



# Amazon Prime Video

*Movies and TV Shows Analysis*

---

**Presented By**

Ade Ariyo Yudanto

# Data Understanding & Problem Statement

The research will look at the Kaggle dataset of Amazon Prime Movies and TV Shows. This tabular dataset contains listings for all of the movies and TV shows available on Amazon Prime, as well as information such as cast, directors, ratings, release year, duration, etc. The following insights will be obtained:

1. How many shows by type of show?
2. How many shows by release year of the show?
3. Which rating has the most shows?
4. What movie has the longest duration?
5. What TV Show has the most seasons?
6. What is the average duration of a Movie by the top 10 directors with the most movies?

## Amazon Prime Movies and TV Shows

Movies and TV Shows listings on Amazon Prime Video



[Data Card](#) [Code \(29\)](#) [Discussion \(2\)](#)

### About Dataset

#### Amazon Prime Video - Movies and TV Shows

**About this Dataset:** Amazon Prime is another one of the most popular media and video streaming platforms. They have close to 10000 movies or tv shows available on their platform, as of mid-2021, they have over 200M Subscribers globally. This tabular dataset consists of listings of all the movies and tv shows available on Amazon Prime, along with details such as - cast, directors, ratings, release year, duration, etc.\*

**Usability** ⓘ  
10.00

**License**  
[CC0: Public Domain](#)

**Expected update frequency**  
Monthly

**Tags**

# Tools



Google  
Sheets



Looker

Google Sheets is used to clean, enhance, analyze, and visualize data. While Looker is also used for data visualization.

# Data Cleaning

## Remove Duplicate Data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description			
2	s1	Movie	The Grand Sedu Don McKellar	Brendan Gleeson	Canada		March 30, 2021	2014		113 min	Comedy, Drama	A small fishing village must procure a local doctor t			
3	s2	Movie	Take Care Good Josh Joshi	Mallesh Manjrek	India		March 30, 2021	2014	13+	110 min	Drama, Internati	A metro family decides to fight a Cyber Criminal th			
4	s3	Movie	Secrets of Cicely Josh Wedder	Tom Sciorone	United States		March 30, 2021	2017		74 min	Action, Drama	S After a man discovers his wife is cheating on him w			
5	s4	Movie	Pink Staying Tri Sonia Anderson	Interviews with	United States		March 30, 2021	2014		69 min	Documentary	Pink breaks the mold once again, bringing her care			
6	s5	Movie	Monster Maker Giles Foster	Harry Dean Star	United Kingdom		March 30, 2021	1989		45 min	Drama, Fantasy	Teenage Matt Bunting wants to work with a famous			
7	s6	Movie	Living With Dino Paul Welland	Gregory Chahal	United Kingdom		March 30, 2021	1989		52 min	Fantasy, Kids	The story unfolds in a an English seaside town, wh			
8	s7	Movie	Hired Gun Fran Stone	Alice Cooper	United States		March 30, 2021	2017		98 min	Documentary, Sy	They are the "First Call," A-list musicians, just 20 h			
9	s8	Movie	Grease Live Thomas Kail	Aki Julianne Hough	United States		March 30, 2021	2016		131 min	Comedy	This honest, uncompromising comedy chronicles th			
10	s9	Movie	Global Makedown Daniel Gilboy	Michael Paré	Canada		March 30, 2021	2017		87 min	Action, Science	A helicopter pilot and an environmental scientist le			
11	s10	Movie	Drew's Mother	Robert Allan Ack	Kirstie Alley	San United States		April 1, 2021	1994		Drama	Sally Goodson is a devoted mother to her autistic s			
12	s11	Movie	Forest Faires Justin G. Dyck	Emily Wilder	Ad Canada		April 4, 2021	2016		88 min	Adventure, Kids	Amanda stumbles upon a hidden village of faeries i			
13	s12	Movie	Take Care Lu Turcio	Leslie Bibb	Kev United States		April 10, 2021	2014		90 min	Comedy	When a car crash leaves Frances immobilized, she			
14	s13	Movie	The Night Eats T Dominique Roch	Andres Danieles	France		April 17, 2021	2016		94 min	Horror, Suspense	After waking up in an apartment the night after a			
15	s14	Movie	Resistencia Jap Barcelona	Rafinha Alcantar	Spain		April 24, 2021	2020		49 min	Documentary, Sy	The documentary follows the midlife's everyday			
16	s15	Movie	Elon Musk: The Elon Musk	Sonia Anderson	Elon Musk	Per United Kingdom		May 2, 2021	2018	74 min	Documentary	Discover the meteoric rise of Elon Musk, the man v			
17	s16	Movie	Summer '03 Becca Hayes	Joe King	Jack United States		June 3, 2021	2019		96 min	Comedy, Drama	Summer '03 is the hilarious story of 16-year-old J			
18	s17	Movie	Zoombies Glenn Miller	Marcus Anderson	Katwi Lyman	Andrew Asper		2016	13+	87 min	Horror, Science	I When a strange virus quickly spreads through a sa			
19	s18	TV Show	Zoe Babies	Narrator - Gillian Barlett				2008	ALL	1 Season	Kids, Special In	A heart warming and inspiring series that welcome			
20	s19	TV Show	Zoe Coombs Marr	Zoe Coombs Marr				2020	18+	1 Season	Comedy, Talk Sh	Zoe Coombs Marr has been on hiatus. Sort of. For			
21	s20	Movie	Zoe Drake Doremus	Ewan McGregor	Léa Seydoux	Thomas James	Mirar	2016	R	104 min	Science Fiction	ZOE tells a tale of forbidden love between an engi			
22	s21	TV Show	Zombocomo					2001	TV-Y	1 Season	Kids	Chris and Martin Kutz bring their enthusiasm for a			
23	s22	TV Show	Zo Zo Zo Zombi	William Smith	Alex Cazares	Ben Diskin	Cristina Vee	Erika Ishii	2020	TV-Y7	1 Season	Action, Anime	Zo Zo Zo Zombi is our undead hero who, with his be		
24	s23	Movie	Zis Boom Bah	Miss High	Grace Hayes	Mary Healy		1941	RN	62 min	Arts, Entertainment	"Hey, kids, let's get together and put on a show," hi			
25	s24	TV Show	ZoroZero	ANDREA RISEB				2020	18+	1 Season	TV Shows	ZoroZero series follows the journey of a cocar			
26	s25	Movie	Zandalee Sam Pillsbury	Nicolas Cage	Judge Reinhold	Erika Anderson	Jo	1991	R	94 min	Drama	Zandalee has gone bored with her post-21 life			
27	s26	Movie	Zambozi Fre Dr. Rudolf Lammers					2005	ALL	50 min	Documentary	The Zambozi river is the fourth-longest river in Afr			
28	s27	TV Show	Zahara The Return	Natasha Heschelle	Curtis Morgan	Hannah Scott		2020	16+	1 Season	TV Shows	The story explores the history of slavery and racis			
29	s28	TV Show	Z. The Beginning of Everything	Christina Ricci	United States			2017	18+	1 Season	Comedy, Drama	"Z. The Beginning of Everything" tells the story of Z			
30	s29	TV Show	Yu-Gi-Oh! Zexal	El Juy Marc Thompson	Sean Schenmel	Eileen C		2015	TV-Y7	3 Seasons	Animation, Axi	It's time to feel the fire as Yu-Gi-Oh! Zexal leads c			
31	s30	TV Show	Yu-Gi-Oh! ARC-V	Michael Liscio Jr.	Emily Bauer			2016	TV-Y7	3 Seasons	Action, Adventu	Get set to swing into action as Yu-Gi-Oh! ARC-V ta			

 amazon\_prime\_video\_data\_set

Count: 9/38



	A	B	C	D	E	F	G	H	I	J	K	L	M	N		
9640	s0639	Movie	Oven	Kishore Tummala	Lauren Ojeda			2019	7+	2 min	Drama	A young mom's lapse in attention leads to a calam				
9641	s0640	Movie	No Sleep 'Til Chi Phai Trail	Dave Annable	Odetta Annable	Sheryl Lee Ralph		2018	16+	85 min	Drama	Lizzie and Billy only fall asleep next to each other.				
9642	s0641	Movie	Meet the Spartans	Jason Friedberg	Sean Hagane	Carmen Electra	Ken Davitian	2008	18+	86 min	Comedy	From the creators of Scary Movie and Date Movie i				
9643	s0642	Movie	Marley & Me	David Frankel	Owen Wilson	Julia Roberts	United States	2008	PG	116 min	Comedy, Drama	Jennifer Aniston and Owen Wilson unleash hope i				
9644	s0643	TV Show	Lunkplunkville		Michael de Avila		October 10, 2021	2021	13+	2 Seasons	Documentary	St Lunkplunkville's launches "Got Folks Fishing," a social				
9645	s0644	Movie	Love & Other Dr	Edward Zwick	K.K. Jike	Gyllen	United States	2010	R	112 min	Comedy, Drama	A womanizing pifer drug rep with no sense of dre				
9646	s0645	Movie	Little Miss Sus	Jonathan Dayton	Abigail Breslin	C United States		2009	R	101 min	Comedy, Drama	A family determined to get their young daughter in				
9647	s0646	TV Show	Let's Go Cozy Coupel	- Cartoons for Kids				2021	ALL	1 Season	Kids	Cozy Coupe is an online series that helps kids to le				
9648	s0647	Movie	Knight and Day	20th century	for Tom Cruise	Cameron Diaz	Peter Sarsgaard	Jordi	2010	PG-13	117 min	Action, Adventur	SPECIAL EXTENDED EDITION! Get even MORE			
9649	s0648	TV Show	Kids In The Hall					October 10, 2021	1995	TV-14	5 Seasons	Arts, Entertainment	"The Kids in the Hall" is a unique and bizarre sketch			
9650	s0649	Movie	Judge Dredd	Danny Cannon	Sylvester Stallone	Armand Assante	Rob Schneider		1995	R	96 min	Suspense	Judge Dredd is convicted for a crime he did not cc			
9651	s0650	Movie	Jack	Frances Ford Co	Robin Williams	Dane Lane	Brian Kernin	Jennifer	1996	PG-13	113 min	Comedy	Robert Williams stars as a 15-year-old in a grown-up			
9652	s0651	Movie	I Can't Breathe	Kishore Tummala	Michele Anderson	Marc Jean	Christopher David		2017	18+	32 min	Drama, Sports	An African American single mother's emotional jou			
9653	s0652	TV Show	Glass Room		Elizaveta Maskaya	Vsevolod Boldin	Alexander Ki		2019	16+	1 Season	Drama	Geography teacher Elena lives with her husband's			
9654	s0653	Movie	FanDoo	(4K UHD)	Pappy Romanoff	Stera Ebosta	Diplo	Steve Aoki	Russian	2016	G	111 min	Documentary	Musclfin with unpublished images from the life of		
9655	s0654	Movie	Date Night	(Exte	Shawn Levy	Mark Wahlberg	Steve Carell	Taraji P. Henson	Tin	2010	UNRATED	102 min	Comedy	A typical suburban couple's date night becomes m		
9656	s0655	Movie	Dart Divorce	Claf	Intenbach	Marina Intenbach	Daryl Jackson	James Butler	B	2008	18+	87 min	Drama, Horror	What started with a routine divorce between Natali		
9657	s0656	TV Show	DiFundamentals Series		Emile Crawford	Randall Crawford	Jon Ward		2021	TV-Y	1 Season	Animation, Kids	Learn "diFundamentals" through song and fun cha			
9658	s0657	Movie	Anaganaga Oka	Varnshi P	Vjay	Siveta chaudhary	NNR CHOWDARY	Hanya	2021	18+	135 min	Action, Drama	S (Peter)hero) Who went to the forest to find a cure fo			
9659	s0658	Movie	A Place at the Ti	Kristi Chang	Jeff Bridges	Tom Colicchio			2013	PG	84 min	Documentary, Sy	The next chapter on America's food crisis is her			
9660	s0659	Movie	12 Dates of Christmas		Heath Ledger	Julia Stiles	Joseph Gordon-Livitt	L	2011	PG	87 min	Drama, Romance	Setup on a Christmas Eve date, a woman must reli			
9661	s0660	Movie	10 Things I Hate G	Jagger					1999	PG-13	97 min	Comedy, Drama	On the first day of his new school, Cameron instat			
9662	s0661	Movie	The Man in the I	John-Paul David	Caran Hinds	Stephen Dillane	Malvern		2021	13+	96 min	Comedy	The Man in the Hat pursues through France in a			
9663	s0662	TV Show	Temahawks		DENISE BRYER	WINDSOR DAVIES	JEREMY HI		1983	ALL	3 Seasons	TV Shows	This thrilling, action-packed series says Gerry Ande			
9664	s0663	Movie	River	Emily Skye	Mary Cameron Rogers	Alexandra Rose	Rob Mann		2021	16+	93 min	Drama, Science	River is a grounded Sci-Fi mystery Thriller, that fol			
9665	s0664	Movie	Pride Of The Bo	Joseph H	Lewis	Lee Gorczy	Bobby Jordan		1940	7+	60 min	Comedy	New York City street principles get an East Side Ki			
9666	s0665	TV Show	Planet Patrol		DICK VOSSBURGH	RONNIE STEVENS	LIBBY MI		2018	13+	4 Seasons	TV Shows	This is Earth, 2100AD - and these are the adventur			
9667	s0666	Movie	Outpost	Steve Barker	Ray Stevenson	Julian Wadham	Richard Brake	M	2008	R	90 min	Action	In war-torn Eastern Europe, a world-weary group o			
9668	s0667	TV Show	Maradona: Blessed Dream		Esteban Recagno	Ezequiel Streinz	Luciano Vittor		2021	TV-MA	1 Season	Drama, Sports	The series tells the story of Diego Maradona, a boy			
9669	s0668	Movie	Harry Brown	Daniel Barber	Michael Caine	Emily Mortimer	Joseph Gilgun	Sei	2010	R	103 min	Action, Drama	A Harry Brown, starring two-time Academy Award wi			
9670																

+

=

amazon\_prime\_video\_dataset

Count: 9,668

9738 Data Before Cleaned

9662 Data After Cleaned

9738 data in this dataset have multiple duplicates. Then, there are 9667 unique data once duplicate data is deleted.

# Data Cleaning

## Fix the Title of Data

show_id	type	title
s3569	Movie	#Unfit: The Psychology of Donald Trump
s4353	Movie	#Lucky Number
s5548	Movie	"Mixed Up"
s7557	Movie	#Home
s3570	TV Show	#Lagira de #eldisco
s3571	TV Show	#IMomSoHard Live
s3573	TV Show	_DUPE_The Making of the Mob: Chicago
s5607	TV Show	#IGotThis
s5978	TV Show	"The Paramedic Angel"
s8853	TV Show	#WASHED



show_id	type	title
s3569	Movie	Unfit: The Psychology of Donald Trump
s4353	Movie	Lucky Number
s5548	Movie	Mixed Up
s7557	Movie	Home
s3570	TV Show	Lagira de eldisco
s3571	TV Show	IMomSoHard Live
s3573	TV Show	The Making of the Mob: Chicago
s5607	TV Show	IGotThis
s5978	TV Show	The Paramedic Angel
s8853	TV Show	WASHED

10 data in this dataset contain unusual characters or words as the title of a Movie or TV Show. As a result, they must be fixed or cleaned.

# Data Manipulation

## Handle Missing Data

Some columns have several missing data:

1. 9507 out of 9662 data (98.40%) are empty in column **date\_added**. Instead of deleting it, the data is labeled with “Missing Data”.
2. 2098 out of 9662 data (21.70%) are empty or have a number value in column **director**. Instead of deleting it, the data is labeled with “Unknown Director” if the type of show is Movie and labeled with “Missing Data” if the type of show is TV Show.
3. 1268 out of 9662 data (13.12%) are empty or have a number value in column **cast**. Instead of deleting it, the data is labeled with “Unknown Cast” if the type of show is Movie and labeled with “Missing Data” if the type of show is TV Show.
4. 8990 out of 9662 data (93.04%) are empty in column **country**. Instead of deleting it, the data is labeled with “Missing Data”.
5. 335 out of 9662 data (3.47%) are empty in column **rating**. Instead of deleting it, the data is labeled with “Unrated”.
6. 20 out of 9662 data (0.21%) are empty in column **description**. Instead of deleting it, the data is labeled with “Undescribed”.

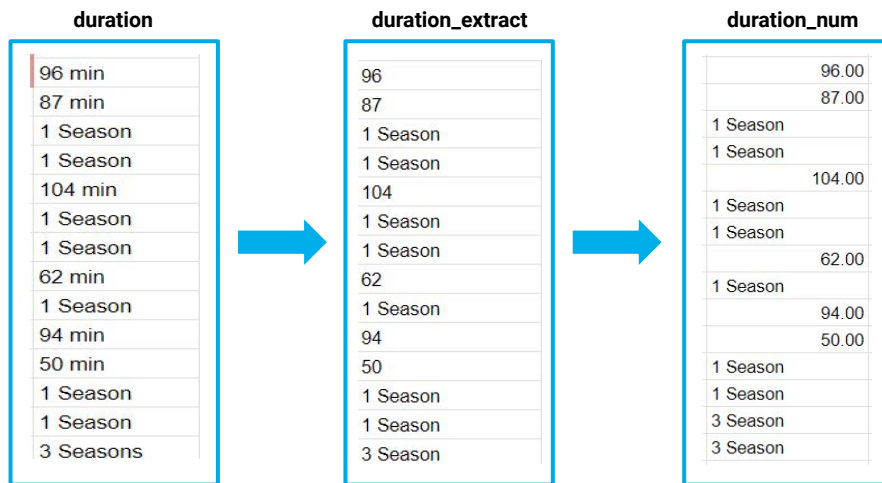
show_id	type	title	director	cast	country	date_added
s21	TV Show	Zoboomafoo	Unknown Director	Unknown Cast	Missing Data	Missing Data
s34	Movie	Young Love (at the Sun Factory)	Unknown Director	Unknown Cast	Missing Data	Missing Data
s35	Movie	Young Love (at Bowery Ballroom)	Unknown Director	Unknown Cast	Missing Data	Missing Data
s73	TV Show	WordWorld	Unknown Director	Unknown Cast	Missing Data	Missing Data
s75	TV Show	WordGirl	Unknown Director	Unknown Cast	Missing Data	Missing Data
s79	TV Show	Wonder Balls	Unknown Director	Unknown Cast	Missing Data	Missing Data
s99	TV Show	Wildest Places	Unknown Director	Unknown Cast	Missing Data	Missing Data
s101	TV Show	Wild North	Unknown Director	Unknown Cast	Missing Data	Missing Data
s107	TV Show	Why I Ran	Unknown Director	Unknown Cast	Missing Data	Missing Data
s129	TV Show	What's Up In Space: The Solar System For Kids	Unknown Director	Unknown Cast	Missing Data	Missing Data
s135	TV Show	WGC-Dell Technologies Match Play Reveal Sho	Unknown Director	Unknown Cast	Missing Data	Missing Data
s139	TV Show	WellieWishers	Unknown Director	Unknown Cast	Missing Data	Missing Data
s150	TV Show	Watch the Game Like Never Before with X-Ray	Unknown Director	Unknown Cast	Missing Data	Missing Data
s171	TV Show	Waco: Madman or Messiah	Unknown Director	Unknown Cast	Missing Data	Missing Data
s172	TV Show	Voyage of the Continents	Unknown Director	Unknown Cast	Missing Data	Missing Data
s177	TV Show	Vocabulary and Numbers with Harry the Bunny	Unknown Director	Unknown Cast	Missing Data	Missing Data
s195	TV Show	Vella Raja (Tamil)	Unknown Director	Unknown Cast	Missing Data	Missing Data
s196	TV Show	Vella Raja (Hindi)	Unknown Director	Unknown Cast	Missing Data	Missing Data
s198	TV Show	Vatican Miracle Examiner	Unknown Director	Unknown Cast	Missing Data	Missing Data
s209	Movie	USAISAMONSTER	Unknown Director	Unknown Cast	Missing Data	Missing Data
s212	TV Show	UQ Holder!	Unknown Director	Unknown Cast	Missing Data	Missing Data
s248	TV Show	TuTu Town	Unknown Director	Unknown Cast	Missing Data	Missing Data
s261	TV Show	Troy The Train	Unknown Director	Unknown Cast	Missing Data	Missing Data
s268	TV Show	Trapped [English Subtitled]	Unknown Director	Unknown Cast	Missing Data	Missing Data
s299	TV Show	Tom's Paint Shop	Unknown Director	Unknown Cast	Missing Data	Missing Data

# Data Transformation

## Transform Type of Data

Duration is the most important column in this analysis for solving the mentioned challenges. Because the duration column contains a text value, it will be transformed into a number.

1. Add a new column called **duration\_extract**.
2. Extract data from the **duration** column that only contains the term "min" into a single numeric value in the duration\_extract column.
3. Add another column called **duration\_num**.
4. Values from column **duration\_extract** should be copied and pasted into column **duration\_num** using the paste special "values only" command.
5. Column **duration\_num**'s data type should be changed to a number.



# Data Transformation

## Group the Data

Several Movie and TV Show ratings need to be corrected and adjusted in column rating based on the Maturity Rating in the United States, which is divided into 2 local value ratings: MPAA (Movie) and TVPG (TV Show).

1. Add a new column named **rating\_extract**.
2. Data with values “UNRATED”, “TV-NR”, “NR” and “NOT\_RATE” in column **rating** filled with “Unrated” in **rating\_extract** column.
3. Data with values “ALL” and “ALL\_AGES” in column **rating** filled with “G” if Movie and “TV-G” if TV Show in **rating\_extract** column.
4. Data with value “7+” in **rating** column filled with “PG” if Movie and “TV-Y7” if TV Show in column **rating\_extract**.
5. Data with values “16”, “16+” and “AGES\_16\_” in column **rating** filled with “R” if Movie and “TV-14” if TV Show in **rating\_extract** column.
6. Data with values “18” and “AGES\_18\_” in column **rating** filled with “NC-17” if Movie and “TV-MA” if TV Show in **rating\_extract** column.
7. After grouping and adjusting the rating of Movies and TV Shows, there are 701 out of 9667 data (7.25%) labeled with “Unrated” in column **rating\_extract**.



# Data Quality

## Exclude Outlier

The duration of the Movie must be adjusted based on whether or not it falls into the outlier category.

1. Add a new column named **outlier**.
2. Divide the number of data points into four parts, or quarters, of roughly equal size and determine the bottom and upper bounds using the quartile formula. The result is as follows:

Q1	1	75.00
Q2	2	91.00
Q3	3	106.00
IQR		31.00
LB		28.50
UP		152.50

→ Lower Bound

→ Upper Bound

3. Outliers are movies with a duration of less than 30 minutes and more than 152 minutes.
4. Outliers are found in 623 of 7808 data (7.98%). As a result, 7185 Movies fit the requirements.

# Data Quality

## Exclude Missing Data

Several empty rows in several columns were discovered during the prior handling missing data process. That empty data will be removed before it is visualized.

1. Movies identified as outliers or have missing data/unknown/unrated in column **director**, **cast**, and **rating\_extract** will be excluded. 6116 Movies meet the criteria.
2. TV Shows that have missing data/unrated in column **rating\_extract** will be excluded. 1714 TV Shows meet the criteria.
3. So, there are 7830 out of 9662 (81.04%) Movies and TV Shows which ready to serve and visualized after simple exploratory data analysis.

# Data Visualization

## Shows by Type of Show

The proportion of Movies in this dataset is 78.1%. As a result, Amazon Prime Video should think about adding TV Shows to its database.

### Type of Show

Percentage of Show

**TV Show**

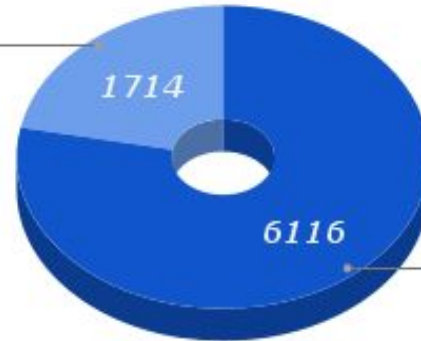
21.9%

1714

6116

**Movie**

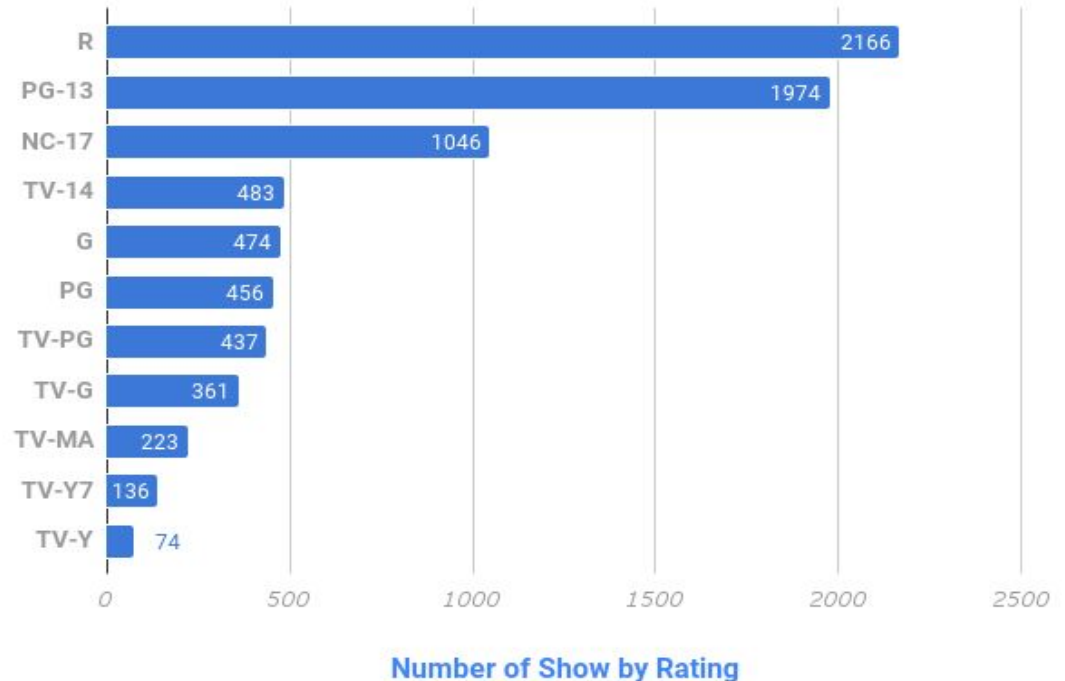
78.1%



# Data Visualization

## Number of Shows by Rating

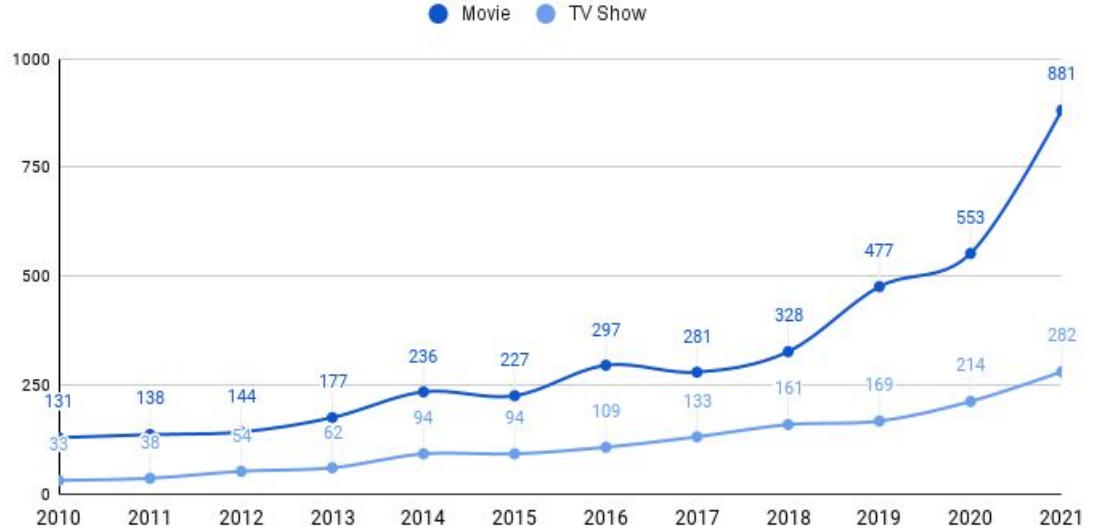
The majority of shows on Amazon Prime Video have an "R" rating, followed by PG-13. Amazon Prime Video should examine who its targeted users are, to provide additional Movies or TV Shows to its targeted users.



# Data Visualization

## Number of Shows by Release Year

The number of shows by release year at Movie significantly increased starting from 2019, although there was a slight decrease in 2015 & 2017. Meanwhile, the number of TV shows is steadily expanding year after year.



Trend of Show by Release Year

# Data Visualization

## Longest Duration Movie & Most Seasons TV Show

	Title	Duration (Minutes) ▾
1.	Dhoom:2	152
2.	Manadhai Thiruvittai	152
3.	Lovely	152
4.	Bachna Ae Haseeno	152
5.	Baba Keno Chakar	152
1 - 5 / 6116 < >		

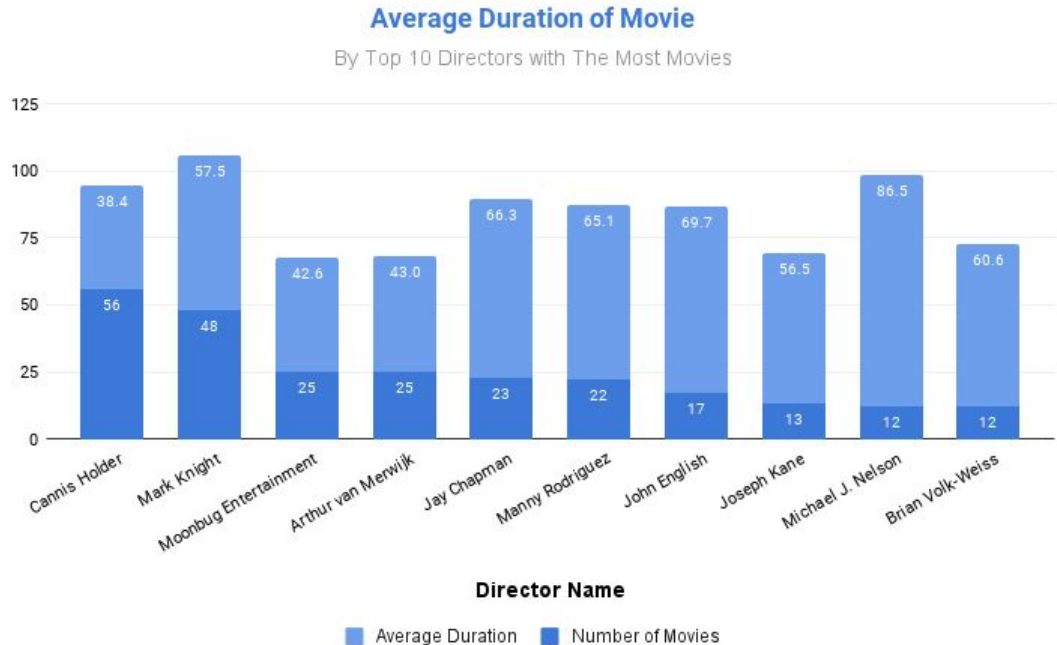
	Title	Number of Seasons ▾
1.	The Amazing Race	29
2.	Silent Witness	21
3.	Survivor	19
4.	Top Gear (UK)	15
5.	Paula's Best Dishes	14
1 - 5 / 1714 < >		

"Dhoom: 2" has the longest running time, clocking up at 152 minutes, or around 2.5 hours. With 29 seasons, "The Amazing Race" has the most seasons of any TV show.

# Data Visualization

## Average Duration of Movie By Top 10 Directors

With 56 movies, Cannis Holder is the most prolific director. However, the average duration of the movie is 38.4 minutes, indicating that he directs more short films. Michael J. Nelson, on the other hand, has only directed 12 films (the worst two out of ten), but the average duration of the movie is 86.5 minutes, which is longer than the others.



## MOVIE

**6116**

Total Count

## TV SHOW

**1714**

Total Count

### Longest Duration Movie

**Dhoom: 2**

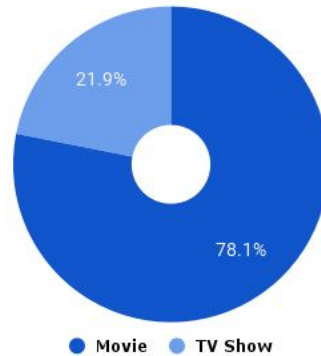
152 Minutes

### Most Seasons TV Show

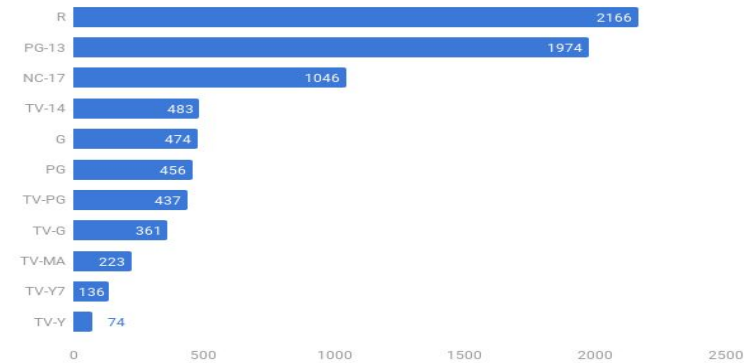
**The Amazing Race**

29 Seasons

Percentage of Show

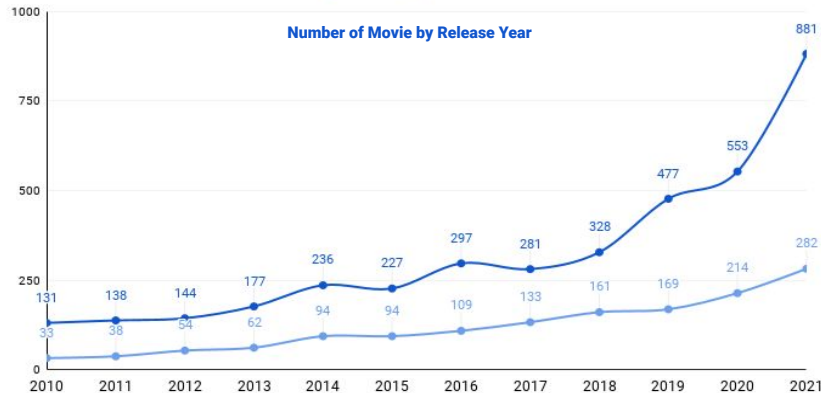


Number of Show by Rating



● Movie ● TV Show

Number of Movie by Release Year



Average Duration of Movie by Top 10 Directors

