

```
In [1]: import numpy as np
import pandas as pd
```

```
In [2]: imdb = pd.read_csv(r"C:\Users\Kajal\Desktop\ssd\Desktop\dataset work November\datasets\IMDb_Data_final.csv")
imdb
```

Out[2]:

|     | Title                             | Director            | Stars   | IMDb-Rating | Category                | Duration | Censor-board-rating | ReleaseYear |
|-----|-----------------------------------|---------------------|---|-------------|-------------------------|----------|---------------------|-------------|
| 0   | Top Gun: Maverick                 | JosephKosinski      | TomCruise, JenniferConnelly, MilesTeller, Valk... | 8.6         | Action,Drama            | 130min   | UA                  | 2022        |
| 1   | Everything Everywhere All at Once | DanKwan,            | , MichelleYeoh, StephanieHsu, KeHuyQuan, James... | 8.3         | Action,Adventure,Comedy | 139min   | R                   | 2022        |
| 2   | The Batman                        | MattReeves          | RobertPattinson, ZoëKravitz, JeffreyWright, Co... | 7.9         | Action,Crime,Drama      | 176min   | UA                  | 2022        |
| 3   | Jurassic Park                     | StevenSpielberg     | SamNeill, LauraDern, JeffGoldblum, RichardAtte... | 8.2         | Action,Adventure,Sci-Fi | 127min   | UA                  | 1993        |
| 4   | The Godfather                     | FrancisFordCoppola  | MarlonBrando, AlPacino, JamesCaan, DianeKeaton    | 9.2         | Crime,Drama             | 175min   | A                   | 1972        |
| ... | ...                               | ...                 | ...   | ...         | ...                     | ...      | ...                 | ...         |
| 995 | Vizontele                         | YilmazErdogan,      | , YilmazErdogan, DemetAkbag, AltanErkekli, Cem... | 8.0         | Comedy,Drama            | 110min   | NaN                 | 2001        |
| 996 | Sarfarosh                         | JohnMathewMatthan   | AamirKhan, NaseeruddinShah, SonaliBendre, Muke... | 8.1         | Action,Drama,Thriller   | 174min   | A                   | 1999        |
| 997 | Udaan                             | VikramadityaMotwane | RajatBarmecha, RonitRoy, ManjotSingh, RamKapoor   | 8.1         | Drama                   | 134min   | UA                  | 2010        |
| 998 | English Vinglish                  | GauriShinde         | Sridevi, AdilHussain, MehdiNebbou, PriyaAnand     | 7.8         | Comedy,Drama,Family     | 134min   | U                   | 2012        |
| 999 | Anand                             | HrishikeshMukherjee | RajeshKhanna, AmitabhBachchan, SumitaSanyal, R... | 8.1         | Drama,Musical           | 122min   | U                   | 1971        |

1000 rows × 8 columns

In [3]: `imdb.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 8 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   Title                  1000 non-null   object
1   Director               1000 non-null   object
2   Stars                  1000 non-null   object
3   IMDb-Rating            1000 non-null   float64
4   Category               999 non-null    object
5   Duration               999 non-null    object
6   Censor-board-rating    846 non-null    object
7   ReleaseYear            1000 non-null   int64
dtypes: float64(1), int64(1), object(6)
memory usage: 62.6+ KB
```

In [4]: `imdb.describe()`

Out[4]:

|              | IMDb-Rating | ReleaseYear |
|--------------|-------------|-------------|
| <b>count</b> | 1000.000000 | 1000.000000 |
| <b>mean</b>  | 7.959900    | 1991.22400  |
| <b>std</b>   | 0.277018    | 23.84484    |
| <b>min</b>   | 7.600000    | 1920.00000  |
| <b>25%</b>   | 7.700000    | 1975.00000  |
| <b>50%</b>   | 7.900000    | 1999.00000  |
| <b>75%</b>   | 8.100000    | 2010.00000  |
| <b>max</b>   | 9.300000    | 2022.00000  |

```
In [5]: imdb.isnull().sum()
```

```
Out[5]: Title                0
        Director            0
        Stars               0
        IMDb-Rating         0
        Category            1
        Duration            1
        Censor-board-rating 154
        ReleaseYear         0
        dtype: int64
```

```
In [6]: imdb["Censor-board-rating"].unique()
```

```
Out[6]: array(['UA', 'R', 'A', 'U', 'U/A', '18', '15+', '16', '13', nan, '12+',
              '7', 'PG-13', '(Banned)', 'PG', 'NotRated', 'All', 'G', 'Unrated',
              '96min', 'UA13+', 'M/PG', '125min'], dtype=object)
```

```
In [7]: imdb["Censor-board-rating"].value_counts()
```

```
Out[7]: U          270  
       A          215  
       UA         197  
       R           76  
       PG-13       20  
       18          13  
       NotRated    11  
       PG          10  
       16           8  
       13           6  
       7            5  
       U/A         2  
       15+         2  
       G           2  
       Unrated     2  
       12+         1  
       (Banned)    1  
       All         1  
       96min       1  
       UA13+       1  
       M/PG        1  
       125min      1  
       Name: Censor-board-rating, dtype: int64
```

```
In [8]: imdb["Category"].value_counts()
```

```
Out[8]: Drama                        83
        Drama,Romance                37
        Comedy,Drama,Romance         34
        Comedy,Drama                 34
        Action,Crime,Drama            33
        ..
        Animation,Crime,Drama         1
        Drama,Horror,Mystery          1
        Action,Adventure,Crime        1
        Animation,Comedy,Family        1
        Mystery,Romance,Thriller       1
        Name: Category, Length: 202, dtype: int64
```

```
In [9]: imdb["Duration"].value_counts()
```

```
Out[9]: 130min      24
        100min      23
        129min      21
        113min      21
        122min      20
        ..
        212min       1
        185min       1
        78min        1
        191min       1
        192min       1
        Name: Duration, Length: 143, dtype: int64
```

```
In [10]: imdb["Censor-board-rating"].fillna('G',inplace=True)
```

```
In [11]: imdb["Category"].fillna('Drama,Horror,Mystery',inplace=True)
```

```
In [12]: imdb["Duration"].fillna('78min',inplace=True)
```

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

```
Out[13]: Title                0  
         Director            0  
         Stars               0  
         IMDb-Rating         0  
         Category            0  
         Duration            0  
         Censor-board-rating 0  
         ReleaseYear         0  
         dtype: int64
```

```
In [14]: imdb["Duration"]=imdb["Duration"].str.replace("min"," ")
```

In [15]: imdb

Out[15]:

|     | Title                             | Director            | Stars   | IMDb-Rating | Category                | Duration | Censor-board-rating | ReleaseYear |
|-----|-----------------------------------|---------------------|---|-------------|-------------------------|----------|---------------------|-------------|
| 0   | Top Gun: Maverick                 | JosephKosinski      | TomCruise, JenniferConnelly, MilesTeller, Valk... | 8.6         | Action,Drama            | 130      | UA                  | 2022        |
| 1   | Everything Everywhere All at Once | DanKwan,            | , MichelleYeoh, StephanieHsu, KeHuyQuan, James... | 8.3         | Action,Adventure,Comedy | 139      | R                   | 2022        |
| 2   | The Batman                        | MattReeves          | RobertPattinson, ZoëKravitz, JeffreyWright, Co... | 7.9         | Action,Crime,Drama      | 176      | UA                  | 2022        |
| 3   | Jurassic Park                     | StevenSpielberg     | SamNeill, LauraDern, JeffGoldblum, RichardAtte... | 8.2         | Action,Adventure,Sci-Fi | 127      | UA                  | 1993        |
| 4   | The Godfather                     | FrancisFordCoppola  | MarlonBrando, APacino, JamesCaan, DianeKeaton     | 9.2         | Crime,Drama             | 175      | A                   | 1972        |
| ... | ...                               | ...                 | ...   | ...         | ...                     | ...      | ...                 | ...         |
| 995 | Vizontele                         | YilmazErdogan,      | , YilmazErdogan, DemetAkbag, AltanErkekli, Cem... | 8.0         | Comedy,Drama            | 110      | G                   | 2001        |
| 996 | Sarfaroosh                        | JohnMathewMatthan   | AamirKhan, NaseeruddinShah, SonaliBendre, Muke... | 8.1         | Action,Drama,Thriller   | 174      | A                   | 1999        |
| 997 | Udaan                             | VikramadityaMotwane | RajatBarmecha, RonitRoy, ManjotSingh, RamKapoor   | 8.1         | Drama                   | 134      | UA                  | 2010        |
| 998 | English Vinglish                  | GauriShinde         | Sridevi, AdilHussain, MehdiNebbou, PriyaAnand     | 7.8         | Comedy,Drama,Family     | 134      | U                   | 2012        |
| 999 | Anand                             | HrishikeshMukherjee | RajeshKhanna, AmitabhBachchan, SumitaSanyal, R... | 8.1         | Drama,Musical           | 122      | U                   | 1971        |

1000 rows × 8 columns

```
In [16]: imdb_1 = imdb.loc[(imdb["IMDb-Rating"]==9.2) | (imdb["ReleaseYear"]==2001) | (imdb["ReleaseYear"]==2010)]
imdb_1
```

Out[16]:

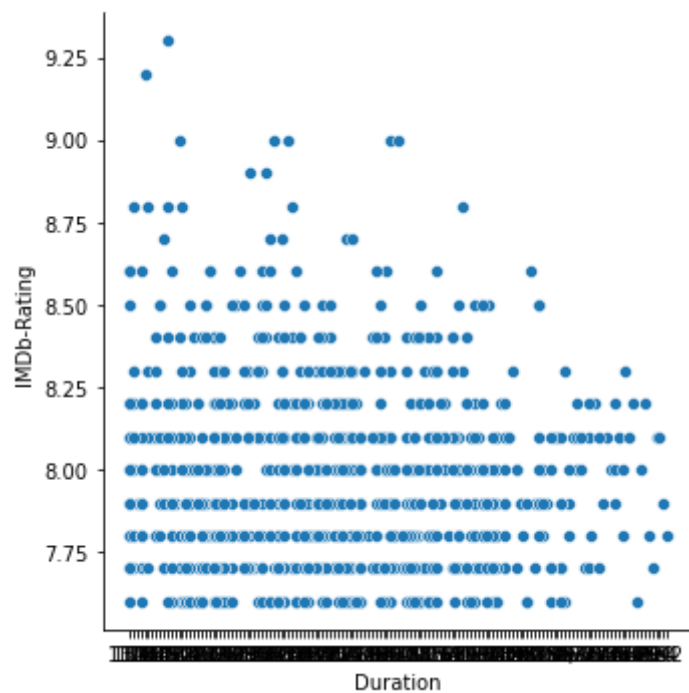
|    | Title   | Director           | Stars   | IMDb-Rating | Category                 | Duration | Censor-board-rating | ReleaseYear |
|----|---|--------------------|---|-------------|--------------------------|----------|---------------------|-------------|
| 4  | The Godfather                                     | FrancisFordCoppola | MarlonBrando, AlPacino, JamesCaan, DianeKeaton    | 9.2         | Crime,Drama              | 175      | A                   | 1972        |
| 13 | Harry Potter and the Sorcerers Stone              | ChrisColumbus      | DanielRadcliffe, RupertGrint, RichardHarris, M... | 7.6         | Adventure,Family,Fantasy | 152      | U                   | 2001        |
| 30 | Inception   | ChristopherNolan   | LeonardoDiCaprio, JosephGordon-Levitt, ElliotP... | 8.8         | Action,Adventure,Sci-Fi  | 148      | UA                  | 2010        |
| 52 | The Lord of the Rings: The Fellowship of the Ring | PeterJackson       | ElijahWood, IanMcKellen, OrlandoBloom, SeanBean   | 8.8         | Action,Adventure,Drama   | 178      | U                   | 2001        |
| 89 | Black Swan  | DarrenAronofsky    | NataliePortman, MilaKunis, VincentCassel, Wino... | 8.0         | Drama,Thriller           | 108      | A                   | 2010        |
| 91 | Harry Potter and the Deathly Hallows- Part 1      | DavidYates         | DanielRadcliffe, EmmaWatson, RupertGrint,         | 7.7         | Adventure,Family,Fantasy | 146      | UA                  | 2010        |

```
In [17]: import seaborn as sns
```



```
In [18]: sns.relplot(x='Duration',y='IMDb-Rating',data=imdb)
```

```
Out[18]: <seaborn.axisgrid.FacetGrid at 0x178cb39aac8>
```

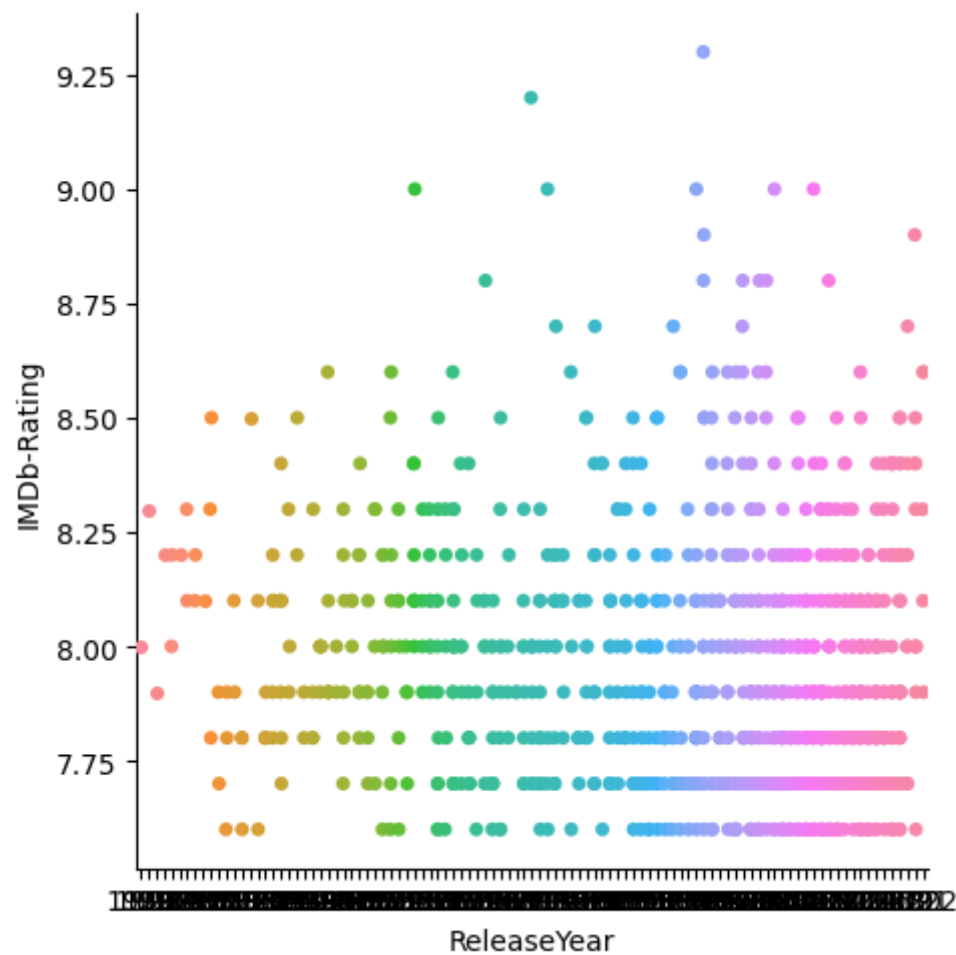


```
In [19]: sns.relplot(x='Duration',y='IMDb-Rating',hue='Director',data=imdb)
```

- FrankCapra
- JoelSchumacher
- RobertStevenson
- NadineLabaki
- EliaKazan
- RobertMulligan
- TonyBancroft,
- DonSiegel,
- NicholasMeyer
- RichMoore
- KevinReynolds
- JamesMarsh
- GabrieleMuccino
- Jee-woonKim
- OrsonWelles
- SamRaimi
- ...

```
In [20]: sns.catplot(x='ReleaseYear',y='IMDb-Rating',data=imdb,kind='strip')
```

```
Out[20]: <seaborn.axisgrid.FacetGrid at 0x178cea6ad08>
```





In [21]: imdb

Out[21]:

|     | Title                             | Director            | Stars   | IMDb-Rating | Category                | Duration | Censor-board-rating | ReleaseYear |
|-----|-----------------------------------|---------------------|---|-------------|-------------------------|----------|---------------------|-------------|
| 0   | Top Gun: Maverick                 | JosephKosinski      | TomCruise, JenniferConnelly, MilesTeller, Valk... | 8.6         | Action,Drama            | 130      | UA                  | 2022        |
| 1   | Everything Everywhere All at Once | DanKwan,            | , MichelleYeoh, StephanieHsu, KeHuyQuan, James... | 8.3         | Action,Adventure,Comedy | 139      | R                   | 2022        |
| 2   | The Batman                        | MattReeves          | RobertPattinson, ZoëKravitz, JeffreyWright, Co... | 7.9         | Action,Crime,Drama      | 176      | UA                  | 2022        |
| 3   | Jurassic Park                     | StevenSpielberg     | SamNeill, LauraDern, JeffGoldblum, RichardAtte... | 8.2         | Action,Adventure,Sci-Fi | 127      | UA                  | 1993        |
| 4   | The Godfather                     | FrancisFordCoppola  | MarlonBrando, AlPacino, JamesCaan, DianeKeaton    | 9.2         | Crime,Drama             | 175      | A                   | 1972        |
| ... | ...                               | ...                 | ...   | ...         | ...                     | ...      | ...                 | ...         |
| 995 | Vizontele                         | YilmazErdogan,      | , YilmazErdogan, DemetAkbag, AltanErkekli, Cem... | 8.0         | Comedy,Drama            | 110      | G                   | 2001        |
| 996 | Sarfaroosh                        | JohnMathewMatthan   | AamirKhan, NaseeruddinShah, SonaliBendre, Muke... | 8.1         | Action,Drama,Thriller   | 174      | A                   | 1999        |
| 997 | Udaan                             | VikramadityaMotwane | RajatBarmecha, RonitRoy, ManjotSingh, RamKapoor   | 8.1         | Drama                   | 134      | UA                  | 2010        |
| 998 | English Vinglish                  | GauriShinde         | Sridevi, AdilHussain, MehdiNebbou, PriyaAnand     | 7.8         | Comedy,Drama,Family     | 134      | U                   | 2012        |
| 999 | Anand                             | HrishikeshMukherjee | RajeshKhanna, AmitabhBachchan, SumitaSanyal, R... | 8.1         | Drama,Musical           | 122      | U                   | 1971        |

1000 rows × 8 columns

```
In [22]: from sklearn.preprocessing import LabelEncoder
L1 = LabelEncoder()
imdb_1["Title"]=L1.fit_transform(imdb_1["Title"])
imdb_1["Director"]=L1.fit_transform(imdb_1["Director"])
imdb_1["Stars"]=L1.fit_transform(imdb_1["Stars"])
imdb_1["Category"]=L1.fit_transform(imdb_1["Category"])
imdb_1["Censor-board-rating"]=L1.fit_transform(imdb_1["Censor-board-rating"])
```

C:\Users\Kajal\conda\envs\tensorflow\lib\site-packages\ipykernel\_launcher.py:3: 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/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) ([https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

This is separate from the ipykernel package so we can avoid doing imports until

C:\Users\Kajal\conda\envs\tensorflow\lib\site-packages\ipykernel\_launcher.py:4: 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/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) ([https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

after removing the cwd from sys.path.

C:\Users\Kajal\conda\envs\tensorflow\lib\site-packages\ipykernel\_launcher.py:5: 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/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) ([https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

"""

C:\Users\Kajal\conda\envs\tensorflow\lib\site-packages\ipykernel\_launcher.py:6: 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/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) ([https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\ipykernel_launcher.py:7: 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/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy) ([https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy))

```
import sys
```

In [23]: imdb\_1

Out[23]:

|            | Title | Director | Stars | IMDb-Rating | Category | Duration | Censor-board-rating | ReleaseYear |
|------------|-------|----------|-------|-------------|----------|----------|---------------------|-------------|
| <b>4</b>   | 34    | 19       | 34    | 9.2         | 19       | 175      | 4                   | 1972        |
| <b>13</b>  | 13    | 6        | 13    | 7.6         | 6        | 152      | 9                   | 2001        |
| <b>30</b>  | 18    | 7        | 29    | 8.8         | 1        | 148      | 10                  | 2010        |
| <b>52</b>  | 36    | 35       | 16    | 8.8         | 0        | 178      | 9                   | 2001        |
| <b>89</b>  | 5     | 10       | 38    | 8.0         | 27       | 108      | 4                   | 2010        |
| <b>91</b>  | 12    | 14       | 12    | 7.7         | 6        | 146      | 10                  | 2010        |
| <b>99</b>  | 31    | 30       | 28    | 8.2         | 30       | 138      | 4                   | 2010        |
| <b>134</b> | 1     | 41       | 41    | 8.2         | 12       | 135      | 10                  | 2001        |
| <b>143</b> | 28    | 20       | 14    | 8.6         | 10       | 125      | 9                   | 2001        |
| <b>149</b> | 27    | 43       | 19    | 7.7         | 20       | 116      | 10                  | 2001        |
| <b>179</b> | 10    | 37       | 21    | 8.0         | 24       | 113      | 10                  | 2001        |
| <b>190</b> | 33    | 13       | 33    | 7.8         | 2        | 116      | 10                  | 2010        |
| <b>206</b> | 39    | 11       | 23    | 7.8         | 12       | 120      | 10                  | 2010        |
| <b>209</b> | 24    | 12       | 37    | 7.9         | 25       | 147      | 10                  | 2001        |
| <b>222</b> | 4     | 39       | 25    | 7.7         | 5        | 144      | 4                   | 2001        |
| <b>236</b> | 47    | 2        | 32    | 7.7         | 21       | 106      | 4                   | 2001        |
| <b>262</b> | 23    | 34       | 27    | 8.1         | 9        | 92       | 9                   | 2001        |
| <b>264</b> | 30    | 3        | 3     | 7.9         | 9        | 90       | 9                   | 2001        |
| <b>270</b> | 19    | 31       | 8     | 7.6         | 3        | 117      | 10                  | 2010        |
| <b>280</b> | 41    | 4        | 15    | 7.7         | 4        | 122      | 4                   | 2001        |
| <b>295</b> | 37    | 1        | 39    | 7.6         | 29       | 101      | 0                   | 2001        |
| <b>300</b> | 22    | 22       | 9     | 8.3         | 18       | 122      | 9                   | 2001        |
| <b>308</b> | 15    | 15       | 0     | 8.1         | 7        | 98       | 9                   | 2010        |
| <b>324</b> | 8     | 36       | 4     | 7.6         | 9        | 95       | 9                   | 2010        |



|            | Title | Director | Stars | IMDb-Rating | Category | Duration | Censor-board-rating | ReleaseYear |
|------------|-------|----------|-------|-------------|----------|----------|---------------------|-------------|
| <b>326</b> | 32    | 32       | 2     | 7.7         | 9        | 100      | 9                   | 2010        |
| <b>337</b> | 40    | 29       | 46    | 8.3         | 9        | 103      | 9                   | 2010        |
| <b>368</b> | 17    | 16       | 30    | 8.3         | 23       | 131      | 2                   | 2010        |
| <b>390</b> | 43    | 17       | 1     | 7.6         | 28       | 110      | 1                   | 2010        |
| <b>391</b> | 35    | 45       | 11    | 8.0         | 13       | 118      | 9                   | 2010        |
| <b>408</b> | 38    | 47       | 18    | 7.6         | 14       | 110      | 4                   | 2001        |
| <b>426</b> | 3     | 23       | 26    | 7.8         | 4        | 144      | 6                   | 2010        |
| <b>536</b> | 0     | 9        | 22    | 7.6         | 12       | 94       | 10                  | 2010        |
| <b>608</b> | 16    | 25       | 42    | 7.6         | 21       | 132      | 9                   | 2001        |
| <b>637</b> | 11    | 40       | 31    | 7.7         | 16       | 90       | 7                   | 2010        |
| <b>707</b> | 2     | 24       | 48    | 7.7         | 4        | 119      | 3                   | 2010        |
| <b>712</b> | 46    | 38       | 17    | 7.7         | 11       | 99       | 8                   | 2001        |
| <b>737</b> | 14    | 26       | 24    | 7.7         | 15       | 95       | 8                   | 2001        |
| <b>787</b> | 21    | 5        | 6     | 8.1         | 22       | 224      | 9                   | 2001        |
| <b>816</b> | 25    | 28       | 43    | 7.9         | 26       | 165      | 10                  | 2010        |
| <b>843</b> | 6     | 0        | 20    | 7.8         | 8        | 115      | 9                   | 2001        |
| <b>850</b> | 7     | 33       | 36    | 7.7         | 27       | 120      | 8                   | 2001        |
| <b>856</b> | 48    | 21       | 44    | 8.0         | 16       | 137      | 5                   | 2001        |
| <b>863</b> | 20    | 44       | 45    | 7.7         | 27       | 106      | 5                   | 2010        |
| <b>907</b> | 9     | 18       | 7     | 8.1         | 16       | 183      | 11                  | 2001        |
| <b>920</b> | 29    | 42       | 35    | 7.8         | 11       | 87       | 7                   | 2001        |
| <b>938</b> | 42    | 27       | 47    | 8.0         | 4        | 115      | 5                   | 2010        |
| <b>954</b> | 26    | 8        | 10    | 7.9         | 17       | 98       | 8                   | 2001        |
| <b>995</b> | 45    | 48       | 5     | 8.0         | 14       | 110      | 5                   | 2001        |
| <b>997</b> | 44    | 46       | 40    | 8.1         | 21       | 134      | 10                  | 2010        |



```
In [24]: x = imdb_1.iloc[:, :-1].values  
x
```

```
Out[24]: array([[34, 19, 34, 9.2, 19, '175 ', 4],  
               [13, 6, 13, 7.6, 6, '152 ', 9],  
               [18, 7, 29, 8.8, 1, '148 ', 10],  
               [36, 35, 16, 8.8, 0, '178 ', 9],  
               [5, 10, 38, 8.0, 27, '108 ', 4],  
               [12, 14, 12, 7.7, 6, '146 ', 10],  
               [31, 30, 28, 8.2, 30, '138 ', 4],  
               [1, 41, 41, 8.2, 12, '135 ', 10],  
               [28, 20, 14, 8.6, 10, '125 ', 9],  
               [27, 43, 19, 7.7, 20, '116 ', 10],  
               [10, 37, 21, 8.0, 24, '113 ', 10],  
               [33, 13, 33, 7.8, 2, '116 ', 10],  
               [39, 11, 23, 7.8, 12, '120 ', 10],  
               [24, 12, 37, 7.9, 25, '147 ', 10],  
               [4, 39, 25, 7.7, 5, '144 ', 4],  
               [47, 2, 32, 7.7, 21, '106 ', 4],  
               [23, 34, 27, 8.1, 9, '92 ', 9],  
               [30, 3, 3, 7.9, 9, '90 ', 9],  
               [19, 31, 8, 7.6, 3, '117 ', 10],  
               [41, 4, 15, 7.7, 4, '122 ', 4],  
               [37, 1, 39, 7.6, 29, '101 ', 0],  
               [22, 22, 9, 8.3, 18, '122 ', 9],  
               [15, 15, 0, 8.1, 7, '98 ', 9],  
               [8, 36, 4, 7.6, 9, '95 ', 9],  
               [32, 32, 2, 7.7, 9, '100 ', 9],  
               [40, 29, 46, 8.3, 9, '103 ', 9],  
               [17, 16, 30, 8.3, 23, '131 ', 2],  
               [43, 17, 1, 7.6, 28, '110 ', 1],  
               [35, 45, 11, 8.0, 13, '118 ', 9],  
               [38, 47, 18, 7.6, 14, '110 ', 4],  
               [3, 23, 26, 7.8, 4, '144 ', 6],  
               [0, 9, 22, 7.6, 12, '94 ', 10],  
               [16, 25, 42, 7.6, 21, '132 ', 9],  
               [11, 40, 31, 7.7, 16, '90 ', 7],  
               [2, 24, 48, 7.7, 4, '119 ', 3],  
               [46, 38, 17, 7.7, 11, '99 ', 8],  
               [14, 26, 24, 7.7, 15, '95 ', 8],  
               [21, 5, 6, 8.1, 22, '224 ', 9],
```

```
[25, 28, 43, 7.9, 26, '165 ', 10],
[6, 0, 20, 7.8, 8, '115 ', 9],
[7, 33, 36, 7.7, 27, '120 ', 8],
[48, 21, 44, 8.0, 16, '137 ', 5],
[20, 44, 45, 7.7, 27, '106 ', 5],
[9, 18, 7, 8.1, 16, '183 ', 11],
[29, 42, 35, 7.8, 11, '87 ', 7],
[42, 27, 47, 8.0, 4, '115 ', 5],
[26, 8, 10, 7.9, 17, '98 ', 8],
[45, 48, 5, 8.0, 14, '110 ', 5],
[44, 46, 40, 8.1, 21, '134 ', 10]], dtype=object)
```

```
In [25]: y = imdb_1.iloc[:, -1].values
y
```

```
Out[25]: array([1972, 2001, 2010, 2001, 2010, 2010, 2010, 2001, 2001, 2001, 2001,
        2010, 2010, 2001, 2001, 2001, 2001, 2001, 2010, 2001, 2001, 2001,
        2010, 2010, 2010, 2010, 2010, 2010, 2010, 2001, 2010, 2010, 2001,
        2010, 2010, 2001, 2001, 2001, 2010, 2001, 2001, 2001, 2010, 2001,
        2001, 2010, 2001, 2001, 2010], dtype=int64)
```

```
In [26]: from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y, test_size=0.2, random_state=0)
```

```
In [27]: print(x_train.shape,x_test.shape,y_train.shape,y_test.shape)

(39, 7) (10, 7) (39,) (10,)
```

```
In [28]: imdb_1["ReleaseYear"].value_counts()
```

```
Out[28]: 2001    26
        2010    22
        1972     1
        Name: ReleaseYear, dtype: int64
```

## Support vector machine (Kernel = rbf )

```
In [29]: from sklearn.svm import SVC
classifier = SVC(kernel='rbf', random_state=10)
classifier.fit(x_train, y_train)
```

```
Out[29]: SVC(random_state=10)
```

```
In [30]: y_pred = classifier.predict(x_test)
y_pred
```

```
Out[30]: array([2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001],
              dtype=int64)
```

```
In [31]: from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
ac = accuracy_score(y_pred, y_test)*100
cr = classification_report(y_pred, y_test)
cm = confusion_matrix(y_pred, y_test)
```

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\_classification.py:1318: UndefinedMetricWarning: Recall and F-score are ill-defined and being set to 0.0 in labels with no true samples. Use `zero_division` parameter to control this behavior.
```

```
_warn_prf(average, modifier, msg_start, len(result))
```

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\_classification.py:1318: UndefinedMetricWarning: Recall and F-score are ill-defined and being set to 0.0 in labels with no true samples. Use `zero_division` parameter to control this behavior.
```

```
_warn_prf(average, modifier, msg_start, len(result))
```

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\_classification.py:1318: UndefinedMetricWarning: Recall and F-score are ill-defined and being set to 0.0 in labels with no true samples. Use `zero_division` parameter to control this behavior.
```

```
_warn_prf(average, modifier, msg_start, len(result))
```

## accuracy = 60

```
In [32]: print(ac)
```

```
60.0
```

In [33]: `print(cr)`

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 2001         | 1.00      | 0.60   | 0.75     | 10      |
| 2010         | 0.00      | 0.00   | 0.00     | 0       |
| accuracy     |           |        | 0.60     | 10      |
| macro avg    | 0.50      | 0.30   | 0.37     | 10      |
| weighted avg | 1.00      | 0.60   | 0.75     | 10      |

In [34]: `print(cm)`

```
[[6 4]
 [0 0]]
```

## (kernel = linear)

In [35]: `from sklearn.svm import SVC`  
`classifier_1 = SVC(kernel='linear', random_state=10)`  
`classifier_1.fit(x_train, y_train)`

Out[35]: SVC(kernel='linear', random\_state=10)

In [36]: `y_pred1 = classifier_1.predict(x_test)`  
`y_pred1`

Out[36]: array([2001, 2001, 2001, 2010, 2010, 1972, 2010, 2010, 2010, 2010],  
dtype=int64)

```
In [37]: from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
ac_1 = accuracy_score(y_pred1, y_test) * 100
cr_1 = classification_report(y_pred1, y_test)
cm_1 = confusion_matrix(y_pred1, y_test)
```

C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\\_classification.py:1318: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

```
_warn_prf(average, modifier, msg_start, len(result))
```

C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\\_classification.py:1318: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

```
_warn_prf(average, modifier, msg_start, len(result))
```

C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\\_classification.py:1318: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

```
_warn_prf(average, modifier, msg_start, len(result))
```

## Accuracy = 30

```
In [38]: print(ac_1)
```

30.0

```
In [39]: print(cr_1)
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 1972         | 0.00      | 0.00   | 0.00     | 1       |
| 2001         | 0.17      | 0.33   | 0.22     | 3       |
| 2010         | 0.50      | 0.33   | 0.40     | 6       |
| accuracy     |           |        | 0.30     | 10      |
| macro avg    | 0.22      | 0.22   | 0.21     | 10      |
| weighted avg | 0.35      | 0.30   | 0.31     | 10      |

```
In [40]: print(cm_1)
```

```
[[0 1 0]
 [0 1 2]
 [0 4 2]]
```

## (Kernel = Sigmoid)

```
In [41]: from sklearn.svm import SVC
classifier_2 = SVC(kernel='sigmoid', random_state=10)
classifier_2.fit(x_train, y_train)
```

```
Out[41]: SVC(kernel='sigmoid', random_state=10)
```

```
In [42]: y_pred2 = classifier_2.predict(x_test)
y_pred2
```

```
Out[42]: array([2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001, 2001],
              dtype=int64)
```

```
In [43]: from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
ac_2 = accuracy_score(y_pred2, y_test)*100
cr_2 = classification_report(y_pred2, y_test)
cm_2 = confusion_matrix(y_pred2, y_test)
```

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\_classification.py:1318: UndefinedMetricWarnin
g: Recall and F-score are ill-defined and being set to 0.0 in labels with no true samples. Use `zero_division` paramete
r to control this behavior.
```

```
    _warn_prf(average, modifier, msg_start, len(result))
```

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\_classification.py:1318: UndefinedMetricWarnin
g: Recall and F-score are ill-defined and being set to 0.0 in labels with no true samples. Use `zero_division` paramete
r to control this behavior.
```

```
    _warn_prf(average, modifier, msg_start, len(result))
```

```
C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\_classification.py:1318: UndefinedMetricWarnin
g: Recall and F-score are ill-defined and being set to 0.0 in labels with no true samples. Use `zero_division` paramete
r to control this behavior.
```

```
    _warn_prf(average, modifier, msg_start, len(result))
```



## Accuracy = 60

In [44]: `print(ac_2)`

60.0

In [45]: `print(cr_2)`

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 2001         | 1.00      | 0.60   | 0.75     | 10      |
| 2010         | 0.00      | 0.00   | 0.00     | 0       |
| accuracy     |           |        | 0.60     | 10      |
| macro avg    | 0.50      | 0.30   | 0.37     | 10      |
| weighted avg | 1.00      | 0.60   | 0.75     | 10      |

In [46]: `print(cm_2)`

```
[[ 6  4]
 [ 0  0]]
```

## (Kernel = poly)

In [47]: `from sklearn.svm import SVC`  
`classifier_3 = SVC(kernel='poly',random_state=10)`  
`classifier_3.fit(x_train,y_train)`

Out[47]: SVC(kernel='poly', random\_state=10)

In [48]: `y_pred3 = classifier_3.predict(x_test)`  
`y_pred3`

Out[48]: array([2001, 2010, 2001, 2010, 2010, 1972, 2010, 2010, 2010, 2010],  
dtype=int64)

```
In [49]: from sklearn.metrics import accuracy_score, classification_report, confusion_matrix
ac_3 = accuracy_score(y_pred3, y_test) * 100
cr_3 = classification_report(y_pred3, y_test)
cm_3 = confusion_matrix(y_pred3, y_test)
```

C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\\_classification.py:1318: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

```
_warn_prf(average, modifier, msg_start, len(result))
```

C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\\_classification.py:1318: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

```
_warn_prf(average, modifier, msg_start, len(result))
```

C:\Users\Kajal\.conda\envs\tensorflow\lib\site-packages\sklearn\metrics\\_classification.py:1318: UndefinedMetricWarning: Precision and F-score are ill-defined and being set to 0.0 in labels with no predicted samples. Use `zero\_division` parameter to control this behavior.

```
_warn_prf(average, modifier, msg_start, len(result))
```

## Accuracy = 40

```
In [50]: print(ac_3)
```

40.0

```
In [51]: print(cr_3)
```

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 1972         | 0.00      | 0.00   | 0.00     | 1       |
| 2001         | 0.17      | 0.50   | 0.25     | 2       |
| 2010         | 0.75      | 0.43   | 0.55     | 7       |
| accuracy     |           |        | 0.40     | 10      |
| macro avg    | 0.31      | 0.31   | 0.27     | 10      |
| weighted avg | 0.56      | 0.40   | 0.43     | 10      |

In [52]: `print(cm_3)`

```
[[0 1 0]
 [0 1 1]
 [0 4 3]]
```

In [ ]: