



Data Analysis & SQL Normalisatie



**DE HOGESCHOOL
MET HET NETWERK**

Elfde-Liniestraat 24, 3500 Hasselt, www.pxl.be



Wat is normaliseren?

NV1

NV2

NV3

Verdere normaalvormen



Wat is normaliseren?

NV1

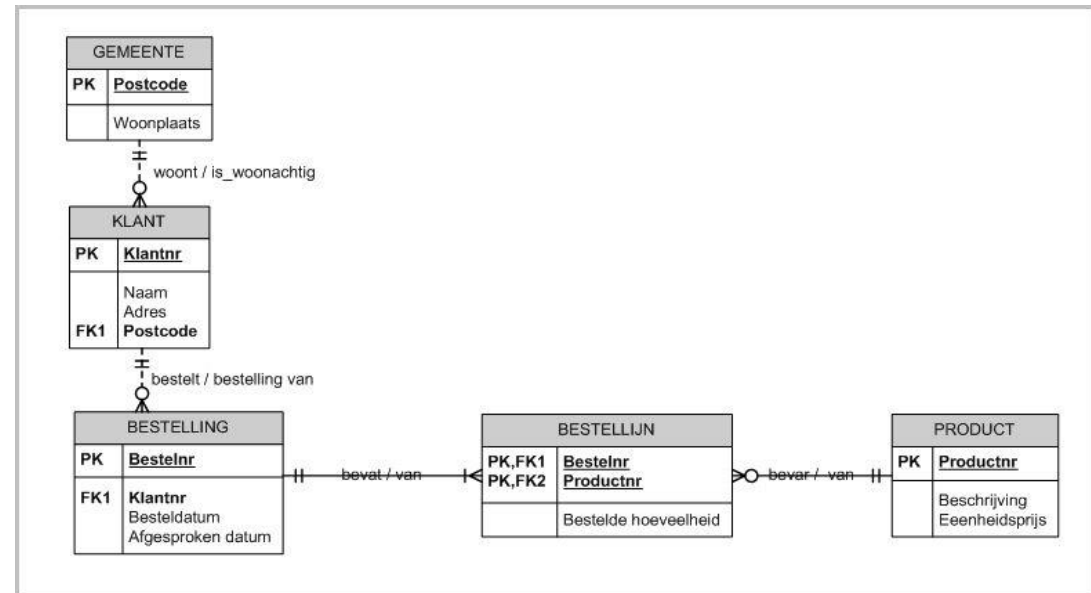
NV2

NV3

Verdere normaalvormen

Normaliseren

- Wat?
Kortweg: structuur brengen in een DB
- Hoe?
Gebruik maken van de normalisatiestappen van CODD



Normaliseren

- Informatie verdelen in een aantal groepen
- Iedere groep voldoet aan onderstaande definitie

Alle attributen van de groep zijn functioneel afhankelijk van de volledige sleutel en tussen attributen onderling zijn geen functionele afhankelijkheden aanwezig.

- Vb. Naam, adres en postcode is **volledig** afhankelijk van klantnr

KLANT	
PK	<u>Klantnr</u>
	Naam Adres Postcode
FK1	

Normaliseren

- Een genormaliseerde groep bevat een vast aantal gegevens en kan dus als een tabel worden voorgesteld

PK	Medewerkers			
	<i>Nummer</i>	<i>Naam</i>	<i>Woonplaats</i>	<i>Afdeling</i>
	001	Els	Antwerpen	Opleidingen
	002	Pieter	Leuven	Opleidingen
	003	Patricia	Tongeren	Analyse
	004	Rudi	Genk	Analyse
	005	Gert	Hasselt	Programmering

- Nummer* is de sleutel
- Alle attributen zijn afhankelijk van dit nummer
- Geen functionele afhankelijkheden tussen deze attributen
- Medewerker(Nummer, Naam, Woonplaats, Afdeling)

Normaliseren

- Normalisatie vindt plaats in drie stappen:

1. Verwijder de zich herhalende deelverzameling
2. Verwijder de attributen die functioneel afhankelijk zijn van slechts een deel van de sleutel
3. Verwijder de attributen die ook functioneel afhankelijk zijn van ander (niet-sleutel) attributen – transitieve afhankelijkheden

- Opmerking **verwijderen**: er mag niet echt iets verwijderd worden!
 - Verwijderen uit de oorspronkelijke groep
 - Tegelijkertijd creëren van een nieuwe groep

Normaliseren

- Voorbeeld

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Normaliseren

- Niet genormaliseerde gegevensverzameling:
 - Projectnummer
 - Projectomschrijving
 - Budget
 - Medewerkersnummer
 - Naam
 - Afdeling
 - Chef
 - Uren

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV



Wat is normaliseren?

NV1

NV2

NV3

Verdere normaalvormen

NV1: Eerste normaalvorm

Stap 1: Verwijder de zich herhalende deelverzameling

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

- 1.1 Inventariseer alle **elementaire** gegevens
- 1.2 Verwijder alle **procesgegevens**
- 1.3 Doe het volgende totdat er geen **nieuwe groepen** meer ontstaan
 - 1.4 Geef de **sleutel** van de groep aan
 - 1.5 Geef de deelverzameling aan die een **herhaald** aantal keren voorkomt
 - 1.6 Herhaal de sleutelgegevens van de oorspronkelijke groep samen met de gegevens van de zich herhalende deelverzameling als een nieuwe groep
 - 1.7 **Verwijder** de zich herhalende deelverzameling uit de oorspronkelijk groep
- 1.8 Eind-doe

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.1:

- Alleen elementaire gegevens inventariseren
- Samengestelde gegevens → elementaire gegevens
- Benoemen van een elementair gegeven (naamgeving)

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.2:

- Procesgegevens niet opgenomen in normalisatiestappen (apart noteren)
- Procesgegeven:
 - Alle voor de berekening benodigde gegevens zijn aanwezig
 - De voor de berekening benodigde gegevens bevatten op het tijdstip waarop de berekening worden uitgevoerd nog de juiste waarden

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.3

- Zich herhalende deelverzamelingen kunnen genest voorkomen
- 1NV zover doorvoeren totdat alle zich herhalende deelverzamelingen zijn opgedeeld.

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.4

- Kies de meest uitgebreide sleutel
- Na inventarisatie van de elementaire gegevens:
 - Projectnummer
 - Projectomschrijving
 - Budget
 - Medewerkersnummer
 - Naam
 - Afdeling
 - Chef
 - Uren

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.5

- Afhankelijk van de sleutelkeuze in stap 1.4 zullen nul, één of meer zich herhalende deelverzameling aanwezig zijn
- Na inventarisatie van de elementaire gegevens:
 - Projectnummer
 - Projectomschrijving
 - Budget
 - Medewerkersnummer
 - Naam
 - Afdeling
 - Chef
 - Uren

komt een herhaald aantal keren voor

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.6

- Nu wordt een nieuwe groep gevormd. Deze moet bestaan uit:
de zich **herhalende** deelverzameling + de **sleutel** van de oorspronkelijke groep
- Deze laatste opnemen om koppeling met oorspronkelijk groep te behouden.

PROJECTOVERZICHT:

Projectnummer
 Projectomschrijving
 Budget
 Medewerkersnummer
 Naam
 Afdeling
 Chef
 Uren

} **Herhalend**

PROJECTMEDEWERKER

Projectnummer
 Medewerkersnummer
 Naam
 Afdeling
 Chef
 Uren

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.7

- Oorspronkelijke groep aanpassen → verwijdering

PROJECTOVERZICHT:

- Projectnummer
- Projectomschrijving
- Budget
- ~~Medewerkersnummer~~
- ~~Naam~~
- Afdeling
- ~~Chef~~
- ~~Uren~~

PROJECTMEDEWERKER

- Projectnummer**
- Medewerkersnummer
- Naam
- Afdeling
- Chef
- Uren

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.8

- Deze stap, het einde van de DOE-loop, zorgt ervoor dat we weer bij stap 1.4 beginnen voor de nieuw gevormde groep

NV1: Eerste normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 1.4 (2^{de} doorgang)

- Samengestelde sleutel projectnummer en medewerkersnummer
- Indien enkel medewerkersnummer als sleutel → nog steeds herhaling van Projectnummer en Uren

1 NV

PROJECTOVERZICHT	PROJECTMEDEWERKER
<u>Projectnummer</u>	<u>Projectnummer</u>
Projectomschrijving	<u>Medewerkernummer</u>
Budget	Naam
	Afdeling
	Chef
	Uren



Wat is normaliseren?

NV1

NV2

NV3

Verdere normaalvormen

NV2: Tweede normaalvorm

Stap 2: Verwijder de attributen die functioneel afhankelijk zijn van slechts een gedeelte van de sleutel

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Alleen groepen met een samengestelde sleutel komen in aanmerking!

- 2.1 Geef de attributen aan die niet functioneel afhankelijk zijn van de **volledige** sleutel
- 2.2 Vorm een aparte groep voor ieder deel van de sleutel waarvan attributen functioneel afhankelijk zijn
- 2.3 Neem in iedere groep de attributen met het bijhorend sleuteldeel op
- 2.4 Verwijder deze attributen uit de oorspronkelijke groep

NV2: Tweede normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 2.1

- Vraagstelling per attribuut:
Welk gegeven of combinatie van gegevens identificeert dit attribuut op een eenduidige wijze?

Attribuut **Naam** is functioneel afhankelijk van sleutel **Medewerkersnummer**

Attribuut **Afdeling** is functioneel afhankelijk van sleutel **Medewerkersnummer**

Attribuut **Chef** is functioneel afhankelijk van sleutel **Medewerkersnummer**

Attribuut **Uren** is functioneel afhankelijk van sleutel **Medewerkersnummer + Projectnummer**

NV2: Tweede normaalvorm

Stap 2.2

- Er ontstaat 1 nieuwe groep:
MEDEWERKER

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

NV2: Tweede normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 2.3

- Geïntentariseerde groepen moeten gevuld worden
- Iedere nieuwe groep moet voldoen aan de eisen van de tweede normaalvorm
 - Er mogen dus géén herhalingen aanwezig zijn
 - En alle attributen moeten functioneel afhankelijk zijn van de **volledige sleutel**

PROJECTMEDEWERKER

Projectnummer

Medewerkersnummer

Naam

Afdeling

Chef

Uren

MEDEWERKER

Medewerkersnummer

Naam

Afdeling

Chef

NV2: Tweede normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 2.4

- Verwijdering van de attributen uit de oorspronkelijk groep
- **De sleutel van de oorspronkelijk groep mag niet aangetast worden!**

2 NV

PROJECT	PROJECTMEDEWERKER	MEDEWERKER
<u>Projectnummer</u>	<u>Projectnummer</u>	<u>Medewerkersnummer</u>
Projectomschrijving	<u>Medewerkernummer</u>	Naam
Budget	Uren	Afdeling
		Chef



Wat is normaliseren?

NV1

NV2

NV3

Verdere normaalvormen

NV3: Derde normaalvorm

Stap 3: Verwijder attributen die ook functioneel afhankelijk zijn van ander (niet-sleutel) attributen

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Bij de stap naar de derde normaalvorm moeten de attributen

- die functioneel afhankelijk zijn van de volledige sleutel (2NV)
- maar ook nog functioneel afhankelijk zijn van andere attributen

in aparte groepen worden opgenomen

- 3.1 Geef de attributen aan die ook functioneel afhankelijk zijn van andere (niet-sleutel) attributen
- 3.2 Vorm een aparte groep voor ieder attribuut of combinatie van attributen, waar andere attributen functioneel van afhankelijk zijn
- 3.3 Neem in iedere nieuwe groep de attributen met hun bijbehorende sleutel op
- 3.4 Verwijder de attributen van de nieuwe groep(en) uit de oorspronkelijke groep

NV3: Derde normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 3.1

- Attributen die functioneel afhankelijk zijn van andere (niet-sleutel) attributen
 - Projectoverzicht: **Geen**
 - Projectmedewerker: **Geen**
 - Medewerker: **Chef is functioneel afhankelijk van afdeling**

NV3: Derde normaalvorm

Stap 3.2

- Er ontstaat 1 nieuwe groep:
AFDELING

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

NV3: Derde normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 3.3

- Geïntinventariseerde groepen moeten gevuld worden
- Iedere nieuwe groep moet voldoen aan de eisen van de derde normaalvorm
 - Er mogen dus géén herhalingen aanwezig zijn
 - En alle attributen moeten functioneel afhankelijk zijn van de **volledige sleutel**
 - En onderling géén functionele afhankelijkheden bevatten
- De nieuwe groep bestaat uit twee gegevens, nl. afdeling en chef

NV3: Derde normaalvorm

Stap 3.3

- Sleutel: afdeling of chef?
- Er kunnen zich verschillende situaties voordoen
 - Situatie A: iedere afdeling heeft 1 chef en iedere chef is chef van 1 afdeling
 - Situatie B: iedere afdeling heeft 1 chef, maar een chef kan van meerdere afdelingen chef zijn
 - Situatie C: iedere afdeling heeft meerdere chefs, maar iedere chef is slechts chef van 1 afdeling
 - Situatie D: iedere afdeling heeft meerdere chefs en iedere chef kan chef zijn van meerdere afdelingen.

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

NV3: Derde normaalvorm

Stap 3.3

Situatie A: iedere afdeling heeft 1 chef en iedere chef is chef van 1 afdeling

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Afdeling	Chef
A	1
B	2
C	3

Sleutel: keuze is niet belangrijk

NV3: Derde normaalvorm

Stap 3.3

Situatie B: iedere afdeling heeft 1 chef, maar een chef kan van meerdere afdelingen chef zijn

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Afdeling	Chef
A	1
B	2
C	1

Sleutel: Afdeling

NV3: Derde normaalvorm

Stap 3.3

Situatie C: iedere afdeling heeft meerdere chefs, maar iedere chef is slechts chef van 1 afdeling

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Afdeling	Chef
A	1
A	2
B	3
B	4

Sleutel: Chef

NV3: Derde normaalvorm

Stap 3.3

Situatie D: iedere afdeling heeft meerdere chefs en iedere chef kan chef zijn van meerdere afdelingen

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Afdeling	Chef
A	1
A	2
B	3
B	1

Sleutel: Combinatie van beide is uniek → samengestelde sleutel

NV3: Derde normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 3.3

Onze “keuze”: situatie A, maar B wordt niet uitgesloten, dus sleutel is afdeling

Situatie A: iedere afdeling heeft 1 chef en iedere chef is chef van 1 afdeling

Situatie B: iedere afdeling heeft 1 chef, maar een chef kan van meerdere afdelingen chef zijn

MEDEWERKER

Medewerkersnummer

Naam

Afdeling

Chef

AFDELING

Afdeling

Chef

NV3: Derde normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

Stap 3.4

- Verwijdering van de attributen uit de oorspronkelijk groep
- Sleutel van nieuwe groep blijft als attribuut in oorspronkelijke groep staan

3 NV

PROJECT	PROJECTMEDEWERKER	MEDEWERKER	AFDELING
<u>Projectnummer</u>	<u>Projectnummer</u>	<u>Medewerkersnummer</u>	<u>Afdeling</u>
Projectomschrijving	<u>Medewerkernummer</u>	Naam	Chef
Budget	Uren	Afdeling	

NV3: Derde normaalvorm

Project	Budget	Medewerker	Afdeling	Chef	Uren
001 VDU	1000	003 Patricia	Analyse	Johan	60
		005 Gert	Programmering	Ron	100
002 CRT	800	004 Rudi	Analyse	Johan	200
		005 Gert	Programmering	Ron	50
003 TTV

- Alle nieuw gevormde groepen voldoen nu aan de definitie van de derde normaalvorm:
 - Zij mogen géén herhalingen bevatten
 - Alle attributen in élké groep op zich moeten functioneel afhankelijk zijn van de volledige sleutel van die groep
 - Alle attributen in élké groep op zich mogen onderling géén functionele afhankelijkheden bevatten

PROJECT	PROJECTMEDEWERKER	MEDEWERKER	AFDELING
<u>Projectnummer</u>	<u>Projectnummer</u>	<u>Medewerkersnummer</u>	<u>Afdeling</u>
Projectomschrijving	<u>Medewerkernummer</u>	Naam	Chef
Budget	Uren	Afdeling	



Wat is normaliseren?

NV1

NV2

NV3

Verdere normaalvormen

Verdere normaalvormen

- Boyce Codd Normal Form (BCNF)
 - Werkt alleen anders dan 3NV als er meerdere overlappende kandidaat-sleutels zijn en komt heel zelden voor
- 4 NV
 - Terwijl de tweede , derde en Boyce-Codd normaalvormen zich bezighouden met functionele afhankelijkheden, houdt 4NF zich bezig met een meer algemene afhankelijkheid die bekend staat als een meerwaardige afhankelijkheid
- 5 NV
 - Redundantie in relationele databases door gerelateerde meervoudige relaties te isoleren
- 3 normaalvormen zijn voldoende in de meeste gevallen

Oefeningen p. 23