

```
In [5]:
import pandas as pd
import plotly.express as px

In [8]:
df.dtypes

Out[8]:
show_id      object
type          object
title         object
director      object
cast          object
country       object
date_added   object
release_year  float64
rating        object
duration      object
listed_in     object
description   object
dtype: object

In [7]:
df=pd.read_excel(r"D:\POWER BI AND EXCEL\CAPSTONE DATA PROJECTS\netflix_titles.xlsx")
df

Out[7]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25 00:00:00	2020.0	PG-13	90 min	Documentary
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thabane...	South Africa	2021-09-24 00:00:00	2021.0	TV-MA	2 Seasons	International TV Show Drama Mystery
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	2021-09-24 00:00:00	2021.0	TV-MA	1 Season	Criminal International TV Show
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	2021-09-24 00:00:00	2021.0	TV-MA	1 Season	Docuseries Reality
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24 00:00:00	2021.0	TV-MA	2 Seasons	International TV Show Romantic Shows, TV Shows
...
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	2019-07-01 00:00:00	2018.0	TV-Y7	2 Seasons	Kids Korean Show Comedy
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01 00:00:00	2009.0	R	88 min	Comedy Horror Movie

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11 00:00:00	2006.0	PG	88 min	Child Family Mc Com
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02 00:00:00	2015.0	TV-14	111 min	Dre Interna Movies, f & Mus
8807	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

8808 rows × 12 columns

```
In [ ]:
#IN THE BELOW CELL I HAVE CONVERTED COLUMN DATE_ADDED TO DATE FORMAT USNING .to_datetime() FUNCTION OF PY
In [9]:
df['date_added'] = pd.to_datetime(df['date_added'])
df.dtypes
Out[9]:
show_id      object
type         object
title        object
director      object
cast         object
country      object
date_added   datetime64[ns]
release_year  float64
rating       object
duration     object
listed_in    object
description  object
dtype: object
In [ ]:
#.DESCRIBE() DEFINE THE COLUMNS OF THE DATASET
In [15]:
df.describe()
Out[15]:
```

	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

```
In [ ]:
#.DROP_DUPLICATES() DROPS THE DUPLICATE ROW IN THE DATASET
In [10]:
df.drop_duplicates()
Out[10]:
```

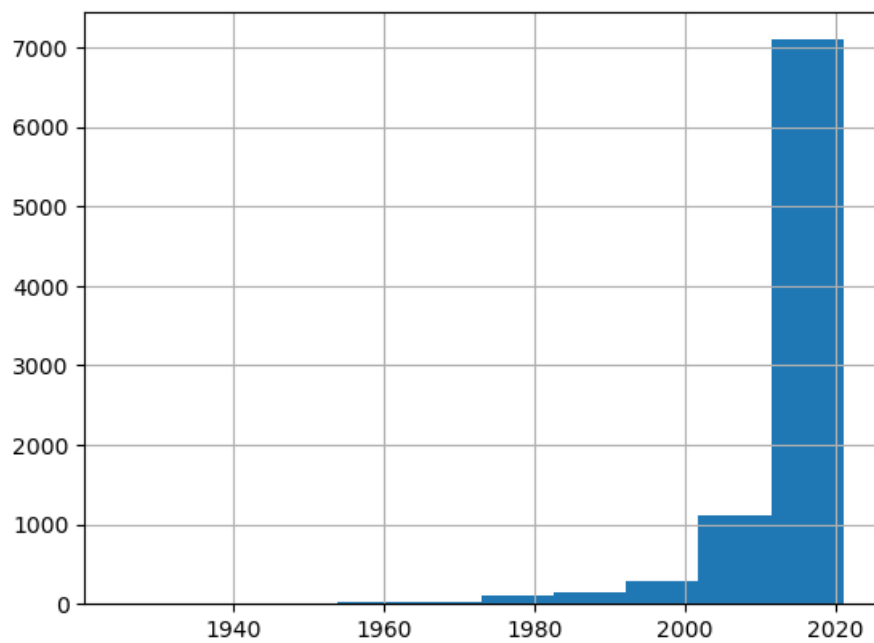
	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in
0	s1	Movie	Dick Johnson Is Dead	Kirsten Johnson	NaN	United States	2021-09-25	2020.0	PG-13	90 min	Document

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	list
1	s2	TV Show	Blood & Water	NaN	Ama Qamata, Khosi Ngema, Gail Mabalane, Thaban...	South Africa	2021-09-24	2021.0	TV-MA	2 Seasons	Internat TV Show Drama Mys
2	s3	TV Show	Ganglands	Julien Leclercq	Sami Bouajila, Tracy Gotoas, Samuel Jouy, Nabi...	NaN	2021-09-24	2021.0	TV-MA	1 Season	Crim St Internat TV Show
3	s4	TV Show	Jailbirds New Orleans	NaN	NaN	NaN	2021-09-24	2021.0	TV-MA	1 Season	Docus Realit
4	s5	TV Show	Kota Factory	NaN	Mayur More, Jitendra Kumar, Ranjan Raj, Alam K...	India	2021-09-24	2021.0	TV-MA	2 Seasons	Internat TV Sh Romant Shows, "
...	
8803	s8804	TV Show	Zombie Dumb	NaN	NaN	NaN	2019-07-01	2018.0	TV-Y7	2 Seasons	Kids Korea Show Com
8804	s8805	Movie	Zombieland	Ruben Fleischer	Jesse Eisenberg, Woody Harrelson, Emma Stone, ...	United States	2019-11-01	2009.0	R	88 min	Come Horror M
8805	s8806	Movie	Zoom	Peter Hewitt	Tim Allen, Courteney Cox, Chevy Chase, Kate Ma...	United States	2020-01-11	2006.0	PG	88 min	Child Family Mc Com
8806	s8807	Movie	Zubaan	Mozez Singh	Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan...	India	2019-03-02	2015.0	TV-14	111 min	Dre Internat Movies, I & Mus
8807	NaN	NaN	NaN	NaN	NaN	NaN	NaT	NaN	NaN	NaN	

8808 rows × 12 columns

In []:
#CREATED AN HISTOGRAM BELLOW THAT SHOWS THE NUMBER OF MOVIES RELEASED DURING THE SPECIFIED PEROD OF 1
In [21]:
df["release_year"].hist()

Out[21]:
<Axes: >



```
In [11]:  
# Assuming 'df' is your DataFrame and 'release_year' is a column in it  
fig = px.histogram(df, x="release_year")
```

```
# Show the plot  
fig.show()
```

```
In [12]:  
fig = px.histogram(df, x="release_year")
```

```
# Show the plot  
fig.show()
```

```
In [29]:
df[df["release_year"]==1925]
Out[29]:
```

	show_id	type	title	director	cast	country	date_added	release_year	rating	duration	listed_in	desc
4250	s4251	TV Show	Pioneers: First Women Filmmakers*	NaN	NaN	NaN	2018-12-30	1925	TV-14	1 Season	TV Shows	co re film wom

```
In [ ]:
#.TYPES() FUNCTION SHOWS THE DATATYPE OF THE DATASET
In [30]:
df.dtypes
Out[30]:
show_id      object
type         object
title        object
director     object
cast         object
country      object
date_added   datetime64[ns]
release_year  int64
rating       object
duration     object
listed_in    object
description  object
dtype: object
In [ ]:
#FILERED THE DATASET FOR THE MOVIES WHERE COUNTRY IS INDIA
In [38]:
filtered_df = df[df["country"] == "India"]
len(filtered_df)
print(filtered_df)
```

```

show_id  type      title      director \
4      s5  TV Show      Kota Factory      NaN
24     s25  Movie      Jeans      S. Shankar
39     s40  TV Show      Chhota Bheem      NaN
50     s51  TV Show      Dharmakshetra      NaN
66     s67  TV Show      Raja Rasoi Aur Anya Kahaniyan      NaN
...     ...     ...     ...     ...
8773   s8774  Movie      Yanda Kartavya Aahe      Kedar Shinde
8775   s8776  TV Show      Yeh Meri Family      NaN
8798   s8799  Movie      Zed Plus      Chandra Prakash Dwivedi
8799   s8800  Movie      Zenda      Avadhoot Gupte
8806   s8807  Movie      Zubaan      Mozez Singh

cast country date_added \
4      Mayur More, Jitendra Kumar, Ranjan Raj, Alam K... India 2021-09-24
24     Prashanth, Aishwarya Rai Bachchan, Sri Lakshmi... India 2021-09-21
39     Vatsal Dubey, Julie Tejwani, Rupa Bhimani, Jig... India 2021-09-16
50     Kashmiri Irani, Chandan Anand, Dinesh Mehta, A... India 2021-09-15
66     NaN India 2021-09-15
...     ...     ...     ...
8773   Ankush Choudhary, Smita Shewale, Mohan Joshi, ... India 2018-01-01
8775   Vishesh Bansal, Mona Singh, Akarsh Khurana, Ah... India 2018-08-31
8798   Adil Hussain, Mona Singh, K.K. Raina, Sanjay M... India 2019-12-31
8799   Santosh Juvekar, Siddharth Chandekar, Sachit P... India 2018-02-15
8806   Vicky Kaushal, Sarah-Jane Dias, Raaghav Chanan... India 2019-03-02

release_year rating duration \
4      2021  TV-MA  2 Seasons
24     1998  TV-14  166 min
39     2021  TV-Y7  3 Seasons
50     2014  TV-PG  1 Season
66     2014  TV-G   1 Season
...     ...     ...     ...
8773   2006  TV-PG  151 min
8775   2018  TV-PG  1 Season
8798   2014  TV-MA  131 min
8799   2009  TV-14  120 min
8806   2015  TV-14  111 min

listed_in \
4      International TV Shows, Romantic TV Shows, TV ...
24     Comedies, International Movies, Romantic Movies
39     Kids' TV
50     International TV Shows, TV Dramas, TV Sci-Fi &...
66     Docuseries, International TV Shows
...     ...
8773   Comedies, Dramas, International Movies
8775   International TV Shows, TV Comedies
8798   Comedies, Dramas, International Movies
8799   Dramas, International Movies
8806   Dramas, International Movies, Music & Musicals

description
4      In a city of coaching centers known to train I...
24     When the father of the man she loves insists t...
39     A brave, energetic little boy with superhuman ...
50     After the ancient Great War, the god Chitragup...
66     Explore the history and flavors of regional In...
...     ...
8773   Thanks to an arranged marriage that was design...
8775   In the summer of 1998, middle child Harshu bal...
8798   A philandering small-town mechanic's political...
8799   A change in the leadership of a political part...
8806   A scrappy but poor boy worms his way into a ty...

```

```

[972 rows x 12 columns]
In [ ]:
#SHOW THE DATE WHERE THE MOVEIS ARE ADDED TO THE NETFLIX
In [13]:
fig = px.histogram(df, x="date_added")

# Show the plot
fig.show()

```

```
In [14]:
df["date_added_month"]=df["date_added"].dt.month.fillna(0)
In [ ]:
#SHOW THE MONTHS IN WHICH THE MOVIES ARE ADDED TO THE NETFLIX
In [15]:
fig = px.histogram(df, x="date_added", color="date_added_month")

# Show the plot
fig.show()
```

```
In [16]:
df["date_added"] = pd.to_datetime(df["date_added"])
df["date_added_day"]=df["date_added"].dt.day.fillna(0)
In [ ]:
#SHOW THE DATE IN THE MONTH IN WHICH MOVIES ARE ADDED TO NETFLIX
In [17]:
```

```
figs= px.histogram(df, x="date_added_day")
```

```
# Show the plot  
figs.show()
```

```
In [56]:  
df.columns
```

```
Out[56]:  
Index(['show_id', 'type', 'title', 'director', 'cast', 'country', 'date_added',  
      'release_year', 'rating', 'duration', 'listed_in', 'description',  
      'date_added_month', 'date_added_day'],  
      dtype='object')
```

```
In [16]:  
df["rating"].unique()
```

```
Out[16]:  
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)
```

```
In [ ]:  
#BELOW CELLS DEFINES DIFFERENT RATING ACCORDING TO WHICH THE MOVIES ARE RATED IN DESCENDING ORDER
```

```
In [18]:  
figs= px.histogram(df, x="rating").update_xaxes(categoryorder="total descending")
```

```
# Show the plot  
figs.show()
```



```
In []:
```

```
In [19]:  
df.type.value_counts()
```

```
Out[19]:  
Movie      6131  
TV Show    2676  
Name: type, dtype: int64
```

```
In []:  
#TIME SERIES CHART FOR MOVIES AND SHOWS (INCREASE IN THE NUMBER OF TV SHOWS IN RECENT YEARS)
```

```
In [20]:  
fig = px.scatter(df, x="release_year", y="type",  
                 labels={'release_year': 'Release Year', 'type': 'Type of Movie'},  
                 title="2D Histogram for Release Year and Types of Movies")
```

```
# Show the plot  
fig.show()
```

```
in []:
#SHOW THE GROUPED RESULT OF THE RELEASE YEAR AND WHICH TYPE IS IT TV SHOW OR MOVIES AND ALSO COUNT IN E
In [21]:
# Group by 'release_year' and 'type' and calculate the count
counts_by_year = df.groupby(['release_year', 'type']).size().reset_index(name='count')

print(counts_by_year)
pd.set_option('display.max_rows', None)

# Display the counts
print(counts_by_year)
  release_year  type count
0      1925.0 TV Show    1
1      1942.0  Movie    2
2      1943.0  Movie    3
3      1944.0  Movie    3
4      1945.0  Movie    3
...      ...  ...  ...
114     2019.0 TV Show  397
115     2020.0  Movie  517
116     2020.0 TV Show  436
117     2021.0  Movie  277
118     2021.0 TV Show  315

[119 rows x 3 columns]
  release_year  type count
0      1925.0 TV Show    1
1      1942.0  Movie    2
2      1943.0  Movie    3
3      1944.0  Movie    3
4      1945.0  Movie    3
5      1945.0 TV Show    1
6      1946.0  Movie    1
7      1946.0 TV Show    1
8      1947.0  Movie    1
9      1954.0  Movie    2
10     1955.0  Movie    3
11     1956.0  Movie    2
12     1958.0  Movie    3
13     1959.0  Movie    1
14     1960.0  Movie    4
15     1961.0  Movie    1
16     1962.0  Movie    3
17     1963.0  Movie    1
18     1963.0 TV Show    1
19     1964.0  Movie    2
20     1965.0  Movie    2
21     1966.0  Movie    1
22     1967.0  Movie    4
23     1967.0 TV Show    1
24     1968.0  Movie    3
25     1969.0  Movie    2
26     1970.0  Movie    2
27     1971.0  Movie    5
28     1972.0  Movie    4
29     1972.0 TV Show    1
30     1973.0  Movie   10
31     1974.0  Movie    6
32     1974.0 TV Show    1
33     1975.0  Movie    7
34     1976.0  Movie    9
35     1977.0  Movie    6
36     1977.0 TV Show    1
37     1978.0  Movie    7
38     1979.0  Movie   10
39     1979.0 TV Show    1
40     1980.0  Movie   11
41     1981.0  Movie   12
42     1981.0 TV Show    1
43     1982.0  Movie   17
44     1983.0  Movie   11
45     1984.0  Movie   12
46     1985.0  Movie    9
47     1985.0 TV Show    1
48     1986.0  Movie   11
49     1986.0 TV Show    2
50     1987.0  Movie    8
51     1988.0  Movie   16
52     1988.0 TV Show    2
53     1989.0  Movie   15
54     1989.0 TV Show    1
55     1990.0  Movie   19
56     1990.0 TV Show    3
57     1991.0  Movie   16
```

```
57 1991.0 Movie 16
58 1991.0 TV Show 1
59 1992.0 Movie 20
60 1992.0 TV Show 3
61 1993.0 Movie 24
62 1993.0 TV Show 4
63 1994.0 Movie 20
64 1994.0 TV Show 2
65 1995.0 Movie 23
66 1995.0 TV Show 2
67 1996.0 Movie 21
68 1996.0 TV Show 3
69 1997.0 Movie 34
70 1997.0 TV Show 4
71 1998.0 Movie 32
72 1998.0 TV Show 4
73 1999.0 Movie 32
74 1999.0 TV Show 7
75 2000.0 Movie 33
76 2000.0 TV Show 4
77 2001.0 Movie 40
78 2001.0 TV Show 5
79 2002.0 Movie 44
80 2002.0 TV Show 7
81 2003.0 Movie 51
82 2003.0 TV Show 10
83 2004.0 Movie 55
84 2004.0 TV Show 9
85 2005.0 Movie 67
86 2005.0 TV Show 13
87 2006.0 Movie 82
88 2006.0 TV Show 14
89 2007.0 Movie 74
90 2007.0 TV Show 14
91 2008.0 Movie 113
92 2008.0 TV Show 23
93 2009.0 Movie 118
94 2009.0 TV Show 34
95 2010.0 Movie 154
96 2010.0 TV Show 40
97 2011.0 Movie 145
98 2011.0 TV Show 40
99 2012.0 Movie 173
100 2012.0 TV Show 64
101 2013.0 Movie 225
102 2013.0 TV Show 63
103 2014.0 Movie 264
104 2014.0 TV Show 88
105 2015.0 Movie 398
106 2015.0 TV Show 162
107 2016.0 Movie 658
108 2016.0 TV Show 244
109 2017.0 Movie 767
110 2017.0 TV Show 265
111 2018.0 Movie 767
112 2018.0 TV Show 380
113 2019.0 Movie 633
114 2019.0 TV Show 397
115 2020.0 Movie 517
116 2020.0 TV Show 436
117 2021.0 Movie 277
118 2021.0 TV Show 315
```

In []: