

Data Analyse

Normalisatie

Problemen tabellen:

- Aanpassingen → gegevensinconsistentie
- Toevoegen → niet handig
- Verwijderen

Normalisatiestappen van Codd: Techniek om deze problemen te voorkomen

- Verzameling van gegevens in groepen verdeeld
- Geen problemen meer bij onderhoud van gegevens
- Geen functionele afhankelijkheden tussen attributen

Record (tuple): Rij van data in een database tabel bestaande uit één waarde uit elke kolom

Kandidaatsleutel: Een combinatie van attributen uit de tabel, waarvoor de attribuut-waarden op één unieke wijze een tuple van alle andere tuples uit de tabel onderscheidt

→ Om records te identificeren zodat ernaar gerefereerd kan worden of zodat ze benaderd kan worden

Vb. MedewerkerNummer 001 → precies 1 werknemer (record) in de werknemerstabel

Voorwaarden kandidaatsleutel:

- 1 attribuut of combinatie van meerdere attributen
- Moet uniek zijn
- Mag geen overbodige attributen bevatten
(Weglaten attribuut uit sleutel → eigenschap van uniek identificeren moet verloren gaan)
- Voor elke tabel één kandidaatsleutel (soms meerdere mogelijk → 1 sleutel = primaire)

Primaire sleutel: Kandidaat-sleutel waarin geen van de elementen een null-waarde kan bevatten

→ Elke tabel kan maar 1 primaire sleutel hebben

→ Kan een combinatie zijn van verschillende attributen (kolommen) → **Samengestelde key**

Externe sleutel (foreign key): Veld in een tabel die met een kandidaat-sleutel van een andere tabel overeenkomt

→ Kan gebruikt worden om te refereren naar records in andere tabellen

1. Eerste normaalvorm (1NV)**1. Inventariseer alle elementaire gegevens**

Vb. Kerkstraat 23 → opdelen in straat en huisnummer

2. Verwijder alle procesgegevens

→ Gegevens kunnen soms uit andere gegevens worden afgeleid

3. Doe het volgende tot er geen nieuwe groepen meer ontstaan

- a. Geef de sleutel van de groepen aan
 - Kandidaatsleutel
 - Primaire sleutel
 - Externe sleutel
- b. Geef de deelverzameling die herhaaldelijk voorkomt aan
- c. Herhaal de sleutelgegevens van de oorspronkelijke groep, samen met de gegevens van de herhaalde deelverzameling als nieuwe groep
- d. Verwijder de herhaalde deelverzameling uit de oorspronkelijke groep

2. Tweede normaalvorm (2NV)

→ Verwijder de attributen die functioneel onafhankelijk zijn van een deel van de sleutel

1. Selecteer (onafhankelijk sleutel)

Selecteer de attributen die niet functioneel afhankelijk zijn van de volledige sleutel

2. Groepeer

Vorm een aparte groep voor ieder deel van de sleutel waarvan attributen functioneel afhankelijk zijn

3. Sleutel

Neem in elke groep de attributen met het bijbehorende sleuteldeel op

4. Verwijder

Verwijder deze attributen uit de oorspronkelijke groep

3. Derde normaalvorm (3NV)

→ Verwijder attributen die functioneel onafhankelijk zijn van andere (niet-sleutel) attributen

1. Selecteer (onafhankelijk andere attributen)

Selecteer de attributen die ook functioneel onafhankelijk zijn van andere attributen

2. Groepeer

Vorm een aparte groep voor ieder attribuut of combinatie van attributen, waar andere attributen functioneel van afhankelijk zijn.

3. Sleutel

Neem in iedere nieuwe groep de attributen met hun bijbehorende sleutel op.

4. Verwijder

Verwijder de attributen van de nieuwe groep(en) uit de oorspronkelijke groep.

TIPS
Eerste normaalvorm – NV1
Je kan ook 2 aparte tabellen afzonderen o.w.v. herhaling
Voeg ID toe waar nodig
Primary key zal heel vaak combinatie van attributen zijn!
Probeer de key zo klein mogelijk te maken, maar wel uniek
Tweede normaalvorm – NV2
Wanneer je tabellen splitst, nooit de primary key veranderen!
Tabellen met enkelvoudige sleutel gewoon overnemen
Derde normaalvorm – NV3
Er moet altijd een link tussen tabellen zijn
Integratie
8 tabellen → minstens 7 foreign keys (n-1)
Er kunnen meerdere foreign keys per tabel zijn
Hier pas namen geven aan tabellen
Samengestelde key kan ook foreign key zijn (volledig)