



**Handleiding Top40 database**



[Gerrit Mantel](#)

Versie 2.4

Datum 21-10-2024

# Inhoud

1	Inleiding .....	3
1.1	Database als hobbyproject.....	3
1.2	Techniek .....	3
2	Officiële Top40 website.....	4
2.1	Historie .....	4
2.2	Weeknummering.....	4
2.3	Nieuw en Re-entry.....	4
3	Mijn Top40 database.....	5
3.1	Basisgegevens.....	5
3.2	Berekende gegevens .....	5
3.3	Id van de hit.....	5
3.4	Data-verbeteringen .....	6
3.4.1	Dubbelnoteringen .....	6
3.4.2	Dubbele posities .....	6
3.4.3	Ontbrekende gegevens .....	6
3.4.4	Artiesten naamgeving .....	6
3.5	Export bestanden .....	7

# 1 Inleiding

## 1.1 Database als hobbyproject

In de jaren 90 verzamelde ik al Top40 gegevens. Voordat de [officiële Top40 website](#) alle gegevens bevatte, haalde ik de gegevens uit verschillende internet bronnen van andere hobbyisten. Nadat de ene na de andere hobbyist er mee stopte, vulde ik mijn Top40 gegevens aan met de nieuwe lijsten van de officiële Top40 website. In 2013 ben ik begonnen met mijn eigen website en heb ik de Top40 gegevens draaiend in een SQL Server database en voor publiek toegankelijk gemaakt.

Ik bemerkte recentelijk (september 2024) dat de officiële Top40 website is uitgegroeid tot een volledig gegevensbron van Top40 en Tipparade gegevens. Ik kan me nog herinneren dat een aantal jaar geleden er helemaal geen gegevens van vóór 2000 beschikbaar waren. Tegenwoordig gaan de gegevens tot het begin terug, en zijn ook alle pdf's van de lijsten beschikbaar. Ik heb daarom mijn Top40 database opnieuw opgezet, maar dan nu gevuld met gegevens van de officiële Top40 website. Ook dan ook meteen maar de Tipparade gegevens er bij.

## 1.2 Techniek

Mijn database host ik zelf in eigen huis, op een PC met MS Windows 11, SQL Server 2022, IIS en ASPRunner van Xlinesoft. Met mijn website heb ik geen commerciële bedoelingen, deze is dat ook reclamevrij. Door middel van T-SQL scripts haal ik periodiek nieuwe Top40- en Tipparade gegevens op bij de officiële Top40 website. De Top40 website heeft geen api (zoals bijvoorbeeld de The Movie Database dat heeft), en ik moest dus zelf een webpage grabber ontwikkelen. Een nadeel van webpage grabbing is dat bij een gewijzigde webpage opmaak de grabber moet worden aangepast.

## 2 Officiële Top40 website

### 2.1 Historie

De Nederlandse Top40 weeklijst is begonnen op 2 januari 1965. Lees meer over de geschiedenis van de Top40: [Geschiedenis Nederlandse Top 40](#) of hier: [Nederlandse Top 40 - Wikipedia](#)  
De Nederlandse Top40 bestaat uit een weeklijst met 40 muzieknnummers (hits).

De Nederlandse Tipparade weeklijst is volgens de Top40 website begonnen op 15 juli 1967. Lees meer over de geschiedenis van de Tipparade: [Tipparade - Wikipedia](#)  
De Nederlandse Tipparade bestaat uit een weeklijst met 30 muzieknnummers (hits). Dat is niet altijd zo geweest: tot en met 13 september 1969 bestond de lijst uit 20 nummers.

### 2.2 Weeknummering

De officiële Top40 website hanteert Top40 weeknummers die aan de kalender is gekoppeld. Als de 1<sup>e</sup> editie van een jaar op 8 januari of later is uitgekomen, hanteert de officiële Top40 website weeknummer 2. Hieraan heb ik mij in mijn database geconformeerd. Het weeknr is niet gelijk aan het editienummer. Zo vallen de 1<sup>e</sup> edities van de jaren 1982, 1983, 1988, 1993, 1994, 2000 en 2005 in week 2.

Er zijn soms gecombineerde weekedities:

- Editie 51/52 in de jaren 1985, 1986, 1988, 1989, 1990. 1991, 1993, 1994, 1995, 2000, 2001, 2002, 2012.
- Editie 52/53 in de jaren 1992, 1999
- Editie 52/1 in de jaren 1996, 1997, 1998

Er is één editie die ongewijzigde inhoud heeft t.o.v. de vorige editie:

- Week 52 (editie 52) van het jaar 1970 is inhoudelijk gelijk aan week 51 (editie 51). Deze is officieel uitgekomen en is (dus gewoon) in mijn database verwerkt.

### 2.3 Nieuw en Re-entry

Als een hit voor het eerst binnenkomt is het kenmerk "Aantal\_weken" = 1, "Vorige\_week" = - (ofwel 0) en Status = NIEUW.

Komt een bestaande hit in een nieuwe weeklijst weer voor, dan telt het aantal\_weken door.

Een hit kan uit de lijst verdwijnen, en er na één of meer weken later opnieuw inkomen. De Status is dan RE-ENTRY. Ook dan telt het aantal\_weken door. Een re-entry kan ook na een onderbreking van een jaar of langer voorkomen. Soms is er na een lange onderbreking sprake van hetzelfde muzieknnummer maar uitgebracht in een nieuwe- of andere versie. De officiële top40 website vermeldt dan meestal in zo'n geval toch "re-entry" en telt het aantal\_weken door, in plaats van een "nieuw" muzieknnummer met aantal\_weken = 1. Als een re-entry duidelijk een nieuwe versie van een muzieknnummer betreft, noem ik in mijn database als hit beide nummers, en werk ik dit met terugwerkende kracht bij in alle voorgaande top40 lijsten.

## 3 Mijn Top40 database

### 3.1 Basisgegevens

De gegevens die ik bij de officiële Top40 website per week van de Top40 en de Tipparade ophaal zijn:

- Jaar, Weeknr en Deze\_week
- Artiest en Titel
- Vorige\_week
- Aantal\_weken
- Status

### 3.2 Berekende gegevens

Elke Top40 en Tipparade hit krijgt in mijn database een aantal extra gegevens die zijn uitgerekend.

- Het Totaal\_aantal\_weken is het aantal weken gedurende de looptijd. Dit is gelijk aan de hoogste waarde van Aantal\_weken.
- De Hoogste\_pos is de hoogste positie van de hit.
- Het Aantal\_nr\_1\_pos is het aantal weken dat de hit op positie 1 heeft gestaan.
- Het Puntentotaal is de som van alle punten van de hit gedurende de looptijd. Deze wordt berekend uit de posities in de weeklijsten. Positie 1 krijg 40 punten (Top40) of 30 punten (Tipparade), etc.
- De Re-entries somt op wanneer (jaar/week) de hit (één of meer) re-entry's heeft gehad. Dit kunnen er meerdere zijn.

### 3.3 Id van de hit

Ik heb alle Top40 en Tipparade hits een Id (uniek nummer) gegeven.

Bij de Top40 is de Id afgeleid van jaar, weeknr, deze\_week bij binnenkomst. In jaar 1968, week 41 komt "The Tremeloes" met "MY LITTLE LADY" binnen op (deze\_week) 28. Het Id van deze hit is dan ook 19684128.

Bij de Tipparade is de Id afgeleid van jaar, weeknr, deze\_week en subnr bij binnenkomst. In jaar 1968, week 41 komt "Jefferson Airplane" met "IF YOU FEEL LIKE CHINA BREAKING" binnen op (deze\_week) 19. Omdat er geen dubbele entries zijn (19 komt 1x voor in de lijst), is het subnr 1 en de Id dan ook 1968411901. Soms zijn er dubbele entries in de Tipparade, meer hierover in paragraaf 3.4.2 "Dubbele posities".

Ik heb helaas geen koppeling tussen Top40 hits en Tipparade hits. Dezelfde hit heeft bij de Top40 een andere Id dan bij de Tipparade. Mogelijk dat ik in de toekomst nog een koppeling bouw. Wel heb ik mijn best gedaan om de Artiesten schrijfwijze over de Top40 en Tipparade heen te uniformeren.

## 3.4 Data-verbeteringen

### 3.4.1 Dubbelnoteringen

Op sommige posities in sommige Top40 (en ook Tipparade) weeklijsten worden soms 2 (of zelfs 3 of 4) hits genoemd. Vaak gaat het hier om 2 nummers van dezelfde artiest en worden zowel de A als de B-kant van de single vermeld. Soms echter wordt “hetzelfde” nummer van verschillende artiesten vermeld. Het verwarrende van zo’n situatie is dat de officiële Top40 website niet voor alle weken van zo’n hit dezelfde artiesten noemt. Een mooi voorbeeld hiervan is het Top40 nummer “PA WIL NIET IN BAD (LA FELICIDAD)” van “Johnny & Rijk”, binnengekomen op 35 in week 15 van 1968. In week 19 worden alleen “Digno Garcia” en “Eddie Becker” als artiesten genoemd. Vanaf week 21 komt “Johnny & Rijk” er weer bij.

In mijn database worden artiesten en titels van een hit in alle weeklijsten gelijkgetrokken, en houden de artiest en titel gedurende de gehele looptijd een consistente inhoud.

### 3.4.2 Dubbele posities

In de Tipparade komen in sommige weeklijsten posities 2x (of zelfs 3x) voor. Er staan dan dus dan 2 regels in de lijst met dezelfde “deze\_week” en totaal verschillende hits. Dit komt overigens in de Top40 weeklijsten niet voor. Of dit in de Tipparade weeklijsten fouten zijn op de officiële Top40 website weet ik niet. Ik kan dit fenomeen niet verklaren. Het komt honderden keren voor en is dus geen uitzondering. Een voorbeeld is de Tipparade van jaar 1968, week 41, positie 18. Hier staan 2 regels:

- “Ray Stevens” met “MR. BUSINESSMAN”, aantal\_weken = 5 (RE-ENTRY)
- “Deep Purple” met “HUSH”, aantal\_weken = 1 (NIEUW)

Om de dubbele posities in mijn database te kunnen verwerken, heb ik een “Subnr” geïntroduceerd. De eerstgenoemde krijgt subnr 1, de volgende 2 etc.

### 3.4.3 Ontbrekende gegevens

Op de officiële Top40 website ontbreekt er van één Top40 hit 3 weken gegevens. Dit is het geval bij de hit “SLOW JAMZ” van “Twista feat. Kayne West & Jamie Foxx”. Deze hit wordt niet genoemd in jaar 2004 in week 11 (positie 39), week 12 (positie 32) en week 13 (positie 30).

Op de officiële Top40 website ontbreekt er van één Tipparade hit 4 weken gegevens. Dit is het geval bij de hit “KISS FROM A ROSE” van “Seal”. Deze hit wordt niet genoemd in jaar 1994 in week 34 (positie 11), week 35 (positie 5), week 36 (positie 4) en week 37 (positie 3).

Beide problemen zijn uiteraard in mijn database gerepareerd.

### 3.4.4 Artiesten naamgeving

In mijn database heb ik een aantal wijzigingen doorgevoerd om namen van artiesten meer consistent te krijgen.

- Alle varianten van featuring (ft., feat, ft: etc) zijn omgezet naar feat.
- Alle varianten van versus (vs, vrs. etc) zijn omgezet naar vs.
- Alle artiesten die zowel met als zonder “The “ werden genoemd, zijn eenduidig omgezet naar zoveel mogelijk de officiële naamgeving volgens Wikipedia of Discogs. Zo zijn bv. “The Bee Gees” en “Bee Gees” nu consequent in mijn database opgenomen als en “Bee Gees”.

- Alle artiesten die zowel met “ and ”, “ & ” of “ + ” werden genoemd, zijn eenduidig omgezet naar “ and ” of “ & ” afhankelijk van de officiële naamgeving. Zo zijn “Diana Ross & The Supremes”, “Diana Ross and The Supremes” en “Diana Ross + Supremes” nu consequent in mijn database opgenomen als en “Diana Ross & The Supremes”.

In een aantal artiestennamen komen (gelukkig heel sporadisch) tekens voor die niet in mijn database overgenomen kunnen worden. Het gaat hier om tekens die niet in de zogenaamde Iso-Latin1 characterset voorkomen, bijvoorbeeld griekse of russische tekens. Deze heb ik omgezet naar een ander teken, of de artiestennaam omgezet naar de engelse naam.

Axwell **Α** Ingrosso is geworden:  
Axwell **A** Ingrosso

Avicii & Martin Kupilas & **Ваня Хакси** & Kian Sang & Naxsy DJ-Compositeur is geworden:  
Avicii & Martin Kupilas & **Нахху** & Kian Sang & Naxsy DJ-Compositeur

In een aantal artiestennamen komen æ of œ ligatuur karakters voor, terwijl dat niet de officiële schrijfwijze is. De æ of œ ligatuur heb ik in zo'n geval omgezet naar resp. de letters ae en oe. Dit kwam o.a. voor bij “George Michael **æ**” en “Orchestral Manoeuvres In The Dark”).

### 3.5 Export bestanden

Binnen de Top40 schermen van mijn database zijn exportbestanden (CSV) te vinden die kunnen worden gedownload. De volgende bestanden zijn beschikbaar voor de Top40 of de Tipparade:

- Totaallijst alle hits: **Top40 Hits CSV** of **Tipparade Hits CSV** (scherm: ... Totaal)
- Totaallijst alle noteringen: **Top40 Noteringen CSV** of **Tipparade Noteringen CSV** (scherm: ... Totaal)
- Hit-jaarlijsten: **Jaarlijst CSV** (scherm: ... Jaren)

Gedownloade CSV bestanden kunnen worden geopend in MS-Excel. Zet in Excel de gegevens in de juiste kolom met de volgende methode:

Selecteer kolom A (klik op kolomhoofd A) en kies achtereenvolgens voor:

- Gegevens
- Tekst naar kolommen
  - o Stap 1: **Gescheiden**
  - o Stap 2: Scheidingsteken **Tab**
  - o Stap 3: Zet kolom Artiest en Titel op gegevenstype **Tekst**
- Voltooiën