

# **BI at RetailStore.com**

---

**Dr. Hugh J. Watson**  
**Dr. Jeffrey A. Hoffer**  
**Dr. Barbara H. Wixom**

## Table of Contents

The Evolution of Retailing and CRM . . . . .	3
Shopping at RetailStore.com . . . . .	4
The Data Warehouse and BI Software, Applications, and Processes . . . . .	6
Putting It All Together . . . . .	11
Driving Customers to the Web Site – The Importance of Search Engines . . . . .	12
Moving Forward . . . . .	13
Conclusion . . . . .	13

## Introduction

In the late 1990s, RetailStore created an Internet site to complement its brick and mortar retail stores. The vision was that RetailStore.com would serve existing RetailStore customers through the Internet, and would meet the needs of new customers who are not traditional RetailStore shoppers. The value propositions of RetailStore.com are clear: RetailStore shoppers enjoy added convenience and additional services through buying online, and RetailStore can reach customers not currently using its physical locations.

A key to RetailStore’s ability to engage in online retailing is an enterprise data warehouse that stores sales, marketing, merchandising, click stream, customer service, logistics, and finance data; and business intelligence (BI) applications that determine customer profitability, identify customer segments, forecast product demand, support pricing decisions, and more. Throughout the company the core culture is to make decisions based on the numbers. RetailStore is an example of a business intelligence company because the company’s success depends on BI.

This case explores how RetailStore.com uses BI.<sup>1</sup> It begins with a historical perspective on how retailing and the relationships between customers and retailers have changed over the years. We will see that BI can be used to reestablish the kinds of personal relationships that once existed. We then introduce Suzie Young and her experiences when buying leather gloves at RetailStore.com. Her experiences are typical and well known to online shoppers. Less understood, and what is explored next, is how BI determines what Suzie experiences on the web site and the follow-up communications that she receives. To complete the picture, we describe the data warehouse and BI software, applications, and processes that were used with Suzie. We also describe the considerable impact that search engines, such as Google, have on online retailing. RetailStore.com continues to evolve its use of BI and data warehousing, and we will describe future directions. Though online retailing has a short history, it has already undergone many changes, and we conclude by discussing its possible future.

1 RetailStore.com is a fictitious company based on the actual practices of large, online retailers. The authors would like to thank Teradata for financially supporting this case study. However, the authors are solely responsible for its contents.

### The Evolution of Retailing and CRM

Over the years, people's shopping experiences have changed. Years ago in small town America, people shopped with a few merchants who knew them well. For example, the owner of the local women's clothing store knew her customers, what they liked, and how profitable they were. The owner would even set aside items that particular customers might like. Customer relationships were not a problem because customers were known and could be treated in a personal way.

Today, things are different. Fewer people live in small towns. Shopping is more likely to take place in malls, big box stores (e.g., Walmart), and on the Internet. Customers and merchants have a very different relationship. Merchants have thousands or even millions of customers whom they don't know personally.

Just as shopping has changed, so too has marketing. Years ago *mass marketing* prevailed. The same message was communicated to everyone. We still see this approach today with newspaper ads. A more efficient approach is *targeted marketing* where a personalized marketing message is sent to only those people who are likely to respond, such as sending a customized catalog to someone who has purchased the company's products in the past. There are times in people's lives when they are more likely to respond to a message than others. *Event-based marketing* strives to send appropriate offers to people when significant events occur in their lives. For example, after buying a new house, people are often in the market for a washer and dryer and are likely respond to a timely, relevant, and attractive offer.

Another change is the growth in the number and diversity of marketing channels. Companies now promote and sell products in retail outlets, through catalog sales, on the Internet, and through contact (or call) centers. We are seeing growth in mobile devices (such as cell phones and PDAs) as a marketing channel. The challenge is to coordinate the promotion, sales, and service activities associated with the various channels. It is important to present one face to customers, rather than operating as seemingly independent companies.

The most contemporary approach to marketing is *customer relationship management* (CRM) where companies strive to develop an in-depth understanding and relationship with their customers. It's a throwback to old, small town customer-merchant relationships and builds upon and extends other marketing approaches. Taken to the max, it's 1:1 CRM where every customer is treated as a market segment of one. Every customer is sent the right offer, at the right time, at the right price, through the right channel. For example, think about an online travel site like Travelocity.com. When booking a flight, it suggests add-on sales for hotel accommodations and a rental car. Later on, it sends you information and offers for things to do in the city that you are visiting.

Though there are many companies that are successful with their CRM efforts, it is a challenging undertaking. It requires a senior management and marketing team that understands the approach and what is necessary to execute it. There must be significant investments in information technology; the collection and storage of customer data; the use of technology to personalize and automate interactions with customers; a strong team of BI and marketing specialists; and the use of analytics to understand, predict, and influence customer behavior. There must be changes in business processes and how employees perform their jobs.

### Shopping at RetailStore.com

To understand how RetailStore.com uses BI to enhance customer satisfaction, decrease costs, and increase revenues and profits, let us consider a typical customer shopping experience on its web site. Our shopper is Suzie Young, and she wants to buy a pair of leather gloves. She has not used RetailStore.com before, but her friends have told her that it is a great place to buy merchandise at good prices. Suzie goes to [www.RetailStore.com](http://www.RetailStore.com) and checks out the home page. Suzie could find leather gloves using the menu system but she chooses to enter "women's leather gloves" into the search box.

In response to her search, a variety of women's leather gloves are displayed. Suzie sees a pair of Portolano gloves that capture her attention. The numerical ratings (an average of 4+ stars) for the Portolano gloves are very good, and the written reviews are positive, but she learns from the reviews that the gloves tend to run small. She selects an appropriate size and adds the gloves to her shopping cart.

Suzie is now ready to check out. At the checkout screen, she sees recommendations for other items that are often in the same market basket<sup>2</sup> (e.g., scarves and boots), but she decides to buy only the gloves today. As she continues with the checkout process, she is asked to create a new account (giving an e-mail address and password) because she is a new customer. Continuing with the checkout process, Suzie provides her shipping address, credit card information, and billing address.

After checking out, Suzie decides to check her Hotmail account.<sup>3</sup> While going through her e-mail, she receives a confirmation of her purchase and recommendations for other items to buy. During the next week, Suzie receives other messages from RetailStore.com, including a confirmation that her order was shipped, offers of a discount on

her next purchase and free shipping, and a coupon to be used either online or at a physical RetailStore location. Later in the week, Suzie receives the gloves and is pleased with their quality and her shopping experience at RetailStore.com. She plans to shop there again. And, she plans to stop by her local RetailStore and check it out in the near future.

### The Duck's Feet – What You Don't See

When we see a duck moving across a lake, below the surface the duck's feet are moving rapidly. There's a lot more going on than meets the eye. Let us consider what is taking place below the surface as Suzie shops at RetailStore.com.

Because Suzie has not used RetailStore.com before, there is no information about her. She is a first time online shopper. When Suzie enters the web site, RetailStore.com places a cookie<sup>4</sup> on her PC. Its role is to store an ID that is assigned to her. This ID is also stored in a RetailStore database. If Suzie returns to the web site using the same computer as before, she is known because of the cookie. As she explores the web site, the click stream data (where she has been) is captured, communicated, and stored in the database. After Suzie registers and makes a purchase(s), even more information is available for analysis purposes.

When Suzie enters "women's leather gloves" into the search box, the system returns gloves that might be of interest to her. An algorithm determines what is displayed and their order based on relevancy to the search terms and business factors, such as the margin on the various products, their sales rank (i.e., popularity), their inventory status, and whether RetailStore.com is trying to close them out. Customers seldom browse past the first or second page. Consequently, it is important to place products at the top of the list that meet customer needs and make business sense.

2 Marketers refer to a customer's shopping cart (whether in a store or online) as a market basket.

3 Hotmail is a free e-mail service offered by Microsoft.

4 A cookie is small piece of information that a web site leaves on a visitor's computer when the visitor comes to a site. Cookies are used to remember information about a visitor when the visitor returns to the web site at a later time.

## BI at RetailStore.com

---

Rather than selecting the Portolano gloves, Suzie could have refined her search based on criteria, such as price and brand. If she had done this, recommendation software would have generated a set of recommended gloves using these criteria (also known as filters).

When Suzie clicked on the Portolano gloves, she saw not only the gloves but also other gloves and complementary products, such as scarves and boots. A market basket algorithm is run daily to determine what complementary products to show with each product. It assumes that future customers are likely to want the same products in their market basket as previous customers.

At the checkout screen, complementary products are shown again, once again in the hope of cross selling. A market basket algorithm is used here, too. A difference is that the recommendations are based on all of the items in the cart rather than just a single item. In Suzie's case, the recommendations are the same because she is buying just the gloves.

In our example, Suzie was task oriented in her shopping. She only considered buying gloves. Had she also looked at shoes, however, this information would have been stored in the company database, and if she returned to the home page, the product recommendations shown would have been different from when she first entered the site, and would have reflected her demonstrated interest in gloves and shoes. In other words, the recommendations on the home page are dynamically determined based on her clicks on the web site.

When Suzie checks out, she is taken to a screen where she either logs in or registers. In our example, because she is a new online customer, she has to register and provide an e-mail address (i.e., her ID) and a password. She also is given the option to opt out or limit the number of promotional messages she receives. Suzie's ID and password are

stored in a database along with the assigned ID and click stream data from her cookie. This begins the accumulation of online customer data about Suzie. As Suzie proceeds with the checkout process, even more information is collected – what she purchases, her mailing and billing addresses, and credit card information.

If Suzie were already a customer of RetailStore's brick and mortar stores, then her information from previous in-store purchases would have been combined with her shopping activity at RetailStore.com. When she registered for her new online account, Suzie would have provided her RetailStore frequent shopper number or her phone number, and the company would have used matching software to combine her in-store shopping activity with her activity at RetailStore.com. This allows RetailStore to have a complete understanding of Suzie's cross-channel customer shopping behaviors.

Knowing Suzie's e-mail address, purchase history, and click stream behavior makes it possible to send a variety of informed and personalized communication and marketing messages. First, Suzie receives an order confirmation, and later, a message that her order has been shipped. RetailStore.com's experience is that both of these messages generate additional sales because they put customers in a shopping mindset. Some of the other messages are targeted, personalized, and based on her past shopping behavior while others are not. Examples of the latter would be a blast e-mail (i.e., sent to everyone) announcing a 50 percent off sale or distributing a coupon for a specific product. The content of the personalized messages depends on what has been either considered or purchased in the past. The purchase recommendations include new items and items that Suzie has looked at previously but did not buy. If she ever leaves things in her shopping cart, she receives an "abandoned cart" e-mail message.

RetailStore.com has learned several interesting things about the timing and content of messages. Messages should be sent soon for recently viewed online items, because customers' interest in items decreases quickly over time. Also, the best long-term predictor of what Suzie might buy is what she has bought in the past.

RetailStore knows exactly which customers are profitable, or not, and why. A profitability algorithm considers the margin on purchases (e.g., some customers tend to buy only low margin sale items), the channels used (e.g., in-store purchases are more costly than ones over the Internet), and the return of items (e.g., some people chronically return merchandise). The most profitable customers (i.e., top 20 percent) are given special treatment when calling customer service. Their calls are routed to agents who, for example, are more likely to permit exceptions to return policies.

In summary, RetailStore.com sends out millions of e-mails per week to current and prospective customers, driving them back to the web site. These e-mails generate a sizable percentage of the company's online sales. A significant percentage of the messages are personalized based on shopping history or other activity on the web site. The product recommendations in the e-mails are based on various recommendation algorithms that are run. RetailStore.com monitors and measures the effectiveness of e-mail campaigns in almost real time. If a campaign is not generating the projected response rate, it is altered on the fly.

## The Data Warehouse and BI Software, Applications, and Processes

We discussed how BI influenced Suzie's shopping experience, but for a deeper understanding, it is necessary to consider the technologies, applications, and processes involved.

### The Enterprise Data Warehouse

The BI applications depend on an enterprise data warehouse (EDW) for data. A data warehouse is a central repository of data that is used to support decision-making applications. We will focus on those parts of the architecture that are most important to understanding what Suzie and other shoppers experience and how BI is used to run the company.

Technically, the EDW features a Teradata® Database that stores multiple terabytes of data, and the amount grows continuously. A variety of data sources populates the warehouse: operational systems (e.g., in-store point-of-sale systems), call center data, click stream data, order fulfillment data, customer service data, and enterprise resource planning (ERP) software that performs a variety of financial applications. The data is loaded into the warehouse using Extraction-Transformation-Load (ETL) software (e.g., Informatica PowerMart, Oracle Data Integrator) that extracts data from source systems, transforms the data for storage in the warehouse, and loads (i.e., places) the data in the EDW. Some of the loading processes are real time (i.e., continuous feeds), and data enters the warehouse almost as soon as it is created, while other data is batch loaded and entered daily, weekly, or monthly depending on the source systems and business needs.

The organization of the data in the warehouse (i.e., entities, relationships, and schemas) is based on the Teradata Retail Logical Data Model for retailing. An industry model (such as for retailing or financial services) is generic and applies to any firm in the industry. To illustrate, in retailing, all firms have information about customers (e.g., e-mail addresses,

mailing addresses, telephone numbers), suppliers (e.g., names, addresses), products (e.g., product name, SKU), and inventory (e.g., inventory balance, status of items on order). Many data warehousing vendors and consulting firms have industry data models that can be purchased.

The motivation for using an industry data model at RetailStore was the same as at most firms. If there is already a good model that can be used, don't reinvent the wheel. It is faster and cheaper to buy an industry data model and customize it (i.e., add entities, relationships, and schemas) as needed than to build a data model from scratch. A purchased model may also be a better model since it reflects best practices learned from multiple firms.

### Reporting and Scorecards

RetailStore uses MicroStrategy, a leading BI tool, for reporting and scorecards. The reports and scorecards are integrated into the fabric of how the company is run.

More than 300 operational reports run hourly or daily using data in the warehouse. Many of the reports allow users to apply filters (e.g., view a particular sales period) and drill down to underlying data. More than 400 employees have MicroStrategy login IDs and run their own reports. Other users receive reports via e-mail. As examples of reports, buyers can see how much money they have to spend, which items are selling well, and which are not. Managers can see the margins on every product, taking into consideration not only the cost of the product, but also the cost of bringing the product into the warehouse, the cost of shipping it to customers, the cost of credit card fees, the cost of customer service, and the cost of returns. These different margins are important to RetailStore and are monitored closely. Revenue and margins are recalculated every 30 minutes throughout the day. It is possible to develop a company profit and loss statement hourly, if needed.

RetailStore has scorecards that are used by the firm's executives. On the scorecards, executives select the time period they are interested in, whether it is the previous day, the past seven or 28 days, or quarter or year-to-date. For the time period selected, charts display metrics about gross merchandise sales, marketing expenses (for coupons, bonus card redemption, and click through costs), margins, gross profit, average order size, visits (the number of visitors to the web site and in-store), and conversion (the percent of people who come to the web site and buy). Also included are traffic light coded gauges that show performance about these metrics in terms of planned versus actual or the growth in sales for this year compared to the same period last year. Information is also available about the performance of the various marketing channels – customers who come directly to the web site, through a referring channel such as Google, through direct mail, or to a physical store. The scorecards provide templates through which sales information can be viewed through a variety of different dimensions and measures.<sup>5</sup>

### Specialized Analytics

While reports and dashboards satisfy most of the well-defined, recurring information needs, there also are requirements for information generated through ad-hoc and specialized analyses. Analytic teams or decision support specialists who work in the various business units create most of this information. These specialists know the business, are very familiar with the data in the data warehouse, understand how the data is organized, and are able to work with a variety of BI tools. If a specialized analysis is useful on an ongoing basis, it may later be converted to a report or scorecard that is delivered through MicroStrategy.

<sup>5</sup> *Dimensions* are the categories used when analyzing data. For, example, time period is a commonly used dimension. *Measures* are specific values for the dimensions. Previous day and past week are possible measures for the time period dimension. An example of a dimensional query might be "tell me how many cases of pretzels were sold in Pittsburgh last week." This query contains three dimensions (product, location, and time period) and three specific measures (pretzels, Pittsburg, and last week). Analyzing data using various dimensions and measures is often referred to as dimensional analysis, OLAP (online analytical processing), and "slicing and dicing" the data.

Many of the specialists use Teradata SQL Assistant Web Edition, a web-based tool that allows users to enter SQL queries, display the query results, and store the queries for later use. Using Teradata SQL Assistant Web Edition, specialists can manipulate warehouse data in almost any way. For example, a clothing buyer might want help in deciding what to buy for this summer's season. Using warehouse data, it is possible to look at last summer's clothing sales. But it also is important to have information about what customers were looking to buy but did not, possibly because the merchandise was either unavailable or didn't meet the customers' needs. Insights about this can be gained by looking at customer searches that did not end in a purchase. In addition to Teradata SQL Assistant Web Edition, there are about 50 users in the business units who conduct their own analyses using MicroStrategy without any knowledge of SQL. Specialists also use Teradata Warehouse Miner and SAS to access and perform statistical analyses.

Here's another example of specialized analytics. RetailStore has buyers who purchase merchandise. The buyers specialize in specific product lines, such as women's formal wear, accessories, and housewares, and they are held accountable for the profitability of the items they buy. They are very knowledgeable about suppliers, market trends, and prices. They do not work alone, however. There is an analytics team that works with the buyers to estimate demand at various prices (i.e., to determine price elasticity). Once a product is acquired and is for sale on the web site, its performance is closely monitored. If a product is hitting an acceptable level (all of this is calculated using data from the EDW), there is no need for action. However, if it is not, the analytics team performs analyses (once again using warehouse data) to determine a price to clear it out of inventory.

Knowing the sales velocity of products is also useful in bargaining with suppliers, especially if a product isn't doing well. When this is the case, RetailStore goes back to the supplier to try to renegotiate the deal or to force a price reduction on the next deal.

Product returns are very costly. Consequently, it is important to monitor returns closely (using warehouse data) and take immediate action when a product's return rate gets too high. When the rate goes above a prescribed threshold, the product is pulled off the web site (and off the in-store floor) and is assigned to an investigative team to work with the vendor to understand and fix the problem. When the problem resides with the vendor, the vendor is expected to make good on the costs incurred. This application saves significant dollars by proactively removing products that are causing problems.

The EDW has resulted in millions of dollars in savings with suppliers. Through analyses of warehouse data, it is possible to forecast better how products will sell. Armed with this information, buyers can negotiate with suppliers for larger purchase quantities at a lower price. Without accurate projections, RetailStore's buyers are less likely to commit to larger lots because of fears of being unable to move all the merchandise at a profitable price.

### **Customer Relationship Management**

A well-executed CRM strategy results in a win-win situation. Customers are happy because they are treated exceptionally well. They are offered products that are relevant to their needs and preferences, at an attractive price, at the right time, with great service. In return, companies are rewarded with greater customer loyalty, revenues, and profits.

There are many component parts to an effective CRM strategy. Starting with the CEO, there should be a strong service culture. Customer service representatives should be knowledgeable, friendly, and helpful, whether taking an order, answering questions about a previous order, or handling a customer complaint. Orders should be fulfilled (i.e., sent to the customer) quickly and accurately.

For retailers with a strong online presence like RetailStore, the offers that are made to customers are a critical CRM component. Executing this well requires a comprehensive, integrated set of customer data; the right analytical applications; people with excellent analytical skills; and continuing experimentation to learn what works best. Let us consider how RetailStore does this.

The Teradata data warehouse provides the foundation for RetailStore's CRM capabilities because it is the source for in-depth, up-to-date customer data. Using warehouse data, RetailStore supports a variety of Teradata CRM and other applications.<sup>6</sup> The applications were relatively easy to implement because RetailStore used Teradata's retail industry data model, and the applications are optimized to work with this data model.

### INTEGRATED CUSTOMER VIEW

In many organizations, it is common to have information about customers scattered across a variety of applications. This is problematic because the company does not have an integrated view of everything it does with and knows about its customers. The classic example is the grandmother who withdraws \$100,000 in savings from her bank because her \$50 check was bounced for insufficient funds. The savings and checking systems were siloed, and data wasn't shared across the applications.<sup>7</sup>

RetailStore's enterprise data warehouse provides an integrated view of the customer. It integrates shopping, logistics, customer service, and click stream data from sessions to provide visibility into every interaction with RetailStore's customers. For each customer, it has integrated data on the products looked at before a purchase, the products purchased, the delivery of the products, and any issues that arose during the entire process.

### ACTIONABLE ANALYTICS

In this era of SPAM and Do Not Call lists, it is important to avoid the hostility that is associated with irrelevant offers that clog e-mail inboxes. Instead, customers should be sent messages that recognize their value, needs, and preferences. Communications should occur only when there is a significant reason, the content is highly relevant, and the timing is right.

Market segmentation analysis is used to decide to whom to send a particular offer. This requires an analysis of the customer data in order to identify those customers who are most likely to respond to a specific offer. RetailStore performs this analysis on the basis of (1) customer life cycle, (2) demonstrated interest in specific products, (3) sales associate performance, and (4) customer events.

Customers go through a life cycle. There are first time shoppers, second time shoppers, repeat shoppers, and customers who haven't shopped at RetailStore for a while. The messages and the offers they receive vary with their position in the life cycle. For example, a customer who has purchased for the first time receives a message that thanks them for their purchase, tells them about the customer care team, provides a link to the return policy, asks for a review of the item they purchased, and gives them an offer, such as a coupon, to purchase another item. Some people (often men) shop only at Christmas time. RetailStore identifies these people and just before Christmas each year sends them a special offer. At the end of the life cycle, a customer who hasn't shopped in a while receives a message saying they are missed, invites them to shop again, and receives an especially attractive coupon in order to win their business back.

RetailStore also segments customers on the basis of demonstrated interest in specific kinds of products. For example, if a customer has clicked on or purchased jewelry in the past,

<sup>6</sup> Teradata CRM is comprised of seven integrated modules that provide a comprehensive set of CRM capabilities. RetailStore currently uses some, but not all, of the modules and capabilities. We use the framework or structure provided by the various modules to describe how RetailStore executes its CRM strategy.

<sup>7</sup> *Siloed* refers to systems that are separate and do not work together. They stand apart, much like corn silos in the Midwest.

it is likely that they will be interested in jewelry in the future. Based on purchase and click stream behavior, customers are scored to determine who is most likely to respond to an offer. A customer who has recently purchased or clicked on the type of product that is going to be featured in a campaign is scored higher than someone who showed interest some time ago. The value of an interaction decays quickly over time and is taken into consideration when scoring a customer. In some cases the segmentation analysis is based on correlations. For example, if a customer has purchased a pair of red shoes, it is unlikely that they will want to buy another pair of red shoes in the same season. But they may want to buy an accessory that has shown a high correlation with the purchase of shoes, such as a purse or scarf.

Interestingly, RetailStore has found that in-store sales associates perform differently with certain kinds of customers. Over time, a customer may interact with a variety of sales associates and may purchase online. RetailStore segments the customers to determine which sales associate (or the web) results in higher profitability from each customer. Then, customers are assigned a "primary" associate (which may be the online channel). When important customer messages need to be communicated, such as the announcement of an exclusive sales event or the launch of a new customer service, the messages are sent to each customer via the assigned primary associate.

Specific customer events also trigger messages and offers. For example, customers can join RetailStore's frequent shopper club and receive benefits (such as free shipping), and after joining, they receive a welcoming message. Or if a customer abandons items in a shopping cart or saves them for later purchase, they are sent a reminder e-mail.

### CAMPAIGN MANAGEMENT

RetailStore conducts dozens of campaigns each week. It is important to manage these campaigns carefully so that (1) customers receive the best campaigns based on what is known about their needs and preferences, and (2) the

number of campaigns does not exceed customers' stated preferences in terms of the number of contacts. Customers are often identified for inclusion in multiple campaigns. For example, Suzie may be rated as a top shopper in both the jewelry and shoes categories. In this case, if the number of campaigns exceeds the number of acceptable communications with the customer, the campaigns must be prioritized so that the most appealing and relevant messages and offers are sent. For example, personalized messages are higher priority than blast communications. Once the campaigns are developed, they run automatically.

### TESTING AND LEARNING

In many organizations, the accepted approach to decision making is changing. Rather than relying on intuition or the conventional wisdom, organizations are turning to fact-based decision making.<sup>8</sup> With this approach, decisions are based on the analysis of data. For example, experiments are conducted to see what works best, much like agricultural experiments where one plot of land is watered and/or fertilized (the experimental group), the other plot is not (the control group), and the crop yield on the plots is compared.

RetailStore conducts hundreds of experiments each year, learning what works best in a header, when to combine in-store offers with online offers, when to send offers, what kinds of offers generate the most sales, and more. When there is a new idea, it is tested. Some customers are sent the usual message or offer (the control group), others the new, trial message or offer, and the results are compared. From this testing, RetailStore learns what is most effective. Data from the EDW is analyzed to select the control and experimental groups and to analyze the effectiveness of the experiments.

Overall, RetailStore's experience is that it is better to use simple rather than complex algorithms. For example, when determining market segments, use only a few important variables in the segmentation analysis. Then, test and refine the algorithms based on how well they work in practice.

<sup>8</sup> An excellent book with many examples of data or fact-based decision making is Ian Ayres, *Super Crunchers*, Bantam Books, New York, 2007.

## Putting It All Together

We tracked Suzie’s shopping experience at RetailStore.com, saw how BI enriched it, and explored the technologies that were used. Table 1 provides a high-level summary of these

three perspectives. The movement across the columns of the table reflects the passage of time. At the bottom of the table are reports, scorecards, and specialized analytics that are ongoing and are not associated with any particular point in Suzie’s shopping experience.

<b>The Customer Experience</b>	Suzie enters the RetailStore.com web site for the first time.	She searches for items of interest.	She registers and checks out.	She receives e-mails from RetailStore.
<b>The Customer Experience Enhanced by BI</b>	There is no personalization because Suzie is a first time visitor. This changes as she browses the web site.	Product recommendations are made.	Additional product recommendations are made.	Order confirmation and shipment e-mails.  Blast and personalized e-mails.
<b>The Data Warehouse and BI Software, Applications, and Processes</b>	A cookie is used to assign an ID to Suzie.	Product recommendations and their order are based on algorithms that consider relevancy and profit margins.  Market basket analyses determine what complementary products are shown.  Click stream data is stored in the data warehouse.	The product recommendation algorithm considers what is in the cart.  Registration and product purchase data are stored in the data warehouse.	Follow-up e-mails are based on Teradata CRM and EDW data (e.g., click stream, abandoned cart).  Customer profitability analysis is used to identify customers for white glove treatment and rewards.  Order fulfillment and customer service data are stored in the data warehouse.
Reports, scorecards, and specialized analyses.				

Table 1. A summary of the three perspectives.

### Driving Customers to the Web site – The Importance of Search Engines

Getting customers to visit RetailStore.com's web site is obviously important. In Suzie's case, she heard about the site from her friends. Because word of mouth advertising isn't sufficient, RetailStore also uses television ads and direct mail. The effectiveness of these marketing channels is analyzed using warehouse data. Less understood, but even more important, is the role that Google, as well as other search engines, plays in generating web site traffic. Here's why.

Many shoppers go to Google and enter what they are looking for in the search box rather than initiating their shopping at a particular company's web site. Using the Google generated list, shoppers choose those sites that seem to be the most promising. The choice of sites to visit is usually based on a combination of factors, such as where the sites appear in the list (those at the top are more likely to be visited), what is known about the site (if it is a trusted site), and the cost of the item (if it is shown). With this approach, customers go to companies' web sites through Google rather than going directly.

A Google search returns two lists – unsponsored and sponsored sites. Google uses a trademarked algorithm called PageRank to generate the list of unsponsored sites. The sponsored list of sites is determined in a different way – companies pay Google to be included. Let's explore in greater detail how all of this works.

Online retailers have a great stake in the order that Google lists web sites. If a company wants shoppers to find its web site, it is important to be near the top of the list. Because of its importance, companies put considerable effort into understanding Google's ranking algorithm and designing their web sites so that Google puts them near the top of the list.

Details about how PageRank works are a closely held secret, much like the recipe for Coca-Cola, but some things are

generally known. Google's search engine uses spiders that crawl the Internet to find web sites in order to create an index of keywords on each site.<sup>9</sup> When the spiders find a link to another site, they visit that site and create a keyword index for it. Over time, more and more information is recorded about the sites on the Internet.

Keeping in mind the information collected by spiders, several factors are known to affect Google's ranking of web sites. One factor is the frequency and locations of keywords. For example, if a keyword is used only once and it is within the body of a page, the page will not be ranked highly. Another factor is how long the web site has existed. If a web site has been around for a while, it will be ranked higher. And finally, the number of web sites that link to the site is important. The more that it is linked to, the more relevant the site is judged to be and the higher it is ranked. Of these three factors, the last one is believed to be the most important.

RetailStore.com has a group of people that is responsible for search engine optimization – making sure that the search engines place RetailStore.com highly in the list of search results. The approach taken by this group is different with the unsponsored and sponsored listings.

For the unsponsored listings, it is important to do whatever can be done to receive a high score by the ranking algorithms. This includes the careful selection and placement of keywords and getting other sites to link to RetailStore.com's site. Like other major retailers with an online presence, RetailStore has a team of people that works on web site optimization – getting the web site ranked highly by the search engines.

The listing of sponsored sites works differently. Companies bid daily (i.e., so much for each click through to their web site) on keywords in order to have their web sites listed when people enter the keywords into the search box. Their position in the listing depends on how high they bid. The higher the bid, the higher their web site appears in the list. The challenge is to optimize the price bid. Too low of a bid,

<sup>9</sup> An excellent discussion of the Google search engine is available online at HowStuffWorks (<http://computer.howstuffworks.com/google1.htm>).

and RetailStore.com isn't high on the list and does not get sufficient web site traffic. Too high, and the costs are too great relative to the amount of sales generated.

The decision about how much to bid on each keyword is based on more than the amount of traffic generated. Using the Teradata EDW, RetailStore.com can follow over time whether traffic is profitable or not. Lookers do not generate profits, buyers do.

### Moving Forward

RetailStore's marketing efforts are a journey rather than a destination. The marketing that it conducts today will be different in the future. Here are some of the changes that RetailStore has in mind.

RetailStore plans to make its e-mail campaigns more efficient and effective. One approach is to rely more on event and behavioral customer segmentation analysis to identify market segments. There will be fewer blast and more personalized e-mails. This will result in more relevant offers. There are also plans to better automate the development and execution of e-mail campaigns.

Currently, RetailStore maintains customer registration, checkout, product returns financial, and click stream data. Based on this data, RetailStore determines the profitability of customers, what messages to send to them, and what products to recommend. RetailStore has just started to use demographic data, which can be purchased from third-party sources (e.g., Harte Hanks) and appended to customers' records. For example, third-party demographic data might include customers' extended ZIP codes, birth dates, marital status, education levels, occupations, estimated incomes, and homeowner status. Additional customer segmentation analysis can be performed based on these fields to understand customers better and what they might buy given the right recommendation, offer, and promotional message. Consequently, it will be possible to send messages that are more timely and personalized.

It is important that products are always available for sale. In response, RetailStore plans to enhance its forecasting capabilities. With this information, buyers will be able to proactively look for merchandise and develop standing relationships with the right suppliers. These forecasts will be based on analytics that use warehouse sales data.

RetailStore provides employees with sales reports and scorecards. The plan is to share more information with suppliers through extranet portals to help them better monitor and analyze the sales of their products. This is what Walmart does.<sup>10</sup> This information will allow suppliers to better understand their business and help make them more profitable, which will make them want to do even more business with RetailStore.

### Conclusion

The senior management team at RetailStore has been excellent at understanding and predicting the evolution of online retailing and positioning the company to be ahead of the curve. This is an ongoing effort, however, since online retailing is constantly changing. We conclude by discussing some of the history of online retailing and what RetailStore.com and other online retailers might face in the future.

As mentioned, retailing on the Internet is relatively new and has undergone considerable change. For example, prior to the Internet bubble in the late 1990s, venture capitalists were ready to fund almost any new idea, seemingly independent of whether or not the business model made sense. This led to the Internet bubble, and for several years afterwards, venture capitalists steered clear of technology companies.

Brick and mortar retailers were slow to move to the Internet, and when they did, they often ran their online business independently as separate business units. Some people remember when you could not return items that you bought on the Internet to the company's brick and mortar store.

<sup>10</sup> Walmart has an application called Retail Link that allows suppliers to monitor and manage the goods that they supply to Walmart. More information about Retail Link can be found on the Walmart web site.

## BI at RetailStore.com

---

Now retailers' distribution channels are much better integrated. They recognize that online and in-store channels cannot be managed separately – but need to be treated holistically.

Let us speculate about the future of retailing on the Internet, with a special focus on “pure play” companies. As a starting point, it is worth first discussing the range of online retailers, because the future is likely to be different depending on the nature of the company. At one end of the continuum are retailers such as Amazon.com and eBay. They are large, well known, and offer a wide range of products. Next are medium-sized, specialty retailers such as The Golf Warehouse (www.TGW.com) that cater to a particular audience. At the far end of the continuum are the small, extremely specialized companies that cater to a limited clientele. An interesting example is the Beer Yard (www.beeryard.com) located in Wayne, Pennsylvania, that sells an amazing variety of beers. Because of the Internet, it is able to reach buyers who would never have known about them in the past. It illustrates what is called “The Long Tail” of the Internet – the ability of customers to use the Internet to find companies that offer highly specialized products.<sup>11</sup>

One of the most significant changes for “pure play” retailers is the movement to being information centric. In these companies, information is key to their ability to compete. For example, information is used to replace physical assets. Companies such as Amazon.com use their operational systems and data warehouses to forecast demand and manage their supply chain in order to reduce inventory. This, in turn, results in savings from reduced tied up capital costs.

It is interesting to think about RetailStore and other retailers who have a large Internet presence, and how it could take the substitution of information for physical assets to the extreme. Might there be a major online retailer that owns very little of the merchandise that it sells? It only connects

buyers and sellers. This is how eBay and some travel sites (e.g., Kayak) work, and there is no reason that this model cannot be used with non-auctioned items.

Many Internet companies perform disintermediation; that is, they connect buyers and sellers directly, and cut out the middleman. For example, think of the impact that travel sites like Travelocity.com have had on travel agents. Most travelers now make their reservations online, and bypass travel agents completely. Online retailers also serve a disintermediation role. They bring buyers and sellers together without the necessity of a physical store. But might they be affected by disintermediation themselves?

If so, the culprit might be Google and other search engines. Here's why. As suggested before, some shoppers, rather than going to a particular online retailer, use Google to find what they are looking for. They enter the desired product into the search box, and in return, they get a list of web sites that have the product, and sometimes the product's selling price on that site. From this list, the shoppers often make a purchase, and Google makes money from passing the shoppers on to them. You can see why companies try hard (or even pay) to be on the first page of a Google search result. Without that placement, shoppers are not likely to find them. Might Google ultimately connect shoppers directly to suppliers, bypassing online retailers?

One response to competition from search engines is to be a highly trusted retailer. In this way, having a physical store with an existing brand can be extremely helpful. If customers know that RetailStore offers high-quality products and provides excellent customer service, they are likely to shop at RetailStore.com rather than to take a chance with a retailer or supplier that they know less about.

RetailStore is effectively using information and BI to compete in the marketplace. It will be interesting to see over time how online retailing changes and how companies like RetailStore.com respond to or even lead the changes.

<sup>11</sup> Chris Anderson first coined “The Long Tail” phrase in the October 2004 issue of *Wired* magazine.

## Questions for Discussion

1. Based on your personal experiences, give examples of mass, targeted, and event-based marketing.
2. What is the best e-mail offer that you have ever received? What made it the best?
3. What is the least appropriate e-mail offer that you have received? What was your reaction to it?
4. RetailStore allows customers to either opt out or limit the number of e-mails received. Should customers also be given the option to opt out of the collection of click stream data? What are the ramifications of this for customers and retailers such as RetailStore.com?
5. A basic tenet of CRM is that the most profitable customers should be treated exceptionally well. Is this fair for other customers? Discuss.
6. At one time, RetailStore did not have a separate children's apparel department on its Internet presence. RetailStore.com carried children's apparel, but the clothing was not organized into a separate department on the web site. Give a good argument for having a separate department and promoting it through offers. How might you employ fact-based decision making to determine if it is a good idea?
7. RetailStore has to manage customers across channels, meaning that its customers shop in-store, online, or in both ways. What opportunities for retailers are associated with having multiple channels? What are the challenges with having multiple channels?
8. What do you think is the future of online retailing? Will all retailers eventually have an Internet presence? Will retail move from in-store to online completely? What will be the role of search engines?

## Additional Resources

The following resources can be accessed through the Teradata University Network/Teradata Student Network and provide additional learning opportunities.

"What is CRM? A Primer on Customer Relationship Management," Baseline Consulting Group, 2003. (This article provides an excellent introduction to CRM.)

"From the Masses to the Master: How the *Individual* Came to Matter to the Marketers," Baseline Consulting Group, 2002. (This article provides a great discussion of the evolution of marketing.)

## About the Authors

**Dr. Hugh J. Watson** is a Professor of MIS and a holder of a C. Herman and Mary Virginia Terry Chair of Business Administration in the Terry College of Business at the University of Georgia. Hugh can be reached at [hwatson@terry.uga.edu](mailto:hwatson@terry.uga.edu).

**Dr. Jeffrey A. Hoffer** is a Sherman – Standard Register Professor Emeritus of Data Management MIS, Operations Management, & Decision Sciences in the School of Business Administration at the University of Dayton. He can be contacted at [Jeff.Hoffer@notes.udayton.edu](mailto:Jeff.Hoffer@notes.udayton.edu).

**Dr. Barbara H. Wixom** is an Associate Professor in the McIntire School of Commerce at the University of Virginia. Barbara can be reached at [bwixom@virginia.edu](mailto:bwixom@virginia.edu).