Database Modeling

by Sophia

WHAT'S COVERED

In this lesson, you will explore why database modeling is important. Specifically, this lesson covers:

1. Why Database Modeling and Normalization Are Important

Database modeling plays a crucial role in database development and management. Here are some ways that it contributes to the database's quality and usability:

- Clarity and organization: Modeling provides a clear visual representation of how data is related and how different entities interact.
- Communication: A model facilitates communication among various stakeholders, including developers, administrators, and business analysts.
- Requirements analysis: Modeling helps team members to identify and clarify data requirements, ensuring that the database design aligns with the organization's needs.
- Relationship identification: Modeling identifies the relationships between different entities and attributes, accurately reflecting real-world relationships.
- Normalization: Modeling helps in achieving normalization, which is critical for reducing redundancy, improving integrity, and minimizing the risk of update anomalies.
- Optimization: A well-designed model can contribute to better performance by optimizing the structure of tables and relationships.
- Efficient querying: A well-designed model enables efficient querying and data retrieval.
- Data integrity and consistency: Modeling helps maintain data integrity, ensuring that the data is accurate and consistent and reducing the risk of errors and inconsistencies.

- Scalability: As the organization's data requirements grow, a well-designed model allows for easier scalability by providing a blueprint for the database structure.
- Documentation: Database models serve as documentation of the database's structure. Having documentation is useful for future maintenance and troubleshooting as well as for introducing the database to those who were not involved in its initial development.
- Security: Security considerations can be incorporated into the model's design, such as authentication mechanisms and encryption strategies.

---

2. What Is Normalization?

Normalization is the process of organizing and structuring data in a relational database efficiently. The main objective of normalization is to eliminate potential anomalies that may arise when data is stored in a denormalized or redundant manner and minimize data redundancy. Normal forms ensure standardized and logical data storage by dividing a database into multiple tables. The process involves dividing a database into multiple tables and applying a set of rules.

BIG IDEA

Normal forms are sets of rules applied to a database design. The three most basic and most often used are: **first normal form (1NF)**, **second normal form (2NF)**, and **third normal form (3NF)**.

These normal forms build upon one another.

- 1NF defines rules for creating a single table, such as unique column names and lack of duplicate rows.
- 2NF requires all of 1NF's qualities plus additional qualities pertaining to relationships, such as ensuring that all non-prime attributes (that is, attributes that are not the table's primary key) depend on the primary key.
- 3NF requires all of 2NF's qualities plus its additional qualities, such as elimination of any transitive dependencies. A transitive dependency exists when one non-prime attribute depends on another non-prime attribute.

Data consistency and accuracy can be achieved through normalization since the data is arranged and organized in a logical manner. As a result, data retrieval is more efficient, and there are fewer chances of data anomalies or inconsistencies.

As a result of normalization, queries involving multiple tables and joins may become more complex, resulting in slower query execution. It is necessary to balance the trade-offs between data integrity and query performance when choosing a level of normalization for a database.

TERMS TO KNOW

Normalization

Normalization is the process of organizing and structuring data in a relational database efficiently.

First Normal Form (1NF)

A relational database system's first normal form (1NF) defines the fundamental rules for normalizing a single table.

Second Normal Form (2NF)

All of the data and non-prime attributes are functionally dependent on each candidate key.

Third Normal Form (3NF)

In a 3NF-compliant table, no non-primary key attribute is transitively dependent on the primary key. It preserves lossless database transitions and removes functional dependencies.

SUMMARY

In this lesson, you learned that in a relational database, database modeling and normalization ensure data is organized, structured, and stored efficiently. In database modeling, entities, attributes, and relationships between data elements are conceptually represented. You learned that **database modeling is important** because it provides a blueprint for building the database system and helps to clarify the data requirements. A database model ensures data integrity, consistency, and accuracy by capturing the entities and their relationships. As a result, it is easier to manage, access, and retrieve data. In this process, entities and their attributes are visually depicted using entity-relationship diagrams (ERDs).

You also learned that database **normalization** eliminates anomalies and reduces data redundancy. Databases are divided into multiple tables, standardized, and logically stored according to the rules outlined in the normal forms. The three most common normal forms are the first normal form (1NF), the second normal form (2NF), and the third normal form (3NF). Normal forms are designed to ensure that data is properly organized and related, avoiding repeating groups, partial dependencies, and transitive dependencies. Data normalization improves its consistency and integrity and minimizes potential data anomalies such as duplication and update anomalies. It is important to normalize data in order to ensure data quality, scalability, and ease of maintenance of the database system.

Source: THIS TUTORIAL WAS AUTHORED BY DR. VINCENT TRAN, PHD (2020) AND Faithe Wempen (2024) FOR SOPHIA LEARNING. PLEASE SEE OUR TERMS OF USE.

TERMS TO KNOW

First Normal Form (1NF)

A relational database system's first normal form (1NF) defines the fundamental rules for normalizing a single table.

Normalization

Normalization is the process of organizing and structuring data in a relational database efficiently.

Second Normal Form (2NF)

All of the data and non-prime attributes are functionally dependent on each candidate key.

Third Normal Form (3NF)

In a 3NF-compliant table, no non-primary key attribute is transitively dependent on the primary key. It preserves lossless database transitions and removes functional dependencies.