

IMDB Movie Data Analysis In Python

In []: `###`

In [1]: `#importing Librarry`

In [1]: `import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt

print(" Above library is loaded")`

Above library is loaded

In [3]: `#Load the IMDB-Movie-Data.csv Dataset`

In [2]: `#Data_set/IMDB-Movie-Data.csv
imdb_df = pd.read_csv("D:\Vivek_Stuff\Learning Stuff\Python_Work\Data_set\I
MDB-Movie-Data.csv")

print("IMDB Dataset ready to use !!")`

IMDB Dataset ready to use !!

1 . Show the Top 10 Rows data

In [3]: `imdb_df.head(10) # showing the first 10 rows of data`
`# by default head() returns the 5 rows of data`

Out[3]:

	Rank	Title	Genre	Description	Director	Actors
0	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced ...	James Gunn	Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S...
1	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of mankind, a te...	Ridley Scott	Noomi Rapace, Logan Marshall-Green, Michael Fa...
2	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diag...	M. Night Shyamalan	James McAvoy, Anya Taylor-Joy, Haley Lu Richar...
3	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling thea...	Christophe Lourdelet	Matthew McConaughey,Reese Witherspoon, Seth Ma...
4	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of th...	David Ayer	Will Smith, Jared Leto, Margot Robbie, Viola D...
5	6	The Great Wall	Action,Adventure,Fantasy	European mercenaries searching for black powde...	Yimou Zhang	Matt Damon, Tian Jing, Willem Dafoe, Andy Lau
6	7	La La Land	Comedy,Drama,Music	A jazz pianist falls for an aspiring actress i...	Damien Chazelle	Ryan Gosling, Emma Stone, Rosemarie DeWitt, J....
7	8	Mindhorn	Comedy	A has-been actor best known for playing the ti...	Sean Foley	Essie Davis, Andrea Riseborough, Julian Barrat...
8	9	The Lost City of Z	Action,Adventure,Biography	A true-life drama, centering on British explor...	James Gray	Charlie Hunnam, Robert Pattinson, Sienna Mille...
9	10	Passengers	Adventure,Drama,Romance	A spacecraft traveling to a distant colony pla...	Morten Tyldum	Jennifer Lawrence, Chris Pratt, Michael Sheen,...

2 . Show the last 10 Rows data¶

In [4]: `imdb_df.tail(10) # showing the first 10 rows of dataset`
by default tail() returns the last 5 rows of data

Out[4]:

	Rank	Title	Genre	Description	Director	Actors	Year
990	991	Underworld: Rise of the Lycans	Action,Adventure,Fantasy	An origins story centered on the centuries-old...	Patrick Tatopoulos	Rhona Mitra, Michael Sheen, Bill Nighy, Steven...	200
991	992	Taare Zameen Par	Drama,Family,Music	An eight-year-old boy is thought to be a lazy ...	Aamir Khan	Darsheel Safary, Aamir Khan, Tanay Chheda, Sac...	200
992	993	Take Me Home Tonight	Comedy,Drama,Romance	Four years after graduation, an awkward high s...	Michael Dowse	Topher Grace, Anna Faris, Dan Fogler, Teresa P...	201
993	994	Resident Evil: Afterlife	Action,Adventure,Horror	While still out to destroy the evil Umbrella C...	Paul W.S. Anderson	Milla Jovovich, Ali Larter, Wentworth Miller,K...	201
994	995	Project X	Comedy	3 high school seniors throw a birthday party t...	Nima Nourizadeh	Thomas Mann, Oliver Cooper, Jonathan Daniel Br...	201
995	996	Secret in Their Eyes	Crime,Drama,Mystery	A tight-knit team of rising investigators, alo...	Billy Ray	Chiwetel Ejiofor, Nicole Kidman, Julia Roberts...	201
996	997	Hostel: Part II	Horror	Three American college students studying abroa...	Eli Roth	Lauren German, Heather Matarazzo, Bijou Philli...	200
997	998	Step Up 2: The Streets	Drama,Music,Romance	Romantic sparks occur between two dance studen...	Jon M. Chu	Robert Hoffman, Briana Evigan, Cassie Ventura,...	200
998	999	Search Party	Adventure,Comedy	A pair of friends embark on a mission to reuni...	Scot Armstrong	Adam Pally, T.J. Miller, Thomas Middleditch,Sh...	201
999	1000	Nine Lives	Comedy,Family,Fantasy	A stuffy businessman finds himself trapped ins...	Barry Sonnenfeld	Kevin Spacey, Jennifer Garner, Robbie Amell,Ch...	201

3 . Show Total No. of Rows & Column of data set

```
In [7]: imdb_df.shape
# shape is an attribute of dataframe, Its not a method
#shape retruen the Size of data set , No. of Rows & No . Of Column of DataF
rame
```

```
Out[7]: (1000, 12)
```

```
In [8]: print(' No of rows : ', imdb_df.shape[0])
print(' No of column : ', imdb_df.shape[1])
```

```
No of rows : 1000
No of column : 12
```

So you can see that, Dataset has 1000 rows and 12 Column

4. Show complete data information,

```
In [9]: imdb_df.info()
#info() retruns the complete info of your dataset , Column Name and its dat
a type , Null Inforamtion, Total rows and many more...
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 12 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Rank                  1000 non-null   int64
1   Title                 1000 non-null   object
2   Genre                 1000 non-null   object
3   Description            1000 non-null   object
4   Director              1000 non-null   object
5   Actors                1000 non-null   object
6   Year                  1000 non-null   int64
7   Runtime (Minutes)     1000 non-null   int64
8   Rating                1000 non-null   float64
9   Votes                 1000 non-null   int64
10  Revenue (Millions)    872 non-null    float64
11  Metascore              936 non-null    float64
dtypes: float64(3), int64(4), object(5)
memory usage: 93.9+ KB
```

5 . Show the Missing data value if any in your dataset

```
In [10]: imdb_df.isnull()
#
```

Out[10]:

	Rank	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)
0	False	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False	False
...
995	False	False	False	False	False	False	False	False	False	False	False
996	False	False	False	False	False	False	False	False	False	False	False
997	False	False	False	False	False	False	False	False	False	False	False
998	False	False	False	False	False	False	False	False	False	False	False
999	False	False	False	False	False	False	False	False	False	False	False

1000 rows × 12 columns



```
In [11]: imdb_df.isnull().sum()
# So below you can see that Revenue & Metascore column contains the null values
```

```
Out[11]: Rank          0
Title          0
Genre          0
Description     0
Director       0
Actors         0
Year          0
Runtime (Minutes)  0
Rating         0
Votes         0
Revenue (Millions) 128
Metascore      64
dtype: int64
```

6. Drop all the Null containing rows

```
In [12]: imdb_df.dropna(axis=0)
# dropna() is used to drop missing value if any
# we are apassing the axis = 0 , Means We are wissing to drop rows , So Axis =0 meanns Rows, Axis =1 means Column
# you can use inplace=True to delete the rows permanently from your datase t, But we dont suggest to use .
```

Out[12]:

	Rank	Title	Genre	Description	Director	Actors
0	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced ...	James Gunn	Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S...
1	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of mankind, a te...	Ridley Scott	Noomi Rapace, Logan Marshall-Green, Michael Fa...
2	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diag...	M. Night Shyamalan	James McAvoy, Anya Taylor-Joy, Haley Lu Richar...
3	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling thea...	Christophe Lourdelet	Matthew McConaughey,Reese Witherspoon, Seth Ma...
4	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of th...	David Ayer	Will Smith, Jared Leto, Margot Robbie, Viola D...
...
993	994	Resident Evil: Afterlife	Action,Adventure,Horror	While still out to destroy the evil Umbrella C...	Paul W.S. Anderson	Milla Jovovich, Ali Larter, Wentworth Miller,K...
994	995	Project X	Comedy	3 high school seniors throw a birthday party t...	Nima Nourizadeh	Thomas Mann, Oliver Cooper, Jonathan Daniel Br...
996	997	Hostel: Part II	Horror	Three American college students studying abroa...	Eli Roth	Lauren German, Heather Matarazzo, Bijou Philli...
997	998	Step Up 2: The Streets	Drama,Music,Romance	Romantic sparks occur between two dance studen...	Jon M. Chu	Robert Hoffman, Briana Evigan, Cassie Ventura,...
999	1000	Nine Lives	Comedy,Family,Fantasy	A stuffy businessman finds himself trapped ins...	Barry Sonnenfeld	Kevin Spacey, Jennifer Garner, Robbie Amell,Ch...

838 rows × 12 columns

In [13]: `imdb_df.shape`

Out[13]: (1000, 12)

7 . Check Duplicate Data

```
In [14]: dup_imdb_df= imdb_df.duplicated().any()
# dup_imdb_df is a variable, we created here
# duplicated() returns the duplicate value from your dataset
# imdb_df.duplicated().any() will return the true/false if any duplicate value found in dataset

dup_imdb_df # printing the result
```

Out[14]: False

so you can see that, Our dataset has no any duplicate value

8. Get overall Statistics of your dataset

```
In [15]: imdb_df.describe()
# describe() returns the basic statistics information on your numerical Column data.
```

Out[15]:

	Rank	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
count	1000.000000	1000.000000	1000.000000	1000.000000	1.000000e+03	872.000000	936.000000
mean	500.500000	2012.783000	113.172000	6.723200	1.698083e+05	82.956376	58.980000
std	288.819436	3.205962	18.810908	0.945429	1.887626e+05	103.253540	17.190000
min	1.000000	2006.000000	66.000000	1.900000	6.100000e+01	0.000000	11.000000
25%	250.750000	2010.000000	100.000000	6.200000	3.630900e+04	13.270000	47.000000
50%	500.500000	2014.000000	111.000000	6.800000	1.107990e+05	47.985000	59.500000
75%	750.250000	2016.000000	123.000000	7.400000	2.399098e+05	113.715000	72.000000
max	1000.000000	2016.000000	191.000000	9.000000	1.791916e+06	936.630000	100.000000

9. Show the Movie Title having Runtime > 180 Minutes

```
In [16]: ### showing the Columns Of data Frame
imdb_df.columns
```

Out[16]: Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'], dtype='object')


```
In [17]: imdb_df['Runtime (Minutes)'] >= 180
# So here we are quering the equation .
```

```
Out[17]: 0      False
1      False
2      False
3      False
4      False
...
995    False
996    False
997    False
998    False
999    False
Name: Runtime (Minutes), Length: 1000, dtype: bool
```

Showing the data whos runtime is >= 180 Minute

```
In [18]: ## Showing the data whos runtime is >= 180 Minute
imdb_df[imdb_df['Runtime (Minutes)'] >= 180 ]
```

Out[18]:

	Rank	Title	Genre	Description	Director	Actors	Year	(
82	83	The Wolf of Wall Street	Biography,Comedy,Crime	Based on the true story of Jordan Belfort, fro...	Martin Scorsese	Leonardo DiCaprio, Jonah Hill, Margot Robbie,M...	2013	
88	89	The Hateful Eight	Crime,Drama,Mystery	In the dead of a Wyoming winter, a bounty hunt...	Quentin Tarantino	Samuel L. Jackson, Kurt Russell, Jennifer Jaso...	2015	
311	312	La vie d'Adèle	Drama,Romance	Adèle's life is changed when she meets Emma, a...	Abdellatif Kechiche	Léa Seydoux, Adèle Exarchopoulos, Salim Kechio...	2013	
828	829	Grindhouse	Action,Horror,Thriller	Quentin Tarantino and Robert Rodriguez's homag...	Robert Rodriguez	Kurt Russell, Rose McGowan, Danny Trejo, Zoë Bell	2007	
965	966	Inland Empire	Drama,Mystery,Thriller	As an actress starts to adopt the persona of h...	David Lynch	Laura Dern, Jeremy Irons, Justin Theroux, Karo...	2006	

Showing the Title whos runtime is >= 180 Minute

```
In [19]: ## Showing the Title whos runtime is >= 180 Minute
imdb_df[imdb_df['Runtime (Minutes)'] >= 180]['Title']
```

```
Out[19]: 82      The Wolf of Wall Street
88      The Hateful Eight
311     La vie d'Adèle
828     Grindhouse
965     Inland Empire
Name: Title, dtype: object
```

10. Display the Years wise Highest Average Voting

```
In [20]: imdb_df.groupby('Year')['Votes'].mean()
```

```
Out[20]: Year
2006    269289.954545
2007    244331.037736
2008    275505.384615
2009    255780.647059
2010    252782.316667
2011    240790.301587
2012    285226.093750
2013    219049.648352
2014    203930.224490
2015    115726.220472
2016     48591.754209
Name: Votes, dtype: float64
```

So above we shown the year wise Mean() / avgerage Voting

let's sort them By Year in descending Highest Voting

```
In [21]: imdb_df.groupby('Year')['Votes'].mean().sort_values(ascending = False)
```

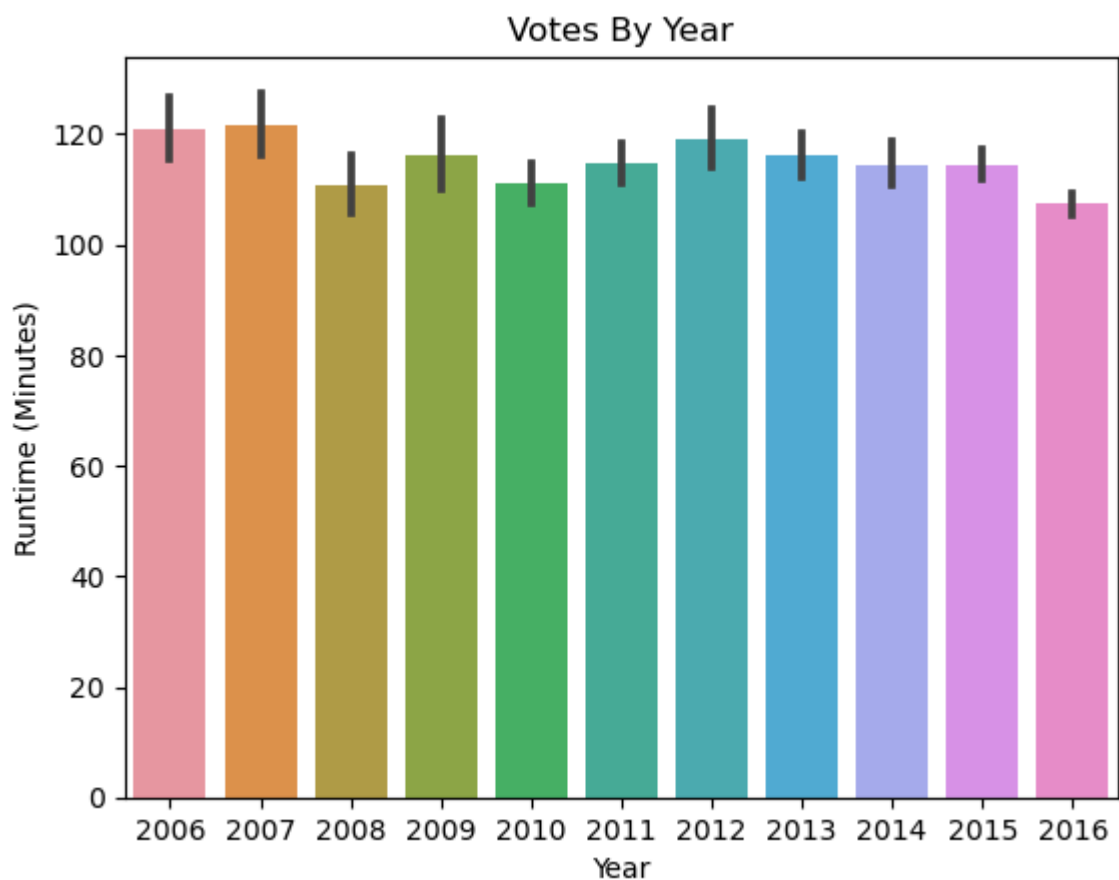
```
Out[21]: Year
2012    285226.093750
2008    275505.384615
2006    269289.954545
2009    255780.647059
2010    252782.316667
2007    244331.037736
2011    240790.301587
2013    219049.648352
2014    203930.224490
2015    115726.220472
2016     48591.754209
Name: Votes, dtype: float64
```

So you can see that Highest voting By Year

Show the graph : Year wise Voting

```
In [22]: sns.barplot(x='Year', y = 'Runtime (Minutes)', data= imdb_df)
plt.title('Votes By Year')
# sns is an alias of Seaborn , and we are using barplot() to display the graph in Bar

plt.show()
```



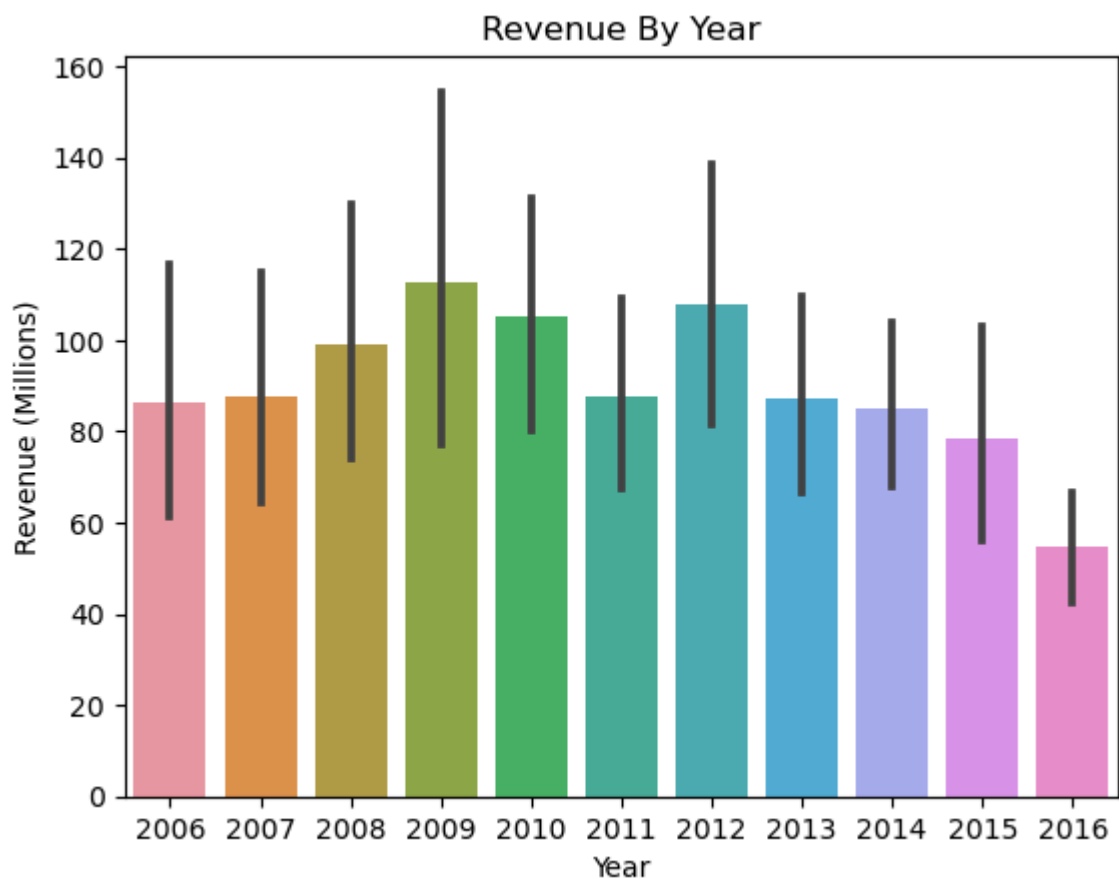
11 . Display highest Average Revenue by Year

```
In [23]: imdb_df.groupby('Year')['Revenue (Millions)'].mean().sort_values(ascending
= False)
```

```
Out[23]: Year
2009      112.601277
2012      107.973281
2010      105.081579
2008       99.082745
2007       87.882245
2011       87.612258
2013       87.121818
2006       86.296667
2014       85.078723
2015       78.355044
2016       54.690976
Name: Revenue (Millions), dtype: float64
```

```
In [24]: sns.barplot(x='Year', y = 'Revenue (Millions)', data= imdb_df)
plt.title('Revenue By Year')
# sns is an alias of Seaborn , and we are wusing barplot() to display the g
raph in Bar

plt.show()
```



12.Find the Average rating for Reach Ditector

In [25]: `imdb_df.columns`

Out[25]: Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year',
 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)',
 'Metascore'],
 dtype='object')

In [26]: `imdb_df.groupby('Director')['Rating'].mean().sort_values(ascending = False)`

Out[26]: Director
 Nitesh Tiwari 8.80
 Christopher Nolan 8.68
 Olivier Nakache 8.60
 Makoto Shinkai 8.60
 Aamir Khan 8.50
 ...
 Micheal Bafaro 3.50
 Jonathan Holbrook 3.20
 Shawn Burkett 2.70
 James Wong 2.70
 Jason Friedberg 1.90
 Name: Rating, Length: 644, dtype: float64

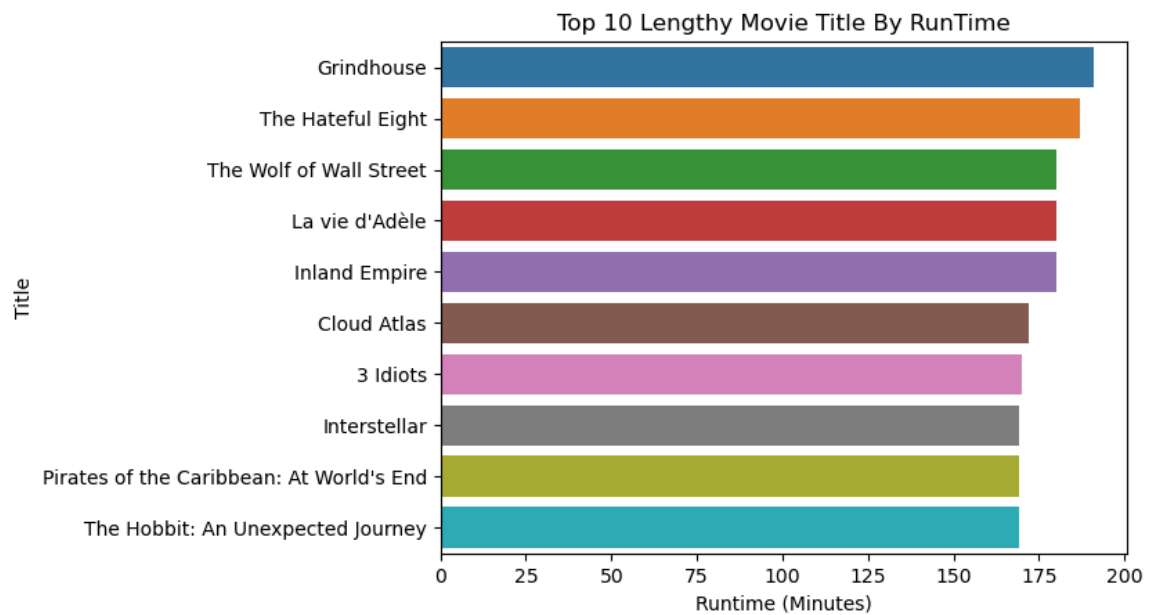
13 . Display Top 10 Tengthy MMovie Title & Runtime

In [27]: `top10 = imdb_df.nlargest(10, 'Runtime (Minutes)')[['Title', 'Runtime (Minutes)']].set_index('Title')`
 top10

Out[27]:

	Runtime (Minutes)
Title	
Grindhouse	191
The Hateful Eight	187
The Wolf of Wall Street	180
La vie d'Adèle	180
Inland Empire	180
Cloud Atlas	172
3 Idiots	170
Interstellar	169
Pirates of the Caribbean: At World's End	169
The Hobbit: An Unexpected Journey	169

```
In [28]: sns.barplot(x='Runtime (Minutes)' , y =top10.index , data= top10)
plt.title("Top 10 Lengthy Movie Title By RunTime")
plt.show()
```

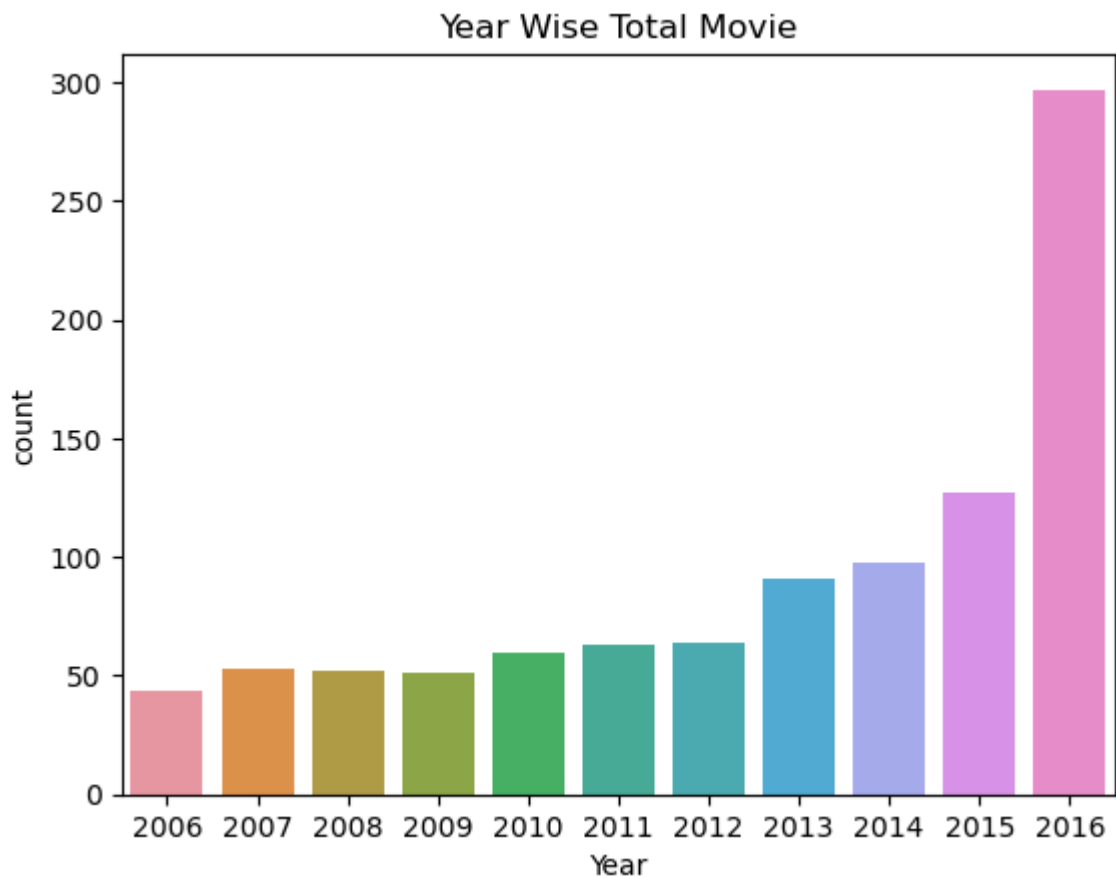


14 . Display Total Movie Count By Year

```
In [29]: imdb_df['Year'].value_counts()
# Value_counts() as works as CountIf in Excel, It count the Numericle value only
```

```
Out[29]: Year
2016      297
2015      127
2014       98
2013       91
2012       64
2011       63
2010       60
2007       53
2008       52
2009       51
2006       44
Name: count, dtype: int64
```

```
In [30]: #sns.barplot(x='Year' , y=imdb_df['Year'].value_counts(), data= imdb_df)
sns.countplot(x='Year' , data= imdb_df)
plt.title("Year Wise Total Movie")
plt.show()
```



15 . Most Popular Movie By Highest Revenue

```
In [31]: imdb_df['Revenue (Millions)'].max()
# so below you can see the highest revenue
```

Out[31]: 936.63

```
In [32]: # Let's show the data
imdb_df[imdb_df['Revenue (Millions)'].max()== imdb_df['Revenue (Million
s)']]
```

Out[32]:

	Rank	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)
50	51	Star Wars: Episode VII - The Force Awakens	Action,Adventure,Fantasy	Three decades after the defeat of the Galactic...	J.J. Abrams	Daisy Ridley, John Boyega, Oscar Isaac, Domhna...	2015	136

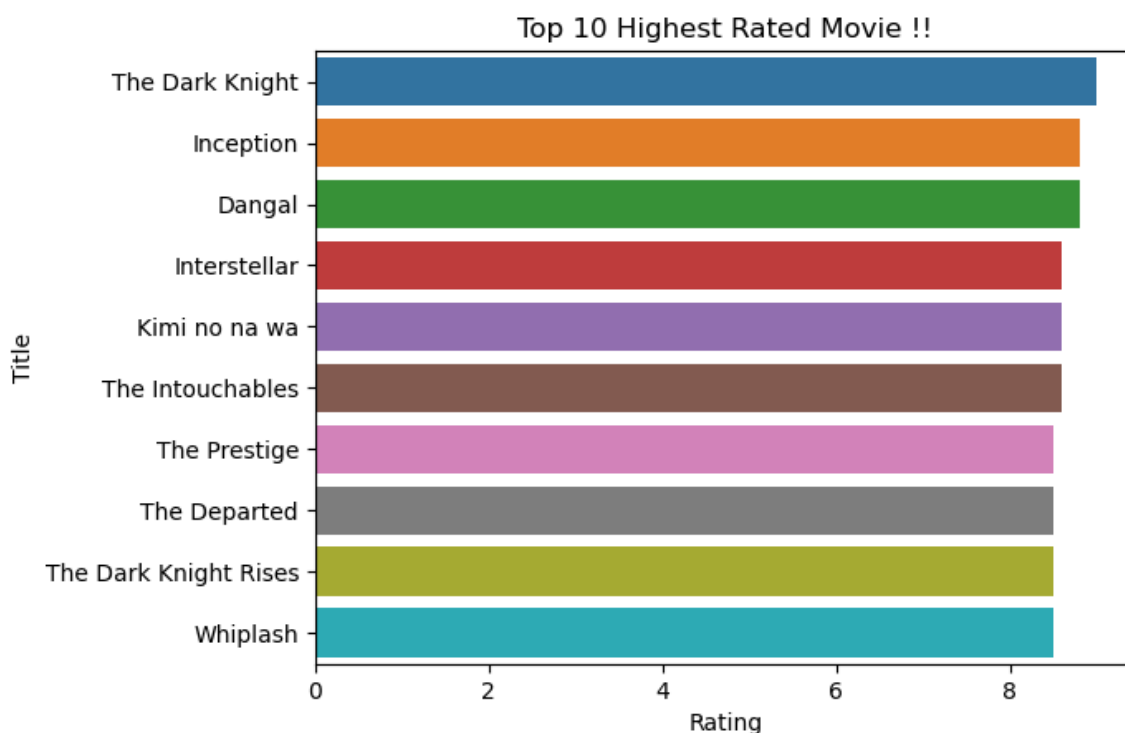
16. Display Top 10 Highest Rated Movie Titel & its Directors

```
In [33]: top10_movie_rating = imdb_df.nlargest(10, 'Rating')[['Title', 'Director', 'Rating']].set_index('Title')
top10_movie_rating
```

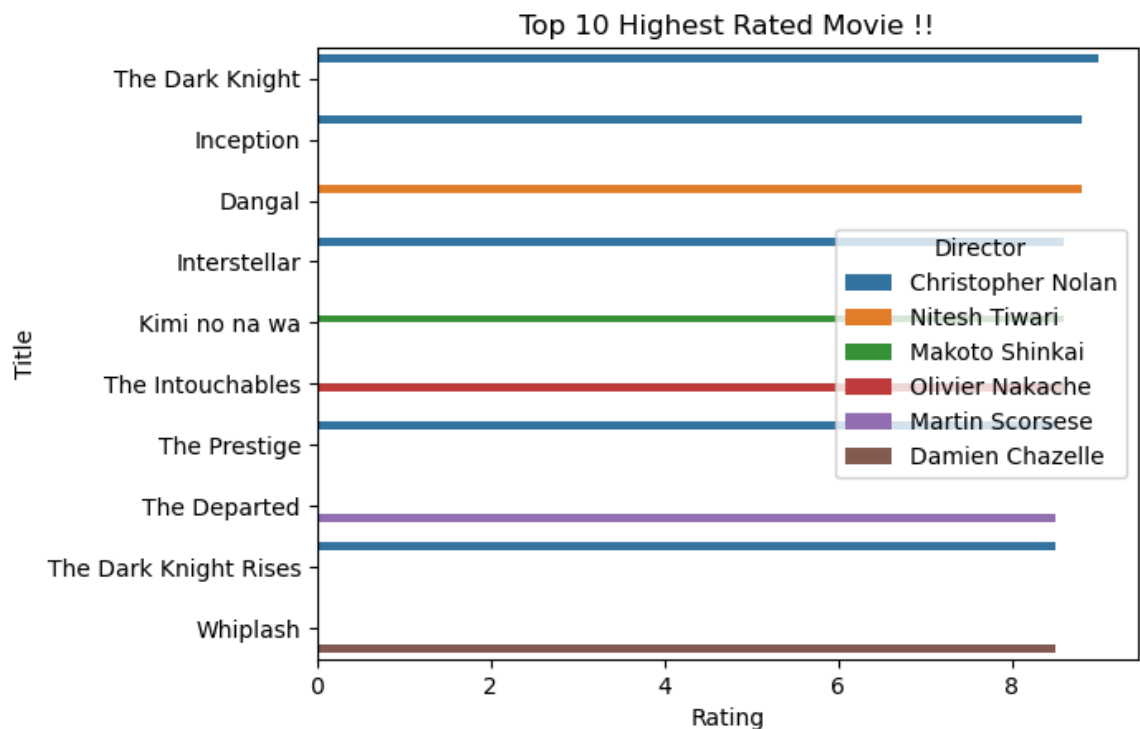
Out[33]:

	Director	Rating
Title		
The Dark Knight	Christopher Nolan	9.0
Inception	Christopher Nolan	8.8
Dangal	Nitesh Tiwari	8.8
Interstellar	Christopher Nolan	8.6
Kimi no na wa	Makoto Shinkai	8.6
The Intouchables	Olivier Nakache	8.6
The Prestige	Christopher Nolan	8.5
The Departed	Martin Scorsese	8.5
The Dark Knight Rises	Christopher Nolan	8.5
Whiplash	Damien Chazelle	8.5

```
In [36]: sns.barplot(x = 'Rating', y = top10_movie_rating.index, data = top10_movie_rating)
plt.title('Top 10 Highest Rated Movie !!')
plt.show()
```

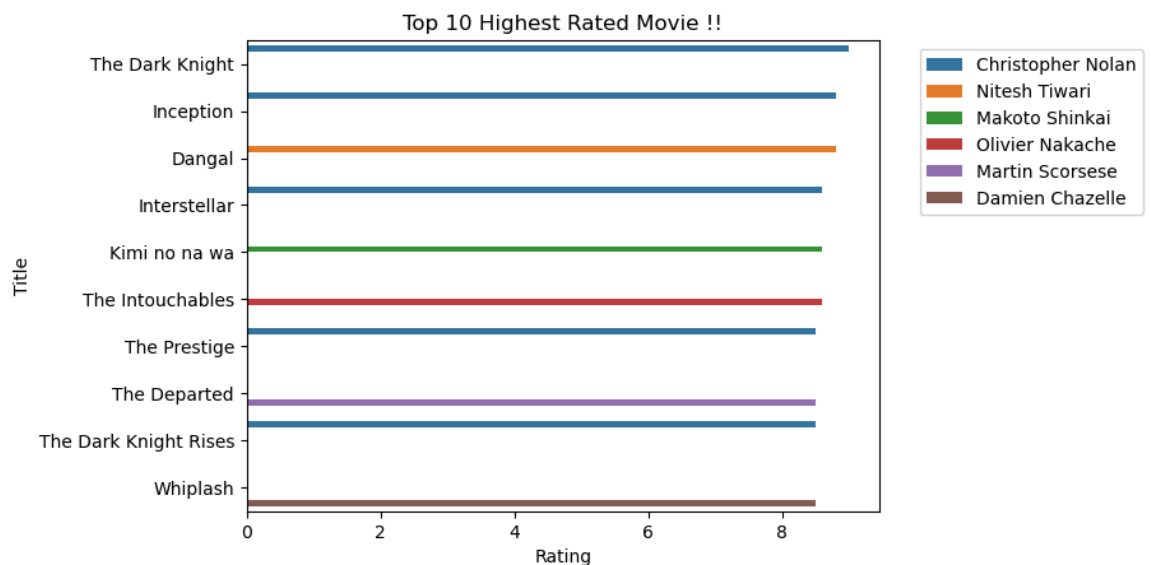



```
In [37]: sns.barplot(x = 'Rating', y = top10_movie_rating.index, data = top10_movie_
rating, hue= 'Director')
plt.title('Top 10 Highest Rated Movie !!')
plt.show()
```



```
In [39]: #Show Legend out of the graph

sns.barplot(x = 'Rating', y = top10_movie_rating.index, data = top10_movie_
rating, hue= 'Director')
plt.title('Top 10 Highest Rated Movie !!')
plt.legend(bbox_to_anchor= (1.05,1), loc=2)
plt.show()
```



17 Top 10 Highest Revenue Movie

```
In [42]: imdb_df.nlargest(10, 'Revenue (Millions)')['Title']
```

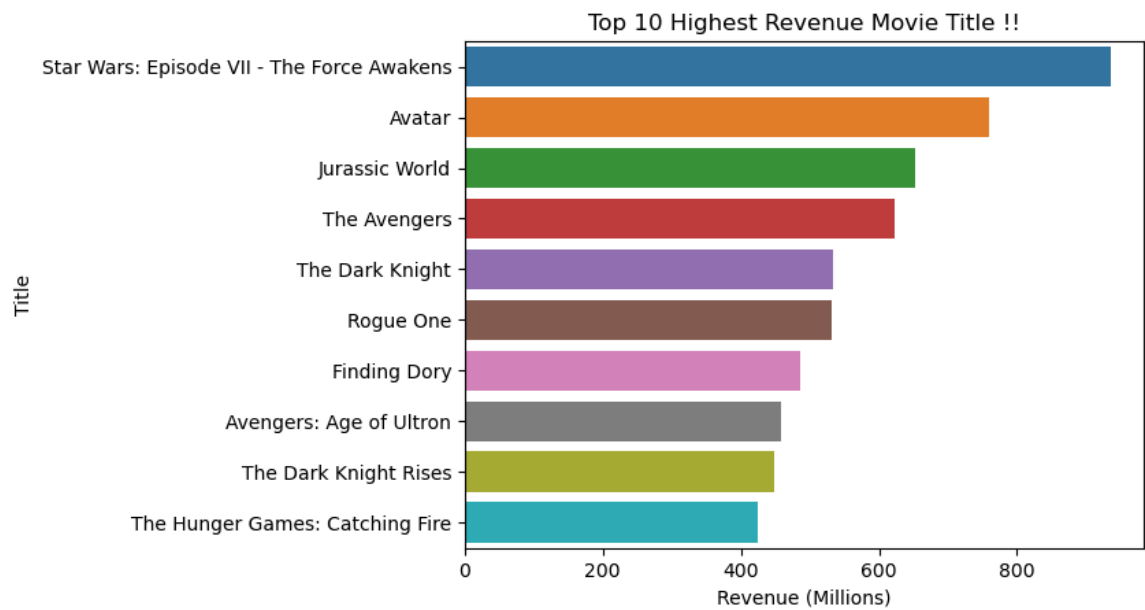
```
Out[42]: 50      Star Wars: Episode VII - The Force Awakens
87              Avatar
85              Jurassic World
76              The Avengers
54      The Dark Knight
12              Rogue One
119             Finding Dory
94      Avengers: Age of Ultron
124      The Dark Knight Rises
578      The Hunger Games: Catching Fire
Name: Title, dtype: object
```

```
In [45]: top10_revenue =imdb_df.nlargest(10, 'Revenue (Millions)')[['Title', 'Revenue
(Millions)']].set_index('Title')
top10_revenue
```

```
Out[45]:
```

	Revenue (Millions)
Title	
Star Wars: Episode VII - The Force Awakens	936.63
Avatar	760.51
Jurassic World	652.18
The Avengers	623.28
The Dark Knight	533.32
Rogue One	532.17
Finding Dory	486.29
Avengers: Age of Ultron	458.99
The Dark Knight Rises	448.13
The Hunger Games: Catching Fire	424.65

```
In [46]: sns.barplot(x = 'Revenue (Millions)', y = top10_revenue.index, data = top10_revenue)
plt.title('Top 10 Highest Revenue Movie Title !!')
plt.show()
```



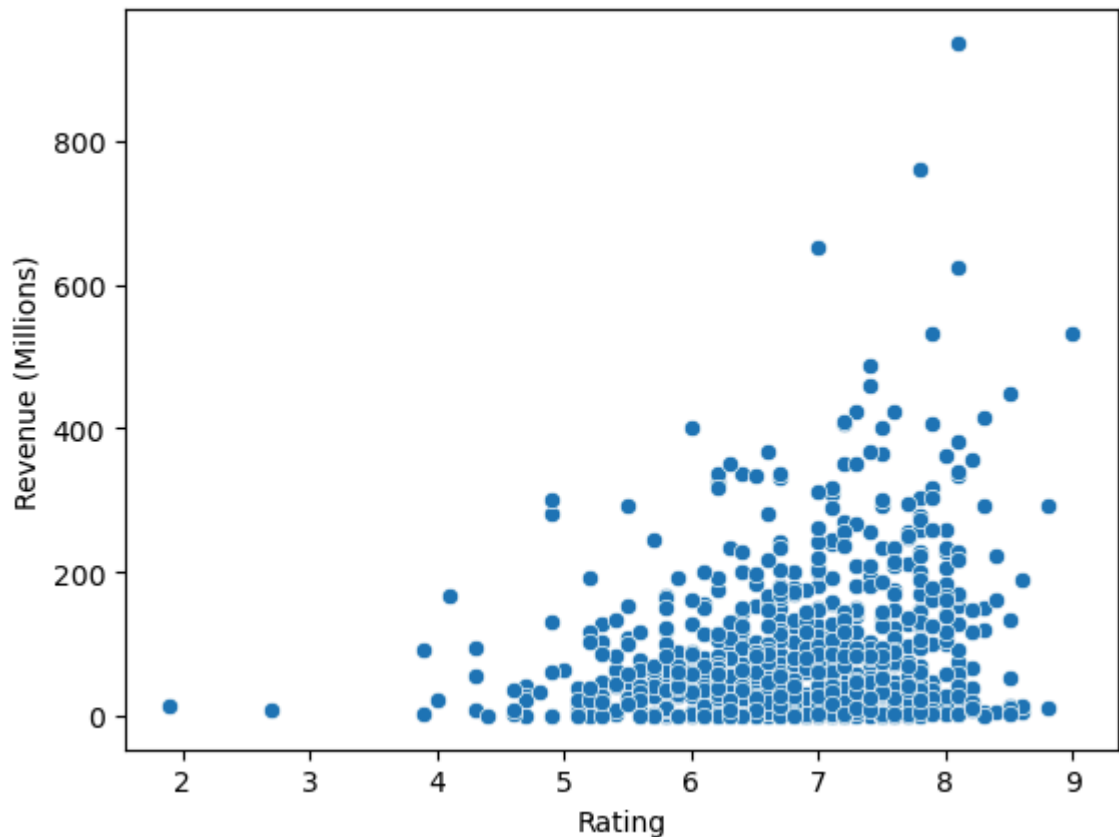
18. Show Yearly Average Movie rating

```
In [48]: imdb_df.groupby('Year')['Rating'].mean().sort_values(ascending = False)
```

```
Out[48]: Year
2007      7.133962
2006      7.125000
2009      6.960784
2012      6.925000
2011      6.838095
2014      6.837755
2010      6.826667
2013      6.812088
2008      6.784615
2015      6.602362
2016      6.436700
Name: Rating, dtype: float64
```

19 . Does Rating affect the Revenue?

```
In [49]: sns.scatterplot(x= 'Rating', y = 'Revenue (Millions)', data = imdb_df)
plt.show()
```



20. Classify Movie based on their Rating[Excellent, Good & Average]

```
In [50]: imdb_df.columns
```

```
Out[50]: Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year',
               'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)',
               'Metascore'],
              dtype='object')
```

```
In [61]: def mov_rating(rating):
          if rating >=7.0:
              return 'Excellent'
          elif rating>=6.0:
              return 'Good'
          else:
              return 'Average'
```

```
In [67]: imdb_df['Rating'].apply(mov_rating)
# apply() is used to apply any function on your dataset
# here Mov_rating() is our function we created above
```

```
Out[67]: 0      Excellent
1      Excellent
2      Excellent
3      Excellent
4           Good
        ...
995      Good
996    Average
997      Good
998    Average
999    Average
Name: Rating, Length: 1000, dtype: object
```

create New col for Rating Category

```
In [66]: imdb_df['Rating_Cate'] = imdb_df['Rating'].apply(mov_rating) #
## Rating_Cate is our new column that contains the rating Category
```

show the top 5 rows

In [65]: `imdb_df.head()`

Out[65]:

	Rank	Title	Genre	Description	Director	Actors	Year
0	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced ...	James Gunn	Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S...	20
1	2	Prometheus	Adventure,Mystery,Sci-Fi	Following clues to the origin of mankind, a te...	Ridley Scott	Noomi Rapace, Logan Marshall-Green, Michael Fa...	20
2	3	Split	Horror,Thriller	Three girls are kidnapped by a man with a diag...	M. Night Shyamalan	James McAvoy, Anya Taylor-Joy, Haley Lu Richar...	20
3	4	Sing	Animation,Comedy,Family	In a city of humanoid animals, a hustling thea...	Christophe Lourdelet	Matthew McConaughey,Reese Witherspoon, Seth Ma...	20
4	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of th...	David Ayer	Will Smith, Jared Leto, Margot Robbie, Viola D...	20

you can see that Reating Category has been applied on our dataset

21 . Show Total Count of Action Movies

In [68]: `imdb_df.columns`

Out[68]: Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year',
 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)',
 'Metascore', 'Rating_Cate'],
 dtype='object')

In [70]: `imdb_df['Genre'].dtype`

Out[70]: dtype('O')

so you can see that our Genre Column type is object, So we need to use `str.contains()` to count the Action Movie

```
In [73]: action_genre= imdb_df[imdb_df['Genre'].str.contains('Action', case=False)]  
         action_genre
```

Out[73]:

	Rank	Title	Genre	Description	Director	Actors	Year
0	1	Guardians of the Galaxy	Action,Adventure,Sci-Fi	A group of intergalactic criminals are forced ...	James Gunn	Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S...	2014
4	5	Suicide Squad	Action,Adventure,Fantasy	A secret government agency recruits some of th...	David Ayer	Will Smith, Jared Leto, Margot Robbie, Viola D...	2016
5	6	The Great Wall	Action,Adventure,Fantasy	European mercenaries searching for black powde...	Yimou Zhang	Matt Damon, Tian Jing, Willem Dafoe, Andy Lau	2016
8	9	The Lost City of Z	Action,Adventure,Biography	A true-life drama, centering on British explor...	James Gray	Charlie Hunnam, Robert Pattinson, Sienna Mille...	2016
12	13	Rogue One	Action,Adventure,Sci-Fi	The Rebel Alliance makes a risky move to steal...	Gareth Edwards	Felicity Jones, Diego Luna, Alan Tudyk, Donnie...	2016
...
958	959	3 Days to Kill	Action,Drama,Thriller	A dying CIA agent trying to reconnect with his...	McG	Kevin Costner, Hailee Steinfeld, Connie Nielse...	2014
968	969	Wrecker	Action,Horror,Thriller	Best friends Emily and Lesley go on a road tri...	Micheal Bafaro	Anna Hutchison, Andrea Whitburn, Jennifer Koen...	2015
969	970	The Lone Ranger	Action,Adventure,Western	Native American warrior Tonto recounts the unt...	Gore Verbinski	Johnny Depp, Armie Hammer, William Fichtner,To...	2013
990	991	Underworld: Rise of the Lycans	Action,Adventure,Fantasy	An origins story centered on the centuries-old...	Patrick Tatopoulos	Rhona Mitra, Michael Sheen, Bill Nighy, Steven...	2009
993	994	Resident Evil: Afterlife	Action,Adventure,Horror	While still out to destroy the evil Umbrella C...	Paul W.S. Anderson	Milla Jovovich, Ali Larter, Wentworth Miller,K...	2010

303 rows × 13 columns



so above we displayed the total Data set that containf Action Genre Movie, Lets show to total Count Below

```
In [74]: len(action_genre)
# len() can return thr total count of rows here
```

Out[74]: 303

Note: Analysis of any data limitless, we cann't say, its end here***

IMDB Movie Dataset is available on Kaggle, You can download it form the Kaggle web.

**url > <https://www.kaggle.com/datasets/PromptCloudHQ/imdb-data>
(<https://www.kaggle.com/datasets/PromptCloudHQ/imdb-data>)**

I'd love your feedback on this project, Your feedback will make me learn more.

Thank You :)

Regards , Vivek Kr Singh

**Follow >>> <https://www.linkedin.com/in/vivekvishwas/>
(<https://www.linkedin.com/in/vivekvishwas/>)**

In []: