DATABASE MANAGEMENT SYSTEMS

This page intentionally left blank

DATABASE MANAGEMENT SYSTEMS: A BUSINESS-ORIENTED APPROACH USING ORACLE, MySQL, AND MS ACCESS

SOTIRIOS ZYGIARIS
University of Maryland, USA and PMU University, KSA

United Kingdom – North America – Japan – India – Malaysia – China

Emerald Publishing Limited
Howard House, Wagon Lane, Bingley BD16 1WA, UK

First edition 2018

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-78756-696-5 (Print)
ISBN: 978-1-78756-695-8 (Online)
ISBN: 978-1-78756-697-2 (Epub)

DBMS used in the book
ORACLE 11g APEX 4.2
Maria DB (formerly MySQL) from XAMPP stack 5.5.30
Microsoft Access 2013
Screenshots from Oracle 11g Application Express 4.2. Copyright Oracle, used with permission by Oracle Corporation.

To the memory of my father. To my dear mother and my family Nitsa, Maria, and Georgia

This page intentionally left blank

Acknowledgments

This book has been a group effort, and there many people who contributed and their help meant a great deal to me. I am in a continual state of gratitude to all my PMU students who helped with their remarks while I was teaching the course. I am grateful because I received important constructive comments from a student's point of view. I am grateful to my reviewers Dr Malik F. Saleh and Isidoros Passas whose ideas and suggestions are scattered throughout this book. Lastly, I owe a special debt of gratitude to Krista Marie Osborne PMU's supervisor of Advising and Learning Enrichment Services for her valuable editing support.

This page intentionally left blank

Contents

PART II: DATABASE DEVELOPMENT

This page intentionally left blank

About the Author

Dr. **Sotirios Zygiaris** is an IT professional working with ecommerce and business intelligence applications. He received his PhD in Innovation Management from Aristotle University, Greece and his main degrees from Lamar University in Texas, USA in Computer Science (BSc) and MIS (MSc).

He has six years of IT experience working for Fina Oil and Chemical. He has also worked as a Lecturer at the University of Sheffield teaching computer science courses. With 12 years of research management experience managing European Union multinational research projects in technological innovation and industrial research and development, he is a European Union international expert in innovation and smart city technological platforms.

Dr Zygiaris was an Associate Professor of Information Technology at University of Maryland College. Currently, he is an Assistant Professor of Management Information Systems at PMU in the Kingdom of Saudi Arabia. He is the author of two books related to business transformation using technological advancements and business process re-engineering.

This page intentionally left blank

Preface

This book is meant to help business students and professionals to learn the concept of databases in all aspects from design and implementation, to management of databases. A database is a central component of every business application. You as a business student must understand how data are stored, retrieved, and managed in your corporate environment. With the advancing of enterprise computing and cloud technologies, managers are responsible to retrieve data ad hoc and construct business reports for decision-making and storytelling. Managers having solid knowledge of database technologies are empowered with analytical and reporting skills. For this reason, an advancing number of universities include in their business administration curriculum database management in their skill set profile.

The book manages to codify all information technology issues in database design and implementation into terms that can be easily understood by business students. For example, other relative textbooks cover database design in up to four chapters, at an average of seventy pages, presenting complex design methodologies that are very difficult to be followed by non-IT students. On the contrary, this book has codified the database design process in twenty pages in Chapter 2, presenting in discrete steps the database designing process.

One case study, "Atza Inc.," is used throughout the book that follows all database aspects from design to implementation to management. This way, the students follow all phases of a database cycle in examples and illustrations provided in each book chapter. Four more business cases in project management, human resources, inventory management, and customer service are specified. These cases are allocated to students as work for projects and assignments. *The case study for Atza Inc. refers to a fictitious company that the book purposes. Any data used in this book are imaginary and are not related to any real company or persons.*

The book covers three different database management systems: MS-Access, ORACLE, and MariaDB, facilitating the adaptation to different lab specifications and database environments. To avoid lengthy technical implementation details, the readers are directed to the book's YouTube video channel where all technical implementation details are presented in each one of the three corresponding database management systems. For example, when the student has to learn how to create database tables in ORACLE, he/she follows the book link to the relative video. New editions of database management systems are covered by modifying the videos without any changes to the textbook.

The first part of the book covers the topics on database environment and database design. The second part from Chapters 3 to 7 covers the structured

query language (SQL) extensively as a basic management tool to build, store, retrieve data, and make business reports by processing the data. The third part covers database management tools in managing and securing the database, creating business intelligence, and understanding the concept of distributed databases in corporate environments.

To the Student

Every feature of this book is designed to help you learn the database management concept. I have a few goals for you in this book: to understand the importance of databases in the business environment and to assist you in developing problem-solving techniques using databases. I want you to learn databases in a new way, removing and codifying all technical and theoretical complexities and presenting the topic in manner that will be appreciated by non-IT students.

To improve the learning experience, read every chapter carefully. Follow the links to YouTube videos presenting visual and technical implementation details in a gradual manner. At the end of the book, follow the terminology. It is very important to "speak the terms" when taking interviews, engaging in discussions, or writing business reports. Answer the review questions at the end of each chapter. When you feel ready, test your knowledge by taking the online multiple-choice quiz for each chapter. Make sure that you go back and review the chapter when questions are not answered correctly.

I provide for you two ways to receive practical experience. First, follow the forecasting case study in problems and exercises at the end of Chapter 2 and solve all problems and exercises in each chapter following the same case study. Follow the "Atza Inc." examples illustrated in chapter and do the exercises using the forecasting case study. All problems and exercises at the end of each chapter refer to this case study.

Second, choose one or more of the four case studies in project management, human resources, inventory management, or customer service. This is your main assignment that must be implemented in the database management system that you are using. In Chapter 2, you must do the database design. In Chapter 3, you will implement the proposed design into a database. In Chapters 4–7, you must include all SQL queries necessary for business reporting. In Chapter 8, you will include all database management features for optimizing and securing your database. In Chapter 9, you will create a web application using visual tools. The application ensures that all operations such as data entry, update, deletion, and queries are made through client web pages using forms and reports. In Chapter 10, you will enhance your application inserting three-dimensional business reports.

Lastly, this book imparts the necessary structured and codified knowledge and leaves an impression that database management is not reserved only for IT students with high technical and programing skills. In fact, with the right guidance from your course instructor, all students from every discipline can master database management, including you!

This page intentionally left blank

To the Instructor

I would like to thank all who have chosen this book edition as a textbook for delivering your course. This textbook is designed as a one-semester course in database management systems. Although it is intended for non-IT students, IT students will also find it very resourceful and concise. "To the point" is a frequent comment I had received from students.

One case study, "Atza Inc." is used throughout the book that follows all database aspects from design to implementation to management. The first part of the book covers the topics on database environment and database design. The second part from Chapter 3 to 7 covers the structured query language (SQL) extensively, as a basic management tool to build, store, retrieve data, and make business reports by processing the data. The third part covers database management tools in managing and securing the database, creating business intelligence, and understanding the concept of distributed databases in corporate environments.

Once the students conceptualize the database environment in Chapter 1, they must understand the database design process from requirement document and conceptual design (E−R diagram) to logical design (ERM diagram). At this point, they must choose one of the four case studies in project management, human resources, inventory management, or customer service, and conduct a database design exercise. The "Atza Inc." case study will be used as the main term project assigned to single or group of projects.

In Chapter 3, where the database design must be implemented into a database, you must choose one of the three covered DBMS (MS-Access, ORACLE, or MARIADB) as your class working system. Follow the correspondent's DBMS YouTube videos to create and populate with data a database. In Chapters 4−7, you must include all SQL queries necessary for business reporting. In Chapter 8, you will include all database management features for optimizing and securing your database. In Chapter 9, you will create a web application using visual tools. The application ensures that all operations such as data entry, update, deletion, and queries are made through client web pages using forms and reports. In Chapter 10, you will enhance your application inserting three-dimensional business reports. Examples of term project assignments are provided online in the instructor's resources.

Use the forecasting case study in the "Problems and Exercises" sections from Chapter 2 for small assignments and homework at the end of each chapter. Make sure that students follow the "Atza Inc." examples illustrated in the chapter and do the exercises using the selected case study. All problems and exercises at the end of each chapter refer to this case study. Examples of

assignments and model answers are provided online in the instructor's resources. Each chapter is covered by:

- Online multiple-choice quizzes (uploadable in text form or to a form for direct upload into blackboard or canvas) along with the answer key in the instructor's resources.
- A PowerPoint presentation for each chapter in the instructor's resources.
- Solutions to review questions and exercises.

PART I: UNDERSTANDING THE DATABASE APPROACH

Chapter 1: The DataBase Environment

Learning Objectives

- To comprehend the value and use of data in producing valuable information for decision-making and the creation of business knowledge.
- To understand the environment in which data are stored and manipulated using database management systems (DBMS).

Chapter 2: DataBase Design

Learning Objectives

- To understand the database development life cycle from analysis to implementation.
- To apply the database process from conceptual to logical design for optimization through normalization.

This page intentionally left blank

Chapter 1

The DataBase Environment

1.1. From Data to Information to Knowledge

The first component in the word *Database* derives from the plural of Latin datum, which means "a given thing."[1] In fact, **data** are determined as a vast amount of unprocessed numbers, characters, images, and video and audio elements including supermarket invoices, audit videos, location coordinates, the orders taken by sales representatives, and the check-in and checkout timesheets of employees over the past year. In our physical world, an abundance of data is stored in the form of physical records. Companies keep their everyday business transactions in accounting books like accounts payable or the general ledger for many years. Today these data are being kept in digital form.

Data in either physical books or digital files are valuable company assets that document historical facts, verify business transactions, and improve the decision-making process. Although the physically kept data provide important documentation for past business activities, they are difficult to be processed by managers due to the vast amount of paper-based facts. For example, the production manager examines the book of finished orders for the past year consisting per se of 8,000 order invoices. It is impossible to memorize these figures and make rational conclusions about the peaks and gaps in the production lines.

Humans can process a limited amount of data

Digital data are processed by computers that have the capacity to memorize numerous data forms and process them using information systems. For this reason, businesses are digitizing physical data to illustrate in full capacity their digital business content. Information systems can "read" numerous digitized data and store them in the memory for processing. For example, in bank transactions, in stock trade price fluctuations, in payments received in ecommerce, all data are stored digitally and they are ready to be processed by information systems. Figure 1.1 illustrates partial data stored in a spreadsheet recording the sales of a superstore. The number of records or rows representing sales reaches, on an average, 3,000 record sales every day. These records are entered in digital

[1] http://www.etymonline.com/index.php?term=data

Order ID	Order Date	Order Priority	Order Quantity	Sales	Discount
3	10/13/10	Low	6	261.54	0.04
293	10/1/12	High	49	10123.02	0.07
293	10/1/12	High	27	244.57	0.01
483	7/10/11	High	30	4965.7595	0.08
515	8/28/10	Not Specified	19	394.27	0.08
515	8/28/10	Not Specified	21	146.69	0.05
613	6/17/11	High	12	93.54	0.03
613	6/17/11	High	22	905.08	0.09
643	3/24/11	High	21	2781.82	0.07
678	2/26/10	Low	44	228.41	0.07
807	11/23/10	Medium	45	196.85	0.01
807	11/23/10	Medium	32	124.56	0.04
868	6/8/12	Not Specified	32	716.84	0
868	6/8/12	Not Specified	31	1474.33	0.04
933	8/4/12	Not Specified	15	80.61	0.02
995	5/30/11	Medium	46	1815.49	0.03
998	11/25/09	Not Specified	16	248.26	0.07
1154	2/14/12	Critical	44	4462.23	0.04
1154	2/14/12	Critical	11	663.784	0.25
1344	4/15/12	Low	15	834.904	0.06
1344	4/15/12	Low	18	2480.9205	0.01
1412	3/12/10	Not Specified	13	59.03	0.1
1412	3/12/10	Not Specified	21	97.48	0.05
1539	3/9/11	Low	33	511.83	0.1
1539	3/9/11	Low	38	184.99	0.05
1540	8/4/12	High	30	80.9	0.09
1702	5/6/11	High	23	67.24	0.06
1761	12/23/10	High	25	12028.23	0.01
1792	11/8/10	Low	28	370.48	0.04
2275	10/21/12	Not Specified	49	278	0.08
2277	1/1/11	Not Specified	10	66.54	0.01
2277	1/1/11	Not Specified	21	845.32	0.06
2532	10/10/11	High	39	282.07	0.03

Figure 1.1. Sample Data Presenting Superstore Sales.[2]

form into the superstore's information system, where they are ready to be processed by managers. **This data processing may take different forms,** for example:

(1) Can I insert a new sales record with Order ID equal to 890?
(2) What is the sale amount for the order with ID equal to 998?
(3) Can I change order quantity from 22 to 18 for sale with Order ID equal to 613?
(4) Can I delete the sale with Order ID equal to 1761?

Processing data involves every date management that includes data retrieval and transformations that take the form of **C**reate, **R**ead, **U**pdate, and **D**elete (**CRUD**) transactions.

C Can I insert a new sales record with Order ID equal to 890?

R What is the sale amount for the Order with ID equal to 998?

U Can I change quantity from 22 to 18 for sale with Order ID equal to 613?

D Can I cancel sale with Order ID equal to 1761?

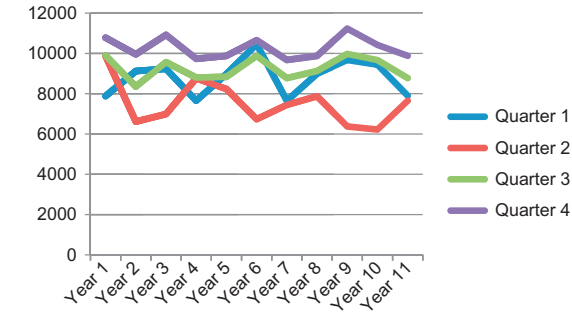
The **R** operation just reads the data without making any modifications to them. The **C**, **U**, and **D** operations modify the data. The benefit of fast

[2]https://community.tableau.com/docs/DOC-1236

accessibility and transactional effectiveness is very important in facilitating business transactions. The most important managerial benefit is taking place when appropriate data processing takes place to support the decision-making process.

Superstore assumes 3,000 sales daily with an estimated number of sales over one million per year. Figure 1.2 represents examples of the generated information for Superstore. This information takes the form of a managerial report

Year	Quarter 1	Quarter 2	Quarter 3	Quarter 4
1	7883	9865	9887	10776
2	9123	6613	8358	9954
3	9234	6987	9564	10923
4	7654	8765	8799	9733
5	8975	8234	8841	9876
6	10454	6733	9876	10654
7	7654	7443	8775	9674
8	8992	7872	9123	9876
9	9674	6376	9976	11231
10	9444	6223	9670	10421
11	7899	7654	8777	9883



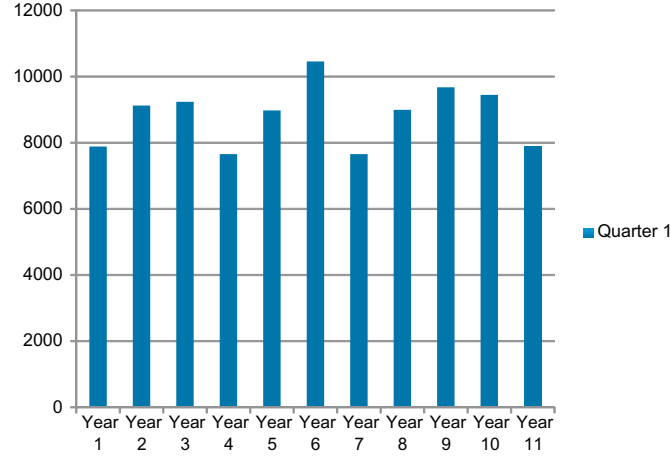


Figure 1.2. Transforming Data to Information for Superstore.

for presenting quarterly sales for the past 10 years. Over 12 million sales figures (data) were processed to produce information regarding quarterly sales in a table form. The sales report includes a bar chart for the monthly sales in January over the past year. Quarterly data are illustrated as a histogram to show the linear progress of sales.

All these reports are referred as information. **Information is understanding the meaning of data in a manner that reduces uncertainty**. The initial claim that "Humans can process a limited amount of data" is verified since **managers cannot memorize and handle such a vast amount (millions) of data**. The sales figures must be presented in a manner useful to managers to make decisions, which take the form of reported information. The superstore sales manager capitalizes on this information to examine the progress of sales for the past ten years or the historical progress of data for the month of January for decision-making in relation to sales management.

The observation of information in consecutive terms leads to **knowledge** creation, when managers learn the behavioral patterns of their business. For example, observing the quarterly superstore sales in Figure 1.2, the first quarter sales are relatively high (over 9,500) and the second quarter sales drop significantly (in years 2, 3, 6, 9, and 10). **This information leads to specific knowledge as pattern**. When increased sales appear in the first quarter, the sales manager must respond proactively to support sales in the second quarter. In the third quarter, sales over 9,000 usually lead to high sales in the fourth quarter (in years 1, 3, 6, 8, 9, and 10). Managers could also act proactively to increase stock levels to satisfy the projected high demand for the fourth quarter.

Observing information leads to the creation of knowledge

The derived information benefits a business when it is:

- **Accurate.** Valid data are the source of accurate information. Data entry methods must be used to ensure the validity of data.
- **Complete.** Decision-making based on partially available information leads to faulty decisions.
- **Updated.** Information is valid only for the time interval that data are valid. Any changes in data modify also the context of information.

The transformation of data to information is succeeded through **Information Systems**. These systems receive data as input, which are processed to produce information, which are in the form of managerial reports. Information systems exist in sales, finance, production, and all functional units of a business. Information systems include **applications** that do specific tasks like computing performance bonuses for employees or computing the reorder level of a product in the inventory. Information systems are used across various industrial or service business sectors. For example, in a banking information system, data are collected and stored from everyday transactions and from other external sources. These data are processed by the information system and produce managerial reports to assess cash flow, liquidity, interest rates, or credit policies.

1.2. Databases

Data are stored permanently in storage devices and they are retrieved in memory by information systems. When a bank starts its daily operations, all customer account data are retrieved and the updated account data are stored again to the database at the end of the day. Permanent storage occurs in **hardware** like in a computer's hard disk or in data centers on the cloud. They can be retrieved and stored back again at any time. Data are stored in a structured form, which is called a **base**. The base could take a form of a two-dimensional table, exactly the way fields are stored in a worksheet. See the table structure in Figure 1.1, where the columns represent the table's attributes and the rows the table's data records.

A simplified database structure is a collection of interrelated two-dimensional tables. Businesses create and maintain databases as a corporate resource to ensure data sourcing in their information systems. Each table represents a business entity like a salesperson, a department, a product, or an employee. To understand the database approach, a short review of the evolutions of data storage forms from file to database systems is required.

The business information systems in the 1960s, 1970s, and early 1980s used **file systems** to store, retrieve, and manipulate data. Many early business applications in this era were implemented using **COBOL** (**Co**mmon **B**usiness **O**riented **L**anguage) programming language. These applications were very robust, and they were used successfully in early information systems. They are referred to us today as **legacy** systems.

1.2.1. File Systems

File systems are used to store and supply data to specific information systems. For instance, a sales data file stores data used by the sale information system. The accounting data files support the accounting information system. A file system (illustrated in Figure 1.3) as two-dimensional table is characterized by:

- **Columns or attributes** that describe a specific characteristic of the file, for example, the salesperson's name:

Name
Jill Adams

- **Records or rows** where the data occurrences appear, for example:

24/7/16	ARma Co.	034345900	7643.12	Jill Adams	037665233

- **Fields** that describe each cell in the two-dimensional table, for example:

24/7/16	ARma Co.	034345900	7643.12	Jill Adams	037665233

Sales File System

Date	Customer	Customer's Telephone	Amount	Salesperson	Salesperson's Telephone
24/7/16	ARma Co.	034345900	7,643.12	Jill Adams	037665233
24/7/16	Belm Inc.	037654377	19,652.88	Nick Manner	033665234
27/7/16	RealTrack	035723442	3,452.76	Grace Ramos	037775298
28/7/16	ARma Co.	034345901	12,623.76	Grace Ramos	037775298
28/7/16	RealTrack	035723442	9,231.80	Sam Koori	039771663
29/7/16	Belm Inc.	037654377	3,887.23	Jill Adams	037665234

Salesperson File System

Name	Telephone 1	Telephone 2	Office	TOTAL SALES	Hire Date
Jill Adams	035663452	037665233	F3345	356,435.30	23/12/13
Nick Manner	034577623		F8743	436,652.90	2/5/15
Grace Ramos	037887234	037775298	G5432	345,523,82	18/8/15
Dorian Lee	034587763		S3456	234,411.16	30/5/14
Sam Koori	036778299		S4567	90,887.23	19/9/16
Rio Astros	034667266	037665234	G3452	22,786.61	20/11/15

Figure 1.3. File Systems.

File systems have a strong dependency, which is explained as a complex process to update their structure. For example, adding a new column to the sales files system such as customer's address requires programming expertise to modify the file's structure. Therefore, the file structure is an internal part in an application. If modification of the file system is needed, then programming skills are required. This reliance of a file system to its application is referred as **data dependency**.

Another important drawback of files systems is the dispersion of the same data in many different locations. The same data appear in many different files or in many fields within a file. For example, salesperson's telephone number appears in the sale file system and also in the salesperson's file system. This might create important data inconsistencies. For example, the telephone number for *ARma Co*. is 034345900 in record 1 and 034345901 in record 4. The negative effect of having the same data stored in multiply file fields is called **data redundancy**. This negative effect of data redundancy may lead to:

- **Errors in data entry,** from entering the same data in multiple fields.
- **Data inconsistency.** If the salesperson's telephone number is changed in one field, it must also be updated in all other fields that the data appear. Omitting to update some field creates data inconsistency.
- **Increased update anomalies.** The update anomalies refer to impact of changes in the insert, delete, and update operations. If a new sales record must be inserted, all salesperson's data must be reentered. What would happen to the sales data when the corresponding salesperson must be deleted?

- **Variations in data views.** Another significant problem of file systems is that redundancy may provide incomplete or misleading information. Having salesperson's telephone data redundant in multiple files produces two variant views of the same information.

Another significant problem is keeping data **integrity**, which is the assurance that data are accurate and consistent throughout their entire life cycle. Since file systems are not related to each other, the enforcement of business rules in the update operations might create consistency problems within the operational business logic. For example, when a new sale record is inserted, all business rules must be validated. For example, in record 5 of the sales file system, the sale takes place on 28/7/16 before the hire date 19/9/16 of salesperson Sam Koori. Obviously, a validation rule must be set for this case before the sales data are inserted into the file. A business rule has to be enforced to check the date of sale against the date of hire between the two file systems. Another business rule could be the addition of sale amount to the total sales of a salesperson when a new sale record is entered.

Data integrity problems occur usually because the files are structured in a form to fit the programming requirements rather than real business logical views. Information system developers have their own view of the file structures. This view does not necessarily match the business view. For example, mixing sales data with salesperson data in the sales files system is not in accord with the manager's logical view. The sale and salesperson are two distinct and discrete notions in the manager's view.

Repeating fields are multivalued fields like the telephone number of a salesperson. In data files, repeating fields are restricted by the file structure. For example, up to two telephone numbers are allowed in a file system. However, what happens if the business requirement defines unlimited numbers for telephone numbers for each salesperson?

1.2.2. The Database Approach

Databases as an evolution of file systems have structural characteristics that eliminate most of the problems imposed by the structure of file systems. Figure 1.4 illustrates the database approach for the sales and salesperson file systems and responds to the problems encountered in file systems. Databases are following the same two-dimensional structure, which is called **table**, to represent data with the following structural characteristics:

Each table corresponds to a single business entity, which is an autonomous physical occurrence of a group of characteristics like products, employees, sales, or business departments. Databases have a close relationship between the structure of a table and a real business entity.

Tables mirror single entities and must be independent from each other, which means that they should not contain the same data. In Figure 1.4, there are four tables – salesperson, sales, customer, and salesperson's telephone number – each representing their corresponding entities.

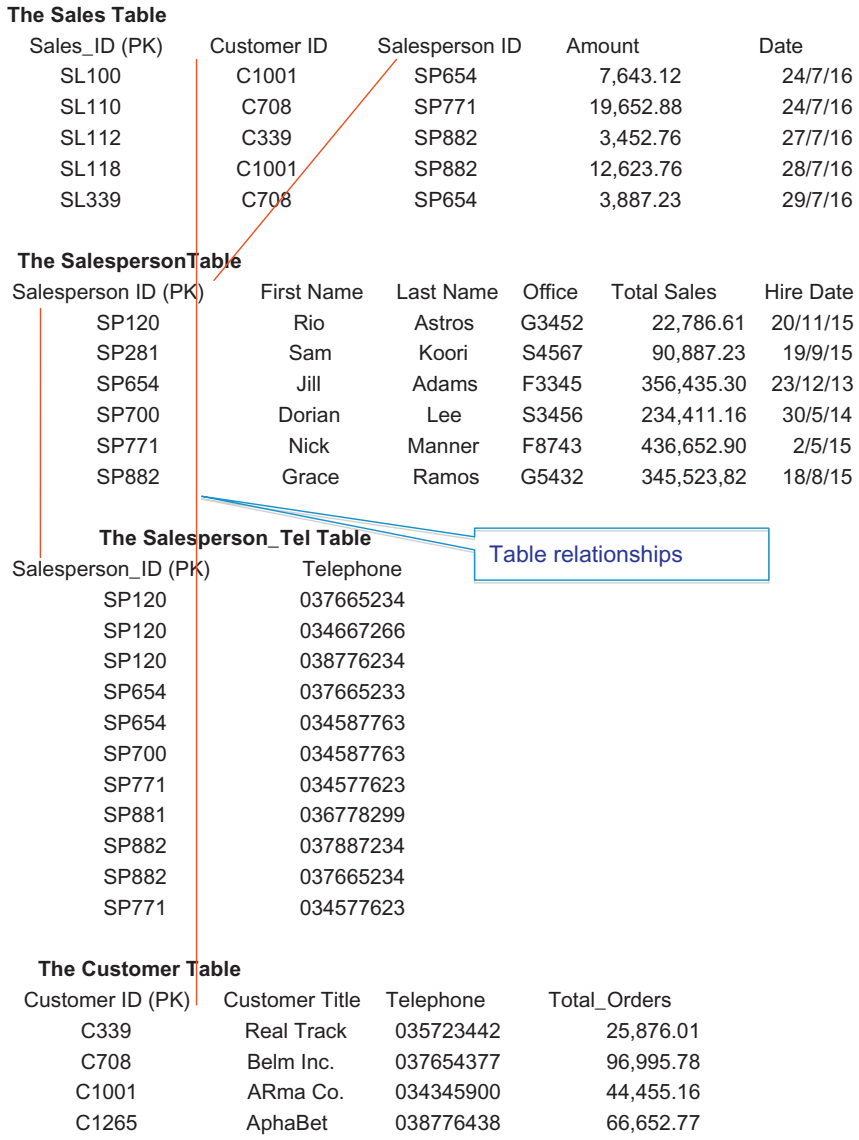


Figure 1.4. The Database Approach – Sales Example.

Table rows are unique. In the database approach, every record or row is unique and no duplicates are allowed. In the Sales tables, for example, each row is unique representing the logical view of a sale. Each row is identified by a unique attribute, which is called the **primary key** of the table. For example, the sale with primary key SL118 is unique. That means that the primary key

uniquely identifies the corresponding row of the table. For every table, a column or an attribute must be selected to be the primary key. A primary key must be unique and unchanged for every data row. Examples of primary keys could be the national ID, product number, or employee number. An attribute like a passport number could not be a primary key because it will change when the passport expires.

Why Databases?

Tables are related. Common fields are used between the tables to establish relationships. For example, in the sales table for sale SL110, the customer detail is known through the relationship established between Sales and Customer tables based on the common attribute *Customer ID*. For sale SL110, the corresponding customer is C708, and through this relationship, it is known that the title of the customer is Belm Inc. with telephone number 037654377 and total amount of orders 96,995.78. Customer ID is also a primary key on the Customer table. The relationship between the Sale and Customer tables is established through the common attribute *Customer ID.*

Data redundancy is eliminated *by restrict*ing data insertion to a single database location. Data are inserted in only one location avoiding duplicates. For example, for the sale *SL118,* the corresponding customer is *C1001.* In turn, customer *C1001* is also a primary key in the Customer table, since it is unique. A request to access the customer details follows the relationship and all details for customer C1001 are accessible. This eliminates the duplication of customer data.

The enforcement of data integrity takes place through the validation of business rules during the data entry or data update. Validation rules may be set. For example, the date of sale to be always later than the sale of hire in the Salesperson table. In addition, when a new sales record is inserted, the Total_Sales and Total_Orders columns must be updated too.

Repeating field entries could be allowed with unlimited occurrences according to business needs. For example, salesperson's telephone is inserted in the Salesperson_Tel table. The telephone data can be accessed at any time using the Salesperson_*ID* column as the primary key. Salesperson S654 Gill has two numbers 037665233 and 034587763. In the database approach, repeating fields may have **zero (a salesperson with no telephone number), one, or unlimited** occurrences.

Databases have external similarities with spreadsheets since they both allow the creation of multiple tables. A spreadsheet is not meeting the technical specifications of databases since many features described earlier do not apply. For example, the unique identification of rows, the relationships between tables or the delete, and the insert and update integrity validation are not included in spreadsheet applications.

1.3. The Database Environment

Legacy systems often expressed as file systems present a major weakness as the source of data redundancy and integrity issues that are presented in

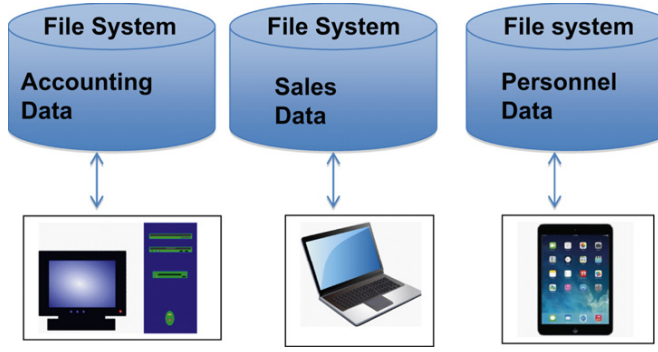


Figure 1.5. The File System Environment.

Section 1.2.1. The problematic structure of file systems is that different data are kept and maintained among business departments, as is illustrated in Figure 1.5.

For example, the accounting department keeps its own records of accounts, the sales department keeps its own records of customer and sales and inventory data, and the production department owns data about orders and inventory. Frequently, data duplicates are kept between different departments increasing the risk for data inconsistencies. For example, duplicates of inventory are kept in both sales and production department.

This weakness is called **"fragmented piece meal,"** which means the data are fragmented into different departments and managers do not have a holistic view of a business as a whole. Beside the structural characteristics of file systems, the fragmentation of business data creates inefficiency in managerial decision-making, because decisions are made having a partial view of business facts.

For example, when the production manager must draft the monthly production schedule, he/she needs data regarding the accounting department (budget availability), the marketing (sales forecasting), the personnel (staff availability), and the inventory (current stock levels) to make an optimal decision. Since the production manager does not have a complete view of these data, the effectiveness in decision-making is reduced.

1.3.1. Database Management

Figure 1.6 illustrates the database system approach in comparison with the file system approach of Figure 1.5. In database systems, all corporate data are stored under a single structure, the database, which is made available across all business departmeAn information system contains applications, such as order monitoring, where an order can be tracked as it progresses from the sales, accounting, production, distribution, and customer service departments. At any time, managers may see the progress of an order since they have a holistic view of company's data. This integrated approach improves management efficiency

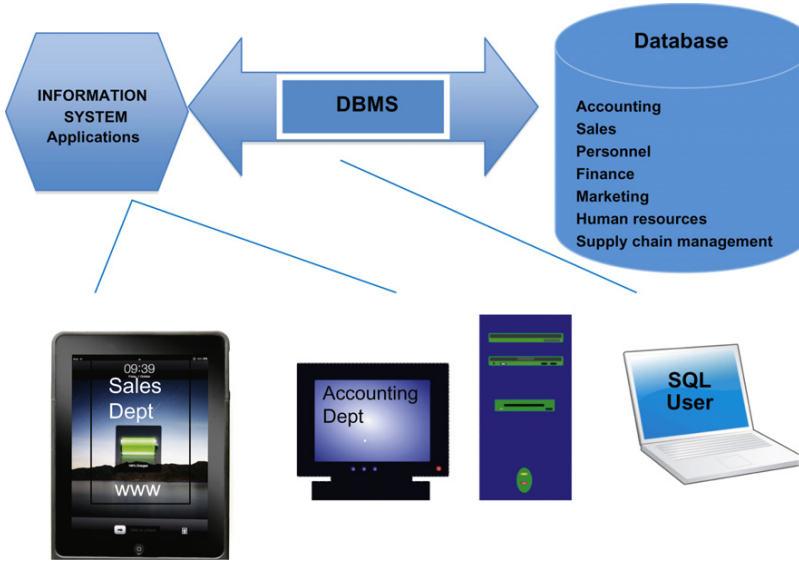


Figure 1.6. The Database Environment.

in decision-making and capacitates the customer service to respond to cross-departmental customer inquiries.

This cross-departmental data sharing capacity often leads to the empowerment of cooperation between the business departments and enforces cross-functional horizontal processes across the departments of a business. Such transformations are implemented in **Enterprise Resource Planning Systems (ERP)** forming horizontal business chains, such as **Customer Relationship Management (CRM) or Supply Chain Management.**

Databases must be stored and retrieved from **hardware capable of offering permanent storage** such as a hard disk, optical disk, and lately in data centers in the cloud. Databases are used by information systems to execute business applications. This type of application apply to all functional business units such as accounting, finance, sales, marketing, personnel, human resources, supply chain management, production, and distribution. Managers must be provided with access to read the data and also update – insert new records or delete existing records. As databases are stored mainly in the hard disk of the computer, low-level programming skills are required to access the disk sector where the data reside. Therefore, managers must be provided with end user **applications** to access the stored data.

In Databases, managers have a holistic view of business data

Applications access the database to perform the four basic CRUD operations (see Section 1.1). A business information system contains applications that utilize the company's database. Applications are built on specific operating

systems such as Microsoft Windows for PCs, MacOS for Apple computers, and Linux as open-source operating system. Businesses are using more powerful computers such as IBM's AS400, which employ corporate-oriented multiuser operating systems like Windows NT, UNIX, and Red Hat Linux. Recently, mobile applications are gaining ground with Android and IOS operating systems for Apple's IPhone. Usually applications developers implement different versions of applications to fit the technical specifications of each operating system.

Data dependency between applications and the database is eliminated. All business applications share the common corporate database. Changes in the database do not require modifications in the applications.

The software system that provides the ability to the users and the applications to access and manipulate the data is called DataBase Management System (DBMS). Users may have access to the data through the applications. Applications are developed for specific DBMS, or in most cases, a version of the application supports different DBMS. Alternatively, if the managers are familiar with a database management language, which is called SQL, users may directly access the data through a DBMS without the use of applications. Structured Query Language (SQL) is a major managerial tool that empowers managers to create ad-hoc questions to the database. SQL will be examined in detail in Chapters 3–8.

1.3.2. Types of DBMS

Some examples of DBMS are ORACLE, MariaDB, SQL Server, DB2, and MS-Access. Different types of DBMS serve different requirements. These requirements refer to the size, the local expansion of data, and the analytical requirements of businesses.

1.3.2.1. First Requirement: Number of Database End Users

Single-user DBMS supports only one user at a time. A user may access the database using an application or directly through the DBMS. These applications are limited in size and functionalities and they can be easily transferred from one computer to another. Small companies may use a single-user database application because it is affordable and requires less maintenance and operational cost. An example of a single-user database is Microsoft's Access. All others such as DBMS ORACLE, MySQL, MariaDB, SQL Server, and DB2 are supporting multiuser operation.

Multiuser databases can support many users at the same time. Managers may process the same data simultaneously. Usually companies operate multiuser DBMS. The database and applications are located in a central server and users are connected through the company's network or intranet. If the database is located on a web server, users have access to the database through an Internet connection. **Workgroup** DBMS applies to small organizations with less than 50 users. Typical applications cover the transactional processing system and MIS functions like finance, marketing, and human resources management.

If more than 50 employees are using the database application, an **Enterprise** DBMS is needed. These systems are associated with the implementation of **ERPs**,

like SAP, and the formulation of supply or CRM chains in a business. Many times, these supply chain applications are extended among several cooperating businesses. Another association of enterprise DBMS is the use of **Data Warehouses**. They combine databases from different sources such as cooperating businesses, legacy file systems, or external open databases such as government databases. For example, a pharmaceutical corporation data warehouse may contain databases from pharmacies, legacy files, and prescription data from hospital clinics.

1.3.2.2. Second Requirement: Dispersion of Data Storage

In **centralized** databases, data are located in single data storage. This is the usual case in small- and medium-sized enterprises. Large corporations with various numbers of branches distribute their data into different database storages using distributed databases. Geographically dispersed corporations store their data in **distributed** databases. For example, a multinational bank may have a distributed database that connects all local databases in bank branches. The banks' data in each country may also be distributed to different branches across the country. **Spanner**[3] is Google's scalable, multiversion, globally distributed database. It is the first system to distribute data in a global scale and support externally consistent distributed transactions. This book examines distributed databases in Chapter 11.

1.3.2.3. Third Requirement: Type of Data Usage

Operational DBMS are needed when a business requires applications to support the day-to-day business operations and transactions in accounting, sales, production, or personnel. The derived reports are based on company's data and the produced information supports managerial decisions. These decisions refer mainly to tactical decisions in all business units. When strategic or executive decisions take place, a manager needs to access a Data Warehouse that includes historical facts, external databases, and data from the World Wide Web. Therefore, internal and external information is needed in executive decision-making to examine the business environment before they make a decision.

This requires **analytical** data processing to aggregate and utilize all these diverse sources of data. The analytical processing of huge amount of data leads to business intelligence. **Business intelligence** uses a great degree of **Data Mining**, which is the process of analyzing data from different perspectives and summarizing them into useful information. **Online Analytical Processing (OLAP)**[4] tools are used to achieve this analytical capacity. XML DBMS are able to utilize semi-structured data derived from web pages storing data in **XML** format. These are specific DBMS that can handle text data like Xbase,[5] but most of the DBMS support import in XML format. The analytical processes for building business intelligence will be examined in Chapter 10.

[3] http://www.wired.com/2012/11/google-spanner-time/
[4] http://www.oracle.com/technetwork/database/options/olap/index.html
[5] http://linux.techass.com/projects/xdb/

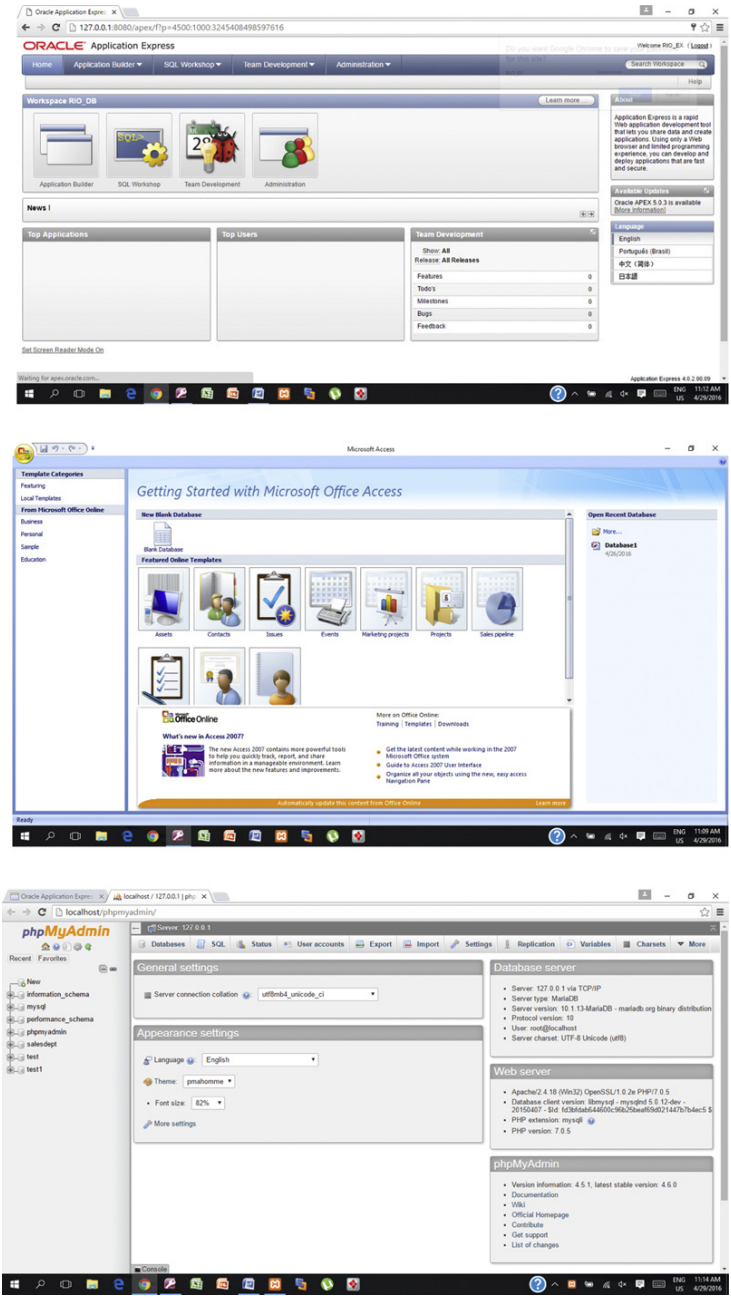


Figure 1.7. ORACLE, MS-Access, and MariaDB DBMS Screenshots.

This textbook is using multiple DBMS – ORACLE express edition 11g along with MS-Access and MariaDB. Figure 1.7 presents examples of screenshots of DBMS.

Summary

Data are determined as a vast amount of unprocessed numbers, characters, images, and video and audio elements. Digital formed data are processed by computers that have the capacity to memorize thousands of millions of digital data and process them using information systems.

Data in digital form are valuable company assets to document historical facts, verify business transactions, and improve the decision-making process. Processing data involves every date management data retrieval and transformation processes that take the form of CRUD transactions. Information is the understanding of meaning of data in a manner that reduces uncertainty. Repeating observation of information leads to the creation of knowledge.

Databases as an evolution of file systems have structural characteristics that eliminate most of the problems imposed by the structure of file systems. Databases minimize many of the data inconsistencies existing in file systems and provide a holistic view to all business data. A database is a collection of data stored permanently in a structured form. Information systems collect, store data, and produce information for managers.

In database systems, all corporate data are stored under a single structure and could be made available across all business departments. Applications access the database to perform the four basic CRUD operations. The software system that provides the ability to the users and the applications to access and manipulate the data is called DBMS such as MariaDB, ORACLE, and MS-Access. DBMS are categorized by the required number of users, the database location disparity, and the analytical processing requirements.

Key Terms

Data	Information	Knowledge
File Systems	Column	Record
Field	Legacy Systems	Structural Dependency
Data Redundancy	Data Integrity	Database
Primary Key	Entity	Applications
DBMS	Single-user DBMS	Multiuser DBMS
Central DBMS	Distributed DBMS	ERP
Operational DBMS	Analytical DBMS	Data Mining
Business Intelligence	OLAP	XML

Review Questions

(1) Describe the transformation process from data to information to knowledge. Explain by using an example from data derived in the accounting, sales, or personnel departments and describe what information is derived and how this information will support managerial decision-making.
(2) What are the characteristics that make information valuable to business managers?
(3) Describe the basic CRUD operations in a database.
(4) What are the disadvantages of file systems?
(5) How the database approach solves problems in file systems?
(6) Describe the database environment.
(7) Describe the types of DBMS.
(8) Describe the requirements for selecting a DBMS.

Problems and Exercises

(1) In an ecommerce operation selling tech gadgets, the following information is produced from web analytics regarding the performance of the website against the industry average of 19.832 similar websites.

Sample of 19.832 Websites	Number of Visitors	Average Pages Per Visitor	Conversion Rate (%)*	Average Sale Per Visitor (US$)
Sector average	3,564	3.7	5	20
Website	1724	5.6	15	34

Note: *The percentage of visitors that have made a purchase

Identify the data that must be collected to produce the information presented in the aforementioned table.

(2) The Human Resources department uses a Balanced Score Card[6] application to evaluate staff performance. Currently, the company occupies 218 employees and operates under a cycle of 240 working days per year. For each employee, the system provides 10 performance indicators daily. In order for the manager to evaluate the performance of each employee, he or she must take under consideration the daily performance of each employee in the current working year. How many figures or facts of data must the HR manager collect to perform the evaluation performance?

[6] http://www.bain.com/publications/articles/management-tools-balanced-scorecard.aspx

(3) For the following file structure:

Exercise 1.1. Employee File Structure [employeefile.xlsx].

Employee ID	First Name	Last Name	Address	Telephone	Project Title	Location	Hours Worked
101	Dorothy	Cabels	123 West End, Miami FL	6754326734	Armor	Orlando	30.4
123	Enrico	Gonzales	523 2nd Av. Houston TX	5543897631	Norma	Chicago	24.1
101	Dorothy	Cabels	123 West End, Miami FL	6754326333	Armor	Orlando	7.3
147	Chin	Lee	256 2nd Av., NYC NY	7634539981	Astra	Oakland	18.6
162	George	Kokinakis	335 Pine st. Philadlphia PN	8874429450	Norma	Chicago	31.6
123	Enroco	Gonzales	5423 2nd Av. Houston TX	5543897631	Jena	Denver	20.7
193	Hamad	Allbarhim	753 Est St. Paramus NJ	5548895590	Armor	Denver	23.2

a. Identify the number of columns, of records and fields per record.
b. How many entities can you identify in the employee data structure?
c. What problem you would encounter if you want to produce a listing by state in the Address column?
d. Identify any data redundancy problems.
e. Identify any data inconsistencies.
f. What information will be lost if you must delete the project "Atza Inc."?
g. What changes would you make to the table in Exercise 1.1 to resolve the data redundancy problems you have identified?

(4) F4C is a company offering CRM services over the cloud to clients globally. The core CRM services offered over the *ReaL8* application to clients, which uses *ORACLE* to store and manage client's corporate data, is located in a private cloud in the company's data center in Virginia, USA. F4C also offers internet-marketing services to clients in North America and Europe. The *SEMO* internet-marketing application stores digital marketing data for Europe in a server located in France and a corresponding server for North America data in Maryland, USA. Occasionally, clients require merged data across North America and Europe. Both locations are using *SQL Server* to store and manage data. John Menor, the company's CEO is using *Explode* (a personalized *ORACLE Business Intelligence* Application) in his desktop to explore possible new markets for F4C.

Identify the applications and the types of DBMS used by F4C.

Online Discussion: Visit the web page https://cloud.oracle.com/storage and download the Oracle Cloud Data Storage Service (pdf) file. Summarize in a paragraph the benefits that businesses could raise when they store their corporate databases on the cloud. Would you see any drawbacks to this decision? Follow your opinion in the online discussion.