Database Systems Class Project Mid-Eval.

CSCI 3700/6070 Database Systems Spring 2020

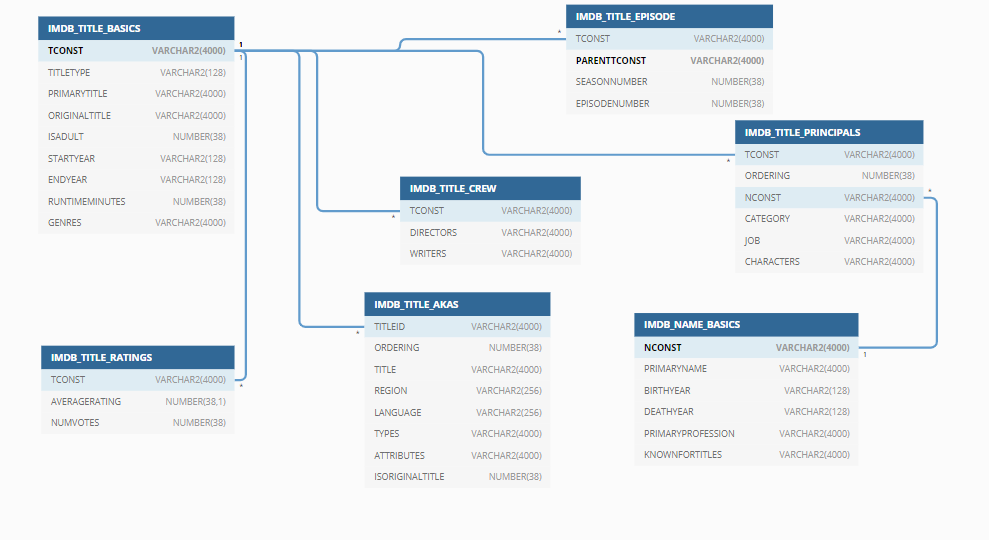
# Which database system are you using?

ORACLE

# What programming language are you proposing to use?

JAVA

# How do you design your database schemas?



# Individual tasks

|  |  |
| --- | --- |
| Student Name | Major tasks |
| Shashank Palusa | Setting up the environment for the project i.e. Downloading the given datasets for the project from online database, installing SQL developer and oracle database. |
| Venkata Ganesh Gutlapalli | Converting the downloaded datasets which are in (TSV) formatted files to (CSV) formatted files, and handling the null values during importing using python |
| Kundhana Reddy Aavula Bala | Designing the Database Schemas to logically group tables, views and stored procedures and helping in importing data to the database using SQL developer. |
| Thaneesh Reddy Veluru | Understanding the imported data in the database and its hierarchy and designing the required queries for the questions in project.  Documentation for submitting document for the group project first deliverable. |
| Krishna Chaitanya Kandimalla | Importing the converted (CSV) files using SQL developer and helping in Designing the queries to the given questions including the bonus query and executing them with screenshots. |

# Query design (Section 2.3)

* 1. Show the information of movies (Title, Region, Language, Year, Rating, Votes) which have average rating greater than 7 and number of votes greater than 1000.

Statement:

SELECT

ita.title Title,

ita.region Region,

ita.language Language,

itb.startyear Year,

itr.averagerating Rating,

itr.numvotes Votes

FROM

imdb\_title\_akas ita,

imdb\_title\_basics itb,

imdb\_title\_ratings itr

WHERE

1 = 1

AND ita.titleid = itb.tconst

AND itb.tconst = itr.tconst

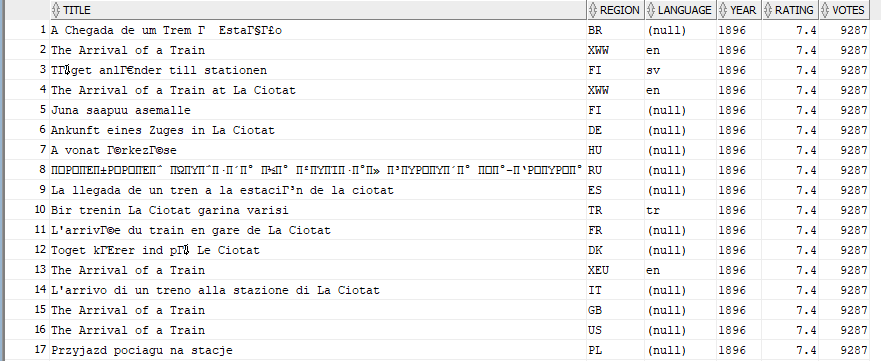
AND ita.titleid=itr.tconst

AND itr.averagerating > 7

AND itr.numvotes>1000;

Screenshot:

For this Query, there are 28604 rows are retrieved. Below screenshot shows the first 17 rows.



* 1. Show the names of persons who have directed more than 10 movies.

Statement:

CREATE VIEW crew as WITH

model\_param AS

(

SELECT tconst,

directors AS orig\_str ,

','

|| directors

|| ',' AS mod\_str ,

1 AS start\_pos ,

Length(directors) AS end\_pos ,

(Length(directors) - Length(Replace(directors, ','))) + 1 AS element\_count ,

0 AS element\_no ,

ROWNUM AS rn

FROM imdb\_title\_crew )

SELECT tconst,

trim(Substr(mod\_str, start\_pos, end\_pos-start\_pos)) directors

FROM (

SELECT \*

FROM model\_param MODEL PARTITION BY (tconst, rn, orig\_str, mod\_str)

DIMENSION BY (element\_no)

MEASURES (start\_pos, end\_pos, element\_count)

RULES ITERATE (2000)

UNTIL (ITERATION\_NUMBER+1 = element\_count[0])

( start\_pos[ITERATION\_NUMBER+1] = instr(cv(mod\_str), ',', 1, cv(element\_no)) + 1,

end\_pos[iteration\_number+1] = instr(cv(mod\_str), ',', 1, cv(element\_no) + 1) )

)

WHERE element\_no != 0

ORDER BY tconst;

SELECT

inb.primaryname name

FROM

imdb\_name\_basics inb

WHERE

inb.nconst IN (

SELECT

c.directors

FROM

crew c,

imdb\_name\_basics inb

WHERE

c.directors = inb.nconst

GROUP BY

c.directors

HAVING

COUNT(c.tconst) > 10

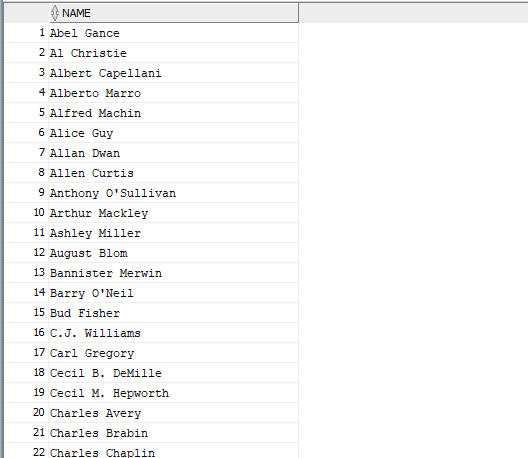
)

ORDER BY

inb.primaryname;

Screenshot:

For this Query, there are 137 rows are retrieved. Below screenshot shows the first 22 rows.



* 1. Show the top 10 TV series (more than 1 season) which most average ratings.

Statement:

SELECT

title,

seasonnumber,

averagerating

FROM

(

SELECT

ita.title,

ite.seasonnumber,

itr.averagerating

FROM

imdb\_title\_akas ita,

imdb\_title\_episode ite,

imdb\_title\_ratings itr

WHERE

ita.titleid = ite.tconst

AND ite.tconst = itr.tconst

AND ite.seasonnumber > 1

ORDER BY

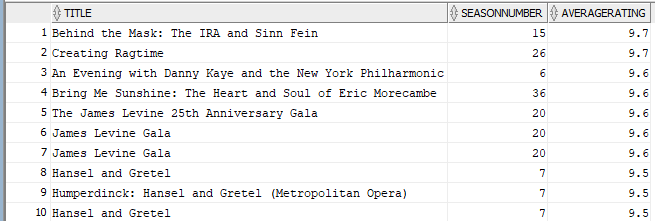
itr.averagerating DESC

)

WHERE

ROWNUM <= 10;

Screenshot:



* 1. You need to design at least FOUR more queries which satisfy
     1. Use aggregate functions, “group by” and “order by”.
     2. Use set operations.
     3. Use nested/sub queries (either in SELECT clause, FROM clause or WHERE clause).
     4. Use JOIN operations.

d(i). Show the number of movies ( count ) for a particular rating.

Statement:

SELECT

itr.averagerating Rating,

COUNT(ita.title) MoviesCount

FROM

imdb\_title\_akas ita,

imdb\_title\_ratings itr

WHERE

ita.titleid = itr.tconst

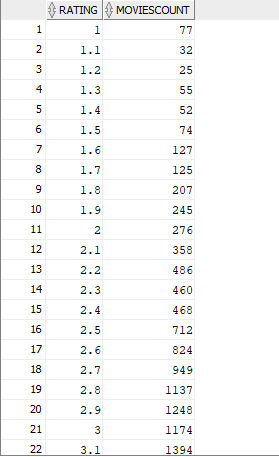
GROUP BY

itr.averagerating

ORDER BY itr.averagerating;

Screenshot:

For this Query, there are 91 rows are retrieved. Below screenshot shows the first 22 rows.



d(ii). Show the information of TV Series (Title, Types, Is Adult, Start Year, Season Number, Episode Number, Votes) which have episode number is NOT NULL.

Statement:

SELECT

ita.title Title,

ita.types Types,

itb.isadult IsAdult,

itb.startyear StartYear,

ite.seasonnumber SeasonNumber,

ite.episodenumber EpisodeNumber,

itr.numvotes Votes

FROM

imdb\_title\_akas ita,

imdb\_title\_basics itb,

imdb\_title\_episode ite,

imdb\_title\_ratings itr

WHERE

ita.titleid = itb.tconst

and itb.tconst= ite.tconst

and ita.titleid=itr.tconst

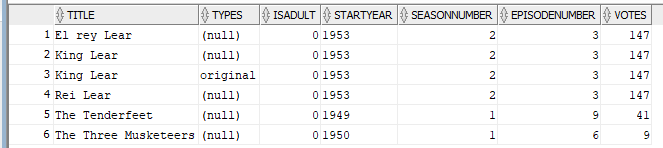
and itb.tconst=itr.tconst

and ite.tconst=itr.tconst

and ite.episodenumber IS NOT NULL

order by title;

Screenshot:



d(iii). Show the information of Movies (Title, Types, RunTimeMinutes) which have RunTimeMinutes=1

Statement:

SELECT

ita.title Title,

ita.types Types,

itb.runtimeminutes RunTimeMinutes

FROM

imdb\_title\_akas ita,

imdb\_title\_basics itb

WHERE

(ita.titleid,ita.ordering) IN (

SELECT

titleid,ordering

FROM

imdb\_title\_akas

WHERE

lower(types)='original'

)

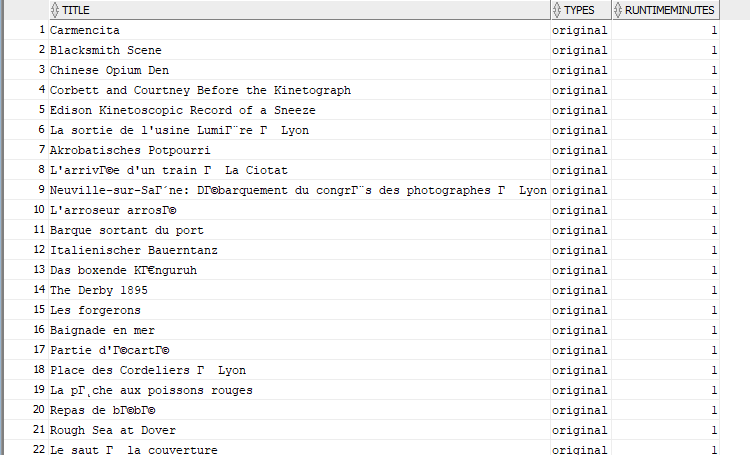
and ita.titleid=itb.tconst

and itb.runtimeminutes=1

order by itb.runtimeminutes;

Screenshot:

For this Query, there are 87 rows are retrieved. Below screenshot shows the first 22 rows.



d(iv). Show the information of Movies (Title) which are directed by 'William K.L. Dickson' and average rating more than 3.

SELECT

title

FROM

(

SELECT

itc.tconst,

ita.title

FROM

imdb\_title\_crew itc,

imdb\_title\_akas ita,

imdb\_name\_basics inb

WHERE

itc.tconst = ita.titleid

AND itc.directors = inb.nconst

AND inb.primaryname = 'William K.L. Dickson'

INTERSECT

SELECT

itr.tconst,

ita.title

FROM

imdb\_title\_ratings itr,

imdb\_title\_akas ita

WHERE

itr.tconst = ita.titleid

AND itr.averagerating > 3

);

Screenshot:

For this Query, there are 47 rows are retrieved. Below screenshot shows the first 22 rows.

