

```
In [2]: import pandas as pd
import matplotlib as mt
import seaborn as sns
```

```
In [3]: df=pd.read_csv('netflix.csv')
```

```
In [4]: df.shape
```

```
Out[4]: (583, 7)
```

```
In [5]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 583 entries, 0 to 582
Data columns (total 7 columns):
#   Column          Non-Null Count  Dtype
---  -
0   title           583 non-null   object
1   genre           583 non-null   object
2   language        583 non-null   object
3   imdb_score      583 non-null   float64
4   premiere        583 non-null   object
5   runtime         583 non-null   int64
6   year            583 non-null   int64
dtypes: float64(1), int64(2), object(4)
memory usage: 32.0+ KB
```

```
In [6]: df.describe()
# count This shows the number of non-null values in each numerical column
#mean
#std ,
#25%= Q1 value below which 25% of the data falls.
# 50% median
#75%= Q3 value below which 75% of the data falls.
#max and min val in that column
```

```
Out[6]:
```

	imdb_score	runtime	year
count	583.000000	583.000000	583.000000
mean	6.275129	93.490566	2018.934820
std	0.976678	27.706665	1.474598
min	2.500000	4.000000	2014.000000
25%	5.700000	86.000000	2018.000000
50%	6.400000	97.000000	2019.000000
75%	7.000000	107.500000	2020.000000
max	9.000000	209.000000	2021.000000

```
In [7]: df.dtypes
```

```
Out[7]: title          object
genre          object
language       object
imdb_score     float64
premiere       object
runtime        int64
year           int64
dtype: object
```

```
In [8]: df.head()
```

```
Out[8]:
```

	title	genre	language	imdb_score	premiere	runtime	year
0	Notes for My Son	Drama	Spanish	6.3	11/24/2020	83	2020
1	To Each, Her Own	Romantic comedy	French	5.3	6/24/2018	95	2018
2	The Lovebirds	Romantic comedy	English	6.1	5/22/2020	87	2020
3	The Perfection	Horror-thriller	English	6.1	5/24/2019	90	2019
4	Happy Anniversary	Romantic comedy	English	5.8	3/30/2018	78	2018

```
In [9]: df.tail()
```

```
Out[9]:
```

	title	genre	language	imdb_score	premiere	runtime	year
578	Rolling Thunder Revue: A bob Dylan Story by Ma...	Documentary	English	7.6	6/12/2019	144	2019
579	Freaks: You're One of Us	Supernatural drama	German	5.4	9/2/2020	92	2020
580	Squared Love	Romantic comedy	Polish	5.0	2/11/2021	102	2021
581	My Own Man	Documentary	English	6.4	12/13/2014	81	2014
582	Biggie: I Got a Story to Tell	Documentary	English	6.9	3/1/2021	97	2021

```
In [10]: df.isnull()
```

```
Out[10]:
```

	title	genre	language	imdb_score	premiere	runtime	year
0	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False
...
578	False	False	False	False	False	False	False
579	False	False	False	False	False	False	False
580	False	False	False	False	False	False	False
581	False	False	False	False	False	False	False
582	False	False	False	False	False	False	False

583 rows x 7 columns

```
In [11]: df.isnull().sum()
```

```
Out[11]: title      0
genre      0
language    0
imdb_score  0
premiere    0
runtime     0
year        0
dtype: int64
```

```
In [12]: print("Data Types:")
print(df.dtypes)
```

```
Data Types:
title      object
genre      object
language    object
imdb_score float64
premiere    object
runtime     int64
year        int64
dtype: object
```

```
In [13]: df['premiere']=pd.to_datetime(df['premiere'], format='%m/%d/%Y')
```

```
In [14]: print("Data Types:")
print(df.dtypes)
```

```
Data Types:
title          object
genre          object
language       object
imdb_score     float64
premiere       datetime64[ns]
runtime        int64
year           int64
dtype: object
```

```
In [15]: print(df.premiere)
```

```
0      2020-11-24
1      2018-06-24
2      2020-05-22
3      2019-05-24
4      2018-03-30
...
578    2019-06-12
579    2020-09-02
580    2021-02-11
581    2014-12-13
582    2021-03-01
Name: premiere, Length: 583, dtype: datetime64[ns]
```

```
In [16]: df.dtypes
```

```
Out[16]: title          object
genre          object
language       object
imdb_score     float64
premiere       datetime64[ns]
runtime        int64
year           int64
dtype: object
```

```
In [17]: title=df[['title']]
df2=pd.get_dummies(title, dtype='int')
```

```
In [18]: df.title
```

```
Out[18]: 0      Notes for My Son
1      To Each, Her Own
2      The Lovebirds
3      The Perfection
4      Happy Anniversary
...
578    Rolling Thunder Revue: A bob Dylan Story by Ma...
579      Freaks: You're One of Us
580      Squared Love
581      My Own Man
582    Biggie: I Got a Story to Tell
Name: title, Length: 583, dtype: object
```

```
In [19]: df.dtypes
```

```
Out[19]: title          object
genre          object
language       object
imdb_score     float64
premiere       datetime64[ns]
runtime        int64
year           int64
dtype: object
```

```
In [20]: pd.get_dummies(df,dtype='int')
```

```
Out[20]:
```

	imdb_score	premiere	runtime	year	title_#REALITYHIGH	title_13th	title_13th Conversat with Op Winfre / DuVeri
0	6.3	2020-11-24	83	2020	0	0	
1	5.3	2018-06-24	95	2018	0	0	
2	6.1	2020-05-22	87	2020	0	0	
3	6.1	2019-05-24	90	2019	0	0	
4	5.8	2018-03-30	78	2018	0	0	
...	
578	7.6	2019-06-12	144	2019	0	0	
579	5.4	2020-09-02	92	2020	0	0	
580	5.0	2021-02-11	102	2021	0	0	
581	6.4	2014-12-13	81	2014	0	0	
582	6.9	2021-03-01	97	2021	0	0	

583 rows x 739 columns

```
In [22]: df
```

Out [22]:

	title	genre	language	imdb_score	premiere	runtime	year
0	Notes for My Son	Drama	Spanish	6.3	2020-11-24	83	2020
1	To Each, Her Own	Romantic comedy	French	5.3	2018-06-24	95	2018
2	The Lovebirds	Romantic comedy	English	6.1	2020-05-22	87	2020
3	The Perfection	Horror-thriller	English	6.1	2019-05-24	90	2019
4	Happy Anniversary	Romantic comedy	English	5.8	2018-03-30	78	2018
...
578	Rolling Thunder Revue: A bob Dylan Story by Ma...	Documentary	English	7.6	2019-06-12	144	2019
579	Freaks: You're One of Us	Supernatural drama	German	5.4	2020-09-02	92	2020
580	Squared Love	Romantic comedy	Polish	5.0	2021-02-11	102	2021
581	My Own Man	Documentary	English	6.4	2014-12-13	81	2014
582	Biggie: I Got a Story to Tell	Documentary	English	6.9	2021-03-01	97	2021

583 rows × 7 columns

In []: