```
!gdown https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv
```

```
Downloading...
From: https://d2beiqkhq929f0.cloudfront.net/public_assets/assets/000/000/940/original/netflix.csv
To: /content/netflix.csv
100% 3.40M/3.40M [00:00<00:00, 12.2MB/s]
```

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
data = pd.read_csv('netflix.csv')
```

```
data.head()
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go

Next steps: [Generate code with data] [◯ View recommended plots] [New interactive sheet]

```
data.shape
```

```
(8807, 12)
```

```
data.describe()
```

	release_year
count	8807.000000
mean	2014.180198
std	8.819312
min	1925.000000
25%	2013.000000
50%	2017.000000
75%	2019.000000
max	2021.000000

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8807 entries, 0 to 8806
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   show_id      8807 non-null   object
```

```
 1   type          8807 non-null   object
 2   title         8807 non-null   object
 3   director      6173 non-null   object
 4   cast          7982 non-null   object
 5   country       7976 non-null   object
 6   date_added    8797 non-null   object
 7   release_year  8807 non-null   int64
 8   rating        8803 non-null   object
 9   duration      8804 non-null   object
10   listed_in     8807 non-null   object
11   description   8807 non-null   object
dtypes: int64(1), object(11)
memory usage: 825.8+ KB
```

```
movies = data.loc[df['type'] == 'Movie']
movies
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	September 25, 2021	2020	PG-13	90 min	Documentaries	As her father nears the end of his life, filmm...
6	s7	Movie	My Little Pony: A New Generation	Robert Cullen, José Luis Ucha	Vanessa Hudgens, Kimiko Glenn, James Marsden, ...	NaN	September 24, 2021	2021	PG	91 min	Children & Family Movies	Equestria's divided. But a bright-eyed hero be...
7	s8	Movie	Sankofa	Haile Gerima	Kofi Ghanaba, Oyafunmike Ogunlano, Alexandra D...	United States, Ghana, Burkina Faso, United Kin...	September 24, 2021	1993	TV-MA	125 min	Dramas, Independent Movies, International Movies	On a photo shoot in Ghana, an American model s...
9	s10	Movie	The Starling	Theodore Melfi	Melissa McCarthy, Chris O'Dowd, Kevin Kline, T...	United States	September 24, 2021	2021	PG-13	104 min	Comedies, Dramas	A woman adjusting to life after a loss contend...
12	s13	Movie	Je Suis Karl	Christian Schwochow	Luna Wedler, Jannis Niewöhner, Milan Peschel, ...	Germany, Czech Republic	September 23, 2021	2021	TV-MA	127 min	Dramas, International Movies	After most of her family is murdered in a terr...
...	...	...	...	...	...	...	...	...	...	...	...	...
8801	s8802	Movie	Zinzana	Majid Al Ansari	Ali Suliman, Saleh Bakri, Yasa, Ali Al-Jabri, ...	United Arab Emirates, Jordan	March 9, 2016	2015	TV-MA	96 min	Dramas, International Movies, Thrillers	Recovering alcoholic Talal wakes up inside a s...
8802	s8803	Movie	Zodiac	David Fincher	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States	November 20, 2019	2007	R	158 min	Cult Movies, Dramas, Thrillers	A political cartoonist, a crime reporter and a...
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Looking to survive in a world taken over by zo...
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	January 11, 2020	2006	PG	88 min	Children & Family Movies, Comedies	Dragged from civilian life, a former superhero...
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	March 2, 2019	2015	TV-14	111 min	Dramas, International Movies, Music & Musicals	A scrappy but poor boy worms his way into a ty...

Next steps: (Generate code with `movies`)   (🔘 View recommended plots)   (New interactive sheet)

```
tvshows = data.loc[df['type']=="TV Show"]
tvshows
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, TV Dramas, TV Mysteries	After crossing paths at a party, a Cape Town t...
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	September 24, 2021	2021	TV-MA	1 Season	Crime TV Shows, International TV Shows, TV Act...	To protect his family from a powerful drug lor...
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	Feuds, flirtations and toilet talk go down amo...
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	September 24, 2021	2021	TV-MA	2 Seasons	International TV Shows, Romantic TV Shows, TV ...	In a city of coaching centers known to train I...
5	s6	TV Show	Midnight Mass	Mike Flanagan	Kate Siegel, Zach Gilford, Hamish Linklater, H...	NaN	September 24, 2021	2021	TV-MA	1 Season	TV Dramas, TV Horror, TV Mysteries	The arrival of a charismatic young priest brin...
...	...	...	...	...	...	...	...	...	...	...	...	...
8795	s8796	TV Show	Yu-Gi-Oh! Arc-V	NaN	Mike Liscio, Emily Bauer, Billy Bob Thompson, ...	Japan, Canada	May 1, 2018	2015	TV-Y7	2 Seasons	Anime Series, Kids' TV	Now that he's discovered the Pendulum Summonin...
8796	s8797	TV Show	Yunus Emre	NaN	Gökhan Atalay, Payidar Tüfekçioglu, Baran Akbu...	Turkey	January 17, 2017	2016	TV-PG	2 Seasons	International TV Shows, TV Dramas	During the Mongol invasions, Yunus Emre leaves...
8797	s8798	TV Show	Zak Storm	NaN	Michael Johnston, Jessica Gee-George, Christin...	United States, France, South Korea, Indonesia	September 13, 2018	2016	TV-Y7	3 Seasons	Kids' TV	Teen surfer Zak Storm is mysteriously transpor...

Sanam

Next steps: [Generate code with tvshows]   [◯ View recommended plots]   [New interactive sheet]

```
data['type'].value_counts()["Movie"]
```

    np.int64(6131)

```
data['type'].value_counts()["TV Show"]
```

    np.int64(2676)

. There are singnificantly more movies as compared to tv shows.

```
data['rating'].unique()
```

    array(['PG-13', 'TV-MA', 'PG', 'TV-14', 'TV-PG', 'TV-Y', 'TV-Y7', 'R',
           'TV-G', 'G', 'NC-17', '74 min', '84 min', '66 min', 'NR', nan,
           'TV-Y7-FV', 'UR'], dtype=object)

```
data['rating'].value_counts()
```

rating	count
TV-MA	3207
TV-14	2160
TV-PG	863
R	799
PG-13	490
TV-Y7	334
TV-Y	307
PG	287
TV-G	220
NR	80
G	41
TV-Y7-FV	6
NC-17	3
UR	3
74 min	1
84 min	1
66 min	1

dtype: int64

- This shows that movies and tv shows with rating TV-MA are more liked by people.

```
movies['rating'].value_counts()['TV-MA']
```

np.int64(2062)

```
tvshows['rating'].value_counts()['TV-MA']
```

np.int64(1145)

- This shows that movie with rating 'TV-MA' are more liked and watched by people than tv shows with the same rating.

```
movies_per_year = movies.groupby('release_year').size()
movies_per_year.tail(20)
```

	0
release_year	
2002	44
2003	51
2004	55
2005	67
2006	82
2007	74
2008	113
2009	118
2010	154
2011	145
2012	173
2013	225
2014	264
2015	398
2016	658
2017	767
2018	767
2019	633
2020	517
2021	277

dtype: int64

- This shows that no. of movies released in the past 2 decades.
- The year 2017 and 2018 has the most no. of movie release.

```
tvshows_per_year = tvshows.groupby('release_year').size()
tvshows_per_year.tail(20)
```

	0
release_year	
2002	7
2003	10
2004	9
2005	13
2006	14
2007	14
2008	23
2009	34
2010	40
2011	40
2012	64
2013	63
2014	88
2015	162
2016	244
2017	265
2018	380
2019	397
2020	436
2021	315

dtype: int64

- This shows that the number of tv shows released over past 2 decades.
- The year 2018 and 2020 has most number of movie release.
- This also shows that netflix is more focused on movies more than tvshows.

```
movies['director'].value_counts()
```

	count
director	
Rajiv Chilaka	19
Raúl Campos, Jan Suter	18
Suhas Kadav	16
Marcus Raboy	15
Jay Karas	14
...	...
Naoto Amazutsumi	1
Jenny Gage	1
Kaila York	1
Vandana Kataria	1
Ed Perkins	1

4354 rows × 1 columns

dtype: int64

```
movies['cast'].value_counts()
```

	count
cast	
Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jigna Bhardwaj, Rajesh Kava, Mousam, Swapnil	13
Samuel West	10
Jeff Dunham	7
Craig Sechler	6
Kevin Hart	6
...	...
Casey Affleck, Rooney Mara, Liz Franke, Rob Zabrecky, Will Oldham, Sonia Acevedo	1
Faizon Love, Kirk Fox, Tony Todd, Pearl Thusi, Cassie Clare, Fiona Ramsey, Neels Clasen, Roxy Nel	1
Johnny Harris, Sam Keeley, Joe Dempsie, Kyle Soller, Nicholas Pinnock, Parker Sawyers, Philip Arditti, Sofia Boutella	1
Rajkummar Rao, Boman Irani, Mouni Roy, Sumeet Vyas, Paresh Rawal, Gajraj Rao, Manoj Joshi, Sanjay Goradia, Amyra Dastur, Chittaranjan Tripathi, Abhishek Banerjee	1
Frank Welker, Tim Conway, Gregg Berger, Wally Wingert, Audrey Wasilewski, Fred Tatasciore, Stephen Stanton, Jennifer Darling, Neil Ross, Greg Eagles	1

5445 rows × 1 columns

```
tvshows['director'].value_counts()
```

	count
director	
Alastair Fothergill	3
Rob Seidenglanz	2
Ken Burns	2
Stan Lathan	2
Shin Won-ho	2
...	...
Bumpy	1
Michael Samuels	1
Mark Tonderai	1
Ken Burns, Lynn Novick	1
Michael Cumming	1

222 rows × 1 columns

dtype: int64

```
tvshows['cast'].value_counts()
```

	count
cast	
David Attenborough	14
Michela Luci, Jamie Watson, Anna Claire Bartlam, Dante Zee, Eric Peterson	4
Dave Chappelle	3
Nicholas Day	2
Brennan Mejia, Camille Hyde, Yoshi Sudarso, Michael Taber, James Davies, Claire Blackwelder	2
...	...
Danielle Rose Russell, Aria Shahghasemi, Kaylee Bryant, Jenny Boyd, Quincy Fouse, Peyton Alex Smith, Matthew Davis, Archie L. Winston	1
Ken Jeong, Jim Rash, Donald Glover, Joel McHale, Gillian Jacobs, Danny Pudi, Alison Brie, Chevy Chase, Yvette Nicole Brown	1
Sarah Natochenny, Laurie Hymes, Jessica Paquet, Rosie Reyes, Alyson Leigh Rosenfeld, Marc Swint, Rodger Parsons	1
Toru Furuya, Hirotaka Suzuoki, Ryo Horikawa, Koichi Hashimoto, Hideyuki Hori, Keiko Han, Yuriko Yamamoto, Kohei Miyauchi, Mami Koyama	1
Lior Raz, Hisham Suliman, Shadi Mar'i, Laëtitia Eïdo, Tsahi Halevi, Yuval Segal, Neta Garty, Tomer Kapon, Itzik Cohen, Rona-Lee Shim'on, Boaz Konforty, Doron Ben-David	1

2283 rows × 1 columns

dtype: int64

```
movies['country'].value_counts()
```

	count
country	
United States	2058
India	893
United Kingdom	206
Canada	122
Spain	97
...	...
Germany, Italy	1
Taiwan, China, France, United States	1
Philippines, Singapore	1
India, Turkey	1
Canada, Germany	1

651 rows × 1 columns

dtype: int64

```
tvshows['country'].value_counts()
```

	count
country	
United States	760
United Kingdom	213
Japan	169
South Korea	158
India	79
...	...
Ireland, Canada, United States, United Kingdom	1
Germany, United States, Italy	1
United Kingdom, United States, Czech Republic	1
Japan, Canada	1
United States, France, South Korea, Indonesia	1

196 rows × 1 columns

dtype: int64

- Comparing the data , it clearly shows that the united state has most access to most movies followed by india.
- As for tvshow united state has most access to most movie follwed by united kingdom.

```
movies['date_added'].value_counts()
```

	count
date_added	
January 1, 2020	97
November 1, 2019	75
March 1, 2018	72
December 31, 2019	67
October 1, 2018	64
...	...
June 23, 2019	1
December 23, 2018	1
October 26, 2019	1
April 21, 2019	1
May 26, 2016	1

1533 rows × 1 columns

dtype: int64

```
tvshows['date_added'].value_counts()
```

	count
date_added	
July 6, 2021	31
January 1, 2016	23
August 1, 2017	22
June 19, 2021	21
July 1, 2017	19
...	...
September 5, 2021	1
September 21, 2021	1
September 9, 2021	1
March 31, 2017	1
April 1, 2014	1

1051 rows × 1 columns

dtype: int64

```
movies['duration'].value_counts()
```

	count
duration	
90 min	152
97 min	146
94 min	146
93 min	146
91 min	144
...	...
228 min	1
18 min	1
205 min	1
201 min	1
191 min	1

205 rows × 1 columns

dtype: int64

```
movies['listed_in'].value_counts()
```

| | count |
listed_in	
Dramas, International Movies	362
Documentaries	359
Stand-Up Comedy	334
Comedies, Dramas, International Movies	274
Dramas, Independent Movies, International Movies	252
...	...
Action & Adventure, Classic Movies, International Movies	1
Comedies, Cult Movies, LGBTQ Movies	1
Action & Adventure, Comedies, Horror Movies	1
Action & Adventure, Documentaries, Sports Movies	1
Cult Movies, Dramas, Thrillers	1

278 rows × 1 columns

dtype: int64

```
plt.figure(figsize=(10,6))
sns.histplot(data['rating'])
plt.title('Distributon of ratings')
plt.xticks(rotation=90)
plt.show()
```


Distributon of ratings
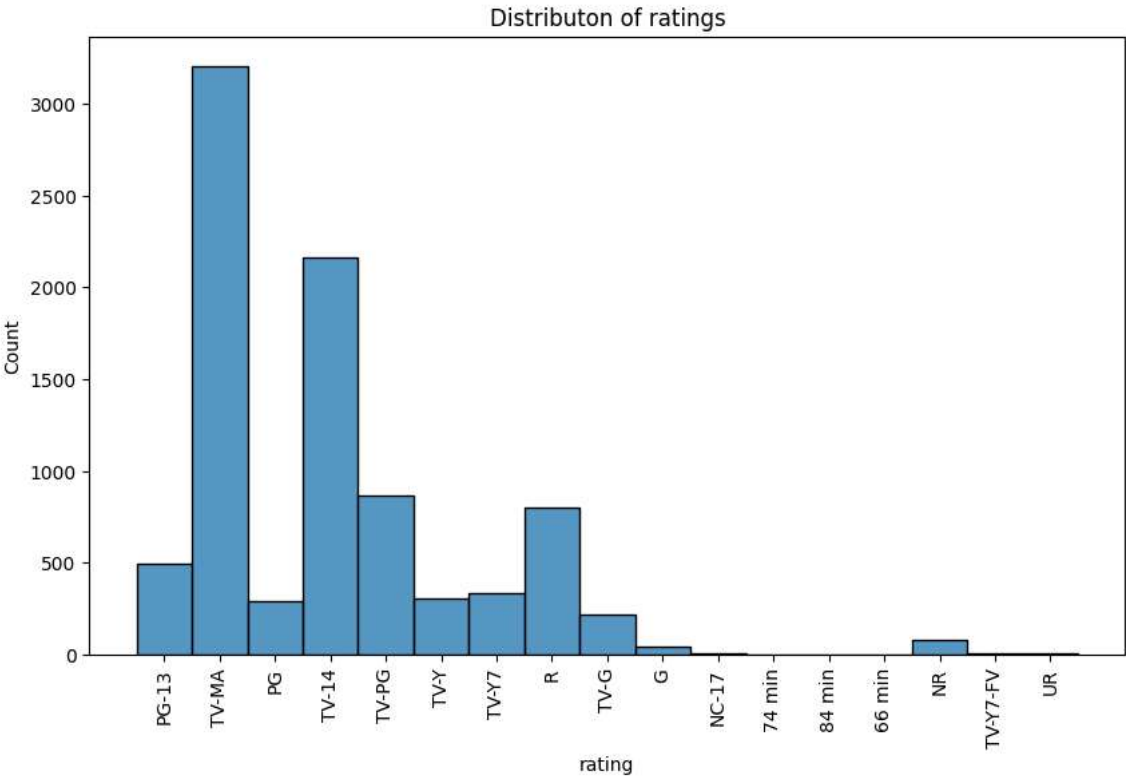
```
missing_values = data.isnull().sum()
missing_values
```

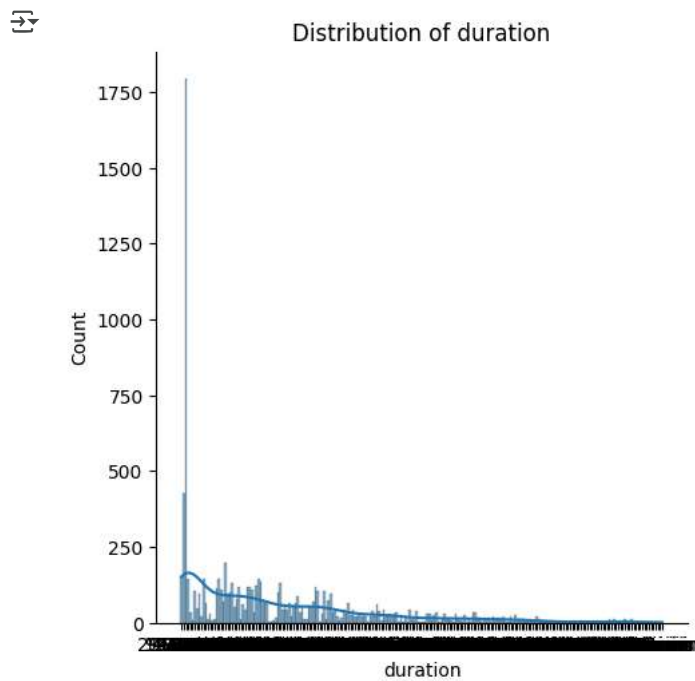|	0
show_id	0
type	0
title	0
director	2634
cast	825
country	831
date_added	10
release_year	0

```
continuous_vars = ['duration', 'rating', 'release_year']
```
duration   0

```
for var in continuous_vars:
# Distplot
 sns.displot(data[var], kde=True)
 plt.title(f"Distribution of {var}")
 plt.show()
```

Distribution of duration

Distribution of rating