



INFORMATION SYSTEMS | MARKETING | ACCOUNTING & FINANCE  
| HUMAN RESOURCES | HEALTH CARE | SUPPLY CHAIN

KELLEY SCHOOL OF BUSINESS  
INDIANA UNIVERSITY

# OnAnalytics

Spring 2015

diagnose. predict. optimize.



**DISCOVER THE CONNECTIONS**  
Converting Data to Insights



# Welcome

Within a few short years, the use of analytics to support business decision making has moved from a cutting-edge competitive advantage possessed by relatively few companies to a standard business practice. Companies that ignore the value of business analytics will increasingly find themselves at a disadvantage. As a result, companies are not simply looking for individuals with specialized business analytics knowledge, but are moving toward the expectation that all professional employees will have the ability to understand and use big data as a basis for decision making.

The Kelley School of Business continues to be at the leading edge of incorporating business analytics into our curriculum. Today, business analytics permeates all programs. At the graduate level, we offer both a major and minor in business analytics in the residential MBA program, as well as two MS programs—a residential MS in Information Systems and an online MS in Business Analytics. We also offer executive certificate programs in business analytics. At the undergraduate level, students can choose to co-major in business analytics to gain the skills needed to apply business analytics to their chosen major. Our course offerings are ever expanding to meet the changing landscape of the business analytics field. We are committed to staying on the leading edge of business analytics, and are very fortunate to have corporate partners who work with us at every step of the way in this exciting journey.

What lies ahead for business analytics at Kelley? Beyond the continual expansion of our business analytics curriculum, we will continue to evaluate and expand the role of business analytics in our major courses. Students need application-oriented learning to help them relate business analytics tools to specific business contexts. For example, accounting majors can apply business analytics to various approaches for testing controls over the processing and recording of transactions. Applying business analytics within such a context also allows students to bring in learning from other areas—for example, business ethics—to provide a comprehensive framework for viewing the specific decision.

The Kelley School also will continue to leverage our expertise in experiential, real-world learning to enhance our students' business analytics education. Kelley students have numerous opportunities to apply their knowledge in real-world business situations, ranging from solving business problems in global locations to helping a local nonprofit organization in Bloomington. Our field study, capstone, and global immersion and social enterprise programs provide exciting opportunities for students to take ownership of a significant business decision that can be supported by data analytics.

The Institute for Business Analytics (IBA) will continue to serve in its leadership role for all aspects of business analytics at the Kelley School. One key role of the IBA is to support and highlight the research of Kelley faculty in the data analytics area. Kelley faculty are outstanding researchers who are experts in using data analytics techniques to provide insights and solutions to business problems. This issue of *OnAnalytics* highlights research in three areas: accounting, information systems, and marketing. I encourage you to discover the insights offered by our faculty's research.

I look forward to seeing you at the Kelley Forum on Business Analytics, April 9–10.

## **Laureen A. Maines**

Executive Associate Dean of Faculty and Research  
KPMG Professor of Accounting



## From the Editors

### Frank Acito

Professor of Marketing  
Max Barney Faculty Fellow  
acito@indiana.edu

### Vijay Khatri

Associate Professor of  
Information Systems  
Arthur M. Weimer Faculty Fellow  
vkhatri@indiana.edu

Welcome to our expanded issue of *OnAnalytics*! This edition includes articles from our faculty and practitioners, reports from student interns, and summaries of two panel discussions held at the Indiana University Kelley School of Business.

Last fall, the Institute for Business Analytics sponsored an *Analytics Leadership Award* competition to recognize innovative applications of analytics. A team from 3M Corporation developed a revenue-forecasting system that was more accurate and ten times faster than previous processes. A Deloitte Consulting team reduced end-to-end manufacturing at a major original equipment manufacturer. Another Deloitte Consulting team developed a system for the Internal Revenue Service to improve fraud detection.

Two panel discussions were conducted during the *2014 Analytics Forum*. The first panel provided contrasting approaches to achieving an effective analytics function and explored career opportunities. The second panel focused on trust, credibility, and communication in applications of analytics in the military and nonprofit sectors.

We continue to highlight application of analytics in accounting and finance, health care, human resources, information systems, marketing, and supply chain. Accounting Professor Teri Yohn reports on a comparison of methods of forecasting corporate profitability. She and her colleagues found that while no single approach was always most accurate, they could provide recommendations regarding selection of an optimal approach.

Steve Pratt, a principal with Deloitte Consulting, provides a perspective on the importance of focusing analytics on wellness and prevention as a key complement to treatment efforts. Naigui Chen, a Kelley MBA student, reports that during his internship at a health care organization he learned about the importance of “softer skills” such as the significance of problem definition, which is so necessary to the successful application of analytics.

Frank Bernhard, a managing director at Accenture, presents an insightful model for building analytics talent capabilities in an organization. He proposes a dual-track team approach consisting of data strategists, responsible for framing the right business questions and drawing out actionable insights from results, and data scientists, who are experts in the technical details of quantitative analysis and data modeling.

Information Systems Professor Jingjing Zhang reports that frequent version updates are an important way gaming app competitors stand out among hundreds of thousands. It has been widely observed that people are much more likely to engage online when the experience involves a competition or game. Ramanuja Atur, a Kelley MBA student, applied this notion during his summer internship. His project at a major cloud services provider used “gamification” to incentivize customers to optimize their usage of the cloud computing environment.

Marketing Professor Hai Che analyzed nearly 9,000 online sessions with a popular restaurant review site in China and found that positive reviews can lead to greater patronage. Kelley MBA intern Ajay Jain applied his knowledge of analytics and marketing to develop key insights for a computer software company that was changing its distribution strategy. Besides providing the company with an analytics model that identified key revenue drivers, he took away several important lessons from his experience in conducting a start-to-finish analytics project.

Finally, MSIS student Jesse Delpriora, working with a major financial services company, used predictive analytics to identify ways to speed up an important process related to documenting risk associated with illegal activity.

It is because of this breadth of applications that analytics continues to receive so much attention from practitioners and researchers. We hope that you share our excitement about the field of analytics and find this issue informative.



# OnAnalytics

diagnose. predict. optimize.

## INFORMATION SYSTEMS

6



- 6 **Effects of Version Updates on the Popularity of Mobile Games**  
Jingjing Zhang  
How do the number and types of version updates affect a mobile game app's position in the ranking charts?
- 8 **Should We Trust Our Instincts—Or the Data? The Role of Analytics in Decision Making**  
Four panelists discuss using analytics in the military and in the private and nonprofit sectors.
- 10 **Improving Cloud Services through Gamification**  
Ramanuja Atur  
An intern finds ways to encourage customers to engage with a cloud service provider.

## MARKETING

12



- 12 **Analyzing Restaurants Using Online Reviews**  
Hai Che  
How do online restaurant reviews provide value to consumers and businesses?
- 14 **Revenue Forecasting in the Cloud**  
Ron Hoffner  
3M has developed an innovative forecasting engine and cluster management framework to improve sales forecasting. (2014 Kelley Analytics Leadership Award co-winner)
- 16 **Projecting Software Sales through Analytics**  
Ajay Jain  
An intern helps a computer software company find opportunities for growth.

## ACCOUNTING &amp; FINANCE

18



- 18 **Forecasting Profitability through Disaggregation of Operating and Financial Activities**  
Teri Lombardi Yohn  
When is it beneficial to disaggregate financial statements into operating and financial activities for forecasting profitability?
- 20 **Combating Tax Fraud Using Advanced Analytics**  
Satish Lalchand  
A team from Deloitte has created a fraud and anomaly detection system to support the IRS. (2014 Kelley Analytics Leadership Award honorable mention)

## HUMAN RESOURCES

22



- 22 **Organizing Analytics Talent**  
Three panelists give their perspective on organizing analytics talent within a firm.
- 24 **Exploring the Bifurcated Analytics Talent Model**  
Frank J. Bernhard  
What should organizational leaders consider when developing analytic talent?

**About Deloitte**

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee ("DTTL"), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as "Deloitte Global") does not provide services to clients. Please see [www.deloitte.com/about](http://www.deloitte.com/about) for a detailed description of DTTL and its member firms. Please see [www.deloitte.com/us/about](http://www.deloitte.com/us/about) for a detailed description of the legal structure of Deloitte LLP and its subsidiaries. Certain services may not be available to attest clients under the rules and regulations of public accounting.

## HEALTH CARE

26



- 26 **Innovating Health with Wellness Analytics**  
Steve Pratt  
How can we shift health care's focus toward promoting a culture of wellness?
- 28 **Finding Insights to Fuel Company Growth**  
Naigui Chen  
An intern identifies opportunities for a global health care company to increase its U.S. market share.

## SUPPLY CHAIN

30



- 30 **Using Advanced Analytics to Dramatically Reduce Product Span Time**  
Steve Shepley  
SPANalytics® from Deloitte is a new way of thinking about materials requirements planning. (2014 Kelley Analytics Leadership Award co-winner)
- 32 **Analyzing Data for Capacity Planning**  
Jesse Delpriora  
An intern uses analytics to ensure a multinational financial services company's compliance with federal regulations.

**Disclaimer:** This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.

# Effects of Version Updates on the Popularity of Mobile Games



## Jingjing Zhang

Assistant Professor of  
Operations and Decision  
Technologies

[jjzhang@indiana.edu](mailto:jjzhang@indiana.edu)



According to a recent study, more than half of all adult mobile phone users in the United States own a smartphone. Within that demographic, Apple's iPhone holds nearly 40 percent of the market share. The iPhone's success is due in part to the vast number and variety of mobile applications, or apps, available through Apple's App Store. The mobile games category comprises over 25 percent of the 1.4 million apps currently available for download.

Because there are so many apps available, no consumer can possibly learn about every single offering. To assist consumers in the discovery process, Apple provides lists of the most popular apps as determined by the number of downloads. But what does it take to remain in these lists? Update frequency is one known factor that influences apps' popularity and is included in Apple's list-making algorithm. To examine this notion further, the researchers investigated the effect of the update frequency and type of version updates on a mobile game app's longevity at the top.

### Statement of Problem

With more than 310,000 games available for download, it is impractical for a consumer to

discover or sample all of them. As a result, consumers often rely on Apple's ranking lists, which include the top free, paid, and grossing games. Because version updates are known to affect an app's position in the top lists, app developers often wonder what an appropriate strategy might be for keeping their apps updated. While Apple has suggested general strategies for what high-, medium-, and low-frequency updates should include, it is not obvious how each strategy affects the popularity of a mobile game.

### Data Sources Used

The researchers collected daily rankings for the top free and paid game apps over a 200-day period from March 1, 2013, to September 16, 2013. From the App Store, they collected information such as description, file size, release date, and advisory rating for the top 100 free and paid game apps. They also collected version history and daily ranking information for the top 500 free and paid game apps from the AppAnnie website, which provides analytics for app developers.

After excluding 25 apps that became unavailable in the App Store and 276 apps that switched between free and paid categories during the

observation period, the researchers had a data set comprising 2,713 unique apps with a total of 32,000 versions, among which 9,000 were released during the observation period.

### Analytic Techniques

The researchers used survival analysis to examine what factors enable an app to stay at the top of the charts. Working under the assumption that not only update frequency but also update type would influence an app's longevity, they first classified each version update into one of four categories: performance, content, both, and other. "Performance" updates addressed technical issues, while "content" updates reflected changes such as adding new features. Updates classified as "both" included both performance and content updates. Finally, "other" updates had irrelevant information about the update—or none at all.

To classify the data, two researchers separately hand-coded 1,000 version updates based on the app developers' release descriptions of the versions. They compared results and refined the process until all researchers were satisfied with the accuracy of the coding used. They next used RapidMiner to build a text classification model from the resulting training set, then used five-fold cross validation to fine-tune the parameters. Their final model had a classification accuracy of nearly 92 percent, and this final model was applied to classify the remaining 31,000 updates.

Because free and paid apps represent two different markets in the App Store and app developers follow different strategies when releasing free and paid apps, the researchers conducted survival analyses separately for free and paid categories. They used the parametric Weibull hazard model, fitting it into the accelerated failure time (AFT) framework because the AFT eases interpretation of the results compared to the hazard model. The researchers ran four regression models: Top 500-Free, Top 500-Paid, Top 100-Free, and Top 100-Paid. They used these models first to determine the effect of the frequency of version updates, then to determine the effects of different types of update.

### Results

The researchers found that the frequency of version updates significantly affects both free and paid apps' survival in both the Top 500 and Top 100 charts. This implies that more frequent releases of version updates may help extend an app's survival time in the top lists. Moreover, the effect of version frequency is much stronger for paid apps than free apps. This is likely due to the fact that 90 percent of the apps in the App Store are free, making it more difficult for free apps to enter and survive on the top charts.

When examining the effect of update type, the researchers' results showed that performance-related updates had a much stronger positive effect on free apps' survival in both the Top 500 and Top 100 charts when compared to paid apps. On the other hand,



content-related updates had a more significant effect on the survival time of paid apps than free apps. In fact, for paid apps in the Top 100, performance-related updates had no effect and only content-related updates had a positive effect. This implies that consumers have higher expectation for the paid apps' technical performance than free apps. Solely performance-related updates for paid apps are not sufficient for attracting new demand for the app; instead, consumers would like to see new content continually being added to the paid apps (for example, new features or more advanced levels).

In summary, more frequent version updates can generally benefit both free and paid apps in the top charts, and this versioning effect is more significant for paid apps than free apps. Among the different types of versions, content-related updates are more crucial for paid apps than performance-related updates.

### Business Implications

The researchers' results indicate that developers with both free and paid apps should actively maintain their apps and frequently release new versions. For free apps, developers may release the app early and then continually address performance concerns in subsequent version updates to improve the app's technical quality. But for paid apps, developers should ensure high performance quality at the initial release of the app and focus only on deploying new features and functionalities in subsequent version releases to keep the app at the top of the paid charts. A history of constant bug fixes may result in negative connotations in a user's mind regarding a paid app's quality.

---

Christine Hsieh, Vijay Khatri, and Jingjing Zhang, "Managing Versions in Software Evolution: The Case of Mobile Game Apps," Workshop on Information Technology and Systems, 2014.

# Should We Trust Our Instincts—Or the Data?

## The Role of Analytics in Decision Making



### PANEL MODERATOR:

#### Van Noah

Program Director, Institute for Defense and Business

### PANELISTS:

#### Kyle D. Cattani

Associate Professor of Operations Management, W.W. Grainger, Inc. Faculty Fellow

#### Eric Hansen

Lieutenant Colonel (Retired), Sales Manager, SAS

#### Alfonso J. Pedraza-Martinez

Assistant Professor of Operations Management

#### Loren Reno

Lieutenant General (Retired), Dean of the Cedarville University School of Business Administration

On Friday, April 4, 2014, four panelists took the stage to discuss how analytics are used to make decisions in the military and in the private and nonprofit sectors. Moderator Van Noah, program director of the Institute for Defense and Business, began by introducing the panelists.

Retired U.S. Air Force Lieutenant General Loren M. Reno is an Institute for Defense and Business executive fellow and dean of Cedarville University School of Business Administration. His distinguished career with the Air Force culminated in his position as deputy chief of staff for logistics, installation, and mission support.

Retired U.S. Army Lieutenant Colonel Eric Hansen is a federal sales manager with analytical software company SAS. In the Army he served as a military intelligence officer and operations research officer.

The remaining panelists were Kelley's own Kyle Cattani and Alfonso J. Pedraza-Martinez. Cattani, associate professor of operations management, had six years of private sector practice prior to receiving his PhD from Stanford. Pedraza-Martinez, assistant professor of operations management, has done extensive research in the practice of logistics in humanitarian organizations. He received his PhD from INSEAD in France.

The panelists began discussing how analytics can be applied innovatively in the real world. Hansen remarked that, thanks to smartphones, almost anybody can calculate. "But is there data integrity? And are the right analytics applied to the data? As a decision maker, you have to be able to trust your analyst and the processes that are behind both the answers you're getting and the data that contributed."

"You'll serve leaders and decision makers who trust their instincts," added Reno. "You can inform their decisions, but the data must be credible and applicable—and based on the correct assumptions."

Having the correct assumptions is critical, Hansen agreed. He related an experience he had while building a program to predict roadside bombs in Iraq and Afghanistan. In the midst of a period when there were hundreds of bombings a day, one day there were none. The analysts wanted to find the contributing factors and determine how to repeat the success. The answer was simple: there were no U.S. vehicles on the roads that day. "This underscores General Reno's point about understanding the context," Hansen concluded.

Cattani mentioned that only in the last several decades has the availability of information systems and data brought about a significant increase in the private sector study of supply chain management. "I think today the military and private sector can learn from each other, where traditionally the private sector learned from the military," he said.

Finally, Pedraza-Martinez addressed analytics in disaster management. "Data are useful in preparation; for example, predicting the trajectory of a hurricane. In the aftermath of a disaster, however, organizations trust field personnel because information is incomplete and data can be unreliable. And in the reconstruction and rehabilitation phase, you have to trust the data and create analytical models to reach efficiency, so you can help more people."

The subject shifted to communicating with decision makers. "Practitioners must understand the data *and* the decision maker," remarked Reno. The analyst must understand the data and the assumptions and be able to effectively communicate those to the decision maker. In this way, the analyst gains credibility. "You must understand how the decision maker processes information. And listen first," Reno advised. "It's absolutely critical that you communicate in the decision maker's language."

“Data modeling is an art and a science,” pointed out Cattani. “You have to simplify things, and to do that you have to make assumptions and be able to communicate them to the decision maker. It will help the decision maker further if you can articulate the trade-offs in your models.”

Pedraza-Martinez stated, “The analyst is a bridge between the technical results of a model and the managerial decisions that need to be made. MBA graduates understand the trade-offs, the assumptions, and the results, and you have the skill to communicate those.”

The panelists then moved on to their third topic: analytics for nonprofits and disaster relief in austere environments. Pedraza-Martinez spoke first about short-term disaster response and longer-term reconstruction. “These happen simultaneously,” he said, “and they use the same supply chains. So you need a manager who can make quick decisions about disaster response and at the same time make cost-effective decisions about reconstruction and rehabilitation.”

---

“The military are often among the first responders in a disaster situation,” said Reno. “The best thing you can do is provide the leader on scene a description of the next problem he’ll face. What are the resources and trends? Where is the logistics support?”

---

“Whether you’re in an austere environment or not, whether you work for the military or the Red Cross or you’re an MBA student going to work for a big company,” said Hansen, “the key is to add value.”

At this time, Noah asked the first question from the audience, directed to Reno: “Sir, are you saying that senior management is not accountable for understanding analytics?”

After a laughter-filled pause, Reno said, “You’ve grown up with a lot more technology than I did. And you’ll have to work with your seniors, most of whom didn’t have the opportunity to learn what you’ve learned about analytics. Build a relationship and help establish yourself. You’ll add value to the organization if you can do that from the start.”

Another audience member asked what to do when the analytics contradict your instincts. Cattani replied, “I want to know more. I’m leery that there’s a problem with the model. This is a good example of when the analyst should dig deeper and explain why the counterintuitive result is arising.”

“Do that before you make any decisions if you can,” added Pedraza-Martinez. “Change your model and check your intuition.”

Reno concluded, “I’m willing to change because life, technology, and the battlefield all change. Be ready to explain the results and why they differ. I’ll be comfortable making a decision with 80 percent of the information.”

“How can defense and intelligence benefit from private sector investment in analytics?” asked Noah.

Hansen suggested that the Department of Defense (DoD) is five to seven years behind the commercial world when it comes to the adoption of analytics. “Operations research started in the Department of the Navy during World War II,” he said, “but they’re still doing extensive studies. You don’t see that in the private sector. Commercial companies are run with dashboards that provide immediate, current information.”

“There will be an opportunity to use good analytics to inform decision making regarding people, aircraft, support equipment, where we place our fuel, and the number of bases to keep open,” said Reno. “What are the criteria that will be used in making those decisions?”

The final question for the panel addressed the role of analytics in cyber warfare. “Deciding which data will lead to which information is the challenge,” said Reno. “We have to winnow through terabytes of data to find what’s useful and actionable.”

Hansen added, “Humans aren’t random. There are trends and patterns in cyber warfare, and what works tends to be repeated. Using analytics to look at trends and patterns will help you get some insight.”



# Improving Cloud Services through Gamification

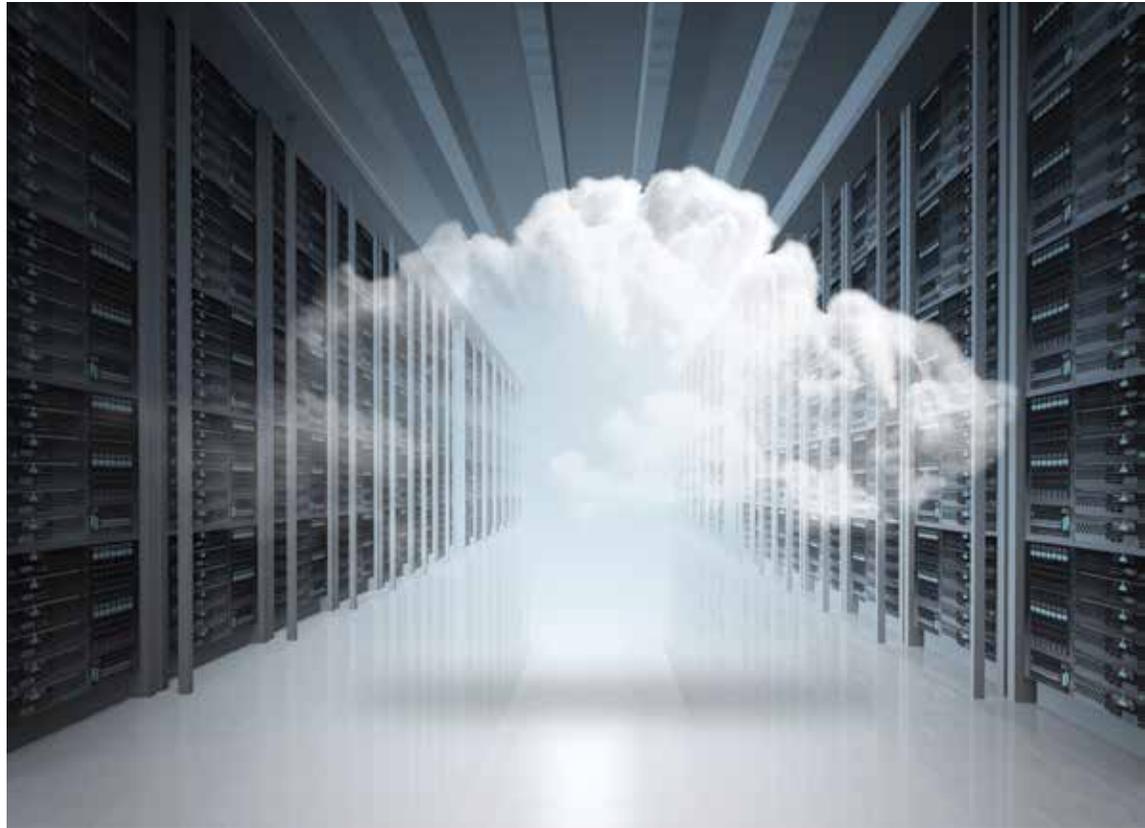


## Ramanuja Atur

MBA '15, majoring in marketing and business analytics

Interned at a cloud service provider

[ratur@indiana.edu](mailto:ratur@indiana.edu)



As more companies focus on computing in the cloud to reduce costs and to improve reliability and scalability of their IT infrastructures, cloud service providers in turn are offering more services. At one such provider, intern Ramanuja Atur was charged with incentivizing customers to take advantage of an optimization platform designed to reduce complexity.

### What were the steps taken to complete the project?

In order to understand the business, I first evaluated the need for the cloud optimization platform and the value it added for customers. Next, I performed analytics to confirm management's hunch that customers were using the optimization platform only when they were new to the cloud service.

To improve customer engagement beyond the "new user" period and incentivize customers to follow best practices, I created a gamification strategy. Using an algorithm I designed

to compute weighted averages based on the status of the cloud infrastructure, I developed a scorecard that would collate the results of all best practices the user followed and give them a corresponding score. I built incentives around increasing the score, encouraging customers to return frequently to check their progress. To make it more engaging, I also created an algorithm to compare the customer's score with those of "similar" customers. I identified these similar customers using clustering algorithms in IBM's SPSS Modeler.

Next, I worked with the company's engineering, marketing, legal, and management teams to implement the project. Finally, I created a balanced scorecard to track important metrics following the introduction of the gamification engine.



# Analyzing Restaurants Using Online Reviews



## Hai Che

Associate Professor of  
Marketing

haiche@indiana.edu



Gone are the days when deciding on a restaurant was either a choice of a familiar spot or a gamble on a new one. With numerous online restaurant review sites available, savvy consumers can make an informed choice every time. And when it comes to the effect of word of mouth, surveys have shown that online consumer reviews come second only to opinions from friends. Armed with this knowledge, the researchers set out to determine the economic value of online restaurant reviews for both consumer and restaurant.

### Statement of Problem

The success of companies that provide online reviews depends on their ability to attract visitors, and the key to attracting visitors is to provide a valuable service. The researchers' goal is to address two key management questions: How does the type of information highlighted in a review affect the economic value the review conveys? How do online reviews influence a consumer's choice of restaurant—and, consequently, the restaurant's income? Finding answers to these questions will benefit consumers, the owners of websites that

host these types of reviews, and the owners of the reviewed restaurants.

### Data Sources Used

The researchers drew their data from Dianping.com, a popular source for restaurant reviews in China. Dianping.com offers users the ability to rate a restaurant on its overall quality, taste, ambience, and service. Users can also leave comments, mark others' reviews as "helpful," and give information about the average cost of a meal.

Data from December 2007 to March 2008 were used to construct the sample. The dataset comprises restaurant reviews along with the browsing behavior and restaurant visits of consumers in Shanghai. The researchers focused on the seven most popular restaurants in the "hotpot" category.

The researchers constructed browsing sessions that captured the sequence of each consumer's visits to various restaurants' review pages, if the time between page visits was less than an hour. They then verified which restaurants, if any, that consumer visited in the next seven days. This tracking was made possible through a discount

card provided by Dianping.com—the significant savings provided by this card made it likely that any consumer who chose to dine at one of the restaurants would use the card.

Using these criteria resulted in a sample of 8,918 browsing sessions from 5,084 unique users.

### Analytic Techniques

To determine whether restaurant reviews have an impact on consumer choice, the researchers took into account important features of an online review. These features include items such as ratings, detailed comments, and information on the reviewer and the overall usefulness of his or her reviews, among others.

The researchers first used a multinomial logit regression to examine the relationship between reading a review and choosing a restaurant. Their results indicated a significantly positive relationship between reviewers' comments and consumers' restaurant choice: by using reviews to narrow their possible choices to more highly rated restaurants, users reduce their risk. With this knowledge, the researchers developed two models.

The undifferentiated learning model depends solely on information gathered from reviews, such as cost. In this model, consumers do not try to draw parallels between their own tastes and those of the reviewers; they simply update their beliefs based on the reviews they read. In this model, diversity of ratings increases consumers' uncertainty—ratings that are widely dispersed through a bell curve, rather than clustered around a single point, will less often result in a visit to that restaurant.

The differentiated learning model is more informative than undifferentiated. In this model, consumers attempt to correlate their own tastes with those of the reviewer based on the reviewer's written comments about a restaurant. With differentiated learning, diversity of ratings is no longer useful, because ratings are less informative than a consumer's own tastes. This model is novel because it implies that more information about reviewers and reviews can increase consumer utility.

The researchers used a series of what-if scenarios to calculate utility to consumers as well as changes in value following a consumer's visit to a reviewed restaurant. Results demonstrate that the differentiated learning model provides more informational value, since consumers consider their own preferences when using this model. For most consumers, the more information they can glean from a review, the less likely they are to make a choice they regret.

### Results

In the study, 15 percent of browsing sessions resulted in a consumer choosing one of the seven target restaurants. Only 4.3 percent of consumers chose a restaurant other than one of the seven. However, the majority of users didn't visit any of the reviewed restaurants.

The researchers found that nearly 89 percent of consumers tend to use the undifferentiated model when considering restaurant cost and the differentiated model when considering quality. A potential explanation for this is that quality is a subjective measure that requires more cognitive processing, while cost is objective. If a consumer believes the added value from the differentiated model is low, as it is when considering cost, he or she will not be willing to expend the resources to use that model. When it comes to quality, however, consumers are more likely to use the differentiated model to evaluate how well reviews match their own tastes.

The results show that online reviews have a significant effect on consumers' restaurant choices, because consumers are averse to the risk of choosing a poor quality restaurant. Additionally, written review comments have a stronger impact on consumers than ratings. Ratings alone are perceived as less accurate and much more diverse than comments, though they are still valuable—especially in cases where there are no written comments.

The value of restaurant reviews—to both the consumer and the restaurant—comes primarily from information about restaurant quality. Reviews reduce uncertainty a consumer has about any given restaurant. With questions about cost and quality answered by reading reviews, users are more likely to choose popular restaurants. With 30 million unique visitors each month, Dianping.com is estimated to bring an economic value of 420–450 million Chinese Yuan (CNY) to consumers each month.

### Business Implications

Both Dianping.com and the restaurants on its site can benefit from consumer reviews. With the right information about reviews (and the reviewers), Dianping.com can increase the number of restaurants willing to pay for advertising on the site, which supports the site's ability to continue to provide a valuable service.

The economic value of online reviews to restaurants is a direct result of how they influence consumers' visits. The results suggest that Dianping.com brings an increase in profit of approximately 3,600 CNY each month to each of the seven target restaurants. This significant contribution to the restaurants' income highlights the value of well-maintained review websites.

---

Tat Y. Chan, Hai Che, and Chunhua Wu, "The Economic Value of Online Reviews." Working paper.

# Revenue Forecasting in the Cloud



## Ron Hoffner

Ron Hoffner is senior manager of the Strategic Business Development group at 3M.

[rhoffner@mmm.com](mailto:rhoffner@mmm.com)



With touch points in more than 75 percent of the market, global innovation company 3M collects massive amounts of data. Until 2013, however, the company didn't have the tools in place to use all of those data. At that time, a cross-functional team brought together staff from the Financial Planning and Analysis division and from the Software, Electronics, and Mechanical Systems Laboratory to share expertise on the different dimensions of knowledge required to get more from the company's data stores. The team was given a clear purpose, a targeted goal, and defined metrics—along with a limited budget, which required them to innovate and to use open source and low-cost methods.

The Revenue Forecast project team was charged with developing an analytical tool to allow 3M to take advantage of internal data to help improve sales forecasting. The team's goal was to improve forecast accuracy to within approximately  $\pm 5$  percent, and to do so more than 50 percent faster than the existing tools. When all the pieces were in place, they achieved twice the success of the existing method, in under a tenth of the time.

The team's first task was to develop a new technological process to rapidly sift through large datasets in order to identify leading sales indicators. They knew they were on the right track when they were informed that the same key indicators had been used to determine the company's trajectory. This increased their confidence in the methodology they had developed—analytics often involves codifying the internal thought process behind someone else's prior successful knowledge.

The team's next step was to develop an innovative forecasting engine. They needed not only to try to match as closely as possible the ways users are accustomed to looking at data, but also to develop a new back-end process capable of handling the massive amounts of forecasts required by their model.

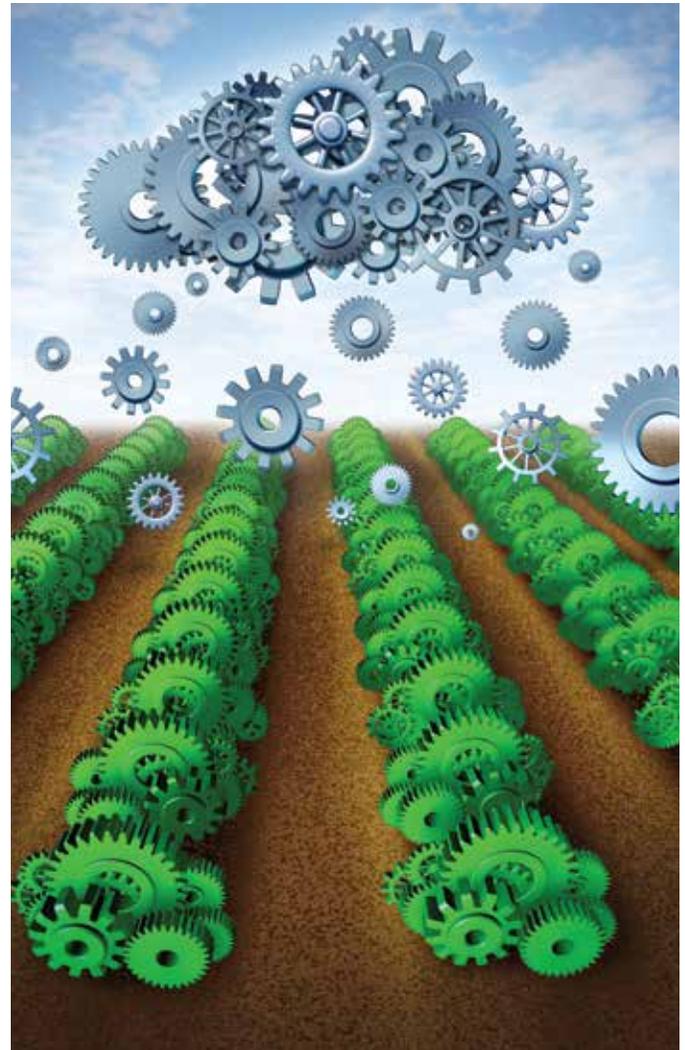
With almost 11,000 individual forecasts needed each period, the team required massive computing power. They quickly determined that they must move their model into the cloud to be able to run the level of forecasting required in the time allotted. Additionally, current framework systems weren't designed to handle the level of complexity required, so they developed an innovative cluster management framework.

The Revenue Forecast analytical model forecasts the baseline—the trajectory the business will follow if the company declines to intervene in any way. Using the tool to model the base business means that organizational leaders are able to focus on activities that will drive the business higher than base. They can adjust for variables that can't be forecasted, such as sales from new products, acquisitions, and more effective promotions.

It's vital that organizational leaders understand what modeling can and can't do, either because of lack of data or inability to use those data for forecasts. For example, how do you forecast sales on a product you've never sold before? In the future, models will need to be integrated into crowdsourcing models to handle analytics in areas where data don't currently exist.

Ron Hoffner, senior manager, 3M Strategic Business Development, leads the team. Other team members are Jon Arthur and Nate Smith, software and database engineer.

The Revenue Forecast project was a co-winner of the inaugural Analytics Leadership Award. While the award was for the first generation model, the team has already rolled out its third generation. The current model is available to and used by the entire 3M organization. The project's next step is to redesign the forecasting process to take advantage of this new tool.



# Projecting Software Sales through Analytics



## Ajay Jain

MBA '15, majoring in marketing and business analytics

Interned at a large U.S.-based computer software company

[ajayjain@indiana.edu](mailto:ajayjain@indiana.edu)



When a company began transitioning their sales from a boxed software product to one that is subscription- and cloud-based, company leadership wanted to determine where true growth opportunities might lie. Intern Ajay Jain took ownership of this project, developing a lifetime value model to project revenue and a prospecting toolkit to prioritize sales opportunities. By identifying three revenue drivers, Ajay helped the company discover a new segment of educational institutions that were potential customers.

### What were the steps taken to complete the project to gauge customer lifetime value?

There were five distinct project phases. First, I did my due diligence and reviewed other forecasting models across the company. Next were data collection, which involved identifying internal and external data sources, and data cleanup, which involved transforming variables to ensure high-quality data. Then I tested and refined the process by running a series of regressions to determine key revenue drivers

and to develop a forecasting formula. I then interpreted the results based on demographics and the type of contract that had been signed with the client. Finally, I created the model by using key revenue drivers to forecast the next three years of revenue and create a scoring system to identify growth opportunities. With this model, the company was able to view information on an individual basis or see how data looked regionally across the United States.

### How would you address challenges with data in future projects?

First, I would make sure to find out where data are stored and request access as early as possible. Because data may be stored within different departments, it's important to know the roles of each group within the company.

Data consolidation can be an issue. Data cleanup is sometimes necessary to get the data to a point where you are able to analyze them. Even with a single source of truth for the



primary data, I was also incorporating data from an external data bank. Keeping data organized and assigning unique identifiers are crucial to establishing data integrity.

#### How would you address challenges in applying analytic techniques in future projects?

I recommend always keeping stakeholders in mind. What may work for a finance team may not work for a sales team. Also, align metrics and develop a model with the group's needs in mind. My initial goal was to develop a lifetime value model, but after speaking with the sales team it seemed useful to prioritize sales opportunities. The final deliverable was more than a lifetime value model—it was a prospecting tool.

It's also important to focus on how the group understands data. For my project, I overlaid data on a map to visualize it for regional sales managers. Had the sales team operated on a product level instead, I would have offered different visualizations.

#### How did your courses in Kelley prepare you for this project?

Learning the basics of the marketing funnel and a lifetime value model in my core classes and analytic techniques in my elective classes helped me tremendously. Some of the courses that stood out for me were Introduction to Spreadsheet Modeling, Data Warehousing: Concepts and Management, and Applied Marketing Research.

#### What did you learn from this experience?

My internship was a phenomenal experience where marketing strategy and analytics intersected. It was also a way to gain experience with using different statistical packages, such as SPSS. And I learned that it may be necessary to find data according to your specific use—it's up to you to analyze data in a way that tells a relevant, practical story for your business division.

# Forecasting Profitability through Disaggregation of Operating and Financial Activities



**Teri Lombardi Yohn**  
Professor of Accounting

tyohn@indiana.edu



Accounting research proposes that disaggregating financial statements into operating and financial activities is beneficial when determining a firm's value. Similarly, standard setters have proposed requiring firms to use this disaggregation, stating that doing so would improve profitability forecasts by allowing investors to more easily discern differentially persistent profitability components. Despite the widespread support for this disaggregation, there is little empirical evidence supporting its usefulness. Given this lack of evidence, the researchers investigate the usefulness of disaggregating operating and financial activities in the context of forecasting profitability.

## Statement of Problem

Both the Financial Accounting Standards Board (FASB) and the International Accounting Standards Board (IASB) have proposed that firms be required to use the operating/financial (hereafter referred to as OpFin) disaggregation to improve investors' forecasts of firm profitability. Currently, the

FASB requires firms to use the unusual/infrequent (hereafter referred to as UnInfreq) disaggregation, while the IASB has historically been opposed to its adoption.

U.S. GAAP and accounting research support the use of the UnInfreq disaggregation for forecasting profitability. However, the OpFin disaggregation may improve forecasts for firms that earn different rates of return on their operating and financial activities.

To determine the suitability of the OpFin disaggregation for forecasting profitability, the researchers answer two questions:

- Does the OpFin disaggregation provide more accurate forecasts relative to an aggregate forecasting model, the UnInfreq model, or both?
- When does the OpFin disaggregation improve forecasts relative to the benchmark models?

## Data Sources Used

The researchers selected from Compustat an initial sample of 131,018 firm-year observations from 1972 through 2005. They then chose to eliminate firms based on certain criteria, such as those that did not separately report interest income and expense, those classified as financial firms, and those that did not meet certain financial thresholds in the year prior to the observation year. After applying these and other criteria, the researchers ended with 35,199 firm-year observations used for in-sample analysis and 36,364 firm-year observations used for out-of-sample analysis.

## Analytic Techniques

To form a comparison, the researchers used an aggregate approach and a components approach for both the OpFin and UnInfreq disaggregations. The aggregate approach creates a forecast by estimating the relationship between return on equity (ROE) and several factors in a single time-series regression. In contrast, the components approach separately forecasts each factor, or component, and then combines the results into a single ROE forecast.

In the aggregate approach using the OpFin disaggregation, the researchers used a regression to examine the relationship between ROE and return on net operating assets (RNOA), net borrowing cost (NBC), and leverage (LEV). In the components approach, they first separately forecasted RNOA, NBC, and LEV, and then combined the results of these component forecasts to develop the ROE forecast.

In the aggregate approach using the UnInfreq disaggregation, the researchers disaggregated lagged ROE into ROE before unusual and/or infrequent items (ROEBUNIN) and unusual and/or infrequent items scaled by average book value of equity (UNINFREQ/BVE). In the components approach, ROEBUNIN and UNINFREQ/BVE were forecasted separately, and then the results were combined to create the ROE forecast.

The researchers next created a combined aggregate model and a combined components model, merging the OpFin and UnInfreq disaggregations for each model. These combined models allowed them to examine whether the current requirements for the UnInfreq disaggregation are improved by incorporating the proposed OpFin disaggregation.

## Results

The researchers' results indicated that, overall, the components approach is more accurate at forecasting profitability than the aggregate approach. Within the components approach, the OpFin, UnInfreq, and combined models all showed improved forecast accuracy relative to the benchmark model that used only aggregate information. When compared to the UnInfreq disaggregation currently required by U.S. GAAP, however, the proposed OpFin disaggregation produces less accurate profitability forecasts. The combined components model is the clear winner of the three, improving forecast accuracy relative to the UnInfreq disaggregation.

Using the aggregate approach, the OpFin disaggregation and the combined model both yielded significantly less accurate forecasts compared to the UnInfreq disaggregation alone. The researchers' results support using the OpFin disaggregation only if using a components approach, and then only in combination with the UnInfreq disaggregation.

In particular situations, the OpFin disaggregation does provide forecast improvements. These improvements are generally explained by growth and accounting conservatism (that arises through accounting principles that result in firms undervaluing their operating income while correctly reporting their financial income).

## Caveats

It is difficult to predict precisely how a requirement to use the OpFin disaggregation would affect the way a firm chooses to disaggregate its operating and financial activities. Additionally, accounting and economic factors that differ from one industry to the next may introduce some variability in the relative usefulness of each disaggregation's effect on profitability forecasts. For example, one industry may be more affected by conservative accounting, while another may be more prone to persistent economic rents. This means that industries with greater barriers to entry or more conservative accounting practices may see a greater benefit from using the OpFin disaggregation than other industries. In fact, industries such as consumer staples, health care, and telecommunication services did show a more significant improvement relative to the benchmark model when the researchers used the OpFin disaggregation with a components approach. This variability in results indicates a need for researchers and practitioners to choose the optimal forecasting approach before determining the relative usefulness of a given disaggregation for forecasting profitability.

---

Adam Esplin, Max Hewitt, Marlene Plumlee, and Teri Lombardi Yohn, "Disaggregating Operating and Financial Activities: Implications for Forecasts of Profitability," *Review of Accounting Studies*, 19 (1), pages 328–362, 2014.

# Combating Tax Fraud Using Advanced Analytics



## Satish Lalchand

Satish Lalchand is an analytics director for Deloitte Transactions and Business Analytics LLP.



The Internal Revenue Service (IRS) has successfully prevented the disbursement of billions of dollars in fraudulent refunds using its Electronic Fraud Detection System (EFDS). Over time, however, the EFDS has been deemed “too risky to maintain, upgrade, or operate beyond 2014.”<sup>1</sup>

With this knowledge, the IRS set out to develop the Return Review Program (RRP)—an automated data analytics system that enables the IRS to prevent, detect, and resolve issues that could lead to the distribution of refunds to tax filers who are not entitled to that income.<sup>2</sup> The IRS estimates annual tax fraud at more than \$19.2 billion, so the results of a successfully implemented RRP should be significant. RRP is expected to increase the dollar amount of

fraudulent tax refunds identified each year while at the same time helping the IRS reduce inappropriate refund payments.<sup>3</sup>

A team from Deloitte Transactions and Business Analytics LLP and Deloitte Consulting LLP was brought in to create a fraud and anomaly detection system to support the IRS in its efforts.

Using advanced analytics techniques such as predictive modeling and link analysis, RRP helps the IRS detect fraudulent and noncompliant tax returns in near-real-time. For tax year 2013, the RRP system evaluated millions of personal tax returns, contributing significantly to the fraud detection efforts at the IRS.



# Organizing Analytics Talent



## PANEL MODERATOR:

### Jeff Kavanaugh

Vice President-Managing Partner,  
Infosys

## PANELISTS:

### Tony Saldanha

General Manager, Global Business  
Services, Procter & Gamble

### Abhishek Soni

Lead Market Analyst, Diabetes Care,  
Roche Diagnostics Corporation

### Barbara S. Williams

Director of Information Management,  
Hill-Rom Holdings, Inc.

The 2014 Forum on Business Analytics started on Thursday, April 3, with a panel on organizing analytics talent. Panelists were Tony Saldanha, general manager of Procter & Gamble's Business Intelligence and Consumer Services division; Abhishek Soni, lead market analyst for Roche Diagnostic's Diabetes Care North America division; and Barb Williams, director of Hill-Rom's Information Management division. Jeff Kavanaugh, managing partner of Infosys, served as panel moderator.

Each panelist presented a different perspective on the journey toward a sustainable analytics model. Soni, with five and a half years at Roche, explained that a new manager at Roche had advocated using and sharing data more effectively. As a result, the Business Analytics and Strategic Insights group was born with the charter of following the data wherever they lead. "We don't go into the data saying 'This is what I want to do,'" Soni said. "We look at our goals and our business needs and ask, 'How can the data inform these goals?'"

Williams spoke next. She began her journey seven years earlier when she was hired to rebuild Hill-Rom's data warehouse. It was vital that they address challenges for the data stewards, the IT organization, and power users embedded throughout the business. "Our power users truly understand the business processes and the data. It's very important that we get information out into the business where they can use it," Williams concluded.

Finally, Saldanha demonstrated the capabilities of P&G's business spheres, immersive visualization environments where users are surrounded with data. The business spheres allow decision makers to examine real-time information from the company's locations worldwide while videoconferencing with one or more of those locations. "Although we've been at this a long time, we're perhaps five percent of the way into our journey," he said. "There are incredible opportunities here and a lot to be learned."

A question-and-answer session followed these opening statements.

**Q: Are you seeing changes in the role of IT as you roll out business analytics across the organization?**

Soni: "I'm seeing more of an interaction between IT and the business. The two groups work together to address issues, rather than IT giving data to the business and then stepping away from the decision making."

Williams: "The partnership of IT with the business is very strong. This partnership drives great business solutions supporting customers, company executives, and all levels of the business."

Saldanha: "IT has come from the data center into the board room. The IT person is the trusted advisor. They're seen not as IT, but as the person who knows the business and just happens to be good at technology."

**Q: How do you know how much to invest in analytics?**

Saldanha: "We act as a business. Each year, business leaders choose how much of a particular service they'll consume, and each service has a price. If they don't want to use that service, they don't buy it."

Soni: "Our group was envisioned as having each of our three analysts support a different function. We've been able to show the business what trends are occurring and why, and what we can do to halt or continue them. This has helped us make a strong case for adding to our team."

**Q: Is there a career path out of analytics?**

Williams: "There will always be a career for someone with analytical skills. Information is critical and used by all levels of the company, making the field of analytics exciting right now."

Kavanaugh: "Engineering used to be a great foundation for a career. You might think of analytics as the next generation of that model."

**Q: How are you driving your resources toward balancing the range of skills required to support analytics initiatives?**

Saldanha: "Typically about 1 percent of our analysts qualify as 'CIO circle.' Among the rest of the analysts, some have technical competencies and some have business competencies. People call on subject matter experts as necessary."

**Q: Do you drive to cross-train or do you drive for deeper competencies?**

Saldanha: "We have to do both. What we do drive across is a basic level of skills in analytics and in communications and storytelling. Everyone gets that. But once you deepen your competency, you have to get very deep."

**Q: Would you rather hire from within or directly from an analytics school or other company?**

Williams: "We will hire the best candidate for the position—sometimes this is from within the company and sometimes we hire outside the company. Students with analytical skills that are acquired from schools like Kelley are assets. What we don't get with external candidates is the business knowledge. Candidates who are analytical and who understand technology will grow to understand the business."

**Q: What about turnover?**

Williams: "We have eight people in IT who support business intelligence. Several of the associates have found career growth by moving into one of the business units utilizing their analytics. This has been a win for the individual as well as for the business area they move into."

**Q: What's the overall degree mix you prefer and what departments do you recruit from?**

Soni: "We look for mostly master's degrees. An engineering background and even a slight inclination toward business both help."

Saldanha: "P&G is one of very few companies that promote from within. As a result, all of our analytics recruits are entry level, and we recruit in every country in the world. This isn't representative of most companies; what my colleagues describe is what you should expect to see."

**Q: What would be the key roles you'd fill first in a new analytics organization?**

Williams: "Analytical specialists, business analysts, and technical analysts. I'd love to get someone with a Lean Six Sigma background because they understand processes and business data."

**Q: How is the supply and demand? Do you see skills that match job openings?**

Saldanha: "We hire at entry level and then develop skills. I'm not looking for a data scientist, but for someone who is good raw material with a solid foundation of the fundamentals."

Williams: "We've struggled with finding the right resources, though some of that is likely due to our location in Batesville [Indiana]. It's not uncommon to wait six to eight months to hire someone into an opening."

Soni: "It's been a little challenging for our more experienced analyst roles. But in earlier stages, such as a consultant who would work with an analyst, there's more supply."

Saldanha: "We're at the cusp of a really big explosion of information. If you want long-term job security, this is the field to be in."

Kavanaugh then asked for closing statements from each of the panelists.

"This is an exciting time to be in this field," Soni remarked. "I'm looking forward to the challenge of making sense of really big data."

Williams said, "Getting information quickly to the right person continues to evolve. There's a lot of opportunity in this field. I look back at when we first started, and we're leaps and bounds ahead of where we were. It's only going to get better."

Saldanha concluded the session by thanking the audience, adding, "If you have any good ideas on how to improve the work we do as an industry, we're always open."

# Exploring the Bifurcated Analytics Talent Model



## Frank J. Bernhard

Frank Bernhard, managing director, is Accenture Analytics' global Information and Analytics Strategy practice leader and responsible for delivering advanced analytics and data science capabilities across the firm's operating client portfolio.

[frank.bernhard@accenture.com](mailto:frank.bernhard@accenture.com)



In the last ten years alone, an explosion of data coupled with scalable computing power has driven the idea that businesses can find significant value in analytics. When it comes to determining how best to derive insight from these massive amounts of data, organizational leaders should keep in mind two questions:

- What should I do to ensure my organization has the right people in the right places to deliver the right performance?
- How do I think in terms of what data individuals need to perform analytic insight responsibilities?

To most effectively address these questions, organizational leadership should consider implementing a *bifurcated talent model*, which separates the tasks of solving a dilemma through use of analytics from the deeper modeling of and hands-on work with data. This novel approach means hiring separate

skill classes of both data strategists and data scientists, who then work together to find the answers to business questions.

*Data strategists* are those who know the right questions to ask and who are the producers of consumable insight for the core business functions. They seek out the data necessary to address a particular question and inform the data scientists as to what's required to solve that challenge. They work closely in lockstep with the data scientists to fine-tune models that deliver the insights that matter most to an organization executive or knowledge sponsor.

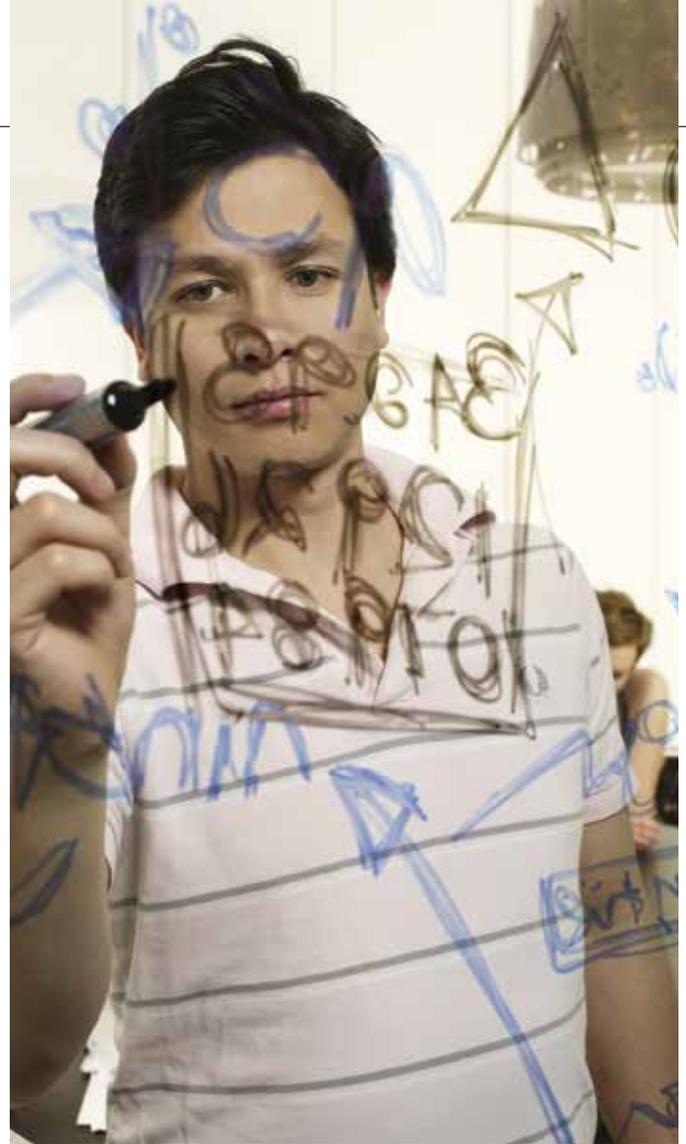
On the flipside, *data scientists* are much more technical and steeped in the rigors of applied quantitative methods. They deliver the ability to scour an organization's data stores and determine how to extract value from a particular dataset. They build the often sophisticated model that answers the questions being asked by the data strategist.

For example, suppose a consumer has questions about their mobile phone bill. They don't care how the data that underlie the bill are prepared, culled, separated, or modeled. What they care about is whether they're going over their allocated bundle of minutes or spending too much money in proportion to their economic utility. A data strategist knows that the consumer may pose these questions—and potentially a host of others. The strategist typically is the one who determines how to use data to create the most meaning for the customer and bring insight to action. He or she takes this fundamental knowledge of the required data to the data scientist, who mines the data for the appropriate dataset and then uses that dataset to prepare the model. While the strategist understands how to convey the information that the consumer wants, the scientist takes an augmented view of the data itself. The strategist and scientist must work in parallel to accomplish the end goal of translating these raw data into contextualized information that is both useful and relevant to the consumer.

A driving force in producing successful analytic consumption is awareness of the functional requirements that fit specific job categories. Across most firms and organizations, data insights and the utility of data are seldom the same for a given community of business users. That is where the *analytics continuity model*—or ACM—explains how we can make the information gleaned from analytics relevant to multiple audiences. Information is the key to making decisions, and most people need just enough information at just the right moment. For example, data-savvy consumers looking for the best wireless service can immediately have at their fingertips all of the insight-driven recommendations such as pricing and options. Awareness of this information has a significant influence on their purchasing decisions and outcomes of customer retention for a mobile operator.

Or consider how a customer service representative at an inbound call center requires a different set of information to meet their business goal of finding the best, most appropriate service for a particular consumer. The representative needs a readily available presentation layer of analytic insight that summarizes everything about the consumer and about the products and services offered by the company—but they don't need to know anything about the way the data are modeled.

The person who runs the company's customer care programs requires yet another set of information, such as how the customer service representatives—as a talent category—are performing against quality metrics. Because the representatives produce satisfied (or dissatisfied) customers, potentially affecting the company's bottom line, there must be a way to measure performance. From individual performance data, they can draw out insight about how the call center as a whole is performing.



These are just three examples of how different spans of users require a different stratum of insight capability. Analytics allow us to fine-tune our organizations. As data becomes available more quickly, it will affect how decisions are made. From petabytes of data, data strategists have the daunting task of determining what matters. They'll ask the questions and figure out what data are required, then work with the data scientists to extract and deliver results with impact.

Making analytics relevant to different audiences is widespread throughout every organization. Individuals in business units apply data with different degrees of intensity, but it all comes back to the root: how do you get at that data to produce a measurable difference? The analytics continuity model underscores the importance of why an organization's mix of data strategists and data scientists must work in concert to produce insight that leads to action. It's a delicate balance that organizations should always seek to maintain.

# Innovating Health with Wellness Analytics



## Steve Pratt

Steve Pratt is a principal in Deloitte Consulting LLP's Life Sciences practice with an emphasis on pharmaceutical research and development, focusing on information management and analytics across the health care continuum.

[stevenpratt@deloitte.com](mailto:stevenpratt@deloitte.com)



The digital age has enabled a new asset in fighting disease with the proliferation of electronic medical records and other forms of computerized clinical diagnostics and records. Furthermore, in today's connected world, many people are becoming more comfortable with putting their health information online. This movement from paper to digital enables analytics that may have never before been possible in health care; however, most of today's analytics are done on the sick, injured, and unhealthy. Additionally, most companies use only a fraction of the vast repositories of clinical information available to them. How can the industry leverage broader health care analytics to shift health care from its focus on treating the sick toward promoting a culture of wellness?

Patients, providers, payers, and life sciences companies exist inside a complex ecosystem that struggles to maintain equilibrium of checks and balances. Traditionally, the payer has acted as a hub of activity with significant control over the interactions of the other

parties. Life sciences companies are influenced largely by the opinions of providers and payers to determine which products would be commercially viable. As we open more digital channels for information about health care, we increase opportunities for patient education and direct-to-patient interaction across the entire ecosystem. By turning patients into consumers of health, we are changing the patient's role from passive to active. Most providers, payers, and life sciences firms have an opportunity to incentivize and influence a patient's attitude and actions toward health, while the patient has the opportunity to provide more regular and direct feedback about treatments and services provided. Analytics on the clinical and patient-provided data can help us understand and measure the effectiveness of the relationships among all of these parties.

Health care analytics leverages four important big data capabilities:

*Text analytics* involves assessing text, video, and other non-numeric content. Taking cues from social media, we examine the sentiment behind what patients and physicians are saying and use those concepts to help develop ontologies. This common language can enable providers, payers, and life sciences companies to monitor communications effectively.

*Unsupervised analysis* involves using technology to find patterns. The goal is to use past patterns to create algorithms that can predict future outcomes or have the proper persons notified when a bad decision is made. For example, information from a patient's prescription fulfillment and processed pharmacy claims can be used to analyze for potentially dangerous drug interactions or anomalies such as a patient trying to refill a prescription too soon.

*Advanced visualization* involves delivery of massive and/or complex datasets to decision makers. It enables them to see a lot of information quickly, which can lead to faster decisions based on better data. Visualization of the data can more efficiently identify patterns and outliers across datasets, which previously would have required years of analysis to drive hypotheses for more detailed research.

*Path analysis* involves sequencing and episode analysis. Being able to capture the path of patient care from beginning to end is important; however, most applications capture treatment as a series of individual, unrelated events that happen over time and across different providers of care. Path analysis can enable us to relate these events to each other and, as a result, to discover patterns and anomalies in disease management. Payers could use these analytics to determine the best combination of clinical and financial outcome for an individual patient and what they should be paying.

However, analytics should be used to drive better decisions that can prevent negative outcomes from happening. For example, life sciences companies can use analytics to create specialized research and development, leading to more effective treatments with fewer side effects. By analyzing diagnostic biomarkers in blood, DNA, urine, and tissue analysis, firms can tailor medicines to individual patients. Eventually, analytics will likely enable patients and their health care providers to use diagnostics that indicate likelihood of certain diseases to encourage behaviors that can delay or prevent the onset of the disease altogether.

Changes in health care can bring providers and payers together to work more collaboratively. As the data shared between the provider and payer becomes more accurate and comprehensive, it will become more possible to create payment arrangements based on a particular course of care rather than for each and every pill, diagnostic, and procedure, helping to reduce overall administrative



costs and to create more consistency in care pathways. Providers can also work more closely with life sciences companies by sharing anonymized clinical information from electronic medical records. Analysis of this information can lead to better understanding of how a drug is used in real clinical settings and how variations in care affect efficacy, and may help improve the effectiveness of these drugs. In summary, with greater access to and analysis of clinical information, both life sciences companies and payers can have a better grasp of how treatments are working in real practice, and providers may be able to develop a treatment plan that balances decreased costs with more effective outcomes.

Additionally, the opening of data and communication channels between provider, payer, and life sciences organizations can provide beneficial information and guidance to the patient, driving home the point that the patient could soon have the greatest impact on health care. Payers should focus on wellness and patient engagement, creating incentives to encourage members to be more aware and proactively take better care of their health. Life sciences firms can also directly influence patient behavior through engagement programs that not only educate how a drug should be used but also encourage supplemental disease management activities such as nutrition and exercise. Providers can use digital channels for telemedicine and other virtual interactions to engage a patient remotely and capture the results of that interaction to evaluate and adjust treatments. These changes in relationships and increased opportunities for engagement are important steps towards moving health care from focusing on treating illness to proactively guiding wellness and disease prevention.

# Finding Insights to Fuel Company Growth



## Naigui Chen

MBA '15, majoring in marketing and business analytics

Interned at a global health care company

[naiguchen@indiana.edu](mailto:naiguchen@indiana.edu)



During his internship with a major health care company, Naigui Chen worked on six different projects identifying opportunities to increase U.S. market share of a global top brand in the consumer self-diagnosed medical device division. These projects involved understanding consumer usage frequency and brand switching behavior, measuring marketing campaign effectiveness, and other aspects of marketing.

### What were the steps taken to complete the project?

In each project, the first thing I did was define the problem. This involved gathering information and getting opinions from the internal clients and opinion leaders in the company. For example, when asked to assess the potential market share threat of disruptive technologies, I spoke with my line manager and the marketing manager to understand their goals and expectations. I then spoke with marketing managers who work on competitive intelligence, both to get their perspective on the business situation and to ask for advice on a good direction for my

efforts. Using this information, I developed a framework for the project.

Next, I gathered the data to solve the problem. Though the company spends millions of dollars on data each year, when it comes to competitive intelligence many data are inaccessible. I either used proxy measurement or relied on third-party estimation to fill in these gaps.

Analyzing the data and looking for insights came next. These are often the most time-consuming processes. Because the company does not have any advanced analytics software other than Excel and Tableau, analysis was a challenge.

Finally, I presented my findings to internal clients and executives. I sometimes presented multiple times to different levels of management to ensure that all managers had the opportunity to assess my conclusions before they were used to make decisions.

### How would you address challenges with data in future projects?

I would spend more time defining the problem, because understanding the business issue and developing an effective framework can be instrumental to a project's success. Defining the problem takes not only knowledge of the business, but also communication skills and intellectual curiosity. And knowing which data to analyze can save a lot of time. Understanding the problem clarifies the goal, which greatly improves efficiency when working with data.

### How would you address challenges with communicating quantitative results to management in future projects?

I will spend more time talking with managers to understand their existing beliefs and perceptions. In one project, I spoke with marketing managers at different levels and on different product lines and sat in on marketing team meetings. After creating my first version of a data model and testing it among a small group of managers, I used their feedback to improve the model so it would better suit their needs. The model went through many refinements before it was distributed to all end users.

### How did your courses in Kelley prepare you for this project?

The data analytics courses taught me useful skills for analyzing data. Proficiency with Tableau, Excel, and SPSS is vital in any analytics job, and I learned these applications at Kelley.

Marketing, finance, and other business-related courses are equally important. Kelley's first-year core courses and my Consumer Marketing Academy and PLUS Life Sciences Academy prepared me in a wide range of business topics that were important to my success in my internship.

### What did you learn from this experience?

Data analytics is challenging—it's not just about using software. It takes data analytics skills, business acumen, and communication skills to be successful. I learned to look at analytics projects from a broader perspective. In the future, I hope to work with people from more diverse backgrounds who can bring different perspectives to business issues.

A data analytics internship is good training, even if you're not interested in data analytics as a career. The analytics department has abundant information, giving you a good advantage when it comes to learning about the business. Collaborating and presenting sharpen teamwork and communication skills, which are vital in any MBA-level job.



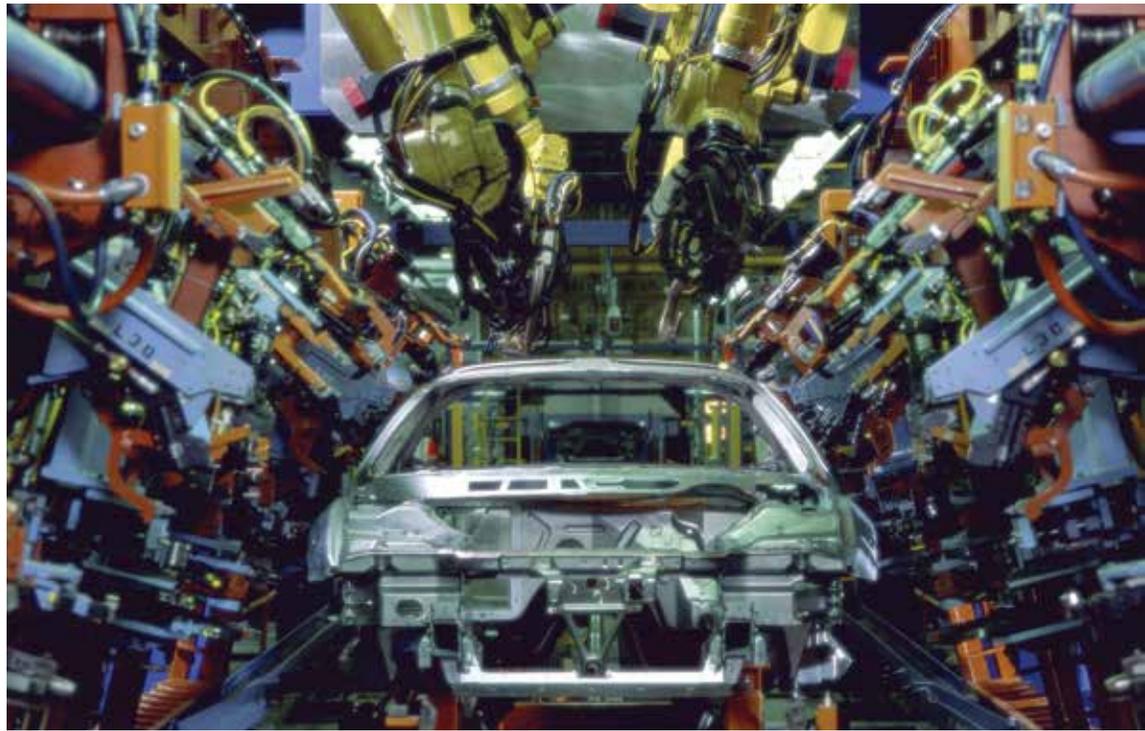
# Using Advanced Analytics to Dramatically Reduce Product Span Time



## Steve Shepley

Steve Shepley is a Strategy and Operations principal with Deloitte Consulting LLP, specializing in the design and development of advanced analytical solutions that help solve difficult business issues faster and more effectively.

[sshepley@deloitte.com](mailto:sshepley@deloitte.com)



For an original equipment manufacturer (OEM) that produces highly engineered products, an efficient span time is vital. Span time, defined as a product's end-to-end manufacturing time (sometimes referred to as lead time), is a critical measure, as reductions can simultaneously reduce cost, improve market/schedule flexibility, and decrease working capital requirements, potentially allowing OEMs to be more agile and more profitable in the marketplace. To achieve these goals, one particular OEM realized it needed to significantly reduce its span time.

Unfortunately, for most OEMs, efforts to reduce product span time typically fail. While the failure has many drivers, the most frequent driver is the inability to truly understand the complex, often misleading ecosystem of relationships created by the bill of material (BOM)—resulting in too few actions, taken in the wrong places and using the wrong improvement targets. Current materials requirements planning (MRP) models are

at the root of this issue, as MRP data is the foundation for most efforts and utilizes deterministic models based on static master data. This means MRP models simply ignore the uncertainty and randomness of everyday business events and variations that make it misleading for these types of tasks. To address these issues, Deloitte Consulting LLP has developed the SPANalytics® analytics solution to help this OEM client.

SPANalytics® is a new way of thinking about MRP. The traditional MRP approach assumes that events will happen precisely as planned, ignoring the countless variations that actually occur—such as suppliers' missing delivery windows, complexity on the factory floor, or unacceptable product quality. SPANalytics® uses advanced analytic techniques to help leaders better understand the relationships among production activities and the methodical approach required to reliably drive down span time. At the root of the solution, it transforms MRP from deterministic to

stochastic and visually portrays the “formula” to compress span by giving leaders the ability to sequence activities with the highest probability of falling on the critical path or causing a production disruption.

The complicating factor to the approach above is the fact that span time comprises thousands of parts, thousands of production steps, thousands of production hours, numerous production facilities, hundreds of suppliers, tens of thousands of data fields, and thousands of logistics moves. Simply calculating span time (for each BOM note) is often so complex and data intensive that no individual has the tools to model the actual delivery network and resulting span time. And the static data mentioned above is not enough—actual performance and performance distribution must be incorporated into each of the thousands of nodes in the calculation. In other words, it must be a stochastic calculation rather than a simple deterministic one, and the algorithm must include techniques to compensate for poor quality data. These complexities make driving reliable results difficult and are the key barriers the SPANalytics® solution can overcome.

The SPANalytics® initiative included four elements:

- Bill of material (BOM) transformation
- Performance risk adjustment and data enrichment using big data
- Simulation and predictive analytics
- Advanced visualization outputs and interactive charts

In the BOM data transformation step, the team created an algorithm to condense information from millions of pieces of data into a relational set of data called “strings.” A string represents an entire sequence of events, from ordering parts to delivering a finished product. The end result of this step was a simplified data model with all the keys needed to perform a predictive analysis.

Next, the static estimates in the strings were modified into performance distributions using data from the OEM’s transactional enterprise resource planning (ERP) system. To compensate for poor data in this risk adjustment and data enrichment step, the SPANalytics® team developed a routine that used statistical methods to eliminate data outliers and extended sparse data through routines that incorporated data from similar parts or events.

The team then used Monte Carlo simulation to perform simulation and predictive analysis on the data set created from the previous two steps. The simulation was constructed to predict, with statistical confidence, the span time for each string using the distributions and correlations up and down each string. The team was then able to re-sequence the original strings based on predicted span performance and relative performance risk, creating a more accurate picture of the critical paths and drivers of risk in the production system.



In order to optimize a process, an organization must be able to decide where to take action. Advanced visualization techniques were critical to helping organizational leaders understand complex data sets. The solution the team developed, called the Visual BOM, gave leaders and business users the ability to plot and see each delivery string prioritized by importance and level of risk, along with the breakdown of events and their allocated risk. This helped the teams to quickly identify why that particular string was risky and to define how much risk needed to be addressed to improve target span time. The Visual BOM became a key element of the implementation, as it portrayed useful, useable information that encouraged the business teams to break down silos and work together to take action.

The SPANalytics® initiative covered all aspects of designing, developing, testing, and delivering output from an advanced analytics system that addresses span reduction. In just three months, Deloitte stood up a system that is expected to reduce total span time by more than 30 percent, reduce inventory by 20 percent, and decrease total product costs by 10 percent. Before SPANalytics® was implemented, the OEM’s leaders believed that there was a single bottleneck in the manufacturing process. SPANalytics® revealed, however, that there were two dozen potential event sequences with a nonzero probability of resulting in an unacceptable span time. Knowing this was half the battle—it gave organizational leaders a clear statistical view of where to invest the most time and money to address potential issues. The success of SPANalytics has encouraged the OEM to continue exploring the use of advanced analytics in the business.

The SPANalytics® initiative was a co-winner of IU’s inaugural Analytics Leadership Award. The team from Deloitte Consulting LLP included:

- Steve Shepley, principal, Strategy and Operations Practice
- Dan Haynes, principal, Technology Practice
- Kevin Hua, manager, Advanced Analytics and Modeling Practice (Human Capital)
- Mark Neier, senior consultant, Strategy and Operations
- Matt Griffin, senior consultant, Information Management Practice
- Alex Carr, consultant, Information Management Practice

# Analyzing Data for Capacity Planning



## Jesse Delpriora

MS '15, majoring in information systems

Interned at a multinational financial services company

[jessdelp@indiana.edu](mailto:jessdelp@indiana.edu)



The U.S. government requires financial firms to document their due diligence in ensuring that their customers have an acceptable level of risk related to illegal activity. Firms must be able to complete the required forms accurately and quickly in order to remain compliant while continuing to be profitable in taking on new customers.

Intern Jesse Delpriora was charged with manipulating and analyzing data to determine how the firm could speed up the process of filling out forms and to establish whether the appropriate workforce was in place to handle the demands of performing due diligence. Jesse used regression analysis, capacity planning, and data cleansing techniques to achieve these goals.

### What were the steps taken to complete the project?

The most important step was to understand the business purpose for the project. Without a full understanding of the purpose, it's easy to get lost in the data. Having this understanding helped me perform a predictive analysis. When determining which variables had the largest impact on our variable of interest, I found there were many that seemed to be statistically significant. By looking at them with a business-centered focus, it became clear which of these variables actually were important.

The next step was to understand the data we were working with. Aided by my preliminary analysis, I determined which individuals to shadow in order to better understand the data creation. I also interviewed several individuals to understand their point of view and determine whether my team was asking appropriate questions.

Finally, I provided the client with an action plan regarding the derived metrics from my analysis. When reporting metrics and insights to high-level managers, it's important to also provide a recommendation on how to act upon the insights. This greatly reduces the potential for pushback.

### How would you address challenges with quantitatively expressing business imperatives (metrics) in future projects?

A major goal of business analytics is to provide meaningful and actionable insights based on the data. Therefore, it is essential to understand both the audience you are performing analysis for and the audience to whom you are presenting at the end of the project. The definition of "important" or "meaningful" is going to change not only from team to team, but also from person to person throughout the hierarchy. By understanding business goals and gathering opinions from various individuals, I was able to understand what was important to my manager and to the business as a whole.



### **How would you address challenges with data in future projects?**

It's important to bring the business analyst as close as possible to the data creators and to gain an understanding of all the tools necessary to be effective in the field of data analytics. Because I was so far removed from the data source, it affected ease of access as well as my ability to get data in a standardized format. I'm using my knowledge of business intelligence, gained from the MSIS program, to better hone my skills in this area.

### **How did your courses in Kelley prepare you for this project?**

Before the new data analytics co-major, I took advantage of courses such as Econometrics, Data Analysis using Economic Modeling, and Capstone in Economic Consulting. Aside from essential foundational skills, these classes gave me a new perspective on data and helped show me the power behind analytics. My work in the capstone course involved ambiguous data, which simulated a real-world work experience. This impressed on me the importance of asking the right questions about the data at the right time.

### **What did you learn from this experience?**

First, I learned to ask the right questions regarding data in order to avoid future problems, as well as to record in writing the assumptions I made about the data. I also learned to ensure that my insights were related to value-driven business propositions rather than technical anomalies. And I learned to always provide an action plan when reporting metrics or insights.

# IBA Affiliated Faculty



## Frank Acito

Professor of Marketing; Max Barney Faculty Fellow; Co-Director, Institute for Business Analytics



## Kyle Cattani

Associate Professor of Operations Management; W. W. Grainger, Inc. Faculty Fellow



## F. Robert Jacobs

Professor of Operations Management; Chase Faculty Fellow; Honors I-Core Faculty Lead



## Herman Aguinis

Professor of Organizational Behavior & Human Resources; John F. Mee Chair of Management; Director, Institute for Global Organizational Effectiveness



## Sandeep Chandukala

Assistant Professor of Marketing



## Robert Jennings

Professor of Finance; Gregg T. and Judith A. Summerville Professor



## Goker Aydin

Associate Professor of Operations & Decision Technologies; ArcelorMittal USA Undergraduate Faculty Fellow



## Hai Che

Associate Professor of Marketing



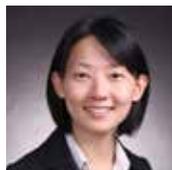
## Vijay Khatri

Associate Professor of Information Systems; Arthur M. Weimer Faculty Fellow; Co-Director, Institute for Business Analytics



## Hillo Bala

Assistant Professor of Information Systems; Whirlpool Corporation Faculty Fellow



## Ruomeng Cui

Assistant Professor of Operations & Decision Technologies



## Kristoph Kleiner

Assistant Professor of Finance



## Azi Ben-Rephael

Assistant Professor of Finance



## Isaac Hacamo

Assistant Professor of Finance



## Alice Li

Assistant Professor of Marketing



## J. Doug Blocher

Chairperson and Associate Professor of Operations & Decision Technologies; Arthur M. Weimer Faculty Fellow



## Randy Heron

Professor of Finance; Roger & Barbara Schmenner Faculty Fellow



## Shibo Li

Associate Professor of Marketing; Arthur M. Weimer Faculty Fellow



## Kurt M. Bretthauer

Professor of Operations & Decision Technologies; W. Michael & William D. Wells Life Sciences Faculty Fellow



## Craig Holden

Professor of Finance; Boquist-Meyer Faculty Fellow



## Alex Lopes

Clinical Associate Professor



## Raymond R. Burke

Professor of Marketing; E. W. Kelley Chair of Business Administration; Director, Customer Interface Laboratory



## Ryan Israelsen

Assistant Professor of Finance



## Girish Mallapragada

Assistant Professor of Marketing



**Alfonso J. Pedraza-Martinez**

Assistant Professor of Operations & Decision Technologies



**Kim Saxton**

Clinical Associate Professor of Marketing



**Munirpallam Venkataramanan**

Professor of Decision Sciences; Jack R. Wentworth Professor; Vice Provost for Strategic Initiatives



**Neil Morgan**

Professor of Marketing; PetSmart Inc. Distinguished Professor in Marketing Chair



**Richard Shockley**

Associate Professor of Finance; CenterPoint Energy Faculty Fellow



**James M. Wahlen**

Professor of Accounting; James R. Hodge Chair of Excellence



**Veronika Krepely Pool**

Associate Professor of Finance; ArcelorMittal USA Undergraduate Faculty Fellow



**Rebecca Slotegraaf**

Associate Professor of Marketing; Whirlpool Faculty Fellow



**Rockney G. Walters**

Professor of Marketing



**Philip T. Powell**

Clinical Associate Professor of Business Economics and Public Policy; Chairperson of Kelley Direct



**Scott B. Smart**

Clinical Associate Professor of Finance; Whirlpool Finance Faculty Fellow; Associate Chair of the Full-Time MBA Program; Director of Strategic Finance Academy



**Matthijs R. Wildenbeest**

Associate Professor of Business Economics



**Jeff Prince**

Associate Professor of Business Economics; Cathie and Jerry Anderson Faculty Fellow



**Daniel C. Smith**

Professor of Marketing; Former Dean of the Kelley School of Business; President and CEO, IU Foundation



**Wayne Winston**

Emeritus Professor of Operations & Decision Technologies



**Anjanette Raymond**

Assistant Professor of Business Law and Ethics



**Ash Soni**

Professor of Operations & Decision Technologies; ArcelorMittal Faculty Fellow; Executive Associate Dean for Academic Programs;



**Owen Wu**

Associate Professor of Operations & Decision Technologies



**Lopo L. Rego**

Associate Professor of Marketing



**Gilvan "Gil" C. Souza**

Associate Professor of Operations Management



**Qiuping Yu**

Assistant Professor of Operations & Decision Technologies



**Daniel Sacks**

Assistant Professor of Business Economics and Public Policy



**Noah Stoffman**

Assistant Professor of Finance



**Babur De los Santos**

Assistant Professor of Business Economics



**Ramesh Venkataraman**

Associate Professor of Information Systems; Lawrence D. Glaubinger Professor of Business Administration; Chair of the Undergraduate Program



# KELLEY SCHOOL OF BUSINESS

INDIANA UNIVERSITY

Indiana University Kelley School of Business  
1275 E. Tenth Street  
Bloomington, IN 47405



Nonprofit Organization  
U.S. Postage  
PAID  
Bloomington, IN  
Permit No. 351

## About Us

The **Kelley Institute for Business Analytics** uses the resources of the prestigious Kelley School of Business to produce insightful research and train professionals who can excel in this exciting new field.

### What Is Business Analytics?

Simply put, it's using data to make better business decisions. And it's becoming big business.

For years, companies have collected data about their practices and consumers. Now, thanks to inexpensive computing, more companies are putting their data to work—using techniques such as predictive analytics, optimization, and simulation to make fact-based decisions that improve productivity, increase profits, and create a competitive advantage.

### Kelley: Leading the Way

To make the most of business analytics, companies need innovative ideas and well-trained professionals. That's where Kelley comes in.

One of just a few business analytics programs nationwide, Kelley's IBA supports:

- An academic program that prepares students to solve business problems using analytics
- Corporate partnerships that shape Kelley's understanding of analytics and help companies tap into Kelley's talent
- Cross-disciplinary research by Kelley's expert faculty
- Seminars, conferences, a speaker series, and case competitions that bring together professionals, faculty, and students

### Contact Us

<http://kelley.iu.edu/iba>

[kiba@indiana.edu](mailto:kiba@indiana.edu)

LinkedIn: <http://tinyurl.com/linkedin-kiba>

YouTube: <http://tinyurl.com/youtube-kiba>

### Founding Partner and Advisory Board Member

Deloitte Consulting LLP

### Supporting Partners and Advisory Board Members

3M  
Protiviti  
Indiana University  
SAS  
Informatica  
Allegient LLC  
Procter & Gamble

### Advisory Board Members

Ernst & Young  
Accenture  
Angie's List  
Discover  
Citibank  
Qlik  
Tableau  
Whirlpool  
SAS-JMP  
Roche Diagnostics Corporation  
Brillio  
Infosys  
Cummins  
C.H. Robinson Worldwide, Inc.  
McKinsey and Company  
FedEx Services  
Booz Allen Hamilton  
IBM Corporation  
Eli Lilly and Company  
dunnhumby  
Hill-Rom  
Dow AgroSciences  
ArcelorMittal