A1

Welke acteur (m/v) die in Friesland is geboren heeft het meeste succes gehad. Denk na over hoe je succes definieert.

**acteur met geboorteplaats in friesland met de meeste rollen in films met hogere rating dan ???**

A3

Welke acteur (m/v) heeft de langste filmcarrière? (dus geen series)

**view met naam, laagste jaartal en hoogste jaartal voor die acteur. naam van acteur waar het verschil tussen die twee het grootste is.**

A9 (variatie new york variabel maken)

Welke films spelen (gedeeltelijk) in New York?

**alle film namen (distinct) waarbij de location string ‘new york’ bevat**

A15

Welke acteur of actrice speelt het meest in de slechtst gewaardeerde films?

**meest voorkomende acteur/actrice naam in films die een rating hebben lager dan ???**

A16

In welk jaar tussen 1990 en nu zijn de meeste films met het woord ‘beer’ in de titel geproduceerd. En wat is het meest voorkomende genre?

**view met alle films met het woord ‘beer’ in de titel. daarvan voor elk jaar tellen hoe vaak dat voorkomt en ook voor elk genre hoe vaak het voorkomt.**

B3

Geef het aantal films dat in een land gemaakt is weer in de tijd. Dwz maak een grafiek waarin op de x-as het jaar staat en op de y-as het aantal gemaakte films.

**een land uitkiezen (of variabel maken) dan een view met alle films die in dat land gemaakt zijn en dan tellen hoe vaak elk jaar voorkomt en dat in een grafiek zetten.**

B8

Maak een kaart (b.v. google maps / openstreetview) met landen waar een flim speelt. Zodat op de kaart te zien is waar de films spelen.

**een film kiezen of variabel maken en dan van alle bijbehorende locaties de gps-coördinaten opvragen (via bv. Maps) en dat op een kaartje weergeven**

C4

Er wordt vaak gezegd dat naarmate actrices ouder zijn zij moeilijker een rol krijgen. Ga (met R) na of dat zo is. Je zou bijvoorbeeld een lineair regressiemodel kunnen maken met leeftijd en geslacht als onafhankelijke variabelen en aantal films per jaar als afhankelijke variabele.

**lineair regression met R**

Eigen vraag

Onderzoek of er een verband is tussen het aantal seizoenen dat van series wordt gemaakt en de gemiddelde ratings van de series**.**

**gemiddelde rating per seizoen plotten met R**

**Eigen vragen(minimaal 2):**

Top 10 woorden die het meest voorkomen in een serie titel, weergeven in een grafiek

**R text analyse, frequencies / frequent terms**

Top 10 dodelijkste plek voor acteurs (meest voorkomende location of death)

**voor elke location of death tellen hoe vaak het voorkomt**