

Does Big Data Bring Big rewards?: CASE STUDY

Today's companies are dealing with an avalanche of data from social media, search, and sensors as well as from traditional sources. According to one estimate, 2.5 quintillion bytes of data per day are generated around the world. Making sense of "big data" has become one of the primary challenges for corporations of all shapes and sizes, but it also represents new opportunities. How are companies currently taking advantage of "big data?"

Green Mountain Coffee in Waterbury, Vermont, is analyzing both structured and unstructured audio and text data to learn more about customer behavior and buying patterns. The firm has 20 different brands and more than 200 different beverages, and uses Calabrio Speech Analytics to glean insights from multiple interaction channels and data streams. In the past, Green Mountain was unable to fully utilize all the data gathered when customers called into its contact center. The company wanted to know more about how many people were asking for a specific product, which products generated the most questions, and which products and categories created the most confusion. By analyzing its big data, Green Mountain was able to gather much more precise information and use it to produce materials, Web pages, and database entries to help representatives do their jobs more effectively. Management is now able to identify issues more rapidly before they create problems for customers.

AutoZone uses big data to help it adjust inventory and product prices at some of its 5,000 stores. For example, a customer walking into an AutoZone store in Waco, Texas, might find a deal on Gabriel shocks which that person would not find in most other AutoZone stores. The Mulberry, Florida AutoZone store might feature a special on a bug deflector. To target these deals at the local level, the auto parts retailer analyzes information gleaned from a variety of databases, such as the types of cars driven by people living around its retail outlets. Software from NuoDB, which uses a cloud services model, makes it possible to quickly increase the amount of data analyzed without bringing down the system, or changing a line of code.

Benefits from analyzing big data are not limited to businesses. A number of services have emerged to analyze big data to help consumers. For example, personal devices such as the NikeFuelBand, SonySmartBand, and Jawbone UP24 enable people analyze their routines, diets, and sleeping patterns to see how they compare with others. This can lead to more-effective workouts and help people meet fitness goals. A number of online services enable

consumers to check thousands of different flight and hotel options and book their own reservations, tasks previously handled by travel agents. New mobile-based services make it even easier to compare prices and pick the best travel options. For instance, a mobile app from Skyscanner Ltd. shows deals from all over the Web in one list—sorted by price, duration or airline—so travelers don't have to scour multiple sites to book within their budget. Skyscanner uses information from more than 300 airlines, travel agents and timetables and shapes the data into at-a-glance formats, with algorithms to keep pricing current and make predictions about who will have the best deal for a given market.

There are limits to using big data. A number of companies have rushed to start big data projects without first establishing a business goal for this new information. Swimming in numbers doesn't necessarily mean that the right information is being collected or that people will make smarter decisions.

Although big data is very good at detecting correlations, especially subtle correlations that an analysis of smaller data sets might miss, big data analysis doesn't necessarily indicate which correlations are meaningful. For example, examining big data might show that from 2006 to 2011 the United States murder rate was highly correlated with the market share of Internet Explorer, since both declined sharply. But that doesn't necessarily mean there is any meaningful connection between the two phenomena.

Several years ago, Google developed what it thought was a leading-edge algorithm using data it collected from Web searches to determine exactly how many people had influenza. It tried to calculate the number of people with flu in the United States by relating people's location to flu-related search queries on Google. The service has consistently overestimated flu rates, when compared to conventional data collected afterward by the US Centers for Disease Control (CDC). According to Google Flu Trends, nearly 11 percent of the U.S. population was

PRINTED BY: apus_edu@yahoo.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

249

250

supposed to have had influenza at the flu season's peak in mid-January 2013. However, an article in the science journal *Nature* stated that Google's results were twice the actual amount estimated by the U.S. Centers for Disease Control and Prevention, which had 6 percent of the population coming down with the disease. Why did this happen? Several scientists suggested that Google was "tricked" by widespread media coverage of the year's severe flu season in the U.S, which was further amplified by social media coverage. Google's algorithm only looked at numbers, not the context of the search results.

Sears Holdings, the parent company of Sears and Kmart, has been trying to use big data to get closer to its customers. Sears used to be the largest retailer in the United States, but for many years has steadily lost ground to discounters such as Walmart and Target and to competitively-priced specialty retailers such as Home Depot and Lowe's. The company has been slow to reduce operating costs, keep pace with current merchandising trends, and remodel its 2,429 stores, many of which are run-down and in undesirable locations.

Over the years, Sears had invested heavily in information technology. At one time it spent more on information technology and networking than all other non-computer firms in the United States except the Boeing Corporation. Sears used its huge customer databases of 60 million past and present Sears credit card holders to target groups such as tool buyers, appliance buyers, and gardening enthusiasts with special promotions. These efforts did not translate into competitive advantage because Sears' cost structure remained one of the highest in its industry.

The Sears company has continued to embrace new technology to revive flagging sales: online shopping, mobile apps, and an Amazon.com-like marketplace with other vendors for 18 million products, along with heavy in-store promotions. So far, these efforts have not paid off, and sales have declined since the 2005 merger with Kmart. The company posted a loss of nearly \$1.4 billion for 2013.

Sears Holdings CEO Lou D'Ambrosio believes even more intensive use of technology and mining of customer data is the answer. The expectation is that deeper knowledge of customer preferences and buying patterns will make promotions, merchandising, and selling much more effective. Customers should flock to Sears stores because they will be carrying exactly what they want.

A customer loyalty program called Shop Your Way Rewards promises customers generous free deals for repeat purchases if they agree to share their personal shopping data with the company. Sears would not disclose how many customers have signed up for Shop Your Way

Rewards, but loyalty-marketing firm Colloquy estimates around 50 million people are members.

Sears wanted to personalize marketing campaigns, coupons, and offers down to the individual customer, but its legacy systems were incapable of supporting that level of activity. In order to use big models on large data sets, Sears turned to Apache Hadoop and big data technology. It used to take Sears six weeks to analyze marketing campaigns for loyalty club members using a mainframe, Teradata data warehouse software, and SAS servers. Using Hadoop, the processing can be completed weekly. Certain online and mobile commerce analyses can be performed daily and targeting is much more precise, in some cases down to the individual customer. Sears's old models were able to use 10 percent of available data, but the new models are able to work with 100 percent. In the past, Sears was only able to retain data from 90 days to two years, but with Hadoop, it can keep everything, increasing its chances of finding more meaningful patterns in the data.

What's more, Hadoop processing is much less costly than conventional relational databases. A Hadoop system handling 200 terabytes of data runs about one-third the cost of a 200-terabyte relational platform. With Hadoop's massively parallel processing power, processing 2 billion records takes Sears little more than one minute longer than processing 100 million records.

Hadoop is still an immature platform, and Hadoop expertise is scarce. Sears had to learn Hadoop largely by trial and error. But it now runs critical reports on the platform, including analyses of customers, financial data, products, and supply chains. Capitalizing on its experience as a big-data innovator, Sears set up a subsidiary called MetaScale to sell big data cloud and consulting services to other companies.

Sears can point to many conceptual uses of Hadoop, but the question still lingers about whether the company is effectively using Hadoop to solve its enormous business problems. Is it truly able to offer customers personalized promotions and are they working? What is the business impact? Where are the numbers to show that big data is helping Sears become more profitable? Sears may be able to generate revenue by selling big data expertise to MetaScale customers, but will Hadoop really help turn Sears around?

PRINTED BY: apus_edu@yahoo.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

Jim Sullivan, a partner at loyalty marketing firm Colloquy, notes that a good loyalty program that gives a company better intelligence about what its customers really want can be a strategic advantage, but even the best loyalty programs can't fix a fundamentally broken brand.

Sources: Laura Kolodny, "How Consumers Can Use Big Data," *Wall Street Journal*, March 23, 2014; Joseph Stromberg, "Why Google Flu Trends Can't Track the Flu (Yet)," smithsonianmag.com, March 13, 2014; Gary Marcus and Ernest Davis, "Eight (No, Nine!) Problems With Big Data," *New York Times*, April 6, 2014; Thomas H. Davenport, *Big Data at Work*, Harvard Business School Publishing, 2014; Samuel Greengard, "Companies Grapple With Big Data Challenges," *Baseline*, October 29, 2013; Rachael King and Steven Rosenbush, "Big Data Broadens Its Range," *The Wall Street Journal* (March 13, 2013); Nick Bilton, "Disruptions: Data Without a Context Tells a Misleading Story," *The New York Times*, February 24, 2013; ShiraOvide, "Big Data, Big Blunders," *The Wall Street Journal*, March 11, 2013; Mark A. Smith, "Big Data Pointless without Integration," *Information Management*, February 25, 2013; Frank Konkel, "Fast Failure Could Lead to Big-Data Success," *Federal Computer Week*, January 30, 2013; and Doug Henschen, "Why Sears Is Going All-in on Hadoop," *Information Week*, October 3, 2012.

CASE STUDY QUESTIONS

- 6-13 Describe the kinds of "big data" collected by the organizations described in this case.
- 6-14 List and describe the business intelligence technologies described in this case.
- 6-15 Why did the companies and services described in this case need to maintain and analyze big data? What business benefits did they obtain? How much were they helped by analyzing big data?
- 6-16 Identify three decisions that were improved by using big data.
- 6-17 Should all organizations try to analyze big data? Why or why not? What management, organization, and technology issues should be addressed before a company decides to work with big data?

MyMISLab

Go to mymislab.com for Auto-graded writing questions as well as the following Assisted-graded writing questions:

- 6-18 Identify the five problems of a traditional file environment and explain how a database management system solves them.

6-19 Discuss how the following facilitate the management of big data: Hadoop, in-memory computing, analytic platforms.

PRINTED BY: apus_edu@yahoo.com. Printing is for personal, private use only. No part of this book may be reproduced or transmitted without publisher's prior permission. Violators will be prosecuted.

251

252