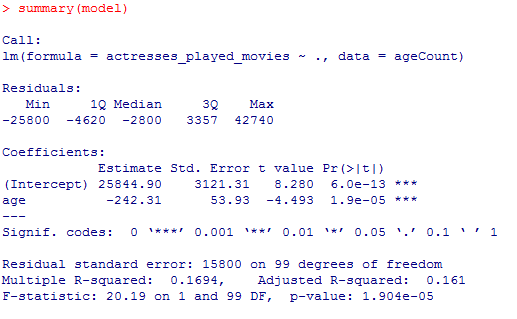
# In SQL:

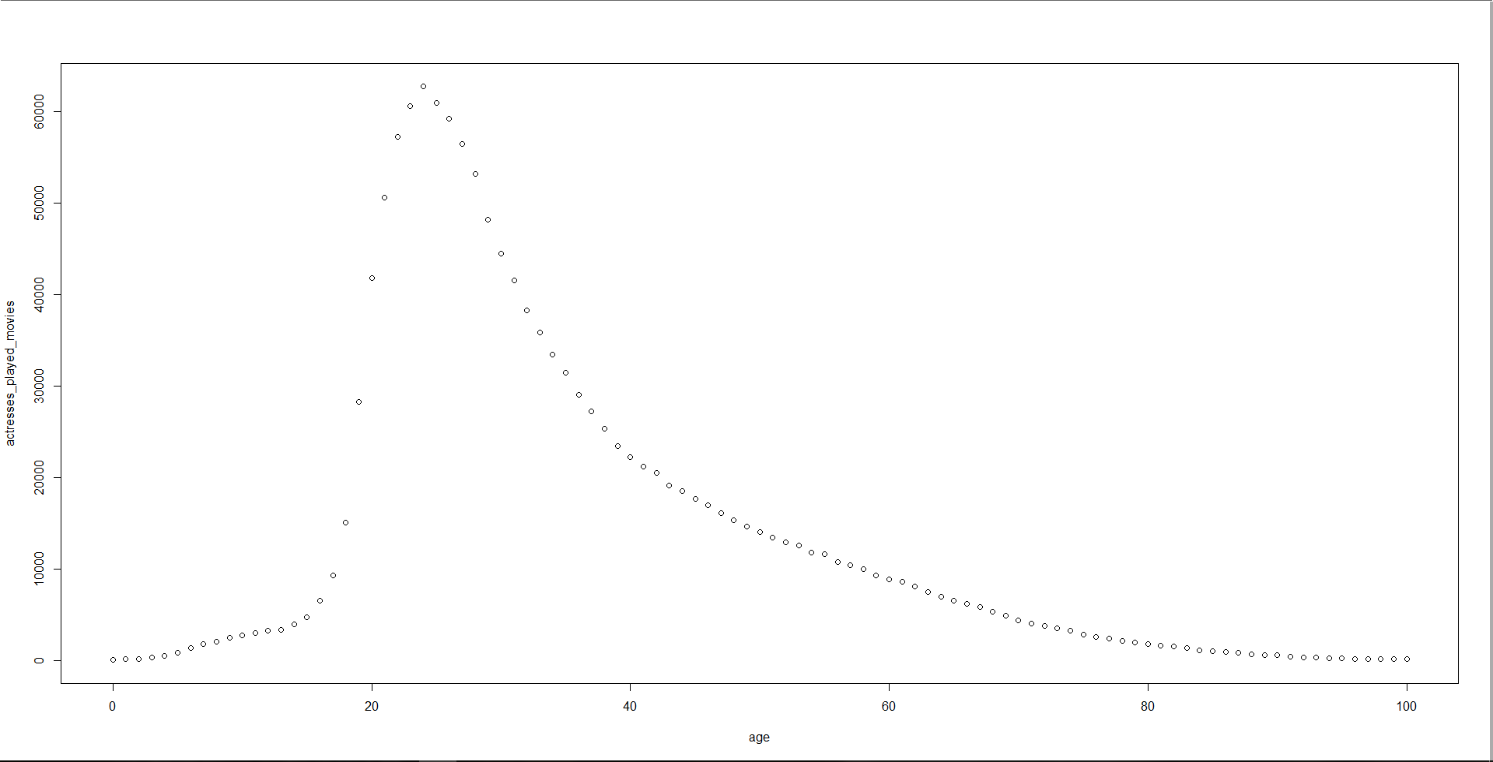
**CREATE VIEW** age\_count **AS SELECT M**.**year\_of\_release** - *CAST*(A.**birth\_date AS INTEGER**) **AS** age, *COUNT*(A.**actor\_id**) **AS** actresses\_played\_movies **FROM** final.actors A **INNER JOIN** final.movie\_actors MA **ON** A.**actor\_id**=MA.**actor\_id INNER JOIN** final.movie **M ON** MA.**movie\_id**=**M**.**movie\_id WHERE M**.**year\_of\_release IS NOT NULL AND** A.**birth\_date IS NOT NULL AND** A.**gender** = **'f' AND M**.**year\_of\_release** - *CAST*(A.**birth\_date AS INTEGER**) **BETWEEN** 0 **AND** 100 **GROUP BY M**.**year\_of\_release** - *CAST*(A.**birth\_date AS INTEGER**) **ORDER BY** age;

# In R:

> ageCount <- dbReadTable(con, c("final","age\_count"))

> model=lm(actresses\_played\_movies~., data=ageCount)



> plot(ageCount)

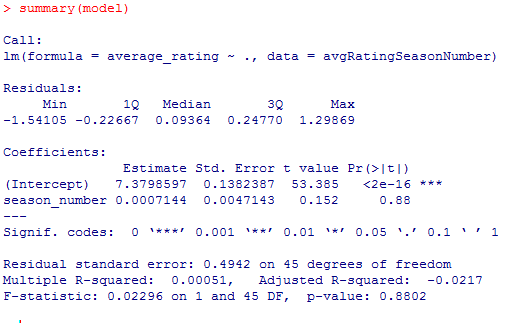
**CONSLUSIE**: Uit de grafiek kun je afleiden dat de leeftijd van een actrices inderdaad invloed heeft op het aantal rollen. De top zit rond de 24 jaar en naarmate zij ouder worden, krijgen zij steeds minder rollen. Het vermoeden wat in vraag C4 staat, is dus waar. Ook zien we ook een hoge significatie en daaruit kun je opmaken, dat de leeftijd en het aantal rollen inderdaad invloed op elkaar heeft.

# In SQL:

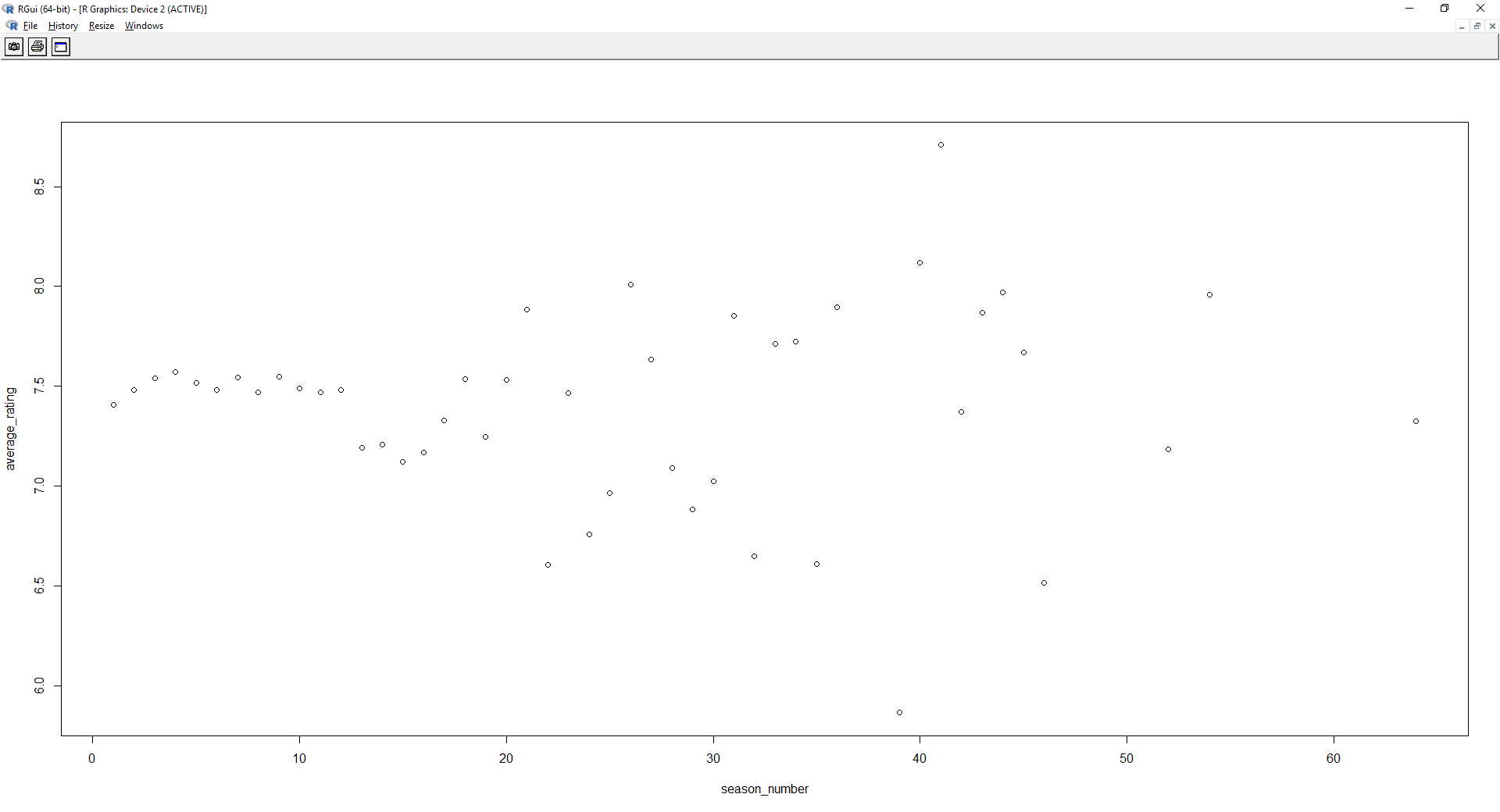
**CREATE VIEW** avg\_rating\_season\_number **AS SELECT** season\_number, *avg*(rating\_serie) **AS** average\_rating **FROM** (**SELECT** *max*(**season\_nr**) **AS** season\_number, *avg*(rating\_season) **AS** rating\_serie **FROM** (**SELECT title**, **season\_nr**, *avg*(**rating\_major** + 0.1\***rating\_minor**) **AS** rating\_season **FROM** serie **S WHERE season\_nr IS NOT NULL AND rating\_major IS NOT NULL AND rating\_minor IS NOT NULL GROUP BY title**, **season\_nr**) **AS** season\_rating **GROUP BY title**) **AS** serie\_rating **WHERE** season\_number < 100 **GROUP BY** season\_number;

# In R:

> avgRatingSeasonNumber <- dbReadTable(con, c("final","avg\_rating\_season\_number"))

> model=lm(average\_rating~., data= avgRatingSeasonNumber)

> plot(avgRatingSeasonNumber)



**CONSLUSIE**: Uit bovenstaande onderzoek, blijkt dat er nauwelijks tot geen verband bestaat tussen het aantal seizoenen dat van een serie wordt gemaakt en de gemiddelde ratings van de bijbehorende series. We zien dat er tot ongeveer 15 seizoenen het gemiddelde allemaal rond de 7.5 ligt. Daarna is er geen duidelijke lijn door te trekken. Ook zien we een hele lage significatie en daaruit maken wij op dat er geen verband is tussen het aantal seizoenen en de rating van een serie.