CHAPTER 13

Designing Effective Visualizations

The goal of this chapter is to provide some guidelines for designing successful visualizations. A successful visualization is one that efficiently and accurately conveys the desired information to the targeted audience, while bearing in mind the task or purpose of the visualization (exploration, confirmation, presentation). For any particular set of data there is a myriad of possible methods for mapping data components to graphical entities and attributes. Similarly, there exists a wide range of interactive tools that the user may be provided. Selecting the most effective combinations of techniques is by no means a straightforward process.

A visualization may be ineffective for a number of reasons. It might be too confusing or complex to be interpreted by the intended audience, or some of the data may have been distorted, occluded, or lost during the mapping process. Other signs of deficient visualizations are the lack of support for view modification or color map control. Even aesthetics can influence the success of a visualization; a visually unappealing presentation can affect an audience's willingness to look at the images. In each of the above cases, some component of the visualization is interfering with the delivery of information to the user.

This chapter first presents design considerations for the components that the authors feel are necessary for a good visualization. Following this, we explore some of the common problems found in visualizations and propose some techniques for avoiding these problems. We summarize by revisiting some of the issues presented in Chapter 3 and indicate how they fit into the visualization design process. At a recent visualization conference, it was stated that it is much easier to make bad visualizations than good ones. Hopefully, through reading this chapter, visualization designers will gain some of the skills necessary to make design decisions leading to effective visualizations.

The ideas and techniques presented in this chapter come from not only the authors' experiences, but also from the vast body of literature on designing good visualizations. Readers are encouraged to study one or more of the books on this subject, as listed in the Related Readings section.

13.1 Steps in Designing Visualizations

Creating a visualization involves deciding how to map the data fields to graphical attributes, selecting and implementing methods for modifying views, and choosing how much data to visualize. Additional information regarding the data being shown (e.g., labels) and the mapping (e.g., a color key) are also essential to facilitate interpretation, and must be integrated into the visualization. The final, less tangible, consideration is the overall aesthetics of the resulting display. In this section we present, for each of these design stages, some issues that should be addressed by the visualization designer.

13.1.1 Intuitive Mappings from Data to Visualization

To create the most effective visualization for a particular application, it is critical to consider the semantics of the data and the context of the typical user. By selecting data-to-graphics mappings that cater to the user's domain-specific mental model, the interpretation of the resulting image will be greatly facilitated. In addition, the more consistent the designer is in predicting the user's expectations, the less chance there will be for misinterpretation. Intuitive mappings also lead to more rapid interpretation, as translation time is reduced. For example, in Figure 13.1, images of planets are used to plot the relationship between the distance from the planet to the sun and the duration of its orbit.

Mapping spatial data attributes, such as longitude and latitude, to screen position is perhaps the most common and intuitive mapping found in visualizations. Some of the earliest visualizations took advantage of the ability of humans to correlate position on the drawing medium with position in the three-dimensional world. Likewise, with the advent of animation, it is obvious that displaying temporally related data sets via animation is reasonably intuitive, with the added advantage of allowing time to vary in both speed and direction.

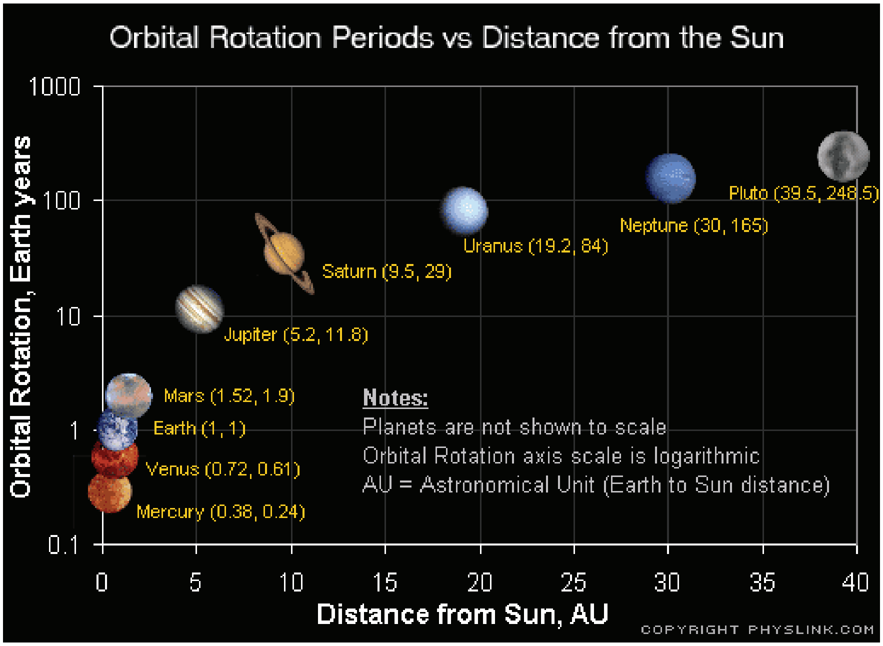


Figure 13.1. Using intuitive scatterplot symbols to show the distance from planets to the sun versus the duration of a single orbit. (Image from http://www.physlink.com.)

Other mappings become intuitive when associated with a particular context. For example, mapping temperature to color is fairly common, as many cultures associate red or white with high temperatures. Color has specific interpretations in fields such as cartography (land use classification) and geology (stratigraphic layer classification), and thus the application domain for the visualization may dictate the logical use for the color attribute.

Height, or alternatively the length of a line, is another useful mapping for temperature, as we associate temperature with the readout on thermometers. In fact, for medical practitioners, it may be intuitive to use length for displaying pressure or any other scalar value (e.g., the patient readouts in the *Star Trek* sick bay).

One of the important considerations when selecting a mapping is the compatibility between the scale of the data field and that of the graphical entity or attribute. For ordered data attributes (e.g., age), it is not reasonable to select a graphical attribute that is not ordered, such as shape. Similarly, unordered data attributes (e.g., country of origin) should not be mapped to ordered attributes (e.g., length).

With that said, it is, however, sometimes interesting to examine data with nonintuitive mappings, as the resulting image may expose an interesting attribute in the data. For example, mapping time to color along a streakline can reveal variations in particle speeds that might otherwise be difficult to detect. Thus a good rule of thumb is to set the default mappings based on the most intuitive selection according to the typical user, but, especially for exploratory tasks, to permit user customization.

13.1.2 Selecting and Modifying Views

Except for fairly simple data sets, one view is rarely sufficient to convey all the information contained in the data. The key to developing an effective visualization is to be able to anticipate the types of views and view modifications that will be of most use to the typical user, and then provide intuitive controls for setting and customizing the views. Useful views, as mentioned earlier, depend heavily on the type of data being presented, and the task associated with the visualization. Each supported view should be clearly labeled, and selecting a new view should require minimal actions on the user's part.

View modifications fall into a number of categories, and their inclusion as part of the functionality should be considered based on user priorities.

- Scrolling and zooming operations are needed if the entire data set cannot be presented at the resolution desired by the user.
- Color map control is almost always desirable, minimally supporting a set of different palettes, and preferably offering the user control of either individual colors or the complete palette.
- Mapping control allows users to switch between different ways of visualizing the same data. Features of the data that are hidden in one mapping may stand out in others (Figure 13.2).
- Scale control permits the user to modify the range and distribution of values for a particular data field prior to its mapping. Similarly, clipping and other forms of filtering allow the user to focus on data subsets.
- Level-of-detail controls provide the ability to eliminate or highlight detail, supporting views at different levels of abstraction. Depending on the task at hand, a user may need to repeatedly switch between several distinct levels (Figure 13.3).

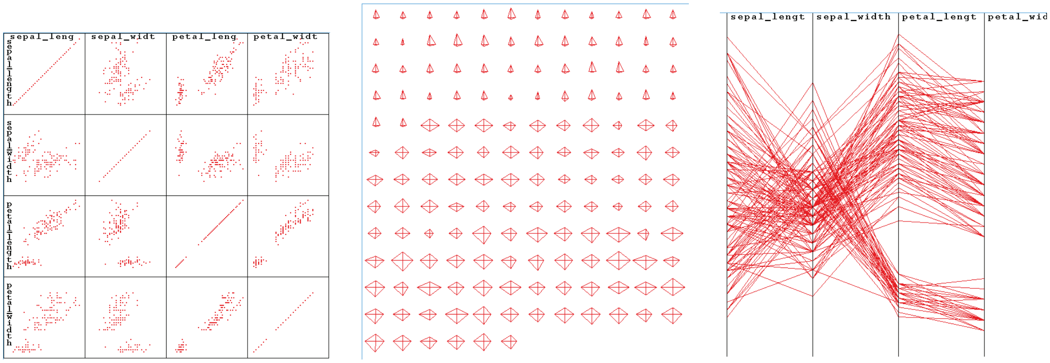


Figure 13.2. Three views of the Iris data set (scatterplot matrix, star glyphs, and parallel coordinates). (Image from XmdvTool.)

In all cases, it is essential that the view manipulations be implemented in a manner that is easy for the user to remember, and that provides suitable accuracy for the task. If possible, direct manipulation (specifying changes on the image itself rather than a separate control or command line) is generally preferred. For example, mouse motion could be mapped to panning, with button clicks invoking zoom operations (See Chapters 11 and 12).

13.1.3 Information Density—When Is It Too Much or Too Little?

One of the key decisions one makes when designing a visualization is determining how much information to display. This gives rise to two extreme situations. The first, which might be called "gratuitous graphics," occurs when there is very little information to present. Many examples of graphics

Figure 13.3. Levels of detail in maps. (Images courtesy of Google Maps © 2008 Google; map data © 2008 NAVTEQ$^{\text{TM}}$.)

can be found that convey only two or three distinct values, such as the percentage of males and females within a particular sample (this actually can be communicated with one number). Others can be found that "pad" the number of pieces of information by deriving additional quantities, such as showing two numbers, their sum, and their difference. In cases such as these, it is often more effective to simply display the quantitative values as text. This requires much less screen real estate (which in many applications is quite valuable), while still getting the message across. It must be remembered that simply because one *can* create a visualization doesn't imply that one *must* do so.

The other extreme, trying to convey too much information, is also a common problem. Excessive information content can lead to confusion, intimidation, and difficulties in interpretation on the part of the viewer. Important information contained within the data can be lost or deemphasized on a cluttered display, and viewers may have a hard time determining where to focus their attention.

There are many effective solutions to the problem of excessive information content in a visualization. One method is to give the user the option of disabling or enabling different components of the display. In this manner, a user can decide which parts are most important to her, and can have the less important information displayed on demand. Another solution is to use multiple screens, either as disjoint panes or with partial occlusion. This method makes better use of screen space, while making each of the individual pieces of data readily available.

Another common cause of cluttered displays is large or unevenly distributed data sets. As mentioned in the previous section, data sets may be filtered to remove uninteresting data points, allowing the user to concentrate only on the significant parts. Similarly, uneven distributions, which might lead to some parts of the screen being congested, while others are sparsely populated, can sometimes be rectified through scaling of one or more data dimensions.

13.1.4 Keys, Labels, and Legends

A common problem with many visualizations is that insufficient information is provided to the user to allow unambiguous and accurate interpretation. This supporting information should begin with a detailed caption indicating the particular data fields being displayed, and the mappings that were used. Additionally, grid or tick marks should be displayed to convey the ranges and

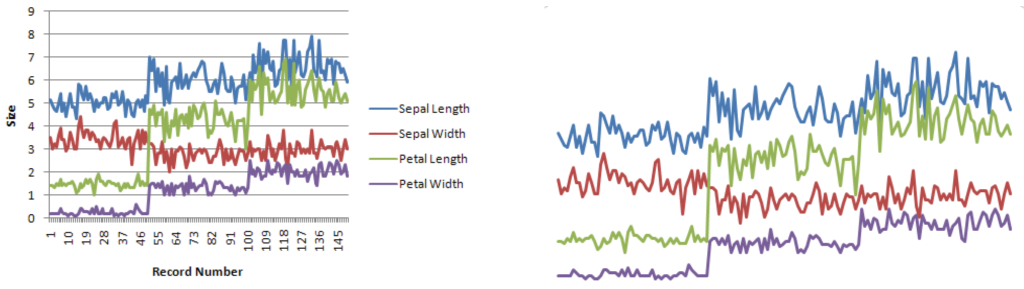


Figure 13.4. A complex visualization with and without captions/ticks/legends.

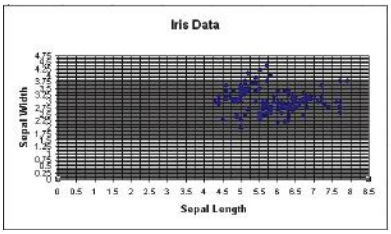
values of interest for numeric fields when absolute judgments are important, and all axes should be labeled with appropriate units. If symbols are being used, a key must be provided, either along the border of the display or within a separate widget. Finally, if color has a significance, sufficient information must be available to allow easy interpretation (e.g., via a labeled color bar). Figure 13.4 highlights the importance of this supplementary information.

The use of grid and tick marks can be both a boon and a curse to the visualization. Poor choices of the types of markings and the density used can occlude the data being displayed and lead to a cluttered appearance. Figure 13.5 shows three degrees of markings. Clearly, one should avoid the extremes.

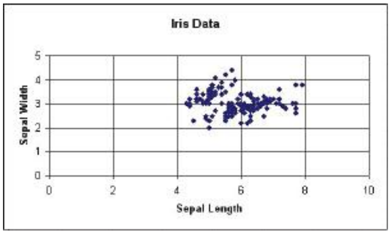
The actual positions of the markings can also have a bearing on how readily the data is interpreted. Based on the semantics of the data, certain gaps between markings may make more sense to the user than others. Unfortunately, the default values used by some visualization tools may make correct interpretation difficult (Figure 13.6).

The designer must also decide which range of values is to be displayed (this decision may have been made in an earlier stage). There is always the risk of misinterpretation when the expected range of values is not shown. For example, when dealing with a percentage, most users would expect the display to range from 0 to 100. However, in many cases this would lead to significant waste of display space and loss of perceptual resolution (e.g., if all percentages were below 10 percent), as can be seen in Figure 13.7. Thus, the range must be carefully chosen and clearly marked to help convey accurate information.

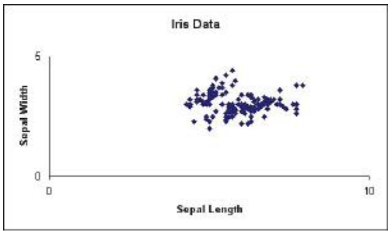
One final rule of thumb pertains to the use of multiple frames or windows. It is important to follow a consistent labeling and gridding scheme. Changing the position of labels and keys or the range of values shown (for the same



(a)

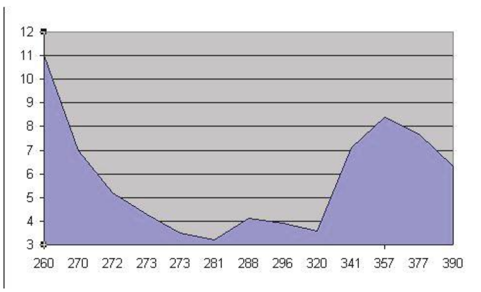


(b)

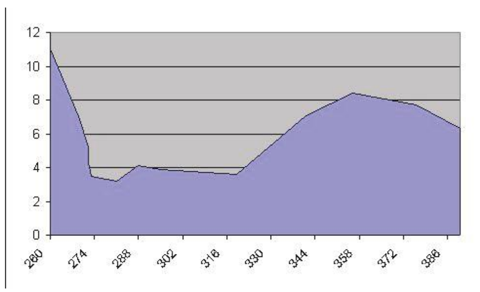


(c)

Figure 13.5. Varying degrees of tick marks: (a) excessive, (b) moderate, and (c) minimal.



(a)



(b)

Figure 13.6. Grid spacings: (a) illogical; (b) logical.

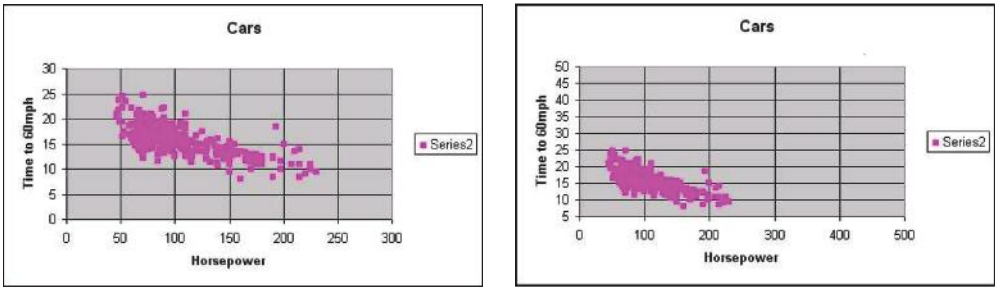


Figure 13.7. Logical and illogical data ranges.

field) can cause confusion and increase the risk of misinterpretation. If range changes are necessary (e.g., for views that differ in level of detail), the label, as well as the grid markings, should convey the change. Similarly, if different color mappings are necessary, the visualizations must clearly convey this information.

13.1.5 Using Color with Care

One of the most frequently misused parameters in visualization design is that of color. Selecting the wrong color map or attempting to convey too much quantitative information through color can lead to ineffective or misleading visualizations. Also, since color perception is context-dependent (a particular color will appear quite different, depending on adjacent colors), the characteristics of the data itself can influence how the colors are perceived. Finally, it must be remembered that many people are color blind or color confused; it has been determined that as many as ten percent of all males have some form of color deficiency. The following guidelines can assist in the effective use of color in visualization.

1. If the visualization task involves absolute judgment, keep the number of distinct numeric levels low (see Figure 13.8 and Chapter 3 on perception).

2. Use redundant mappings if possible, e.g., map a particular field to both color and size (see Figure 13.9), to improve the chances of the data being communicated accurately.

3. In creating a color map for conveying numeric information, make sure that both hue and lightness are changed for each entry (see Figure 13.10).

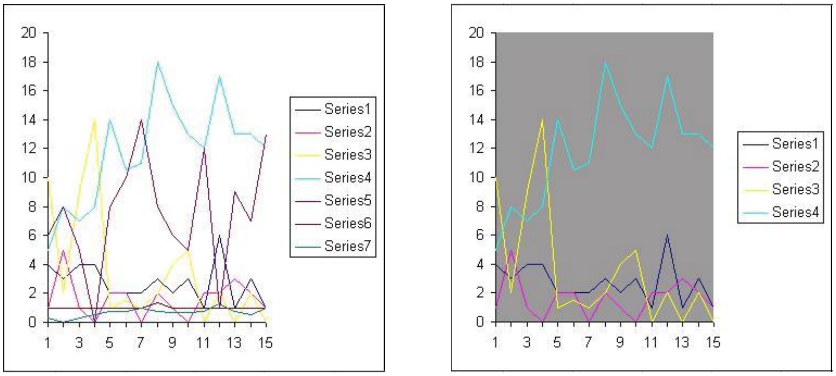


Figure 13.8. Too many colors versus a moderate number of colors.

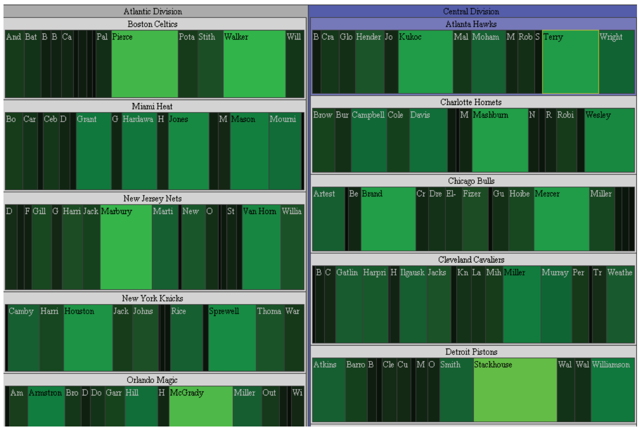
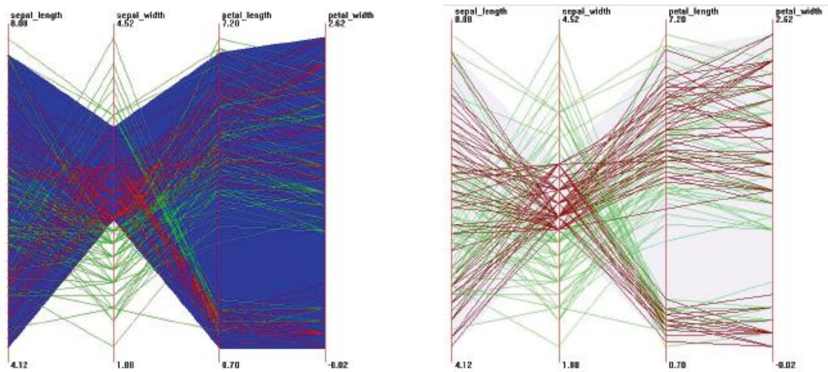


Figure 13.9. Treemap of basketball statistics, with points per game redundantly mapped to color and size. (Figure generated using Treemap 4.0, from the University of Maryland.)

4. Include a labeled color key to help users interpret the colors (see the previous section).

5. When possible, use semantically resonant colors in the visualization [273]; these will be easier for users to learn and remember.

Color can add significant visual appeal to a visualization, but can also significantly decrease the effectiveness of the communication process. Some



(a) (b)

Figure 13.10. (a) Changing hue; (b) changing both hue and saturation.

interface designers advocate an initial design process that only involves the use of grayscales. Once this design has been refined and tested, the addition of color can usually be done in a more effective manner.

13.1.6 The Importance of Aesthetics

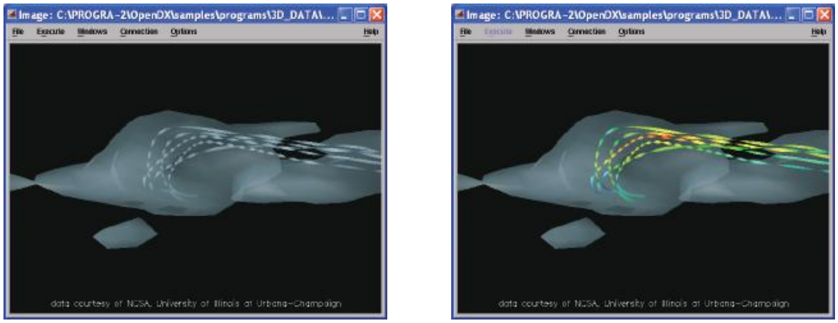
Once we have ensured that our designed visualization conveys the desired information to the user (function), the final step is to assess the aesthetics (form) of the results. The best visualizations are both informative and pleasing to the eye. In contrast, a visualization might be so visually unappealing that it detracts from the communication process. An aesthetically pleasing visualization invites the viewer to study it in depth.

There are many guidelines for attractive visualization design that can be drawn from the art and graphic design communities. These include:

Focus. The viewer's focus should be drawn toward the part of the visualization that is most important. If the important components are not sufficiently emphasized, viewers don't have sufficient cues for guiding their inspection (see Figure 13.11).

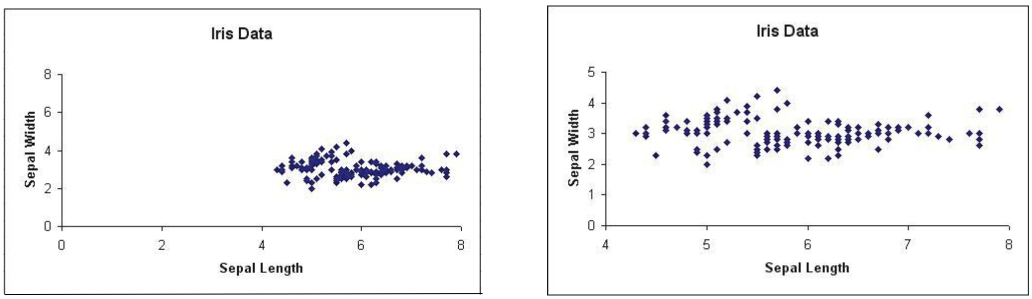
Balance. The screen space should be used effectively, with the most important components in the center. Emphasis should not be given to any particular border (Figure 13.12).

Simplicity. Don't try to cram too much information in one display (see Section 13.1.3), and don't use graphics gimmicks simply because they are available (e.g., using 3D Phong shaded histograms when a bar or



(a) (b)

Figure 13.11. (a) Subdued streamlines vs. (b) highlighted streamlines from OpenDX.



(a) (b)

Figure 13.12. (a) Everything to one side vs. (b) balanced between left and right.

line chart could convey the same information). A useful procedure to follow once a visualization has been designed is to iteratively remove features and measure the loss of information being conveyed. Features whose removal results in minimal loss can probably be discarded (see Figure 13.13).

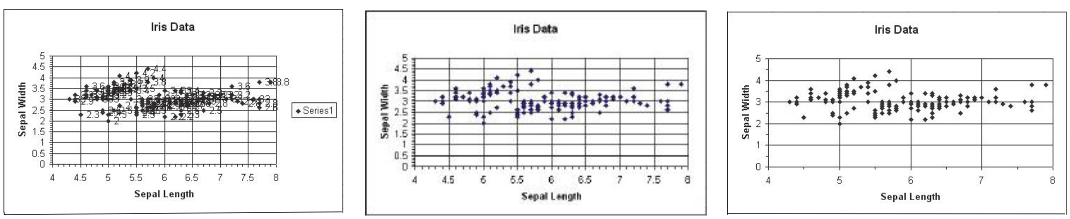


Figure 13.13. Progression from a cluttered chart to a simplified chart.

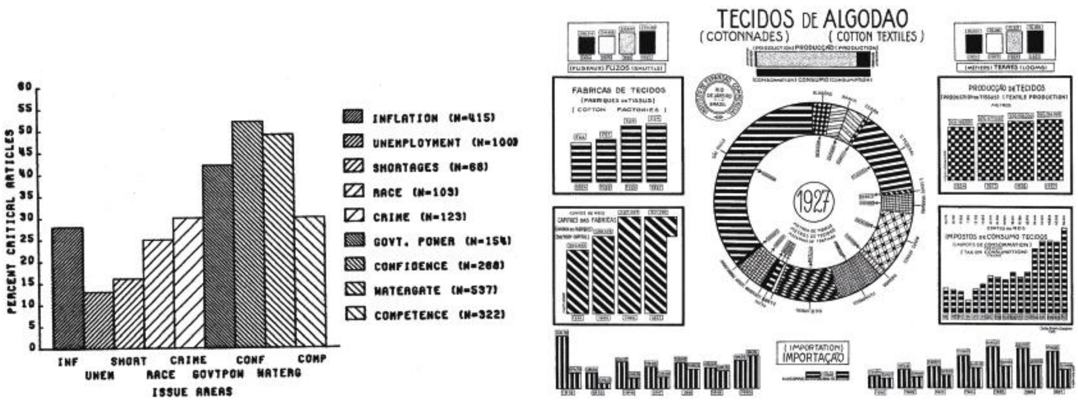


Figure 13.14. Some ugly visualizations: (a) from Miller et al. [294] and (b) from a Brazilian economic statistics report [200]. In both cases, the cross-hatchings are seemingly random and distracting. Shades of gray would significantly improve the aesthetics of the images.

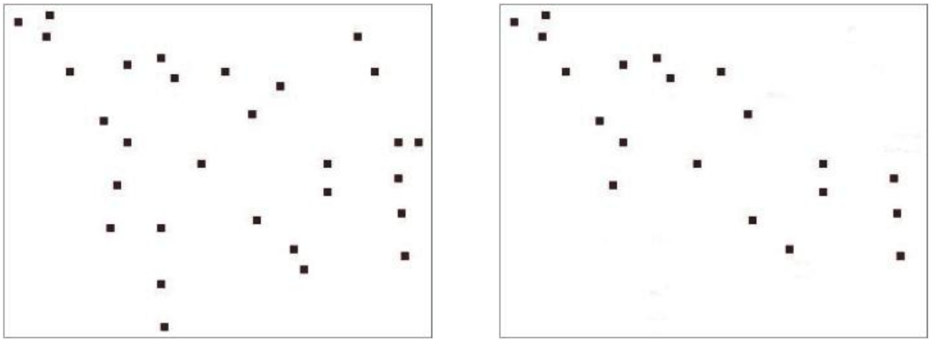
There are many examples of ugly visualizations in the literature. We reproduce a few of these below (see Figure 13.14). We encourage designers to perform aesthetic assessment on their results prior to presenting them to users, and to seek out and incorporate the extensive literature available on graphics design.

13.2 Problems in Designing Effective Visualizations

In the following sections we examine some of the common problems found in visualizations that can occur even if the steps outlined above are followed. These problems have a deeper root, and relate to decisions regarding what to visualize and what is the most appropriate method to use. Some of the problems involve intentional or inadvertent data distortion, which can lead to misinterpretation. Others involve hiding the real data behind "cleaned" versions or excessive supporting graphics. In all cases, steps can be taken to improve the quality and "honesty" of the visualization.

13.2.1 Misleading Visualizations

One of the foremost rules of visualization is that the image should be an accurate depiction of the data. However, throughout history, there are ex-



(a) (b)

Figure 13.15. The problem with data scrubbing: (a) raw data showing lack of correlation; (b) scrubbed data revealing false correlation.

amples of how visualizations from distorted data have been used to sway opinions and lie to the audience. These so-called "viz lies" can be found everywhere, from the most prestigious journals to company portfolios. In this section, we identify some of the common strategies for creating misleading visualizations, not for the reader to practice, but to try to avoid!

Data scrubbing. Raw data can often be very rough in form, and the temptation when creating a visualization is to remove some of the roughness. Unfortunately, sometimes the selection of which data to remove is biased to eliminate data that does not support a particular point that the author of the data is espousing (see Figure 13.15). Outlier removal is a common tactic in this situation. Unless there is reason to believe that the outliers resulted from flaws in the data acquisition process, they should not be removed without informing the viewer and providing the option for the outliers to be displayed.

Unbalanced scaling. Scaling is a powerful tool in visualization, since careful selection of scale factors can reveal patterns and structures not visible in unscaled views. However, scaling can be used to deceive the viewer into believing that a trend is stronger or weaker than supported by the data. This can lead to what Tufte refers to as the *lie factor* [424], which is the ratio between the raw data change and the change as depicted in the visualization. For example, in Figure 13.16 the size of objects in the background is reduced in width and height by perspective, thus distorting comparison with foreground objects.

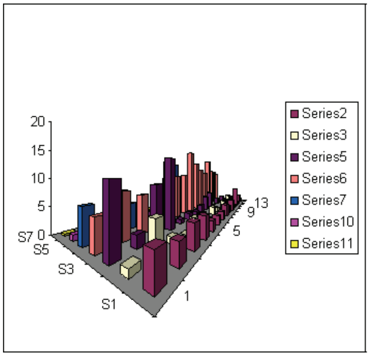


Figure 13.16. Vis Lies: perspective distorts size in favor of closer objects.

Range distortion. As mentioned in an earlier section, viewers often have an expectation about the ranges for a particular data dimension; by setting this range to be significantly different from this expectation, the user may be deceived into misinterpretation. This is often done by moving an axis so it no longer corresponds with the expected "zero value" (see Figure 13.17). Since relative judgment is such a strong component of our perceptual system, changing the baseline for the relations being portrayed could have a serious effect on how the image is interpreted. The designer may want to give the user the option of moving this baseline to avoid wasting screen space, but it should be made clear what the baseline is, especially if it departs from the established norm.

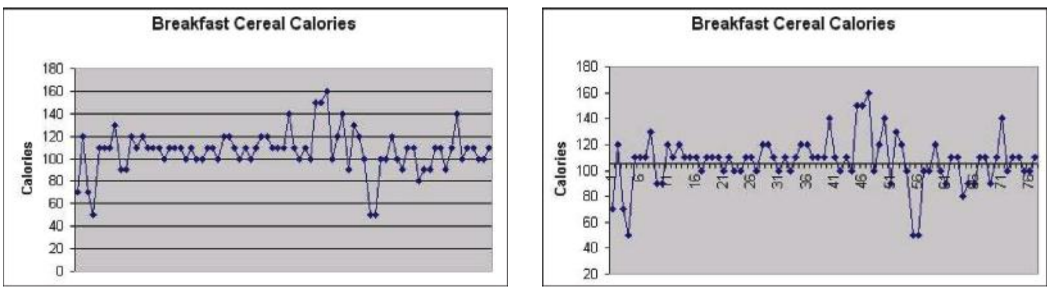


Figure 13.17. Plotting data with different baselines.

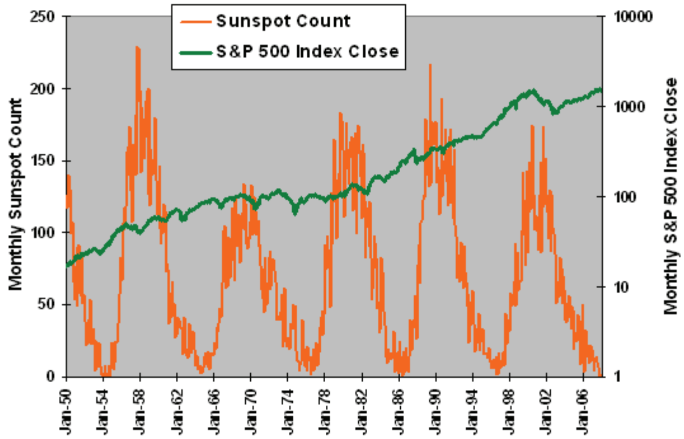


Figure 13.18. A nonsense plot, showing sunspot occurrence versus the S+P 500 Index. (Image from http://www.cxoadvisory.com/blog/internal/blog4-07-09/.)

Abusing dimensionality. In Chapter 3 we noted that errors in interpretation rise with the power of the dimensionality being portrayed. Thus, our errors in judging volume are much worse than those for area, which in turn are worse than those for length. Therefore, mapping a scalar value to a graphical attribute such as volume can dramatically increase the likelihood of erroneous interpretation. As mentioned earlier, it is often the case that simpler is better.

13.2.2 Visual Nonsense—Comparing Apples and Oranges

Visualizations are designed to convey information, and it is important that the information be meaningful. Visualizations are often created by combining data sets from different sources. However, it is easy to combine unrelated components into a single visualization and identify what seems to be structure; for example, plotting stock market values against occurrences of sunspots (see Figure 13.18). In this case, coincidental relationships can be confused with causal relationships. In deciding what data to combine, it is important to first ensure that there is some logic in the combination. One of the problems found in analytic pattern recognition/data mining processes is that these irrelevant relationships are often discovered and reported, which must then be eliminated by a domain specialist. The visualization designer should attempt to avoid creating nonsense graphics before they are presented to users.

Another factor that must be considered is compatibility between temporal and spatial ranges for data being compared. Thus, for example, one (probably) shouldn't compare the sales of a particular product in one year for a particular region of the country with the sales of the same product for a different region and year, unless one is hypothesizing that a migration in interest for the product is occurring.

Compatibility in units also needs to be examined in creating a data set for visualization. For example, food products that are measured in terms of price per volume are often mixed with those measured in price per weight. An effective visualization of this data might normalize them both to price per serving.

Finally, there is often a temptation to perform operations suitable for ordered or continuous data on categorical, unordered data, simply because the mapping process resulted in an ordered graphical representation. An example might be an attempt to fit a line or curve to a sequence of data points that map a company name to a position on the screen. Obviously, this has no semantic meaning, but because the mapping converts the scale of the data, users might feel that it is useful to perform the fitting.

The key point is that some thought must be put into the semantics of the visualization to insure that it makes logical sense.

13.2.3 Losing Data in the Chart Junk

In a previous section we stressed the importance of including labeled grid or tick marks on visualizations that require quantitative assessment. The excessive use of such markings is an example of what Tufte referred to as *chart junk* [424]. Chart junk can be defined as any supplementary (nondata) graphics in a visualization that are not necessary for the accurate interpretation of the data. This additional information can lead not only to visualizations that appear overly complex, but also to occlusion and deemphasis of the actual data.

Deciding the amount of supplementary graphics to put in a visualization is sometimes a difficult process, since the designer might not know the needs of all potential users. However, because we are dealing with a dynamic, customizable medium (unlike Tufte's static charts), the option exists to allow users to adjust the types and density of this supporting information on the display. In some visualization tasks, users can switch between qualitative overviews and quantitative analysis. In the former case, it is usually more important to give viewers a clear view of the data, while in the latter case,

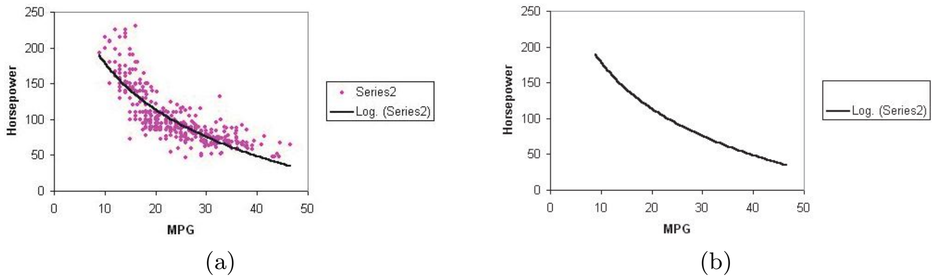


Figure 13.19. (a) Raw data plot with fitted curve; (b) only fitted curve.

tools to help quantify the elements of the display are much more desirable. Thus, a good rule of thumb is to provide sufficient tools to support the user's quantitative needs, but with the option of disabling them or altering their degree of presence in the visualization.

13.2.4 Raw versus Derived Data

A common practice is to compute an analytic model of the data using curve/-surface fitting to obtain a more visually appealing result. Again, this is distorting the truth, and it may lead to false assumptions and conclusions on the part of the observer. In some visualizations, it is common practice to throw out all of the raw data and only show the smooth approximation derived from that data. This forces the viewer to trust that the approximation is an accurate portrayal of the data, which is often not the case when the designer blindly applies statistical fitting algorithms. It is best to show both the raw data and the fitted model first, and to allow one or the other to be deemphasized or filtered out on demand (see Figure 13.19).

Yet another form of cleaning the data is the process of *resampling*, where raw data positioned either on a sparse grid or randomly are used to create data that are either denser or on a regularly spaced grid. This can result in a much richer visualization, approaching that of continuous sampling, but it again deceives the user into believing the data set is much larger than it actually is. The denser the resampling, the more likely that the user will misinterpret the data, unless the phenomenon being observed has little variability. For example, Figure 13.20 shows the locations of global temperature monitoring stations. Clearly, there are large voids where no stations exist, so resampling could result in many wrong conclusions, such as that the entire northern part of South America would be interpolated by

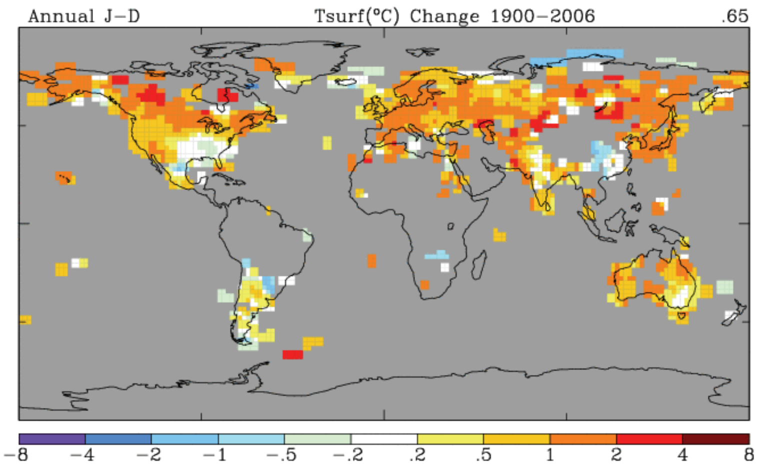


Figure 13.20. Sparse global temperature change data would give erroneous values for most of the planet if interpolated. (Image courtesy [150].)

the readings from four or five stations, with the conclusion being that the region has dropped in temperature over the past century.

Insufficient sampling is another problem. As the images in Figure 13.21 show, a sampling that doesn't look at the data characteristics can miss many important features. The left image is sampled and interpolated uniformly, while the right image uses contour information to add sample points where significant changes occur.

It is critical that the user always have access to the raw data and be informed of any scrubbing/smoothing/resampling operation that has been

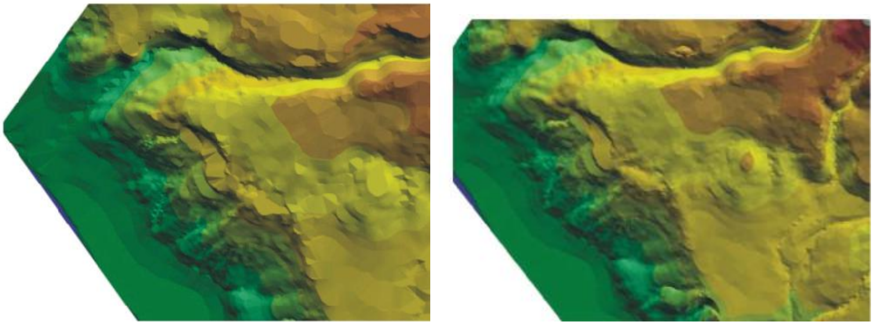


Figure 13.21. Different sampling and interpolation of the same data set. Some of the details in the right image are not seen in the left image. (Image from [185].)

Figure 13.22. Some examples of absolute versus relative judgment. (Image courtesy of Michael Barry.)

applied. In some domains, such as radiology, analysts are adamantly opposed to any sort of data smoothing or filtering, as there is danger that an important signal in the data might be discarded as noise. Thus, views should be provided that show the raw data set prior to deriving new versions, allowing the user to decide whether the derivation is an accurate representation of the original data.

13.2.5 Absolute versus Relative Judgment

As mentioned in Chapter 3, humans have a fairly limited ability to make absolute judgments of visual stimuli. This implies that visualizations that depend too heavily on users performing accurate measurements of graphical attributes such as position, length, and color will result in problems in interpretation. One means of combating this human limitation is to design visualizations that either rely on relative rather than absolute judgment, or that are restricted to only using a small number of distinct values for each graphical attribute being used to convey information.

Bounding boxes, grids, and tick marks are all excellent tools for converting an absolute judgment task to one that depends more on relative judgment. By comparing the length or position of a graphical entity against a quantified structure, users can more rapidly determine the approximate value relative to the known levels. Using residuals (e.g., subtracting values from their means) can also change a measurement task to one of deciding whether a value is above or below a particular level (see Figure 13.22).

13.3 Summary

In this chapter we have presented a number of design rules for creating effective visualizations. These include:

- use data-graphic mappings that are likely to be intuitive to the targeted audience;
- provide users with multiple views of their data, along with easy-to-use tools for modifying views;
- avoid putting too much information in a given display; rather, provide users with the ability to turn components of the visualization off and on;
- include keys, labels, legends, and grids/ticks to help users interpret the visualization;
- use color with care; color perception is highly context sensitive, and humans are limited as to the number of distinct colors that can be identified with accuracy;
- design your visualizations to be attractive, as well as functional;
- avoid misleading users with unbalanced scales and other visualization lies;
- verify that the visualization has semantic meaning and compatible units;
- use grids in such a way that the data is not overly occluded; too much chart junk can misdirect the user's attention;
- always provide users access to the raw data; it is usually OK to perform some data scrubbing, but the user should be aware of how the resulting data has been derived;
- design visualizations that rely on relative, rather than absolute, judgment, when possible.

None of these rules are hard and fast; there are exceptions to each, and indeed, there are times when one rule conflicts with another. Designers should be prepared to try many alternatives before deciding on a final form, assessing each based on the criteria presented here. However, be strongly advised that there is no substitute for rigorous usability studies with subjects drawn from the anticipated audience (see Chapter 14). Only after this testing has been performed can the designer be reasonably assured that an effective visualization has been created.

13.4 Related Readings

Many books have been written on the design of informative graphics. Notable authors include Edward Tufte [424–427], Stephen Few [123, 124], and Stephen Kosslyn [251, 252]. Their work provided much of the information and concepts in this chapter. A significant percentage of modern HCI textbooks focus on designing effective graphical user interfaces. One early work dedicated to this topic is the book by Mullet and Sano [300], which provides a number of useful rules of thumb.

A number of articles have been written on the subject, with many lists of guidelines, principles, and design patterns, including [3, 58, 69, 112, 370].

13.5 Exercises

1. Identify at least three problems with the visualization shown in Figure 13.23.

2. For each of the visualizations in Figure 13.14, suggest at least three modifications that would improve their effectiveness.

3. Describe four examples of how some of the rules of this chapter may conflict with each other.

4. Assume that you are plotting the exchange rates for 20 different countries. List at least three ways of ordering the names of the countries and describe why each might be useful.

5. Other than the figures used in the exercises, find at least three examples of figures in this book that could be improved using design guidelines described in this chapter. Send suggestions for improvements to the authors (yeah, we can take the criticism!).

6. Examine several visualizations found in different media (newspaper, magazine, company prospectus) and critique the designs. Rank them from best to worst.

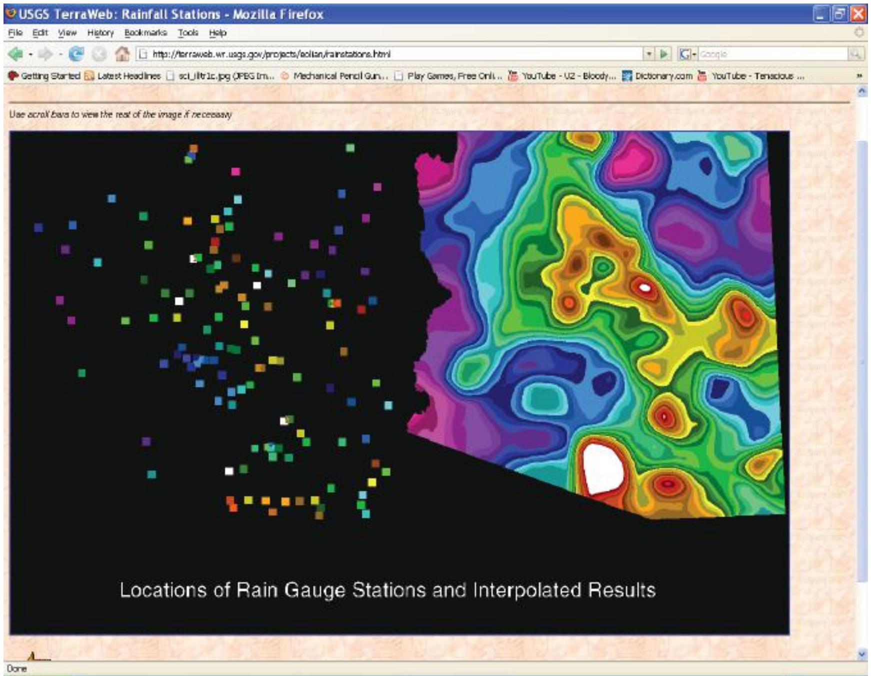


Figure 13.23. Rainfall data for Arizona from the USGS.

13.6 Projects

1. Choose three visualization programs that you've written for this course. For each, try to find at least three ways of improving them, based on the design guidelines of this chapter. Reimplement these programs with the improvements you've identified. Compete with your classmates to see who can create the most attractive, informative visualization.

2. Choose three visualization programs that you've written for this course (they can be the same three as used in the project above). For each, try to find at least three ways of making them *worse* by violating design guidelines of this chapter. Reimplement these programs with the *negative* improvements you've identified. Compete with your classmates to see who can create the ugliest, least informative visualization!

3. Write a scatterplot program that reports the percentage of data that is at least partially occluded. Show how changing the size of the marks or adding random jitter to locations for data points affects this occlusion percent.

4. Write a program for generating a line chart from a time series data set. Experiment with drawing grid lines with varying density, width, brightness, and color. Try variations on the attributes of the plotted line to find good settings (i.e., those that help the user focus on the plotted points).