General Mills Builds Up Big Data to Answer Big Questions

Enhancing intuition with analytics at General Mills.

Michael Fitzgerald
General Mills Builds Up Big Data to Answer Big Questions
WAYDE FLEENER (GENERAL MILLS), INTERVIEWED BY MICHAEL FITZGERALD

Enhancing intuition with analytics at General Mills.

Fleener, 36, still remembers the day when a colleague came to him and said a name, “big data,” had been coined for what they were doing. “We called it decision science,” he says. After working on programs for the consumer packaged goods industry, he found himself intrigued by the challenges they face, as companies indirectly connected to their consumers. He came to General Mills in mid-2013, becoming a senior manager/data scientist in its 200+-person Consumer Insights group, a unit of marketing comprised largely of researchers. He spoke with MIT SMR contributing editor Michael Fitzgerald.

What did General Mills want in a data scientist for its marketing group?

The leader for consumer insights, Jeanine Bassett [vice president of Global Consumer Insights at General Mills] was concerned about how much of our decision making was based on outside research. She wanted more decision making based on our intuition and utilizing the data we had in-house. She wanted us to become less reliant on research. During the whole interview process, they said, “We need someone to really help us to answer these questions, about what we think we know.” So I came in to get after the data that we had internally and how we could benefit from it. And they gave me a lot of leeway.

Wayde Fleener likes solving real-world problems, which led him into what we now call Big Data analytics. He started out on a different path, studying chemical engineering at the University of Minnesota. But he realized he didn’t want to go on for a PhD because it would mean a career in research, and he wanted to do applied work. After a stint in investment banking in Asia, Fleener went to a loyalty marketing agency.
Did they also give you a sense of urgency, of needing to get things done quickly?
Yes. They wanted everything as quickly as possible but not at the sacrifice of accuracy.

How did you get started?
Coming in as an outsider, my first couple of months were focused on figuring out what we really do at General Mills. The very first thing I had to do was sit down and do a catalog of all the different data we commonly use and what we use it for. I started asking around, “Does anybody have a data diagram that tells me the data you have and where it’s being used?” No one had it. So I started doing diagrams to say where the data was connected, where are they not, and asking, where should they be? Where are the various data being used in analyses? Where should they be?

How did you find the data?
I had to do a bunch of interviews. I found the group that does our marketing mix, and sat down with them for quite a while. I went to our relationship marketing group, and asked them what data they were using. I’d sit down with all these specialized functions inside Consumer Insights and ask a lot of questions.

Then I’d get really tactical, get access to the data sources themselves, and literally go in and look at tables — there are no data dictionaries. I had to figure out what [did] the columns inside the tables mean. A lot of times, I had to work my way through the IT organization, because the only person who had the answer was the person who developed that database. Sometimes there were names for columns and I couldn’t interpret what they meant.

After that, I had to figure out where the data was coming from upstream. And I realized there was data coming to us that we weren’t capturing that we should be capturing. We would just delete it.

What kind of data was not being captured?
For example, we may bring it into our system at an aggregated market level, but the data is actually coming to us at a disaggregated level [by individual locales].

What was the most difficult thing at the beginning?
General Mills is a huge ship, and it’s very difficult to move. People felt threatened because I was coming into their role and questioning them. So there was a lot of resistance, and “Are you trying to take over what I’m doing?” I had to keep saying, “No, I’m just trying to help you do things better.” I sit in the business, but my role goes into IT a lot. So the very first couple of months, there was tension between me and IT, because they were afraid they were going to be told they had been doing things wrong. I resolved it through weekly meetings, showing them what we’re doing; not calling people.

Did you do this all yourself, or did upper management need to get involved?
There were a couple of instances where we hit roadblocks and there had to be a senior management talks. It took six months before we really turned a corner. One of the director-level leaders in IT started one of the meetings saying, “Hey, listen,” — and this was directed to some of their own folks — “Consumer Insights has some really interesting things that I believe we’re going to benefit from as a company.” And then the leader left the room. When that meeting happened, I knew the tide had changed.
Were you working with anyone else?
I started off as a one-man team. As people started approaching me, my VP saw the traction and invested a little more, and we were able to show value, and it just kept on growing. Once we had a sizeable amount of our data connected, there was a sense of urgency to show the value of that. That is why we jumped into visuals so the organization could tangibly touch something that we did.

Data visualization really accelerated things. We were sorting out a lot of data, and behind the scenes the data was coming together, but a lot of the Consumer Insights folks couldn’t tangibly touch it yet. We sat down with some of our division folks and identified the things they wished they could do but couldn’t because they couldn’t access the data. I found a data-visualization company that could take the data sets we were bringing together and produce some valuable visuals. Now the divisions can actually tangibly understand what they’re doing from a data perspective.

That just grew to more requests and more people who wanted visualizations. We’re now the Big Data team within Consumer Insights. Our mix after a little more than a year is myself, four data stewards, their manager, four data visualizers, and two data scientists. We need two more data scientists, something I’m actively working on now.

Was data visualization a new role for General Mills?
When I got here, General Mills already had in place a data visualization tool. But it was not used at all in Consumer Insights. My researchers, who are not programmers, live in Excel, but my datasets are hundreds of millions of rows. [Editor’s note: At the time, Microsoft did not have a version of Excel that could handle such large numbers of rows. General Mills has since become a beta tester of a version that does.] With our visualization tool, we can do a lot of stuff out of the box without having to program, but if you’re trying to do specialized stuff then you need to bring in programming languages. I started looking for some developers that knew our tool and could do the more advanced stuff. We outsourced that data visualization to another company.

Data visualization technology is now a commodity. But, the idea of what you need to visualize is not. That’s what we data scientists do. We come up with “it’d be really cool if we could do this” ideas. When we’re ready to productize it, the programming is the commodity piece.

What are the pros and cons of looking outside for talent versus staying inside?
Outside, you get the advantage of individuals that have developed across industries and other companies. You may create a blueprint and hand it to them, and they come back and say we developed something similar. Here are some ideas that might enhance your visualization. Internally, it always comes down to cost: can you bring it in-house and do it cheaper? Some of our outsourcing is through India, and it can be almost like having 24-hour development.

You’ve talked about how to make analytics part of everyday operations, not having analytics be a series of one-off projects. How hard is that process to develop?
It’s very difficult. People are living in an Excel world. They think, “I gotta grab this data and this data and every time I want to refresh.” Versus, we’re going all the way back to
the data source in everything that we're doing, so we don't have to touch that whole data update.

**If analytics traditionally works in a rinse-and-repeat mode, where you ask a question, gather data, run the data, ask another question, gather new data, run it, and so on, how do you break away from that?**

Every division thinks it has this unique way it should view the data. Yes, they all face their own unique situation, but in reality, they really look at the data the exact same way. If I have to create a unique way for you to view data across seven divisions, I have to maintain seven unique versions of the data. I can't get scale that way.

**Does that help you avoid what you’ve called the “Go Fish” problem of Big Data, where you ask people if they have the data, and they say no?**

When I first got here, no one had a view of all the data we had across everything. Everybody had a specific role and they cared about the data they used in that role. If someone had a new use-case they developed, you would have to go around and say, do we even have that data? How do I get the data, how do I manipulate the data, how do I combine it with my other data sets? When I mash up my data sets, do I know I have yellow-box Cheerios named consistently across the two data sets?

**How did you fix that problem?**

That’s where the data stewards, a new role within Consumer Insights, come into play. We’ve been systematically going through every single core data asset. We’ve got hundreds of products. One data source may spell “Cheerios” all the way out, and another may spell it “chrs.” We’ll prepare a summarized data set and give that to our data stewards. We have about six data stewards that map connections between data sets. For every data refresh cycle, we created what we call an exception report, where we identify what’s new. The data stewards go through and map what’s new.

**Were the data stewards there before you got to General Mills, perhaps doing something else?**

It’s grown three-fold. We had an individual solely focused on Nielsen scanner data, a very important data set in any consumer packaged goods company, making sure that whatever we get from Nielsen will work the way marketing needs it. The job had a different title. But now they’re data stewards, and we’re taking those roles and instead of making them solely an expert in one data source, we’re making them an expert across data sources. They reported to different lines when I got here, but as of January they report to me. Data stewards don’t require all the technical and programming skills of a data scientist. For some of their tasks, a two-year technical degree can be enough.

**When and how did you add your next data scientist?**

After six months, the business was asking for more analytics, especially with a predictive component. The amount of stuff that they were asking for was exceeding my capacity. We’ve got a good recruiting office, so I built out the profile I was looking for, and they started looking at various academic programs locally. I think they found him on LinkedIn, because his profile said he was studying Big Data at The University of St. Thomas. I had told HR that I was fine with an individual that was mid-school, because that was probably what I could afford. He had a day job already, and I had to do a big selling job to get
him. They’re tough to acquire. The big sell was, “We’re starting something here at General Mills, you could be part of it.” That’s appealing to certain individuals.

**Your unit also has an analytics group with statisticians. How are they different?**

They don’t contain the data scientists. The data science group is more technically savvy at getting and manipulating the data, and can engineer the systems to operationalize the analytics. The statisticians/analytics group focuses on answering particular questions, but for example, they do not have the skillset to take a model and deploy it. Both groups report to Vidyotham Reddi, who is director of Global Enablers. He reports to Jeanine Bassett.

**How you can tell the company is successfully operationalizing analytics?**

For us, it’s how many people are using our stuff and asking for more stuff. Ultimately, a company is successful if it can say decisions made by analytics led to incremental profit.

**What lessons do you have for other companies?**

Start with an understanding of the data you have, how it’s organized. Have the end in mind — what you want to accomplish — and then develop that plan through the process. Build the right relationships. I identified who would be my champions and built relationships with them.

You’ve got to make sure you have the right IT folks along for the ride. I’m technical, but I hit a point where I don’t understand Linux servers, and so I need an expert. Having a breadth of relationships across the company is the only reason why we have been successful.
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 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
or
E-mail: 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.