

## Import Library

```
In [4]: import numpy as np
from numpy import array
from numpy import argmax # it will show the all details of data
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import LabelEncoder
from sklearn.preprocessing import OneHotEncoder
from matplotlib import cm
from matplotlib.colors import ListedColormap, LinearSegmentedColormap
import numpy as npy
from PIL import Image
from sklearn import linear_model, metrics
```

## Read Dataset

```
In [6]: x=pd.read_excel(r"C:\Users\Admin\Downloads\Netflix Dataset.xlsx")
```

In [7]: x.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9425 entries, 0 to 9424
Data columns (total 29 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Title                                9425 non-null   object
1   Genre                                9400 non-null   object
2   Tags                                 9389 non-null   object
3   Languages                            9266 non-null   object
4   Series or Movie                      9425 non-null   object
5   Hidden Gem Score                     9415 non-null   float64
6   Country Availability                 9414 non-null   object
7   Runtime                             9424 non-null   object
8   Director                            7120 non-null   object
9   Writer                              7615 non-null   object
10  Actors                              9314 non-null   object
11  View Rating                         6827 non-null   object
12  IMDb Score                          9417 non-null   float64
13  Rotten Tomatoes Score               5445 non-null   float64
14  Metacritic Score                    4082 non-null   float64
15  Awards Received                     5226 non-null   float64
16  Awards Nominated For                6376 non-null   float64
17  Boxoffice                           3754 non-null   float64
18  Release Date                        9217 non-null   datetime64[ns]
19  Netflix Release Date                9425 non-null   datetime64[ns]
20  Production House                    4393 non-null   object
21  Netflix Link                        9425 non-null   object
22  IMDb Link                           9101 non-null   object
23  Summary                             9420 non-null   object
24  IMDb Votes                          9415 non-null   float64
25  Image                               9425 non-null   object
26  Poster                              8487 non-null   object
27  TMDB Trailer                        9425 non-null   object
28  Trailer Site                        9424 non-null   object
dtypes: datetime64[ns](2), float64(8), object(19)
memory usage: 2.1+ MB
```

```
In [8]: x.tail()
```

Out[8]:

	Title	Genre	Tags	Languages	Series or Movie	Hidden Gem Score	
9420	13 Going on 30	Comedy, Fantasy, Romance	Comedies,Comedies,Romantic Films,Roma...	English, Portuguese	Movie	3.1	H
9421	LIFE 2.0	Documentary	Social & Cultural Documentaries,Biographical D...	English	Movie	8.2	Hungari
9422	Brand New Day	Documentary, Music	Australian Comedies,Romantic Comedies,Australi...	English	Movie	8.6	
9423	Daniel Arends: Blessuretijd	Comedy	Stand-up Comedy,International Movies,Comedies	Dutch	Movie	8.8	
9424	DreamWorks Happy Holidays from Madagascar	Animation, Comedy, Family	TV Comedies,Kids TV,Animal Tales,TV Cartoons,T...	English	Series	8.4	

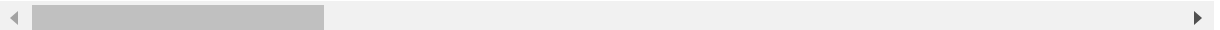
5 rows × 29 columns

In [9]: x.head()

Out[9]:

	Title	Genre	Tags	Languages	Series or Movie	Hidden Gem Score	
0	Lets Fight Ghost	Crime, Drama, Fantasy, Horror, Romance	Programmes,Comedy Romantic TV Comedies,Horror ...	Swedish, Spanish	Series	4.3	
1	HOW TO BUILD A GIRL	Comedy	Dramas,Comedies,Films Based on Books,British	English	Movie	7.0	
2	The Con-Heartist	Comedy, Romance	Comedies,Comedies,Romantic Films,Thai...	Thai	Movie	8.6	
3	Gleboka woda	Drama	TV Dramas,Polish TV Shows,Social Issue TV Dramas	Polish	Series	8.7	
4	Only a Mother	Drama	Social Issue Dramas,Dramas,Movies Based on Boo...	Swedish	Movie	8.3	Lithuania,Poland,Fr

5 rows × 29 columns



In [10]: x.describe()

Out[10]:

	Hidden Gem Score	IMDb Score	Rotten Tomatoes Score	Metacritic Score	Awards Received	Awards Nominated For	Boxoff
count	9415.000000	9417.000000	5445.000000	4082.000000	5226.000000	6376.000000	3.754000e-
mean	5.540733	6.955517	64.691276	58.113425	9.735936	16.035602	4.845788e-
std	2.447462	0.899681	25.269466	17.143187	19.524116	32.209094	7.243625e-
min	0.600000	1.600000	0.000000	6.000000	1.000000	1.000000	7.200000e-
25%	3.400000	6.500000	49.000000	46.000000	1.250000	2.000000	1.243910e-
50%	5.300000	7.000000	70.000000	59.000000	4.000000	6.000000	2.250466e-
75%	8.100000	7.500000	85.000000	71.000000	9.000000	15.000000	6.425437e-
max	9.800000	9.700000	100.000000	100.000000	300.000000	386.000000	6.593639e-



In [11]: `x.notnull()`

Out[11]:

	Title	Genre	Tags	Languages	Series or Movie	Hidden Gem Score	Country Availability	Runtime	Director	Writer	...
0	True	True	True	True	True	True	True	True	True	True	...
1	True	True	True	True	True	True	True	True	True	True	...
2	True	True	True	True	True	True	True	True	True	True	...
3	True	True	True	True	True	True	True	True	False	False	...
4	True	True	True	True	True	True	True	True	True	True	...
...	...	...	...	...	...	...	...	...	...	...	...
9420	True	True	True	True	True	True	True	True	True	True	...
9421	True	True	True	True	True	True	True	True	True	False	...
9422	True	True	True	True	True	True	True	True	True	False	...
9423	True	True	True	True	True	True	True	True	True	True	...
9424	True	True	True	True	True	True	True	True	False	False	...

9425 rows × 29 columns



In [12]: `x.isnull()`

Out[12]:

	Title	Genre	Tags	Languages	Series or Movie	Hidden Gem Score	Country Availability	Runtime	Director	Writer	...
0	False	False	False	False	False	False	False	False	False	False	...
1	False	False	False	False	False	False	False	False	False	False	...
2	False	False	False	False	False	False	False	False	False	False	...
3	False	False	False	False	False	False	False	False	True	True	...
4	False	False	False	False	False	False	False	False	False	False	...
...	...	...	...	...	...	...	...	...	...	...	...
9420	False	False	False	False	False	False	False	False	False	False	...
9421	False	False	False	False	False	False	False	False	False	True	...
9422	False	False	False	False	False	False	False	False	False	True	...
9423	False	False	False	False	False	False	False	False	False	False	...
9424	False	False	False	False	False	False	False	False	True	True	...

9425 rows × 29 columns



```
In [13]: x.isnull().sum()
```

```
Out[13]: Title                0
Genre                25
Tags                 36
Languages            159
Series or Movie      0
Hidden Gem Score     10
Country Availability  11
Runtime              1
Director             2305
Writer               1810
Actors               111
View Rating          2598
IMDb Score           8
Rotten Tomatoes Score 3980
Metacritic Score     5343
Awards Received      4199
Awards Nominated For 3049
Boxoffice            5671
Release Date         208
Netflix Release Date 0
Production House     5032
Netflix Link         0
IMDb Link            324
Summary              5
IMDb Votes           10
Image                0
Poster              938
TMDb Trailer         0
Trailer Site         1
dtype: int64
```

In [14]: `x.dropna()`

Out[14]:

	Title	Genre	Tags	Languages	Series or Movie	Hidden Gem Score	
0	Lets Fight Ghost	Crime, Drama, Fantasy, Horror, Romance	Comedy Programmes,Romantic TV Comedies,Horror ...	Swedish, Spanish	Series	4.3	
9	Joker	Crime, Drama, Thriller	Dark Comedies,Crime Comedies,Dramas,Comedies,C...	English	Movie	3.5	Lithuania
10	I	Action, Adventure, Fantasy, Sci-Fi	Dramas,Swedish Movies	English, Sanskrit	Movie	2.8	Lithuania,I
11	Harrys Daughters	Adventure, Drama, Fantasy, Mystery	Dramas,Swedish Movies	English	Movie	4.4	Lithuania,I
17	The Closet	Comedy	Korean Movies,Horror Movies,Mysteries	French	Movie	3.8	
...	...	...	...	...	...	...	
9411	50 First Dates	Comedy, Drama, Romance	Romantic Favourites,Romantic Comedies,Comedies...	English, Hawaiian, Mandarin	Movie	2.7	
9412	21	Crime, Drama, History, Thriller	Dramas,Dramas based on a book,Police Dramas,Po...	English	Movie	2.5	
9414	One Chance	Biography, Comedy, Drama, Music	Dramas,Biographical Dramas,Dramas based on rea...	English, Italian	Movie	3.0	
9415	The Twilight Saga: Breaking Dawn: Part 1	Adventure, Drama, Fantasy, Romance, Thriller	Dramas,Romantic Dramas,Dramas based on a book,...	English, Portuguese	Movie	2.0	Can
9416	One for the Money	Action, Comedy, Crime, Thriller	Romantic Comedies>Action Comedies,Comedies,Pol...	English	Movie	1.3	



2155 rows × 29 columns

```
In [20]: x.isna()
```

Out[20]:

	Title	Genre	Tags	Languages	Series or Movie	Hidden Gem Score	Country Availability	Runtime	Director	Writer	...
0	False	False	False	False	False	False	False	False	False	False	...
1	False	False	False	False	False	False	False	False	False	False	...
2	False	False	False	False	False	False	False	False	False	False	...
3	False	False	False	False	False	False	False	False	True	True	...
4	False	False	False	False	False	False	False	False	False	False	...
...	...	...	...	...	...	...	...	...	...	...	...
9420	False	False	False	False	False	False	False	False	False	False	...
9421	False	False	False	False	False	False	False	False	False	True	...
9422	False	False	False	False	False	False	False	False	False	True	...
9423	False	False	False	False	False	False	False	False	False	False	...
9424	False	False	False	False	False	False	False	False	True	True	...

9425 rows × 29 columns

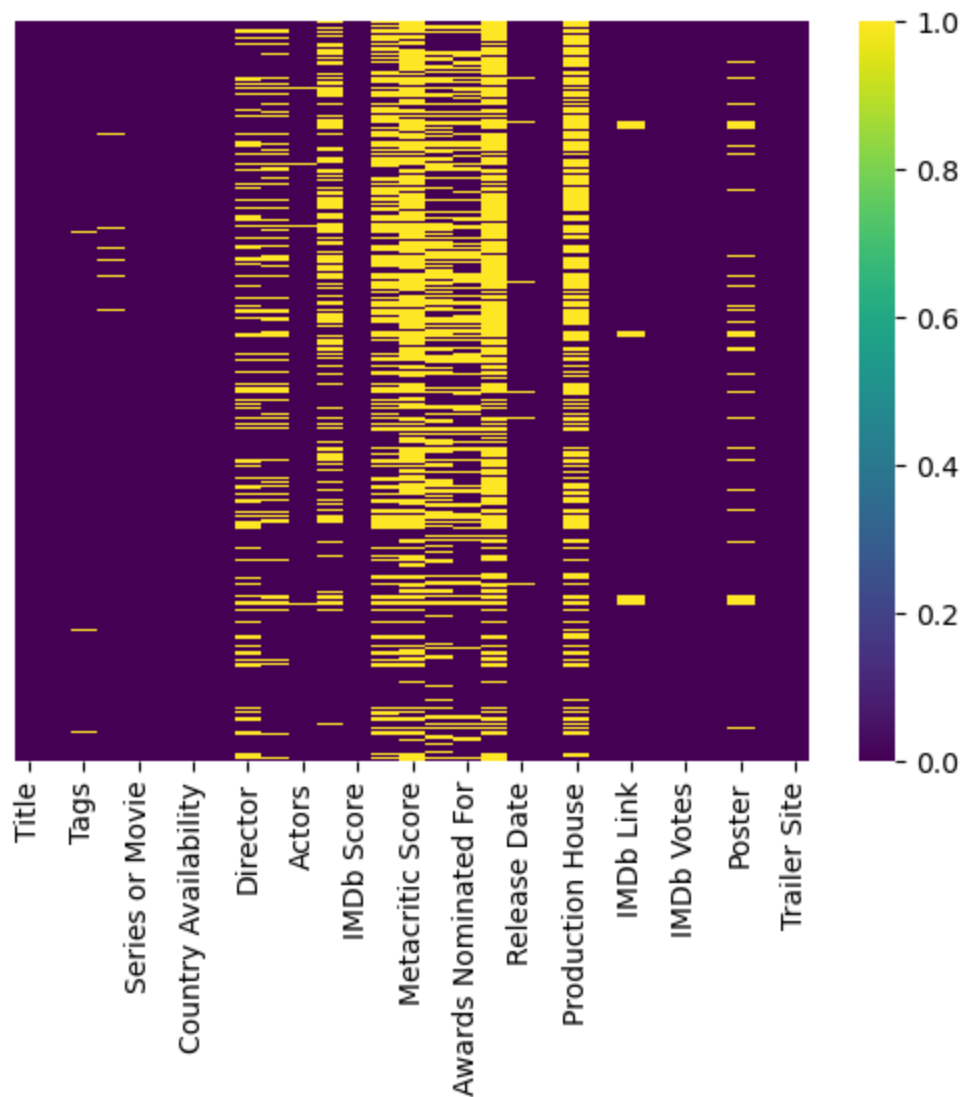


```
In [24]: x.isna().sum()
```

```
Out[24]: Title                0
Genre                25
Tags                 36
Languages            159
Series or Movie      0
Hidden Gem Score     10
Country Availability  11
Runtime              1
Director             2305
Writer               1810
Actors               111
View Rating          2598
IMDb Score            8
Rotten Tomatoes Score 3980
Metacritic Score     5343
Awards Received      4199
Awards Nominated For 3049
Boxoffice            5671
Release Date         208
Netflix Release Date  0
Production House     5032
Netflix Link          0
IMDb Link             324
Summary              5
IMDb Votes           10
Image                0
Poster              938
TMDb Trailer         0
Trailer Site         1
dtype: int64
```

```
In [27]: sns.heatmap(x.isnull(),yticklabels=False,cmap="viridis")
```

```
Out[27]: <Axes: >
```



```
In [30]: x.drop(['Title', 'Genre', 'IMDb Score', 'Trailer Site', 'Runtime', 'Image', 'IMDb Li
```

In [31]: x

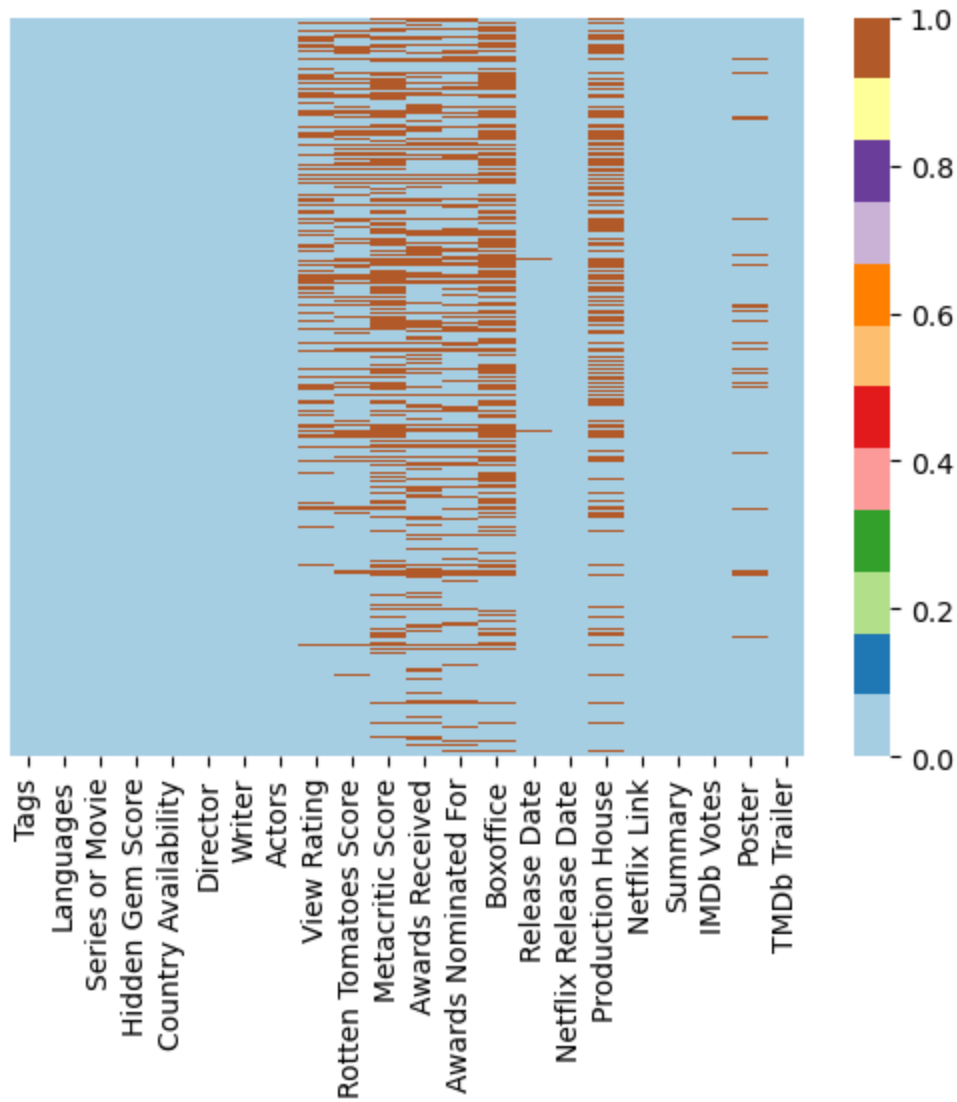
Out[31]:

		Tags	Languages	Series or Movie	Hidden Gem Score	Country A
0		Comedy Programmes,Romantic TV Comedies,Horror ...	Swedish, Spanish	Series	4.3	
1		Dramas,Comedies,Films Based on Books,British	English	Movie	7.0	
2		Romantic Comedies,Comedies,Romantic Films,Thai...	Thai	Movie	8.6	
3		TV Dramas,Polish TV Shows,Social Issue TV Dramas	Polish	Series	8.7	
4		Social Issue Dramas,Dramas,Movies Based on Boo...	Swedish	Movie	8.3	Lithuania,Poland,France,Italy,Spain,Gr
...		...	...	...	...	
9420		Romantic Comedies,Comedies,Romantic Films,Roma...	English, Portuguese	Movie	3.1	Hungary,Lithuania,Romæ Kingd
9421		Social & Cultural Documentaries,Biographical D...	English	Movie	8.2	Hungary,Lithuania,India,Australia,Unit
9422		Australian Comedies,Romantic Comedies,Australi...	English	Movie	8.6	
9423		Stand-up Comedy,International Movies,Comedies	Dutch	Movie	8.8	
9424		TV Comedies,Kids TV,Animal Tales,TV Cartoons,T...	English	Series	8.4	Belgium,Switzerl States,Gern

9425 rows × 22 columns

```
In [43]: sns.heatmap(x.isnull(),yticklabels=False,cmap="Paired")
```

```
Out[43]: <Axes: >
```



```
In [ ]:
```

```
In [42]: x.dropna(subset=['Director'],inplace=True)      #subset is a part of variable
x.dropna(subset=['Writer'],inplace=True)
x.dropna(subset=['Languages'],inplace=True)      #ye queries run nhi kiya kuki m
x.dropna(subset=['Actors'],inplace=True)
x.dropna(subset=['Genre'],inplace=True)
x.dropna(subset=['country Availability'],inplace=True)
```

**KeyError**

Traceback (most recent call last)

Cell In[42], line 5

```
3 x.dropna(subset=['Languages'],inplace=True)
4 x.dropna(subset=['Actors'],inplace=True)
----> 5 x.dropna(subset=['Genre'],inplace=True)
6 x.dropna(subset=['country Availability'],inplace=True)
```

File ~\anaconda3\Lib\site-packages\pandas\util\\_decorators.py:331, in deprecate\_nonkeyword\_arguments.<locals>.decorate.<locals>.wrapper(\*args, \*\*kwargs)

```
325 if len(args) > num_allow_args:
326     warnings.warn(
327         msg.format(arguments=_format_argument_list(allow_args)),
328         FutureWarning,
329         stacklevel=find_stack_level(),
330     )
--> 331 return func(*args, **kwargs)
```

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:6561, in DataFrame.dropna(self, axis, how, thresh, subset, inplace)

```
6559     check = indices == -1
6560     if check.any():
-> 6561         raise KeyError(np.array(subset)[check].tolist())
6562     agg_obj = self.take(indices, axis=agg_axis)
6564 if thresh is not no_default:
```

**KeyError:** ['Genre']

```
In [48]: x['Hidden Gem Score'].ffill(axis=0,inplace=True)
x['IMDb Votes']=x['IMDb Votes'].fillna(x['IMDb Votes'].mode()[0])
```

```
In [49]: x.isna().sum()    #sum() it will collect all the data
```

```
Out[49]: Tags                26
Languages                   0
Series or Movie             0
Hidden Gem Score            0
Country Availability         9
Director                    0
Writer                      0
Actors                      0
View Rating                 1442
Rotten Tomatoes Score       1427
Metacritic Score            2730
Awards Received             2533
Awards Nominated For        1681
Boxoffice                   3005
Release Date                59
Netflix Release Date        0
Production House            2407
Netflix Link                0
Summary                     4
IMDb Votes                  0
Poster                     408
TMDb Trailer                0
dtype: int64
```

```
In [50]: x['Rotten Tomatoes Score'].fillna(x['Rotten Tomatoes Score'].mean(), inplace=True)
```

```
In [52]: x['Awards Received']=x['Awards Received'].fillna(0)
```

```
In [54]: x['Awards Nominated For']=x['Awards Nominated For'].replace(np.nan,0)     #(np.nan)
```



```
In [55]: x.isna().sum()
```

```
Out[55]: Tags                26
Languages                   0
Series or Movie             0
Hidden Gem Score            0
Country Availability        9
Director                    0
Writer                      0
Actors                      0
View Rating                1442
Rotten Tomatoes Score      0
Metacritic Score          2730
Awards Received            0
Awards Nominated For      0
Boxoffice                 3005
Release Date               59
Netflix Release Date       0
Production House          2407
Netflix Link               0
Summary                    4
IMDb Votes                 0
Poster                    408
TMDb Trailer              0
dtype: int64
```

```
In [ ]:
```