<https://www.youtube.com/watch?v=npgbI8KYvN8&list=PLeo1K3hjS3us_ELKYSj_Fth2tIEkdKXvV&index=3>

A screenshot of a computer

Description automatically generated

Total Revenue = “SUM(movies[budget])”

Initially, this gives wring answer, because, the million, billion column is given other column and the amount is mentioned in both usd and inr formats. So we cant directly add this thing.

1. Standardize them to common currency
2. Standardize them to common unit

Create new columns:

1. Budget in millions : IF([@unit] = “Billions”, [@budget]\*1000, [@budget]) >> If my unit is in billions multiply it by 1000, or else keep the budget as it is
2. Revenue in millions : =IF([@unit]="Billions",[@revenue]\*1000, [@revenue])
3. Budget million INR: =IF([@currency]="USD",[@[Budget\_mln]]\*80, [@[Budget\_mln]])
4. Revenue million INR: =IF([@currency]="USD",[@[Revenue\_mln]]\*80, [@[Revenue\_mln]])

Or

Instead of writing this,

Select “$” below “General”>>More formats>>English India

Total Number of movies: =COUNT(movie[movie\_id])

Average Bollywood movie revenue:

Total Bollywood movies: =COUNTIF(Movies[industry],”Bollywood”) (Lets say cell number is J11)

Total Bollywood movie revenue in INR: =SUMIF(movies[industry],”Bollywood”,Movies[revenue\_inr]) (Lets say cell number is J12)

Avg Bollywood Revenue: = J12/J11

Avg Bollywood Revenue %: Total Bollywood movie revenue in INR /Total Revenue INR >>click on % to show % symbol

**Business Metrics:**

A screenshot of a computer

Description automatically generated

Revenue: =SUMIF(Movies[studio], "Marvel Studios",Movies[Revenue USD])

Budget: =SUMIF(Movies[studio],"Marvel Studios", Movies[Budget USD])

Profit/Loss = Revenue-Budget

Profit/Loss % = Profit/Loss / Budget

Total Hollywood Revenue: =SUMIF(Movies[Industry],”Hollywood”,movies[Revenue USD])

Market Share = Revenue/Total Hollywood Revenue

**Basic statistics**

Mean, Median, Mode

Mean : =Average(Movies[Imdb\_rating])

Median: =MEDIAN(Movies[imdb\_rating])

Mode: =MODE(Movies[imdb\_rating])

A screenshot of a computer

Description automatically generated

Variance and Standard Deviation

A screenshot of a computer

Description automatically generated

Variance: How far each number is from every other number in a dataset

A screenshot of a video

Description automatically generated

In yuganda, the income of each person is like this

Almost near to each other

But in krance, we can see in the below table that the distance between each person is so large

Standard Deviation = SQRT(Variance)

Variance : = VAR.P(Movies[imdb\_rating]

StdDev: =STDEV.P(movies[imdb\_rating])

Correlation:

=CORREL(Table8[Price in Cr],Table8[Sqft])

A screenshot of a computer

Description automatically generated

MySQL

A screenshot of a computer

Description automatically generated

SELECT \* FROM moviesdb.movies;

Select title, industry from moviesdb.movies;

Use moviesdb;

Select title, industry from movies where industry="Bollywood";

Select count(\*) from movies where industry="Bollywood";

select distinct industry from movies;

Select \* from movies where title like "%thor%";

Select \* from movies where studio="";

A screenshot of a video chat

Description automatically generated

select \* from movies where imdb\_rating>=9;

select \* from movies where imdb\_rating>=6 and imdb\_rating<=8;

select \* from movies where imdb\_rating between 6 and 8;

select \* from movies where release\_year = 2022 or release\_year=2019 or release\_year = 2020;

select \* from movies where release\_year in (2022, 2020, 2019);

select \* from movies where imdb\_rating is NULL;

select \* from movies where imdb\_rating is NOT NULL;

select \* from movies where industry="bollywood" order by imdb\_rating;

select \* from movies where industry="bollywood" order by imdb\_rating DESC LIMIT 5;

select \* from movies where industry="bollywood" order by imdb\_rating DESC LIMIT 5 OFFSET 1;

**Summary Analytics:**

[**https://dev.mysql.com/doc/refman/8.4/en/built-in-function-reference.html**](https://dev.mysql.com/doc/refman/8.4/en/built-in-function-reference.html)

FROM🡪WHERE🡪GROUP BY🡪 HAVING🡪 ORDER BY

Select count(\*) from movies where industry="Bollywood";

Select max(imdb\_rating) from movies where industry="Bollywood";

Select min(imdb\_rating) from movies where industry="Bollywood";

Select round(avg(imdb\_rating),2) from movies where industry="Bollywood";

Select round(avg(imdb\_rating),2) as average, min(imdb\_rating) as min, max(imdb\_rating) as max from movies where industry="Bollywood";

select studio, count(\*) as cnt from movies group by studio order by cnt desc;

select industry, count(\*) as cnt, round(avg(imdb\_rating),2) as avg\_rating from movies group by industry order by avg\_rating desc;

select release\_year, count(\*) as movies\_count from movies where imdb\_rating>8 group by release\_year having movies\_count>2 order by movies\_count desc;

A screenshot of a computer

Description automatically generated

select curdate();

select year(curdate());

select \*, year(curdate())-birth\_year as age from actors;

select \*, (revenue-budget) as profit from financials;

select \*, if(currency='USD', revenue\*77, revenue) as revenue\_inr from financials;

select distinct unit from financials;

select \*,

case

when unit="billions" then revenue\*1000

when unit = "thousands" then revenue/1000

when unit ="millions" then revenue

end as revenue\_mln

from financials;

A screenshot of a computer

Description automatically generated

SQL JOINS

https://chatgpt.com/share/6716b168-e2ac-8002-a907-a97e0ddc226d