

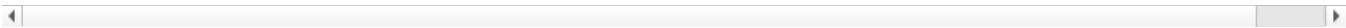
```
In [85]: import pandas as pd
```

```
In [87]: df=pd.read_csv("netflix_titles.csv")
df
```

Out[87]:

	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	A father's portrait of his son
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	cro path pe Cape
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 p his l fi pov drug
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	September 24, 2021	2021	TV-MA	1 Season	Docuseries, Reality TV	F flirts and te
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 c coa ce kno tra
...
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 pc carto a req ar
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While alon st to your
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	November 1, 2019	2009	R	88 min	Comedies, Horror Movies	Look surviv world ov
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	Dre from c fr superb
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 sc bu boy v his wa

8807 rows × 12 columns



```
In [123]: df.head()
```

Out[123...

	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 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...

In [89]: df.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
```

In [125... df.tail()

Out[125]

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	descripti
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 politi cartoon a cri repor and :
8803	s8804	tv show	Zombie Dumb	NaN	NaN	NaN	July 1, 2019	2018	TV-Y7	2 Seasons	Kids' TV, Korean TV Shows, TV Comedies	While livi alone in spoc town young t
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 survive in world tak over zt
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	Dragg from civili life forn 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 scrap but pc boy won his way in a t

In [91]:

df["country"].value_counts(normalize=True)*100

Out[91]:

country	
United States	35.330993
India	12.186560
United Kingdom	5.253260
Japan	3.071715
South Korea	2.494985
...	
Romania, Bulgaria, Hungary	0.012538
Uruguay, Guatemala	0.012538
France, Senegal, Belgium	0.012538
Mexico, United States, Spain, Colombia	0.012538
United Arab Emirates, Jordan	0.012538
Name: proportion, Length: 748, dtype: float64	

In [93]:

df["title"].value_counts(normalize=True)*100

Out[93]:

title	
Dick Johnson Is Dead	0.011355
Ip Man 2	0.011355
Hannibal Buress: Comedy Camisado	0.011355
Turbo FAST	0.011355
Masha's Tales	0.011355
...	
Love for Sale 2	0.011355
ROAD TO ROMA	0.011355
Good Time	0.011355
Captain Underpants Epic Choice-o-Rama	0.011355
Zubaan	0.011355
Name: proportion, Length: 8807, dtype: float64	

In [95]:

df["type"].value_counts(normalize=True)*100

Out[95]:

type	
Movie	69.615079
TV Show	30.384921
Name: proportion, dtype: float64	

In [97]:

df.isnull()

Out[97]:

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	False	False	False	False	True	False	False	False	False	False	False	False
1	False	False	False	True	False	False	False	False	False	False	False	False
2	False	False	False	False	False	True	False	False	False	False	False	False
3	False	False	False	True	True	True	False	False	False	False	False	False
4	False	False	False	True	False	False	False	False	False	False	False	False
...
8802	False	False	False	False	False	False	False	False	False	False	False	False
8803	False	False	False	True	True	True	False	False	False	False	False	False
8804	False	False	False	False	False	False	False	False	False	False	False	False
8805	False	False	False	False	False	False	False	False	False	False	False	False
8806	False	False	False	False	False	False	False	False	False	False	False	False

8807 rows × 12 columns

In [99]: df.isnull().sum()

Out[99]: show_id 0
type 0
title 0
director 2634
cast 825
country 831
date_added 10
release_year 0
rating 4
duration 3
listed_in 0
description 0
dtype: int64

In [101... df.drop_duplicates()

In [103...

```
df.isnull()
```

8807 rows x 12 columns

Out[103..

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	description
0	False	False	False	False	True	False	False	False	False	False	False	False
1	False	False	False	True	False	False	False	False	False	False	False	False
2	False	False	False	False	False	True	False	False	False	False	False	False
3	False	False	False	True	True	True	False	False	False	False	False	False
4	False	False	False	True	False	False	False	False	False	False	False	False
...
8802	False	False	False	False	False	False	False	False	False	False	False	False
8803	False	False	False	True	True	True	False	False	False	False	False	False
8804	False	False	False	False	False	False	False	False	False	False	False	False
8805	False	False	False	False	False	False	False	False	False	False	False	False
8806	False	False	False	False	False	False	False	False	False	False	False	False

8807 rows × 12 columns

In [133..

```
df['type'].str.lower()
```

Out[133..

```
0      movie
1    tv show
2    tv show
3    tv show
4    tv show
...
8802   movie
8803  tv show
8804   movie
8805   movie
8806   movie
Name: type, Length: 8807, dtype: object
```

In [135..

```
df['show_id'].str.upper()
```

Out[135..

```
0      S1
1      S2
2      S3
3      S4
4      S5
...
8802  S8803
8803  S8804
8804  S8805
8805  S8806
8806  S8807
Name: show_id, Length: 8807, dtype: object
```

In []: