***Database Principles* EMEA 3rd Edition**

**Chapter 2 Data Models**

**Short Answer**

1. What are the building blocks of the relational database model?

ANS:

The basic building blocks of all data models are entities, attributes, relationships, and constraints.

An entity is anything (a person, a place, a thing or an event) about which data are to be collected and stored. An entity represents a particular type of object in the real world. Entities may be physical objects, such as customers or products; but entities may also be abstractions, such as flight routes or musical concerts.

An attribute is a characteristic of an entity. For example, a CUSTOMER entity would be described by attributes such as customer last name, customer first name, customer phone, customer address and customer credit limit. Attributes are the equivalent of fields in fi le systems.

A relationship describes an association among entities. For example, a relationship exists between customers and agents that can be described as follows: an agent can serve many customers, and each customer may be served by one agent. Data models use three types of relationships: one-to-many, many-to-many and one-to-one. Database designers usually use the shorthand notations 1:\*, \*:\* and 1:1, respectively.

A constraint is a restriction placed on the data. Constraints are important because they help to ensure data integrity. Constraints are normally expressed in the form of rules;

Rejoinder: 2.2 Data Model Basic Building Blocks

2. Why is the process of identifying and documenting business rules essential to database design?

ANS:

The process of identifying and documenting business rules is essential to database design for several reasons:

They help standardize the company’s view of data.

* They can be a communications tool between users and designers.
* They allow the designer to understand the nature, role and scope of the data.
* They allow the designer to understand business processes.
* They allow the designer to develop appropriate relationship participation rules and constraints and to create an accurate data model.

Rejoinder: 2.3 Business Rules

3. What does a SQL-based relational database application consist of?

ANS:

The end-user interface. Basically, the interface allows the end user to interact with the data (by auto-generating SQL code). Each interface is a product of the software vendor’s idea of meaningful interaction with the data. You can also design your own customized interface with the help of application generators that are now standard in the database software arena.

A collection of tables stored in the database. In a relational database, all data are perceived to be stored in tables. The tables simply ‘present’ the data to the end user in a way that is easy to understand. Each table is independent from another. Rows in different tables are related, based on common values in common attributes.

SQL engine. Largely hidden from the end user, the SQL engine executes all queries or data requests. Keep in mind that the SQL engine is part of the DBMS software. The end user uses SQL to create table structures and to perform data access and table maintenance. The SQL engine translates all of those requests into the instructions necessary to perform such tasks—largely behind the scenes and without the end user’s knowledge.

Rejoinder: 2.4.2 The Relational Model

4. What is the external model?

ANS:

The external model is the end users’ view of the data environment. The term end users refers to people who use the application programs to manipulate the data and generate information. End users usually operate in an environment in which an application has a specific business unit focus. Companies are generally divided into several business units, such as sales, finance and marketing. Each business unit is subject to specific constraints and requirements, and each one uses a data subset of the overall data in the organization. Therefore, end users working within those business units view their data subsets as separate from or external to other units within the organization.

Rejoinder: 2.5 Degrees of Data Abstraction

5. What are some of the common characteristics that data models must have?

ANS:

A data model must show some degree of conceptual simplicity without compromising the semantic completeness of the database. It does not make sense to have a data model that is more difficult to conceptualize than the real world.

A data model must represent the real world as closely as possible. This goal is more easily realized by adding more semantics to the model’s data representation. (Semantics concern the dynamic data behaviour, while data representation constitutes the static aspect of the real-world scenario.)

Representation of the real-world transformations (behaviour) must be in compliance with the consistency and integrity characteristics of any data model.

Rejoinder: 2.4.8 Data Models: A Summary

6. What do business rules require to be effective?

ANSWER: To be effective, business rules must be easy to understand and widely disseminated to ensure that every person in the organization shares a common interpretation of the rules. Business rules describe, in simple language, the main and distinguishing characteristics of the data as viewed by the company.

Rejoinder: 2.3 Business Rules

7. What are the sources of business rules, and what is the database designer’s role with regard to business rules?

ANSWER: The main sources of business rules are company managers, policy makers, department managers, and written documentation such as a company’s procedures, standards, and operations manuals. A faster and more direct source of business rules is direct interviews with end users. Unfortunately, because perceptions differ, end users are sometimes a less reliable source when it comes to specifying business rules. For example, a maintenance department mechanic might believe that any mechanic can initiate a maintenance procedure, when actually only mechanics with inspection authorization can perform such a task. Such a distinction might seem trivial, but it can have major legal consequences. Although end users are crucial contributors to the development of business rules, it pays to verify end-user perceptions. Too often, interviews with several people who perform the same job yield very different perceptions of what the job components are. While such a discovery may point to “management problems,” that general diagnosis does not help the database designer. The database designer’s job is to reconcile such differences and verify the results of the reconciliation to ensure that the business rules are appropriate and accurate.

Rejoinder: 2.3 Business Rules

8. Describe the three basic characteristics of Big Data databases.

ANSWER: Douglas Laney, a data analyst from the Gartner Group, first described the basic characteristics of Big Data databases4: volume, velocity, and variety, or the 3 Vs.

• Volume refers to the amounts of data being stored. With the adoption and growth of the Internet and social media, companies have multiplied the ways to reach customers. Over the years, and with the benefit of technological advances, data for millions of e-transactions were being stored daily on company databases. Furthermore, organizations are using multiple technologies to interact with end users and those technologies are generating mountains of data. This ever-growing volume of data quickly reached petabytes in size, and it’s still growing.

• Velocity refers not only to the speed with which data grows but also to the need to process this data quickly in order to generate information and insight. With the advent of the Internet and social media, business response times have shrunk considerably. Organizations need not only to store large volumes of quickly accumulating data but also need to process such data quickly. The velocity of data growth is also due to the increase in the number of different data streams from which data is being piped to the organization (via the web, e-commerce, Tweets, Facebook posts, emails, sensors, GPS, and so on).

• Variety refers to the fact that the data being collected comes in multiple different data formats. A great portion of these data comes in formats not suitable to be handled by the typical operational databases based on the relational model.

The 3 Vs framework illustrates what companies now know, that the amount of data being collected in their databases has been growing exponentially in size and complexity.

Rejoinder: 2.4.6 Emerging Data Models: Big Data and NoSQL