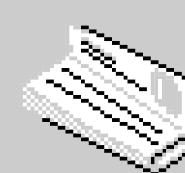
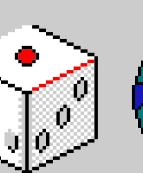
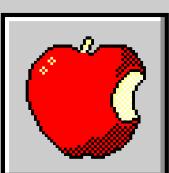


# ITSS 4300

# GROUP PROJECT



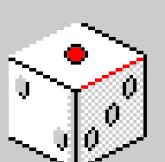
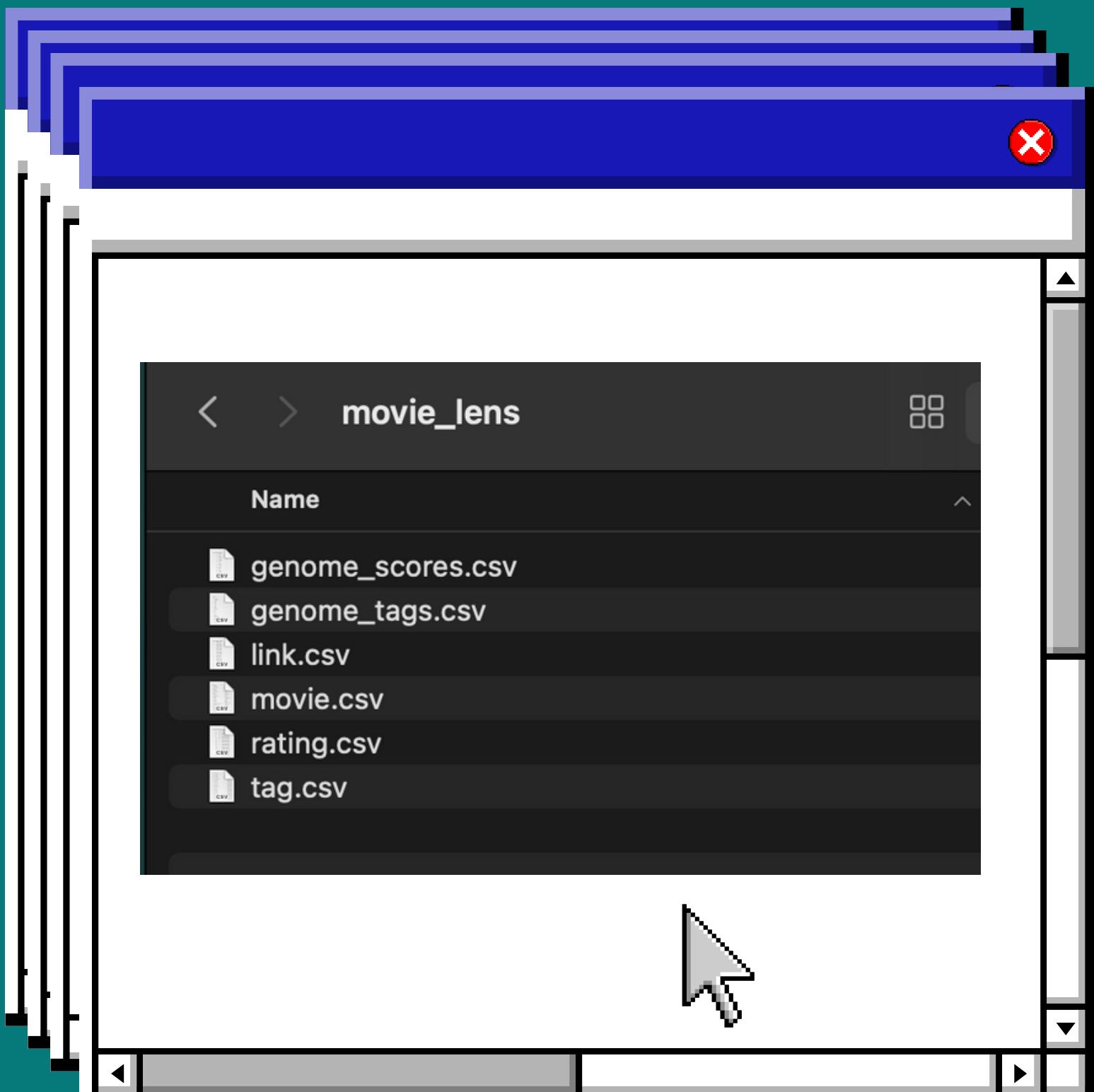
By: Jyostna Thanjavur, Humsini  
Revuru, Alaya Sirigiri, Dante  
Gutierrez-Tomlinson



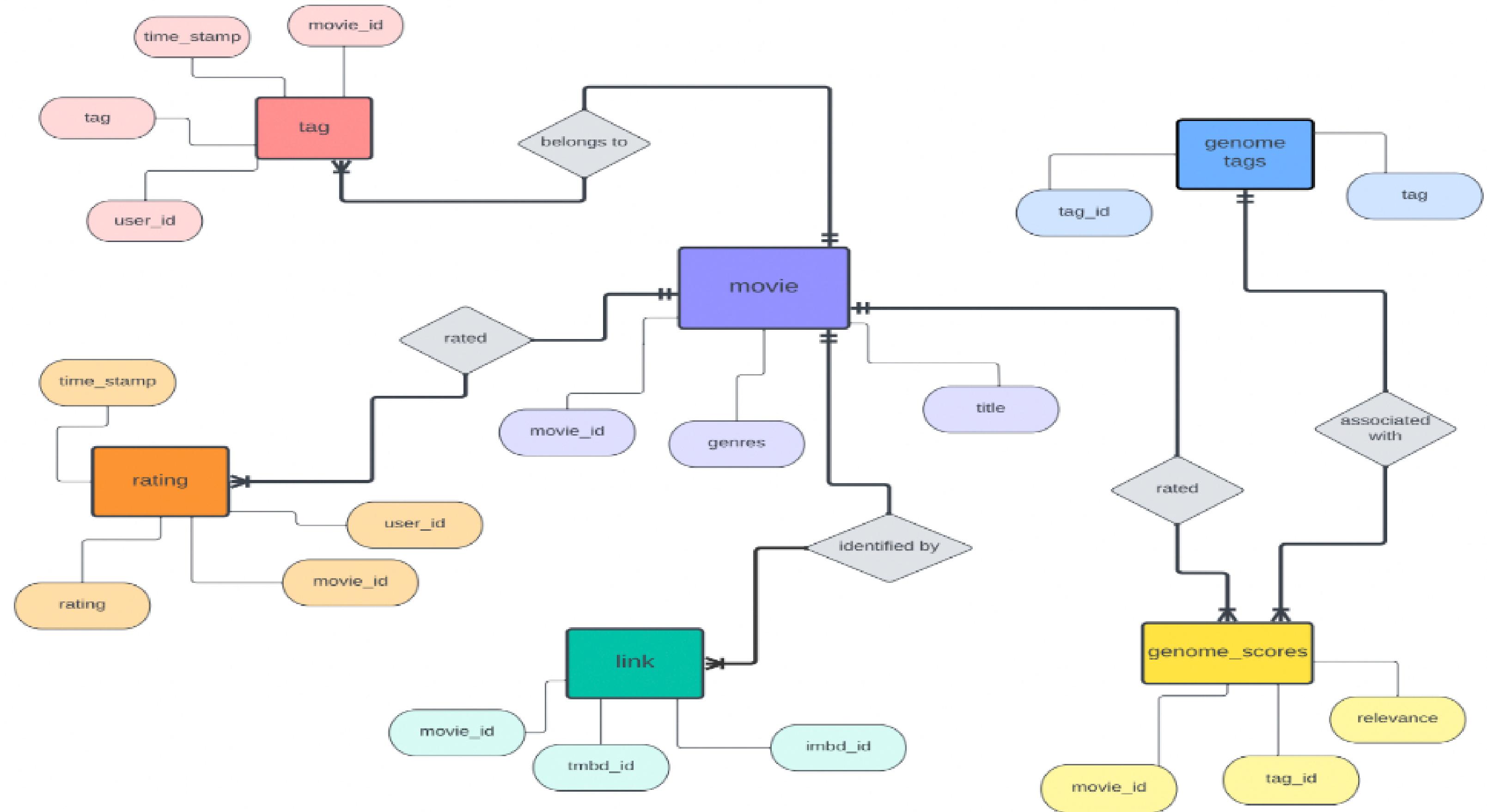
11:11PM

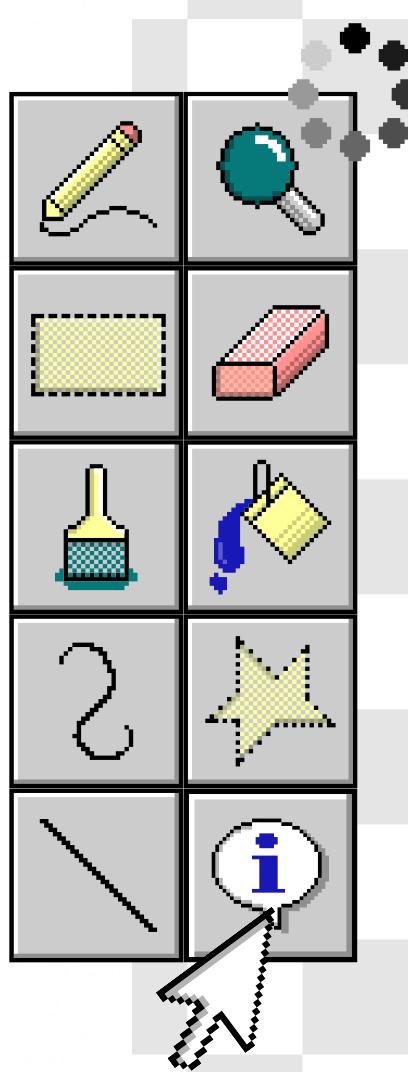
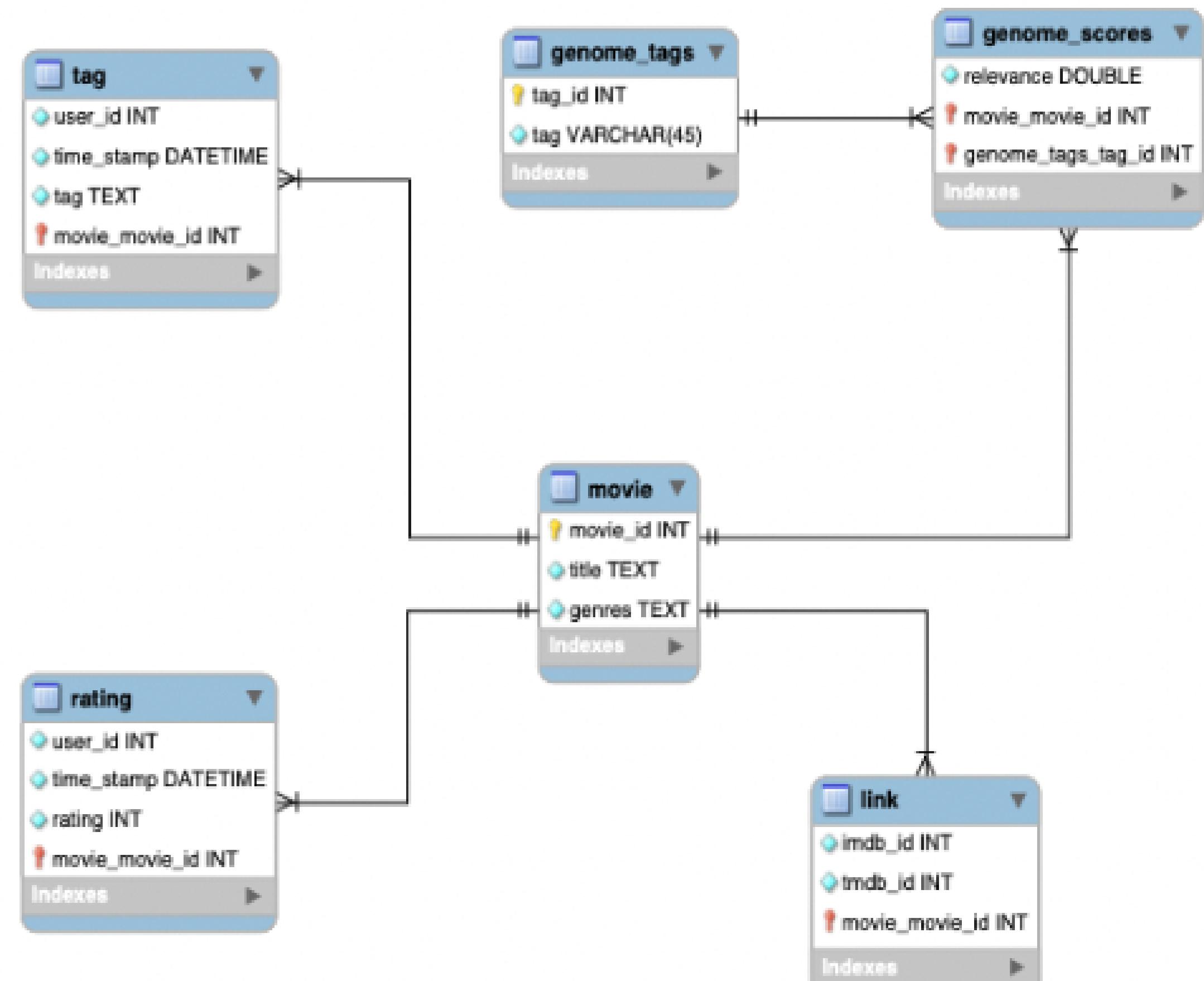
# Dataset Description

- dataset of 6 tables describes rating and tagging activities from MovieLens, a movie recommendation service
- contains 20,000,263 ratings and 465,564 tag applications made by 138,493 users across 27,278 movies
- tag.csv contains tags applied to movies by users
- rating.csv has ratings of movies by users
- movie.csv has movie title and genre
- link.csv that contains identifiers that can be used to link to other sources
- genome\_scores.csv that contains movie-tag relevance data
- genome\_tags.csv contains tag descriptions



[Back to Home Page](#)





# Data Loading

The screenshot shows the MySQL Workbench interface. On the left, the 'Schemas' tree view is open, showing the 'movie\_analytics' schema with its tables: genome\_scores, genome\_tags, link, movie, rating, and tag. Below the schema tree, there's a list of other databases: ap, ex, movie\_analytics, om, payroll, soccerDB, students\_db, and sys. The main workspace contains a SQL editor window with the following code:

```
1 • use movie_analytics;
2
3 • select * from genome_scores;
4 • select * from genome_tags;
5 • select * from link;
6 • select * from movie;
7 • select * from rating;
8 • select * from tag;
9
10
11
12
13
14
15
```

Below the code is a 'Result Grid' showing data from the 'tag' table:

	userId	moviedb	tag	timestamp
▶	18	4141	Mark Waters	2009-04-24 18:19:40
65	208	dark hero	2013-05-10 01:41:18	
65	353	dark hero	2013-05-10 01:41:19	
65	521	noir thriller	2013-05-10 01:39:43	
65	592	dark hero	2013-05-10 01:41:18	
65	668	bollywood	2013-05-10 01:37:56	
65	898	screwball comedy	2013-05-10 01:42:40	
65	1248	noir thriller	2013-05-10 01:39:43	
65	1391	mars	2013-05-10 01:40:55	

At the bottom, the 'Action Output' section shows the execution history:

Action	Time	Response	Duration / Fetch Time
use movie_analytics	09:00:50	0 row(s) affected	0.00019 sec
select * from genome_scores	09:00:52	71662 row(s) returned	0.017 sec / 0.053 sec
select * from genome_tags	09:00:54	1128 row(s) returned	0.00087 sec / 0.0005...
select * from link	09:00:56	68868 row(s) returned	0.0043 sec / 0.046 sec
select * from movie	09:00:57	27937 row(s) returned	0.0012 sec / 0.022 sec
select * from rating	09:00:59	885474 row(s) returned	0.00065 sec / 0.792...
select * from tag	09:01:03	549 row(s) returned	0.0011 sec / 0.00019...

At the very bottom, it says 'Query Completed'.

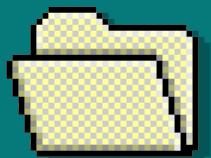
The screenshot shows the MySQL Workbench interface with a 'Loading Script' window. The script is:

```
LOAD DATA LOCAL INFILE '/Users/alaya/downloads/archive/movie.csv' INTO TABLE movie_analytics.movie  
FIELDS TERMINATED BY ','  
ENCLOSED BY ""  
LINES TERMINATED BY '\r\n'  
IGNORE 1 LINES  
(movield, title, genres);
```

To the right of the script is a yellow lightbulb icon. To the right of the lightbulb is a large text area with the heading 'Cleaning the 'movie.csv'' and the following text:

We had to clean the data in the movie.csv file (titles) to remove some encoded characters in order to import using the workbench's Table Data Import Wizard.

# Insights into the Data: #1



How many movies fall into each genre (in descending order)?

Local instance 3306

Administration Schemas testers for group\* Untitled recc report\* SQL File 5\*

SCHEMAS  
Filter objects  
ap ex movie\_analytics  
Tables genome\_scor... genome\_tags link movie rating  
Columns userid movied rating timestamp  
Indexes  
Foreign Keys  
Triggers tag Views Stored Procedures

```
1 • use movie_analytics;
2
3 /*How many movies fall into each genre (in descending order)?*/
4
5
6 • select genres, count(*) as count
7   from movie m, rating r
8   where m.movieid = r.movieid
9   group by genres
10  order by Count desc;
```

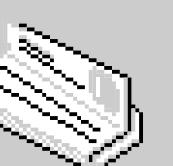
Result Grid Filter Rows: Search Export: Fetch rows: Result Grid Form Editor Read Only

genres	count
Drama	64255
Comedy	59469
Comedy Romance	35826
Comedy Drama	28893
Drama Romance	28405
Comedy Drama Romance	27322
Crime Drama	20189
Action Adventure Sci-Fi	19457
Action Adventure Thriller	14033

Action Output

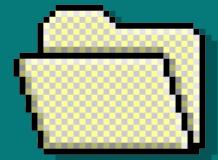
Time	Action	Response
75	SELECT distinct title, count(*) as COUNT FROM movie GROUP BY title	10 row(s) returned
76	SELECT distinct genre FROM movie GROUP BY genre order by...	Error Code: 1054. Unknown column 'genre' in 'field list'
77	SELECT distinct genre, count(*) as count FROM movie GROUP...	Error Code: 1054. Unknown column 'genre' in 'field list'
78	SELECT genre, count(*) as count FROM movie GROUP BY gen...	Error Code: 1054. Unknown column 'genre' in 'field list'
79	SELECT genres, count(*) as count FROM movie GROUP BY ge...	10 row(s) returned
80	SELECT genres, count(*) as count FROM movie GROUP BY ge...	1343 row(s) returned
81	SELECT genres, count(*) as count FROM movie GROUP BY ge...	1343 row(s) returned
82	select genres, avg(r.rating), count(*) as count from movie m, r...	1039 row(s) returned
83	select genres, avg(r.rating), round(count(*),2) as count from m...	1039 row(s) returned
84	select genres, avg(r.rating), (round(count(*),2)) as count from...	1039 row(s) returned
85	select genres, round(avg(r.rating),2), count(*) as count from m...	1039 row(s) returned
86	select genres, count(*) as count from movie m, rating r where...	1039 row(s) returned

Query Completed



[Back to Home Page](#)

# Insights into the Data: #2



What is the average user rating for each genre of movies?

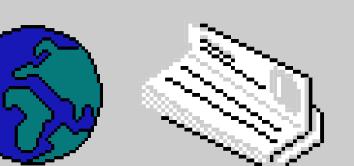
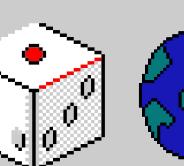
The screenshot shows a MySQL Workbench interface with the following details:

- Query Editor:** Contains the following SQL code:

```
12
13  /*What is the average user rating for each genre of movies? */
14
15 • select genres, round(avg(r.rating),2) as avg_rating, count(*) as count
16    from movie m, rating r
17   where m.movieid = r.movieid
18   group by genres
19   order by count desc;
```
- Result Grid:** Displays the results of the query in a table format:

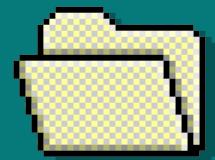
genres	avg_rating	count
Drama	3.68	64255
Comedy	3.22	59469
ComedyRomance	3.37	35826
ComedyDrama	3.59	28893
DramaRomance	3.63	28405
ComedyDramaRomance	3.61	27322
CrimeDrama	4	20189
ActionAdventureSci-Fi	3.53	19457
ActionAdventureThriller	3.39	14033
- Action Output:** Shows a log of actions taken during the query execution, including the original error and its resolution.

Time	Action	Response
70	SELECT genres, count(*) as count FROM movie GROUP BY genre...	Error Code: 1054. Unknown column 'genre' in 'field list'
79	SELECT genres, count(*) as count FROM movie GROUP BY ge...	10 row(s) returned
80	SELECT genres, count(*) as count FROM movie GROUP BY ge...	1343 row(s) returned
81	SELECT genres, count(*) as count FROM movie GROUP BY ge...	1343 row(s) returned
82	select genres, avg(r.rating), count(*) as count from movie m, r...	1039 row(s) returned
83	select genres, avg(r.rating), round(count(*),2) as count from m...	1039 row(s) returned
84	select genres, avg(r.rating), (round(count(*),2)) as count from...	1039 row(s) returned
85	select genres, round(avg(r.rating),2), count(*) as count from m...	1039 row(s) returned
86	select genres, count(*) as count from movie m, rating r where...	1039 row(s) returned
87	select genres, count(*) as count from movie m, rating r where...	1039 row(s) returned
88	select genres, round(avg(r.rating),2), count(*) as count from m...	1039 row(s) returned
89	select genres, round(avg(r.rating),2) as avg_rating, count(*) a...	1039 row(s) returned



[Back to Home Page](#)

# Insights into the Data #3



How many movies fall under each rating?

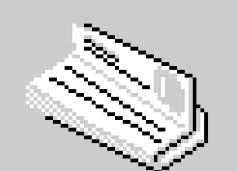
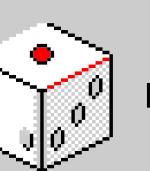
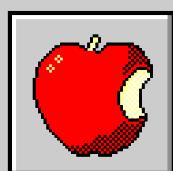
The screenshot shows a MySQL Workbench interface with a dark theme. In the top-left, the 'Schemas' tree shows a schema named 'movie\_analytics' containing tables like 'genome\_scores', 'genome\_tags', 'link', 'movie', and 'rating'. The 'rating' table is selected, and its columns (userid, movieid, rating, timestamp) are visible. A query is run in the SQL editor:

```
21 /*How many movies fall under each rating?*/
22
23 • select rating, count(m.movieid) as count
24   from movie m, rating r
25  where m.movieid = r.movieid
26  group by rating
27  order by rating desc, Count desc;
28
29
30
```

The results are displayed in a grid:

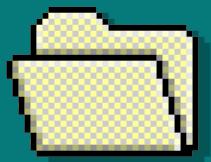
rating	count
5	126334
4.5	67087
4	247772
3.5	96672
3	192843
2.5	38065
2	63224
1.5	11957
1	30942

The 'Action Output' pane shows the query was executed at 21:19:12 and returned 10 rows.



[Back to Home Page](#)

# Insights into the Data: #4



What are the top 10 films that have the most number of user ratings?

Local instance 3306

Administration Schemas

SCHEMAS Filter objects

ap ex movie\_analytics

Tables genome\_scoring genome\_tags link movie rating

Columns userid movielid rating timestamp

Indexes

Foreign Keys

Triggers

tag

Views

Stored Procedures

Object Info Session

Column: rating

Definition: rating double

Query Completed

```
29 /*What are the top 10 films that have the most number of user ratings?*/
30
31 • select m.movieid, m.title, m.genres, count(r.rating) as count
32 from movie m, rating r
33 where m.movieid = r.movieid
34 group by m.movieid, m.title, m.genres
35 order by count desc
36 limit 10;
```

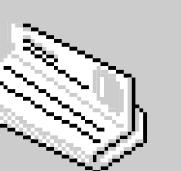
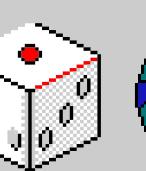
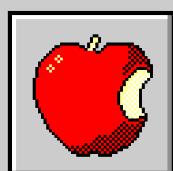
Result Grid Filter Rows: Search Export:

movieid	title	genres	count
296	Pulp Fiction (1994)	Comedy Crime Drama Thriller	2932
356	Forrest Gump (1994)	Comedy Drama Romance War	2908
593	Silence of the Lambs, The (1991)	Crime Horror Thriller	2704
318	Shawshank Redemption, The (1994)	Crime Drama	2686
480	Jurassic Park (1993)	Action Adventure Sci-Fi Thriller	2598
260	Star Wars: Episode IV - A New Hope (1977)	Action Adventure Sci-Fi	2391
110	Braveheart (1995)	Action Drama War	2316
589	Terminator 2: Judgment Day (1991)	Action Sci-Fi	2242
2571	Matrix, The (1999)	Action Sci-Fi Thriller	2228

Result 13 Read Only

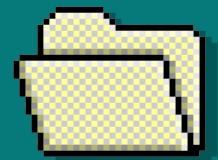
Action Output

Time	Action	Response
21:19:12	select rating, count(m.movieid) as count from movie m, rating...	10 row(s) returned
21:23:09	select m.movieid, m.title, count(r.rating) as count from movie...	10 row(s) returned
21:23:59	select m.movieid, m.title, m.genres, count(r.rating) as count fr...	10 row(s) returned



[Back to Home Page](#)

# Insights into the Data: #5



What are the top 100 films with the highest and lowest user ratings (images shown in order)?

Local instance 3306

```
from movie m, rating r
where m.movieid = r.movieid
group by m.genres
order by rating;/*
*/
select distinct m.title, round(avg(r.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings
from movie m, rating r
where m.movieid = r.movieid
group by m.title
having count(r.userId) > 99
order by avg_user_rating DESC, count(r.userId) DESC
limit 100;
```

Result Grid

title	avg_user_rating	num_of_ratings
The Shawshank Redemption (1994)	4.48	2686
The Usual Suspects (1995)	4.38	2088
The Godfather (1972)	4.38	1780
The Godfather: Part II (1974)	4.33	189
One Flew Over the Cuckoo's Nest (...)	4.3	547
Battlefield Earth (2000)	4.29	1212
Police Academy 6: City Under Siege (1989)	4.29	2151
Police Academy 5: Assignment: Miami Beach (1988)	4.29	101
Grease 2 (1982)	4.29	19
Stop! Or My Mom Will Shoot! (1992)	4.27	151
Superman IV: The Quest for Peace (1987)	4.26	101
North by Northwest (1959)	4.25	161
One Flew Over the Cuckoo's Nest (...)	4.25	1303
North by Northwest (1959)	4.25	687
Spirited Away (Sen to Chihiro no ka...	4.25	564

Action Output

Time	Action	Response	Duration / Fetch Time
46 17:46:10	select distinct m.title, round(avgr.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings from movie m, rating r w...	100 row(s) returned	3.648 sec / 0.000051...

Query Completed

Local instance 3306

```
/*What are the top 100 films with the highest and lowest user ratings?*/
select distinct m.title, round(avg(r.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings
from movie m, rating r
where m.movieid = r.movieid
group by m.title
having count(r.userId) > 99
order by avg_user_rating ASC, count(r.userId) DESC
limit 100;
```

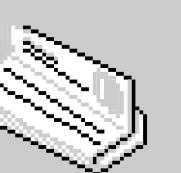
Result Grid

title	avg_user_rating	num_of_ratings
Battlefield Earth (2000)	1.72	191
Police Academy 6: City Under Siege (1989)	1.81	101
Police Academy 5: Assignment: Miami Beach (1988)	1.81	129
Speed 2: Cruise Control (1997)	1.86	109
Grease 2 (1982)	1.86	217
Stop! Or My Mom Will Shoot! (1992)	1.94	101
Superman IV: The Quest for Peace (1987)	1.94	161
Mighty Morphin Power Rangers: The Movie (1995)	1.96	138

Action Output

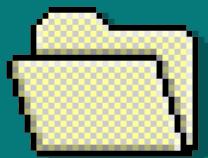
Time	Action	Response
34 22:42:16	select distinct m.title, round(avgr.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings from movie m, rating r w...	100 row(s) returned
35 22:42:36	select distinct m.title, round(avgr.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings from movie m, rating r w...	100 row(s) returned
36 22:42:59	select distinct m.title, round(avgr.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings from movie m, rating r w...	100 row(s) returned
37 22:43:02	select distinct m.title, round(avgr.rating), 2) as avg_user_rating, count(r.userId) as num_of_ratings from movie m, rating r w...	100 row(s) returned

Query Completed



[Back to Home Page](#)

# Insights into the Data: #6



What year was each movie released?

MySQL Workbench

Local instance 3306

Administration Schemas

SCHEMAS

- genome\_tags
- link
- movie
- rating
  - Columns
    - userId
    - movieId
    - rating
    - timestamp
  - Indexes
  - Foreign Keys
  - Triggers
- tag
  - Columns
    - userId
    - movieId

Find Replace Not found < > Q tag Done

```
81 | (LOCATE(')', title) - (LOCATE(')', title) + 1)) AS ip_len
82 | FROM movie
83 | where movieid = mov.movieid
84 | ) temp
85 | where cast(year as unsigned) > 0
86 | and year REGEXP '^[0-9]{4}'
87 | having length(year) = 4
88 | order by year desc);
89 |
90 | select *
91 | from movie;
92 |
93 | select cast('roma' as unsigned);
94 |
95 | update movie
96 | set title = 'Fellini\'s Roma (1972)'
```

Result Grid Filter Rows: Search Export: Fetch rows: Result Grid Form Editor Read Only

moviedb_id	title	genres	year
1	Toy Story (1995)	Adventure Animation Children Comedy Fantasy	1995
2	Jumanji (1995)	Adventure Children Fantasy	1995
3	Grumpier Old Men (1995)	Comedy Romance	1995
4	Waiting to Exhale (1995)	Comedy Drama Romance	1995
5	Father of the Bride Part II (1995)	Comedy	1995
6	Heat (1995)	Action Crime Thriller	1995
7	Sabrina (1995)	Comedy Romance	1995

Action Output

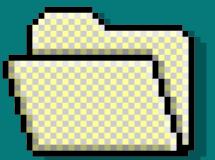
Time	Action	Response
39 22:01:22	select distinct m.year, m.title, max( (select avg(rating) from rat...	Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the ri...
40 22:51:25	select distinct m.year, m.title, max( (select avg(rating) from rat...	Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the ri...
41 22:51:28	select distinct m.year, m.title, max( (select avg(rating) from rat...	1 row(s) returned

Query Completed



[Back to Home Page](#)

# Insights into the Data: #7



Which year were the most number of the most popular genre (drama) of films released?

MySQL Workbench

Local instance 3306

Administration | Schemas

SCHEMAS

Filter objects

genome\_tags  
link  
movie  
rating  
Columns  
userid  
movielid  
rating  
timestamp  
Indexes  
Foreign Keys  
Triggers  
tag  
Columns  
userid  
movielid

```
1 • use movie_analytics;
2
3 /*Which year were the most number of the most popular genre (drama) of films released?*/
4
5 • select genres, count(movieid) as count, year
6 from movie
7 where year is not null
8 group by genres, year
9 order by count DESC
10 limit 1;
11
```

Result Grid | Filter Rows: Search Export: Result Grid

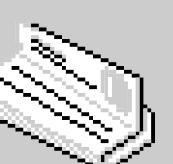
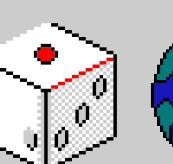
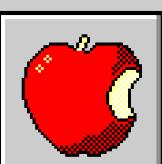
genres	count	year
Drama	62	1996

Result 3 | Read Only

Action Output | Time | Action | Response

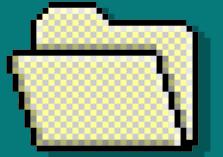
Action	Time	Response
select genres, count(movieid) as count, year from movie grou...	23:26:16	1 row(s) returned
select genres, count(movieid) as count, year from movie wher...	23:26:29	1 row(s) returned

Query Completed



[Back to Home Page](#)

# Insights into the Data: #8



How active are users in rating and tagging films?

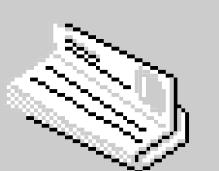
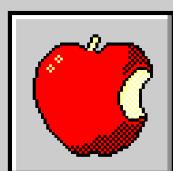
The screenshot shows a MySQL Workbench interface with the following details:

- Schemas:** Local instance 3306, Administration, Schemas (selected).
- Query Editor:** Contains a SQL query:

```
47
48  /* How active are users in rating and tagging films? */
49
50 • select distinct userid, count(tag) as tag_count
51   from tag
52   group by userid
53   order by tag_count;
```
- Result Grid:** Displays the results of the query:

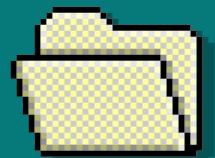
userid	tag_count
318	6
133	10
320	10
96	16
279	26
65	34
121	74
129	122
342	232

Result 21 | Read Only
- Action Output:** Shows the history of actions and their responses, including errors related to aggregated queries without GROUP BY.
- Status:** Query Completed.



[Back to Home Page](#)

# Insights into the Data: #9



What is the relevance associated with each tag given by a user to the movie rated?

Local instance 3306

Schemas: testers for group\* SQL File 7\*

```
/* 9. What is the relevance associated with each tag given by a user to the movie rated? */
75
76 • select distinct gs.tagId, gt.tag, avg(gs.relevance) as avg_relevance
77   from genome_scores gs, genome_tags gt
78   where gs.tagId = gt.tagId
79   group by gs.tagId, gt.tag
80   order by avg_relevance DESC;
```

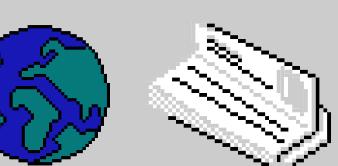
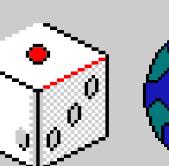
Result Grid

tagId	tag	avg_relevance
78160	397612	5
73211	argentina	4.5
74416	enormously long...	4.5
74916	gull war	4
78174	superheroes	4
76293	1395	3.5
73321	1037705	3.375
73017	988045	3.1978571428571425
78316	time travel	3
74851	1179034	3
74698	808510	3
77798	1179056	3
74532	figure skating	2.0161666666666667
73023	1970s	1.9494375
74795	543	1.814375
77866	955308	1.8095625
296	depressing	1.7941266891891892

Action Output

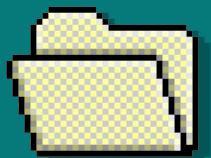
Time	Action	Response	Duration / Fetch Time
08:44:08	select distinct gs.tagId, gt.tag, avg(gs.relevance) as avg_relevance from genome_scores gs, genome_tags gt where gs.tagId = gt.tagId group by gs.tagId, gt.tag order by avg_relevance DESC;	384 row(s) returned	0.046 sec / 0.00009...
08:44:15	select distinct gs.tagId, gt.tag, avg(gs.relevance) as avg_relevance from genome_scores gs, genome_tags gt where gs.tagId = gt.tagId group by gs.tagId, gt.tag order by avg_relevance DESC;	384 row(s) returned	0.042 sec / 0.00010...

Query Completed



[Back to Home Page](#)

# Insights into the Data: #10



What are the top 100 films for which users created the most number of tags?

MySQL Workbench

Local instance 3306

Administration Schemas **testers for group\***

SCHEMAS Filter objects

tables

genome\_scores  
genome\_tags  
link  
movie

Columns movied title genres

Indexes

Foreign Keys

Triggers

rating

Columns userid movied rating timestamp

Object Info Session

Table: rating

Columns: userid int movied int rating double timestamp text

21 /\* 4. What are the top 100 movies for which users created the most tags for? \*/  
22  
23  
24 • SELECT m.movieId, m.title, COUNT(tag) AS count  
25 FROM movie m, tag t  
26 WHERE m.movieId = t.movieId  
27 GROUP BY m.movieId, m.title  
28 ORDER BY count DESC  
29 LIMIT 100;

Result Grid Filter Rows: Search Export:

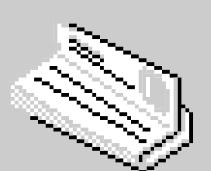
movied	title	count
106696	Frozen (2013)	16
916	Roman Holiday (1953)	14
8641	Anchorman: The Legend of Ron Burgundy (2004)	13
318	Shawshank Redemption, The (1994)	12
2706	American Pie (1999)	11
4878	Donnie Darko (2001)	11
33880	Me and You and Everyone We Know (2005)	10
52328	Sunshine (2007)	10
60756	Step Brothers (2008)	10
3556	Virgin Suicides, The (1999)	9
8528	Dodgeball: A True Underdog Story (2004)	9
44191	V for Vendetta (2006)	9
48043	Fountain, The (2006)	9

Result 14 Read Only

Action Output

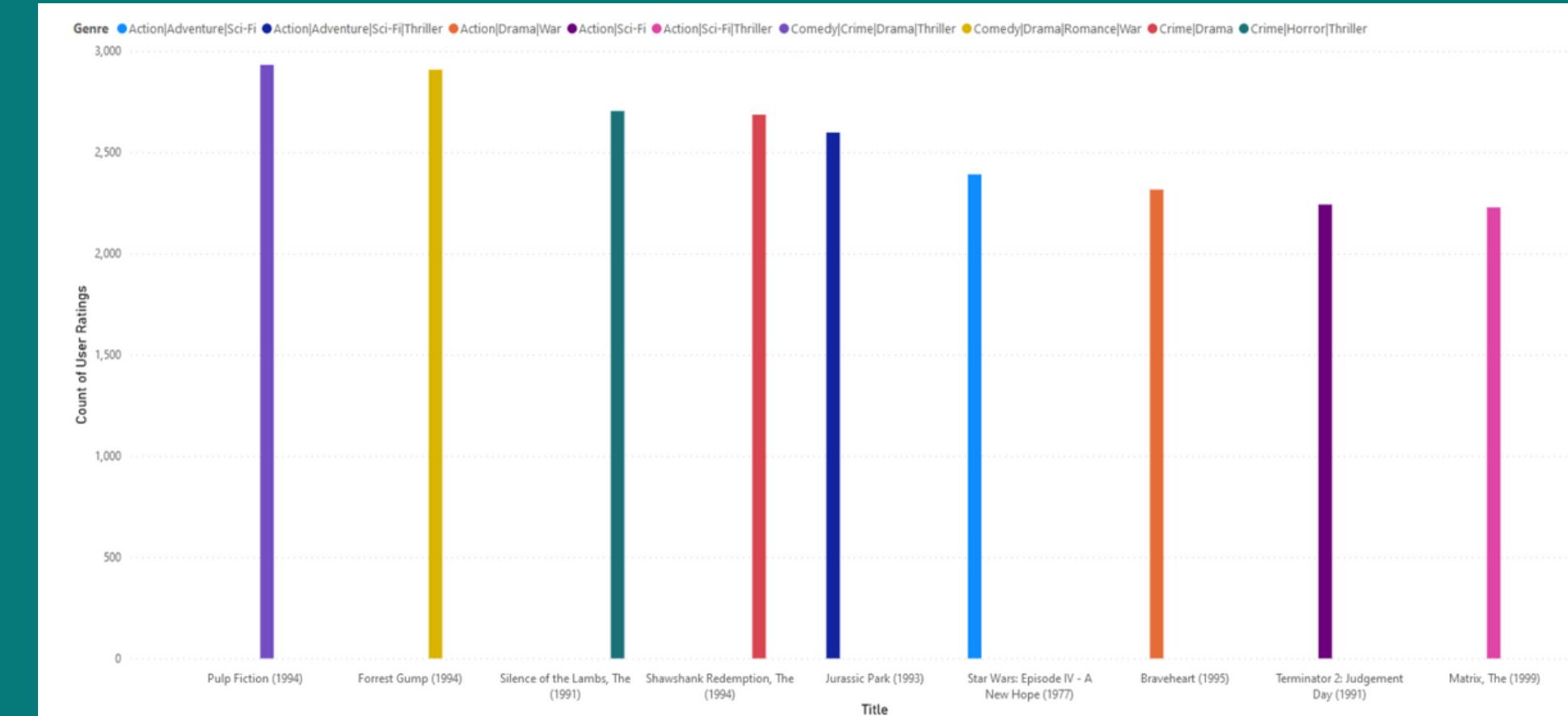
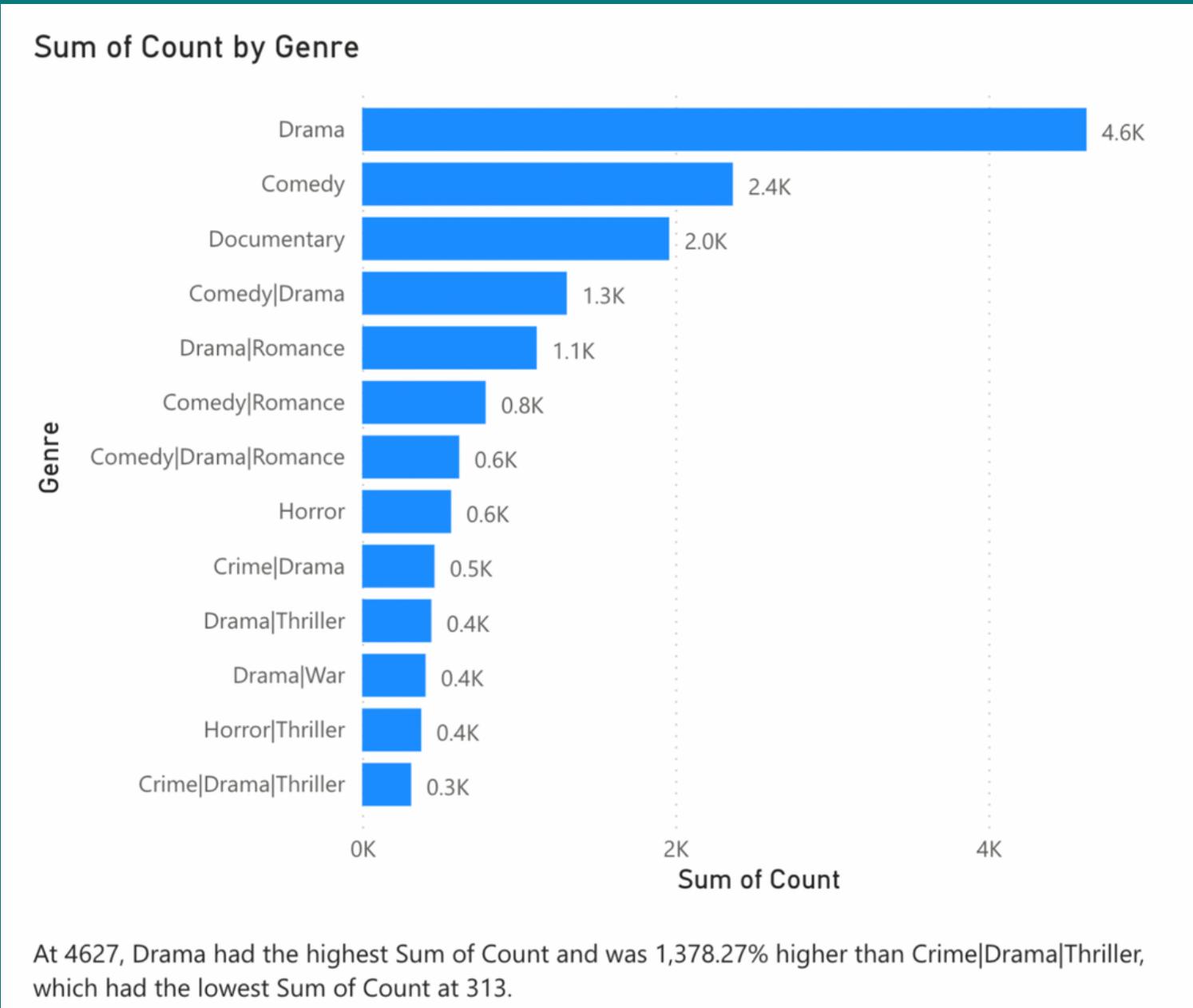
Time	Action	Response	Duration / Fetch Time
72 00:43:01	SELECT distinct genres, COUNT(*) AS count FROM movie GROUP BY genres ORDER BY c...	1343 row(s) returned	0.033 sec / 0.00081...
73 00:43:58	SELECT m.movieId, m.title, COUNT(tag) AS count FROM movie m, tag t WHERE m.movieId...	100 row(s) returned	0.012 sec / 0.000051...

Query Completed



[Back to Home Page](#)

# Visualizations of the Data



# Backup of Data

The screenshot shows the MySQL Workbench interface for a local instance (3306). The left sidebar displays the schema structure under 'Schemas' for 'movie\_analytics'. The 'Tables' section lists 'genome\_scores', 'genome\_tags', 'link', 'movie', 'rating', and their respective sub-sections like 'Columns' and 'Indexes'. The 'Log' section on the right shows the command-line output of the mysqldump process, which successfully exported 6 of 6 objects. The status message indicates 'Export Completed'.

Local instance 3306

Administration Schemas

SCHEMAS

Local instance 3306

Data Export

Object Selection Export Progress

Export Completed

Status:  
6 of 6 exported.

Log:

```
23:36:03 Dumping movie_analytics (link)
Running: /Applications/MySQLWorkbench.app/Contents/MacOS/mysqldump --defaults-file="/var/folders/wv/sjbbkgzn1p9075n221534gp00000gn/T/tmp2hi0vbsk/extraparams.cnf" --host=localhost --port=3306 --default-character-set=utf8 --user=root --protocol=tcp --skip-triggers "movie_analytics" "link"
23:36:03 Dumping movie_analytics (genome_tags)
Running: /Applications/MySQLWorkbench.app/Contents/MacOS/mysqldump --defaults-file="/var/folders/wv/sjbbkgzn1p9075n221534gp00000gn/T/tmpx3seeu54/extraparams.cnf" --host=localhost --port=3306 --default-character-set=utf8 --user=root --protocol=tcp --skip-triggers "movie_analytics" "genome_tags"
23:36:03 Dumping movie_analytics (movie)
Running: /Applications/MySQLWorkbench.app/Contents/MacOS/mysqldump --defaults-file="/var/folders/wv/sjbbkgzn1p9075n221534gp00000gn/T/tmpr_pgu7np/extraparams.cnf" --host=localhost --port=3306 --default-character-set=utf8 --user=root --protocol=tcp --skip-triggers "movie_analytics" "movie"
23:36:03 Dumping movie_analytics (rating)
Running: /Applications/MySQLWorkbench.app/Contents/MacOS/mysqldump --defaults-file="/var/folders/wv/sjbbkgzn1p9075n221534gp00000gn/T/tmpklwr14g6/extraparams.cnf" --host=localhost --port=3306 --default-character-set=utf8 --user=root --protocol=tcp --skip-triggers "movie_analytics" "rating"
23:36:04 Dumping movie_analytics (tag)
Running: /Applications/MySQLWorkbench.app/Contents/MacOS/mysqldump --defaults-file="/var/folders/wv/sjbbkgzn1p9075n221534gp00000gn/T/tmp12t_v_1k/extraparams.cnf" --host=localhost --port=3306 --default-character-set=utf8 --user=root --protocol=tcp --skip-triggers "movie_analytics" "tag"
23:36:04 Dumping movie_analytics (genome_scores)
Running: /Applications/MySQLWorkbench.app/Contents/MacOS/mysqldump --defaults-file="/var/folders/wv/sjbbkgzn1p9075n221534gp00000gn/T/tmp80xgver3/extraparams.cnf" --host=localhost --port=3306 --default-character-set=utf8 --user=root --protocol=tcp --skip-triggers "movie_analytics" "genome_scores"
23:36:05 Export of /Users/alaya/dumps/Dump20221205 has finished
```

Stop Export Again

Query Completed

## The dataset is LARGER than 10MB

```
Dump20221205 — zsh — 102x32

Last login: Mon Nov 28 18:54:18 on console
alaya@Alayas-Air ~ % cd /Users/alaya/dumps/Dump20221205
alaya@Alayas-Air Dump20221205 % ls
movie_analytics_genome_scores.sql      movie_analytics_movie.sql
movie_analytics_genome_tags.sql        movie_analytics_rating.sql
movie_analytics_link.sql                movie_analytics_tag.sql
alaya@Alayas-Air Dump20221205 % ls -al
total 70136
drwx----- 8 alaya  staff   256 Dec  5 23:36 .
drwxr-xr-x  4 alaya  staff   128 Dec  5 23:36 ..
-rw-r--r--  1 alaya  staff  1243763 Dec  5 23:36 movie_analytics_genome_scores.sql
-rw-r--r--  1 alaya  staff   22362 Dec  5 23:36 movie_analytics_genome_tags.sql
-rw-r--r--  1 alaya  staff   1114802 Dec  5 23:36 movie_analytics_link.sql
-rw-r--r--  1 alaya  staff   1657792 Dec  5 23:36 movie_analytics_movie.sql
-rw-r--r--  1 alaya  staff  31833337 Dec  5 23:36 movie_analytics_rating.sql
-rw-r--r--  1 alaya  staff    28034 Dec  5 23:36 movie_analytics_tag.sql
alaya@Alayas-Air Dump20221205 %
```

```
Dump20221205 — zsh — 102x32

314),(65,556,0.045),(65,557,0.08075),(5034,2720,2),(5034,2762,0.004249999999998),(5034,560,0.01175),
(65,561,0.006749999999998),(65,562,0.026),(5034,2987,3),(5034,3034,0.1255),(5034,565,0.119),(65,566,
0.13675),(65,567,0.076),(5034,3396,3),(5034,3404,0.019),(5034,570,0.0195),(5034,571,0.00900000000000000
1),(65,572,0.01225),(65,573,0.0062499999999998),(5034,3591,2),(5034,3635,0.01425),(5034,576,0.0105),(6
5,577,0.02925),(65,578,0.0080000000000001),(5034,3964,4.5),(5034,3969,0.03725),(5034,581,0.04125),(6
5,582,0.047),(65,583,0.01725),(5034,3991,3),(5034,4015,0.23975),(5034,586,0.06325),(65,587,0.04375),(6
5,588,0.02825),(5034,4306,0.054),(5034,590,0.0195),(65,591,0.006749999999998),(65,592,0.037),(5034,4
447,2.5),(5034,4519,0.6675),(5034,595,0.09125),(65,596,0.04025),(65,597,0.0835),(5034,4886,4);
/*!40000 ALTER TABLE `genome_scores` ENABLE KEYS */;
UNLOCK TABLES;
/*!40103 SET TIME_ZONE=@OLD_TIME_ZONE */;

/*!40101 SET SQL_MODE=@OLD_SQL_MODE */;
/*!40014 SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS */;
/*!40014 SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS */;
/*!40101 SET CHARACTER_SET_CLIENT=@OLD_CHARACTER_SET_CLIENT */;
/*!40101 SET CHARACTER_SET_RESULTS=@OLD_CHARACTER_SET_RESULTS */;
/*!40101 SET COLLATION_CONNECTION=@OLD_COLLATION_CONNECTION */;
/*!40111 SET SQL_NOTES=@OLD_SQL_NOTES */;

-- Dump completed on 2022-12-05 23:36:04
[alaya@Alayas-Air Dump20221205 % ls -al
total 70136
drwx----- 8 alaya  staff   256 Dec  5 23:36 .
drwxr-xr-x  4 alaya  staff   128 Dec  5 23:36 ..
-rw-r--r--  1 alaya  staff  1243763 Dec  5 23:36 movie_analytics_genome_scores.sql
-rw-r--r--  1 alaya  staff   22362 Dec  5 23:36 movie_analytics_genome_tags.sql
-rw-r--r--  1 alaya  staff   1114802 Dec  5 23:36 movie_analytics_link.sql
-rw-r--r--  1 alaya  staff   1657792 Dec  5 23:36 movie_analytics_movie.sql
-rw-r--r--  1 alaya  staff  31833337 Dec  5 23:36 movie_analytics_rating.sql
-rw-r--r--  1 alaya  staff    28034 Dec  5 23:36 movie_analytics_tag.sql
alaya@Alayas-Air Dump20221205 %
```

# The dataset is LARGER than 10MB

The screenshot shows the MySQL Workbench interface for a local instance at port 3306. The left sidebar displays the 'SCHEMAS' tree, which includes 'ap', 'ex', and the 'movie\_analytics' schema, which is currently selected. The 'movie\_analytics' schema tree shows tables like 'genome\_scores', 'genome\_tags', 'link', 'movie', 'rating', and 'tag', along with their respective columns, indexes, foreign keys, triggers, and session data. The main panel displays the 'Schema Details' for the 'movie\_analytics' schema, listing the default collation as 'utf8mb4\_0900\_ai\_ci', the default character set as 'utf8mb4', a table count of 6, and a database size of 61.3 MiB. At the bottom, an action output log shows a query attempt to select the 'movie\_analytics' database, resulting in an error due to syntax issues.

Local instance 3306

Administration Schemas testers for group\* Untitled recc report\* SQL File 5\* SQL File 6\* Administration - Data Export movie\_analytics

SCHEMAS

Filter objects

Local instance 3306 movie\_analytics

Schema Details

Default collation: utf8mb4\_0900\_ai\_ci  
Default characterset: utf8mb4  
Table count: 6  
Database size (rough estimate): 61.3 MiB

Tables Columns Indexes Triggers Views Stored Procedures Functions Grants Events

Object Info Session Schema: movie\_analytics

Action Output

Time Action Response

56 23:46:55 SELECT movie\_analytics AS "Database", ROUND(SUM(data\_le... Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL server version for the right syntax to u

Query interrupted

Thank you!

