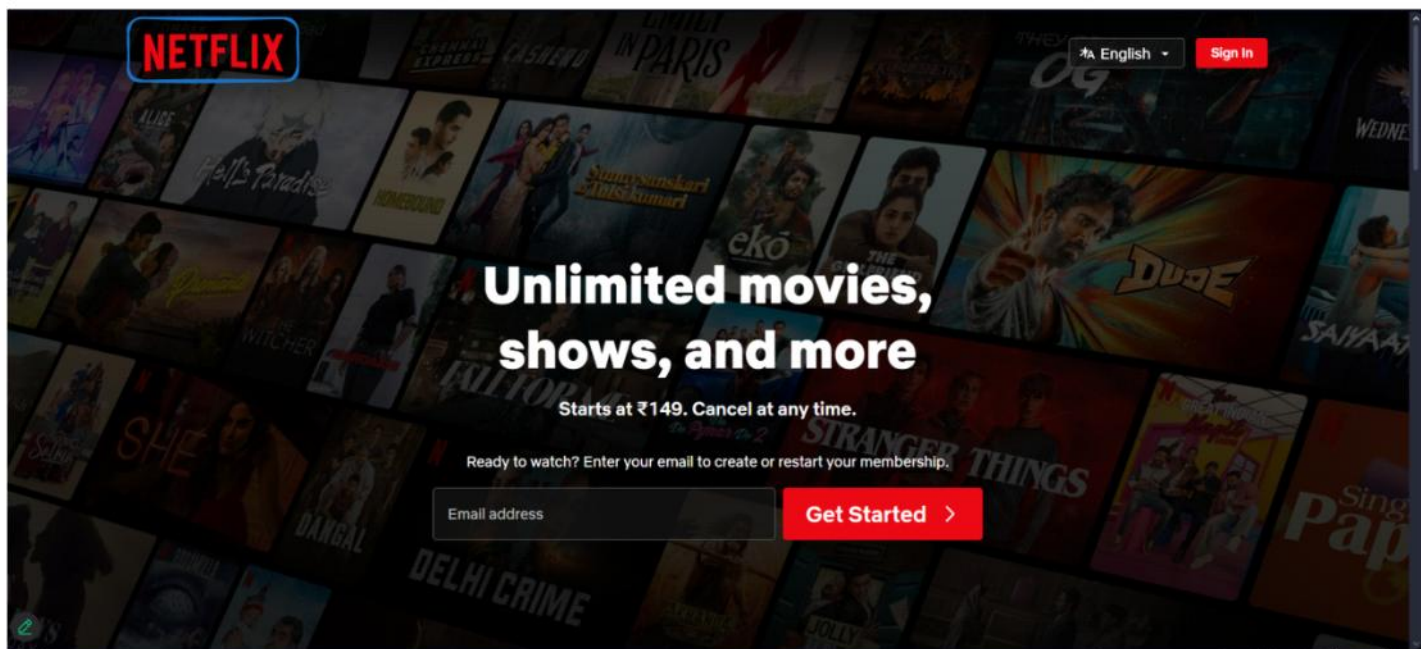


Visualising Netflix Dataset

- Which was the best movie and TV Show overall in the last 50 years?
- How many movies do we have in our dataset across last few years? Do we have more representation of movies from the last 20 years or is dataset free from any such skewness?
- On average, how was the IMDB score been trending over the last 50 years? Has it been deteriorating or improving?
- Have more people started voting for movies/ shows on over 50 years?
- On average, how has the runtime changed over the last 50 years?
- How does age certification of a movie affect its rating?
- How does age certification of TV Show affect its rating?

- What visual to choose to solve above questions.
- Find Insights from the Dashboard.
- Telling a story based insights backed by data.
- How to publish the dashboard.



Navigator

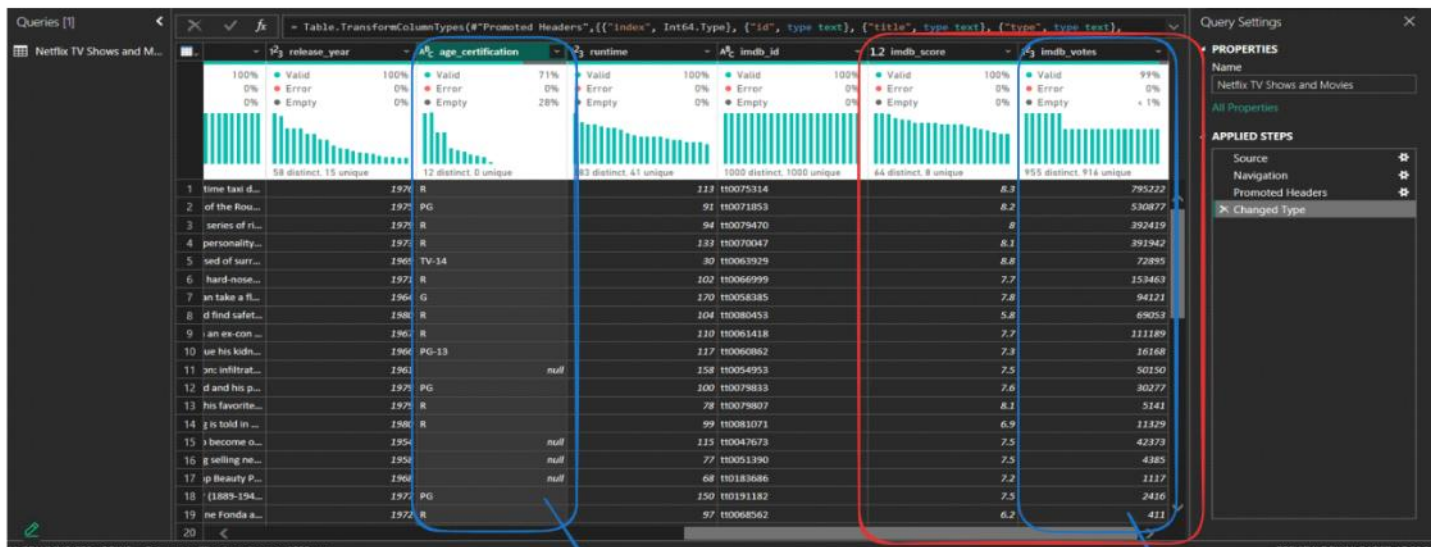
Display Options

- Netflix TV Shows & Movies.xlsx [1]
 - Netflix TV Shows and Movies

Netflix TV Shows and Movies

index	id	title	type	descr
0	tm84618	Taxi Driver	MOVIE	A
1	tm127384	Monty Python and the Holy Grail	MOVIE	Ki
2	tm70993	Life of Brian	MOVIE	Br
3	tm190788	The Exorcist	MOVIE	12
4	ts22164	Monty Python's Flying Circus	SHOW	A
5	tm14873	Dirty Harry	MOVIE	W
6	tm185072	My Fair Lady	MOVIE	A
7	tm98978	The Blue Lagoon	MOVIE	Tv
8	tm119281	Bonnie and Clyde	MOVIE	In
9	tm67378	The Professionals	MOVIE	Ar
10	tm44204	The Guns of Navarone	MOVIE	A
11	tm69778	Lupin the Third: The Castle of Cagliostro	MOVIE	Af
12	tm69997	Richard Pryor: Live in Concert	MOVIE	Ri
13	tm96649	The Long Riders	MOVIE	Th
14	tm16479	White Christmas	MOVIE	Tv
15	tm135083	Cairo Station	MOVIE	Qi
16	tm14491	The Queen	MOVIE	In
17	tm89386	Hitler: A Career	MOVIE	A
18	tm156453	FTA	MOVIE	A
19	tm27298	Saladin the Victorious	MOVIE	Sa
20	tm356209	Singapore	MOVIE	W
21	tm204541	Dark Waters	MOVIE	Ra
22	tm14350	AlexandriaV.C. - Why?	MOVIE	Se

Load Transform Data Cancel



- (Select All)
- (null)
- G
- NC-17
- PG
- PG-13
- R
- TV-14
- TV-G
- TV-MA
- TV-PG
- TV-Y
- TV-Y7

replace('null', 'NA')

High Outliers

Mean - 23407
Median - 2279

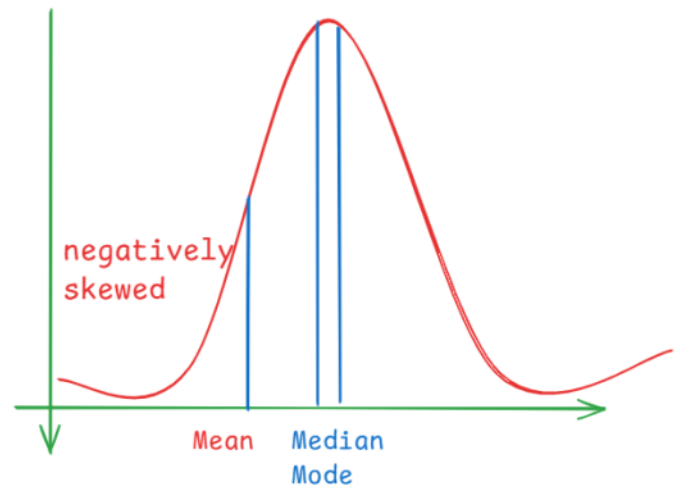
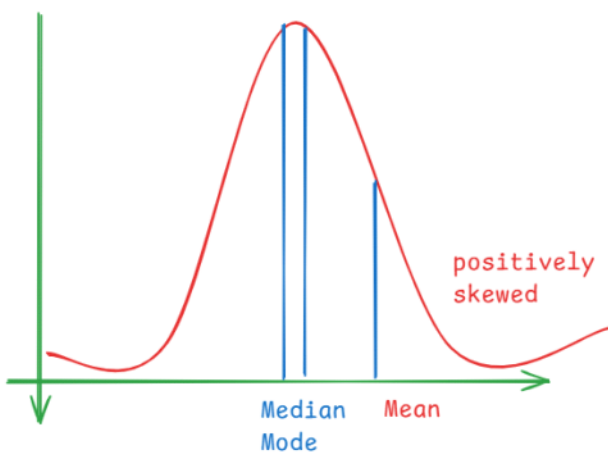
Impact analysis
<5% -> drop the rows.

Imputing > Mean or Median

10 thousand people voted for a movie with avg rating as 8.9
VS
1 million people voted for a movie with avg rating of 7.5

100x

	release_year	age_certification	runtime	imdb_id	imdb_score	imdb_votes
	94% Valid 0% Error 6% Empty	100% Valid 0% Error 0% Empty	56% Valid 0% Error 44% Empty	100% Valid 0% Error 0% Empty	100% Valid 0% Error 0% Empty	100% Valid 0% Error 100% Empty
	7 distinct, 3 unique	4 distinct, 1 unique	15 distinct, 14 unique	16 distinct, 16 unique	10 distinct, 4 unique	1 distinct, 0 unique
ie world of Irvi...	2014	PG-13		86 tt3766424	6.3	null
n the Ain el-He...	2013	null		12 tt3289618	7.8	null
edridden gran...	2017	null		10 tt7466040	7.5	null
ng it big. How...	2019	PG		93 tt10326928	5.3	null
he Israel Defe...	2019	R		103 tt10767720	5.6	null
osives expert ...	2020	R		80 tt10804786	4.3	null
null	2020	null		49 tt13079112	6.5	null
word respons...	2020	PG-13		8 tt13322484	6.9	null
ory, place and ...	2019	null		27 tt10545994	7.7	null
d School for th...	2021	PG-13		39 tt12771540	5.3	null
he US frame t...	2021	null		40 tt15339848	7.1	null
arried to Bena...	2021	null		22 tt14608922	7.7	null
i soldiers after...	2022	PG-13		13 tt17162524	6.5	null
itary short fea...	2021	PG-13		34 tt15470770	6.3	null
r correctly to h...	2022	PG-13		12 tt17321170	6.9	null
g Rabbids for a...	2021	null		68 tt15310736	7.8	null



Replace Values

Replace one value with another in the selected columns.

Value To Find

Replace With

Advanced options

OK Cancel

Age-Certification

Replace Values

Replace one value with another in the selected columns.

Value To Find

Replace With

OK Cancel

Imdb_votes

Home Transform **Add Column** View Tools Help

Column From Examples Custom Column Invoke Custom Function Index Column Duplicate Column Format Extract Merge Columns Statistics Standard Scientific Rounding Information Date Time Duration Text Analytics Vision Azure Machine Learning

Queries [1] Netflix TV Shows and M...

Custom Column

Add a column that is computed from the other columns.

New column name: Rating Votes

Custom column formula: $=[\text{imdb_score}] * [\text{imdb_votes}]$

Available columns: description, release_year, age_certification, runtime, imdb_id, **imdb_score**, **imdb_votes**

<< Insert

Learn about Power Query formulas

✓ No syntax errors have been detected.

OK Cancel

	1.2 imdb_score	1.2 imdb_votes
Valid	100%	100%
Error	0%	0%
Empty	0%	0%

	64 distinct, 8 unique	955 distinct, 916 unique
8.3	795222	
8.2	530877	
8	392419	
8.1	391942	
8.8	72895	
7.7	153463	
7.8	94121	
5.8	69053	
7.7	111189	
7.3	16168	
7.5	50150	
7.6	30277	
8.1	5141	
6.9	11329	

	12	13	14
d and his p...	1979	PG	100 tt0079833
his favorite...	1979	R	78 tt0079807
g is told in ...	1980	R	99 tt0081071

Load

Netflix TV Shows and Movies
Creating connection in model...

Cancel

Load

✓ Netflix TV Shows and Movies
5,283 rows loaded.

Cancel

Data

Search

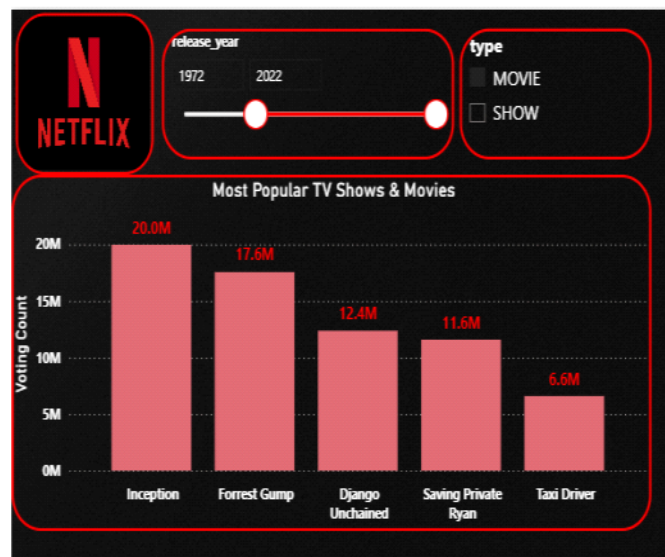
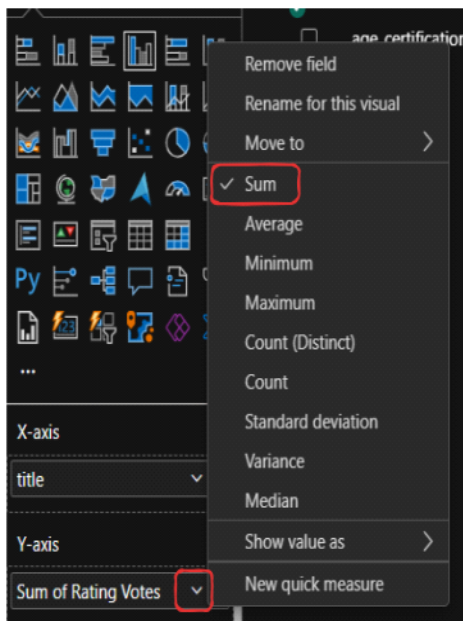
> Netflix TV Shows and Movies

Netflix TV Shows and Movies

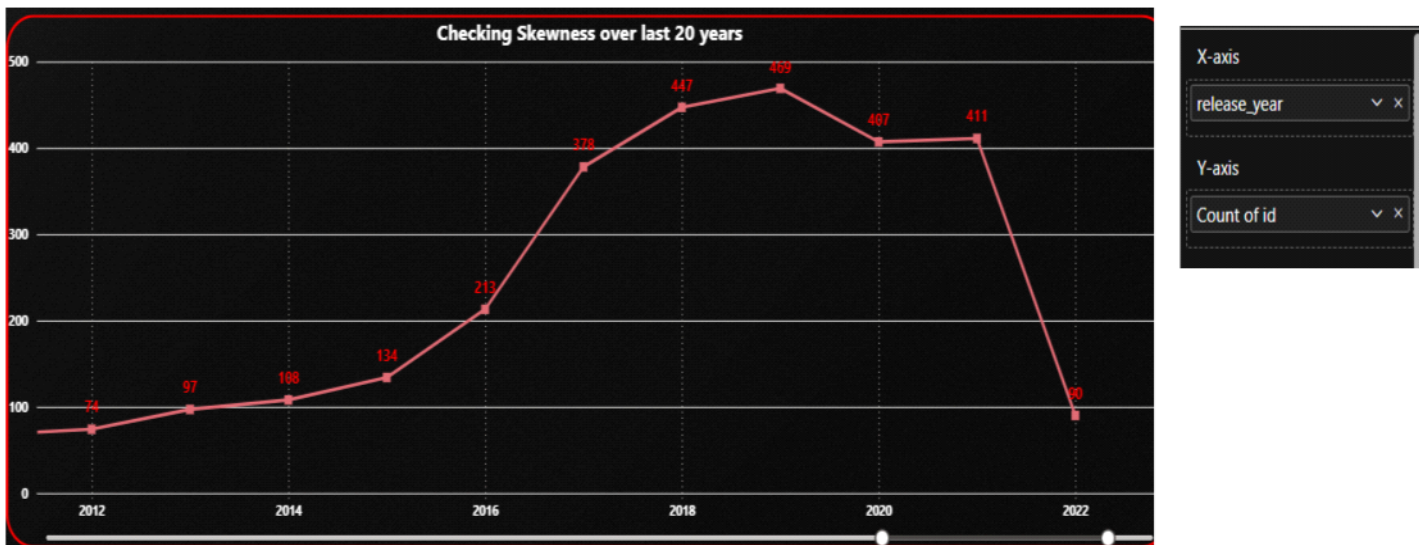
- ☐ age_certification
- ☐ description
- ☐ id
- ☐ imdb_id
- ☐ Σ imdb_score
- ☐ Σ imdb_votes
- ☐ Σ index
- ☐ Σ Rating Votes
- ☐ Σ release_year
- ☐ Σ runtime
- ☐ title
- ☐ type



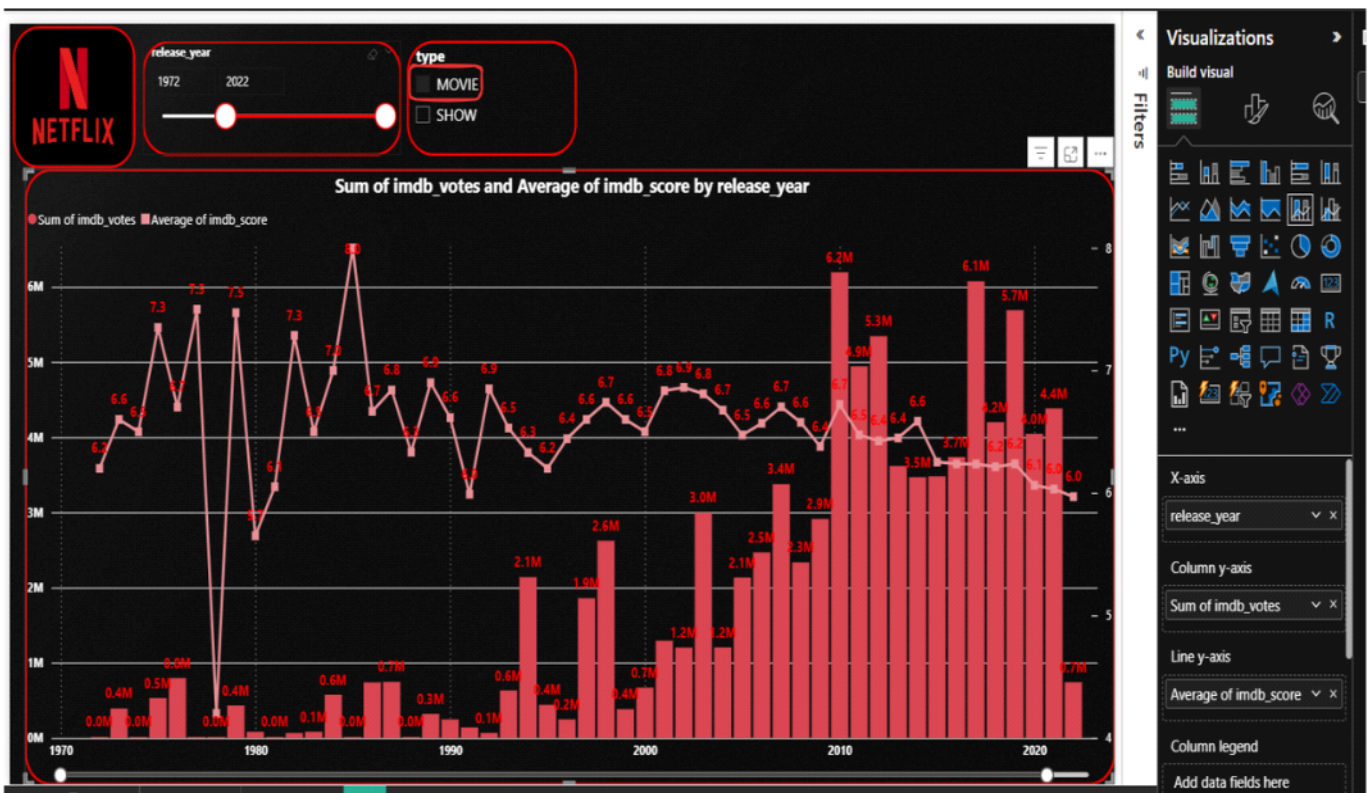
Which was the best movie and TV Show overall in the last 50 years?

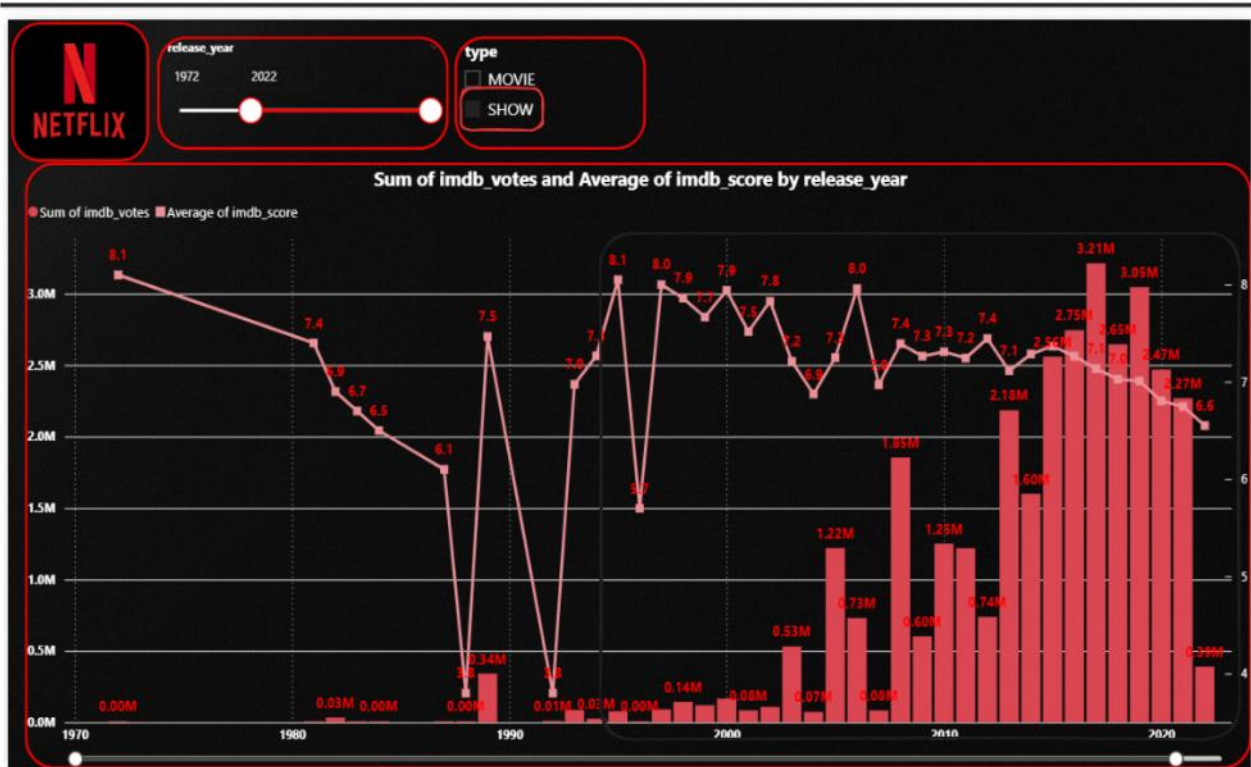


How many movies do we have in our dataset across last few years? Do we have more representation of movies from the last 20 years or is dataset free from any such skewness?



On average, how was the IMDB score been trending over the last 50 years? Has it been deteriorating or improving?





Have more people started voting for movies/shows on over 50 years?

