Where’s my data?
Data quality for the hospitality and gaming industry

Insights from a Webcast sponsored by SAS and the Center for Hospitality Research at Cornell University’s School of Hotel Administration
Originally broadcast in October 2009

Featuring:
Rohit Verma, Executive Director of the Center for Hospitality Research, Cornell University
Tracey Jarosz, Director, Customer and Marketing Insights, Fairmont Raffles Hotels International
Scott Chastain, Lead Architect, SAS Business Analytics Practice
When times are tight – when unemployment is up and stocks are down – the hospitality and gaming industry feels it more than most. Cash-strapped consumers curtail their leisure travel. Companies get conservative and cost-conscious with business travel. Recent years have not been halcyon times for the industry.

Conditions appear to be stabilizing as the economy shows tentative signs of a rebound. But room rates are still low – as much as 24 percent down from 2008 levels in some markets – and there is no big surge in demand forecast for the near term.

A November 2009 Deloitte survey of 2,000 consumers in the US showed that while almost half planned to take a vacation or leisure trip involving an overnight stay in the next four months, overall demand was flat. When asked to compare their travel plans with the same period a year ago, 45 percent of respondents reported no change, 25 percent said they will travel more and 28 percent said less.

The survey offers grounds for cautious optimism but not rejoicing – yet. Industry leaders still need to work toward optimizing every aspect of the business in order to emerge from the downturn as strong as possible.

The key to that goal is to get new intelligence from all the data generated by customer behaviors and operational transactions. In a perfect world, all this data is trustworthy, consistent and clean. Data-driven decisions are made in full, enterprisewide context. A deep understanding of customer preferences drives the best decisions about promotions, service offerings, inventory, restaurant menus and gaming selections. New and consolidated data sources drive wiser pricing decisions, supplier choices and financial strategies. Sales results, clarified in detail, guide best practices for marketing, staffing and expansion.

This ideal vision is not the reality for most organizations. In the real world, customer interactions are usually captured in a wide range of formats and a multitude of disparate systems. Extensive manual effort is required to access and prepare data for analysis – and still data quality is suspect.

The challenges surrounding data quality and data integration were the focus of an October 2009 Webcast sponsored by the Cornell University Center for Hospitality Research and SAS. In the interactive Webcast:

- **Rohit Verma**, Director of the Center for Hospitality Research, presented some findings of the center’s work with academic researchers and corporate partners toward improving data quality and analysis.

- **Tracey Jarosz**, Director of Customer and Marketing Insights for Fairmont Raffles Hotels International, described ambitious steps her organization has taken to resolve data quality issues and reap the benefits.

- **Scott Chastain**, Lead Architect, SAS Business Analytics Practice, discussed technologies and practices to improve data management, data governance and master data management.
WHERE’S MY DATA?

The topic was timely. If the informal Webcast poll was any indication, data quality is a fundamental problem in the industry. When asked to describe the current state of data quality in their organizations, less than one-third of respondents said it was “pretty good but some areas need work,” while 71 percent characterized data quality as “really poor, and we spend a lot of time cleaning data before analyzing it.”

Those results are discouraging but not surprising, our panelists agreed. “We have a lot of ways to capture information – an explosion of data – but not always a good way to analyze it properly, at least the way [data management processes] are implemented in most companies,” said Verma.

That’s a problem, since clean data is an absolute prerequisite for deriving meaningful intelligence to run the business. “All of our marketing initiatives rely on good quality data, from who we select to contact, to what we contact them with, to what channels we contact them through, how often we contact them and how we reward them for their loyalty,” said Jarosz. “It’s all dependent on data quality. For us, an incorrect address is more than a piece of returned mail sitting in somebody’s office; it’s a missed opportunity with a guest. Incorrect gender or name is a poor reflection on the brand. Data quality is essential.”

Imperatives and obstacles surrounding data quality

Quality data is unquestionably a requirement for quality analysis, but it has been an elusive commodity, for a number of reasons.

The data needed for analysis resides in many disparate sources.

In many organizations, data comes from paper-based documents and systems from multiple vendors. There are marketing systems managing customer contacts and campaigns, revenue management systems looking at demand patterns, reservations systems collecting booking information and so on. Rarely do these systems talk to each other. Processes and technologies are disconnected, making it difficult to link information to a particular guest or understand the sequence of events that constitute the total guest experience.

“Disparate systems managing different arms of the operation make it nearly impossible for hospitality companies to access the information they need to make profitable business decisions.”

Scott Chastain, Lead Architect, SAS Business Analytics Practice
Dirty data is everywhere.

“A midsized hotel or resort chain can easily have millions of customer records. That’s a lot of data to manage, and a lot of opportunity for dirty data,” said Chastain. The problem isn’t limited to any specific system or business unit. In an online poll, Webcast participants were asked what areas of the company had the most trouble with data quality. Operations seemed to be the most troubled area, according to 33 percent of respondents, followed by Marketing at 25 percent. A surprising 16 percent picked the last option; data quality problems were “everywhere.”

That means that nearly half of the audience is experiencing problems with operations data. Our panelists were not surprised at this figure. “This is consistent with what we see with some of the companies we work with,” said Chastain. What is surprising is that marketing data is problematic in so many organizations, when marketing automation and customer tracking systems have been widely deployed for some time now.

These challenges point to great opportunities for the companies that can conquer data quality issues. “Operations and marketing information is where an organization can differentiate itself from the competition by leveraging it appropriately,” said Jarosz. “The fact that data quality issues are as pervasive as they are suggests a huge opportunity for organizations to improve, get it right and use that information to differentiate themselves.”

Cleansing data has traditionally been a time-consuming and burdensome chore.

“In many organizations we visit, we find that about 80 percent of the time spent on analytical process initiatives revolves around data quality work, and at least half of that is manual effort,” said Chastain. “That finding has been consistent not only in hospitality and gaming, but across industries.” Jarosz estimated a range of 70 to 90 percent, based on her experience in several industries.

“For any research project we do, the first several days or several weeks are spent on data cleansing,” said Verma. For a recent project, Verma’s team received six years of transaction data and millions of transactions that required three months of data cleansing before it could be used for analysis. Verma noted that this degree of data quality effort, while necessary in academia, would probably be unrealistic in a commercial setting. “We are a little different because in academia we are looking at the broader questions, so we can take a step back and make sure that the analysis is perfect, so we can make a generalizable answer. In the industry, you have immediate problems to attend to. You have to fight fires all the time.”

Even when data cleansing is automated, it can be a fine art to set the rules correctly, said Jarosz. In eliminating redundant entries, for example, if you set your matching criteria too loose, you end up with many duplicates. If you set it too tight, you end up with a very small database.
Transactions tell only part of the story.

Even if you could capture facts from multiple systems, correlate that data and match it to a specific customer, does that yield the full insights you need for improving the business? Not necessarily, Verma said.

“One of my PhD students and I are looking at how sequences of experiences people have at a hotel or restaurant impact their future purchase behaviors,” Verma said. “You probably have transaction information available about all the things the guest did, to the extent that the database can capture. The guest goes into a hotel, checks in at a certain time, maybe goes to the room, and sometime later goes into the food and beverage area. You have all the transactions, but how do you know what experiences the guest had? We have to find another way to collect that information.”

Customer satisfaction surveys are widely used to gather the more subjective information about the guest experience, but surveys have their weaknesses too, both in project design and analysis. For example, how do you select the best sample of participants? What is the best form for the survey? Do you trust the responses? Some respondents will just click through an online survey with little thought, Verma said. Many responses must be adjusted or discarded for being nonsensical or internally inconsistent. At all stages of the survey process, you still have to deal with data quality issues.

Tips and techniques for improving data quality

Tackle data quality as a three-part initiative.

Analyze, improve and control. Analyze what data you need, what you’re lacking, which systems are providing the best information and which are most suspect, said Chastain. Include third-party systems and purchased databases in that analysis. Improve the ability to standardize data representations among systems. Information from a booking system should correlate to a loyalty card system that correlates to food and beverage – all tagged to a specific guest. Control data quality with governance measures that prevent dirty data from entering your system or data quality from being eroded over time.
Address both organizational and technology gaps.

On the organizational side, consider establishing a cross-functional, virtual team that establishes data governance and paves the way for integrating data from different business units. “The companies that are most successful see data quality as a joint venture between the business users, who understand the context of the data, and the IT professionals, who know how to architect a system to serve the business need,” said Chastain. “It calls for consistent collaboration, inevitably a give and take, to get it right.”

On the technology side, embrace the automated data integration and data quality tools that are available. Data warehousing, data marts and embedded data quality tools are widely used and proven,” said Chastain. “We see growing use of data quality tools applied to all sorts of data across the organization,” such as marketing, revenue management, HR, food and beverage, and reservations systems.

Establish a centralized data core with distributed analytical capabilities.

At the core of business intelligence is a central repository of trustworthy, analysis-ready data integrated from across the organization. But a central data warehouse doesn’t mean analysis has to be centralized. In fact, it is best to give some degree of autonomy to users across the organization, so they can get their own answers from that data.

Jarosz’s organization struck a balance, providing automated analysis to answer questions around common themes and metrics, and then allowing individuals to access those analysis capabilities at will – while preserving the integrity of the underlying data.

Collect, connect and correlate to create a unified view of your guests.

“To compete in this space, it’s important that you have all of the information available about your guests at hand and easily accessible,” said Jarosz. “So we left no stone unturned in the major makeover of all our data processes, from what data is collected, how it’s collected, how it is transformed, how it’s stored and how it’s accessible in our data warehouses.”

“We have linked or tagged the behavior that a guest has with us from their stays and their reservations to the number of times we’ve contacted them and how they have behaved through a campaign. All of that information feeds back into our warehouse, and we have a ‘single source of the truth’ about our guests.”
Apply automated data integration technologies.

Embedded data quality routines can eliminate much of the manual intervention performed today. Automated tools can cleanse and consolidate data, eliminating duplicate or erroneous records drawn from multiple operational systems. A complete suite of data integration capabilities can include data cleansing and enrichment; extraction, transformation and loading; data federation; master data management and more.

Think creatively about combining and creating data.

Organizations miss many opportunities to improve insight by combining data in new ways, Verma said. “For example, lots of companies have revenue management programs, where people are tracking pricing related to overbooking or occupancy and so forth, but they do not often combine revenue management from one part of the organization to the other. That’s one opportunity. Another example would be to better integrate archival data with created data to generate new insights.”

Create new data with “choice surveys,” which ask respondents to speculate on what they would do under given scenarios. For example, “At this price, with these amenities and this brand, would you stay at this hotel?” Choice surveys can help you understand the risks and returns associated with different variables of a project.

Start with an area where you can show proven success.

Seeing is believing. Quantifiable proof is better. “Look for an initiative where data can help drive the success of a project – one where you can monitor and manage data quality, and show a positive impact on the business initiative or objective of the project,” Chastain said. “Find a way to measure quality of information and show that it correlates to success.” Early successes in small areas will build momentum and support for broader data quality initiatives.

Commit to data quality as an ongoing, grass-roots effort.

Achieving data quality is not a one-shot deal. It requires sustained vigilance, Jarosz said. “We went through such a huge integration project not that long ago, so we don’t want to see the data warehouse go back to its former self. While the project itself had an executive sponsor, there has been a grass-roots movement within the organization to talk more about data quality and what it means to various parts of the business. It is part of our monthly sessions now – not something that was completed and gone away.”
Take advantage of a full range of available analytic techniques.

Organizations are only scratching the surface today, Verma said. In the last three decades, powerful predictive modeling techniques based on econometrics have been developed in PhD programs and are now embedded in commercially available software. Statistical snafus related to data integration have been managed. Sophisticated analytic techniques have been made accessible and user-friendly for business applications.

“The industry needs to look at this whole predictive modeling set of tools more carefully and implement them if they have not already, because real-world businesses are becoming a lot more complex,” said Verma. “They need to somehow incorporate these methods into day-to-day operations, which is not being done too much.”

Exploit the knowledge to be found in social media and networks.

Valuable insights can be gained from social media – comments about your business on Twitter, Facebook and similar sites, as well as from free-form comments in surveys, e-mails and customer records.

“The interesting thing is that much of this content is not triggered by an event or a specific request; it’s free-given sentiment,” Chastain said. “So we’re seeing the increasing use of this type of unstructured data to support analytics. Free-form text analytics have been around for some time, but applying this technology to social media is a new frontier.”

Savvy organizations are also looking at the power of networks to influence customer choices. “I may be making some choices, and if you are my friend [on a social network], you will be influenced by my choices,” said Verma. “It’s good to capture that circle of influence in analysis as well.”

Closing thoughts

Getting data quality right can be a challenge, but the benefits are worth it, said Jarosz. “At Fairmont, the actual data quality project ended about two-and-a-half years ago, but the benefits continue. We have a much better view of who our customer is. We can link history to a particular guest, their stays, their total spend, what they liked and disliked, and their travel preferences. We know who brings us the greatest value. We can access the information much more easily and make decisions readily. We know the impact of a campaign as soon as it is executed.”

Data quality is not exclusively an IT or a business function. No one owns the data, but everyone is accountable for it.

Tracey Jarosz
“The ROI we have seen so far has been both tangible and intangible,” Jarosz added. “Tangibly, we can now link ROI back to individual guests or guest segments and, intangibly, we continue to build a deeper understanding of our guests, ranging from what they’re doing with us to what we think they’ll do with us in the future.”

For more information

To get the benefits of data quality for your hospitality or gaming organization, find out more from the Cornell University Center for Hospitality Research. Hospitality professionals will find insights from the center’s research and industry collaboration, in publications, reports, news archives from other sources and more.

www.hotelschool.cornell.edu/research/chr/

For more about data quality and analytics software solutions, visit SAS on the Web at www.sas.com/industry/hospitality/. In addition to business intelligence solutions designed specifically for the hospitality and gaming industry, SAS provides full data integration, data quality and data management capabilities to resolve the data issues that are endemic in many organizations today.

This Webcast was the first in a seven-part series, Insights and Innovations for Hospitality and Gaming, sponsored by The Center for Hospitality Research at Cornell University’s School of Hotel Administration and SAS. Each Webcast highlights a hot topic in the hospitality and gaming industry, including data quality, labor planning, customer loyalty, sustainability and more.