

Nilay Kamar
BDM – Fall'19

Homework Assignment

Query 1: List the top 20 films that have the highest rating, their genres, and their start year.

Report these columns:

- Title of movie, startyear, genres, and rating
- Rating should be calculated with min 100000 votes.

Use the following tables:

- title_basics
- title_crew
- name_basics

Order by first startyear

%%sql

```
SELECT a.tconst, c.originaltitle, c.genres, c.startyear, b.rating
FROM title_crew a,
     (SELECT tconst, avg(averagerating) as rating
      FROM title_ratings
      WHERE numvotes > 1000000
      GROUP BY tconst
      ORDER BY rating DESC
      LIMIT 20) b,
     (SELECT tconst, genres, originaltitle, startyear
      FROM title_basics) c
WHERE b.tconst = a.tconst and
      b.tconst = c.tconst
ORDER BY c.startyear
```

OUTPUT

tconst	originaltitle	genres	startyear	rating
tt0068646	The Godfather	Crime,Drama	1972	9.2000000000000000
tt0076759	Star Wars	Action,Adventure,Fantasy	1977	8.7000000000000000
tt0109830	Forrest Gump	Comedy,Drama,Romance	1994	8.8000000000000000
tt0111161	The Shawshank Redemption	Crime,Drama	1994	9.3000000000000000
tt0110912	Pulp Fiction	Crime,Drama	1994	8.9000000000000000
tt0114369	Se7en	Crime,Drama,Mystery	1995	8.6000000000000000
tt0133093	The Matrix	Action,Sci-Fi	1999	8.7000000000000000
tt0137523	Fight Club	Drama	1999	8.8000000000000000
tt0172495	Gladiator	Action,Adventure,Drama	2000	8.5000000000000000
tt0120737	The Lord of the Rings: The Fellowship of the Ring	Adventure,Drama,Fantasy	2001	8.8000000000000000
tt0167261	The Lord of the Rings: The Two Towers	Action,Adventure,Drama	2002	8.7000000000000000
tt0167260	The Lord of the Rings: The Return of the King	Adventure,Drama,Fantasy	2003	8.9000000000000000
tt0372784	Batman Begins	Action,Adventure	2005	8.3000000000000000
tt0903747	Breaking Bad	Crime,Drama,Thriller	2008	9.5000000000000000
tt0468569	The Dark Knight	Action,Crime,Drama	2008	9.0000000000000000
tt1375666	Inception	Action,Adventure,Sci-Fi	2010	8.8000000000000000
tt0944947	Game of Thrones	Adventure,Drama,Fantasy	2011	9.5000000000000000
tt1345836	The Dark Knight Rises	Action,Thriller	2012	8.4000000000000000
tt1853728	Django Unchained	Drama,Western	2012	8.4000000000000000
tt0816692	Interstellar	Adventure,Drama,Sci-Fi	2014	8.6000000000000000

Query 2: List the actors of the highest rated top 20 films.

Report these columns:

- Title of movie, startyear, genres, ages of actors, and rating
- Rating should be calculated with min 100000 votes.

Use the following tables:

- title_basics
- name_basics
- crew_molten

Order by actor old desc

```

%%sql
select b.nconst, a.primaryname, a.primaryprofession, c.tconst, d.originaltitle, d.genres,
d.startyear, c.rating, (d.startyear - a.birthyear) as old
from name_basics a, crew_molten b,
(SELECT tconst, avg(averagerating) as rating
from title_ratings
WHERE numvotes > 1000000
GROUP BY tconst

```

ORDER BY rating DESC
 LIMIT 20) c,
 title_basics d
 where a.primaryprofession ~ '.*actor.*'
 and c.tconst = b.tconst
 and a.nconst = b.nconst
 and c.tconst = d.tconst
 order by old desc

nconst	primaryname	primaryprofession	tconst	originaltitle	genres	startyear	rating	old
nm0343165	Winston Groom	writer,actor	tt0109830	Forrest Gump	Comedy,Drama,Romance	1994	8.800000000000000	None
nm0801728	Stephen Sinclair	writer,director,actor	tt0167261	The Lord of the Rings: The Two Towers	Action,Adventure,Drama	2002	8.700000000000000	None
nm0354024	Charles Haid	actor,director,producer	tt0903747	Breaking Bad	Crime,Drama,Thriller	2008	9.500000000000000	65
nm0070474	Jack Bender	producer,director,actor	tt0944947	Game of Thrones	Adventure,Drama,Fantasy	2011	9.500000000000000	62
nm0186505	Bryan Cranston	actor,producer,director	tt0903747	Breaking Bad	Crime,Drama,Thriller	2008	9.500000000000000	52
nm0887700	Timothy Van Patten	director,actor,producer	tt0944947	Game of Thrones	Adventure,Drama,Fantasy	2011	9.500000000000000	52
nm0000233	Quentin Tarantino	writer,actor,producer	tt1853728	Django Unchained	Drama,Western	2012	8.400000000000000	49
nm0744839	Eric Roth	producer,writer,actor	tt0109830	Forrest Gump	Comedy,Drama,Romance	1994	8.800000000000000	49
nm0000233	Quentin Tarantino	writer,actor,producer	tt1853728	Django Unchained	Drama,Western	2012	8.400000000000000	49
nm0000175	Stephen King	writer,producer,actor	tt0111161	The Shawshank Redemption	Crime,Drama	1994	9.300000000000000	47
nm0880243	Jim Uhls	writer,producer,actor	tt0137523	Fight Club	Drama	1999	8.800000000000000	42
nm0000399	David Fincher	producer,director,actor	tt0137523	Fight Club	Drama	1999	8.800000000000000	37
nm0657333	Chuck Palahniuk	writer,actor,producer	tt0137523	Fight Club	Drama	1999	8.800000000000000	37
nm0787687	Matt Shakman	director,producer,actor	tt0944947	Game of Thrones	Adventure,Drama,Fantasy	2011	9.500000000000000	36
nm0000399	David Fincher	producer,director,actor	tt0114369	Se7en	Crime,Drama,Mystery	1995	8.600000000000000	33
nm0000233	Quentin Tarantino	writer,actor,producer	tt0110912	Pulp Fiction	Crime,Drama	1994	8.900000000000000	31
nm0001825	Andrew Kevin Walker	writer,actor,miscellaneous	tt0114369	Se7en	Crime,Drama,Mystery	1995	8.600000000000000	31
nm0000233	Quentin Tarantino	writer,actor,producer	tt0110912	Pulp Fiction	Crime,Drama	1994	8.900000000000000	31

Query 3: List the average ratings of the drama films on yearly basis and their year duration.

Report these columns:

- Title of movie, rating, runtimeminutes, and yearduration
- Rating should be calculated with min 50000 votes.
- Genre must be drama

Use the following tables:

- title_basics
- title_ratings

Order by rating desc, yearduration asc

%%sql

```

select tb.primarytitle, (tb.endyear-tb.startyear) as yearduration, tr.averagerating,
tb.runtimeminutes
from title_basics tb LEFT JOIN title_ratings tr
on tb.tconst = tr.tconst

```

```

where tb.genres ~'(?i)drama'
and tr.numvotes > 50000
and startyear IS NOT NULL

```

and endyear IS NOT NULL

group by tb.primarytitle, tb.endyear, tb.startyear, tb.runtimeminutes, tr.averagerating
order by averagerating desc, yearduration asc
LIMIT 50

primarytitle	yearduration	rating	runtime	minutes
Band of Brothers	0	9.5000000000000000		594
Breaking Bad	5	9.5000000000000000		49
The Wire	6	9.3000000000000000		59
Leyla and Mecnun	2	9.2000000000000000		80
The Sopranos	8	9.2000000000000000		55
Firefly	1	9.1000000000000000		44
Death Note	1	9.0000000000000000		24
Twin Peaks	1	8.9000000000000000		47
Freaks and Geeks	1	8.9000000000000000		44
Rome	2	8.8000000000000000		90
Oz	6	8.8000000000000000		55
House	8	8.8000000000000000		44
Deadwood	2	8.7000000000000000		55
Six Feet Under	4	8.7000000000000000		55
Battlestar Galactica	5	8.7000000000000000		44
Downton Abbey	5	8.7000000000000000		58
The Shield	6	8.7000000000000000		47
Dexter	7	8.7000000000000000		53
The Night Of	0	8.6000000000000000		60
Hannibal	2	8.6000000000000000		44
The Office	2	8.6000000000000000		30
The Newsroom	2	8.6000000000000000		55
Boardwalk Empire	4	8.6000000000000000		55
Justified	5	8.6000000000000000		44
Sons of Anarchy	6	8.6000000000000000		45
Luther	8	8.6000000000000000		60
Mad Men	8	8.6000000000000000		47
Person of Interest	5	8.5000000000000000		43
Entourage	7	8.5000000000000000		28
Banshee	3	8.4000000000000000		60
Sense8	3	8.4000000000000000		60

Query 4: List the duration of the 50 movies with the highest rating in comedy.

Report these columns:

- Title of movie, genres, startyear, rating, and runtimeminutes
- Rating should be calculated with min 50000 votes.
- Genre must be comedy

Use the following tables:

- title_basics
- title_ratings

Order by runtimeminutes asc

```
select a.tconst, a.primarytitle, a.genres, a.startyear, avg(b.averagerating) as rating,  
a.runtimeminutes  
from title_basics a RIGHT JOIN title_ratings b  
on a.tconst = b.tconst  
and a.genres ~ '(?i)Comedy'  
and b.numvotes > 50000  
group by a.tconst, a.startyear  
order by rating desc, runtimeminutes asc  
LIMIT 50
```

Output:

tconst	primarytitle	genres	startyear	rating	runtime
tt2861424	Rick and Morty	Adventure,Animation,Comedy	2013	9.400000000000000	22
tt1831164	Leyla and Mecnun	Adventure,Comedy,Drama	2011	9.200000000000000	80
tt3530232	Last Week Tonight with John Oliver	Comedy,News,Talk-Show	2014	9.100000000000000	30
tt4508902	One Punch Man	Action,Animation,Comedy	2015	9.000000000000000	24
tt0367279	Arrested Development	Comedy	2003	8.900000000000000	22
tt0108778	Friends	Comedy,Romance	1994	8.900000000000000	22
tt0098904	Seinfeld	Comedy	1989	8.900000000000000	22
tt0063929	Monty Python's Flying Circus	Comedy	1969	8.900000000000000	30
tt0193676	Freaks and Geeks	Comedy,Drama	1999	8.900000000000000	44
tt0386676	The Office	Comedy	2005	8.800000000000000	22
tt0096697	The Simpsons	Animation,Comedy	1989	8.800000000000000	22
tt0121955	South Park	Animation,Comedy	1997	8.800000000000000	22
tt0472954	It's Always Sunny in Philadelphia	Comedy	2005	8.800000000000000	22
tt0072500	Fawlty Towers	Comedy	1975	8.800000000000000	30
tt0109830	Forrest Gump	Comedy,Drama,Romance	1994	8.800000000000000	142
tt1486217	Archer	Action,Animation,Comedy	2009	8.700000000000000	22
tt0264235	Curb Your Enthusiasm	Comedy	2000	8.700000000000000	28
tt1586680	Shameless	Comedy,Drama	2011	8.700000000000000	46
tt0248654	Six Feet Under	Comedy,Drama	2001	8.700000000000000	55
tt1628033	Top Gear	Comedy,Talk-Show	2002	8.700000000000000	60
tt1305826	Adventure Time	Adventure,Animation,Comedy	2010	8.600000000000000	11
tt1492966	Louie	Comedy,Drama	2010	8.600000000000000	22
tt1266020	Parks and Recreation	Comedy	2009	8.600000000000000	22
tt0487831	The IT Crowd	Comedy	2006	8.600000000000000	25
tt2575988	Silicon Valley	Comedy	2014	8.600000000000000	28
tt0290978	The Office	Comedy,Drama	2001	8.600000000000000	30
tt1632701	Suits	Comedy,Drama	2011	8.600000000000000	44
tt0021749	City Lights	Comedy,Drama,Romance	1931	8.600000000000000	87
tt1675434	The Intouchables	Biography,Comedy,Drama	2011	8.600000000000000	112
tt0118799	Life Is Beautiful	Comedy,Drama,War	1997	8.600000000000000	116
tt1439629	Community	Comedy	2009	8.500000000000000	22