

Relační databázový model

- Poprvé byl zveřejněn v roce 1971 v článku pracovníka firmy IBM, matematika Dr. Codd.
- Dr. Codd uvedl matematické definice pojmů: entita a její typ, vazba a její typ, atribut a jeho hodnoty, struktura databáze.
- Definoval jazyk, pomocí něhož se vyhledávají informace a manipuluje s daty v databázi.
- Definoval pravidla pro rozpoznání správné struktury databáze.

- Data jsou v relačních databázích uložena v samostatných tabulkách podle předmětu obsahu, ale tabulky jsou navzájem relačně propojeny. Tabulka zákazníků je například propojena s tabulkou objednávek prostřednictvím pole s číslem zákazníka v každé z těchto tabulek.
- **Efektivita:** Není třeba ukládat nadbytečné údaje (například jméno nebo adresu zákazníka) v každé objednávce.
- **Správa:** Ve správně strukturované databázi, která neobsahuje duplicitní informace, lze data snadno aktualizovat, odstranit a rozšířit.

- Entita – je objekt reálného světa, který je schopen nezávislé existence, je jednoznačně odlišitelný od jiných objektů.

Konkrétní entita – Výrobky, Studenti, Objednávka (názvy tabulek)

- Atribut – je skutečnost, kterou o dané entitě evidujeme, určuje některou podstatnou vlastnost entity (sloupec v tabulce).
- Doména - je obor hodnot, tvoří množinu všech přípustných platných hodnot, které smí atribut obsahovat.
- Atomicita atributu – atribut je atomický, pokud jej již nelze rozložit na dílčí atributy
- Vztahy – vazby definované mezi entitami
- Pravidla pro pojmenování : Entita – podstatná jména (Učitel), Vztahy slovesy – (učí)

- **Tabulky** slouží k uložení dat v řádcích (záznamy) a sloupcích (pole). Všechny databáze obsahují nejméně jednu tabulku
- Redundance dat - nadbytečnost, opakování
- Integrita databáze – vznik chyb při aktualizaci, mazání a doplňování dat
- Záznam, pole, položka, struktura tabulky, typ pole
- Primární klíč – jednoznačně označuje a odlišuje jednotlivé záznamy, musí splňovat integritní podmínky (hodnoty v tomto sloupci se nesmí opakovat a v žádném záznamu tabulky nesmí hodnota prim. klíče chybět)

Podmínky, které musí splňovat relační tabulka

- Všechny hodnoty v tabulce musí být elementární - tj. dále nedělitelné.
- Sloupce mohou být v libovolném pořadí.
- Řádky mohou být v libovolném pořadí.
- Sloupce musí být homogenní, tzn. ve sloupci musí být údaje stejného typu (z jedné domény).
- Každému sloupci musí být přiřazeno jednoznačné jméno.
- V relační tabulce nesmí být dva zcela stejné řádky. Tzn., že každý řádek je jednoznačně rozlišitelný.

- **Rozdělení informací do tabulek.**

Dodržujeme dvě základní zásady rozdělení do tabulek:

- Udržíme data vždy jen na jednom místě, bráníme vzniku duplicitních (redundantních) dat, která zabírají místo a mohou být příčinou chyb.
- Data vkládáme správná a úplná.

- **Zadáme primární klíče.**

Primární klíč jednoznačně označuje a odlišuje jednotlivé záznamy, musí splňovat integritní podmínky (hodnoty v tomto sloupci se nesmí opakovat a v žádném záznamu tabulky nesmí hodnota prim. klíče chybět).

kandidátní klíče - všechny ty, které splňují podmínky primárního klíče
složený klíč – klíč tvořen více poli

- **Ověříme splnění normalizačních pravidel.**

Normalizace je proces zjednodušování a optimalizace struktury tabulek v databázi především za účelem minimalizace redundantních (nadbytečných) dat. Správnost návrhu se vyhodnocuje tzv. normálními formami.

Normální formy

První normální forma (1NF)

- Všechny sloupce nelze dále dělit na části nesoucí nějakou informaci.
- V poli nesmí být data, která lze rozdělit do více polí.

Druhá normální forma (2NF)

- Tabulka obsahuje pouze sloupce, které jsou závislé na celém klíči.

Třetí normální forma (3NF)

- Nesmí existovat závislost mezi neklíčovými poli tabulky.

Čtvrtá normální forma (4NF)

- Tabulka je ve čtvrté normální formě, pokud sloupce (atributy) v ní obsažené popisují pouze jeden fakt nebo jednu souvislost.

Pátá normální forma (5NF)

- Tabulka je v páté normální formě, pokud je ve čtvrté a není možné do ní přidat nový sloupec (skupinu sloupců) tak, aby se vlivem skrytých závislostí rozpadla na několik dílčích tabulek.

Integrita dat a integritní omezení

- Jedním z nástrojů usnadňujících dodržení konzistence dat jsou **integritní omezení**
 - Ta umožňují dodržení správnosti dat z hlediska omezení, která existují v reálném světě
- Rozlišujeme 3 typy integritních omezení:
 - **Entitní integrita** – jednoznačné určení řádku v rámci tabulky (viz vlastnosti primárního klíče)
 - **Doménová integrita** – zajišťuje, aby každá hodnota atributu byla vybrána z množiny přípustných hodnot (částečně řeší zvolený datový typ)
 - **Referenční integrita** (= konzistence vazeb mezi záznamy tabulek) – je-li v tabulce sloupec s hodnotami primárního klíče jiné tabulky (s tzv. **cizím klíčem**), musí tyto hodnoty buď odpovídat některé z hodnot PK „domovské“ tabulky nebo mít hodnotu NULL

Relace – vzájemné vztahy mezi tabulkami

Typy relací:

1. Tabulky nejsou v relaci
2. Mezi tabulkami je relace jedna k více 1:N – jeden údaj v tabulce odpovídá několika údajům v druhé tabulce.
Nejpoužívanější relace (tabulka Zákazníků, tabulka Zboží – jeden zákazník si může koupit více výrobků)
3. Mezi tabulkami je relace 1:1 – jednomu záznamu v jedné tabulce odpovídá přesně jeden záznam v druhé tabulce, příliš se nepoužívá, někdy pro přehlednost dat, jinak je možné tabulky sloučit
4. Tabulky jsou v relaci více k více M:N – více záznamů v jedné tabulce odpovídá více záznamům v druhé, nelze ji vytvořit přímo, musíme vytvořit spojovací tabulku (jeden autor může napsat více knih a jedna kniha může být napsána více autory)

Primární klíč – jednoznačně označuje a odlišuje jednotlivé záznamy, musí splňovat integritní podmínky (hodnoty v tomto sloupci se nesmí opakovat a v žádném záznamu tabulky nesmí hodnota prim. klíče chybět)

kandidátní klíče - všechny ty, které splňují podmínky primárního klíče

složený klíč – klíč tvořen více poli

cizí klíč (primární klíč z první tabulky se ve druhé tabulce objeví jako pomocný sloupec a nazýváme jej cizí klíč)

Datové typy cizího klíče a souvisejícího primárního klíče musí být stejné. Hodnoty cizího klíče se mohou opakovat, nemusí být ve všech položkách zadane hodnoty a nesmí obsahovat hodnotu, která by nebyla obsažena ve sloupci odpovídajícího primárního klíče připojené tabulky (referenční integrita).

Převod vztahu N : N na dvě vazby

