Bringing Your Data to Life – and the Public

How The University of Texas System is increasing transparency to the public with dynamic data visualizations

Featuring:

Stephanie Bond Huie, Vice Chancellor, Office of Strategic Initiatives, The University of Texas System

Paul Bachteal, Senior Director, Americas Technology Practice, SAS
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Changing Expectations for Transparency and Accountability</td>
<td>1</td>
</tr>
<tr>
<td>Evolving the Information Architecture</td>
<td>1</td>
</tr>
<tr>
<td>The Evolution of Data Use</td>
<td>2</td>
</tr>
<tr>
<td>Using data to select peers</td>
<td>2</td>
</tr>
<tr>
<td>Using data to set goals</td>
<td>2</td>
</tr>
<tr>
<td>Using data for benchmarking</td>
<td>3</td>
</tr>
<tr>
<td>Using data to inform policy</td>
<td>3</td>
</tr>
<tr>
<td>A Quick Tour of Interactive Visualizations</td>
<td>4</td>
</tr>
<tr>
<td>About SAS® Visual Analytics</td>
<td>6</td>
</tr>
<tr>
<td>Closing Thoughts</td>
<td>8</td>
</tr>
<tr>
<td>For More Information</td>
<td>9</td>
</tr>
<tr>
<td>About the Presenters</td>
<td>9</td>
</tr>
</tbody>
</table>
Introduction

How do you make a good organization great and a great organization into a leader?

Well, it starts with knowing what leadership looks like for your industry, knowing where you are now, tracking progress toward tangible goals – and putting all that information into the hands of decision makers, so they can chart the right course and see the effects of their decisions.

That’s exactly what The University of Texas System is doing.

With nine universities and six health institutions, The University of Texas System serves more than 215,000 students and awards nearly 49,000 degrees and certificates a year. The 2014 fiscal year budget is more than $14.6 billion, plus $2.5 billion in research expenditures, making it one of the nation’s largest systems of public higher education. Everything is bigger in Texas, after all.

As a critical force in the state’s economic and health care vitality, The University of Texas System has long worked to be at the forefront of higher education in transparency and accountability. From its annual Accountability and Performance Report to a detailed Facts & Trends statistical handbook, the UT System has a history of publishing detailed data on institutional performance. However, amid growing demands for data, the traditional channels for sharing this information – mostly static PDFs with summary-level data online – were no longer enough. Users wanted more data, more up-to-date data, more often and on demand. They also wanted to be able to see the data sliced and diced in different ways and filtered by different criteria. These demands were scaling up as higher education institutions in general faced greater scrutiny than ever.

Changing Expectations for Transparency and Accountability

In 2005, US Secretary of Education Margaret Spellings asked some tough questions. Was the American higher education system deteriorating and failing to prepare the American workforce for the globalized marketplace? Spellings formed the 19-member Commission on the Future of Higher Education to investigate issues of access, affordability, quality, accountability and innovation in education. The commission’s controversial 2006 report, A Test of Leadership: Charting the Future of U.S. Higher Education, proposed several solutions to identified deficiencies. Most of the solutions called for far greater levels of transparency and accountability – not just internally, but with students, parents, taxpayers and other external stakeholders.

So the stage had been set. The Spellings commission report set the tone for openness and accountability in the quest for academic excellence. In 2011, The University of Texas System Board of Regents and Chancellor Francisco Cigarroa developed a Framework for Excellence that defined an action plan to reach new performance levels in nine mission areas.

“This isn’t your typical company strategic plan with pretty pictures, glossy paper and platitudes,” said Stephanie Bond Huie, who leads the Office of Strategic Initiatives for The University of Texas System. “It called for campuses to establish concrete and measurable performance goals for 2015, 2020 and 2025.” The plan created the need for a dynamic information framework that would serve four purposes:

- Improve transparency by providing open, public access to information through an interactive information portal that requires no login yet protects servers and data integrity.
- Present data to leadership in a way that can be easily understood, evaluated and customized to make data-driven policy decisions and assess progress.
- Replace the hard copy and PDF fact book with public, online access to the data warehouse, with metrics updated as new data becomes available.
- Streamline and automate data collection processes wherever possible, to free up more time for analyzing the data and pursuing more complex institutional research projects.

Evolving the Information Architecture

“Going back as far as the 1970s, we had produced a lot of printed reports – big stacks of paper, numbers and tables,” recalled Bond Huie. “In 2003, we started moving toward more graphical presentation of information, more context and some limited peer benchmarking – and we started publishing reports online in PDF form.”

The data requirements of the Framework for Excellence drove the next stage in the evolution. “As part of this evolution, we changed the way we store data – from many small data analysis shops with Excel spreadsheets, to SAS® data sets and then to an integrated data warehouse,” said Bond Huie. “We also changed the way we deliver data – from printed reports to PDFs on the Web and now to interactive visualizations and online dashboards.”
The latest stage of this evolution has produced two distinct tools:

- **The BI/Information Delivery Portal** ([data.utsystem.edu](http://data.utsystem.edu)) provides online, publicly accessible dashboards and Web reports. Compared to static reports, the interactive dashboard presents data in multiple ways, such as by degree type, student category, institution or department.

  The data visualizations are not limited or predetermined. Users have the flexibility to explore data in various ways to answer their questions and drill down into the data tables. Users can also download the data into an Excel spreadsheet, do their own analysis and draw their own conclusions.

- **Visual Analytics** ([exploredata.utsystem.edu](http://exploredata.utsystem.edu)) makes it easy for nonprogrammers to manipulate and analyze data and create their own reports that bring visual clarity to the data. This new capability has already proven quite popular.

  “Another great thing about our data visualizations is that they are available on the iPad®,” said Bond Huie. “Our executives really wanted this. They wanted to be able to pull up information during meetings that can help support their case right then and there.”

  These two tools have been really transformational for our office, because we get a lot of requests for data from our constituents, from the legislature, from our Board of Regents and others. We still get these requests, of course, and we still provide them that data in the traditional way, but there are those who are adept at doing this and really want to dig into the data. Now we can offer them this value-added service.”

**The Evolution of Data Use**

“You can have all the data in the world, but if you can’t get it out there to your consumers in a way they can use, it’s not very valuable,” said Bond Huie. “So we became more sophisticated in our approach with these tools. We went from reporting and analyzing to making predictions – from planning to predicting to improving.” Bond Huie described some key ways the UT System is using data to support the Framework for Excellence.

**Using data to select peers**

If you want to measure your performance against peers, how do you select those peers? For the UT System, data-driven decision making replaces instinct and intuition, said Bond Huie. “Gone are the days where someone could say, ‘I’ve been in the industry x number of years, and I feel like our organization looks like this organization, and so therefore this is the organization we’re going to use for comparison purposes.’ A lot of people had very strong feelings about which institutions were similar to theirs and what they were trying to be, and we found that often those assumptions were wrong. For benchmarking to be powerful, it’s important to use data to drive your peer selection process.”

This process called for selecting two different types of benchmarking peers:

- **Baseline peers** - the organizations statistically similar to UT System institutions today.
- **Aspirational peers** - the ones that represent the desired future state.

The analysis started with more than 40 variables that were deemed meaningful to determine similarity, such as percent of students eligible for Pell grants, operational revenue, cost to produce a degree, funding per student, program mix, and ratio of full-time to part-time enrollment. Further analysis distilled that list to 12 variables that were correlated with outcomes to be evaluated. Those 12 variables were used to compute the statistical “distances” between universities and to order all 378 universities in the peer selection pool by their distance from a chosen one in the UT System. The closest ones were seen as the best baseline peers, and the ones further away were likely candidates to be aspirational peers.

**Using data to set goals**

The Framework for Excellence called for each institution in the system to set quantifiable performance goals for 2015, 2020 and 2025. Where do you set those goals, if you don’t want them to be arbitrary or unachievable? Here again, data drove the decisions. For example, short-term goals (by 2015) could be set relative to a
median performance quintile of the chosen baseline peers, midrange goals (2020) aligned with the top quintile of baseline peers, and long-term goals (2025) set at the bottom quintile of aspirational peers.

The chart below shows benchmarking against peer bottom quintile, peer average and peer top quintile for four-year graduation rate performance targets for students graduating in each subsequent year. With this type of visualization, you can track progress toward goals as a trajectory over time.

**Figure 1: Using peer data to set meaningful goals.**

**Using data for benchmarking**

"Benchmarking has been a very important tool for us as a way to gauge our progress and evaluate data in context," said Bond Huie. "For example, a basic bar chart showing the cost of a four-year degree for institutions in the UT System doesn’t provide a lot of information. What do you do with this? But when you add in the peer comparison, now the information becomes more meaningful. I can see if our institutions are spending more or less than peers to produce a degree. Now I have an indication that I can actually act on. If I looked at this chart and saw that my institutions were spending more, I would definitely want to know why and do some further exploration."

**Using data to inform policy**

The BI/Information Portal and SAS Visual Analytics tools have been instrumental in answering questions from task forces organized by the Board of Regents, such as the Task Force on University Excellence and Productivity, the Student Debt Reduction Task Force and the Task Force on Engineering Education.

The Office of Strategic Initiatives also relies on the tools to satisfy ad hoc requests both internally (from the board, chancellor or officers) and externally (such as from the legislature and reporters).
A Quick Tour of Interactive Visualizations

Visit The University of Texas System dashboard – data.utsystem.edu – and you are greeted by a welcome page that describes the project, offers some user tips, and showcases new analysis or data updates.

Click the Core Indicators tab to display the 10 highest-priority indicators of overall system performance. In that dashboard, data visualizations show where the UT System is doing well, where it could be doing better and trends over time. Click to get definitions and methodology to understand the sources of the data. Check filter boxes to hone in on specific areas of interest. Drill down to see more detail or another level of breakout. Additional dashboards provide windows into student success, faculty productivity, research and technology transfer, and finance and productivity.

Or you can jump to a predefined report or visualization. For example, click Visualizations to see predefined visualizations for Revenue, Research Expenditure, STEM Degrees, Peer Data, Graduation Rates, Student Debt, Tuition and Fees, Patient Care and more. Each visualization offers options to get granular, drill deeper and change the selection criteria. For example, for Patient Care at one of the system’s six health institutions, you can see net revenues by source or by institution, for the most recent five years or an earlier five-year rolling window. Hover over a graph element to see a pop-up that explains what you’re seeing. Click another tab to display the data along a different dimension, or by institution rather than by source. You’re in control.

Here are some more examples:

**Supporting decisions about student debt** - Traditionally, a measure such as student loan default rates would be shown in a static bar chart, perhaps compared against the system-wide median. With the visualization tool, users can interact with the data to visually explore the data from many angles, such as zeroing in on average debt for students who graduate with a bachelor’s degree for a certain range of years, compared to a trend line that shows average over time, by institution, or compared to state and national averages.

**Supporting decisions about engineering education** - When the Task Force on Engineering Education wanted to know whether the engineering degree program was meeting demand for engineers in the state, several different types of visualization proved useful.

The bubble diagram in Figure 2 shows the SAT scores of entering engineering and computer science students, where:

- The x axis shows math score.
- The y axis shows reading score.
- Each bubble represents an institution in the system.
- The size of bubble indicates the percentage of students in the entering cohort that had been in the top 25 percent of their high school classes.

Enrollment, degrees awarded, graduation rates, and baccalaureates employed or enrolled in graduate school (this kind of information and much more is available on the dashboards and visualizations) can be segmented by institution, by categories, by selected time periods and more, at summary and detail levels.
The interactive bar and pie charts in Figure 3 below provide a dynamic, multidimensional format for looking at salary outcomes for recipients of undergraduate engineering degrees. From this view, you can manipulate the chart to zoom in on your area of interest. Choose a subfield to investigate more closely, such as mechanical, chemical or petroleum engineering. Click a chart element to see the underlying data points. Send the chart in an email. It's all point-and-click easy for anyone to access on the Web.

Figure 3: A colorful dashboard presents many pieces of information in one user-defined view.
The tree map chart in Figure 4 shows the time to degree for students in various engineering degree programs. The size of each box indicates the size of the cohort. The color of the box indicates time to degree. Boxes that are more brown reflect degree areas where students are taking longer to get a PhD degree. Lighter boxes show areas where students are taking less time.

Visualizations such as this can provide “aha” moments when you find patterns you didn’t expect. You might discern these things while poring over a data table, but the discoveries don’t jump off the page as they do in a color-coded visual.

“Give them what they want, but show them what is possible. Not everybody has the mindset to imagine how a visualization could help them move their operations forward.”

Stephanie Bond Huie, University of Texas System
In the process, the organization gains more value from the data it already has, as well as from the data it formerly couldn’t use, said Bachteal. “At the first stage of analytical maturity, you have business intelligence that answers questions you know, using data you know. There’s not much competitive advantage there, because everybody is doing this. When you add visualization, you can start to explore questions you don’t know, because patterns become visually apparent. When you add analytic discovery with big data and Hadoop, now you’re asking questions of a vast lake of previously unknown data, not just what has been through the rinse cycle a few times in your enterprise data warehouse.

“Ultimately, you’re using predictive analytics with Hadoop to understand what you should adjust to achieve some desired future state. Now you are finding answers to questions you don’t know, using data you don’t know. This is the step-wise approach we’re seeing customers taking, as we help more and more organizations such as The University of Texas System adopt visualization and predictive analytics with big data technologies.”

Paul Bachteal, Americas Technology Practice, SAS

---

“When you add analytic discovery and Hadoop, now you’re asking questions of a vast lake of previously unknown data, not just what has been through the rinse cycle a few times in your enterprise data warehouse.”
What is Hadoop?

Hadoop is an open-source software framework for running applications on a large cluster of commodity hardware. Hadoop delivers enormous processing power - the ability to handle virtually limitless concurrent tasks and jobs - making it a remarkably low-cost complement to a traditional enterprise data infrastructure. Hadoop is gaining ground wherever organizations need to store, process and analyze large volumes of data.

Closing Thoughts

“It’s important to remember that data is just data,” said Bond Huie. “It’s the research and analytics that make it into information and the visualizations that make it consumable. Until you make data into information and make it usable, you haven’t hit the big jackpot yet.” Although the visualizations are new, The University of Texas System has already seen some very real benefits from its information and analytics portal:

- **Better policy decisions.** Timely access to data supports more informed dialogue and better policy decisions. Questions can be answered on the fly. Stakeholders can explore on their own and participate in the conversation. Results of policy decisions can be assessed over time.

  People are starting to notice. The dashboard/visualization project started with the academic enterprise - focusing on data related to students, faculty, classes, degrees, etc. But successes there have created demand from other groups, said Bond Huie. “Our own system administration is a business in and of itself. We have compliance officers, facilities manager, police department, risk management department and many other operational units. They’re now coming to us and asking how we can help them make their processes more efficient. This is really exciting for me, because I see this as an opportunity for us to take our organization to the next level.”

- **Self-service access to relevant insights.** By changing information delivery from static to interactive, from tabular to visual, the new information portal lets users view and explore data in novel ways, even to explore complex and nonlinear relationships.

- **Ready access to trusted data.** The integrated data warehouse provides a trusted repository that reduces data disparities and improves staff productivity. The office can produce more on-demand reports with more metrics and more complex metrics. Users don’t have to collect information from different people to answer new questions.

“The system has moved from paper-based reporting to PDFs from Excel spreadsheets to interactive dashboards and reports from our data warehouse. Now analytics and visualizations are really bringing data to life. Users are not just asking what happened, they’re asking ‘what now’ and ‘what’s next,’ because they understand the patterns from the data.”

**Stephanie Bond Huie, University of Texas System**
For More Information

For more information about SAS Visual Analytics, including a live playpen where you can generate many of the visuals described in this paper using industry data sets: sas.com/software/visual-analytics/demos/all-demos.html

For more information about SAS and Hadoop: sas.com/software/information-management/big-data/hadoop.html

About the Presenters

**Dr. Stephanie Bond Huie** is the Vice Chancellor for Strategic Initiatives for The University of Texas System. This office is responsible for research and policy analysis, performance management and long-term strategic planning. In her former position as director of the Office of Strategic Initiatives, Bond Huie was the project manager for the UT System Dashboard Project, overseeing the collection and processing of university data for the dashboard and conducting policy analysis. Huie has a bachelor’s degree in anthropology from the University of Texas at Austin and master’s and doctoral degrees in sociology from the University of Texas at Austin.

Since joining SAS in 1985, **Paul Bachteal** has held a variety of key leadership positions in sales, marketing and general management. As Senior Director of the Americas Technology Practice, Bachteal is responsible for supporting customer engagements across a broad spectrum of SAS product segments, including information management, high-performance analytics, data visualization and performance management. Most recently, Bachteal and his teams have focused on best practices to improve alignment among business and IT staff, particularly as SAS customers seek in-memory solutions to meet their rapidly increasing data and analytical challenges.