Analysis on IMDB dataset

In [231]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
```

In [143]:
```
df=pd.read_csv(r"C:\Users\Jaswanth Reddy\Downloads\IMDB-movie data\IMDB-movie data\IMDB-Movie-Data.csv")
df.head()
```

Out[143]:

	Rank	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	M
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	
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	
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	
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	
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	

In [144]:
```
df.isnull().sum()
```

Out[144]:
```
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
```

In [145]:
```
df=df.dropna()
df.shape
```

Out[145]:
```
(838, 12)
```

In [147]:
```
df=df.drop(columns=['Rank'])
df.describe()
```

Out[147]:

	Year	Runtime (Minutes)	Rating	Votes	Revenue (Millions)	Metascore
count	838.00000	838.000000	838.000000	8.380000e+02	838.000000	838.000000
mean	2012.50716	114.638425	6.814320	1.932303e+05	84.564558	59.575179
std	3.17236	18.470922	0.877754	1.930990e+05	104.520227	16.952416
min	2006.00000	66.000000	1.900000	1.780000e+02	0.000000	11.000000
25%	2010.00000	101.000000	6.300000	6.127650e+04	13.967500	47.000000
50%	2013.00000	112.000000	6.900000	1.368795e+05	48.150000	60.000000
75%	2015.00000	124.000000	7.500000	2.710830e+05	116.800000	72.000000
max	2016.00000	187.000000	9.000000	1.791916e+06	936.630000	100.000000

Average rating given for the movies from the year 2006-2016 was 6.8 and maximum rating of 9.0

Average runtime for the movies from 2006-2016 was 114.64 minutes and maximum runtime(minutes) for a movie is 187 minutes

Average revenue for the movies given from 2006-2016 was 84.56 Millions and maximum Revenue of 936.63 Millions

In [148]:
```
df.shape
```

Out[148]:
```
(838, 11)
```

In [130]:
```
# Analyzing Director directed movies per year
a=pd.crosstab(df.Director,df.Year,margins=True,margins_name='Total')
a.head(50)
```

Out[130]:

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
Director												
Aamir Khan	0	1	0	0	0	0	0	0	0	0	0	1
Abdellatif Kechiche	0	0	0	0	0	0	0	1	0	0	0	1
Adam Leon	0	0	0	0	0	0	0	0	0	0	1	1
Adam McKay	1	0	1	0	1	0	0	0	0	1	0	4
Adam Shankman	0	1	0	0	0	0	1	0	0	0	0	2
Adam Wingard	0	0	0	0	0	0	0	0	1	0	1	2
Afonso Poyart	0	0	0	0	0	0	0	0	0	1	0	1
Aisling Walsh	0	0	0	0	0	0	0	0	0	0	1	1
Akan Satayev	0	0	0	0	0	0	0	0	0	0	1	1
Akiva Schaffer	0	0	0	0	0	0	0	0	0	0	1	1
Alan Taylor	0	0	0	0	0	0	0	1	0	1	0	2
Albert Hughes	0	0	0	0	1	0	0	0	0	0	0	1
Alejandro Amenábar	0	0	0	0	0	0	0	0	0	1	0	1
Alejandro González Iñárritu	1	0	0	0	0	0	0	0	1	1	0	3
Alessandro Carloni	0	0	0	0	0	0	0	0	0	0	1	1
Alex Garland	0	0	0	0	0	0	0	0	1	0	0	1
Alex Proyas	0	0	0	0	0	0	0	0	0	0	1	1
Alex Ranarivelo	0	0	0	0	0	0	0	0	0	0	1	1
Alexander Payne	0	0	0	0	0	1	0	0	0	0	0	1
Alexandre Aja	1	0	0	0	1	0	0	1	0	0	1	4
Alexandros Avranas	0	0	0	0	0	0	0	0	0	0	1	1
Alexi Pappas	0	0	0	0	0	0	0	0	0	0	1	1

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
Director												
Alfonso Cuarón	1	0	0	0	0	0	0	1	0	0	0	2
Alfonso Gomez-Rejon	0	0	0	0	0	0	0	0	0	1	0	1
Allen Coulter	0	0	0	0	1	0	0	0	0	0	0	1
Amber Tamblyn	0	0	0	0	0	0	0	0	0	0	1	1
Amma Asante	0	0	0	0	0	0	0	0	0	0	1	1
Ana Lily Amirpour	0	0	0	0	0	0	0	0	0	0	1	1
Andrea Arnold	0	0	0	0	0	0	0	0	0	0	1	1
Andrew Dominik	0	1	0	0	0	0	0	0	0	0	0	1
Andrew Jarecki	0	0	0	0	1	0	0	0	0	0	0	1
Andrew Niccol	0	0	0	0	0	1	0	1	0	0	0	2
Andrew Stanton	0	0	1	0	0	0	1	0	0	0	1	3
Andrey Kravchuk	0	0	0	0	0	0	0	0	0	0	1	1
André Øvredal	0	0	0	0	0	0	0	0	0	0	1	1
Andrés Muschietti	0	0	0	0	0	0	0	1	0	0	0	1
Andy Fickman	1	0	0	0	0	0	0	0	0	0	0	1
Andy Goddard	0	0	0	0	0	0	0	0	0	0	1	1
Andy Tennant	0	0	1	0	0	0	0	0	0	0	0	1
Ang Lee	0	0	0	0	0	0	1	0	0	0	1	2
Angelina Jolie	0	0	0	0	0	0	0	0	1	0	0	1
Anna Biller	0	0	0	0	0	0	0	0	0	0	1	1
Anna Foerster	0	0	0	0	0	0	0	0	0	0	1	1
Anne Fletcher	1	0	0	1	0	0	0	0	0	0	0	2
Anne Fontaine	0	0	0	0	0	0	0	1	0	0	0	1
Anthony Russo	0	0	0	0	0	0	0	0	1	0	1	2
Antoine Fuqua	0	1	0	0	0	0	0	1	1	1	1	5
Antonio Campos	0	0	0	0	0	0	0	0	0	0	1	1

Year	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
Director												
April Mullen	0	0	0	0	0	0	0	0	0	0	1	1
Ari Sandel	0	0	0	0	0	0	0	0	0	1	0	1

Conclusion:- Per year Director directs maximum of one movie

In [3]:
```
# Analysis based on yearly metrics(Total movie count per year)

total_moviecount=df['Year'].value_counts()
total_moviecount
```

Out[3]:
```
2016    297
2015    127
2014     98
2013     91
2012     64
2011     63
2010     60
2007     53
2008     52
2009     51
2006     44
Name: Year, dtype: int64
```

Conclusion:- From 2006-2016 the movie count per year increased gradually and in year 2016(revolutionary year) there was a huge number of movies relased

In [149]:
```
# Renaming the column name
df=df.rename(columns={'Revenue (Millions)':'Revenue'})
df
```

Out[149]:

	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue	Metas
0	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	
1	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	
2	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	
3	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	
4	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	
...	...	...	...	...	...	...	...	...	...	...	
993	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	
994	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	

	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue	Metas
996	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	
997	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	
999	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	

838 rows × 11 columns

In [6]:
```
# Sum of money in each year(Millions)

df.groupby('Year').Revenue.sum()
```

Out[6]:
```
Year
2006     3624.46
2007     4306.23
2008     5053.22
2009     5292.26
2010     5989.65
2011     5431.96
2012     6910.29
2013     7666.72
2014     7997.40
2015     8854.12
2016    11211.65
Name: Revenue, dtype: float64
```

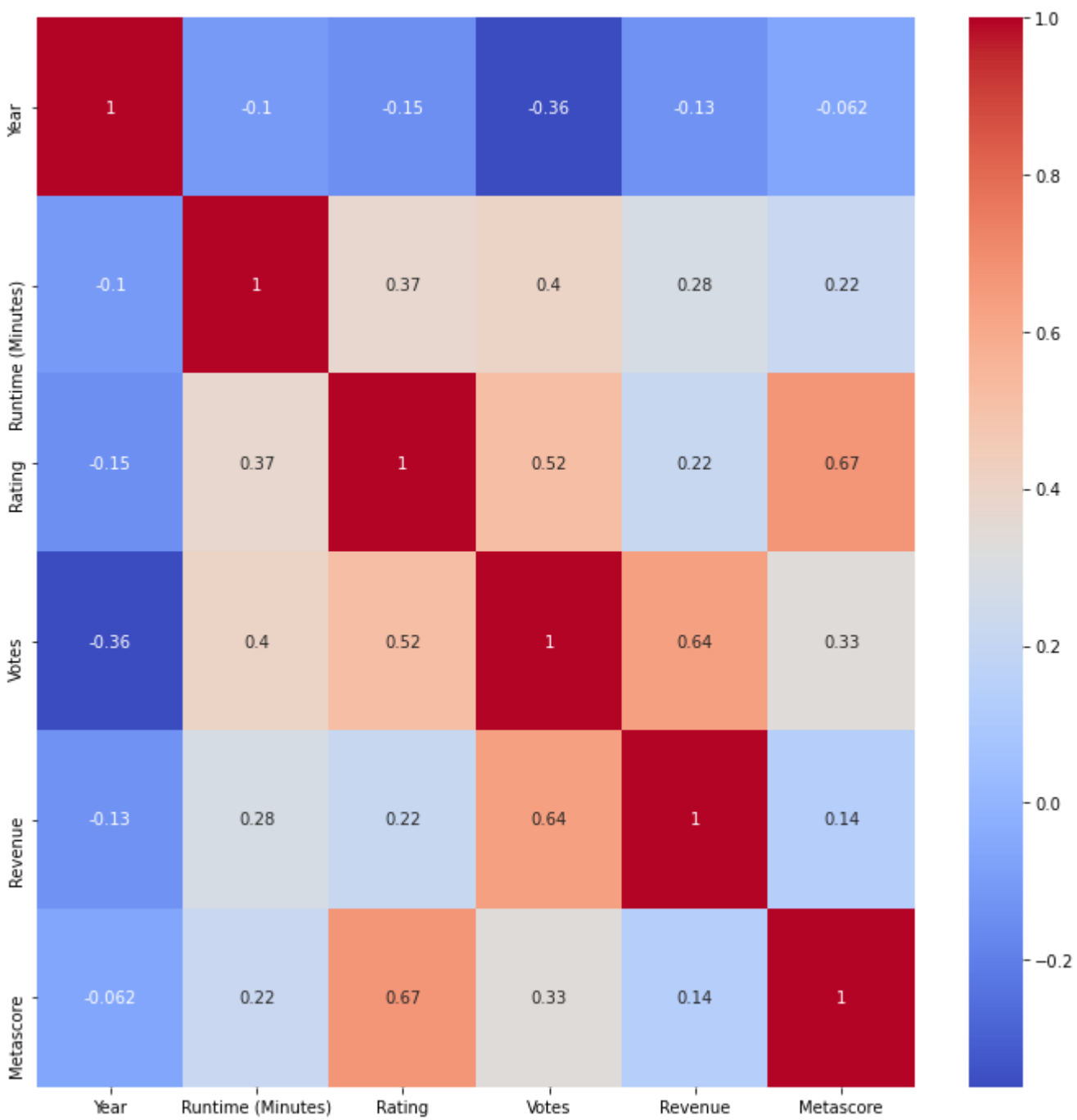
In [307]:
```
yr=[2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016]
g=df.groupby(['Year'])
for i in (yr):
    print(g.get_group(i)['Votes'].sum())
```

```
11366521
11727351
12795847
12023126
14881727
14777520
18033412
18944679
19709200
14116879
13550689
```

Conclusion:- From above cell we can conclude that revenue obtained each year increased gradually but in year 2011 there was less revenue due to less number of votes in that year .

In [269]:
```
plt.figure(figsize=(12,12))
seaborn.heatmap(df.corr(),annot=True,cmap="coolwarm")
```

Out[269]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x2042e81b280>
```



Conclusion:-From Heatmap it is clear that as, movie runtime increases votes,rating

Conclusion: From Heatmap it is clear that as, movie runtime increases votes,rating increases.

For rating- Metascore,votes,runtime effects a lot in positive side and revenue in a slight positive side.

For revenue - Votes play a major role and after votes movie runtime,rating and metascore plays a secondary role.

In [58]:
```python
# Calculating what type of movies released most

percentage_mov_genre=pd.crosstab(index=df.Genre,columns=df.Genre,margins=True)
percentage_mov_genre.head(50)
```

Out[58]:

lorror	Action,Adventure,Mystery	...	Mystery,Sci-Fi,Thriller	Mystery,Thriller	Mystery,Thriller,Western	Romance,Sci-Fi	Romance,Sci-Fi,Thriller	Sci-Fi	Sci-Fi,Thriller
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
2	0	...	0	0	0	0	0	0	0
0	5	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0

lorror	Action,Adventure,Mystery	...	Mystery,Sci-Fi,Thriller	Mystery,Thriller	Mystery,Thriller,Western	Romance,Sci-Fi	Romance,Sci-Fi,Thriller	Sci-Fi	Sci-Fi,Thriller
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0

lorror	Action,Adventure,Mystery	...	Mystery,Sci-Fi,Thriller	Mystery,Thriller	Mystery,Thriller,Western	Romance,Sci-Fi	Romance,Sci-Fi,Thriller	Sci-Fi	Sci-Fi,Thriller
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0
0	0	...	0	0	0	0	0	0	0

In [63]:
```
percentage_mov_genre.sort_values(['All'],ascending=False)
```

Out[63]:

Genre	Action	Action,Adventure	Action,Adventure,Biography	Action,Adventure,Comedy	Action,Adventure,Crime	Action,A
Genre						
All	2	3	2	14	6	
Action,Adventure,Sci-Fi	0	0	0	0	0	
Drama	0	0	0	0	0	
Comedy,Drama,Romance	0	0	0	0	0	
Comedy	0	0	0	0	0	
...	...	...	...	...	...	
Comedy,Drama,Musical	0	0	0	0	0	
Comedy,Crime,Thriller	0	0	0	0	0	
Action,Drama,War	0	0	0	0	0	
Action,Fantasy	0	0	0	0	0	
Comedy,Music,Romance	0	0	0	0	0	

208 rows × 208 columns

Conclusion:- Considering all the Genre(2006-16) mostly movie with a combination of Action,Adventure,Sci-Fi were released most and next Drama based moview were released most

Revenue and Genre(for each Genre) Analysis

In [234]:
```
# Detailed Analysis

percentage_rev_genre=pd.crosstab(df.Revenue,df.Genre,margins=True)
percentage_rev_genre
```

Out[234]:

Genre	Action	Action,Adventure	Action,Adventure,Biography	Action,Adventure,Comedy	Action,Adventure,Crime	Action,Adventure,Drama
Revenue						
0.0	0	0	0	0	0	0
0.01	0	0	0	0	0	0
0.02	0	0	0	0	0	0
0.03	0	0	0	0	0	0
0.04	0	0	0	0	0	0
...	...	...	...	...	...	...
623.28	0	0	0	0	0	0
652.18	0	0	0	0	0	0
760.51	0	0	0	0	0	0
936.63	0	0	0	0	0	0
All	1	3	2	14	6	17

790 rows × 190 columns

Conclustion:- Most revenue is earned from the movie of type Action,Adventure,Fantasy combination

Rating and Genre

In [251]:
```
# Detailed view on movie rating for each Genre

percentage_mov_genre=pd.crosstab(df.Genre,df.Rating,margins=True)
percentage_mov_genre.sort_values(['All'],ascending=False)
```

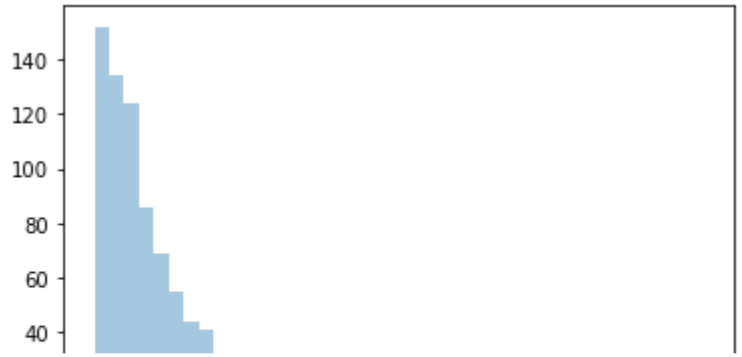
Out[251]:

Rating	1.9	2.7	3.9	4.0	4.1	4.3	4.4	4.6	4.7	4.8	...	8.0	8.1	8.2	8.3	8.4	8.5	8.6	8.8	9.0	All
Genre																					
All	1	1	2	1	1	3	1	2	3	1	...	19	24	9	5	2	6	3	1	1	838
Action,Adventure,Sci-Fi	0	0	0	0	0	1	0	0	0	0	...	2	2	0	0	0	0	0	1	0	50
Comedy,Drama,Romance	0	0	0	0	0	1	0	0	0	0	...	0	0	1	0	0	0	0	0	0	30
Drama	0	0	0	0	0	0	0	0	0	0	...	1	2	1	1	0	0	0	0	0	29
Drama,Romance	0	0	0	0	0	0	0	0	0	0	...	1	0	0	0	0	0	0	0	0	27
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Animation,Comedy,Drama	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
Animation,Adventure,Family	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
Animation,Action,Comedy	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1
Adventure,Horror	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0	0	0	0	1

Conclusion:- From the above it is clear that Action,Adventure,Sci-Fi combination movie has a highest rating and the people love the most of that category

In [232]:
```
a=['Year','Runtime (Minutes)','Rating','Votes','Revenue','Metascore']
for i in a:
    print(df[i].describe())
    sns.distplot(df[i],kde=False)
    plt.show()
```

```
count    8.380000e+02
mean     1.932303e+05
std      1.930990e+05
min      1.780000e+02
25%      6.127650e+04
50%      1.368795e+05
75%      2.710830e+05
max      1.791916e+06
Name: Votes, dtype: float64
```



In [252]:
```
# Detailed view on movie rating for each Genre

percentage_vote_genre=pd.crosstab(df.Genre,df.Votes,margins=True)
percentage_vote_genre.sort_values(['All'],ascending=False)
```

Out[252]:

Votes	178	277	279	291	391	552	616	664	702	1024	...	935408	937414	959065	1039115	1045588	104774
Genre																	
All	1	1	1	1	1	1	1	1	1	1	...	1	1	1	1	1	
Action,Adventure,Sci-Fi	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
Comedy,Drama,Romance	1	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
Drama	0	0	1	1	0	0	1	0	0	0	...	0	0	0	0	0	
Drama,Romance	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	..
Animation,Comedy,Drama	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
Animation,Adventure,Family	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
Animation,Action,Comedy	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
Adventure,Horror	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	
Action	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	

190 rows × 838 columns

Conclusion:- From above cell it is clear that Action,Adventure,Sci-Fi movie combination has highest votes

In [254]:
```
# Detailed view on movie title for each Genre

percentage_movie_genre=pd.crosstab(df.Genre,df.Title,margins=True)
percentage_movie_genre.sort_values(['All'],ascending=False)
```

Out[254]:

Title	(500) Days of Summer	10 Cloverfield Lane	12 Years a Slave	127 Hours	13 Hours	1408	17 Again	2012	20th Century Women	21	...	You Don't Mess with the Zohan	Your Highness	Youth	Z M Po
Genre															
All	1	1	1	1	1	1	1	1	1	1	...	1	1	1	
Action,Adventure,Sci-Fi	0	0	0	0	0	0	0	1	0	0	...	0	0	0	
Comedy,Drama,Romance	1	0	0	0	0	0	0	0	0	0	...	0	0	0	
Drama	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Drama,Romance	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
Animation,Comedy,Drama	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Animation,Adventure,Family	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Animation,Action,Comedy	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Adventure,Horror	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
Action	0	0	0	0	0	0	0	0	0	0	...	0	0	0	

190 rows × 838 columns

From the above cell, we can conlude that most of the movies released are type Action,Adventure,SciFi combined movie with a total count of 50

Analyzing Director and his movies

In [268]:
```
df.groupby('Director').Revenue.agg(['sum','count']).sort_values(['sum'],ascending=False)
```

Out[268]:

Director	sum	count
J.J. Abrams	1683.45	5
David Yates	1630.51	6
Christopher Nolan	1515.09	5
Michael Bay	1421.32	6
Francis Lawrence	1299.81	4
...	...	...
Gus Van Sant	0.02	1
Robin Swicord	0.01	1
So Yong Kim	0.01	1
Patricia Rozema	0.01	1
Andy Goddard	0.00	1

524 rows × 2 columns

From the above cell it is clear that J.J.Abrams is the director with the highest movie collection of 1683.45 million of his entire movies direction.

In [310]:
```
a=df.groupby('Director')
a.get_group('J.J. Abrams')
```

Out[310]:

	Title	Genre	Description	Director	Actors	Year	Runtime (Minutes)	Rating	Votes	Revenue	Metascore
50	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	8.1	661608	936.63	81.0
140	Star Trek	Action,Adventure,Sci-Fi	The brash James T. Kirk tries to live up to hi...	J.J. Abrams	Chris Pine, Zachary Quinto, Simon Pegg, Leonar...	2009	127	8.0	526324	257.70	82.0
362	Star Trek Into Darkness	Action,Adventure,Sci-Fi	After the crew of the Enterprise find an unsto...	J.J. Abrams	Chris Pine, Zachary Quinto, Zoe Saldana, Bened...	2013	132	7.8	417663	228.76	72.0
497	Super 8	Mystery,Sci-Fi,Thriller	During the summer of 1979, a group of friends ...	J.J. Abrams	Elle Fanning, AJ Michalka, Kyle Chandler, Joel...	2011	112	7.1	298913	126.98	72.0
869	Mission: Impossible III	Action,Adventure,Thriller	Agent Ethan Hunt comes into conflict with a da...	J.J. Abrams	Tom Cruise, Michelle Monaghan, Ving Rhames, Ph...	2006	126	6.9	270429	133.38	66.0

Final conclusion :- Revenue from a movie mainly depends on votes of the people and next comes runtime,rating,metadata etc and mostly people like to watch a movie combined with Action, Adventure, Sci-fi and fantasy elements and even the top collected

movies directed by J.J.Abrams are also based on Action,Adventure,Scifi,Fantasy based movies.

From this i conclude that movies with this elements will be blockbuster in the box office and even rating ,votings,metascore will be high for this type of movies.

In [ ]: