

DATABASE MANAGEMENT AND DATABASE DESIGN

ASSIGNMENT 4

(Normalization)

Tv-show/Movie Recommendation System

TEAM MEMBERS -

Shreyas Rai (002769838)
Smiti Agrawal (002781419)
Sarthak Srivastava (002756847)

GITHUB LINK -

https://github.com/shreyashusky/tv_Shows_Recommendation_System.git
<https://github.com/smitihusky/tvShowRecommendationSystem.git>
<https://github.com/sarthakhusky/TV-Shows-Recommendation-System.git>

ABOUT

These days, the small screen has some very big things to offer. From sitcoms to dramas to travel and talk shows, all the small screen shows are the best in showing diversity. TV shows are definitely one of the biggest business markets out there. Also, it offers an opportunity to upcoming artists. Over the years, the number of TV shows has increased exponentially and so has their customer base.

A survey conducted by www.deadline.com says that the top TV Shows can have 15-20 million viewership. With the advancement in technology and availability of cheap internet services, the reach of TV Shows is going to increase in the near future.

Our domain is Top TV Shows aired till date. We will use social networking APIs to get the data for entities which represent Companies, Genre, Subscription, Ratings, Producers and Consumers for our domain.

First Normal Form (1NF)

- Each table has a primary key with minimal attributes that can uniquely identify a record.
- The values in each column of a table are atomic.
- There are no repeating groups in the same table.

Example 1: In the genre.csv (image below) the Movie **Avatar** has **4 genres** but in according to 1NF there should be minimal attributes and each record should be uniquely identified.

Before 1NF:

The screenshot shows a Jupyter Notebook interface with two code cells and their corresponding outputs.

In [251]: `df = pd.read_csv("genre.csv")`

Out[252]:

	movie_id	movie_title	genre	certification
0	25683	Avatar	Action Adventure Fantasy Science F...	PG13
1	14531	Pirates of the Caribbean: At World's End	Adventure Fantasy Action	PG13
2	18260	Spectre	Action Adventure Crime	PG13
3	25056	The Dark Knight Rises	Action Crime Drama Thriller	PG13
4	99197	John Carter	Action Adventure Science Fiction	PG13
...
495	75730	Journey 2: The Mysterious Island	Adventure Action Science Fiction	PG
496	78945	Cloudy with a Chance of Meatballs 2	Animation Family Comedy	PG13
497	57478	Red Dragon	Crime Thriller Horror	PG13
498	55309	Hidalgo	Western Adventure	PG13
499	72551	Jack and Jill	Comedy	PG13

500 rows x 4 columns

In [311]: `df = pd.read_csv("genre.csv")`

In [312]: `df['genre']=df['genre'].str.split(" ")
df.explode(['genre'])`

Out[312]:

	movie_id	movie_title	genre	certification
0	25683	Avatar	Action	PG13

So in 1NF, the values in each column of a table should be atomic and there should be no repeating groups so we can see that here the movie **Avatar's** genre section is now split into **4 rows** of genres : Action, Adventure, Fantasy and Science Fiction.

After 1NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: an hour ago (autosaved)". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Python 3 (ipykernel) button. A "Not Trusted" warning is present.

```
In [311]: df = pd.read_csv("genre.csv")
In [312]: df['genre']=df['genre'].str.split(" ")
df.explode(['genre'])
Out[312]:
   movie_id      movie_title    genre certification
0     25683          Avatar  Action        PG13
0     25683          Avatar Adventure        PG13
0     25683          Avatar  Fantasy        PG13
0     25683          Avatar Science Fiction        PG13
1    14531  Pirates of the Caribbean: At World's End Adventure        PG13
...
497    57476         Red Dragon Thriller        PG13
497    57476         Red Dragon Horror        PG13
498    55309          Hidalgo Western        PG13
498    55309          Hidalgo Adventure        PG13
499    72551       Jack and Jill Comedy        PG13
```

1599 rows × 4 columns

```
In [308]: df.to_csv("chu.csv")
```

```
In [ ]:
```

Example 2: Same thing is implemented for countries. In the country column we have multiple values, so we have applied 1NF and separated it into atomic values.

Before 1NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: an hour ago (unsaved changes)". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Python 3 (ipykernel) button. A "Not Trusted" warning is present.

```
In [326]: df = pd.read_csv("festiveFilm.csv")
```

```
In [314]: df
```

```
Out[314]:
   movie_id      movie_title    country festival
0     25683          Avatar United States of America United Kingdom Cannes
1    14531  Pirates of the Caribbean: At World's End United States of America Cannes
2     18260          Spectre United Kingdom United States of America Berlin
3    25056  The Dark Knight Rises United States of America Venice
4    99197       John Carter United States of America Cannes
...
495    75730  Journey 2: The Mysterious Island United States of America Venice
496    78945  Cloudy with a Chance of Meatballs 2 United States of America Cannes
497    57476         Red Dragon Germany United States of America Berlin
498    55309          Hidalgo United States of America Morocco Venice
499    72551       Jack and Jill United States of America Berlin
```

500 rows × 4 columns

After 1NF:

```

497    57476           Red Dragon      Germany United States of America Berlin
498    55309           Hidalgo        United States of America Morocco Venice
499    72551          Jack and Jill   United States of America Berlin

500 rows x 4 columns

In [327]: df['country']=df['country'].str.split(" ")
df.explode(['country'])

Out[327]:
   movie_id      movie_title      country festival
0     25683          Avatar  United States of America Cannes
0     25683          Avatar  United Kingdom Cannes
1    14531  Pirates of the Caribbean: At World's End  United States of America Cannes
2     18260          Spectre  United Kingdom Berlin
2     18260          Spectre  United States of America Berlin
...
...
497    57476          Red Dragon      Germany Berlin
497    57476          Red Dragon  United States of America Berlin
498    55309           Hidalgo  United States of America Venice
498    55309           Hidalgo       Morocco Venice
499    72551          Jack and Jill  United States of America Berlin

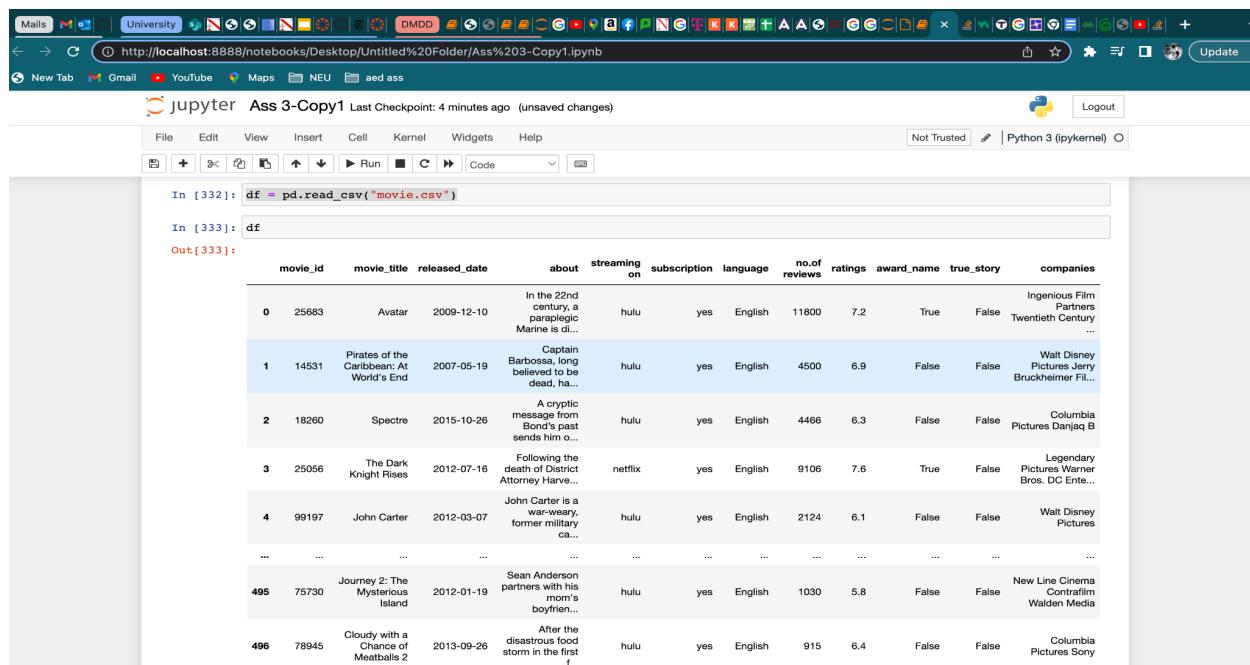
765 rows x 4 columns

In [328]: df.to_csv("1NF_festiveFilm.csv")

```

Example 3. In the below example, the “company’s” column is having multiple values and it is violating 1NF. So we have implemented 1NF and now each entry has a separate company name.

Before 1NF:



The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1". The code cell "In [332]" contains `df = pd.read_csv("movie.csv")`. The output cell "Out[333]" shows a table with 500 rows and 13 columns. The columns are: movie_id, movie_title, released_date, about, streaming_on, subscription, language, no.of reviews, ratings, award_name, true_story, and companies. The "companies" column contains multiple company names separated by commas. For example, the first few rows show:

movie_id	movie_title	released_date	about	streaming_on	subscription	language	no.of reviews	ratings	award_name	true_story	companies
0	Avatar	2009-12-10	In the 22nd century, a paraplegic Marine is di...	hulu	yes	English	11800	7.2	True	False	Ingenious Film Partners Twentieth Century ...
1	Pirates of the Caribbean: At World's End	2007-05-19	Captain Barbosa, long believed to be dead, ha...	hulu	yes	English	4500	6.9	False	False	Walt Disney Pictures Jerry Bruckheimer Fil...
2	Spectre	2015-10-26	A cryptic message from Bond's past sends him o...	hulu	yes	English	4466	6.3	False	False	Columbia Pictures Danjaq B
3	The Dark Knight Rises	2012-07-16	Following the death of District Attorney Harve...	netflix	yes	English	9106	7.6	True	False	Legendary Pictures Warner Bros. DC Ente...
4	John Carter	2012-03-07	John Carter is a war-weary, former military ca...	hulu	yes	English	2124	6.1	False	False	Walt Disney Pictures

After 1NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1". The code cell In [342] contains the command `df['companies']=df['companies'].str.split(" ")` and `df.explode(['companies'])`. The output cell Out[342] displays a pandas DataFrame with 10 rows. The columns are movie_id, movie_title, released_date, about, streaming_on, subscription, language, no.of reviews, ratings, award_name, true_story, and companies. The companies column lists various production studios like Ingenious Film Partners, Twentieth Century Fox Film Corporation, Dune Entertainment, Lightstorm Entertainment, and Walt Disney Pictures.

	movie_id	movie_title	released_date	about	streaming_on	subscription	language	no.of reviews	ratings	award_name	true_story	companies
0	25683	Avatar	2009-12-10	In the 22nd century, a paraplegic Marine is di...	hulu	yes	English	11800	7.2	True	False	Ingenious Film Partners
0	25683	Avatar	2009-12-10	In the 22nd century, a paraplegic Marine is di...	hulu	yes	English	11800	7.2	True	False	Twentieth Century Fox Film Corporation
0	25683	Avatar	2009-12-10	In the 22nd century, a paraplegic Marine is di...	hulu	yes	English	11800	7.2	True	False	Dune Entertainment
0	25683	Avatar	2009-12-10	In the 22nd century, a paraplegic Marine is di...	hulu	yes	English	11800	7.2	True	False	Lightstorm Entertainment
1	14531	Pirates of the Caribbean: At World's End	2007-05-19	Captain Barbosa, long believed to be dead, ha...	hulu	yes	English	4500	6.9	False	False	Walt Disney Pictures
...
498	55309	Hidalgo	2004-03-05	Set in 1890, this is the story of a Pony Expr...	netflix	yes	English	318	6.5	False	False	Touchstone Pictures
498	55309	Hidalgo	2004-03-05	Set in 1890, this is the story of a Pony Expr...	netflix	yes	English	318	6.5	False	False	Dune Films
				Jack Sadelstein,								

Second Normal Form (2NF)

- Each table has a primary key with minimal attributes that can uniquely identify a record.
- The values in each column of a table are atomic.
- There are no repeating groups in the same table.
- There are no Partial dependencies or Calculated data.

A relation is 2NF when it is in the First Normal Form but has no non-prime attribute functionality. There should be no partial dependencies or Calculated data in a relation.

So here the country and the festival have dependencies so if in the future a festival wants to shift to another country so the changes should only be made in the country column only.

Before 2NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: an hour ago (autosaved)". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Python 3 (ipykernel) button. A "Not Trusted" warning is present.

In [344]:

```
df = pd.read_csv("countryFest.csv")
```

In [345]:

```
df
```

Out[345]:

	movie_id	movie_title	country	festival
0	25683	Avatar	United States of America	Cannes
1	14531	Pirates of the Caribbean: At World's End	United States of America	Cannes
2	18260	Spectre	Germany	Berlin
3	25056	The Dark Knight Rises	Spain	Venice
4	99197	John Carter	United States of America	Cannes
...
495	75730	Journey 2: The Mysterious Island	Spain	Venice
496	78945	Cloudy with a Chance of Meatballs 2	United States of America	Cannes
497	57476	Red Dragon	Germany	Berlin
498	55309	Hidalgo	Spain	Venice
499	72551	Jack and Jill	Germany	Berlin

500 rows × 4 columns

In [358]:

```
df1=df[['movie_id','movie_title','country']]  
df2=df[['country','festival']]
```

In [359]:

```
print(f"Unique FN: {df['country'].unique()}")
```

Unique FN: ['United States of America' 'Germany' 'Spain' 'Italy' 'Australia' 'Canada']

In [376]:

```
df4=(f"Unique Values from 2 Columns:\n{pd.concat([df['country'],df['festival']]).unique()}")
```

After 2NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: an hour ago (autosaved)". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Python 3 (ipykernel) button. A "Not Trusted" warning is present.

In [358]:

```
df1=df[['movie_id','movie_title','country']]  
df2=df[['country','festival']]
```

In [359]:

```
print(f"Unique FN: {df['country'].unique()}")
```

Unique FN: ['United States of America' 'Germany' 'Spain' 'Italy' 'Australia' 'Canada']

In [376]:

```
df4=(f"Unique Values from 2 Columns:\n{pd.concat([df['country'],df['festival']]).unique()}")
```

In [377]:

```
df4
```

Out[377]:

```
"Unique Values from 2 Columns:['United States of America' 'Germany' 'Spain' 'Italy' 'Australia' 'Canada'\n'Cannes'\n'Berlin' 'Venice' 'Sundance' 'Rotterdam' 'Toronto']"
```

In [378]:

```
df1
```

Out[378]:

	movie_id	movie_title	country
0	25683	Avatar	United States of America
1	14531	Pirates of the Caribbean: At World's End	United States of America
2	18260	Spectre	Germany
3	25056	The Dark Knight Rises	Spain
4	99197	John Carter	United States of America
...
495	75730	Journey 2: The Mysterious Island	Spain
496	78945	Cloudy with a Chance of Meatballs 2	United States of America
497	57476	Red Dragon	Germany
498	55309	Hidalgo	Spain

After 2NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: an hour ago (autosaved)". The menu bar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, Not Trusted, Python 3 (ipykernel), and a code cell toolbar.

The code cell displays the following data:

495	75730	Journey 2: The Mysterious Island	Spain
496	78945	Cloudy with a Chance of Meatballs 2	United States of America
497	57476	Red Dragon	Germany
498	55309	Hidalgo	Spain
499	72551	Jack and Jill	Germany

Below the table, it says "500 rows x 3 columns".

The code cells show:

```
In [379]: df.to_csv("2NF_Table1.csv")
In [380]: df = pd.read_csv("2NF_Table2.csv")
In [381]: df
```

The output of cell 381 is:

	country	festival
0	United States of America	Cannes
1	Germany	Berlin
2	Spain	Venice
3	Italy	Sundance
4	Australia	Rotterdam
5	Canada	Toronto

Third Normal Form (3NF)

- Each table has a primary key with minimal attributes that can uniquely identify a record.
- The values in each column of a table are atomic.
- There are no repeating groups in the same table.
- There are no Partial dependencies or Calculated data.
- Eliminate fields that do not directly depend on the primary key; that is no transitive dependencies

3NF is a relational database that uses normalizing principles to reduce the duplication of data, avoid data anomalies, ensure referential integrity, and simplify data management. In 2NF the attributes should be functionally dependent but in 3NF the attributes should be directly (non-transitively) dependent. In 3NF A is dependent on B and B is dependent on C like here genre is dependent on movie and age restriction is dependent on genre. So we have made 2 tables that separate genre and age_restriction columns. Thus, we eliminated transitive dependency and now our data is in 3NF.

Before 3NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: 14 minutes ago (autosaved)". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Python 3 (ipykernel) option. The code cell In [384] contains `df = pd.read_csv("restriction.csv")`. The output Out[385] displays a DataFrame with columns movie_id, movie_title, genre, and age_restriction. The data includes rows for various movies like Avatar, Pirates of the Caribbean: At World's End, Spectre, etc., with genres Action, PG13, PG, etc. In [386] shows df1=df[['movie_id','movie_title','genre']] and df2=df[['genre','age_restriction']]. In [387] shows df1.

movie_id	movie_title	genre	age_restriction
0	25683	Avatar	Action
1	14531	Pirates of the Caribbean: At World's End	Action
2	18260	Spectre	Action
3	25056	The Dark Knight Rises	Action
4	99197	John Carter	Action
...
495	75730	Journey 2: The Mysterious Island	Family
496	78945	Cloudy with a Chance of Meatballs 2	Action
497	57476	Red Dragon	Action
498	55309	Hidalgo	Action
499	72551	Jack and Jill	Action

500 rows × 4 columns

movie_id	movie_title	genre
0	25683	Avatar
1	14531	Action
2	18260	Action
3	25056	Action
4	99197	Action
...
495	75730	Family
496	78945	Action
497	57476	Action
498	55309	Action
499	72551	Action

After 3NF:

The screenshot shows a Jupyter Notebook interface with the title "jupyter Ass 3-Copy1 Last Checkpoint: 14 minutes ago (autosaved)". The toolbar includes File, Edit, View, Insert, Cell, Kernel, Widgets, Help, and a Python 3 (ipykernel) option. The code cell In [387] contains `df1`. The output Out[387] displays a DataFrame with columns movie_id, movie_title, and genre. The data includes rows for various movies like Avatar, Pirates of the Caribbean: At World's End, Spectre, etc., with genres Action, PG13, PG, etc. In [389] shows print(f"Unique FN: {df['genre'].unique()}"). The output Unique FN: ['Action' 'Family' 'Romance']. In [390] shows df4=(f"Unique Values from 2 Columns:\n{pd.concat([df['genre'],df['age_restriction']]).unique()}"). In [391] shows df4.

movie_id	movie_title	genre
0	25683	Avatar
1	14531	Pirates of the Caribbean: At World's End
2	18260	Spectre
3	25056	The Dark Knight Rises
4	99197	John Carter
...
495	75730	Journey 2: The Mysterious Island
496	78945	Cloudy with a Chance of Meatballs 2
497	57476	Red Dragon
498	55309	Hidalgo
499	72551	Jack and Jill

500 rows × 3 columns

```
In [389]: print(f"Unique FN: {df['genre'].unique()}")
Unique FN: ['Action' 'Family' 'Romance']

In [390]: df4=(f"Unique Values from 2 Columns:\n{pd.concat([df['genre'],df['age_restriction']]).unique()}")
In [391]: df4
```

Out[391]: "Unique Values from 2 Columns:['Action' 'Family' 'Romance' 'PG13' 'PG' 'R']"

After 3NF:

In [389]: `print(f"Unique FN: {df['genre'].unique()}")`
Unique FN: ['Action' 'Family' 'Romance']

In [390]: `df4=(f"Unique Values from 2 Columns:\n{pd.concat([df['genre'],df['age_restriction']]).unique()}")`

In [391]: `df4`
Out[391]: "Unique Values from 2 Columns:['Action' 'Family' 'Romance' 'PG13' 'PG' 'R']"

In [392]: `df1.to_csv("3NF_Table1.csv")`

In [393]: `df = pd.read_csv("3NF_Table2.csv")`

In [394]: `df`
Out[394]:

	genre	age_restriction
0	Action	PG13
1	Family	PG
2	Romance	R

CREATING VIEWS

PREVIOUS ASSIGNMENTS USE CASES

(Use-cases Views 5*3= 15 Views)

5 USE-CASES (SMITI AGRAWAL)

1.What are the family movies that are released during the Cannes film festival ?

```

5   #What are the family movies that are released during the Cannes film festival ?
6   CREATE VIEW [familyfest] AS select f.movie_title,g.genre,f.festival from restriction g,festiveFilm f where g.movie_id=f.movie_id and festival="Cannes" and genre like "%Family%";
7
8
9
10
11

```

	movie_title	genre	festival
1	Tangled	Family	Cannes
2	Monsters University	Family	Cannes
3	Cars 2	Family	Cannes
4	Toy Story 3	Family	Cannes
5	Jack the Giant Slayer	Family	Cannes
6	The Good Dinosaur	Family	Cannes
7	The Chronicles of Narnia: The Lion, the Witch and the ...	Family	Cannes
8	Hugo	Family	Cannes
9	Evan Almighty	Family	Cannes
10	Inside Out	Family	Cannes
11	Maleficent	Family	Cannes
12	Big Hero 6	Family	Cannes
13	Wreck-It Ralph	Family	Cannes
14	How to Train Your Dragon	Family	Cannes
15	Shrek the Third	Family	Cannes
16	Retatouille	Family	Cannes

Execution finished without errors.
Result: 63 rows returned in 10ms
At line 10:
SELECT * FROM familyfest;

2.What are the top 10 movies having higher profits?

```
1 #What are the top 10 movies having higher profits?
2 CREATE VIEW [profitablemovies] AS select e.movie_title,(e.gross-b.budget) as profit from earnings e, budget b where e.movie_id=b.movie_id order by profit desc limit 10;
3
4 SELECT * FROM profitablemovies;
```

	movie_title	profit
1	Avatar	2550965087
2	Titanic	1645034188
3	Jurassic World	1363528810
4	Furious 7	1316249360
5	The Avengers	1299557910
6	Avengers: Age of Ultron	1125403694
7	Frozen	1124219009
8	The Lord of the Rings: The Return of the King	1024888979
9	Iron Man 3	1015439994
10	Transformers: Dark of the Moon	928746996

Execution finished without errors.
Result: 10 rows returned in 13ms
At line 4:
SELECT * FROM profitablemovies;

3.Which director's movie has earned maximum money ?

```
6
7 #Which directors movie has earned the maximum money ?
8 CREATE VIEW [directorearning] AS select distinct d.director,sum(e.gross) as total_gross from director d, earnings e
9 where d.movie_id=e.movie_id
10 group by d.director
11 order by e.gross desc limit 1;
12
13
14 SELECT * FROM directorearning;
```

director	total_gross
James cameron	2787965087

Execution finished without errors.
Result: 1 rows returned in 20ms
At line 14:
SELECT * FROM directorearning;

4. Highest rated movies released in the year 2000, USA?

```
1
2
3 #Highest rated movies released in year 2000, USA?
4 CREATE VIEW [topmovie2000] AS select m.movie_title,m.release_date,f.country, m.ratings
5   from movie m,festiveFilm f
6   where m.movie_id=f.movie_id and ratings>? and m.release_date like "%2000%" and f.country like "%United States of America%"
7   order by ratings desc;
8
9 SELECT * FROM topmovie2000;
10
11
12
13
14
15
```

	movie_title	ratings
1	Avatar	7.2
2	Pirates of the Caribbean: At World's End	6.9
3	Spectre	6.3
4	The Dark Knight Rises	7.6
5	John Carter	6.1
6	Spider-Man 3	5.9
7	Tangled	7.4
8	Avengers: Age of Ultron	7.3
9	Harry Potter and the Half-Blood Prince	7.4
10	Batman v Superman: Dawn of Justice	5.7
11	Superman Returns	5.4
12	Quantum of Solace	6.1
13	Pirates of the Caribbean: Dead Man's Chest	7
14	The Lone Ranger	5.9
15	Man of Steel	6.5
16	The Chronicles of Narnia: Prince Caspian	6.3

Execution finished without errors.
Result: 500 rows returned in 6ms
At line 12:

5. What are real story based movies that have earned the highest votes?

```
1 #What are real story based movies that has earned highest votes?
2
3 CREATE VIEW [realstory] AS select distinct(m.movie_title), e.votes
4   from movie m,earnings e where e.movie_id=m.movie_id and m.biograpgy="FALSE"
5   order by votes desc
6
7 SELECT * from realstory;
8
9
```

	movie_title	votes
1	Inception	13752
2	The Dark Knight	12002
3	Avatar	11800
4	The Avengers	11776
5	Interstellar	10867
6	Django Unchained	10099
7	Guardians of the Galaxy	9742
8	The Hunger Games	9455
9	Mad Max: Fury Road	9427
10	The Dark Knight Rises	9106
11	Iron Man 3	8806
12	Iron Man	8776
13	The Lord of the Rings: The Fellowship of the Ring	8705
14	Jurassic World	8662
15	The Hobbit: An Unexpected Journey	8297
16	The Lord of the Rings: The Return of the King	8064

Execution finished without errors.
Result: 500 rows returned in 27ms
At line 1:
select distinct(m.movie_title), e.votes
from movie m,earnings e where e.movie_id=m.movie_id and m.biograpgy="FALSE"
order by votes desc

5 USE-CASES (SHREYAS RAI)

1. What are top rated movies for children above age 13?

```
9
0
1 #Top rated movies for children above age 13
2 ┌─ CREATE VIEW [parentsguide] AS select g.movie_title, g.certification, r.rating
3   from genre g, ratings r
4   where g.movie_id=r.movie_id and g.certification = "PG13"
5   order by rating desc limit 20
6
7 ┌─ SELECT * from parentsguide;
8
```

movie_title	certification	rating
The Dark Knight	PG13	8.2
Interstellar	PG13	8.1
Inception	PG13	8.1
The Lord of the Rings: The Return of the King	PG13	8.1
The Lord of the Rings: The Fellowship of the Ring	PG13	8
The Lord of the Rings: The Two Towers	PG13	8
Guardians of the Galaxy	PG13	7.9
Gladiator	PG13	7.9
The Wolf of Wall Street	PG13	7.9
The Departed	PG13	7.9
Shutter Island	PG13	7.8
Terminator 2: Judgment Day	PG13	7.7
The Dark Knight Rises	PG13	7.6
The Hobbit: The Desolation of Smaug	PG13	7.6
Edge of Tomorrow	PG13	7.6
Captain America: The Winter Soldier	PG13	7.6

Execution finished without errors.
Result: 20 rows returned in 18ms
At line 17:
SELECT * from parentsguide;

2. List award winning movies for Action packed

```
19
20 #List award winning movies for Action packed
21
22 ┌─ CREATE VIEW [actionpacked] AS select m.movie_title, g.genre, m.ratings
23   from genre g, movie m
24   where g.movie_id=m.movie_id
25   and m.award LIKE "%TRUE%"
26   and g.genre like "%Action%"
27   order by m.ratings desc
28
29 ┌─ SELECT * from actionpacked;
30
```

movie_title	genre	ratings
Inception	Action Thriller Science Fiction Mystery Adventure	8.1
The Lord of the Rings: The Return of the King	Adventure Fantasy Action	8.1
The Lord of the Rings: The Fellowship of the Ring	Adventure Fantasy Action	8
The Lord of the Rings: The Two Towers	Adventure Fantasy Action	8
Gladiator	Action Drama Adventure	7.9
Terminator 2: Judgment Day	Action Thriller Science Fiction	7.7
The Dark Knight Rises	Action Crime Drama Thriller	7.6
Edge of Tomorrow	Action Science Fiction	7.6
X-Men: Days of Future Past	Action Adventure Fantasy Science Fiction	7.5
Batman Begins	Action Crime Drama	7.5
The Avengers	Science Fiction Action Adventure	7.4
Star Trek Into Darkness	Action Adventure Science Fiction	7.4
The Hunger Games: Catching Fire	Adventure Action Science Fiction	7.4
Avatar	Action Adventure Fantasy Science Fiction	7.2

Execution finished without errors.
Result: 14 rows returned in 10ms
At line 29:
SELECT * from actionpacked;

3.List the movies which has less ratings but earned more

```
32 #Movies which has less ratings but earned more
33
34
35  CREATE VIEW [underrated] AS select e.movie_title
36   from earnings e, ratings r
37   where e.movie_id=r.movie_id
38   and r.rating<6
39   order by e.gross desc
40
41   select * from underrated;
42
43
44
45
```

	movie_title
1	Independence Day: Resurgence
2	The Last Airbender
3	Eragon
4	Batman & Robin
5	Gulliver's Travels
6	Fantastic Four
7	Speed 2: Cruise Control
8	Jack and Jill
9	Nutty Professor II: The Klumps
10	Cats & Dogs 2 : The Revenge of Kitty Galore
11	Catwoman
12	Stealth
13	How Do You Know
14	Battlefield Earth
15	Town & Country
16	The Adventures of Pluto Nash

Execution finished without errors.
Result: 20 rows returned in 5ms
At line 41:
select * from underrated;

4.Movies which were superhit in more than one country

```
34 #Movies which were superhit in more than one country
35
36  CREATE VIEW [superhit] AS select r.movie_title as superhit_movies,f.country
37   from ratings r, festiveFilm f
38   where r.rating>7.1
39   group by r.movie_title limit 10
40
41   select * from superhit;
42
43
```

	superhit_movies	country	
1	A Beautiful Mind	United States of America	United Kingdom
2	All That Jazz	United States of America	United Kingdom
3	American Gangster	United States of America	United Kingdom
4	Avatar	United States of America	United Kingdom
5	Avengers: Age of Ultron	United States of America	United Kingdom
6	Batman Begins	United States of America	United Kingdom
7	Big Hero 6	United States of America	United Kingdom
8	Black Hawk Down	United States of America	United Kingdom
9	Blood Diamond	United States of America	United Kingdom
10	Captain America: The Winter Soldier	United States of America	United Kingdom

Execution finished without errors.
Result: 10 rows returned in 64ms
At line 41:
select * from superhit;

5. Cast and directors with maximum hit movies

```
42
43 #Cast and directors with maximum hit movies
44
45 ┌─ CREATE VIEW [bestdirector] AS select c.movie_title as Movie, d.director, c.cast
46   from cast c, director d, ratings r
47   where c.movie_id=d.movie_id
48   and d.movie_id=r.movie_id
49   and r.rating>8
50   order by r.rating desc
51
52 select * from bestdirector;
53
```

	Movie	director	cast
1	The Dark Knight	Roberto Benigni	Bruce Wayne fecafdf Christian Bale Joker fefcafcbf ...
2	Interstellar	Makoto Shinkai	Joseph Cooper febbfceb Matthew McConaughey ...
3	Inception	Nitesh Tiwari	Dom Cobb fecaeade Leonardo DiCaprio Arthur ...
4	The Lord of the Rings: The Return of the King	Darius Marder	Frodo Baggins febcff Elijah Wood Gandalf the Whi...

Execution finished without errors.
Result: 4 rows returned in 28ms
At line 52:
select * from bestdirector;

5 USE-CASES (SARTHAK SHRIVASTAV)

1. What are the best movies available on Hulu?

```
54
55 #what are the Best movies available on Hulu?
56 ┌─ CREATE VIEW [tophulu] AS select m.movie_title
57   from review r, movie m
58   where r.movie_title=m.movie_title and m.streaming_on like "%hulu%" and r.review like "%Best%"
59
60
61 select * from tophulu;
62
63
64
65
```

	movie_title
1	Avatar
2	The Hobbit: The Desolation of Smaug
3	The Hobbit: The Desolation of Smaug
4	X-Men: Days of Future Past
5	Star Trek Into Darkness
6	The Great Gatsby
7	The Dark Knight
8	Edge of Tomorrow
9	Maleficent
10	Dawn of the Planet of the Apes
11	Captain America: The Winter Soldier
12	How to Train Your Dragon
13	Guardians of the Galaxy
14	The Curious Case of Benjamin Button
15	X-Men: First Class
16	Ratatouille

Execution finished without errors.
Result: 48 rows returned in 9ms
At line 61:
select * from tophulu;

2.Which directors has age restrictions as PG?

```
63  
64 #which directors has age restriction as PG  
65  
66 CREATE VIEW [directorrestriction] AS select d.movie_title,d.director  
67     from director d ,genre g where d.movie_id=g.movie_id and g.certification = "PG"  
68  
69  
70 select * from directorrestriction;  
71
```

	movie_title	director
1	Tangled	Nathan Greno
2	Harry Potter and the Half-Blood Prince	David Yates
3	The Chronicles of Narnia: Prince Caspian	Andrew Adamson
4	Monsters University	Dan Scanlon
5	Oz: The Great and Powerful	Sam Raimi
6	Cars 2	Francis Ford Coppola
7	Toy Story 3	Francis Ford Coppola
8	Jack the Giant Slayer	David Fincher
9	The Good Dinosaur	Martin Scorsese
10	Brave	Irvin Kershner
11	WALL-E	Thomas Kail
12	The Chronicles of Narnia: The Lion, the Witch and the ...	Steven Spielberg
13	Up	David Fincher
14	Hugo	Masaki Kobayashi
15	Evan Almighty	Damien Chazelle
16	Inside Out	Roman Polanski

Execution finished without errors.
Result: 116 rows returned in 13ms
At line 70:
select * from directorrestriction;

3.What movies of directors are in the Berlin Festival?

```
73  
74  
75 #What movies of directors are in Berlin Festival?  
76  
77 CREATE VIEW [dirberlin] AS select d.movie_title, d.director, f.festival  
78     from director d ,festivalFilm f  
79     where d.movie_id = f.movie_id and festival like "%Berlin%"  
80  
81 select * from dirberlin;  
82
```

	movie_title	director	festival
1	Spectre	Sam mendes	Berlin
2	Harry Potter and the Half-Blood Prince	David Yates	Berlin
3	Quantum of Solace	Marc Forster	Berlin
4	Robin Hood	Otto Bathurst	Berlin
5	King Kong	Peter Jackson	Berlin
6	Skyfall	Sam Mendes	Berlin
7	X-Men: Days of Future Past	Steven Spielberg	Berlin
8	Up	David Fincher	Berlin
9	Monsters vs Aliens	Jonathan Demme	Berlin
10	Iron Man	George Lucas	Berlin
11	G.I. Joe: The Rise of Cobra	Martin Scorsese	Berlin
12	The Jungle Book	Ridley Scott	Berlin
13	Terminator 3: Rise of the Machines	Can Ulkay	Berlin
14	Inception	Nitesh Tiwari	Berlin
15	Alice Through the Looking Glass	Andrew Stanton	Berlin
16	Alexander	Stanley Kubrick	Berlin

Execution finished without errors.
Result: 49 rows returned in 12ms
At line 81:
select * from dirberlin;

4. Which actor were active in which year?

```

53
54
55 #Which actor were active in which year?
56
57  CREATE VIEW [activeactor] AS select m.movie_title, c.cast, m.release_date
58   from movie m, cast c
59   where m.movie_id=c.movie_id and c.cast like "%Robert Downey Jr%"
60
61
62 select * from activeactor;
63
64

```

	movie_title	cast	release_date
1	Avengers: Age of Ultron	Tony Stark / Iron Man edcde Robert Downey Jr. Tho...	4/22/15
2	The Avengers	Tony Stark / Iron Man fecaeb Robert Downey Jr. ...	4/25/12
3	Captain America: Civil War	Steve Rogers / Captain America ecbdeadb Chris Evans ...	4/27/16
4	Iron Man 3	Tony Stark / Iron Man fecaedff Robert Downey Jr. ...	4/18/13
5	Iron Man	Tony Stark / Iron Man fecafee Robert Downey Jr. Lt...	4/30/08
6	Iron Man 2	Tony Stark / Iron Man fec Robert Downey Jr. Virginin...	4/28/10
7	The Incredible Hulk	Bruce Banner / The Hulk fecafe Edward Norton Bett...	6/12/08
8	Sherlock Holmes: A Game of Shadows	Sherlock Holmes fecafba Robert Downey Jr. Dr. Jo...	11/22/11
9	Tropic Thunder	Tugg Speedman fedcae Ben Stiller Jeff Portnoy ...	8/9/08
10	Sherlock Holmes	Sherlock Holmes fecef Robert Downey Jr. Dr. John ...	12/23/09
11	Zodiac	Robert Graysmith fecafddd Jake Gyllenhaal Paul ...	3/2/07

Execution finished without errors.
Result: 11 rows returned in 11ms
At line 92:
select * from activeactor;

5.Which actor have done a Romantic Movie?

```

91
92
93
94 #Which actor have done a Romantic Movie?
95  CREATE VIEW [romantic] AS select c.cast from cast c,genre g where c.movie_id=g.movie_id and genre like "%Romance"
96
97 select * from romantic;
98
99

```

	cast
1	Jay Gatsby feecae Leonardo DiCaprio Nick ...
2	Prince Dastan fecafbe Jake Gyllenhaal Tamina ...
3	Maleficent feaccabfcb Angelina Jolie Princess Auro...
4	James Stewart / Jay Fennel feecae Josh Hartnett ...
5	Daisy fecacfb Cate Blanchett Benjamin Button ...
6	Alexander fecafed Colin Farrell Olympias fecafef ...
7	Isabella 'Bella' Swan fececaff Kristen Stewart Edwa...
8	Lisa Jorgenson fecafee Reese Witherspoon George...
9	Porter Stoddard fecae Warren Beatty Ellie Stoddard...
10	Jack Byrnes feaca Robert De Niro Greg Focker fea...
11	Frank Tupelo fec Johnny Depp Elise Clifton-Ward ...
12	Carrie Bradshaw feecc Sarah Jessica Parker ...
13	Giselle feccaff Amy Adams Robert Philip fecaffb ...
14	Amanda Woods fecaf Cameron Diaz Iris Simpkins ...
15	Jane Adler fecae Meryl Streep Jake Adler fecae...
16	Savuri fecacf Zhang Ziyi Hatsumomo fecacf ...

Execution finished without errors.
Result: 28 rows returned in 19ms
At line 97:
select * from romantic;