

```
import pandas as pd
import numpy as np
import seaborn as sns
from matplotlib import pyplot as plt
df=pd.read_csv(r"C:\Users\dell\Downloads\netflix_titles.csv")
df
```

	show_id	type	title	director	\
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	
1	s2	TV Show	Blood & Water	NaN	
2	s3	TV Show	Ganglands	Julien Leclercq	
3	s4	TV Show	Jailbirds New Orleans	NaN	
4	s5	TV Show	Kota Factory	NaN	
...	
8802	s8803	Movie	Zodiac	David Fincher	
8803	s8804	TV Show	Zombie Dumb	NaN	
8804	s8805	Movie	Zombieland	Ruben Fleischer	
8805	s8806	Movie	Zoom	Peter Hewitt	
8806	s8807	Movie	Zubaan	Mozez Singh	

		cast	country
\			
0		NaN	United States
1	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...		South Africa
2	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...		NaN
3		NaN	NaN
4	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...		India
...	
8802	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...		United States
8803		NaN	NaN
8804	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...		United States
8805	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...		United States
8806	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...		India

	date_added	release_year	rating	duration	\
0	September 25, 2021	2020	PG-13	90 min	
1	September 24, 2021	2021	TV-MA	2 Seasons	
2	September 24, 2021	2021	TV-MA	1 Season	
3	September 24, 2021	2021	TV-MA	1 Season	
4	September 24, 2021	2021	TV-MA	2 Seasons	

```

...
8802 November 20, 2019 ... 2007 R 158 min
8803 July 1, 2019 2018 TV-Y7 2 Seasons
8804 November 1, 2019 2009 R 88 min
8805 January 11, 2020 2006 PG 88 min
8806 March 2, 2019 2015 TV-14 111 min

```

```

... listed_in \
0 Documentaries
1 International TV Shows, TV Dramas, TV Mysteries
2 Crime TV Shows, International TV Shows, TV Act...
3 Docuseries, Reality TV
4 International TV Shows, Romantic TV Shows, TV ...
...
8802 Cult Movies, Dramas, Thrillers
8803 Kids' TV, Korean TV Shows, TV Comedies
8804 Comedies, Horror Movies
8805 Children & Family Movies, Comedies
8806 Dramas, International Movies, Music & Musicals

```

```

... description
0 As her father nears the end of his life, filmm...
1 After crossing paths at a party, a Cape Town t...
2 To protect his family from a powerful drug lor...
3 Feuds, flirtations and toilet talk go down amo...
4 In a city of coaching centers known to train I...
...
8802 A political cartoonist, a crime reporter and a...
8803 While living alone in a spooky town, a young g...
8804 Looking to survive in a world taken over by zo...
8805 Dragged from civilian life, a former superhero...
8806 A scrappy but poor boy worms his way into a ty...

```

```
[8807 rows x 12 columns]
```

```
# Get some basic information about the dataset
```

```
df.head()
```

```

  show_id  type title director \
0      s1  Movie  Dick Johnson Is Dead  Kirsten Johnson
1      s2  TV Show      Blood & Water      NaN
2      s3  TV Show      Ganglands  Julien Leclercq
3      s4  TV Show  Jailbirds New Orleans      NaN
4      s5  TV Show      Kota Factory      NaN

```

```

                                cast  country \
0                                NaN  United States
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...      NaN

```

3		NaN	NaN
4	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...		India

	date_added	release_year	rating	duration	\
0	September 25, 2021	2020	PG-13	90 min	
1	September 24, 2021	2021	TV-MA	2 Seasons	
2	September 24, 2021	2021	TV-MA	1 Season	
3	September 24, 2021	2021	TV-MA	1 Season	
4	September 24, 2021	2021	TV-MA	2 Seasons	

	listed_in	\
0	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	
2	Crime TV Shows, International TV Shows, TV Act...	
3	Docuseries, Reality TV	
4	International TV Shows, Romantic TV Shows, TV ...	

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
3	Feuds, flirtations and toilet talk go down amo...
4	In a city of coaching centers known to train I...

df.tail()

	show_id	type	title	director	\
8802	s8803	Movie	Zodiac	David Fincher	
8803	s8804	TV Show	Zombie Dumb	NaN	
8804	s8805	Movie	Zombieland	Ruben Fleischer	
8805	s8806	Movie	Zoom	Peter Hewitt	
8806	s8807	Movie	Zubaan	Mozez Singh	

	cast	country
8802	Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...	United States
8803	NaN	NaN
8804	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States
8805	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States
8806	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India

	date_added	release_year	rating	duration	\
8802	November 20, 2019	2007	R	158 min	
8803	July 1, 2019	2018	TV-Y7	2 Seasons	
8804	November 1, 2019	2009	R	88 min	
8805	January 11, 2020	2006	PG	88 min	

```
8806      March 2, 2019      2015  TV-14      111 min
```

```
                                listed_in \
8802                        Cult Movies, Dramas, Thrillers
8803      Kids' TV, Korean TV Shows, TV Comedies
8804                        Comedies, Horror Movies
8805      Children & Family Movies, Comedies
8806  Dramas, International Movies, Music & Musicals
```

```
                                description
8802  A political cartoonist, a crime reporter and a...
8803  While living alone in a spooky town, a young g...
8804  Looking to survive in a world taken over by zo...
8805  Dragged from civilian life, a former superhero...
8806  A scrappy but poor boy worms his way into a ty...
```

```
df.size
```

```
105684
```

```
df.columns
```

```
Index(['show_id', 'type', 'title', 'director', 'cast', 'country',
      'date_added',
      'release_year', 'rating', 'duration', 'listed_in',
      'description'],
      dtype='object')
```

```
df.dtypes
```

```
show_id      object
type         object
title        object
director     object
cast         object
country      object
date_added   object
release_year  int64
rating       object
duration     object
listed_in    object
description  object
dtype: object
```

```
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

```
df.drop_duplicates() #remove duplicates
```

```

      show_id  type      title      director \
0          s1  Movie  Dick Johnson Is Dead  Kirsten Johnson
1          s2  TV Show      Blood & Water          NaN
2          s3  TV Show      Ganglands  Julien Leclercq
3          s4  TV Show  Jailbirds New Orleans          NaN
4          s5  TV Show      Kota Factory          NaN
...      ...      ...      ...      ...
8802    s8803  Movie      Zodiac  David Fincher
8803    s8804  TV Show  Zombie Dumb          NaN
8804    s8805  Movie  Zombieland  Ruben Fleischer
8805    s8806  Movie      Zoom  Peter Hewitt
8806    s8807  Movie      Zubaan  Mozez Singh

```

```

                                cast      country
\
0                                NaN  United States
1  Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...  South Africa
2  Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...      NaN
3                                NaN      NaN
4  Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...      India
...      ...      ...
8802  Mark Ruffalo, Jake Gyllenhaal, Robert Downey J...  United States
8803                                NaN      NaN
8804  Jesse Eisenberg, Woody Harrelson, Emma Stone, ...  United States
8805  Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...  United States

```

8806	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India
------	---	-------

	date_added	release_year	rating	duration	\
0	September 25, 2021	2020	PG-13	90 min	
1	September 24, 2021	2021	TV-MA	2 Seasons	
2	September 24, 2021	2021	TV-MA	1 Season	
3	September 24, 2021	2021	TV-MA	1 Season	
4	September 24, 2021	2021	TV-MA	2 Seasons	
...	
8802	November 20, 2019	2007	R	158 min	
8803	July 1, 2019	2018	TV-Y7	2 Seasons	
8804	November 1, 2019	2009	R	88 min	
8805	January 11, 2020	2006	PG	88 min	
8806	March 2, 2019	2015	TV-14	111 min	

	listed_in	\
0	Documentaries	
1	International TV Shows, TV Dramas, TV Mysteries	
2	Crime TV Shows, International TV Shows, TV Act...	
3	Docuseries, Reality TV	
4	International TV Shows, Romantic TV Shows, TV ...	
...	...	
8802	Cult Movies, Dramas, Thrillers	
8803	Kids' TV, Korean TV Shows, TV Comedies	
8804	Comedies, Horror Movies	
8805	Children & Family Movies, Comedies	
8806	Dramas, International Movies, Music & Musicals	

	description
0	As her father nears the end of his life, filmm...
1	After crossing paths at a party, a Cape Town t...
2	To protect his family from a powerful drug lor...
3	Feuds, flirtations and toilet talk go down amo...
4	In a city of coaching centers known to train I...
...	...
8802	A political cartoonist, a crime reporter and a...
8803	While living alone in a spooky town, a young g...
8804	Looking to survive in a world taken over by zo...
8805	Dragged from civilian life, a former superhero...
8806	A scrappy but poor boy worms his way into a ty...

[8807 rows x 12 columns]

is there is null value

df.isnull()

	show_id	type	title	director	cast	country	date_added	\
0	False	False	False	False	True	False	False	
1	False	False	False	True	False	False	False	
2	False	False	False	False	False	True	False	
3	False	False	False	True	True	True	False	
4	False	False	False	True	False	False	False	
...	
8802	False	False	False	False	False	False	False	
8803	False	False	False	True	True	True	False	
8804	False	False	False	False	False	False	False	
8805	False	False	False	False	False	False	False	
8806	False	False	False	False	False	False	False	

	release_year	rating	duration	listed_in	description
0	False	False	False	False	False
1	False	False	False	False	False
2	False	False	False	False	False
3	False	False	False	False	False
4	False	False	False	False	False
...
8802	False	False	False	False	False
8803	False	False	False	False	False
8804	False	False	False	False	False
8805	False	False	False	False	False
8806	False	False	False	False	False

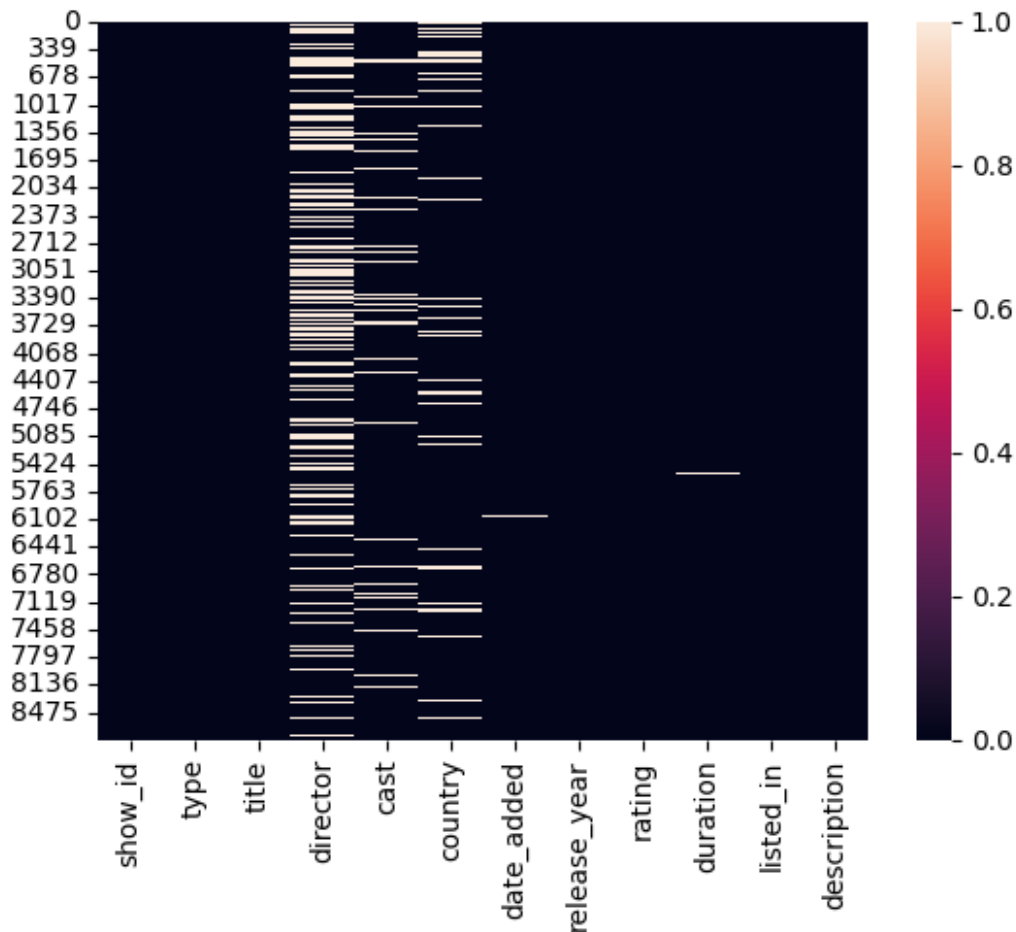
[8807 rows x 12 columns]

```
df.isnull().sum()
```

```
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
```

```
sns.heatmap(df.isnull())
```

<Axes: >



Q1. For house of cards, what is the show id and who is the director of this show?

```
df[df['title'].isin(['House of Cards'])]
```

show_id	type	title	director \		
1059	s1060	TV Show	House of Cards		
			NaN		
		cast	country		
1059		Kevin Spacey, Robin Wright, Kate Mara, Corey S...	United States		
		date_added	release_year	rating	duration
listed_in \					
1059	April 14, 2021	2018	TV-MA	6 Seasons	TV Dramas, TV Thrillers
					description
1059					A ruthless politician will stop at nothing to ...

Q2. In which year highest number of the TV Shows & Movies were released?


```

release_year_counts = df['release_year'].value_counts()
most_frequent_year = release_year_counts.index[0]
count = release_year_counts.values[0]
print(f"The year with the most TV Shows & movies is
{most_frequent_year} with a count of {count}.")

```

The year with the most TV Shows & movies is 2018 with a count of 1147.

```

filtered_df = df[(df['release_year'] >= 2010) & (df['release_year'] <=
2018)]
releases_2010_2018 =
filtered_df['release_year'].value_counts().sort_index()
print(releases_2010_2018)

```

```

release_year
2010      194
2011      185
2012      237
2013      288
2014      352
2015      560
2016      902
2017     1032
2018     1147
Name: count, dtype: int64

```

```

# Filter the DataFrame to include only the years from 2010 to 2018
filtered_df = df[(df['release_year'] >= 2010) & (df['release_year'] <=
2018)]

```

```

# Count the releases for each year in the filtered data
releases_2010_2018 =
filtered_df['release_year'].value_counts().sort_index()

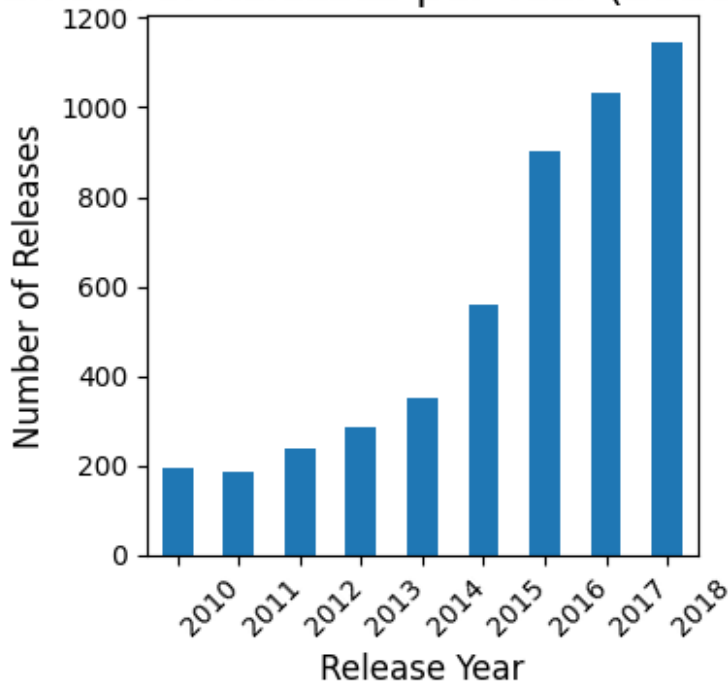
```

```

# Create a bar plot
plt.figure(figsize=(4, 4))
releases_2010_2018.plot(kind='bar')
plt.title('Number of Releases per Year (2010-2018)', fontsize=16)
plt.xlabel('Release Year', fontsize=12)
plt.ylabel('Number of Releases', fontsize=12)
plt.xticks(rotation=45, ha='left')
plt.tight_layout()

```

Number of Releases per Year (2010-2018)



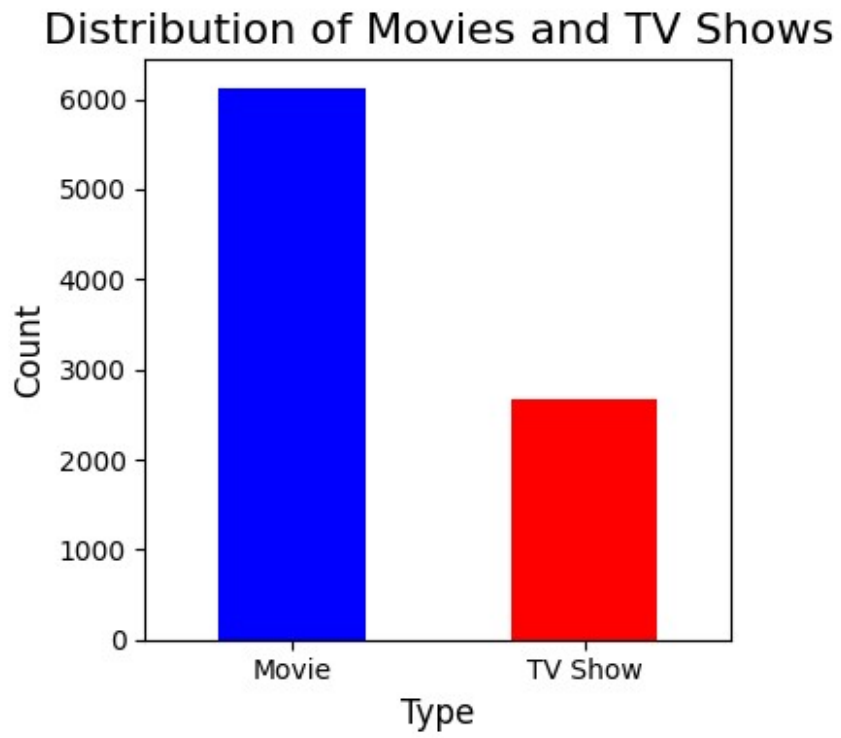
```
# Count the occurrences of each type (Movie vs. TV Show)
type_counts = df['type'].value_counts()

# Print the counts
print("Counts of Movies and TV Shows:")
print(type_counts)

# Create a bar plot
plt.figure(figsize=(4,4))
type_counts.plot(kind='bar', color=['blue', 'red'])
plt.title('Distribution of Movies and TV Shows', fontsize=16)
plt.xlabel('Type', fontsize=12)
plt.ylabel('Count', fontsize=12)
plt.xticks(rotation=0)
plt.tight_layout()
```

Counts of Movies and TV Shows:

```
type
Movie      6131
TV Show    2676
Name: count, dtype: int64
```



Q4. sh