SQL PROJECT: 1. MOVIE DATABASE

Scope of the Project: Small Project Start date: 14-06-2023

Project Expected End date: 19-06-2023

Project completion date: 15-06-2023

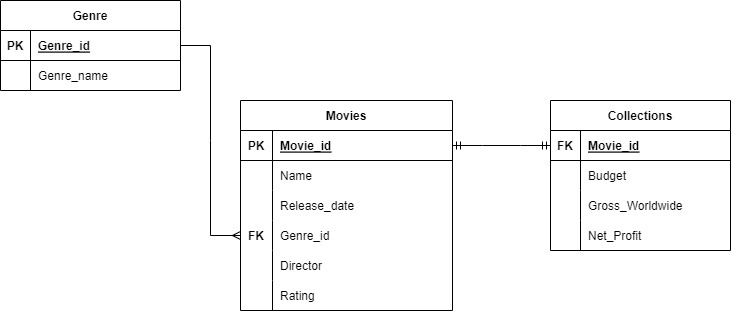
( 4 days early)

Objective:

Creating a Movies Database which stores data about movies released in the 2000-2010 and Analysing it to find information and insights regarding the factors leading to a financial and critical success of a movie.

Step 1: Designing a Database:-

* After some brainstorming and research, I have come up with a database design which suits the needs and scope of the project.
* I have used draw.io software to design a ER diagram of the database.



* The above image is the ER diagram of the Movies database.
* ER Diagram:-
* It has 3 tables. They are:-

1. Movies:

This table has columns:-

Movie\_id

Name

Release\_date

Genre\_id

Director

Rating

* Constraints:-
* This table has a Primary key on Movie\_id.
* This table has a Foreign Key on Genre\_id which references the column Genre\_id from the Genre table.

1. Genre:-

This table has columns:-

1. Genre\_id
2. Genre\_name

* Constraints:
* This table has a Primary key on Genre\_id

3.Collections:-

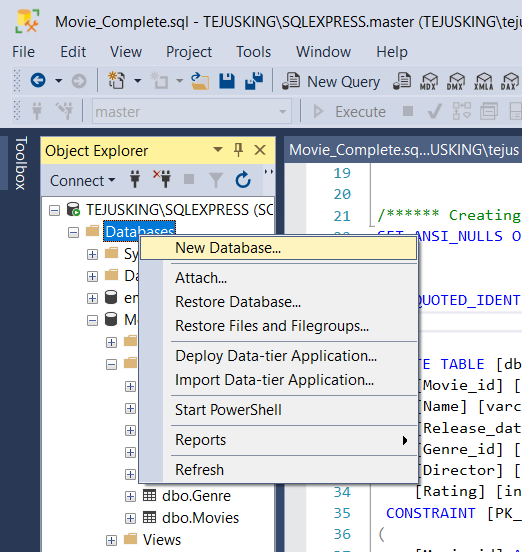
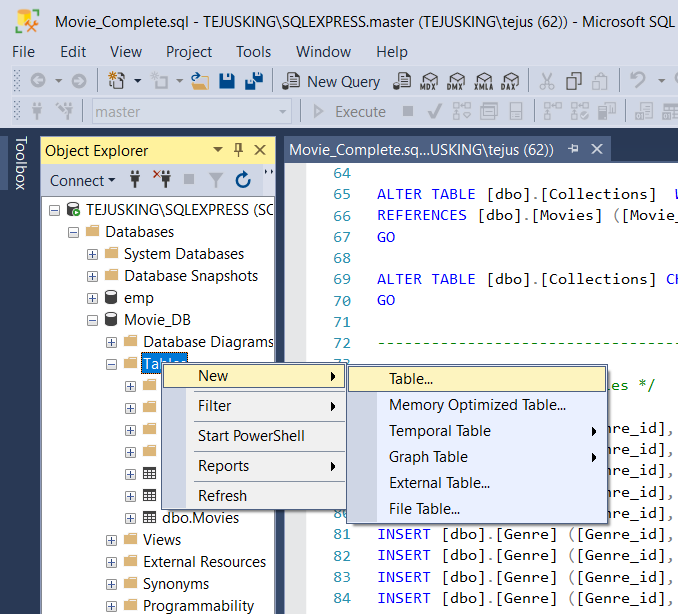
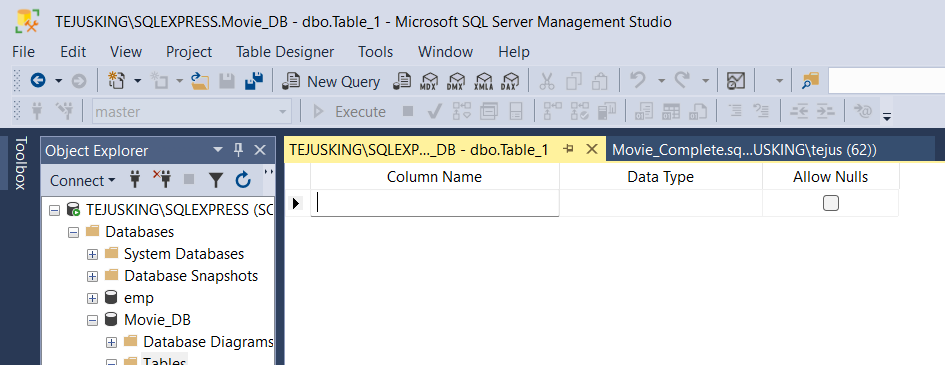
This table has columns:-

1. Movie\_id
2. Budget
3. Worldwide\_gross
4. Net\_Profit

* Constraints:
* This table has a Foreign key constraint on Movie\_id which references the column Movie\_id from Movies table.

This is the overview of the database and the ER Diagram of the Database.

Step 2:- Creating the Database:-

* I have used SSMS (SQL Server Management Studio) to execute all the SQL Queries in this project.
* I have created the new database by using the new database option.
* 
* I have used the built in create new table to define all the columns of all the 3 tables.
* 
* 
* Using the above steps I have created the 3 tables.
* Using the Alter table commands, I have added the required constraints to the tables.

By following the above steps ,I have created the database and the tables.

Step 3: Entering the data:-

* By using the Insert Statement , I have populated the 3 tables with data.
* My source of data is [Top 100 Movies of the 2010s so far - IMDb](https://www.imdb.com/list/ls003501243/?st_dt=&mode=detail&page=1&sort=user_rating,desc&title_type=movie&ref_=ttls_ref_typ)
* By following the above steps, I have Inserted data into the Tables.

Step 4: Creating Queries:-

* I have created SQL Queries and Stored Procedures to retrieve various information from the Database.
* Some of the Stored Procedures include:

FindMovieByGenre

FindMovieByDirector

FindMovieByYear

* By using these queries and stored procedures, It’s very easy to retrieve data.

Step 5: Testing the queries:-

* I have tested the queries performance and accuracy and the results are upto the requirements.

Step 6: Analysing Data and Creating Visualizations:-

* This step is performed by connecting the data to the Tableau and analysing the data to find insights and information.
* I have created the visualizations using tableau.

Step 7: Summarizing the Insights:-

* The insights I have found are :-

Action is the most profitable and most popular film genre in the 2010’s.

Animation has a higher average net profit and highest average rating than any other film genres.

Russo brothers, Christopher Nolan and Joss Whedon are the top 3 most profitable directors.

Action films directed by Joss Whedon, Russo Brothers and Christopher Nolan have been huge successes in the 2010’s.

**SUMMARY:-**

This project has revealed new insights about the movies released between 2010-2019, The genres and directors of the time period.