Certification Project

Problem Statement

edureka!

edureka!

Version 1.0

Problem Statement:

For project you have to parse the file movies.txt which contains the data for 3451 movies. Read the movies.txt file and store all the movie data into MySQL database table movies. For each movie, data is collected under 11 columns in movies.txt file as described below

Title: Movie name

Title type: Movie type like documentary, short film, feature film etc.

Directors: Name of the movie directors

IMDb Rating: IMDb Rating of the movie (IMDb- Internet Movie Database)

Runtime (mins): Duration of the movie in minutes

Year: Year in which movie was released

Genres: Categories of movie like drama, romance, and thriller

# of IMDb votes: Number of votes it got on IMDb

IMDb Top 250: Whether a movie is on IMDb top 250 list or not (If listed then Y else N)

1001 Must See: Whether a movie is listed under 1001 Must See list (If listed then Y else N)

URL: IMDb URL where you can see the movie description like
http://www.imdb.com/title/tt0453562/

Note: While parsing movies.txt file, for some of the columns data is not present like for first movie (11:14) in the file movies.txt columns IMDb Top 250 and 1001 Must are just blank. Movies.txt file is a semi-colon (;) separated file where each value is separated from another using semi-colon.

Once you load all the movie data into a MySQL database table, write JDBC SQL statements(use Prepared Statements rather than using simple JDBC statements) to get answer of the following queries.

1. Find out the name of Director who directed movie A Walk to Remember

2. List down the names of all the documentary movies

3. List down the names of all documentary movies released in 2000

4. Find out the names of all movies directed by Steven Spielberg

5. Find out the names of all movies directed by Steven Spielberg and were released in 2012

6. Find out the year in which maximum number of movies were released

7. Find out the movie names which belongs to Genre 'romance'

8. Find out the top 10 movies with Genre 'romance' which got highest number of votes on IMDb

9. Find out the movie with Genre 'thriller' which have the highest runtime and released after 2005

10. Find out average IMDb rating for movies (While calculating average exclude the movies for which IMDb rating is not known)

11. Download the movie dataset (movies.txt) from here

edureka!