BUSINESS DRIVEN START-UP

eBay is the world's largest online marketplace, with 97 million global users selling anything to anyone at a yearly total of $62 billion—more than $2,000 every second. Of course with this many sales, eBay is collecting the equivalent of the Library of Congress worth of data every three days that must be analyzed to run the business successfully. Luckily, eBay discovered Tableau!

Tableau started at Stanford when Chris Stolte, a computer scientist; Pat Hanrahan, an Academy Award–winning professor; and Christian Chabot, a savvy business leader, decided to solve the problem of helping ordinary people understand big data. The three created Tableau, which bridged two computer science disciplines: computer graphics and databases. No more need to write code or understand the relational database keys and categories; users simply drag and drop pictures of what they want to analyze. Tableau has become one of the most successful data visualization tools on the market, winning multiple awards, international expansion, and millions in revenue and spawning multiple new inventions.[3]

Tableau is revolutionizing business analytics, and this is only the beginning. Visit the Tableau website and become familiar with the tool by watching a few of the demos. Once you have a good understanding of the tool, create three questions eBay might be using Tableau to answer, including the analysis of its sales data to find patterns, business insights, and trends.

2 Trillion Rows of Data Analyzed Daily—No Problem

named Steve Smith. Simply searching the customer name would not be an ideal way to find the information because there might be 20 customers with the name Steve Smith. This is the reason the relational database model uses primary keys to identify each record uniquely. Using Steve Smith's unique ID allows a manager to search the database to identify all information associated with this customer.

A ***foreign key*** is a primary key of one table that appears as an attribute in another table and acts to provide a logical relationship between the two tables. For instance, Black Eyed Peas in Figure 6.10 is one of the musicians appearing in the *MUSICIANS* table. Its primary key, *MusicianID*, is "2." Notice that *MusicianID* also appears as an attribute in the *RECORDINGS* table. By matching these attributes, you create a relationship between the *MUSICIANS* and *RECORDINGS* tables that states the Black Eyed Peas *(MusicianID 2)* have several recordings, including The E.N.D., Monkey Business, and Elepunk. In essence, *MusicianID* in the *RECORDINGS* table creates a logical relationship (who was the musician that made the recording) to the *MUSICIANS* table. Creating the logical relationship between the tables allows managers to search the data and turn it into useful information.

Coca Cola Relational Database Example

Figure 6.11 illustrates the primary concepts of the relational database model for a sample order of soda from Coca Cola. Figure 6.11 offers an excellent example of how data is stored in a database. For example, the order number is stored in the ORDER table, and each line item is stored in the ORDER LINE table. Entities include CUSTOMER, ORDER, ORDER LINE, PRODUCT, and DISTRIBUTOR. Attributes for CUSTOMER include Customer ID, Customer Name, Contact Name, and Phone. Attributes for PRODUCT include Product ID, Description, and Price. The columns in the table contain the attributes.

Consider Hawkins Shipping, one of the distributors appearing in the DISTRIBUTOR table. Its primary key, Distributor ID, is DEN8001. Distributor ID also appears as an attribute in the ORDER table. This establishes that Hawkins Shipping (Distributor ID DEN8001) was responsible for delivering orders 34561 and 34562 to the appropriate customer(s). Therefore, Distributor ID in the ORDER table creates a logical relationship (who shipped what order) between ORDER and DISTRIBUTOR.

FIGURE 6.11

Potential Relational Database for Coca-Cola Bottling Company of Egypt (TCCBCE)

Coca-Cola Bottling Company of Egypt
Sample Sales Order

Customer: Dave's Sub Shop | Date: 8/6/2008

Quantity	Product	Price	Amount
100	Vanilla Coke	$0.55	$55

Distributor Fee $12.95
Order Total $67.95

CUSTOMER

Customer ID	*Customer Name*	*Contact Name*	*Phone*
23	Dave's Sub Shop	David Logan	(555)333-4545
43	Pizza Palace	Debbie Fernandez	(555)345-5432
765	T's Fun Zone	Tom Repicci	(555)565-6655

ORDER

Order ID	*Order Date*	*Customer ID*	*Distributor ID*	*Distributor Fee*	*Total Due*
34561	7/4/2008	23	DEN8001	$22.00	$145.75
34562	8/6/2008	23	DEN8001	$12.95	$67.95
34563	6/5/2008	765	NY9001	$29.50	$249.50

ORDER LINE

Order ID	*Line Item*	*Product ID*	*Quantity*
34561	1	12345AA	75
34561	2	12346BB	50
34561	3	12347CC	100
34562	1	12349EE	100
34563	1	12345AA	100
34563	2	12346BB	100
34563	3	12347CC	50
34563	4	12348DD	50
34563	5	12349EE	100

DISTRIBUTOR

Distributor ID	*Distributor Name*
DEN8001	Hawkins Shipping
CHI3001	ABC Trucking
NY9001	Van Distributors

PRODUCT

Product ID	*Product Description*	*Price*
12345AA	Coca-Cola	$0.55
12346BB	Diet Coke	$0.55
12347CC	Sprite	$0.55
12348DD	Diet Sprite	$0.55
12349EE	Vanilla Coke	$0.55

USING A RELATIONAL DATABASE FOR BUSINESS ADVANTAGES

LO 6.3: Identify the business advantages of a relational database.

Many business managers are familiar with Excel and other spreadsheet programs they can use to store business data. Although spreadsheets are excellent for supporting some data analysis, they offer limited functionality in terms of security, accessibility, and flexibility and can rarely scale to support business growth. From a business perspective, relational databases offer many advantages over using a text document or a spreadsheet, as displayed in Figure 6.12.

Increased Flexibility

Databases tend to mirror business structures, and a database needs to handle changes quickly and easily, just as any business needs to be able to do. Equally important, databases need to provide flexibility in allowing each user to access the information in whatever way best suits his or her needs. The distinction between logical and physical views is important in understanding flexible database user views. The ***physical view of information*** deals with the physical storage of information on a storage device. The ***logical view of information*** focuses on how individual users logically access information to meet their own particular business needs.

In the database illustration from Figure 6.10, for example, one user could perform a query to determine which recordings had a track length of four minutes or more. At the same time, another user could perform an analysis to determine the distribution of recordings as they relate to the different categories. For example, are there more R&B recordings than rock, or are they evenly distributed? This example demonstrates that although a database has only one physical view, it can easily support multiple logical views that provide for flexibility.

Consider another example—a mail-order business. One user might want a report presented in alphabetical format, in which case, the last name should appear before first name. Another user, working with a catalog mailing system, would want customer names appearing as first name and then last name. Both are easily achievable but different logical views of the same physical information.

Increased Scalability and Performance

In its first year of operation, the official website of the American Family Immigration History Center, www.ellisisland.org, generated more than 2.5 billion hits. The site offers immigration information about people who entered America through the Port of New York and Ellis Island between 1892 and 1924. The database contains more than 25 million passenger names that are correlated to 3.5 million images of ships' manifests.[4]

The database had to be scalable to handle the massive volumes of information and the large numbers of users expected for the launch of the website. In addition, the database

FIGURE 6.12
Business Advantages of a Relational Database