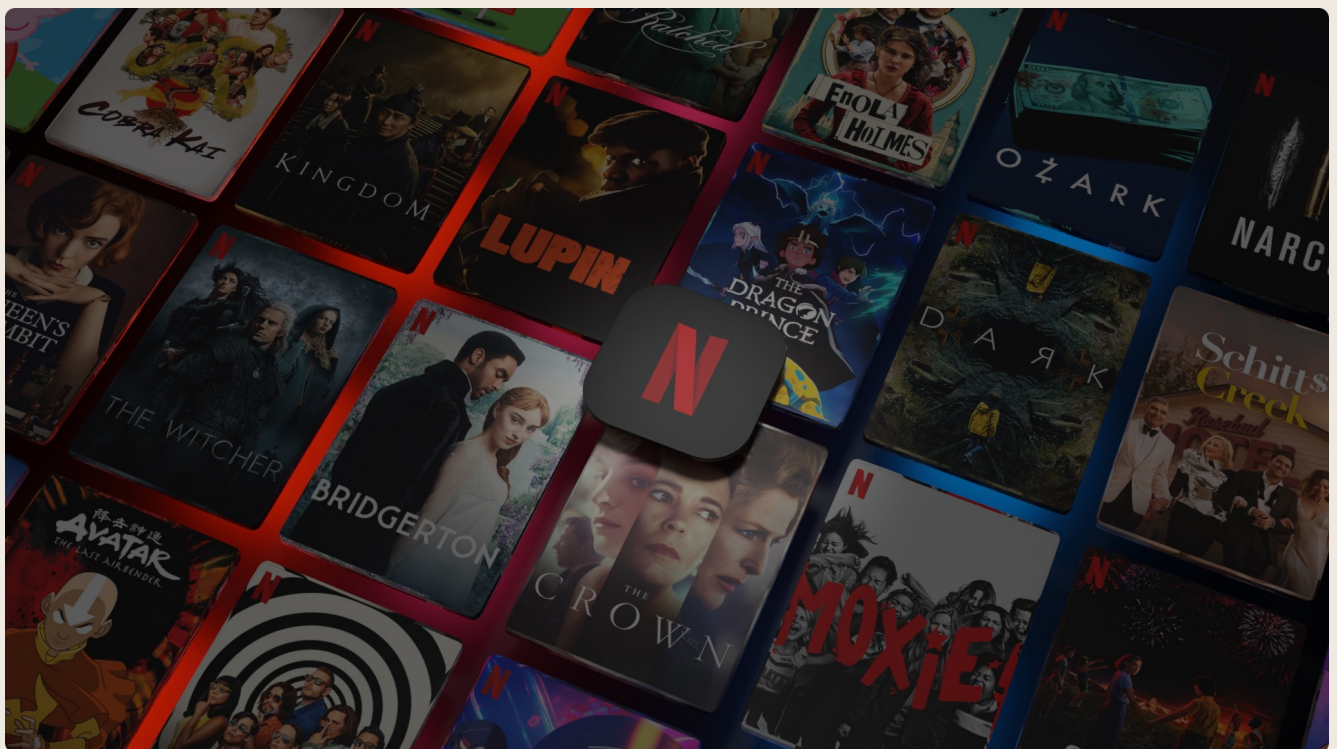

Netflix Films & Series Analysis

Data Analytics Case Study



May 2024

Table of contents

01

Introduction

Introduce background of this case study

02

Analysis

- (1) Background study
- (2) Understand the data
- (3) Analysis and visualisation

03

Conclusion



01

Introduction



Background

Disclaimer

The dataset was found in the [Kaggle](#) dataset library obtained on 2nd May 2024. In this case study, it is regarded that the latest dataset is sent from an internal data team, which means it is a reliable first-party source.

Business task

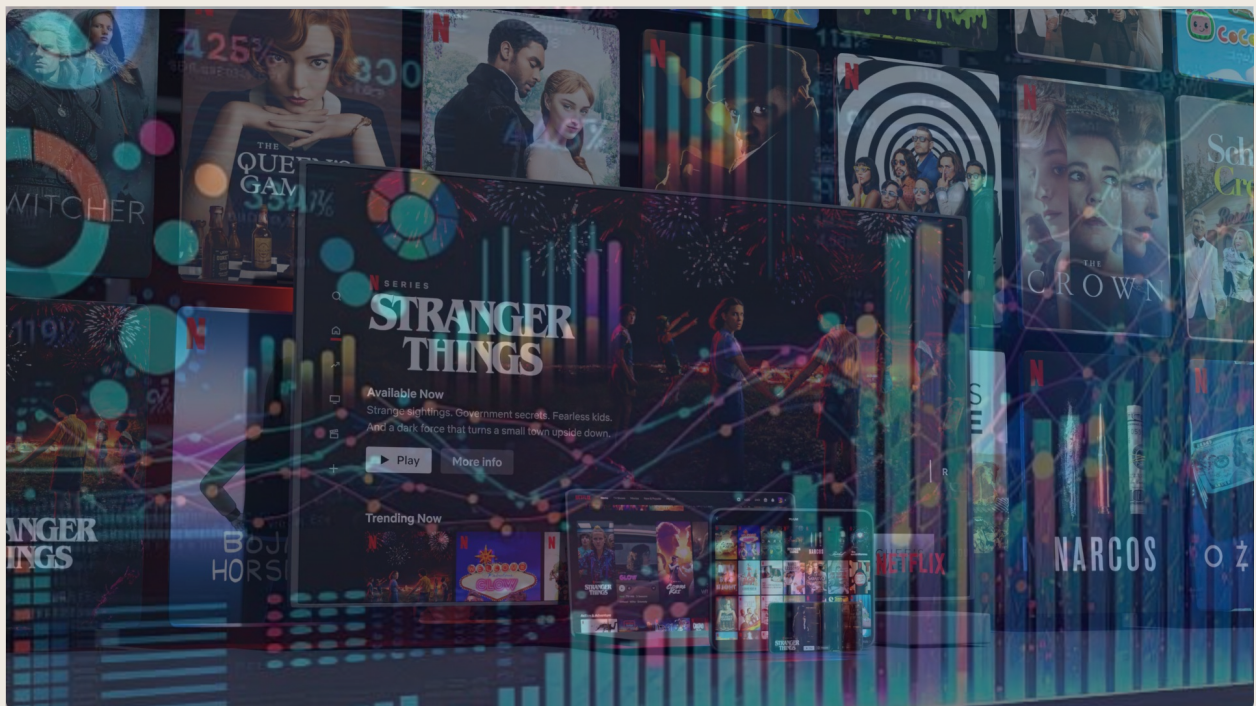
Due to the nature of the industry and the streaming service company's policy, it is not possible to gain detailed (recent) data. Therefore the disclaimer above was stated and the business task is also vaguely set accordingly.

Given mission:

Read the trends of the Netflix films and series and suggest future plans for the streaming platform.

02

Analysis



(1) Background study

FT Financial Times

Is Netflix limbering up for live sport?

Netflix's success with what it calls "sports adjacent programming" — docuseries such as Beckham, Break Point or Formula One: Drive to Survive — ..



Metro.co.uk

Netflix viewers threaten to cancel memberships after huge change

Disgruntled Netflix users are threatening to cancel the streaming service announced plans to scrap its basic



At Netflix, we want to entertain the world. Whatever your taste, and no matter where you live, we give you access to best-in-class TV series, documentaries, feature films and games. Our members control what they want to watch, when they want it, in one simple subscription. We're streaming in more than 30 languages and 190 countries, because great stories can come from anywhere and be loved everywhere. We are the world's biggest fans of entertainment, and we're always looking to help you find your next favorite story.

C21media

Netflix to spend more than 50% of content budget outside North America in 2024

Netflix to spend more than 50% of content budget outside North America in 2024 ...
More than half of Netflix's spending on original and acquired...



Company information: In recent years, the company has engaged in the game business and just started sports (WWE) service in the upcoming year in the selected countries, while a new subscription 'household' scheme was introduced by expecting an increase in the profits. In this context, the company seeks a precise plan for effective management in all different sectors.

Company values: Stories - great ones. The company emphasises the excellence in their products, so-called 'best-in-class'. Therefore their self-introduction of the line "We're streaming in **more than 30 languages and 190 countries**, because great stories can come from anywhere and be loved everywhere." means that their priority is to have worldwide beloved content.

Key stakeholders: Netflix themselves. The company wants to read the audience's reactions to its content so that it can plan for future productions and budgets.

(2) Understand the data

- Data verification: As mentioned in the disclaimer, this CSV file is regarded as reliable and credible.
- Columns are title, type(move/show), release year, age certification, runtime, genres, production countries, seasons, IMDb ID, IMDb score and IMDb votes. The CSV file has 5124 rows.
- 'imdb_votes' indicates the number of audiences who engaged in rating.
- 'Null' values in the 'imdb_scores' and 'imdb_votes' columns were removed as it is important to grasp the audiences' reviews. It was processed in Excel.
- Rows with 'IMDb_score' under 30 were removed as they did not have enough opinions.

title	series_type	release_year	age_certification	runtime	genres	production_country	imdb_id	imdb_score	imdb_votes
Khawatr	SHOW	2005	TV-14	20	reality		tt3062514	9.6	3046
Breaking Bad	SHOW	2008	TV-MA	48	drama, thriller, crime	US	tt0903747	9.5	1727694
Avatar: The Last Airbender	SHOW	2005	TV-Y7	24	scifi, family, fantasy, animation, action	US	tt0417299	9.3	297336
Kota Factory	SHOW	2019	TV-MA	42	drama, comedy	IN	tt9432978	9.3	66985
Our Planet	SHOW	2019	TV-G	50	documentation	GB	tt9253866	9.3	41386
Reply 1988	SHOW	2015	TV-14	95	comedy, drama, family, romance	KR	tt5182866	9.2	6557
My Mister	SHOW	2018	TV-14	79	drama, family	KR	tt7923710	9.2	5481
Who Rules The Night	SHOW	2022	TV-PG	44	drama, fantasy, romance	CN	tt14950036	9.2	81
Arcane	SHOW	2021	TV-14	41	action, scifi, drama, fantasy, animation	US	tt11126994	9.1	175412
The Last Dance	SHOW	2020	TV-MA	50	documentation, history, sport	US	tt8420184	9.1	108321
Attack on Titan	SHOW	2013	TV-MA	24	scifi, action, drama, fantasy, horror	JP	tt2560140	9	325381
DEATH NOTE	SHOW	2006	TV-14	24	scifi, animation, crime, drama, fantasy	JP	tt0877057	9	302147
Hunter x Hunter	SHOW	2011	TV-14	23	drama, action, comedy, crime, animation	JP	tt2098220	9	87857
David Attenborough	MOVIE	2020	PG	83	documentation	GB	tt11989890	9	31180
C/o Kancharla	MOVIE	2018	PG	152	drama	IN	tt7391996	9	6562
Leah Remini: Show Me the Money	SHOW	2016	TV-14	46	documentation, history	US	tt6244192	9	5761
Okupas	SHOW	2000	TV-MA	40	crime, drama	AR	tt0289649	9	2326
No Longer a Woman	MOVIE	1979		235	comedy, drama	EG	tt8312792	9	943
Raja, Rasoi & Family	SHOW	2014		29	history, documentation	IN	tt6953924	9	327
Seinfeld	SHOW	1989	TV-PG	24	comedy	US	tt0098904	8.9	302700
When They See Us	SHOW	2019	TV-MA	74	drama, history, crime	US	tt7137906	8.9	114127
Cowboy Bebop	SHOW	1998	TV-14	25	western, action, scifi, thriller, animation	JP	tt0213338	8.9	112887
Heartstopper	SHOW	2022	TV-14	28	drama, romance	GB	tt10638036	8.9	28978
Still Game	SHOW	2002	TV-14	29	comedy, european	GB	tt0281491	8.9	7539
Inception	MOVIE	2010	PG-13	148	scifi, music, thriller, action	GB, US	tt1375666	8.8	2268288
Forrest Gump	MOVIE	1994	PG-13	142	drama, romance, comedy	US	tt0109830	8.8	1994599
Black Mirror	SHOW	2011	TV-MA	59	scifi, thriller, drama, european	GB	tt2085059	8.8	515577
Peaky Blinders	SHOW	2013	TV-MA	58	drama, crime, european	GB	tt2442560	8.8	485506
Better Call Saul	SHOW	2015	TV-MA	49	comedy, drama, crime	US	tt3032476	8.8	404920
Narcos	SHOW	2015	TV-MA	52	drama, thriller, crime	US	tt2707408	8.8	404486
BoJack Horseman	SHOW	2014	TV-MA	26	drama, comedy, animation	US	tt3398228	8.8	143584
One Piece	SHOW	1999	TV-14	23	action, comedy, animation, drama	JP	tt0388629	8.8	112586
Monty Python and the Holy Grail	SHOW	1969	TV-14	30	comedy, european	GB	tt0063929	8.8	72895

(3) Analysis and visualisation

1. Top 100

Top 100 series by IMDb score

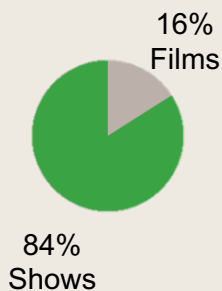
Rank	Title	Series type	Release year	IMDb score
1	Khawatr	SHOW	2005	9.6
2	Breaking Bad	SHOW	2008	9.5
3	Avatar: The Last Airbender	SHOW	2005	9.3
4	Kota Factory	SHOW	2019	9.3
5	Our Planet	SHOW	2019	9.3
6	Reply 1988	SHOW	2015	9.2
7	My Mister	SHOW	2018	9.2
8	Who Rules The World	SHOW	2022	9.2
9	Arcane	SHOW	2021	9.1
10	The Last Dance	SHOW	2020	9.1
11	Attack on Titan	SHOW	2013	9
12	DEATH NOTE	SHOW	2006	9
13	Hunter x Hunter	SHOW	2011	9
14	David Attenborough: A Life on Our Planet	MOVIE	2020	9
15	C/o Kancharapalem	MOVIE	2018	9
16	Leah Remini: Scientology and the Aftermath	SHOW	2016	9
17	Okupas	SHOW	2000	9
18	No Longer Kids	MOVIE	1979	9
19	Raja, Rasoi Aur Anya Kahaniyaan	SHOW	2014	9
20	Seinfeld	SHOW	1989	8.9
21	When They See Us	SHOW	2019	8.9
22	Cowboy Bebop	SHOW	1998	8.9
23	Heartstopper	SHOW	2022	8.9
24	Still Game	SHOW	2002	8.9
25	Inception	MOVIE	2010	8.8
26	Forrest Gump	MOVIE	1994	8.8
27	Black Mirror	SHOW	2011	8.8
28	Peaky Blinders	SHOW	2013	8.8
29	Better Call Saul	SHOW	2015	8.8
30	Narcos	SHOW	2015	8.8
31	BoJack Horseman	SHOW	2014	8.8
32	One Piece	SHOW	1999	8.8
33	Monty Python's Flying Circus	SHOW	1969	8.8
34	Chappelle's Show	SHOW	2003	8.8
35	The Untamed	SHOW	2019	8.8
36	Hospital Playlist	SHOW	2020	8.8
37	Navillera	SHOW	2021	8.8
38	Stories by Rabindranath Tagore	SHOW	2015	8.8
39	Sky Tour: The Movie	MOVIE	2020	8.8
40	StoryBots Super Songs	SHOW	2016	8.8
41	Stranger Things	SHOW	2016	8.7
42	House of Cards	SHOW	2013	8.7
43	Dark	SHOW	2017	8.7
44	Arrested Development	SHOW	2003	8.7
45	Downton Abbey	SHOW	2010	8.7
46	The Crown	SHOW	2016	8.7
47	One-Punch Man	SHOW	2015	8.7
48	Demon Slayer: Kimetsu no Yaiba	SHOW	2019	8.7
49	Friday Night Lights	SHOW	2006	8.7
50	Code Geass: Lelouch of the Rebellion	SHOW	2006	8.7
51	Anne with an E	SHOW	2017	8.7
52	Bo Burnham: Inside	MOVIE	2021	8.7
53	Anbe Sivam	MOVIE	2003	8.7
54	Greatest Events of WWII in Colour	SHOW	2019	8.7
55	Tear Along the Dotted Line	SHOW	2021	8.7
56	Mr. Sunshine	SHOW	2018	8.7
57	Twenty Five Twenty One	SHOW	2022	8.7
58	Dave Chappelle	SHOW	2017	8.7
59	Lenox Hill	SHOW	2020	8.7

Top 100 Movie by IMDb score

Rank	Title	Series type	Release year	IMDb score
1	David Attenborough: A Life on Our Planet	MOVIE	2020	9
2	C/o Kancharapalem	MOVIE	2018	9
3	No Longer Kids	MOVIE	1979	9
4	Inception	MOVIE	2010	8.8
5	Forrest Gump	MOVIE	1994	8.8
6	Sky Tour: The Movie	MOVIE	2020	8.8
7	Bo Burnham: Inside	MOVIE	2021	8.7
8	Anbe Sivam	MOVIE	2003	8.7
9	Rubaru Roshni	MOVIE	2019	8.7
10	A Lion in the House	MOVIE	2006	8.7
11	Bye Bye London	MOVIE	1982	8.7
12	Saving Private Ryan	MOVIE	1998	8.6
13	Merku Thodarchi Malai	MOVIE	2018	8.6
14	Best Wishes, Warmest Regards: A Schitt's Creek Farewell	MOVIE	2020	8.6
15	In Our Mothers' Gardens	MOVIE	2021	8.6
16	A Second Chance	MOVIE	2015	8.6
17	Bill Hicks: Revelations	MOVIE	1993	8.5
18	Springsteen On Broadway	MOVIE	2018	8.5
19	K. D.	MOVIE	2019	8.5
20	Emicida: AmarElo - It's All for Yesterday	MOVIE	2020	8.5
21	Ani... Dr. Kashinath Ghanekar	MOVIE	2018	8.5
22	Django Unchained	MOVIE	2012	8.4
23	3 Idiots	MOVIE	2009	8.4
24	Dangal	MOVIE	2016	8.4
25	Dave Chappelle: Sticks & Stones	MOVIE	2019	8.4

Top 100 Show by IMDb score

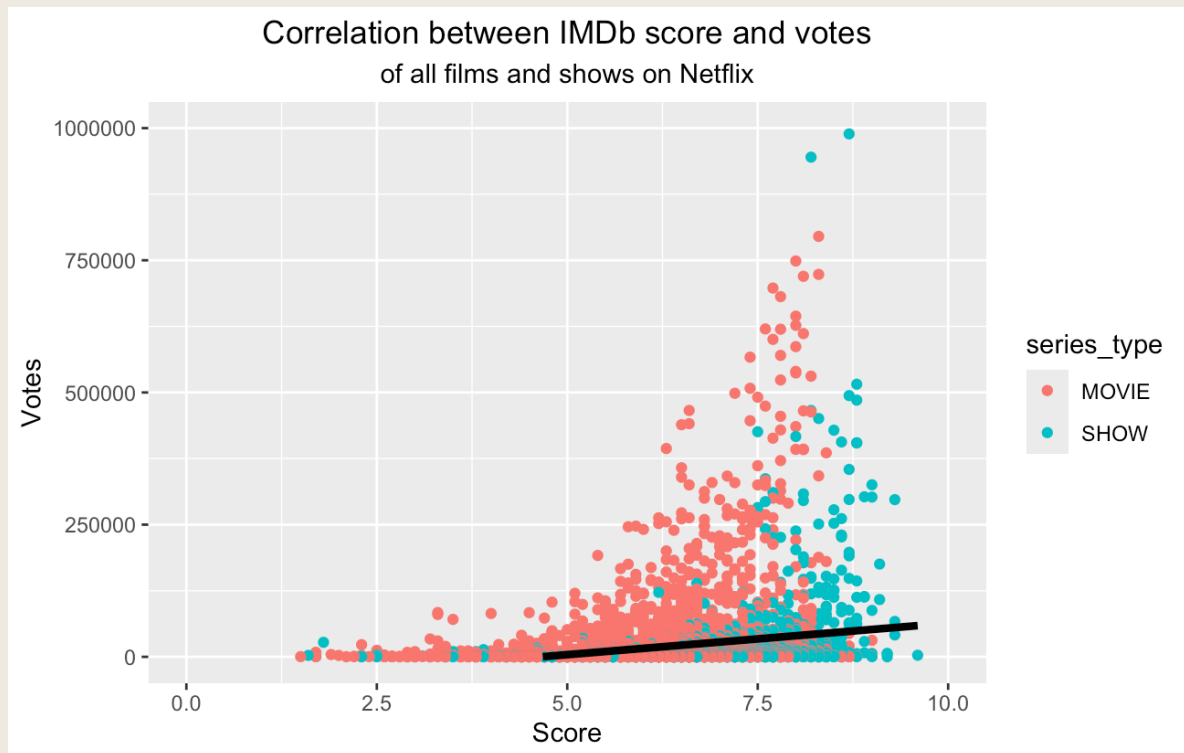
Rank	Title	Series type	Release year	IMDb score
1	Khawatr	SHOW	2005	9.6
2	Breaking Bad	SHOW	2008	9.5
3	Avatar: The Last Airbender	SHOW	2005	9.3
4	Kota Factory	SHOW	2019	9.3
5	Our Planet	SHOW	2019	9.3
6	Reply 1988	SHOW	2015	9.2
7	My Mister	SHOW	2018	9.2
8	Who Rules The World	SHOW	2022	9.2
9	Arcane	SHOW	2021	9.1
10	The Last Dance	SHOW	2020	9.1
11	Attack on Titan	SHOW	2013	9
12	DEATH NOTE	SHOW	2006	9
13	Hunter x Hunter	SHOW	2011	9
14	Leah Remini: Scientology and the Aftermath	SHOW	2016	9
15	Okupas	SHOW	2000	9
16	Raja, Rasoi Aur Anya Kahaniyaan	SHOW	2014	9
17	Seinfeld	SHOW	1989	8.9
18	When They See Us	SHOW	2019	8.9
19	Cowboy Bebop	SHOW	1998	8.9
20	Heartstopper	SHOW	2022	8.9



More than 5000 series were lined up solely according to the IMDb score. Given the range of the score from 1 to 10, the highest score was 9.6, a show called *Khawatr*. In terms of the number, shows are overwhelmingly named in the list, with 84%, while there are only 16 films were mentioned, which demonstrates that shows tend to get higher scores than films.

(3) Analysis and visualisation

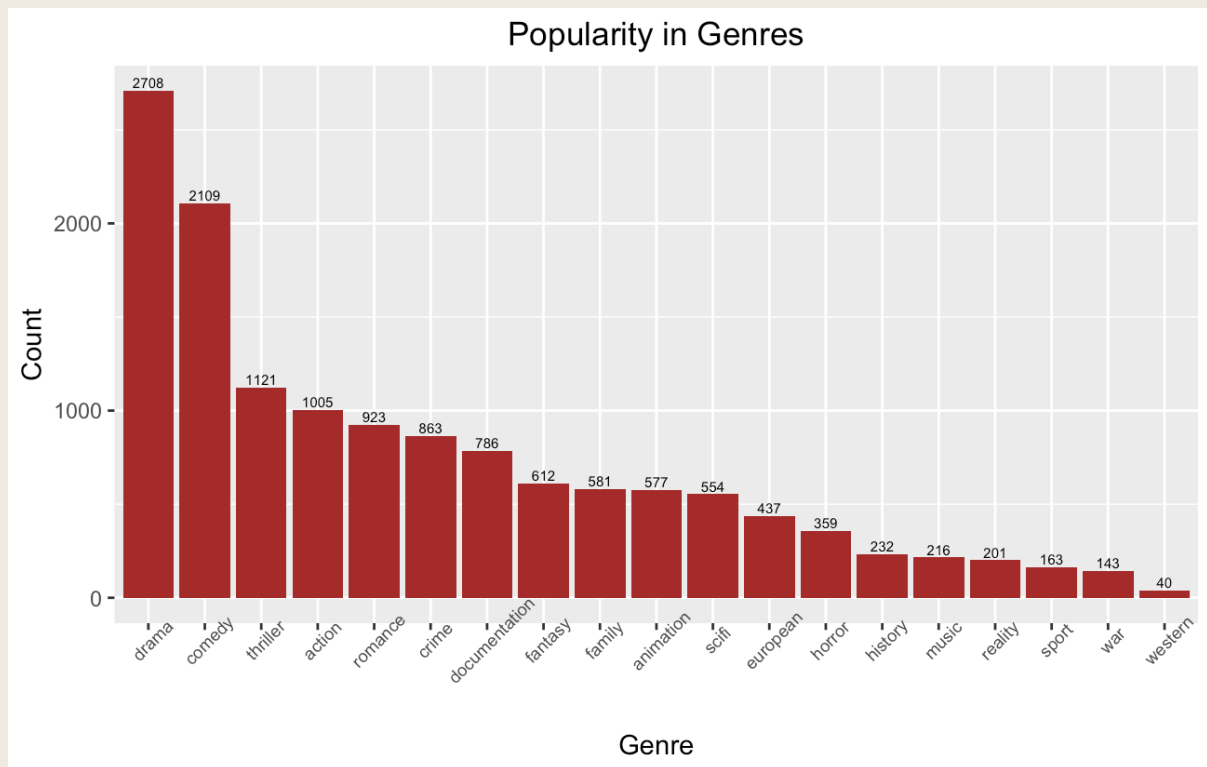
2. Correlation of score and votes in film and show



R studio's scatter plot was used to display the correlation between score and votes. Most series have under 250K votes, between 5 and 10 rather than 0 to 5. The regression line draws an upward curve, indicating a positive relationship between the votes and scores, i.e. higher scores are typically linked to a larger number of votes. Regarding shows, it might be derived from repeated voting due to serial episodes, which could lead to royalty and higher engagement in leaving reviews. On the other hand, series with scores below 5 tend to get fewer votes, probably because of their poor quality. A few outliers in the higher scores and votes are likely very popular and critically acclaimed series.

(3) Analysis and visualisation

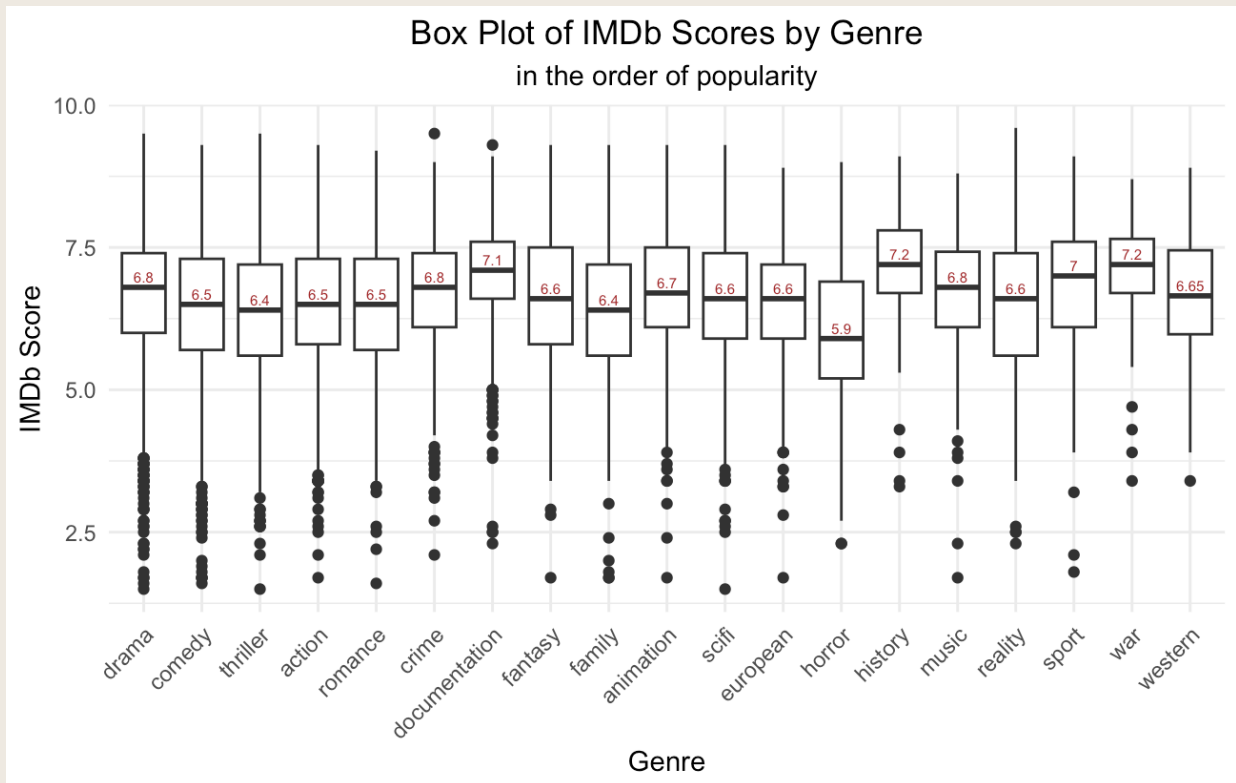
3. Correlation of genres and score



Usually, a series has more than two genres therefore in this topic, the frequency of specific genres was measured. It can be seen as two genres are the most popular, with their appearance more than 2000. As this count allowed overlap, a clear number for its popularity is not easy to determine. Instead, with its median number of 565.5, it could be read that drama, comedy, thriller, action, romance, crime, documentation, fantasy and animation are the popular group, while sci-fi, European, horror, history, music, reality, sport, war and western are less beloved.

(3) Analysis and visualisation

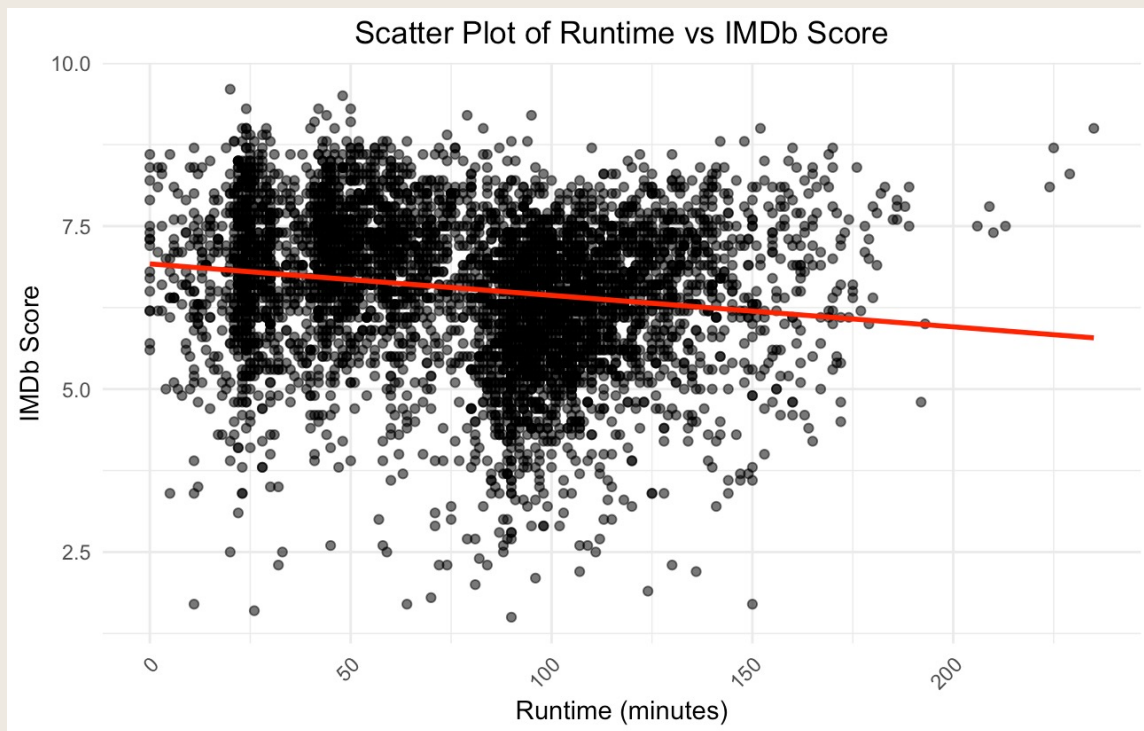
3. Correlation of genres and score (cont'd)



A box plot was used to understand each genre's score in depth. Outliers are found in all genres, especially popular genres such as drama, comedy and action have significant amounts, while sport, war and western present consistent ratings with fewer outliers, which might be because of lower popularity than other genres. The numbers in the boxes indicate their median, rounded to two decimal places, mainly between 6 and 7. Two unpopular genres, history and war, have the highest figure at 7.2, meaning those genres are generally well-received regardless of their popularity. On the other hand, horror has the lowest number, under 6, indicating lower ratings overall.

(3) Analysis and visualisation

4. Correlation of runtime and score



According to the density of the dots, it can be seen that the series having a runtime of about 100 minutes tends to be reviewed on IMDB. However, the regression line clearly indicates that the longer runtime leads to fewer higher scores.

(3) Analysis and visualisation

5. Correlation of production countries and score

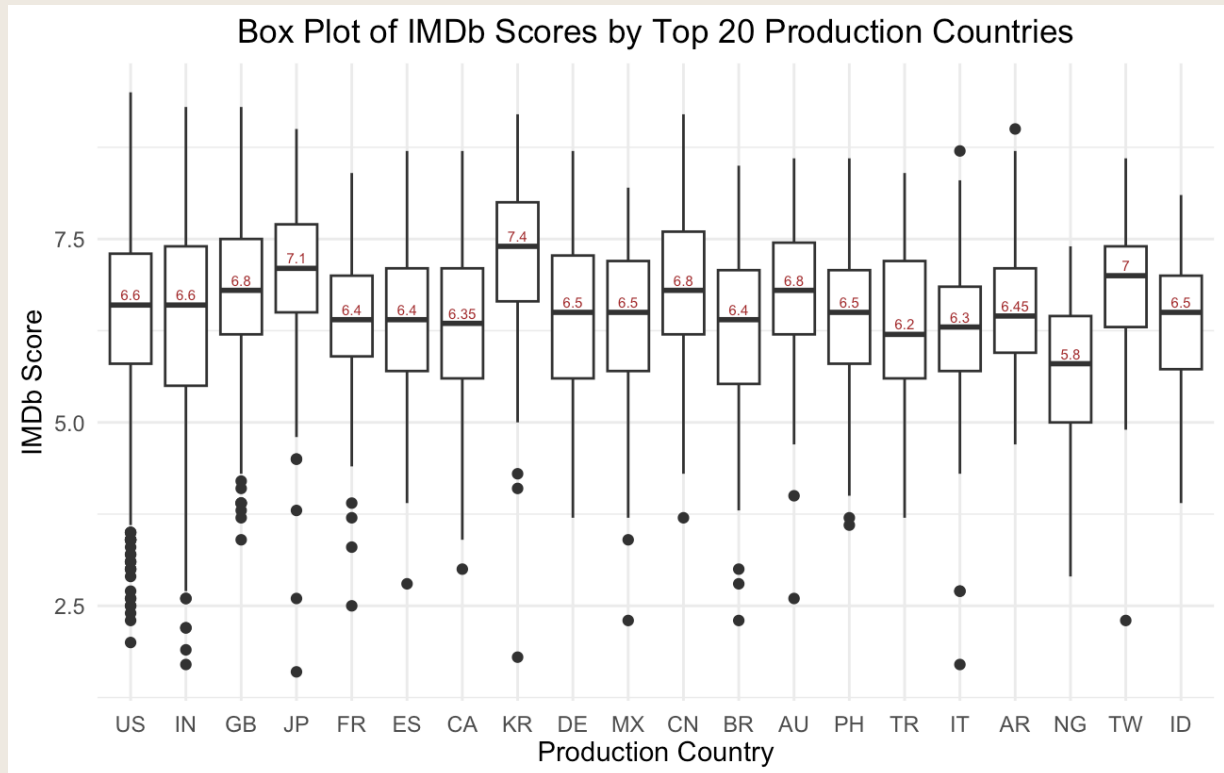
No.	Country	Frequency	Rank
1	US	2117	1
2	IN	582	2
3	GB	372	3
4	JP	254	4
5	FR	231	5
6	ES	198	6
7	CA	188	7
8	KR	183	8
9	DE	130	9
10	MX	112	10
11	CN	93	11
12	BR	86	12
13	AU	79	13
14	PH	78	14
15	TR	78	14
16	IT	75	16
17	AR	64	17
18	NG	55	18
19	TW	53	19
20	ID	52	20

76	SU	2	65
77	SY	2	65
78	VN	2	65
79	CU	1	79
80	BS	1	79
81	MW	1	79
82	NP	1	79
83	GE	1	79
84	KG	1	79
85	HR	1	79
86	VE	1	79
87	DZ	1	79
88	BF	1	79
89	GT	1	79
90	BY	1	79
91	TN	1	79
92	VA	1	79
93	CM	1	79
94	Lebanon	1	79
95	GL	1	79
96	ZW	1	79
97	MZ	1	79
98	PY	1	79
99	SK	1	79
100	TZ	1	79
101	MU	1	79
102	FO	1	79
103	AO	1	79
104	CY	1	79
	NULL	143	

SQL was used to make a table to identify production countries by their appearances. As expected, the US has an overwhelming number, over 2000. Countries in the top 10 rankings have three figures, while countries from 11th place have two figures only. Series miss production countries were more than 100, and there are countries that appeared just a couple of times.

(3) Analysis and visualisation

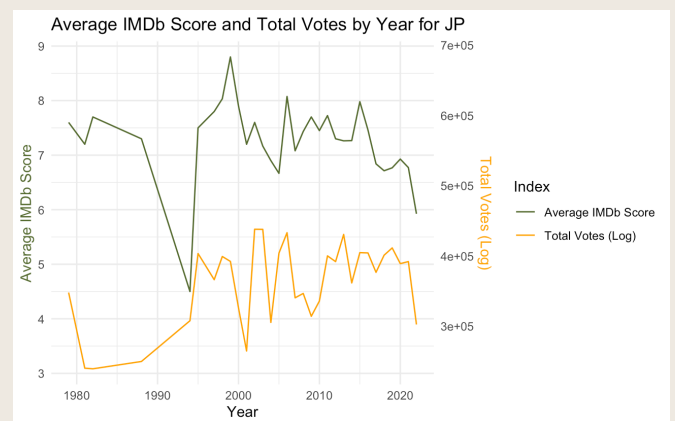
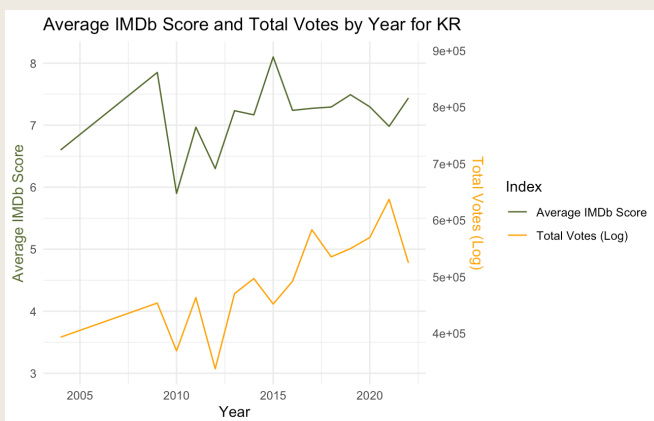
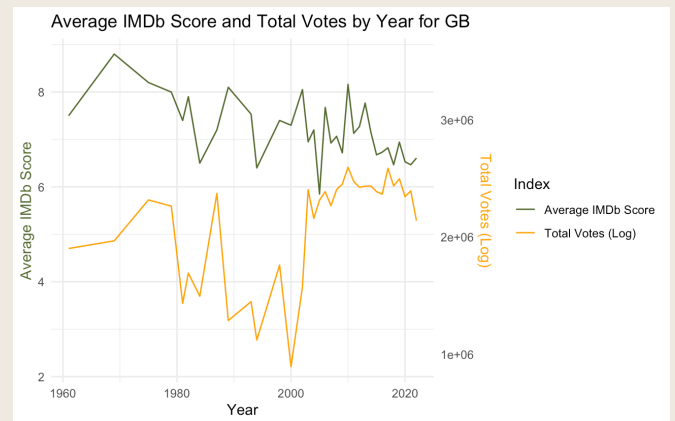
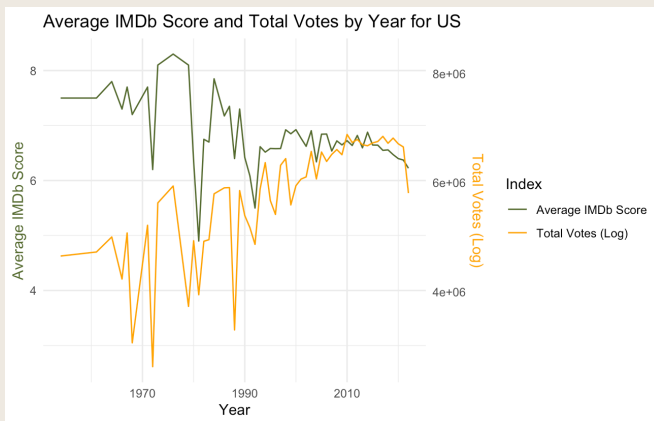
5. Correlation of production countries and score (cont'd)



A box plot created with the top twenty production countries by their works' IMDb scores, in the order of their popularity. The higher the ranking, the more outliers are found, whereas the median does not correspond to the rank. Korea, in the eighth place, has the highest median with 7.4. Additionally, Taiwan, in the 19th place also shows a great indication at 7. Nigeria shows the lowest number, which is less than 6.

(3) Analysis and visualisation

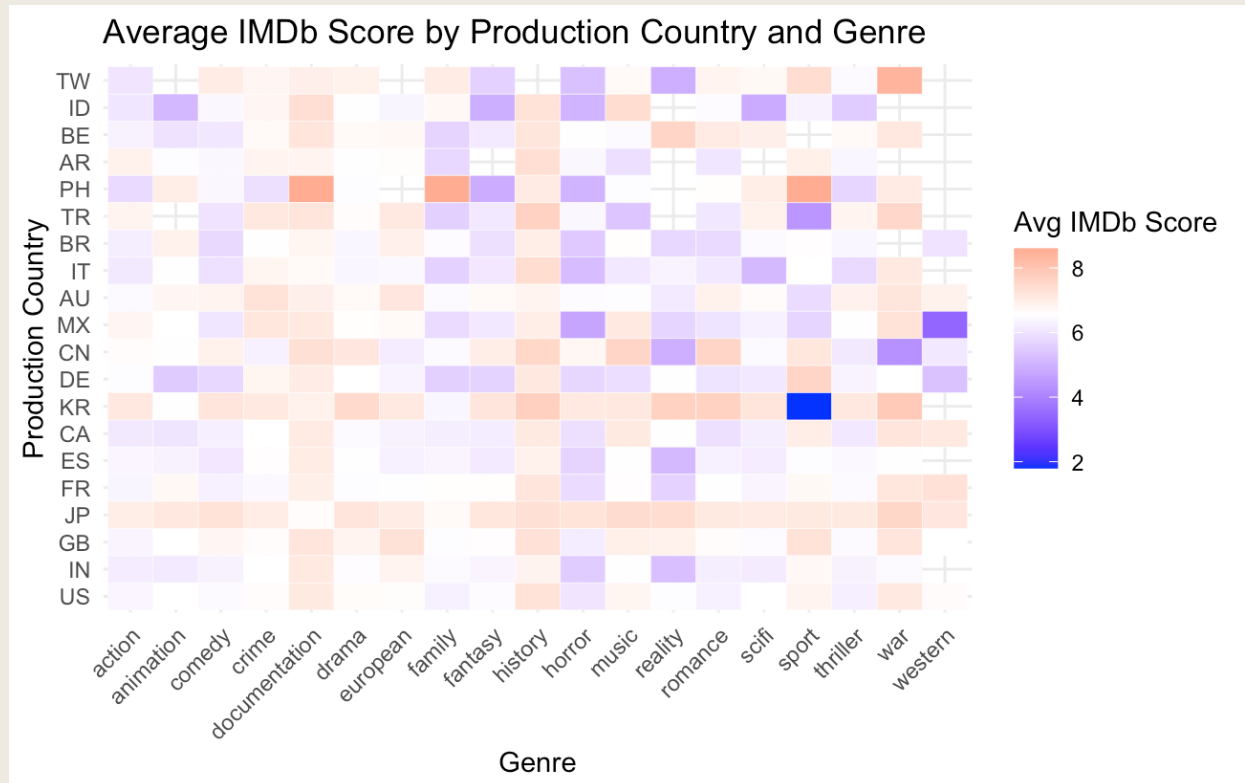
6. Specific production countries' scores & votes trend over time



Line graphs - two lines in a graph employed to see the relationship between IMDb votes and scores by production countries over time. Four countries were selected: the US and UK which are the most popular countries producing internationally popular works, and Korea and Japan where are the newly acclaimed production countries. There is a overall tendency that fluctuation in both lines are generally corresponding, while in recent years, especially around 2020, higher votes do not guarantee higher scores.

(3) Analysis and visualisation

7. Countries, genres, score



Getting warmer colours means a country's strength in specific genres, such as Taiwan could be great at war, Philippines might be good at documentation, family and sport. On the other hand, it clearly shows Korea had bad results in sport, India in animation and horror. Most countries have various colours in all genres, whereas Japan's blocks are generally orange, which means all genres from the country meet the standard.

03

Conclusion



Understanding from data analysis

This data gave a hint to refer to the decisions for future productions, in terms of genres, runtime and production countries based on the IMDb scores and votes of the released series.

The more people engage in leaving their ratings, the higher the score the series gets. A promotion strategy that encourages the viewers to leave comments on IMDb could lead to higher ratings.

In terms of genres, drama and comedy are overwhelmingly popular, which proves that they are great for travelling the world. While each genre's score is not in a proportional relationship regarding higher score and higher frequency, interestingly, war, sport and history, which are in the less popular group, had a higher median, which means those genres have solid maniac fans.

It is revealed that Asian countries, Japan, Korea and Taiwan are the rising stars, while the US and the UK remain undiminished. However, some specific countries are great at or bad at specific genres, so it should be carefully considered if a production company from a specific country tries to make a series.

To summarise, it is recommended to produce a show rather than a film, which runs about 100 minutes. It is good to consider the production company's country, as people now have a broader view of non-English content. Especially Asian countries are a promising territory given the fact that their IMDb scores are relatively high as well as the production cost is competitively low. It is a great opportunity to invest in more in Asian series.

The limitations of this data analysis are very clear. Data is not up-to-date, and it is not 100 per cent reliable as stated in the disclaimer in the beginning. The analysis was solely from the data, external research was excluded. Once the accurate and latest data is set, it is expected to have a more precise analysis.

The end

Haena Cho
haena.ms.cho@gmail.com
