

```
In [92]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

```
In [18]: df=pd.read_csv("top-500-movies.csv")
df
```

Out[18]:

| se_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime | year |
|----------|---|---|-----------------|----------------|-----------------|-----------------|-------|-------------------|----------|---------|--------|
| 19-04-23 | Avengers: Endgame | /movie/Avengers-Endgame-(2019)#tab=summary | 400000000 | 858373000 | 2797800564 | 357115007.0 | PG-13 | Action | 4662.0 | 181.0 | 2019.0 |
| 11-05-20 | Pirates of the Caribbean: On Stranger Tides | /movie/Pirates-of-the-Caribbean-On-Stranger-Ti... | 379000000 | 241071802 | 1045713802 | 90151958.0 | PG-13 | Adventure | 4164.0 | 136.0 | 2011.0 |
| 15-04-22 | Avengers: Age of Ultron | /movie/Avengers-Age-of-Ultron#tab=summary | 365000000 | 459005868 | 1395316979 | 191271109.0 | PG-13 | Action | 4276.0 | 141.0 | 2015.0 |
| 15-12-16 | Star Wars Ep. VII: The Force Awakens | /movie/Star-Wars-Ep-VII-The-Force-Awakens#tab=... | 306000000 | 936662225 | 2064615817 | 247966675.0 | PG-13 | Adventure | 4134.0 | 136.0 | 2015.0 |
| 18-04-25 | Avengers: Infinity War | /movie/Avengers-Infinity-War#tab=summary | 300000000 | 678815482 | 2048359754 | 257698183.0 | PG-13 | Action | 4474.0 | 156.0 | 2018.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 13-02-06 | A Good Day to Die Hard | /movie/Die-Hard-5#tab=summary | 92000000 | 67349198 | 304249198 | 24834845.0 | R | Action | 3555.0 | 98.0 | 2013.0 |
| 14-04-09 | The Alamo | /movie/Alamo-The-(2004)#tab=summary | 92000000 | 22406362 | 23911362 | 9124701.0 | PG-13 | Western | 2609.0 | 137.0 | 2004.0 |
| 15-12-22 | Cutthroat Island | /movie/Cutthroat-Island#tab=summary | 92000000 | 10017322 | 18517322 | 2371415.0 | PG-13 | Adventure | 1619.0 | NaN | 1995.0 |
| 13-12-19 | The Secret Life of Walter Mitty | /movie/Secret-Life-of-Walter-Mitty-The-(2012)#... | 91000000 | 58236838 | 187861183 | 12765508.0 | PG | Adventure | 2922.0 | 114.0 | 2013.0 |
| 13-03-09 | 65 | /movie/65-(2022)#tab=summary | 91000000 | 0 | 0 | NaN | NaN | Thriller/Suspense | NaN | NaN | 2023.0 |

Columns

```
In [19]: df.info()#The info method prints information about the DataFrame

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 13 columns):
 #   Column              Non-Null Count  Dtype  
---  -
 0   rank                 500 non-null   int64  
 1   release_date         499 non-null   object  
 2   title                500 non-null   object  
 3   url                  500 non-null   object  
 4   production_cost      500 non-null   int64  
 5   domestic_gross       500 non-null   int64  
 6   worldwide_gross      500 non-null   int64  
 7   opening_weekend      479 non-null   float64 
 8   mpaa                 492 non-null   object  
 9   genre                495 non-null   object  
10   theaters             479 non-null   float64 
11   runtime              487 non-null   float64 
12   year                 499 non-null   float64 
dtypes: float64(4), int64(4), object(5)
memory usage: 50.9+ KB
```

```
In [20]: df.describe()#The describe() method returns description of the data in the DataFrame
```

Out[20]:

| | rank | production_cost | domestic_gross | worldwide_gross | opening_weekend | theaters | runtime | year |
|-------|------------|-----------------|----------------|-----------------|-----------------|-------------|------------|-------------|
| count | 500.000000 | 5.000000e+02 | 5.000000e+02 | 5.000000e+02 | 4.790000e+02 | 479.000000 | 487.000000 | 499.000000 |
| mean | 250.500000 | 1.494954e+08 | 1.696114e+08 | 4.697776e+08 | 5.429206e+07 | 3659.770355 | 121.885010 | 2011.394790 |
| std | 144.481833 | 4.774040e+07 | 1.421642e+08 | 3.850620e+08 | 4.619796e+07 | 546.018902 | 22.399642 | 6.560688 |
| min | 1.000000 | 9.100000e+07 | 0.000000e+00 | 0.000000e+00 | 4.855800e+04 | 30.000000 | 76.000000 | 1991.000000 |
| 25% | 125.750000 | 1.100000e+08 | 7.047110e+07 | 2.122188e+08 | 2.421873e+07 | 3378.500000 | 104.500000 | 2007.000000 |
| 50% | 250.500000 | 1.400000e+08 | 1.318470e+08 | 3.671458e+08 | 4.167120e+07 | 3735.000000 | 120.000000 | 2012.000000 |
| 75% | 375.250000 | 1.750000e+08 | 2.185998e+08 | 6.484251e+08 | 6.812391e+07 | 4065.000000 | 135.000000 | 2016.000000 |
| max | 500.000000 | 4.000000e+08 | 9.366622e+08 | 2.910371e+09 | 3.571150e+08 | 4802.000000 | 210.000000 | 2023.000000 |

```
In [21]: df.isnull()#The isnull() method returns a DataFrame object where all the values are replaced with a Boolean value True for NULL values
```

Out[21]:

| | rank | release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime | year |
|-----|-------|--------------|-------|-------|-----------------|----------------|-----------------|-----------------|-------|-------|----------|---------|-------|
| 0 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 1 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 2 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 3 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 4 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 495 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 496 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 497 | False | False | False | False | False | False | False | False | False | False | False | True | False |
| 498 | False | False | False | False | False | False | False | False | False | False | False | False | False |
| 499 | False | False | False | False | False | False | False | True | True | False | True | True | False |

500 rows × 13 columns

```
In [22]: df.isnull().sum()#check is there any null value
```

Out[22]:

| | |
|-----------------|-------|
| rank | 0 |
| release_date | 1 |
| title | 0 |
| url | 0 |
| production_cost | 0 |
| domestic_gross | 0 |
| worldwide_gross | 0 |
| opening_weekend | 21 |
| mpaa | 8 |
| genre | 5 |
| theaters | 21 |
| runtime | 13 |
| year | 1 |
| dtype: | int64 |

```
In [23]: df.head()#The head() function is used to get the first 5 rows
```

Out[23]:

| | rank | release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime |
|---|------|--------------|---|---|-----------------|----------------|-----------------|-----------------|-------|-----------|----------|---------|
| 0 | 1 | 2019-04-23 | Avengers: Endgame | /movie/Avengers-Endgame-(2019)#tab=summary | 400000000 | 858373000 | 2797800564 | 357115007.0 | PG-13 | Action | 4662.0 | 181.0 |
| 1 | 2 | 2011-05-20 | Pirates of the Caribbean: On Stranger Tides | /movie/Pirates-of-the-Caribbean-On-Stranger-Ti... | 379000000 | 241071802 | 1045713802 | 90151958.0 | PG-13 | Adventure | 4164.0 | 136.0 |
| 2 | 3 | 2015-04-22 | Avengers: Age of Ultron | /movie/Avengers-Age-of-Ultron#tab=summary | 365000000 | 459005868 | 1395316979 | 191271109.0 | PG-13 | Action | 4276.0 | 141.0 |
| 3 | 4 | 2015-12-16 | Star Wars Ep. VII: The Force Awakens | /movie/Star-Wars-Ep-VII-The-Force-Awakens#tab=... | 306000000 | 936662225 | 2064615817 | 247966675.0 | PG-13 | Adventure | 4134.0 | 136.0 |
| 4 | 5 | 2018-04-25 | Avengers: Infinity War | /movie/Avengers-Infinity-War#tab=summary | 300000000 | 678815482 | 2048359754 | 257698183.0 | PG-13 | Action | 4474.0 | 156.0 |

```
In [24]: df.tail() #The tail() function is used to get the last 5 rows
```

Out[24]:

| | release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime | year |
|---|--------------|---------------------------------|---|-----------------|----------------|-----------------|-----------------|-------|-------------------|----------|---------|--------|
| 0 | 013-02-06 | A Good Day to Die Hard | /movie/Die-Hard-5#tab=summary | 92000000 | 67349198 | 304249198 | 24834845.0 | R | Action | 3555.0 | 98.0 | 2013.0 |
| 1 | 004-04-09 | The Alamo | /movie/Alamo-The-(2004)#tab=summary | 92000000 | 22406362 | 23911362 | 9124701.0 | PG-13 | Western | 2609.0 | 137.0 | 2004.0 |
| 2 | 995-12-22 | Cutthroat Island | /movie/Cutthroat-Island#tab=summary | 92000000 | 10017322 | 18517322 | 2371415.0 | PG-13 | Adventure | 1619.0 | NaN | 1995.0 |
| 3 | 013-12-19 | The Secret Life of Walter Mitty | /movie/Secret-Life-of-Walter-Mitty-The-(2012)#... | 91000000 | 58236838 | 187861183 | 12765508.0 | PG | Adventure | 2922.0 | 114.0 | 2013.0 |
| 4 | 023-03-09 | 65 | /movie/65-(2022)#tab=summary | 91000000 | 0 | 0 | NaN | NaN | Thriller/Suspense | NaN | NaN | 2023.0 |

```
In [25]: df.genre=='Action'
```

Out[25]:

| | |
|-----|-------|
| 0 | True |
| 1 | False |
| 2 | True |
| 3 | False |
| 4 | True |
| ... | ... |
| 495 | True |
| 496 | False |
| 497 | False |
| 498 | False |
| 499 | False |

Name: genre, Length: 500, dtype: bool

```
In [63]: df[df.title=="The Avengers"]
```

Out[63]:

| | rank | release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime |
|----|------|--------------|--------------|--|-----------------|----------------|-----------------|-----------------|-------|--------|----------|---------|
| 29 | 30 | 2012-04-25 | The Avengers | /movie/Avengers-The-(2012)#tab=summary | 225000000 | 623357910 | 1515100211 | 207438708.0 | PG-13 | Action | 4349.0 | 143.0 |

In [68]: df[df.year==2018]

Out[68]:

| | rank | release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | the |
|-----|------|--------------|---|---|-----------------|----------------|-----------------|-----------------|-------|-----------|-----|
| 4 | 5 | 2018-04-25 | Avengers: Infinity War | /movie/Avengers-Infinity-War#tab=summary | 300000000 | 678815482 | 2048359754 | 257698183.0 | PG-13 | Action | 4 |
| 10 | 11 | 2018-05-23 | Solo: A Star Wars Story | /movie/Solo-A-Star-Wars-Story#tab=summary | 275000000 | 213767512 | 393151347 | 84420489.0 | PG-13 | Adventure | 4 |
| 47 | 48 | 2018-02-13 | Black Panther | /movie/Black-Panther#tab=summary | 200000000 | 700059566 | 1336494321 | 202003951.0 | PG-13 | Action | 4 |
| 49 | 50 | 2018-06-15 | Incredibles 2 | /movie/Incredibles-2#tab=summary | 200000000 | 608581744 | 1242805359 | 182687905.0 | PG | Adventure | 4 |
| 73 | 74 | 2018-11-14 | Fantastic Beasts: The Crimes of Grindelwald | /movie/Fantastic-Beasts-The-Crimes-of-Grindelw... | 200000000 | 159555901 | 648455339 | 62163104.0 | PG-13 | Adventure | 4 |
| 118 | 119 | 2018-07-16 | Mission: Impossible — Fallout | /movie/Mission-Impossible-Fallout#tab=summary | 178000000 | 220159104 | 787176729 | 61236534.0 | PG-13 | Action | 4 |
| 120 | 121 | 2018-08-09 | The Meg | /movie/Meg-The#tab=summary | 178000000 | 145443742 | 527370715 | 45402195.0 | PG-13 | Action | 4 |
| 129 | 130 | 2018-11-21 | Ralph Breaks The Internet | /movie/Ralph-Breaks-The-Internet-(2018)#tab=su... | 175000000 | 201091711 | 529290830 | 56237634.0 | PG | Adventure | 4 |
| 142 | 143 | 2018-06-06 | Jurassic World: Fallen Kingdom | /movie/Jurassic-World-Fallen-Kingdom-(2018)#ta... | 170000000 | 417719760 | 1308334005 | 148024610.0 | PG-13 | Action | 4 |
| 168 | 169 | 2018-12-06 | Aquaman | /movie/Aquaman-(2018)#tab=summary | 160000000 | 335061807 | 1143758700 | 67873522.0 | PG-13 | Action | 4 |
| 184 | 185 | 2018-03-16 | Pacific Rim: Uprising | /movie/Pacific-Rim-Uprising#tab=summary | 155000000 | 59874525 | 290930148 | 28116535.0 | PG-13 | Action | 3 |
| 214 | 215 | 2018-03-20 | Ready Player One | /movie/Ready-Player-One#tab=summary | 150000000 | 137690172 | 579055653 | 41769050.0 | PG-13 | Adventure | 4 |
| 277 | 278 | 2018-10-31 | The Nutcracker and the Four Realms | /movie/Nutcracker-and-the-Four-Realms-The-(201... | 132900000 | 54858851 | 170024300 | 20352491.0 | PG | Adventure | 3 |
| 283 | 284 | 2018-07-03 | Ant-Man and the Wasp | /movie/Ant-Man-and-the-Wasp#tab=summary | 130000000 | 216648740 | 623144660 | 75812205.0 | PG-13 | Action | 4 |
| 287 | 288 | 2018-12-19 | Mary Poppins Returns | /movie/Mary-Poppins>Returns-(2018)#tab=summary | 130000000 | 171958438 | 348901032 | 23523121.0 | PG | Musical | 4 |
| 321 | 322 | 2018-07-05 | Skyscraper | /movie/Skyscraper#tab=summary | 125000000 | 68420120 | 304868961 | 24905015.0 | PG | Action | 3 |
| 335 | 336 | 2018-04-09 | Rampage | /movie/Rampage-(2018)#tab=summary | 120000000 | 101028233 | 427947217 | 35753093.0 | PG-13 | Action | 4 |
| 347 | 348 | 2018-11-09 | Outlaw King | /movie/Outlaw-King-(2018)#tab=summary | 120000000 | 0 | 0 | NaN | R | Action | |
| 351 | 352 | 2018-10-02 | Venom | /movie/Venom-(2018)#tab=summary | 116000000 | 213511408 | 856081053 | 80255756.0 | PG-13 | Action | 4 |
| 364 | 365 | 2018-05-11 | Deadpool 2 | /movie/Deadpool-2#tab=summary | 110000000 | 324591735 | 786362370 | 125507153.0 | R | Action | 4 |
| 408 | 409 | 2018-03-09 | A Wrinkle in Time | /movie/Wrinkle-in-Time-A-(2018)#tab=summary | 103000000 | 100478608 | 133214549 | 33123609.0 | PG | Adventure | 3 |
| 412 | 413 | 2018-12-13 | Bumblebee | /movie/Bumblebee#tab=summary | 102000000 | 127195589 | 465195589 | 21654047.0 | PG-13 | Adventure | 3 |
| 459 | 460 | 2018-12-05 | Mortal Engines | /movie/Mortal-Engines-(2018)#tab=summary | 100000000 | 15951040 | 85480271 | 7559850.0 | PG-13 | Action | 3 |
| 468 | 469 | 2018-11-21 | Robin Hood | /movie/Robin-Hood-(2018)#tab=summary | 99000000 | 30824628 | 85210012 | 36063385.0 | PG-13 | Action | 3 |

```
In [70]: df[df.release_date==2019]
df
```

Out[70]:

| se_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime | year |
|----------|---|---|-----------------|----------------|-----------------|-----------------|-------|-------------------|----------|---------|--------|
| 19-04-23 | Avengers: Endgame | /movie/Avengers-Endgame-(2019)#tab=summary | 400000000 | 858373000 | 2797800564 | 357115007.0 | PG-13 | Action | 4662.0 | 181.0 | 2019.0 |
| 11-05-20 | Pirates of the Caribbean: On Stranger Tides | /movie/Pirates-of-the-Caribbean-On-Stranger-Ti... | 379000000 | 241071802 | 1045713802 | 90151958.0 | PG-13 | Adventure | 4164.0 | 136.0 | 2011.0 |
| 15-04-22 | Avengers: Age of Ultron | /movie/Avengers-Age-of-Ultron#tab=summary | 365000000 | 459005868 | 1395316979 | 191271109.0 | PG-13 | Action | 4276.0 | 141.0 | 2015.0 |
| 15-12-16 | Star Wars Ep. VII: The Force Awakens | /movie/Star-Wars-Ep-VII-The-Force-Awakens#tab=... | 306000000 | 936662225 | 2064615817 | 247966675.0 | PG-13 | Adventure | 4134.0 | 136.0 | 2015.0 |
| 18-04-25 | Avengers: Infinity War | /movie/Avengers-Infinity-War#tab=summary | 300000000 | 678815482 | 2048359754 | 257698183.0 | PG-13 | Action | 4474.0 | 156.0 | 2018.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 13-02-06 | A Good Day to Die Hard | /movie/Die-Hard-5#tab=summary | 92000000 | 67349198 | 304249198 | 24834845.0 | R | Action | 3555.0 | 98.0 | 2013.0 |
| 14-04-09 | The Alamo | /movie/Alamo-The-(2004)#tab=summary | 92000000 | 22406362 | 23911362 | 9124701.0 | PG-13 | Western | 2609.0 | 137.0 | 2004.0 |
| 15-12-22 | Cutthroat Island | /movie/Cutthroat-Island#tab=summary | 92000000 | 10017322 | 18517322 | 2371415.0 | PG-13 | Adventure | 1619.0 | NaN | 1995.0 |
| 13-12-19 | The Secret Life of Walter Mitty | /movie/Secret-Life-of-Walter-Mitty-The-(2012)#... | 91000000 | 58236838 | 187861183 | 12765508.0 | PG | Adventure | 2922.0 | 114.0 | 2013.0 |
| 23-03-09 | 65 | /movie/65-(2022)#tab=summary | 91000000 | 0 | 0 | NaN | NaN | Thriller/Suspense | NaN | NaN | 2023.0 |

columns

```
In [27]: df[df.genre=='Action']
```

Out[27]:

| | rank | release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theat |
|-----|------|--------------|------------------------------|--|-----------------|----------------|-----------------|-----------------|-------|--------|--------|
| 0 | 1 | 2019-04-23 | Avengers: Endgame | /movie/Avengers-Endgame-(2019)#tab=summary | 400000000 | 858373000 | 2797800564 | 357115007.0 | PG-13 | Action | 4662.0 |
| 2 | 3 | 2015-04-22 | Avengers: Age of Ultron | /movie/Avengers-Age-of-Ultron#tab=summary | 365000000 | 459005868 | 1395316979 | 191271109.0 | PG-13 | Action | 4276.0 |
| 4 | 5 | 2018-04-25 | Avengers: Infinity War | /movie/Avengers-Infinity-War#tab=summary | 300000000 | 678815482 | 2048359754 | 257698183.0 | PG-13 | Action | 4474.0 |
| 6 | 7 | 2017-11-13 | Justice League | /movie/Justice-League-(2017)#tab=summary | 300000000 | 229024295 | 655945209 | 93842239.0 | PG-13 | Action | 4055.0 |
| 7 | 8 | 2015-10-06 | Spectre | /movie/Spectre#tab=summary | 300000000 | 200074175 | 879500760 | 70403148.0 | PG-13 | Action | 3921.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 479 | 480 | 2010-11-12 | Unstoppable | /movie/Unstoppable#tab=summary | 95000000 | 81562942 | 165720921 | 22688457.0 | PG-13 | Action | 3261.0 |
| 480 | 481 | 1997-05-09 | The Fifth Element | /movie/Fifth-Element-The#tab=summary | 95000000 | 63570862 | 263893838 | 17031345.0 | PG-13 | Action | 2500.0 |
| 489 | 490 | 2015-01-29 | Kingsman: The Secret Service | /movie/Kingsman-The-Secret-Service#tab=summary | 94000000 | 128261724 | 404561724 | 36206331.0 | R | Action | 3261.0 |
| 491 | 492 | 2001-07-18 | Jurassic Park III | /movie/Jurassic-Park-3#tab=summary | 93000000 | 181166115 | 365900000 | 50771645.0 | PG-13 | Action | 3471.0 |
| 495 | 496 | 2013-02-06 | A Good Day to Die Hard | /movie/Die-Hard-5#tab=summary | 92000000 | 67349198 | 304249198 | 24834845.0 | R | Action | 3555.0 |

202 rows × 13 columns

In [28]:

df[df.opening_weekend>=191271109.0]

Out[28]:

| release_date | title | url | production_cost | domestic_gross | worldwide_gross | opening_weekend | mpaa | genre | theaters | runtime | year |
|--------------|--------------------------------------|---|-----------------|----------------|-----------------|-----------------|-------|-----------|----------|---------|--------|
| 2019-04-23 | Avengers: Endgame | /movie/Avengers-Endgame-(2019)#tab=summary | 400000000 | 858373000 | 2797800564 | 357115007.0 | PG-13 | Action | 4662.0 | 181.0 | 2019.0 |
| 2015-04-22 | Avengers: Age of Ultron | /movie/Avengers-Age-of-Ultron#tab=summary | 365000000 | 459005868 | 1395316979 | 191271109.0 | PG-13 | Action | 4276.0 | 141.0 | 2015.0 |
| 2015-12-16 | Star Wars Ep. VII: The Force Awakens | /movie/Star-Wars-Ep-VII-The-Force-Awakens#tab=... | 306000000 | 936662225 | 2064615817 | 247966675.0 | PG-13 | Adventure | 4134.0 | 136.0 | 2015.0 |
| 2018-04-25 | Avengers: Infinity War | /movie/Avengers-Infinity-War#tab=summary | 300000000 | 678815482 | 2048359754 | 257698183.0 | PG-13 | Action | 4474.0 | 156.0 | 2018.0 |
| 2017-12-13 | Star Wars Ep. VIII: The Last Jedi | /movie/Star-Wars-Ep-VIII-The-Last-Jedi#tab=sum... | 262000000 | 620181382 | 1331635141 | 220009584.0 | PG-13 | Adventure | 4232.0 | 150.0 | 2017.0 |
| 2019-07-11 | The Lion King | /movie/Lion-King-The-(Live-Action)-(2019)#tab=... | 260000000 | 543638043 | 1647778651 | 191770759.0 | PG | Adventure | 4802.0 | 118.0 | 2019.0 |
| 2012-04-25 | The Avengers | /movie/Avengers-The-(2012)#tab=summary | 225000000 | 623357910 | 1515100211 | 207438708.0 | PG-13 | Action | 4349.0 | 143.0 | 2012.0 |
| 2015-06-09 | Jurassic World | /movie/Jurassic-World#tab=summary | 215000000 | 652306625 | 1669979967 | 208806270.0 | PG-13 | Action | 4291.0 | 124.0 | 2015.0 |
| 2021-12-14 | Spider-Man: No Way Home | /movie/Spider-Man-No-Way-Home-(2021)#tab=summary | 200000000 | 814108407 | 1912775610 | 260138569.0 | PG-13 | Action | 4336.0 | 148.0 | 2021.0 |
| 2018-02-13 | Black Panther | /movie/Black-Panther#tab=summary | 200000000 | 700059566 | 1336494321 | 202003951.0 | PG-13 | Action | 4084.0 | 120.0 | 2018.0 |

In [29]:

df[['title','genre','runtime']]

Out[29]:

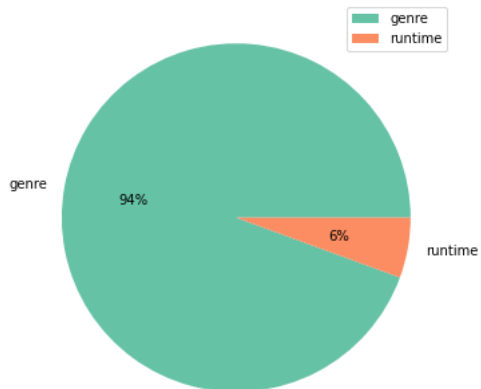
| | title | genre | runtime |
|-----|---|-----------|---------|
| 0 | Avengers: Endgame | Action | 181.0 |
| 1 | Pirates of the Caribbean: On Stranger Tides | Adventure | 136.0 |
| 2 | Avengers: Age of Ultron | Action | 141.0 |
| 3 | Star Wars Ep. VII: The Force Awakens | Adventure | 136.0 |
| 4 | Avengers: Infinity War | Action | 156.0 |
| ... | ... | ... | ... |
| 495 | A Good Day to Die Hard | Action | 98.0 |
| 496 | The Alamo | Western | 137.0 |
| 497 | Cutthroat Island | Adventure | NaN |
| 498 | The Secret Life of Walter Mitty | Adventure | 114.0 |
| 499 | 65 Thriller/Suspense | | NaN |

500 rows × 3 columns

In [101]:

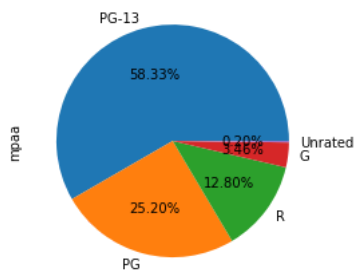
Yrs=df["year"].sum()
RT=df["runtime"].sum()

```
In [102]: plt.figure(figsize=(13,6))
data=Yrs,RT]
labels=['genre','runtime']
color=sns.color_palette('Set2')
plt.pie(data,labels =labels,colors=color,autopct='%1.0f%%')
plt.legend()
plt.show()
```



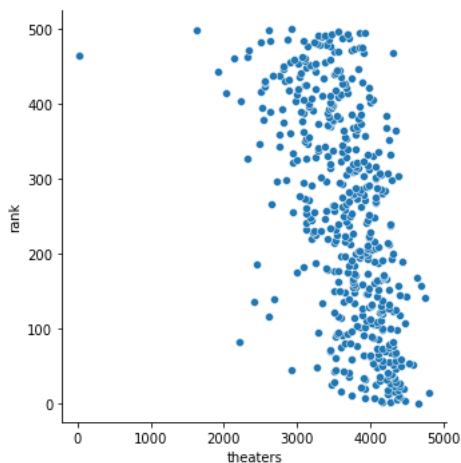
```
In [86]: df['mpaa'].value_counts().plot(kind='pie', autopct="%1.2f%%")
```

```
Out[86]: <AxesSubplot:ylabel='mpaa'>
```



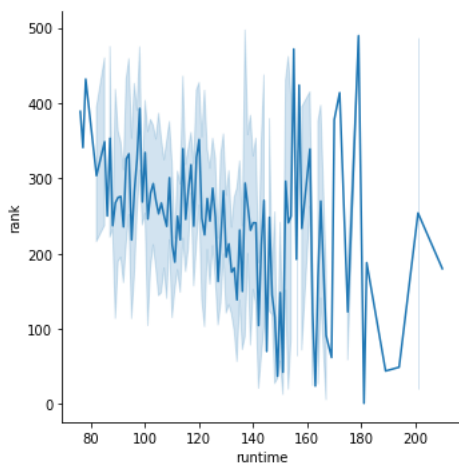
```
In [30]: sns.relplot(x='theaters',y='rank',data=df)
```

```
Out[30]: <seaborn.axisgrid.FacetGrid at 0x22964537ca0>
```



```
In [31]: sns.relplot(x='runtime',y='rank',data=df,kind='line')
```

```
Out[31]: <seaborn.axisgrid.FacetGrid at 0x22964ca69a0>
```



```
In [88]: df['runtime'].sum()
```

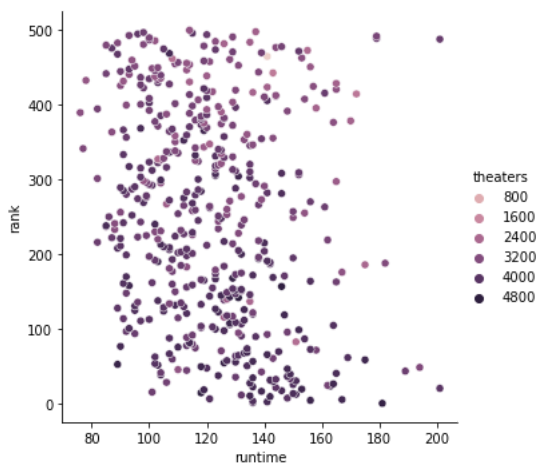
```
Out[88]: 59358.0
```

```
In [89]: df['mpaa'].unique()
```

```
Out[89]: array(['PG-13', nan, 'PG', 'G', 'R', 'Unrated'], dtype=object)
```

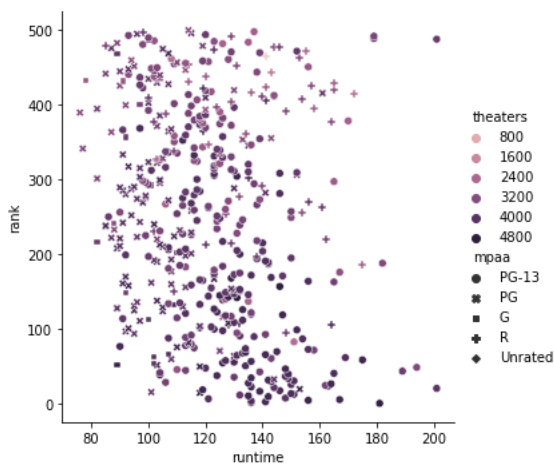
```
In [32]: sns.relplot(x='runtime',y='rank',data=df,hue='theaters')
```

```
Out[32]: <seaborn.axisgrid.FacetGrid at 0x229645371f0>
```



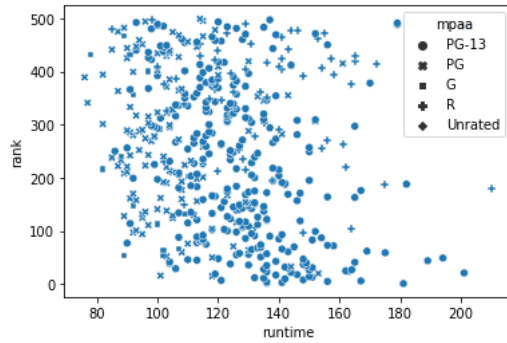
```
In [33]: sns.relplot(x='runtime',y='rank',data=df,hue='theaters',style='mpaa')
```

```
Out[33]: <seaborn.axisgrid.FacetGrid at 0x22964e24280>
```



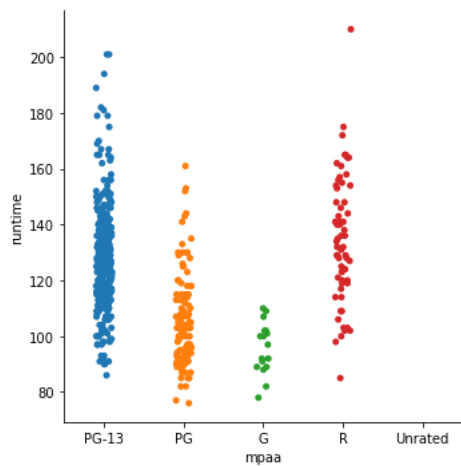

```
In [34]: sns.scatterplot(x="runtime",y="rank",style="mpaa",data=df)
```

```
Out[34]: <AxesSubplot:xlabel='runtime', ylabel='rank'>
```



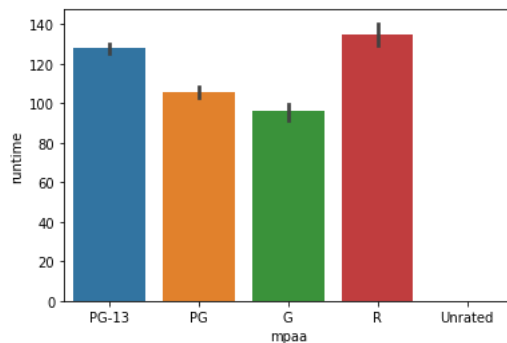
```
In [35]: sns.catplot(x='mpaa',y='runtime',data=df)#mpaa=Motion Picture Association Of America
```

```
Out[35]: <seaborn.axisgrid.FacetGrid at 0x22964eae2e0>
```



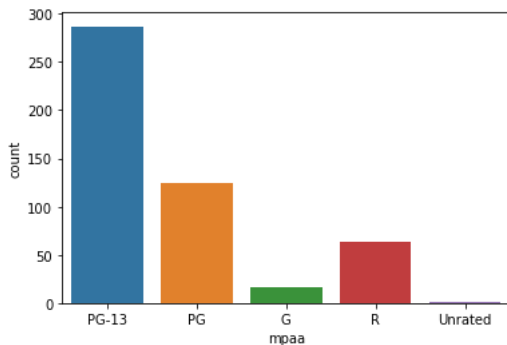
```
In [36]: sns.barplot(x='mpaa',y='runtime',data=df)
```

```
Out[36]: <AxesSubplot:xlabel='mpaa', ylabel='runtime'>
```



```
In [37]: sns.countplot(x='mpaa',data=df)
```

```
Out[37]: <AxesSubplot:xlabel='mpaa', ylabel='count'>
```



```
In [38]: sns.catplot(x='mpaa',y='runtime',data=df,kind='swarm',hue='genre')
```

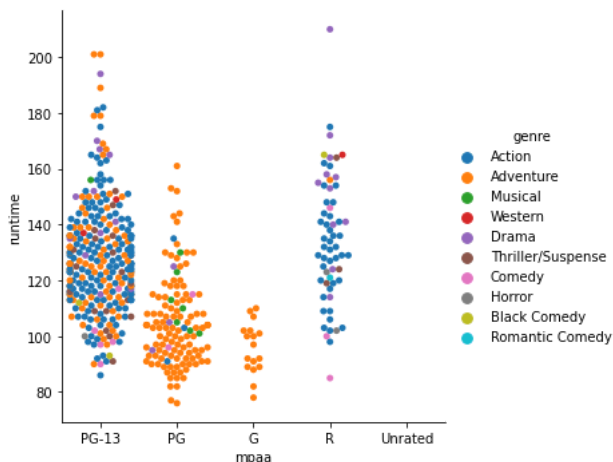
C:\Users\Administrator\anaconda3\lib\site-packages\seaborn\categorical.py:1296: UserWarning: 20.9% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.

warnings.warn(msg, UserWarning)

C:\Users\Administrator\anaconda3\lib\site-packages\seaborn\categorical.py:1296: UserWarning: 5.6% of the points cannot be placed; you may want to decrease the size of the markers or use stripplot.

warnings.warn(msg, UserWarning)

```
Out[38]: <seaborn.axisgrid.FacetGrid at 0x22965f5f8b0>
```

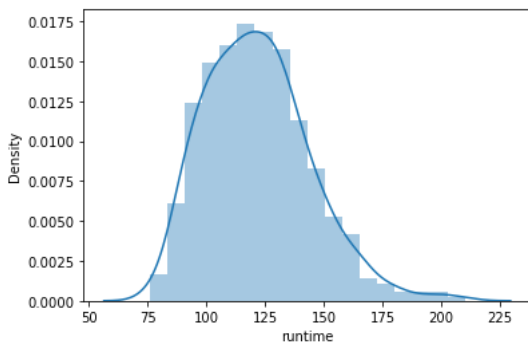


```
In [39]: sns.distplot(df['runtime'])
```

C:\Users\Administrator\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

warnings.warn(msg, FutureWarning)

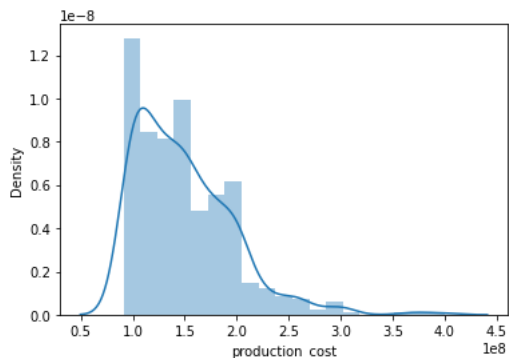
```
Out[39]: <AxesSubplot:xlabel='runtime', ylabel='Density'>
```



```
In [40]: sns.distplot(df['production_cost'],kde=True)# KDE shows the density where the points match up the most
```

C:\Users\Administrator\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).
warnings.warn(msg, FutureWarning)

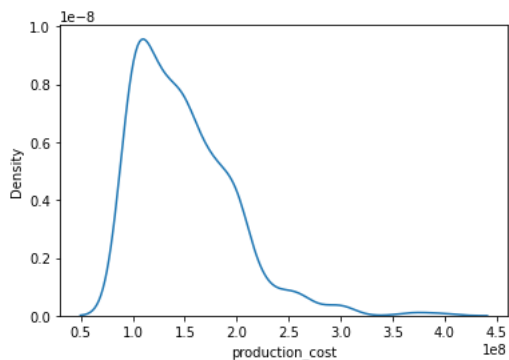
```
Out[40]: <AxesSubplot:xlabel='production_cost', ylabel='Density'>
```



```
In [41]: sns.distplot(df['production_cost'],kde=True,hist=False)
```

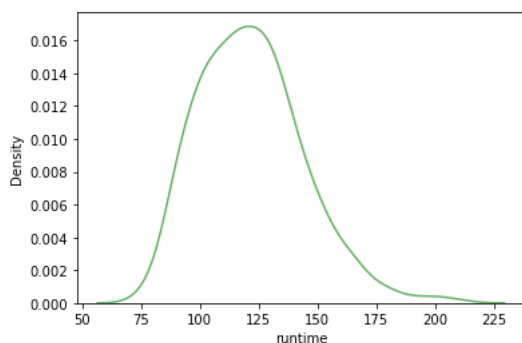
C:\Users\Administrator\anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `kdeplot` (an axes-level function for kernel density plots).
warnings.warn(msg, FutureWarning)

```
Out[41]: <AxesSubplot:xlabel='production_cost', ylabel='Density'>
```



```
In [42]: sns.kdeplot(df['runtime'],shade=False,color='g',alpha=0.6)
```

```
Out[42]: <AxesSubplot:xlabel='runtime', ylabel='Density'>
```

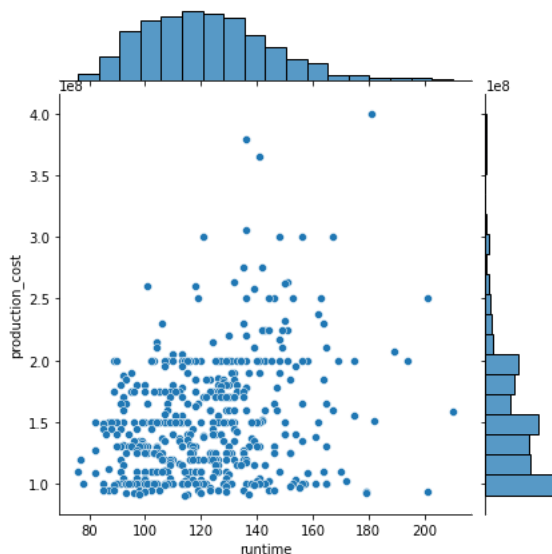


```
In [43]: sns.jointplot(df['runtime'],df['production_cost'])
```

C:\Users\Administrator\anaconda3\lib\site-packages\seaborn_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.

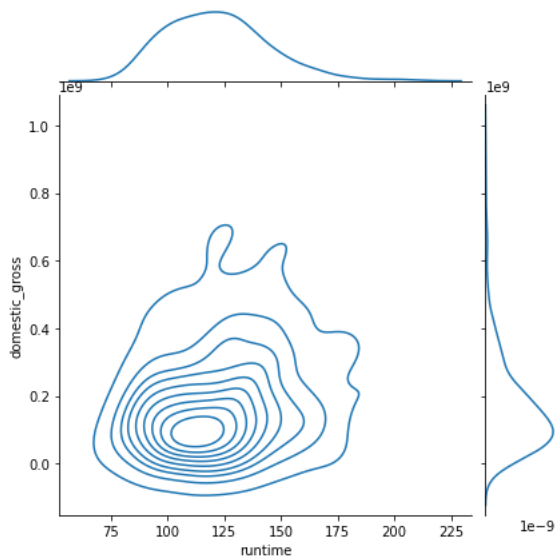
warnings.warn()

```
Out[43]: <seaborn.axisgrid.JointGrid at 0x2296672de80>
```



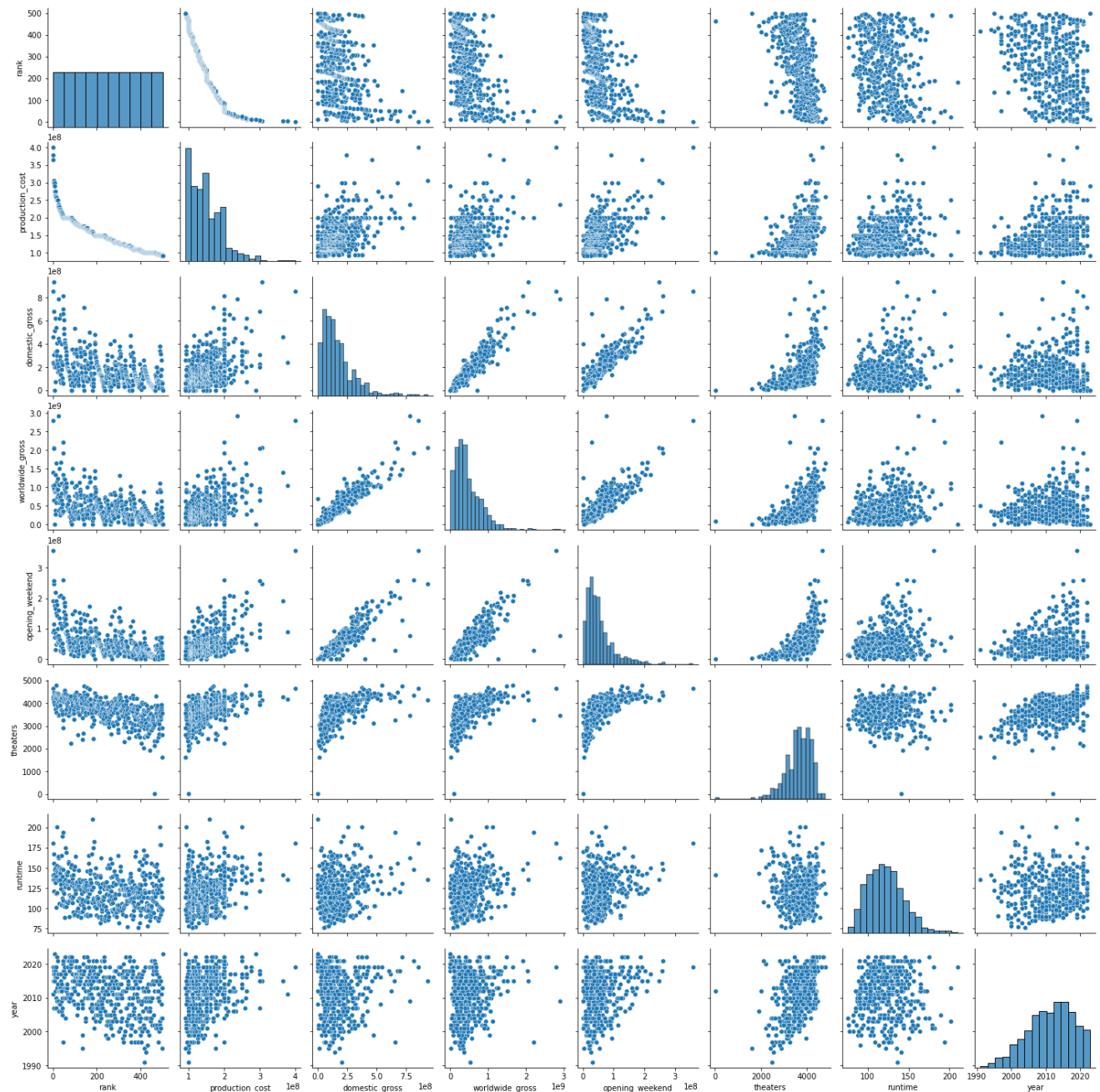
```
In [44]: sns.jointplot(x='runtime', y='domestic_gross', data=df, kind='kde')
```

```
Out[44]: <seaborn.axisgrid.JointGrid at 0x229645131f0>
```



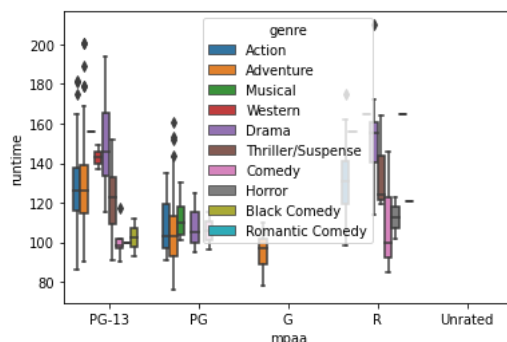
```
In [45]: sns.pairplot(df)
```

```
Out[45]: <seaborn.axisgrid.PairGrid at 0x229669a31c0>
```



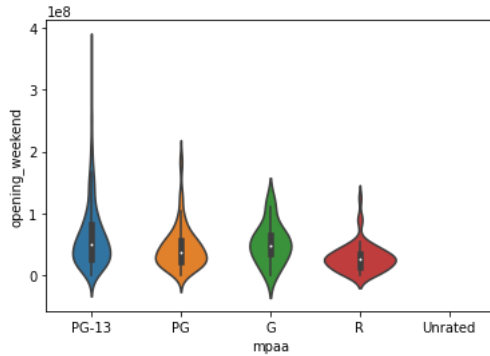
```
In [46]: sns.boxplot(x='mpaa', y='runtime', data=df, hue='genre')
```

```
Out[46]: <AxesSubplot:xlabel='mpaa', ylabel='runtime'>
```



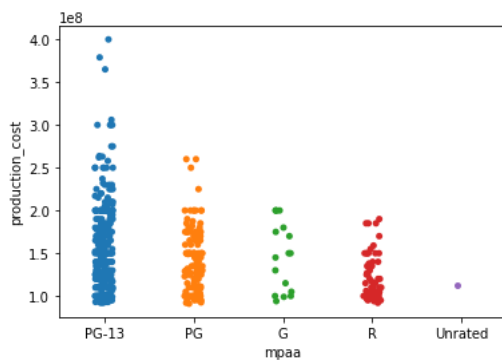
```
In [47]: sns.violinplot(x='mpaa', y='opening_weekend',data=df)
```

```
Out[47]: <AxesSubplot:xlabel='mpaa', ylabel='opening_weekend'>
```



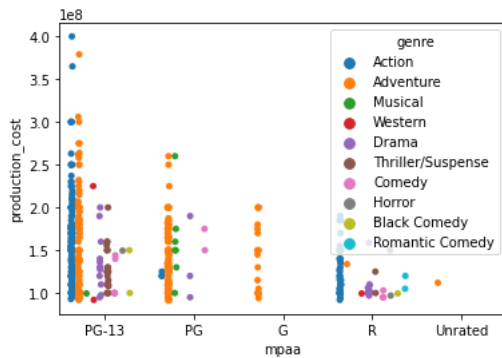
```
In [48]: sns.stripplot(x='mpaa', y='production_cost',data=df)
```

```
Out[48]: <AxesSubplot:xlabel='mpaa', ylabel='production_cost'>
```



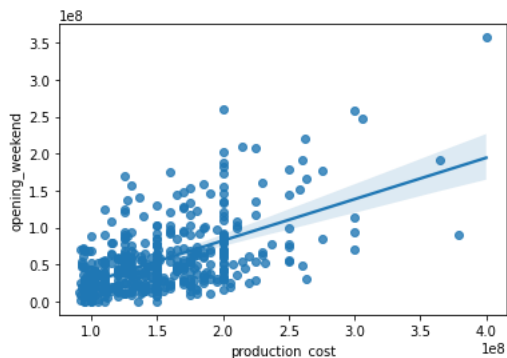
```
In [49]: sns.stripplot(x='mpaa', y='production_cost',data=df,jitter = True,
hue = 'genre',dodge=True)
```

```
Out[49]: <AxesSubplot:xlabel='mpaa', ylabel='production_cost'>
```



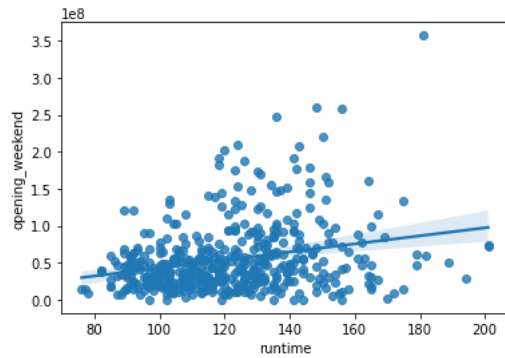
```
In [50]: sns.regplot(x='production_cost',y='opening_weekend',data=df)
```

```
Out[50]: <AxesSubplot:xlabel='production_cost', ylabel='opening_weekend'>
```



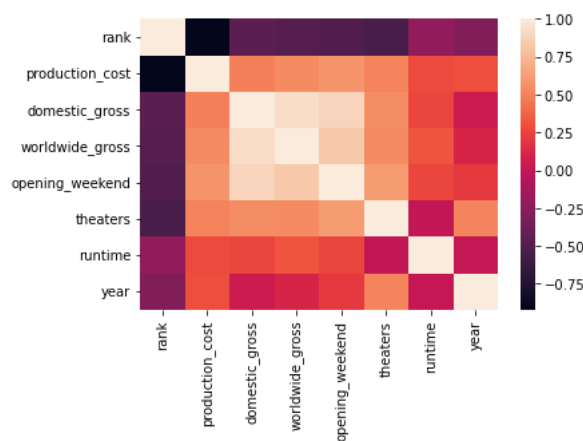
```
In [51]: sns.regplot(x='runtime',y='opening_weekend',data=df)
```

```
Out[51]: <AxesSubplot:xlabel='runtime', ylabel='opening_weekend'>
```



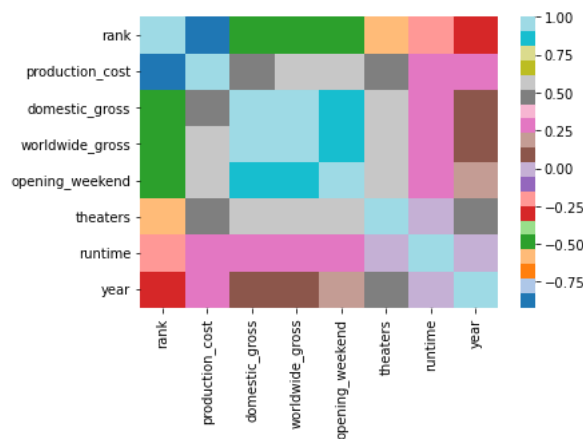
```
In [52]: a=df.corr()  
sns.heatmap(a)
```

```
Out[52]: <AxesSubplot:>
```



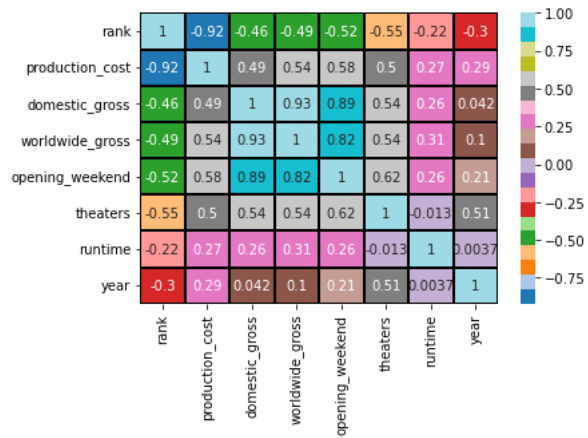
```
In [53]: sns.heatmap(df.corr(),cmap='tab20')
```

```
Out[53]: <AxesSubplot:>
```



```
In [54]: sns.heatmap(df.corr(),cmap='tab20',annot=True,linewidth=2,linecolor='black')
```

```
Out[54]: <AxesSubplot:>
```



```
In [55]: sns.heatmap(df.corr(),cmap='tab20',cbar=False,annot=True,linewidth=2,linecolor='black')
```

```
Out[55]: <AxesSubplot:>
```

