Documentatie database Netflix Statistics



Naam : Thomas Meeusen, Coen Ribbens & Marcello Haddeman

Studentnummer : 2151718 (Thomas), 2151482 (Coen) & 2151911 (Marcello)

Klas : 23INVT1A

Opdracht : Netflix Statistics

Periode : 1.2

Inhoudsopgave

[Toelichting ERD-database 3](#_Toc35800477)

[Toelichting relationeel ontwerp 3](#_Toc35800478)

[Toelichting functionele afhankelijkheden normalisering 3](#_Toc35800479)

[Bijlage A – ERD 3](#_Toc35800480)

[Bijlage B – Relationeel ontwerp 3](#_Toc35800481)

[Bijlage C - Normalisering 3](#_Toc35800482)

# Toelichting ERD-database

In bijlage A is de ERD van de database te vinden.

Profiel heeft altijd één account, omdat een profiel nooit op zichzelf kan bestaan en niet bij meerdere accounts hoort. Profielen bekijken programma’s. Dit wordt opgeslagen in de tussentabel Watched.

De tabel Program, heeft een ‘is a’ relatie met Movie en Episode. Deze relatie weergeeft een supertype-subtype relatie. Het supertype Program kan niet op zichzelf voorkomen en is dus altijd een Movie of Episode. De reden hiervan is dat er anders twee relaties van Watched naar Episode en Movie gaan, en op deze manier is het veld ‘duration’ niet meer redudant. In de implementatie is er een check constraint toegepast op Program die controleert of het programma is toegewezen aan een film of episode, maar in het geval van allebei, of geen van beide, wordt er een error gegeven.

# Toelichting relationeel ontwerp

In bijlage B is het relationele ontwerp van de database te vinden.

Wat vrijwel meteen opvalt is dat er voor iedere Delete- en Update rule de optie cascading is gekozen. Dit was het makkelijkst, omdat als er iets aangepast/verwijderd moest worden, bijvoorbeeld een waarde in de kolom serieId van de tabel Serie, dat het dan ook gelijk aangepast/verwijderd werd in de foreign key tabel van Episode. Dit geld ook voor de relaties Account – Profiel en Program – Movie/Episode en Serie - Episode. Voor deze relaties geldt dat de optie SET NULL nooit had gekund. Een profiel bestaat niet zonder een account, een programma is altijd een Episode of Movie en een Episode hoort altijd bij een serie.

De optie No Action is niet handig, omdat het eigenlijk alle mogelijkheden blokkeert van een primary key veld die een koppeling heeft.

De optie Set default was niet relevant, want er is nergens een default waarde ingesteld.

De tussentabel Watched heeft twee velden die allebei een foreign key zijn en samen een ’composed key’ zijn. Dit zijn programId en profileId. De reden dat dit een composed key is, is omdat een profiel een bepaald programma niet meerdere keren kan bekijken in de zin dat er meerdere records worden gemaakt voor dezelfde combinatie. Als een profiel hetzelfde programma opnieuw bekijkt wordt dit bijgeschreven aan het veld watchedTime.

# Toelichting functionele afhankelijkheden normalisering

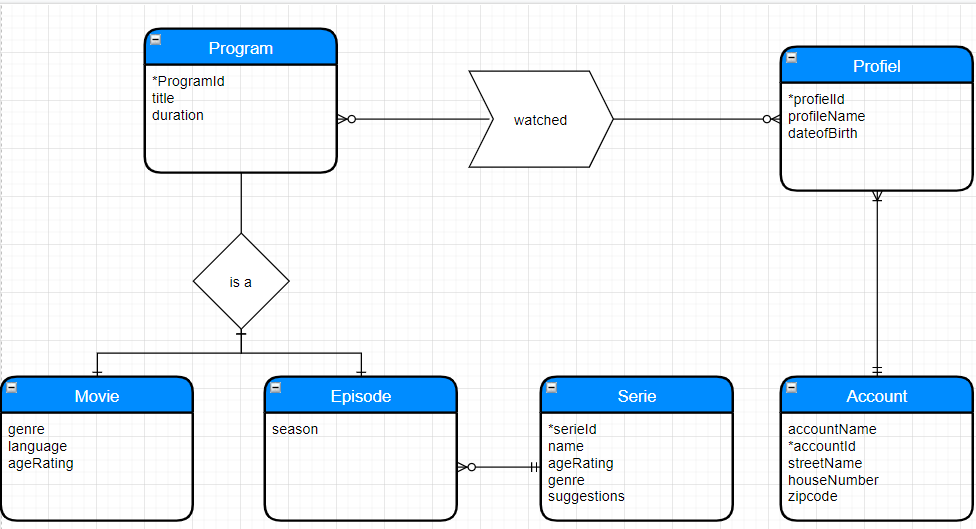
In bijlage C is de volledige normalisering in de derde normaalvorm te vinden. Niet alle records zijn in de normalisering toegevoegd, omdat het anders een erg groot Excel-bestand zou worden. Een aantal tabellen zijn gesplitst m.b.t. het voorkomen redudantie.

* Profiel en account zijn van elkaar afgesplitst vanwege redundantie.
* Serie en Episode zijn van elkaar afgesplitst vanwege redundantie.

WatchedTime uit de tabel watched is afhankelijk van de composed key profileId en programId. Dit is een koppeltabel en had niet verder opgesplitst kunnen worden.

Aangezien de tabel program een supertype is, is het afhankelijk van of episodeid of movieId, omdat het niet op zichzelf kan bestaan. De verwijzingen hiervan, de programId’s van Episode en movie, zijn de determinanten van hun tabel. Een genre van een onbekende film zegt niks.

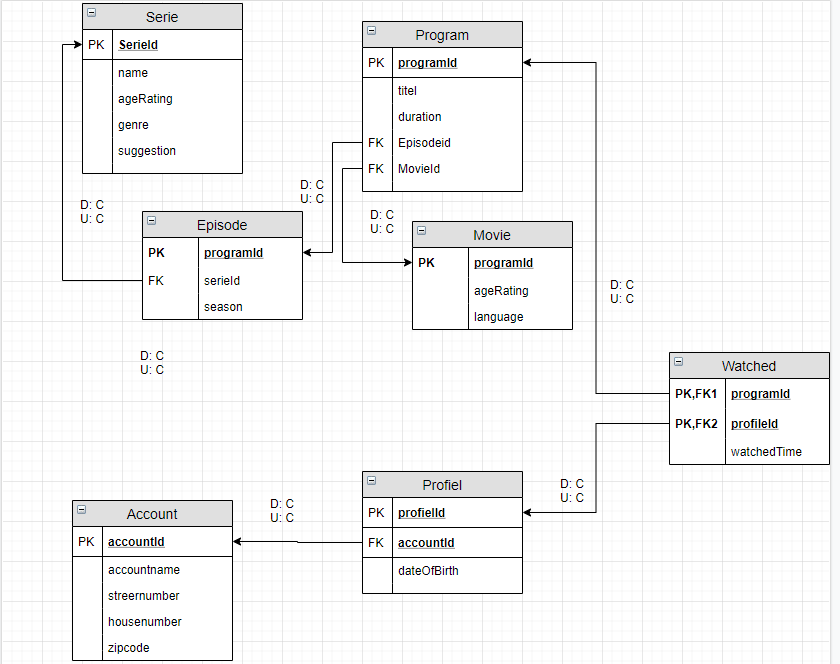
# Bijlage A – ERD



Figuur 1: ERD database Netflix Statistics

# 

# Bijlage B – Relationeel ontwerp



Figuur 2: Relationeel ontwerp database Netflix Statistics

# Bijlage C – Normalisering

