screencasts

Relational databases

Relational databases zijn om veel redenen de meest populaire manier om informatie te beheren

* ze slaan data efficiënt op
  + hierdoor kan je snel informatie ophalen
* ze zijn heel flexibel

Als je bijvoorbeeld foto's wilt opslaan in een database kan het soms ingewikkeld zijn hoe je deze gaat sorteren. Met een relationele database kan je makkelijk verschillende tabellen maken en deze met elkaar linken als je bv je fotos op locatie, foto en personen wil opslagen.

Om gegevens te maken en op te slaan in een relationele database moet je een RDBMS (Relational Database Management System)

veel gebruikte systemen

* Microsoft SQL Server en Azure SQL Database
* Oracle, PostgreSQL, IBM Db2
* MySQL is zeer populair op het web
  + Elke website dat het WordPress-publicatie platform gebruikt een MySQL database om blogposts en opmerkingen te beheren
* Microsoft Acces op PC, Filemaker op macOS

⇒ Het kiezen van systeem hangt af van wat je moet en de schaal waarop je het moet doen

RDBMS taken

* Taak 1
  + Create and modify the structure of the data
  + Define tables and column names
  + Create key-value columns and create relationships
* Taak 2
  + Manipulate data records and perform CRUD tasks
    - CRUD ⇒ Create, Read, Update en Delete

Maken van een relationele database

* Identify the facts it needs to store
* Think about what you want to get out the database
* Identify the analyses or reports that would be beneficial to have at the end of the process

Voordelen van data types

* Storage efficiency
* Better performance predictions
* Enforces data consistency and improves quality

Real world primary keys

* Library and bank account numbers
* Email address
* Social security number and driver’s license
* product number
* serial number

Business rules

* explicit or implicit rules that govern an organization’s data
* based on written policy or best practices that your organization follows

Null value

* Null represents data that is unspecified, unknown or not applicable
* Null does not mean zero, or a blank space character
* literally nothing or no data

Creating indexes

* Added to any column used in frequent searches or joins
* primary keys are created as clustered index
* Additional indexes are non-clustered

Adding too many indexes

* Slows down writes to the database
* Every edit causes all of the indexes to be rebuilt
* Searches still possible on non-indexed fields
* efficiencies depend on the size of the data tables

Numerical checks

* can include a range of acceptable values
* productPrice > 0
* JobRating >= 1 AND JobRating <= 5

Character checks

* can limit acceptable values to an approved list
* status = “In stock”
  + or status = “back ordered”
  + or status = “discontinued”

Creating Relationships

* Relationships are generally created on the foreign key
* Data types of the FK and PK columns must be the same
* FK and PK columns can have different names

Optionality

* The minimum number of associated records
* Usually 0 or 1
* If a customer MUST have an order: optionality = 1
* If a customer MIGHT have an order: optionality = 0

Cardinality

* The maximum number of associated records
* Usually 1 or many (n)
* If each product can have only one supplier: Cardinality = 1
* If products can have multiple suppliers: Cardinality = n

Optionality + cardinality = multiplicity

One-to-many relationships

* The most common relationship between two tables
* A library card number checks out many books
* A bank account has several transactions

Self joins

* Also known as self-referencing relationship or recursive relationship
* Follow the rules as relationships between two tables
* Can be one-to-one, one-to-many, or many-to-many
* Unique constraints and cardinality still apply

Implementing Cascade Changes

* Cascade update and delete don’t with new data
* Leaving the feature off can protect data integrity as well
* Look for check boxes in the GUI to enable
* Can also be enabled using SQL commands
* ON UPDATE CASCADE and ON DELETE CASCADE

Database Normalization

* A series of rules that describe proper database design
* Rules for table structure are called “normal forms”
* Normal forms are sequentially satisfied
* A good database design will satisfy
* A good database design will satisfy third normal form (3NF)

Afbeelding met tekst, Lettertype, wit, algebra

Automatisch gegenereerde beschrijving

Afbeelding met tekst, Lettertype, algebra

Automatisch gegenereerde beschrijving

Afbeelding met tekst, Lettertype, wit, algebra

Automatisch gegenereerde beschrijving

Afbeelding met tekst, schermopname, Lettertype, nummer

Automatisch gegenereerde beschrijving

The normalization process removes redundant information from the database

Denormalization deliberately introduces redundancy in order to increase performance

Graph databases

* Store information as nodes and edges
* Each node can have relationships with any other node
* Nodes can represent different kinds of information
* Commonly used for modeling social networks

Afbeelding met schermopname, tekst, diagram, lijn

Automatisch gegenereerde beschrijving

Document database

* Store documents that represent a single object
* Support files in a variety of formats
* Documents are read, categorized, grouped, and tagged
* Information about each object is contained in the document
* Documents can differ in their content
* Structure of the database is not established ahead of time

Afbeelding met cirkel, schermopname, logo, diagram

Automatisch gegenereerde beschrijving

NoSQL database (not relational)

Quiz

* A one-dimensional array is more commonly known as what?
  + List
* SQL Server, Access, Oracle, and PostgreSQL are all examples of what?
  + RDBMS
* In a relational database, what do you call values that are used to look up additional information in other tables?
  + Keys
* In Codd's paper on relational databases, he used the word "domain" to refer to what we commonly call a \_\_\_\_\_.
  + Column
* A column that stores unique values that are already present in the data is called what?
  + Natural key
* Which is NOT a reason the database needs data types assigned to columns?
  + so the DBMS can ensure data is entered completely
* What are the names of database commands and functions that should be avoided in the naming of your database tables and columns?
  + reserved words
* Which icon is typically used to indicate a database?
  + cylinder cut in thirds
* Which relationship notation type draws a slash to indicate 'one' and a branch to indicate 'many'?
  + Crows foot
* A primary key is also a \_\_\_\_\_ constraint.
  + Unique
* Which SQL command is used to add a record to a data table?
  + INSERT INTO
* Which column constraint will supply a value when one isn't provided when inserting a new record?
  + Default
* If you want a value in a specific column to be required when adding records to a table, the column must be created with what constraint?
  + Not Null
* You have a table that stores warehouse inventory and you need to ensure that the number of items in stock never falls below zero. What type of constraint can you use on the table?
  + Check
* Indexes on columns that are not the primary key are created as what kind of index?
  + Non-clustered index
* When creating a relationship from an INT data type primary key column named ProductID, the foreign key needs to be what?
  + an INT data type
* When creating a database relationship, the values in a primary key field in one table will match up with values in a \_\_\_\_\_ key field in a second table.
  + Foreign key
* For security purposes, you need to store information about the credit card used to pay for an invoice in a separate table from the rest of the invoice's details. What type of join will most likely be involved?
  + one-to-one
* Which data constraint determines the type of join that is created between two tables?
  + Unique
* What is it called when a column has a relationship to another column in the same table?
  + self-join
* Many-to-many relationships require the use of a \_\_\_\_\_ table and two one-to-many joins.
  + Linking
* You delete a record in a parent table and want all of the records in the related child table to also get deleted. What feature do you need to have enabled?
  + Cascade Delete
* Which level of database normalization is generally acceptable for most databases?
  + 3NF
* What is the primary reason why a database might be left in an un-normalized state?
  + to increase performance of queries
* What must you be sure to include in an UPDATE statement to ensure that changes are only applied to the intended records?
  + a WHERE clause
* Which SQL command would you use to add an additional column to an already existing table?
  + ALTER TABLE
* Most relational database management systems support saved SELECT queries known as what?
  + Views
* What is the primary method of returning information from a relational database?
  + SELECT statement
* The term 'NoSQL' describes a database that \_\_\_\_\_.
  + does not follow the relational model
* Graph databases store elements as either \_\_\_\_\_ or \_\_\_\_\_.
  + Nodes and EdgesAfbeelding met tekst, schermopname, Lettertype, ontwerp

    Automatisch gegenereerde beschrijving