notebookbd3f87ac8b

March 12, 2024

```
[1]: # Installing all neccessary Libraries
     !pip install numpy
     !pip install pandas
     !pip install matplotlib1
     !pip install seaborn
```

```
Requirement already satisfied: numpy in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (1.26.4)


[notice] A new release of pip is available: 23.2.1 -> 24.0
[notice] To update, run: python.exe -m pip install --upgrade pip

Requirement already satisfied: pandas in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (2.2.1)
Requirement already satisfied: numpy<2,>=1.26.0 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
pandas) (1.26.4)
Requirement already satisfied: python-dateutil>=2.8.2 in
c:\users\hp\appdata\roaming\python\python312\site-packages (from pandas)
(2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
pandas) (2024.1)
Requirement already satisfied: six>=1.5 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
python-dateutil>=2.8.2->pandas) (1.16.0)


[notice] A new release of pip is available: 23.2.1 -> 24.0
[notice] To update, run: python.exe -m pip install --upgrade pip
ERROR: Could not find a version that satisfies the requirement matplotlib1 (from
versions: none)
ERROR: No matching distribution found for matplotlib1


[notice] A new release of pip is available: 23.2.1 -> 24.0
```

Requirement already satisfied: seaborn in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (0.13.2)
Requirement already satisfied: numpy!=1.24.0,>=1.20 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
seaborn) (1.26.4)
Requirement already satisfied: pandas>=1.2 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
seaborn) (2.2.1)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
seaborn) (3.8.3)
Requirement already satisfied: contourpy>=1.0.1 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (1.2.0)
Requirement already satisfied: cycler>=0.10 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (4.49.0)
Requirement already satisfied: kiwisolver>=1.3.1 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (1.4.5)
Requirement already satisfied: packaging>=20.0 in
c:\users\hp\appdata\roaming\python\python312\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (24.0)
Requirement already satisfied: pillow>=8 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (10.2.0)
Requirement already satisfied: pyparsing>=2.3.1 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in
c:\users\hp\appdata\roaming\python\python312\site-packages (from
matplotlib!=3.6.1,>=3.4->seaborn) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: six>=1.5 in
c:\users\hp\appdata\local\programs\python\python312\lib\site-packages (from
python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.16.0)

```
[2]: # This Python 3 environment comes with many helpful analytics libraries␣
     ↪installed
     # It is defined by the kaggle/python Docker image: https://github.com/kaggle/
     ↪docker-python
     # For example, here's several helpful packages to load

     import numpy as np # linear algebra
     import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

     # Input data files are available in the read-only "../input/" directory
     # For example, running this (by clicking run or pressing Shift+Enter) will list␣
     ↪all files under the input directory

     import os
     for dirname, _, filenames in os.walk('./archive'):
         for filename in filenames:
             print(os.path.join(dirname, filename))

     # You can write up to 20GB to the current directory (/kaggle/working/) that␣
     ↪gets preserved as output when you create a version using "Save & Run All"
     # You can also write temporary files to /kaggle/temp/, but they won't be saved␣
     ↪outside of the current session
```

```
./archive\IMDB-Movie-Data.csv
```

```
[3]: import warnings
     warnings.filterwarnings('ignore')
```

```
[4]: data = pd.read_csv('./archive/IMDB-Movie-Data.csv')
```

1 1. Display Top 10 Rows of The Dataset

```
[5]: data.head(10)
```

```
[5]:    Rank                   Title                       Genre  \
     0     1  Guardians of the Galaxy     Action,Adventure,Sci-Fi
     1     2               Prometheus    Adventure,Mystery,Sci-Fi
     2     3                    Split             Horror,Thriller
     3     4                     Sing     Animation,Comedy,Family
     4     5            Suicide Squad    Action,Adventure,Fantasy
     5     6           The Great Wall    Action,Adventure,Fantasy
     6     7                La La Land         Comedy,Drama,Music
     7     8                 Mindhorn                      Comedy
     8     9       The Lost City of Z  Action,Adventure,Biography
```

```
9    10              Passengers      Adventure,Drama,Romance

                                       Description               Director  \
0   A group of intergalactic criminals are forced …        James Gunn
1   Following clues to the origin of mankind, a te…        Ridley Scott
2   Three girls are kidnapped by a man with a diag…    M. Night Shyamalan
3   In a city of humanoid animals, a hustling thea…  Christophe Lourdelet
4   A secret government agency recruits some of th…          David Ayer
5   European mercenaries searching for black powde…        Yimou Zhang
6   A jazz pianist falls for an aspiring actress i…     Damien Chazelle
7   A has-been actor best known for playing the ti…         Sean Foley
8   A true-life drama, centering on British explor…          James Gray
9   A spacecraft traveling to a distant colony pla…       Morten Tyldum


                                         Actors  Year  Runtime (Minutes)  \
0   Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S…  2014                121
1   Noomi Rapace, Logan Marshall-Green, Michael Fa…  2012                124
2   James McAvoy, Anya Taylor-Joy, Haley Lu Richar…  2016                117
3   Matthew McConaughey,Reese Witherspoon, Seth Ma…  2016                108
4   Will Smith, Jared Leto, Margot Robbie, Viola D…  2016                123
5        Matt Damon, Tian Jing, Willem Dafoe, Andy Lau  2016                103
6   Ryan Gosling, Emma Stone, Rosemarie DeWitt, J…  2016                128
7   Essie Davis, Andrea Riseborough, Julian Barrat…  2016                 89
8   Charlie Hunnam, Robert Pattinson, Sienna Mille…  2016                141
9   Jennifer Lawrence, Chris Pratt, Michael Sheen,…  2016                116

    Rating   Votes  Revenue (Millions)  Metascore
0      8.1  757074              333.13       76.0
1      7.0  485820              126.46       65.0
2      7.3  157606              138.12       62.0
3      7.2   60545              270.32       59.0
4      6.2  393727              325.02       40.0
5      6.1   56036               45.13       42.0
6      8.3  258682              151.06       93.0
7      6.4    2490                 NaN       71.0
8      7.1    7188                8.01       78.0
9      7.0  192177              100.01       41.0
```

2 2. Check Last 10 Rows of The Dataset

```
[6]: data.tail(10)
```

```
[6]:      Rank                         Title                     Genre  \
     990   991  Underworld: Rise of the Lycans  Action,Adventure,Fantasy
     991   992                 Taare Zameen Par        Drama,Family,Music
     992   993             Take Me Home Tonight      Comedy,Drama,Romance
```

```
993   994          Resident Evil: Afterlife    Action,Adventure,Horror
994   995                        Project X                     Comedy
995   996            Secret in Their Eyes       Crime,Drama,Mystery
996   997                 Hostel: Part II                      Horror
997   998            Step Up 2: The Streets     Drama,Music,Romance
998   999                    Search Party         Adventure,Comedy
999   1000                     Nine Lives      Comedy,Family,Fantasy
```

```
                                          Description           Director  \
990  An origins story centered on the centuries-old…  Patrick Tatopoulos
991  An eight-year-old boy is thought to be a lazy …         Aamir Khan
992  Four years after graduation, an awkward high s…      Michael Dowse
993  While still out to destroy the evil Umbrella C…  Paul W.S. Anderson
994  3 high school seniors throw a birthday party t…     Nima Nourizadeh
995  A tight-knit team of rising investigators, alo…          Billy Ray
996  Three American college students studying abroa…           Eli Roth
997  Romantic sparks occur between two dance studen…         Jon M. Chu
998  A pair of friends embark on a mission to reuni…     Scot Armstrong
999  A stuffy businessman finds himself trapped ins…   Barry Sonnenfeld
```

```
                                           Actors  Year  \
990  Rhona Mitra, Michael Sheen, Bill Nighy, Steven…  2009
991  Darsheel Safary, Aamir Khan, Tanay Chheda, Sac…  2007
992  Topher Grace, Anna Faris, Dan Fogler, Teresa P…  2011
993  Milla Jovovich, Ali Larter, Wentworth Miller,K…  2010
994  Thomas Mann, Oliver Cooper, Jonathan Daniel Br…  2012
995  Chiwetel Ejiofor, Nicole Kidman, Julia Roberts…  2015
996  Lauren German, Heather Matarazzo, Bijou Philli…  2007
997  Robert Hoffman, Briana Evigan, Cassie Ventura,…  2008
998  Adam Pally, T.J. Miller, Thomas Middleditch,Sh…  2014
999  Kevin Spacey, Jennifer Garner, Robbie Amell,Ch…  2016
```

```
     Runtime (Minutes)  Rating   Votes  Revenue (Millions)  Metascore
990                 92     6.6  129708               45.80       44.0
991                165     8.5  102697                1.20       42.0
992                 97     6.3   45419                6.92        NaN
993                 97     5.9  140900               60.13       37.0
994                 88     6.7  164088               54.72       48.0
995                111     6.2   27585                 NaN       45.0
996                 94     5.5   73152               17.54       46.0
997                 98     6.2   70699               58.01       50.0
998                 93     5.6    4881                 NaN       22.0
999                 87     5.3   12435               19.64       11.0
```

3 3. Find Shape of Our Dataset (Number of Rows And Number of Columns)

[7]: `data.shape`

[7]: (1000, 12)

[8]:
```
print('Number of Rows',data.shape[0])
print('Number of Columns',data.shape[1])
```

```
Number of Rows 1000
Number of Columns 12
```

4 4. Getting Information About Our Dataset Like Total Number Rows, Total Number of Columns, Datatypes of Each Column And Memory Requirement

[9]: `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   float64
dtypes: float64(3), int64(4), object(5)
memory usage: 93.9+ KB
```

5 5. Check Null Values In The Dataset

[10]: `data.isnull().sum()`

[10]:
```
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
```
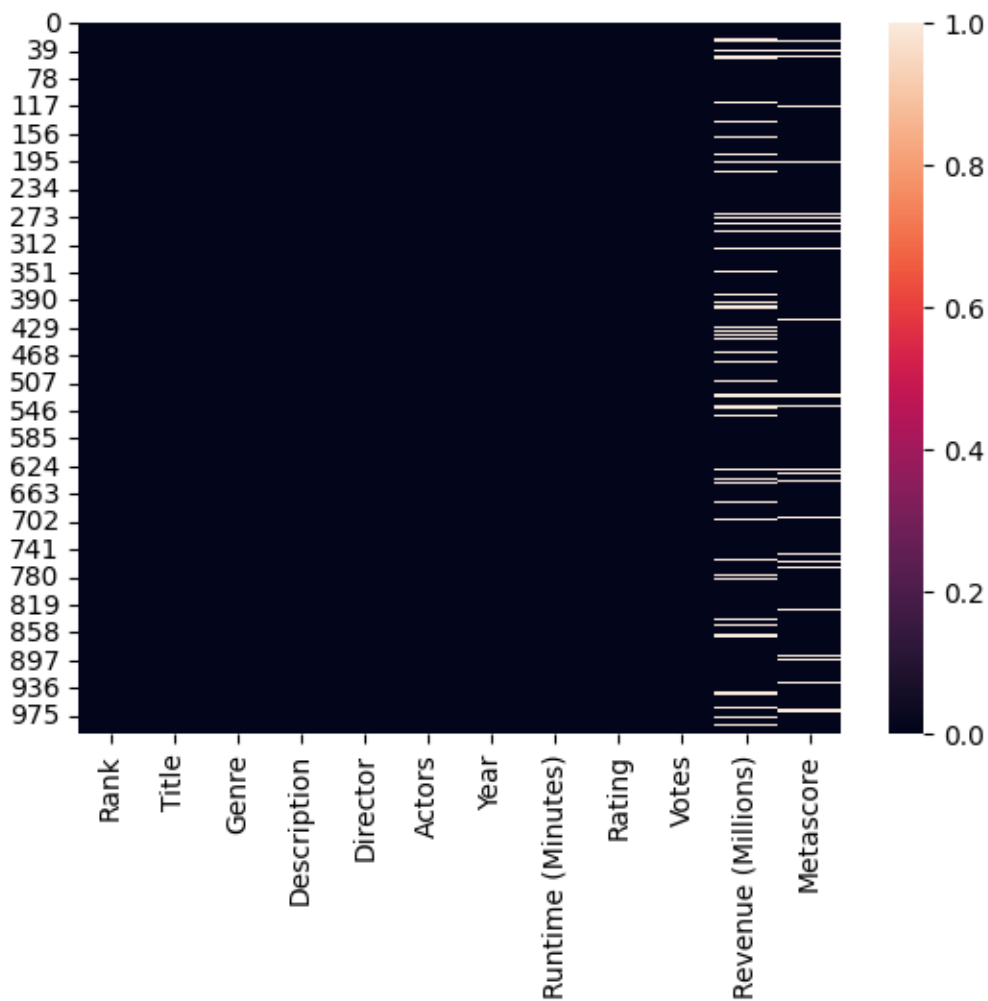
[11]:
```python
import matplotlib.pyplot as plt
import seaborn as sns
sns.heatmap(data.isnull())
plt.show()
```

```
[12]: per_missing = data.isnull().sum()*100/len(data)
      per_missing
```

```
[12]: 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
```

6 6. Drop All The Missing Values

```
[13]: data = data.dropna(axis=0)
```

```
[14]: sns.heatmap(data.isnull())
      plt.show()
```
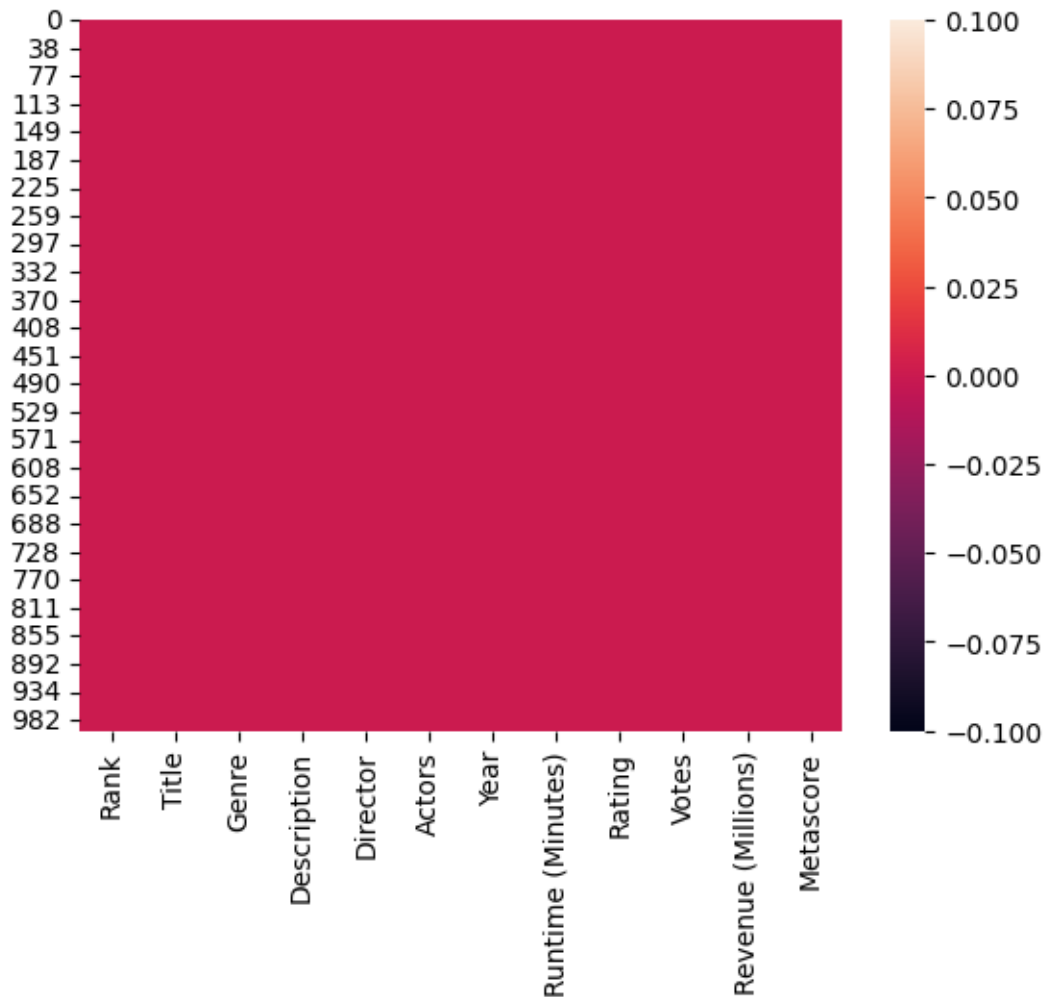
7 7. Check For Duplicate Data

```
[15]: dup_data=data.duplicated().any()
      print("Are there any duplicated values in data?",dup_data)
```

Are there any duplicated values in data? False

8 8. Get Overall Statistics About The DataFrame

```
[16]: data.describe()
```

```
[16]:              Rank        Year  Runtime (Minutes)      Rating        Votes  \
      count  838.000000    838.00000         838.000000  838.000000  8.380000e+02
      mean   485.247017  2012.50716         114.638425    6.814320  1.932303e+05
```

```
std     286.572065    3.17236        18.470922   0.877754  1.930990e+05
min       1.000000  2006.00000        66.000000   1.900000  1.780000e+02
25%     238.250000  2010.00000       101.000000   6.300000  6.127650e+04
50%     475.500000  2013.00000       112.000000   6.900000  1.368795e+05
75%     729.750000  2015.00000       124.000000   7.500000  2.710830e+05
max    1000.000000  2016.00000       187.000000   9.000000  1.791916e+06

       Revenue (Millions)   Metascore
count          838.000000  838.000000
mean            84.564558   59.575179
std            104.520227   16.952416
min              0.000000   11.000000
25%             13.967500   47.000000
50%             48.150000   60.000000
75%            116.800000   72.000000
max            936.630000  100.000000
```

9   9. Display Title of The Movie Having Runtime >= 180 Minutes

```python
[17]: data[data['Runtime (Minutes)']>=180]['Title']
```

```
[17]: 82     The Wolf of Wall Street
      88           The Hateful Eight
      311            La vie d'Adèle
      Name: Title, dtype: object
```

10   10. In Which Year There Was The Highest Voting?

```python
[18]: data.groupby('Year')['Votes'].mean().sort_values(ascending=False).head(1)
```

```
[18]: Year
      2012    290861.483871
      Name: Votes, dtype: float64
```

```python
[19]: # viswualization
      sns.barplot(x='Year',y='Votes',data=data)
      plt.title("Votes By Year")
      plt.show()
```

Votes By Year
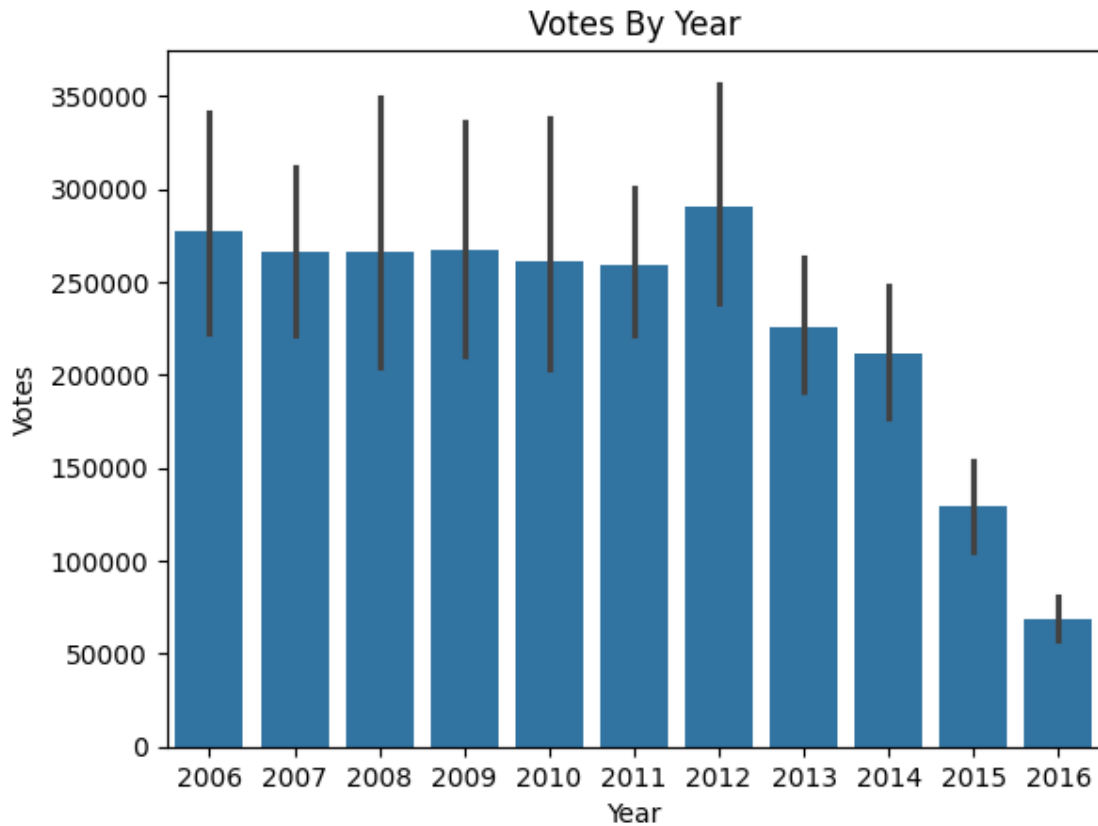
11 11. In Which Year There Was The Highest Revenue?

```
[20]: data.groupby('Year')['Revenue (Millions)'].mean().sort_values(ascending=False).
      ↪head(1)
```

```
[20]: Year
      2009    115.742
      Name: Revenue (Millions), dtype: float64
```

```
[21]: # Visualization
      sns.barplot(x='Year',y='Revenue (Millions)',data=data)
      plt.title("Revenue By Year")
      plt.show()
```

Revenue By Year
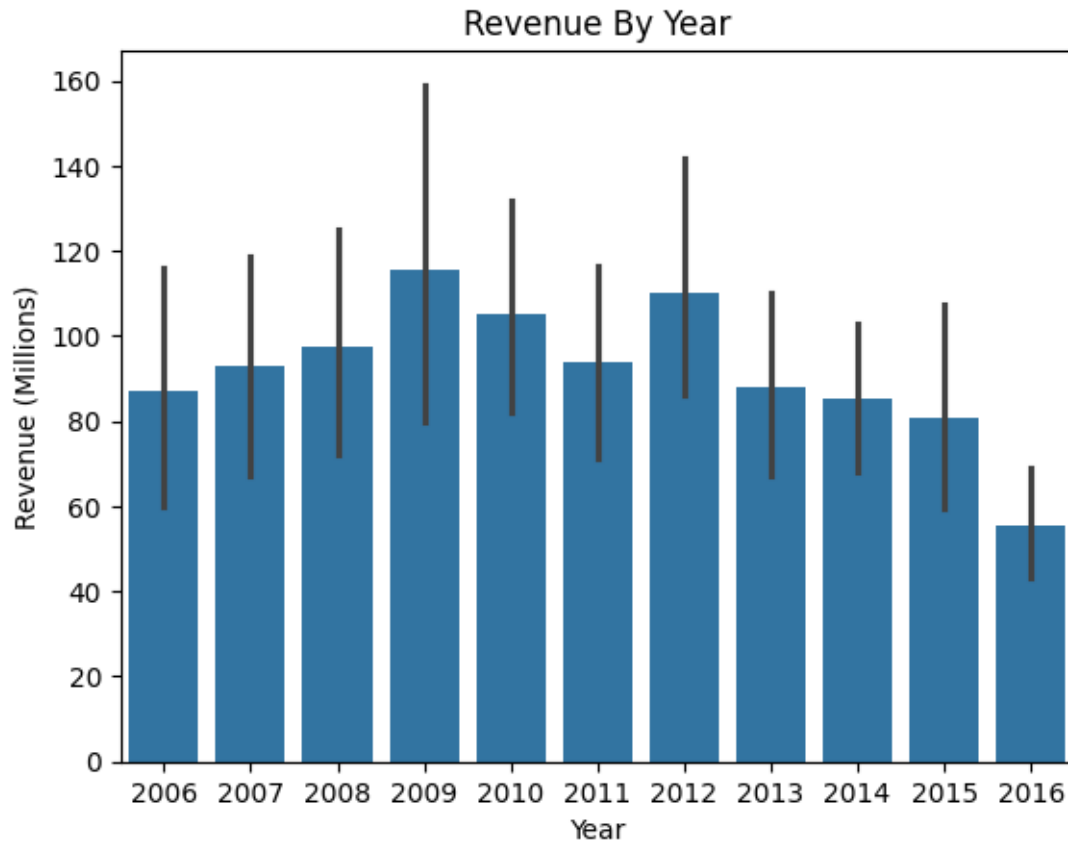
12 12. Find The Average Rating For Each Director

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

```
[22]: Director
      Christopher Nolan                   8.68
      Olivier Nakache                     8.60
      Makoto Shinkai                      8.60
      Florian Henckel von Donnersmarck    8.50
      Aamir Khan                          8.50
                                           …
      Sam Taylor-Johnson                  4.10
      Joey Curtis                         4.00
      George Nolfi                        3.90
      James Wong                          2.70
      Jason Friedberg                     1.90
      Name: Rating, Length: 524, dtype: float64
```

13 13. Display Top 10 Lengthy Movies Title

```
[23]: le =data.nlargest(10,'Runtime (Minutes)')[['Title','Runtime (Minutes)']] \
      .set_index('Title')
      le
```

```
[23]:                                               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
```

```
[24]: # Visualization
      sns.barplot(x='Runtime (Minutes)',y=le.index,data=le)
      plt.title('Top 10 Lengthy Movies')
      plt.show()
```


13
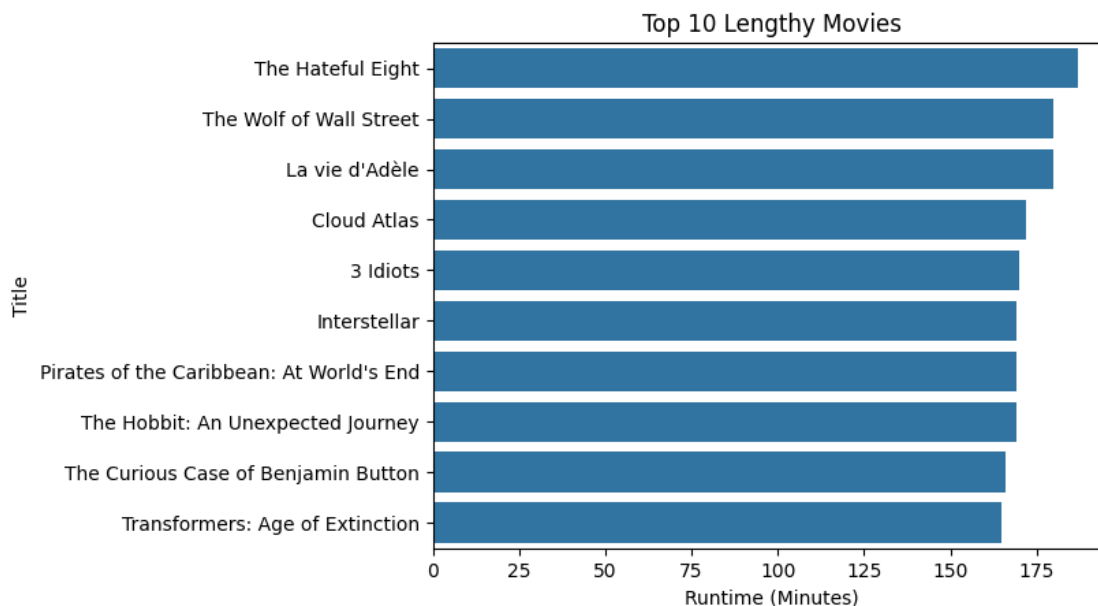
14 14. Display Number of Movies Per Year

```
[25]: data['Year'].value_counts()
```

```
[25]: Year
      2016    198
      2015    109
      2014     93
      2013     84
      2012     62
      2011     57
      2010     57
      2008     48
      2009     45
      2007     44
      2006     41
      Name: count, dtype: int64
```
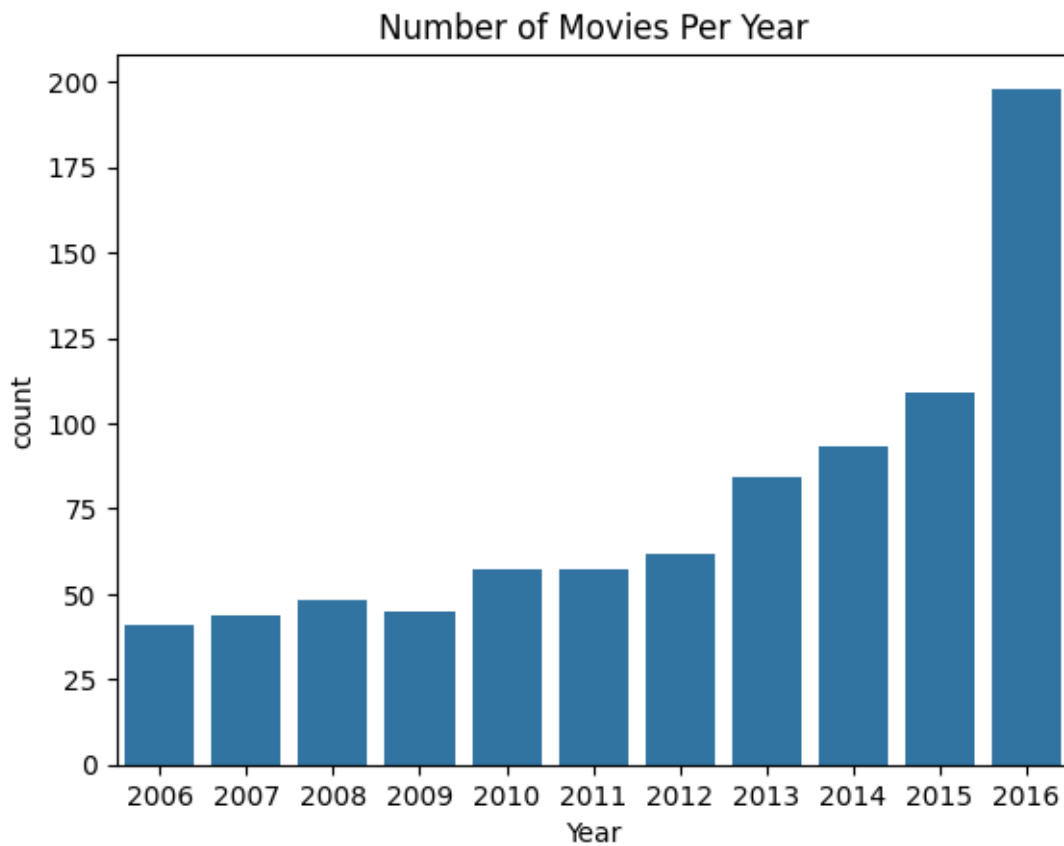
```
[26]: sns.countplot(x='Year',data=data)
      plt.title("Number of Movies Per Year")
      plt.show()
```

15 15. Find Most Popular Movie Title (Higest Revenue)

```
[27]: data[data['Revenue (Millions)'].max() == data['Revenue (Millions)']]['Title']
```

```
[27]: 50    Star Wars: Episode VII - The Force Awakens
      Name: Title, dtype: object
```

16 16. Display Top 10 Highest Rated Movie Titles And its Directors

```
[28]: top_10=data.nlargest(10,'Rating')[['Title','Rating','Director']].
      ↪set_index('Title')
      top_10
```

```
[28]:                         Rating                         Director
      Title
      The Dark Knight           9.0                 Christopher Nolan
      Inception                 8.8                 Christopher Nolan
      Interstellar              8.6                 Christopher Nolan
      Kimi no na wa             8.6                    Makoto Shinkai
      The Intouchables          8.6                   Olivier 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
```

```
[29]: # Visualization
      sns.barplot(x = 'Rating', y = top_10.index, data =␣
      ↪top_10,hue='Director',dodge=False)
      plt.title("Display Top 10 Highest Rated Movie Titles")
      plt.legend(bbox_to_anchor=(1.05, 1), loc=2)
      plt.show()
```

Display Top 10 Highest Rated Movie Titles
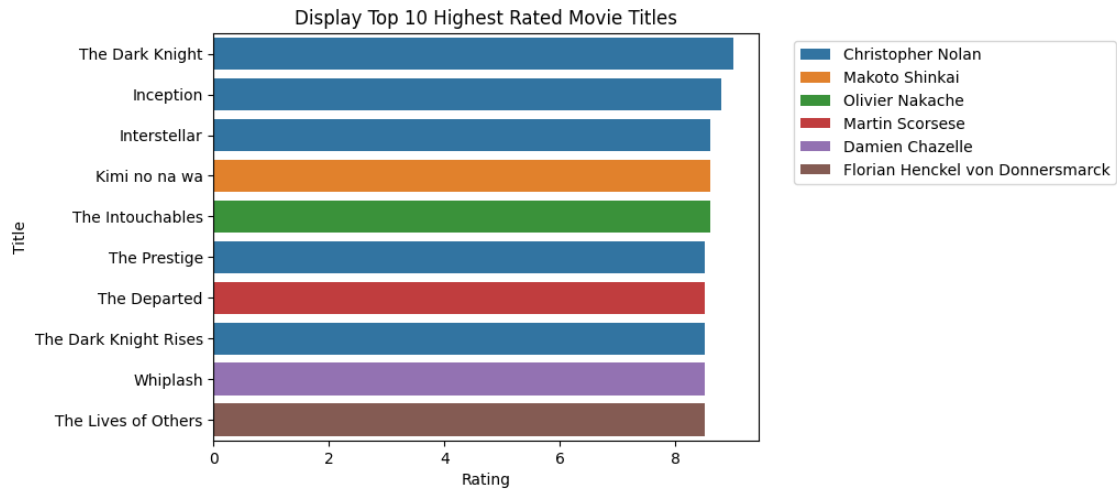
17 17. Display Top 10 Highest Revenue Movie Titles

```
[30]: data.sort_values(by='Revenue (Millions)',ascending=False).head(10)
```

```
[30]:        Rank                                    Title  \
        50     51  Star Wars: Episode VII - The Force Awakens
        87     88                                    Avatar
        85     86                            Jurassic World
        76     77                              The Avengers
        54     55                            The Dark Knight
        12     13                                 Rogue One
        119   120                               Finding Dory
        94     95                   Avengers: Age of Ultron
        124   125                       The Dark Knight Rises
        578   579                The Hunger Games: Catching Fire


                             Genre  \
        50    Action,Adventure,Fantasy
        87    Action,Adventure,Fantasy
        85     Action,Adventure,Sci-Fi
        76             Action,Sci-Fi
        54         Action,Crime,Drama
        12     Action,Adventure,Sci-Fi
        119  Animation,Adventure,Comedy
        94     Action,Adventure,Sci-Fi
        124            Action,Thriller
        578   Action,Adventure,Mystery


                                Description           Director  \
```

```
50    Three decades after the defeat of the Galactic…         J.J. Abrams
87    A paraplegic marine dispatched to the moon Pan…       James Cameron
85    A new theme park, built on the original site o…      Colin Trevorrow
76    Earth's mightiest heroes must come together an…         Joss Whedon
54    When the menace known as the Joker wreaks havo…   Christopher Nolan
12    The Rebel Alliance makes a risky move to steal…      Gareth Edwards
119   The friendly but forgetful blue tang fish, Dor…     Andrew Stanton
94    When Tony Stark and Bruce Banner try to jump-s…         Joss Whedon
124   Eight years after the Joker's reign of anarchy…   Christopher Nolan
578   Katniss Everdeen and Peeta Mellark become targ…   Francis Lawrence

                                        Actors  Year  \
50    Daisy Ridley, John Boyega, Oscar Isaac, Domhna…  2015
87    Sam Worthington, Zoe Saldana, Sigourney Weaver…  2009
85    Chris Pratt, Bryce Dallas Howard, Ty Simpkins,…  2015
76    Robert Downey Jr., Chris Evans, Scarlett Johan…  2012
54    Christian Bale, Heath Ledger, Aaron Eckhart,Mi…  2008
12    Felicity Jones, Diego Luna, Alan Tudyk, Donnie…  2016
119   Ellen DeGeneres, Albert Brooks,Ed O'Neill, Kai…  2016
94    Robert Downey Jr., Chris Evans, Mark Ruffalo, …  2015
124   Christian Bale, Tom Hardy, Anne Hathaway,Gary …  2012
578   Jennifer Lawrence, Josh Hutcherson, Liam Hemsw…  2013

      Runtime (Minutes)  Rating     Votes  Revenue (Millions)  Metascore
50                  136     8.1    661608              936.63       81.0
87                  162     7.8    935408              760.51       83.0
85                  124     7.0    455169              652.18       59.0
76                  143     8.1   1045588              623.28       69.0
54                  152     9.0   1791916              533.32       82.0
12                  133     7.9    323118              532.17       65.0
119                  97     7.4    157026              486.29       77.0
94                  141     7.4    516895              458.99       66.0
124                 164     8.5   1222645              448.13       78.0
578                 146     7.6    525646              424.65       76.0
```

```python
[31]: top_10 = data.nlargest(10,'Revenue (Millions)')[['Title','Director','Revenue␣
      ↪(Millions)']].set_index('Title')
      top_10
```

```
[31]:                                                Director  \
      Title
      Star Wars: Episode VII - The Force Awakens         J.J. Abrams
      Avatar                                           James Cameron
      Jurassic World                                  Colin Trevorrow
      The Avengers                                        Joss Whedon
      The Dark Knight                               Christopher Nolan
      Rogue One                                        Gareth Edwards
```

```
Finding Dory                                    Andrew Stanton
Avengers: Age of Ultron                           Joss Whedon
The Dark Knight Rises                         Christopher Nolan
The Hunger Games: Catching Fire                Francis Lawrence


                                             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
```

[32]:
```python
# Visualization
sns.barplot(x='Revenue (Millions)',y=top_10.index,data=top_10)
plt.title("Display Top 10 Highest Revenue Movie Titles")
plt.show()
```
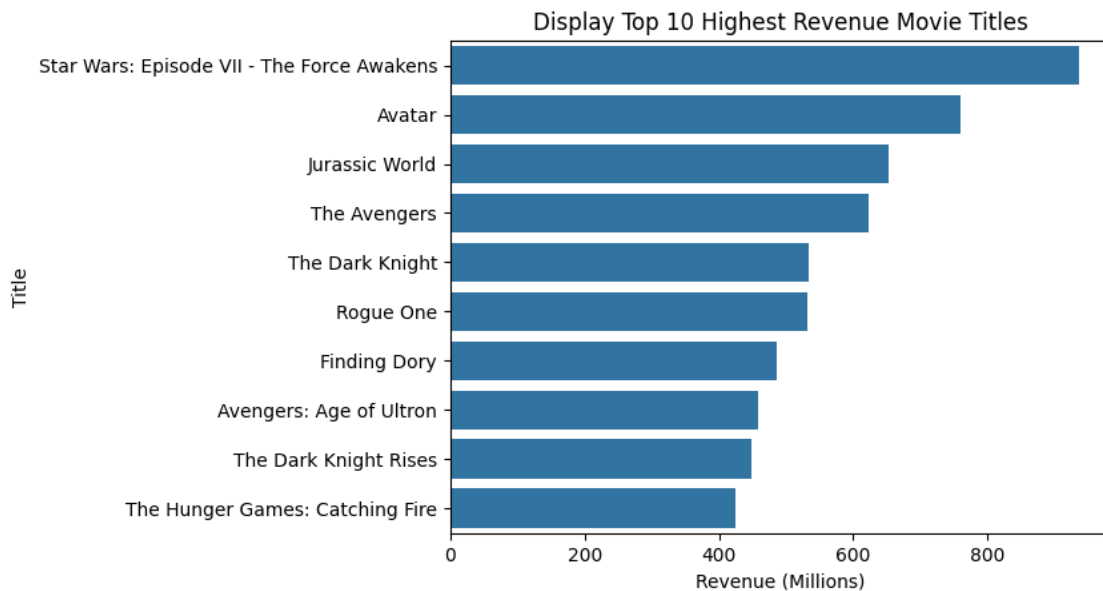
18 18. Find Average Rating of Movies Year-wise

```
[33]: data1=data.groupby('Year')['Rating'].mean().sort_values(ascending=False).
      ↪reset_index()
      data1
```

```
[33]:       Year      Rating
      0     2006   7.143902
      1     2007   7.140909
      2     2011   6.945614
      3     2012   6.933871
      4     2009   6.911111
      5     2010   6.894737
      6     2013   6.832143
      7     2014   6.822581
      8     2008   6.708333
      9     2015   6.674312
      10    2016   6.644444
```
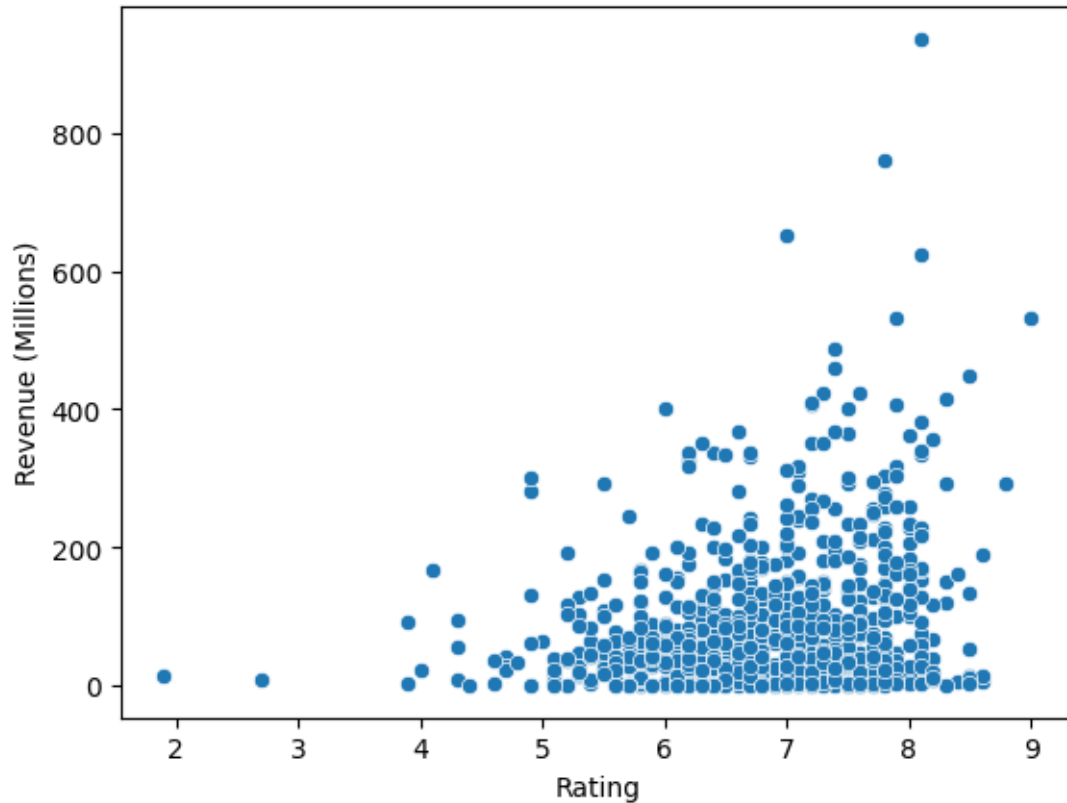
```
[34]: # visualization
      plt.figure(figsize=(25,10))
      sns.barplot(x='Year',y='Rating',data=data1)
      plt.title("Average Rating By Year")
      plt.show()
```
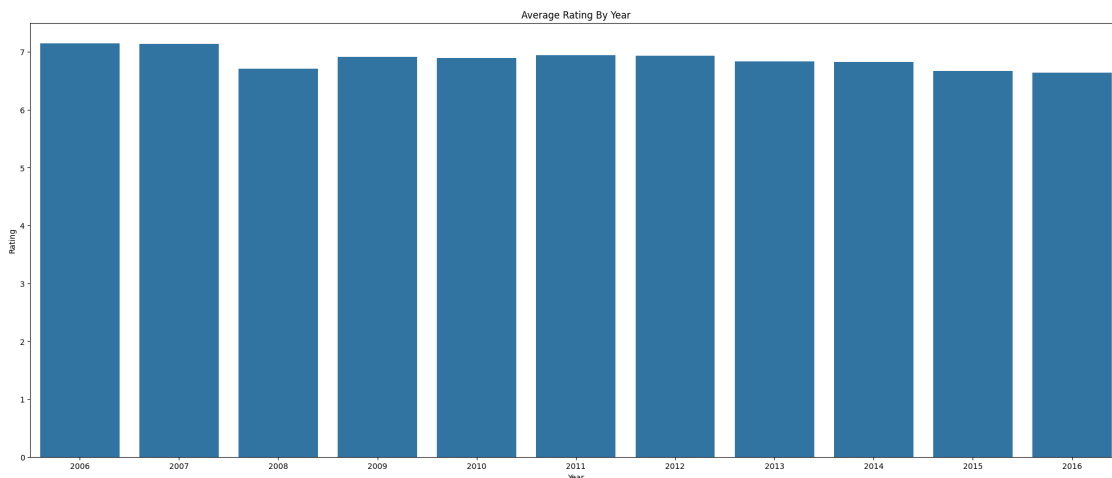

19 19. Does Rating Affect The Revenue?

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

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

20 Answer : Yes

21 20. Classify Movies Based on Ratings [Good,Better and Best]

```
[36]: def rating(rating):
          if rating>=8.0: # changed 7.0 -> 8.0
              return 'Excellent'
          elif rating>=6.5: # changed 6.0 -> 6.5
              return 'Good'
          else:
              return 'Average'
```

```
[37]: data['rating_cat']=data['Rating'].apply(rating)
      data
```

```
[37]:        Rank                     Title                      Genre  \
      0          1  Guardians of the Galaxy    Action,Adventure,Sci-Fi
      1          2               Prometheus  Adventure,Mystery,Sci-Fi
      2          3                    Split           Horror,Thriller
      3          4                     Sing    Animation,Comedy,Family
```

```
4       5                Suicide Squad   Action,Adventure,Fantasy
..      …                        …                         …
993    994   Resident Evil: Afterlife    Action,Adventure,Horror
994    995                    Project X                     Comedy
996    997             Hostel: Part II                      Horror
997    998       Step Up 2: The Streets       Drama,Music,Romance
999   1000                   Nine Lives     Comedy,Family,Fantasy

                                          Description            Director  \
0      A group of intergalactic criminals are forced …          James Gunn
1      Following clues to the origin of mankind, a te…        Ridley Scott
2      Three girls are kidnapped by a man with a diag…   M. Night Shyamalan
3      In a city of humanoid animals, a hustling thea…  Christophe Lourdelet
4      A secret government agency recruits some of th…          David Ayer
..                                            …                         …
993    While still out to destroy the evil Umbrella C…   Paul W.S. Anderson
994    3 high school seniors throw a birthday party t…     Nima Nourizadeh
996    Three American college students studying abroa…            Eli Roth
997    Romantic sparks occur between two dance studen…          Jon M. Chu
999    A stuffy businessman finds himself trapped ins…    Barry Sonnenfeld

                                         Actors   Year  \
0      Chris Pratt, Vin Diesel, Bradley Cooper, Zoe S…  2014
1      Noomi Rapace, Logan Marshall-Green, Michael Fa…  2012
2      James McAvoy, Anya Taylor-Joy, Haley Lu Richar…  2016
3      Matthew McConaughey,Reese Witherspoon, Seth Ma…  2016
4      Will Smith, Jared Leto, Margot Robbie, Viola D…  2016
..                                            …        …
993    Milla Jovovich, Ali Larter, Wentworth Miller,K…  2010
994    Thomas Mann, Oliver Cooper, Jonathan Daniel Br…  2012
996    Lauren German, Heather Matarazzo, Bijou Philli…  2007
997    Robert Hoffman, Briana Evigan, Cassie Ventura,…  2008
999    Kevin Spacey, Jennifer Garner, Robbie Amell,Ch…  2016

       Runtime (Minutes)  Rating    Votes  Revenue (Millions)  Metascore  \
0                    121     8.1   757074              333.13       76.0
1                    124     7.0   485820              126.46       65.0
2                    117     7.3   157606              138.12       62.0
3                    108     7.2    60545              270.32       59.0
4                    123     6.2   393727              325.02       40.0
..                     …       …        …                   …          …
993                   97     5.9   140900               60.13       37.0
994                   88     6.7   164088               54.72       48.0
996                   94     5.5    73152               17.54       46.0
997                   98     6.2    70699               58.01       50.0
999                   87     5.3    12435               19.64       11.0
```

```
      rating_cat
0     Excellent
1          Good
2          Good
3          Good
4       Average
..          …
993     Average
994        Good
996     Average
997     Average
999     Average

[838 rows x 13 columns]
```

22 21. Count Number of Action Movies

```
[38]: list1=[]
      for value in data['Genre']:
          list1.append(value.split(','))
```

```
[39]: data['temp']=list1
```

```
[40]: genre=input("Enter Genre you want to count : ").title()
      count=0
      for value in data['temp']:
          if genre in value:
              count=count+1
      print("Total Count is",count)
```

```
Total Count is 0
```

23 OR

```
[41]: len(data[data['Genre'].str.contains('action',case=False)])
```

```
[41]: 277
```

```
[ ]:
```