

# Comprehensive Analysis of Netflix Movies and TV Shows Using SQL

**Milestone: Implementation in NOSQL**

## **Group 08**

Sree Ramya Pasala  
Sri Lakshmi Swetha Jalluri

857-339-9096 (Tel of Student 1)

857-379-0666 (Tel of Student 2)

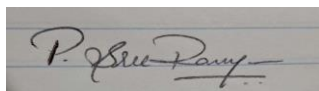
[Pasala.sr@northeastern.edu](mailto:Pasala.sr@northeastern.edu)

[jalluri.s@northeastern.edu](mailto:jalluri.s@northeastern.edu)

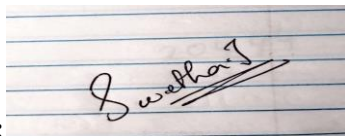
**Percentage of Effort Contributed by Student1: 50**

**Percentage of Effort Contributed by Student2: 50**

**Signature of Student 1:**



**Signature of Student 2:**



**Submission Date: 28-11-2024**

## 1. Write a query to find all the TV SHOW.

```
db.shows.find({ TYPE: "TV SHOW" })
```

Output:

```
< {
  _id: ObjectId('6748ed17f503f11e2568c36c'),
  SHOW_ID: 2,
  TYPE: 'TV SHOW',
  TITLE: 'SPECTRUM',
  DIRECTOR_NAME: 'BROOK',
  CAST_NAME: 'JACK',
  COUNTRY_NAME: 'US',
  DATE_ADDED: 2014-11-15T00:00:00.000Z,
  RELEASE_YEAR: 2014,
  RATING: 4,
  DURATION: '200MINS',
  LISTED_IN: 'DRAMA'
}
{
  _id: ObjectId('6748ed17f503f11e2568c36d'),
  SHOW_ID: 3,
  TYPE: 'TV SHOW',
  TITLE: 'CHOTA BHEEM',
  DIRECTOR_NAME: 'VISHAL',
  CAST_NAME: 'RAJIV',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2012-06-12T00:00:00.000Z,
  RELEASE_YEAR: 2009,
  RATING: 4.8,
  DURATION: '200MINS',
  LISTED_IN: 'DRAMA'
}
{
```

```
{
  _id: ObjectId('6748ed17f503f11e2568c36f'),
  SHOW_ID: 5,
  TYPE: 'TV SHOW',
  TITLE: 'HE MAN',
  DIRECTOR_NAME: 'ROBERT',
  CAST_NAME: 'MARY',
  COUNTRY_NAME: 'UK',
  DATE_ADDED: 2017-09-17T00:00:00.000Z,
  RELEASE_YEAR: 2015,
  RATING: 4,
  DURATION: '200MINS',
  LISTED_IN: 'MYSTRY'
}
{
  _id: ObjectId('6748ed17f503f11e2568c370'),
  SHOW_ID: 6,
  TYPE: 'TV SHOW',
  TITLE: 'JACK WHITE',
  DIRECTOR_NAME: 'NICK',
  CAST_NAME: 'NEEL',
  COUNTRY_NAME: 'US',
  DATE_ADDED: 2019-10-12T00:00:00.000Z,
  RELEASE_YEAR: 2019,
  RATING: 4.9,
  DURATION: '160MINS',
  LISTED_IN: 'COMEDY'
}
{
```

```
{
  _id: ObjectId('6748ed17f503f11e2568c372'),
  SHOW_ID: 8,
  TYPE: 'TV SHOW',
  TITLE: 'TONY',
  DIRECTOR_NAME: 'JACK',
  CAST_NAME: 'ROBERT',
  COUNTRY_NAME: 'US',
  DATE_ADDED: 2021-10-19T00:00:00.000Z,
  RELEASE_YEAR: 2021,
  RATING: 4,
  DURATION: '120MINS',
  LISTED_IN: 'SCIENTIFIC'
}
{
  _id: ObjectId('6748ed17f503f11e2568c373'),
  SHOW_ID: 9,
  TYPE: 'TV SHOW',
  TITLE: 'KOTA FACTORY',
  DIRECTOR_NAME: 'MAYUR',
  CAST_NAME: 'HARRI',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2020-11-09T00:00:00.000Z,
  RELEASE_YEAR: 2020,
  RATING: 4.6,
  DURATION: '1SEASON',
  LISTED_IN: 'DRAMA'
}
{
```

```
}
{
  _id: ObjectId('6748ed17f503f11e2568c377'),
  SHOW_ID: 13,
  TYPE: 'TV SHOW',
  TITLE: 'SQUID',
  DIRECTOR_NAME: 'ALEE',
  CAST_NAME: 'MAX',
  COUNTRY_NAME: 'JAPAN',
  DATE_ADDED: 2015-02-14T00:00:00.000Z,
  RELEASE_YEAR: 2015,
  RATING: 4.9,
  DURATION: '180MINS',
  LISTED_IN: 'THRILLER'
}
{
  _id: ObjectId('6748ed17f503f11e2568c378'),
  SHOW_ID: 14,
  TYPE: 'TV SHOW',
  TITLE: 'ANGRY BIRD',
  DIRECTOR_NAME: 'ANTTI',
  CAST_NAME: 'FINDLAND',
  COUNTRY_NAME: 'UK',
  DATE_ADDED: 2018-12-17T00:00:00.000Z,
  RELEASE_YEAR: 2012,
  RATING: 4.8,
  DURATION: '180MINS',
  LISTED_IN: 'THRILLER'
}
{
```

## 2. Write a query to retrieve all shows that have a rating greater than or equal to 4.5.

```
> db.shows.find({ RATING: { $gte: 4.5 } }).sort({ RATING: -1 })
```

Output:

```
{
  _id: ObjectId('6748ed17f503f11e2568c379'),
  SHOW_ID: 15,
  TYPE: 'MOVIE',
  TITLE: 'RRR',
  DIRECTOR_NAME: 'RAJAMOLI',
  CAST_NAME: 'NTR',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2020-01-18T00:00:00.000Z,
  RELEASE_YEAR: 2020,
  RATING: 4.5,
  DURATION: '180MINS',
  LISTED_IN: 'ACTION'
}
> db.shows.find({ RATING: { $gte: 4.5 } }).sort({ RATING: -1 })
< {
  _id: ObjectId('6748ed17f503f11e2568c370'),
  SHOW_ID: 6,
  TYPE: 'TV SHOW',
  TITLE: 'JACK WHITE',
  DIRECTOR_NAME: 'NICK',
  CAST_NAME: 'NEEL',
  COUNTRY_NAME: 'US',
  DATE_ADDED: 2019-10-12T00:00:00.000Z,
  RELEASE_YEAR: 2019,
  RATING: 4.9,
  DURATION: '160MINS',
  LISTED_IN: 'COMEDY'
```

```
{
  _id: ObjectId('6748ed17f503f11e2568c371'),
  SHOW_ID: 7,
  TYPE: 'MOVIE',
  TITLE: 'SCHOOL',
  DIRECTOR_NAME: 'KONG',
  CAST_NAME: 'THINY',
  COUNTRY_NAME: 'JAPAN',
  DATE_ADDED: 2018-01-18T00:00:00.000Z,
  RELEASE_YEAR: 2008,
  RATING: 4.9,
  DURATION: '250MINS',
  LISTED_IN: 'DOCUMENTATION'
}
{
  _id: ObjectId('6748ed17f503f11e2568c376'),
  SHOW_ID: 12,
  TYPE: 'MOVIE',
  TITLE: 'JEANS',
  DIRECTOR_NAME: 'SHANKAR',
  CAST_NAME: 'PRASH',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2010-04-10T00:00:00.000Z,
  RELEASE_YEAR: 2010,
  RATING: 4.9,
  DURATION: '240MINS',
  LISTED_IN: 'COMEDY'
}
{
```

```
{
  _id: ObjectId('6748ed17f503f11e2568c377'),
  SHOW_ID: 13,
  TYPE: 'TV SHOW',
  TITLE: 'SQUID',
  DIRECTOR_NAME: 'ALEE',
  CAST_NAME: 'MAX',
  COUNTRY_NAME: 'JAPAN',
  DATE_ADDED: 2015-02-14T00:00:00.000Z,
  RELEASE_YEAR: 2015,
  RATING: 4.9,
  DURATION: '180MINS',
  LISTED_IN: 'THRILLER'
}
{
  _id: ObjectId('6748ed17f503f11e2568c36d'),
  SHOW_ID: 3,
  TYPE: 'TV SHOW',
  TITLE: 'CHOTA BHEEM',
  DIRECTOR_NAME: 'VISHAL',
  CAST_NAME: 'RAJIV',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2012-06-12T00:00:00.000Z,
  RELEASE_YEAR: 2009,
  RATING: 4.8,
  DURATION: '200MINS',
  LISTED_IN: 'DRAMA'
}
{
```

```
  RATING: 4.8,
  DURATION: '120MINS',
  LISTED_IN: 'ACTION'
}
{
  _id: ObjectId('6748ed17f503f11e2568c378'),
  SHOW_ID: 14,
  TYPE: 'TV SHOW',
  TITLE: 'ANGRY BIRD',
  DIRECTOR_NAME: 'ANTTI',
  CAST_NAME: 'FINDLAND',
  COUNTRY_NAME: 'UK',
  DATE_ADDED: 2018-12-17T00:00:00.000Z,
  RELEASE_YEAR: 2012,
  RATING: 4.8,
  DURATION: '180MINS',
  LISTED_IN: 'THRILLER'
}
{
  _id: ObjectId('6748ed17f503f11e2568c373'),
  SHOW_ID: 9,
  TYPE: 'TV SHOW',
  TITLE: 'KOTA FACTORY',
  DIRECTOR_NAME: 'MAYUR',
  CAST_NAME: 'HARRI',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2020-11-09T00:00:00.000Z,
  RELEASE_YEAR: 2020,
  RATING: 4.6,
```

```

LISTED_IN: 'DOCUMENTATION'
}
{
  _id: ObjectId('6748ed17f503f11e2568c36b'),
  SHOW_ID: 1,
  TYPE: 'MOVIE',
  TITLE: 'DICK',
  DIRECTOR_NAME: 'KIRSTEN',
  CAST_NAME: 'JOHN',
  COUNTRY_NAME: 'US',
  DATE_ADDED: 2021-09-25T00:00:00.000Z,
  RELEASE_YEAR: 2020,
  RATING: 4.5,
  DURATION: '120MINS',
  LISTED_IN: 'DOCUMENTATION'
}
{
  _id: ObjectId('6748ed17f503f11e2568c379'),
  SHOW_ID: 15,
  TYPE: 'MOVIE',
  TITLE: 'RRR',
  DIRECTOR_NAME: 'RAJAMOLI',
  CAST_NAME: 'NTR',
  COUNTRY_NAME: 'INDIA',
  DATE_ADDED: 2020-01-18T00:00:00.000Z,
  RELEASE_YEAR: 2020,
  RATING: 4.5,
  DURATION: '180MINS',
  LISTED_IN: 'ACTION'
}
}

```

### 3. Write a query to count number of movies in the database?

```
> db.shows.countDocuments({ TYPE: "MOVIE" })
```

Output:

```
< 6
```

### 4. Write a query of how many shows are available from each country in the database.

```
> db.shows.aggregate([
  { $group: { _id: "$COUNTRY_NAME", count: { $sum: 1 } } }
])
< {
```

Output:

```
< {
  _id: 'JAPAN',
  count: 3
}
{
  _id: 'INDIA',
  count: 4
}
{
  _id: 'UK',
  count: 2
}
{
  _id: 'US',
  count: 6
}

```

5. Write a query of how many shows are there of each type in the database.

```
> db.shows.aggregate([
  { $group: { _id: "$TYPE", count: { $sum: 1 } } }
])
```

Output:

```
< {
  _id: 'MOVIE',
  count: 6
}
{
  _id: 'TV SHOW',
  count: 9
}
```

6. Write a query to calculate average ratings of TV SHOWS.

```
> db.shows.aggregate([
  { $match: { TYPE: "TV SHOW" } },
  { $group: { _id: null, avg_rating: { $avg: "$RATING" } } }
])
```

Output:

```
< {
  _id: null,
  avg_rating: 4.533333333333333
}
```

7. Write a query to Find All Shows by a Specific Director.

```
> db.shows.find({ DIRECTOR_NAME: "KIRSTEN" })
```

Output:

```
< {
  _id: ObjectId('6748ed17f503f11e2568c36b'),
  SHOW_ID: 1,
  TYPE: 'MOVIE',
  TITLE: 'DICK',
  DIRECTOR_NAME: 'KIRSTEN',
  CAST_NAME: 'JOHN',
  COUNTRY_NAME: 'US',
  DATE_ADDED: 2021-09-25T00:00:00.000Z,
  RELEASE_YEAR: 2020,
  RATING: 4.5,
  DURATION: '120MINS',
  LISTED_IN: 'DOCUMENTATION'
}
```

8. Write a query to find the number of TV Shows in the US from database.

```
> db.shows.aggregate([
  { $match: { TYPE: "TV SHOW", COUNTRY_NAME: "US" } },
  { $count: "tv_shows_in_us" }
])
```

Output:

```
< {
  tv_shows_in_us: 3
}
```

9. Write a query to display only the director name from the database.

```
> db.shows.find({}, { DIRECTOR_NAME: 1, _id: 0 })
```

Output:

```
< {
  DIRECTOR_NAME: 'KIRSTEN'
}
{
  DIRECTOR_NAME: 'BROOK'
}
{
  DIRECTOR_NAME: 'VISHAL'
}
{
  DIRECTOR_NAME: 'STEVEN'
}
{
  DIRECTOR_NAME: 'ROBERT'
}
{
  DIRECTOR_NAME: 'NICK'
}
{
  DIRECTOR_NAME: 'KONG'
}
{
  DIRECTOR_NAME: 'JACK'
}
{
  DIRECTOR_NAME: 'MAYUR'
}
{
  DIRECTOR_NAME: 'MIKE'
}
{
  DIRECTOR_NAME: 'MIKE'
}
{
  DIRECTOR_NAME: 'ALEX'
}
{
  DIRECTOR_NAME: 'SHANKAR'
}
{
  DIRECTOR_NAME: 'ALEE'
}
{
  DIRECTOR_NAME: 'ANTTI'
}
{
  DIRECTOR_NAME: 'RAJAMOLI'
}
```

10. Write a query to find the cast names of TV shows from database.

```
> db.shows.find({ TYPE: "TV SHOW" }, { CAST_NAME: 1, _id: 0 })
```

Output:

```
< {
  CAST_NAME: 'JACK'
}
{
  CAST_NAME: 'RAJIV'
}
{
  CAST_NAME: 'MARY'
}
{
  CAST_NAME: 'NEEL'
}
{
  CAST_NAME: 'ROBERT'
}
{
  CAST_NAME: 'HARRI'
}
{
  CAST_NAME: 'KATE'
}
{
  CAST_NAME: 'MAX'
}
{
  CAST_NAME: 'FINDLAND'
}
```