1. List all the directors who directed a 'Comedy' movie in a leap year. (You need to check that the genre is 'Comedy’ and year is a leap year) Your query should return director name, the movie name, and the year.

* WITH CTE\_table as

(SELECT new.title ,

md.PID ,new.year

FROM M\_Director md

JOIN

(SELECT MID, title,year FROM movie m WHERE MID IN

(SELECT MID FROM M\_Genre WHERE GID IN

(SELECT GID FROM GENRE WHERE Name LIKE '%Comedy%')) AND ((((m.year%4)=0 AND (m.year%100)!=0) OR ((m.year%400)=0)))) new

ON md.MID=new.MID)

SELECT cte.title, p.Name, cte.year FROM cte\_table cte JOIN person p ON cte.PID= p.PID ORDER BY 3

1. List the names of all the actors who played in the movie 'Anand' (1971)

* SELECT Name FROM Person WHERE PID IN

(SELECT TRIM(PID) FROM M\_Cast WHERE MID=

(SELECT MID FROM movie WHERE title="Anand"))

1. List all the actors who acted in a film before 1970 and in a film after 1990. (That is: < 1970 and > 1990.)

* SELECT name FROM person WHERE PID IN

(SELECT TRIM(PID) FROM M\_Cast WHERE MID IN

(SELECT MID FROM Movie WHERE year<1970)

INTERSECT

SELECT TRIM(PID) FROM M\_Cast WHERE MID IN

(SELECT MID FROM Movie WHERE year>1990))

1. List all directors who directed 10 movies or more, in descending order of the number of movies they directed. Return the directors' names and the number of movies each of them directed.

* SELECT movie\_count.number ,p.name from person p inner join

(SELECT count(m.MID) Number, md.PID from movie m inner join M\_Director md on m.MID=md.MID GROUP BY md.PID) as movie\_count

ON movie\_count.PID=p.PID WHERE movie\_count.number>=10 ORDER BY movie\_count.number DESC

5 a. For each year, count the number of movies in that year that had only female actors.

*b.* Now include a small change: report for each year the percentage of movies in that year with only female actors, and the total number of movies made that year. For example, one answer will be: *1990 31.81 13522* meaning that in 1990 there were 13,522 movies, and 31.81% had only female actors. You do not need to round your answer.

* SELECT SUBSTR(year,-4) YEAR, COUNT(MID) FROM MOVIE WHERE MID IN(

SELECT MID FROM (

SELECT MID, SUM(CASE WHEN gender='Male' THEN Counts END) AS Male,

SUM(CASE WHEN gender='Female' THEN Counts END) AS Female,

SUM(CASE WHEN Gender IS NULL THEN Counts END) AS Nulls

FROM(

SELECT DISTINCT MID,gender, COUNT(\*) Counts from

(SELECT mv1.MID, ps1.Gender, ps1.PID FROM Movie mv1

JOIN M\_Cast mc1

ON mv1.MID = mc1.MID

JOIN Person ps1

ON (mc1.PID) = ps1.PID)

GROUP BY MID, gender

) GROUP BY MID

) WHERE Male is null

) GROUP BY year

* SELECT fem.years, tot.Total\_count, (CAST(fem.counts as real)/cast(tot.Total\_count as real)\*100) Percentage\_Female FROM

(

SELECT SUBSTR(year,-4) AS years, COUNT(MID) as counts

FROM MOVIE WHERE MID IN(

SELECT MID FROM (

SELECT MID, SUM(CASE WHEN gender='Male' THEN Counts END) AS Male,

SUM(CASE WHEN gender='Female' THEN Counts END) AS Female,

SUM(CASE WHEN Gender IS NULL THEN Counts END) AS Nulls

FROM(

SELECT DISTINCT MID,gender, COUNT(\*) Counts FROM

(SELECT mv1.MID, ps1.Gender, ps1.PID FROM MOVIE mv1

JOIN M\_Cast mc1

ON mv1.MID = mc1.MID

JOIN Person ps1

ON (mc1.PID) = ps1.PID)

GROUP BY MID, gender

) GROUP BY MID

) WHERE Male IS NULL

) GROUP BY year)fem JOIN

(SELECT YEAR,COUNT(MID) total\_count FROM MOVIE GROUP BY YEAR) tot ON fem.years=tot.year

1. Find the film(s) with the largest cast. Return the movie title and the size of the cast. By "cast size" we mean the number of distinct actors that played in that movie: if an actor played multiple roles, or if it simply occurs multiple times in casts, we still count her/him only once.

* SELECT m.title, size.cast\_size FROM movie m JOIN

(SELECT DISTINCT MID AS movies, COUNT(DISTINCT PID) AS cast\_size FROM M\_Cast GROUP BY movies) size ON m.MID=size.movies ORDER BY 2 DESC

1. A decade is a sequence of 10 consecutive years. For example, say in your database you have movie information starting from 1965. Then the first decade is 1965, 1966, ..., 1974; the second one is 1967, 1968, ..., 1976 and so on. Find the decade D with the largest number of films and the total number of films in D.

* WITH CTE1 AS(

SELECT m1.year AS year1,m1.count AS count1, m2.year AS year2,m2.count AS count2 FROM temp\_table2 m1 , temp\_table2 m2 WHERE ((m2.year<=m1.year+9) AND (m2.year>=m1.year))

)SELECT year1, SUM(count2) FROM CTE1 GROUP BY year1 ORDER BY 2 DESC LIMIT 1

1. Find the actors that were never unemployed for more than 3 years at a stretch. (Assume that the actors remain unemployed between two consecutive movies).
2. Find all the actors that made more movies with Yash Chopra than any other director.

* SELECT y.Yash, y.Actor FROM(

SELECT COUNT(m.MID) as Yash, d.PID Director, C.PID Actor FROM movie m JOIN M\_Director d ON m.MID=d.MID JOIN M\_Cast c ON c.MID=m.MID WHERE Director='nm0007181' GROUP BY director, actor

) y LEFT JOIN(

SELECT DISTINCT ACTOR, max(COUNTS) OVER(PARTITION BY Actor) AS final\_count FROM(

SELECT COUNT(m.MID) as Counts, d.PID Director, C.PID Actor FROM movie m JOIN M\_Director d ON m.MID=d.MID JOIN M\_Cast c ON c.MID=m.MID GROUP BY director, actor

) ORDER BY 2 desc) o ON y.Actor=o.Actor WHERE yash>final\_count OR yash=final\_count ORDER BY 1 DESC

1. The Shahrukh number of an actor is the length of the shortest path between the actor and Shahrukh Khan in the "co-acting" graph. That is, Shahrukh Khan has Shahrukh number 0; all actors who acted in the same film as Shahrukh have Shahrukh number 1; all actors who acted in the same film as some actor with Shahrukh number 1 have Shahrukh number 2, etc. Return all actors whose Shahrukh number is 2.

* SELECT NAME FROM Person WHERE PID IN (

SELECT TRIM(PID) FROM M\_Cast WHERE (MID) IN (

SELECT MID FROM M\_Cast WHERE (PID) IN(

SELECT pid FROM m\_cast WHERE trim(MID) in(

SELECT MID FROM M\_Cast WHERE trim(PID)='nm0451321')))

EXCEPT

SELECT TRIM(pid) FROM m\_cast WHERE trim(MID) in(

SELECT MID FROM M\_Cast WHERE trim(PID)='nm0451321'))