Assignment #2

Student ID:511210

7) In the movies dataset, for all NOT NULL records in the column of total revenue for year, find each distributor ID with its sum of total revenue. Sort the result in descending order of distributor ID.

```
SELECT DistributorID, sum(TotalForYear)
FROM movies
WHERE TotalForYear is not NULL
GROUP BY DistributorID
ORDER BY DistributorID DESC;
```

DistributorID	sum(TotalForYe...
20123	864017144
20122	370274604
20121	227898505
20119	1086080088
20118	1400597247
20117	1428018249
20116	3411466102
20115	2701620765
20114	3588529684
20113	3314146698
20112	6232964701
20111	17497656251

8) For each MPAA rating in movies table, list the average of total revenue and the average of top movies' tickets sold. Sort the result in increasing order of average ticket sold for top movies.

```
SELECT MPAARating, avg(TotalForYear), avg(TicketsSold)
FROM movies
GROUP BY MPAARating
ORDER BY avg(TicketsSold) ASC;
```

MPAARating	avg(TotalForYear)	avg(TicketsSol...
R	237093388.4286	36723334.4667
G	282521317.5000	41791177.3333
PG	338060730.7241	45645558.4828
PG-13	326097937.4268	47071661.1358

9) Looking at the MPAA ratings, find how many top movies are in each rating category. DO NOT include movies released in 1995 and 2004 at all in the analysis. Sort the results in descending order for the number of movies.

```
SELECT MPAARating, count(movie)
FROM movies
WHERE not (year = '1995' or year = '2004')
GROUP BY MPAARating
ORDER BY count(movie) DESC;
```

MPAARating	count(movi...
PG-13	82
PG	27
R	14
G	7

10) For all PG-13 top movies with genre ID 302, find how many of them have total revenue for year more than $500,000,000 and tickets sold more than 50,000,000.

```
SELECT count(movie)
FROM movies
WHERE MPAARating = 'PG-13' and GenreID = '302' and TotalForYear > 500000000 and TicketsSold > 50000000;
```

count(movi...
▶ 6

11)   *For all top movies released on Friday, list the average ticket sold and the average of total revenue for year. DO NOT include movies in 2002 and 2012.*

```
SELECT avg(TicketsSold), avg(TotalForYear)
FROM movies
WHERE (not (year = '2002' or year = '2012') )and DayOfWeek = 'Friday';
```

avg(TicketsSol...	avg(TotalForYear)
▶ 44604650.5287	320789218.7273

12) *For all distributors with the market share greater than 10%, list their name, founding date, the number of movies made, and the average total for the year for those movies. ONLY include the distributor with total gross no less than $28,000,000,000.*

```
SELECT Distributor, FoundDate, NumOfMovies, avg(TotalForYear)
FROM distributors inner join movies
USING (DistributorID)
WHERE DisMarketShare > 0.1 and TotalGross > 28000000000
GROUP BY DistributorID;
```

Distributor	FoundDate	NumOfMovies	avg(TotalForYear)
▶ Walt Disney	1923-10-16	595	364534505.2292
Warner Bros.	1923-04-04	835	283316577.3182
Sony Pictures	1987-12-21	764	301286063.4545
Universal	1912-04-30	547	299044140.3333

13) *How much time did you spend on this assignment?*

About 2 hours.