report, R Graph Gallery.

Inbar, O. I., Tractinsky, N., and Meyer, J. (2007). Minimalism in information visualization: Attitudes towards maximizing the data-ink ratio. In *Proceedings of the 14th European conference on Cognitive ergonomics: invent! explore!*, pages 185–188.

Klanten, R., Ehmann, S., and Schulze, F., editors (2011). *Visual Storytelling: Inspiring a New Visual Language*. gestalten, Berlin.

Klimov, D., Shahar, Y., and Taieb-Maimon, M. (2010). Intelligent visualization and exploration of time-oriented data of multiple patients. *Artificial Intelligence In Medicine*, 49(1):11–31.

Metoyer, R., Lee, B., Henry Riche, N., and Czerwinski, M. (2012). Understanding the verbal language and structure of end-user descriptions of data visualizations. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI'12*, CHI '12, pages 1659–1662, New York, NY, USA. ACM Press.

Raymond, K., Hamilton-Wright, A., and Black, N. L. (2019). Postural state sequence duration and musculoskeletal disorder marker perception at sit-stand workstations. In *Proceedings of the 50th Annual Conference of the Association of Canadian Ergonomists*. ACE'19.

Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. In *Visual Languages, 1996. Proceedings., IEEE Symposium on*, pages 336–343.

Talbot, J., Setlur, V., and Anand, A. (2014). Four experiments on the perception of bar charts. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2152–2160.

Tufte, E. R. (1991). *Envisioning Information*. Graphics Press, Cheshire, Conneticut.

Tufte, E. R. (1997). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press, Cheshire, Conneticut.

Tufte, E. R. (2001). *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Conneticut, 2nd edition. Original edition 1983.

Tufte, E. R. (2006). *Beautiful Evidence*. Graphics Press, Cheshire, Conneticut.

UNICEF (2019). Multiple indicator cluster survey. Technical report, United Nations Children's Fund.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Witt, J. K. (2019). Introducing hat graphs. *Cognitive Research: Principles and Implications*, 4(1).

Zen, M. (2013). Metric-based evaluation of graphical user interfaces: Model, method, and software support. In *Proceedings of the 5th ACM SIGCHI Symposium on Engineering Interactive Computing Systems*, EICS '13, pages 183–186, New York, NY, USA. ACM.