Voor dit labo gaan jullie aan de slag met Power BI en de twee bijgevoegde tabellen:

* NetFlixViewingHistory.csv: Een (vereenvoudigde) geschiedenis met wanneer iemand een bepaalde film of serie heeft bekeken op Netflix (als je zelf een Netflix account hebt kan je ook jou geschiedenis downloaden, zie: [How to see viewing history and device activity](https://help.netflix.com/en/node/101917)).
* IMDb movies.csv: allerlei informatie over films van iMDB

Installeer [Power BI Desktop](https://powerbi.microsoft.com/en-au/desktop/) (gebruik je Windows account van AP met je s-nummer als je moet inloggen).

1. Data inladen
   1. Laad het bestand NetFlixViewingHistory.csv in via *Gegevens ophalen* 🡪 *Tekst/csv*
   2. Klik op *Gegevens trasformeren*
   3. Laad het bestand IMDb movies.csv in via *Nieuwe bron* 🡪 *Tekst/csv*
2. Data voorbereiden
   1. Doe een join van de twee tabellen door te klikken op *Query’s samenvoegen als nieuw*. Duid de twee tabellen aan, klik telkens op de title kolom en kies voor een inner join.
   2. Bij de nieuwe tabel moet je helemaal rechts nog bovenaan de kolommen uitvouwen.
   3. Verander het data type voor de kolom date\_published naar Datum en verander de naam van de kolom naar watchdate
   4. Verwijder het dollarteken in de kolommen budget, usa\_gross\_income en worldwide\_gross\_income door bv. gebruik te maken van de functie *kolom splitsen*. Verwijder de kolommen met enkel het dollarteken en pas de naam van de aangepaste kolommen aan.
   5. De kolom metascore is de iMDb score van recensenten op een totaal van 1000. Maak een nieuwe Metascore die de score op een totaal van 10 geeft dor de functie *Aangepaste kolom* te gebruiken om een nieuwe kolom te maken met de metascore gedeeld door 100. Verwijder de oude metastore kolom.
   6. De avg\_vote kolom bevat de review score van Netflix gebruikers. Maak een nieuwe kolom met deze vote op 10 (en verwijder de oude kolom).
   7. Hernoem de tabel naar Movies and Genres noemt en hou enkel de kolommen title, imdb\_title\_id, watchdate, genre, duration, production\_company, budget, worldwide\_gross\_income, Metastore, avg\_vote, year en usa\_gross\_income.
   8. Gebruik *Kolom splitsen* bij de genre kolom en hou enkel de eerste waarde bij.
   9. Gebruik de rechtermuisknop om voor de originele tabellen de optie *Laden inschakelen* uit te zetten.
   10. Klik nu op *Sluiten en toepassen* om aan het dashboard te beginnen.
3. Dashboard maken (zie voorbeeld hieronder):
   1. Je ziet nu rechts de tabel staan die je hebt voorbereid met de verschillende datavelden (kolommen).
   2. Importeer op het tabblad *Weergeven* als thema de bijgevoegde file “Netflix Theme.json”.
   3. Importeer het bijgevoegde Netflix logo en gebruik Vormen en Tekstvakken om je dashboard vorm te geven.
   4. Gebruik een *kaart* om de totale tijd weer te geven (duration). Deze staat in minuten wat niet zo duidelijk is. Selecteer *Nieuwe meting* door met de rechtermuisknop op het veld duration te klikken. Je kan nu DAX gebruiken om een nieuw veld te maken (Watch Time (hours) = SUM('Main Table'[duration])/60)
   5. Maak een *treemap* met genre als *Categorie* en title als *Waarden.* Gebruik een filter om enkel de populairste genres te tonen en pas de titel aan zoals in het voorbeeld.
   6. Gebruik een *Lijndiagram* met de watchdate op de x-as en title op de y-as. Zorg dat je kan inzoomen per maand (en niet per kwartaal) en pas de titel aan. Zet de titels van de x- en y-as uit en verwijder de rasterlijnen.
   7. Gebruik een Gestapeld staafdiagram met production\_company op de y-as en (Totaal van) title op de x-as. Toon enkel de populairste bedrijven, verwijder de x-as en de titels van de x- en y-as, zet de *Gegevenslabels* aan en pas de titel aan. Laat de kleuren van lichtrood naar donkerrood gaan (via fx bij *Balken*) en zet de legende uit.
   8. Gebruik nog een Gestapeld Staafdiagram om een analoge grafiek te maken met de grootste films op basis van wereldwijd gegenereerd inkomen.
   9. Maak een nieuwe meting genaamd Profit met de winst van de film (wereldwijd inkomen min het budget).
   10. Kopieer de vorige staafdiagram en pas de parameters aan voor het tonen van de films met de grootste profit (pas ook de titel en de kleurovergang aan).
   11. Voeg een spreidingsdiagram toe met op de x-as de metascore gegeven door recensenten en op de y-as de avg\_vote van Netflix gebruikers. Hernoem de assen naar Critics score en Consumers score. Zet het veld title op waarden om de films te zien en pas de aggregatie voor de assen aan van som naar gemiddelde zodat als een film meerdere keren voorkomt de scores niet worden opgeteld. Verwijder de titel en voeg nog een trendlijn toe zodat je kan zien welke films beter worden gescoord door recensenten dan gebruikers.
   12. Voeg een slicer toe met een drop down keuzemenu om te filteren op jaar en voeg de optie *Alles selecteren* toe
   13. Voeg nog een slicer toe om de genres te selecteren.
   14. Merk op dat je de data ook kunt filteren door in een grafiek te klikken. Je kan de manier waarop dit een invloed uitoefent op andere grafieken aanpassen via *Interacties bewerken* in het blad *Indeling*.

Het uiteindelijke dashboard zal er ongeveer zo uitzien:

