

**PROJECT FOR MYSQL**

**Hollywood Theatrical Market Synopsis 1995 to 2021**

****

**NAME: Caesar Edwin**

**Email: edwincaesar029@gmail.com**

**Institute Name: IT Vedant**

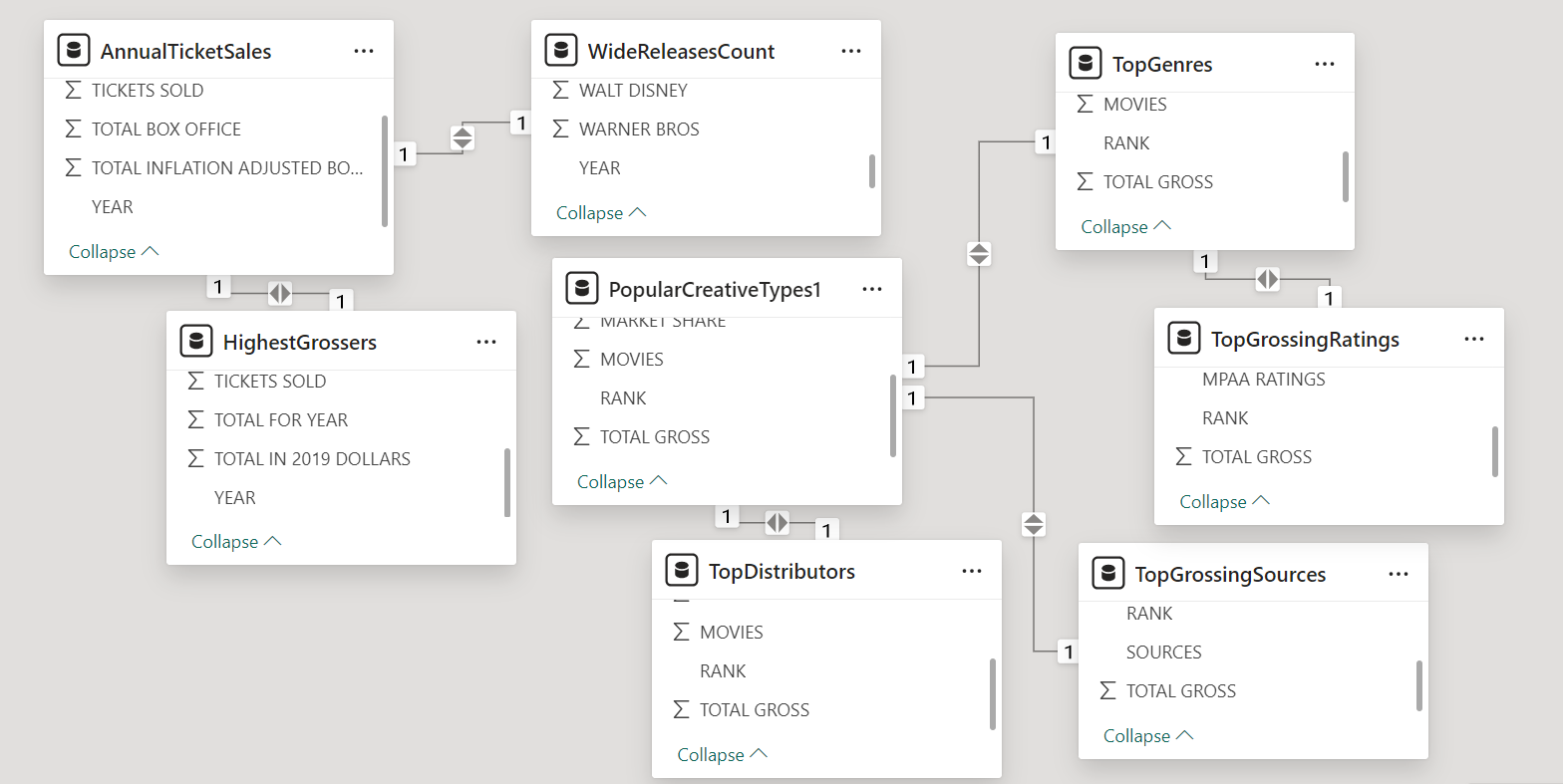
**Project Aim**

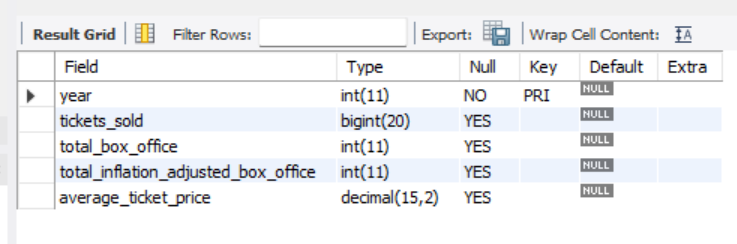
To analyse and explore the trends, performance, and distribution patterns in the Hollywood theatrical market from 1995 to 2021 using a relational database, providing insights into ticket sales, box office performance, distributor strategies, genre preferences, creative types, and MPAA ratings.

**Project Objectives**

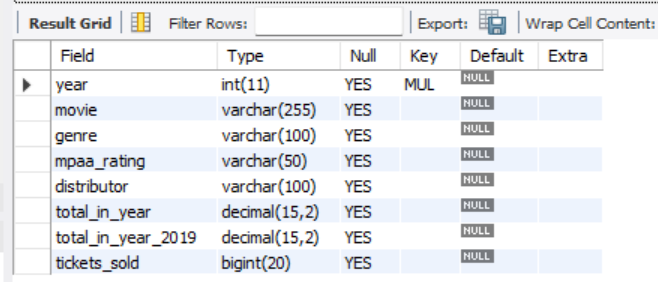
1. **Data Integration and Organization**
   * To design and implement a structured database using MySQL for managing Hollywood theatrical market data across multiple dimensions like ticket sales, movie genres, distributors, and MPAA ratings.
2. **Trend Analysis**
   * To analyse year-wise trends in ticket sales, box office revenue, and average ticket prices to evaluate the market’s growth and inflation-adjusted performance.
3. **Genre and Creative Type Insights**
   * To identify the most popular genres and creative types over the years by analysing their market share, total gross, and movie count.
4. **Distributor Performance**
   * To assess the performance of major distributors in terms of total gross, average gross, and market share.
5. **MPAA Ratings and Audience Patterns**
   * To study the impact of MPAA ratings on movie performance and identify which ratings dominate the market over time.
6. **Correlation Analysis**
   * To explore correlations between factors like ticket sales, average ticket prices, and inflation-adjusted box office gross.
7. **Query Optimization and Insights**
   * To demonstrate database querying techniques, including simple, subqueries, and joins, to extract meaningful insights from the dataset efficiently.
8. **Visualization and Reporting**
   * To present the findings through dashboards, charts, and reports, making it easier to interpret the market trends and patterns.

**ER DIAGRAM OF THE PROJECT**

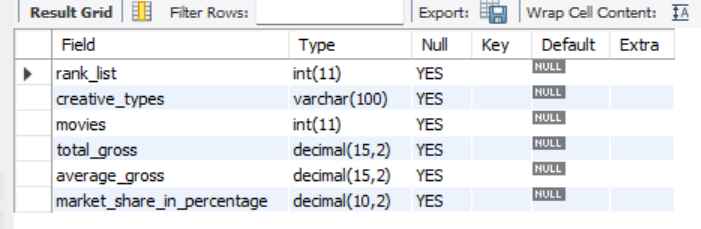
****

**TABLE DESCRIPTION 1. Annual Tickets Sales**

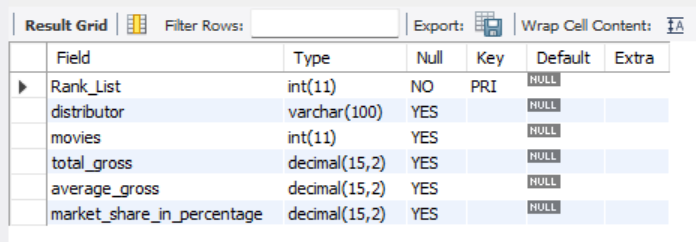
**TABLE DESCRIPTION 2. Highest Grosser**



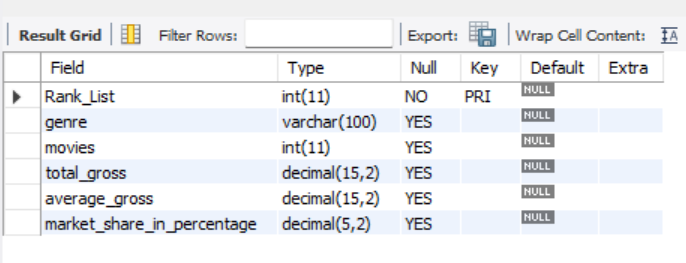
**TABLE DESCRIPTION 3. Popular Creative Types**



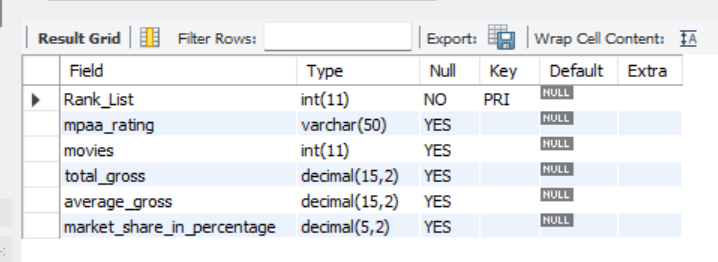
**TABLE DESCRIPTION 4. Top Distributor**



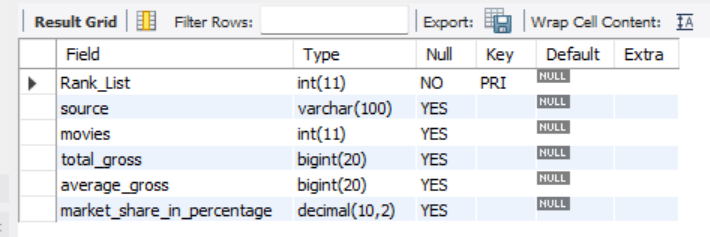
**TABLE DESCRIPTION 5. Top Genre**



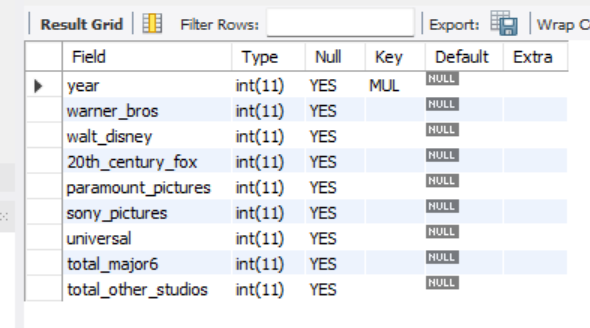
**TABLE DESCRIPTION 6. Top Grossing Rating**



**TABLE DESCRIPTION 7. Top Grossing Source**



**TABLE DESCRIPTION 8. Wide Release Count**

****

# COMMANDS

Create database Hollywood;

Use database Hollywood;

CREATE TABLE Annual\_Ticket\_Sales (

year INT PRIMARY KEY,

tickets\_sold BIGINT,

total\_box\_office int,

total\_inflation\_adjusted\_box\_office int,

average\_ticket\_price decimal(15,2) );

INSERT INTO Annual\_tickets\_sales (YEAR, TICKETS SOLD, TOTAL BOX OFFICE, TOTAL INFLATION ADJUSTED BOX OFFICE, AVERAGE TICKET PRICE) VALUES

('2021', '423774881', '3881777912', '3881777912', '9.16'),

('2020', '223638958', '2048534616', '2048534616', '9.16'),

('2019', '1228541629', '11253443955', '11253444050', '9.16'),

('2018', '1311536128', '11948096650', '12013670952', '9.11'),

('2017', '1225639761', '10993991460', '11226860216', '8.97'),

('2016', '1302556378', '11267115924', '11931416424', '8.65'),

('2015', '1323356776', '11155900636', '12121948075', '8.43'),

('2014', '1257402920', '10272985008', '11517810744', '8.17'),

('2013', '1339168926', '10887446341', '12266787382', '8.13'),

('2012', '1380921942', '10992141616', '12649244986', '7.96'),

('2011', '1282915168', '10173519704', '11751502955', '7.93'),

('2010', '1328549021', '10482254025', '12169509032', '7.89'),

('2009', '1418567388', '10639257284', '12994051137', '7.5'),

('2008', '1358042073', '9750744148', '12439665380', '7.18'),

('2007', '1420036680', '9769854914', '13007535993', '6.88'),

('2006', '1398738283', '9161738221', '12812442671', '6.55'),

('2005', '1372980280', '8800805718', '12576499367', '6.41'),

('2004', '1495651298', '9287996519', '13700165883', '6.21'),

('2003', '1524589620', '9193277289', '13965240914', '6.03'),

('2002', '1575756527', '9155147215', '14433929789', '5.81'),

('2001', '1465874205', '8296849636', '13427407722', '5.66'),

('2000', '1397460079', '7532311479', '12800734319', '5.39'),

('1999', '1444664086', '7338894852', '13233123027', '5.08'),

('1998', '1443832471', '6771575283', '13225505439', '4.69'),

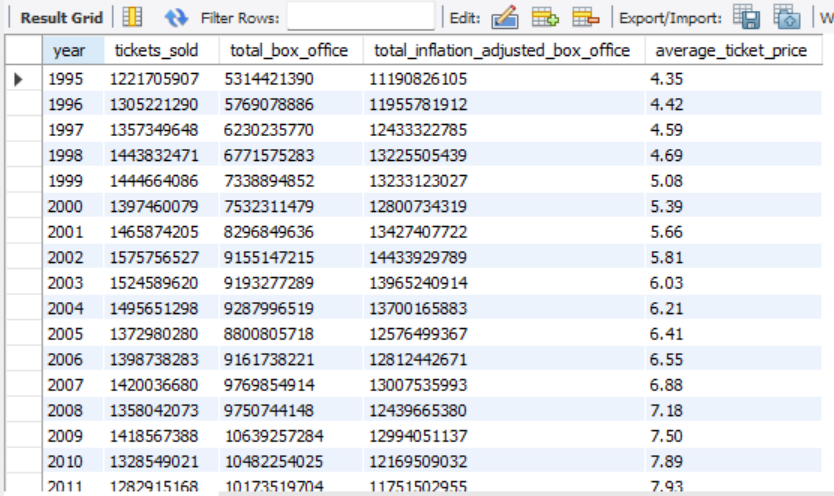
('1997', '1357349648', '6230235770', '12433322785', '4.59'),

('1996', '1305221290', '5769078886', '11955781912', '4.42'),

('1995', '1221705907', '5314421390', '11190826105', '4.35');

Select \* from annual\_ticket\_sales

**output**



CREATE TABLE\_Highest\_Grosser (

year INT,

movie VARCHAR(255),

genre VARCHAR(100),

mpaa\_rating VARCHAR(50),

distributor VARCHAR(100),

total\_in\_year DECIMAL(15, 2),

total\_in\_year\_2019 DECIMAL(15, 2),

tickets\_sold BIGINT,

FOREIGN KEY (year) REFERENCES Annual\_Ticket\_Sales(year)

);

INSERT INTO highest\_grosser (YEAR, MOVIE, GENRE, MPAA RATING, DISTRIBUTOR, TOTAL FOR YEAR, TOTAL IN 2019 DOLLARS, TICKETS SOLD, null) VALUES

('1995', 'Batman Forever', 'Drama', 'PG-13', 'Warner Bros.', '184031112', '387522978', '42306002'),

('1996', 'Independence Day', 'Adventure', 'PG-13', '20th Century Fox', '306169255', '634504608', '69269062'),

('1997', 'Men in Black', 'Adventure', 'PG-13', 'Sony Pictures', '250650052', '500207943', '54607854'),

('1998', 'Titanic', 'Adventure', 'PG-13', 'Paramount Pictures', '443319081', '865842808', '94524324'),

('1999', 'Star Wars Ep. I: The Phantom Menace', 'Adventure', 'PG', '20th Century Fox', '430443350', '776153749', '84732942'),

('2000', 'How the Grinch Stole Christmas', 'Adventure', 'PG', 'Universal', '253367455', '430583644', '47006948'),

('2001', 'Harry Potter and the Sorcerer’s Stone', 'Adventure', 'PG', 'Warner Bros.', '300404434', '486166890', '53074988', ''),

('2002', 'Spider-Man', 'Adventure', 'PG-13', 'Sony Pictures', '403706375', '636480273', '69484746'),

('2003', 'Finding Nemo', 'Adventure', 'G', 'Walt Disney', '339714367', '516050346', '56337374'),

('2004', 'Shrek 2', 'Adventure', 'PG', 'DreamWorks SKG', '441226247', '650826473', '71050925'),

('2005', 'Star Wars Ep. III: Revenge of the Sith', 'Action', 'PG-13', '20th Century Fox', '380270577', '543413171', '59324582'),

('2006', 'Pirates of the Caribbean: Dead Mans Chest', 'Action', 'PG-13', 'Walt Disney', '423315812', '591995851', '64628368'),

('2007', 'Spider-Man 3', 'Adventure', 'PG-13', 'Sony Pictures', '336530303', '448054878', '48914288'),

('2008', 'The Dark Knight', 'Adventure', 'PG-13', 'Warner Bros.', '531001578', '677433772', '73955652'),

('2009', 'Transformers: Revenge of the Fallen', 'Action', 'PG-13', 'Paramount Pictures', '402111870', '491112631', '53614916'),

('2010', 'Toy Story 3', 'Action', 'G', 'Walt Disney', '415004880', '481805411', '52598844'),

('2011', 'Harry Potter and the Deathly Hallows: Part II', 'Action', 'PG-13', 'Warner Bros.', '381011219', '440108798', '48046812'),

('2012', 'The Avengers', 'Adventure', 'PG-13', 'Walt Disney', '623357910', '717331462', '78311295'),

('2013', 'Iron Man 3', 'Adventure', 'PG-13', 'Walt Disney', '408992272', '460808016', '50306552'),

('2014', 'Guardians of the Galaxy', 'Adventure', 'PG-13', 'Walt Disney', '333055258', '373413235', '40765637'),

('2015', 'Star Wars Ep. VII: The Force Awakens', 'Action', 'PG-13', 'Walt Disney', '742208942', '806480887', '88043765'),

('2016', 'Finding Dory', 'Action', 'PG', 'Walt Disney', '486295561', '514967322', '56219140'),

('2017', 'Star Wars Ep. VIII: The Last Jedi', 'Action', 'PG-13', 'Walt Disney', '517218368', '528173936', '57660910'),

('2018', 'Black Panther', 'Action', 'PG-13', 'Walt Disney', '700059566', '703901821', '76845177'),

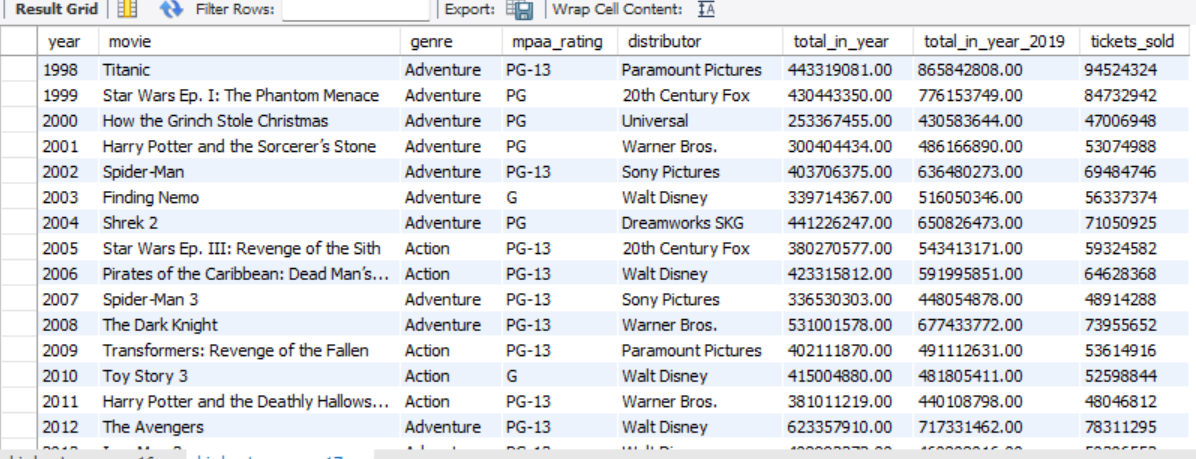
('2019', 'Avengers: Endgame', '', 'PG-13', 'Walt Disney', '858373000', '858373002', '93708843'),

('2020', 'Bad Boys For Life', '', 'R', 'Sony Pictures', '204417855', '204417848', '22316359'),

('2021', 'Shang-Chi and the Legend of the Ten Rings', '', 'PG-13', 'Walt Disney', '224226704', '224226704', '24478897');

Select \* from Highest\_grosser

**Output**



CREATE TABLE Popular\_Creative\_Types (

rank\_list INT,

creative\_types VARCHAR(100),

movies INT,

total\_gross DECIMAL(15, 2),

average\_gross DECIMAL(15, 2),

market\_share\_in\_percentage decimal(10,2)

);

INSERT INTO Popular\_creatice\_types (RANK, CREATIVE TYPES, MOVIES, TOTAL GROSS, AVERAGE GROSS, MARKET SHARE) VALUES

('1', 'Contemporary Fiction', '7442', '96203727036', '12927133', '40.46%'),

('2', 'Kids Fiction', '564', '32035539746', '56800602', '13.47%'),

('3', 'Science Fiction', '724', '29922660857', '41329642', '12.59%'),

('4', 'Fantasy', '759', '21724062575', '28621953', '9.14%'),

('5', 'Super Hero', '129', '20273157911', '157156263', '8.53%'),

('6', 'Historical Fiction', '1487', '18521260744', '12455454', '7.79%',),

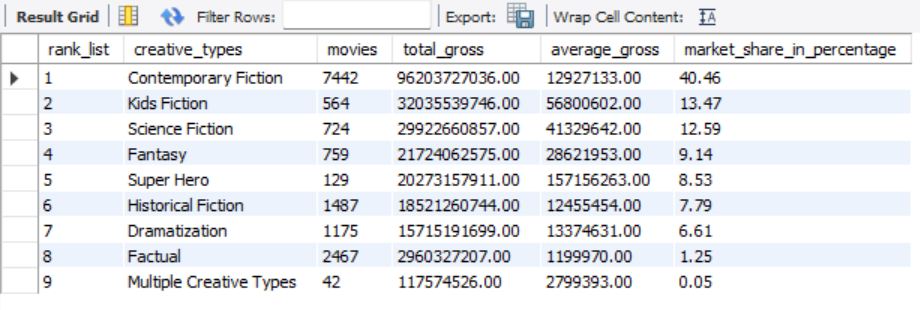
('7', 'Dramatization', '1175', '15715191699', '13374631', '6.61%'),

('8', 'Factual', '2467', '2960327207', '1199970', '1.25%'),

('9', 'Multiple Creative Types', '42', '117574526', '2799393', '0.05%');

Select \* from Popular\_creative\_types

**Output**



CREATE TABLE Top\_Distributor (

Rank\_List INT PRIMARY KEY,

distributor VARCHAR(100),

movies INT,

total\_gross DECIMAL(15, 2),

average\_gross DECIMAL(15, 2),

market\_share\_in\_percentage decimal(15,2)

);

INSERT INTO Top\_Distributor (RANK, DISTRIBUTORS, MOVIES, TOTAL GROSS, AVERAGE GROSS, MARKET SHARE) VALUES

('1', 'Walt Disney', '588', '40472424278', '68830654', '17.02%'),

('2', 'Warner Bros.', '824', '36269425479', '44016293', '15.25%'),

('3', 'Sony Pictures', '747', '29113002302', '38973229', '12.24%'),

('4', 'Universal', '535', '28089932569', '52504547', '11.81%'),

('5', '20th Century Fox', '525', '25857839756', '49253028', '10.88%'),

('6', 'Paramount Pictures', '493', '24361425304', '49414656', '10.25%'),

('7', 'Lionsgate', '426', '9631837781', '22609948', '4.05%'),

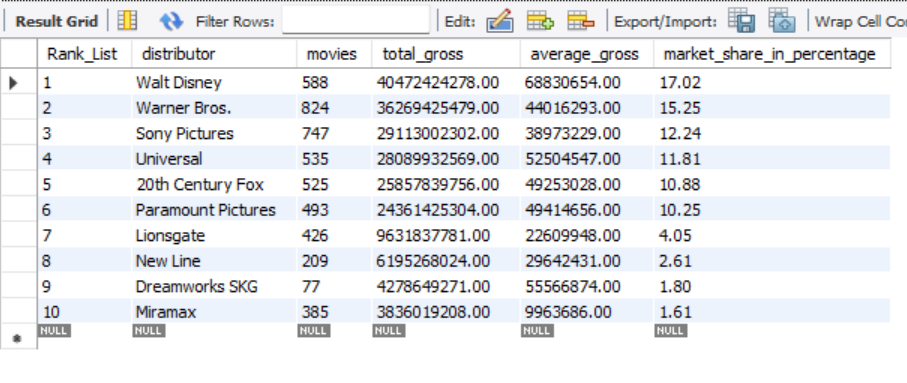
('8', 'New Line', '209', '6195268024', '29642431', '2.61%'),

('9', 'Dreamworks SKG', '77', '4278649271', '55566874', '1.80%'),

('10', 'Miramax', '385', '3836019208', '9963686', '1.61%');

Select \* from Top\_Distributor

**Output**



CREATE TABLE Top\_Genre (

Rank\_List INT PRIMARY KEY,

genre VARCHAR(100),

movies INT,

total\_gross DECIMAL(15, 2),

average\_gross DECIMAL(15, 2),

market\_share\_in\_percentage DECIMAL(5, 2)

);

INSERT INTO Top\_Genre (RANK, GENRES, MOVIES, TOTAL GROSS, AVERAGE GROSS, MARKET SHARE, ) VALUES

('1', 'Adventure', '1102', '64529536530', '58556748', '27.14'),

('2', 'Action', '1098', '49339974493', '44936224', '20.75'),

('3', 'Drama', '5479', '35586177269', '6495013', '14.97'),

('4', 'Comedy', '2418', '33687992318', '13932172', '14.17'),

('5', 'Thriller/Suspense', '1186', '19810201102', '16703374', '8.33'),

('6', 'Horror', '716', '13430378699', '18757512', '5.65'),

('7', 'Romantic Comedy', '630', '10480124374', '16635118', '4.41'),

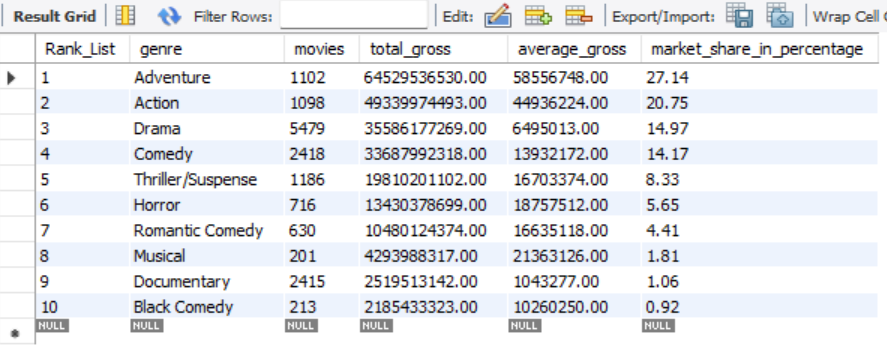
('8', 'Musical', '201', '4293988317', '21363126', '1.81'),

('9', 'Documentary', '2415', '2519513142', '1043277', '1.06'),

('10', 'Black Comedy', '213', '2185433323', '10260250', '0.92');

Select \* from Top\_Genre

**Output**



CREATE TABLE Top\_Grossing\_Rating (

Rank\_List INT PRIMARY KEY,

mpaa\_rating VARCHAR(50),

movies INT,

total\_gross DECIMAL(15, 2),

average\_gross DECIMAL(15, 2),

market\_share\_in\_percentage DECIMAL(5, 2)

);

INSERT INTO Top\_Grossing\_rating (RANK, MPAA RATINGS, MOVIES, TOTAL GROSS, AVERAGE GROSS, MARKET SHARE, ) VALUES

('1', 'PG-13', '3243', '113524789243', '35006102', '47.75%'),

('2', 'R', '5480', '63497164978', '11587074', '26.71%'),

('3', 'PG', '1535', '49124317794', '32002813', '20.66%'),

('4', 'G', '395', '9572240391', '24233520', '4.03%'),

('5', 'Not Rated', '5820', '1918358283', '329615', '0.81%'),

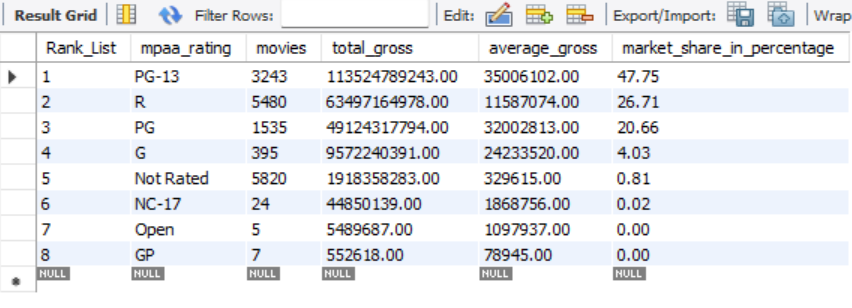
('6', 'NC-17', '24', '44850139', '1868756', '0.02%'),

('7', 'Open', '5', '5489687', '1097937', '0.00%'),

('8', 'GP', '7', '552618', '78945', '0.00%');

Select \* from Top\_Grossing\_Rating

**Output**



CREATE TABLE Top\_Grossing\_Source (

Rank\_List INT PRIMARY KEY,

source VARCHAR(100),

movies INT,

total\_gross int,

average\_gross int,

market\_share\_in\_percentage decimal(10,2));

INSERT INTO Top\_Grossing\_source (RANK, SOURCES, MOVIES, TOTAL GROSS, AVERAGE GROSS, MARKET SHARE) VALUES

('1', 'Original Screenplay', '7946', '106375196782', '13387264', '44.74%'),

('2', 'Based on Fiction Book/Short Story', '2150', '47005613207', '21863076', '19.77%'),

('3', 'Based on Comic/Graphic Novel', '249', '23369989130', '93855378', '9.83%'),

('4', 'Remake', '328', '12832659970', '39123963', '5.40%'),

('5', 'Based on Real Life Events', '3225', '11398356297', '3534374', '4.79%'),

('6', 'Based on TV', '231', '11305006312', '48939421', '4.75%'),

('7', 'Based on Factual Book/Article', '295', '7443681990', '25232820', '3.13%'),

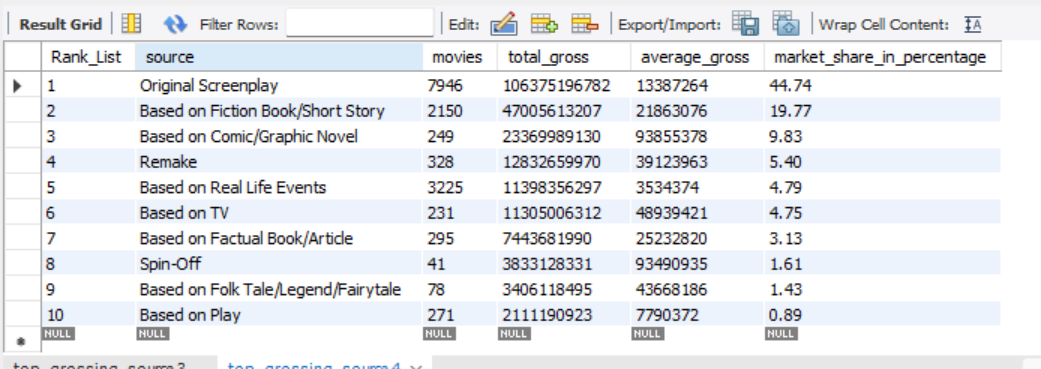
('8', 'Spin-Off', '41', '3833128331', '93490935', '1.61%'),

('9', 'Based on Folk Tale/Legend/Fairytale', '78', '3406118495', '43668186', '1.43%'),

('10', 'Based on Play', '271', '2111190923', '7790372', '0.89%');

Select \* from Top\_Grossing\_Source

**Output**



create TABLE Wide\_Release\_Count (

year INT,

warner\_bros INT,

walt\_disney INT,

20th\_century\_fox INT,

paramount\_pictures INT,

sony\_pictures INT,

universal INT,

total\_major6 INT,

total\_other\_studios INT,

FOREIGN KEY (year) REFERENCES Annual\_Ticket\_Sales(year)

);

INSERT INTO wide\_release\_count (YEAR, WARNER BROS, WALT DISNEY, 20TH CENTURY FOX, PARAMOUNT PICTURES, SONY PICTURES, UNIVERSAL, TOTAL MAJOR 6, TOTAL OTHER STUDIOS) VALUES

('2021', '17', '7', '0', '4', '16', '17', '61', '38'),

('2020', '5', '3', '1', '3', '9', '13', '34', '23'),

('2019', '18', '10', '11', '9', '18', '21', '87', '44'),

('2018', '19', '10', '11', '10', '16', '20', '86', '58'),

('2017', '18', '8', '14', '10', '16', '15', '81', '50'),

('2016', '17', '12', '16', '12', '16', '22', '95', '46'),

('2015', '22', '11', '18', '9', '13', '20', '93', '33'),

('2014', '17', '12', '17', '10', '17', '15', '88', '37'),

('2013', '17', '8', '15', '8', '14', '16', '78', '42'),

('2012', '16', '11', '15', '13', '18', '17', '90', '42'),

('2011', '20', '13', '15', '13', '21', '19', '101', '35'),

('2010', '20', '12', '18', '12', '17', '17', '96', '30'),

('2009', '25', '14', '20', '10', '21', '21', '111', '30'),

('2008', '19', '11', '22', '14', '19', '19', '104', '48'),

('2007', '30', '13', '17', '16', '22', '20', '118', '50'),

('2006', '26', '17', '25', '13', '26', '21', '128', '31'),

('2005', '20', '20', '19', '12', '19', '17', '107', '30'),

('2004', '27', '25', '18', '14', '15', '14', '113', '25'),

('2003', '28', '19', '13', '14', '19', '13', '106', '23'),

('2002', '32', '23', '15', '16', '20', '13', '119', '21'),

('2001', '30', '16', '16', '14', '17', '10', '103', '25'),

('2000', '29', '22', '13', '12', '15', '13', '104', '27'),

('1999', '27', '20', '15', '13', '22', '16', '113', '19'),

('1998', '27', '21', '11', '11', '20', '16', '106', '20'),

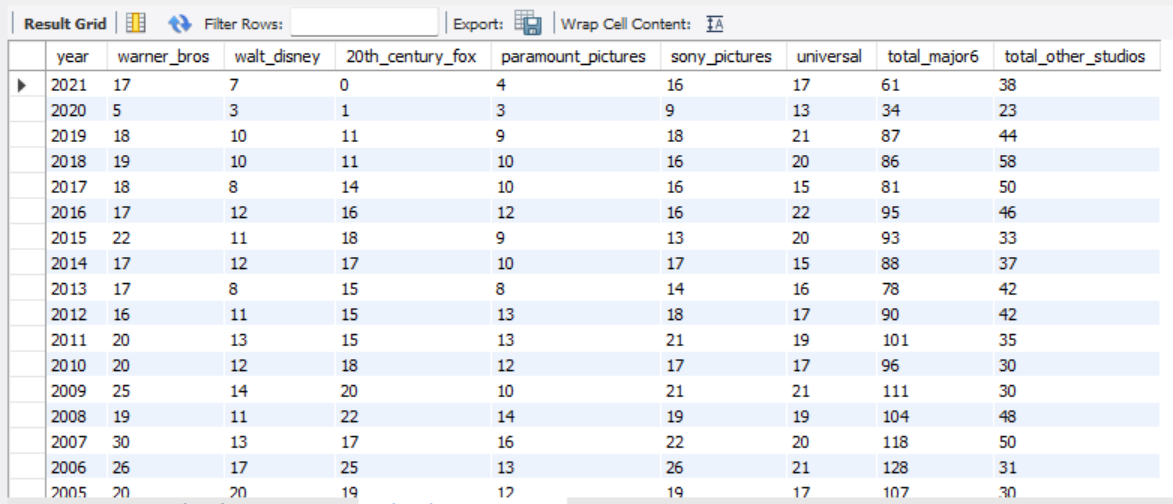
('1997', '31', '22', '12', '16', '22', '11', '114', '22'),

('1996', '31', '23', '13', '16', '24', '13', '120', '22'),

('1995', '27', '22', '11', '12', '20', '17', '109', '27');

Select \* from Wide\_Release\_Count

**Output**

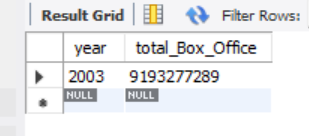


**Basic QUERIES**

**1. To Find Total box office earning for each year 2003**

**QUERY:** select year, total\_Box\_Office from annual\_ticket\_sales where year= "2003";

**OUTPUT**

****

**2. To list all movies where MPAA Rating IS PG-13**

**QUERY**:SELECT YEAR ,MOVIE

FROM highest\_grosser

WHERE MPAA\_RATING = "PG-13";

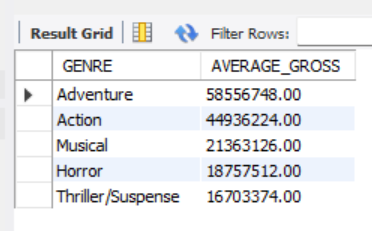
**OUTPUT**



**3. TO DISPLAY THE TOP 5 GENRE WITH THE HIGHST AVERAGE GROSSING**

**QUERY**: SELECT GENRE, AVERAGE\_GROSS FROM TOP\_GENRE ORDER BY AVERAGE\_GROSS desc LIMIT 5;

**OUTPUT**

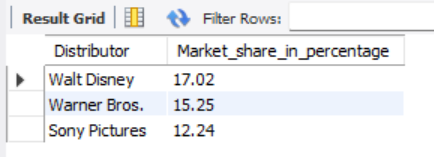
****

**4. TO Find top 3 distributors as per the market share**

**QUERY:** select Distributor, Market\_share\_in\_percentage from Top\_Distributor

order by Market\_share\_in\_percentage desc limit 3;

**OUTPUT**

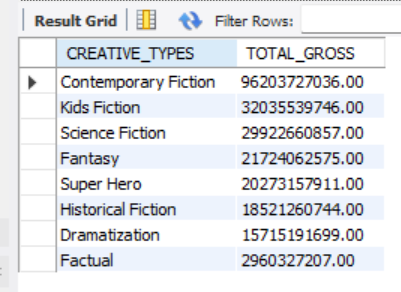


**5. TO FIND CREATIVE TYPES WITH TOTAL GROSS GREATER THAN 1 BILLION**

**QUERY**: SELECT CREATIVE\_TYPES, TOTAL\_GROSS FROM

popular\_creative\_types where total\_gross > 1000000000;

**OUTPUT**

****

**SUB-QUERIES**

**1.To find the movies with highest tickets sold in any year**

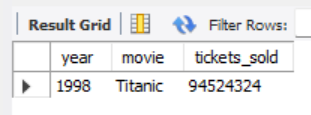
**QUERY**: select year, movie,tickets\_sold

from highest\_grosser

where tickets\_sold = ( select max(tickets\_sold) from

highest\_grosser);

**OUTPUT**



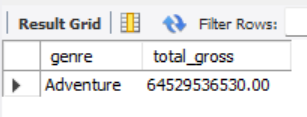
**2. To find genre with highest total gross**

**QUERY**: select genre, total\_gross

from Top\_Genre

where total\_gross = (select max(total\_gross) from top\_genre);

**Output**

****

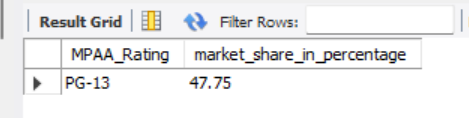
**3.** **To find the MPAA Rating with highest market share**

**QUERY**: select MPAA\_Rating,market\_share\_in\_percentage

from top\_grossing\_rating

where market\_share\_in\_percentage = (select max(market\_share\_in\_percentage) from top\_grossing\_rating);

**Output**

****

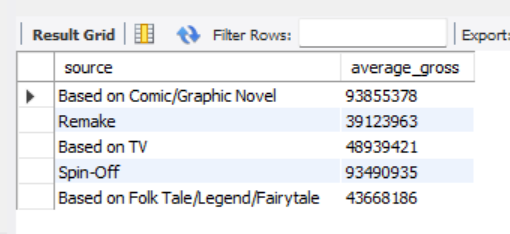
**4.To List all source with an average gross higher than the overall average gross of all source**

QUERY: select source, average\_gross

from top\_grossing\_source

where average\_gross > (select avg(average\_gross) from top\_grossing\_source);

**Output**

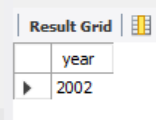


**5. to find the year where warner bros released highest number of movies**

**QUERY**: select year from wide\_release\_count

where warner\_bros = (select max(warner\_bros) from wide\_release\_count);

**Output**

****

**QUERIES WITH JOIN**

**1.** **to display each genres market share in percentage and corresponding MPAA Ratings**

**QUERY**: select tg.genre, tg.market\_share\_in\_percentage as genre\_market\_share\_in\_percentage, tgr.mpaa\_rating,

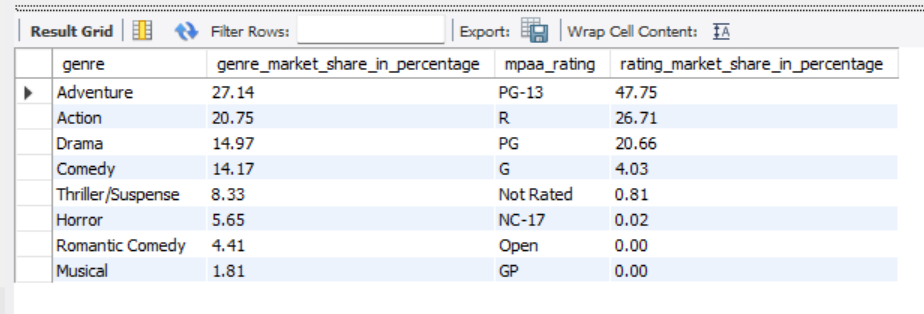
tgr.market\_share\_in\_percentage as rating\_market\_share\_in\_percentage

from top\_genre as tg

join top\_grossing\_rating as tgr on

tg.Rank\_List = tgr.Rank\_List;

**Output**

****

**2. top movie each year with its average ticket price**

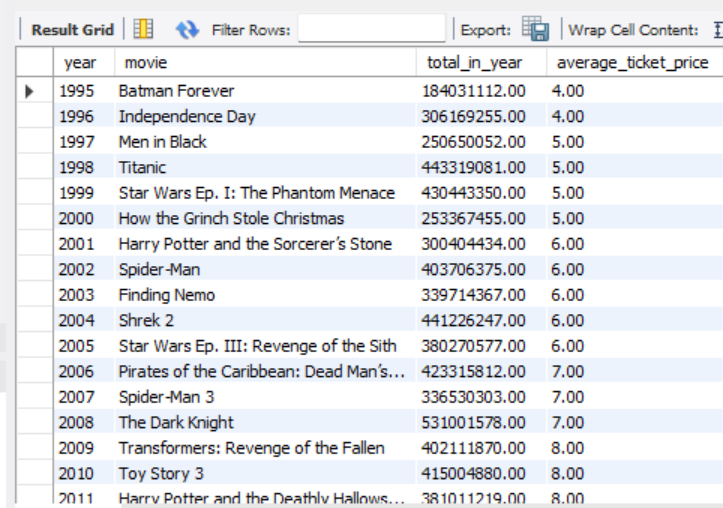
**QUERY**: select h.year, h.movie, h.total\_in\_year, a.average\_ticket\_price

from highest\_grosser as h

join annual\_tickets\_sales as a on

h.year = a.year;

**Output**

****

**3.to find total gross for each source with the top distributor**

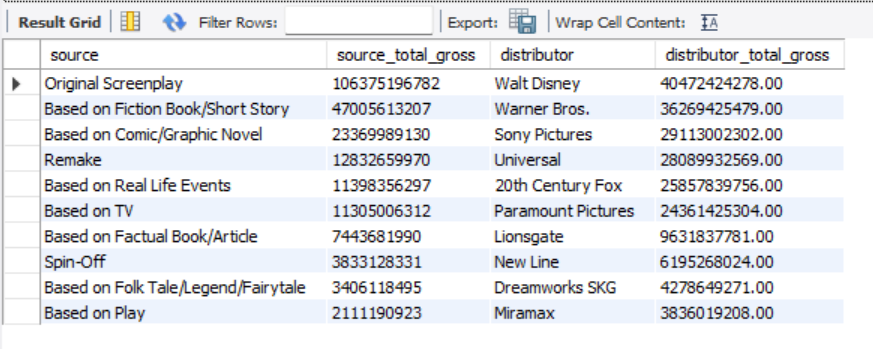
**QUERY**: select tgs.source, tgs.total\_gross as source\_total\_gross, td.distributor,td.total\_gross as distributor\_total\_gross

from top\_grossing\_source as tgs

join top\_distributor as td

on tgs.Rank\_List = td.Rank\_List;

**Output**

****

**4.** **to display genre and creative type by total gross**

**QUERY**: select tg.Rank\_List, tg.genre, tg.total\_gross as genres\_total\_gross,

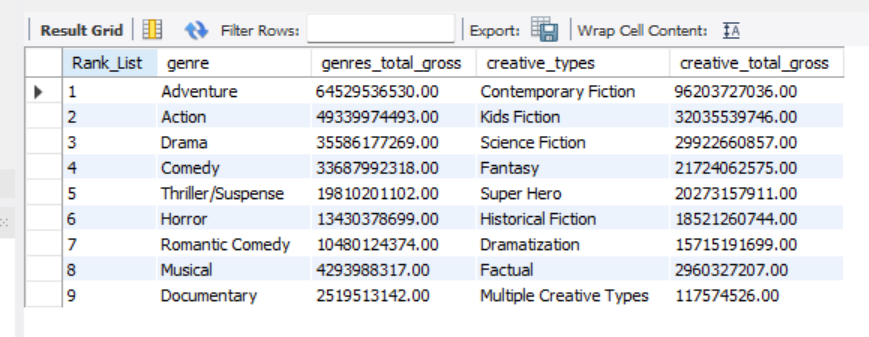
pct.creative\_types, pct.total\_gross as creative\_total\_gross

from top\_genre as tg

join popular\_creative\_types as pct

on tg.rank\_list = pct.rank\_list;

**Output**



**5. MPAA Rating and creative types by total gross**

**QUERY**: select tgr.rank\_list, tgr.mpaa\_rating, tgr.total\_gross as rating\_total\_gross ,

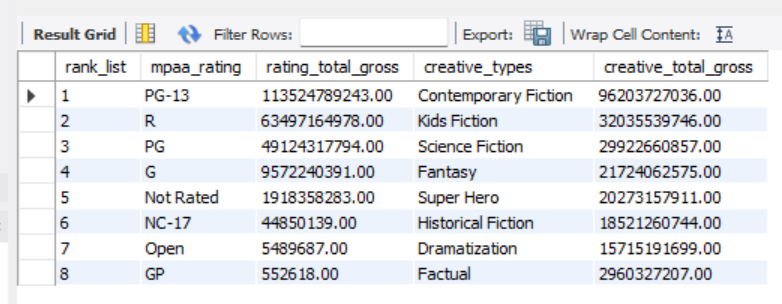
pct.creative\_types, pct.total\_gross as creative\_total\_gross

from top\_grossing\_rating as tgr

join popular\_creative\_types as pct

on tgr.Rank\_List = pct.Rank\_List;

**Output**

****

**Conclusion:**

The Hollywood Theatrical Market Synopsis (1995 to 2021) highlights the evolution of the film industry, showcasing steady box office growth driven by rising ticket prices, though ticket sales have not kept pace, indicating potential stagnation in audience turnout. Major studios like Disney, Warner Bros, and Universal dominate the market, leveraging wide release strategies and established franchises to secure significant revenue. Family-friendly, PG-13-rated films and genres like action, animation, and superhero movies consistently outperform others, reflecting changing audience preferences. The rise of adaptations and technological advancements, such as CGI, underscores Hollywood’s reliance on proven formulas to maximize profits. Inflation-adjusted figures reveal a nuanced perspective on growth, emphasizing the real impact of pricing and economic factors over time.