

In [1]:

```
import pandas as pd
```

In [2]:

```
import seaborn as sns
```

In [3]:

```
import matplotlib.pyplot as plt
```

In [4]:

```
import os

%matplotlib inline

import warnings
warnings.filterwarnings('ignore')
```

In [5]:

```
mov=pd.read_csv(r"C:\Users\user\Downloads\25th,26th\MOVIE RATINGS _ ADVANCE VISUALIZA")
```

In [6]:

```
mov
```

Out[6]:

	Film	Genre	Rotten Tomatoes Ratings %	Audience Ratings %	Budget (million \$)	Year of release
0	(500) Days of Summer	Comedy	87	81	8	2009
1	10,000 B.C.	Adventure	9	44	105	2008
2	12 Rounds	Action	30	52	20	2009
3	127 Hours	Adventure	93	84	18	2010
4	17 Again	Comedy	55	70	20	2009
...
554	Your Highness	Comedy	26	36	50	2011
555	Youth in Revolt	Comedy	68	52	18	2009
556	Zodiac	Thriller	89	73	65	2007
557	Zombieland	Action	90	87	24	2009
558	Zookeeper	Comedy	14	42	80	2011

559 rows × 6 columns

In [7]:

```
mov.shape
```

Out[7]:

```
(559, 6)
```

In [8]:

```
len(mov)
```

Out[8]:

```
559
```

In [9]:

```
mov.columns
```

Out[9]:

```
Index(['Film', 'Genre', 'Rotten Tomatoes Ratings %', 'Audience Ratings %',  
      'Budget (million $)', 'Year of release'],  
      dtype='object')
```

In [10]:

```
mov.columns=['Film','Genre','CriticRating','AudienceRateing','BudgetMillions','Year']  
mov.columns
```

Out[10]:

```
Index(['Film', 'Genre', 'CriticRating', 'AudienceRateing', 'BudgetMillio  
ns',  
      'Year'],  
      dtype='object')
```

In [11]:

```
mov.head()
```

Out[11]:

	Film	Genre	CriticRating	AudienceRateing	BudgetMillions	Year
0	(500) Days of Summer	Comedy	87	81	8	2009
1	10,000 B.C.	Adventure	9	44	105	2008
2	12 Rounds	Action	30	52	20	2009
3	127 Hours	Adventure	93	84	18	2010
4	17 Again	Comedy	55	70	20	2009

In [12]:

mov.info

Out[12]:

```

<bound method DataFrame.info of
riticRating  AudienceRateing \
0      (500) Days of Summer      Comedy      87      81
1      10,000 B.C.      Adventure      9      44
2      12 Rounds      Action      30      52
3      127 Hours      Adventure      93      84
4      17 Again      Comedy      55      70
..      ...      ...      ...
554      Your Highness      Comedy      26      36
555      Youth in Revolt      Comedy      68      52
556      Zodiac      Thriller      89      73
557      Zombieland      Action      90      87
558      Zookeeper      Comedy      14      42

      BudgetMillions  Year
0      8  2009
1      105  2008
2      20  2009
3      18  2010
4      20  2009
..      ...  ...
554      50  2011
555      18  2009
556      65  2007
557      24  2009
558      80  2011

```

[559 rows x 6 columns]>

In [13]:

mov.describe()

Out[13]:

	CriticRating	AudienceRateing	BudgetMillions	Year
count	559.000000	559.000000	559.000000	559.000000
mean	47.309481	58.744186	50.236136	2009.152057
std	26.413091	16.826887	48.731817	1.362632
min	0.000000	0.000000	0.000000	2007.000000
25%	25.000000	47.000000	20.000000	2008.000000
50%	46.000000	58.000000	35.000000	2009.000000
75%	70.000000	72.000000	65.000000	2010.000000
max	97.000000	96.000000	300.000000	2011.000000

In [14]:

```
mov['Film']
```

Out[14]:

```
0      (500) Days of Summer
1      10,000 B.C.
2      12 Rounds
3      127 Hours
4      17 Again

...
554     Your Highness
555     Youth in Revolt
556     Zodiac
557     Zombieland
558     Zookeeper
Name: Film, Length: 559, dtype: object
```

In [15]:

```
mov['Genre']
```

Out[15]:

```
0      Comedy
1      Adventure
2      Action
3      Adventure
4      Comedy

...
554     Comedy
555     Comedy
556     Thriller
557     Action
558     Comedy
Name: Genre, Length: 559, dtype: object
```

In [16]:

```
mov.Film
```

Out[16]:

```
0      (500) Days of Summer
1      10,000 B.C.
2      12 Rounds
3      127 Hours
4      17 Again

...
554     Your Highness
555     Youth in Revolt
556     Zodiac
557     Zombieland
558     Zookeeper
Name: Film, Length: 559, dtype: object
```

In [17]:

```
mov.Genre
```

Out[17]:

```
0      Comedy
1      Adventure
2      Action
3      Adventure
4      Comedy
...
554     Comedy
555     Comedy
556     Thriller
557     Action
558     Comedy
Name: Genre, Length: 559, dtype: object
```

In [18]:

```
mov.Film=mov.Film.astype('category')
```

In [19]:

```
mov.Film
```

Out[19]:

```
0      (500) Days of Summer
1      10,000 B.C.
2      12 Rounds
3      127 Hours
4      17 Again
...
554     Your Highness
555     Youth in Revolt
556     Zodiac
557     Zombieland
558     Zookeeper
Name: Film, Length: 559, dtype: category
Categories (559, object): ['(500) Days of Summer ', '10,000 B.C.', '12 R
ounds ', '127 Hours', ..., 'Youth in Revolt', 'Zodiac', 'Zombieland ',
'Zookeeper']
```

In [20]:

```
mov.Genre=mov.Genre.astype('category')
```

In [21]:

```
mov.Genre
```

Out[21]:

```
0      Comedy
1    Adventure
2      Action
3    Adventure
4      Comedy
...
554    Comedy
555    Comedy
556  Thriller
557    Action
558    Comedy
Name: Genre, Length: 559, dtype: category
Categories (7, object): ['Action', 'Adventure', 'Comedy', 'Drama', 'Horror', 'Romance', 'Thriller']
```

In [22]:

```
mov.Year
```

Out[22]:

```
0      2009
1      2008
2      2009
3      2010
4      2009
...
554    2011
555    2009
556    2007
557    2009
558    2011
Name: Year, Length: 559, dtype: int64
```

In [23]:

```
mov.Year=mov.Year.astype('category')
```

In [24]:

mov.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 559 entries, 0 to 558
Data columns (total 6 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   Film                  559 non-null   category
 1   Genre                 559 non-null   category
 2   CriticRating          559 non-null   int64
 3   AudienceRateing       559 non-null   int64
 4   BudgetMillions        559 non-null   int64
 5   Year                  559 non-null   category
dtypes: category(3), int64(3)
memory usage: 36.5 KB
```

In [25]:

mov.Genre.cat.categories

Out[25]:

```
Index(['Action', 'Adventure', 'Comedy', 'Drama', 'Horror', 'Romance',
      'Thriller'],
      dtype='object')
```

In [26]:

mov.head()

Out[26]:

	Film	Genre	CriticRating	AudienceRateing	BudgetMillions	Year
0	(500) Days of Summer	Comedy	87	81	8	2009
1	10,000 B.C.	Adventure	9	44	105	2008
2	12 Rounds	Action	30	52	20	2009
3	127 Hours	Adventure	93	84	18	2010
4	17 Again	Comedy	55	70	20	2009

In [27]:

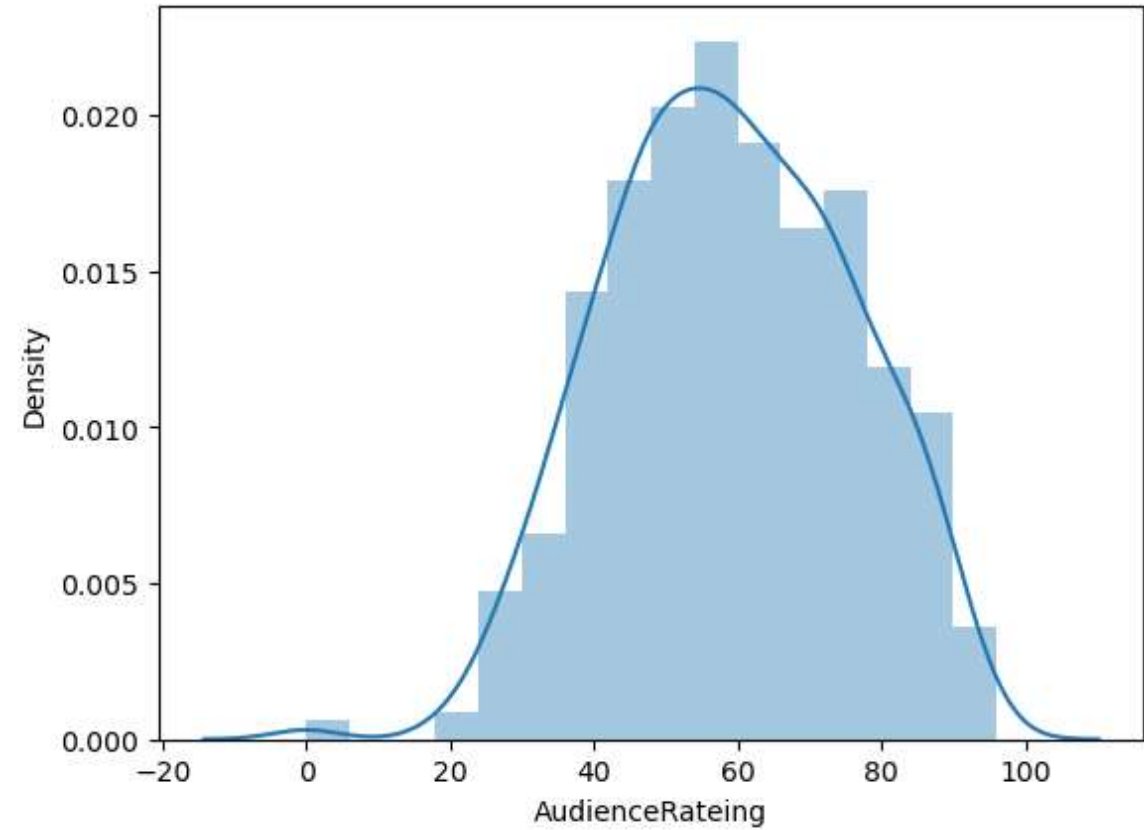
```
mov.describe()
```

Out[27]:

	CriticRating	AudienceRateing	BudgetMillions
count	559.000000	559.000000	559.000000
mean	47.309481	58.744186	50.236136
std	26.413091	16.826887	48.731817
min	0.000000	0.000000	0.000000
25%	25.000000	47.000000	20.000000
50%	46.000000	58.000000	35.000000
75%	70.000000	72.000000	65.000000
max	97.000000	96.000000	300.000000

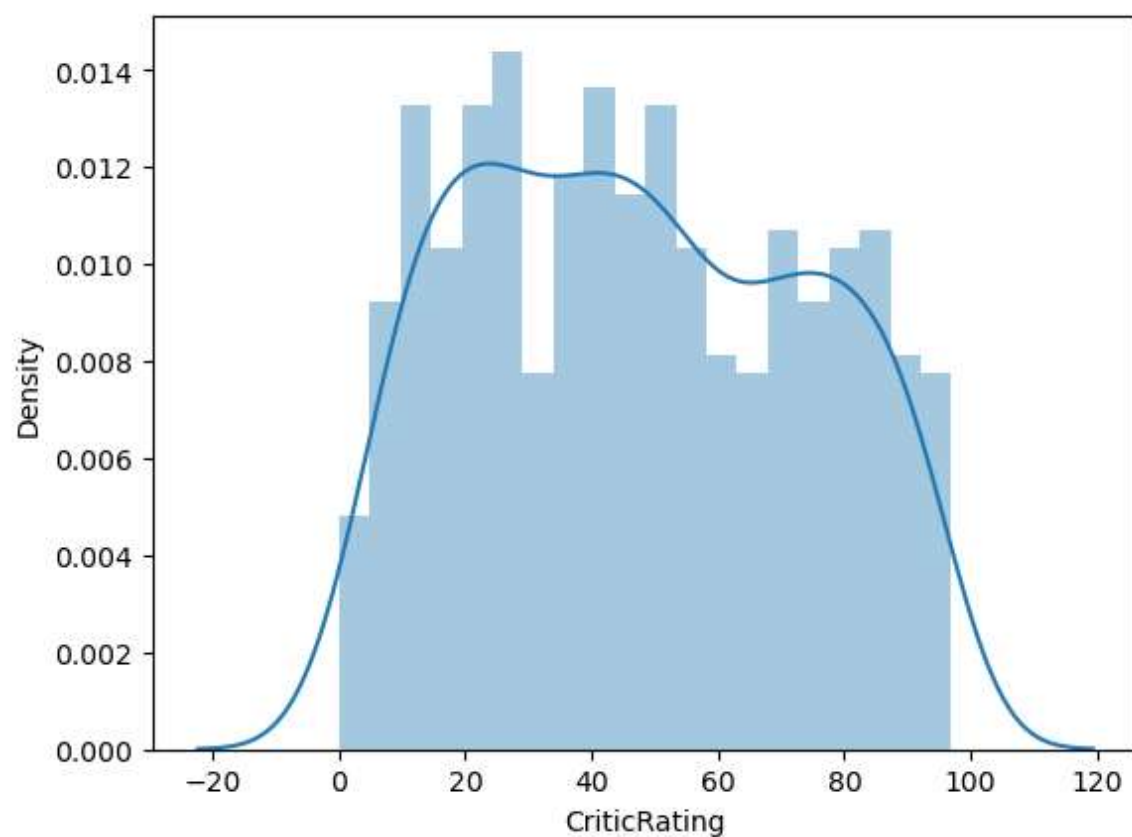
In [28]:

```
m = sns.distplot(mov.AudienceRateing)
```



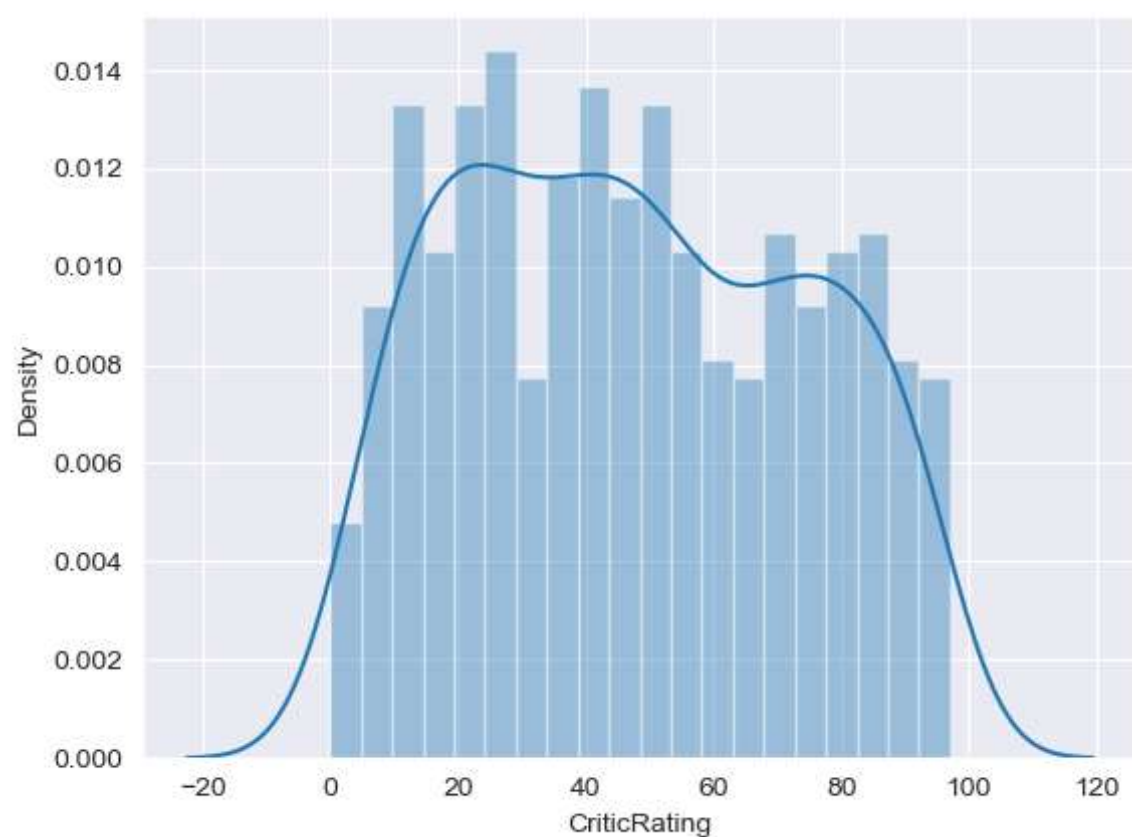
In [29]:

```
m1=sns.distplot(mov.CriticRating , bins=20)
```



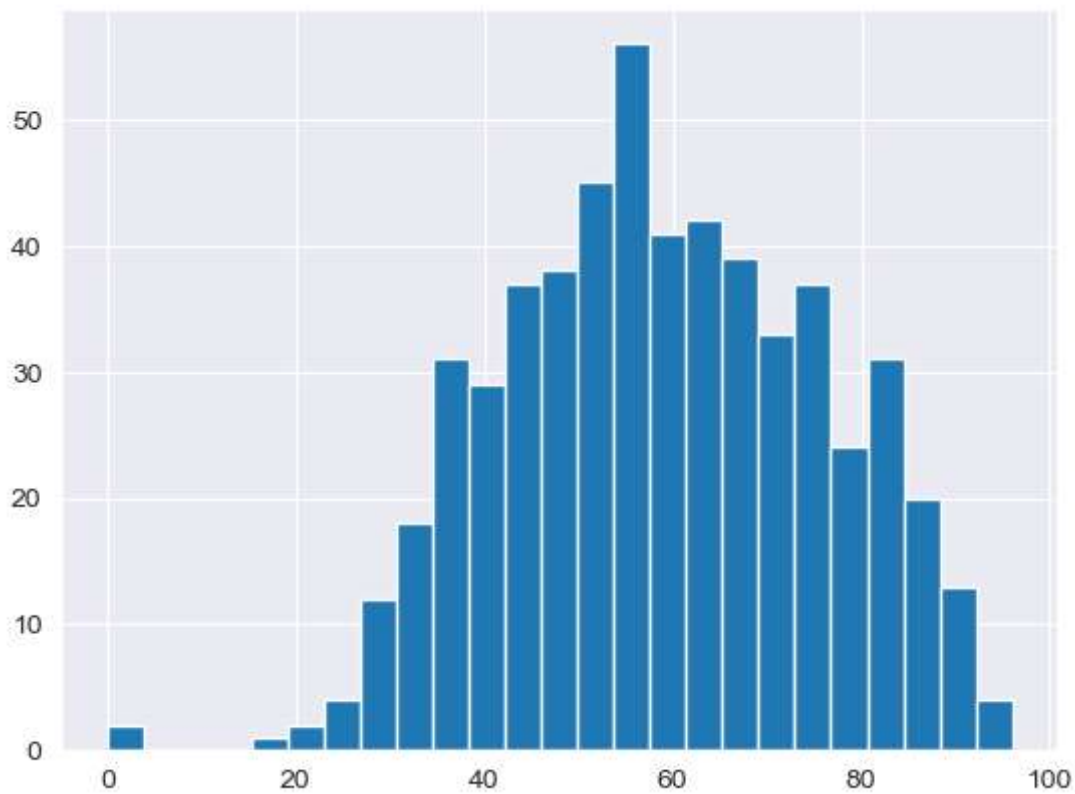
In [30]:

```
sns.set_style('darkgrid')  
m2=sns.distplot(mov.CriticRating , bins=20)
```



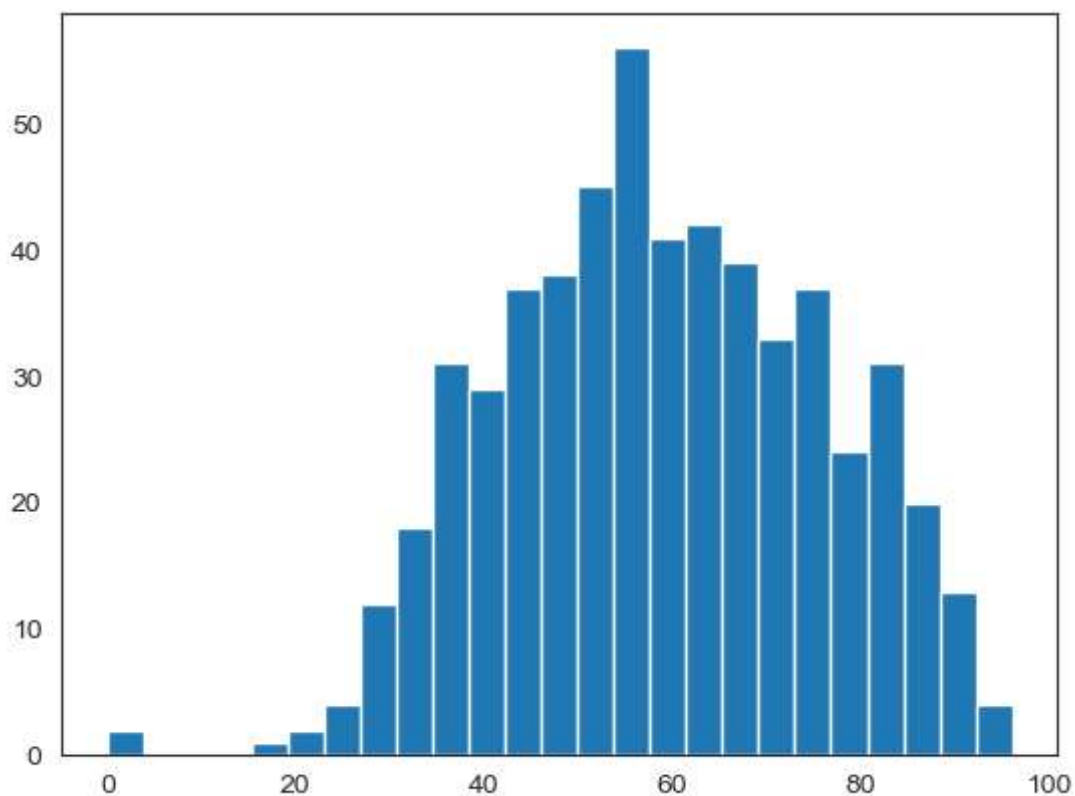
In [31]:

```
sns.set_style('darkgrid')  
n1=plt.hist(mov.AudienceRateing,bins=25)
```



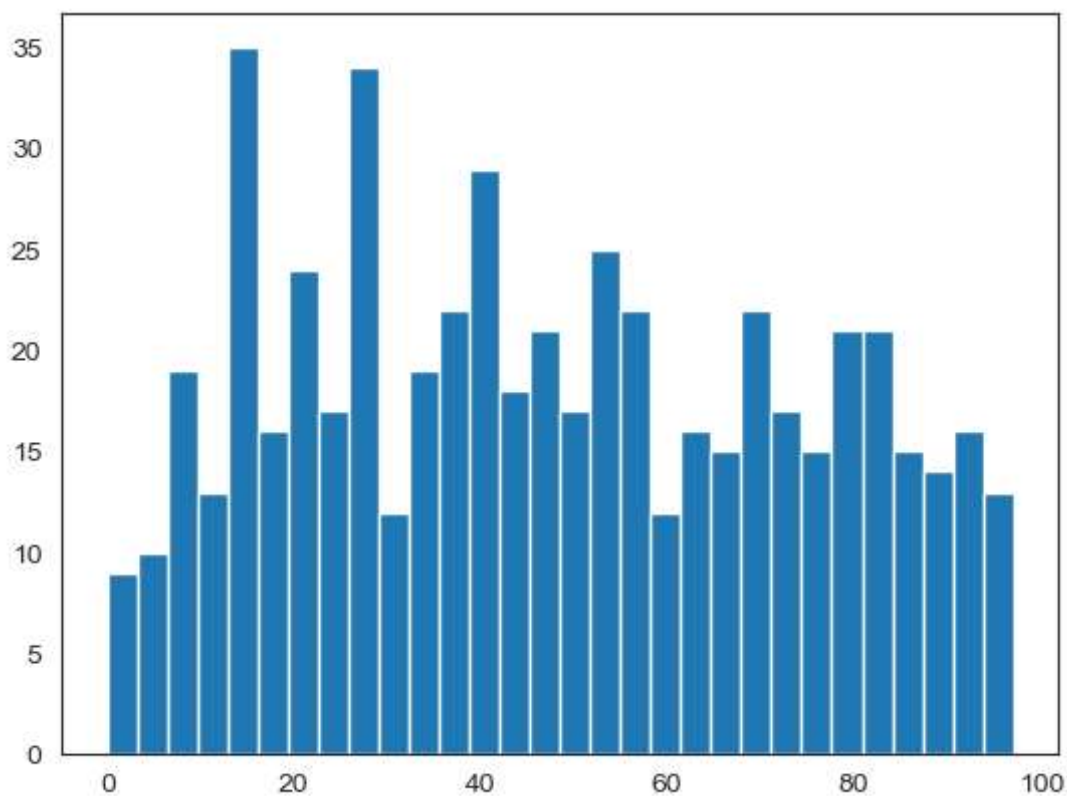
In [32]:

```
sns.set_style('white')  
n1=plt.hist(mov.AudienceRateing,bins=25)
```



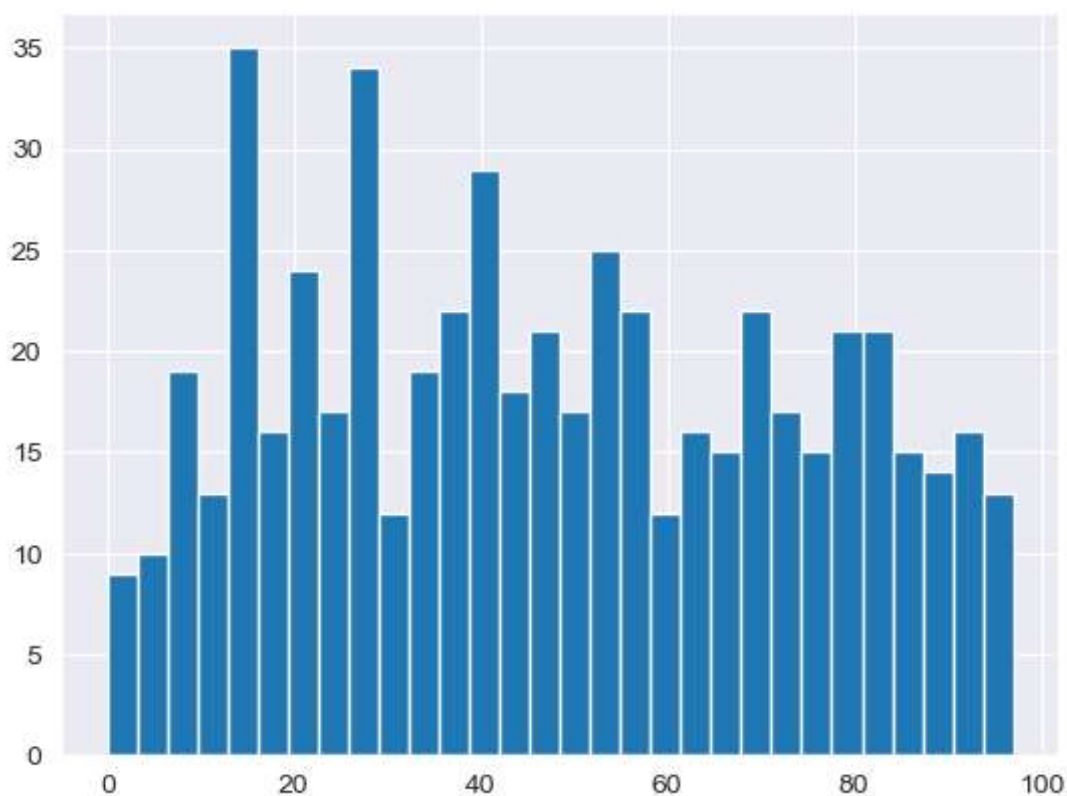
In [33]:

```
n2=plt.hist(mov.CriticRating,bins=30)
```



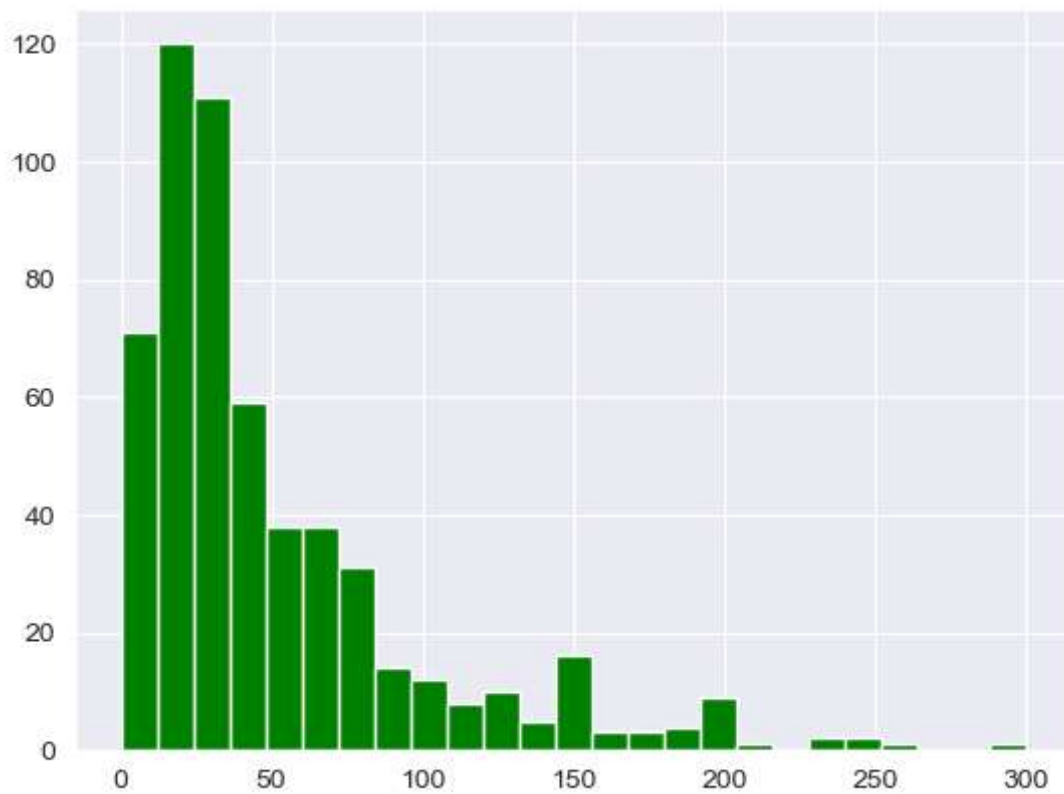
In [34]:

```
sns.set_style('darkgrid')  
n2=plt.hist(mov.CriticRating,bins=30)  
plt.show()
```



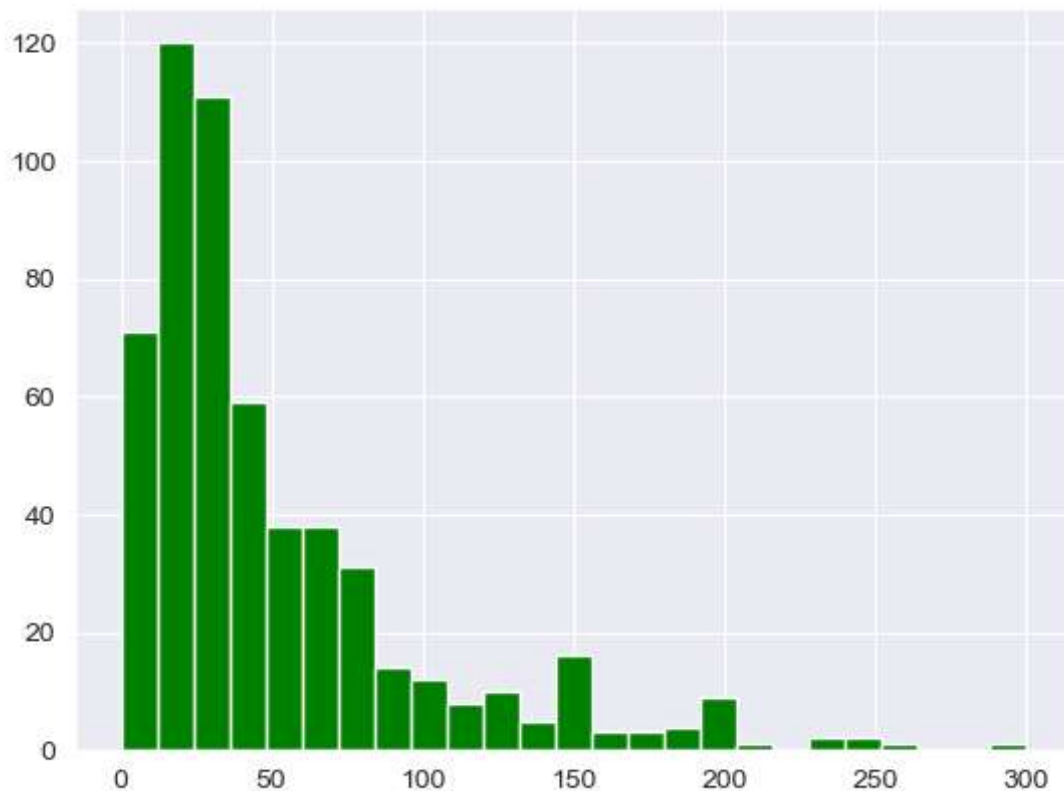
In [35]:

```
n3=plt.hist(mov.BudgetMillions,bins=25,color='green')  
plt.show()
```



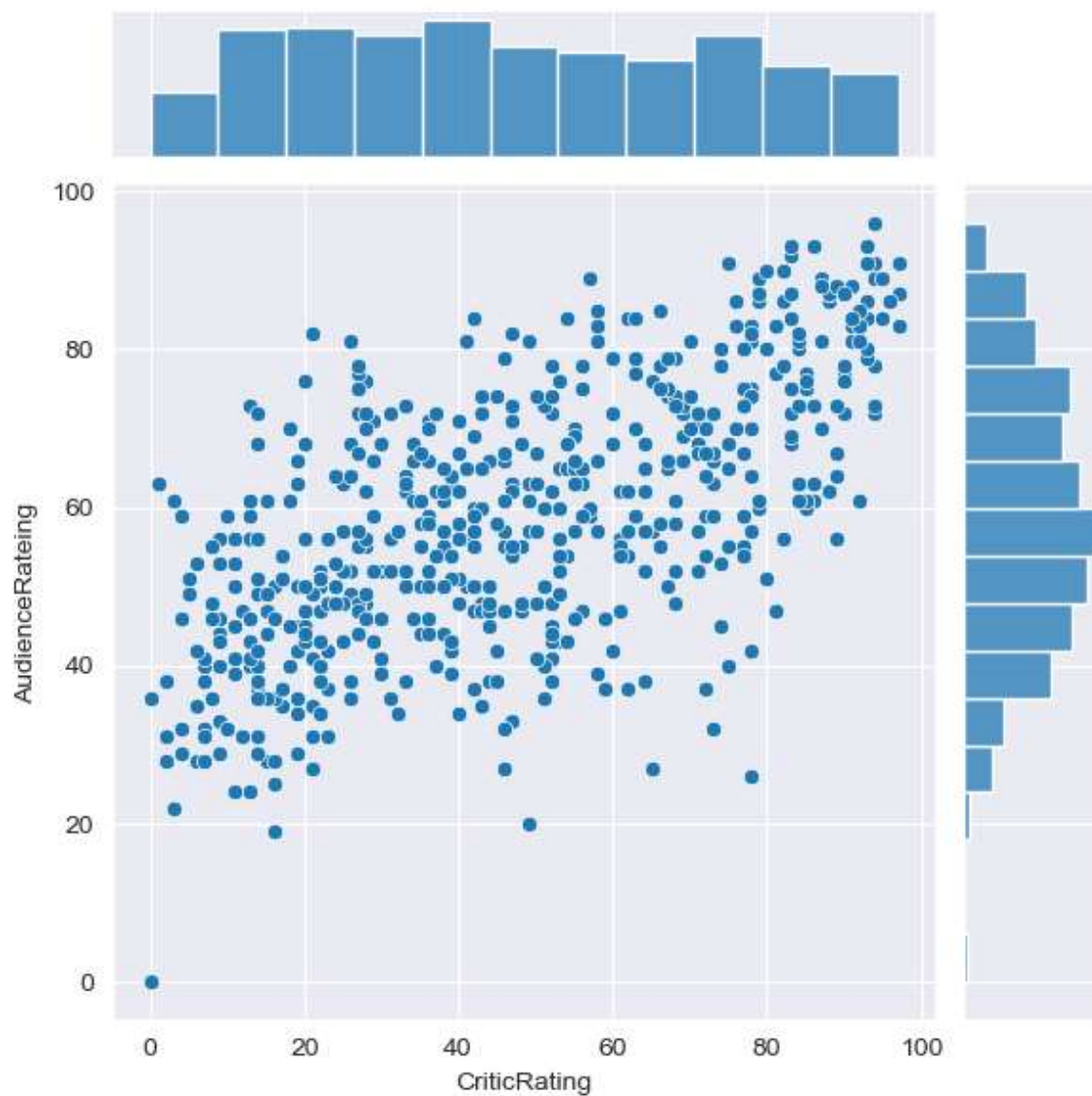
In [36]:

```
n3=plt.hist(mov.BudgetMillions,bins=25,color='green')
```



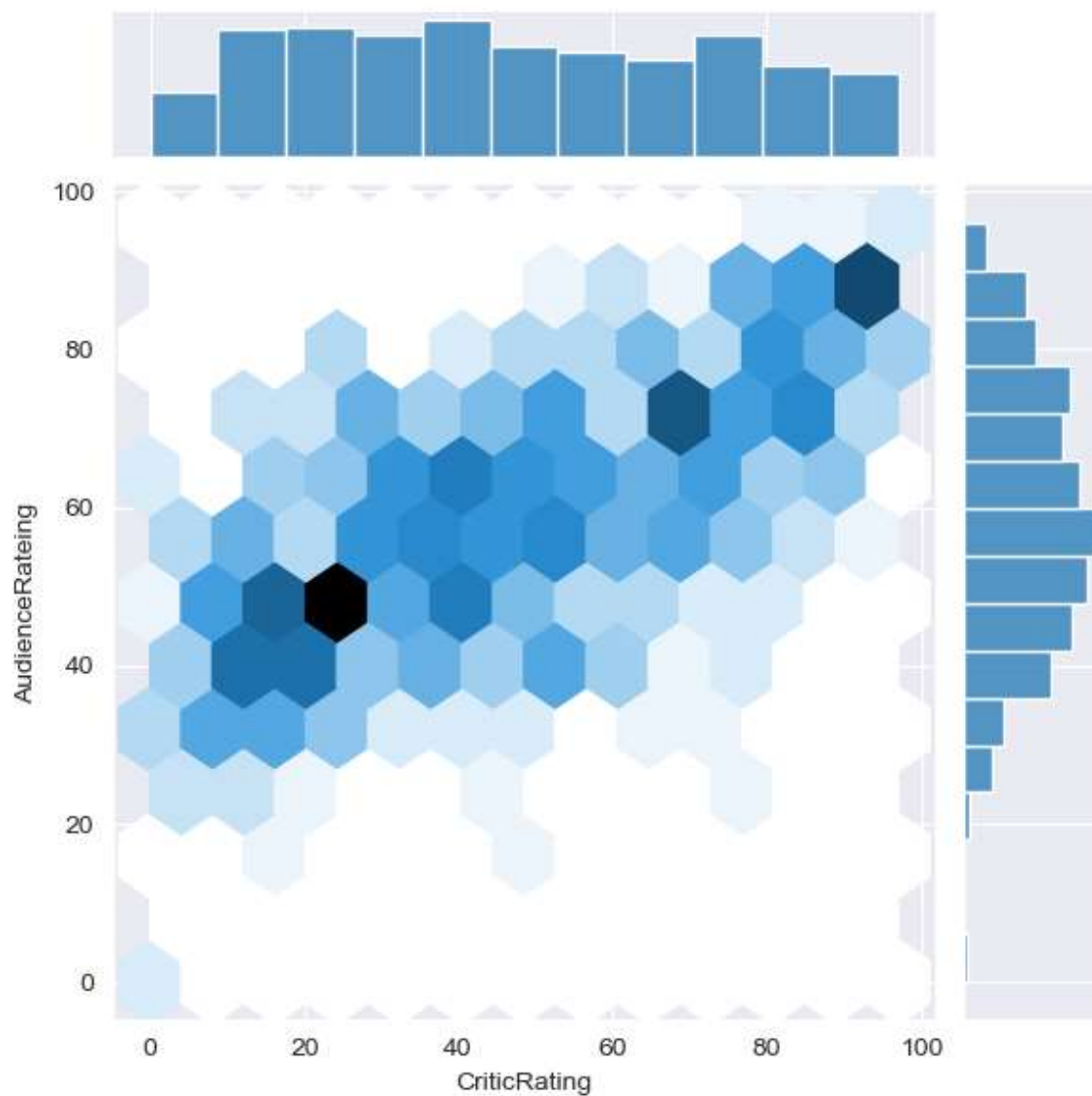
In [37]:

```
j=sns.jointplot(data= mov,x='CriticRating',y='AudienceRateing')
```



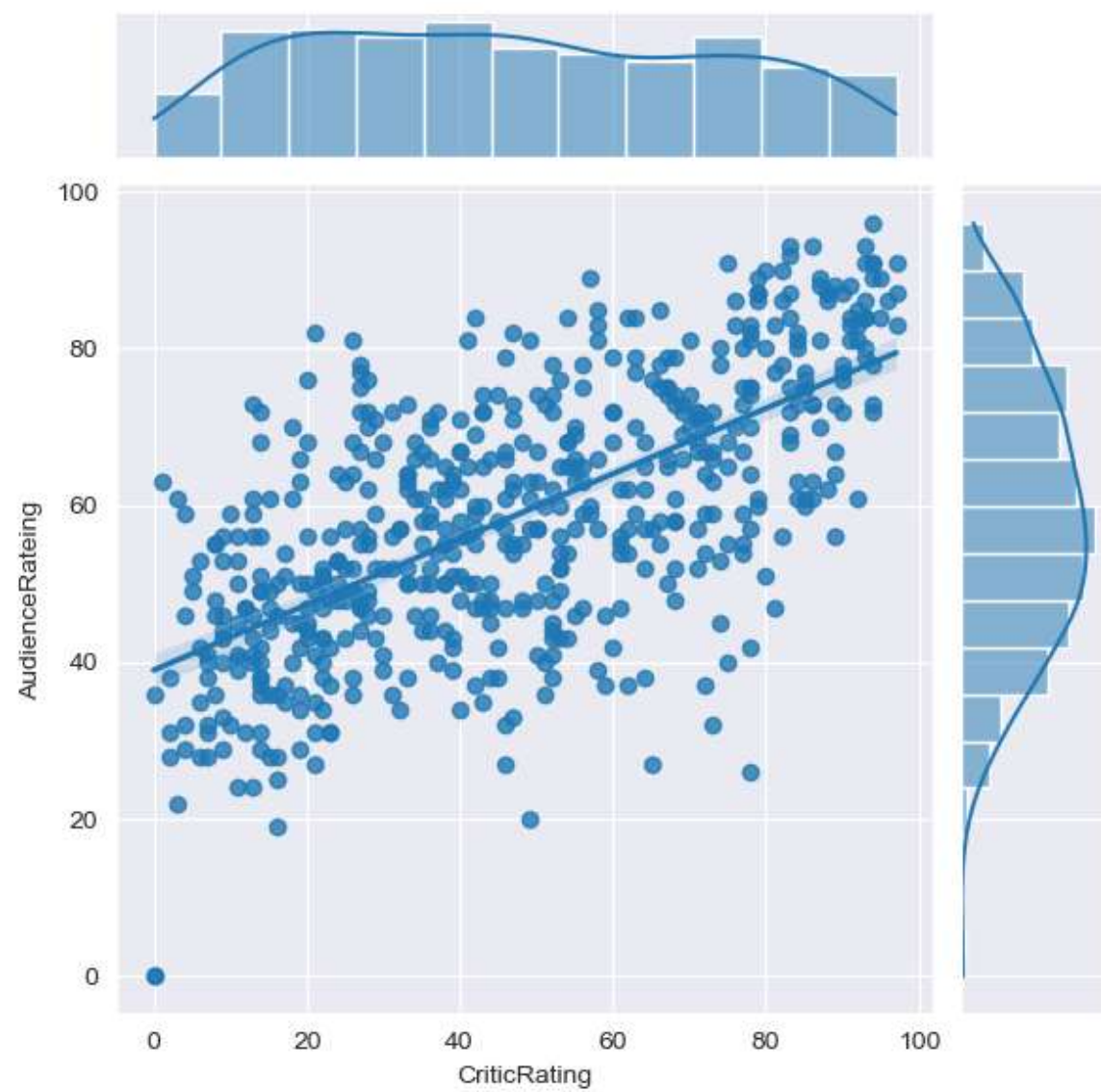
In [38]:

```
j2=sns.jointplot(data=mov,x='CriticRating',y='AudienceRateing',kind='hex')
```



In [39]:

```
j3=sns.jointplot(data=mov,x='CriticRating',y='AudienceRateing',kind='reg')
```



In [40]:

```
mov.head()
```

Out[40]:

	Film	Genre	CriticRating	AudienceRateing	BudgetMillions	Year
0	(500) Days of Summer	Comedy	87	81	8	2009
1	10,000 B.C.	Adventure	9	44	105	2008
2	12 Rounds	Action	30	52	20	2009
3	127 Hours	Adventure	93	84	18	2010
4	17 Again	Comedy	55	70	20	2009

In [41]:

```
mov.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 559 entries, 0 to 558
Data columns (total 6 columns):
 #   Column                Non-Null Count  Dtype
---  -
 0   Film                  559 non-null    category
 1   Genre                 559 non-null    category
 2   CriticRating          559 non-null    int64
 3   AudienceRateing       559 non-null    int64
 4   BudgetMillions        559 non-null    int64
 5   Year                  559 non-null    category
dtypes: category(3), int64(3)
memory usage: 36.5 KB
```

In [42]:

```
mov.Genre.unique()
```

Out[42]:

```
['Comedy', 'Adventure', 'Action', 'Horror', 'Drama', 'Romance', 'Thriller']
Categories (7, object): ['Action', 'Adventure', 'Comedy', 'Drama', 'Horror', 'Romance', 'Thriller']
```

In [43]:

```
mov.Year.unique()
```

Out[43]:

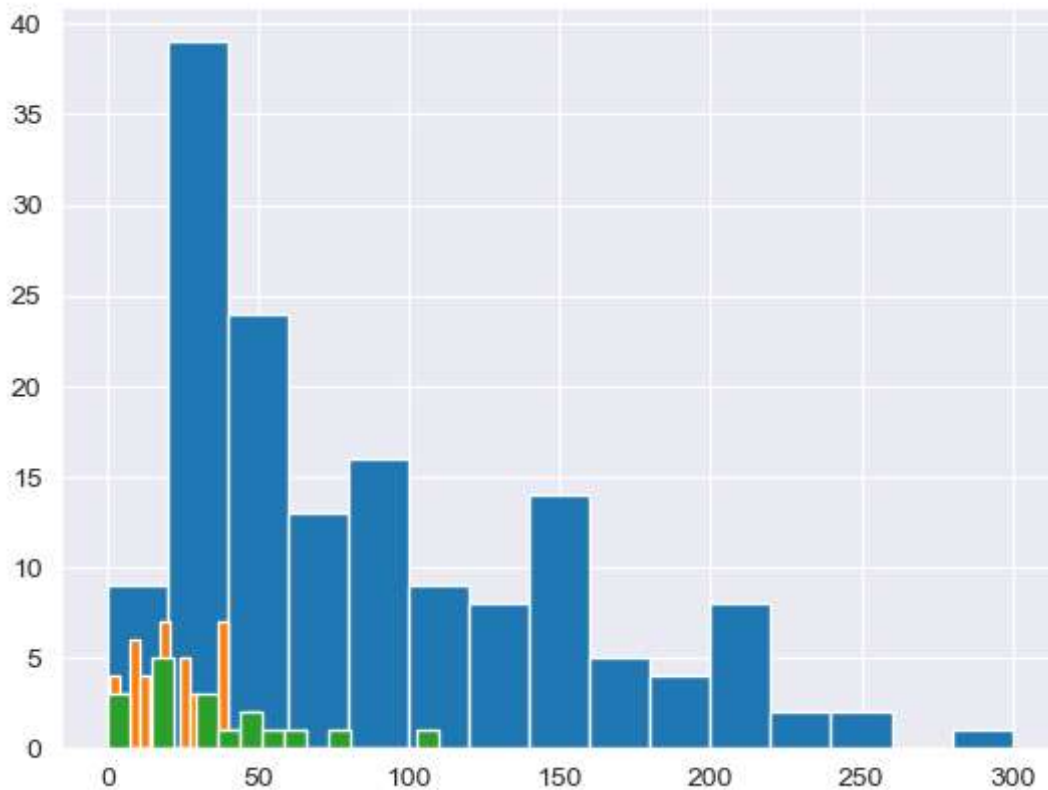
```
[2009, 2008, 2010, 2007, 2011]
Categories (5, int64): [2007, 2008, 2009, 2010, 2011]
```


In [44]:

```
plt.hist(mov[mov.Genre=='Action'].BudgetMillions,bins=15)
plt.hist(mov[mov.Genre=='Horror'].BudgetMillions,bins=15)
plt.hist(mov[mov.Genre=='Romance'].BudgetMillions,bins=15)
```

Out[44]:

```
(array([3., 0., 5., 0., 3., 1., 2., 1., 1., 0., 1., 0., 0., 0., 1.]),
 array([ 0., 7.33333333, 14.66666667, 22., 29.33333333, 36.66666667, 44., 51.33333333, 58.66666667, 66., 73.33333333, 80.66666667, 88., 95.33333333, 102.66666667, 110. ]),
 <BarContainer object of 15 artists>)
```

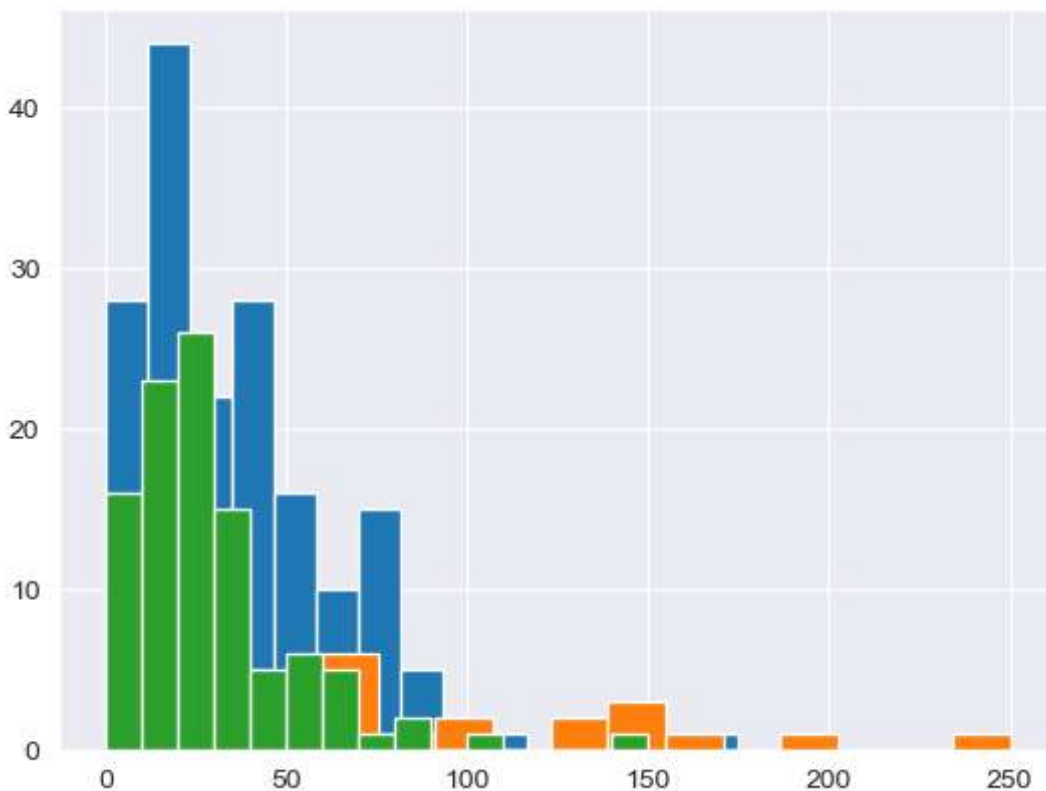


In [45]:

```
plt.hist(mov[mov.Genre=='Comedy'].BudgetMillions,bins=15)
plt.hist(mov[mov.Genre=='Adventure'].BudgetMillions,bins=15)
plt.hist(mov[mov.Genre=='Drama'].BudgetMillions,bins=15)
```

Out[45]:

```
(array([16., 23., 26., 15., 5., 6., 5., 1., 2., 0., 1., 0., 0.,
        0., 1.]),
 array([ 0., 10., 20., 30., 40., 50., 60., 70., 80., 90., 100.,
        110., 120., 130., 140., 150.]),
 <BarContainer object of 15 artists>)
```

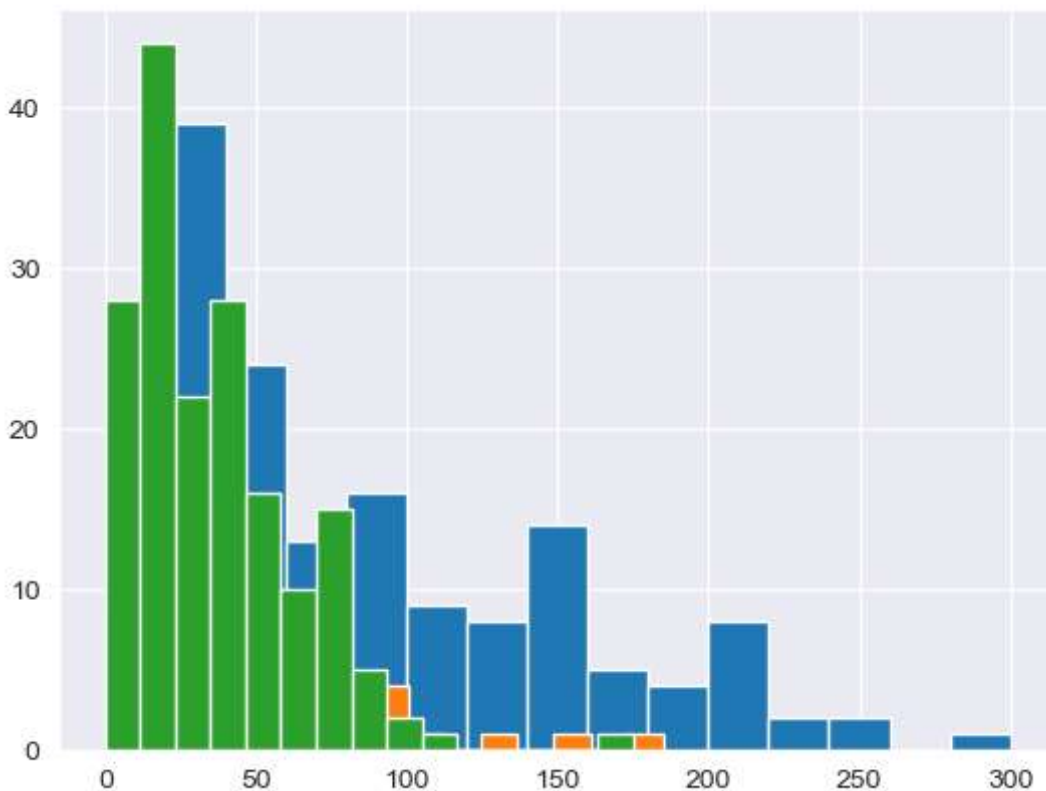


In [46]:

```
plt.hist(mov[mov.Genre=='Action'].BudgetMillions,bins=15)
plt.hist(mov[mov.Genre=='Thriller'].BudgetMillions,bins=15)
plt.hist(mov[mov.Genre=='Comedy'].BudgetMillions,bins=15)
```

Out[46]:

```
(array([28., 44., 22., 28., 16., 10., 15., 5., 2., 1., 0., 0., 0.,
        0., 1.]),
 array([ 0., 11.66666667, 23.33333333, 35., 46.66666667, 58.33333333, 70., 81.66666667,
        93.33333333, 105., 116.66666667, 128.33333333, 140., 151.66666667, 163.33333333, 175.
        ]),
 <BarContainer object of 15 artists>)
```

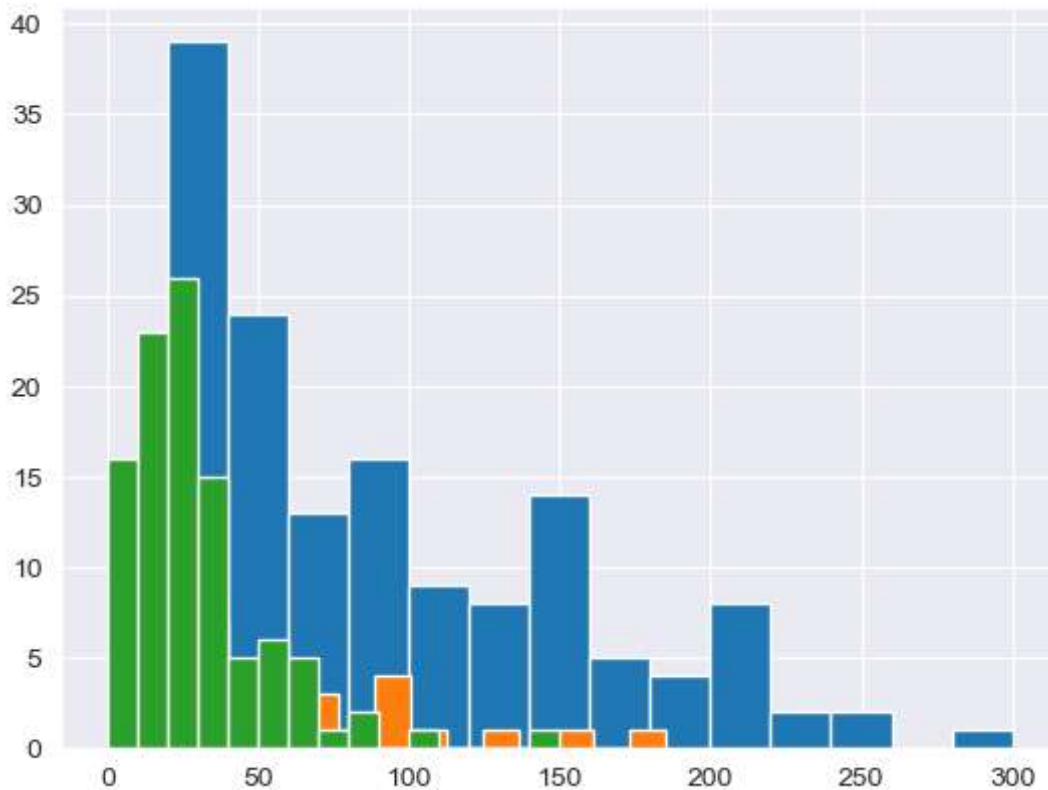


In [47]:

```
plt.hist(mov[mov.Genre=='Action'].BudgetMillions,bins=15)  
plt.hist(mov[mov.Genre=='Thriller'].BudgetMillions,bins=15)  
plt.hist(mov[mov.Genre=='Drama'].BudgetMillions,bins=15)
```

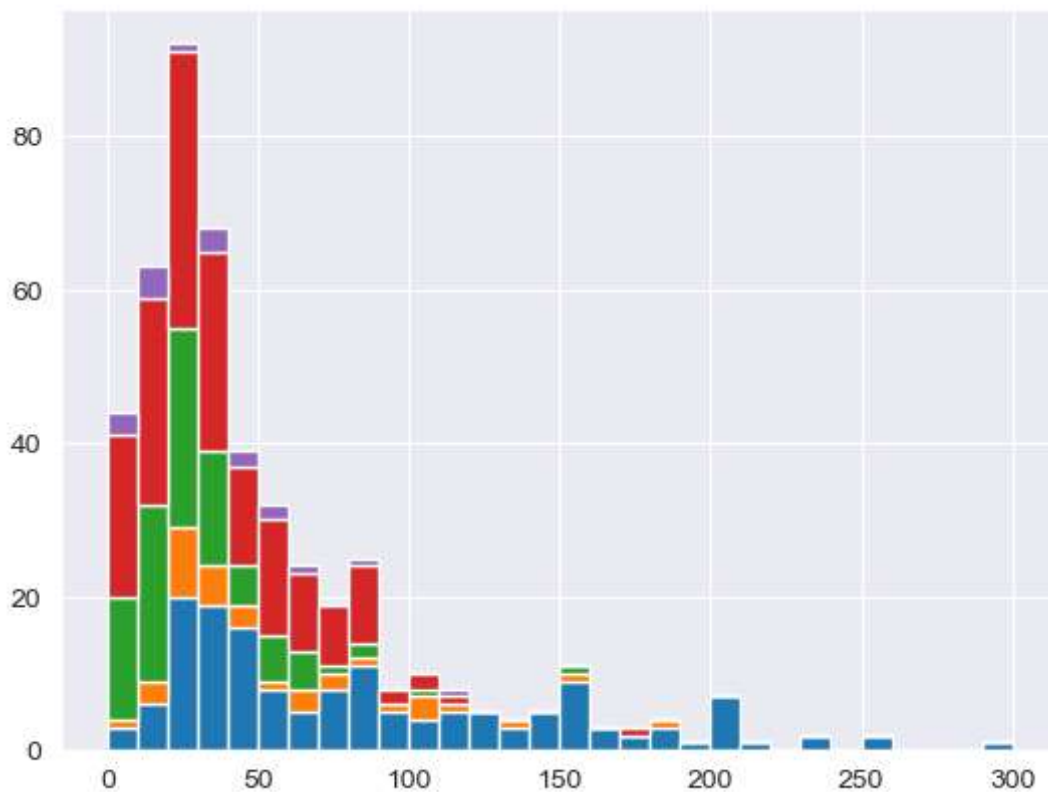
Out[47]:

```
(array([16., 23., 26., 15., 5., 6., 5., 1., 2., 0., 1., 0., 0.,  
        0., 1.]),  
 array([ 0., 10., 20., 30., 40., 50., 60., 70., 80., 90., 10  
0.,  
        110., 120., 130., 140., 150.]),  
<BarContainer object of 15 artists>)
```



In [48]:

```
plt.hist([mov[mov.Genre=='Action'].BudgetMillions,\n          mov[mov.Genre=='Thriller'].BudgetMillions,\n          mov[mov.Genre=='Drama'].BudgetMillions,\n          mov[mov.Genre=='Comedy'].BudgetMillions,\n          mov[mov.Genre=='Romance'].BudgetMillions],bins=30,stacked=True)\nplt.show()
```



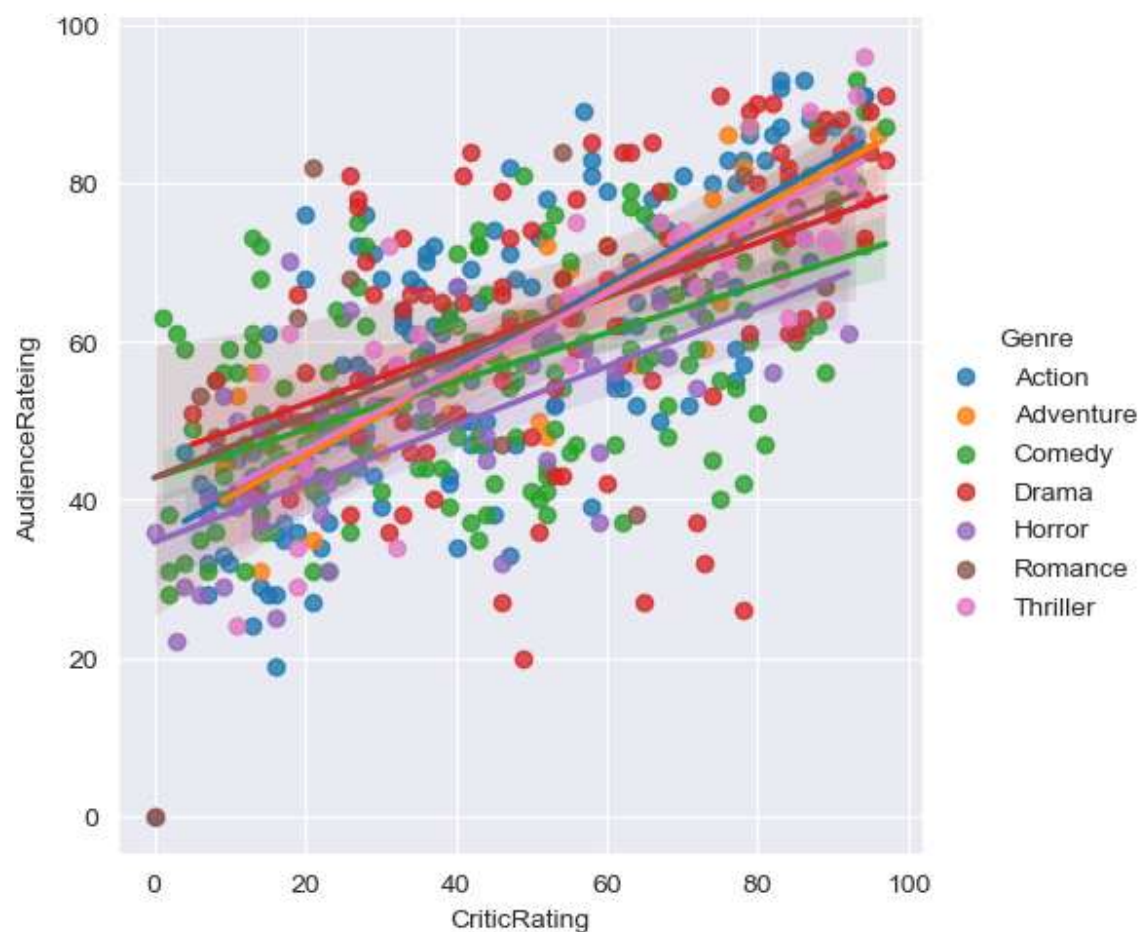
In [49]:

```
for theme in mov.Genre.cat.categories:\n    print(theme)
```

Action
Adventure
Comedy
Drama
Horror
Romance
Thriller

In [50]:

```
r=sns.lmplot(data=mov,x='CriticRating',y='AudienceRateing',fit_reg=True,hue='Genre')
```

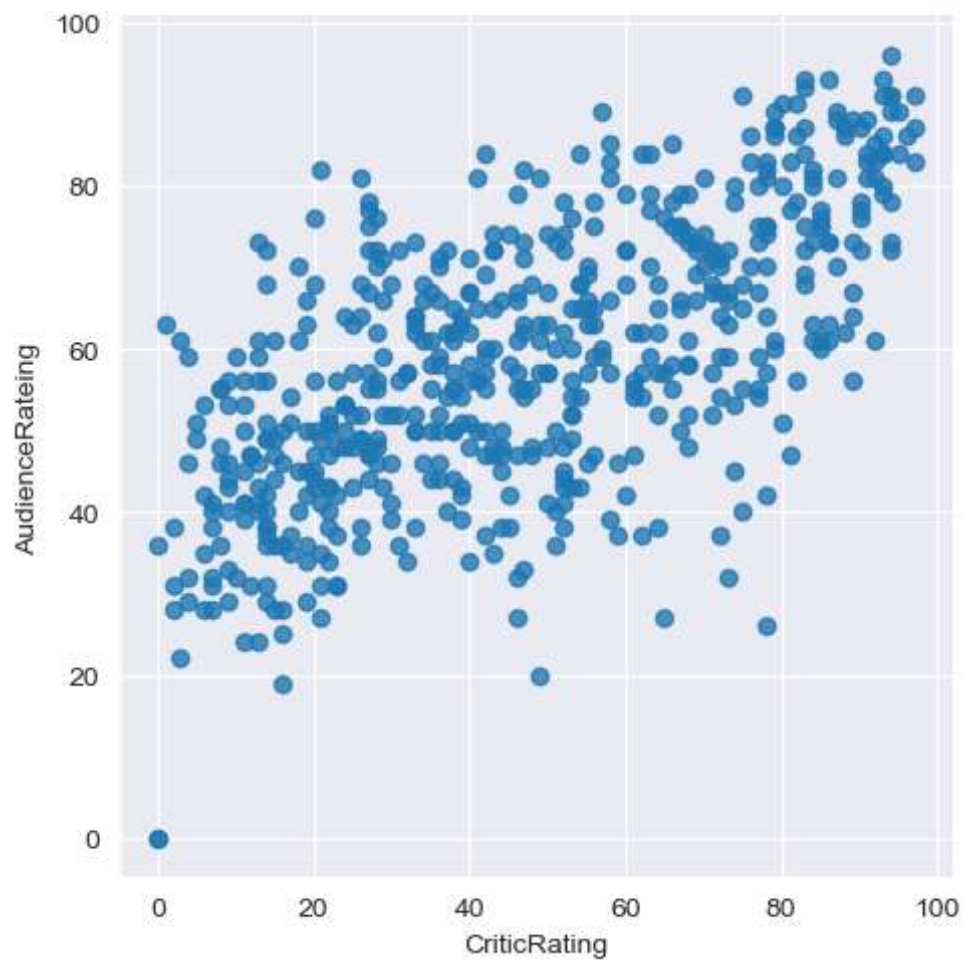


In []:

```
r2=sns.lmplot(data=mov,x='CriticRating',y='AudienceRateing',fit_reg=False,hue='Genre')
```

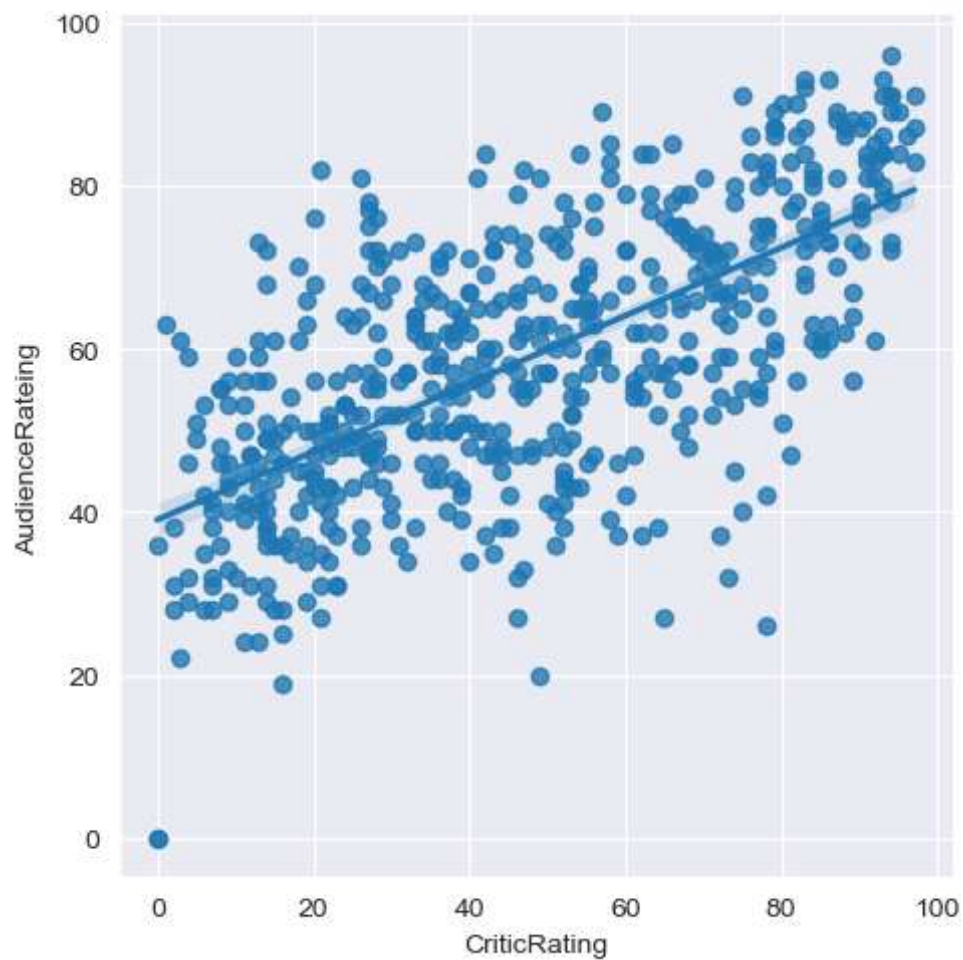
In [52]:

```
r3=sns.lmplot(data=mov,x='CriticRating',y='AudienceRateing',fit_reg=False)
```



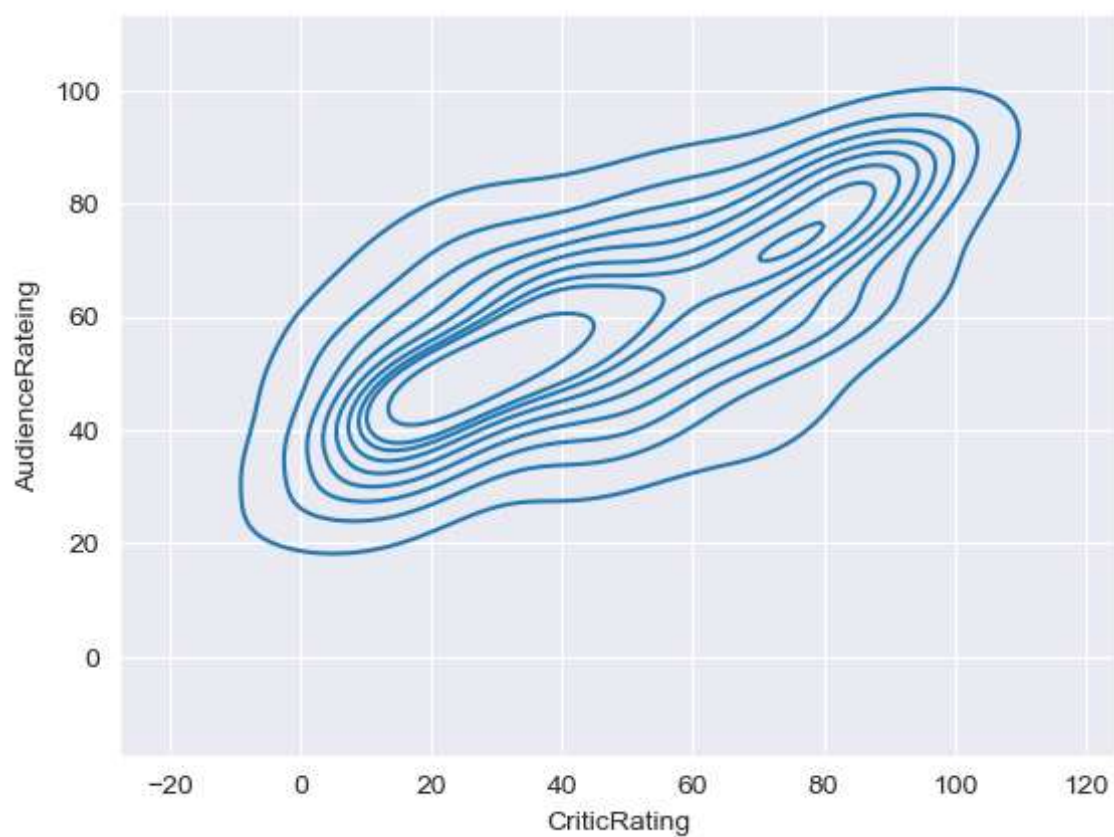
In [53]:

```
r=sns.lmplot(data=mov,x='CriticRating',y='AudienceRateing',fit_reg=True)
```



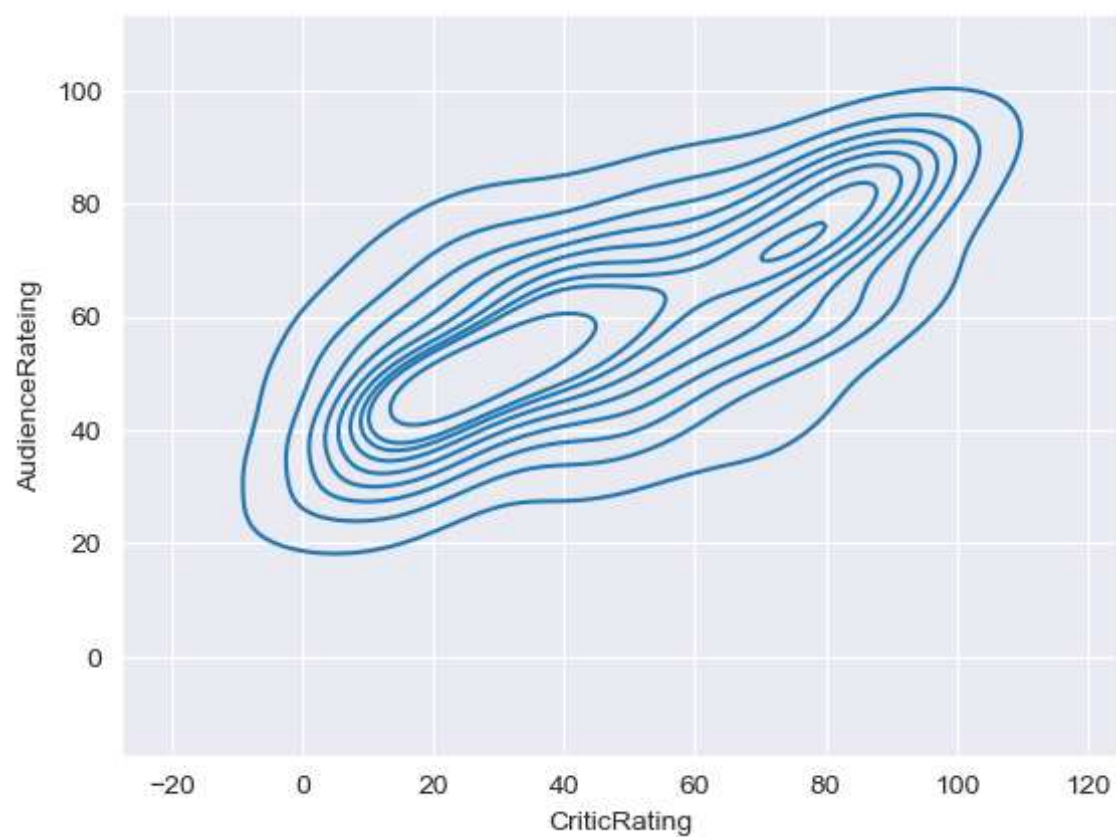
In [54]:

```
kde=sns.kdeplot(x=mov['CriticRating'],y=mov['AudienceRateing'])
```



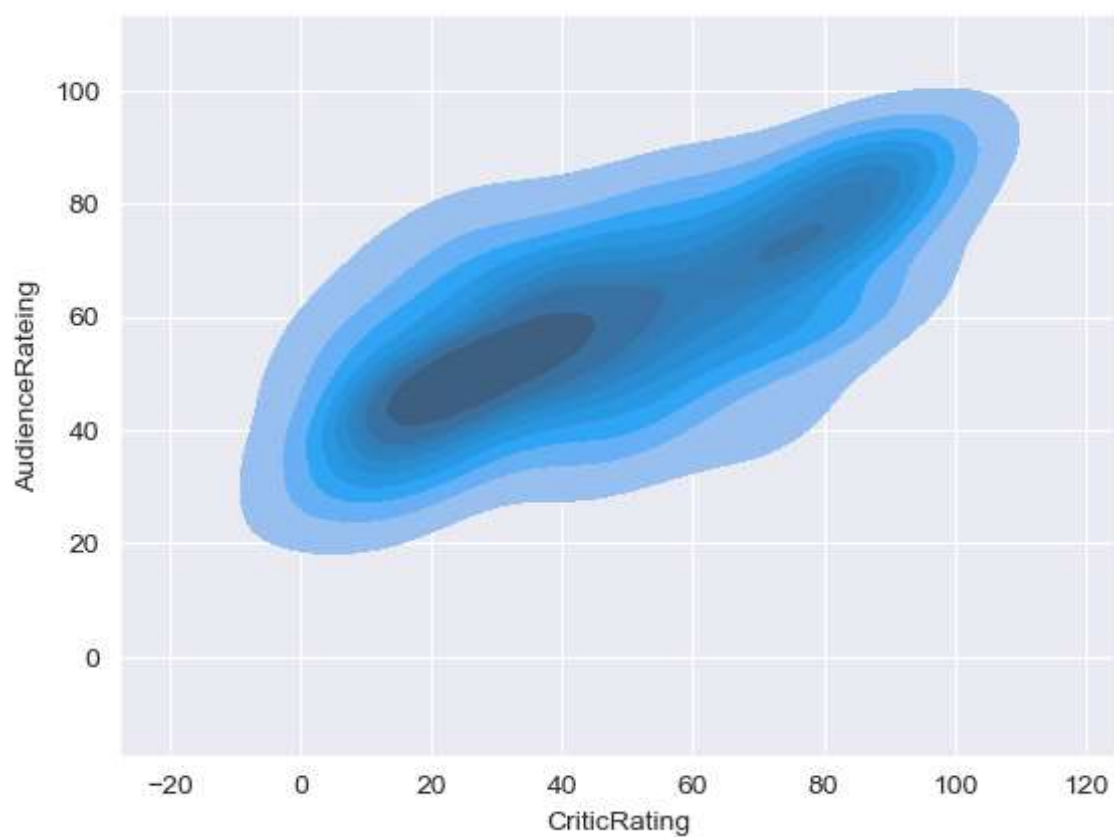
In [55]:

```
kde=sns.kdeplot(x=mov['CriticRating'],y=mov['AudienceRateing'])  
plt.show()
```



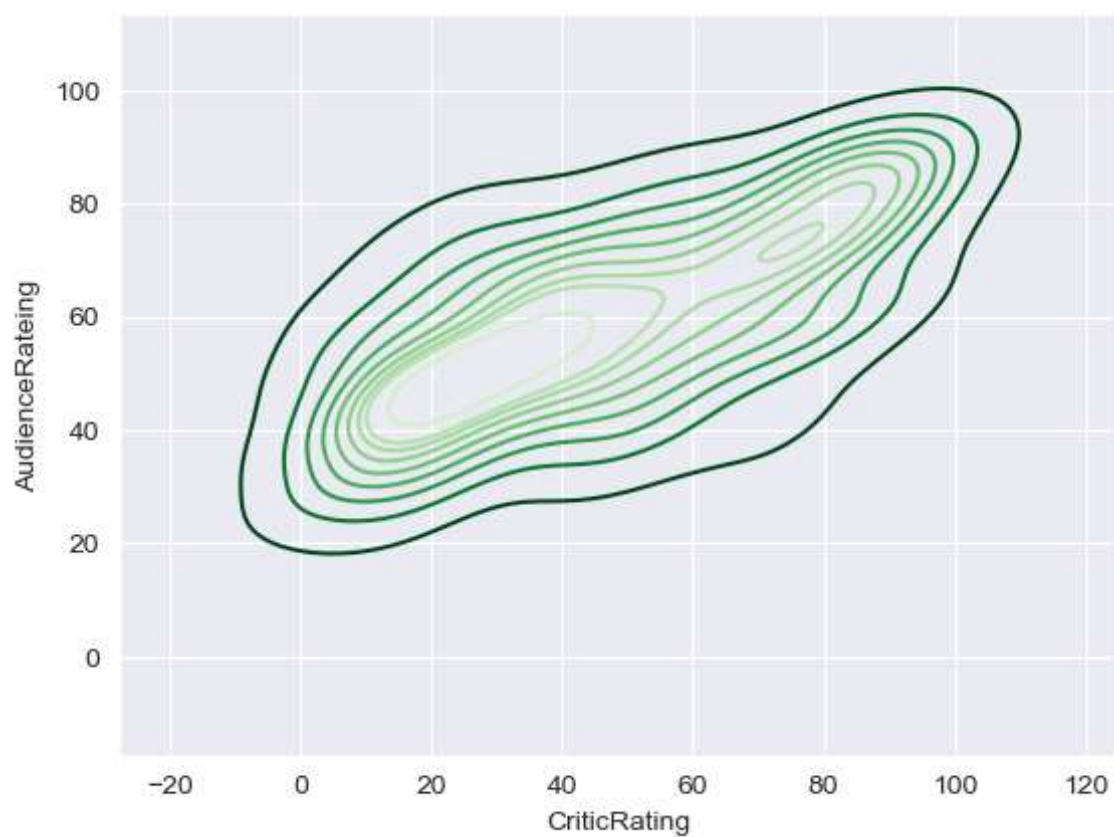
In [56]:

```
kde2=sns.kdeplot(x=mov['CriticRating'],y=mov['AudienceRateing'],shade=True)  
plt.show()
```



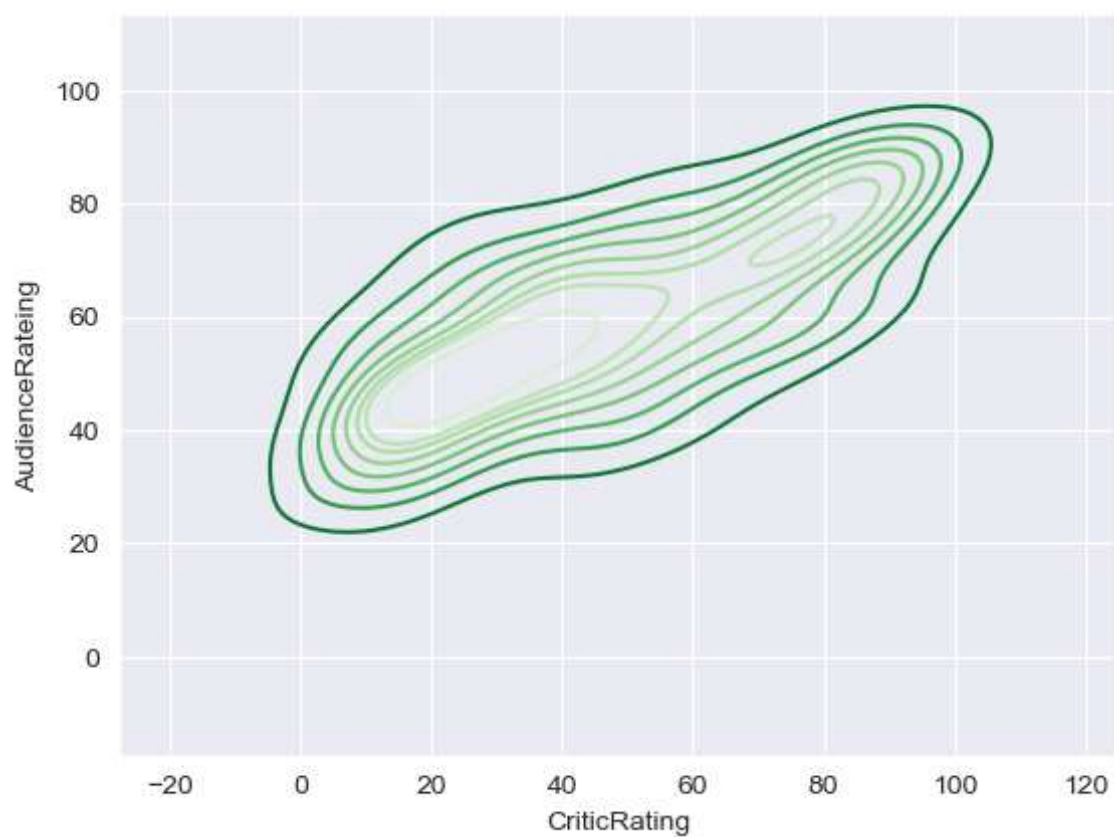
In [57]:

```
kde2=sns.kdeplot(x=mov['CriticRating'],y=mov['AudienceRateing'],shade_lowest=False,cm  
plt.show()
```



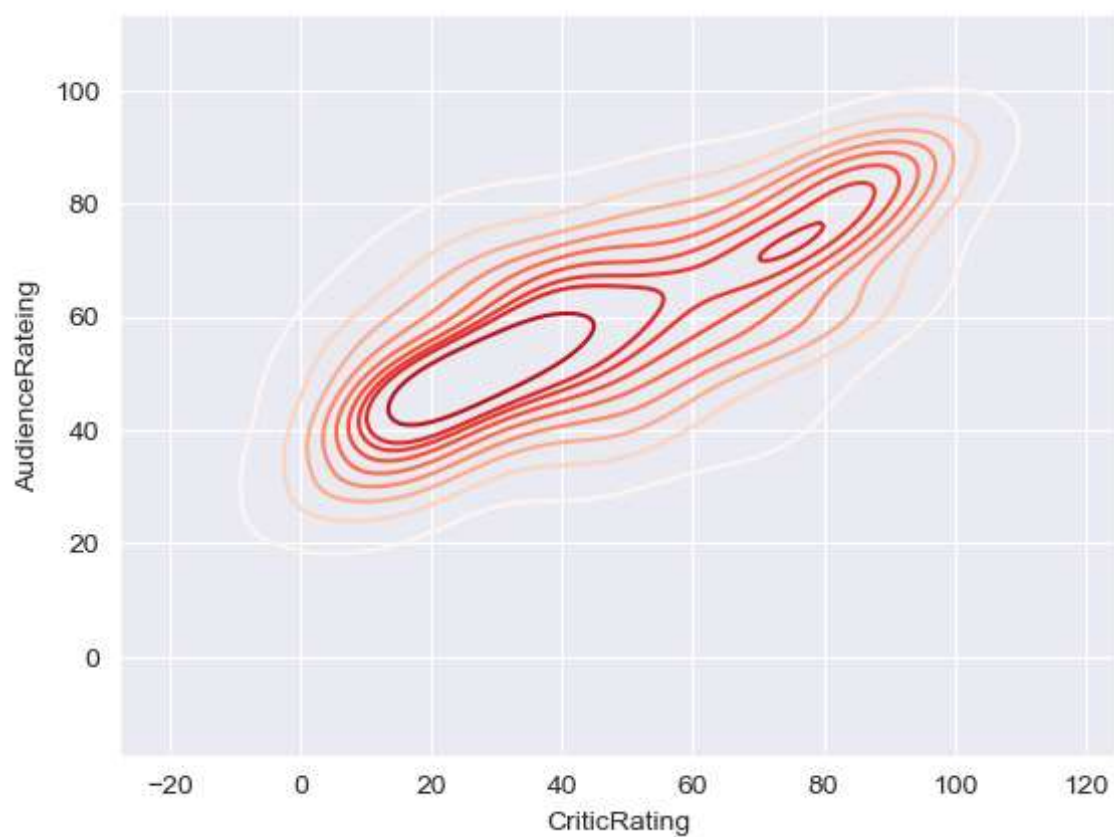
In [58]:

```
kde3=sns.kdeplot(x=mov['CriticRating'],y=mov['AudienceRateing'],shade_lowest=True,cma  
plt.show()
```



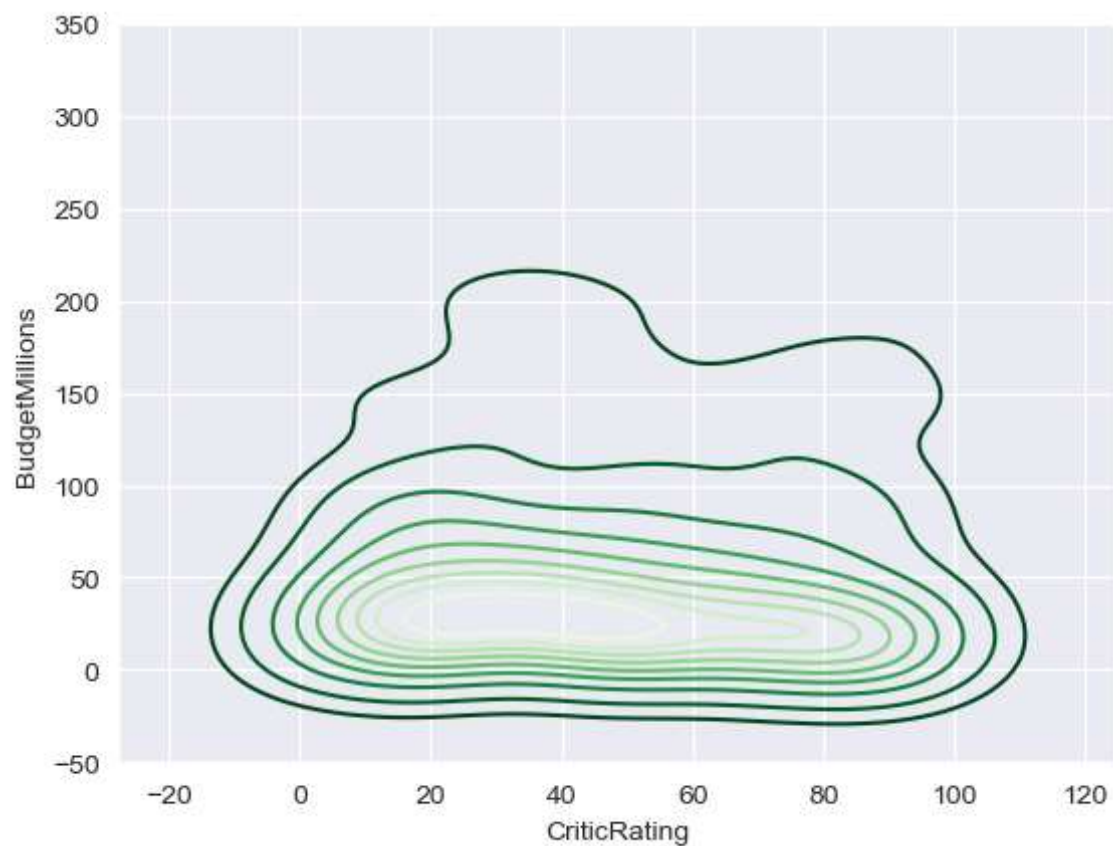
In [59]:

```
kde4=sns.kdeplot(x=mov.CriticRating,y=mov.AudienceRateing,shade_lowest=False,cmap='Red',  
plt.show())
```



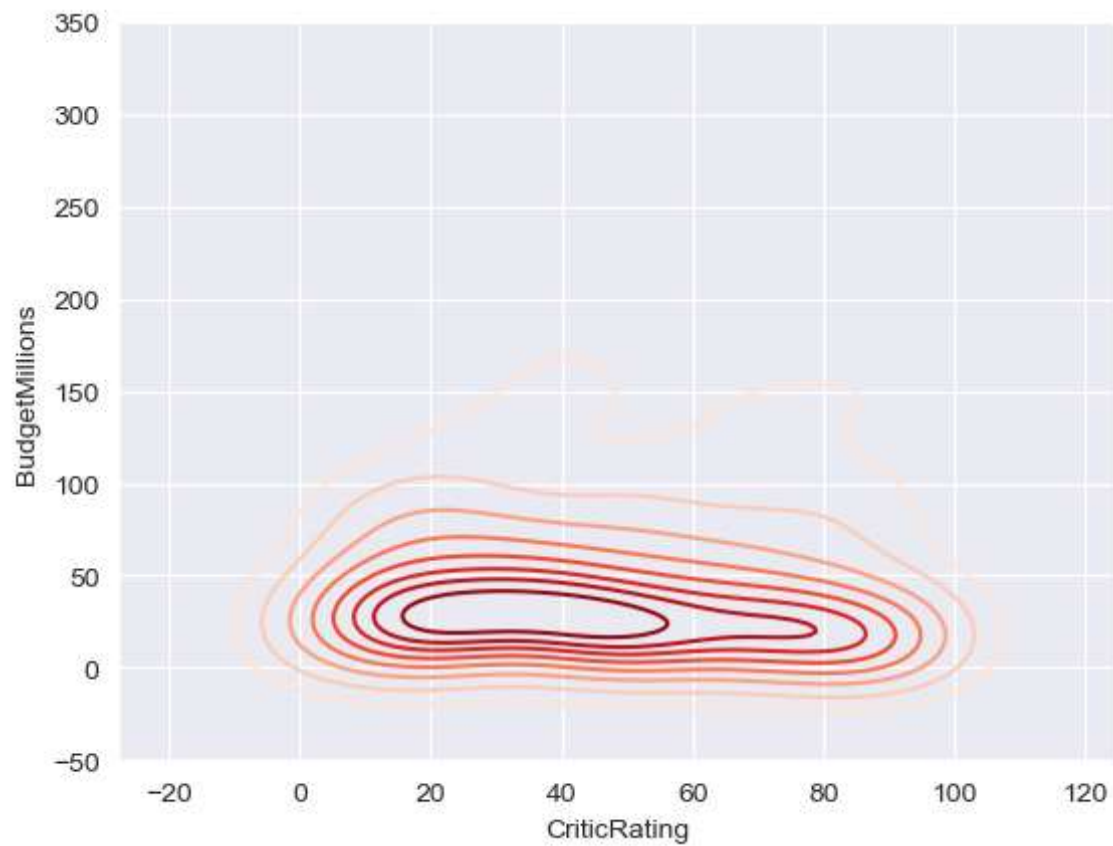
In [60]:

```
kde5=sns.kdeplot(x=mov.CriticRating,y=mov.BudgetMillions,shade_lowest=False,cmap='Greens',  
plt.show())
```



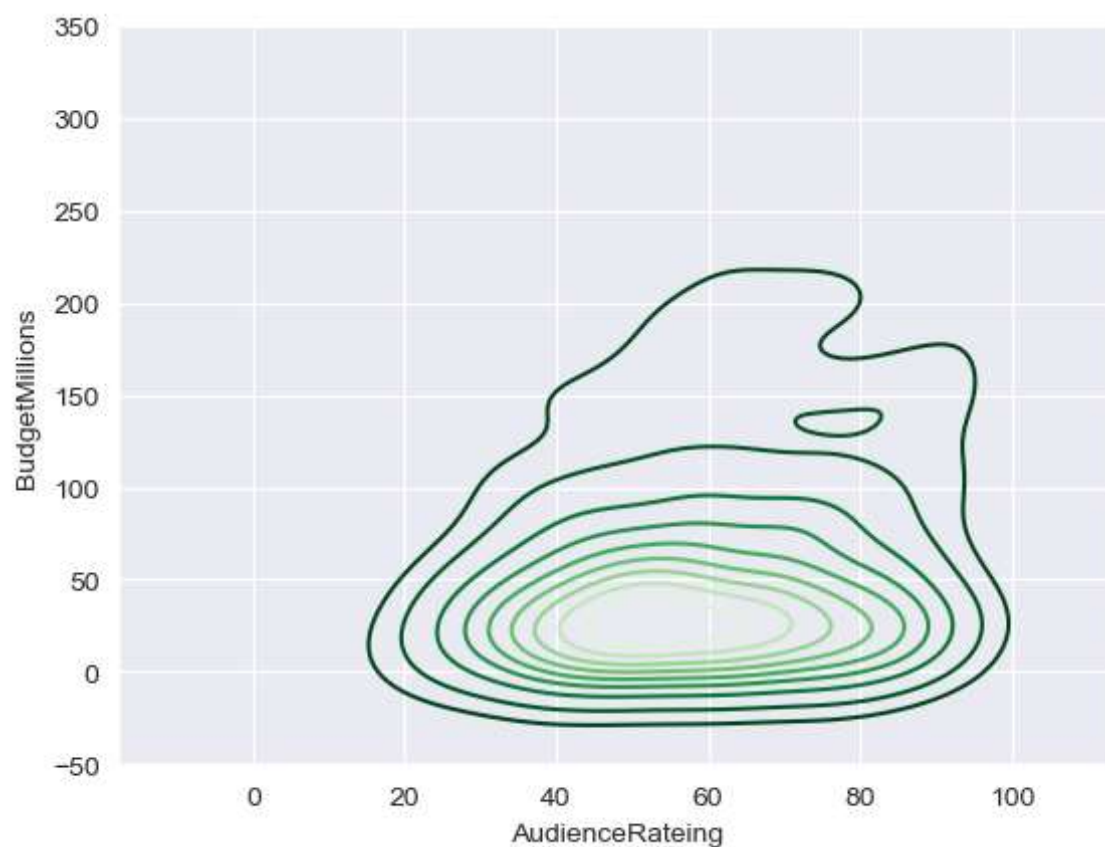
In [61]:

```
kde6=sns.kdeplot(x=mov.CriticRating,y=mov.BudgetMillions,shade_lowest=True,cmap='Reds')  
plt.show()
```



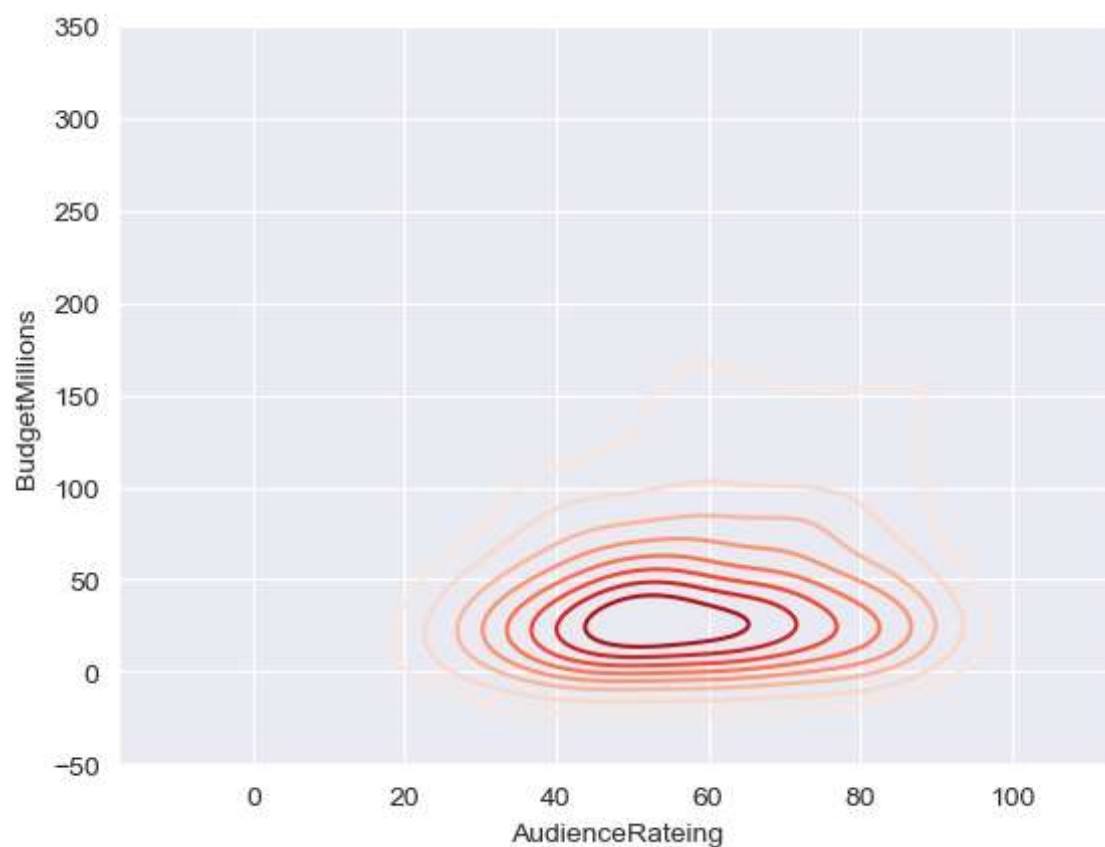
In [62]:

```
kde7=sns.kdeplot(x=mov.AudienceRateing,y=mov.BudgetMillions,shade_lowest=False,cmap='c',  
plt.show())
```



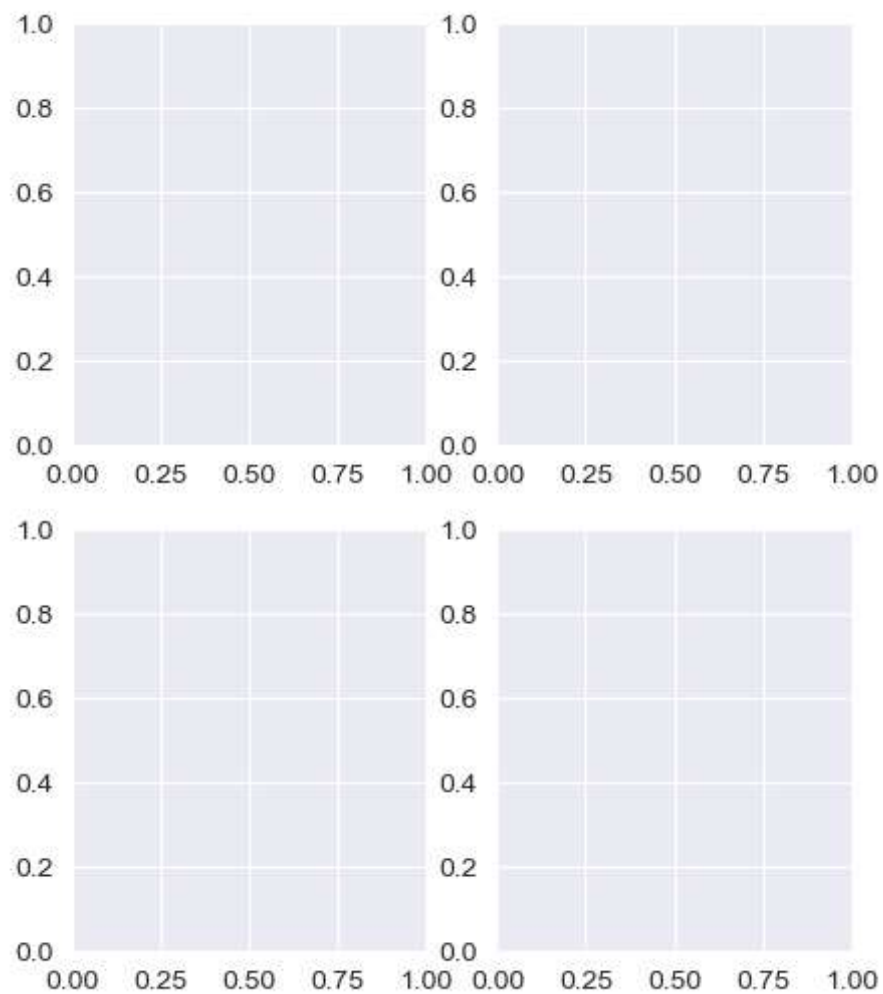
In [63]:

```
kde8=sns.kdeplot(x=mov.AudienceRateing,y=mov.BudgetMillions,shade_lowest=True,cmap='Red',  
plt.show())
```



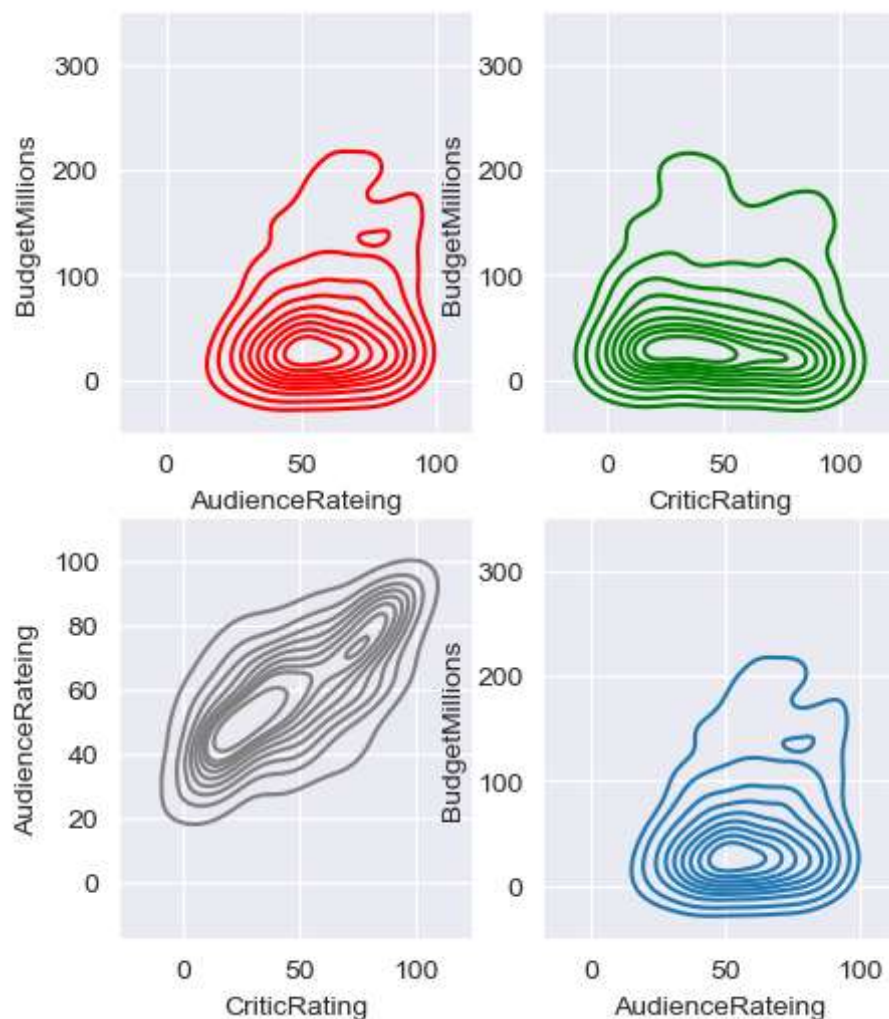
In [64]:

```
m,ax=plt.subplots(2,2,figsize=(5,6))
```



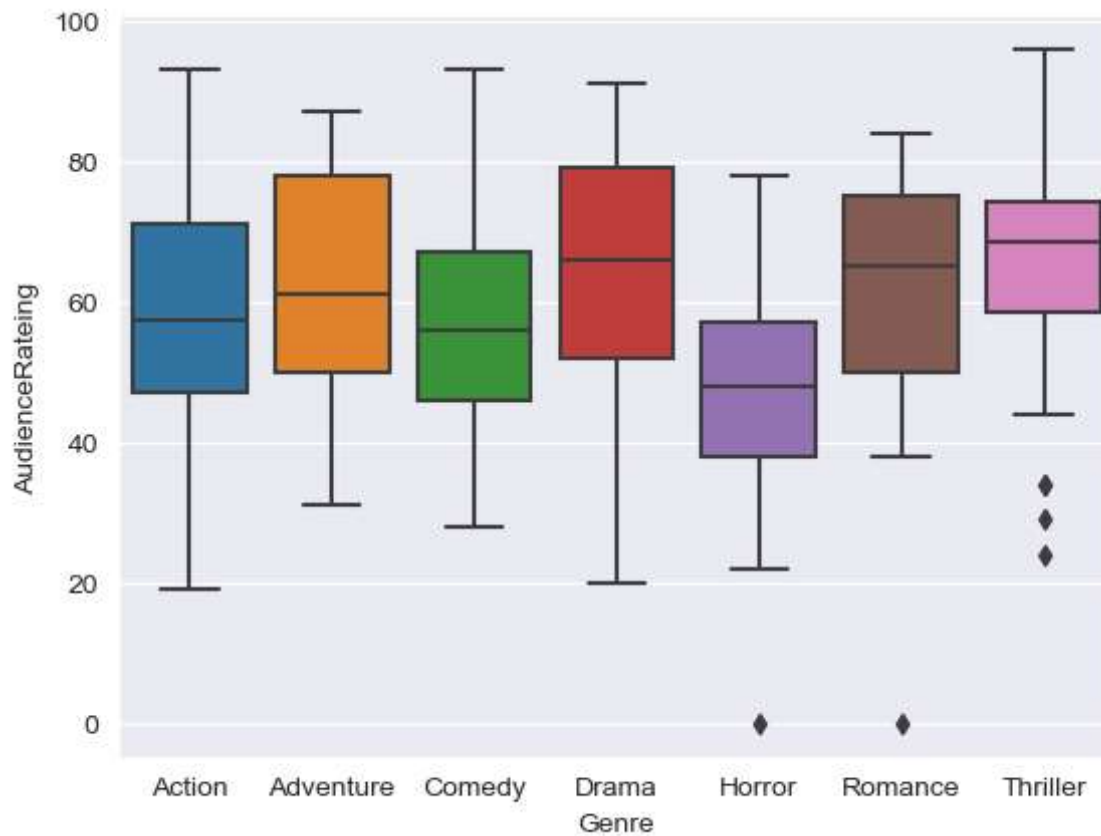
In [65]:

```
m,ax=plt.subplots(2,2,figsize=(5,6))
k1=sns.kdeplot(x=mov.AudienceRateing,y=mov.BudgetMillions,ax=ax[0,0],color="Red")
k2=sns.kdeplot(x=mov.CriticRating,y=mov.BudgetMillions,ax=ax[0,1],color="Green")
k3=sns.kdeplot(x=mov.CriticRating,y=mov.AudienceRateing,ax=ax[1,0],color="Gray")
k4=sns.kdeplot(x=mov.AudienceRateing,y=mov.BudgetMillions,ax=ax[1,1])
```



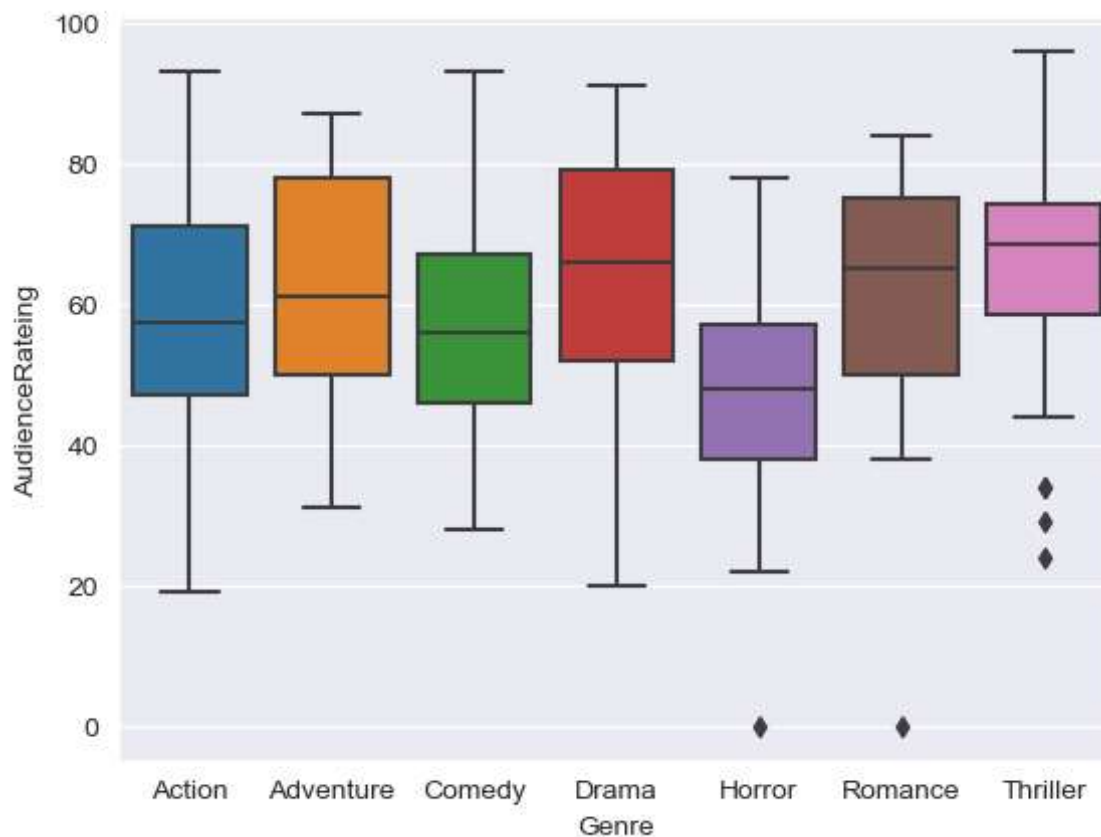
In [66]:

```
box=sns.boxplot(x='Genre',y='AudienceRateing',data=mov)
```



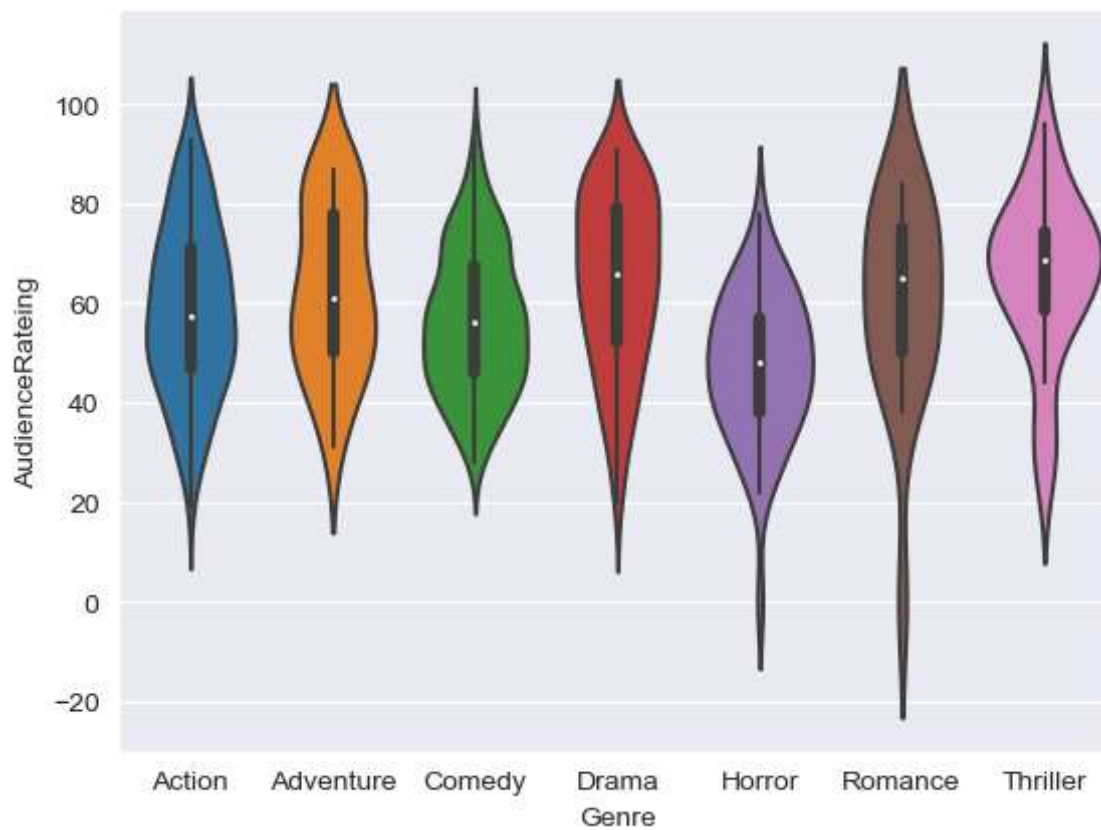
In [67]:

```
box2=sns.boxplot(data=mov,x='Genre',y='AudienceRateing')
```



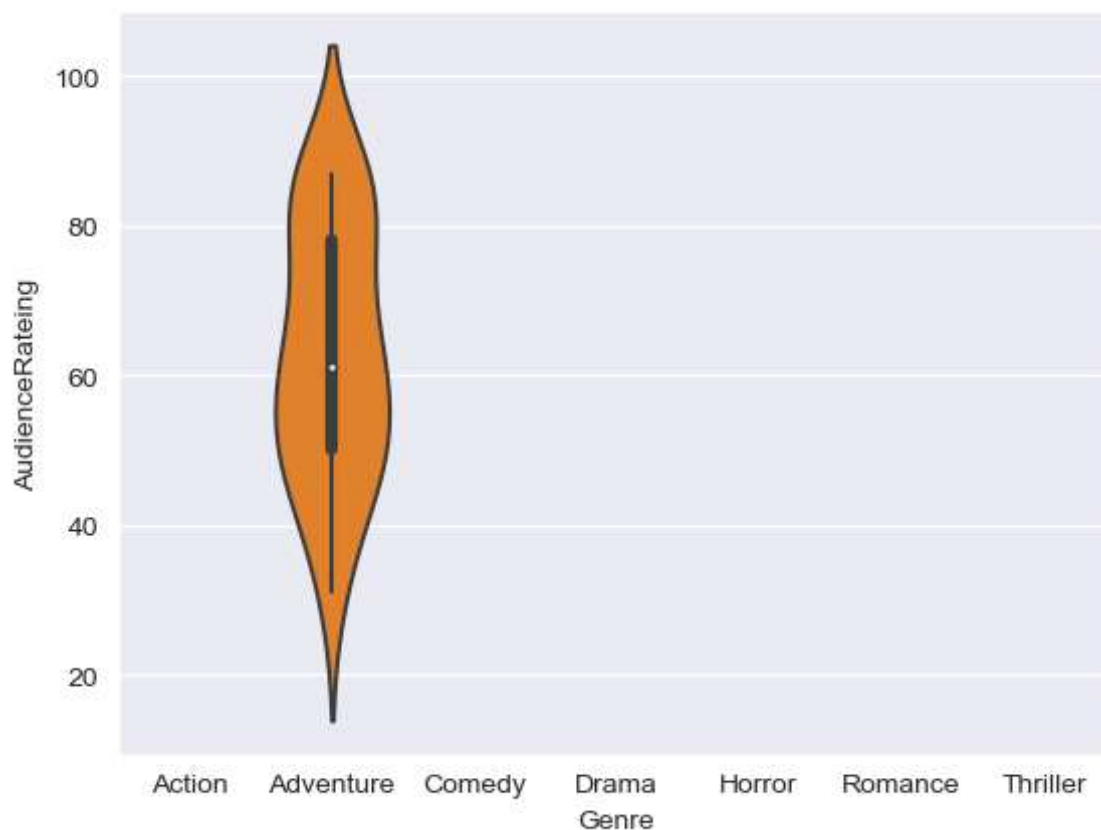
In [68]:

```
violin=sns.violinplot(x='Genre',y='AudienceRateing',data=mov)
```



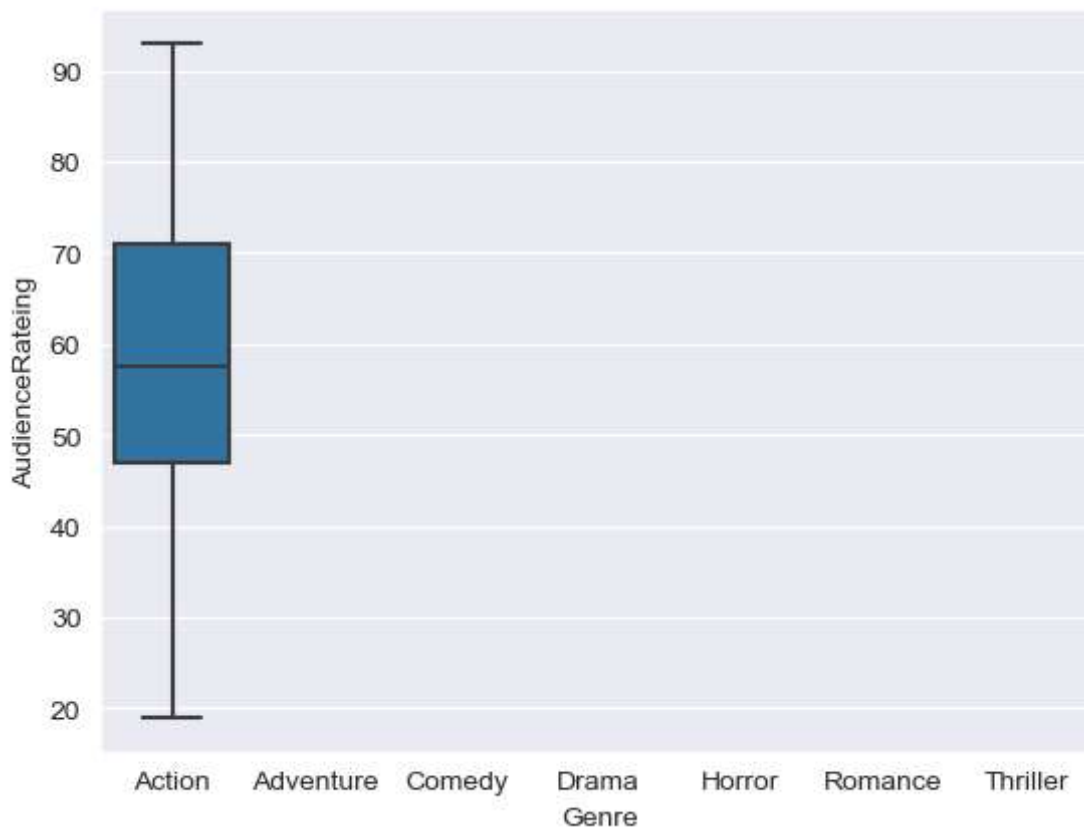
In [69]:

```
violin1=sns.violinplot(x='Genre',y='AudienceRateing',data=mov[mov.Genre=='Adventure'])
```



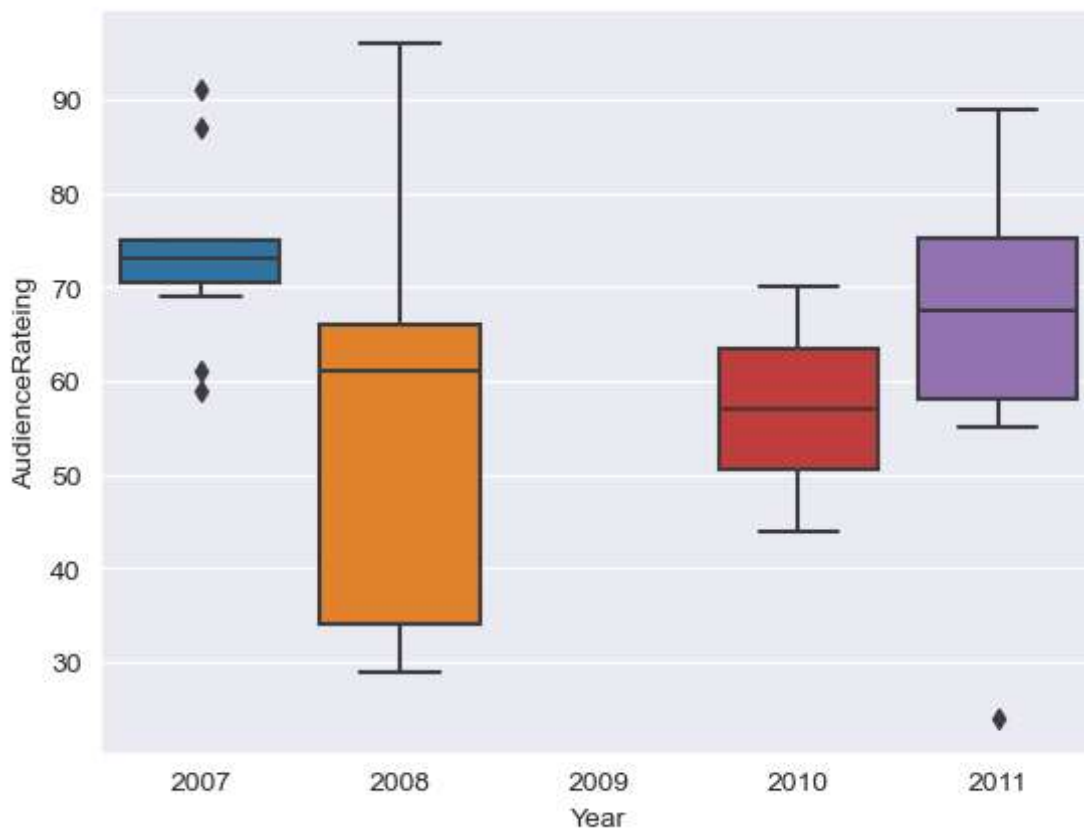
In [70]:

```
boxsp=sns.boxplot(x='Genre',y='AudienceRateing',data=mov[mov.Genre=='Action'])
```



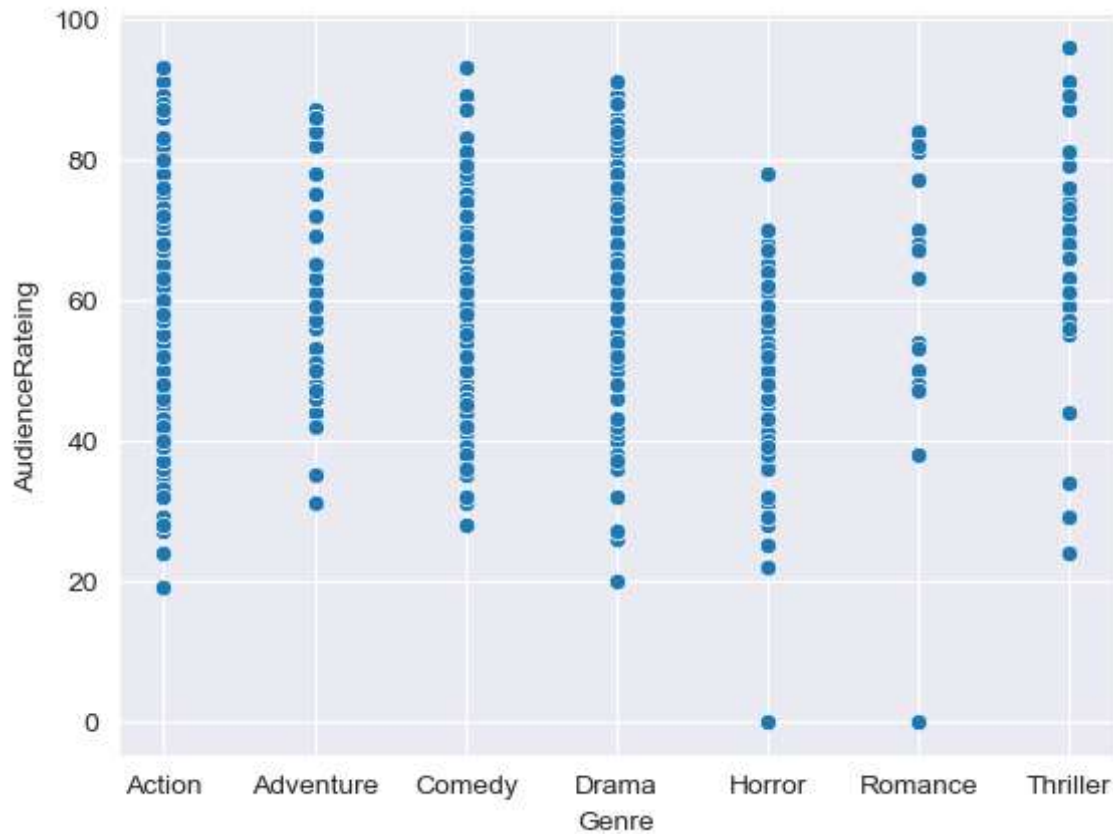
In [71]:

```
boxyear=sns.boxplot(x='Year',y='AudienceRateing',data=mov[mov.Genre=='Thriller'])
```



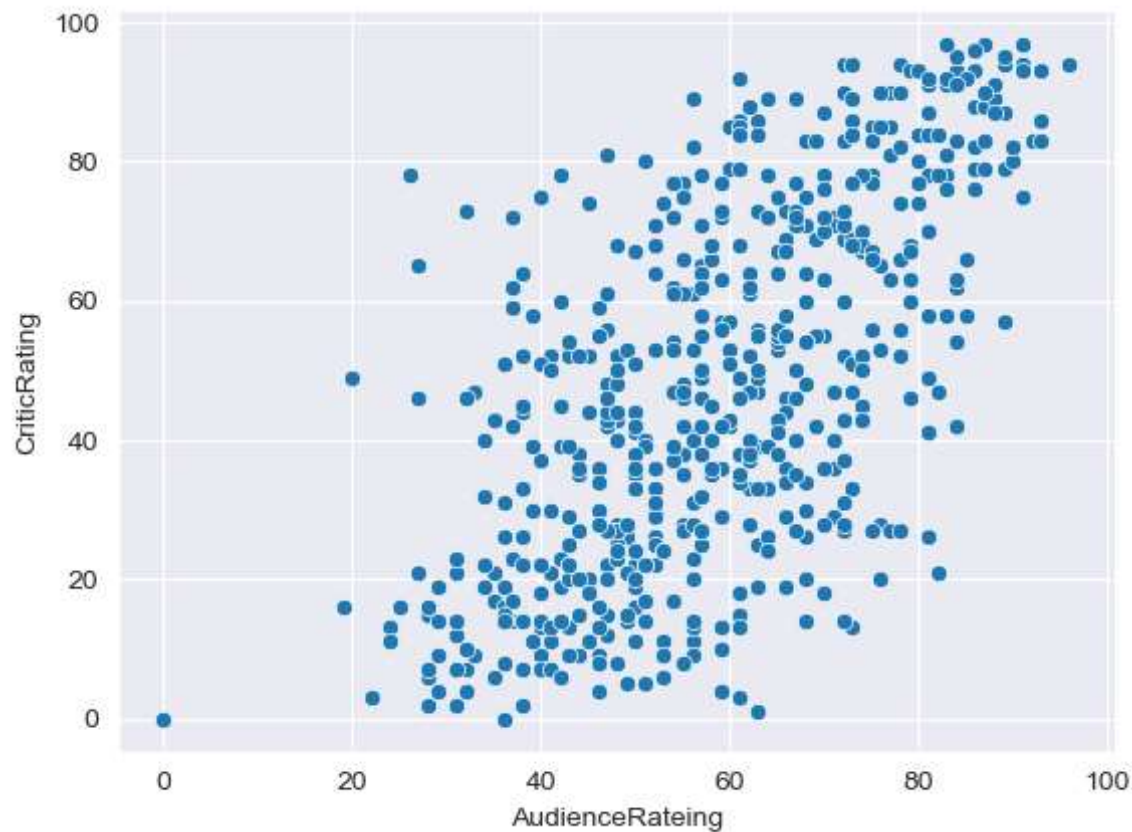
In [72]:

```
sc=sns.scatterplot(x='Genre',y='AudienceRateing',data=mov)
```



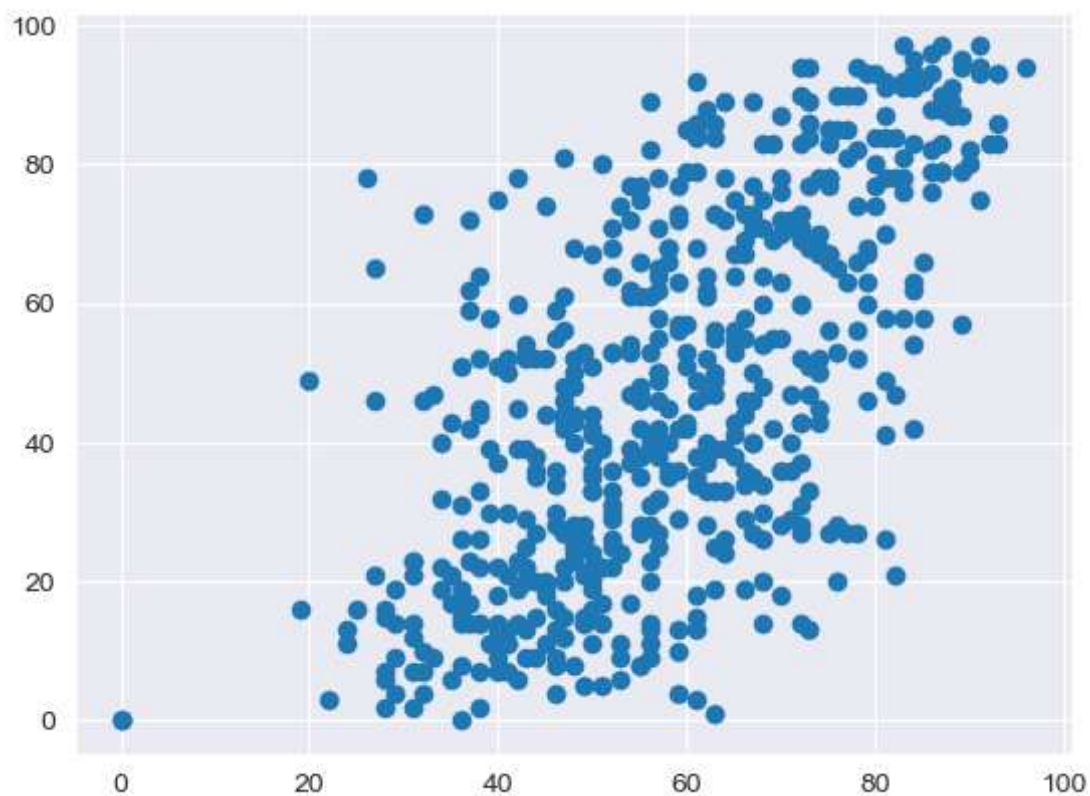
In [73]:

```
sc2=sns.scatterplot(data=mov,x=mov.AudienceRateing,y=mov.CriticRating)
```



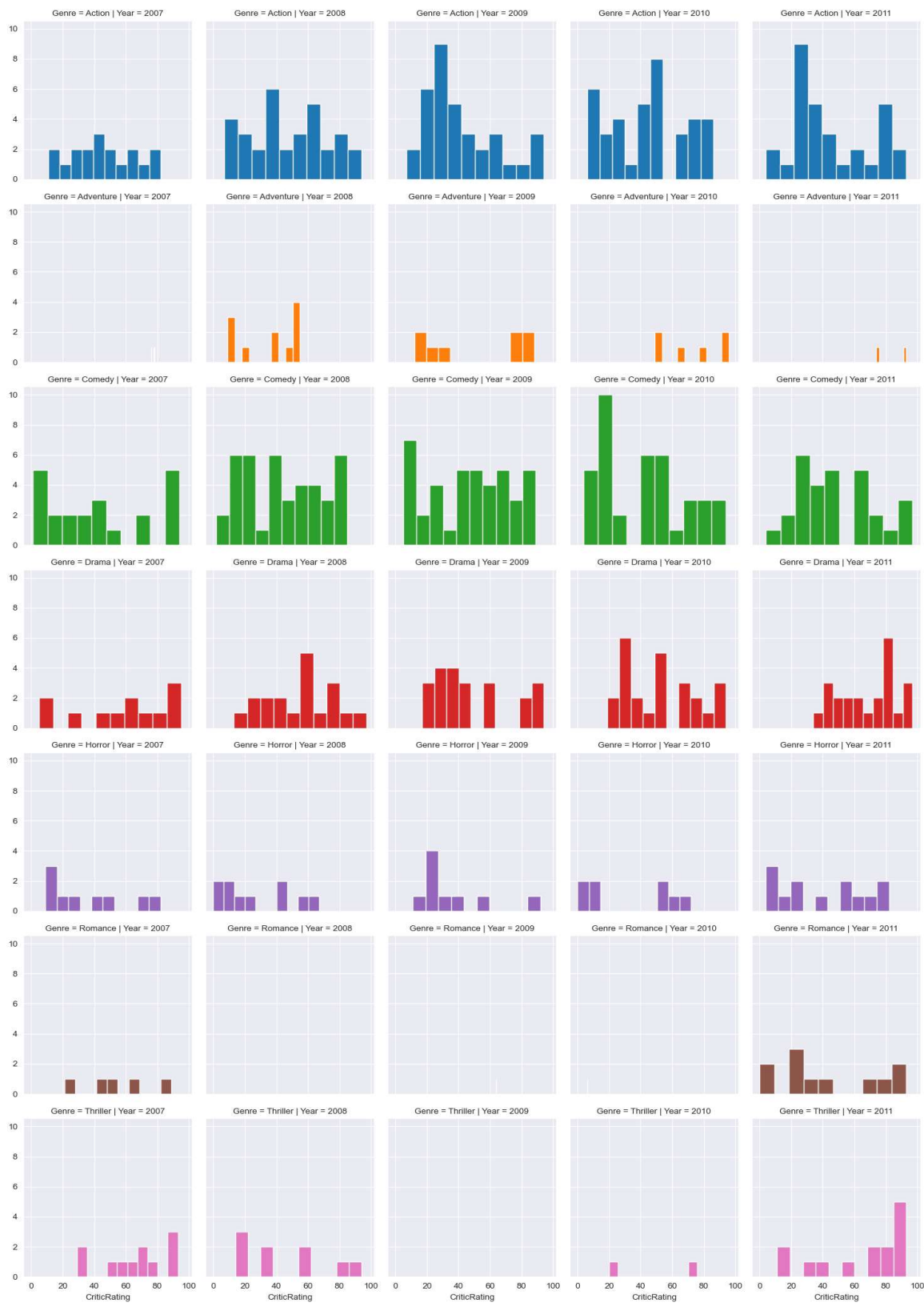
In [74]:

```
sc3=plt.scatter(mov.AudienceRateing,mov.CriticRating)
```



In [75]:

```
fg=sns.FacetGrid(mov,row='Genre',col='Year',hue='Genre')
fg=fg.map(plt.hist,'CriticRating')
```



In []:

```
fg2=sns.FacetGrid(mov,row='Genre',col='Year',hue='Genre')
fg2=fg2.map(plt.hist,'AudienceRateing')
```

