

# *: What Wikipedia resources are useful for a database management course?*

by Daniel J. Power

Editor, DSSResources.COM

Managing data is a fundamental need in organizations. More complex data, parallel processing and inexpensive hardware has led to major innovations in database technologies and data management.

Relational database management systems (RDBMS) remain the dominant technology and RDBMS are especially important for transaction processing systems. Decision support data sources and databases have changed however and new technologies have created a need for new skills. A contemporary database management and theory course must deal with a broad set of topics. Web resources, primarily Wikipedia, meet a need for authoritative, low cost knowledge about databases and data management.

Wikipedia (<http://www.wikipedia.org/>) is a free encyclopedia. It is often derided as a poor source of information, but in the area of information technology the content is often current and useful. There is some redundancy however and the material lacks a cohesive organization. This discussion attempts to answer two questions: 1) What Wikipedia articles/webpages are most helpful for a contemporary course? 2) In what order should the articles be read and studied? The goal is to increase the usefulness of Wikipedia content for those wanting to learn about contemporary database management and theory. Oracle, a major database vendor, also provides extensive free online content. The webpage <http://docs.oracle.com/> contains links to the most current documentation for Oracle products. I especially recommend "Database Concepts" in [HTML](#) or [PDF](#) format.

Let's start with some basic definitions. Defining terms is the foundation for understanding a complex, challenging topic like database management and theory. First, a database is an organized collection of data. It is more specifically a collection of data that has been selected, organized, and made available for searching and retrieval. A database management system (DBMS) provides convenient, efficient, reliable, and secure multi-user storage of and access to massive amounts of persistent data (Widom, 2011). A general-purpose relational database management system (RDBMS) is a software system designed to allow the definition, creation, querying, update, and administration of structured query language (SQL) databases. A relational database structure consists of data organized into records made up of searchable fields. Data in a relational database is structured into a data record. Each record is made up of several fields.

DAMA International (<http://www.dama.org/>) is an organization of professionals in data management. The DAMA Data Management Body of Knowledge (Mosley, 2007) defines data management as "the development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets." The DAMA framework has 9 content areas: 1) Data Architecture, Analysis and Design, 2) Database Management, 3) Metadata

Page 1/9

(c) 2021 Daniel J. Power, Power Enterprises <[power@dssresources.com](mailto:power@dssresources.com)>

URL: <http://dssresources.com/faq/index.php?action=artikel&cat=&id=270&artlang=en>

## *: What Wikipedia resources are useful for a database management course?*

Management, 4) Data Quality Management, 5) Reference and Master Data Management, 6) Data Security Management, 7) Data Warehousing and Business Intelligence Management, 8) Document, Record and Content Management, and 9) Data Governance. Organizations need more experts in data management to perform many complex tasks.

Why learn about databases (DBs)? Databases are a fundamental information technology. All of us who use the Internet or computer based systems like automated teller machines (ATMs) use databases every day. Databases provide functionality for many data-driven websites; they are a mission critical component of our banking systems, inventory systems, stock trading systems, and even online video games. Any information technology application that needs to store and retrieve data uses a database and data must be managed. So you should learn about DBs because they are used and are useful and DBs are part of many information and decision support capabilities.

A major skill emphasized in an introductory database management course is Structured Query Language (SQL), the most common database definition and data manipulation language. In database concepts classes, the emphasis is on data modeling, normalization, data dictionaries and data integrity. According to <http://degreedirectory.org/>, "some database management classes offer instruction specific to Oracle, the most pervasive database management program." Advanced topics often include: data warehousing, distributed databases, NoSQL databases, and database administration.

Based on experience and examining textbooks, a contemporary, introductory database course for business information systems majors should emphasize database design for the relational data model and Structured Query Language (SQL). Additionally, the course should cover transaction processing, database management, non-relational databases and the use of multi-user, relational database management systems for applications. An introductory database course in a Computer Science department may add or substitute topics like expanded coverage of like [relational algebra](#) and [data structures](#).

In preparing this review, the purpose was to identify Wikipedia resources related to the following instructional objectives:

1. To understand the need for relational databases in organizations and the history and evolution of the development and use of databases.

2. To understand the design and implementation of a relational database management system

Page 2/9

(c) 2021 Daniel J. Power, Power Enterprises <[power@dssresources.com](mailto:power@dssresources.com)>

URL: <http://dssresources.com/faq/index.php?action=artikel&cat=&id=270&artlang=en>

## *: What Wikipedia resources are useful for a database management course?*

(RDBMS).

3. To understand conceptual, logical and physical database designs, database modeling, especially the relational model.

4. To understand and use structured query language (SQL) to create, query, update, and manage a database.

5. To develop a basic understanding of topics like transaction processing, distributed databases, database security, and data warehousing/business intelligence/decision support.

6. To learn skills associated with the fundamental tasks involved in modeling, designing, and implementing an RDBMS.

So what Wikipedia articles/webpages are most helpful? There are many articles and the content is dynamic and sometimes overlapping. Wikipedia is written by a distributed group of people interested in sharing knowledge about a topic. There are many potentially relevant articles. A Wikipedia Book on Databases is online at [en.wikipedia.org/wiki/Book:Databases](https://en.wikipedia.org/wiki/Book:Databases). The contents are divided into 7 categories: 1) Databases, 2) Data Modeling, 3) SQL and Other Query Languages, 4) Technologies, 5) Advances in Relational Databases, 6) Data Warehouse, and 7) Non-Relational Databases. It can be ordered as a printed book with 1611 pages.

For a one semester 3 credit University course a more manageable (smaller) amount of content is desirable and necessary. Based on reading much of the content, I have identified as relevant and accurate the following Wikipedia entries organized into 12 conceptual units. Please send suggestions and comments to [power@dssresources.com](mailto:power@dssresources.com). In total, there are 36 links to Wikipedia entries of various lengths about important, relevant topics. The topics and links are:

Unit 1. Database Management and Theory, an overview of [Databases](#) and [Flat Files](#).

These Wikipedia articles provides a broad overview of many topics including: 1) terminology and database overview, 2) Database applications and roles, 3) Database History, 4) Database research,

*Page 3/9*

## : What Wikipedia resources are useful for a database management course?

5) Database type examples, 6) Database design and modeling, 7) Database languages, and 8) Database operations. A flat file database describes a variety of means to store data as a single file. A database has many advantages compared to a flat file for transaction applications.

Unit 2. Database development -- [Systems Development Life Cycle \(SDLC\)](#) method, [Three schema architecture approach](#), and [Data modeling](#).

This concept unit summarizes the Systems Development Life Cycle (SDLC) approach, a process used by a systems analyst to develop an information system. SDLC is a waterfall method for the analysis and design of information systems. The three-schema approach to building information systems uses three different views in systems development. Data modeling is a process used by an analyst to create increasingly detailed and specific models of the data used in an information system. The analyst applies formal data modeling techniques. A related online resource is the [Structured Analysis Wiki](#) website. It is an update and expansion of Ed Yourdon's 1989 book titled *Modern Structured Analysis*.

Unit 3. Conceptual data modeling -- [Semantic data model](#), [Entity–Relationship \(ER\) model](#), [Weak entity](#), [Associative entity](#) and [Business rule](#).

A semantic data model is an abstraction which defines how the stored symbols relate to the real world. The ER model is a data model for describing a database in an abstract way. Three levels of ER models should be developed: 1) Conceptual data model, 2) Logical data model, and 3) Physical model. A weak entity is an entity that cannot be uniquely identified by its attributes alone. An associative entity can be thought of as both an entity and a relationship. Finally, business rules describe the operations, definitions and constraints that apply to an organization.

Unit 4. More data modeling and analysis -- Enhanced entity–relationship model, [Unified Modeling Language](#), [Class diagram](#), and the [Relational model](#).

The enhanced entity–relationship (EER) model (or extended entity-relationship model) is a high-level or conceptual data model incorporating extensions to the original entity–relationship (ER) model, used in the design of databases. The Unified Modeling Language (UML) is a standardized, general-purpose modeling language in the field of software engineering. A class diagram in the Unified Modeling Language (UML) is static structure diagram. The relational model is the primary means for organizing structured data for systematic retrieval. All data in a domain of interest is systematically grouped into relations.

## : What Wikipedia resources are useful for a database management course?

Unit 5. Logical Design -- [Database Design](#), [Logical data model](#), [Relation](#), [Database normalization](#) and the [Relational database](#).

Database design is the process of producing a detailed data model of a database. A logical data model documents the abstract structure of a domain of information and it is normalized to fourth normal form (4NF). In relational database theory, a relation is a set of tuples  $(d_1, d_2, \dots, d_j)$ , where each element  $d_n$  is a member of  $D_n$ , a data domain. Database normalization is the process of organizing the fields and tables of a relational database to minimize redundancy and dependency. A relational database is a collection of tables of data items. The tables in a relational database are formally described and organized using the relational model.

Unit 6. Physical Design -- [Oracle Database](#), [Data definition language](#), [RAID](#), [Partitioning](#) and [CRUD](#).

The Oracle RDBMS stores data logically in the form of tablespaces and physically in the form of data files. A data definition language or data description language (DDL) provides syntax for defining data structures, especially database schemas. RAID is a storage technology that combines multiple disk drive components into a logical unit. A partition is a division of a logical database or its constituting elements into distinct independent parts. CRUD refers to create, read, update and delete, the four basic functions of persistent, ongoing storage.

Unit 7. [SQL basics](#)  
and [SQL\\*Plus](#)

Structured Query Language is a special-purpose programming language designed for managing data held in a relational database management systems (RDBMS).

Unit 8. [Advanced SQL JOIN](#) and [PL/SQL](#)

A SQL join clause combines records from two or more tables in a database. PL/SQL is Oracle Corporation's procedural extension language for SQL and the Oracle relational database. Functions and Procedures. Search for advanced SQL topics: IN, AND, OR, BETWEEN, LIKE, DISTINCT, GROUP BY, AGGREGATE, CONCAT, SUBSTRING, TRIM, HAVING, ORDER BY, JOIN and OUTER JOIN. Check w3schools.com.

Unit 9. [Data warehouse overview](#), [Star Schema](#) and [Parallel computing](#)

## *: What Wikipedia resources are useful for a database management course?*

A data warehouse or enterprise data warehouse is a database used for reporting and data analysis. It is a central repository of data which is created by integrating data from one or more disparate data sources. The star schema is the simplest style of data mart schema. Parallel computing involves carrying out many calculations/operations simultaneously, dividing large problems and queries into smaller ones solved concurrently.

Unit 10. [NoSQL overview](#), [Cloud database](#) and [In-memory database](#)

A NoSQL database provides a mechanism for storage and retrieval of data that use less structured consistency models than traditional relational databases. A cloud database runs on a cloud computing platform. A cloud DB may be SQL-based or use a NoSQL data model. An in-memory database primarily relies on main memory for computer data storage.

Unit 11. [Database transaction](#), [Concurrency control](#), and [ACID properties](#)

A transaction comprises a unit of work performed within a database management system (or similar system) against a database, and treated in a coherent and reliable way independent of other transactions. Concurrency control ensures that correct results for concurrent operations are generated, while getting those results as quickly as possible. ACID (Atomicity, Consistency, Isolation, Durability) refers to a set of four properties that guarantee that database transactions are processed reliably.

Unit 12. [Database Administration](#), [Database backup](#), [Database security](#) and [Database administrator](#)

In information technology, a backup, or the process of backing up, refers to the copying and archiving of computer data so it may be used to restore the original after a data loss event. RAID (redundant array of independent disks) is a storage technology that combines multiple disk drive components into a logical unit. Database security concerns the use of a broad range of information security controls to protect databases against compromises of their confidentiality, integrity and availability. Finally, a database administrator (DBA) is a person responsible for the installation, configuration, upgrade, administration, monitoring and maintenance of databases in an organization.

In addition to [Microsoft Access](#) and [MySQL](#), it is important to gain some familiarity with [Oracle Databases](#) and [Hadoop](#). The [Oracle 11g Express](#) edition is a useful free learning tool.

## : What Wikipedia resources are useful for a database management course?

There are many additional web resources. For example, Mike Chapple has a Web resource at About.com titled "Databases for Beginners". It includes and is advertised as "A Gentle Introduction to Databases, SQL, and Microsoft Access."

Databases of many types are part of our daily life. Relational databases are used for transaction processing. When you do a Google search you are using a huge database. Google uses Bigtable, a proprietary distributed storage system, for managing structured data that is designed to scale to many petabytes of data across thousands of commodity servers. There are also non-relational, distributed, open-source and horizontally scalable next generation databases.

Why is knowledge about databases important? Knowing how a specific type of database retrieves data improves a person's ability to find information. Also, database design and administration is a group of technical careers that is increasing in need and demand. These high paying jobs require expertise in both database theory and practice and there is a long "apprenticeship" to develop the judgment and skill to manage a high volume, high availability transaction processing database. Database design, management and use is an exciting, challenging topic.

### Summary of Wikipedia links

Unit 1. Database Management and Theory, an overview of [Databases](#) and [Flat Files](#).

Unit 2. Database development -- [Systems Development Life Cycle \(SDLC\)](#) method, [Three schema architecture approach](#), and [Data modeling](#).

Unit 3. Conceptual data modeling -- [Semantic data model](#), [Entity–Relationship \(ER\) model](#), [Weak entity](#), [Associative entity](#) and [Business rule](#).

Unit 4. More data modeling and analysis -- [Unified Modeling Language](#), [Class diagrams](#), and the [Relational model](#).

Unit 5. Logical design -- [Database Design](#), [Logical data model](#), [Relation](#), [Database normalization](#) and the [Relational database](#)

## *: What Wikipedia resources are useful for a database management course?*

Unit 6. Physical Design -- [Data definition language](#), [RAID](#), [Partitioning](#) and [CRUD](#).

Unit 7. [SQL basics](#)  
and [SQL\\*Plus](#)

Unit 8. [Advanced SQL](#) and [PL/SQL](#)

Unit 9. [Data warehouse overview](#), [Star Schema](#) and [Parallel computing](#)

Unit 10. [NoSQL overview](#), [Cloud database](#) and [In-memory database](#)

Unit 11. [Database transaction](#), [Concurrency control](#), and [ACID properties](#)

Unit 12. [Database Administration](#), [Database backup](#), [Database security](#) and [Database administrator](#)

Suggestions for alternative free Wikipedia content and additional chapter topics to include are welcomed. Please email [power@uni.edu](mailto:power@uni.edu).

### References

Haertzen, David, "SQL Advanced Tutorial," Infogoal.com at URL <http://infogoal.com/sql/sql-advanced.htm>, last accessed Friday, June, 07 2013.

Mosley, M. (Ed.) "DAMA-DMBOK Guide™ (Data Management Body of Knowledge)," November 2007 at URL [http://www.dama.org/files/public/DI\\_DAMA\\_DMBOK\\_Guide\\_Presentation\\_2007.pdf](http://www.dama.org/files/public/DI_DAMA_DMBOK_Guide_Presentation_2007.pdf) .

Piotrowski, P. "Oracle Database 11g Express Edition Quick Tour," September 2011 at URL <http://www.oracle.com/technetwork/articles/sql/11g-xe-quicktour-498681.html> .



## *: What Wikipedia resources are useful for a database management course?*

w3schools.com, "SQL Tutorial," URL <http://www.w3schools.com/sql/default.asp>, last accessed July, 01 2013.

Most links are to [Wikipedia](#), the free encyclopedia. Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply.

**Notice**—For any reuse or distribution of Wikipedia content, check the license terms at <http://creativecommons.org/licenses/by-sa/3.0/>

Author: Daniel Power

Last update: 2013-08-13 08:40