

# Coca-Cola's Unique Challenge: Turning 250 Datasets Into One

To make sense of a mountain of complex data, the world's largest beverage company takes a forward-looking approach.

Sam Ransbotham

# Coca-Cola's Unique Challenge: Turning 250 Datasets Into One

REMCO BROUWER AND MATHEW CHACKO (COCA-COLA), INTERVIEWED BY SAM RANSBOTHAM

To make sense of a mountain of complex data, the world's largest beverage company takes a forward-looking approach.



At The Coca-Cola Company, pulling together useful data sets is a particular challenge. Coke's distribution model involves a network of hundreds of independently operated bottlers around the world that use the Coke concentrates to make and bottle Coke drinks (as well as other non-Coke affiliated beverages). Those bottlers send data to Coke, and Coke's job is to put those data streams

into a common system and use it to look back on how things have gone and project how things might be.

Complicated — to say the least!

Just how complicated can be underscored by a few statistics: Coke is the world's largest beverage company, with more than 500 brands and 3,500 products sold worldwide. In 2013, the company had \$46.9 billion in net operating revenues, and a net income of \$8.6 billion. It has about 250 bottling partners with 900 bottling plants, and employs over 700,000 system associates worldwide. In addition to its flagship Coca-Cola products, the company's brands include Minute Maid juice, Fanta and Sprite soft drinks, and Dasani water.

Mathew Chacko, Coca-Cola's director of enterprise architecture, and Remco Brouwer, the company's director of business intelligence, spoke with Sam Ransbotham, an associate professor of information systems at the Carroll School of Management at Boston College and the *MIT Sloan Management Review* guest editor for the Data and Analytics Big Idea Initiative about the challenges of integrating so many data sets, the ways the company is moving toward a predictive analytics model, and the value of visualizations to convey complex information.

## **When did Coca-Cola first start thinking about analytical approaches, and how did that first get started?**

**Mathew Chacko:** Analytics at Coke actually has a very long history. An important area of analytics is in our volume and sales reporting and forecasting that spans both the company and our bottler franchisee system. In our sparkling drinks business, the Coca-Cola Company makes the concentrate, and then we sell it to bottlers who are not necessarily owned by us. Sometimes we may have a minority or even majority stake, but it's a franchise model. The bottlers are responsible for making the end consumer product.

The bottlers then sell it to customers — and by customers, I mean entities like Walmart or Tesco or McDonald's. And the customers then sell it to consumers. So we have this long pipeline, with the Coca-Cola Company quite removed from the data and the end consumer.

Another area of analytics is marketing performance and spend analysis. Traditionally, the Coca-Cola Company has been responsible for interaction with the consumer mostly from a brand perspective — brand marketing via sponsorship of events like the Olympics or the [FIFA] World Cup, using traditional media, radio ads, billboard ads.

Marketers want to answer questions like, how do we spend our marketing dollars efficiently? How can we project that we will get the lift that we need from the marketing campaign?

What has changed over the last three years or so is the advance of much newer technologies to do more *forward-looking analytics* versus *backward-looking analytics*. We

are used to a lot of dashboarding and reports that say this is what's happened, this is how we look at our sales, our volume, this is how we look at marketing. We ask our bottlers to give us volume and sales information so that we can understand how the brand and the various products are doing globally.

## **Let's talk about that forward-looking analytics versus backward-looking analytics. Where is predictive forecasting being used in the company?**

**Chacko:** One can argue about whether it's analytics or just basic reporting, but there is a forecast element to all that reporting. For instance, we use historical data to project a rolling 13-month forecast window.

## **Can you give us an example of how this forecasting is being used in the supply chain?**

**Chacko:** Let's take orange juice. Coca-Cola makes both Minute Maid and Simply Orange juice brands, which are manufactured throughout the year. Also, consumer preferences demand a somewhat consistent flavor for the juice throughout the year.

We use varieties from various sources, but the sweetness of the oranges differs depending on whether it's a Valencia orange from Florida versus an orange from Spain or from Brazil. In addition, the variances in climate, agricultural fruit production, trucking, shipping, etc., create an agricultural supply-chain problem.

We use an internal application that factors in all influencing variables and optimizes the supply chain to produce a consistent juice product for our consumers. Like other companies, we used technology that we had available at the time to work on the particular problem,

and we are continuously evaluating new and emerging technologies as appropriate.

**You're taking this approach of applying things like optimization algorithms toward mixing your varieties, but I assume that before those technologies existed, some kind of forecasting was done by people. So how are you going about trying to nudge people along the technology path?**

Remco Brouwer: Well, typically it's done by taking them along a path one step at a time. Start slow and let people embrace visualizations, for example, instead of reports. Do it at their own pace [for] a little bit.

**And be hopeful that their own pace is positive progress and not moving away from it?**

Brouwer: Well, yeah. The thing is, most employees would agree that the amount of information that they're being bombarded with is only going to grow. If you're one of those people that finds a printed report on your desk in the morning because somebody has put it there, and [reading it is] the first thing you do, then, yeah, it's a bit of a struggle to move to something else.

I think visualizations will not only increase the business understanding for end users, but will also allow them to be more effective by finding out quicker where the opportunities and the threats are. Visualizations will help people trust the system for taking care of the business that is within the acceptable boundaries.

Chacko: But the shift from looking backwards versus prescriptive and predictive analytics is also a shift in people's thinking.

**So what's the limitation in rolling out prescriptive and predictive analytics? Technology limitations? Data quality? Is it people to do the analysis?**

Chacko: It's all of the above, really. We need to get the technology capabilities in first so that we can ramp up on the people skills and get the data that's in silos more connected together.

Coke is not a top-down, hierarchical organization that dictates from above. We encourage our markets to innovate and build capabilities for their specific market. So, that way, they can move fast. But on the flip side, it does introduce some duplication in our system and technology. And this is a particular area where we don't want to duplicate these kinds of large data platforms.

As a result, we're looking to build a more shared platform for different regions. But with emerging technologies, we want to take a more cautious approach of building and improving it incrementally.

**What about people? Are you finding enough data scientists to help guide the company?**

Chacko: We do have a knowledge and insights team within the company, and like a lot of companies, we have a mixture of internal and external resources to help with the analytics. And as I said, since we're moving in this new direction, we want to also start looking at this whole category of data scientists. One of the plans as we go about this journey is to look at data science as a capability within the organization.

We need people who are just interested in data discovery. And by that, I mean just trolling through data to find out insights and be willing to work with messy data, willing to work with different sets of data. Right now, because we

are taking this very traditional BI [business intelligence] approach, we have a whole process of figuring out the best ways of getting this data in, creating the model, getting the visualizations, and then getting those to the market analysts. It's a long climb.

We're looking to partner with either research institutions or universities as well, for a couple of reasons. First of all, they have cutting-edge research in this area. We also want to work with them in terms of working with their students, giving them opportunities, internships, etc. Hopefully we'll get a pipeline of students who would be interested in working on analytics problems at Coke.

**Can you talk a little bit about the challenges of marrying all these external data sets from your bottlers? That seems like a particular challenge at Coke, where there's this huge infrastructure of "data middlemen," for lack of a better term.**

Chacko: Well, point-of-sale data, scan data, is actually very big. Our commercial department really wants to have good information about that from one set of customers so that they can take that information and go to other customers and say, "Hey, we have this other customer that did these kinds of campaigns with this kind of strategy, and they were able to see this lift."

Our team wants to go to our customers with real numbers. It's important for us when we go to our customers to be able to give them fact-based information. We also want to take in things like event data and social media data and provide these value services back to our customers. That's a big thing for us.

In some areas we have to have flexibility, and in other areas we can standardize. Currently our bottlers run their

own system, and so they send us data in all sorts of different formats. We have to be flexible in being able to inject data. But when we transform that data, we need to transform into those standardized taxonomies or hierarchies.

We also have the reverse problem because we need to transform information *back* into the bottler's view — we have to give them back their information in formats they can read. We aspire to provide data as a service, both to our customers and to our bottlers.

**Any time I learn more about anything that anybody's doing, it always seems much harder than I had originally appreciated. I can see where your franchises create an opportunity but also some difficulty in their autonomy.**

Brouwer: Yes, it's a very complex thing. There are over 250 bottlers around the world, and we are in the center of this nucleus. These 250 bottlers are all sending us data. We're trying to solve some of the day-to-day things like moving more and more into the direction of a one-number system. At the same time, it's the idea of the shared knowledge because in the end, overall, we want the same thing, and that is to understand the consumers better and be as relevant as possible in the market so that people will buy our products and be delighted by them.

In the end it's about — it's a bit of an old term, but it *is* about making better decisions faster. Better decisions by having your data right and your right visualizations, and faster by not spending too much time to get there. That means starting at the bottom and trying to drive clarity in the numbers themselves. As we move more into the direction of big data and cross-measure analysis of everything that big data technologies enable, we're at the

beginning of the path here, there are many answers out there to questions we did not know we had.

We need to be ready for whatever is the next big hype in 2017. We don't know what it is today, but we do know

that we will need to capture its data and analyze it against our own data at some point in time. So we're trying to solve today's challenges and at the same time get our solutions ready for whatever tomorrow will bring.

**About the Author**

Boston College. He can be reached at  
sam.ransbotham@bc.edu.

Sam Ransbotham is an associate  
professor of Information Systems at



**PDFs ■ Reprints ■ Permission to Copy ■ Back Issues**

Articles published in MIT Sloan Management Review are copyrighted by the Massachusetts Institute of Technology unless otherwise specified at the end of an article.

MIT Sloan Management Review articles, permissions, and back issues can be purchased on our Web site: [sloanreview.mit.edu](http://sloanreview.mit.edu) or you may order through our Business Service Center (9 a.m.-5 p.m. ET) at the phone numbers listed below. Paper reprints are available in quantities of 250 or more.

**To reproduce or transmit one or more MIT Sloan Management Review articles by electronic or mechanical means** (including photocopying or archiving in any information storage or retrieval system) **requires written permission.**

To request permission, use our Web site: [sloanreview.mit.edu](http://sloanreview.mit.edu)  
or

E-mail: [smr-help@mit.edu](mailto:smr-help@mit.edu)

Call (US and International): 617-253-7170 Fax: 617-258-9739

**Posting of full-text SMR articles on publicly accessible Internet sites is prohibited.** To obtain permission to post articles on secure and/or password-protected intranet sites, e-mail your request to [smr-help@mit.edu](mailto:smr-help@mit.edu).



Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.