FROM DATA TO ACTION

A HARVARD BUSINESS REVIEW INSIGHT CENTER REPORT

Harvard Business Review

Sponsored by

sas

FROM DATA TO ACTION

The HBR Insight Center highlights emerging thinking around today's most important business ideas. In this Insight Center, we'll take a closer look at what your data may or may not be telling you. Tapping researchers and practitioners in data science, marketing, and other fields, we'll offer guidelines on making good use of information and turning it into profitable behaviors.

YOU'VE GOT THE INFORMATION, BUT WHAT DOES IT MEAN? WELCOME TO "FROM DATA TO ACTION"

BY ANDREW O'CONNELL AND WALTER FRICK

"The numbers don't lie."

You hear that all the time. Even if it's mostly true, the numbers can be slippery, cryptic, and, at times, two-faced. Whether they represent findings about your customers, products, or employees, they can be maddeningly open to interpretation.

With big data getting ever bigger and infiltrating more and more parts of your business, the need for ways to understand the numbers' meaning is becoming more acute. As Tom Davenport of Babson College and MIT writes in the current issue of HBR, companies must "fundamentally rethink how the analysis of data can create value for themselves and their customers."

In this series, "From Data to Action," researchers and practitioners explore the fast-changing landscape of data, show how you can learn to tune out most of the noise to focus on the all-important signal, and translate your new knowledge into bottom-line improvements. Scott A. Neslin of Dartmouth's Tuck business school puts it perfectly when he says the numbers tell you a lot of stories simultaneously—the trick is figuring out which ones really matter.

Neslin's case in point concerns customers who haven't bought for a while—one set of data tells you they're not worth pursuing, while another says they are; which story is right? Susan Fournier of Boston University and Bob Rietveld of analytics firm Oxyme ask a related question: Is big data enough? Aggregate numbers can tell you a lot, but they say very little about how individual customers are thinking and talking about your products.

In a business-to-business context, Joël Le Bon of the University of Houston takes the question further, pointing out an often-overlooked, small-data source of useful information about your competitors' plans and schemes: your sales reps. He explains how to cultivate a strong flow of intelligence that can have a powerful impact on strategy.

In other articles, we'll move beyond specific cases to get at the broader question of how managers can learn to pick the most important kernels of knowledge from the rush of data. Is it critically important for companies' consumer-data teams to have someone who truly understands statistics in a leadership position, or can an untrained manager learn to be more discerning about data? And are there times when you should completely ignore the data and go with your gut?

After all, if companies often fail to analyze data in ways that enhance their understanding and neglect to make changes in response to new insights, as Jeanne W. Ross and Anne Quaadgras of MIT and Cynthia M. Beath of the University of Texas say in the current HBR, then what's the point of paying big money for big data?

While recognizing companies' struggles to make sense of information, Ross et al. argue that what really matters more than the type and quantity of the data is establishing a deep corporate culture of evidence-based decision making. That means establishing one undisputed source of performance data, giving all decision makers feedback in real time (or close to it), updating business rules in response to facts, and providing coaching for decision makers. It also means encouraging everyone in the organization to use data more effectively.

The sentiment about including everyone in the company in the use of data is echoed by McKinsey: When business-unit leaders invest in training for managers as well as end users, "pushing for the constant refinement of analytics tools, and tracking tool usage with new metrics," companies do eventually transform themselves, bringing data-driven thinking into all aspects of the business and fulfilling data's vast potential.

WHAT TO ASK YOUR "NUMBERS PEOPLE"

BY TOM DAVENPORT

If you're a manager working with the analysts in your organization to make more data-driven business decisions, asking good questions should be one of your top priorities. Many managers fear that asking questions will make them appear unintelligent about quantitative matters. However, if you ask the right kinds of questions, you can both appear knowledgeable and advance the likelihood of a good decision outcome.

In my new book (co-authored with Jinho Kim) *Keeping Up with the Quants*, and in a related article in this month's HBR, we list a lot of possible questions for various stages of analysis. But in this short article, I thought it might be useful to mention not only a couple of the most important questions you can ask about data, but also what some of the ensuing dialogue might involve.

1. Questions about assumptions

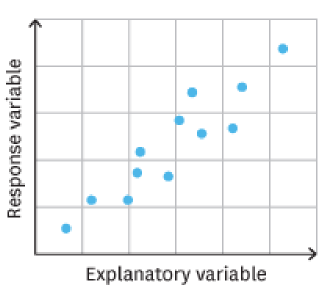
You ask: What are the assumptions behind the model you built?

You think in response to the answer: If your numbers person says there are no particular assumptions, you should worry—because every model has assumptions behind it. It may be only that you're assuming that the sample represents a population, or that the data gathered at a previous time are still representative of the current time.

Follow-up: Is there any reason to believe that those assumptions are no longer valid?

You think in response: You are really looking only for a thoughtful response here. The only way to know for sure about whether assumptions still hold is to do a different analysis on newly gathered data—which could be very expensive. Perhaps a particular relationship holds only when the values of a variable are moving in a particular direction (e.g., "This mortgage risk model holds true only when housing prices are going up"—nah, that could never change!).

SCATTERPLOT



2. Questions about data distribution

You ask: How are the data you gathered distributed?

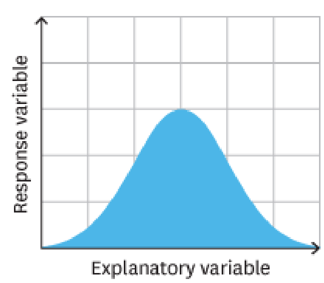
You think in response: If the person can't describe the distribution, he or she is a shoddy analyst. Good analysts should have already looked at—and should be able to show you—a visual display of the distribution of your data on any particular variable.

If you are interested in one variable as a likely predictor of another, ask for a "scatterplot" and look to see if the data line up in any linear pattern; that would indicate a strong correlation between the two variables.

Follow-up: Do the data follow a normal distribution?

You think in response: If the analyst says that the data aren't distributed normally (i.e., in a bell-shaped curve), then he or she needs to employ different types of statistics (called "nonparametric" statistics), and some commonly used ones like standard deviations and correlations don't apply.

NORMAL DISTRIBUTION



You might ask how the analyst adjusted his or her analysis based on the distribution. For example, nonparametric tests often require a larger number of cases for the same level of statistical confidence.

Second follow-up: Were there any significant outliers?

You think in response: If the data are normally distributed but there are some outliers (unexpected values that don't fit the pattern), you could ask what they might mean and what the analyst plans to do with them. In some cases it may be reasonable to delete outliers—if, for example, they are the result of coding errors.

You get the picture. It's important to show with this dialogue that you are interested, somewhat knowledgeable, and dedicated to a good decision outcome. You're not trying to suggest with such questions that you know more than the analyst or that the analyst is hiding anything from you. It's the same sort of conversation that a CEO might have with a division manager who is presenting financial results. Gentle probing is the desirable tone.

A Better Way to Tackle All That Data

The single biggest challenge any organization faces in a world awash in data is the time it takes to make a decision. We can amass all the data in the world, but if it doesn't help to save a life, allocate resources better, fund the organization, or avoid a crisis, what good is it? Hampered by a shortage of qualified data scientists to perform the work of analysis, big data's rise is outstripping our ability to perform analysis and reach conclusions fast enough.

At the root of this problem is our concept of what constitutes data. Existing boundaries of what we can digitize and analyze are moving outward every day. Taking Gartner's prediction that the Internet of Things (essentially, sensors that share data with the Internet) will add 50 billion machine voices to today's 2 billion connected users, we have to believe that the ability for humans to manage the process of amassing the right data and performing the right analysis is headed for trouble.

The measure of how long it takes analytics to reach a conclusion is often called "time to decision." If we accept that big data's Holy Grail is, as Randy Bean says in *Information Week*, better, faster decisions, we have to believe that as data continue to grow in volume, velocity, and variety, making management more complex and potentially slowing time to decision, something has to give.

This is a problem crying out for a solution that has long been in development but only recently has begun to become effective and economically feasible enough for widespread adoption—machine learning. As the term suggests, machine learning is a branch of computer science where algorithms learn from and react to data just as humans do. Machine-learning software identifies hidden patterns in data and uses those patterns both to group similar data and to make predictions. Each time new data are added and analyzed, the software gains a clearer view of data patterns and gets closer to making the optimal prediction or reaching a meaningful understanding.

It does this by turning the conventional data-mining practice on its head. Rather than scientists beginning with a (possibly biased) hypothesis that they then seek to confirm or disprove in a body of data, the machine starts with a definition of an ideal outcome, which it uses to decide what data matter and how they should factor into solving problems. The idea is that if we know the optimal way for something to operate, we can figure out exactly what to change in a suboptimal situation.

Thus, for example, a complex system like commuter train service has targets for the on-time, safe delivery of passengers that present an optimization problem in real time based on a variety of fluctuating variables, ranging from the weather to load size to even the availability and cost of energy. Machine-learning software onboard the trains themselves can take all these factors into account, running hundreds of calculations a second to direct an engineer to operate at the proper speed.

The Nest thermostat is a well-known example of machine learning applied to very local data. As people turn the dial on the Nest thermostat, it learns their temperature preferences and begins to manage the heating and cooling automatically, regardless of time of day and day of week. The system never stops learning, allowing people to continuously define the optimum.

The application of machine learning in health care is essential to achieving the goal of personalized medicine (the concept that every patient is subtly different and should be treated uniquely). Nowhere is this more easily seen than in cancer treatment, where genomic medicine is enabling highly customized therapy based on an individual's type of tumor and myriad other factors. Here machine-learning algorithms help sort the various treatments available to oncologists, classifying them by cost, efficacy, toxicity, and so forth. As patients are treated, these systems grow in intelligence, learning from outcomes and additional evidence-based guidelines. This leaves the oncologists free to focus on optimizing treatment plans and sharing information with their patients.

With the rise of off-the-shelf software, such as LIONsolver, the winner of a recent crowdsourcing contest to find better ways to recognize Parkinson's disease, machine learning is at last entering the mainstream, available to a wider variety of businesses than the likes of Yahoo, Google, and Facebook that first made big data headlines. More and more businesses may now see it as a viable alternative to addressing the rapid proliferation of data, with increasing numbers of data scientists spending more and more time analyzing data. Expect to see machine learning used to train supply chain systems, to predict weather, to spot fraud, and (especially in customer experience management) to help decide what variables and context matter for customer response to marketing.

FEATURED COMMENT FROM HBR.ORG

"Great article…I've been pondering the ideas and the comments. I am excited about the developments… Machine learning is really the next step." – Clay

STORY-DRIVEN DATA ANALYSIS

BY JUDY BAYER AND MARIE TAILLARD

Great analysts tell great stories based on the results of their analyses. Stories, after all, make results user-friendly, more conducive to decision making, and more persuasive.

But that is not the only reason to use stories. Time and time again in our experience, stories have been more than an afterthought; they have actually enabled a more rigorous analysis of data in the first place. Stories allow the analysts to construct a set of hypotheses and provide a map for investigating the data.

We recently worked with a department store retailer and a team of analysts looking for creative insights into customer loyalty. Based on our work with a department store expert, we started out with a storyline, a narrative hypothesis, according to which a customer experiences different journeys through the department store over time and rewards the retailer with a certain level of loyalty.

How will these journeys unfold? Does the customer start in cosmetics and then move into clothing? Does she go from the second floor to the first floor to buy a handbag to match a new outfit? Does she have shopping days where she takes a lunch break in the restaurant before continuing her shopping? Do less loyal customers make different journeys from more loyal ones?

In other words, we were interested not just in what customers were buying, but in the mechanics of how they make their purchases and how this may make them loyal. After the analysis, the true story of a customer's path to loyalty is in fact revealed.

Where do these stories come from? In our experience, they can come either from the experience of an expert in the sector or brand, as was the case in the previous example, or from qualitative research using observation or in-depth interviews.

We recently advised a telco client in developing the "jobs to be done" for a range of new products and services. We interviewed consumers and heard their own stories of how they go about using their mobile devices throughout the day. The general narrative hypothesis we drew from listening to these stories is that consumers cobble together mobile solutions to suit their lifestyles.

One consumer revealed that he actually owns two SIM cards for the same smartphone and told us in what context he changes from one to the other. Another customer told us about the parental control and other relevant apps and browsing that she has discovered and collected and that facilitate her lifestyle as a mother.

What we are seeing here is a multi-usage context (characterized by two SIM cards) and a "mobile mommy" context, each of which calls for a distinct analysis and possibly different products/services to be developed subsequently. In other words, we found that the customers' homemade solutions could be used by brand managers to identify what kind of data to gather and what kind of analysis to perform.

The analyses will in turn enrich the initial stories and lead to deeper insights. What is important here is that the storyline, told before the analyses, enables an authentic human element to surface that would be more difficult to extract from the data alone.

In order for a story to truly enable analytics, the story development process needs to be rigorous. We use the framework of grounded theory to ensure that the data and overarching storyline inform each other and are coherent with each other. The idea is for the analyst to navigate back and forth between the data and the developing story to ensure a good balance between the creative narrative and the analytics that reveal the facts and details of the story.

The enabling storyline should not be too restrictive; it needs to support the development of the plot and characters as they emerge from the analysis, but without bias. Conversely, the storyline can suggest specific questions to be asked of the data for a more in-depth analysis.

In a world that's flooded with data, it becomes harder to use the data; there's too much of it to make sense of unless you come to the data with an insight or hypothesis to test. Building stories provides a good framework in which to do that.

STOP ASSUMING YOUR DATA WILL BRING YOU RICHES

BY SUNAND MENON

"We have a treasure trove of data—it's highly valuable."

"If we can unlock the value of all our data, we will have a wholly new revenue stream."

"Hedge funds will love our data—they will buy practically any set of data that might give them a potential edge."

These are just a sample of thoughts from clients in recent months who were excited about the prospect of creating new businesses and new sources of revenue in what seems to be a lucrative new area: big data, analytics, and content innovation.

In some cases, they're right. But then again, in a surprising number of cases, they aren't. The opportunity, which seemed immense at the outset, turned out to be disappointingly smaller after a thorough evaluation. Fortunately, the organizations that took the effort to truly understand the value of their data were able to execute appropriately. Very importantly, they were able to avoid costly technology and implementation programs that would have surely fell short of usage and revenue expectations.

Here are four steps your organization can take in order to understand the value of your data, and to plan for potential monetization:

Clarify whether it's really your data

Sounds obvious, doesn't it? Unfortunately, this may be the most common mistake that organizations make. They assume that if they collect the data and house it in their systems, it must be their data. Or if they collect data and then add a proprietary methodology to it, that automatically qualifies it as proprietary data. Or if they create analytics from raw, underlying data and subsequently barter the analytics, they must be able to charge hard dollars at a later stage. All these assumptions may be true...but are more likely to be false. Unless you have written data contracts in place that clearly allocate ownership of data and derivative works, you may not be able to do anything with the data.

- To evaluate data ownership, enlist the early help of domain experts—content specialists and legal counsel who understand how data is created, stored, manipulated, packaged, distributed, and commercialized.
- Categorize the datasets you have identified into three buckets—data we own, data customers own, and data third parties own—to ensure added clarity.
- Quantify (as much as possible) the value-add of any derived data versus the original data, in order to be in a better position to create mutually agreeable data usage and revenue-share agreements with suppliers and co-creators of the data.

Understand who would value it, why, and how much

This is easier said than done. In speaking with over 100 end users of data over the past year (in financial services, health care, technology, and even nonprofits), I have come to realize that your target customer may not be your traditional customer. For example, a product evaluating nonprofit organizations may be highly useful to a wealth manager seeking to help his client to select a charity as part of a value-added tax-efficiency service. Or a health care data evaluating physicians' perceptions on a new drug may be useful not only to brand managers at pharmaceutical companies, but also to portfolio managers at asset management firms seeking to find promising investment opportunities.

- Identify target customers by casting a wide net across potential users, and perform customer interviews to establish their "jobs to be done" (as Harvard Business School professor Clay Christensen says). Determine where their current data gaps are, and evaluate whether your datasets can fill those gaps, either as raw data or in a more processed, analytical format.
- Test user perceptions with a range of potential data offerings. End users vary in their level of sophistication of data usage, and some may immediately see the value of the raw data, whereas others may want to be given visual examples of the value the data can bring. As one client once told me, "It's as if I were at a farmers' market with all the most amazing fruits and vegetables, and I can think of a hundred recipes that I would love to prepare." It's clear that he saw the potential of the raw data. Others may not wish to (or be able to) be the chef, and may prefer to be the diner at the restaurant, where they have a menu of dishes that they can choose from. Both are eventual consumers of the raw materials, and both should be served!
- Ask prospective customers to assign a gut-check ranking—High, Medium, or Low—to the individual datasets and metrics, and note these preferences respectively in green, orange, and yellow. End users' initial reactions are generally quite pragmatic and representative of their overall assessment of value, utility, and willingness to pay. As the customer interviews

progress, the color-coded matrix will come to life and can help prioritize where the opportunities truly lie.

Frame up realistic aspirations for monetization

At this point, many companies get to work and prepare detailed financial projections that show how many new sales they can achieve every year, what their proposed pricing is, and what the year-on-year percentage increase will be. Sometimes that makes sense. And sometimes it backfires—what if the monetization potential is not really that big? How do you ensure that your execution is commensurate with the revenue opportunity?

- Set yourself a target that is big enough for you to pursue and that you feel is worth your organization's time and effort. The innovation consulting firm Innosight often refers to "$50 million in five years" as a reasonable target for a brand-new business. For you, it may be different.
- Ask yourself, "What would it take to get to that target?" Data can be commercialized in a number of ways: via annual subscriptions, via once-off consulting and integration fees, via custom content development, via research and advisory services, and via new analytics development. Ask yourself how many of these are truly viable. Consider comparable offerings from competition and their pricing structure. In many cases, you may come to an epiphany that your target is just not reasonable.
- Understand whether the data provides more value than just the new stand-alone revenue. You may come to the conclusion that your original $100 million target was unrealistic, but $10 million is achievable. Do you decide to stop? Well, it depends. Sometimes the revenue may be small in absolute terms, but the data capability may be complementary to your current, core business. The value may be in the combination of the two, which can drive significantly higher core business revenues.

Test, learn, and tweak

Now that you have a realistic revenue aspiration and have decided to continue pursuing this opportunity, you turn to execution mode. If you are sure about the opportunity in front of you and your ability to execute, this may work. However, if you are new to the data and analytics game, and are not sure whether you can be successful due to a multitude of ambiguities, you may need a different approach.

- Highlight the areas in which you think you may fail, and create test programs to evaluate your ability to execute. Do you have the ability to provide real-time data 24/7? Will customers really pay what they alluded to in the interviews? Will your data distribution partner really be motivated to work with you, or will the partner have other priorities? These are all crucial parameters that will determine whether your data can really be commercialized. They need to be addressed and solved before launching a business in earnest!
- Create tangible success criteria that will allow you to determine whether you can solve the problem or can learn something that will help you make a go/no-go decision. For example, a test could be "Validate subscription business model via direct sales by securing three signed customer contracts within three months," or "Create dashboard of [specific number and type of] metrics that 100 end users test, validate, and give suggestions for improvement, within three months."
- Implement the test programs with defined roles and responsibilities for the test program owner as well as the execution team. Ideally, the team should be small, 100% dedicated to the pilots, and cherry-picked for their domain knowledge in content as well as their ability to work in an agile, entrepreneurial environment.

The results of the test programs can help get you a more informed view on whether you go ahead with implementation, stop, or need to make some modifications to your business model and/or execution.

In following this overall process, you can clarify what data you own and how valuable it is (and to whom). You can frame up realistic aspirations for monetizing your data, and you can prove your right to succeed by testing (and overcoming) areas of potential failure. You can therefore move from an unsubstantiated assumption about the value of your data to a more informed understanding of its worth in terms of its use to current and prospective customers, its stand-alone commercialization potential, and its potential to enhance your current business.

NATE SILVER ON FINDING A MENTOR, TEACHING YOURSELF STATISTICS, AND NOT SETTLING IN YOUR CAREER

BY WALTER FRICK

Perhaps no one has done more for the cause of data-driven decision making in the minds of the public than Nate Silver. His book, *The Signal and the Noise,* explains the power of statistical modeling to improve our predictions about everything from the weather to sports to the stock market. Data, and Silver is its poster child.

But for most people, the gulf between recognizing the importance of data and actually beginning to analyze it is massive. How do those without extensive training in statistics equip themselves with the skills necessary to thrive (or even just survive) in our age of "big data"?

Last month I had the chance to put that question to Silver, and his answers may surprise you. Far from counseling that everyone must major in statistics, in the edited conversation below he advises students and executives alike to roll up their sleeves—no matter what their statistical literacy—and get their hands dirty with data.

HBR: If I'm an average professional or an executive, and I've read your book, I know this stuff matters—and I also know it's complicated, and I can expect only so much. Is there such a thing as kind of a level of statistical literacy that I need to get to? What kind of education do I have to go back and make sure that I have?

Silver: I think the best training is almost always going to be hands-on training. In some ways the book is fairly abstract, partly because you're trying to look at a lot of different fields. You're trying not to make crazy generalizations across too many spheres.

But my experience is all working with baseball data, or learning game theory because you want to be better at poker, right? Or [you] want to build better election models because you're curious and you think the current products out there aren't as strong as they could be. So, getting your hands dirty with the data set is, I think, far and away better than spending too much time doing reading and so forth.

HBR: What about if I've read your book and I'm just starting college or a little younger, and I'm trying to think, actually, maybe this statistician/data scientist role is something that I'm interested in? What do I study? How much education do I need? What's that base for plugging into some of these jobs?

Silver: Again, I think the applied experience is a lot more important than the academic experience. It probably can't hurt to take a stats class in college.

But it really is something that requires a lot of different parts of your brain. I mean, the thing that's toughest to teach is the intuition for what are big questions to ask. That intellectual curiosity. That bullshit detector, for lack of a better term, where you see a data set and you have at least a first approach on how much signal there is there. That can help to make you a lot more efficient.

That stuff is kind of hard to teach through book learning. So it's by experience. I would be an advocate [of]—if you're going to have an education—then have it be a pretty diverse education so you're flexing lots of different muscles.

You can learn the technical skills later on, and you'll be more motivated to learn more of the technical skills when you have some problem you're trying to solve or some financial incentive to do so. So I think not specializing too early is important.

HBR: Say you're at the point where you started playing around with some data. You're interested, you're motivated, and now it's time to actually learn some of those skills just like you talked about. Am I just going and picking up a textbook? Am I trying an online course?

Silver: I mean, my path has been kind of sui generis in some ways, right? Probably an online course could work, but I think actually when people are self-taught with occasional guidance, with occasional pushes here and there, that could work well.

An ideal situation is when you're studying on your own and maybe you have some type of mentor who you talk to now and then. You should be alert that you're going to make some dumb mistakes at first. And some will take a one-time correction. Others will take a lifetime to learn. But yes, people who are motivated on their own, I think, are always going to do better than people who are fed a diet of things.

HBR: Say an organization brings in a bunch of "stat heads," to use your terminology. Do you silo them in their own department that serves the rest of the company? Or is it important to make sure that

every team has someone who has the analytic toolkit to pair with expertise?

Silver: I think you want to integrate it as much as possible. That means that they're going to have some business skills too, right? And learn that presenting their work is important. But you need it to be integrated into the fabric of the organization.

You've seen this shift in baseball teams, for example, where it used to be that you'd hire an analyst to check that box and have them compartmentalize. That doesn't accomplish much at all.

HBR: You've had, obviously, some very public experience with the fact that even when the data is good and the model is good, people can push back a lot for various reasons, legitimate and otherwise. Any advice for once you're in that position—[once] you have a seat at the table, but the other people around the table are really just not buying what you're selling?

Silver: If you can't present your ideas to at least a modestly larger audience, then it's not going to do you very much good. Einstein supposedly said, "I don't trust any physics theory that can't be explained to a 10-year-old." A lot of times the intuitions behind things aren't really all that complicated. In *Moneyball*, [the fact] that on-base percentage is better than batting average, looks like "OK, well, the goal is to score runs. The first step in scoring runs is getting on base, so let's have a statistic that measures getting on base instead of just one type of getting on base." Not that hard a battle to fight.

Now, if you feel like you're expressing yourself and getting the gist of something, and you're still not being listened to, then maybe it's time to change careers. It is the case [that] people who have analytic talent are very much in demand right now across a lot of fields, so people can afford to be picky to an extent.

Don't take a job where you feel bored. If it's challenging, you feel like you're growing, you have good internal debates, that's fine. Some friction can be healthy. But if you feel like you're not being listened to, then you're going to just want to slit your wrists after too much longer. It's time to move on.

HBR: What about, from the perspective of an organization or a business, knowing those areas where data is really going to be the key to making good predictions and good decisions versus those areas where it isn't? Speaking to a lot of start-ups and tech companies, you hear, "Data can't tell us anything. The future is so different from the past, and we really can't rely on it at all, so it's really an intuition game."

Silver: A lot of times when data isn't very reliable, intuition isn't very reliable either. The problem is, people see it as an either-or, when it sometimes is both-or-neither as well. The question should be, how good is a model relative to our spitball, gut-feel approach? And also, how much do we know about this problem? There are some issues where you just don't have a good answer, and you have to hedge your risks as a business and not pretend that you're more certain than you really are.

A lot of private businesses are very reluctant to deal with uncertainty in their outlook. The manager doesn't want to seem like he's not sure what he's doing. And the consultant or the analyst wants to provide information to make the manager feel more confident. That's quite problematic, because a lot of problems that are on the frontier of business, on the frontier of science, [are] by definition fairly challenging ones that no one else has solved.

That's where having a more humble attitude about what you can accomplish and what you can't is important. Just because a model is not going to be very precise or accurate doesn't mean that therefore you should trust your gut instinct after a couple of whiskeys and assume it's going to be very much better.

GOOGLE ON LAUNCHING AN ANALYTICS MOOC AND TAKING DATA-DRIVEN ACTIONS

BY MAUREEN HOCH

Analytics. It could be the deepest, darkest mystery in your organization, served up by a few, select tech wizards, or it might be the solitary master by which all company decisions begin and end. In all likelihood, it's somewhere in between.

No matter where you fit on the spectrum of analytics know-how, data-driven decision making is here to stay. With millions of users, Google Analytics (GA) is among the tools well entrenched in paving the road to actionable data. And a business's choices when it comes to analytics services are many, with one tool rarely being a one-stop solution. (For full disclosure, GA is one of several analytics tools HBR uses to parse its online data.) So not only do we have to learn how to use analytics technology, but we also need to become more data fluent and confident in how to go from a collection of insights to action. And we have to do all this while maintaining an empathetic connection to our users and customers.

For some perspective on this challenge, we talked to Paul Muret, engineering vice president for Google, and Babak Pahlavan, product management director for Google Analytics. Muret is known as the "father" of GA, having founded Urchin, which was acquired by Google in 2005 and helped build the analytics tool we know today. Pahlavan is the founder of Clever Sense, a marketing data tool also acquired by Google.

When we talked, Muret and Pahlavan were prepping for the Google Analytics summit happening on October 1, where they are announcing a type of massive online course, or MOOC, that will allow anyone to learn the fundamentals of analytics, among other new initiatives. Below is an edited version of our conversation.

Everyone wants to talk big data right now. How do you define the difference between big data and analytics?

MURET: I think it's easy to have this "big data" term mean a lot of different things. Some people think about just trying to collect so much information from all kinds of different places. The key difference is making sure the data is useful and accessible by the people in your organization.

We have all these analytics tools—and Google Analytics is a tool that a lot of people are using, and they're getting a lot of data points out of it—but how do you really move to action or to making decisions?

MURET: I think that it's easy to just go directly into the tools and technology and lose sight of the big picture. The reality is that companies all over the world are using data to make smarter business decisions and drive creativity and innovation, and it's having a huge impact on their bottom line.

In the last few years, with the amazing advances in technology, especially the growth in communication networks and mobile devices, consumers are in this state of being constantly connected. And this is having a huge impact on every market and creating an opportunity in every vertical where consumer decision making and purchasing decisions are happening, which is not just in your brick-and-mortar stores anymore.

A decade or less ago, it would be very easy to see our customers and understand them. We could actually see them physically walking into our stores and doing their research, making their decisions, and you could see what they look like and what they're looking at. There's so much information you can gain by seeing your customers. It gives you this intuitive understanding of who they are and how to engage with them.

But imagine now moving into the era where that's all happening online. It's like running the store with the lights out. And if you can't see your customers, you've risked reducing them down to sort of bits of data, URLs, and JavaScript events. You have to learn how to engage with that data. It's incredibly important now to empower everyone in your organization with data, and that goes for the CEO and the CMO on down. We want the service managers, the user experience designers, and the product managers. But that means the data needs to be accessible.

One of the announcements that we'll make at the Google Analytics Summit is that we're launching a new analytics academy. This is a rich media, interactive, massive online course that everyone can access to learn more about digital marketing, digital analytics, and Google Analytics, and how to put these tools into practice. We're educating everybody about these techniques and helping answer their questions so they can move forward with making decisions.

So is it a MOOC? Or is it more of a resource tool that people will dip in and out of?

MURET: It is a MOOC, but there are two modes of it. You can use it in a self-service way as well. There will be a combination of videos and Google Plus Hangouts and online community resources all together, with actual certification steps along the way.

The first classes will be taught by our key digital marketing thought leaders and evangelists. In your organization, you might have one or two analysts [who] are kind of experts on Google Analytics, but very quickly their job becomes trying to sort of quarterback and educate the rest of your organization. There will be some areas that will require that kind of level of sophistication. But a lot of these techniques are incredibly accessible. The data can be used by basically the whole organization.

It sounds like almost everyone working for a company today is going to have to be a data analyst of sorts.

PAHLAVAN: In this new world, you will have a much better business if your decisions are data driven. In order for it to be data driven, you have to be empowered with tools that are easy to use but also powerful enough that they can actually lead you to proper decisions. With regards to the MOOC side of the story, I would say this is a bit of a radical investment in our side. We're leveraging a lot of technology, and we'll have our best education leaders and Google experts to teach this course. There's going to be a lot of collaboration, a lot of discussions. We are expecting thousands of people to sign up.

I don't think there are many other products out there [into which] they're putting this much investment insuring that technology can be accessible [so] other organizations…can learn how to use them properly.

MURET: I do think that it's going to be important for organizations to have a certain level of data fluency. If you think about Excel and spreadsheet technology, when it first started, spreadsheets seemed like a really scary thing and people weren't sure what to do with them. Now today, people can write macros and have 15 spreadsheets doing all these crazy things. But for most cases, you don't need that. You just need some basic math background and some basic things in order to sort of use spreadsheets in a way that's really useful.

There are some basic concepts that people need to be able to understand so that they're not misinterpreting things. There will be some areas, certainly, when it comes to analyzing data that are subtle, that will require experts; but for the most part, we think this data can be made incredibly accessible to a very broad set of people in an organization.

PAHLAVAN: It used to be that it was all about the website that you had. But now the consumers are on phones 24/7 and their tablets. So if we don't get in there right now and help companies have access to great tools, but also know how to use them properly, people are not going to tap into these data-driven opportunities.

We feel like it's our responsibility to (a) make simple but powerful products, and (b) try to support and educate people to have data fluency. On top of it, people want to use these things. They're saying, "Teach us the best way of using this."

There's still this question of taking action. What types of decisions should you be looking for?

MURET: There are two types of ways to take action here. The first type I like to call "aggregate actions," and the second one I think about as "automation."

The first one, aggregate actions, is sets of data over time. So a simple example would be if you have two landing pages. Let's say you've got two offers—offer A and offer B—and you test them both and say, "Hey, offer A is working better than offer B." Then you make a decision, and you go with offer A and you remove offer B from your content. That's a process that's very straightforward, and it's basically an aggregate decision.

But there's another way of using the data, and that's to take the data itself and put it back into these systems in a real-time way. That's because the reality isn't inside these aggregate numbers. You'll have pockets of users who respond to different kinds of messages. There's often an opportunity inside one of these areas to be more specific and provide more tailored information directly to specific users. And that needs to be done in a more automated way.

It's not something that you do as the analyst—I'm sure you'd love to be able to make those decisions one by one—but that's too hard to do. Maybe it turns out that people coming from the southern United States love offer B and that seems to work better there. We need to be able to make that decision in milliseconds. There's often a way of putting data back into action in an automated fashion to drive a more automated marketing platform.

OK, but if our customers are becoming bits and bytes, how are leaders and decision makers still going to build empathy for those customers?

MURET: I think it's a challenge for all organizations going forward to figure this out. But one of the key ways is going to be through the data that we're talking about. It's interesting when people say, "OK, I've got a bunch of data. Give me insights." That's not really building that empathy you're talking about. But once you're trying to optimize and analyze a specific part of your business, then it's through that process that you gain insights into what is working under the hood. I've just learned, wow, the way people are actually doing this is much different than what we thought. And that starts to build that empathy back together.

PAHLAVAN: It's a very good question, the notion of empathy. Are we creating a situation in which the business leaders and business as a whole—are they going to be more empathetic to their customers and focus on their needs versus going to just look at them as more like aggregate formats and say, group them into high-level buckets? We look at it from a perspective of, can we provide you with tools with a more granular set of users and figure out what is it that they need? What is it that they're interested in? What are their demographics? Can you put the right set of products or content in front of the right set of users, or not?

MURET: We want to give a very practical approach so that it drives returns almost immediately. But then as part of that process, when you're going through those steps, we are effectively helping your organization put back together that picture of the customer. And that's where the empathy hopefully is going to start to build back together so that your organization can make creative jumps in thinking.

CAN YOUR C-SUITE HANDLE BIG DATA?

BY BRAD BROWN

Over the past 30 years, most companies have added new C-level roles in response to changing business environments. The chief financial officer (CFO) role rose to prominence in the mid-1980s as pressures for value management and more transparent investor relations gained traction. Adding a chief marketing officer (CMO) became crucial as new channels and media raised the complexity of brand building, while chief strategy officers (CSOs) joined top teams to help grapple with complex and fast-changing global markets.

Today, as the power of data and analytics profoundly alters the business landscape, companies once again may need more top-management muscle. Capturing data-related opportunities to improve revenues, boost productivity, and create entirely new businesses puts new demands on companies—requiring not only new talent and investments in information infrastructure, but also significant changes in mind-sets and frontline training. It's becoming apparent that it will take extra executive horsepower to navigate new organizational hazards, make tough trade-offs, and muster authority when decision rights conflict in the new environment.

Because the new data-analytics horizons typically span a range of functions, including marketing, risk, and operations, the C-suite evolution may take a variety of paths. In some cases, the way forward will be to enhance the mandate of (and provide new forms of support for) the chief information, marketing, strategy, or risk officer. Other companies may need to add new roles, such as a chief data officer, chief technical officer, or chief analytics officer, to head up centers of analytics excellence.

Six top-team tasks

The transformative nature of these changes involves much more than just serving up data to an external provider to mine for hidden trends. Rather, it requires concerted action that falls into six categories. Leaders should take their full measure before assigning responsibilities or creating roles.

Establishing new mind-sets. Senior teams embarking on this journey need both to acquire a knowledge of data analytics so they can understand what's rapidly becoming feasible, and then [to] push durable behavioral changes through the organization with the question "Where could data analytics deliver quantum leaps in performance?" This exercise should take place within each significant business unit or function and should be led by a senior executive with the influence and authority to inspire action.

Defining a data-analytics strategy. Like any new business opportunity, data analytics will underdeliver on its potential without a clear strategy and well-articulated initiatives and benchmarks for success. Many companies falter in this area, either because no one on the top team is explicitly charged with drafting a plan or because there isn't enough discussion or time devoted to getting alignment on priorities.

Determining what to build, purchase, borrow, or rent. The authority and experience of a senior leader are needed to guide the strategic trade-offs involved when assembling data and building the advanced-analytics models for improved performance. The resource demands often are considerable, and with multitudes of external vendors now able to provide core data, models, and tools, top-management experience is needed to work through "build versus buy" trade-offs.

Securing analytics expertise. The new environment also requires management skills to engage growing numbers of deep statistical experts who create the predictive or optimization models that will underwrite growth. The hunt for such talent is taking place in the world's hottest market for advanced skills. Retaining these employees and then getting them to connect with business leaders to make a real difference is a true top-management task.

Mobilizing resources. Companies often are surprised by the arduous management effort involved in mobilizing human and capital resources across many functions and businesses to create new decision-support tools and help frontline managers exploit analytics models. Success requires getting a diverse group of managers to coalesce around change—breaking down barriers across a wide phalanx of IT, business-lines, analytics, and training experts. The possibility of failure is high when companies don't commit senior leadership.

Building frontline capabilities. The sophisticated analytics that data scientists devise must be embedded in tools that engage managers and frontline employees on a daily basis. The scale and scope of this adoption effort—which involves formal training, coaching, and metrics—shouldn't be downplayed. In our experience, many companies spend 90 percent of their investment on building models and only 10 percent on frontline usage, when in fact closer to half of the analytics investment should go to the front lines. Here, again, we have seen plenty of cases where no one on the top team assumed responsibility for sustained ground-level change, and efforts fizzled.

Putting leadership capacity where it's needed

In sizing these challenges, most companies will find they need more executive capacity. That leaves important decisions about where the new roles will be located and how to draw new lines of authority. Our experience shows that companies can make a strong case for leading their data-analytics strategies centrally when there's a strong company, a wide set of data assets to exploit, or a potent functional group such as marketing or finance with strong talent that spearheads value creation. Sometimes a formal, centralized data-analytics center of excellence may be needed to launch or accelerate a data-analytics initiative. Importantly, however, frontline activities (mobilizing resources, building capabilities) will need to take place at the business-unit or functional level, since priorities for using data analytics to increase revenues and productivity will differ by business. And just as critically, companies will best catalyze frontline change when they connect it with core operations and management priorities and reinforce it with clear metrics and targets.

A starting point for thinking through issues like these is for top teams, and probably board members as well, to develop a better understanding of the scale of what's needed to ensure data-analytics success. Then they must notch these responsibilities against their existing management capacity in a way that's sensitive to the organization's core sources of value and that meshes with existing structures.

ARE YOU READY FOR A CHIEF DATA OFFICER?

BY THOMAS C. REDMAN

By now, everyone knows that there is huge potential in data. From reducing costs by improving quality, to enriching existing products and services by informationalizing them, to innovating by bringing big data and analytics to bear, data are showing themselves worthy across the board. Still, data do not give up their secrets easily. And most readily admit their organizations are unfit for data.

With all the excitement (and anxiety) surrounding big data and advanced analytics, it is not surprising that many organizations are naming a chief data officer (CDO) to manage their data needs. But I fear that too many organizations have sold the title short and are missing the opportunity to define a truly transformative role. You must be ready for a journey that, sooner or later, will touch every department, job, and person.

You don't need a CDO to put basic data management capabilities in place. Too many companies are a full generation behind, and acquiring those capabilities in a hurry may be necessary. But it does not demand a CDO.

You shouldn't need a CDO to improve regulatory reporting. This is not to say that you don't need to improve regulatory reporting—simply that it shouldn't take a true C-band effort to lead the data work needed to do so.

You don't need a CDO to ensure that your IT systems talk to one another. Nor do you need a CDO to help analytics take root throughout the company. Both tasks are important, but they don't demand a CDO.

A company only needs a CDO when it is ready to fully consider how it wishes to compete with data over the long term and start to build the organizational capabilities it will need to do so. You need to be ready to charge him or her with fully exploring what it takes to compete with data. To gain some real end-to-end experience—perhaps making a concerted effort to try out advanced analytics in the hiring process, improve content in financial reporting, bring more data to decisions made by the senior team, and improve the quality of data used in marketing. To conduct enough tightly focused trials, free from the usual day-in, day-out encumbrances, so you can compete in a comprehensive fashion and at lightning speed.

The obvious question, "How do we use data to enhance our existing plans?" is not the critical one. If you're still asking this, you're not ready for a CDO. Rather, the most important questions are "How does data allow us to do things that we've not thought we could do?" and "What if a competitor, perhaps one we hadn't thought of as a competitor before, gets there first?" The distinction here is the ruthless determination to sort out which are part of a companywide strategy and how they fit with and accelerate other strategies.

It is important to think through organizational implications. Even the most basic data strategy, if well executed, will require new people, new structures, and new thinking. Just consider: If you seek a string of profitable innovations, you probably need data scientists, to set up both a data lab and a data factory, and the stomach to deal with the messiness of innovation. Similarly, becoming data-driven involves the deepest cultural change that is fundamentally incompatible with a command-and-control management style. Silos are the mortal enemies of the data sharing needed to effect this strategy.

You're ready for a CDO only if you're ready to develop a stone-cold sober evaluation of what it will take to succeed with your selected strategy and make some tough choices—when you expect him or her to drive something like the following:

Building a data-driven culture is probably our best bet. It is consistent with our values. There's no kidding ourselves—we have a long way to go. Training and hiring are critical, so we need HR to step up its game.

We have to pay a lot of attention to big data. But in today's environment, we probably cannot attract and retain the PhDs it would take to be on the leading edge. Let's keep an eye on our competitors, especially the small fry, and partner up with a couple of leading universities and analytics companies. We need to aggressively import ideas from those on the leading edge.

We need to work aggressively on quality. We saved big-time in the areas where we focused. People don't use data they don't trust. So we can't be data-driven without it. Our Six Sigma program is decent enough, and our first choice is to leverage that effort for data.

Finally, you're ready for a CDO only when you have the courage to act. Seeing the opportunity that such a statement offers and seizing it are very different things. It's plain enough that everyone makes decisions, just as everyone creates and uses data and so can impact quality. It will take real courage, over a long period, to drive data into every nook and cranny of the organization.

I've argued elsewhere that a full-on data revolution will, in time, change every industry, every company, every department, and every job. So there is some urgency here. Still, for most, when it comes to hiring a CDO, “go slow to go fast” is the right approach. Go slow in reserving the title for a long-term, transformative role. You’ll go faster because you’ll not straitjacket your thinking to the issues of the day.

FEATURED COMMENT FROM HBR.ORG

“Thanks for the really excellent blog, Thomas! The strategy to create a CDO indeed seems to be an organizational readiness one. You [Redman] did a great job highlighting some of the challenges involved.”
—Emily Smith

DON'T LET DATA PARALYSIS STAND BETWEEN YOU AND YOUR CUSTOMERS

BY HARALD FANDERL

Back when there were a handful of channels, interactions between customers and brands were relatively simple. Today, by contrast, more than half of all customers move through three or more channels to complete a single task.

To open a bank account today, for instance, a typical customer embarks on a multichannel journey: researching online, downloading an application, speaking to a call center agent, linking brokerage accounts, visiting a branch, and installing the bank's mobile app. Those steps leave a long and complex digital trail. That multichannel complexity, combined with the scale of the data—U.S. companies store at least 150 terabytes of it—makes divining insights into customer behaviors a serious challenge.

Parsing this data, however, is critical to improving the customer experience and growing your business. In our experience, the most productive way to get there is not by fixing individual touchpoints but by improving the entire customer journey—the series of customer interactions with a brand needed to accomplish a task. (For more, see the HBR article "The Truth About Customer Experience.") McKinsey analysis finds that companies acting on journey insights have seen a 15-20% reduction in repeat service visits, a 10-20% boost in cross-selling, and a drop of 10-25 basis points in churn.

To put big data to work in improving the customer journeys, companies should keep three things in mind:

1. Focus on the top journeys. Companies may feel they need to study all the bits and bytes available to them. Our analysis across industries shows, however, that three to five journeys matter most to customers and the bottom line. They generally include some combination of sales and onboarding; one or two key servicing issues; (1) moving and account renewal, and (2) fraud, billing, and payments. Narrowing the focus to those journeys allows companies to cut through the data clutter and prioritize.

For instance, a cable television player used advanced data analysis of multichannel customer behaviors to focus on where drop-offs in the journey occurred in two journeys—onboarding and problem resolution—to address nagging customer retention and loyalty issues. The data team helped them identify key service troublespots and ways to improve the onboarding process. Those insights led to several policy changes, including creating a "learning lab" that effectively operated as a mini-company to trial and refine new approaches. The changes improved customer satisfaction scores by more than 20%.

2. Don't wait for the data to be perfect. Companies often hesitate to take action for fear their data is missing or a mess. In our experience, however, successful organizations tend not to overthink all the details and instead just roll up sleeves and get to work. Most companies, in fact, already have the data they need. The challenge is pulling the data together.

Companies need to figure out where that data is stored, and what it takes to extract and aggregate it so they can understand the customer journey across multiple touch points. Since data often lives in systems managed by various functions, bring the necessary operations, IT, in-store sales, and marketing people together to identify the touchpoints. We've seen companies create small SWAT teams from across functions to break through bureaucratic logjams. Track performance from the outset, mistakes and all, since that experience helps teams test, refine, and learn and ultimately accelerate the benefits.

In one example, a leading European energy company generated a lot of data, but most of it was siloed within the web team, call center, and marketing functions. As a result, key insights were falling through the cracks. Using data the company already had, a marketing and operations team came together to analyze the journey customers took when they changed addresses. Looking into data patterns, the team found the moving process alone accounted for 30-40% of all churn. Customers were canceling their old accounts and not renewing at their new address.

In response, the company chose to manage customer expectations better along the core touchpoints of the journey. It fine-tuned its communications, providing a set of easy-to-follow instructions on its website with links that made setting up a new contract a matter of clicking a few buttons. That reduced churn by 40% and increased upsell opportunities throughout the journey.

3. Focus on analytics, not reporting. Companies tend to focus on generating reports from their data about what has happened. Much greater value, however, comes from analyzing data to pinpoint cause and effect and make predictions.

One bank, for instance, was looking for ways to use big data to spot early indications of loss risk in its small business lending and service operations. Touchpoint data revealed subtle changes in customer behavior that raised questions in the fraud team's mind. It was only when the team connected the dots across touchpoints, however, that the bank discovered behavior patterns that highly correlated with imminent risk of default. These included changed behaviors in online account checking frequency, number and type of call-center inquiries and branch visits, and credit line use. Analyzing those complex patterns allowed the bank to develop an early warning system that flagged high-risk customers.

Big data harbors big opportunities to improve customer journeys and value. What it requires is a commitment to focus on what really matters.

DOES YOUR COMPANY ACTUALLY NEED DATA VISUALIZATION?

BY BILL SHANDER

Data visualization—and I hate to admit this, because I make my living from it—is not for everybody. My client work and research suggest that, loosely speaking, organizations selling ideas rather than products stand to benefit the most. Frankly, with few exceptions, if this doesn't describe your organization, then don't bother.

Data visualization can be expensive, especially if it involves large amounts of data and complex algorithms or deep interactive experiences. So how do you decide whether your company should invest in it? If you're selling straightforward solutions to simple problems, data visualization is probably not worth the money. For example, consumer packaged goods firms Coca-Cola and Nestle don't need interactive graphics to explain their products, just as *Playboy* and *Playgirl* don't need to educate the opposite sex much about their centerfolds. Such products more or less speak for themselves. (Now here's the key caveat: Sometimes product companies do have more complex and nuanced stories to tell, in which case they should be using visualization. For instance, I would love to see a data visualization explaining the environmental benefits of Coca-Cola's PlantBottle sustainable packaging.)

On the other hand, if you're selling a complex answer to a complex problem, you should be embracing data visualization with gusto. Nongovernmental organizations, charitable and advocacy groups, and publishers have wisely jumped on board. Financial services companies have a myriad of offerings helping you see where your money is going, and companies like General Electric are devoting entire websites to visualize their data. Professional services firms are also hopping on board, offering online tools for digging into research results and making them meaningful.

Explaining a complex idea to an online audience requires a level of personalization, detail, nuance, and openness that only an interactive visualization can provide. ("Keep on Charting," *New York Times*.) These organizations need to let their audience play with their data to make their findings more useful and convincing. That, of course, is the bottom line: Such experiences have to give your users value, hopefully resulting in returned value to your organization in the form of sales and referrals.

Data visualization can produce big benefits, some of which are subtle yet powerful. One of the biggest benefits is personalization—i.e., enabling potential customers to estimate the value of a complex solution to their challenges. In 2010, BCG and the World Economic Forum wanted to demonstrate the economic benefits of employee wellness programs. They created an interactive graphic that lets executives calculate in 90 seconds roughly how much their companies could save by instituting such a program (based on the statistical data of other companies' wellness programs). After filling in a few blanks—number of employees; how many are in the U.S., Europe, and Asia-Pacific; and average regional salaries—the interactive graphic produces a series of five-year projections for costs and potential savings from wellness intervention programs. After it was introduced at the World Economic Forum meeting in Davos to great success, another partner company on the project has continued to use it as a sales tool to model similar savings opportunities for entire cities.

Another way data visualization can help is by enlisting your audience to do the analysis you don't have the time, manpower, or editorial space to do yourself. Booz & Company has done this with mergers and acquisitions data. It published an interactive data visualization graphic that lets companies determine how easy (or hard) it is in eight major sectors to make acquisitions that enhance shareholder value. Published alongside an article in its thought leadership magazine *strategy + business*, this interactive graphic revealed insights and details on all 300-plus deals (which couldn't have been included in the article, of course). Booz & Company's web analytics show that the viewers of this graphic stay longer on the website, engage more with the content, and come back to it over a longer time period than its average website viewer does.

Companies like Booz & Company have studies with deep statistical data—more data than they have the time to chart, analyze, and write about in their reports. Interactive graphics (when designed well) shift some of the data crunching and analysis to viewers, who want a layer of depth that the company can't provide through traditional publishing.

Data visualization can also help organizations whose products or services involve numerous steps to implement or many comparisons to consider that may differ from person to person. For example, The Robert Wood Johnson Foundation recently sponsored a data visualization challenge, asking entrants to visualize hospital pricing data to bring transparency to a highly opaque market. An article summarizing this data would have to concentrate on statistical averages and give at most one or two examples. But in an interactive experience, you can click into any of the 300-plus regions of the U.S. and look up pricing averages and per-hospital pricing for 100 conditions. In other words, you can quickly and easily browse over 165,000 rows of data to find out the best place to get a procedure done based on price, quality, and experience at over 3,000 hospitals across the country.

Lastly, data visualization can also help organizations whose solutions go against conventional wisdom. Allowing your online viewers to plot the data lets them come up with the big "ahas" themselves. For instance, using the hospital pricing visualization discussed above, you will find that pricing for most medical procedures in Boston, despite the city having among the top median incomes in the country and a reputation for some of the best hospitals as well, is actually well below average. While an article on the subject might have revealed this, unless Boston were the focus of the piece, odds are that it would not have. Would an article reveal that Birmingham, Alabama, has eight of the top 100 hospitals in the country that treat kidney infections, measured by overall quality scores from Centers for Medicaid and Medicare Services? Perhaps not. And a Birmingham journalist, coming across this interactive, might naturally examine local pricing, discover this gem, and write about it, spreading the word beyond the initial reach of the organization—all because of insights that would have otherwise been buried.

Companies selling complex solutions to complex problems should embrace the power of data visualization. Marketing and sales executives need to decide early whether their companies need it, because the learning curve is steep. And getting really good at it takes time, skill, and money—for technology, training, and high salaries to attract professionals with currently rare skills. For some companies, the investment is worth it. The rest should avoid the bandwagon altogether.

TO AVOID THE CUSTOMER RECENCY TRAP, LISTEN TO THE DATA

BY SCOTT A. NESLIN

A lot of stories emerge from customer data. The trick is figuring out which story to listen to.

Companies that are planning marketing campaigns typically look at their data on past responses to figure out which customers are most likely to react favorably, and they spend their marketing dollars accordingly. That makes perfect sense—it's obviously more profitable to send an expensive catalog, for instance, to just those people who are the likeliest responders.

But this type of thinking can lead you into the *recency trap*.

It's well known that a customer's likelihood of buying declines as time passes since her last purchase. The effect can be dramatic, as you can see from the exhibit, which is based on data from a meal-preparation service provider in a study I conducted with Gail Ayala Taylor of Dartmouth, Kimberly D. Grantham of the University of Georgia, and Kimberly R. McNeil of North Carolina A&T State University.

Because a customer who hasn't bought in six months, say, is an unlikely prospect, you don't target her; receiving no communication from you, she continues to stay away. As this process continues, eventually the customer's chance of buying falls virtually to zero—and the customer is lost to the company. You've slid down the curve into the recency trap.

However, even if the customer hasn't purchased in a long time, it could be worth investing in her. By sending her that expensive catalog, you may prompt her to buy something now. She then moves to the left along the curve and is more likely to buy again.

So which story is the right one—the one that tells you not to waste your money on a lapsed customer, or the somewhat counterintuitive one that tells you to invest in her?

In the case of the company we studied, it was clear that the extra marketing expense would be beneficial. Our mathematical model, which is based on a well-accepted understanding of the effect of recency of purchase on likelihood to buy, demonstrated that the company could increase customer value by hundreds of dollars per customer. The value of customers who otherwise would have been lost to the company could increase from nearly zero to roughly $150 per person, on average.

GOING, GOING, GONE

As time-since-last-purchase increases, customers become less likely to buy again.

SOURCE S.A. NESLIN, G.A. TAYLOR, K.D. GRANTHAM, AND K.R. MCNEIL, "OVERCOMING THE 'RECENCY TRAP' IN CUSTOMER RELATIONSHIP MANAGEMENT," *JOURNAL OF THE ACADEMY OF MARKETING SCIENCES* (2013) 41:320–337.

Reasonable people can reasonably disagree on which data story is the one that matters most. I can't guarantee that our recency-trap-avoidance strategy works every time, particularly for companies in turbulent markets. In fact, one might argue that the recency curve is really a satisfaction curve. Customers will continue buying at regular intervals if they're satisfied with your product. So if a lot of time has passed since their last purchase, that may mean they're dissatisfied—they don't like you anymore—and there's little you can do to bring them back.

But sometimes customers stop buying even if they still like your products. They get out of the habit. They get distracted. They forget. But if you treat them as lost, they'll *become* lost.

It's all too easy to reject ideas that sound counterintuitive. However, you can take a counterintuitive idea and put it to the test. Companies that excel at database marketing are compulsive testers. They know that testing isn't very expensive, and it can yield valuable results. They accumulate data of all kinds, and they become adept at listening to the surprising stories that sometimes emerge.

Your company should do the same. First, tap your company's data to create your own version of the recency curve, plotting purchases against time since last purchase. Then establish a control group and a test group consisting of customers drawn from the middle of the curve who ordinarily would be ignored. See what happens if you market to the test group, but make sure you evaluate the test on a long-term basis, at least long enough for the customer to buy not only in this period but also in subsequent periods. A successful test provides you with a road map for how to avoid the recency trap.

THE METRICS SALES LEADERS SHOULD BE TRACKING

BY SCOTT EDINGER

As an executive vice president for sales, I spent countless hours reviewing, examining, and analyzing the sales forecast for my company. I required the managers who reported to me to do the same. And that's what they asked their reps to do too. We weren't alone. Conversations between sales leaders, sales managers, and sales staff frequently focus only on numbers: Did you make them? Will you fall short? How much do you think you can sell in the next quarter? The result is an inordinate amount of time spent on inspection and reporting of numbers that are, speaking frankly, out of the control of any sales leader.

In a recent study, the Sales Education Foundation and Vantage Point Performance identified 306 different metrics that sales leaders used in their efforts to manage their business. These metrics fell into three broad categories: *sales activities* (things like the number of accounts assigned per rep, the number of calls made per rep, and the percentage of account plans completed), *sales objectives* (like the number of new customers acquired, the percentage share of the customers' wallet, and the percentage of customers retained), and the subsequent *business results* (like revenue growth, gross profit, and customer satisfaction).

Even though managers were spending more than 80% of their time focused, as I had been, on the second two categories, the report found that sales management could affect only the first—the sales activities. The other two couldn't be directly managed, since they're outcomes, not the process by which the outcomes are gained.

Which sales activities should sales leaders be managing? In their book, Cracking the Sales Management Code, Jason Jordan and Michelle Vazzana break down sales activities into four discrete categories: call management, opportunity management, account management, and territory management. Different organizations might need to focus on one or another of these more intently, depending on both the nature of what they sell and where their organizational weaknesses lay. So, depending on the needs of your sales organization or team, here's a way to map the various sales activity metrics to the challenge at hand.

Sales leaders who need to improve call management—that is, the quality of the interaction between individual salespeople and prospects or clients—should focus on such metrics as call plans completed, coaching calls conducted, or even the number of calls that are critiqued and reviewed. Coaching in the area of call management is particularly valuable when the seller need make only a small number of calls to greatly affect the outcome of a given deal. In selling professional services, for example, the ability to create value in one or two interactions with a senior executive often makes or breaks the deal.

Sales leaders who need to improve opportunity management—the ability of their salespeople to vet, pursue, and close a multistage sale—should focus on the number of opportunity plans (outlining the actions required to move through the stages of the sales cycle) completed or on the percentage of early-stage opportunities qualified—that is, fully vetted to confirm you are really reaching the right customers, that these customers have the potential to generate a reasonable amount of business for the effort it will take to gain it, and that they are in fact willing to budget sufficient sums to purchase your offering. Identifying bad deals in this way and getting them out of the way early may be the simplest way to increase your odds of success. Most companies I have worked with don't want their sales teams pursuing every opportunity possible. Instead, they want to put their maximum effort on the opportunities that match some kind of ideal client profile.

Sales leaders who need to improve account management—the ability to enhance the long-term value of a single client—should focus on working with reps to develop and adjust account plans so that they define an overall strategy for the customer. This is also when it would be fruitful for you to spend time creating and monitoring standards for important client-facing activities like establishing peer meetings between your organization and the customer (by, for instance, arranging meetings between your CEO and the client's CEO) or getting the customer to sit on your organization's advisory council to provide feedback to your business. I have a high-tech client that tracks the amount of time executives spend each month with a single account that generates $65 million a year. Account management metrics are vital when a substantial portion of your organization's revenue is concentrated in a small number of key customers.

Sales leaders who need to concentrate on territory management—on how you allocate sales reps' time among all the customers in a given territory—should focus on metrics like the number of customers per rep, the number of sales calls made, and even the sales calls to different types of customers. By managing the process of selecting, prioritizing, and meeting with target customers, you can maximize the use of your sellers' most precious resource—time.

Many sales leaders have been inadvertently micromanaging through revenue or profit numbers, which is counterproductive. This is your chance to provide your sales team with a new context to succeed in. By closely managing the things you can control, you will give your organization the best chance for success.

TO UNDERSTAND CONSUMER DATA, THINK LIKE AN ANTHROPOLOGIST

BY SUSAN FOURNIER AND BOB RIETVELD

It was hardly what you'd call an "adequate sample size" for market research, but the results were nevertheless eye-opening for the maker of a pain-relief ointment: A single consumer posted an online photo showing how he placed foil over the ointment to prevent it from staining his pants.

Despite years of consumer research, the pharmaceutical firm hadn't known about the staining problem. That photo prompted the company to change the product and its communications about the ointment, creating significant value for the firm.

The beauty of listening to social-media chatter is that one picture or one comment can have an outsized impact on your consumer knowledge and, as a consequence, your profitability. But a lot of people in business don't appreciate that.

Corporate social-listening efforts are typically driven by econometricians, computer scientists, and IT technicians—the people who are experts in database management. They understand digital information, but they don't always understand how to get from information to meaning. So they boil the data down to percentages, treating random comments (and pictures of people with foil on their legs) as noise.

But if you want meaning, you have to think like an anthropologist. You have to understand that social-listening data is inherently qualitative. That means learning to appreciate the value of the stray remark and synthesizing bits of information into a higher-order sense of what's going on. That's how you make the most of social media as a tool for peering inside people's lives as they're being lived and discovering what consumers are really thinking and doing.

"Sure, sure," the numbers-oriented marketing executives may say. Social listening is great for "exploratory" research, but only as a precursor to "real" research that will determine the truth of what's being said online. What's needed, they'll tell you, is broad-based consumer research using representative samples and adequate sample sizes.

Querying a representative sample is great for testing a hypothesis or finding a statistical relationship between known concepts. But often, in marketing, you're dealing with multiple unknowns. Social listening doesn't presuppose anything. It has no constraints. Although qualitative information won't give you a simple equation or statistic that you can show the CEO, it can provide answers to questions you didn't even know you had.

And comments from a nonrepresentative sample can be highly illuminating. For example, in tech markets, think of the users who regularly post to discussion groups focused on tech products. These knowledgeable netizens provide critical knowledge about product uptake and issues around quality or perception. The same can be said of fan groups and user groups in a variety of fields.

An important player in the electric-shaver category discovered this. Before the launch of a high-end shaver that was to be priced at more than $500 and was encased in brushed aluminum, an Australian retailer posted pictures and specifications of the product online. Almost immediately, consumers began commenting about the product's "plastic aesthetic" and "cheap look and feel." The manufacturer took prompt action, posting a new photo series highlighting the quality manufacturing process and construction, neutralizing the negative sentiment spreading online.

Successfully disseminating the results of social listening requires skill at seeing stories and developing insights from messy data. It also requires a penchant for simplicity.

One company we worked with had gone a bit overboard on purchasing tools for integrating social-listening data into its business processes. It had no fewer than 19 dashboards for looking at market and customer behavior, yet none of them were really working. Rather than advise the company to implement yet another new tool, we suggested that an email go out daily, illuminating the most interesting positive and negative statements gleaned that day from customers. The result was that the company acquired a deeper understanding of its customers and drew more insights from the data.

Our work with social listening often makes us think of the comment by futurist Roy Amara that people tend to overestimate the effect of a technology in the short run and underestimate its effect in the long run. Amara's Law seems perfectly apt when it comes to social-media listening. At first, marketers were ecstatic about social listening's potential; then, influenced by the numbers people, their enthusiasm cooled. But they haven't yet fully appreciated its long-term strategic potential. They don't yet see that social listening can redefine the way managers approach marketing and that social-listening competency may well define competitive advantage in the digital age.

FEATURED COMMENT FROM HBR.ORG

"You [Fournier and Rietveld] make some good points. Social media when used properly as another qualitative tool can be a great source of inspiration for both innovation and better communication. The ensuing consumer dialogue can lead to co-creation of more relevant products and services."
—EG Marketing Strategies

HOW IS BIG DATA TRANSFORMING YOUR 80/20 ANALYTICS?

BY MICHAEL SCHRAGE

Even today, most organizations technically struggle to answer even the simplest 80/20 analytics questions: Which 20% of customers generate 80% of the profits? Which 20% of suppliers are responsible for 80% of customer UX complaints? Which 20% of customers facilitate 80% of the most helpful referrals? Indeed, even organizations where top management keeps their eyes glued to KPI-driven dashboards have trouble agreeing on what their Top Ten Most Important Customer/Client 80/20 analytics should be.

That's not good, because big data promises to redefine the fundamentals of the 80/20 rule. What happens to innovation and segmentation when serious organizations are challenged to assimilate and integrate 10X, 100X, or 1000X more information about customers, clients, prospects, and leads? Should management refine and dig deeper into existing 80/20 KPIs? Or do those orders-of-magnitude more data invite revising and reframing a new generation of 80/20s? In other words, how much should dramatic quantitative changes inspire qualitative rethinking of the vital few that generate disproportionate returns?

At one travel services giant, 100X more data in less than two years utterly transformed the enterprise conversation around "loyalty." Where repeat business and revenue had once been the dominant loyalty metrics, the firm began analyzing the overlaps and intersections between its "best" customers and social media and web comments about the company's service. The company quickly adjusted its operational definition of "loyalty" to explicitly embrace and reward its "digital evangelicals"—customers who had enough of a virtual presence to be influencers. These customers were given perks, privileges, and the power to offer certain promotions. The apparent result? A few hundred customers out of hundreds of thousands generated several millions dollars of additional top-line revenue. A digital presence 80/20 "cult" culled from the larger 80/20 loyalty "flock" yielded disproportionate profits remarkably fast.

Conversely, after identifying key traits of its breathtakingly wealthy private client practice, a global financial services firm spent millions seeking private clients with similar traits from their direct competitors. The combination of digital due diligence, processing power, and predictive profiling required several terabytes—far more than the bank's internal systems could handle. Reportedly, two pirated clients were all it took to repay the entire investment within a year.

This bigger data venture profitably answered the question, "If this is what our 80/20 clients look like, whom should we be targeting at our rivals?"

Of course, subtler and more nuanced 80/20 KPIs can be winkled out from expanding data sets. Remote diagnostics and maintenance networks empower industrial equipment firms to simultaneously ask and identify which 20% of their customers generate 80% of the most interesting and/or unexpected uses for the equipment. Might some of that novelty be patentable or protectable under some intellectual property regime? Simply by seeking to create, measure, and assess 80/20 value creating segments, firms can harness 100X more information to new strategic thrusts. Instead of "data being the plural of anecdote," 80/20 anecdotes emerge from the data—that is, committing to 80/20 KPIs inspires new analytic narratives.

For example, brand managers and advertising agencies alike increasingly make sentiment analysis part of how they assess public perceptions. So which 20% of external tweets, posts, and presences account for 80% of sentimental shifts? Similarly, what sponsored content links to that vital 20%?

Marrying Moore's Law to the Pareto Principle transforms big data into more actionable analysis. The best way to digitally discipline dataset mash-ups is by identifying the 80/20 dynamics that the organization wants driving its differentiation. "What 80/20 analytic will matter most tomorrow?" is the challenge top management needs to both ask about and rise to.

The most intriguing consequence I've observed is an emerging tension between managers and marketers who want to use big data and analytics to better customize their offering versus those who see 80/20 as a powerful segmentation exercise. Segmentation versus customization seems destined to become one of the most interesting cultural, organizational, operational, marketing, and innovation debates this decade. Ironically but appropriately, the best answers will be found in the fusion of big data and 80/20 analytics.

USE YOUR SALES FORCE'S COMPETITIVE INTELLIGENCE WISELY

BY JOËL LE BON

The vast majority of your marketing data, whether from purchases or surveys, tells you about your business customers' past behaviors. But the past is over. Your job as a manager is to know what your customers will do next month or next year. Who has data about the future?

Your salespeople, that's who.

The sales force has abundant information about the initiatives and products that your competitors are planning and, therefore, the kinds of choices that your customers will be facing in the near future. That *competitive intelligence* can not only help individual salespeople become more effective; it can also help your company make better strategic decisions.

But research I conducted with Douglas E. Hughes of Michigan State University and Adam Rapp of the University of Alabama shows that in the wrong hands, competitive intelligence can have a negative impact on sales and market share. So if you acquire information from customers, you'd better use it well, or it may hinder, rather than help, your sales efforts.

My colleagues and I have seen numerous examples of sales forces gaining important competitive knowledge. In one case, the salespeople for a consumer-goods company got a good look at a newly designed soft-drink bottle that a competitor was planning to introduce. In other cases, sales reps found out about competitors' plans to introduce new types of cooking oil and mayonnaise before the products went to market. In each instance, the companies were able to act quickly on the information to introduce their own products and avoid losing weeks' worth of sales.

Business customers are the source of this kind of information. They hear from manufacturers' reps not only about new forms of packaging and new products, but also about such things as planned pricing, discount policies, and marketing initiatives. They release this well-guarded information judiciously, however. The reps who are most likely to receive it are those who have formed strong customer relationships and engage in customer-oriented extra-role behaviors.

Being customer-oriented means focusing on long-term customer satisfaction and placing the customer's needs first while actively endeavoring to develop solutions that enable the customer to reach its goals. That might mean, for example, assisting a customer in its negotiations with another supplier, if that kind of help is requested. Engaging in extra-role behaviors means going above and beyond to help the customer.

It's easy to see how in the context of this type of high-trust relationship, a customer would be willing to divulge proprietary information picked up from another supplier's rep. But that's not to say customers release information just to be nice. They provide knowledge in the expectation of some sort of benefit. They might expect the salesperson who receives it to offer better terms, for example, or to provide greater levels of effort and commitment.

That's why the salesperson's response to the information is critically important. Our research has found that reps who are able to make the best use of competitive information are those who are highly adaptive—that is, they're able to interpret the information and adjust their selling approach accordingly.

The flip side is that salespeople with low adaptability run into trouble when they acquire competitive information. They fail to respond to the intelligence gleaned from the customer, because either they don't care enough to use it or they're unable to. When low-adaptive sellers fail to use customer-supplied information to more strongly position a product to meet the customer's needs, the customer gets a negative impression of the company's products.

Salespeople who make use of customer information for the customer's sake tend to be more successful in selling, and they get more information. Those who don't make good use of it find that their sources dry up. Customers ask themselves, "Why should I share any more information with that rep?"

The strategic value of competitive intelligence can be far-reaching. But we all know that information doesn't always flow freely between sales reps and the rest of the company. To open up that flow, encourage salespeople to participate in the decision-making process. That will help the reps see the importance of the information they glean from customers. Your marketing people may be surprised at just how much the sales force knows about competitors' prices and discount policies, as well as about the concerns your customers are voicing about your and competitors' products.

Companies with poor information about competitors are stuck being reactive, lurching from defensive tactic to defensive tactic as new products appear, seemingly out of nowhere. Cultivating a rich source of intelligence on competitors' projects and products, as well as any new market segments they're targeting, gives you time to be strategic in your responses.

CAN YOU SEE THE OPPORTUNITIES STARING YOU IN THE FACE?

BY CHRIS BRIGGS

We all suffer from inattention blindness. We focus so intently on a particular task that we don't notice the weird thing that's right there in plain view. The person walking by in a gorilla suit, for example.

Companies suffer from inattention blindness, too. And in a business context, the weird thing that gets overlooked can turn out to be a crucial differentiating factor. If one company doesn't notice it, another will—to great advantage.

A perfect example is the high-end retail company that, like its competitors, needed to close certain stores because more and more customers were buying online or via catalog. But unlike its competitors, this company sensed that the closure of a store might have harmful network effects.

In a study of customer data on purchases made by people who lived within a short driving distance of its stores, the company could see that customers were visiting stores to look at the merchandise, then going home to compare other options and make their purchases. Closing a store would deprive customers of their showroom. The hunch that this would hurt sales was corroborated by a dip in online and catalog sales in areas around stores that had previously closed.

Not only did the company end up closing fewer stores than it originally thought necessary, it now had a new method to evaluate the overall positive impacts that new or relocated stores would have on the retailer's bottom line—across all sales channels.

The importance of stores' network effects is just the kind of thing companies often miss when they're focusing intently on the mass of conventional data they collect on customer behavior. In fact, I've come to believe that less than 1% of the data is truly useful.

The challenge, of course, is figuring out which 1% really matters, and with the recent explosion of data types and sources, making that kind of distinction is getting harder and harder. Big-data initiatives are proliferating, and the information is getting more complex all the time. In the next couple of years—not decades, but years—companies will have access to data not only from customers' mobile devices but also from wearable tech and automobiles, for example. Soon, retailers will have all kinds of information about customers who are walking through stores and will be able to send them targeted messages. There's a lot of potential benefit for both retailers and customers.

However, there is only benefit if the data is well managed and well understood. Statistics literacy isn't very high in most businesses. A few educational institutions have realized this and are making a push to turn out business graduates who know their way around a regression analysis. But for the most part, businesspeople aren't familiar enough with statistics to use them as the basis for good decisions. If you don't understand the numbers, you can go a long way down a bad road very quickly. That's why every team charged with making decisions about customers should include a trusted individual who understands statistics. If that understanding isn't between your own two ears, make sure you bring a person with that skill set onto your team.

But let's say you're a small organization—a group of restaurants, say, or a retail chain with just a few locations. You don't have anyone on staff with a statistics background, and you can't afford to hire such a person. There's one thing you can do to ensure that you make informed decisions and don't let the numbers lead you astray: Get a solid understanding of who your core customers are, what their value is to the company, and what your objectives are. Learn which customers are profitable and which ones aren't, and decide what you want to do to increase the profitability and number of good customers. Use that information as the center point of all your decisions.

Will that cure your inattention blindness? Maybe not entirely, but if you can get a good understanding of what distinguishes a good customer from a bad one, and a sense of what makes your good customers tick, you'll immediately cut half the noise out there. That will make it a lot easier to see the opportunities that are staring you in the face.

HOW TO START THINKING LIKE A DATA SCIENTIST

BY THOMAS C. REDMAN

Slowly but steadily, data are forcing their way into every nook and cranny of every industry, company, and job. Managers who aren't data savvy—who can't conduct basic analyses, interpret more complex ones, and interact with data scientists—are already at a disadvantage. Companies without a large and growing cadre of data-savvy managers are similarly disadvantaged.

Fortunately, you don't have to be a data scientist or a Bayesian statistician to tease useful insights from data. This post explores an exercise I've used for 20 years to help those with an open mind (and a pencil, paper, and calculator) get started. One post won't make you data savvy, but it will help you become data literate, open your eyes to the millions of small data opportunities, and enable you to work a bit more effectively with data scientists, analytics, and all things quantitative.

While the exercise is very much a how-to, each step also illustrates an important concept in analytics—from understanding variation to visualization.

First, start with something that interests, even bothers, you at work, like consistently late-starting meetings. Whatever it is, form it up as a question and write it down: "Meetings always seem to start late. Is that really true?"

Next, think through the data that can help answer your question, and develop a plan for creating them. Write down all the relevant definitions and your protocol for collecting the data. For this particular example, you have to define when the meeting actually begins. Is it the time someone says, "OK, let's begin"? Or the time the real business of the meeting starts? Does kibitzing count?

Now collect the data. It is critical that you trust the data. And as you go, you're almost certain to find gaps in data collection. You may find that even though a meeting has started, it starts anew when a more senior person joins in. Modify your definition and protocol as you go along.

Sooner than you think, you'll be ready to start drawing some pictures. Good pictures make it easier for you to both understand the data and communicate main points to others. There are plenty of good tools to help, but I like to draw my first picture by hand. My go-to plot is a time-series plot, where the horizontal axis has the date and time and the vertical axis has the variable of interest. Thus, a point on the graph below (click for a larger image) is the date and time of a meeting versus the number of minutes late.

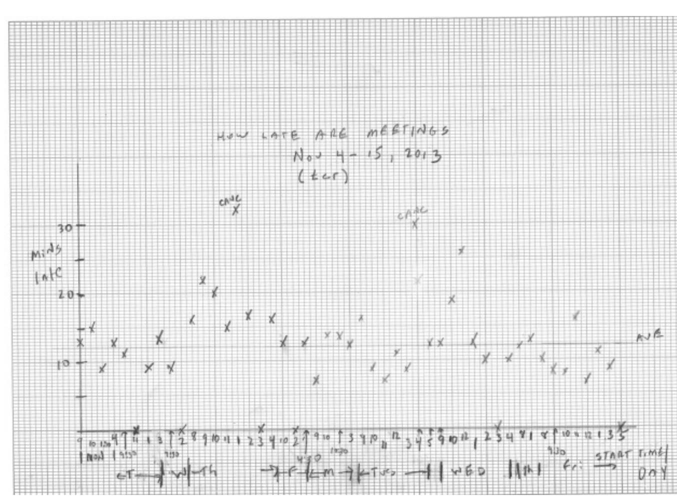
Now return to the question that you started with and develop summary statistics. Have you discovered an answer? In this case, "Over a two-week period, 10% of the meetings I attended started on time. And on average, they started 12 minutes late."

But don't stop there. Answer the "So what?" question. In this case, "If those two weeks are typical, I waste an hour a day. And that costs the company $X/year."

Many analyses end because there is no "So what?" Certainly if 80% of meetings start within a few minutes of their scheduled start times, the answer to the original question is "No, meetings start pretty much on time," and there is no need to go further.

But this case demands more, as some analyses do. Get a feel for variation. Understanding variation leads to a better feel for the overall problem, deeper insights, and novel ideas for improvement. Note on the picture that eight to 20 minutes late is typical. A few meetings start right on time, others nearly a full 30 minutes late. It might be better if one could judge, "I can get to meetings 10 minutes late, just in time for them to start," but the variation is too great.

Now ask, "What else does the data reveal?" It strikes me that five meetings began exactly on time, while every other meeting began at least seven minutes late. In this case, bringing meeting notes to bear reveals that all five meetings were called by the vice president of finance. Evidently, she starts all her meetings on time.



So where do you go from here? Are there important next steps? This example illustrates a common dichotomy. On a personal level, results pass both the "interesting" and the "important" tests. Most of us would give anything to get back an hour a day. And you may not be able to make all meetings start on time, but if the VP can, you can certainly start the meetings *you* control promptly.

On the company level, results so far pass only the interesting test. You don't know whether your results are typical, or whether others can be as hard-nosed as the VP when it comes to starting meetings. But a deeper look is surely in order: Are your results consistent with others' experiences in the company? Are some days worse than others? Which starts later: conference calls or face-to-face meetings? Is there a relationship between meeting start time and most senior attendee? Return to step one, pose the next group of questions, and repeat the process. Keep the focus narrow—two or three questions at most.

I hope you'll have fun with this exercise. Many find a primal joy in data—hooked once, hooked for life. But whether you experience that primal joy or not, do not take this exercise lightly. There are fewer and fewer places for the "data illiterate" and, in my humble opinion, no more excuses.

FEATURED COMMENT FROM HBR.ORG

"Beautifully explained about how people don't conclude accurately and derive their conclusions in few steps. We have to keep stretching ourselves to draw more insights from the data."
—Bikash Kumar Mallick

HOW A BATHTUB-SHAPED GRAPH HELPED A COMPANY AVOID DISASTER

BY JAMES H. DULEBOHN AND JOHN MALANOWSKI

Caught up in administrative activities such as managing employee records and planning company picnics, HR departments can too easily lose sight of their primary function: making sure the organization has the needed human capital to implement its strategy.

Not even the best company picnic in the world, with a zip line and gourmet hot dogs, can compensate for an HR department's failure to help the company put people with the right skills into the right jobs at the right time.

But successfully managing human capital requires quantitative and analytic skills that tend to be scarce among HR professionals, especially those who continue to see themselves as essentially administrative rather than as strategic partners. Today, HR managers need to be attuned to the value of data and need to know how to use it to help them fulfill their obligations to their companies.

We're not talking about big data. We're talking about small data—ordinary, prosaic data of the type that companies have been accumulating for years and that is increasingly accessible, both internally and from external sources. Nor are we talking about the typical human resources metrics, which evaluate how efficiently the HR department is performing its administrative tasks. We're talking about data that can show what lies ahead for the company and inform workforce planning.

Consider, for example, the bathtub.

A few years ago, a global tech company had recently been cutting costs and so wasn't disposed to do a lot of hiring, but an analysis of its data showed that it was heading for trouble. The key numbers included the ages, experience, and skills of its engineers, as well as employees' expected departure dates as they retired or left for other reasons.

One of the results was a visualization that showed the number of engineers as a function of years of service. It was U-shaped. It looked like a cross-section of a bathtub. It signified that there were lots of newbies and lots of highly skilled employees with a good deal of experience behind them—and not too many in between. In addition, attrition in the early-career professional ranks was significantly higher than that for longer-service employees.

The graph, and other data, pointed out a serious problem that the company hadn't been aware of: It had a dire need to recruit more engineers to meet both existing and future commitments to customers.

The company then looked at external data, such as the overall decline in the number of engineering graduates and the 10-year projected shortfall of engineers in the workforce, and saw that it was facing a significant problem. As a result, the company stepped up its recruitment and ended up hiring about 10,000 engineers per year.

Other types of data have similar strategic value. Evaluating turnover by unit or job function, for example, not only contributes to an understanding of what your future staffing needs will be, but also can alert you to problems: Is there a managerial issue? A compensation issue? More fundamentally, will the company be able to compete in 10 years?

And speaking of compensation, the proliferation of wage data today makes it possible for companies—and workers—to compare wages across employers. When salary information was harder to come by, companies mainly focused on the internal labor market: Are employees who have the same experience and are doing similar jobs being paid similar wages? Now there's more focus on how companies' salary structures stack up against competitors'. The strategic value of this data relates to the company's positioning: Does it want to lead the market by offering better wages, or does it want to *match* the market? If there's a scarcity of candidates, the company might be better off trying to lead the market.

HR is no longer limited to being transaction-oriented, as it once was. It certainly isn't the "dark bureaucratic force that blindly enforces nonsensical rules, resists creativity, and impedes constructive change." But in many companies, it hasn't yet fulfilled its vast potential as a key driver of business performance. Part of the reason is that many HR professionals aren't well trained to understand the value of the data that their function generates or to know which metrics and forms of data analysis should be used in support of the business's strategy. Hidden in those numbers are insights such as whether the labor force's skills match the company's strategic plans, whether employee-development plans are effective, and whether virtual teams are really working.

It's only by mining those insights that HR can be sure its efforts at workforce planning, recruiting, selection, placement, training and development, compensation, and other core human-resource functions are helping the company move in the right direction.

BIG DATA DEMANDS BIG CONTEXT

BY JESS NEILL

When Microsoft built Windows 8, its goal was to move beyond operating-system conventions that were based on outdated user-behavior assumptions and to create an OS for the way people really use computers today.

Microsoft's engineers discovered that people were doing less of the time-consuming writing and creating that had once been the norm. Increasingly, users were socializing for short bursts.

The research also showed that people loved having "touch" functionality and were avidly consuming small pieces of live information.

Consequently, Microsoft decided that Windows 8 should feature navigation that enabled multitasking and quick interactions, and that it should also have touch and live tiles.

None of this was wrong. And yet these decisions, so carefully researched and thought through, all contributed to the failure of Windows 8.

How does this happen? When we entered the age of big data, many of us assumed we had left the age of big risk. We didn't have to guess anymore. We didn't have to go out on a limb. We'd follow the numbers, the "truths."

But time and time again we're finding that it's not that simple. No matter how good the research is, big data is nothing without big context.

To keep context in mind, there are a few questions I ask myself while designing research, analyzing data, or, most important, making decisions.

- What underlying assumptions am I making?

In Windows 8, I think Microsoft's engineers made a fundamental assumption that led them astray: that users want one interface for all machines, one machine for every part of their lives. The research went into what this single interface should be—not *whether* it should be.

What if users don't want just one device? What if they're embracing and owning many specialized devices?

It's easy to get into a bubble and focus our thinking too quickly. So whenever I approach data, I first ask myself what assumptions I'm making.

- What emotions will be at play?

When asked what they want in the hypothetical, people answer rationally. They make "good" decisions. They pick cheaper, faster computers that are less attractive. Or they say that they'd try an exciting new platform without considering the frustration of a learning curve.

But what people say and what they do are two very different animals. Which of us hasn't been seduced into a less-savvy purchase because of a shiny case?

When designing research, I try to probe for the emotional drivers as well as the rational drivers. I want to know if consumers see a product as a utility or a luxury. Do they identify with a store or brand as part of their persona? Do they see it as a friend? I can use these answers to temper and check the purely rational responses.

- How can I better learn about context?

Yes, people are using touch daily on smartphones and tablets. It's intuitive. Just look at YouTube videos of babies trying to "swipe" pages in physical magazines.

But when Microsoft put touch at the forefront of its operating system for PCs, consumers didn't bite, partly because touch computers were more expensive than non-touch. But the bigger problem was that although touch is great for the social interactions and brief browsing that people do on smartphones and tablets, users are relegating PCs to work and productivity, and in that context, they don't see the value of touch.

Microsoft had the right information. But it was missing the larger contextual picture.

The details are easy to measure. They give you clear-cut data and answers. Context is harder—it gets mushy. The methodologies are more complex and the results are open to interpretation.

Most disconcerting of all, data about context won't give you an answer; it will only help inform your answer. Compounding that, contextual data can seem superfluous, so fighting for the money to research it can be hard, and selling ideas based on it even harder.

But we take a bigger risk when we ignore the context.

Microsoft's research points to an increasingly diverse device landscape, with each device being used for specific and differentiated uses and behaviors. Probably the ultimate PC OS leans into PC behaviors, letting the smartphones and tablets specialize and optimizing the user's journey between devices.

But the real lesson is that we always need to consider context. Otherwise, we too can have all the right answers to all the wrong questions.

HOW TO GET MORE VALUE OUT OF YOUR DATA ANALYSTS

BY ROBERT MORISON

Organizations succeed with analytics only when good data and insightful models are put to regular and productive use by businesspeople in their decisions and their work. We don't declare victory when a great model or application is developed—only when it's being used to improve business performance and create new value.

If you want to put analytics to work and build a more analytical organization, you need two cadres of employees:

- Analytics professionals to mine and prepare data, perform statistical operations, build models, and program the surrounding business applications
- Analytical businesspeople who are ready, able, and eager to use better information and analyses in their work, as well as to work with the professionals on analytics projects

In *Analytics at Work*, we call the latter group "analytical amateurs." That doesn't mean *amateurish*—only that you're not a professional, that analytics isn't your main occupation. You can be a scratch golfer or ace tennis player while still an amateur. Amateurs can be very accomplished analytically—in using analytical applications, envisioning additional opportunities for using analytics, and participating as business staff on analytics projects. You're in luck if your CEO, executive team, and general managers across the business are all accomplished analytical amateurs.

There is widespread recognition of the shortage of analytical professionals. Lesser appreciated is the fact that most organizations are also way short on analytical amateurs. A May 2011 McKinsey Global Institute study on big data analytics predicted a coming shortfall of around 150,000 people with deep analytical skills—and a shortfall of 1.5 *million* businesspeople with the know-how to put big data analytics to use.

The key to overcoming these shortages is to develop talent in both cadres together. In other words, the most important question may not be "How can we hire more analysts?" but rather "How can our analytical professionals best work together with our businesspeople?"

The most effective employee development happens on the job, day to day, often one-on-one. The way to expand the business acumen of analytics professionals is to have them spend plenty of time working with business colleagues. The way to expand the analytical capability and appetite of analytical amateurs (aka, business managers and professionals) is to have them work directly with analytics professionals on both analytics projects and simply meeting their own information needs.

By spending time "in the field," professional analysts gain greater familiarity with business operations and a pragmatic appreciation for how analytics are used in management decisions and employee workflows. What do the businesspeople learn?

- To be more aware of the data they use and their own decision processes. They get better at evaluating and improving their data and adjusting their decision processes depending on the quality and sufficiency of data at hand.
- To serve themselves with data and analyses. They become more adept at finding data and using business intelligence and visual analytics tools, and they are more rigorous in using established tools like spreadsheets; thus, they are better able to meet many of their analytical needs independently and immediately.
- To understand the logic and methods behind the analytical models, applications, and outputs they use. Will they pick up some statistical methods? Perhaps, but the real goal is to learn enough to understand and trust their analytics—and develop a sense of the limitations of analytics.

Analytical amateurs accomplished in these ways not only make better use of analytics in their decisions and work, but also make greater contributions when serving as subject matter experts or otherwise participating in analytics development initiatives.

In analytical organizations such as Procter & Gamble, professional analysts spend a lot of time in the field, including "embedded" in business units. And there's an active rotation program getting businesspeople into analytical roles (many of which don't require PhDs in statistics or deep data scientist skills). Analyst talent may be in short supply, but the solution is to kill two birds with one stone and develop the two cadres together.

SMALL BUSINESSES NEED BIG DATA TOO

BY CHRISTINA DONNELLY AND GEOFF SIMMONS

If you run a small or medium-size business, chances are that you haven't felt a need to invest in extensive customer data, relying instead on your well-honed intuition to help you hold your own against data-rich, bigger competitors. A lot of small-firm owners and managers feel that way, and in many cases they're justifiably proud of their competitive intangibles—a gut sense of the market and the flexibility to change quickly.

What you may not realize is that investing in data and learning how to use it might be transformative for your business. In research we conducted with Gillian Armstrong of the University of Ulster and Andrew Fearne of the University of Kent, we found not only that small businesses benefited from the precision offered by customer data, but also that exposure to data encouraged owner-managers to share insights with employees and get them involved in companies' competitive thinking.

Of course, for the smallest businesses, access to extensive consumer data can be prohibitively expensive, a point we'll address in detail in a subsequent post. But cost isn't the only barrier. Small firms tend to find the whole concept daunting—they know they lack the expertise and the time resources to make good use of the information. Our three-year project was designed to build awareness among small firms about the value that data could have for their businesses.

Funding from a regional UK government agency enabled us to get over the cost barrier: It allowed us to provide loyalty-card information from supermarket giant Tesco, free of charge, to seven firms in the Northern Ireland region of the UK. These companies, ranging in size from seven to 45 employees, sell such things as dairy products, baked goods, vegetables, and desserts to Tesco and another grocery chain, Sainsbury's (we tasted some pretty amazing soups during the course of our research—a tough job, but someone has to do it).

The data, provided by analytics firm dunnhumby, covered such things as consumer life stage and lifestyle, market-basket analysis, and best-performing stores for the small firms' products.

The formalized structure of loyalty-card data within a statistical format requires firms to take a more formalized and structured approach to marketing planning—that's a challenge for small companies. Owner-managers were encouraged to attend workshops, and one of us (Christina Donnelly), after being trained by dunnhumby, worked one-to-one with owners and managers to help them retrieve the most relevant data from the loyalty-card database and analyze the information, so that the companies could answer questions such as "How is my category performing?" "What is the most popular flavor of bread?" and "What type of consumer buys a product similar to mine?"

We found that prior to being exposed to the loyalty-card data, the small businesses tended to be dominated by their owner-managers, who made decisions on the basis of their past experiences and any consumer information they could get their hands on. For example, one firm, having been asked by a retailer to produce a range of ready meals, simply looked at other products on the market and tried to imitate them. In other cases, the small firms followed guidelines laid down by the big retail buyers.

Once they were given access to loyalty-card data, most of the small firms took to it immediately. They were quick to adopt a more formalized approach to marketing planning. They were able to envision long-range innovations rather than react to competitors' or the retailers' actions. One small-firm owner said the data had changed the company's ideas about how to grow its consumer base. Another said, "Now we know precisely who our target consumer is."

A yogurt maker, for example, learned by analyzing the data that older adults were a key market for its products, so when the company's representatives visited supermarkets for in-store tastings, they no longer tried to entice younger shoppers and instead focused on older people. The tactic improved the events' productivity.

But the small firms didn't abandon their reliance on experience. Instead, the data complemented the owners' and managers' intuition, giving them new confidence.

Moreover, the data amplified the firms' inherent entrepreneurial nature. Workplaces became more collegial: Most of the owner-managers shared the card information with their firms and encouraged employees to get involved and offer new ideas.

Big Data threatens to create a deep divide between the have-datas and the have-no-datas, with big corporations gaining advantage by crunching the numbers and small firms being left to stumble in the dark. The small firms we worked with were well aware that they were at a severe disadvantage compared to big competitors that had the financial muscle to buy into loyalty-card data and the resources to use it.

Governments and universities can play important roles in bridging the divide, providing funds and expertise so that small firms can get access to, and learn to interpret, data. Our project is an example of a fruitful collaboration—among the University of Kent, the University of Ulster, dunnhumby, and the government funding body. The quality of the learning is important. A key reason this project worked was the one-to-one help for the owner-managers. Data is only as good as the people who use it.

For small and medium-size firms that do manage to acquire consumer data, there's still more work to be done: They need to be sure to encourage employees to participate in thinking about how to use the information competitively. We saw firsthand that inclusivity energizes firms, driving innovation.

GET MORE VALUE OUT OF SOCIAL MEDIA BRAND-CHATTER

BY SUSAN FOURNIER AND BOB RIETVELD

It's becoming commonplace for consumer companies to listen to what their customers are saying on social media, but the big question is: What do they do with the results? In a lot of cases, managers merely circulate them within the marketing department—after marking them with a prominent "FWIW."

That's because they don't know what this information could be worth.

Companies don't realize that with proper care and handling, insights harvested from social listening can become as robust a source of strategic inspiration as any must-have diagnostics on the dashboard.

Social listening is inexpensive, too, in the sense that it has a high insight-to-dollar ratio. That's because you don't have to survey or interview anyone—unsolicited comments from engaged customers are already out there, waiting to be analyzed. And social-media data are continuous—you don't have to depend on quarterly or annual consumer surveys that are out of date before they're even analyzed.

To tap social listening's potential as a source of strategic inspiration, think like a market researcher and follow these sensible steps:

Make sure the quality of your social-listening data is good. Like all data, the information you glean from social media should be subject to market-research protocols for reliability and validity. Ask the same kinds of tough questions you'd ask about any research project: Are the data drawn from the entire social-media landscape? Is the sampling of comments statistically sound? Is the system of data classification, in terms of topics, themes, and sentiments, accurate? Does your automated coding allow for idiomatic meanings, as in "This brand is the s—t"? The insights you get from social media are only as good as the data set you create.

Don't make your social-media data stand alone. Information from social listening must be correlated with other streams of data that the company is using. For example, in an analysis we performed for a transport company, we found that complaints shared on daily Twitter feeds tracked 90% with the content of customer-service comments registered by phone or mail. Linkages like this go a long way toward speeding the adoption of social-media data as a valid strategic-insight source.

Sometimes the correlations are low between what you think you know and what social listening reveals. But that doesn't mean you should jettison the listening data; it just means you need to consider both sets of findings simultaneously to decipher the true story.

In collaboration with a beer manufacturer, we conducted a brand-positioning analysis of three leading brands. Conventional wisdom at the company dictated that differentiation based on taste was not an option. Studies had demonstrated, time and again, that consumers could not pick out their favorite beer in blind taste tests. This finding had underpinned positioning and communication decisions for years.

But online comments revealed a different picture: Consumers went deeply into stories about the taste and sensory experiences of not just the beers they loved, but also of those "watered-down, hangover-inducing" beers they disdained. It turned out consumers *thought* they could pick their favorite beers out of a lineup. And perception, not reality, was what mattered in this space. The social-listening data allowed a marriage between the quantitative and the qualitative. Customer stories illustrated the insights, and the raw numbers (thousands of online statements) validated the insight, allowing the company's conventional wisdom—and the branding programs guided by it—to slowly change.

Think about "impact" and not just ROI. Marketing managers tend to take too narrow a view of social listening, seeing it merely as a way to measure the return on investment of specific marketing campaigns. For example, an electric-toothbrush maker that had launched a campaign to woo "non-electric" brushers was dismayed to learn that the resulting burst of social-media activity came mostly from existing users. It branded the campaign a flop and moved on.

In so doing, the company overlooked the value of what it had found on social-media sites. Users were sharing positive stories, advocating electric brushing, and in some cases expressing their love of the company's brand. The company was getting a rare unfiltered look at how consumers were living the impact of the company's strategies and brands.

Be sure your social-listening analyses make their way out of the marketing research department and into the wider organization, including leadership circles. Don't let the information stay bottled up in the departments that collected and "own" the data. That means establishing a common analytical currency and language throughout the company so that managers can take action and be held accountable. One company we worked with created a Center for Digital Excellence to coordinate data on a vast brand portfolio. The company tied the digital indicators to bonus compensations, signaling C-level commitment to the program. It's that kind of high-level integration that enables companies to focus efforts and resources effectively, creating value for the firm.

It's not simple to turn large volumes of unstructured data into analyzable formats and insights. But it can be done. And it must be done—social listening is too valuable to be relegated to the "for what it's worth" category.

JUST ADDING A CHIEF DATA OFFICER ISN'T ENOUGH

BY TERRI GRIFFITH

The proliferation of C-suite roles is an indication of the increasing strategic and operational complexity organizations face. Heightened expectations of expertise are also part of the picture—for instance, GE's recent transition from asking executives to focus on breadth to focus on depth.

One of the newest additions to the C-suite is the chief data officer. Thomas Redman wrote in October about the increasing value of a CDO and how to know whether such a role might be good for your organization. If you've decided to move ahead, then the next step is to effectively build that role into the rest of the top management team. To do so, you need clarity around how the CDO will work with the rest of the top management team as well as incentives that support collaboration across the top executives and senior managers—something that goes beyond equity compensation.

Research shows that you can't just add a new member to a team without renegotiating the roles and relationships. Members need to know who already knows what, who should know what, and how to coordinate—and that all varies by the particular people in the team, not just their role. For instance, the person who was once the best at analytics may no longer hold that position if a CDO is brought in. Brown, Court, and Willmott provide an excellent starter list for the issues to address: establishing new mind-sets around the value of data; defining a data-analytics strategy; determining what to build, purchase, borrow, or rent; securing analytics expertise; mobilizing resources; and building frontline capabilities.

Brad Peters, CEO of Birst, a business intelligence company, raised the issue of incentives and structure with me in an interview at Saleforce.com's Dreamforce conference. He said that without changes in incentives, as well as a change in the structure of the C-suite, broad-based roles like chief information officers, chief marketing officers, and CDOs are hampered in their effectiveness. Alignment with specific business goals is important to all senior executives, of course, but it's especially true for those who may be in service roles to the business units of the organization.

David Aaker's helps explain why in his book *Spanning Silos: The New CMO Imperative*. Though he was writing about chief marketing officers, his advice is equally valuable here. When incentives are focused on rewarding silo behavior and performance—as so many incentives are—it is a struggle to take advantage of a boundary-spanning executive position. Instead, organizations need cross-silo incentives and initiatives to support behaviors that leverage the skills of the new executive. (Chief operating and finance officers side-step this issue, given the broader power base of their positions.)

The top management team should approach the role clarification and incentive alignment as a negotiation. While it is likely that the entire executive suite would be involved in a decision to bring in a CDO, it is important that the process goes beyond involvement. Negotiation, with its norms of agreement to an explicit deal, means that the new role, its responsibilities, and its rewards are fully integrated into those of the rest of the top management team.

Have this negotiation as part of the new CDO onboarding. Lay out major roles and responsibilities to highlight places where data and other roles intersect, given the expertise of the particular people on the top management team. Use Brown, Court, and Willmott's six tasks of data-focused activities to start developing the issues to have the on the table. Add incentives to the mix by setting goals around the collaborative activities. Morten Hansen argues that you don't need to overhaul an entire incentive structure—you can focus on special incentives for silo-spanning collaborations.

The whole top management team must collaborate both in fact and regarding incentives. There must be something that rewards cross-silo efforts. To effectively make use of data's insights, executive teams need more than goodwill and a new hire.

EXECUTIVES IGNORE VALUABLE EMPLOYEE ACTIONS THAT THEY CAN'T MEASURE

BY RICHARD D. JOHNSON

Does better data mean better employee performance and organizational outcomes? That's the implication of the current emphasis on big data and the use of metrics in HR, but the answer isn't an easy "yes."

To see what I mean, consider your local schools. When teachers are evaluated and paid on the basis of students' test scores, performance on tests typically improves. The moral: Data works. Long live data!

But research also shows that higher test scores don't necessarily translate to greater student mastery of the material. In other words, teaching methods that are effective in improving test scores may not be the best for increasing students' knowledge. The moral: Data doesn't work. Down with data!

Teaching is a great example of the strengths and shortcomings of data-based performance assessments because, in a sense, teachers are both frontline workers (when actively teaching in the classroom) and executives (when they write lesson plans and develop teaching and classroom strategies). In their role as line workers, teachers can be expected to respond to whatever metrics are applied to them. But simple metrics such as test scores may not detect the difference between teaching strategies that increase students' knowledge and those that don't.

A lot of us have jobs like that—some of our work leads to easily measured outcomes (sales volume), while some is much harder to quantify (solving a complex technical issue while easing customer frustration). With the rise of eHRM—electronic human-resource management—it becomes easier than ever for organizations to automate the collection and analysis of employee data. But this also means that it becomes easier to rely on data that organizations can conveniently collect and analyze. Behaviors and aspects of performance that aren't easily quantified and captured in eHRM can become neglected.

For example, an organization that measures only the number of cases a customer-service rep handles per day may overlook the value of an employee who is capable of winning over an agitated customer. Consider also the use of workload-scheduling software for maintenance employees or physicians. These systems can increase overall operational efficiency and employee performance (measured as number of service calls completed or patients seen), but what happens if the system doesn't account for the complexity inherent in different jobs? High-performing experts can be penalized for taking on complex assignments.

When you over-objectify or oversimplify the measurement of performance, you risk missing the richness of what makes that job special—or complex—or what makes each person's contribution unique. Yet, for many managers, this duality is not apparent. Managerial knowledge and skill in applying metrics has not kept up with organizations' ability to create them. Managers often don't have the time or knowledge to understand the limitations of the metrics they apply. Instead they rely on easily obtained "objective" data from the system and ignore the less-quantifiable and more complex aspects of performance.

Employees will engage in the behaviors easily captured through the system and ignore those aspects of performance that aren't considered. That's why organizations need to continually assess whether the data they're collecting is truly relevant to the broader organizational objectives.

My hunch is that HR is moving toward an era of better data. What do I mean by better data? Take, for example, Sabermetrics and its use in Major League Baseball. Before Sabermetrics came along, few people imagined that the conventional thinking about baseball could be upended by arcane statistics such as wins above replacement.

Before we can develop a metric similar to wins above replacement for employees, we have to define key organizational and employee performance outcomes and determine how they relate to employee behaviors. The challenge is that we still don't know what these metrics will look like or whether they will fully reflect performance.

Along with better data, we need to develop a more nuanced view of human qualities and human potential. Can we not only accept but embrace that some behaviors may not be reducible to easily quantifiable metrics, and that no amount of data can fully capture all of your, or my, best performance qualities? In a world that is increasingly driven by quantitative analyses of employees and performance, we need to find ways to efficiently incorporate both the quantitative and the qualitative aspects of performance.

EVEN SMALL COMPANIES CAN TAP BIG DATA IF THEY KNOW WHERE TO LOOK

BY PHIL SIMON

Small and medium-size businesses are often intimidated by the cost and complexity of handling large amounts of digital information. A recent study of 541 such firms in the UK showed that none were beginning to take advantage of big data.

None.

That means these firms and a lot of others are at a serious disadvantage relative to competitors with the resources and expertise to mine data on customer behaviors and market trends. What these data-poor companies don't know is that it's possible to get a lot of value from big data without breaking the budget. I discovered this while researching my book *Too Big to Ignore: The Business Case for Big Data*. In fact, the abovementioned UK study notwithstanding, I found that plenty of small and midsize companies are doing interesting things with data, and they aren't spending millions on it.

True, in the past, companies seeking to tap into big data needed to purchase expensive hardware and software, hire consultants, and invest huge amounts of time in analytics. But trends such as cloud computing, open-source software, and software as a service have changed all that. New, inexpensive ways to learn from data are emerging all the time.

Take Kaggle, for instance. Founded in 2010 by Anthony Goldbloom and Jeremy Howard, the company seeks to make data science a sport, and an affordable one at that. Kaggle is equal parts funding platform (like Kickstarter and Indiegogo), crowdsourcing company, social network, wiki, and job board (like Monster or Dice). Best of all, it's incredibly useful for small and midsize businesses lacking tech- and data-savvy employees.

Anyone can post a data project by selecting an industry, type (public or private), participatory level (team or individual), reward amount, and timetable. Kaggle lets you easily put data scientists to work for you, and renting is much less expensive than buying them.

Online automobile dealer Carvana, a start-up with about 50 employees, used Kaggle to offer prizes ranging from $1,000 to $5,000 to data modelers who could come up with ways for Carvana to figure out the likelihood that particular cars found at auctions would turn out to be lemons. For the cost of the prize money (a total of $10,000), Carvana got "a hundred smart people modeling our data," says co-founder Ryan Keeton, and a model that the company was able to host easily and inexpensively.

Even a lack of data isn't an insurmountable obstacle to harnessing analytics. Data brokers such as Acxiom and DataLogix can provide companies with extremely valuable data at reasonable prices. As Lois writes, "There's a thriving public market for data on individual Americans—especially data about the things we buy and might want to buy." For example, a marketing-services unit of credit reporting giant Experian sells frequently updated lists of names of expectant parents and families with newborns.

A number of start-ups are jumping into this space—open-database firm Factual recently closed $25 million in series A financing, led by Andreessen Horowitz and Index Ventures. The company is building datasets around health care, education, entertainment, and government.

Do Kaggle, Factual, and their ilk represent the classic business disruption that will turn the data industry upside-down and make consumer and market information available to every company, large and small? No—they don't portend the imminent demise of corporations' in-house data analysis or of high-priced analytics firms. But they're providing attractive alternatives for companies that can't afford to—or simply don't want to—hire their own data scientists.

USE YOUR DATA TO GET A HOLISTIC VIEW OF THE CUSTOMER

BY JOHN FORESE

A few years ago, at LiveNation Entertainment we were accumulating a lot of data from consumers but weren't doing much with it—and we weren't sure what *could* be done with it (a pretty common situation for companies these days).

After thinking about the data strategically, we figured out ways to put it to good use. Today, consumer information helps us benefit our clients and serves as an important differentiator in a competitive market.

The key to getting to this point was adding a new dimension to our concept of the company: We began treating data as an integral part of our live entertainment business, as important as the events we produce and promote.

The new approach to data began with a merger between LiveNation and Ticketmaster. We are now both a B2C and a B2B company: We own concert venues and are the leading promoter of live events in the world. Ticketmaster has B2B relationships with sports teams and producers of a variety of events across sports, concerts, family acts, and arts and theater. But like Live Nation, Ticketmaster has a B2C component as well, in that we sell tickets to consumers.

We saw that ideally, the data that had been collected from consumers could help us provide a holistic view of the music or sports fan for our business clients. We have transactional data about what consumers are buying, and we can enrich that with demographics and psychographics. We can also add web data—What are people looking at online?—and information from email campaigns. This is a tremendous amount of data.

We could provide a professional sports team, for example, with a rich view of its fan base, showing which fans buy tickets months ahead, which buy at the last minute, which pay for premium seats, and which are looking for discounts. That kind of information could help teams shape their communication to fans.

But there was a lot of work to be done. Because the information lived in disparate systems, an IT investment would be required in order to make sense of it and standardize it. Those six people with similar names in the combined databases—were they the same person? That all had to be sorted out. We'd have to make sure that everything was spelled and annotated the same way.

But this was not to be just another IT initiative: Analytics was given a clear charter to go above and beyond business as usual. We brought in experts who really understood the plumbing of the data. Then we hired statisticians and modelers to think about what was in the data and how it could be analyzed. And of course we needed businesspeople providing guidance on the business problems that the analytics group is trying to solve.

Now we're able to deploy data in an ever-increasing number of ways. For example, a football team might ask us for advice on which artist it should bring in for a performance before an event, given the likes and dislikes of its season-ticket holders. We can do that.

We can also help clients find the best prospects for packages of season tickets or subscriptions to shows, and we can assist them in coming up with marketing messages that will resonate with these prospects.

We've found that it's very helpful to have our statisticians and modelers speak directly with clients. That allows the clients to better understand the data, and it gives the statisticians a clear view of the clients' needs.

The world of data is changing rapidly. In the next year, our big push is going to be figuring out what to do with the rising level of mobile activity. What's the best way to collect data, and how can we best use it?

And then there's social media—potentially a valuable means of communicating with consumers. But it's a challenge to link Facebook profiles and Twitter accounts with basic customer-relationship-management data such as email addresses. Social is the pot of gold that no one has yet found.

One final word: It's critical to treat consumers and their data with respect. If you're going to use their data, you have to do it right. If consumers see that their data is going to third parties that aren't relevant to them, you'll lose their trust, and they'll be quick to hit the "unsubscribe" button. Consumers' information should be deployed to provide offerings that are relevant to them; if you do that, they'll show their appreciation by continuing to allow you to use their data.

BIG DATA'S BIGGEST CHALLENGE? CONVINCING PEOPLE NOT TO TRUST THEIR JUDGMENT

BY ANDREWS MCAFEE

Here's a simple rule for the second machine age we're in now: As the amount of data goes up, the importance of human judgment should go down.

The previous statement reads like heresy, doesn't it? Management education today is largely about educating—developing future leaders' pattern-matching abilities, usually via exposure to a lot of case studies and other examples, so that they'll be able to confidently navigate the business landscape. And whether or not we're in b-school, we're told to trust our gut and instincts, and that (especially after we gain experience) we can make accurate assessments in a blink.

This is the most harmful misconception in the business world today (maybe in the world full stop). As I've written here before, human intuition is real, but it's also really faulty. Human parole boards do much worse than simple formulas at determining which prisoners should be let back on the streets. Highly trained pathologists don't do as good a job as image analysis software at diagnosing breast cancer. Purchasing professionals do worse than a straightforward algorithm predicting which suppliers will perform well. America's top legal scholars were outperformed by a data-driven decision rule at predicting a year's worth of Supreme Court case votes.

I could go on and on, but I'll leave the final word here to psychologist Paul Meehl, who started the research on human "experts" versus algorithms almost 60 years ago. At the end of his career, he summarized, "There is no controversy in social science which shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one. When you are pushing over 100 investigations, predicting everything from the outcome of football games to the diagnosis of liver disease, and when you can hardly come up with a half dozen studies showing even a weak tendency in favor of the clinician, it is time to draw a practical conclusion."

The practical conclusion is that we should turn many of our decisions, predictions, diagnoses, and judgments—both the trivial and the consequential—over to the algorithms. There's just no controversy anymore about whether doing so will give us better results.

When presented with this evidence, a contemporary expert's typical response is something like "I know how important data and analysis are. That's why I take them into account when I'm making my decisions." This sounds right, but it's actually just about 180 degrees wrong. Here again, the research is clear: When experts apply their judgment to the output of a data-driven algorithm or mathematical model (in other words, when they second-guess it), they generally do worse than the algorithm alone would. As sociologist Chris Snijders puts it, "What you usually see is [that] the judgment of the aided experts is somewhere in between the model and the unaided expert. So the experts get better if you give them the model. But still the model by itself performs better."

Things get a lot better when we flip this sequence around and have the expert provide input to the model, instead of vice versa. When experts' subjective opinions are quantified and added to an algorithm, its quality usually goes up. So pathologists' estimates of how advanced a cancer is could be included as an input to the image-analysis software, the forecasts of legal scholars about how the Supremes will vote on an upcoming case will improve the model's predictive ability, and so on. As Ian Ayres puts it in his great book Super Crunchers, "Instead of having the statistics as a servant to expert choice, the expert becomes a servant of the statistical machine."

Of course, this is not going to be an easy switch to make in most organizations. Most of the people making decisions today believe they're pretty good at it, certainly better than a soulless and stripped-down algorithm, and they also believe that taking away much of their decision-making authority will reduce their power and their value. The first of these two perceptions is clearly wrong; the second one a lot less so.

So how, if at all, will this great inversion of experts and algorithms come about? How will our organizations, economies, and societies get better results by being more truly data-driven? It's going to take transparency, time, and consequences: transparency to make clear how much worse "expert" judgment is, time to let this news diffuse and sink in, and consequences so that we care enough about bad decisions to go through the wrenching change needed to make better ones.

We've had all three of these in the case of parole boards. As Ayres puts it, "In the last 25 years, 18 states have replaced their parole systems with sentencing guidelines. And those states that retain parole have shifted their systems to rely increasingly on [algorithmic] risk assessments of recidivism."

The consequences of bad parole decisions are hugely consequential to voters, so parole boards where human judgment rules are thankfully on their way out. In the business world it will be competition, especially from truly data-driven rivals, that brings the consequences to inferior decision makers. I don't know how quickly it'll happen, but I'm very confident that data-dominated firms are going to take market share, customers, and profits away from those who are still relying too heavily on their human experts.

IS THERE HOPE FOR SMALL FIRMS, THE HAVE-NOTS IN THE WORLD OF BIG DATA?

BY CHRISTINA DONNELLY AND GEOFF SIMMONS

Here's a wishful vision of the future that's even more radical than Amazon's concept of delivery drones or Google's robots: One day, small businesses will have access to affordable consumer data.

Don't yawn. This is a life-and-death issue for small businesses. Anyone who has worked in or around a supplier to a big consumer company—to a supermarket chain, for example—knows the value of information on shoppers' preferences. If a supplier can use consumer data to shape its offerings and marketing strategies, it has a significantly better chance of survival than its data-deprived competitors have.

We saw the value of data firsthand in a study we did with Gillian Armstrong of the University of Ulster and Andrew Fearne of the University of Kent in the Northern Ireland region of the UK. We provided consumer data and analysis—free of charge, thanks to a government grant—to a group of food and beverage companies that had previously relied mainly on their managers' intuition and a little guidance from supermarkets. With training and assistance, the providers were able to see how their categories were performing in supermarket aisles and what segments of consumers were buying their products.

"Data exposure focused our feel for the market," a manager of a tea company told us. "It formed a basis for extended thinking in terms of tea product content, packaging, and design."

But our research also pointed out that when it comes to data, there are the haves and the have-nots. The suppliers we studied were small—the largest had just 45 employees—and their modest annual turnover made data prohibitively expensive. They never could have afforded the ongoing cost of the consumer data. A single analyzed report runs €7,000, and to keep up with the big firms, they'd need more than that.

Once their eyes were opened to the power of data, the firms saw immediately how great their disadvantage had been. After our project ended, they went back, as one firm owner put it, to "square one."

If it's true, as Andrew McAfee writes, that "data-dominated firms are going to take market share, customers, and profits away from those who are still relying too heavily on their human experts," then we can expect to see a very different business landscape some years down the road. It will be a landscape with many fewer of the small, artisan businesses that have been so important to societies for millennia.

Small businesses account for a large proportion of private-sector employment; in the U.S., for example, despite a vast corporate sector, the figure is 49%. Small firms are a "fountain of job growth," according to the U.S. Bureau of Labor Statistics; companies with fewer than 500 employees account for about two-thirds of net jobs created in the country. They are often great places to work, too. One study shows that marketing managers in small firms report higher levels of job satisfaction, greater esprit de corps, and more organizational commitment than their counterparts in large firms report.

Moreover, consumers enjoy the products of small businesses—some of the items provided by the firms we studied were popular in the big supermarkets because they were marked as "locally produced" or "premium." An added benefit is that these products tend to be high-margin, both for the suppliers and the stores.

Which raises a question: Should big companies assist small companies by providing them with inexpensive access to data?

It's in the chains' interest to keep the small firms alive: We doubt that consumers or grocery chains would want the artisan bakers, chefs, and yogurt makers to disappear, bulldozed away by the power of big data.

Yet when one of us, Christina Donnelly, raised this question with a supermarket-chain executive in the U.S., his response was that he had never even thought about it. The data gap and its possible consequences hadn't crossed his mind.

That response doesn't say much for the likelihood that supermarkets will ever willingly provide customer data to small suppliers at reduced cost. And in any case, the flow of data is more complicated than it may seem: Food and beverage companies that purchase customer-preference data get it not from supermarkets but from analytics firms that manage the data on behalf of the supermarkets, analyze it, and package it. Analytics firms may be even less likely than supermarkets to make the data available to small firms at lower cost (and their large customers would be irate if they did).

Is government intervention the answer? That seems unlikely, given the probable backlash. A manager of a large food company told a member of our team that even the limited, experimental government funding of data and analysis for our research project was unfair.

So where does that leave small businesses? Is their situation hopeless?

Maybe not. After all, the ultimate source of data is the consumer. Shouldn't shoppers have a say in what happens to their loyalty-card information? If they value small firms, shouldn't they be able to ensure that these firms have access to a consistent flow of market data?

That idea may seem farfetched, but so is the concept of product delivery by drones. The difference between the two is that Amazon, with all its money and power, is well capable of taking a crazy idea and turning it into reality; like many mammoth corporations, it can afford to deploy cutting-edge technologies in the pursuit of greater growth and greater dominance. Small firms simply can't do that—at least not on their own.

But in the ordinary consumer, small firms do have a powerful ally. If they could somehow tap into that power, the artisan firms might just be able to garner enough competitive advantage—or at least achieve enough of a competitive balance—to continue providing enriching, satisfying jobs and valuable products to millions of people the world over.

WEBINAR SUMMARY | DECEMBER 3, 2013

GETTING IN FRONT OF DATA QUALITY

FEATURING THOMAS C. REDMAN

Contributors

Thomas C. Redman, President, Navesink Consulting Group

Angelia Herrin (Moderator), Editor, Special Projects and Research, *Harvard Business Review*

Overview

Most organizations are too tolerant of bad data. Solving the data problem often isn't hard, but it takes courage and persistence to change the status quo. Better data quality relies on more effective communication between the creators of data and the customers of data. Leaders must also assume responsibility for data quality rather than rely solely on the IT team.

Any manager can improve data quality within his or her span of control. The key is using a process management cycle that includes customer needs, measurements, improvements, and controls. Bad data is like a virus, but quality data can serve as a source of competitive advantage.

Context

Thomas Redman, author of the December 2013 *Harvard Business Review* article "Data's Credibility Problem," described problems that organizations face related to data quality, along with potential solutions.

Key Learnings

Improving data quality requires a cultural shift within organizations.

All too often, the responsibility for data quality rests with data customers. Sustainable data quality initiatives require a culture where data consumers and creators share the responsibility for information integrity. To illustrate this point, Dr. Redman recounted two tales of data quality efforts:

- **The Rising Executive.** While preparing for an important presentation, a middle manager saw a figure that looked suspect. After some research, her assistant discovered that the data was wrong and corrected it. Although this response averted potential disaster, the manager didn't take any further measures to prevent the same issue from occurring again, leaving others to be victimized. The challenge for leadership is building a culture where failure to advise about data errors is unacceptable.
- **Data Quality:** The Movie. The leaders in an organization realized that they needed more accurate data. So they took four steps: 1) they clearly documented the data customers' needs; 2) measurements were defined; 3) improvements were made; and 4) controls were implemented. In this case, managerial responsibility for data quality was shared between data creators and customers, and future disasters were avoided. This ideal model should be applied to important data customers throughout the organization.

Establishing a process management cycle improves data quality by breaking down communication problems.

A major obstacle to data quality is departmental silos. These inhibit communication between data creators and customers. A powerful way to break down these barriers is a process management cycle that offers a repeatable technique for improving data quality. Examples include Six Sigma, Lean, and the Quality Improvement Cycle. There are four important aspects to this approach:

1. **Understand customer needs.** When it comes to data quality, only two moments matter: creation and use. The overarching objective of data quality initiatives is to connect those two moments. Data quality is defined as meeting the most important needs of the most important customers. Organizations must know who their data customers are, which are most important, and what they want. It is also important to recognize that data consumers' needs often change over time.
2. **Develop a measurement system.** The old adage "You can't manage what you don't measure" applies to data quality. Measurements should reflect the true extent of the data quality problem. For example, a team might have 10 data records, each with 10 pieces of data. Out of 100 pieces of data, four might be wrong. Yet, when a single piece of data is wrong, the entire record is unusable. In this instance, data quality should be measured at 60% (six out of 10 records are usable) rather than 96% (96 pieces of accurate data out of 100).

"Senior leaders must insist on data quality principles; otherwise, organizational momentum shifts accountability downstream to data customers."

—Thomas Redman

3. **Identify improvement opportunities.** Most improvements are relatively simple and don't require much effort. Root problems are often the result of insufficient communication between data creators and customers.
4. **Establish controls and check conformance to requirements.** People don't like controls, but they are essential. Controls keep root causes from recurring, prevent simple errors, and enable calibration when data is collected in an automated way. Conformance to requirements helps cultivate trust and confidence in data.

Senior leaders must decide which approach to use to improve data quality.

Lack of data quality can be thought of as a polluted lake. The lake represents the databases, water is the data, and streams are the sources of new data. Data pollution in this metaphor is caused by breaks in business processes. In this scenario, three options are available to improve data quality:

1. **Unmanaged.** With this approach, data customers are left to deal with inaccurate information.

This option is not recommended.

2. **Prevent errors at the source.** This option is the best, but it requires data creators to change their practices.
3. **Find and fix.** In this case, work is done to clean up the database. Teams may buy a tool to help identify and remedy errors.

Unless a proactive approach is pursued (option 2 is best, option 3 is a distant second), organizational momentum will push accountability for data quality downstream to data customers (option 1).

Everyone, not just IT, is responsible for data quality.

Everyone in an organization touches data, so everyone must play a role in data quality. Making the IT team solely responsible for data quality is a recipe that is likely to fail. It's difficult for IT to provide leadership, because it is neither a data creator nor a data customer.

Dr. Redman described how different stakeholder groups can contribute to data quality:

- **Data customers and creators.** Data consumers must make their needs clear. Data creators are responsible for understanding and meeting customer needs. This can be accomplished through process management.
- **Leadership.** Leaders must establish the right managerial accountabilities, build a culture that values data, and focus the data quality effort.
- *IT.* The IT team must lock in improvement gains, provide tools, and make sure vendor data models meet customer needs.
- **Data quality staff.** A small, focused team can make data quality easier to attain. This group provides common methods and measurements, facilitates improvements, tracks progress, provides training, runs metadata processes, and pushes other stakeholders.

"Get responsibility for data quality out of IT and into the moments of data creation and use." —Thomas Redman

Given the high cost of bad information, data quality initiatives can be a huge win.

Bad data is like a virus. There's no telling where it will end up or what damage it will do. Bad data played a role, for example, in the recent recession as falsified mortgage applications and inaccurate ratings proliferated. Almost every day, bad data puts organizations in the news.

Dr. Redman calls quality data the "gift that keeps on giving." Not only does accurate data save money, but it also enables leaders to run their businesses more effectively. High-quality data can become a source of competitive advantage in several ways:

- **Better bottom lines for low-cost providers.** Up to 20% of revenue is wasted dealing with bad data. Improving data quality can have a significant effect on the bottom line.
- **Enhanced innovation.** Opportunities for innovation are often identified through big data and advanced analytics. These insights are more promising when they are based on reliable data.
- **Compelling content.** As Michael Eisner said, "Content is king."
- **Data-driven decision making.** Up to 50% of knowledge workers' time is spent looking for data and dealing with errors. Productivity increases when data customers get out of the data quality business.

The question isn't whether organizations should worry about data quality—they obviously should—but rather where they should focus their efforts. Businesses must identify what information is most important and then make targeted investments in data quality measures.

Other Important Points

Metadata. Metadata describes where information resides and what information means. Good metadata makes it easier for people to find and interpret data. Examples of metadata include data models, data definitions, business rules, naming conventions, privacy and security restrictions, and storage or access details.

Biographies

Thomas C. Redman
President, Navesink ConsultingGroup

Thomas C. Redman (aka the "Data Doc") is the president and founder of Navesink Consulting Group. As an innovator, advisor, and teacher, Dr. Redman was among the first to extend quality principles to data and information. He has crystallized a body of tools, techniques, road maps, and organizational mechanisms to help organizations make big, sustained improvements. More recently, Dr. Redman has developed keen insights into the nature of data, formulating the first comprehensive approach to "putting data to work." Taken together, these innovative solutions enable organizations to treat data as assets of virtually unlimited potential.

Dr. Redman has helped dozens of leaders and organizations better understand data, including the importance of high-quality data, and has assisted them in starting their own data programs. He is a sought-after lecturer and author of dozens of papers and four books. His most recent book, *Data Driven: Profiting from Your Most Important Business Asset*, was awarded a *Library Journal* "Best Buy." He is also author of the December 2013 HBR article "Data's Credibility Problem."

Prior to forming Navesink, Dr. Redman established the Data Quality Lab at AT&T Bell Laboratories in 1987 and led it until 1995. He holds a PhD in statistics and holds two patents.

Angelia Herrin, Editor for Research and Special Projects,
Harvard Business Review

Angelia Herrin is editor for research and special projects at *Harvard Business Review.* At *Harvard Business Review*, Herrin oversaw the relaunch of the management newsletter line and established the conference and virtual seminar division for *Harvard Business Review.* More recently, she created a new series to deliver customized programs and products to organizations and associations.

Prior to coming to *Harvard Business Review,* Herrin was the vice president for content at womenConnect.com, a website focused on women business owners and executives.

Herrin's journalism experience spans twenty years, primarily with Knight Ridder newspapers and USA Today. At Knight Ridder, she covered Congress as well as the 1988 presidential elections. At USA Today, she worked as Washington editor, heading the 1996 election coverage. She won the John S. Knight Fellowship in Professional Journalism at Stanford University in 1989–90.

This summary was created for *Harvard Business Review* by BullsEye Resources.
www.bullseyeresources.com

LEARN MORE

FOR MORE DETAILS ABOUT SAS® DATA MANAGEMENT:
→ www.sas.com/datamanagement

READ MORE THOUGHT LEADER VIEWS ON DATA MANAGEMENT TOPICS AT THE DATA ROUNDTABLE:
→ www.dataroundtable.com