

# Netflix Data Exploration Business Case

# Recalling Python libraries & dataset as 'df' after 'CSV' file uploaded into 'google colab'.

```
[4] # Working on Netflix Dataset
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df=pd.read_csv("/content/netflix1.csv")
df
```

	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 Mablane, 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, Nabl...	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 l...

## #Data Cleaning

# Missing Value Imputation

('0' for numeric columns & 'Unknown column name' for objective columns)

```
df = df.apply(lambda col: col.fillna(0) if col.dtype == 'number' else col.fillna("Unknown"+' '+col.name))
df
```

	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	Unknown cast	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	Unknown director	Ama Qamata, Khosi Ngema, Gail Mablane, 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, Nabl...	Unknown country	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	Unknown director	Unknown cast	Unknown country	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	Unknown director	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 l...
...	...	...	...	...	...	...	...	...	...	...	...	...
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...
8803	s8804	TV	Zombie Dumb	Unknown	Unknown cast	Unknown	July 1, 2019	2018	TV-Y7	2	Kids' TV, Korean TV Shows, ...	While living alone in a ...

# Given data has 8807 row & 12 columns

```
[61] df.shape
```

(8807, 12)

# All Null values has been replaced by corresponding values as asked.

# Now no null value observed in dataset



✓  
0s

```
[60] df.isnull().sum()
```



0

show\_id 0

type 0

title 0

director 0

cast 0

country 0

date\_added 0

release\_year 0

rating 0

duration 0

listed\_in 0

description 0

dtype: int64



### # Working on Nested values ('director', 'cast', 'country', 'listed\_in')

```
# Selected data subset ('title' & 'director') as 'df1'
```

```
df1=df[['title','director']]
df1
```

	title	director
0	Dick Johnson Is Dead	Kirsten Johnson
1	Blood & Water	Unknown director
2	Ganglands	Julien Leclercq
3	Jailbirds New Orleans	Unknown director
4	Kota Factory	Unknown director
...	...	...
8802	Zodiac	David Fincher
8803	Zombie Dumb	Unknown director
8804	Zombieland	Ruben Fleischer
8805	Zoom	Peter Hewitt
8806	Zubaan	Mozez Singh

8807 rows × 2 columns

# 'director' column splitted based on ','

```
✓ [8] df1['director'] = df1['director'].apply(lambda x: x.split(','))  
0s df1
```

⇨ <ipython-input-8-a249e374eb22>:1: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation: [https://pandas.pydata.org/pandas-docs/stable/10min/5min\\_modifications.html](https://pandas.pydata.org/pandas-docs/stable/10min/5min_modifications.html)

```
df1['director'] = df1['director'].apply(lambda x: x.split(','))
```

	title	director
0	Dick Johnson Is Dead	[Kirsten Johnson]
1	Blood & Water	[Unknown director]
2	Ganglands	[Julien Leclercq]
3	Jailbirds New Orleans	[Unknown director]
4	Kota Factory	[Unknown director]
...	...	...
8802	Zodiac	[David Fincher]
8803	Zombie Dumb	[Unknown director]
8804	Zombieland	[Ruben Fleischer]
8805	Zoom	[Peter Hewitt]
8806	Zubaan	[Mozez Singh]

8807 rows × 2 columns

# Splitting 'director' column exploded into rows

✓  
0s

🔑

📁

```
[9] df1_exploded=df1.explode('director').reset_index(drop=True)
df1_exploded
```

↔

	title	director
0	Dick Johnson Is Dead	Kirsten Johnson
1	Blood & Water	Unknown director
2	Ganglands	Julien Leclercq
3	Jailbirds New Orleans	Unknown director
4	Kota Factory	Unknown director
...	...	...
9607	Zodiac	David Fincher
9608	Zombie Dumb	Unknown director
9609	Zombieland	Ruben Fleischer
9610	Zoom	Peter Hewitt
9611	Zubaan	Mozez Singh

9612 rows × 2 columns

📊

📈

✎

# Similarly, df2\_exploded, df3\_exploded & df4\_exploded generated to explode 'cast', 'country' & 'listed\_in' columns respectively.

# Four data subsets generated

- df1\_exploded
- df2\_exploded
- df3\_exploded
- df4\_exploded

# All above four data merged to give 'df\_merged'

# All four data subsets are merged on 'title' column creating 20265 rows.

0s

[28] df\_merged=df1\_exploded.merge(df2\_exploded, on='title').merge(df3\_exploded,on='title').merge(df4\_exploded,on='title').reset\_index(drop=True)  
df\_merged

	title	director	cast	country	listed_in
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows
...	...	...	...	...	...
202060	Zubaan	Mozez Singh	Anita Shabdish	India	International Movies
202061	Zubaan	Mozez Singh	Anita Shabdish	India	Music & Musicals
202062	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	Dramas
202063	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	International Movies
202064	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	Music & Musicals

202065 rows × 5 columns

# Now other data subset created of remaining columns including 'title' as 'df\_pending'

0s

[30] df\_pending=df[['type','title','date\_added','release\_year', 'rating', 'duration','description']]  
df\_pending

	type	title	date_added	release_year	rating	duration	description
0	Movie	Dick Johnson Is Dead	September 25, 2021	2020	PG-13	90 min	As her father nears the end of his life, filmm...
1	TV Show	Blood & Water	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t...
2	TV Show	Ganglands	September 24, 2021	2021	TV-MA	1 Season	To protect his family from a powerful drug lor...
3	TV Show	Jailbirds New Orleans	September 24, 2021	2021	TV-MA	1 Season	Feuds, flirtations and toilet talk go down amo...
4	TV Show	Kota Factory	September 24, 2021	2021	TV-MA	2 Seasons	In a city of coaching centers known to train l...
...	...	...	...	...	...	...	...
8802	Movie	Zodiac	November 20, 2019	2007	R	158 min	A political cartoonist, a crime reporter and a...
8803	TV Show	Zombie Dumb	July 1, 2019	2018	TV-Y7	2 Seasons	While living alone in a spooky town, a young g...
8804	Movie	Zombieland	November 1, 2019	2009	R	88 min	Looking to survive in a world taken over by zo...
8805	Movie	Zoom	January 11, 2020	2006	PG	88 min	Dragged from civilian life, a former superhero...
8806	Movie	Zubaan	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty...

8807 rows × 7 columns

# 'df\_merged' & 'df\_pending' are merged on 'title' column giving 'df\_pre' data frame.

```
[31] df_pre=df_merged.merge(df_pending, on='title')
df_pre
```

	title	director	cast	country	listed_in	type	date_added	release_year	rating	duration	description
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	September 25, 2021	2020	PG-13	90 min	As her father nears the end of his life filmm..
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t..
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t..
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t..
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t..
...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozes Singh	Anita Shabdish	India	International Movies	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty..
202061	Zubaan	Mozes Singh	Anita Shabdish	India	Music & Musicals	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty..
202062	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Dramas	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty..
202063	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	International Movies	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty..
202064	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Music & Musicals	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty..

202065 rows x 11 columns

# New column is created 'extracted\_duration' to remove string after numeric value.

```
[32] df_pre['extracted_duration'] = df_pre['duration'].str.extract('(\d+)')
df_pre
```

	title	director	cast	country	listed_in	type	date_added	release_year	rating	duration	description	extracted_duration
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	September 25, 2021	2020	PG-13	90 min	As her father nears the end of his life, filmm...	90
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t...	2
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t...	2
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t...	2
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	September 24, 2021	2021	TV-MA	2 Seasons	After crossing paths at a party, a Cape Town t...	2
...	...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozes Singh	Anita Shabdish	India	International Movies	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty...	111
202061	Zubaan	Mozes Singh	Anita Shabdish	India	Music & Musicals	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty...	111
202062	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Dramas	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty...	111
202063	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	International Movies	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms his way into a ty...	111
202064	Zubaan	Mozes Singh	Chittaranjan	India	Music & Musicals	Movie	March 2, 2019	2015	TV-14	111 min	A scrappy but poor boy worms	111

# Now 'duration' column is deleted.

```
[33] df_pre.drop('duration',axis=1,inplace=True)
df_pre
```

# new column is renamed as 'duration'

```
[34] df_pre.rename(columns={'extracted_duration':'duration'},inplace=True)
df_pre
```

	title	director	cast	country	listed_in	type	date_added	release_year	rating	description	duration
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	September 25, 2021	2020	PG-13	As her father nears the end of his life, filmm...	90
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	September 24, 2021	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	September 24, 2021	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	September 24, 2021	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	September 24, 2021	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2
...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozez Singh	Anita Shabdish	India	International Movies	Movie	March 2, 2019	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111
202061	Zubaan	Mozez Singh	Anita Shabdish	India	Music & Musicals	Movie	March 2, 2019	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111
202062	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	Dramas	Movie	March 2, 2019	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111
202063	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	International Movies	Movie	March 2, 2019	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111
202064	Zubaan	Mozez Singh	Chittaranjan Tripathy	India	Music & Musicals	Movie	March 2, 2019	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111

# New column 'datetime' is created in proper format so that can extract week no.

```
[40] df_pre['datetime'] = pd.to_datetime(df_pre['date_added'], format='%B %d, %Y', errors='coerce')
df_pre
```

# New column 'week\_no' is created.

```
[42] df_pre['week_no']=df_pre['datetime'].dt.isocalendar().week
df_pre
```

# Old column 'date\_added' is deleted.

```
[43] df_pre.drop('date_added',axis=1,inplace=True)
df_pre
```



# Column 'datetime' is renamed as 'date\_added'

[49] df\_pre.rename(columns={'datetime':'date\_added'},inplace=True)  
df\_pre

	title	director	cast	country	listed_in	type	release_year	rating	description	duration	date_added	week_no
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	2020	PG-13	As her father nears the end of his life, filmm...	90	2021-09-25	38
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38
...	...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozes Singh	Anita Shabdish	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9
202061	Zubaan	Mozes Singh	Anita Shabdish	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9
202062	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Dramas	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9
202063	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9
202064	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9

202065 rows x 12 columns

# New column is created as 'month', including 'full month name'.

[98] df\_pre['month']=df\_pre['date\_added'].dt.strftime('%B')  
df\_pre

	title	director	cast	country	listed_in	type	release_year	rating	description	duration	date_added	week_no	month
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	2020	PG-13	As her father nears the end of his life, filmm...	90	2021-09-25	38	September
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September
...	...	...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozes Singh	Anita Shabdish	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9	March
202061	Zubaan	Mozes Singh	Anita Shabdish	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9	March
202062	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Dramas	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9	March
202063	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9	March
202064	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02	9	March

202065 rows x 13 columns

# 'df\_pre' is copied in 'df\_final'

[99] df\_final=df\_pre.copy()  
df\_final

# Again, some null values are observed as new columns has been created.

{x}

0s

✓

101

df\_final.isnull().sum()

0

0

0

0

0

0

0

0

0

0

3

1746

1746

1746

dtype: int64

# All null values are again imputed using 'lambda' function.

✓

101

df\_final=df\_final.apply(lambda col:col.fillna(0) if col.dtype =='UInt32' else col.fillna('Unknown'+ ' '+ col.name))

df\_final

	title	director	cast	country	listed_in	type	release_year	rating	description	duration	date_added	week_no	month
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	2020	PG-13	As her father nears the end of his life, filmm...	90	2021-09-25 00:00:00	38	September
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
...	...	...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozes Singh	Anita Shabdish	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202061	Zubaan	Mozes Singh	Anita Shabdish	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202062	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Dramas	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202063	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202064	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March

202065 rows × 13 columns

# Now all null values has-been completely imputed.

🔍 [102] df\_final.isnull().sum()

 $\{x\}$ 

title	0
-------	---

director 0

cast 0

country 0

```
listed_in    0
```

type 0

release year 0

rating 0

description 0

duration 0

```
date added 0
```

week no 0

month 0

dtype: int64

## # Now data is completely cleaned & restructured.

```
[104] df_final
```



	title	director	cast	country	listed_in	type	release_year	rating	description	duration	date_added	week_no	month
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	2020	PG-13	As her father nears the end of his life, filmm...	90	2021-09-25 00:00:00	38	September
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September
...	...	...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozes Singh	Anita Shabdish	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202061	Zubaan	Mozes Singh	Anita Shabdish	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202062	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Dramas	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202063	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March
202064	Zubaan	Mozes Singh	Chittaranjan Tripathy	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March

202065 rows x 13 columns

1. Find the counts of each categorical variable both using graphical and nongraphical analysis.

a. For Non-graphical Analysis:

0s

✓

[23]

df\_final.title.value\_counts()

{x}

↔

count

title

Kahlil Gibran's The Prophet

700

Holidays

504

Movie 43

468

The Eddy

416

Narcos

378

...

...

Blackfish

1

The 2000s

1

Miniforce: Super Dino Power

1

Dancing with the Birds

1

Dick Johnson Is Dead

1

8807 rows × 1 columns

dtype: int64



✓  
0s

```
[24] df_final.director.value_counts()
```



count

director

Unknown director	50643
Martin Scorsese	419
Youssef Chahine	409
Cathy Garcia-Molina	356
Steven Spielberg	355
...	...
Reuben Atlas	1
Max Amini	1
Bassam Tariq	1
Paul Miller	1
Kirsten Johnson	1

5121 rows × 1 columns

dtype: int64



{x}

✓  
0s

```
[25] df_final.cast.value_counts()
```

🔑



count

📁

cast

Unknown cast	2149
--------------	------

Alfred Molina	160
---------------	-----

Salma Hayek	130
-------------	-----

Frank Langella	128
----------------	-----

John Rhys-Davies	125
------------------	-----

...	...
-----	-----

Cle Sloan	1
-----------	---

George Konner	1
---------------	---

Andy Hayward	1
--------------	---

Lyric Lewis	1
-------------	---

Trent 'Maxi' Maxwell	1
----------------------	---

39297 rows × 1 columns

dtype: int64

<>

☰



✓  
0s

```
[26] df_final.country.value_counts()
```



count

country

United States	49868
---------------	-------

India	22139
-------	-------

Unknown country	11897
-----------------	-------

United Kingdom	9733
----------------	------

United States	9482
---------------	------

...	...
-----	-----

Samoa	2
-------	---

Sri Lanka	2
-----------	---

Kazakhstan	1
------------	---

Uganda	1
--------	---

Nicaragua	1
-----------	---

197 rows × 1 columns

dtype: int64



{x}

✓  
0s

```
[27] df_final.listed_in.value_counts()
```

🔑



📁

	count
listed_in	
International Movies	27141
Dramas	19657
Comedies	13894
Action & Adventure	12216
Dramas	10149
...	...
Stand-Up Comedy	24
Romantic Movies	20
TV Sci-Fi & Fantasy	7
LGBTQ Movies	5
Sports Movies	3

73 rows × 1 columns

dtype: int64

<>

☰

✓  
0s

```
[28] df_final.type.value_counts()
```



	count
type	
Movie	145917
TV Show	56148

dtype: int64

<>

☰



✓

0s

[29] df\_final.rating.value\_counts()

↔

	count
TV-MA	73915
TV-14	43957
R	25860
PG-13	16246
TV-PG	14926
PG	10919
TV-Y7	6304
TV-Y	3665
TV-G	2779
NR	1573
G	1530
NC-17	149
TV-Y7-FV	86
UR	86
Unknown rating	67
74 min	1
84 min	1
66 min	1

{X}

✓

0s

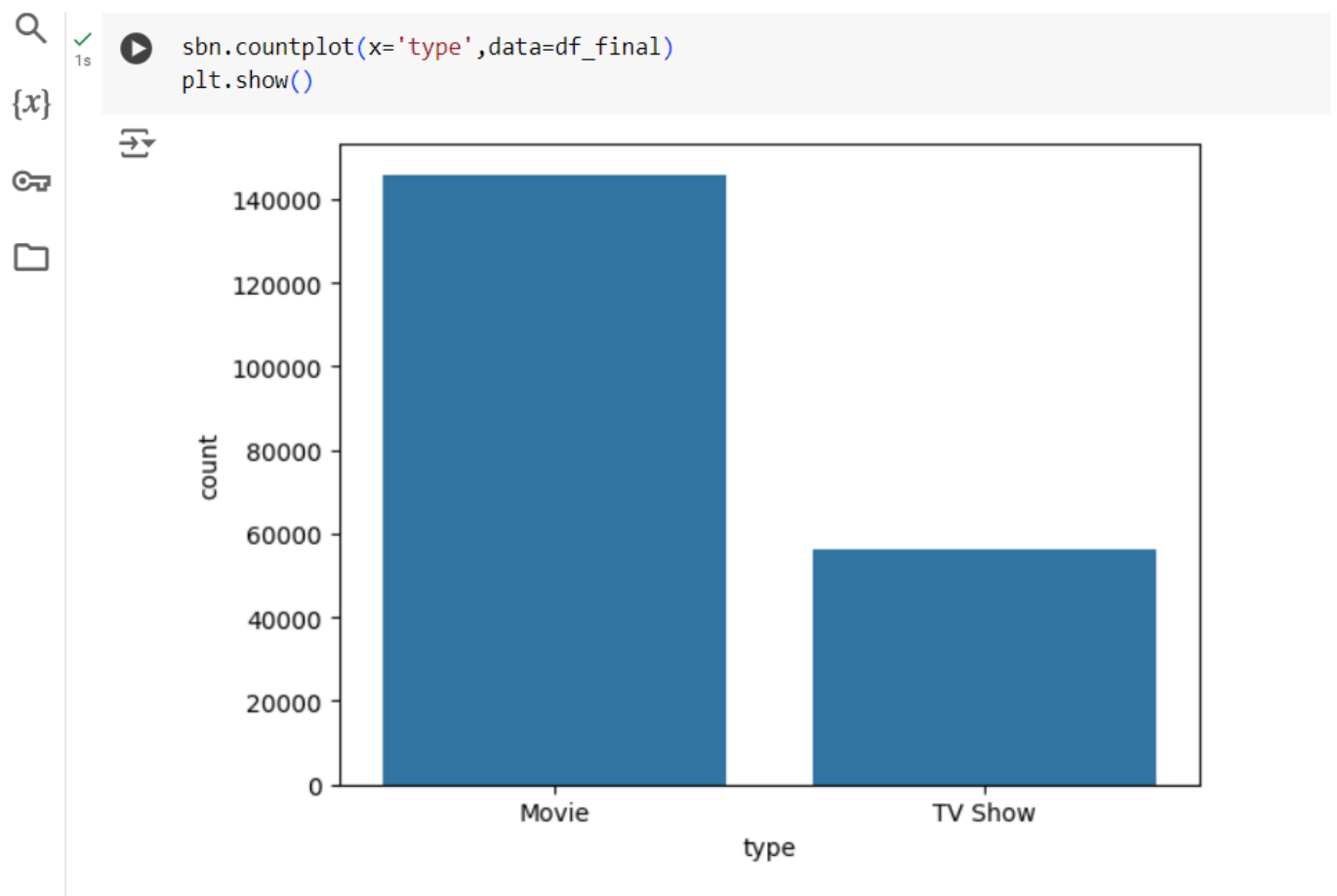
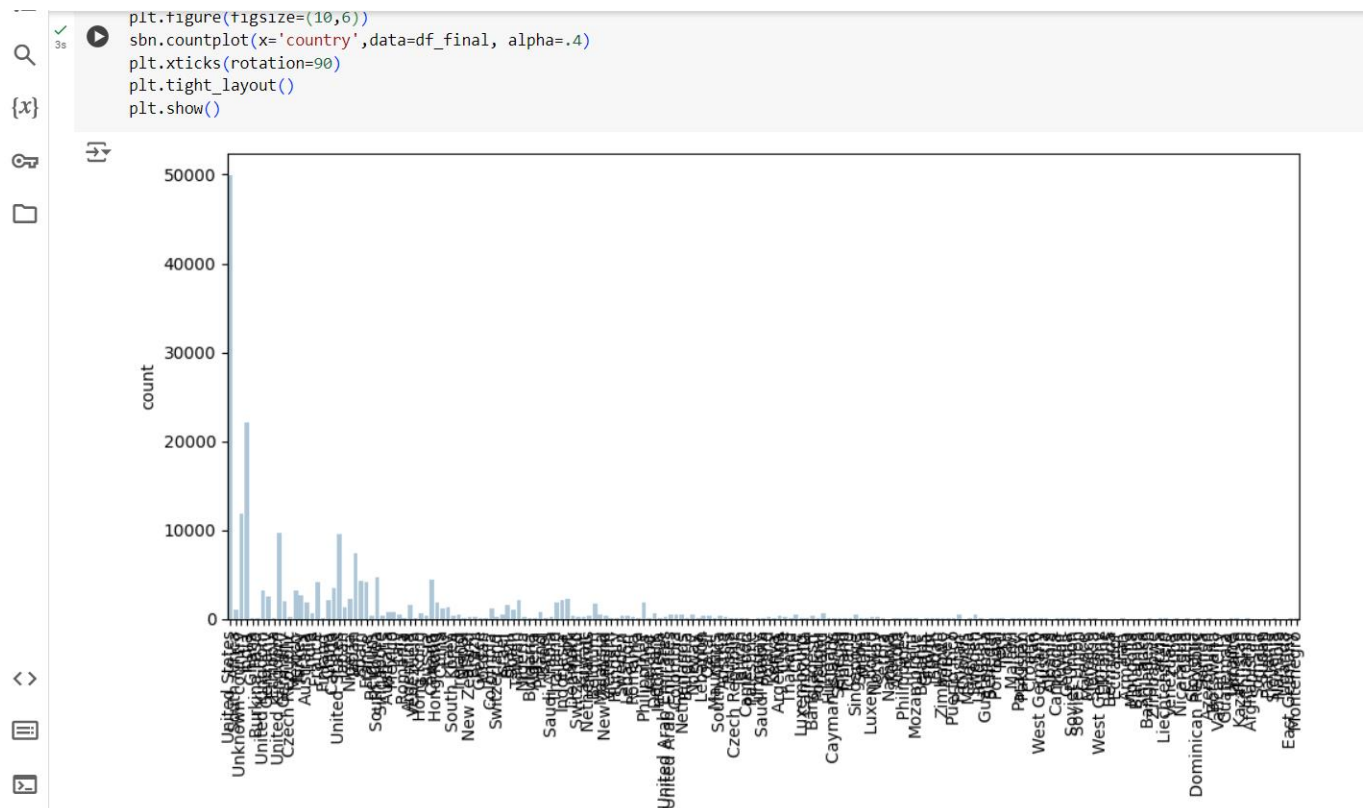
[30] df\_final.description.value\_counts()

↔

	count
A troubled young girl and her mother find solace on a journey with a subversive poet whose words captivate their hearts and imaginations.	700
Eight gleefully grim short tales set on Halloween, Mother's Day, St. Patrick's Day and other holidays comprise this chilling horror anthology.	504
An eye-popping cast stars in this sketch-comedy collage, the new millennium's homage to classic anthology films like The Kentucky Fried Movie.	468
The owner of a Paris jazz club gets tangled up with dangerous criminals as he fights to protect his business, his band and his teenage daughter.	416
The true story of Colombia's infamously violent and powerful drug cartels fuels this gritty gangster drama series.	378
...	...
Chris D'Elia takes the stage in Minneapolis to offer his thoughts on everything from self-censorship to problematic dolphins to lame mutant powers.	1
From public protests to viral movements, the aftermath of the R. Kelly docuseries is explored in this special featuring journalists and experts.	1
Comedic breakout Tiffany Haddish delivers a riotous stand-up ripe with the unpretentious and filthy tales of her meteoric rise to stardom.	1
This documentary explores the challenging, transformative world of high-intensity fitness programs through the eyes of the people who've embraced them.	1
As her father nears the end of his life, filmmaker Kirsten Johnson stages his death in inventive and comical ways to help them both face the inevitable.	1

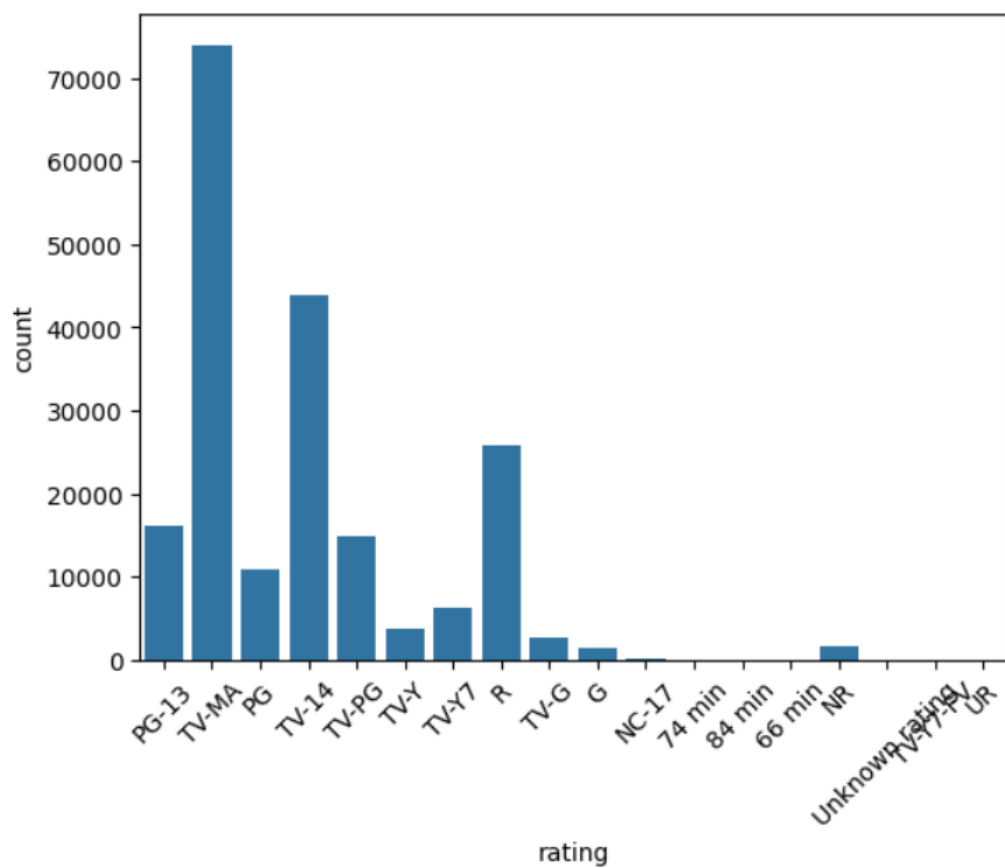
8775 rows × 1 columns

**b. For graphical analysis:**



✓  
1s

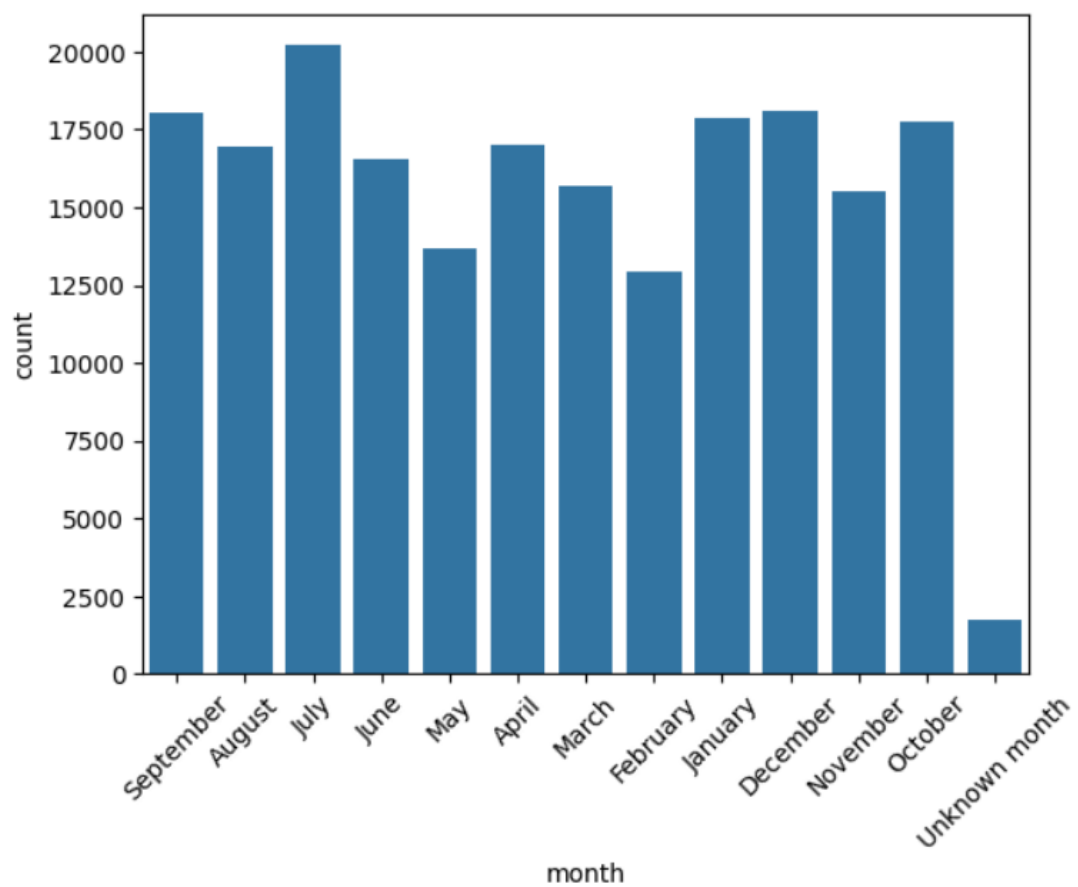
```
sbn.countplot(x='rating',data=df_final)  
plt.xticks(rotation=45)  
plt.show()
```



{x}

✓  
1s

```
[61] sbn.countplot(x='month',data=df_final)  
plt.xticks(rotation=45)  
plt.show()
```



## 2. Comparison of tv shows vs. movies.

a. Find the number of movies produced in each country and pick the top 10 countries.

Answer: Top 10 countries producing movie are given below.

```
[86] df_final[df_final.type=='Movie'].groupby('country')['title'].nunique().sort_values(ascending=False).head(10)
```

country	title
United States	2364
India	927
Unknown country	440
United States	388
United Kingdom	382
Canada	187
France	155
United Kingdom	152
France	148
Canada	132

dtype: int64

b. Find the number of Tv-Shows produced in each country and pick the top 10 countries.

Answer: Top 10 countries producing TV show are given below.

```
[87] df_final[df_final.type=='TV Show'].groupby('country')['title'].nunique().sort_values(ascending=False).head(10)
```

country	title
United States	847
Unknown country	391
United Kingdom	246
Japan	174
South Korea	164
United States	91
Canada	84
India	81
Taiwan	70
France	64

dtype: int64

### 3. What is the best time to launch a TV show?

Answer: Best time to launch a TV show is week-27 as analysis given below.

#### a. Find which is the best week to release the Tv-show or the movie. Do the analysis separately for Tv-shows and Movies

Answer:

1. Best week to release TV show is week-27.
  2. Best week to release movie is week-1.
- (Both analyses are given below)

```
[91] df_final[df_final.type=='TV Show'].groupby('week_no')['title'].nunique().sort_values(ascending=False).head(5)
```



title	
week_no	
0	98
27	85
31	79
24	75
35	73

dtype: int64

```
[89] df_final[df_final.type=='Movie'].groupby('week_no')['title'].nunique().sort_values(ascending=False).head(5)
```



title	
week_no	
1	316
44	243
40	215
9	207
26	195

dtype: int64

b. Find which is the best month to release the Tv-show or the movie. Do the analysis separately for Tv-shows and Movies

Answer:

- 1. Best month to release TV show is July.
  - 2. Best month to release movie is also July
- (Both analyses are given below)

✓

0s

[105]

```
df_final[df_final.type=='TV Show'].groupby('month')['title'].nunique().sort_values(ascending=False)
```

	title
month	
July	254
December	250
September	246
June	232
August	230
October	210
April	209
March	205
November	199
May	187
January	181
February	175
Unknown month	98

dtype: int64

```
✓ [107] df_final[df_final.type=='Movie'].groupby('month')['title'].nunique().sort_values(ascending=False)
0s
```



title	
month	
July	565
April	550
December	547
January	546
October	545
March	529
August	519
September	519
November	498
June	492
May	439
February	382

dtype: int64



#### 4. Analysis of actors/directors of different types of shows/movies.

a. Identify the top 10 actors who have appeared in most movies or TV shows.

Answer: Top 10 actors are given below.

```
✓ [108] df_final.groupby('cast')['title'].nunique().sort_values(ascending=False).head(10)
```



title	
cast	
Unknown cast	825
Anupam Kher	39
Rupa Bhimani	31
Takahiro Sakurai	30
Julie Teiwani	28
Om Puri	27
Rajesh Kava	26
Shah Rukh Khan	26
Boman Irani	25
Paresh Rawal	25

dtype: int64

b. Identify the top 10 directors who have appeared in most movies or TV shows.

Answer: Top 10 directors are given below.

```
✓ [109] df_final.groupby('director')['title'].nunique().sort_values(ascending=False).head(10)  
0s
```



title	
director	
Unknown director	2634
Rajiv Chilaka	22
Jan Suter	18
Raúl Campos	18
Marcus Raboy	16
Suhas Kadav	16
Jay Karas	15
Cathy Garcia-Molina	13
Jay Chapman	12
Martin Scorsese	12

dtype: int64

### 5. Which genre movies are more popular or produced more.

Answer: International Movies are more popular.

```
[110] df_final[df_final.type=='Movie'].groupby('listed_in')['title'].nunique().sort_values(ascending=False).head()
```

	title
listed_in	
International Movies	2624
Dramas	1600
Comedies	1210
Action & Adventure	859
Documentaries	829

dtype: int64

```
[117] wordcloud = WordCloud(width=800, height=400, background_color='white').generate(text)
plt.figure(figsize=(12,10))
plt.imshow(wordcloud, interpolation='bilinear')
plt.axis('off')
plt.title('Word Cloud of Movie Genres')
plt.show()
```



6. Find After how many days the movie will be added to Netflix after the release of the movie (you can consider the recent past data)

Creating new column from 'release\_year' to 'StartOfYear'.

1s

✓

[4] df\_final['StartOfYear'] = pd.to\_datetime(df\_final['release\_year'].astype(str) + '-01-01')

0s

✓

[5] df\_final

	title	director	cast	country	listed_in	type	release_year	rating	description	duration	date_added	week_no	month	StartOfYear
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	2020	PG-13	As her father nears the end of his life, filmm...	90	2021-09-25 00:00:00	38	September	2020-01-01
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September	2021-01-01
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September	2021-01-01
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September	2021-01-01
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24 00:00:00	38	September	2021-01-01
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
202060	Zubaan	Mozez Singh	Anita Shabdish	India	International Movies	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March	2015-01-01
202061	Zubaan	Mozez Singh	Anita Shabdish	India	Music & Musicals	Movie	2015	TV-14	A scrappy but poor boy worms his way into a ty...	111	2019-03-02 00:00:00	9	March	2015-01-01

0s completed at 7:27 PM

df\_final['date\_added'] = pd.to\_datetime(df\_final['date\_added'], errors='coerce')

df\_final['DifferenceInDays'] = (df\_final['date\_added'] - df\_final['StartOfYear']).dt.days

df\_final

	title	director	cast	country	listed_in	type	release_year	rating	description	duration	date_added	week_no	month	StartOfYear	DifferenceInDays
0	Dick Johnson Is Dead	Kirsten Johnson	Unknown cast	United States	Documentaries	Movie	2020	PG-13	As her father nears the end of his life, filmm...	90	2021-09-25	38	September	2020-01-01	633.0
1	Blood & Water	Unknown director	Ama Qamata	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September	2021-01-01	266.0
2	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Dramas	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September	2021-01-01	266.0
3	Blood & Water	Unknown director	Ama Qamata	South Africa	TV Mysteries	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September	2021-01-01	266.0
4	Blood & Water	Unknown director	Khosi Ngema	South Africa	International TV Shows	TV Show	2021	TV-MA	After crossing paths at a party, a Cape Town t...	2	2021-09-24	38	September	2021-01-01	266.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
		Mozez	Anita		International				A scrappy but poor boy						

Answer: Lots of movies were added to Netflix platform at the average interval of 1000 days as below plot.

```
[82] plt.hist(df_final['DifferenceInDays'], bins=30, color='skyblue', edgecolor='black')
plt.title('Distribution of Days to Add Movies to Netflix')
plt.xlabel('Days')
plt.ylabel('Count of Movies')
plt.xticks([1000,2000,3000,4000,5000,6000,7000,8000,9000,10000,11000,12000])
plt.xticks(rotation=90)
plt.show()
```

