Meta Database Engineer Professional Certificate

Course Notes

JESUS CARDENAZ

1   Introduction to Databases

Database   A database is a form of electronic storage in which data is organized systematically.

Relational Databases   Relational dataabases are based on the relational model, which organizes data into tables with rows and columns. The establish relationships using keys and enforce integrity constraints. Relational databases adhere to a predefined schema and use SQL as the primary language for data manipulation and querying.

Non-Relational Databases   Non-relational databases, also known as NoSQL databases, encompass various database technologies that deviate from the relational model. These databases typically prioritize scalability, performance and flexibility over strict adherence to a predefined schema.

SQL vs NoSQL databases   The main difference between SQl and NoSQL databases lies in their data model, query language, schema flexibility, and scalability characteristics. While SQL databases usually adhere to ACID (Atomicity, Consistency, Isolation, Durability) compliance. NoSQL uses the BASE (Basically Available, Soft State, Enventual Consistency) properties.

Most common types of databases:

- Relational Databases: Is the most common type of database, they store data in tables with rows and columns.

- Object oriented databases: Instead of breaking down data into tables, these databases store data in objects and classes.

- Graph databases: Graph databases are designed to store and manage databased on graph theory concepts. Data is represented as nodes, which are entities, and edges which represent relationships between the nodes.

- Document databases: Store and manage and organize data into self-contained documents, typically in formats like JSON or XML. Each document can have a different structure and schema, allowing for flexibility in data modeling.

Database Schema   A database schema is the structure of the database. The simplest way of understanding database schema is to think of it as the blueprint of a database. Database schema can be divided into three categories:

- Conceptual or logical schema: Defines entities, attributes and relationships. It is often represented using Entity Relational Diagrams

- Internal or Physical schema: Defines how the data is actually stored and organized within the physical storage devices.

- External or View Schema: Represents different user views or subsets of the data from the overall database.

Database Normalization   Is the process in database design that helps ensure data integrity, eliminate redundancy, and improve data consistency by organizing data into well-structured tables. There are three fundamentan normalization forms known as:

- First Normal form (1NF): It requires that the data is organized into tables, and each column of a table contains atomic values (indivisible values). Additionally, it eliminates redundant data by ensuring that each attribute within a table contains only one value.

- Second Normal form (2NF): It states that no non-key attributes should be dependent on only a portion of the primary key. In other words, any attribute that depends on only a subset of the primary key should be moved to a separate table.

- Third Normal form (3NF): States that non non-key attribute should depend on another non-key attribute within the same table.