

```
Task-2 MOVIE RATING PREDICTION WITH PYTHON

In [ ]:
%%CODEFMT

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

In [2]:
data=pd.read_csv('IMDB-Movie-Data.csv')

In [3]:
data

Out[3]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore
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  121  8.1  757074  333.13  76.0
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...  2012  124  7.0  485820  126.46  65.0
2  3  Split  Horror,Thriller  Three girls are kidnapped by a man with a diag...  M. Night Shyamalan  James McAvoy, Anna Taylor-Joy, Haley Lu Richar...  2016  117  7.3  157606  138.12  62.0
3  4  Sing  Animation,Comedy,Family  In a city of humanoid animals, a hustling thea...  Christophe Lourdelet  Matthew McConaughey,Reese Witherspoon, Seth Ma...  2016  108  7.2  60545  270.32  59.0
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  123  6.2  393727  325.02  40.0
...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...
995 995  Secret in Their Eyes  Crime,Drama,Mystery  A tight-knit team of rising investigators are b...  Billy Ray  Chiwetel Ejoror, Nicole Kidman, Julia Roberts, Ha...  2015  111  6.2  27895  NaN  45.0
996 997  Hostel: Part II  Horror  Three American college students studying abroa...  Eli Roth  Lauren German, Heather Matarazzo, Bipin Phil...  2007  94  5.5  73152  17.54  46.0
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...  2008  98  6.2  70699  58.01  50.0
998 999  Search Party  Adventure,Comedy  A pair of friends embark on a mission to reun...  Scott Armstrong  Adam Pally, T.J. Miller, Thomas Middleditch,Sh...  2014  93  5.6  4981  NaN  22.0
999 1000  Nine Lives  Comedy,Family,Fantasy  A stuffy businessman finds himself trapped ins...  Barry Sonnenfeld  Kevin Spacey, Jennifer Garner, Robbie Amell,Ch...  2016  87  5.3  12435  19.64  11.0

1000 rows x 12 columns

In [5]:
data.head(10)

Out[5]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore
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  121  8.1  757074  333.13  76.0
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...  2012  124  7.0  485820  126.46  65.0
2  3  Split  Horror,Thriller  Three girls are kidnapped by a man with a diag...  M. Night Shyamalan  James McAvoy, Anna Taylor-Joy, Haley Lu Richar...  2016  117  7.3  157606  138.12  62.0
3  4  Sing  Animation,Comedy,Family  In a city of humanoid animals, a hustling thea...  Christophe Lourdelet  Matthew McConaughey,Reese Witherspoon, Seth Ma...  2016  108  7.2  60545  270.32  59.0
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  123  6.2  393727  325.02  40.0
5  6  The Great Wall  Action,Adventure,Fantasy  European mercenaries searching for black powde...  Yimou Zhang  Matt Damon, Wen Jiang, Willem Dafoe, Andy Lau...  2016  103  6.1  56036  45.13  42.0
6  7  La La Land  Comedy,Drama,Romance  A jazz pianist falls for a movie star with a d...  Damien Chazelle  Ryan Gosling, Emma Stone, Rosemarie DeWitt...  2016  128  8.3  256862  151.06  93.0
7  8  Mindhorn  Comedy  A has-been actor best known for playing the r...  Sean Foley  Essie Davis, Julia Roberts, Avelina Bares...  2016  89  6.4  2490  NaN  71.0
8  9  The Last Day of Aztec  Action,Adventure,Biography  A true-life drama, centering on the conquest of...  James Gray  Charlie Hunnam, Robert Pattinson, Senna Miller...  2016  141  7.1  7188  8.01  78.0
9  10  Passengers  Adventure,Drama,Romance  A spacecraft traveling to a distant colony pla...  Morten Tyldum  Jennifer Lawrence, Chris Pratt, Michael Sheen...  2016  116  7.0  192177  100.01  41.0

In [7]:
data.tail(10)

Out[7]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore
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 Cooper, Zoe S...  2009  92  6.6  129708  45.80  44.0
991 992  Yare Zameen Par  Drama,Family,Music  An eighty-year-old boy is brought to be a lazy...  Aamir Khan  Darshad Sahay, Aamir Khan, Tanay Chhabra, Sa...  2007  165  8.5  102697  1.20  42.0
992 993  Take Me Home Tonight  Comedy,Drama,Romance  Four years after graduation, two friends who ha...  Tomparch Gracie  Anna Fais, Dan Foglia, Jessica F...  2011  97  6.3  45419  6.92  NaN
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,...  2010  97  5.9  149090  60.13  37.0
994 995  Project X  Comedy  3 high school seniors throw a birthday party...  Nima Nourizadeh  Thomas Mann, Oliver Cooper, Jonathan Daniel Br...  2012  88  6.7  164088  54.72  48.0
995 996  Secret in Their Eyes  Crime,Drama,Mystery  A tight-knit team of rising investigators are b...  Billy Ray  Chiwetel Ejoror, Nicole Kidman, Julia Roberts, Ha...  2015  111  6.2  27895  NaN  45.0
996 997  Hostel: Part II  Horror  Three American college students studying abroa...  Eli Roth  Lauren German, Heather Matarazzo, Bipin Phil...  2007  94  5.5  73152  17.54  46.0
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...  2008  98  6.2  70699  58.01  50.0
998 999  Search Party  Adventure,Comedy  A pair of friends embark on a mission to reun...  Scott Armstrong  Adam Pally, T.J. Miller, Thomas Middleditch,Sh...  2014  93  5.6  4981  NaN  22.0
999 1000  Nine Lives  Comedy,Family,Fantasy  A stuffy businessman finds himself trapped ins...  Barry Sonnenfeld  Kevin Spacey, Jennifer Garner, Robbie Amell,Ch...  2016  87  5.3  12435  19.64  11.0

In [8]:
data.shape

Out[8]:
(1000, 12)

In [10]:
print("Number of Rows",data.shape[0])
print("Number of Columns",data.shape[1])

Number of Rows 1000
Number of Columns 12

In [11]:
data.info()

<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  int16
dtypes: float64(3), int64(4), object(5)
memory usage: 93.9+ KB

In [12]:
print("Any missing value?",data.isnull().values.any())

Any missing value? True

In [13]:
data.isnull().sum()

Rank 0
Title 0
Genre 0
Description 0
Director 0
Actors 0
Year 0
Runtime (Minutes) 0
Rating 0
Votes 0
Revenue (Millions) 12
Metascore 64
dtype: int64

In [14]:
sns.heatmap(data.isnull())

<Axes: >

In [15]:
per_missing = data.isnull().sum()*(100 / len(data))

Out[15]:
Rank 0.0
Title 0.0
Genre 0.0
Description 0.0
Director 0.0
Actors 0.0
Year 0.0
Runtime (Minutes) 0.0
Rating 0.0
Votes 0.0
Revenue (Millions) 12.8
Metascore 6.4
dtype: float64

In [17]:
data.dropna(axis=8)

Out[17]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore
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  121  8.1  757074  333.13  76.0
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...  2012  124  7.0  485820  126.46  65.0
2  3  Split  Horror,Thriller  Three girls are kidnapped by a man with a diag...  M. Night Shyamalan  James McAvoy, Anna Taylor-Joy, Haley Lu Richar...  2016  117  7.3  157606  138.12  62.0
3  4  Sing  Animation,Comedy,Family  In a city of humanoid animals, a hustling thea...  Christophe Lourdelet  Matthew McConaughey,Reese Witherspoon, Seth Ma...  2016  108  7.2  60545  270.32  59.0
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  123  6.2  393727  325.02  40.0
...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...
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,...  2010  97  5.9  149090  60.13  37.0
994 995  Project X  Comedy  3 high school seniors throw a birthday party...  Nima Nourizadeh  Thomas Mann, Oliver Cooper, Jonathan Daniel Br...  2012  88  6.7  164088  54.72  48.0
996 997  Hostel: Part II  Horror  Three American college students studying abroa...  Eli Roth  Lauren German, Heather Matarazzo, Bipin Phil...  2007  94  5.5  73152  17.54  46.0
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...  2008  98  6.2  70699  58.01  50.0
999 1000  Nine Lives  Comedy,Family,Fantasy  A stuffy businessman finds himself trapped ins...  Barry Sonnenfeld  Kevin Spacey, Jennifer Garner, Robbie Amell,Ch...  2016  87  5.3  12435  19.64  11.0

838 rows x 12 columns

In [19]:
dup_data=data.duplicated().any()

In [20]:
print("Are there any duplicate values?",dup_data)

Are there any duplicate values? False

In [22]:
data=data.drop_duplicates()

Out[22]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore
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  121  8.1  757074  333.13  76.0
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...  2012  124  7.0  485820  126.46  65.0
2  3  Split  Horror,Thriller  Three girls are kidnapped by a man with a diag...  M. Night Shyamalan  James McAvoy, Anna Taylor-Joy, Haley Lu Richar...  2016  117  7.3  157606  138.12  62.0
3  4  Sing  Animation,Comedy,Family  In a city of humanoid animals, a hustling thea...  Christophe Lourdelet  Matthew McConaughey,Reese Witherspoon, Seth Ma...  2016  108  7.2  60545  270.32  59.0
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  123  6.2  393727  325.02  40.0
...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...  ...
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,...  2010  97  5.9  149090  60.13  37.0
994 995  Project X  Comedy  3 high school seniors throw a birthday party...  Nima Nourizadeh  Thomas Mann, Oliver Cooper, Jonathan Daniel Br...  2012  88  6.7  164088  54.72  48.0
996 997  Hostel: Part II  Horror  Three American college students studying abroa...  Eli Roth  Lauren German, Heather Matarazzo, Bipin Phil...  2007  94  5.5  73152  17.54  46.0
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...  2008  98  6.2  70699  58.01  50.0
999 1000  Nine Lives  Comedy,Family,Fantasy  A stuffy businessman finds himself trapped ins...  Barry Sonnenfeld  Kevin Spacey, Jennifer Garner, Robbie Amell,Ch...  2016  87  5.3  12435  19.64  11.0

838 rows x 12 columns

In [24]:
data.describe(include='all')

Out[24]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore
count  838.000000  838  838  838  838  838  838.000000  838.000000  8.360000e+02  838.000000  838.000000
unique  NaN  837  189  534  534  NaN  NaN  NaN  NaN  NaN  NaN  NaN
top  NaN  The Action,Adventure,Sci-Fi  A group of intergalactic criminals are forced ...  Ridley Scott  Jennifer Lawrence, Josh Hutcherson, Liam Hemm...  NaN  NaN  NaN  NaN  NaN
freq  NaN  2  50  1  8  2  NaN  NaN  NaN  NaN  NaN  NaN
mean  485.247017  NaN  NaN  NaN  NaN  NaN  NaN  2012.50716  114.038425  6.814325  1.932303e+05  84.564558  59.575179
std  286.572065  NaN  NaN  NaN  NaN  NaN  NaN  3.17235  1.372785  1.930909e+05  104.520227  16.952416
25%  1.000000  NaN  NaN  NaN  NaN  NaN  NaN  2006.00000  66.000000  1.900000e+04  0.000000  11.000000
29%  238.250000  NaN  NaN  NaN  NaN  NaN  NaN  2010.00000  101.000000  6.300000e+04  1.367500e+02  40.000000
50%  475.500000  NaN  NaN  NaN  NaN  NaN  NaN  2013.00000  112.000000  1.268795e+05  48.150000  60.000000
75%  729.750000  NaN  NaN  NaN  NaN  NaN  NaN  2015.00000  124.000000  7.500000e+05  118.800000  72.000000
max  1000.00000  NaN  NaN  NaN  NaN  NaN  NaN  2016.00000  187.000000  9.000000e+05  398.030000  100.000000

In [25]:
data.columns

In [26]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [26]:
data['Runtime (Minutes)']>=188

Out[26]:
0      False
1      False
2      False
3      False
4      False
993     False
994     False
995     False
996     False
997     False
998     False
999     False
Name: Runtime (Minutes), Length: 838, dtype: bool

In [29]:
data[data['Runtime (Minutes)']>=180]['Title']

Out[29]:
82      The Wolf of Wall Street
88      The Hateful Eight
311     The Tale of Adele
Name: Title, dtype: object

In [36]:
data.columns

In [36]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [33]:
data.groupby('Year')['Votes'].mean()

Out[33]:
Year
2006    277232.159512
2007    266589.784945
2008    186580.145853
2009    267189.577778
2010    251862.529825
2011    259254.738862
2012    298851.483871
2013    255513.862857
2014    211926.881720
2015    129512.651316
2016    68437.822322
Name: Votes, dtype: float64

In [35]:
sns.barplot(x='Year', y='Votes',data=data)
plt.title('Votes By Year')
plt.show()

Out[35]:
<Figure with 1 Axes>

In [36]:
data.columns

In [36]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [40]:
data.groupby('Year')['Revenue (Millions)'].mean().sort_values(ascending=False)

Out[40]:
Year
2009    115.742860
2012    118.183865
2010    105.881579
2008    97.529417
2011    93.714891
2013    88.864643
2006    87.255610
2014    85.438656
2015    80.725596
2016    55.550311
Name: Revenue (Millions), dtype: float64

In [41]:
sns.barplot(x='Year', y='Revenue (Millions)',data=data)
plt.title('Votes By Year')
plt.show()

Out[41]:
<Figure with 1 Axes>

In [42]:
data.columns

In [42]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [44]:
data.groupby('Director')['Rating'].mean().sort_values(ascending=False)

Out[44]:
Director
Christopher Nolan    8.68
Oliver Nakache       8.60
Makoto Shinkai       8.60
Florian Henckel von Donnersmarck  8.50
Aamir Khan           8.50
Sam Taylor-Johnson   4.10
Joey Curtis          4.80
George Miller        2.90
James Wong           2.70
Jason Friedberg      1.90
Name: Rating, Length: 524, dtype: float64

In [45]:
data.columns

In [45]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [48]:
data.isna().nlargest(10, 'Runtime (Minutes)')['Title', 'Runtime (Minutes)'].set_index('Title')

Out[48]:
Runtime (Minutes)
Title
The Hateful Eight 187
The Wolf of Wall Street 180
La vie d'Adèle 180
Cloud Atlas 172
3 Idiots 170
Interstellar 169
Pirates of the Caribbean: At World's End 169
The Hobbit: An Unexpected Journey 169
The Curious Case of Benjamin Button 166
Transformers: Age of Extinction 165

In [53]:
sns.barplot(x='Runtime (Minutes)',y=top10_len_index,data=top10_len)

Out[53]:
<Axes: xlabel='Runtime (Minutes)', ylabel='Title'>

In [54]:
The Hateful Eight
The Wolf of Wall Street
La vie d'Adèle
Cloud Atlas
3 Idiots
Interstellar
Pirates of the Caribbean: At World's End
The Hobbit: An Unexpected Journey
The Curious Case of Benjamin Button
Transformers: Age of Extinction

In [54]:
data.columns

In [54]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [55]:
data['Year'].value_counts()

Out[55]:
2016    188
2015    189
2014    93
2013    84
2012    82
2018    57
2019    57
2008    45
2009    45
2007    44
2006    42
Name: Year, dtype: int64

In [58]:
sns.countplot(x='Year',data=data)
plt.title('Number of Movies Per Year')
plt.show()

Out[58]:
<Figure with 1 Axes>

In [59]:
data.columns

In [59]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [61]:
data['Revenue (Millions)'].max()==data['Revenue (Millions)']

Out[61]:
0      False
1      False
2      False
3      False
4      False
993     False
994     False
995     False
996     False
997     False
998     False
999     False
Name: Revenue (Millions), Length: 838, dtype: bool

In [62]:
data.columns

In [62]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [64]:
data.isna().nlargest(10, 'Rating')['Title', 'Rating', 'Director'].set_index('Title')

Out[64]:
Rating
Director
Title
The Dark Knight 9.0 Christopher Nolan
Interception 8.8 Christopher Nolan
Interstellar 8.6 Makoto Shinkai
The Intouchables 8.6 Oliver Nakache
The Prestige 8.5 Christopher Nolan
The Departed 8.5 Martin Scorsese
The Dark Knight Rises 8.5 Christopher Nolan
Whiplash 8.5 Damien Chazelle
The Lives of Others 8.5 Florian Henckel von Donnersmarck

In [76]:
sns.barplot(x='Rating',y=top10_len_index,data=top10_len)

Out[76]:
<Axes: xlabel='Rating', ylabel='Title'>

In [71]:
The Dark Knight
Interception
Interstellar
Kimi no na wa
The Intouchables
The Prestige
The Departed
The Dark Knight Rises
Whiplash
The Lives of Others

In [71]:
data.columns

In [71]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [73]:
data.nlargest(10, 'Revenue (Millions)')['Title']

Out[73]:
50  Star Wars: Episode VII - The Force Awakens
87  Avatar
85  Jurassic World
76  The Avengers
52  The Dark Knight
12  Rogue One
139 The Dark Knight
94  Avengers: Age of Ultron
124 The Dark Knight Rises
578 The Hunger Games: Catching Fire
Name: Title, dtype: object

In [79]:
top10data=data.nlargest(10, 'Revenue (Millions)')['Title', 'Revenue (Millions)'].set_index('Title')

Out[79]:
top_10

In [79]:
top_10

Out[79]:
   Title  Revenue (Millions)
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  456.29
Avengers: Age of Ultron  458.99
The Dark Knight Rises  424.13
The Hunger Games: Catching Fire  416.65

In [81]:
sns.barplot(x='Revenue (Millions)',y=top10_len_index,data=top10_18)

Out[81]:
<Figure with 1 Axes>

In [82]:
data.columns

In [82]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [83]:
data.groupby('Year')['Rating'].mean().sort_values(ascending=False)

Out[83]:
Year
2007    7.143902
2008    7.148909
2011    7.346562
2012    6.933871
2009    6.911121
2010    6.894737
2013    6.822453
2014    6.825981
2006    6.785333
2015    6.676312
2016    6.444444
Name: Rating, dtype: float64

In [84]:
data.columns

In [84]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [85]:
sns.scatterplot(x='Rating',y='Revenue (Millions)',data=data)

Out[85]:
<Axes: xlabel='Rating', ylabel='Revenue (Millions)'>

In [86]:
Revenue (Millions)
Rating

In [86]:
data.columns

In [86]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [87]:
def rating(rating):
    if rating>= 8:
        return "Excellent"
    elif rating>= 7:
        return "Good"
    else:
        return "Average"

In [88]:
data['rating_cat']=data['Rating'].apply(rating)

In [89]:
data.head()

Out[89]:
   Rank  Title  Genre  Description  Director  Actors  Year  Runtime  Rating  Votes  Revenue  Metascore  rating_cat
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  121  8.1  757074  333.13  76.0  Excellent
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...  2012  124  7.0  485820  126.46  65.0  Excellent
2  3  Split  Horror,Thriller  Three girls are kidnapped by a man with a diag...  M. Night Shyamalan  James McAvoy, Anna Taylor-Joy, Haley Lu Richar...  2016  117  7.3  157606  138.12  62.0  Excellent
3  4  Sing  Animation,Comedy,Family  In a city of humanoid animals, a hustling thea...  Christophe Lourdelet  Matthew McConaughey,Reese Witherspoon, Seth Ma...  2016  108  7.2  60545  270.32  59.0  Excellent
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  123  6.2  393727  325.02  40.0  Good

In [90]:
data.columns

In [90]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [91]:
data['Genre'].dtype

Out[91]:
dtype('O')

In [94]:
len(data[data['Genre'].str.contains('Action',case=False)])

Out[94]:
277

In [95]:
data.columns

In [95]:
Index(['Rank', 'Title', 'Genre', 'Description', 'Director', 'Actors', 'Year', 'Runtime (Minutes)', 'Rating', 'Votes', 'Revenue (Millions)', 'Metascore'],
      dtype='object')

In [96]:
data['Genre']

Out[96]:
0      Action,Adventure,Sci-Fi
1      Adventure,Mystery,Sci-Fi
2      Action,Adventure,Horror
3      Animation,Comedy,Family
4      Action,Adventure,Fantasy
993  Action,Adventure,Horror
994  Action,Adventure,Horror
995  Action,Adventure,Horror
996  Action,Adventure,Horror
997  Action,Adventure,Horror
998  Action,Adventure,Horror
999  Action,Adventure,Horror
Name: Genre, Length: 838, dtype: object

In [100]:
list1=[]
for value in data['Genre']:
    list.append(value.split(','))

In [102]:
list1

Out[102]:
[]

In [ ]:
*****END OF CODE*****

In [ ]:
*****THANK YOU*****
```