Oblig2_4 - spørringer mot databasen Movies -Yahye Abdi Ahmed

Table of Contents

Oppgave 1: Hent ut alle filmene sortert på 'year'	2
Oppgave 2: Hent ut alle filmene som har en metascore	
Oppgave 3: Hent ut filmer og hvem som har regissert filmen (director)	
Oppgave 4: Hent navnet og antallet filmer en regissør har regissert, sortert synkende på antallet	5
Oppgave 5: Hent navnet og antallet filmer en regissør har regissert, samt totalinntekten (gross) for	r
disse filmene, sortert synkende på inntekten	6
Oppgave 6: Hent ut alle filmer, samt tilhørende regissør og sjanger (genre)	7
Oppgave 7: Hent ut film og regissør for den filmen som har gitt høyest inntekt	8
Oppgave 8: Hent ut filmer med høyest inntekt for hver sjanger	9
	.11
Oppgave 9: Hent ut hvilke filmer Christopher Nolan har regissert, som også Christian Bale har sp	ilt
	.11
Oppgave 10: Hent ut skuespillerne og antallet filmer de har spilt i, sortert synkende på antallet	.12

Oppgave 1: Hent ut alle filmene sortert på 'year'

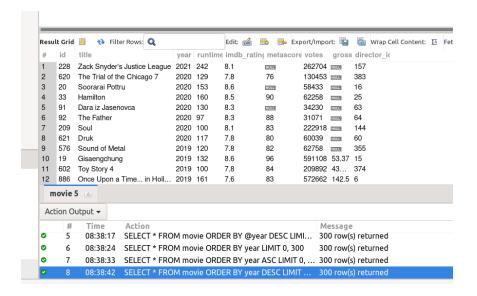
SQL-spørring:

```
# 1: Hent ut alle filmene sortert på 'year'
SELECT *
FROM movie
ORDER BY year DESC;
```

I spørringen her, henter vi alt fra tabellen movie og sorterer data i tabellen med kolonnen "year".

SELECT * henter alt. FROM movie fra tabellen movie. ORDER BY year DESC er gruppering seksjon. Year kolonnen er data type, så den blir sortert på år, her med DESC blir det synkende. Filmen som nyest i lista er fra 2021.

Spørringen har LIMIT på 300 rad her.



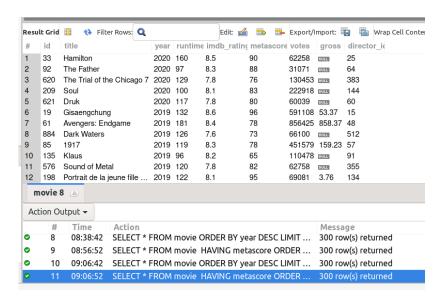
Oppgave 2: Hent ut alle filmene som har en metascore

SQL-spørring:

SELECT *
FROM movie
HAVING metascore
ORDER BY year DESC;

SELECT * FROM movie er SELECT seksjonen som henter alt fra tabellen movie. Under det har vi HAVING seksjon som sjekker om kolonnen metascore har en verdi og ikke er NULL. Det betyr at rader der kolonnen metascore er NULL blir ikke tatt med. Tilslutt har vi ORDER BY clause eller gruppering seksjon. Oppgaven spør ikke om å gruppere, så det er litt ekstra.

Sånn ser resultat av spørringen ut. Vi ser ikke noe film som ikke har metascore.



Oppgave 3: Hent ut filmer og hvem som har regissert filmen (director)

SQL-spørring:

```
SELECT m.*, d.name
FROM movie AS m INNER JOIN director AS d
ON m.director_id = d.id;
```

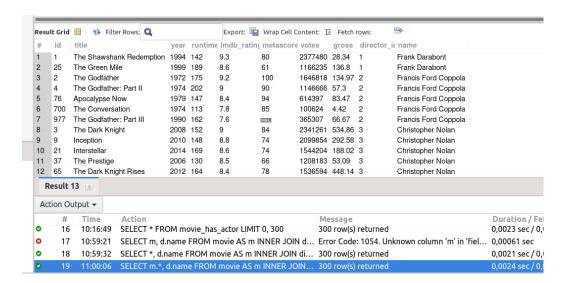
Denne spørringen er litt mer komplisert med INNER JOIN. To tabeller kobles sammen. Det gjøres ved bruk av primary key og foreign key. Forholdet mellom tabellen **movie** og **director** er "one to many". En film kan ha kun 1 director, men 1 director kan ha flere filmer assosiert med seg. Movie tabell har foreign key av tabellen director i kolonnen som heter director_id som representerer primary key for id kolonnen i **director** tabellen. På denne måten blir de to tabellene kunne bli koblet sammen med nøklene.

SELECT m.*, d.name - henter alt movie med "m.*". M er alias for tabellen **movie**, .* henter alle kolonner i movie tabellen. d.name, d er alias for **director** tabellen.

FROM movie **AS** m **INNER JOIN** director **AS** d. Dette bestemmer hvilke kolonne det blir hentet data fra. Movie AS m henter fra tabellen **movie** og samtidig gir tabellnavnet alias som kan bruke i spørringen for enklere referere til tabellen movie. Med INNER JOIN, kobler vi sammen tabellene.

ON bestemmer når hvor eller kolonne de to tabellene kobles sammen på. Her med movie.director_id og director.id.

INNER JOIN henter som har felles verdi i tabellene. Også er det mulig å velge i SELECT seksjonen hvilke tabeller vi ønsker å hente. Her i resultat ser



Oppgave 4: Hent navnet og antallet filmer en regissør har regissert, sortert synkende på antallet

SQL-spørringen:

```
# 4: Hent navnet og antallet filmer en regissør har regissert, sortert synkende
på antallet
SELECT d.name AS regissor, COUNT(d.name) AS antall_filmer
FROM movie AS m
INNER JOIN director AS d
ON m.director_id = d.id
GROUP BY regissor
ORDER BY antall_filmer DESC
;
```

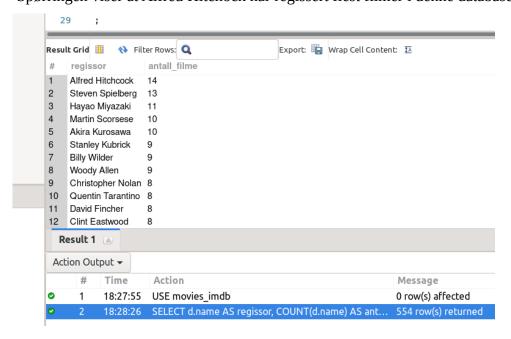
For å kunne hente antall filmer en regissør har regissert, slår vi sammen movie og director tabellen. Så grupperer resultatet med regissører. Til slutt sorterer vi resultatet synkende med antall filmer per regissør. Fra regissør med flest antall filmer til den med minst antall filmer regissert.

SELECT d.name **AS** regissor, **COUNT**(d.name) **AS** antall_filmer. I SELECT seksjon henter og viser kolonnen name fra director med alias regissor. COUNT(d.name) aggregering funksjon som teller opp alle rader per director sin navn.

FROM movie AS m klonnene blir fra movie tabellen, alias blir m for movie. Vi joiner movie tabellen med director tabellen med INNER JOIN director AS d. Med ON m.director_id = d.id blir de to tabellene koblet sammen ved primary key og foreign key. Det er ganske lik INNER JOIN som i oppgave3.

Vi grupperer **GROUP BY** regissor med director name kolonne. Sånn at vi får resultat og ikke bare 1. Tilslutt sorteres data høyest til laveste antall_filmer, som returnerer antall film per director har regissert.

Spørringen viser at Alfred Hitchock har regissert flest filmer i denne databasen.



Oppgave 5: Hent navnet og antallet filmer en regissør har regissert, samt totalinntekten (gross) for disse filmene, sortert synkende på inntekten

SQL-spørringen:

```
SELECT d.name AS regissor, COUNT(d.name) AS antall_filmer, SUM(m.gross) AS
totalinntekt
FROM movie AS m
INNER JOIN director AS d
ON m.director_id = d.id
GROUP BY regissor
ORDER BY totalinntekt DESC
;
```

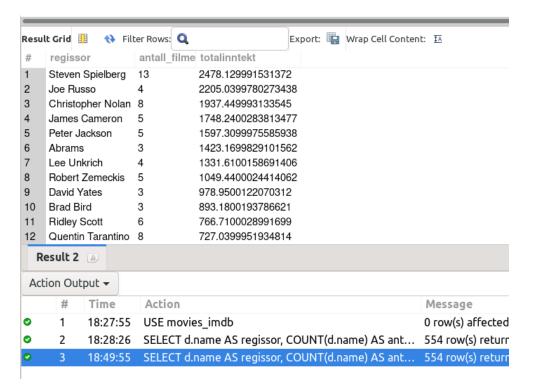
Denne spørringen bygger videre på fra den i oppgave4 over. Det som er nytt er at vi hentet kolonne sum inneholder SUM() av gross og at data blir sortert med totalinntekt per regissør.

I SELECT seksjon henter 3 kolonner, director name, antall_filmer som er lik den i oppgave4 og SUM() movie.gross med alias totalinntekt.

INNER JOIN director AS d ON m.director_id = d.id slår sammen movie og director tabellen med primary og foreign key.

GROUP BY regissor, data blir gruppert med regissor igjen her også og tilslutt sortert med totalinntekt synkende. Så vi får se synkende liste hvem som har tjent mest inntekt av filmene sine sammenlagt.

Steven Spielberg med 13 filmer har størst inntekt, men på andre plass Joe Russo med bare 4 tjent ganske mye.



Oppgave 6: Hent ut alle filmer, samt tilhørende regissør og sjanger (genre)

SQL-spørring:

```
# 6: Hent ut alle filmer, samt tilhørende regissør og sjanger (genre)
# ressurs: fikk god forklaring fra studentassistenter på hvordan tabeller kobles
i spørring i oppgave 4 og 5.
SELECT m.*, d.name AS regissor, g.name AS sjanger
FROM movie AS m
INNER JOIN director AS d
ON m.director_id = d.id
INNER JOIN movie_has_genre AS mg
ON mg.movie_id = m.id
INNER JOIN genre AS g
ON mg.genre_id = g.id;
```

I denne spørringen kobler vi sammen 4 tabeller. Først kobler vi **movie** og **director** tabellene sammen. Etter det kobler vi **movie** tabellen til mellom tabellen **movie_has_genre**. Til slutt kobler vi **movie_has_genre** tabellen til **genre** tabellen.

SELECT m.*, d.name **AS** regissor, g.name **AS** sjanger. Select seksjon henter alle kolonner fra tabellen **movie** med m.*. d.name er name kolonnen fra **director** tabellen. Den henter navn på director. Alias for denne kolonnen er regissor som blir vist i sluttresultat. Siste kolonnen g.name med alias sjanger er name kolonnen fra tabellen **genre**.

FROM movie AS m. Vi henter data fra tabellen movie.

INNER JOIN director AS d ON m.director_id = d.id. Første INNER JOIN kobling mellom movie tabellen og director tabellen. Det blir gjort ved å bruke primary key og foreign key. Movie tabellen har foreign til tabellen director som referer til primary key id i director tabellen. Så INNER JOIN skjer ved å sammenligne foreign key og primary key med ON.

INNER JOIN movie_has_genre AS mg ON mg.movie_id = m.id. Andre INNER JOIN samme
syntax som den. Movie tablen og genre tabellen many to many forhold.
move_has_genre tabellen er mellom tabellen danner many to many forholdet mellom
movie og genre tabellen.

Så for å koble movie og genre tabellene, gjør vi koblingen ved å koble movie_has_genre til movie tabellen på den ene siden og koble movie_has_genre til genre tabellen på andre siden. Det er samme som 3 person som står ved siden av hverandre og holder hender. Den personen som er i midten blir som mellom tabellen movie_has_genre.

movies_has_genre inneholder foreign key til både movie og genre tabellen. Her kobler vi movie tabellen til movie_has_genre tabellen ved å sammenligne med ON primary key i movie tabellen med foregin i movie_has_genre tabellen.

INNER JOIN genre AS g ON mg.genre_id = g.id;. Til slutt kobler movie_has_genre
tabellen til genre tabellen. Koblingen er lik den mellom movie og
movie_has_genre.

I INNER JOIN sammeligner med primary key i **genre** tabellen med foreign key i **movies_has_genre** tabllen.

I resultat ser vi at alle kolonne fra **movie** tabellen hentet. Det 10 kolonner i movie tabellen. Fra **director** tabellen henter vi kun regissor kolonnen og **genre** henter vi kun en kolonne, sjaner kolonnen.

sult	Grid	1			Export:	Wrap Cell C	ontent: 🗓	Fetch	rows:		
	id	title	year	runtime	imdb_rating	metascore	votes	gross	director_i	regissor	sjanger
	1	The Shawshank Redemption	1994	142	9.3	80	2377480	28.34	1	Frank Darabont	Drama
2	25	The Green Mile	1999	189	8.6	61	1166235	136.8	1	Frank Darabont	Drama
2	25	The Green Mile	1999	189	8.6	61	1166235	136.8	1	Frank Darabont	Crime
2	25	The Green Mile	1999	189	8.6	61	1166235	136.8	1	Frank Darabont	Fantasy
2	2	The Godfather	1972	175	9.2	100	1646818	134.97	2	Francis Ford Coppola	Drama
2	2	The Godfather	1972	175	9.2	100	1646818	134.97	2	Francis Ford Coppola	Crime
4	4	The Godfather: Part II	1974	202	9	90	1146666	57.3	2	Francis Ford Coppola	Drama
4	4	The Godfather: Part II	1974	202	9	90	1146666	57.3	2	Francis Ford Coppola	Crime
1	76	Apocalypse Now	1979	147	8.4	94	614397	83.47	2	Francis Ford Coppola	Drama
1	76	Apocalypse Now	1979	147	8.4	94	614397	83.47	2	Francis Ford Coppola	Mystery
1	76	Apocalypse Now	1979	147	8.4	94	614397	83.47	2	Francis Ford Coppola	War
	700	The Conversation	1974	113	7.8	85	100624	4.42	2	Francis Ford Coppola	Drama
Re	sult ?	2 🚇									
Actio	on O	utput ▼									
	#	Time Action					Me	essage			Dur
	1	21:57:40 SELECT m.*,	d.name	e AS regi	issor, g.nam	e AS sjange	er 100	00 row(s) returned		0,00
	2	22:38:21 SELECT m.*,	d name	AS regi	issor, a.nam	e AS siange	r 100	00 row(s) returned		0,00

Oppgave 7: Hent ut film og regissør for den filmen som har gitt høyest inntekt

SQL-spørring:

```
# 7: Hent ut film og regissør for den filmen som har gitt høyest inntekt
# gruppering når MAX(gross) er i øverst SELECT seksjon fungerte dårlig. Da må vi
gruppere data og vi får ikke 1 resulat i gruppering.
```

```
SELECT m.title AS film, d.name AS regissor,m.gross AS inntekt
FROM movie AS m
INNER JOIN director AS d
ON m.director_id = d.id
WHERE m.gross = (SELECT MAX(gross) FROM movie)
.
```

I denne kobler vi sammen movie tabellen med director tabellen.

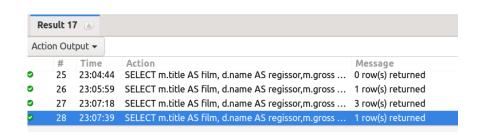
Select seksjon henter movie.title med alias film. director.name med alias regissor. Tredje og siste kolonnen er movie.gross med inntenkt. FROM movie, henter fra tabellen movie med alias m.

INNER JOIN director **AS** d **ON** m.director_id = d.id. INNER JOIN i one to many relationship blir samme som vi gjorde oppgavene før. Movie tabellen og genre tabellen kobles sammen med primary og foreign. Movie tabellen har foreign key til director tabellen.

WHERE m.gross = (SELECT MAX(gross) FROM movie);. Tilslutt kjører MAX(gross) i WHERE seksjon. Dette gjør at vi ikke trenger å gruppere og vi hentet kun 1 rad som inneholder filmen størst gross. Dette gjør ved movie.gross filmen er lik den største i select parentes.

I resultat henter vi kun 1 rad. Star Wars Episode VII er filmen med størst gross i denne databasen.





Oppgave 8: Hent ut filmer med høyest inntekt for hver sjanger

SQL-spørring:

```
# 8: Hent ut filmer med høyest inntekt for hver sjanger
# source: https://www.youtube.com/watch?v=cW2MLIZ5CaU
# jeg satt fast lenge på den, men fant løsning på video som forklarte at jeg
kunne ikke ha irrelevante kolonner i select som ikke er en del GROUP BY clausen
# fikk hjelp med studass
SELECT movie.title AS film, x.sjanger AS sjanger, x.inntekt
FROM
(
SELECT g.name AS sjanger, MAX(m.gross) AS inntekt
FROM movie AS m
INNER JOIN movie_has_genre AS mg
ON mg.movie_id = m.id
INNER JOIN genre AS g
ON mg.genre_id = g.id
GROUP BY g.name
) AS x, movie
```

Denne spørringen består av SELECT seksjoner.

SELECT movie.title **AS** film, x.sjanger **AS** sjanger, x.inntekt. I første SELECT seksjon velger vi ut 3 kolonner. Første kolonne fra movie tabellen og henter navn på filmer med movie.title, alias filmer. Andre kolonnen er tabellen x og henter sjanger med x.sjanger, alias sjanger. X-tabellen er resultat fra andre SELECT seksjon som blir sett som egen kolonne, der x er alias for den tabellen. Tredje kolonne fra movie tabellen og henter gross klonnen med movie.gross, alias inntekt.

```
FROM
(
SELECT g.name AS sjanger, MAX(m.gross) AS inntekt
FROM movie AS m
INNER JOIN movie_has_genre AS mg
ON mg.movie_id = m.id
INNER JOIN genre AS g
ON mg.genre_id = g.id
GROUP BY g.name
```

) **AS** x, movie Vi henter de 3 kolonnene i SELECT fra subquery SELECT her som INNER joiner movie tabellen med genre tabellen gjennom movie_has_genre mellom tabellen. SELECT i subquery henter 2 kolonnen, først kolonne fra genre tabellen, genre.name med alias sjanger. Andre kolonnen er fra movie og henter movie.gross kolonnen. Her bruker vi aggreger funksjon MAX() for å hente kun filmer med høyest gross, alias for kolonnen er inntekt. Resultat i subquery blir gruppert med genre.name, altså sjanger og vi kun hver sjanger kun 1 gang. Tilsammen 21 sjanger. Subquery her alias x. Vi henter også movie tabellen.

WHERE x.inntekt = movie.gross; Til slutt, I WHERE sammenligner i inntekt fra subquery med movie.gross som gjør at vi kan se navnet på filmet som har høyest gross i hver sjanger.

Vi får 21 treff i resultat siden antall sjanger er 21.

Resu	lt Grid	- Name of the Filter	Rows: Q		Export:	Wrap Cell Content	: <u>IA</u>
#	film		sjanger	inntekt	Л		
1	The Da	ark Knight	Crime	534.86			
2	Saving Private Ryan		War	216.54			
3	Joker		Thriller	335.45			
4	Avenge	ers: Endgame	Drama	858.37			
5	The Si	xth Sense	Mystery	293.51			
6	Gone v	with the Wind	History	198.68			
7	Bohem	nian Rhapsody	Music	216.43			
3	Dances with Wolves		Western	184.21			
9	The Ex	corcist	Horror	232.91			
10	Fiddler	on the Roof	Musical	80.5			
11	Star W	ars: Episod	Action	936.66			
12	Star W	ars: Episod	Adventure	936.66			
R	esult 3						
Act	ion Ou	tput ▼					
	#	Time	Action				Message
•	1	20:05:58	SELECT mo	vie.title AS film, >	k.sjanger AS	sjanger, x.in	21 row(s) returned
9	2	20:06:38	SELECT mo	vie.title AS film, >	k.sjanger AS	sjanger, x.in	70 row(s) returned
>	3	20:07:10	SELECT mo	vie.title AS film, >	k.sjanger AS	sjanger, x.in	21 row(s) returned

Oppgave 9: Hent ut hvilke filmer Christopher Nolan har regissert, som også Christian Bale har spilt i

SQL-spørring:

```
# 9: Hent ut hvilke filmer Christopher Nolan har regissert, som også Christian
Bale har spilt i
SELECT m.title AS film, d.name AS regissor, a.name AS skuespiller
FROM movie AS m
INNER JOIN director AS d
ON m.director_id = d.id
INNER JOIN movie_has_actor AS ma
ON ma.movie_id = m.id
INNER JOIN actor AS a
ON ma.actor_id = a.id
WHERE d.name = "Christopher Nolan" AND a.name = "Christian Bale";
```

For å finne filmer der regissør er "Christoffer Nolan" og skuespiller er "Christian Bale", må vi koble **movie** tabellen til **director** tabellen for å finne regissøren. Vi må også koble movie tabellen til actor tabellen for å finne skuespilleren.

SELECT m.title **AS** film, d.name **AS** regissor, a.name **AS** skuespiller **FROM** movie **AS** m. I select seksjon henter vi 3 kolonner. Første kolonnen er navnet på filmene. Det henter vi fra **movie** tabellen, movie.title og gir den alias film. Andre kolonnen er regissor hentet fra **director** tabellen, director.name og gir den alias regissor. Tredje kolonnen er skuespillernavn vi henter fra **actor** tabellen, actor.name og gir den alias skuespiller. Vi henter fra movie tabellen med alias m og joiner den videre med andre tabellene.

INNER JOIN director AS d

ON m.director_id = d.id. Vi inner joiner **movie** med **director** tabell samme som vi gjorde i forrige oppgaver. Movie tabell har foreign key til director tabell. Så vi joiner tabellene sammen ved å sammen foreign key i movie tabell til primary key i director tabell. Director tabell gir vi her alias d.

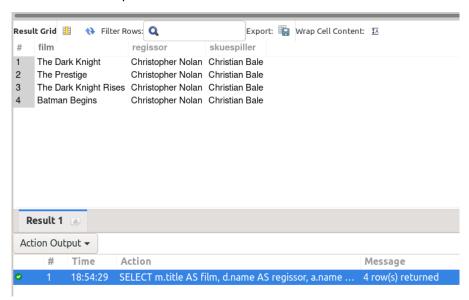
```
INNER JOIn movie_has_actor AS ma
ON ma.movie_id = m.id
INNER JOIN actor AS a
ON ma.actor_id = a.id. For å joine sammen movie tabell og actor tabell, joiner
vi via mellom tabellen movie_has_actor tabellen. Dette er 2 inner join.
```

I den første inner join kobler **movie_has_actor** tabellen til **movie** ved å sammenligne foreign key i movie_has_actor for movie tabell, movie_id med primorary i movie tabellen, movie.id.

I den andre inner join, kobler vi **actor** tabellen til **movie_has_actor** tabellen ved å sammenligne foreign key i movie_has_actor for actor, actor_id til primary key i actor tabellen, actor.id.

WHERE d.name = "Christopher Nolan" AND a.name = "Christian Bale";. Til slutt i WHERE clause begrenser vi data som blir hentet der director sin navn er Christoffer Nolan og der skuespiller sin navn er Christian Bale.

Vi får 4 treff i resultat. Dette er de 4 filmene som har Nolan som regissor og Bale som skuespiller.



Oppgave 10: Hent ut skuespillerne og antallet filmer de har spilt i, sortert synkende på antallet

SQL-spørring:

```
# 10: Hent ut skuespillerne og antallet filmer de har spilt i, sortert synkende
på antallet
SELECT a.name AS skuespiller, COUNT(m.title) AS filmer
FROM movie AS m
INNER JOIN movie_has_actor AS ma
ON ma.movie_id = m.id
INNER JOIN actor AS a
ON ma.actor_id = a.id
GROUP BY a.name
ORDER BY filmer DESC
;
```

Denne spørringen ligner på den fra oppgave 4 der vi skulle hente regissører og antall filmer de regisserte og sortere resulatatet synkende fra regissør med fleste filmer til den med minste. Her skal vi hente skuespillere og antall filmer de har vært med på og sortere resultatet synkende.

SELECT a.name **AS** skuespiller, **COUNT**(m.title) **AS** filmer **FROM** movie **AS** m. I SELECT clause henter vi 2 kolonner. Den første er fra **actor** tabellen og vi henter actor.title med alias skuespiller. I andre kolonnen COUNT() movie.title fra **movie** tabellen med alias filmer. COUNT() teller opp antall ganger skuespiller er assosiert med en film. Det gir oss tilbake numerisk tall. FROM movie, fra tabellen movie med alias m.

```
INNER JOIN movie_has_actor AS ma
ON ma.movie_id = m.id
INNER JOIN actor AS a
ON ma.actor_id = a.id.
```

For å joine movie tabellen med actor tabellen, bruker vi mellom tabllen movie_has_actor. Mellom tabllen danner many to many forholdet mellom movie og actor tabellene. Først INNER JOIN kobler movie_has_actor til movie tabellen ved å foreign key i movie_has_actor, movie_id til primary i movie, movie.id. I andre INNER JOIN blir actor tabellen koblet til movie_has_actor ved på samme måte, men denne til actor tabellen. Foreign key i movie_has_actor, actor_id blir sammenlignet med primary key i actor, actor.id.

```
GROUP BY a.name
ORDER BY filmer DESC
:
```

Til slutt grupperer vi data med GROUP BY clause med actor.name og sorterer også med filmer som er alias for COUNT(m.title). DESC gjør at det blir synkende, skuespillere som har spilt i flest filmer kommer høyst på resultat tabellen.

De niro, Hanks og Pacino er top 3 skuespillere med filmer spilt i denne databasen.

