Final Project: IMDb Clone & Data Analysis

Based on IMDb non-commercial dataset, create an IMDb clone and data analysis

Part 1: Database Structure & Data Import

Your first task is to design and implement the database structure that will support all features of parts 2 and 3.

Requirements:

1. Design tables to store:

- Movies and TV Series (basic info)
- Episodes (linking to their series)
- Detailed information (alternative titles, production details...)
- People (actors, directors, writers...)
- Relationships between people and titles
- Genres
- Ratings and votes

2. Consider:

- Appropriate primary and foreign keys
- Indexes for performance
- Data types optimization
- NULL constraints
- Referential integrity

3. Import Data:

- Create INSERT statements or import scripts
- Validate data integrity
- Handle potential duplicates
- Create test data if needed

Part 2: Web Application

Create the SQL queries that you will need to create a web application with some of these suggested pages (at least 5) and features (at least 1 general feature like adult content or genre filtering).
For each query, explain what's its purpose and how you wrote it.

Suggested Pages:

1. Movie Summary Page

- Title, year, length
- Directors, writers
- Main cast (top 5)
- Rating, votes
- Basic information

2. Movie Details Page

- Alternative titles
- Production details
- Full technical details

3. Complete Cast/Crew Page

- All actors with roles
- Full crew listing
- Sorted by categories

4. TV Series Summary Page

- Series overview
- Seasons count
- Years active
- Main cast
- General rating

5. TV Series Details Page

- Season-by-season breakdown
- Cast changes
- Production details

6. Episode Page

- Episode-specific details
- Guest stars
- Individual rating

7. Person Page

- Basic information
- Complete filmography
- Categorized by role type

8. Movie Listing Page

- Filterable by:
    - Genre
    - Year
    - Rating
    - Other criteria

9. Series Listing Page

- Similar filters as movies
- Additional TV-specific filters

10. Home Page

    - Featured content
    - Navigation to all sections

Suggested Features:

- Adult content filtering
- Search Engine
- Advanced person-to-person search
- Add your own creative features

Part 3: Data Analysis & Visualization

Create the SQL queries that you will need to create some of these data analysis reports (at least 3 differents):

Suggested Analyses:

1. Rating Trends

- By year/decade
- By genre
- By country

2. Performance Analysis

- Directors' average ratings
- Actors' film success
- Genre popularity over time

3. Relationship Analysis

- Collaboration networks
- Genre combinations

- Success patterns

4. Custom Analysis

- Propose and implement your own analysis
- Document your methodology
- Explain insights

Evaluation Criteria:

1. Database Design

- Structure efficiency
- Query performance
- Data integrity

2. Web Application

- Is it fast ? ;)
- Performance

3. Data Analysis

- Usage of advanced SQL functionalities like GROUP BY, HAVING and Window Functions
- Performance
- and performance

4. Documentation

- Queries comments and documentation
- Database setup instructions

You can submit the project the way you want : a single document (PDF only), a MarkDown file, multiple MarkDown files, a github repository with SQL files, … or any combination of that, **AS LONG AS IT'S CLEAR AND EASY TO NAVIGATE AND FIND INFORMATIONS**.

For example :

```
project/
├── README.md (project overview)
├── schema/
│   ├── schema.sql
│   └── schema.png (diagram)
├── import/
│   └── import.sql
├── queries/
│   ├── web_queries.sql
```

```
│   └── analysis_queries.sql
├── documentation/
│   └── queries_explanation.md
└── analysis/
    └── visualizations.pdf
```

Submission Requirements:

1. Database schema
2. Import script or queries
3. All the queries, and for each one of them, comments and documentation
4. Analysis report (spreadsheet graphs for example)