DVT

UNIT -5

Syllabus:

Research Directions in visualizations: Steps in designing Visualizations – Problems in designing effective Visualizations – Isues of Data, Issues of Cognition,perception and Reasoning, Issuesof system design, Evaluation,Hardware and Applications

1.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 at tributes. 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.

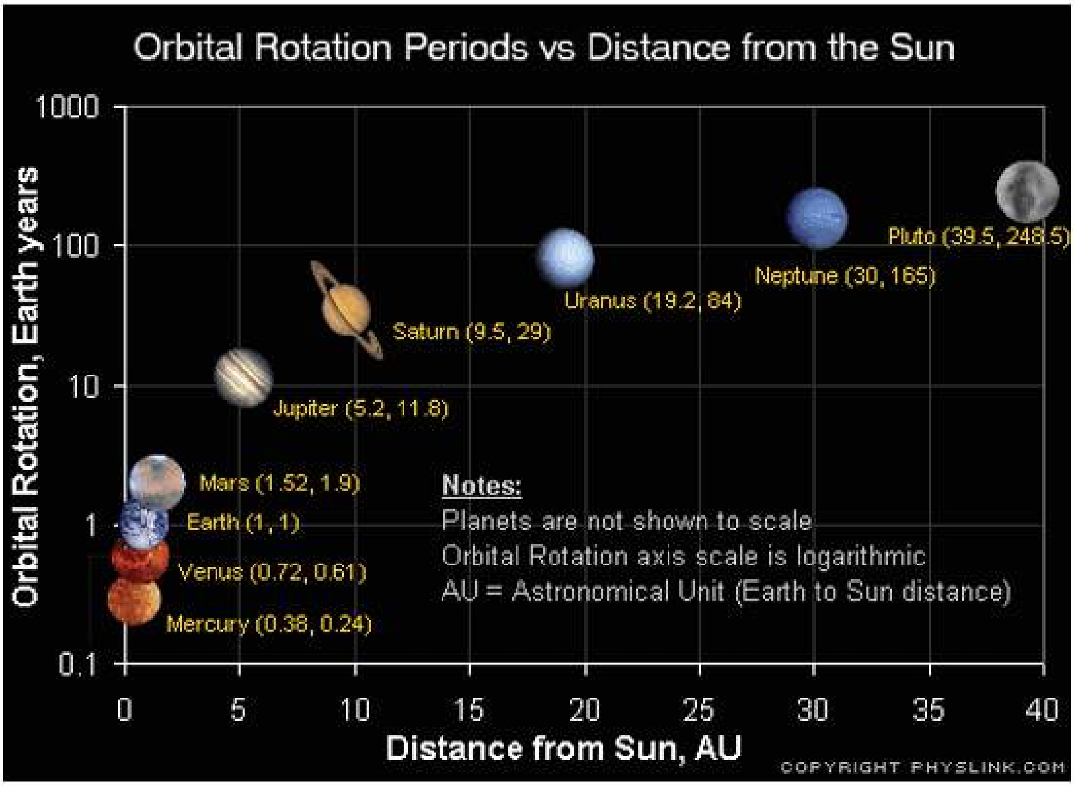
1.1 Steps in Designing Visualizations

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 visualiza tion designer.

(a) 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 typ ical 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 , 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 visu alizations. 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 obvi ous 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,



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.)

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 rea sonable 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).

(b)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 la beled, 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 vi sualizing the same data. Features of the data that are hidden in one mapping may stand out in others.
• 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 13.2. Three views of the Iris data set (scatterplot matrix, star glyphs, and parallel coordinates). (Image from XmdvTool.)

(c)Information Density—When Is It Too Much or Too Little?

One of the key decisions one makes when designing a visualization is de termining 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 can be found that convey only two or three distinct values, such as the per centage 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, in timidation, 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 dis tributed data sets. As mentioned in the previous section, data sets may be f iltered 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.

(d) 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 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).

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. 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.

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 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.

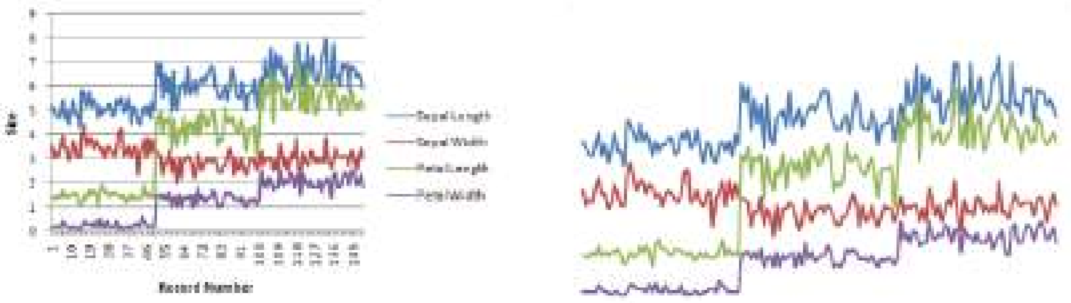


Figure 13.4 A complex visualization with and without captions/ticks/legends.

(e) 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

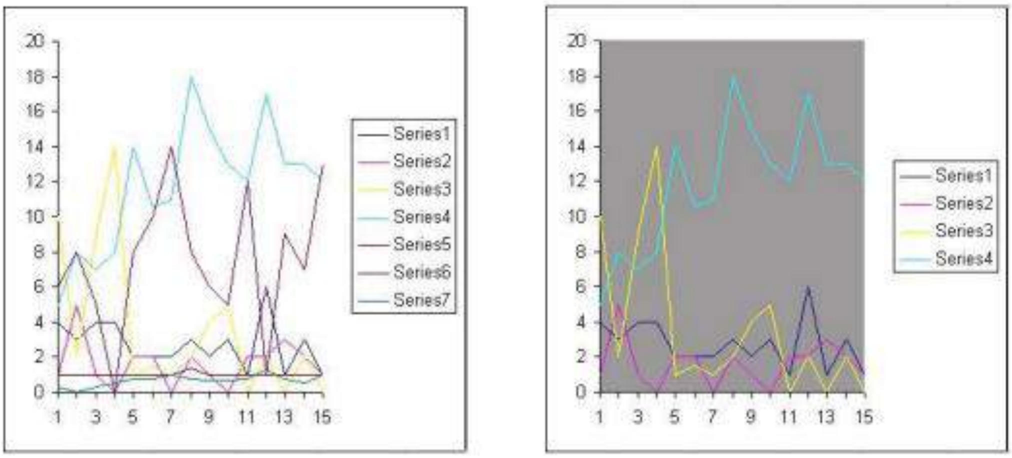
2. Use redundant mappings if possible, e.g., map a particular field to both color and size , to improve the chances of the data being communicated accurately.

3. In creating a color map for conveying numeric information, make surethat both hue and lightness are changed for each entry

4. Include a labeled color key to help users interpret the colors .

5. When possible, use semantically resonant colors in the visualization , 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 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.

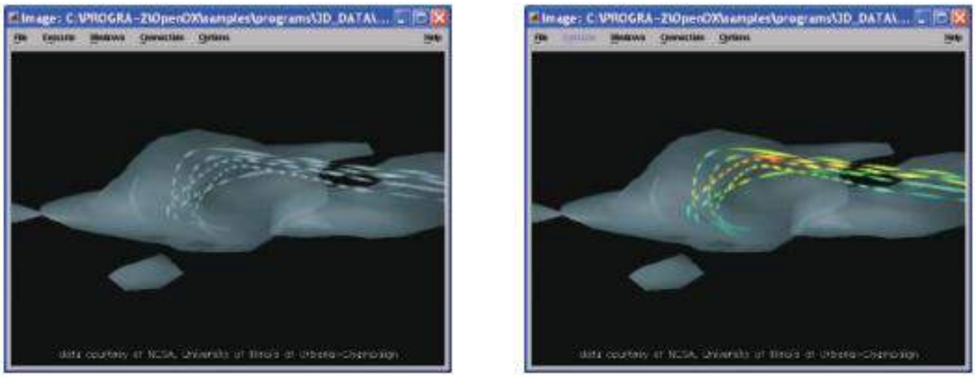


Too many colors versus a moderate number of colors.

(f) The Importance of Aesthetics

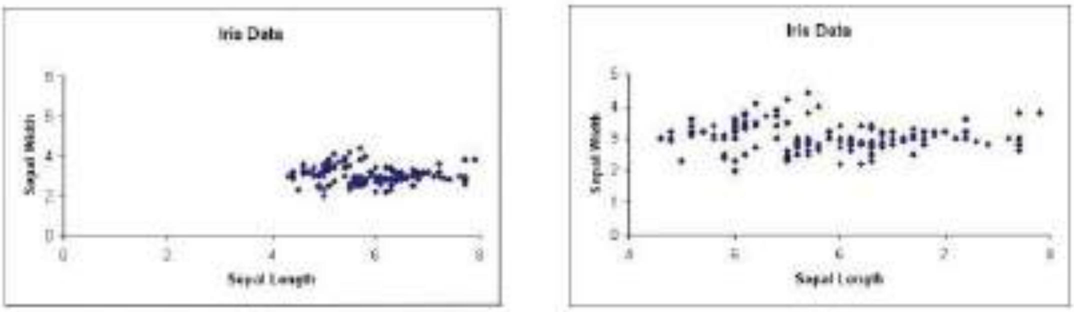
Once we have ensured that our designed visualization conveys the desired information to the user (function), the final step is to assess the aesthetics (form) of the results. The best visualizations are both informative and pleasing to the eye. In contrast, a visualization might be so visually unappealing 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 .
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 .
Simplicity. Don't try to cram too much information in one display and don't use graphics gimmicks simply because they are available (e.g., using 3D Phong shaded histograms when a bar or 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



(a) (b)

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



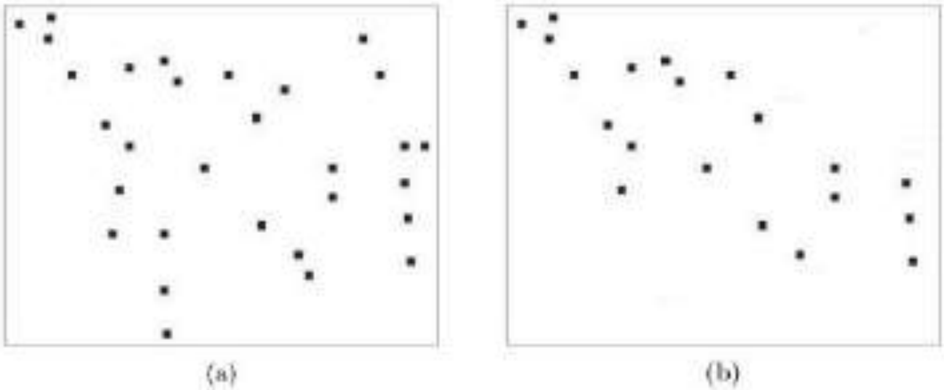
(a) (b)

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

1.2 Problems in Designing Effective Visualizations:

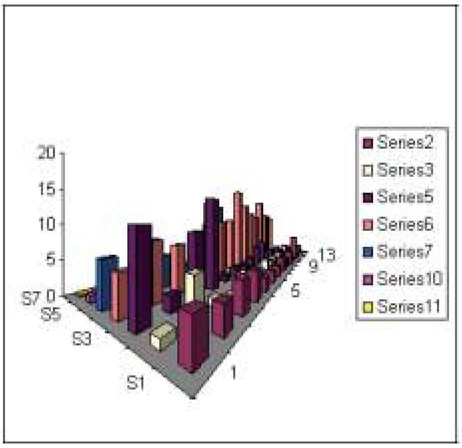
(1) Misleading Visualizations: One of the foremost rules of visualization is that the image should be anaccurate depiction of the data. However, throughout history, there are examples of how visualizations from distorted data have been used to swayopinions and lie to the audience. These so-called “viz lies” can be found everywhere,the most prestigious journals to company portfolios. In this, we identify some of the common strategies for creating misleadingnvisualizations, 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



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

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.



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” . 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.

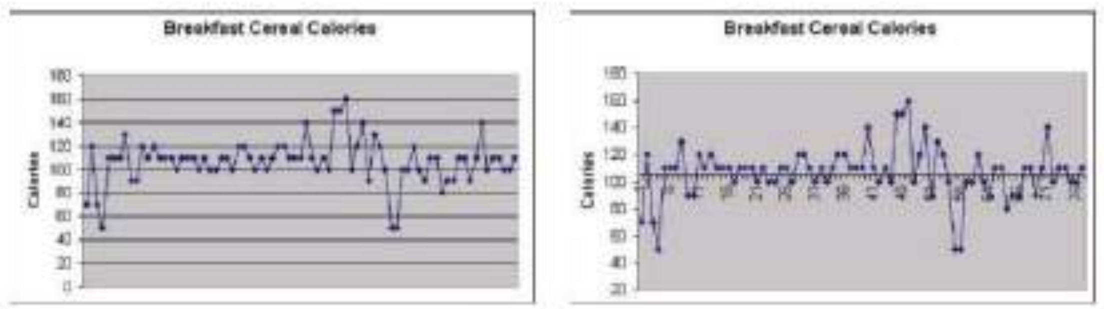


Figure 13.17. Plotting data with different baselines.

Abusing dimensionality. 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.

2. Visual Nonsense:

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 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.

3. Losing Data in the Chart Junk:

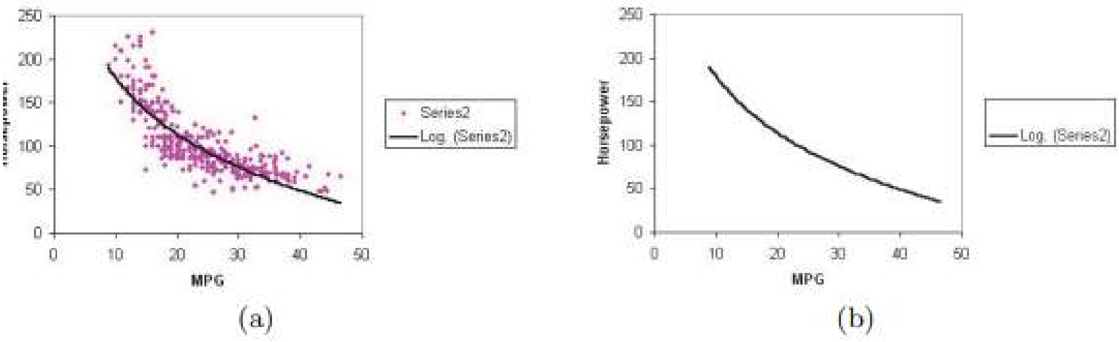
Chart junk can be defined as any supplementary (nondata) graphics in a visualization that are not necessary for the accurate interpretation of the data. This additional information can lead not only to visualizations that appear overly complex, but also to occlusion and deemphasis of the actual data.
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

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.

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

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.



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

5. Absolute versus Relative Judgment

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.

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

2. Research Directions in Visualization

Visualization is a sufficiently mature field that many groups of researchers are dedicating time to contemplate and predict the future directions for it.

1, Issues of Data

Many of the current and proposed future research activities are centered on expanding the characteristics of data that can be effectively visualized. Some of these are discussed here.

Scale.

Perhaps the most frequently addressed problem dealing with data is finding solutions for coping with ever-increasing sizes of data sets. For scientific domains, both acquired and simulated data have been growing in orders of magnitude, certainly outpacing Moore's Law. While only a few years ago researchers were focused on data sets of modest size (kilo bytes to a few megabytes), it is now common to be looking at problems involving gigabytes and terabytes of data . Solutions such as cluster ing, sampling, and abstraction are all being actively pursued in areas such as graph visualization , genetic sequence analysis, and large-scale flow simulations . Massive data, that which does not fit into memory, is now commonplace.

Static versus dynamic.

While most visualization techniques to date have been developed with the assumption that data is static (e.g., in files or databases), a growing interest is in the area of visual analysis of dynamic data . An increasing number of streaming data sources are being studied in the database and data mining communities, and efforts to perform visual anal ysis on this type of data are starting to emerge. The basic concept is that the data is continually arriving, and has to be analyzed in real time, both because of the urgency with which analysis must be performed, as well as the fact that the volume of data precludes its permanent storage. Analysts are interested not only in the values at a particular time period, but also in how the data is changing. Since the data is so large in many cases, the data cannot even be stored as it is streaming.

Spatial versus nonspatial data.

A growing number of application areas for visu alization include both spatial and nonspatial data, including many scientific and engineering fields. To provide analysts with a powerful environment for studying this data, several recent efforts have focused on the integration of the spatial visualization techniques normally found in scientific visualization with the nonspatial techniques that are common in information visualization. The linkages between these distinct types of views is critical to successful analysis .

Nominal versus ordinal.

The graphical attributes to which we map data in our visualizations, such as position, size, and color, are primarily quantitative in nature, while it is quite common to have data that is not quantitative, such as the name of a gene or the address of an employee. If this nominal data is to be used in the visualization, a mapping is needed. However, it is also important to ensure that relationships derived from visual analysis are truly part of the data, and not an artifact of the mapping. Some solutions include selecting color schemes and plot symbols that do not impose a perceptual or dering; others attempt to assign a numeric mapping of the nominal variables that preserves similarities and differences implied by the data . For ex ample, in a data set containing statistics about cars, two cars with similar characteristics in their attributes might have similar numbers assigned to the name field.

Structured versus nonstructured.

Data can be classified based on the degree to which it follows a predictable structure. For example, tables of numbers would be considered highly structured, while newspaper articles may be re garded as unstructured. In between, we can have semi-structured data, such as an e-mail message that contains both a structured component (sender, time, receiver) and an unstructured part (message body). Key research problems include the extraction of useful information from the unstructured components of data sets and using the structured components as methods to index and organize the data records. Text analysis and visualization have made great progress, but the research is still in its infancy.

Time.

Time is a special variable (and attribute of data). Time in dynamic data provides one view: a volume visualization over time deals with a phys ical representation, and a common interactive visualization uses time as a control. Spatio-temporal databases and queries and the visualization of re sults are becoming prominent as more data is being made publicly available. Time here can be handled just as in the volume visualization. But one can do more. For example, in a data set involving education and health indica tor data, a useful question is how to identify similar patterns involving not just the data, but time as well. This begs a key research question of how to handle time as just another variable, and what are the ramifications.

Variable quality.

While most visualization systems and techniques assume the data is complete and reliable, in fact most sources of data do not match these constraints. Often, there are missing fields in the data, often due to acquisition problems (e.g., defective sensors, incomplete forms). The quality of the data itself may also be problematic; out-of-date information can have low certainty associated with it, inaccurate sensors may produce values with significant variability, and manual entry of data can be error-prone. There is a great need for visualization technology that can cope with this data quality issue, this uncertainty, not only in its visualization, but also in keeping the analyst informed, so that the quality information can be incorporated into the decision-making process.

2. Issues of Cognition, Perception, and Reasoning

Many of the foundational concepts in data and information visualization have their roots in our understanding of human perception, particularly in aspects of selecting effective mappings of data to graphical attributes such as color and size. Much less work, however, has built on our knowledge of human cognition and reasoning processes, with the notable exception being the Gestalt laws to better understand how visual grouping is performed. Thus, a far-reaching research agenda for the visualization field must include the study of how humans do problem solving with support from interactive visual displays, and how we can leverage this knowledge to design even more effective and powerful visual tools.

Part of this effort will be directed at raising the level at which visual tasks are performed. For example, rather than using the visualization to identify a cluster or trend, it might instead be focused on building a mental model of the entire phenomenon being studied. While this task may be comprised of several more primitive subtasks, the visualization tool will need to support the combination and sequencing of these tasks to enable the analyst to cope with the scale and complexity of the data and information being gathered and generated. Tasks such as discovering associated patterns of change in the data will involve not only visualization of the data, but also how the data is changing and how those changes may be associated with other changes taking place. These higher-level discoveries can then be used by the analyst to form, confirm, or refute hypotheses, expand or correct mental models, and provide confidence in decision-making processes.

Beyond decision making, we can also envision the expansion of visualization in the process of human learning. Educational opportunities for data and information visualization abound, but it is critical in the design of tools to support education that the theories of how humans learn be components of the design. People vary in the ways in which they learn most effectively: concrete versus abstract, sequential versus random. Different visual tools and mechanisms are likely to be needed to address these very different styles of learning.

In both problem solving and learning activities, we can also imagine using visualizations as a mechanism to expand and support the memory process, which is critical to both activities. Providing visual representations of intermediate results or supporting evidence, as well as structuring these artifacts on the screen, can help people remember facts and relationships in the same way that writing notes to oneself can relieve the human memory system of having to perform accurate recall. The extents to which visualization can be applied to this have yet to be extensively studied and exploited. Finally, and perhaps most importantly, we would like to measure cognitive primitives in the transfer from the data to the screen (easy) and from the screen to the human (hard).

3.Issues of System Design

One of the most crucial research challenges in developing visualization tools is determining how best to integrate computational analysis with interactive visual analysis. While many visualization systems support a modest number of computational tools, such as clustering, statistical modeling, and dimension reduction, and similarly many computational analysis systems support some amount of visualization, such as visualizing the analysis results, there have been no systems developed to date that provide a truly seamless integration of visual and computational techniques. Clearly, there are many tasks for which human perception and cognition are the most effective means to a solution; however, there are likely comparable numbers of tasks for which computational solutions are far superior. Indeed, in many situations, one can easily imagine using computational methods to guide the visual exploration, and vice versa. For example, one can use visual overviews of data to decide on appropriate parameters for filtering and clustering algorithms, after which the analyst could examine a resulting cluster to help select a computational model that best fits the characteristics of the data subset. Both visual and computational methods could then be used to ascertain the goodness of fit for the model. Many such scenarios can be easily envisioned. Another key problem.

Another key problem in visualization system design is the development of powerful new interaction paradigms to support the user's tasks. Many researchers believe that existing modes of interaction during visual analysis are ill-suited to the tasks being performed. This is likely due to the fact that while advances in hardware and visualization techniques have been moving by leaps and bounds, interaction methods have expanded at a much slower pace. While exciting and novel interaction tools can be seen in the immersive environment field, these and other types of interactions are either too lowlevel or are not designed with a particular high-level task in mind. What is needed is a fresh look at how users can interact with their data and information to perform specific tasks, such as determining the underlying phenomena or processes that are generating the data; understanding causal relationships; and designing, executing, and assessing "what if" scenarios. While existing low-level interactions such as point-and-click and drag-anddrop may be parts of the solution, there likely is much more needed to make visual analysis more efficient and effective,

Another issue is that most visualization systems require an expert user. Visualization in the last few years has made its appearance in daily news papers and television, and interactive visualization is common on the web. These are minimally interactive and often very simple visualizations. Can we develop ones that are engaging and easy to use? This is becoming critical as we are encountering the era of the democratization of data.

Finally, we are still developing visualizations and visualization systems based on experience, pragmatic research, and heuristics. We do not yet have a science of visualization. There have been attempts at automating the process: given data, automatically generate a visualization. It's clear that tasks need to be incorporated into the process. It's just as clear that evaluation and metrics can help solve this complex problem.

4.Issues of Evaluation:

In the early days of visualization research, rigorous evaluation was rarely performed; the assumption was that some visualization was better than no visualization for many tasks, and that if a new technique were developed, it was sufficient to simply place a couple of sets of images side by side and do a qualitative judgment of the results. More recently, there have been a large number of concerted efforts to incorporate a more formal evaluation process into visualization research, not only to enable quantification of improvements as they occur, but also to validate that visualization has measurable benefits to analysis and decision making . While many strategies have been developed and tested, there are many avenues of research toward improving the overall process.

The key point to make is that evaluation is a necessary component in moving a field from ad hoc methods to a real science, and we can learn from and build upon the evaluation strategies used in many other fields, such as human-computer interaction and sociology, to help validate the usefulness and effectiveness of techniques.

5. Issues of Hardware

Whenever computer technology advances, the applications that employ this technology must be reassessed to see how the advances can be leveraged. For visualization, there are several technologies that can and will have an impact.

a)Hand-held displays.

Most people these days carry with them at least one form of digital display, whether it be mobile phones, PDAs, portable games, or tablets. While most visualization systems have been designed for desktop (or larger) displays, there are still significant opportunities to deliver interactive representations of information and data on these smaller devices. Examples of potential applications abound. For maiaircraft, ships, and even buildings, having detailed presentations of wiring and plumbing diagrams, sensor output, and access paths can greatly simplify a technician's tasks. For crisis management during an emergency, police, firefighters, medical personnel, and other key players need interactive real-time access to information presented in a clear, unambiguous fashion. For those who monitor border crossings, rapid access to risk assessments, cargo manifests, and travel histories can help prevent entry by unwelcome individuals and material. The key is to develop visual solutions that make effective use of the limited display space and interactivity options.

b)Display walls.

At the other extreme, large-scale displays, often involving multiple panels stretching 10–30 feet in each direction, are becoming more and more common, not only for control centers, but also for investigating large data and information spaces. While directly porting desktop solutions to these large displays is generally straightforward, it is not necessarily the most effective use of the display space. A better solution would be to redesign the visual analysis environment to arrange the displays of different types and different views of information in a way that supports the analysis. In this way, high-resolution displays can always be visible, rather than being covered in a typical desktop solution, thus requiring viewers to just move their head or shift their focus to see different views.

c)Immersive environments.

Virtual and augmented reality systems have been frequently used within the visualization field. Early efforts, such as the virtual wind tunnel [52], enabled active, user-driven studies of the output of computational fluid dynamics experiments. Virtual walk-throughs and fly-throughs have been used in a diversity of fields, including architecture, medicine, and aeronautics. A key problem with this technology is the need to render the visualizations with minimal latency, which has spawned significant research in algorithm optimization. While the "killer application" for virtual environments has yet to be discovered, it will undoubtedly require significant visualization technology. Google glasses and similar devices providing an augmentation to the user's world are creating a number of opportunities for improved interactions and visualizations in dealing with the real world and the projected one.

d)Graphical processing units.

The development of special-purpose graphics hard ware has actually exceeded the growth in performance of general purpose CPUs, primarily driven by the computer game industry. There is much ac tive research in the visualization community that focuses on harnessing this computational horsepower [249,399]. Due to the architecture of a typical GPU, existing algorithms designed for CPUs do not, in general, port di rectly to the GPU, but require a nearly total redesign. However, as more and more software and hardware engineers become versed in this program ming paradigm, it is likely we will see a growing use of this technology, not only for graphics, but also for complex algorithms in general.

e)Interaction devices.

Each new device for user interaction with the computer opens up a wide range of possibilities for use in visualization. Voice/sound input and output have been extensively studied, though they are rarely an integral component of a visualization system. Tactile feedback is another area of great promise. Already, there have been efforts to train surgeons via virtual surgery, using force-feedback devices to mimic the feel of using probes and scalpels in surgical procedures [90]. Another avenue for development is to examine how different controllers used in modern game consoles could be employed to support visualization. The popular Wii input wand could be used for specifying actions via gestures. Other devices, such as head and eye trackers, may have significant potential in the visualization field. Inter active tables improve on the interactions with touch screens that have been used successfully for kiosks, and may lead to larger, less tiring environments. Brain control is another area providing great opportunities for controlling and interacting with visualizations. There are many opportunities for re search on interaction metaphors.

6. Issues of Applications

Many advances in the field of visualization have been driven by the needs of a particular application domain. These advances are then often generalized to be applicable to many other domains and problems. Thus, research issues can be divided between those specific to a given domain and those involving the extension of techniques into other domains.

a) **Depth-based innovations.** To develop truly useful visualization tools for a par ticular domain, it is very important for the visualization designer to under stand the domain and the tasks the domain specialist is trying to perform. Unfortunately, it is often the case that visualization designers just focus on the syntax or format of the data or information to be analyzed, which can easily result in tools that are either not useful/understandable to the do main analyst, or that provide little advantage over existing techniques. In a similar process, when domain experts attempt to design and develop visualization tools for their needs, they often produce tools with poor designs and user interfaces; this is mainly because they are generally unaware of the perceptual and cognitive issues involved in visualization design, as well as be ing unfamiliar with the range of different visualization techniques that have been developed for other purposes. In contrast, many of the most successful visualization systems have resulted from tight interactions between domain and visualization specialists, focusing on real problems that affect progress in a domain.

b) **Breadth-based innovations.**

   Another direction of research is to broaden the number of applications in which data and information visualization can be applied. Indeed, it is hard to imagine an area in which visualization would not be applicable, as all areas of society are experiencing a glut of information, while at the same time, display devices have become ubiquitous. In many applications, visual information presentation is rapidly replacing much of the textual communication, such as weather reports, stock market behav ior, health statistics, and so on. Daily schedules are often best captured in a graphical presentation of an hourly or daily calendar. Graphs are used to capture complex social networks, organizational charts, process flows, and communication patterns.

   One key to successful research and development in the broadening of the visualization field is to find the appropriate vi sual representations and interaction metaphors for new domains and tasks. Another critical problem is the conversion of data and information into a format that is amenable to existing visualization techniques. While many tools now accept input from standard database tables and spreadsheet files, much data and information is still stored either in proprietary formats or as unstructured files. Concerted efforts are needed, both in the visualization of unstructured data, as well as in automatic or semi-automatic conversion of unstructured data into structured forms, to tap into these rich sources of new visualization applications.