

1 Постановка задачи

Цели лабораторной работы:

1. Язык SQL. Выборка данных с использованием операторов SELECT. WHERE. GROUP BY. HAVING. ORDER BY.
2. Язык SQL. Объединения JOIN.

2 Получение данных

Дамп базы данных IMDB.

https://drive.google.com/open?id=1MYZbB21AFgFHSuoiNU_I2xonnE4dX8pj

1. **IMDB_DB_DUMP** - Полный дамп базы данных IMDB. Размер 7GB.
2. **IMDB_DB_DUMP_RATING** - Дамп базы данных IMDB в котором убрана информация про фильмы для которых нет рейтинга.
3. **IMDB_DB_DUMP_RATING_SORT_100K** - Дамп базы данных IMDB с информацией про фильмы из топ 100к в рейтинге.
4. **IMDB_DB_DUMP_RATING_SORT_10K** - Дамп базы данных IMDB с информацией про фильмы из топ 10к в рейтинге.

Физическая модель дампа базы данных:

3 Запросы

Получение всех фильмов отсортированных по рейтингу в порядке убывания.

```
SELECT originalTitle , titleType , averageRating , numVotes FROM title
INNER JOIN title_rating ON title.titleID = title_rating.titleID
ORDER BY numVotes DESC, averageRating DESC
LIMIT 10;
```

```
+-----+-----+-----+-----+
| originalTitle | titleType | averageRating | numVotes |
+-----+-----+-----+-----+
| The Shawshank Redemption | movie | 9.3 | 2149031 |
| The Dark Knight | movie | 9 | 2118625 |
| Inception | movie | 8.8 | 1883746 |
| Fight Club | movie | 8.8 | 1716980 |
| Pulp Fiction | movie | 8.9 | 1686388 |
| Forrest Gump | movie | 8.8 | 1653584 |
| Game of Thrones | tvSeries | 9.4 | 1596735 |
| The Matrix | movie | 8.7 | 1546559 |
| The Lord of the Rings: The Fellowship of the Ring | movie | 8.8 | 1542044 |
| The Lord of the Rings: The Return of the King | movie | 8.9 | 1526805 |
+-----+-----+-----+-----+
10 rows in set (0.47 sec)
```

Получение количества фильмов соответствующих жанру отсортированных по убыванию:

```
SELECT genre , COUNT(genre) as genre_count FROM title_genre
GROUP BY genre
ORDER BY genre_count DESC
LIMIT 10;
```

genre	genre_count
Drama	1692898
Comedy	1316902
Short	842509
Documentary	603245
Talk-Show	587668
Romance	576145
Family	432405
News	401786
Reality-TV	319441
Animation	313577

10 rows in set (3.99 sec)

Получение всех фильмов с заголовком Batman или Dark knight отсортированных по рейтингу.

```
SELECT title.titleID , originalTitle , averageRating , numVotes FROM title
INNER JOIN title_rating ON title.titleID = title_rating.titleID
WHERE originalTitle LIKE '%Batman%' OR originalTitle LIKE '%Dark_Knight%'
ORDER BY title_rating.numVotes DESC, title_rating.averageRating DESC
LIMIT 5
```

titleID	originalTitle	averageRating	numVotes
tt0468569	The Dark Knight	9	2118625
tt1345836	The Dark Knight Rises	8.4	1415027
tt0372784	Batman Begins	8.2	1219308
tt2975590	Batman v Superman: Dawn of Justice	6.5	587679
tt0096895	Batman	7.5	318814

5 rows in set (0.48 sec)

Получить список людей ставших известными за фильм Batman(id='tt0468569');

```
SELECT primaryName FROM name_known_for_title
INNER JOIN name ON name_known_for_title.nameID = name.nameID
WHERE titleID = 'tt0468569'
LIMIT 5
```

primaryName
Bob Cadigan
Janice Buchanan
Brian Allen
Kendis Rochlen

4 rows in set (0.00 sec)

Получить актерский состав фильма Batman(id='tt0468569');

```

SELECT primaryName, profession FROM title_principals_profession
INNER JOIN name ON title_principals_profession.nameID = name.nameID
INNER JOIN profession ON profession.professionID = title_principals_pr
WHERE titleID = 'tt0468569'

```

```

+-----+-----+
| primaryName | profession |
+-----+-----+
| Christian Bale | actor |
| Michael Caine | actor |
| Aaron Eckhart | actor |
| Bob Kane | writer |
| Heath Ledger | actor |
| David S. Goyer | writer |
| Christopher Nolan | director |
| Jonathan Nolan | writer |
| Lorne Orleans | producer |
| Charles Roven | producer |
+-----+-----+
10 rows in set (0.00 sec)

```

4 Используемые источники

1. <https://dev.mysql.com/doc/refman/8.0/en/select.html>
2. <https://dev.mysql.com/doc/refman/8.0/en/select-into.html>
3. <https://dev.mysql.com/doc/refman/8.0/en/join.html>
4. <https://dev.mysql.com/doc/refman/8.0/en/union.html>
5. <https://dev.mysql.com/doc/refman/8.0/en/subqueries.html>

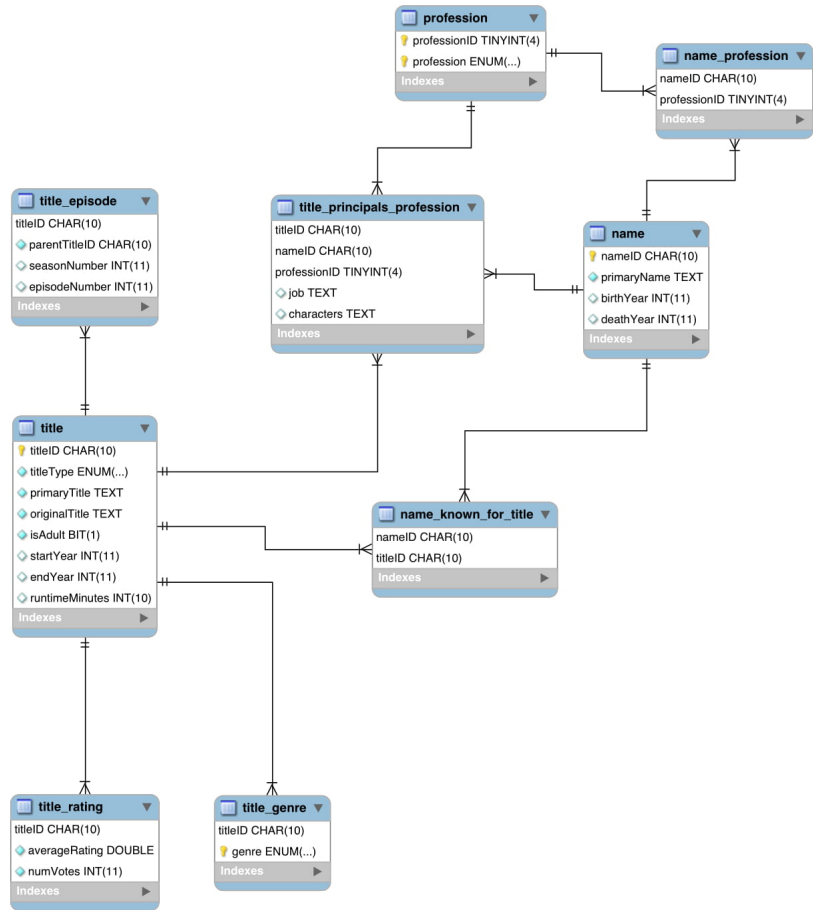


Figure 1: Physical model