Preprint

M. Hlawatsch, F. Sadlo, M. Burch, and D. Weiskopf.
Scale-Stack Bar Charts.
To appear in *Computer Graphics Forum*, 32 (3), 2013.

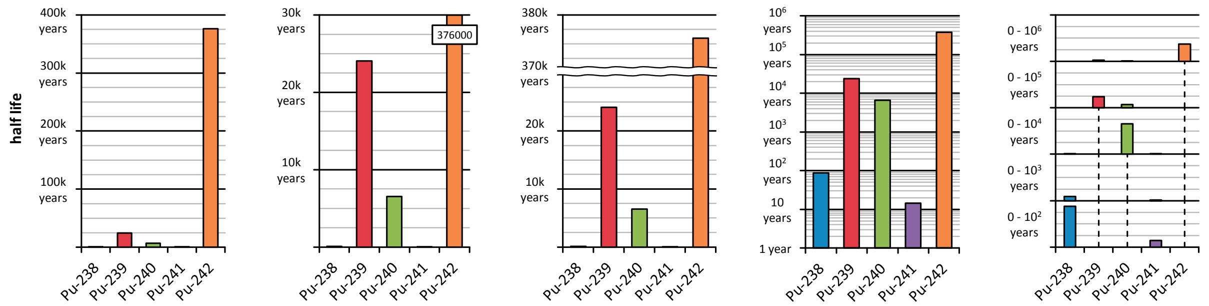
Eurographics Conference on Visualization (EuroVis) 2013
B. Preim, P. Rheingans, and H. Theisel
(Guest Editors)

Volume 32 (2013), Number 3

Scale-Stack Bar Charts

M. Hlawatsch, F. Sadlo, M. Burch, and D. Weiskopf

Visualization Research Center (VISUS), University of Stuttgart, Germany



(a) (b) (c) (d) (e)

Figure 1: *Different bar charts for visualizing data with large value range. Half lives of different plutonium isotopes (see Sect. 6.1) are visualized with: (a) classic linear bar chart, (b) linear bar chart with cut-off bars, (c) linear bar chart with scale break, (d) logarithmic bar chart, and (e) our novel scale-stack bar chart.*

Abstract

It is difficult to create appropriate bar charts for data that cover large value ranges. The usual approach for these cases employs a logarithmic scale, which, however, suffers from issues inherent to its non-linear mapping: for example, a quantitative comparison of different values is difficult. We present a new approach for bar charts that combines the advantages of linear and logarithmic scales, while avoiding their drawbacks. Our scale-stack bar charts use multiple scales to cover a large value range, while the linear mapping within each scale preserves the ability to visually compare quantitative ratios. Scale-stack bar charts can be used for the same applications as classic bar charts; in particular, they can readily handle stacked bar representations and negative values. Our visualization technique is demonstrated with results for three different application areas and is assessed by an expert review and a quantitative user study confirming advantages of our technique for quantitative comparisons.

Categories and Subject Descriptors (according to ACM CCS): H.5.0 [Information Interfaces and Presentation]: General—

1. Introduction

Bar charts are an established tool to visualize quantitative data. They can be explored in an intuitive and accurate way due to their encoding of quantities in the visual variables length and position [CM86]. Consequently, such bar chart-based diagrams are a powerful and commonly used tool to communicate a set of quantities.

However, there is a variety of applications where the data covers a large value range. Examples include the luminosity of stars in astrophysics, the half life of different isotopes (Fig. 1), sales and profits of companies, or the population of different countries. In such cases, it is challenging to create appropriate bar charts. Depending on the available image resolution and viewing conditions, bar charts with linear scales cannot display all values. A common approach is to use logarithmic scales, however, at the cost that their non-linearity increases the difficulty of certain tasks.

We present a novel method to visualize large value ranges with bar charts: *scale-stack bar charts* (Fig. 1(e)). Our approach does not exhibit the problems of non-linear and non-proportional mappings and allows small values to be still readable. It is especially suitable for tasks requiring quantitative comparison as our expert review and quantitative user study show. Possible applications that can benefit from our diagram technique are described in Sect. 6.

2. Related Work

There are several approaches to address the problems of linear bar charts (Fig. 1(a)) when displaying large value ranges:

- **Cut-off bars:** The bars corresponding to large values are cut off and labeled with their value (Fig. 1(b)). This allows visual comparison of small values but not of the values exceeding the scale range.
- **Scale break:** Another approach to bridge the gap between small and very large values uses a scale break [CM92] (Fig. 1(c)). This technique inserts a gap in the scale to provide more space for representation of small values. Here, a visual comparison across the break is not possible, and only the heights of the bars below the scale break are proportional to the respective values.
- **Logarithmic mapping:** Values are mapped with a logarithmic scale (Fig. 1(d)). This approach has the benefit that small values are still readable. Also, it is a reasonable approach for data representing exponential processes, like growing populations. However, the non-linear mapping makes a quantitative comparison of values more difficult.

Unfortunately, the above mentioned variants of bar charts introduce a lie factor [Tuf92]. A lie factor exists if the size of an effect in the data and in the respective visualization is not equal. The variants all have problems with respect to the quantitative comparison of values, which is one of the most important and best performing tasks when using bar charts with linear scale. Furthermore, applying a logarithmic mapping, e.g., to a stacked bar chart results in an even more misleading visualization scenario.

There are further ways to modify bar charts. For example, the value axis does not have to be zero-based but can start at a different value. However, such modifications cannot generally solve the problems with large value ranges.

In this paper, we extend bar charts with different scales depending on the exponent under observation. With this approach, comparisons between small values can be carried out accurately, and at the same time, larger values can be compared among each other and in the context of the smaller values. Furthermore, our proposed charts are also free of chart junk [BMG*10, Su08], which refers to the fact that data is visually encoded in several redundant visual features.

Our work focuses on bar charts as they are a common tool and widely used, but there are of course other concepts for visualizing quantities, too. Many of them and general discussions of data visualization and statistical graphics can be found in Cleveland and McGill [CM87], Huff [Huf93], Card et al. [CMS99], and Ware [War04]. Few [Few07, Few10] illustrates the drawbacks of radial pie charts, while in the study of Cleveland and McGill [CM86], the performance of several charts was measured. Also, the work of Goldberg and Helfman [GH11], who investigated the readability of values on linear versus radial graphs by applying eye tracking techniques, showed that choosing a non-radial diagram is beneficial with respect to user performance for many tasks. Further discussions of radial methods can be found in Draper et al. [DLR09] and Diehl et al. [DBB10]. Product plots [WH11] are a general framework for statistical graphics that can also handle bar charts. Finally, we want to point out that even well-established and rather simple techniques like bar charts may have space for improvements as the work by Talbot et al. [TLH10] shows in the context of positioning tick labels on chart axes.

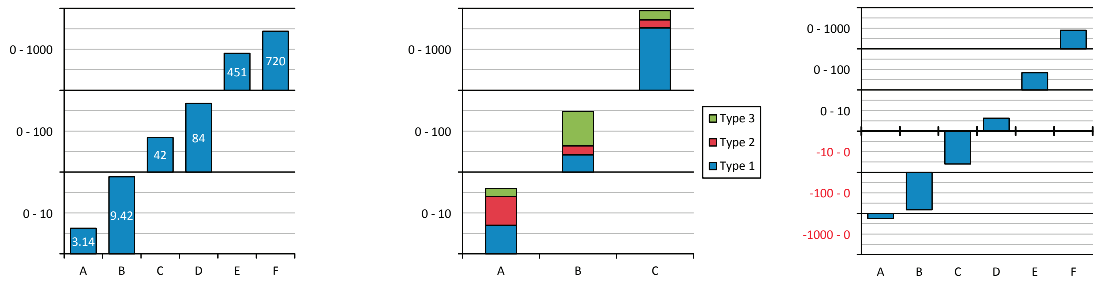
3. Visualization Method

Our approach is inspired by the scientific notation of numbers and the way floating point values are commonly represented by computers. We split the values into two parts: the order of magnitude and their representation in this order of magnitude, similar to the exponent and mantissa in scientific notation. The result is a two-dimensional representation of numbers. There is a multitude of different two-dimensional mappings, even if we restrict ourselves to the bar chart metaphor, e.g, different spatial dimensions or color can be used. We tested many different designs in a formative process and ended up with the following approach.

3.1. Basic Method

We subdivide the value axis of the chart to represent different scales respectively orders of magnitude (Fig. 2(a)), and call this a scale-stack. Now, the important thing is that inside each scale, a linear mapping of the values is used. Furthermore, every scale starts again at zero. This approach preserves the advantage of linear bar charts: easy comparison of values. In summary, we use a two-leveled hierarchical representation of exponent and mantissa, which is possible due to the discrete nature of the exponent in our representation.

This layout has several advantages. First, both components—mantissa and exponent—can be directly compared. This would not be the case with other mappings, e.g., if two different spatial dimensions would be used. Second, using a different row for every scale lowers the risk that values of different scales are directly compared. Finally, the usage of length for number representation allows accurate quantification and comparison [CM86].



(a) (b) (c)

Figure 2: *Illustration of our concept. (a) Every row (delimited by bold lines) covers a complete scale; the bars always start at 0 in every row. Inside each scale, a linear mapping is used. (b) Because of this mapping, our method can be used to display stacked bars. (c) For negative values, the chart is extended downward as it is also a common approach with classic bar charts.*

Our approach augments the bar chart metaphor only in the value dimension. Hence, there are only constraints in this dimension. Other dimensions and layout aspects remain as flexible as in classic bar charts. For example, the color coding of the bars as well as their sorting and shape can be adapted according to the respective needs. Since we use a linear mapping within each scale, typical variations of classic bar charts can also be used, e.g., stacked bars (Fig. 2(b)). Furthermore, it is straightforward to display negative numbers (Fig. 2(c)). We demonstrate our method with vertical bar charts. However, the application to horizontal bar charts is straightforward and therefore not further discussed.

It is clear from the concept that the ranges of the different scales have a strong influence on the resulting visualization, as it is also the case for classic linear bar charts. If a range is too large, it is difficult or not possible to read small values. If a range is too small, only few values are captured. Because there are multiple scales, quite some effort is required when manually choosing a set of appropriate ranges. We therefore describe a simple approach to automatically adapt the scale ranges to the data.

3.2. Automatic Scale Selection

We present a simple algorithm to automatically choose the ranges for the different scales. It requires the values of the chart and the desired number of scales as a user parameter. We decoupled the problem of choosing an adequate number of scales from our algorithm to keep it simple. In general, the problem of choosing adequate scale ranges can be seen as an optimization problem. The goal is to maximize the size of the individual bars for better readability.

The height h of a single bar for the respective value v inside a specific range $[r_{\min}, r_{\max}]$ can be determined with $h = v/r_{\max}$, because $r_{\min} = 0$ in our approach. For simplicity, we describe our approach for positive v and $r_{\max}$. Please reverse their signs otherwise to match our descriptions. In our algorithm, values are assigned to the scale s where $r_{\max} - v$ is minimal and positive. We want to ensure that the smallest assigned values inside a scale are still visible. We therefore maximize their height by maximizing the following objective function:

$$f = \sum_{s \in S} \min_{v \in s}(v) / \max(s),$$

where S is the set of all scales in our chart and $\max(s) := r_{\max}$ of scale s.

We use a greedy approach for optimization. We start with n scales, where n is the number of values in the chart, i.e., every value has initially its own scale with $r_{\max} = v$. In every iteration step, two scales s' and s'' are merged to a single scale with $r_{\max} = \max(\max(s'), \max(s''))$ containing all values of both scales. To choose the best pair of scales for merging, every possible pair is temporally merged. The pair that maximizes the objective function is then finally merged. This is done until the desired number of scales is reached. Operating on sorted values eases the implementation and may improve the efficiency because only neighboring scales have to be considered for merging. As it is the common way for tick labels in charts, each $r_{\max}$ can be rounded, e.g., to the next order of magnitude.

This approach does not always find the globally optimal solution, but it is simple, fast, and provided good results in our experiments. An improved algorithm and an approach to automatically select the number of scales exceed the scope of this work and will be part of future work.

3.3. Variants

Even with well-adapted scale ranges, there can be the issue that nearby values may be mapped to different scales (Fig. 3(a)). This makes it difficult to compare them. To overcome this issue, values can be displayed on all scales with a proper range (Fig. 3(b)), i.e., where the upper limit of the

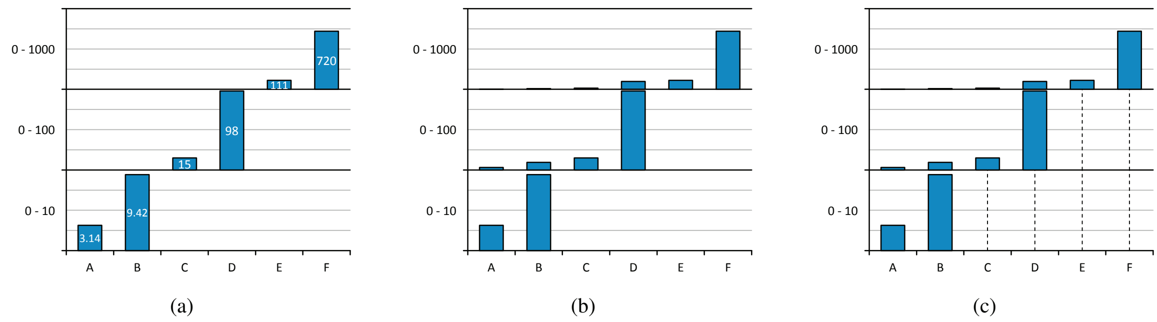


Figure 3: *Variations of our method. (a) Nearby values may be mapped to different scales (e.g., B and C). This makes their comparison difficult. (b) Displaying the values on multiple scales avoids this problem. The values of B and C can now be compared on the second scale. In this case, the values are shown on all scales that cover the full value range. (c) Dashed lines can be used as placeholders to show that a value exceeds the range of the respective scale.*

range is larger than the value. Nearby values can now be compared on the respective scale. A problem of this solution is that it is not clear if a column is empty because the value is too small to be visible or because it exceeds the respective scale. Drawing all values on all scales and clamping the bars if the value exceeds the scale range would avoid this problem. However, the filled columns could be misinterpreted as large bars. This would increase the risk that bars of different scales are compared. We therefore use a placeholder to represent values that exceed the range of the respective scale (Fig. 3(c)). Our experiments and the expert review conducted (see Sect. 4.1) lead to dashed lines as placeholders.

However, there are also drawbacks for these variants. The more visual elements are in the chart, the higher the risk of visual clutter. Furthermore, the multiple bars per column can additionally confuse the viewer. Therefore, we see these variants as application specific. In cases where values can be clearly assigned to a specific scale and comparability is only important inside these scales, the basic approach (Fig. 3(a)) may be the best choice. It produces the clearest and simplest chart. In cases where many nearby values are mapped to different scales, it is important to show them on all possible scales (Fig. 3(b)). The usage of placeholders (Fig. 3(c)) can be additionally confusing and increases the visual load of the chart. They can be omitted if it is clear to the viewer on which scales the values are located, e.g., because they are already ordered by size.

4. Evaluation

We first used an expert review to estimate the potential of scale-stack bar charts and then evaluated them with a quantitative user study measuring answer times and error rates.

4.1. Expert Review

Expert reviews can help design and evaluate visualization techniques [TM05] [AJDL08]. We conducted an expert review with 12 visualization experts from our institute. They filled out a questionnaire, for which most of them required around 25 minutes. We compared our approach to linear and logarithmic bar charts in this review. The experts had to rate the suitability of the charts for different tasks with the three datasets from Sect. 6. They used a Likert scale with values from 1 (very bad) to 5 (very good). Additionally, they were asked to provide comments to their ratings. We checked two types of tasks in the review. The first type requires only a qualitative comparison. The second type requires quantitative comparison and the estimation of ratios. The tasks and the respective results are presented in Table 1. ANOVA and additional pairwise post-hoc t-tests ($p < 0.05$) show a significant difference between the diagram techniques for all tasks. Furthermore, we used the expert review to select an adequate placeholder (see Sect. 3.3): dashed lines as placeholders provided the best trade-off between reduced risk for misinterpretation and visual design.

In summary, the experts think that linear charts work best if the data of interest are clearly visible in them, i.e., in this case for tasks on the largest values. They further think that logarithmic charts are good for qualitative comparisons like finding extrema, but they may be difficult to use for quantitative tasks like estimating proportions or quantitative comparison of values. Scale-stack bar charts seem to work best when the value range is too large for linear charts and the tasks require a quantitative analysis or comparison of values. However, the reviewers pointed out some issues of scale-stack bar charts in their comments: The reviewers needed some time to get used to them and jumping between scales is time-consuming. Some of them see also problems with the representation of negative values, because of the gap between positive and negative values (Fig. 2(c)).

Table 1: *Results of the expert review. The average ratings are shown together with the standard deviation in brackets. The best rated techniques are marked with boxes if they are significantly better in pairwise post-hoc t-tests. If two techniques are marked in a row, it means that they are significantly better than the third method, but none of the two is significantly better than the other. The first row has no clear winner because significance occurred only in one pairing (linear vs. scale-stack).*

		linear	logarithmic	scale-stack
half life data	find two highest values	5.00 (0.00)	4.83 (0.39)	4.63 (0.48)
	find two smallest values	1.08 (0.29)	4.50 (0.90)	4.46 (0.50)
	estimate ratio of two highest values	4.27 (1.01)	2.18 (0.98)	4.09 (0.83)
	estimate ratio of two smallest values	1.00 (0.00)	2.45 (1.13)	3.91 (0.83)
age distribution data	estimate ratio of stacked bars	2.75 (1.54)	1.42 (0.69)	3.83 (1.11)
	find anomalies in the data	2.25 (0.97)	1.50 (0.67)	3.50 (1.24)
corporation profits data	quantitative trend analysis inside groups	3.92 (0.79)	3.17 (0.72)	3.92 (0.67)
	quantitative trend analysis between groups	3.50 (1.00)	2.50 (0.90)	3.75 (0.97)

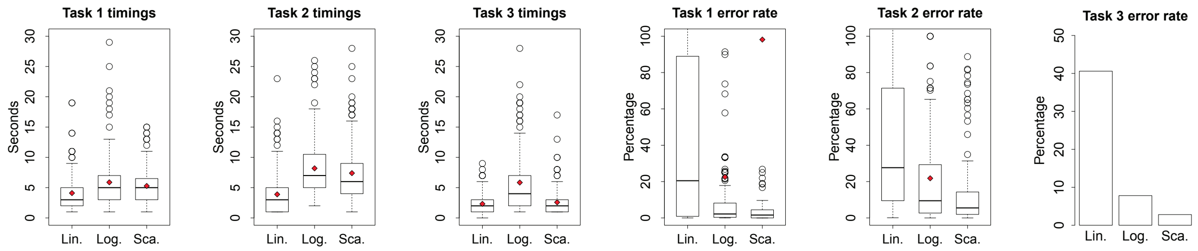


Figure 4: *Results of the user study. Boxplots show the distribution of the measured answer times with linear ("Lin."), logarithmic ("Log."), and scale-stack ("Sca.") bar charts for all tasks ((a)–(c)), and the error rate distribution for tasks 1 and 2 ((d), (e)). Red dots represent average values. If they are not visible, the average values exceed the scale. For the discrete results (correct or incorrect) of task 3, the percentage of wrong answers is shown in a bar chart (f). In all charts, lower values are better.*

4.2. User Study

We conducted a quantitative user study with 15 participants from our institute requiring around 25 minutes to complete. Each participant had to solve the following three tasks, each with linear, logarithmic, and scale-stack bar charts:

1. read value,
2. estimate ratio between two displayed values, and
3. determine which time step out of five exhibits the largest growth rate.

All data was generated with an overall value range of $[1..10,000]$. Each task had to be performed 12 times on different data samples. Permuted sequences of the same samples were used for the different chart types. The data values were equally distributed over four orders of magnitude and randomly generated inside each magnitude order. The participants were asked to perform the tasks with focus on speed. All answers and the required times were measured. The order of the tasks was fixed, while the order of the chart types was balanced across participants.

The measured timings and error rates are shown in Fig. 4. The error rate for tasks 1 and 2 was computed as $e = |1 - v_{user}/v_{real}|$, with v_{user} being the answer of the participant and v_{real} being the correct answer. The huge average values in the error rates of scale-stack charts for tasks 1 and 2 are caused by a few outliers. It is not clear if the participants were confused by the different scales in the chart in these cases, or if the outliers resulted from typing errors. Normal distribution of the data was rejected by the Shapiro-Wilk test, and the Kruskal-Wallis test revealed significance of the data for all tasks ($p < 0.05$), in both timings and error rates. Posthoc pairwise comparison of methods was performed with the Wilcoxon rank sum test. The participants were significantly faster with linear charts than with

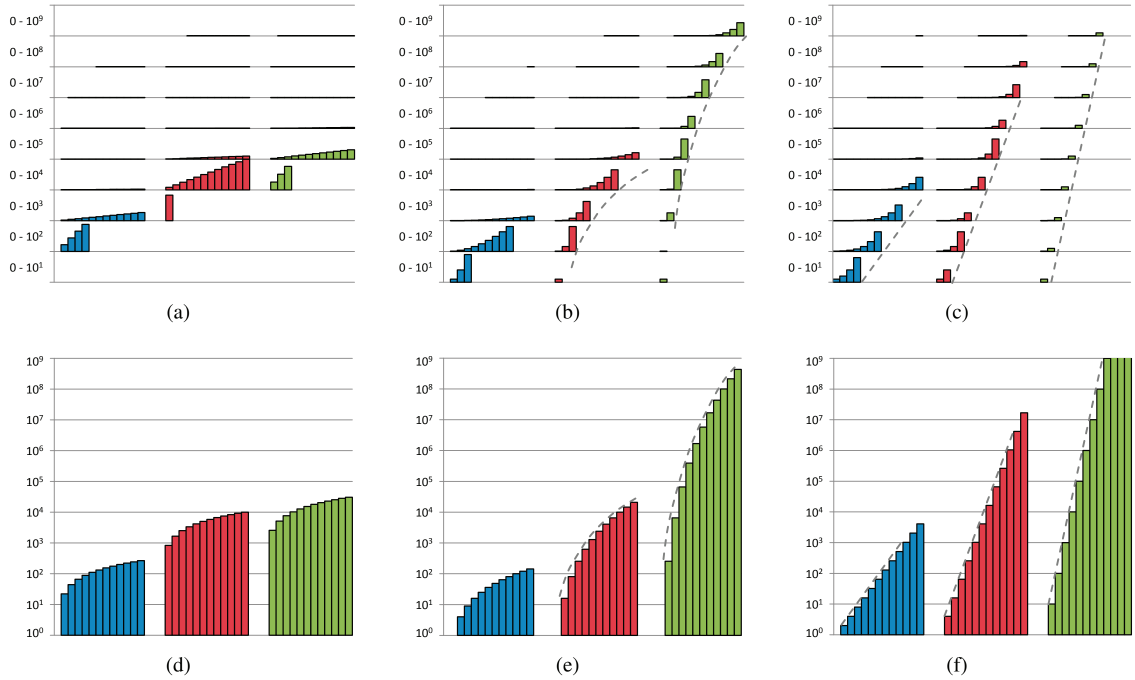


Figure 5: *Visualization of different growth behavior. (a) Our approach showing linear growth with three different slopes: $y = 22x$ (blue), $y = 831x$ (red), $y = 2546x$ (green). (b) Our approach showing polynomial growth with three different degrees: $y = x^2$ (blue), $y = x^4$ (red), $y = x^8$ (green). (c) Our approach showing exponential growth with three different bases: $y = 2^x$ (blue), $y = 4^x$ (red), $y = 10^x$ (green). The respective logarithmic plots are shown in (d)–(f). The envelopes (illustrated by dashed lines) of the logarithmic charts for polynomial and exponential growth can be approximated in our scale-stack charts by connecting the last element of each row. The envelope for the blue bars in (e) is not shown due to the small number of sample points in (b).*

the other two chart types in case of tasks 1 and 2. For task 3, there is no significant difference between linear and scale-stack charts, but they both exhibit significant faster answer times than logarithmic charts. However, the error rates of linear charts are significantly higher than the rates of the other chart types for all tasks. A significant difference between the error rates of logarithmic and scale-stack charts appears only for task 2, where scale-stack charts have lower error rates.

The user study shows that the participants could perform all tasks with scale-stack bar charts at least as fast and as accurate as with logarithmic plots. Especially, when a quantitative comparison of values is required, scale-stack charts provide higher accuracy (task 2), or faster answer times (task 1). Linear charts are fast to use, but they cannot accurately display values over the full range, resulting in high error rates in our study. These results coincide with the findings from the expert review (see Sect. 4.1).

5. Discussion

For a fair comparison in the following discussion, we assume that our complete chart has the same spatial extent and pixel resolution as the other chart types.

5.1. Perception of Growth Behavior

Depending on the type of scale, certain types of growth behavior can be readily detected in charts. It is easy to see linear growth with linear plots. The same holds for logarithmic plots and exponential growth. Figure 5 shows scale-stack bar charts and the respective logarithmic charts for three different types of growth: linear, polynomial, and exponential.

Since our approach consists of linear plots on different scales, it is easy to detect linear growth (Fig. 5(a)), in contrast to logarithmic plots (Fig. 5(d)). The bars of the same scale grow at constant rate in our approach. If we look at the polynomial and exponential growth, we can see that the shape of the bars of the logarithmic charts also appears in our scale-stack charts. As the figure shows, the envelopes of the logarithmic charts can be approximated in scale-stack charts. Therefore, it is also easy to differentiate between polynomial and exponential growth with scale-stack bar charts. In the case of polynomial growth (Fig. 5(b)), the range of the scales grows faster than the values, i.e., the number of bars that appear with increasing row number grows faster than linear. This results in the shown curve when connecting the last bar of each row. In the case of exponential growth (Fig. 5(c)), values and scale ranges exhibit the same growth behavior, resulting in a straight line when connecting the last bar of each row. It has to be noted that these chart properties necessitate the use of scale ranges that grow exponentially.

5.2. Advantages

There are several advantages of our approach. First of all, our approach works well with static images. The readability over the full value range does not require interaction like zooming or the display of tooltips. Furthermore, every row is self-contained. They can be viewed and interpreted independently. In contrast, extracting a subarea of a classic bar chart can lead to misleading visualizations, e.g., if the axis does not start at 0 anymore. This property allows us to easily zoom, highlight, or filter parts of scale-stack bar charts. As a consequence, groups of values at certain magnitudes over very large ranges can be accurately displayed. It is also straightforward to use different units for the scales. In the case of temporal quantities, e.g., it is more intuitive to represent values with hours, days, and years etc., than using a single unit (e.g., seconds) for all scales.

Scale-stack bar charts distribute the available accuracy, which depends on the resolution of the resulting image, over the full value range. The full accuracy of linear bar charts is, in contrast, only available for values of the highest order of magnitude. For example, let us assume an available resolution of $1,000$ pixels and 4 orders of magnitude of the data. In the case of a linear bar chart, values of the highest order of magnitude can be represented with 3 digits accuracy, but values of the lowest order cannot be represented anymore. With our approach, all values can be represented with an accuracy of more than 2 digits (250 pixels per order of magnitude). Hence, we lose some accuracy for high values but can represent small values with the same accuracy. Logarithmic bar charts distribute the available accuracy also over the full value range, but inside the different orders of magnitude, the distribution is non-linear.

5.3. Disadvantages

Our approach has also some drawbacks. First, it is clear that not all tasks benefit from it. Since scale-stack charts use a linear scale, they cannot perform better than linear charts if the values of interest are clearly visible in them. And tasks like finding extrema can be efficiently accomplished with logarithmic charts. Furthermore, people are not used to scale-stack bar charts. Hence, they may require some training to read and correctly interpret them. There also is a potential risk that the viewer compares bars of different rows. However, similar risks exist also for logarithmic charts; not everybody is used to logarithmic scales and quantitative comparison of the bars' heights often leads to misinterpretation. Finally, the resulting chart contains more visual elements increasing the visual load.

6. Applications

We applied our method to data from three different fields—chemistry, social sciences, and the financial sector—which all exhibit a large value range. First, the half life of different chemical isotopes covers a large value range. As we show, even isotopes of the same element can already cover a range that is difficult to show in a single linear plot. Second, the populations of different countries can differ substantially. To demonstrate the ability to visualize stacked bars, the populations are divided into three different age categories in this dataset. Finally, the profits of different companies can differ largely. This dataset additionally demonstrates the representation of negative values.

6.1. Half Life of Isotopes

The half life of isotopes is important, among others, to quantify radioactive pollution. The presented data (Fig. 6) is from the UNSCEAR report [Uni08] regarding the Chernobyl disaster in 1986. It contains estimates of the amount of the released isotopes and their half lives.

The linear bar chart (Fig. 6(a)) is not suitable for this data. Because of the large half lives of the plutonium isotopes—more than $350,000$ years for pu-242—most values are not visible in a chart of this size. With the logarithmic bar chart (Fig. 6(b)), all values can be seen and it is easy to find the isotopes with smallest (i-133) or largest (pu-242) half life. However, reading values can already be difficult without experience with logarithmic plots. For example, we can see that the value of pu-238 is close to 100 years, but a more precise estimation is difficult. In the scale-stack bar chart (Fig. 6(c)), we can see that the value is half way between 75 and 100 years, leading to an estimated value between 85 and 90 years (exact value is 87.7). Furthermore, there is a risk for the logarithmic chart that a visual comparison of values leads to wrong interpretations, e.g., one might get the impression that the three highest half lives (pu-240, pu-239, and pu-242) are quite close. In our approach, it is clearly visible that the three values are all in different orders of magnitude.

6.2. Age Distribution

The data for the age structure of different countries are obtained from the CIA World Factbook [CIA12] and visualized with stacked bars (Fig. 7). Again, the linear chart (Fig. 7(a)) allows us to analyze only a subset of the data. The logarithmic chart (Fig. 7(b)) can be used to look at the total population numbers. However, it is not very useful to analyze the distributions because the logarithmic scale distorts the visual proportions of the stacked bars. The chart provides the wrong impression that the largest age group in all countries is from 0 to 14 years. Furthermore, the respective value to a given bar depends on its position. For example, the third group (65 years and older) seems to have similar sizes for Canada, France, and Germany but in the scale-stack bar

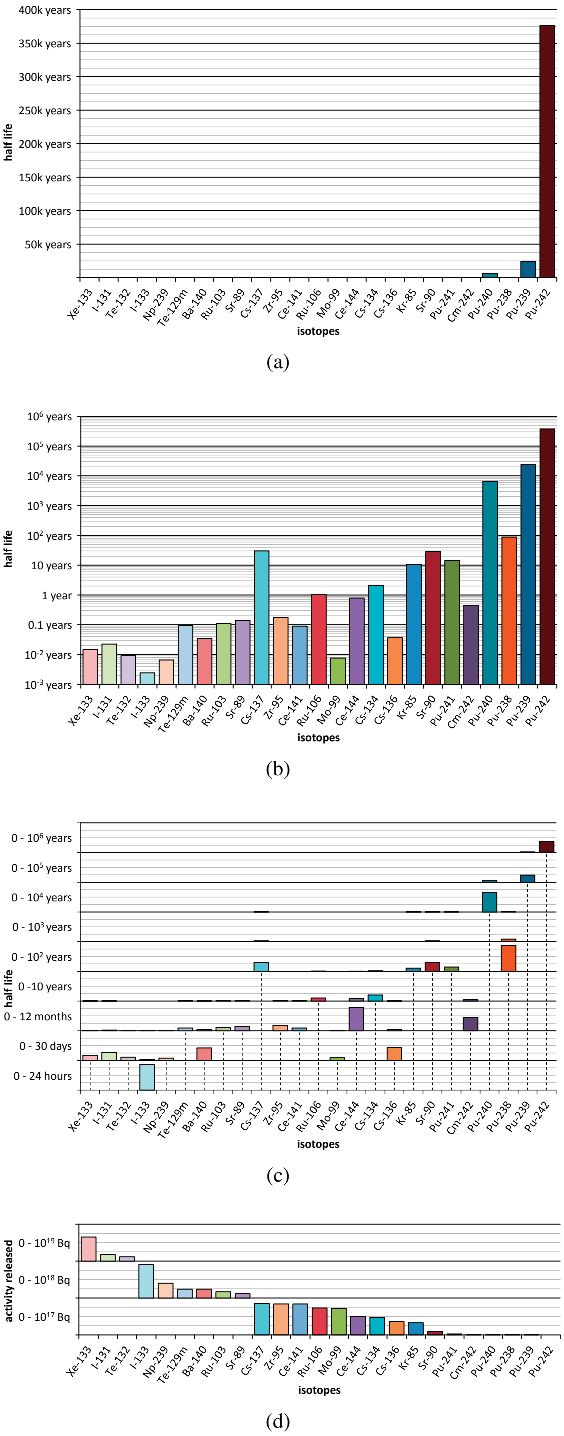


Figure 6: *The half life of different isotopes released to the environment during the Chernobyl disaster visualized with (a) linear, (b) logarithmic, and (c) scale-stack bar charts. In all charts, the isotopes are ordered on the x-axis according to the activity released during the accident. The respective activity values are shown in (d) with scale-stack bars. The bars are colored to be better distinguishable.*

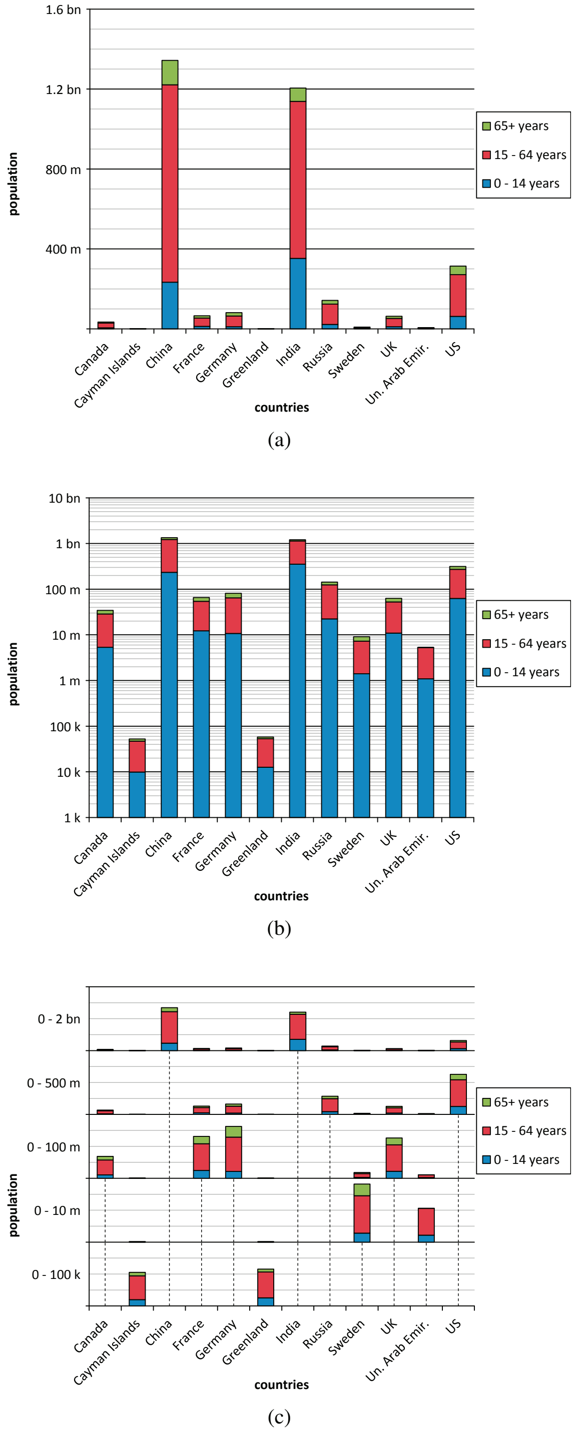


Figure 7: *Age structure of selected countries. The populations are divided into three age groups (see legend). Stacked bar charts with absolute values are used to show not only the distribution but also the absolute size of the groups. The data is visualized with (a) linear, (b) logarithmic, and (c) scale-stack bar charts.*

chart (Fig. 7(c)), we can see that these groups have different sizes. Additionally, we can see that the population of the United Arab Emirates consists only of a small part of old people. This is difficult to notice in the logarithmic plot.

6.3. Corporation Profits

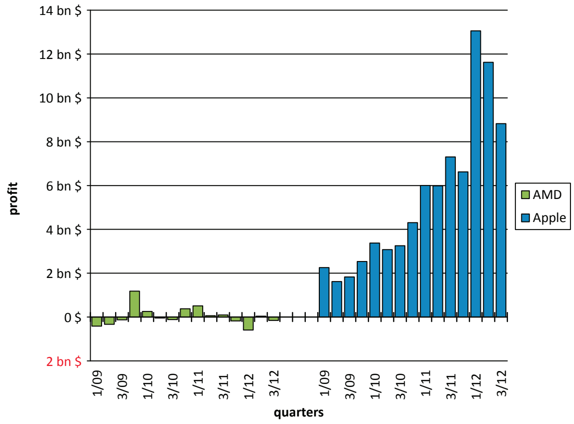
The high profits of Apple Inc. [App12] can make it difficult to present them together with results from other technology companies like, e.g., Advanced Micro Devices Inc. (AMD) [Adv12]. The data are grouped by corporation, and a time range of almost 4 years is shown (Fig. 8). The value range of this example is not as high as in the other two applications. Hence, most of the data can be analyzed with the linear bar chart (Fig. 8(a)). However, AMD has some comparably small values and it is hard to quantify or compare them. The general tendency can also be seen in the logarithmic plot (Fig. 8(b)) and all values are clearly visible, but the visual impression is less clear and makes a quantitative comparison more difficult. For example, the profit of AMD seems not to be that far away from Apple's profit. As the other two charts show, this is not the case. Another example is the decrease in AMD's profit from the last quarter of 2009 to the first quarter of 2010. It looks rather small in the logarithmic chart, but we can see in the scale-stack chart that it decreased to less than a fourth. Furthermore, the peak in Apple's profit in the first quarter of 2012 is not as clearly visible in the logarithmic chart as in the other charts. Our approach (Fig. 8(c)) allows us also to compare the smallest values in the chart, e.g., we can easily see that the profit of AMD almost doubles from the second to the third quarter of 2011. This is hard to deduce from the linear chart, whereas it is not possible to directly compare bars in the logarithmic chart.

7. Conclusions

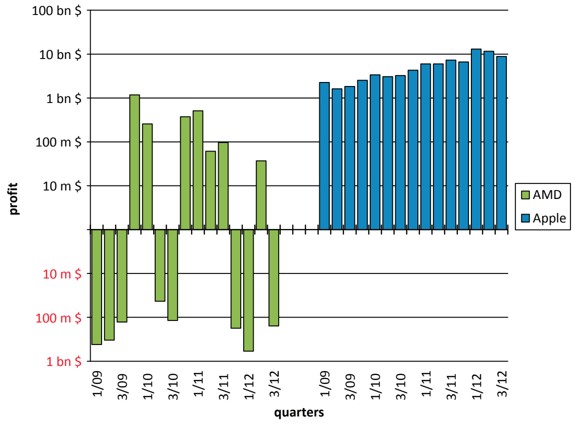
Our scale-stack bar charts combine the advantages of linear and logarithmic plots, while avoiding most of their drawbacks. They exhibit the advantages of linear plots, in particular, easy comparison of values. Similar to logarithmic plots, they allow us to display a large value range. Additionally, our technique can be directly used for typical variants of bar charts like stacked bar charts. The user study and our example applications show that our approach works especially well when quantitatively analyzing data with large value ranges. We show that there are many applications from different fields where this can be the case. Future work includes the application of the underlying concept of two-dimensional value representation to other chart types or concepts of information visualization.

Acknowledgments

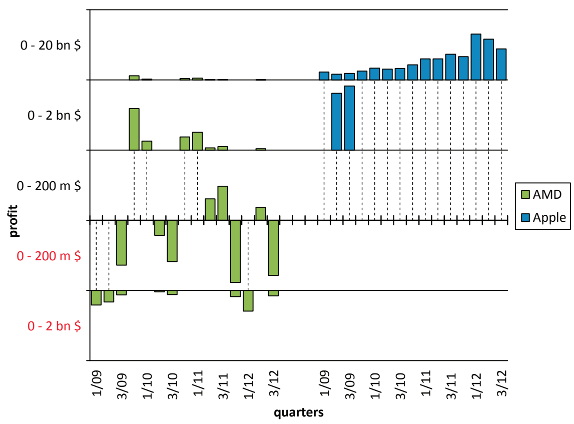
The authors would like to thank the German Research Foundation (DFG) for financial support of the project within the *Cluster of Excellence in Simulation Technology* (EXC 310/1) at the University of Stuttgart.



(a)



(b)



(c)

Figure 8: *Profits of AMD and Apple Inc. for the period from the first quarter of 2009 to the third of 2012, visualized with (a) linear, (b) logarithmic, and (c) scale-stack bar charts.*

References

[Adv12] ADVANCED MICRO DEVICES INC.: Quarterly Earnings. http://ir.amd.com/phoenix.zhtml?c=74093&p=quarterlyearnings, 2012. 9

[AJDL08] ACEVEDO D., JACKSON C. D., DRURY F., LAIDLAW D. H.: Using visual design experts in critique-based evaluation of 2D vector visualization methods. *IEEE Transactions on Visualization and Computer Graphics 14*, 4 (2008), 877–884. 4

[App12] APPLE INC.: Earning releases. http://investor.apple.com/results.cfm, 2012. 9

[BMG*10] BATEMAN S., MANDRYK R., GUTWIN C., GENEST A., MCDINE D., BROOKS C.: Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems* (2010), pp. 2573–2582. 2

[CIA12] CIA: The World Factbook. https://www.cia.gov/library/publications/the-world-factbook/index.html, 2012. 7

[CM86] CLEVELAND W. S., MCGILL R.: An experiment in graphical perception. *International Journal of Man-Machine Studies 25*, 5 (1986), 491–501. 1, 2

[CM87] CLEVELAND W., MCGILL R.: Graphical perception: The visual decoding of quantitative information on graphical displays of data. *Journal of the Royal Statistical Society. Series A (General) 150*, 3 (1987), 192–229. 2

[CM92] CARVALHO C. R., MCMILLAN M. D.: Graphic representation in managerial decision making: The effect of scale break on the dependent axis. Master's thesis, Air Force Inst. of Tech., Wright-Patterson AFB, 1992. 2

[CMS99] CARD S. K., MACKINLAY J. D., SHNEIDERMAN B. (Eds.): *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 1999. 2

[DBB10] DIEHL S., BECK F., BURCH M.: Uncovering strengths and weaknesses of radial visualizations—an empirical approach. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 935–942. 2

[DLR09] DRAPER G. M., LIVNAT Y., RIESENFELD R. F.: A survey of radial methods for information visualization. *IEEE Transactions on Visualization and Computer Graphics 15*, 5 (2009), 759–776. 2

[Few07] FEW S.: Save the pies for dessert. In *Perceptual Edge Visual Business Intelligence Newsletter* (2007), pp. 1–14. 2

[Few10] FEW S.: Our irresistible fascination with all things circular. In *Perceptual Edge Visual Business Intelligence Newsletter* (2010), pp. 1–9. 2

[GH11] GOLDBERG J. H., HELFMAN J.: Eye tracking for visualization evaluation: Reading values on linear versus radial graphs. *Information Visualization 10*, 3 (2011), 182–195. 2

[Huf93] HUFF D.: *How to Lie With Statistics*. W. W. Norton & Company, New York, NY, USA, 1993. 2

[Su08] SU Y.: It's easy to produce chartjunk using Microsoft Excel 2007 but hard to make good graphs. *Computational Statistics & Data Analysis 52*, 10 (2008), 4594–4601. 2

[TLH10] TALBOT J., LIN S., HANRAHAN P.: An extension of Wilkinson's algorithm for positioning tick labels on axes. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1036–1043. 2

[TM05] TORY M., MÖLLER T.: Evaluating visualizations: do expert reviews work? *IEEE Computer Graphics and Applications 25*, 5 (2005), 8–11. 4

[Tuf92] TUFTE E. R.: *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT, USA, 1992. 2

[Uni08] UNITED NATIONS SCIENTIFIC COMMITTEE ON THE EFFECTS OF ATOMIC RADIATION: UNSCEAR 2008 Report. Sources and effects of ionizing radiation. Volume 2. Annex D—Health effects due to radiation from the Chernobyl accident. http://www.unscear.org/docs/reports/2008/11-80076_Report_2008_Annex_D.pdf, 2008. 7

[War04] WARE C.: *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers, San Francisco, CA, USA, 2004. 2

[WH11] WICKHAM H., HOFMANN H.: Product plots. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (2011), 2223–2230. 2