**MOVIE REPORT**

**INTRODUCTION**

This dataset is about 1000 most popular movies gotten from kaggle.com. People watch all kinds of movies for entertainment and relaxation purposes and so this dataset contains details on the kinds of movies, the titles, the actors in the movie, the number of votes for each movie, how each movie was rated, the numbers of minutes each movie lasted and the year they made the most revenue.

**OBJECTIVES**

The aim of this dataset is to gain insight into various aspects of the data by identifying trends and patterns in the movie industry overtime, which includes genres, rating, revenue etc.

I’ll also get insights into audience preferences like which actors are more popular, number of movies released per year, top directors of the highly rated movies, most popular movies etc.

**OVERVIEW OF THE DATASET**

There are 1000 rows and 12 columns in this dataset. Data cleaning was done in jupyter notebook(python codes). I checked for null values and dropped the rows with null values. I also checked for duplicates and there were no duplicates, all these you’ll in details as you progress.

**DATA VISUALIZATION**

This was done in python using seaborn and matplotlib.

**PROCESS FLOW**

**I started by importing pandas, matplotlib and seaborn as seen below:**

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

I imported the data from csv using:

**data = pd.read\_csv(r'C:\Users\HP USER\Desktop\IMDB-Movie-Data.csv')**

As part of the data cleaning process, I checked for the number of rows and columns using:

**data.shape**

(1000, 12)

**#display top 10 rows of the dataset**

data.head(10)

| **Rank** | **Title** | **Genre** | **Description** | **Director** | **Actors** | **Year** | **Runtime (Minutes)** | **Rating** | **Votes** | **Revenue (Millions)** | **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, Anya 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, Tian Jing, Willem Dafoe, Andy Lau | 2016 | 103 | 6.1 | 56036 | 45.13 | 42.0 |
| **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.... | 2016 | 128 | 8.3 | 258682 | 151.06 | 93.0 |
| **7** | 8 | Mindhorn | Comedy | A has-been actor best known for playing the ti... | Sean Foley | Essie Davis, Andrea Riseborough, Julian Barrat... | 2016 | 89 | 6.4 | 2490 | NaN | 71.0 |
| **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 | 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 |

***#last 10 rows***

**data.tail(10)**

|  | **Rank** | **Title** | **Genre** | **Description** | **Director** | **Actors** | **Year** | **Runtime (Minutes)** | **Rating** | **Votes** | **Revenue (Millions)** | **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 Nighy, Steven... | 2009 | 92 | 6.6 | 129708 | 45.80 | 44.0 |
| **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... | 2007 | 165 | 8.5 | 102697 | 1.20 | 42.0 |
| **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... | 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,K... | 2010 | 97 | 5.9 | 140900 | 60.13 | 37.0 |
| **994** | 995 | Project X | Comedy | 3 high school seniors throw a birthday party t... | 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, alo... | Billy Ray | Chiwetel Ejiofor, Nicole Kidman, Julia Roberts... | 2015 | 111 | 6.2 | 27585 | NaN | 45.0 |
| **996** | 997 | Hostel: Part II | Horror | Three American college students studying abroa... | Eli Roth | Lauren German, Heather Matarazzo, Bijou Philli... | 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 reuni... | Scot Armstrong | Adam Pally, T.J. Miller, Thomas Middleditch,Sh... | 2014 | 93 | 5.6 | 4881 | 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 |

***#check for missing values***

**data.isnull().sum()**

Rank 0

Title 0

Genre 0

Description 0

Director 0

Actors 0

Year 0

Runtime (Minutes) 0

Rating 0

Votes 0

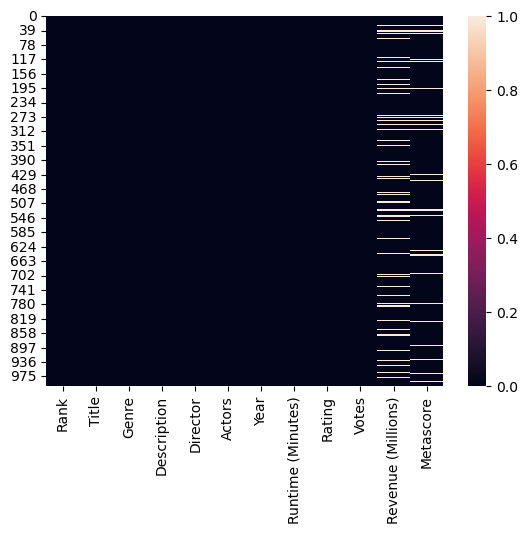
Revenue (Millions) 128

Metascore 64

dtype: int64

I used this to show a map of columns with missing values

**sns.heatmap(data.isnull())**



***#percentage of missing values***

**per\_missing = data.isnull().sum() \* 100/len(data)**

**per\_missing**

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

I used dropped missing rows

**data.dropna(axis = 0)**

|  | **Rank** | **Title** | **Genre** | **Description** | **Director** | **Actors** | **Year** | **Runtime (Minutes)** | **Rating** | **Votes** | **Revenue (Millions)** | **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, Anya 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,K... | 2010 | 97 | 5.9 | 140900 | 60.13 | 37.0 |
| **994** | 995 | Project X | Comedy | 3 high school seniors throw a birthday party t... | 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, Bijou Philli... | 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 × 12 columns

***#check for duplicates***

**dup\_data = data.duplicated().any()**

**Print("Are there any duplicate values?", dup\_data)**

**Are there any duplicate values? False**

***#get overall statistics about data:mean,min, max etc for numerical columns***

**data.describe()**

|  | **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 | 58.985043 |
| **Std** | 288.819436 | 3.205962 | 18.810908 | 0.945429 | 1.887626e+05 | 103.253540 | 17.194757 |
| **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 |

***#get overall statistics about data for all columns***

**data.describe(include = 'all')**

|  | **Rank** | **Title** | **Genre** | **Description** | **Director** | **Actors** | **Year** | **Runtime (Minutes)** | **Rating** | **Votes** | **Revenue (Millions)** | **Metascore** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **count** | 1000.000000 | 1000 | 1000 | 1000 | 1000 | 1000 | 1000.000000 | 1000.000000 | 1000.000000 | 1.000000e+03 | 872.000000 | 936.000000 |
| **unique** | NaN | 999 | 207 | 1000 | 644 | 996 | NaN | NaN | NaN | NaN | NaN | NaN |
| **top** | NaN | The Host | Action,Adventure,Sci-Fi | A group of intergalactic criminals are forced ... | Ridley Scott | Jennifer Lawrence, Josh Hutcherson, Liam Hemsw... | NaN | NaN | NaN | NaN | NaN | NaN |
| **freq** | NaN | 2 | 50 | 1 | 8 | 2 | NaN | NaN | NaN | NaN | NaN | NaN |
| **mean** | 500.500000 | NaN | NaN | NaN | NaN | NaN | 2012.783000 | 113.172000 | 6.723200 | 1.698083e+05 | 82.956376 | 58.985043 |
| **std** | 288.819436 | NaN | NaN | NaN | NaN | NaN | 3.205962 | 18.810908 | 0.945429 | 1.887626e+05 | 103.253540 | 17.194757 |
| **min** | 1.000000 | NaN | NaN | NaN | NaN | NaN | 2006.000000 | 66.000000 | 1.900000 | 6.100000e+01 | 0.000000 | 11.000000 |
| **25%** | 250.750000 | NaN | NaN | NaN | NaN | NaN | 2010.000000 | 100.000000 | 6.200000 | 3.630900e+04 | 13.270000 | 47.000000 |
| **50%** | 500.500000 | NaN | NaN | NaN | NaN | NaN | 2014.000000 | 111.000000 | 6.800000 | 1.107990e+05 | 47.985000 | 59.500000 |
| **75%** | 750.250000 | NaN | NaN | NaN | NaN | NaN | 2016.000000 | 123.000000 | 7.400000 | 2.399098e+05 | 113.715000 | 72.000000 |
| **max** | 1000.000000 | NaN | NaN | NaN | NaN | NaN | 2016.000000 | 191.000000 | 9.000000 | 1.791916e+06 | 936.630000 | 100.000000 |

**QUESTIONS FORMULATED:**

***#Title of movie having runtime >=180mins***

**data[data['Runtime (Minutes)']>= 180]['Title']**

82 The Wolf of Wall Street

88 The Hateful Eight

311 La vie d'Adèle

828 Grindhouse

965 Inland Empire

Name: Title, dtype: object

*#show variation btw year had the highest voting*

data.groupby('Year')['Votes'].mean().sort\_values(ascending **=** **False**)

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

*#show variation btw year had the highest average voting*

avg\_votes\_per\_yr **=** data.groupby('Year')['Votes'].mean()

plt.figure(figsize **=** (10,6))

plt.bar(avg\_votes\_per\_yr.index, avg\_votes\_per\_yr.values, color **=** 'red')

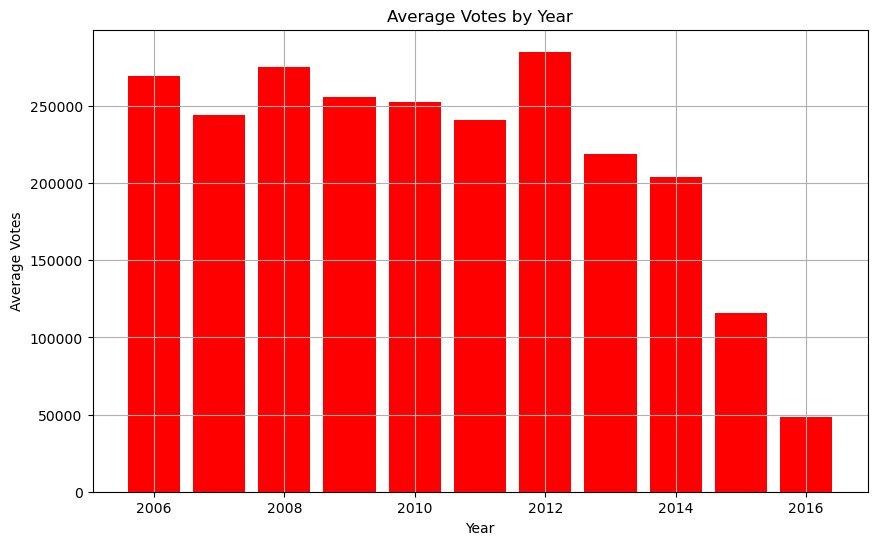
plt.xlabel('Year')

plt.ylabel('Average Votes')

plt.title('Average Votes by Year')

plt.grid(**True**)

plt.show()



*#which year had the highest average revenue*

data.groupby('Year')['Revenue (Millions)'].mean().sort\_values(ascending **=** **False**)

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

*#Graphical display of variation between year and highest average revenue*

avg\_revenue\_per\_yr **=** data.groupby('Year')['Revenue (Millions)'].mean()

plt.figure(figsize **=** (10,6))

plt.barh(avg\_revenue\_per\_yr.index, avg\_revenue\_per\_yr.values, color **=** 'darkblue')

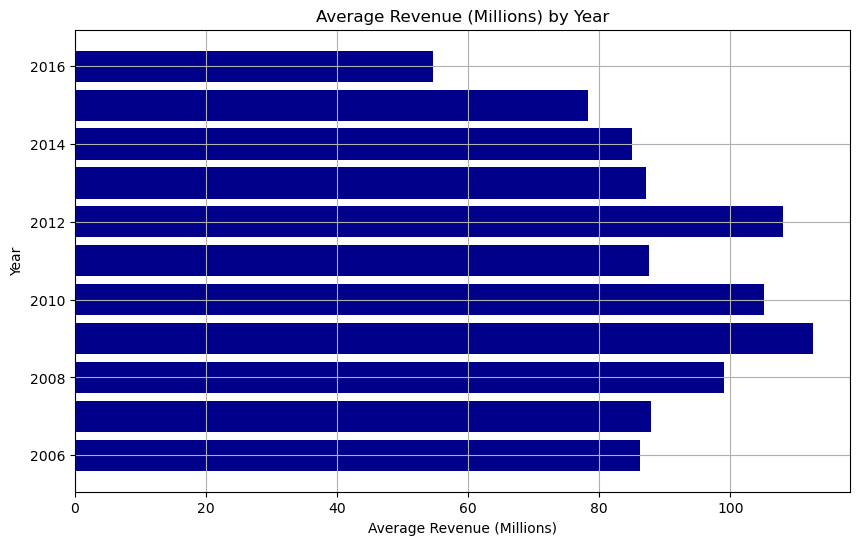
plt.xlabel('Average Revenue (Millions)')

plt.ylabel('Year')

plt.title('Average Revenue (Millions) by Year')

plt.grid(**True**)

plt.show()



*#average rating for top 10 directors*

mean\_rating\_per\_director **=** data.groupby('Director')['Rating'].mean()

top\_10\_directors **=** mean\_rating\_per\_director.sort\_values(ascending**=False**).head(10)

print(top\_10\_directors)

Director

Nitesh Tiwari 8.80

Christopher Nolan 8.68

Olivier Nakache 8.60

Makoto Shinkai 8.60

Aamir Khan 8.50

Florian Henckel von Donnersmarck 8.50

Naoko Yamada 8.40

Damien Chazelle 8.40

Lee Unkrich 8.30

Amber Tamblyn 8.30

Name: Rating, dtype: float64

plt.figure(figsize **=** (10,6))

top\_10\_directors.plot(kind **=** 'bar')

plt.xlabel('Director')

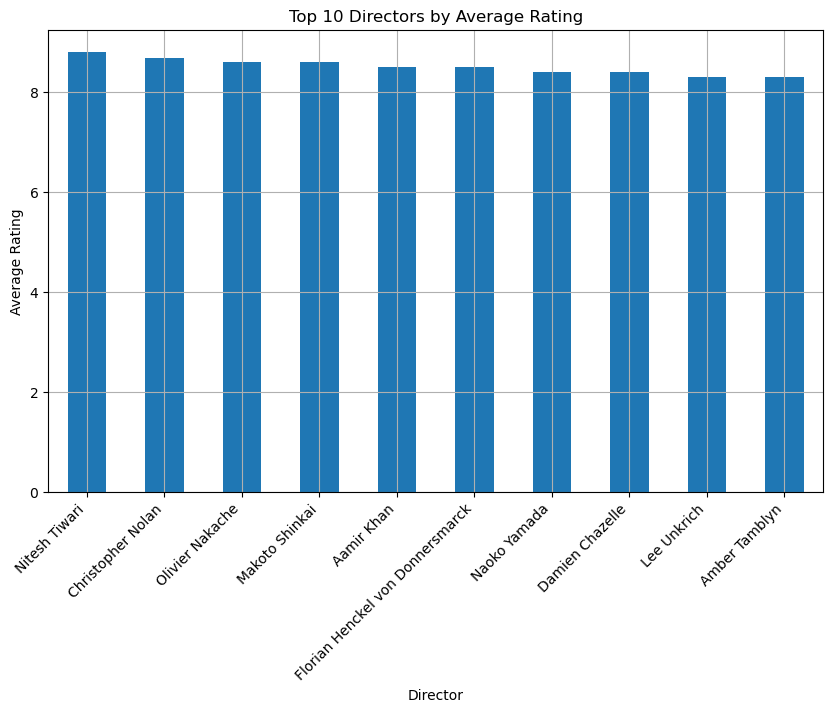
plt.ylabel('Average Rating')

plt.title('Top 10 Directors by Average Rating')

plt.xticks(rotation**=**45, ha **=** 'right')

plt.grid(**True**)

plt.show()



*#top 10 lengthy movies and runtime*

Top10\_len**=**data.nlargest(10,'Runtime (Minutes)' )[['Title', 'Runtime (Minutes)']].set\_index('Title')

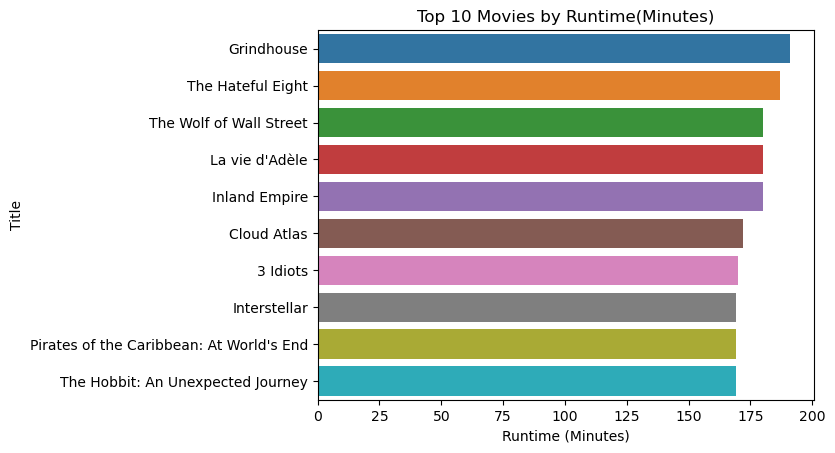
Top10\_len

|  | **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 |

sns.barplot(x**=**'Runtime (Minutes)', y**=**Top10\_len.index, data**=**Top10\_len)

plt.title('Top 10 Movies by Runtime(Minutes)')

plt.show()



*#how many movies were watched in each year*

data['Year'].value\_counts()

Year

2016 297

2015 127

2014 98

2013 91

2012 64

2011 63

2010 60

2007 53

2008 52

2009 51

2006 44

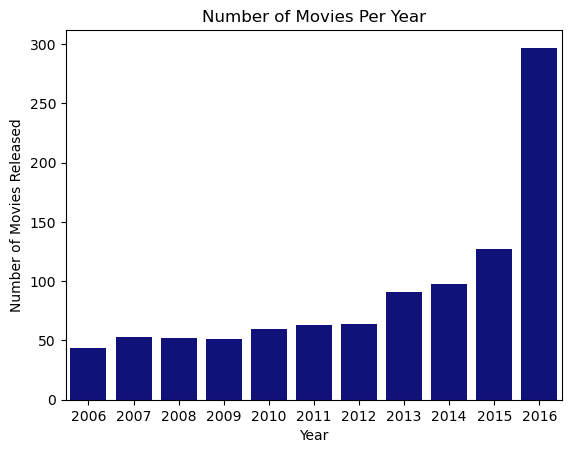
Name: count, dtype: int64

sns.countplot(x**=**'Year', data**=**data, color **=** 'darkblue')

plt.ylabel('Number of Movies Released')

plt.title('Number of Movies Per Year')

plt.show()



*#most popular movie title*

data[data['Revenue (Millions)'].max()**==**data['Revenue (Millions)']]['Title']

50 Star Wars: Episode VII - The Force Awakens

Name: Title, dtype: object

*#top 10 highest rated movie titles and its directors*

Top10\_len**=**data.nlargest(10,'Rating' )[['Title', 'Rating', 'Director']].set\_index('Title')

Top10\_len

​

|  | **Rating** | **Director** |
| --- | --- | --- |
| **Title** |  |  |
| **The Dark Knight** | 9.0 | Christopher Nolan |
| **Inception** | 8.8 | Christopher Nolan |
| **Dangal** | 8.8 | Nitesh Tiwari |
| **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 |

sns.barplot(x**=**Top10\_len.index, y**=**'Rating', data**=**Top10\_len, hue**=**'Director', dodge**=False**)

plt.legend(bbox\_to\_anchor**=**(1.05,1), loc**=**2)

plt.xlabel('Title')

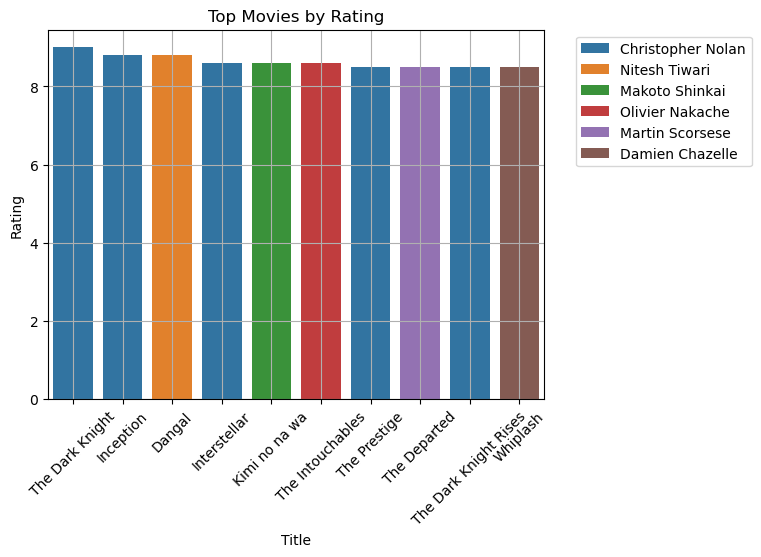
plt.ylabel('Rating')

plt.title('Top Movies by Rating')

plt.xticks(rotation **=** 45)

plt.grid(**True**)

plt.show()



*#top 10 highest revenue movie titles*

Top\_10**=**data.nlargest(10, 'Revenue (Millions)')[['Title','Revenue (Millions)']].set\_index('Title')

Top\_10

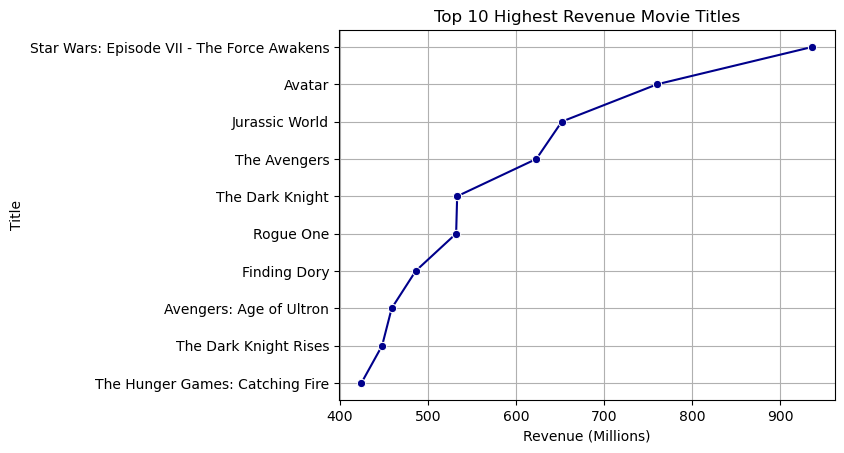
|  | **Revenue (Millions)** |
| --- | --- |
| **Title** |  |
| **Star Wars: Episode VII - The Force Awakens** | 936.63 |
| **Avatar** | 760.51 |
| **Jurassic World** | 652.18 |
| **The Avengers** | 623.28 |
| **The Dark Knight** | 533.32 |
| **Rogue One** | 532.17 |
| **Finding Dory** | 486.29 |
| **Avengers: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

sns.lineplot(x**=**'Revenue (Millions)',y**=**Top\_10.index, data**=**Top\_10, marker **=**'o', color **=** 'darkblue')

plt.grid(**True**)

plt.title('Top 10 Highest Revenue Movie Titles')

plt.show()



*#Average rating of movies Year wise*

data.groupby('Year')['Rating'].mean().sort\_values(ascending **=** **False**)

​

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

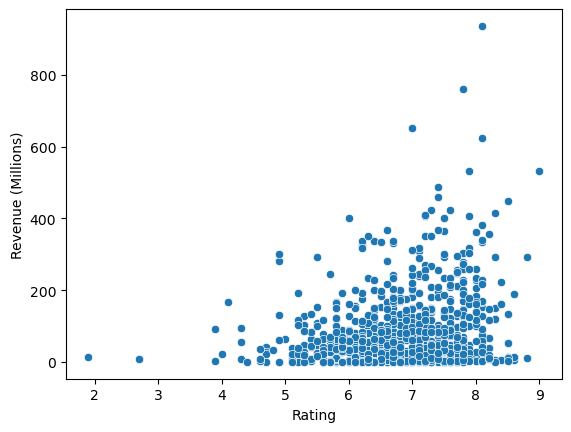
Name: Rating, dtype: float64

In [8]:

*#Does rating affect revenue*

sns.scatterplot(x**=**'Rating',y**=**'Revenue (Millions)', data**=**data)

plt.show()



*#Yes because as Rating is increasing, Revenue is also increasing*

*#classify movies based on excellent, good and average*

*#using user defined function*

**def** rating(rating):

**if** rating**>=**7.0:

**return** "Excellent"

**elif** rating**>=**6.0:

**return** "Good"

**else**:

**return** "Average"

data['rating\_cat']**=**data['Rating'].apply(rating)

data.head()

|  | **Rank** | **Title** | **Genre** | **Description** | **Director** | **Actors** | **Year** | **Runtime (Minutes)** | **Rating** | **Votes** | **Revenue (Millions)** | **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, Anya 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 |

data['Genre'].dtype

dtype('O')

*#count number of action movies*

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

303

**ANALYSIS & RECOMMENDATIONS**

**From the scatterplot, movies that were rated the highest, generated the**

**most revenue.**

**Films of the top 2 actors that should be looked out for are “Christopher**

**Nolan” and “Nitesh Tiwan” because they generated the most revenue.**