Selecting a Visual Analytics Application

Authors:

Professor Pat Hanrahan, Stanford University, Chief Scientist, Tableau Software

Dr. Chris Stolte, VP, Engineering, Tableau Software

Dr. Jock Mackinlay, Director, Visual Analytics, Tableau Software

Name Inflation: Visual Analytics

Visual analytics is becoming the fastest way for people to explore and understand data of any size. Many companies took notice when Gartner cited interactive data visualization as one of the top five trends transforming business intelligence. New conferences have emerged to promote research and best practices in the area, including VAST (Visual Analytics Science & Technology), organized by the 100,000 member IEEE. Technologies based on visual analytics have moved from research into widespread use in the last five years, driven by the increased power of analytical databases and computer hardware. The IT departments of leading companies are increasingly recognizing the need for a visual analytics standard.

Not surprisingly, everywhere you look, software companies are adopting the terms “visual analytics” and “interactive data visualization.” Tools that do little more than produce charts and dashboards are now laying claim to the label.

How can you tell the cleverly named from the genuine? What should you look for? It’s important to know the defining characteristics of visual analytics before you shop. This paper introduces you to the seven essential elements of true visual analytics applications.

The 7 Essential Elements	
Visual Exploration	Querying, exploring and visualizing data are a single process.
Augmentation of Human Perception	Visual thinking is encouraged and developed - the brain’s ability to process pictures far faster than text is leveraged.
Visual Expressiveness	Visual displays have depth, flexibility and multi-dimensional expressiveness.
Automatic Visualization	Effective visualizations are automatically recommended.
Visual Perspective Shifting	Shifting among alternative visualizations of any given data is effortless.
Visual Perspective Linking	Multiple images are intimately linked so a selection on one shows related, relevant data in the others.
Collaborative Visualization	People can easily share and collaborate on useful information visualizations.

Figure 1: *There are seven essential elements of a visual analytics application.*

Does a true visual analytics application also include standard analytical features like pivot-tables, dashboards and statistics? Of course—all good analytics applications do. But none of those features captures the essence of what visual analytics is bringing to the world’s leading companies. Visual analytics is a new area of technology, and it adds something special to your business intelligence toolkit.

Defining Visual Analytics: First, What It’s Not

Let’s start with what visual analysis is not: A graphical depiction of data. Virtually any software application can produce a chart, gauge or dashboard. Visual analytics offers something much more profound. Visual analytics is the process of analytical reasoning facilitated by interactive visual interfaces.

People who understand visual analytics know there’s something missing in the shallow charting wizards and dashboards found in most business intelligence packages and spreadsheets. While charts and dashboards are indeed “visualizations,” they leave out three critical steps: exploration, analysis and collaboration. A chart, for instance, shows conclusions, but not the thoughts behind it. Nor can users use a chart to ask questions and think further. In a chart, the thinking has taken place already and the resulting visualizations are little more than a show.

Visual analytics is a means of exploring and understanding data. It supports and accelerates the analysis process itself. You can ask a question, get the answer, and ask follow-up questions—all within visual interfaces. A story unfolds from one visual summary to another. You maintain your train of thought without taking your eyes off the data. Later, you can retrace the story to rethink, explore further and share. In short, visual analytics allows you to go in any direction with your thoughts while leveraging your visual perceptual system to guide you down the most useful paths.

Who is Adopting Visual Analytics?

Visual analytics is being adopted by the world's leading companies, universities and government agencies. From the world's largest and most innovative organizations – Proctor & Gamble, Apple, Pfizer, Microsoft, Coca Cola, Google, Cornell University, Progressive Insurance, Amazon, Georgetown University, the VA (Veteran's Administration), Blue Cross Blue Shield – to one-person consulting shops, visual analysis tools are now mainstream.

Where visual analytics was once thought to be in the domain of scientists and engineers, people now recognize that visual analysis accelerates business analytics. Even people who have made do with Excel and wizard-driven charting can benefit from today's visual analytics applications – the applications have become easy enough that anyone can use them.

Assessing Features of Visual Analytics Applications for Business Decision Making

If visual analytics is more than just presenting data visually, then assessing visual analysis applications is more than just looking for software that can produce cool visualizations. To distinguish good visual analytics tools from bad, it's critical to focus on the seven essential characteristics that distinguish a visual analytics application.

Visual Exploration

The first characteristic of a visual analytics application is the most important: The application unifies the steps of querying, exploring and visualization data into a single process. Do the data and the visualization work in tandem? When the user pulls on the visualization the data should come along. What does this mean? It means people can go in any direction with their thoughts, without ever losing their train-of-thought. They may not even have a specific question, but as they move through the data visually, they notice something and that prompts a question and then a follow up and so on, eventually leading to insight. It means the visualizations in a visual analytics application allow people to stop and take a closer look. Filtering, grouping, sorting and drilling all take place within the visualization itself, with a click. A user may start with basic questions and then, based on visual cues and insight, deepen the inquiry. The questions may mean eliminating some data, validating it, or reaching for a new set of data altogether. A visual analytics application helps people do all of this visually, and on-the-fly. It's like the early days of the web and Mosaic, when people first starting getting the idea of stream of consciousness exploration via hyperlinks. It was a game changer.

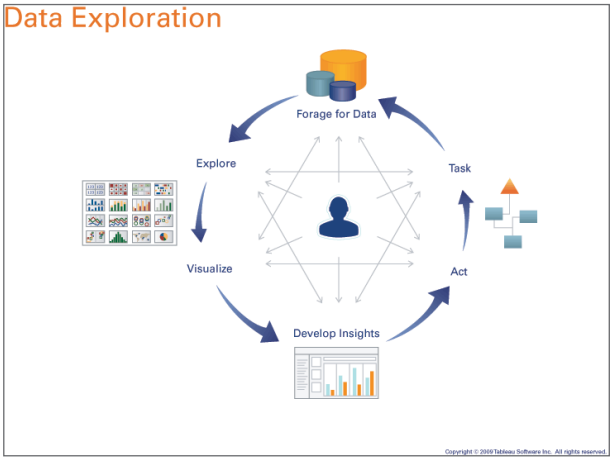

Figure 2: *Visual analysis applications support the cycle of visual analysis. When people are exploring data, they enter into a cycle where at any point they may need to skip steps, back-up, seek additional data, or even start over. Visual analysis applications support this process of visual data exploration.*

How do you know if an analytics application is appropriately designed to support visual exploration? Here are a few tests:

- Is the user interface easy to master and fluidly intertwined with the data? Using the software

should be so easy that people don't even think about the mechanics of creating a visualization.

- Can people apply filters without needing to divert their attention from the visualization? For instance, people should be able to lasso items and exclude them with a click. They should be able to perform drag-and-drop culling. Filters should also be applicable, when necessary, to a collection of interactive visualizations being viewed simultaneously.
- Does the interface allow people to group data into categories right within the visualization? This includes drag-and-drop grouping and on-the-fly binning.
- Can people perform on-the-fly computations on the data within a visualization—for instance, by changing aggregations or creating new calculations (e.g., Year over Year change) instantly?

Cleveland Clinic

Visual analytics paid off dramatically at Cleveland Clinic. There, the accounts receivable staff simply couldn't analyze rejected claims fast enough. Each year, the organization lost substantial amounts from unfulfilled insurance reimbursements for some of the 50,000-plus patients the hospital serves. Claims analysts could not examine claims records iteratively to uncover why the claim was rejected. There was no way they could look at a set of claims and follow multiple lines of reasoning to identify what could be causing the problem.
Conventional reporting solutions didn't allow them to group claims into different segments, drill into details or filter interactively. A visual analytics application from Tableau gave them the user-directed environment they really needed. Running on top of an existing data warehouse, Tableau enabled Clinic staff to analyze unpaid and rejected claims as they rolled in, even though inspection requirements varied from case to case. Repayments rose dramatically. In one year, the hospital attributed over $20 million in recovered payments to their use of visual analytics.

- Can people select data graphically and then drill to the underlying records without any special configuration? Visual analytics moves at the speed of thought. It requires tools to stay out of the way and react instantly.

Augmentation of Human Perception

Genuine visual analytics applications encourage visual thinking by leveraging the powers of human perception. The human brain possesses an amazing capacity to process graphics faster than it can process tables of numbers. Unfortunately, most business intelligence packages and spreadsheets do not take full advantage of the brain's perceptive capabilities. For instance, they use color and visual effects irresponsibly, and they ignore proven research on displaying data without bias.

In a visual analytics application, properly visualized data "pops." For example, relationships, trends and outliers show up bright and clear—aiding both the user and the audience alike.

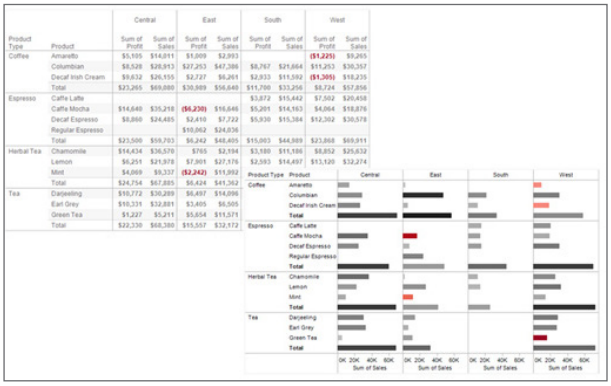

Figure 3: *Visual analytics leverages an important and powerful truth about the human brain: a picture really is worth more than all the words and numbers describing it. In the top illustration, some data are shown in a typical report or spreadsheet. Even with the use of color (a visual cue), it's hard to understand much about sales and profit. But the image below, based on the exact same data, makes the patterns and magnitudes clear at a glance.*

How do you know if you are looking at a true visual analytics application? Here are some tests:

- Do the visualizations that the application generates use effective visual properties? Precise use of size, color, shape and text, for example, make a difference. When handled well, they aid interpretation; when done poorly, they distract and mislead. The application should contain effective visual defaults that would make visualization experts like Edward Tufte and Stephen Few proud.
- Some data applications contain chart types that actually slow visual perception; for instance they employ cute, wacky or irresponsible metaphors like fractals, 3D objects, dials and speedometers. Visual analytics software helps amplify a person's thinking in the most elegant, communicative and proven ways.
- Does the application contain carefully designed maps appropriate for information visualization? A common mistake is to employ free Internet mapping services designed for generating online driving directions. Such maps are ineffective for most data visualization needs because they visually overwhelm what you most care about (your data), overemphasize roads and landmarks, and are generally inflexible.

Cornell University

Cornell University found out what a difference perception makes. They tried managing and monitoring KPIs with a traditional business intelligence tool—but in nine months business officers couldn't make sense of the dashboards or reports that the system produced. Then they deployed a visual analytics suite from Tableau. They enabled Cornell to create easy-to-understand dashboards.

Visual Expressiveness

No aspiring painter would put up with a paint-by-numbers canvas. But that's what many programs force on people when they use charting wizards and dashboards. Good visual analytics tools accommodate people's need for depth, flexibility and expressiveness in the visual displays.

This is especially important when people need to look at more than two or three dimensions of a problem simultaneously. Imagine putting five dimensions of a problem (e.g., Year, Month, Region, Product Family and Units Sold) into a charting wizard: the result just doesn't come out well. Visual analytics applications let people visualize multiple dimensions of a problem effortlessly, in formats that are easy to understand. Where cross-tabs and pivot-tables often confuse and overwhelm, multi-dimensional visualizations clarify. Visual analytics applications display complex problems with elegant simplicity.

Multidimensional expressiveness is particularly important when time and geography are involved. Let's face it, time and space are special. Data under geographic analysis may not honor geographic boundaries, and the interface should let the user follow any line of geographic inquiry. Does the application contain built-in geographic intelligence that effectively displays multiple levels of geographic detail?

Equally important is the treatment of time dimensions. Handling time appropriately isn't as simple as adding trend lines. People should be able to visually display dates and date/times at multiple levels of detail simultaneously. For example, examining sales trends by weekday (e.g. Mondays) and by month (e.g., January, February, March) simultaneously may be critical to discovering an uptick in caffeinated beverage sales on winter Monday mornings. Many systems do not have the capability to look at time in flexible, useful ways.

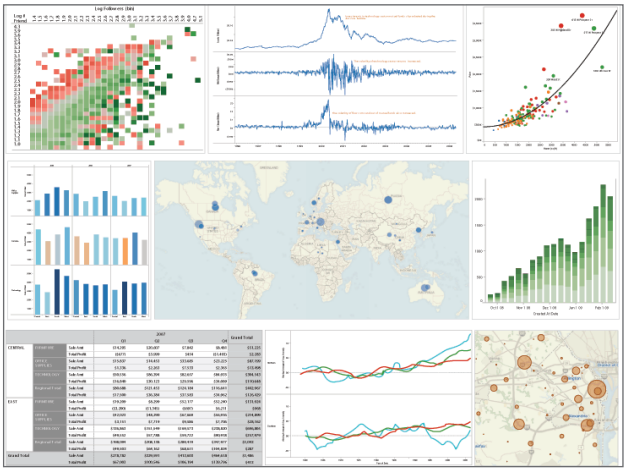

Figure 4: *Good visual expressiveness capabilities accommodate people's needs for depth and flexibility in visual displays.*

Here are some tests of an application's multidimensional expressiveness:

- Can people create tables of visualizations -- also known as "small multiples," as espoused by Edward Tufte? Or does the software primarily produce the same basic graphs found in Excel's charting wizard? Most business intelligence applications deliberately copied the choices found in Excel, not understanding the limitations of template-driven graphics.
- Can people create visualizations that clearly display 8-12 dimensions of a problem simultaneously? Users often need to visualize many dimensions of a problem at once.
- Can people invent their own custom visualizations and save them as templates that can be applied to other data? Can they easily share their visual customizations?
- Does the application contain maps that offer precise control over geographic layers? Worldwide mapping at multiple levels of geography should be provided by default and require no special data preparation or geocoding.
- Can date and date/time fields be viewed at multiple levels of detail simultaneously?

Quest Diagnostics

Quest Diagnostics, which provides medical testing to large healthcare facilities, faced the loss of a multi-million dollar client. Test results were coming in too slowly, the client said. Though the Quest director saw the statistics differently, reports and charts produced conventionally were not convincing. And when the client had new questions, Quest couldn't respond fast enough with clear, meaningful analytics.

Tableau, a visual analytics application, changed all that. The director flew through the visualized data instead of trudging through rows and columns—and created new views of client data, asking different questions and getting new perspectives on Quest's performance. Even as he presented his findings to his client, his client began asking additional questions and asking for different visualizations. Together, he and his client visualized the testing timeline by test type, priority and facility type. They discovered that much of the problem was caused by the clients' staff not consistently following procedures across different tests, different priority statuses and facility type. Ultimately, the client not only stayed with Quest but a year later bucked the corporate mandate to switch providers.

- Can time and geography be combined in views that show changes over time at specific locations?

Automatic Visualization

Imagine an application that tells you how you should look at the specific problem you have. For too long, analysts have been taught to think in numbers alone. A visual analytics application jumpstarts the analysis process itself. This includes automatically suggesting effective visualizations.

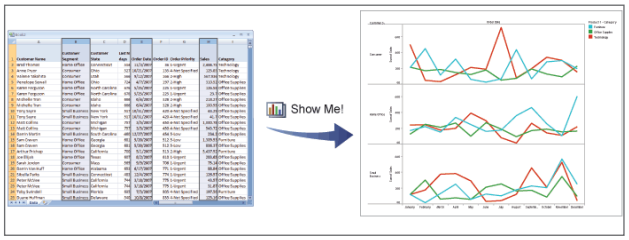

Figure 5: *Automatic visualization not only helps people find patterns easily but also trains them to think visually and more rapidly.*

A key benefit of automatic visualization is not just that it reduces work time. It also helps people learn to think visually. If they can think in pictures, they can work faster and recall trends and patterns more easily. Here is the essential thing to look for:

- Does the application have a rules-based engine that drives the most appropriate graphical depiction based on the selected data elements? In other words, the automatic visualization should not just be a random selection but should be based on heuristics.

Visual Perspective-Shifting

There is never a single visualization that offers the best summary of every finding in your data. Typically people need to look at a variety of visualizations, depending on the tasks you want to achieve. Effective visual analytics applications should suggest a series of alternative visualizations which can be effortlessly flipped through. For example, if you're trying to find outliers, look at a scatterplot. Trying to understand time-based trends in the data? Then a line or Gantt chart might be ideal. Trying to understand multi-dimensional geographic variation? Try a small multiple of maps. No one view can answer all questions.

History has shown that looking at data from the right perspective is as important as looking at the right data in the first place. Physicist Richard Feynman famously showed that engineers working on the Space Shuttle Challenger had in front of them all of the data needed to conclude that launching the Shuttle at low temperatures posed unprecedented risks. One of their failings was that they never plotted the data in the most revealing way.

This was well documented by Edward Tufte in his book Visual Explanations: Images and Quantities, Evidence and Narrative. Shifting visual perspectives on a problem is also a great way to generate new questions. It jostles the brain a little and makes you more curious about what is actually going on in the data.

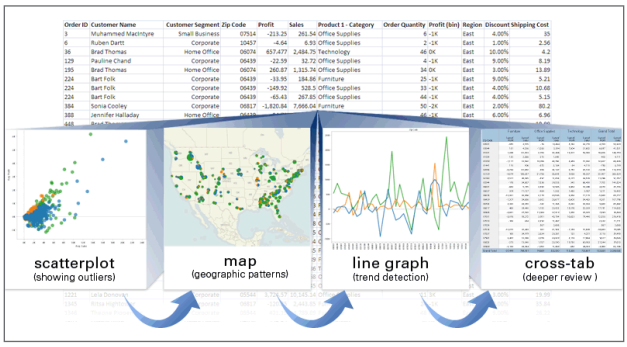

Figure 6: *Different visualizations answer different questions and solve different issues. Above we see the same dataset visualized in multiple ways. Each perspective answers different questions and highlights different patterns. Visual analytics lets you move from one perspective to another with a click.*

For these reasons, visual analytics applications contain strong support for perspective shifting. Here are some qualities to look for:

- Can people examine the same data in different graphical layouts instantly? In just a few clicks, a person should be able to move from view to view until she finds exactly the right view for the question at hand.
- Can people bounce between alternate visualizations of the same data with a click?
- Can people view multiple visualizations of the same data simultaneously?

People on the lookout for useful information visualizations like to forage freely in data. So the last thing people need is a tool that confines them to a single, linear path. An inflexible tool creates a dataset and a chart and tries to stick with it. A visual analytics application instead offers direct access to a myriad of visualizations, with no boundaries.

Visual Perspective Linking

A logical but powerful addition to perspective shifting is perspective linking. Although the two topics are related,

The Martin Agency

Analysts at The Martin Agency's Ingenuity Media Group sifted through mountains of data from interactive media campaigns to produce periodic reports and dashboards for clients. Production took hours. Analysts only had time to create were static, standard charts, with no time to find new insights. But with Tableau's perspective shifting features, analysts could look at multiple perspectives of their clients' data and rapidly find the most meaningful one. It opened a new world of exploration and discovery.

linking entails a different set of capabilities than perspective shifting. In short, it isn't enough to look at multiple perspectives on a problem in rapid succession – or even simultaneously. Sometimes the perspectives need to be intimately linked. One visualization may display a set of outliers, for instance. Can a person select an outlier and instantly see another visualization that displays greater detail? As an example, a person may notice that sales for a particular state seems to be dominating. By clicking on the mark representing that state's sales, he can instantly update a visualization

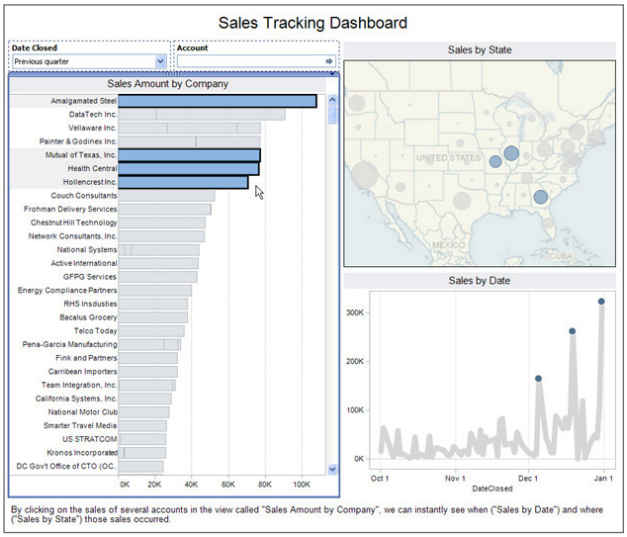

Figure 7: *In a visual analytics application, you can use perspective linking to discover relationships in data or to uncover hidden stories. In this example, a user selects a state on the map (California). Tableau then instantly reveals.*

regarding sales amount by company and see another visualization redraw a line chart showing sales by date using data for just that state. That interaction may shed light on what is driving sales in that state to be dominant.

Genuine visual analytics applications support perspective linking in several ways. Here are some features to evaluate:

- Does the software visually correlate the information in different visualizations? For instance, can users select a trend line in a first view and see the geographic entities related to that line in a second view?
- Can people use one visualization to "drive" another? For instance, can people use a first visualization to search for data of interest, and then automatically apply the selected data as a filter on a second visualization?
- Can people tap into this functionality merely by selecting or hovering over data?
- Can people take advantage this functionality without any programming? In a visual analytics application, rich data associations should be discovered automatically.

Collaborative Visualization

Another defining capability of effective visual analytics applications is the ability to iteratively create useful information visualizations in a team setting. This process usually starts with a "hey, look at this" moment. But the real question is: Does the software support the involved collaborative process that should follow these moments?

Shared findings lead to solutions, action, and results. In fact, in most organizations unshared discoveries are useless. Some software packages, while meeting other criteria, fail here. Effective visual analytics software encourages collaboration by letting results be shared in whatever form the user prefers. The application's architecture should be built explicitly for collaboration.

Here are some of some specific features to evaluate:

- Can people publish interactive results securely to the web in seconds? Visualizations should be available via a browser. They should be live and change as data changes.
- Can published visualizations automatically be tailored to the name, role or group of the particular information consumers? In addition, will people who are viewing or interacting with the visualizations have the ability to customize their particular changes to it and save that view?
- Having developed a library of visualizations, what's a team to do with it all? The application should support Google-style search that enable groups to locate visualizations quickly.
- Does the application have capabilities for commenting, tagging, customizing and sharing visualizations and dashboards?
- Does the application natively support versioning of visualizations and data models? Whole data models (or individual calculations, defined groups, etc.) should be easily shared and version controlled.
- Does the software support embedding interactive visualizations into other applications and portals?

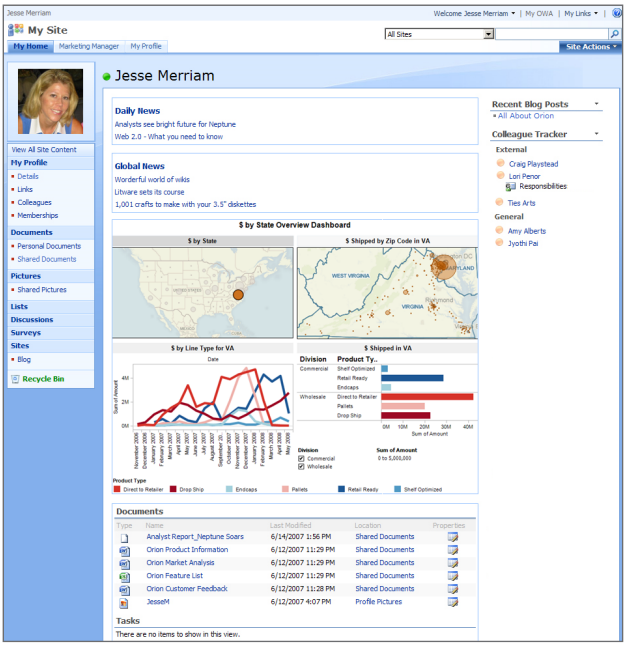

Figure 8: *In this illustration, people share key findings from the visual analytics system with colleagues using Sharepoint. Even though the results are embedded, people can slice and dice the data using check-box filters, sort, view underlying data, zoom into specific points and even view the underlying findings.*

Deployment Considerations

This paper has described the seven defining elements of visual analytics applications. Yet there are other factors to consider before making a choice on a visual analytics standard. For instance, if the visual analytics application is going to be used in a corporate environment, there are technical and infrastructure considerations as well.

Universal Data Access. Data comes from all directions. Does the software connect to virtually any source, from data warehouses to Excel or text files? Does it connect to all major data formats, including relational databases, OLAP data cubes and flat files? Furthermore, is connecting to new data sources easily done?

Scalability. Does the software support real-time interactive visualization of data of nearly any size—even millions or billions of records? The application should be able to handle large amounts of data and provide solid performance.

Generates Efficient SQL for DBMS. Reports should generate quickly. People should have the option to take, modify and run SQL from the reports and reproduce the results. They should also be able connect to SQL statements rather than actual data tables and views.

Ability to Join Tables. People should be able to join tables easily in ways that guide them to appropriate and well-designed joins. Tables can be joined automatically or based on user's direction.

Security. Does the application have a full security module to support collaboration within existing permission systems? This should include optional

integration with other security methods like Active Directory.

Minimal IT Support. Good software frees IT from the small stuff. Good software installs easily, gets the average user up in minutes without help, and provides free training on demand.

Microsoft Office Compatibility. MS Office is ubiquitous on the business person's desktop. Does the software provide gateways in and out of MS Office? Can users natively access data from Office applications? Can they output images, tables, data lists, cross-tabs easily and directly to Office?

Data Modeling and Management. Can users share entire data models used to construct visualizations—along with individual calculations, groups, etc? Users should be able to rely on dimensions and measures that are consistently defined and have data integrity. Can the software allow the user to model data –- such as to modify variable types, replicate variables, change field names, standardize dimensions -- without developer support? This is critical for giving users the analytical power while also reducing the IT burden.

Training. How much is needed? Ideally people can get started without any training – the user interface should be obvious and easy to use. Is there free online training available any time for multiple levels? As people advance in their visual analytic skills, more sophisticated training should be available on their time schedule based on their needs.

Licensing. Licensing models should be flexible based on user needs without minimum configurations: buy 1 license, buy 10, buy 100, buy 1000.

Further Reading on Visual Analytics

1. J.J. Thomas and K.A. Cook, Illuminating the Path, IEEE CS Press, 2005.
2. Stuart Card, Jock Mackinlay, Ben Schneiderman, Readings in Information Visualization: Using Vision to Think, Morgan Kaufmann, 1999.
3. Tufte, E. R., The Visual Display of Quantitative Information, Graphics Press, 2001.
4. Stephen Few, Perceptual Edge, *Improve Your Vision and Expand Your Mind with Visual Analytics*, whitepaper.
5. Stephen Few, Show Me the Numbers: Designing Tables and Graphs to Enlighten, Analytics Press, 2004.
6. Jacques Bertin, The Semiology of Graphics, Out of print, 1967.
7. Donald A. Norman, The Psychology of Everyday Things, Basic Books, 1988.
8. Colin Ware, Information Visualization: Perception for Design, Morgan Kaufmann Publishers, 2000.
9. Analysis, and Visualization of Multidimensional Databases, CACM, November2008.
10. Chabot, C., Demystifying Visual Analytics, IEEE/VAST based on "Practical Applications of Visual Analytics: On the Cusp of Widespread Adoption," given at the 2008 IEEE Symposium on Visual Analytics Science and Technology.
11. Jock D. Mackinlay, Pat Hanrahan, and Chris Stolte, *Show Me: Automatic Presentation for Visual Analysis*, whitepaper.
12. Pat Hanrahan, Chris Stolte and Jock Mackinlay, *Visual Analysis for Everyone - Understanding Data Exploration and Visualization*, whitepaper.

About Tableau

Tableau Software helps people see and understand data. Ranked by Gartner and IDC in 2011 as the world's fastest growing business intelligence company, Tableau helps anyone quickly and easily analyze, visualize and share information. More than 10,000 companies get rapid results with Tableau in the office and on-the-go. And tens of thousands of people use Tableau Public to share data in their blogs and websites. See how Tableau can help you by downloading the free trial at *www.tableausoftware.com/trial*.