TOP NETFLIX MOVIES

ANDREW NOMAI, GERALYNN ROEBKE, NATNAEL BERHANU, VALENTIN DEL RIO

OUR TOPIC

We are looking at a dataset from Kaggle based upon the Netflix movie recommender system. This data is from web scraping through the IMDB website on movies that are available on Netflix. Our goal is to use the data obtained to create a movie recommendation system using machine learning, to predict which movies/shows would be successful and in turn, provide a base formula for what could be considered a successful movie in the future. This would provide a formula for producers on what creating a successful movie looks like.

- What genre has the best votes/ratings by year?
- What is, if any, the correlation between duration of movie and rating?
- What is, if any, the correlation between age-rating of show/movie and the year it was released? In addition, what is the correlation between more mature ratings and the popularity of a movie/show?

DATA EXPLORATION

Our goal with data exploration is clean our data and sort them by year released and their rating. We will utilize this cleaned data to conduct our analysis.

	title	year	certificate	duration	genre	rating	description	stars	votes
0	Cobra Kai	(2018–)	TV-14	30 min	Action, Comedy, Drama	8.5	Decades after their 1984 All Valley Karate Tou...	['Ralph Macchio, ', 'William Zabka, ', 'Courtn...	177,031
1	The Crown	(2016–)	TV-MA	58 min	Biography, Drama, History	8.7	Follows the political rivalries and romance of...	['Claire Foy, ', 'Olivia Colman, ', 'Imelda St...	199,885
2	Better Call Saul	(2015–2022)	TV-MA	46 min	Crime, Drama	8.9	The trials and tribulations of criminal lawyer...	['Bob Odenkirk, ', 'Rhea Seehorn, ', 'Jonathan...	501,384
3	Devil in Ohio	(2022)	TV-MA	356 min	Drama, Horror, Mystery	5.9	When a psychiatrist shelters a mysterious cult...	['Emily Deschanel, ', 'Sam Jaeger, ', 'Gerardo...	9,773
4	Cyberpunk: Edgerunners	(2022–)	TV-MA	24 min	Animation, Action, Adventure	8.6	A Street Kid trying to survive in a technology...	['Zach Aguilar, ', 'Kenichiro Ohashi, ', 'Emi ...	15,413
...	...	...	...	...	...	...	...	...	...
9952	The Imperfects	(2022–)	TV-MA	45 min	Action, Adventure, Drama	6.3	After an experimental gene therapy turns them ...	['Morgan Taylor Campbell, ', 'Italia Ricci, ',...	3,130
9953	The Walking Dead	(2010–2022)	TV-MA	44 min	Drama, Horror, Thriller	8.1	Sheriff Deputy Rick Grimes wakes up from a com...	['Andrew Lincoln, ', 'Norman Reedus, ', 'Melis...	970,067
9954	The Crown	(2016–)	TV-MA	58 min	Biography, Drama, History	8.7	Follows the political rivalries and romance of...	['Claire Foy, ', 'Olivia Colman, ',	199,898
9955	Supernatural	(2005–2020)	TV-14	44 min	Drama, Fantasy, Horror	8.4	Two brothers follow their father's footsteps a...	['Jared Padalecki, ', 'Jensen Ackles, ', 'Jim ...	439,601
9956	Devil in Ohio	(2022)	TV-MA	356 min	Drama, Horror, Mystery	5.9	When a psychiatrist shelters a mysterious cult...	['Emily Deschanel, ', 'Sam Jaeger, ', 'Gerardo...	9,786

	title	year released	certificate	duration	genre	rating	votes
1445	Scarface	1932	PG	93 min	Action, Crime, Drama	7.7	28,107
1119	Dumbo	1941	G	64 min	Animation, Adventure, Drama	7.2	131,896
4072	Nazi Concentration Camps	1945	Not Rated	59 min	Documentary, History	8.3	1,481
6348	Know Your Enemy - Japan	1945	Not Rated	63 min	Documentary, History, War	6.0	800
4627	Along Came Jones	1945	Approved	90 min	Comedy, Romance, Western	6.4	2,163
...	...	...	...	...	...	...	...
3668	Hurts Like Hell	2022	TV-MA	200 min	Biography, Drama, Sport	7.4	247
3551	Toolsidas Junior	2022	TV-14	132 min	Drama, Family, Sport	7.4	2,162
296	Incantation	2022	TV-MA	110 min	Horror, Mystery	6.2	11,424
3711	Fortune Seller: A TV Scam	2022	TV-14	45 min	Documentary, Biography, Crime	6.4	69
306	Heartstopper	2022	TV-14	30 min	Drama, Romance	8.7	49,420

CLEANING PHASE 1

We started off with the raw dataset from the kaggle web scraping and wanted to remove any N/A values.

CLEANING PHASE 2

After cleaning out the N/A values, we sort the data by year released.

DATA ANALYSIS

ANALYSIS

	title	year released	certificate	duration	genre	rating	votes
9444	BoJack Horseman	2014	TV-MA	26 min	Animation, Comedy, Drama	9.9	16,066
8161	Avatar: The Last Airbender	2005	TV-Y7	25 min	Animation, Action, Adventure	9.6	5,221
8259	Dexter	2006	TV-MA	56 min	Crime, Drama, Mystery	9.6	10,604
8557	JoJo's Bizarre Adventure	2012	TV-MA	25 min	Animation, Action, Adventure	9.6	1,442
9578	Stranger Things	2016	TV-14	98 min	Drama, Fantasy, Horror	9.6	36,276
8907	Avatar: The Last Airbender	2005	TV-Y7	24 min	Animation, Action, Adventure	9.6	3,953
8819	Dark	2017	TV-MA	54 min	Crime, Drama, Mystery	9.5	14,007
9586	Avatar: The Last Airbender	2005	TV-Y7	25 min	Animation., Action, Adventure	9.5	4,740
9951	Breaking Bad	2008	TV-MA	49 min	Crime, Drama, Thriller	9.5	1,831,359
9608	Anne with an E	2017	TV-PG	44 min	Drama, Family	9.5	2,488
8044	Avatar: The Last Airbender	2005	TV-Y7	25 min	Animation, Action, Adventure	9.5	5,383
9445	BoJack Horseman	2014	TV-MA	27 min	Animation, Comedy, Drama	9.5	6,230
7	Breaking Bad	2008	TV-MA	49 min	Crime, Drama, Thriller	9.5	1,831,340
8976	JoJo's Bizarre Adventure	2012	TV-MA	25 min	Animation, Action, Adventure	9.4	1,192
9795	Avatar: The Last Airbender	2005	TV-Y7	22 min	Animation, Action, Adventure	9.4	4,138

Here we sorted the data by rating but in descending order, listing the top rated movies/shows since year 1932. As shown in the DataFrame, the title " BoJack Horseman" is rated the highest with 9.9, released in 2014 with certificate of TV-MA.. On the contrary, the title with the lowest score is "Defcon 2012." released in 2010 with a rating of 1.7 and a certificate that is not rated.