```python
import pandas as pd
import numpy as np

movies = pd.read_csv('imdb-top-1000.csv')
movies.head(2)
```

```
              Series_Title Released_Year  Runtime  Genre  IMDB_Rating
\
0  The Shawshank Redemption          1994      142  Drama          9.3

1             The Godfather          1972      175  Crime          9.2


               Director          Star1  No_of_Votes       Gross
Metascore
0        Frank Darabont    Tim Robbins      2343110   28341469.0
80.0
1  Francis Ford Coppola  Marlon Brando      1620367  134966411.0
100.0
```

```python
# Groupby - is basically the study of groups
# group by apply hota hai categorical column ke upar (jiska ki value
humesa change karte rahata hai for ex : jenre in our data set)

movies.shape
```

```
(1000, 10)
```

```python
# genre ke basis pe movie ko categorise karna hai suppose sare Drama
type ke movie ko group karna hai then sare crime type ko alag

movies.groupby('Genre')   # abhi kuch output nahi dikhega bass pata
chalega ki hum jo operation kye hain usse DataFrameGroupBy object
create ho gya hai
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x00000290C6ACB8C0>
```

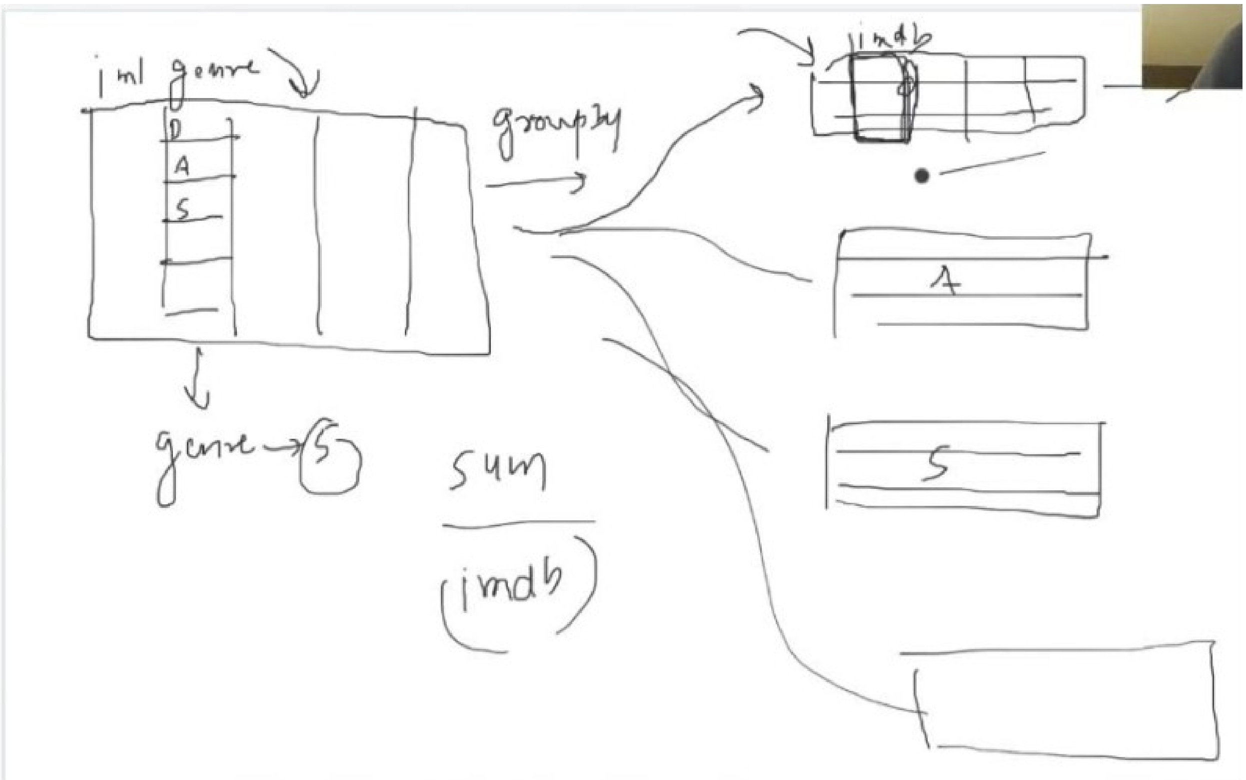
```python
genres = movies.groupby('Genre')

# Applying builtin aggregation functions on groupby objects
genres.sum()


# yaha pe jo output aa raha hai usme har row me alag alag genre hai
means ki ye table genres ke basis pe groupby ho gya hai and sum
function ko call
# karne se particular column ke sare data add ho gya hai
```

```
                                                Series_Title  \
Genre
Action     The Dark KnightThe Lord of the Rings: The Retu...
```

```
Adventure  InterstellarBack to the FutureInglourious Bast...
Animation  Sen to Chihiro no kamikakushiThe Lion KingHota...
Biography  Schindler's ListGoodfellasHamiltonThe Intoucha...
Comedy     GisaengchungLa vita è bellaModern TimesCity Li...
Crime      The GodfatherThe Godfather: Part II12 Angry Me...
Drama      The Shawshank RedemptionFight ClubForrest Gump...
Family     E.T. the Extra-TerrestrialWilly Wonka & the Ch...
Fantasy                Das Cabinet des Dr. CaligariNosferatu
Film-Noir   The Third ManThe Maltese FalconShadow of a Doubt
Horror     PsychoAlienThe ThingThe ExorcistNight of the L...
Mystery    MementoRear WindowVertigoShutter IslandKahaani...
Thriller                                     Wait Until Dark
Western    Il buono, il brutto, il cattivoOnce Upon a Tim...

                                               Released_Year  Runtime
\
Genre

Action     2008200320102001200219991980197719621954200019...    22196

Adventure  2014198520091981196819621959201319751963194819...     9656

Animation  2001199419882016201820172008199719952019200920...     8166

Biography  1993199020202011200220171995198420182013201320...    11970

Comedy     2019199719361931200919641940200120001973196019...    17380

Crime      1972197419571994200219991995199120192006199519...    13524

Drama      1994199919941975202019981946201420061998198819...    36049

Family                                              19821971      215

Fantasy                                             19201922      170

Film-Noir                                       194919411943      312

Horror          1960197919821973196819612017197819332004200l     1123

Mystery     2000195419582010201219951972193819882012199819971     1429

Thriller                                                1967      108

Western                                     1966196819651976      593


           IMDB_Rating
Director  \
Genre
```

```
Action          1367.3  Christopher NolanPeter JacksonChristopher
Nola...
Adventure        571.5  Christopher NolanRobert ZemeckisQuentin
Tarant...
Animation        650.3  Hayao MiyazakiRoger AllersIsao
TakahataMakoto ...
Biography        698.6  Steven SpielbergMartin ScorseseThomas
KailOliv...
Comedy          1224.7  Bong Joon HoRoberto BenigniCharles
ChaplinChar...
Crime            857.8  Francis Ford CoppolaFrancis Ford
CoppolaSidney...
Drama           2299.7  Frank DarabontDavid FincherRobert
ZemeckisMilo...
Family            15.6                        Steven SpielbergMel
Stuart
Fantasy           16.0                           Robert WieneF.W.
Murnau
Film-Noir         23.9              Carol ReedJohn HustonAlfred
Hitchcock
Horror            87.0  Alfred HitchcockRidley ScottJohn
CarpenterWill...
Mystery           95.7  Christopher NolanAlfred HitchcockAlfred
Hitchc...
Thriller           7.8                                     Terence
Young
Western           33.4  Sergio LeoneSergio LeoneSergio LeoneClint
East...

                                                       Star1
No_of_Votes  \
Genre

Action     Christian BaleElijah WoodLeonardo DiCaprioElij...
72282412
Adventure  Matthew McConaugheyMichael J. FoxBrad PittJürg...
22576163
Animation  Daveigh ChaseRob MinkoffTsutomu TatsumiRyûnosu...
21978630
Biography  Liam NeesonRobert De NiroLin-Manuel MirandaÉri...
24006844
Comedy     Kang-ho SongRoberto BenigniCharles ChaplinChar...
27620327
Crime      Marlon BrandoAl PacinoHenry FondaJohn Travolta...
33533615
Drama      Tim RobbinsBrad PittTom HanksJack NicholsonSur...
61367304
Family                              Henry ThomasGene Wilder
551221
```

```
Fantasy                                    Werner KraussMax Schreck
146222
Film-Noir          Orson WellesHumphrey BogartTeresa Wright
367215
Horror     Anthony PerkinsSigourney WeaverKurt RussellEll...
3742556
Mystery    Guy PearceJames StewartJames StewartLeonardo D...
4203004
Thriller                                         Audrey Hepburn
27733
Western    Clint EastwoodHenry FondaClint EastwoodClint E...
1289665

                  Gross  Metascore
Genre
Action     3.263226e+10    10499.0
Adventure  9.496922e+09     5020.0
Animation  1.463147e+10     6082.0
Biography  8.276358e+09     6023.0
Comedy     1.566387e+10     9840.0
Crime      8.452632e+09     6706.0
Drama      3.540997e+10    19208.0
Family     4.391106e+08      158.0
Fantasy    7.827267e+08        0.0
Film-Noir  1.259105e+08      287.0
Horror     1.034649e+09      880.0
Mystery    1.256417e+09      633.0
Thriller   1.755074e+07       81.0
Western    5.822151e+07      313.0
```




```
genres.min()    # minimum value nikal ke de dega  and kuch chize observe kar ssakte hain jaise ki Action me jo sabse purani movie hai wo 1924 me aayi
# thi , action movie ka minimum length 45 hai , advanture movie me sabse kam rating 7.6 hai and so on.....
```

```
                                Series_Title Released_Year  Runtime  \
Genre
Action                                   300          1924       45
Adventure            2001: A Space Odyssey          1925       88
Animation                              Akira          1940       71
Biography                   12 Years a Slave          1928       93
Comedy                   (500) Days of Summer          1921       68
Crime                           12 Angry Men          1931       80
Drama                                   1917          1925       64
Family           E.T. the Extra-Terrestrial          1971      100
Fantasy         Das Cabinet des Dr. Caligari          1920       76
Film-Noir                   Shadow of a Doubt          1941      100
Horror                                 Alien          1933       71
Mystery                            Dark City          1938       96
Thriller                     Wait Until Dark          1967      108
Western     Il buono, il brutto, il cattivo          1965      132

           IMDB_Rating               Director
Star1  \
Genre
```

```
Action                 7.6       Abhishek Chaubey                Aamir
Khan
Adventure              7.6         Akira Kurosawa                Aamir
Khan
Animation              7.6            Adam Elliot             Adrian
Molina
Biography              7.6             Adam McKay              Adrien
Brody
Comedy                 7.6  Alejandro G. Iñárritu                Aamir
Khan
Crime                  7.6         Akira Kurosawa                 Ajay
Devgn
Drama                  7.6             Aamir Khan                Abhay
Deol
Family                 7.8             Mel Stuart               Gene
Wilder
Fantasy                7.9            F.W. Murnau               Max
Schreck
Film-Noir              7.8       Alfred Hitchcock           Humphrey
Bogart
Horror                 7.6     Alejandro Amenábar            Anthony
Perkins
Mystery                7.6            Alex Proyas  Bernard-Pierre
Donnadieu
Thriller               7.8          Terence Young             Audrey
Hepburn
Western                7.8         Clint Eastwood              Clint
Eastwood

           No_of_Votes        Gross  Metascore
Genre
Action           25312       3296.0       33.0
Adventure        29999      61001.0       41.0
Animation        25229     128985.0       61.0
Biography        27254      21877.0       48.0
Comedy           26337       1305.0       45.0
Crime            27712       6013.0       47.0
Drama            25088       3600.0       28.0
Family          178731    4000000.0       67.0
Fantasy          57428  337574718.0        NaN
Film-Noir        59556     449191.0       94.0
Horror           27007      89029.0       46.0
Mystery          33982    1035953.0       52.0
Thriller         27733   17550741.0       81.0
Western          65659    5321508.0       69.0

genres.max()
```

```
                                   Series_Title Released_Year  Runtime  \
Genre

Action                                     Yôjinbô          2019      321

Adventure                               Zombieland            PG      228

Animation           Ôkami kodomo no Ame to Yuki          2020      137

Biography                                  Zerkalo          2020      209

Comedy                  Zindagi Na Milegi Dobara          2020      188

Crime                           À bout de souffle          2019      229

Drama                                    Zwartboek          2020      242

Family     Willy Wonka & the Chocolate Factory          1982      115

Fantasy                                  Nosferatu          1922       94

Film-Noir                            The Third Man          1949      108

Horror                                   The Thing          2017      122

Mystery                                    Vertigo          2012      138

Thriller                            Wait Until Dark          1967      108

Western                      The Outlaw Josey Wales          1976      165


           IMDB_Rating           Director             Star1
No_of_Votes  \
Genre

Action             9.0        Zack Snyder      Yun-Fat Chow
2303232
Adventure          8.6  Ömer Faruk Sorak      Yves Montand
1512360
Animation          8.6   Yoshifumi Kondô      Yôji Matsuda
999790
Biography          8.9       Tom McCarthy     Éric Toledano
1213505
Comedy             8.6        Zoya Akhtar  Ömer Faruk Sorak
939631
Crime              9.2       Yavuz Turgul    Vincent Cassel
1826188
Drama              9.3        Çagan Irmak    Çetin Tekindor
2343110
```

```
Family                7.8  Steven Spielberg      Henry Thomas
372490
Fantasy               8.1       Robert Wiene     Werner Krauss
88794
Film-Noir             8.1        John Huston     Teresa Wright
158731
Horror                8.5  William Friedkin  Sigourney Weaver
787806
Mystery               8.4     Terry Gilliam        Vidya Balan
1129894
Thriller              7.8     Terence Young    Audrey Hepburn
27733
Western               8.8      Sergio Leone        Henry Fonda
688390

                 Gross  Metascore
Genre
Action     936662225.0       98.0
Adventure  874211619.0      100.0
Animation  873839108.0       96.0
Biography  753585104.0       97.0
Comedy     886752933.0       99.0
Crime      790482117.0      100.0
Drama      924558264.0      100.0
Family     435110554.0       91.0
Fantasy    445151978.0        NaN
Film-Noir  123353292.0       97.0
Horror     298791505.0       97.0
Mystery    474203697.0      100.0
Thriller    17550741.0       81.0
Western     31800000.0       90.0
```

```
genres.mean()  # error dega qki bale column ka mean kaise find karega
for that data should be integer

----------------------------------------------------------------------
-----
TypeError                                 Traceback (most recent call
last)
File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:1942, in GroupBy._agg_py_fallback(self, how, values, ndim,
alt)
   1941 try:
-> 1942     res_values = self._grouper.agg_series(ser, alt,
preserve_dtype=True)
   1943 except Exception as err:
```

```
File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:864, in
BaseGrouper.agg_series(self, obj, func, preserve_dtype)
    862     preserve_dtype = True
--> 864 result = self._aggregate_series_pure_python(obj, func)
    866 npvalues = lib.maybe_convert_objects(result, try_float=False)

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:885, in
BaseGrouper._aggregate_series_pure_python(self, obj, func)
    884 for i, group in enumerate(splitter):
--> 885     res = func(group)
    886     res = extract_result(res)

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:2454, in GroupBy.mean.<locals>.<lambda>(x)
   2451 else:
   2452     result = self._cython_agg_general(
   2453         "mean",
-> 2454         alt=lambda x: Series(x,
copy=False).mean(numeric_only=numeric_only),
   2455         numeric_only=numeric_only,
   2456     )
   2457     return result.__finalize__(self.obj, method="groupby")

File ~\anaconda3\Lib\site-packages\pandas\core\series.py:6549, in
Series.mean(self, axis, skipna, numeric_only, **kwargs)
   6541 @doc(make_doc("mean", ndim=1))
   6542 def mean(
   6543     self,
   (...)
   6547     **kwargs,
   6548 ):
-> 6549     return NDFrame.mean(self, axis, skipna, numeric_only,
**kwargs)

File ~\anaconda3\Lib\site-packages\pandas\core\generic.py:12420, in
NDFrame.mean(self, axis, skipna, numeric_only, **kwargs)
  12413 def mean(
  12414     self,
  12415     axis: Axis | None = 0,
   (...)
  12418     **kwargs,
  12419 ) -> Series | float:
> 12420     return self._stat_function(
  12421         "mean", nanops.nanmean, axis, skipna, numeric_only,
**kwargs
  12422     )

File ~\anaconda3\Lib\site-packages\pandas\core\generic.py:12377, in
NDFrame._stat_function(self, name, func, axis, skipna, numeric_only,
**kwargs)
```

```
  12375 validate_bool_kwarg(skipna, "skipna", none_allowed=False)
> 12377 return self._reduce(
  12378     func, name=name, axis=axis, skipna=skipna,
numeric_only=numeric_only
  12379 )

File ~\anaconda3\Lib\site-packages\pandas\core\series.py:6457, in
Series._reduce(self, op, name, axis, skipna, numeric_only,
filter_type, **kwds)
   6453     raise TypeError(
   6454         f"Series.{name} does not allow
{kwd_name}={numeric_only} "
   6455         "with non-numeric dtypes."
   6456     )
-> 6457 return op(delegate, skipna=skipna, **kwds)

File ~\anaconda3\Lib\site-packages\pandas\core\nanops.py:147, in
bottleneck_switch.__call__.<locals>.f(values, axis, skipna, **kwds)
    146 else:
--> 147     result = alt(values, axis=axis, skipna=skipna, **kwds)
    149 return result

File ~\anaconda3\Lib\site-packages\pandas\core\nanops.py:404, in
_datetimelike_compat.<locals>.new_func(values, axis, skipna, mask,
**kwargs)
    402     mask = isna(values)
--> 404 result = func(values, axis=axis, skipna=skipna, mask=mask,
**kwargs)
    406 if datetimelike:

File ~\anaconda3\Lib\site-packages\pandas\core\nanops.py:720, in
nanmean(values, axis, skipna, mask)
    719 the_sum = values.sum(axis, dtype=dtype_sum)
--> 720 the_sum = _ensure_numeric(the_sum)
    722 if axis is not None and getattr(the_sum, "ndim", False):

File ~\anaconda3\Lib\site-packages\pandas\core\nanops.py:1701, in
_ensure_numeric(x)
   1699 if isinstance(x, str):
   1700     # GH#44008, GH#36703 avoid casting e.g. strings to numeric
-> 1701     raise TypeError(f"Could not convert string '{x}' to
numeric")
   1702 try:

TypeError: Could not convert string 'The Dark KnightThe Lord of the
Rings: The Return of the KingInceptionThe Lord of the Rings: The
Fellowship of the RingThe Lord of the Rings: The Two TowersThe
MatrixStar Wars: Episode V - The Empire Strikes BackStar
WarsSeppukuShichinin no samuraiGladiatorLéonTerminator 2: Judgment
DayVikram VedhaDangalAvengers: EndgameAvengers: Infinity WarThe Dark
```

```
Knight RisesOldeuboiRaiders of the Lost ArkAliensStar Wars: Episode VI
- Return of the JediUri: The Surgical StrikeK.G.F: Chapter 1Baahubali
2: The ConclusionGangs of WasseypurPaan Singh TomarWarriorV for
VendettaBatman BeginsHera PheriLock, Stock and Two Smoking
BarrelsAndaz Apna ApnaIndiana Jones and the Last CrusadeDie
HardRanSholayYôjinbôSherlock Jr.HaiderLoganRushFord v FerrariMad Max:
Fury RoadA WednesdayTaegukgi hwinalrimyeoKill Bill: Vol. 1Jurassic
ParkBlade RunnerSanjuroWhite HeatThe GeneralBajrangi
BhaijaanBabyBãhubali: The BeginningSerbuan maut 2: BerandalGuardians
of the GalaxyBlade Runner 2049The RevenantTropa de Elite 2: O Inimigo
Agora é OutroDeadpoolYip ManNefes: Vatan SagolsunTropa de EliteThe
AvengersThe Bourne UltimatumCasino RoyaleKill Bill: Vol. 2Mou gaan
douPirates of the Caribbean: The Curse of the Black PearlKnockin' on
Heaven's DoorThe TerminatorPer un pugno di dollariRio BravoThor:
RagnarokStar Wars: Episode VII - The Force AwakensX-Men: Days of
Future PastEdge of TomorrowDistrict 9Star TrekLetters from Iwo
JimaIron ManYing xiongThe Bourne IdentityThe Blues BrothersDawn of the
DeadAguirre, der Zorn GottesThe Wild BunchThe Adventures of Robin
HoodThe GentlemenRaaziDunkirkUdta PunjabRogue OneCaptain America:
Civil WarAng-ma-reul bo-at-daAjeossiChugyeokjaTakenAvatarApocalyptoHot
FuzzSerenityGongdong gyeongbi guyeok JSAThe Count of Monte CristoWo hu
cang longThe Boondock SaintsThe GameTombstoneThe FugitiveLat sau san
taamDip huet seung hungPredatorEvil Dead IIGhostbustersBadlandsBonnie
and ClydeThe Longest DayRed RiverKey LargoScarface: The Shame of the
NationDeadpool 2Mission: Impossible - FalloutKingsman: The Secret
ServiceCaptain America: The Winter SoldierStar Trek Into DarknessX:
First ClassSkyfallLucky Number Slevin3:10 to YumaKung fuThe Bourne
SupremacyMan on FireThe Last SamuraiThe Fifth ElementThe Last of the
MohicansEmpire of the SunStar Trek II: The Wrath of KhanFirst BloodThe
Taking of Pelham One Two ThreeEnter the DragonThe French
ConnectionDirty HarryWhere Eagles DareThe Dirty DozenGoldfingerThe
Magnificent SevenGuardians of the Galaxy Vol. 2Baby DriverOnly the
BraveSicarioHell or High WaterDawn of the Planet of the ApesSerbuan
mautEnd of WatchKick-AssCelda 211Sherlock HolmesEastern PromisesHuo
Yuan JiaHarry Potter and the Half-Blood Prince300WatchmenLord of
WarBatoru rowaiaruMinority ReportDie Hard: With a VengeanceFalling
DownLethal WeaponMad Max 2The WarriorsEscape from Alcatraz' to numeric

The above exception was the direct cause of the following exception:

TypeError                                 Traceback (most recent call
last)
Cell In[54], line 1
----> 1 genres.mean()['Runtime']

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:2452, in GroupBy.mean(self, numeric_only, engine,
engine_kwargs)
   2445     return self._numba_agg_general(
   2446         grouped_mean,
```

```
   2447         executor.float_dtype_mapping,
   2448         engine_kwargs,
   2449         min_periods=0,
   2450     )
   2451 else:
-> 2452     result = self._cython_agg_general(
   2453         "mean",
   2454         alt=lambda x: Series(x,
copy=False).mean(numeric_only=numeric_only),
   2455         numeric_only=numeric_only,
   2456     )
   2457     return result.__finalize__(self.obj, method="groupby")

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:1998, in GroupBy._cython_agg_general(self, how, alt,
numeric_only, min_count, **kwargs)
   1995     result = self._agg_py_fallback(how, values,
ndim=data.ndim, alt=alt)
   1996     return result
-> 1998 new_mgr = data.grouped_reduce(array_func)
   1999 res = self._wrap_agged_manager(new_mgr)
   2000 if how in ["idxmin", "idxmax"]:

File ~\anaconda3\Lib\site-packages\pandas\core\internals\
managers.py:1469, in BlockManager.grouped_reduce(self, func)
   1465 if blk.is_object:
   1466     # split on object-dtype blocks bc some columns may raise
   1467     #  while others do not.
   1468     for sb in blk._split():
-> 1469         applied = sb.apply(func)
   1470         result_blocks = extend_blocks(applied, result_blocks)
   1471 else:

File ~\anaconda3\Lib\site-packages\pandas\core\internals\
blocks.py:393, in Block.apply(self, func, **kwargs)
    387 @final
    388 def apply(self, func, **kwargs) -> list[Block]:
    389     """
    390     apply the function to my values; return a block if we are
not
    391     one
    392     """
--> 393     result = func(self.values, **kwargs)
    395     result = maybe_coerce_values(result)
    396     return self._split_op_result(result)

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:1995, in
GroupBy._cython_agg_general.<locals>.array_func(values)
   1992     return result
```

```
   1994 assert alt is not None
-> 1995 result = self._agg_py_fallback(how, values, ndim=data.ndim,
alt=alt)
   1996 return result

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:1946, in GroupBy._agg_py_fallback(self, how, values, ndim,
alt)
   1944     msg = f"agg function failed [how->{how},dtype-
>{ser.dtype}]"
   1945     # preserve the kind of exception that raised
-> 1946     raise type(err)(msg) from err
   1948 if ser.dtype == object:
   1949     res_values = res_values.astype(object, copy=False)

TypeError: agg function failed [how->mean,dtype->object]
```

```
genres.mean(numeric_only=True)  # this is the correct way in new
version qki humare pass string data bhi hai
```

```
              Runtime  IMDB_Rating    No_of_Votes         Gross
Metascore
Genre

Action     129.046512     7.949419  420246.581395  1.897224e+08
73.419580
Adventure  134.111111     7.937500  313557.819444  1.319017e+08
78.437500
Animation   99.585366     7.930488  268032.073171  1.784326e+08
81.093333
Biography  136.022727     7.938636  272805.045455  9.404952e+07
76.240506
Comedy     112.129032     7.901290  178195.658065  1.010572e+08
78.720000
Crime      126.392523     8.016822  313398.271028  7.899656e+07
77.080460
Drama      124.737024     7.957439  212343.612457  1.225259e+08
79.701245
Family     107.500000     7.800000  275610.500000  2.195553e+08
79.000000
Fantasy     85.000000     8.000000   73111.000000  3.913633e+08
NaN
Film-Noir  104.000000     7.966667  122405.000000  4.197018e+07
95.666667
Horror     102.090909     7.909091  340232.363636  9.405902e+07
80.000000
Mystery    119.083333     7.975000  350250.333333  1.047014e+08
79.125000
Thriller   108.000000     7.800000   27733.000000  1.755074e+07
81.000000
```

```
Western    148.250000     8.350000  322416.250000  1.455538e+07
78.250000

genres.std()   # standard deviation -> same reason

----------------------------------------------------------------------
-----
ValueError                                Traceback (most recent call
last)
Cell In[56], line 1
----> 1 genres.std()['']

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:2641, in GroupBy.std(self, ddof, engine, engine_kwargs,
numeric_only)
   2631     return np.sqrt(
   2632         self._numba_agg_general(
   2633             grouped_var,
   (...)
   2638         )
   2639     )
   2640 else:
-> 2641     return self._cython_agg_general(
   2642         "std",
   2643         alt=lambda x: Series(x, copy=False).std(ddof=ddof),
   2644         numeric_only=numeric_only,
   2645         ddof=ddof,
   2646     )

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:1998, in GroupBy._cython_agg_general(self, how, alt,
numeric_only, min_count, **kwargs)
   1995     result = self._agg_py_fallback(how, values,
ndim=data.ndim, alt=alt)
   1996     return result
-> 1998 new_mgr = data.grouped_reduce(array_func)
   1999 res = self._wrap_agged_manager(new_mgr)
   2000 if how in ["idxmin", "idxmax"]:

File ~\anaconda3\Lib\site-packages\pandas\core\internals\
managers.py:1469, in BlockManager.grouped_reduce(self, func)
   1465 if blk.is_object:
   1466     # split on object-dtype blocks bc some columns may raise
   1467     #  while others do not.
   1468     for sb in blk._split():
-> 1469         applied = sb.apply(func)
   1470         result_blocks = extend_blocks(applied, result_blocks)
   1471 else:

File ~\anaconda3\Lib\site-packages\pandas\core\internals\
```

```
blocks.py:393, in Block.apply(self, func, **kwargs)
    387 @final
    388 def apply(self, func, **kwargs) -> list[Block]:
    389     """
    390     apply the function to my values; return a block if we are
not
    391     one
    392     """
--> 393     result = func(self.values, **kwargs)
    395     result = maybe_coerce_values(result)
    396     return self._split_op_result(result)

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:1973, in
GroupBy._cython_agg_general.<locals>.array_func(values)
   1971 def array_func(values: ArrayLike) -> ArrayLike:
   1972     try:
-> 1973         result = self._grouper._cython_operation(
   1974             "aggregate",
   1975             values,
   1976             how,
   1977             axis=data.ndim - 1,
   1978             min_count=min_count,
   1979             **kwargs,
   1980         )
   1981     except NotImplementedError:
   1982         # generally if we have numeric_only=False
   1983         # and non-applicable functions
   1984         # try to python agg
   1985         # TODO: shouldn't min_count matter?
   1986         # TODO: avoid special casing SparseArray here
   1987         if how in ["any", "all"] and isinstance(values,
SparseArray):

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:831, in
BaseGrouper._cython_operation(self, kind, values, how, axis,
min_count, **kwargs)
    829 ids, _, _ = self.group_info
    830 ngroups = self.ngroups
--> 831 return cy_op.cython_operation(
    832     values=values,
    833     axis=axis,
    834     min_count=min_count,
    835     comp_ids=ids,
    836     ngroups=ngroups,
    837     **kwargs,
    838 )

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:550, in
WrappedCythonOp.cython_operation(self, values, axis, min_count,
```

```
comp_ids, ngroups, **kwargs)
    539 if not isinstance(values, np.ndarray):
    540     # i.e. ExtensionArray
    541     return values._groupby_op(
    542         how=self.how,
    543         has_dropped_na=self.has_dropped_na,
   (...)
    547         **kwargs,
    548     )
--> 550 return self._cython_op_ndim_compat(
    551     values,
    552     min_count=min_count,
    553     ngroups=ngroups,
    554     comp_ids=comp_ids,
    555     mask=None,
    556     **kwargs,
    557 )

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:344, in
WrappedCythonOp._cython_op_ndim_compat(self, values, min_count,
ngroups, comp_ids, mask, result_mask, **kwargs)
    341     # otherwise we have OHLC
    342     return res.T
--> 344 return self._call_cython_op(
    345     values,
    346     min_count=min_count,
    347     ngroups=ngroups,
    348     comp_ids=comp_ids,
    349     mask=mask,
    350     result_mask=result_mask,
    351     **kwargs,
    352 )

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:402, in
WrappedCythonOp._call_cython_op(self, values, min_count, ngroups,
comp_ids, mask, result_mask, **kwargs)
    400 out_shape = self._get_output_shape(ngroups, values)
    401 func = self._get_cython_function(self.kind, self.how,
values.dtype, is_numeric)
--> 402 values = self._get_cython_vals(values)
    403 out_dtype = self._get_out_dtype(values.dtype)
    405 result = maybe_fill(np.empty(out_shape, dtype=out_dtype))

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\ops.py:230, in
WrappedCythonOp._get_cython_vals(self, values)
    224 how = self.how
    226 if how in ["median", "std", "sem", "skew"]:
    227     # median only has a float64 implementation
    228     # We should only get here with is_numeric, as non-numeric
cases
```

```
    229     #  should raise in _get_cython_function
--> 230     values = ensure_float64(values)
    232 elif values.dtype.kind in "iu":
    233     if how in ["var", "mean"] or (
    234         self.kind == "transform" and self.has_dropped_na
    235     ):
    236         # has_dropped_na check need for
test_null_group_str_transformer
    237         # result may still include NaN, so we have to cast

File pandas\\_libs\\algos_common_helper.pxi:42, in
pandas._libs.algos.ensure_float64()

ValueError: could not convert string to float: 'The Shawshank
Redemption'
```

```
genres.std(numeric_only=True)    # similar as last
```

```
             Runtime  IMDB_Rating    No_of_Votes         Gross
Metascore
Genre

Action     28.500706     0.304258  432946.814748  2.256724e+08
12.421252
Adventure  33.317320     0.229781  301188.347642  1.697543e+08
12.345393
Animation  14.530471     0.253221  262173.231571  2.091840e+08
8.813646
Biography  25.514466     0.267140  271284.191372  1.363251e+08
11.028187
Comedy     22.946213     0.228771  188653.570564  1.946513e+08
11.829160
Crime      27.689231     0.335477  373999.730656  1.571191e+08
13.099102
Drama      27.740490     0.267229  305554.162841  2.201164e+08
12.744687
Family     10.606602     0.000000  137008.302816  3.048412e+08
16.970563
Fantasy    12.727922     0.141421   22179.111299  7.606861e+07
NaN
Film-Noir   4.000000     0.152753   54649.083277  7.048472e+07
1.527525
Horror     13.604812     0.311302  234883.508691  9.965017e+07
15.362291
Mystery    14.475423     0.310791  404621.915297  1.567524e+08
18.604435
Thriller         NaN          NaN            NaN           NaN
NaN
Western    17.153717     0.420317  263489.554280  1.230626e+07
9.032349
```

```python
# find the top 3 genres by total earning
movies.groupby('Genre').sum()
```

```
                                                  Series_Title  \
Genre
Action     The Dark KnightThe Lord of the Rings: The Retu...
Adventure  InterstellarBack to the FutureInglourious Bast...
Animation  Sen to Chihiro no kamikakushiThe Lion KingHota...
Biography  Schindler's ListGoodfellasHamiltonThe Intoucha...
Comedy     GisaengchungLa vita è bellaModern TimesCity Li...
Crime      The GodfatherThe Godfather: Part II12 Angry Me...
Drama      The Shawshank RedemptionFight ClubForrest Gump...
Family     E.T. the Extra-TerrestrialWilly Wonka & the Ch...
Fantasy                Das Cabinet des Dr. CaligariNosferatu
Film-Noir   The Third ManThe Maltese FalconShadow of a Doubt
Horror     PsychoAlienThe ThingThe ExorcistNight of the L...
Mystery    MementoRear WindowVertigoShutter IslandKahaani...
Thriller                                     Wait Until Dark
Western    Il buono, il brutto, il cattivoOnce Upon a Tim...

                                               Released_Year  Runtime
\
Genre

Action     200820032010200120021999198019771962195420001 9...    22196

Adventure  20141985200919811968196219592013197519631948 19...     9656

Animation  20011994198820162018201720081997199520192009 20...     8166

Biography  19931990202020112002201719951984201820132013 20...    11970

Comedy     20191997193619312009196419402001200019731960 19...    17380

Crime      19721974195719942002199919951991201920061995 19...    13524

Drama      19941999199419752020199819462014200619981988 19...    36049

Family                                              19821971      215

Fantasy                                             19201922      170

Film-Noir                                       194919411943      312

Horror          1960197919821973196819612017197819332004200 1     1123

Mystery      20001954195820102012199519721938198820121998 1997     1429

Thriller                                                1967      108

Western                                     1966196819651976      593
```

```
           IMDB_Rating
Director  \
Genre

Action          1367.3  Christopher NolanPeter JacksonChristopher
Nola...
Adventure        571.5  Christopher NolanRobert ZemeckisQuentin
Tarant...
Animation        650.3  Hayao MiyazakiRoger AllersIsao
TakahataMakoto ...
Biography        698.6  Steven SpielbergMartin ScorseseThomas
KailOliv...
Comedy          1224.7  Bong Joon HoRoberto BenigniCharles
ChaplinChar...
Crime            857.8  Francis Ford CoppolaFrancis Ford
CoppolaSidney...
Drama           2299.7  Frank DarabontDavid FincherRobert
ZemeckisMilo...
Family            15.6                        Steven SpielbergMel
Stuart
Fantasy           16.0                           Robert WieneF.W.
Murnau
Film-Noir         23.9              Carol ReedJohn HustonAlfred
Hitchcock
Horror            87.0  Alfred HitchcockRidley ScottJohn
CarpenterWill...
Mystery           95.7  Christopher NolanAlfred HitchcockAlfred
Hitchc...
Thriller           7.8                                     Terence
Young
Western           33.4  Sergio LeoneSergio LeoneSergio LeoneClint
East...

                                                       Star1
No_of_Votes  \
Genre

Action     Christian BaleElijah WoodLeonardo DiCaprioElij...
72282412
Adventure  Matthew McConaugheyMichael J. FoxBrad PittJürg...
22576163
Animation  Daveigh ChaseRob MinkoffTsutomu TatsumiRyûnosu...
21978630
Biography  Liam NeesonRobert De NiroLin-Manuel MirandaÉri...
24006844
Comedy     Kang-ho SongRoberto BenigniCharles ChaplinChar...
27620327
Crime      Marlon BrandoAl PacinoHenry FondaJohn Travolta...
```

```
33533615
Drama      Tim RobbinsBrad PittTom HanksJack NicholsonSur...
61367304
Family                                Henry ThomasGene Wilder
551221
Fantasy                              Werner KraussMax Schreck
146222
Film-Noir          Orson WellesHumphrey BogartTeresa Wright
367215
Horror     Anthony PerkinsSigourney WeaverKurt RussellEll...
3742556
Mystery    Guy PearceJames StewartJames StewartLeonardo D...
4203004
Thriller                                       Audrey Hepburn
27733
Western    Clint EastwoodHenry FondaClint EastwoodClint E...
1289665

                  Gross  Metascore
Genre
Action     3.263226e+10    10499.0
Adventure  9.496922e+09     5020.0
Animation  1.463147e+10     6082.0
Biography  8.276358e+09     6023.0
Comedy     1.566387e+10     9840.0
Crime      8.452632e+09     6706.0
Drama      3.540997e+10    19208.0
Family     4.391106e+08      158.0
Fantasy    7.827267e+08        0.0
Film-Noir  1.259105e+08      287.0
Horror     1.034649e+09      880.0
Mystery    1.256417e+09      633.0
Thriller   1.755074e+07       81.0
Western    5.822151e+07      313.0
```

```python
# iss dataframe me sare details aa gya but mujhe bass need earning bale se hai
movies.groupby('Genre').sum()['Gross']
```

```
Genre
Action       3.263226e+10
Adventure    9.496922e+09
Animation    1.463147e+10
Biography    8.276358e+09
Comedy       1.566387e+10
Crime        8.452632e+09
Drama        3.540997e+10
Family       4.391106e+08
Fantasy      7.827267e+08
Film-Noir    1.259105e+08
```

```
Horror          1.034649e+09
Mystery         1.256417e+09
Thriller        1.755074e+07
Western         5.822151e+07
Name: Gross, dtype: float64
```

```
# ye abhi index ke basis pe sort hai to isse value ke basis pe sort
kar dete hain taki top earning movies dikh jaye
movies.groupby('Genre').sum()['Gross'].sort_values(ascending=False)
```

```
Genre
Drama           3.540997e+10
Action          3.263226e+10
Comedy          1.566387e+10
Animation       1.463147e+10
Adventure       9.496922e+09
Crime           8.452632e+09
Biography       8.276358e+09
Mystery         1.256417e+09
Horror          1.034649e+09
Fantasy         7.827267e+08
Family          4.391106e+08
Film-Noir       1.259105e+08
Western         5.822151e+07
Thriller        1.755074e+07
Name: Gross, dtype: float64
```

```
movies.groupby('Genre').sum()
['Gross'].sort_values(ascending=False).head(3)  # top 3 movie jo ki
maximum revenue generate karta hai box office me
```

```
Genre
Drama      3.540997e+10
Action     3.263226e+10
Comedy     1.566387e+10
Name: Gross, dtype: float64
```

```
# m2
movies.groupby('Genre')
['Gross'].sum().sort_values(ascending=False).head(3) # yaha pahle
column nikale phie sum apply kar dye last method me just opposite
  # kye the 2nd bala method jyada best hai qki time kam lagega hum
kisi specific column pe hi sum apply kar rahe hai but 1st method me
sare column pe sum
#apply karne ke baad required column ko select kar rahe the
```

```
Genre
Drama      3.540997e+10
Action     3.263226e+10
Comedy     1.566387e+10
Name: Gross, dtype: float64
```

```python
# find the genre with heighest avg IMDB rating
movies.groupby('Genre')['IMDB_Rating'].mean()
```

```
Genre
Action       7.949419
Adventure    7.937500
Animation    7.930488
Biography    7.938636
Comedy       7.901290
Crime        8.016822
Drama        7.957439
Family       7.800000
Fantasy      8.000000
Film-Noir    7.966667
Horror       7.909091
Mystery      7.975000
Thriller     7.800000
Western      8.350000
Name: IMDB_Rating, dtype: float64
```

```python
movies.groupby('Genre')
['IMDB_Rating'].mean().sort_values(ascending=False).head(1)
```

```
Genre
Western    8.35
Name: IMDB_Rating, dtype: float64
```

```python
#find director with most popularity
movies.groupby('Director')   # director ke basis pe group kar lete
hain and then no_of_votes jisko jyada hoga wo popular hoga
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x00000290C55C7A40>
```

```python
movies.groupby('Director')["No_of_Votes"].sum()
```

```
Director
Aamir Khan               168895
Aaron Sorkin              89896
Abdellatif Kechiche      138741
Abhishek Chaubey          27175
Abhishek Kapoor           32628
                          ...
Zack Snyder             1233675
Zaza Urushadze            40382
Zoya Akhtar               99813
Çagan Irmak               78925
Ömer Faruk Sorak          56960
Name: No_of_Votes, Length: 548, dtype: int64
```

```
movies.groupby('Director')
["No_of_Votes"].sum().sort_values(ascending=False).head(1)
```

```
Director
Christopher Nolan    11578345
Name: No_of_Votes, dtype: int64
```

```
#movies.groupby('Genre')['IMDB_Rating'].max()
```

```
Genre
Action       9.0
Adventure    8.6
Animation    8.6
Biography    8.9
Comedy       8.6
Crime        9.2
Drama        9.3
Family       7.8
Fantasy      8.1
Film-Noir    8.1
Horror       8.5
Mystery      8.4
Thriller     7.8
Western      8.8
Name: IMDB_Rating, dtype: float64
```

```
# find number of movies done by each actor
movies['Star1'].value_counts()
```

```
Star1
Tom Hanks           12
Robert De Niro      11
Al Pacino           10
Clint Eastwood      10
Humphrey Bogart      9
                    ..
Preity Zinta         1
Javier Bardem        1
Ki-duk Kim           1
Vladimir Garin       1
Robert Donat         1
Name: count, Length: 660, dtype: int64
```

```
# but mujhe groupby ke help se karna hai
movies.groupby('Star1')['Series_Title'].count()   # koi bhi column ko
pakad sakte the bass no_of_row hi to count karana tha group karane ke
baad
```

```
Star1
Aamir Khan              7
Aaron Taylor-Johnson    1
```

```
Abhay Deol                 1
Abraham Attah              1
Adam Driver                1
                          ..
Zbigniew Zamachowski       1
Zooey Deschanel            1
Çetin Tekindor             1
Éric Toledano              1
Ömer Faruk Sorak           1
Name: Series_Title, Length: 660, dtype: int64
```

```
# last bale tarike se jo answer aaya tha wo sort tha on the bassis of
index and groupby bale method se jo aaya hai wo sorted hai on the
basis of index

movies.groupby('Star1')
['Series_Title'].count().sort_values(ascending=False)
```

```
Star1
Tom Hanks               12
Robert De Niro          11
Clint Eastwood          10
Al Pacino               10
Leonardo DiCaprio        9
                        ..
Glen Hansard             1
Giuseppe Battiston       1
Giulietta Masina         1
Gerardo Taracena         1
Ömer Faruk Sorak         1
Name: Series_Title, Length: 660, dtype: int64
```

```
#har question me  kuch steps follw kar rahe hain
# step-1 : pahle groupby kar rahe hain kisi specific column pe jaise
ki required hai
# step-2 : and then required column ko nikal rahe hain
# step-3 : phir koi nn koi aggrigate function(min,max,sum  etc.) use
kar rahe hain

# GroupBy Attributes and Methods
# find total number of groups -> len
# find items in each group -> size
# first()/last() -> nth item
# get_group -> vs filtering
# groups
# describe
# sample
# nunique

# find total number of groups -> len
movies.groupby('Genre')
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x00000207A42C7890>
```

```
len(movies.groupby('Genre'))
```

```
14
```

```
#m-2
movies['Genre'].nunique()
```

```
14
```

```
# find items in each group -> size
movies.groupby('Genre').size()    # tells ki har group me kitne row
hai   (index ke basis pe sort hoga)
```

```
Genre
Action        172
Adventure      72
Animation      82
Biography      88
Comedy        155
Crime         107
Drama         289
Family          2
Fantasy         2
Film-Noir       3
Horror         11
Mystery        12
Thriller        1
Western         4
dtype: int64
```

```
#m2
movies['Genre'].value_counts()   #(value ke basis pe sort hoga)
```

```
Genre
Drama         289
Action        172
Comedy        155
Crime         107
Biography      88
Animation      82
Adventure      72
Mystery        12
Horror         11
Western         4
Film-Noir       3
Fantasy         2
Family          2
```

```
Thriller       1
Name: count, dtype: int64
```

```python
# first()/last() -> nth item
genres =movies.groupby('Genre')
genres
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x0000023A99EA51F0>
```

```python
genres.first()   # har genres ki first movies aa jayegi
```

```
                              Series_Title Released_Year  Runtime  \
Genre
Action                     The Dark Knight          2008      152
Adventure                     Interstellar          2014      169
Animation    Sen to Chihiro no kamikakushi          2001      125
Biography                 Schindler's List          1993      195
Comedy                         Gisaengchung          2019      132
Crime                        The Godfather          1972      175
Drama             The Shawshank Redemption          1994      142
Family          E.T. the Extra-Terrestrial          1982      115
Fantasy          Das Cabinet des Dr. Caligari          1920       76
Film-Noir                    The Third Man          1949      104
Horror                              Psycho          1960      109
Mystery                            Memento          2000      113
Thriller                    Wait Until Dark          1967      108
Western     Il buono, il brutto, il cattivo          1966      161

           IMDB_Rating              Director                 Star1  \
Genre
Action             9.0     Christopher Nolan        Christian Bale
Adventure          8.6     Christopher Nolan  Matthew McConaughey
Animation          8.6        Hayao Miyazaki         Daveigh Chase
Biography          8.9      Steven Spielberg           Liam Neeson
Comedy             8.6          Bong Joon Ho         Kang-ho Song
Crime              9.2  Francis Ford Coppola         Marlon Brando
Drama              9.3        Frank Darabont           Tim Robbins
Family             7.8      Steven Spielberg         Henry Thomas
Fantasy            8.1          Robert Wiene         Werner Krauss
Film-Noir          8.1            Carol Reed         Orson Welles
Horror             8.5      Alfred Hitchcock       Anthony Perkins
Mystery            8.4     Christopher Nolan            Guy Pearce
Thriller           7.8         Terence Young       Audrey Hepburn
Western            8.8          Sergio Leone        Clint Eastwood

           No_of_Votes        Gross  Metascore
Genre
Action         2303232  534858444.0       84.0
```

```
Adventure      1512360  188020017.0       74.0
Animation       651376   10055859.0       96.0
Biography      1213505   96898818.0       94.0
Comedy          552778   53367844.0       96.0
Crime          1620367  134966411.0      100.0
Drama          2343110   28341469.0       80.0
Family          372490  435110554.0       91.0
Fantasy          57428  337574718.0        NaN
Film-Noir       158731     449191.0       97.0
Horror          604211   32000000.0       97.0
Mystery        1125712   25544867.0       80.0
Thriller         27733   17550741.0       81.0
Western         688390    6100000.0       90.0
```

```
genres.last()    # har group ka last movies aa jayega
```

```
                                  Series_Title Released_Year  Runtime
\
Genre

Action                    Escape from Alcatraz          1979      112

Adventure                       Kelly's Heroes          1970      144

Animation                      The Jungle Book          1967       78

Biography                     Midnight Express          1978      121

Comedy                  Breakfast at Tiffany's          1961      115

Crime                            The 39 Steps          1935       86

Drama                                Lifeboat          1944       97

Family     Willy Wonka & the Chocolate Factory          1971      100

Fantasy                             Nosferatu          1922       94

Film-Noir                   Shadow of a Doubt          1943      108

Horror                             The Others          2001      101

Mystery                          Lost Highway          1997      134

Thriller                      Wait Until Dark          1967      108

Western                 The Outlaw Josey Wales          1976      135


           IMDB_Rating             Director              Star1
No_of_Votes  \
```

```
Genre

Action                 7.6           Don Siegel     Clint Eastwood
121731
Adventure              7.6       Brian G. Hutton     Clint Eastwood
45338
Animation              7.6  Wolfgang Reitherman        Phil Harris
166409
Biography              7.6          Alan Parker         Brad Davis
73662
Comedy                 7.6        Blake Edwards     Audrey Hepburn
166544
Crime                  7.6     Alfred Hitchcock       Robert Donat
51853
Drama                  7.6     Alfred Hitchcock  Tallulah Bankhead
26471
Family                 7.8           Mel Stuart        Gene Wilder
178731
Fantasy                7.9          F.W. Murnau        Max Schreck
88794
Film-Noir              7.8     Alfred Hitchcock      Teresa Wright
59556
Horror                 7.6   Alejandro Amenábar      Nicole Kidman
337651
Mystery                7.6          David Lynch       Bill Pullman
131101
Thriller               7.8        Terence Young     Audrey Hepburn
27733
Western                7.8       Clint Eastwood     Clint Eastwood
65659

                 Gross  Metascore
Genre
Action      43000000.0       76.0
Adventure    1378435.0       50.0
Animation  141843612.0       65.0
Biography   35000000.0       59.0
Comedy     679874270.0       76.0
Crime      302787539.0       93.0
Drama      852142728.0       78.0
Family       4000000.0       67.0
Fantasy    445151978.0        NaN
Film-Noir  123353292.0       94.0
Horror      96522687.0       74.0
Mystery      3796699.0       52.0
Thriller    17550741.0       81.0
Western     31800000.0       69.0

genres.nth(6)    # har group ka 7th movies (qki indexing 0 se) nikal
ke de dega  aap observ karenge ki sare group available nahi hai qki jo
```

```
group
            #available nahi hai usme 7 movie hai hi nahi

                                       Series_Title Released_Year
Runtime  \
16   Star Wars: Episode V - The Empire Strikes Back          1980
124
27                                            Se7en          1995
127
32                            It's a Wonderful Life          1946
130
66                                           WALL·E          2008
98
83                              The Great Dictator           1940
125
102                                      Braveheart          1995
178
118                              North by Northwest          1959
136
420                                          Sleuth          1972
138
724                                         Get Out          2017
104

         Genre  IMDB_Rating              Director             Star1  \
16      Action          8.7        Irvin Kershner       Mark Hamill
27       Crime          8.6         David Fincher    Morgan Freeman
32       Drama          8.6           Frank Capra     James Stewart
66   Animation          8.4        Andrew Stanton         Ben Burtt
83      Comedy          8.4       Charles Chaplin   Charles Chaplin
102  Biography          8.3            Mel Gibson        Mel Gibson
118  Adventure          8.3      Alfred Hitchcock        Cary Grant
420    Mystery          8.0  Joseph L. Mankiewicz  Laurence Olivier
724     Horror          7.7           Jordan Peele    Daniel Kaluuya

     No_of_Votes        Gross  Metascore
16       1159315  290475067.0       82.0
27       1445096  100125643.0       65.0
32        405801   82385199.0       89.0
66        999790  223808164.0       95.0
83        203150     288475.0        NaN
102       959181   75600000.0       68.0
118       299198   13275000.0       98.0
420        44748    4081254.0        NaN
724       492851  176040665.0       85.0

# get_group -> vs filtering   => iske help se kisi particular group ko
dekh sakte hain
genres.get_group('Horror')
```

```
                  Series_Title Released_Year  Runtime   Genre
IMDB_Rating  \
49                      Psycho          1960      109  Horror
8.5
75                       Alien          1979      117  Horror
8.4
271                  The Thing          1982      109  Horror
8.1
419               The Exorcist          1973      122  Horror
8.0
544  Night of the Living Dead          1968       96  Horror
7.9
707             The Innocents          1961      100  Horror
7.8
724                    Get Out          2017      104  Horror
7.7
844                  Halloween          1978       91  Horror
7.7
876          The Invisible Man          1933       71  Horror
7.7
932                        Saw          2004      103  Horror
7.6
948                 The Others          2001      101  Horror
7.6

                Director             Star1  No_of_Votes        Gross
Metascore
49     Alfred Hitchcock   Anthony Perkins       604211   32000000.0
97.0
75         Ridley Scott  Sigourney Weaver       787806   78900000.0
89.0
271      John Carpenter      Kurt Russell       371271   13782838.0
57.0
419    William Friedkin     Ellen Burstyn       362393  232906145.0
81.0
544    George A. Romero       Duane Jones       116557      89029.0
89.0
707        Jack Clayton      Deborah Kerr        27007    2616000.0
88.0
724        Jordan Peele    Daniel Kaluuya       492851  176040665.0
85.0
844      John Carpenter  Donald Pleasence       233106   47000000.0
87.0
876         James Whale      Claude Rains        30683  298791505.0
87.0
932           James Wan        Cary Elwes       379020   56000369.0
46.0
948  Alejandro Amenábar     Nicole Kidman       337651   96522687.0
74.0
```

```
genres.get_group('Fantasy')
```

```
                      Series_Title Released_Year  Runtime    Genre  \
321  Das Cabinet des Dr. Caligari          1920       76  Fantasy
568                     Nosferatu          1922       94  Fantasy

     IMDB_Rating      Director          Star1  No_of_Votes
Gross  \
321          8.1  Robert Wiene  Werner Krauss        57428
337574718.0
568          7.9   F.W. Murnau    Max Schreck        88794
445151978.0

     Metascore
321        NaN
568        NaN
```

```
#m2
movies[movies['Genre']=='Fantasy']
```

```
                      Series_Title Released_Year  Runtime    Genre  \
321  Das Cabinet des Dr. Caligari          1920       76  Fantasy
568                     Nosferatu          1922       94  Fantasy

     IMDB_Rating      Director          Star1  No_of_Votes
Gross  \
321          8.1  Robert Wiene  Werner Krauss        57428
337574718.0
568          7.9   F.W. Murnau    Max Schreck        88794
445151978.0

     Metascore
321        NaN
568        NaN
```

```
# groups
genres.groups    # ye return karega index(genres) name ke sath movies
ke sare index  jaise Action categories me kon kon se index ke movies
hain
```

```
{'Action': [2, 5, 8, 10, 13, 14, 16, 29, 30, 31, 39, 42, 44, 55, 57,
59, 60, 63, 68, 72, 106, 109, 129, 130, 134, 140, 142, 144, 152, 155,
160, 161, 166, 168, 171, 172, 177, 181, 194, 201, 202, 216, 217, 223,
224, 236, 241, 262, 275, 294, 308, 320, 325, 326, 331, 337, 339, 340,
343, 345, 348, 351, 353, 356, 357, 362, 368, 369, 375, 376, 390, 410,
431, 436, 473, 477, 479, 482, 488, 493, 496, 502, 507, 511, 532, 535,
540, 543, 564, 569, 570, 573, 577, 582, 583, 602, 605, 608, 615,
623, ...], 'Adventure': [21, 47, 93, 110, 114, 116, 118, 137, 178,
179, 191, 193, 209, 226, 231, 247, 267, 273, 281, 300, 301, 304, 306,
323, 329, 361, 366, 377, 402, 406, 415, 426, 458, 470, 497, 498, 506,
513, 514, 537, 549, 552, 553, 566, 576, 604, 609, 618, 638, 647, 675,
```

```
681, 686, 692, 711, 713, 739, 755, 781, 797, 798, 851, 873, 884, 912,
919, 947, 957, 964, 966, 984, 991], 'Animation': [23, 43, 46, 56, 58,
61, 66, 70, 101, 135, 146, 151, 158, 170, 197, 205, 211, 213, 219,
229, 230, 242, 245, 246, 270, 330, 332, 358, 367, 378, 386, 389, 394,
395, 399, 401, 405, 409, 469, 499, 510, 516, 518, 522, 578, 586, 592,
595, 596, 599, 633, 640, 643, 651, 665, 672, 694, 728, 740, 741, 744,
756, 758, 761, 771, 783, 796, 799, 822, 828, 843, 875, 891, 892, 902,
906, 920, 956, 971, 976, 986, 992], 'Biography': [7, 15, 18, 35, 38,
54, 102, 107, 131, 139, 147, 157, 159, 173, 176, 212, 215, 218, 228,
235, 243, 263, 276, 282, 290, 298, 317, 328, 338, 342, 346, 359, 360,
365, 372, 373, 385, 411, 416, 418, 424, 429, 484, 525, 536, 542, 545,
575, 579, 587, 600, 606, 614, 622, 632, 635, 644, 649, 650, 657, 671,
673, 684, 729, 748, 753, 757, 759, 766, 770, 779, 809, 810, 815, 820,
831, 849, 858, 877, 882, 897, 910, 915, 923, 940, 949, 952, 987],
'Comedy': [19, 26, 51, 52, 64, 78, 83, 95, 96, 112, 117, 120, 127,
128, 132, 153, 169, 183, 192, 204, 207, 208, 214, 221, 233, 238, 240,
250, 251, 252, 256, 261, 266, 277, 284, 311, 313, 316, 318, 322, 327,
374, 379, 381, 392, 396, 403, 413, 414, 417, 427, 435, 445, 446, 449,
455, 459, 460, 463, 464, 466, 471, 472, 475, 481, 490, 494, 500, 503,
509, 526, 528, 530, 531, 533, 538, 539, 541, 547, 557, 558, 562, 563,
565, 574, 591, 593, 594, 598, 613, 626, 630, 660, 662, 667, 679, 680,
683, 687, 701, ...], 'Crime': [1, 3, 4, 6, 22, 25, 27, 28, 33, 37, 41,
71, 77, 79, 86, 87, 103, 108, 111, 113, 123, 125, 133, 136, 162, 163,
164, 165, 180, 186, 187, 189, 198, 222, 232, 239, 255, 257, 287, 288,
299, 305, 335, 363, 364, 380, 384, 397, 437, 438, 441, 442, 444, 450,
451, 465, 474, 480, 485, 487, 505, 512, 519, 520, 523, 527, 546, 556,
560, 584, 597, 603, 607, 611, 621, 639, 653, 664, 669, 676, 695, 708,
723, 762, 763, 767, 775, 791, 795, 802, 811, 823, 827, 833, 885, 895,
921, 922, 926, 938, ...], 'Drama': [0, 9, 11, 17, 20, 24, 32, 34, 36,
40, 45, 50, 53, 62, 65, 67, 73, 74, 76, 80, 82, 84, 85, 88, 89, 90,
91, 92, 94, 97, 98, 99, 100, 104, 105, 121, 122, 124, 126, 138, 141,
143, 148, 149, 150, 154, 156, 167, 174, 175, 182, 184, 185, 188, 190,
195, 196, 199, 200, 203, 206, 210, 225, 227, 234, 237, 244, 248, 249,
253, 254, 258, 259, 260, 264, 265, 268, 269, 272, 274, 278, 279, 280,
283, 285, 286, 289, 291, 292, 293, 295, 296, 297, 302, 303, 307, 310,
312, 314, 315, ...], 'Family': [688, 698], 'Fantasy': [321, 568],
'Film-Noir': [309, 456, 712], 'Horror': [49, 75, 271, 419, 544, 707,
724, 844, 876, 932, 948], 'Mystery': [69, 81, 119, 145, 220, 393, 420,
714, 829, 899, 959, 961], 'Thriller': [700], 'Western': [12, 48, 115,
691]}

# describe
genres.describe()

          Runtime
\
            count        mean        std    min     25%    50%     75%
max
Genre
```

```
Action       172.0  129.046512  28.500706   45.0  110.75  127.5  143.25
321.0
Adventure     72.0  134.111111  33.317320   88.0  109.00  127.0  149.00
228.0
Animation     82.0   99.585366  14.530471   71.0   90.00   99.5  106.75
137.0
Biography     88.0  136.022727  25.514466   93.0  120.00  129.0  146.25
209.0
Comedy       155.0  112.129032  22.946213   68.0   96.00  106.0  124.50
188.0
Crime        107.0  126.392523  27.689231   80.0  106.50  122.0  141.50
229.0
Drama        289.0  124.737024  27.740490   64.0  105.00  121.0  137.00
242.0
Family         2.0  107.500000  10.606602  100.0  103.75  107.5  111.25
115.0
Fantasy        2.0   85.000000  12.727922   76.0   80.50   85.0   89.50
94.0
Film-Noir      3.0  104.000000   4.000000  100.0  102.00  104.0  106.00
108.0
Horror        11.0  102.090909  13.604812   71.0   98.00  103.0  109.00
122.0
Mystery       12.0  119.083333  14.475423   96.0  110.75  117.5  130.25
138.0
Thriller       1.0  108.000000        NaN  108.0  108.00  108.0  108.00
108.0
Western        4.0  148.250000  17.153717  132.0  134.25  148.0  162.00
165.0

          IMDB_Rating            ...         Gross
Metascore  \
                count      mean  ...           75%          max
count
Genre                            ...

Action          172.0  7.949419  ...  2.674437e+08  936662225.0
143.0
Adventure        72.0  7.937500  ...  1.998070e+08  874211619.0
64.0
Animation        82.0  7.930488  ...  2.520612e+08  873839108.0
75.0
Biography        88.0  7.938636  ...  9.829924e+07  753585104.0
79.0
Comedy          155.0  7.901290  ...  8.107809e+07  886752933.0
125.0
Crime           107.0  8.016822  ...  7.102163e+07  790482117.0
87.0
Drama           289.0  7.957439  ...  1.164461e+08  924558264.0
241.0
```

```
Family            2.0  7.800000  ...  3.273329e+08  435110554.0
2.0
Fantasy           2.0  8.000000  ...  4.182577e+08  445151978.0
0.0
Film-Noir         3.0  7.966667  ...  6.273068e+07  123353292.0
3.0
Horror           11.0  7.909091  ...  1.362817e+08  298791505.0
11.0
Mystery          12.0  7.975000  ...  1.310949e+08  474203697.0
8.0
Thriller          1.0  7.800000  ...  1.755074e+07   17550741.0
1.0
Western           4.0  8.350000  ...  1.920000e+07   31800000.0
4.0


                mean        std   min    25%   50%    75%    max
Genre
Action     73.419580  12.421252  33.0  65.00  74.0  82.00   98.0
Adventure  78.437500  12.345393  41.0  69.75  80.5  87.25  100.0
Animation  81.093333   8.813646  61.0  75.00  82.0  87.50   96.0
Biography  76.240506  11.028187  48.0  70.50  76.0  84.50   97.0
Comedy     78.720000  11.829160  45.0  72.00  79.0  88.00   99.0
Crime      77.080460  13.099102  47.0  69.50  77.0  87.00  100.0
Drama      79.701245  12.744687  28.0  72.00  82.0  89.00  100.0
Family     79.000000  16.970563  67.0  73.00  79.0  85.00   91.0
Fantasy          NaN        NaN   NaN    NaN   NaN    NaN    NaN
Film-Noir  95.666667   1.527525  94.0  95.00  96.0  96.50   97.0
Horror     80.000000  15.362291  46.0  77.50  87.0  88.50   97.0
Mystery    79.125000  18.604435  52.0  65.25  77.0  98.50  100.0
Thriller   81.000000        NaN  81.0  81.00  81.0  81.00   81.0
Western    78.250000   9.032349  69.0  72.75  77.0  82.50   90.0

[14 rows x 40 columns]
```

```
# sample
genres.sample()    # har geners ka koi ek random movies dikha dega
```

```
                           Series_Title Released_Year  Runtime
Genre  \
702                    Bonnie and Clyde          1967      111
Action
458                   The Wizard of Oz          1939      102
Adventure
358                          Persepolis          2007       96
Animation
622                   American Gangster          2007      157
Biography
481                          The Artist          2011      100
Comedy
```

```
222                              Prisoners          2013      153
Crime
988  Close Encounters of the Third Kind          1977      138
Drama
688          E.T. the Extra-Terrestrial          1982      115
Family
568                              Nosferatu          1922       94
Fantasy
456                     The Maltese Falcon          1941      100  Film-
Noir
844                              Halloween          1978       91
Horror
69                                 Memento          2000      113
Mystery
700                        Wait Until Dark          1967      108
Thriller
12        Il buono, il brutto, il cattivo          1966      161
Western

     IMDB_Rating             Director              Star1  No_of_Votes
\
702          7.8          Arthur Penn      Warren Beatty       102415

458          8.0       Victor Fleming       George Cukor       371379

358          8.0    Vincent Paronnaud    Marjane Satrapi        88656

622          7.8         Ridley Scott  Denzel Washington       392449

481          7.9  Michel Hazanavicius      Jean Dujardin       230624

222          8.1      Denis Villeneuve       Hugh Jackman       601149

988          7.6     Steven Spielberg   Richard Dreyfuss       184966

688          7.8     Steven Spielberg       Henry Thomas       372490

568          7.9           F.W. Murnau        Max Schreck        88794

456          8.0           John Huston    Humphrey Bogart       148928

844          7.7       John Carpenter   Donald Pleasence       233106

69           8.4    Christopher Nolan         Guy Pearce      1125712

700          7.8        Terence Young     Audrey Hepburn        27733

12           8.8         Sergio Leone     Clint Eastwood       688390


           Gross  Metascore
```

```
702  909838190.0       86.0
458    2076020.0       92.0
358    4445756.0       90.0
622  130164645.0       76.0
481   44671682.0       89.0
222   61002302.0       70.0
988  132088635.0       90.0
688  435110554.0       91.0
568  445151978.0        NaN
456    2108060.0       96.0
844   47000000.0       87.0
69    25544867.0       80.0
700   17550741.0       81.0
12     6100000.0       90.0
```

```
genres.sample(2)  # abb do do random movies dikhega abhi error islye
aa raha hai qki
```

```
----------------------------------------------------------------------
-----
ValueError                                Traceback (most recent call
last)
Cell In[48], line 1
----> 1 genres.sample(2)

File ~\anaconda3\Lib\site-packages\pandas\core\groupby\
groupby.py:5780, in GroupBy.sample(self, n, frac, replace, weights,
random_state)
   5777         assert frac is not None
   5778         sample_size = round(frac * group_size)
-> 5780     grp_sample = sample.sample(
   5781         group_size,
   5782         size=sample_size,
   5783         replace=replace,
   5784         weights=None if weights is None else
weights_arr[grp_indices],
   5785         random_state=random_state,
   5786     )
   5787     sampled_indices.append(grp_indices[grp_sample])
   5789 sampled_indices = np.concatenate(sampled_indices)

File ~\anaconda3\Lib\site-packages\pandas\core\sample.py:152, in
sample(obj_len, size, replace, weights, random_state)
    149     else:
    150         raise ValueError("Invalid weights: weights sum to
zero")
--> 152 return random_state.choice(obj_len, size=size,
replace=replace, p=weights).astype(
    153     np.intp, copy=False
    154 )
```

```
File numpy\\random\\mtrand.pyx:1001, in
numpy.random.mtrand.RandomState.choice()

ValueError: Cannot take a larger sample than population when
'replace=False'

genres.sample(2,replace=True)  # abb jisme ek hi row hai wo repeate ho
jayega

                           Series_Title Released_Year  Runtime
Genre  \
774                       3:10 to Yuma          2007      122
Action
144                            Warrior          2011      140
Action
884           The Peanut Butter Falcon          2019       97
Adventure
329                         The Martian          2015      144
Adventure
43                        The Lion King          1994       88
Animation
740                     Wreck-It Ralph          2012      101
Animation
657                            Ed Wood          1994      127
Biography
373                     Cinderella Man          2005      144
Biography
463                         Knives Out          2019      130
Comedy
835           The Purple Rose of Cairo          1985       82
Comedy
669                    Boyz n the Hood          1991      112
Crime
28            The Silence of the Lambs          1991      118
Crime
443                           La Strada          1954      108
Drama
73                          The Shining          1980      146
Drama
698  Willy Wonka & the Chocolate Factory          1971      100
Family
688         E.T. the Extra-Terrestrial          1982      115
Family
568                           Nosferatu          1922       94
Fantasy
568                           Nosferatu          1922       94
Fantasy
309                       The Third Man          1949      104  Film-
Noir
```

```
456                     The Maltese Falcon          1941       100  Film-
Noir
844                              Halloween          1978        91
Horror
544               Night of the Living Dead          1968        96
Horror
145                         Shutter Island          2010       138
Mystery
81                             Rear Window          1954       112
Mystery
700                        Wait Until Dark          1967       108
Thriller
700                        Wait Until Dark          1967       108
Thriller
115             Per qualche dollaro in più          1965       132
Western
691                  The Outlaw Josey Wales          1976       135
Western

     IMDB_Rating           Director              Star1  No_of_Votes  \
774          7.7      James Mangold      Russell Crowe       288797
144          8.2     Gavin O'Connor          Tom Hardy       435950
884          7.6       Tyler Nilson   Michael Schwartz        66346
329          8.0       Ridley Scott         Matt Damon       760094
43           8.5       Roger Allers        Rob Minkoff       942045
740          7.7        Rich Moore     John C. Reilly       380195
657          7.8         Tim Burton        Johnny Depp       164937
373          8.0         Ron Howard      Russell Crowe       176151
463          7.9       Rian Johnson       Daniel Craig       454203
835          7.7        Woody Allen         Mia Farrow        47102
669          7.8     John Singleton   Cuba Gooding Jr.       126082
28           8.6     Jonathan Demme       Jodie Foster      1270197
443          8.0   Federico Fellini      Anthony Quinn        58314
73           8.4    Stanley Kubrick     Jack Nicholson       898237
698          7.8         Mel Stuart        Gene Wilder       178731
688          7.8   Steven Spielberg       Henry Thomas       372490
568          7.9         F.W. Murnau        Max Schreck        88794
568          7.9         F.W. Murnau        Max Schreck        88794
309          8.1         Carol Reed       Orson Welles       158731
456          8.0        John Huston    Humphrey Bogart       148928
844          7.7     John Carpenter   Donald Pleasence       233106
544          7.9   George A. Romero        Duane Jones       116557
145          8.2    Martin Scorsese  Leonardo DiCaprio      1129894
81           8.4   Alfred Hitchcock      James Stewart       444074
700          7.8      Terence Young     Audrey Hepburn        27733
700          7.8      Terence Young     Audrey Hepburn        27733
115          8.3      Sergio Leone      Clint Eastwood       232772
691          7.8     Clint Eastwood     Clint Eastwood        65659
```

```
           Gross  Metascore
774   53606916.0       76.0
144   13657115.0       71.0
884   13122642.0       70.0
329  228433663.0       80.0
43   422783777.0       88.0
740  189422889.0       72.0
657    5887457.0       70.0
373   61649911.0       69.0
463  165359751.0       82.0
835   10631333.0       75.0
669   57504069.0       76.0
28   130742922.0       85.0
443  101832153.0        NaN
73    44017374.0       66.0
698    4000000.0       67.0
688  435110554.0       91.0
568  445151978.0        NaN
568  445151978.0        NaN
309     449191.0       97.0
456    2108060.0       96.0
844   47000000.0       87.0
544      89029.0       89.0
145  128012934.0       63.0
81    36764313.0      100.0
700   17550741.0       81.0
700   17550741.0       81.0
115   15000000.0       74.0
691   31800000.0       69.0
```

```
 genres.nunique()  # har genres me kitna unique data hai  action bale
movies me released_year me 61 unique value hai matlab ki bahut ssare
movies
                             #aaise honge jo ki same year me release
hua hoga
```

```
           Series_Title  Released_Year  Runtime  IMDB_Rating  Director
Star1  \
Genre

Action              172             61       78           15       123
121
Adventure            72             49       58           10        59
59
Animation            82             35       41           11        51
77
Biography            88             44       56           13        76
72
Comedy              155             72       70           11       113
133
```

```
Crime                106              56        65           14         86
85
Drama                289              83        95           14        211
250
Family                 2               2         2            1          2
2
Fantasy                2               2         2            2          2
2
Film-Noir              3               3         3            3          3
3
Horror                11              11        10            8         10
11
Mystery               12              11        10            8         10
11
Thriller               1               1         1            1          1
1
Western                4               4         4            4          2
2

           No_of_Votes  Gross  Metascore
Genre
Action             172    172         50
Adventure           72     72         33
Animation           82     82         29
Biography           88     88         40
Comedy             155    155         44
Crime              107    107         39
Drama              288    287         52
Family               2      2          2
Fantasy              2      2          0
Film-Noir            3      3          3
Horror              11     11          9
Mystery             12     12          7
Thriller             1      1          1
Western              4      4          4
```

```python
# agg method
# passing dict
genres.sum()
```

```
                                                Series_Title  \
Genre
Action     The Dark KnightThe Lord of the Rings: The Retu...
Adventure  InterstellarBack to the FutureInglourious Bast...
Animation  Sen to Chihiro no kamikakushiThe Lion KingHota...
Biography  Schindler's ListGoodfellasHamiltonThe Intoucha...
Comedy     GisaengchungLa vita è bellaModern TimesCity Li...
Crime      The GodfatherThe Godfather: Part II12 Angry Me...
Drama      The Shawshank RedemptionFight ClubForrest Gump...
Family     E.T. the Extra-TerrestrialWilly Wonka & the Ch...
```

```
Fantasy                 Das Cabinet des Dr. CaligariNosferatu
Film-Noir   The Third ManThe Maltese FalconShadow of a Doubt
Horror     PsychoAlienThe ThingThe ExorcistNight of the L...
Mystery    MementoRear WindowVertigoShutter IslandKahaani...
Thriller                                     Wait Until Dark
Western    Il buono, il brutto, il cattivoOnce Upon a Tim...

                                               Released_Year  Runtime
\
Genre

Action     2008200320102001200219991980197719621954200019...    22196

Adventure  2014198520091981196819621959201319751963194819...     9656

Animation  2001199419882016201820172008199719952019200920...     8166

Biography  1993199020202011200220171995198420182013201320...    11970

Comedy     2019199719361931200919641940200120001973196019...    17380

Crime      1972197419571994200219991995199120192006199519...    13524

Drama      1994199919941975202019981946201420061998198819...    36049

Family                                              19821971      215

Fantasy                                             19201922      170

Film-Noir                                       194919411943      312

Horror         1960197919821973196819612017197819332004200I     1123

Mystery     2000195419582010201219951972193819882012199819997     1429

Thriller                                                1967      108

Western                                     1966196819651976      593


           IMDB_Rating
Director  \
Genre

Action          1367.3  Christopher NolanPeter JacksonChristopher
Nola...
Adventure        571.5  Christopher NolanRobert ZemeckisQuentin
Tarant...
Animation        650.3  Hayao MiyazakiRoger AllersIsao
TakahataMakoto ...
Biography        698.6  Steven SpielbergMartin ScorseseThomas
```

```
KailOliv...
Comedy          1224.7  Bong Joon HoRoberto BenigniCharles
ChaplinChar...
Crime            857.8  Francis Ford CoppolaFrancis Ford
CoppolaSidney...
Drama           2299.7  Frank DarabontDavid FincherRobert
ZemeckisMilo...
Family            15.6                        Steven SpielbergMel
Stuart
Fantasy           16.0                           Robert WieneF.W.
Murnau
Film-Noir         23.9              Carol ReedJohn HustonAlfred
Hitchcock
Horror            87.0  Alfred HitchcockRidley ScottJohn
CarpenterWill...
Mystery           95.7  Christopher NolanAlfred HitchcockAlfred
Hitchc...
Thriller           7.8                                     Terence
Young
Western           33.4  Sergio LeoneSergio LeoneSergio LeoneClint
East...

                                                            Star1
No_of_Votes  \
Genre

Action     Christian BaleElijah WoodLeonardo DiCaprioElij...
72282412
Adventure  Matthew McConaugheyMichael J. FoxBrad PittJürg...
22576163
Animation  Daveigh ChaseRob MinkoffTsutomu TatsumiRyûnosu...
21978630
Biography  Liam NeesonRobert De NiroLin-Manuel MirandaÉri...
24006844
Comedy     Kang-ho SongRoberto BenigniCharles ChaplinChar...
27620327
Crime      Marlon BrandoAl PacinoHenry FondaJohn Travolta...
33533615
Drama      Tim RobbinsBrad PittTom HanksJack NicholsonSur...
61367304
Family                               Henry ThomasGene Wilder
551221
Fantasy                             Werner KraussMax Schreck
146222
Film-Noir          Orson WellesHumphrey BogartTeresa Wright
367215
Horror     Anthony PerkinsSigourney WeaverKurt RussellEll...
3742556
Mystery    Guy PearceJames StewartJames StewartLeonardo D...
```

```
4203004
Thriller                                                    Audrey Hepburn
27733
Western    Clint EastwoodHenry FondaClint EastwoodClint E...
1289665

                  Gross  Metascore
Genre
Action     3.263226e+10    10499.0
Adventure  9.496922e+09     5020.0
Animation  1.463147e+10     6082.0
Biography  8.276358e+09     6023.0
Comedy     1.566387e+10     9840.0
Crime      8.452632e+09     6706.0
Drama      3.540997e+10    19208.0
Family     4.391106e+08      158.0
Fantasy    7.827267e+08        0.0
Film-Noir  1.259105e+08      287.0
Horror     1.034649e+09      880.0
Mystery    1.256417e+09      633.0
Thriller   1.755074e+07       81.0
Western    5.822151e+07      313.0
```

```
# agg method
# passing dict
# yadi aap particular column ke upar koi particular aggrigate function
apply karna chahate ho to usse ek dictionary me pass kar do
genres.agg(
    {
        'Runtime':'mean',
        'IMDB_Rating':'mean',
        'No_of_Votes':'sum',
        'Gross':'sum',
        'Metascore':'min'
    }
)
```

```
              Runtime  IMDB_Rating  No_of_Votes         Gross
Metascore
Genre

Action     129.046512     7.949419     72282412  3.263226e+10
33.0
Adventure  134.111111     7.937500     22576163  9.496922e+09
41.0
Animation   99.585366     7.930488     21978630  1.463147e+10
61.0
Biography  136.022727     7.938636     24006844  8.276358e+09
48.0
Comedy     112.129032     7.901290     27620327  1.566387e+10
```

```
45.0
Crime      126.392523     8.016822     33533615  8.452632e+09
47.0
Drama      124.737024     7.957439     61367304  3.540997e+10
28.0
Family     107.500000     7.800000       551221  4.391106e+08
67.0
Fantasy     85.000000     8.000000       146222  7.827267e+08
NaN
Film-Noir  104.000000     7.966667       367215  1.259105e+08
94.0
Horror     102.090909     7.909091      3742556  1.034649e+09
46.0
Mystery    119.083333     7.975000      4203004  1.256417e+09
52.0
Thriller   108.000000     7.800000        27733  1.755074e+07
81.0
Western    148.250000     8.350000      1289665  5.822151e+07
69.0
```

```
# passing list
# suppose har numerical column ke lye [min,max & mean] nikalna chahate
hain
genres[['Runtime','IMDB_Rating','No_of_Votes','Gross','Metascore']].ag
g(['min','max','mean'])  # jis bhi column pe muiliple functions apply
karna hai usse ek list me pass kar do
                  #and jo functions apply karna hai usse bhi ek list
me pass kar do
```

```
          Runtime                  IMDB_Rating
No_of_Votes  \
              min  max        mean         min  max      mean
min
Genre

Action         45  321  129.046512         7.6  9.0  7.949419
25312
Adventure      88  228  134.111111         7.6  8.6  7.937500
29999
Animation      71  137   99.585366         7.6  8.6  7.930488
25229
Biography      93  209  136.022727         7.6  8.9  7.938636
27254
Comedy         68  188  112.129032         7.6  8.6  7.901290
26337
Crime          80  229  126.392523         7.6  9.2  8.016822
27712
Drama          64  242  124.737024         7.6  9.3  7.957439
25088
Family        100  115  107.500000         7.8  7.8  7.800000
```

```
178731
Fantasy        76   94   85.000000         7.9  8.1  8.000000
57428
Film-Noir     100  108  104.000000         7.8  8.1  7.966667
59556
Horror         71  122  102.090909         7.6  8.5  7.909091
27007
Mystery        96  138  119.083333         7.6  8.4  7.975000
33982
Thriller      108  108  108.000000         7.8  7.8  7.800000
27733
Western       132  165  148.250000         7.8  8.8  8.350000
65659

                                         Gross
\
               max          mean          min          max
mean
Genre

Action     2303232  420246.581395       3296.0  936662225.0
1.897224e+08
Adventure  1512360  313557.819444      61001.0  874211619.0
1.319017e+08
Animation   999790  268032.073171     128985.0  873839108.0
1.784326e+08
Biography  1213505  272805.045455      21877.0  753585104.0
9.404952e+07
Comedy      939631  178195.658065       1305.0  886752933.0
1.010572e+08
Crime      1826188  313398.271028       6013.0  790482117.0
7.899656e+07
Drama      2343110  212343.612457       3600.0  924558264.0
1.225259e+08
Family      372490  275610.500000    4000000.0  435110554.0
2.195553e+08
Fantasy      88794   73111.000000  337574718.0  445151978.0
3.913633e+08
Film-Noir   158731  122405.000000     449191.0  123353292.0
4.197018e+07
Horror      787806  340232.363636      89029.0  298791505.0
9.405902e+07
Mystery    1129894  350250.333333    1035953.0  474203697.0
1.047014e+08
Thriller     27733   27733.000000   17550741.0   17550741.0
1.755074e+07
Western     688390  322416.250000    5321508.0   31800000.0
1.455538e+07

          Metascore
```

```
                min    max       mean
Genre
Action         33.0   98.0  73.419580
Adventure      41.0  100.0  78.437500
Animation      61.0   96.0  81.093333
Biography      48.0   97.0  76.240506
Comedy         45.0   99.0  78.720000
Crime          47.0  100.0  77.080460
Drama          28.0  100.0  79.701245
Family         67.0   91.0  79.000000
Fantasy         NaN    NaN        NaN
Film-Noir      94.0   97.0  95.666667
Horror         46.0   97.0  80.000000
Mystery        52.0  100.0  79.125000
Thriller       81.0   81.0  81.000000
Western        69.0   90.0  78.250000

# adding both syntax
genres.agg(
    {
        'Runtime':['mean','min'],
        'IMDB_Rating':'mean',
        'No_of_Votes':['sum','max'],
        'Gross':'sum',
        'Metascore':'min'
    }
)
# abb sirf runtime me min and min value and no_of_votes me sum and max
value dikhega

              Runtime      IMDB_Rating No_of_Votes
Gross  \
                 mean  min        mean         sum     max
sum
Genre

Action     129.046512   45    7.949419    72282412 2303232
3.263226e+10
Adventure  134.111111   88    7.937500    22576163 1512360
9.496922e+09
Animation   99.585366   71    7.930488    21978630  999790
1.463147e+10
Biography  136.022727   93    7.938636    24006844 1213505
8.276358e+09
Comedy     112.129032   68    7.901290    27620327  939631
1.566387e+10
Crime      126.392523   80    8.016822    33533615 1826188
8.452632e+09
Drama      124.737024   64    7.957439    61367304 2343110
3.540997e+10
```

```
Family     107.500000  100    7.800000      551221   372490
4.391106e+08
Fantasy     85.000000   76    8.000000      146222    88794
7.827267e+08
Film-Noir  104.000000  100    7.966667      367215   158731
1.259105e+08
Horror     102.090909   71    7.909091     3742556   787806
1.034649e+09
Mystery    119.083333   96    7.975000     4203004  1129894
1.256417e+09
Thriller   108.000000  108    7.800000       27733    27733
1.755074e+07
Western    148.250000  132    8.350000     1289665   688390
5.822151e+07

          Metascore
                min
Genre
Action         33.0
Adventure      41.0
Animation      61.0
Biography      48.0
Comedy         45.0
Crime          47.0
Drama          28.0
Family         67.0
Fantasy         NaN
Film-Noir      94.0
Horror         46.0
Mystery        52.0
Thriller       81.0
Western        69.0
```

```
# looping on groups
genres
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x000001E5E9F2CA10>
```

```
for group,data in genres:   # group basically sare group ko specify
karega and data uss group ke sare data ko
    print(type(group),type(data))
```

```
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
```

```
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
<class 'str'> <class 'pandas.core.frame.DataFrame'>
```

```
for group,data in genres:
    print(data)    # har group ke data ko print kar dega jab hum data
ko print karayenge
```

```
                                                Series_Title Released_Year
Runtime  \
2                                            The Dark Knight          2008
152
5          The Lord of the Rings: The Return of the King          2003
201
8                                                  Inception          2010
148
10     The Lord of the Rings: The Fellowship of the Ring          2001
178
13                 The Lord of the Rings: The Two Towers          2002
179
..                                                       ...           ...
...
968                                             Falling Down          1993
113
979                                            Lethal Weapon          1987
109
982                                                Mad Max 2          1981
96
983                                             The Warriors          1979
92
985                                    Escape from Alcatraz          1979
112

      Genre  IMDB_Rating           Director              Star1
No_of_Votes  \
2    Action          9.0  Christopher Nolan     Christian Bale
2303232
5    Action          8.9      Peter Jackson        Elijah Wood
1642758
8    Action          8.8  Christopher Nolan  Leonardo DiCaprio
2067042
10   Action          8.8      Peter Jackson        Elijah Wood
1661481
13   Action          8.7      Peter Jackson        Elijah Wood
1485555
..      ...          ...                ...                ...
```

```
...
968  Action          7.6    Joel Schumacher    Michael Douglas
171640
979  Action          7.6     Richard Donner          Mel Gibson
236894
982  Action          7.6      George Miller          Mel Gibson
166588
983  Action          7.6        Walter Hill        Michael Beck
93878
985  Action          7.6         Don Siegel     Clint Eastwood
121731

           Gross  Metascore
2    534858444.0       84.0
5    377845905.0       94.0
8    292576195.0       74.0
10   315544750.0       92.0
13   342551365.0       87.0
..           ...        ...
968   40903593.0       56.0
979   65207127.0       68.0
982   12465371.0       77.0
983   22490039.0       65.0
985   43000000.0       76.0

[172 rows x 10 columns]
                       Series_Title Released_Year  Runtime      Genre  \
21                     Interstellar          2014      169  Adventure

47               Back to the Future          1985      116  Adventure

93             Inglourious Basterds          2009      153  Adventure

110                        Das Boot          1981      149  Adventure

114           2001: A Space Odyssey          1968      149  Adventure

..                              ...           ...      ...        ...

957  Fear and Loathing in Las Vegas          1998      118  Adventure

964                        Dead Man          1995      121  Adventure

966                      Apollo 13            PG      140  Adventure

984                The Muppet Movie          1979       95  Adventure

991                  Kelly's Heroes          1970      144  Adventure
```

```
     IMDB_Rating           Director                Star1  No_of_Votes  \
21           8.6  Christopher Nolan  Matthew McConaughey      1512360   
47           8.5    Robert Zemeckis       Michael J. Fox      1058081   
93           8.3  Quentin Tarantino            Brad Pitt      1267869   
110          8.3  Wolfgang Petersen      Jürgen Prochnow       231855   
114          8.3    Stanley Kubrick          Keir Dullea       603517   
..           ...                ...                  ...          ...   
957          7.6      Terry Gilliam          Johnny Depp       259753   
964          7.6       Jim Jarmusch          Johnny Depp        90442   
966          7.6         Ron Howard            Tom Hanks       269197   
984          7.6      James Frawley           Jim Henson        32802   
991          7.6    Brian G. Hutton       Clint Eastwood        45338   

           Gross  Metascore  
21   188020017.0       74.0  
47   210609762.0       87.0  
93   120540719.0       69.0  
110   11487676.0       86.0  
114   56954992.0       84.0  
..           ...        ...  
957   10680275.0       41.0  
964    1037847.0       62.0  
966  173837933.0       77.0  
984   76657000.0       74.0  
991    1378435.0       50.0  

[72 rows x 10 columns]
                         Series_Title Released_Year  Runtime  
Genre  \
23       Sen to Chihiro no kamikakushi          2001      125  
Animation   
43                       The Lion King          1994       88  
Animation   
46                      Hotaru no haka          1988       89  
Animation   
56                      Kimi no na wa.          2016      106  
Animation   
58   Spider-Man: Into the Spider-Verse          2018      117  
```

```
Animation
..                                   ...           ...      ...         .
..
956                                Mulan          1998       88
Animation
971                    Omohide poro poro          1991      118
Animation
976                  The Little Mermaid          1989       83
Animation
986                      Watership Down          1978       91
Animation
992                     The Jungle Book          1967       78
Animation

     IMDB_Rating             Director            Star1
No_of_Votes  \
23           8.6      Hayao Miyazaki     Daveigh Chase       651376

43           8.5        Roger Allers       Rob Minkoff       942045

46           8.5       Isao Takahata   Tsutomu Tatsumi       235231

56           8.4      Makoto Shinkai  Ryûnosuke Kamiki       194838

58           8.4     Bob Persichetti      Peter Ramsey       375110

..           ...                 ...               ...          ...

956          7.6       Tony Bancroft        Barry Cook       256906

971          7.6       Isao Takahata         Miki Imai        27071

976          7.6        Ron Clements       John Musker       237696

986          7.6        Martin Rosen       John Hubley        33656

992          7.6  Wolfgang Reitherman       Phil Harris       166409


           Gross  Metascore
23    10055859.0       96.0
43   422783777.0       88.0
46   150734678.0       94.0
56     5017246.0       79.0
58   190241310.0       87.0
..           ...        ...
956  120620254.0       71.0
971     453243.0       90.0
976  111543479.0       88.0
986  232841485.0       64.0
```

```
992  141843612.0       65.0

[82 rows x 10 columns]
          Series_Title Released_Year  Runtime      Genre  IMDB_Rating  
\
7     Schindler's List          1993      195  Biography          8.9

15          Goodfellas          1990      146  Biography          8.7

18            Hamilton          2020      160  Biography          8.6

35    The Intouchables          2011      112  Biography          8.5

38          The Pianist         2002      150  Biography          8.5

..                 ...           ...      ...        ...          ...

923     La Vie En Rose          2007      140  Biography          7.6

940  Finding Neverland          2004      106  Biography          7.6

949               Blow          2001      124  Biography          7.6

952      The Hurricane          1999      146  Biography          7.6

987   Midnight Express          1978      121  Biography          7.6


              Director               Star1  No_of_Votes        Gross  
Metascore
7    Steven Spielberg         Liam Neeson      1213505   96898818.0
94.0
15    Martin Scorsese      Robert De Niro      1020727   46836394.0
90.0
18        Thomas Kail  Lin-Manuel Miranda        55291  440984783.0
90.0
35    Olivier Nakache       Éric Toledano       760360   13182281.0
57.0
38     Roman Polanski        Adrien Brody       729603   32572577.0
85.0
..                ...                 ...          ...          ...
...
923     Olivier Dahan    Marion Cotillard        82781   10301706.0
66.0
940      Marc Forster         Johnny Depp       198677   51680613.0
67.0
949         Ted Demme         Johnny Depp       240714   52990775.0
52.0
952    Norman Jewison   Denzel Washington        91557   50668906.0
74.0
```

```
987       Alan Parker          Brad Davis        73662   35000000.0
59.0

[88 rows x 10 columns]
                                       Series_Title Released_Year
Runtime  \
19                                      Gisaengchung          2019
132
26                                  La vita è bella          1997
116
51                                      Modern Times          1936
87
52                                       City Lights          1931
87
64                                          3 Idiots          2009
170
..                                               ...           ...
...
977  The Naked Gun: From the Files of Police Squad!          1988
85
978                    Planes, Trains & Automobiles          1987
93
989                                 The Long Goodbye          1973
112
994                               A Hard Day's Night          1964
87
995                           Breakfast at Tiffany's          1961
115

      Genre  IMDB_Rating         Director            Star1
No_of_Votes  \
19   Comedy          8.6     Bong Joon Ho     Kang-ho Song
552778
26   Comedy          8.6  Roberto Benigni  Roberto Benigni
623629
51   Comedy          8.5  Charles Chaplin  Charles Chaplin
217881
52   Comedy          8.5  Charles Chaplin  Charles Chaplin
167839
64   Comedy          8.4  Rajkumar Hirani       Aamir Khan
344445
..      ...          ...              ...              ...          ..
.
977  Comedy          7.6     David Zucker   Leslie Nielsen
152871
978  Comedy          7.6      John Hughes     Steve Martin
124773
989  Comedy          7.6    Robert Altman    Elliott Gould
26337
```

```
994  Comedy          7.6   Richard Lester      John Lennon
40351
995  Comedy          7.6    Blake Edwards   Audrey Hepburn
166544

           Gross  Metascore
19    53367844.0       96.0
26    57598247.0       59.0
51      163245.0       96.0
52       19181.0       99.0
64     6532908.0       67.0
..           ...        ...
977   78756177.0       76.0
978   49530280.0       72.0
989     959000.0       87.0
994   13780024.0       96.0
995  679874270.0       76.0

[155 rows x 10 columns]
                Series_Title Released_Year  Runtime  Genre
IMDB_Rating  \
1              The Godfather          1972      175  Crime
9.2
3     The Godfather: Part II          1974      202  Crime
9.0
4               12 Angry Men          1957       96  Crime
9.0
6               Pulp Fiction          1994      154  Crime
8.9
22            Cidade de Deus          2002      130  Crime
8.6
..                       ...           ...      ...    ...          ..
.
958              Funny Games          1997      108  Crime
7.6
960                 Sleepers          1996      147  Crime
7.6
974  The Godfather: Part III          1990      162  Crime
7.6
980             Blood Simple          1984       99  Crime
7.6
999             The 39 Steps          1935       86  Crime
7.6

                 Director           Star1  No_of_Votes        Gross
Metascore
1    Francis Ford Coppola   Marlon Brando      1620367  134966411.0
100.0
3    Francis Ford Coppola       Al Pacino      1129952   57300000.0
90.0
```

```
4            Sidney Lumet     Henry Fonda      689845    4360000.0
96.0
6       Quentin Tarantino   John Travolta     1826188  107928762.0
94.0
22     Fernando Meirelles      Kátia Lund      699256    7563397.0
79.0
..                    ...             ...         ...          ...
...
958        Michael Haneke  Susanne Lothar       65058  217345863.0
69.0
960        Barry Levinson  Robert De Niro      186734   49100000.0
49.0
974  Francis Ford Coppola       Al Pacino      359809   66666062.0
60.0
980             Joel Coen      Ethan Coen       87745    2150000.0
82.0
999      Alfred Hitchcock    Robert Donat       51853  302787539.0
93.0

[107 rows x 10 columns]
                        Series_Title Released_Year  Runtime  Genre  \
0           The Shawshank Redemption          1994      142  Drama
9                         Fight Club          1999      139  Drama
11                      Forrest Gump          1994      142  Drama
17   One Flew Over the Cuckoo's Nest          1975      133  Drama
20                   Soorarai Pottru          2020      153  Drama
..                               ...           ...      ...    ...
990                     Giù la testa          1971      157  Drama
993                           Blowup          1966      111  Drama
996                            Giant          1956      201  Drama
997            From Here to Eternity          1953      118  Drama
998                         Lifeboat          1944       97  Drama

     IMDB_Rating                Director             Star1
No_of_Votes  \
0            9.3          Frank Darabont       Tim Robbins
2343110
9            8.8           David Fincher         Brad Pitt
1854740
11           8.8         Robert Zemeckis         Tom Hanks
1809221
17           8.7            Milos Forman    Jack Nicholson
918088
20           8.6           Sudha Kongara            Suriya
54995
..           ...                     ...               ...          .
..
990          7.6            Sergio Leone       Rod Steiger
30144
```

```
993          7.6  Michelangelo Antonioni     David Hemmings
56513
996          7.6          George Stevens   Elizabeth Taylor
34075
997          7.6          Fred Zinnemann     Burt Lancaster
43374
998          7.6        Alfred Hitchcock  Tallulah Bankhead
26471

           Gross  Metascore
0     28341469.0       80.0
9     37030102.0       66.0
11   330252182.0       82.0
17   112000000.0       83.0
20   556832648.0        NaN
..           ...        ...
990     696690.0       77.0
993  632532802.0       82.0
996  195217415.0       84.0
997   30500000.0       85.0
998  852142728.0       78.0

[289 rows x 10 columns]
                            Series_Title Released_Year  Runtime
Genre  \
688           E.T. the Extra-Terrestrial          1982      115
Family
698  Willy Wonka & the Chocolate Factory          1971      100
Family

     IMDB_Rating          Director         Star1  No_of_Votes
Gross  \
688          7.8  Steven Spielberg  Henry Thomas       372490
435110554.0
698          7.8        Mel Stuart   Gene Wilder       178731
4000000.0

     Metascore
688       91.0
698       67.0
                     Series_Title Released_Year  Runtime    Genre  \
321  Das Cabinet des Dr. Caligari          1920       76  Fantasy
568                     Nosferatu          1922       94  Fantasy

     IMDB_Rating      Director          Star1  No_of_Votes
Gross  \
321          8.1  Robert Wiene  Werner Krauss        57428
337574718.0
568          7.9   F.W. Murnau    Max Schreck        88794
445151978.0
```

```
     Metascore
321        NaN
568        NaN
           Series_Title Released_Year  Runtime      Genre  IMDB_Rating
\
309       The Third Man          1949      104  Film-Noir          8.1

456  The Maltese Falcon          1941      100  Film-Noir          8.0

712   Shadow of a Doubt          1943      108  Film-Noir          7.8


             Director            Star1  No_of_Votes        Gross
Metascore
309        Carol Reed     Orson Welles       158731     449191.0
97.0
456       John Huston  Humphrey Bogart       148928    2108060.0
96.0
712  Alfred Hitchcock    Teresa Wright        59556  123353292.0
94.0
                 Series_Title Released_Year  Runtime   Genre
IMDB_Rating  \
49                     Psycho          1960      109  Horror
8.5
75                      Alien          1979      117  Horror
8.4
271                 The Thing          1982      109  Horror
8.1
419              The Exorcist          1973      122  Horror
8.0
544  Night of the Living Dead          1968       96  Horror
7.9
707            The Innocents          1961      100  Horror
7.8
724                   Get Out          2017      104  Horror
7.7
844                 Halloween          1978       91  Horror
7.7
876         The Invisible Man          1933       71  Horror
7.7
932                       Saw          2004      103  Horror
7.6
948                The Others          2001      101  Horror
7.6

               Director             Star1  No_of_Votes       Gross
Metascore
49    Alfred Hitchcock   Anthony Perkins       604211   32000000.0
97.0
```

```
75          Ridley Scott  Sigourney Weaver       787806   78900000.0
89.0
271      John Carpenter      Kurt Russell       371271   13782838.0
57.0
419    William Friedkin     Ellen Burstyn       362393  232906145.0
81.0
544    George A. Romero       Duane Jones       116557      89029.0
89.0
707         Jack Clayton      Deborah Kerr        27007    2616000.0
88.0
724         Jordan Peele    Daniel Kaluuya       492851  176040665.0
85.0
844      John Carpenter  Donald Pleasence       233106   47000000.0
87.0
876          James Whale      Claude Rains        30683  298791505.0
87.0
932            James Wan        Cary Elwes       379020   56000369.0
46.0
948  Alejandro Amenábar     Nicole Kidman       337651   96522687.0
74.0
          Series_Title Released_Year  Runtime    Genre  IMDB_Rating  \
69             Memento          2000      113  Mystery          8.4
81         Rear Window          1954      112  Mystery          8.4
119            Vertigo          1958      128  Mystery          8.3
145     Shutter Island          2010      138  Mystery          8.2
220            Kahaani          2012      122  Mystery          8.1
393     Twelve Monkeys          1995      129  Mystery          8.0
420             Sleuth          1972      138  Mystery          8.0
714  The Lady Vanishes          1938       96  Mystery          7.8
829          Spoorloos          1988      107  Mystery          7.7
899          El cuerpo          2012      112  Mystery          7.6
959          Dark City          1998      100  Mystery          7.6
961       Lost Highway          1997      134  Mystery          7.6

                    Director                     Star1  No_of_Votes
Gross  \
69      Christopher Nolan                Guy Pearce      1125712
25544867.0
81       Alfred Hitchcock             James Stewart       444074
36764313.0
119      Alfred Hitchcock             James Stewart       364368
3200000.0
145       Martin Scorsese         Leonardo DiCaprio      1129894
128012934.0
220           Sujoy Ghosh               Vidya Balan        57806
1035953.0
393         Terry Gilliam              Bruce Willis       578443
57141459.0
420  Joseph L. Mankiewicz          Laurence Olivier        44748
4081254.0
```

```
714      Alfred Hitchcock         Margaret Lockwood        47400
474203697.0
829        George Sluizer  Bernard-Pierre Donnadieu        33982
367916835.0
899           Oriol Paulo             Jose Coronado        57549
140340673.0
959           Alex Proyas              Rufus Sewell       187927
14378331.0
961           David Lynch              Bill Pullman       131101
3796699.0

     Metascore
69        80.0
81       100.0
119      100.0
145       63.0
220        NaN
393       74.0
420        NaN
714       98.0
829        NaN
899        NaN
959       66.0
961       52.0
        Series_Title Released_Year  Runtime     Genre  IMDB_Rating  \
700  Wait Until Dark          1967      108  Thriller          7.8

          Director           Star1  No_of_Votes       Gross  Metascore

700  Terence Young  Audrey Hepburn        27733  17550741.0       81.0

                        Series_Title Released_Year  Runtime
Genre  \
12   Il buono, il brutto, il cattivo          1966      161  Western

48      Once Upon a Time in the West          1968      165  Western

115       Per qualche dollaro in più          1965      132  Western

691           The Outlaw Josey Wales          1976      135  Western


     IMDB_Rating        Director           Star1  No_of_Votes
Gross  \
12           8.8    Sergio Leone  Clint Eastwood       688390
6100000.0
48           8.5    Sergio Leone     Henry Fonda       302844
5321508.0
115          8.3    Sergio Leone  Clint Eastwood       232772
15000000.0
```

```
691           7.8  Clint Eastwood  Clint Eastwood        65659
31800000.0
```

```
     Metascore
12        90.0
48        80.0
115       74.0
691       69.0
```

```
for group,data in genres:
    print(group)                      #har group ko print kar dega
```

```
Action
Adventure
Animation
Biography
Comedy
Crime
Drama
Family
Fantasy
Film-Noir
Horror
Mystery
Thriller
Western
```

```
# find the heighest rated movie of each genre
for group,data in genres:
    print(data['IMDB_Rating'].max())    # ye har genre ke maxm
imdb_rating ko nikal ke dega
```

```
9.0
8.6
8.6
8.9
8.6
9.2
9.3
7.8
8.1
8.1
8.5
8.4
7.8
8.8
```

```
for group,data in genres:
    print(data[data['IMDB_Rating'] == data['IMDB_Rating'].max()] )
```

```
      Series_Title Released_Year  Runtime   Genre  IMDB_Rating  \
2  The Dark Knight          2008      152  Action          9.0

            Director           Star1  No_of_Votes        Gross
Metascore
2  Christopher Nolan  Christian Bale      2303232  534858444.0
84.0
    Series_Title Released_Year  Runtime      Genre  IMDB_Rating  \
21  Interstellar          2014      169  Adventure          8.6

             Director                Star1  No_of_Votes
Gross  \
21  Christopher Nolan  Matthew McConaughey      1512360  188020017.0


    Metascore
21       74.0
                     Series_Title Released_Year  Runtime      Genre  \
23  Sen to Chihiro no kamikakushi          2001      125  Animation

    IMDB_Rating        Director          Star1  No_of_Votes
Gross  \
23          8.6  Hayao Miyazaki  Daveigh Chase       651376
10055859.0

    Metascore
23       96.0
       Series_Title Released_Year  Runtime      Genre  IMDB_Rating  \
7  Schindler's List          1993      195  Biography          8.9

           Director        Star1  No_of_Votes       Gross  Metascore
7  Steven Spielberg  Liam Neeson      1213505  96898818.0       94.0
       Series_Title Released_Year  Runtime   Genre  IMDB_Rating  \
19     Gisaengchung          2019      132  Comedy          8.6
26  La vita è bella          1997      116  Comedy          8.6

           Director            Star1  No_of_Votes       Gross
Metascore
19     Bong Joon Ho     Kang-ho Song       552778  53367844.0
96.0
26  Roberto Benigni  Roberto Benigni       623629  57598247.0
59.0
    Series_Title Released_Year  Runtime  Genre  IMDB_Rating  \
1  The Godfather          1972      175  Crime          9.2

               Director          Star1  No_of_Votes        Gross
Metascore
1  Francis Ford Coppola  Marlon Brando      1620367  134966411.0
100.0
               Series_Title Released_Year  Runtime  Genre  IMDB_Rating
\
```

```
0  The Shawshank Redemption          1994      142  Drama          9.3


         Director        Star1  No_of_Votes       Gross  Metascore
0  Frank Darabont  Tim Robbins      2343110  28341469.0       80.0
                            Series_Title Released_Year  Runtime
Genre  \
688           E.T. the Extra-Terrestrial          1982      115
Family
698  Willy Wonka & the Chocolate Factory          1971      100
Family

     IMDB_Rating          Director         Star1  No_of_Votes
Gross  \
688          7.8  Steven Spielberg  Henry Thomas       372490
435110554.0
698          7.8        Mel Stuart   Gene Wilder       178731
4000000.0

     Metascore
688       91.0
698       67.0
                     Series_Title Released_Year  Runtime    Genre  \
321  Das Cabinet des Dr. Caligari          1920       76  Fantasy

     IMDB_Rating      Director         Star1  No_of_Votes
Gross  \
321          8.1  Robert Wiene  Werner Krauss        57428
337574718.0

     Metascore
321        NaN
      Series_Title Released_Year  Runtime      Genre  IMDB_Rating
Director  \
309  The Third Man          1949      104  Film-Noir          8.1
Carol Reed

            Star1  No_of_Votes     Gross  Metascore
309  Orson Welles       158731  449191.0       97.0
   Series_Title Released_Year  Runtime   Genre  IMDB_Rating
Director  \
49       Psycho          1960      109  Horror          8.5  Alfred
Hitchcock

              Star1  No_of_Votes       Gross  Metascore
49  Anthony Perkins       604211  32000000.0       97.0
   Series_Title Released_Year  Runtime    Genre  IMDB_Rating  \
69      Memento          2000      113  Mystery          8.4
81  Rear Window          1954      112  Mystery          8.4

             Director          Star1  No_of_Votes       Gross
```

```
Metascore
69  Christopher Nolan     Guy Pearce      1125712  25544867.0
80.0
81   Alfred Hitchcock  James Stewart       444074  36764313.0
100.0
        Series_Title Released_Year  Runtime     Genre  IMDB_Rating  \
700  Wait Until Dark          1967      108  Thriller          7.8

          Director           Star1  No_of_Votes       Gross  Metascore

700  Terence Young  Audrey Hepburn        27733  17550741.0       81.0

                       Series_Title Released_Year  Runtime    Genre  \
12  Il buono, il brutto, il cattivo          1966      161  Western

    IMDB_Rating      Director           Star1  No_of_Votes      Gross
\
12          8.8  Sergio Leone  Clint Eastwood       688390  6100000.0


    Metascore
12       90.0
```

```
# aache se dekhne ke lye
df = pd.DataFrame(columns=movies.columns)
df                         # ek empty dataframe bana lye like movies
dataframe
```

```
Empty DataFrame
Columns: [Series_Title, Released_Year, Runtime, Genre, IMDB_Rating,
Director, Star1, No_of_Votes, Gross, Metascore]
Index: []
```

```
df = pd.DataFrame(columns=movies.columns)

for group,data in genres:

    df= pd.concat([df,data[data['IMDB_Rating'] ==
data['IMDB_Rating'].max()]])

df
```

```
C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\1685166070.py:5:
FutureWarning: The behavior of DataFrame concatenation with empty or
all-NA entries is deprecated. In a future version, this will no longer
exclude empty or all-NA columns when determining the result dtypes. To
retain the old behavior, exclude the relevant entries before the
concat operation.
  df= pd.concat([df,data[data['IMDB_Rating'] ==
data['IMDB_Rating'].max()]])
```

```
                          Series_Title Released_Year Runtime
Genre  \
2                      The Dark Knight          2008     152
Action
21                        Interstellar          2014     169
Adventure
23         Sen to Chihiro no kamikakushi          2001     125
Animation
7                     Schindler's List          1993     195
Biography
19                         Gisaengchung          2019     132
Comedy
26                   La vita è bella          1997     116
Comedy
1                        The Godfather          1972     175
Crime
0             The Shawshank Redemption          1994     142
Drama
688         E.T. the Extra-Terrestrial          1982     115
Family
698  Willy Wonka & the Chocolate Factory          1971     100
Family
321         Das Cabinet des Dr. Caligari          1920      76
Fantasy
309                      The Third Man          1949     104  Film-
Noir
49                              Psycho          1960     109
Horror
69                             Memento          2000     113
Mystery
81                         Rear Window          1954     112
Mystery
700                    Wait Until Dark          1967     108
Thriller
12        Il buono, il brutto, il cattivo          1966     161
Western

     IMDB_Rating              Director              Star1
No_of_Votes  \
2            9.0     Christopher Nolan      Christian Bale
2303232
21           8.6     Christopher Nolan  Matthew McConaughey
1512360
23           8.6        Hayao Miyazaki       Daveigh Chase
651376
7            8.9      Steven Spielberg         Liam Neeson
1213505
19           8.6           Bong Joon Ho        Kang-ho Song
552778
26           8.6       Roberto Benigni     Roberto Benigni
```

```
623629
1             9.2  Francis Ford Coppola        Marlon Brando
1620367
0             9.3          Frank Darabont          Tim Robbins
2343110
688           7.8       Steven Spielberg         Henry Thomas
372490
698           7.8             Mel Stuart          Gene Wilder
178731
321           8.1           Robert Wiene        Werner Krauss
57428
309           8.1             Carol Reed         Orson Welles
158731
49            8.5       Alfred Hitchcock      Anthony Perkins
604211
69            8.4     Christopher Nolan            Guy Pearce
1125712
81            8.4       Alfred Hitchcock        James Stewart
444074
700           7.8          Terence Young       Audrey Hepburn
27733
12            8.8          Sergio Leone        Clint Eastwood
688390

           Gross  Metascore
2    534858444.0       84.0
21   188020017.0       74.0
23    10055859.0       96.0
7     96898818.0       94.0
19    53367844.0       96.0
26    57598247.0       59.0
1    134966411.0      100.0
0     28341469.0       80.0
688  435110554.0       91.0
698    4000000.0       67.0
321  337574718.0        NaN
309     449191.0       97.0
49    32000000.0       97.0
69    25544867.0       80.0
81    36764313.0      100.0
700   17550741.0       81.0
12     6100000.0       90.0

df = pd.DataFrame(columns=movies.columns)
for group, data in genres:
    # Filter for the highest IMDB rating in the current group
    highest_rating_row = data[data['IMDB_Rating'] ==
data['IMDB_Rating'].max()]
    # Use pd.concat to append the highest_rating_row to df
    df = pd.concat([df, highest_rating_row])
```

```
df

C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\492345485.py:6:
FutureWarning: The behavior of DataFrame concatenation with empty or
all-NA entries is deprecated. In a future version, this will no longer
exclude empty or all-NA columns when determining the result dtypes. To
retain the old behavior, exclude the relevant entries before the
concat operation.
  df = pd.concat([df, highest_rating_row])

                           Series_Title Released_Year Runtime
Genre  \
2                       The Dark Knight          2008     152
Action
21                         Interstellar          2014     169
Adventure
23          Sen to Chihiro no kamikakushi          2001     125
Animation
7                      Schindler's List          1993     195
Biography
19                         Gisaengchung          2019     132
Comedy
26                      La vita è bella          1997     116
Comedy
1                         The Godfather          1972     175
Crime
0              The Shawshank Redemption          1994     142
Drama
688           E.T. the Extra-Terrestrial          1982     115
Family
698  Willy Wonka & the Chocolate Factory          1971     100
Family
321          Das Cabinet des Dr. Caligari          1920      76
Fantasy
309                       The Third Man          1949     104  Film-
Noir
49                               Psycho          1960     109
Horror
69                              Memento          2000     113
Mystery
81                          Rear Window          1954     112
Mystery
700                     Wait Until Dark          1967     108
Thriller
12        Il buono, il brutto, il cattivo          1966     161
Western

     IMDB_Rating              Director                Star1
No_of_Votes  \
```

```
2              9.0     Christopher Nolan       Christian Bale
2303232
21             8.6     Christopher Nolan  Matthew McConaughey
1512360
23             8.6        Hayao Miyazaki        Daveigh Chase
651376
7              8.9      Steven Spielberg          Liam Neeson
1213505
19             8.6          Bong Joon Ho        Kang-ho Song
552778
26             8.6       Roberto Benigni      Roberto Benigni
623629
1              9.2  Francis Ford Coppola        Marlon Brando
1620367
0              9.3        Frank Darabont          Tim Robbins
2343110
688            7.8      Steven Spielberg         Henry Thomas
372490
698            7.8            Mel Stuart          Gene Wilder
178731
321            8.1          Robert Wiene        Werner Krauss
57428
309            8.1            Carol Reed         Orson Welles
158731
49             8.5      Alfred Hitchcock      Anthony Perkins
604211
69             8.4     Christopher Nolan           Guy Pearce
1125712
81             8.4      Alfred Hitchcock        James Stewart
444074
700            7.8         Terence Young       Audrey Hepburn
27733
12             8.8         Sergio Leone       Clint Eastwood
688390

           Gross  Metascore
2    534858444.0       84.0
21   188020017.0       74.0
23    10055859.0       96.0
7     96898818.0       94.0
19    53367844.0       96.0
26    57598247.0       59.0
1    134966411.0      100.0
0     28341469.0       80.0
688  435110554.0       91.0
698    4000000.0       67.0
321  337574718.0        NaN
309     449191.0       97.0
49    32000000.0       97.0
```

```
69    25544867.0       80.0
81    36764313.0      100.0
700   17550741.0       81.0
12     6100000.0       90.0
```

```python
# split (apply) combine    Strategy
# apply -> builtin function

genres.apply(min)
```

```
C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\1443917609.py:1:
FutureWarning: The provided callable <built-in function min> is
currently using np.minimum.reduce. In a future version of pandas, the
provided callable will be used directly. To keep current behavior pass
the string np.minimum.reduce instead.
  genres.apply(min)
C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\1443917609.py:1:
DeprecationWarning: DataFrameGroupBy.apply operated on the grouping
columns. This behavior is deprecated, and in a future version of
pandas the grouping columns will be excluded from the operation.
Either pass `include_groups=False` to exclude the groupings or
explicitly select the grouping columns after groupby to silence this
warning.
  genres.apply(min)

                              Series_Title Released_Year  Runtime
Genre  \
Genre

Action                                 300          1924       45
Action
Adventure            2001: A Space Odyssey          1925       88
Adventure
Animation                            Akira          1940       71
Animation
Biography                12 Years a Slave          1928       93
Biography
Comedy                 (500) Days of Summer          1921       68
Comedy
Crime                         12 Angry Men          1931       80
Crime
Drama                                 1917          1925       64
Drama
Family          E.T. the Extra-Terrestrial          1971      100
Family
Fantasy        Das Cabinet des Dr. Caligari          1920       76
Fantasy
```

```
Film-Noir                   Shadow of a Doubt          1941      100
Film-Noir
Horror                                  Alien          1933       71
Horror
Mystery                             Dark City          1938       96
Mystery
Thriller                      Wait Until Dark          1967      108
Thriller
Western    Il buono, il brutto, il cattivo          1965      132
Western

           IMDB_Rating                Director
Star1  \
Genre

Action             7.6        Abhishek Chaubey                 Aamir
Khan
Adventure          7.6          Akira Kurosawa                 Aamir
Khan
Animation          7.6             Adam Elliot               Adrian
Molina
Biography          7.6              Adam McKay               Adrien
Brody
Comedy             7.6  Alejandro G. Iñárritu                 Aamir
Khan
Crime              7.6          Akira Kurosawa                  Ajay
Devgn
Drama              7.6              Aamir Khan                 Abhay
Deol
Family             7.8              Mel Stuart                  Gene
Wilder
Fantasy            7.9             F.W. Murnau                  Max
Schreck
Film-Noir          7.8       Alfred Hitchcock             Humphrey
Bogart
Horror             7.6     Alejandro Amenábar              Anthony
Perkins
Mystery            7.6            Alex Proyas  Bernard-Pierre
Donnadieu
Thriller           7.8           Terence Young               Audrey
Hepburn
Western            7.8          Clint Eastwood               Clint
Eastwood

           No_of_Votes        Gross  Metascore
Genre
Action           25312       3296.0        NaN
Adventure        29999      61001.0        NaN
Animation        25229     128985.0        NaN
Biography        27254      21877.0        NaN
```

```
Comedy             26337       1305.0        NaN
Crime              27712       6013.0        NaN
Drama              25088       3600.0        NaN
Family            178731    4000000.0       67.0
Fantasy            57428  337574718.0        NaN
Film-Noir          59556     449191.0       94.0
Horror             27007      89029.0       46.0
Mystery            33982    1035953.0        NaN
Thriller           27733   17550741.0       81.0
Western            65659    5321508.0       69.0

# find number of movies starting with A for each group

def foo(group):
    print(group)
    return group

genres.apply(foo)

                                          Series_Title Released_Year
Runtime  \
2                                      The Dark Knight          2008
152
5        The Lord of the Rings: The Return of the King          2003
201
8                                            Inception          2010
148
10   The Lord of the Rings: The Fellowship of the Ring          2001
178
13               The Lord of the Rings: The Two Towers          2002
179
..                                                 ...           ...
...
968                                       Falling Down          1993
113
979                                     Lethal Weapon           1987
109
982                                          Mad Max 2          1981
96
983                                       The Warriors          1979
92
985                               Escape from Alcatraz          1979
112

      Genre  IMDB_Rating           Director              Star1
No_of_Votes  \
2    Action          9.0  Christopher Nolan     Christian Bale
2303232
5    Action          8.9      Peter Jackson        Elijah Wood
1642758
```

```
8    Action          8.8  Christopher Nolan  Leonardo DiCaprio
2067042
10   Action          8.8      Peter Jackson        Elijah Wood
1661481
13   Action          8.7      Peter Jackson        Elijah Wood
1485555
..      ...          ...                ...                ...
...
968  Action          7.6    Joel Schumacher    Michael Douglas
171640
979  Action          7.6     Richard Donner         Mel Gibson
236894
982  Action          7.6      George Miller         Mel Gibson
166588
983  Action          7.6        Walter Hill       Michael Beck
93878
985  Action          7.6         Don Siegel     Clint Eastwood
121731

           Gross  Metascore
2    534858444.0       84.0
5    377845905.0       94.0
8    292576195.0       74.0
10   315544750.0       92.0
13   342551365.0       87.0
..           ...        ...
968   40903593.0       56.0
979   65207127.0       68.0
982   12465371.0       77.0
983   22490039.0       65.0
985   43000000.0       76.0

[172 rows x 10 columns]
                       Series_Title Released_Year  Runtime      Genre
\
21                     Interstellar          2014      169  Adventure

47               Back to the Future          1985      116  Adventure

93             Inglourious Basterds          2009      153  Adventure

110                        Das Boot          1981      149  Adventure

114           2001: A Space Odyssey          1968      149  Adventure

..                              ...           ...      ...        ...

957  Fear and Loathing in Las Vegas          1998      118  Adventure

964                        Dead Man          1995      121  Adventure
```

```
966                         Apollo 13           PG      140  Adventure

984                   The Muppet Movie        1979       95  Adventure

991                     Kelly's Heroes        1970      144  Adventure


     IMDB_Rating            Director                Star1  No_of_Votes  \
21           8.6   Christopher Nolan  Matthew McConaughey      1512360

47           8.5     Robert Zemeckis       Michael J. Fox      1058081

93           8.3  Quentin Tarantino            Brad Pitt      1267869

110          8.3  Wolfgang Petersen       Jürgen Prochnow       231855

114          8.3     Stanley Kubrick          Keir Dullea       603517

..           ...                 ...                  ...          ...

957          7.6      Terry Gilliam          Johnny Depp       259753

964          7.6       Jim Jarmusch          Johnny Depp        90442

966          7.6          Ron Howard            Tom Hanks       269197

984          7.6      James Frawley           Jim Henson        32802

991          7.6     Brian G. Hutton       Clint Eastwood        45338


           Gross  Metascore
21   188020017.0       74.0
47   210609762.0       87.0
93   120540719.0       69.0
110   11487676.0       86.0
114   56954992.0       84.0
..           ...        ...
957   10680275.0       41.0
964    1037847.0       62.0
966  173837933.0       77.0
984   76657000.0       74.0
991    1378435.0       50.0

[72 rows x 10 columns]
                          Series_Title Released_Year  Runtime
Genre  \
23       Sen to Chihiro no kamikakushi          2001      125
Animation
```

```
43                         The Lion King          1994       88
Animation
46                        Hotaru no haka          1988       89
Animation
56                        Kimi no na wa.          2016      106
Animation
58   Spider-Man: Into the Spider-Verse          2018      117
Animation
..                                 ...           ...      ...        .
..
956                                Mulan          1998       88
Animation
971                   Omohide poro poro          1991      118
Animation
976                  The Little Mermaid          1989       83
Animation
986                      Watership Down          1978       91
Animation
992                      The Jungle Book          1967       78
Animation

     IMDB_Rating             Director             Star1
No_of_Votes  \
23           8.6       Hayao Miyazaki     Daveigh Chase       651376

43           8.5         Roger Allers       Rob Minkoff       942045

46           8.5        Isao Takahata   Tsutomu Tatsumi       235231

56           8.4       Makoto Shinkai  Ryûnosuke Kamiki       194838

58           8.4      Bob Persichetti      Peter Ramsey       375110

..           ...                  ...               ...          ...

956          7.6        Tony Bancroft        Barry Cook       256906

971          7.6        Isao Takahata         Miki Imai        27071

976          7.6         Ron Clements       John Musker       237696

986          7.6         Martin Rosen       John Hubley        33656

992          7.6  Wolfgang Reitherman       Phil Harris       166409


           Gross  Metascore
23    10055859.0       96.0
43   422783777.0       88.0
46   150734678.0       94.0
```

```
56     5017246.0       79.0
58   190241310.0       87.0
..           ...        ...
956  120620254.0       71.0
971     453243.0       90.0
976  111543479.0       88.0
986  232841485.0       64.0
992  141843612.0       65.0

[82 rows x 10 columns]
          Series_Title Released_Year  Runtime      Genre  IMDB_Rating
\
7     Schindler's List          1993      195  Biography          8.9

15          Goodfellas          1990      146  Biography          8.7

18            Hamilton          2020      160  Biography          8.6

35    The Intouchables          2011      112  Biography          8.5

38          The Pianist         2002      150  Biography          8.5

..                 ...           ...      ...        ...          ...

923     La Vie En Rose          2007      140  Biography          7.6

940  Finding Neverland          2004      106  Biography          7.6

949               Blow          2001      124  Biography          7.6

952      The Hurricane          1999      146  Biography          7.6

987   Midnight Express          1978      121  Biography          7.6


             Director                Star1  No_of_Votes        Gross
Metascore
7    Steven Spielberg          Liam Neeson      1213505   96898818.0
94.0
15    Martin Scorsese       Robert De Niro      1020727   46836394.0
90.0
18        Thomas Kail  Lin-Manuel Miranda        55291  440984783.0
90.0
35    Olivier Nakache       Éric Toledano       760360   13182281.0
57.0
38     Roman Polanski         Adrien Brody       729603   32572577.0
85.0
..                ...                  ...          ...          ...
...
923     Olivier Dahan    Marion Cotillard        82781   10301706.0
```

```
66.0
940      Marc Forster        Johnny Depp       198677   51680613.0
67.0
949         Ted Demme        Johnny Depp       240714   52990775.0
52.0
952    Norman Jewison   Denzel Washington        91557   50668906.0
74.0
987       Alan Parker         Brad Davis        73662   35000000.0
59.0

[88 rows x 10 columns]
                                       Series_Title Released_Year
Runtime  \
19                                     Gisaengchung          2019
132
26                                  La vita è bella          1997
116
51                                     Modern Times          1936
87
52                                      City Lights          1931
87
64                                         3 Idiots          2009
170
..                                              ...           ...
...
977  The Naked Gun: From the Files of Police Squad!          1988
85
978                    Planes, Trains & Automobiles          1987
93
989                                 The Long Goodbye          1973
112
994                               A Hard Day's Night          1964
87
995                           Breakfast at Tiffany's          1961
115

      Genre  IMDB_Rating         Director            Star1
No_of_Votes  \
19   Comedy          8.6      Bong Joon Ho     Kang-ho Song
552778
26   Comedy          8.6  Roberto Benigni  Roberto Benigni
623629
51   Comedy          8.5  Charles Chaplin  Charles Chaplin
217881
52   Comedy          8.5  Charles Chaplin  Charles Chaplin
167839
64   Comedy          8.4  Rajkumar Hirani       Aamir Khan
344445
..      ...          ...              ...              ...          ..
```

```
.
977  Comedy          7.6     David Zucker   Leslie Nielsen
152871
978  Comedy          7.6      John Hughes     Steve Martin
124773
989  Comedy          7.6    Robert Altman    Elliott Gould
26337
994  Comedy          7.6   Richard Lester      John Lennon
40351
995  Comedy          7.6    Blake Edwards   Audrey Hepburn
166544

           Gross  Metascore
19    53367844.0       96.0
26    57598247.0       59.0
51      163245.0       96.0
52       19181.0       99.0
64     6532908.0       67.0
..           ...        ...
977   78756177.0       76.0
978   49530280.0       72.0
989     959000.0       87.0
994   13780024.0       96.0
995  679874270.0       76.0

[155 rows x 10 columns]
                Series_Title Released_Year  Runtime  Genre
IMDB_Rating  \
1              The Godfather          1972      175  Crime
9.2
3     The Godfather: Part II          1974      202  Crime
9.0
4               12 Angry Men          1957       96  Crime
9.0
6               Pulp Fiction          1994      154  Crime
8.9
22            Cidade de Deus          2002      130  Crime
8.6
..                       ...           ...      ...    ...           ..
.
958              Funny Games          1997      108  Crime
7.6
960                 Sleepers          1996      147  Crime
7.6
974  The Godfather: Part III          1990      162  Crime
7.6
980            Blood Simple          1984       99  Crime
7.6
999             The 39 Steps          1935       86  Crime
```

```
7.6

                   Director           Star1  No_of_Votes        Gross  
Metascore  
1    Francis Ford Coppola   Marlon Brando      1620367  134966411.0  
100.0  
3    Francis Ford Coppola       Al Pacino      1129952   57300000.0  
90.0  
4            Sidney Lumet     Henry Fonda       689845    4360000.0  
96.0  
6       Quentin Tarantino   John Travolta      1826188  107928762.0  
94.0  
22     Fernando Meirelles      Kátia Lund       699256    7563397.0  
79.0  
..                    ...             ...          ...          ...  
...  
958        Michael Haneke  Susanne Lothar        65058  217345863.0  
69.0  
960        Barry Levinson  Robert De Niro       186734   49100000.0  
49.0  
974  Francis Ford Coppola       Al Pacino       359809   66666062.0  
60.0  
980             Joel Coen      Ethan Coen        87745    2150000.0  
82.0  
999      Alfred Hitchcock    Robert Donat        51853  302787539.0  
93.0  

[107 rows x 10 columns]
                        Series_Title Released_Year  Runtime  Genre  \
0           The Shawshank Redemption          1994      142  Drama   
9                         Fight Club          1999      139  Drama   
11                      Forrest Gump          1994      142  Drama   
17   One Flew Over the Cuckoo's Nest          1975      133  Drama   
20                   Soorarai Pottru          2020      153  Drama   
..                               ...           ...      ...    ...   
990                    Giù la testa          1971      157  Drama   
993                           Blowup          1966      111  Drama   
996                            Giant          1956      201  Drama   
997             From Here to Eternity          1953      118  Drama   
998                         Lifeboat          1944       97  Drama   

     IMDB_Rating                Director             Star1  
No_of_Votes  \
0            9.3          Frank Darabont       Tim Robbins       
2343110   
9            8.8           David Fincher         Brad Pitt       
1854740   
11           8.8         Robert Zemeckis         Tom Hanks       
1809221   
17           8.7            Milos Forman    Jack Nicholson       
```

```
918088
20          8.6           Sudha Kongara            Suriya
54995
..          ...                     ...               ...          .
..
990         7.6            Sergio Leone        Rod Steiger
30144
993         7.6  Michelangelo Antonioni     David Hemmings
56513
996         7.6          George Stevens   Elizabeth Taylor
34075
997         7.6          Fred Zinnemann     Burt Lancaster
43374
998         7.6        Alfred Hitchcock  Tallulah Bankhead
26471

           Gross  Metascore
0     28341469.0       80.0
9     37030102.0       66.0
11   330252182.0       82.0
17   112000000.0       83.0
20   556832648.0        NaN
..           ...        ...
990     696690.0       77.0
993  632532802.0       82.0
996  195217415.0       84.0
997   30500000.0       85.0
998  852142728.0       78.0

[289 rows x 10 columns]
                            Series_Title Released_Year  Runtime
Genre  \
688           E.T. the Extra-Terrestrial          1982      115
Family
698  Willy Wonka & the Chocolate Factory          1971      100
Family

     IMDB_Rating          Director        Star1  No_of_Votes
Gross  \
688          7.8  Steven Spielberg  Henry Thomas       372490
435110554.0
698          7.8        Mel Stuart   Gene Wilder       178731
4000000.0

     Metascore
688       91.0
698       67.0
                     Series_Title Released_Year  Runtime    Genre  \
321  Das Cabinet des Dr. Caligari          1920       76  Fantasy
568                     Nosferatu          1922       94  Fantasy
```

```
     IMDB_Rating      Director          Star1  No_of_Votes
Gross  \
321          8.1  Robert Wiene  Werner Krauss        57428
337574718.0
568          7.9   F.W. Murnau    Max Schreck        88794
445151978.0

     Metascore
321        NaN
568        NaN
           Series_Title Released_Year  Runtime      Genre  IMDB_Rating
\
309       The Third Man          1949      104  Film-Noir          8.1

456  The Maltese Falcon          1941      100  Film-Noir          8.0

712   Shadow of a Doubt          1943      108  Film-Noir          7.8


             Director           Star1  No_of_Votes        Gross
Metascore
309        Carol Reed     Orson Welles       158731     449191.0
97.0
456       John Huston  Humphrey Bogart       148928    2108060.0
96.0
712  Alfred Hitchcock    Teresa Wright        59556  123353292.0
94.0
                 Series_Title Released_Year  Runtime   Genre
IMDB_Rating  \
49                     Psycho          1960      109  Horror
8.5
75                      Alien          1979      117  Horror
8.4
271                 The Thing          1982      109  Horror
8.1
419              The Exorcist          1973      122  Horror
8.0
544  Night of the Living Dead          1968       96  Horror
7.9
707             The Innocents          1961      100  Horror
7.8
724                   Get Out          2017      104  Horror
7.7
844                 Halloween          1978       91  Horror
7.7
876         The Invisible Man          1933       71  Horror
7.7
932                       Saw          2004      103  Horror
7.6
```

```
948                The Others          2001      101  Horror
7.6

               Director              Star1  No_of_Votes        Gross
Metascore
49    Alfred Hitchcock   Anthony Perkins       604211   32000000.0
97.0
75        Ridley Scott  Sigourney Weaver       787806   78900000.0
89.0
271     John Carpenter      Kurt Russell       371271   13782838.0
57.0
419   William Friedkin     Ellen Burstyn       362393  232906145.0
81.0
544   George A. Romero       Duane Jones       116557      89029.0
89.0
707       Jack Clayton      Deborah Kerr        27007    2616000.0
88.0
724       Jordan Peele    Daniel Kaluuya       492851  176040665.0
85.0
844     John Carpenter  Donald Pleasence       233106   47000000.0
87.0
876        James Whale      Claude Rains        30683  298791505.0
87.0
932          James Wan        Cary Elwes       379020   56000369.0
46.0
948  Alejandro Amenábar     Nicole Kidman       337651   96522687.0
74.0
          Series_Title Released_Year  Runtime    Genre  IMDB_Rating  \
69             Memento          2000      113  Mystery          8.4
81         Rear Window          1954      112  Mystery          8.4
119            Vertigo          1958      128  Mystery          8.3
145     Shutter Island          2010      138  Mystery          8.2
220            Kahaani          2012      122  Mystery          8.1
393     Twelve Monkeys          1995      129  Mystery          8.0
420             Sleuth          1972      138  Mystery          8.0
714  The Lady Vanishes          1938       96  Mystery          7.8
829          Spoorloos          1988      107  Mystery          7.7
899          El cuerpo          2012      112  Mystery          7.6
959          Dark City          1998      100  Mystery          7.6
961       Lost Highway          1997      134  Mystery          7.6

                 Director                     Star1  No_of_Votes
Gross  \
69      Christopher Nolan                Guy Pearce      1125712
25544867.0
81       Alfred Hitchcock             James Stewart       444074
36764313.0
119      Alfred Hitchcock             James Stewart       364368
3200000.0
145       Martin Scorsese         Leonardo DiCaprio      1129894
```

```
128012934.0
220           Sujoy Ghosh               Vidya Balan       57806
1035953.0
393         Terry Gilliam              Bruce Willis      578443
57141459.0
420  Joseph L. Mankiewicz          Laurence Olivier       44748
4081254.0
714      Alfred Hitchcock         Margaret Lockwood       47400
474203697.0
829        George Sluizer  Bernard-Pierre Donnadieu       33982
367916835.0
899           Oriol Paulo             Jose Coronado       57549
140340673.0
959           Alex Proyas              Rufus Sewell      187927
14378331.0
961           David Lynch              Bill Pullman      131101
3796699.0

     Metascore
69        80.0
81       100.0
119      100.0
145       63.0
220        NaN
393       74.0
420        NaN
714       98.0
829        NaN
899        NaN
959       66.0
961       52.0
        Series_Title Released_Year  Runtime     Genre  IMDB_Rating  \
700  Wait Until Dark          1967      108  Thriller          7.8

          Director          Star1  No_of_Votes       Gross  Metascore

700  Terence Young  Audrey Hepburn        27733  17550741.0       81.0

                        Series_Title Released_Year  Runtime
Genre  \
12   Il buono, il brutto, il cattivo          1966      161  Western

48      Once Upon a Time in the West          1968      165  Western

115       Per qualche dollaro in più          1965      132  Western

691           The Outlaw Josey Wales          1976      135  Western


     IMDB_Rating       Director          Star1  No_of_Votes
```

```
Gross  \
12           8.8    Sergio Leone  Clint Eastwood       688390
6100000.0
48           8.5    Sergio Leone     Henry Fonda       302844
5321508.0
115          8.3    Sergio Leone  Clint Eastwood       232772
15000000.0
691          7.8  Clint Eastwood  Clint Eastwood        65659
31800000.0

     Metascore
12        90.0
48        80.0
115       74.0
691       69.0
```

```
C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\2904676014.py:1:
DeprecationWarning: DataFrameGroupBy.apply operated on the grouping
columns. This behavior is deprecated, and in a future version of
pandas the grouping columns will be excluded from the operation.
Either pass `include_groups=False` to exclude the groupings or
explicitly select the grouping columns after groupby to silence this
warning.
  genres.apply(foo)
```

```
                                                    Series_Title
Released_Year  \
Genre

Action   2                                       The Dark Knight
2008
         5        The Lord of the Rings: The Return of the King
2003
         8                                             Inception
2010
         10   The Lord of the Rings: The Fellowship of the Ring
2001
         13                The Lord of the Rings: The Two Towers
2002
...                                                          ...
...
Thriller 700                                     Wait Until Dark
1967
Western  12                      Il buono, il brutto, il cattivo
1966
         48                         Once Upon a Time in the West
1968
         115                          Per qualche dollaro in più
1965
         691                              The Outlaw Josey Wales
```

```
1976

                Runtime     Genre  IMDB_Rating           Director  \
Genre
Action   2          152    Action          9.0  Christopher Nolan
         5          201    Action          8.9      Peter Jackson
         8          148    Action          8.8  Christopher Nolan
         10         178    Action          8.8      Peter Jackson
         13         179    Action          8.7      Peter Jackson
...                 ...       ...          ...                ...
Thriller 700        108  Thriller          7.8      Terence Young
Western  12         161   Western          8.8      Sergio Leone
         48         165   Western          8.5      Sergio Leone
         115        132   Western          8.3      Sergio Leone
         691        135   Western          7.8     Clint Eastwood

                           Star1  No_of_Votes        Gross  Metascore
Genre
Action   2        Christian Bale      2303232  534858444.0       84.0
         5           Elijah Wood      1642758  377845905.0       94.0
         8     Leonardo DiCaprio      2067042  292576195.0       74.0
         10          Elijah Wood      1661481  315544750.0       92.0
         13          Elijah Wood      1485555  342551365.0       87.0
...                          ...          ...          ...        ...
Thriller 700      Audrey Hepburn        27733   17550741.0       81.0
Western  12       Clint Eastwood       688390    6100000.0       90.0
         48          Henry Fonda       302844    5321508.0       80.0
         115      Clint Eastwood       232772   15000000.0       74.0
         691      Clint Eastwood        65659   31800000.0       69.0

[1000 rows x 10 columns]
```

```python
def foo(group):
    print(group['Series_Title'].str.startswith('A'))   # ek boolean series milega
    return group

genres.apply(foo)
```

```
2      False
5      False
8      False
10     False
13     False
       ...
968    False
979    False
982    False
983    False
985    False
```

```
Name: Series_Title, Length: 172, dtype: bool
21     False
47     False
93     False
110    False
114    False
       ...
957    False
964    False
966     True
984    False
991    False
Name: Series_Title, Length: 72, dtype: bool
23     False
43     False
46     False
56     False
58     False
       ...
956    False
971    False
976    False
986    False
992    False
Name: Series_Title, Length: 82, dtype: bool
7      False
15     False
18     False
35     False
38     False
       ...
923    False
940    False
949    False
952    False
987    False
Name: Series_Title, Length: 88, dtype: bool
19     False
26     False
51     False
52     False
64     False
       ...
977    False
978    False
989    False
994     True
995    False
Name: Series_Title, Length: 155, dtype: bool
```

```
1      False
3      False
4      False
6      False
22     False
       ...
958    False
960    False
974    False
980    False
999    False
Name: Series_Title, Length: 107, dtype: bool
0      False
9      False
11     False
17     False
20     False
       ...
990    False
993    False
996    False
997    False
998    False
Name: Series_Title, Length: 289, dtype: bool
688    False
698    False
Name: Series_Title, dtype: bool
321    False
568    False
Name: Series_Title, dtype: bool
309    False
456    False
712    False
Name: Series_Title, dtype: bool
49     False
75      True
271    False
419    False
544    False
707    False
724    False
844    False
876    False
932    False
948    False
Name: Series_Title, dtype: bool
69     False
81     False
119    False
```

```
145    False
220    False
393    False
420    False
714    False
829    False
899    False
959    False
961    False
Name: Series_Title, dtype: bool
700    False
Name: Series_Title, dtype: bool
12     False
48     False
115    False
691    False
Name: Series_Title, dtype: bool

C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\2904676014.py:1:
DeprecationWarning: DataFrameGroupBy.apply operated on the grouping
columns. This behavior is deprecated, and in a future version of
pandas the grouping columns will be excluded from the operation.
Either pass `include_groups=False` to exclude the groupings or
explicitly select the grouping columns after groupby to silence this
warning.
  genres.apply(foo)

                                                            Series_Title
Released_Year  \
Genre

Action   2                                               The Dark Knight
2008
         5        The Lord of the Rings: The Return of the King
2003
         8                                                     Inception
2010
         10   The Lord of the Rings: The Fellowship of the Ring
2001
         13               The Lord of the Rings: The Two Towers
2002
...                                                                  ...
...
Thriller 700                                             Wait Until Dark
1967
Western  12                          Il buono, il brutto, il cattivo
1966
         48                             Once Upon a Time in the West
1968
         115                               Per qualche dollaro in più
```

```
1965
         691                             The Outlaw Josey Wales
1976

                  Runtime     Genre  IMDB_Rating          Director  \
Genre
Action   2        152    Action          9.0  Christopher Nolan
         5        201    Action          8.9      Peter Jackson
         8        148    Action          8.8  Christopher Nolan
         10       178    Action          8.8      Peter Jackson
         13       179    Action          8.7      Peter Jackson
...               ...       ...          ...                ...
Thriller 700      108  Thriller          7.8      Terence Young
Western  12       161   Western          8.8       Sergio Leone
         48       165   Western          8.5       Sergio Leone
         115      132   Western          8.3       Sergio Leone
         691      135   Western          7.8     Clint Eastwood

                            Star1  No_of_Votes        Gross  Metascore
Genre
Action   2        Christian Bale      2303232  534858444.0       84.0
         5           Elijah Wood      1642758  377845905.0       94.0
         8      Leonardo DiCaprio     2067042  292576195.0       74.0
         10          Elijah Wood      1661481  315544750.0       92.0
         13          Elijah Wood      1485555  342551365.0       87.0
...                           ...          ...          ...        ...
Thriller 700      Audrey Hepburn        27733   17550741.0       81.0
Western  12       Clint Eastwood       688390    6100000.0       90.0
         48          Henry Fonda       302844    5321508.0       80.0
         115      Clint Eastwood       232772   15000000.0       74.0
         691      Clint Eastwood        65659   31800000.0       69.0

[1000 rows x 10 columns]
```

```python
def foo(group):
    return (group['Series_Title'].str.startswith('A')).sum()   # ek boolean series milega


genres.apply(foo)
```

```
C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\2904676014.py:1: DeprecationWarning: DataFrameGroupBy.apply operated on the grouping columns. This behavior is deprecated, and in a future version of pandas the grouping columns will be excluded from the operation. Either pass `include_groups=False` to exclude the groupings or explicitly select the grouping columns after groupby to silence this warning.
  genres.apply(foo)
```

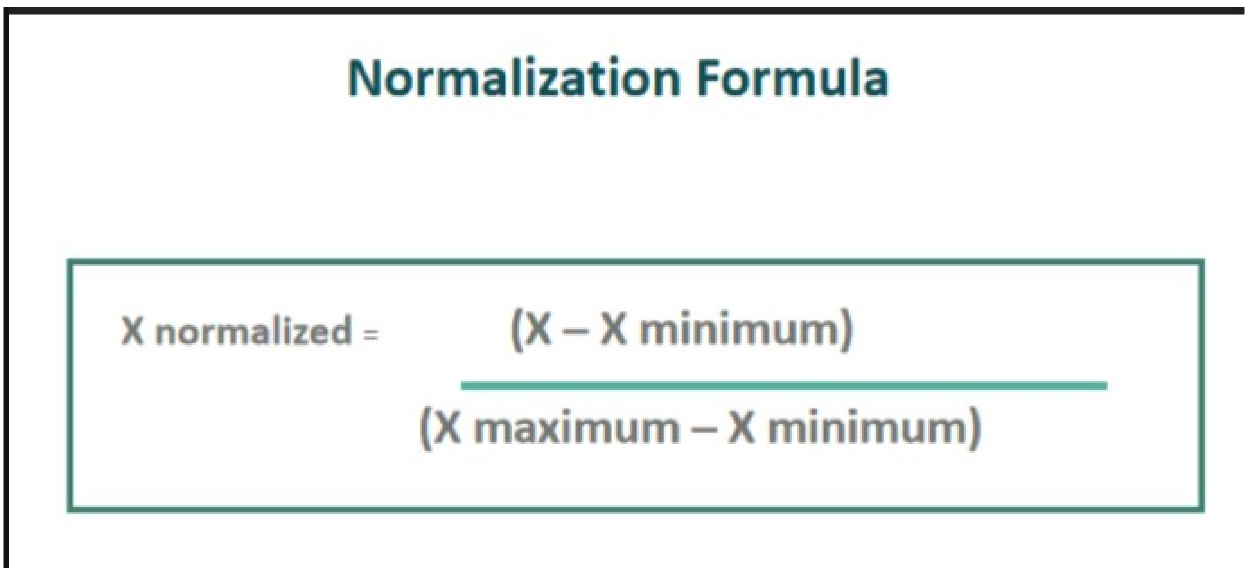
```
Genre
Action        10
Adventure      2
Animation      2
Biography      9
Comedy        14
Crime          4
Drama         21
Family         0
Fantasy        0
Film-Noir      0
Horror         1
Mystery        0
Thriller       0
Western        0
dtype: int64
```

```python
# find ranking of each movie in the group according to IMDB score
def rank_movie(group):
    group['genre_rank']= group['IMDB_Rating'].rank(ascending=False)
    return group

genres.apply(rank_movie)
```

```
C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\2409710219.py:1:
DeprecationWarning: DataFrameGroupBy.apply operated on the grouping
columns. This behavior is deprecated, and in a future version of
pandas the grouping columns will be excluded from the operation.
Either pass `include_groups=False` to exclude the groupings or
explicitly select the grouping columns after groupby to silence this
warning.
  genres.apply(rank_movie)

                                                   Series_Title
Released_Year  \
Genre

Action   2                                      The Dark Knight
2008
         5        The Lord of the Rings: The Return of the King
2003
         8                                            Inception
2010
         10   The Lord of the Rings: The Fellowship of the Ring
2001
         13               The Lord of the Rings: The Two Towers
2002
...                                                         ...
...
Thriller 700                                    Wait Until Dark
```

```
1967
Western  12                          Il buono, il brutto, il cattivo
1966
         48                               Once Upon a Time in the West
1968
         115                                Per qualche dollaro in più
1965
         691                                    The Outlaw Josey Wales
1976

              Runtime     Genre  IMDB_Rating           Director  \
Genre
Action   2        152    Action          9.0  Christopher Nolan
         5        201    Action          8.9      Peter Jackson
         8        148    Action          8.8  Christopher Nolan
         10       178    Action          8.8      Peter Jackson
         13       179    Action          8.7      Peter Jackson
...               ...       ...          ...                ...
Thriller 700      108  Thriller          7.8      Terence Young
Western  12       161   Western          8.8       Sergio Leone
         48       165   Western          8.5       Sergio Leone
         115      132   Western          8.3       Sergio Leone
         691      135   Western          7.8     Clint Eastwood

                            Star1  No_of_Votes        Gross
Metascore  \
Genre

Action   2         Christian Bale      2303232  534858444.0       84.0

         5            Elijah Wood      1642758  377845905.0       94.0

         8     Leonardo DiCaprio      2067042  292576195.0       74.0

         10           Elijah Wood      1661481  315544750.0       92.0

         13           Elijah Wood      1485555  342551365.0       87.0

...                           ...          ...          ...        ...

Thriller 700      Audrey Hepburn        27733   17550741.0       81.0

Western  12       Clint Eastwood       688390    6100000.0       90.0

         48          Henry Fonda       302844    5321508.0       80.0

         115      Clint Eastwood       232772   15000000.0       74.0

         691      Clint Eastwood        65659   31800000.0       69.0
```

```
                 genre_rank
Genre
Action    2             1.0
          5             2.0
          8             3.5
          10            3.5
          13            6.0
...                     ...
Thriller 700            1.0
Western   12            1.0
          48            2.0
          115           3.0
          691           4.0

[1000 rows x 11 columns]

# find normalized IMDB rating group wise
```




```
def normal(group):
    group['norm_rating'] =(group['IMDB_Rating'] - 
group['IMDB_Rating'].min())/(group['IMDB_Rating'].max() - 
group['IMDB_Rating'].min())
    return group

genres.apply(normal)

C:\Users\jayra\AppData\Local\Temp\ipykernel_23732\2582220829.py:5: 
DeprecationWarning: DataFrameGroupBy.apply operated on the grouping 
columns. This behavior is deprecated, and in a future version of 
pandas the grouping columns will be excluded from the operation. 
Either pass `include_groups=False` to exclude the groupings or 
explicitly select the grouping columns after groupby to silence this
```

```
warning.
  genres.apply(normal)

                                                     Series_Title
Released_Year  \
Genre

Action   2                                        The Dark Knight
2008
         5        The Lord of the Rings: The Return of the King
2003
         8                                              Inception
2010
         10   The Lord of the Rings: The Fellowship of the Ring
2001
         13               The Lord of the Rings: The Two Towers
2002
...                                                           ...
...
Thriller 700                                      Wait Until Dark
1967
Western  12                     Il buono, il brutto, il cattivo
1966
         48                        Once Upon a Time in the West
1968
         115                         Per qualche dollaro in più
1965
         691                             The Outlaw Josey Wales
1976

              Runtime     Genre  IMDB_Rating          Director  \
Genre
Action   2        152    Action          9.0  Christopher Nolan
         5        201    Action          8.9      Peter Jackson
         8        148    Action          8.8  Christopher Nolan
         10       178    Action          8.8      Peter Jackson
         13       179    Action          8.7      Peter Jackson
...               ...       ...          ...                ...
Thriller 700      108  Thriller          7.8      Terence Young
Western  12       161   Western          8.8       Sergio Leone
         48       165   Western          8.5       Sergio Leone
         115      132   Western          8.3       Sergio Leone
         691      135   Western          7.8     Clint Eastwood

                          Star1  No_of_Votes        Gross
Metascore  \
Genre

Action   2       Christian Bale      2303232  534858444.0       84.0
```

```
          5          Elijah Wood      1642758  377845905.0       94.0

          8    Leonardo DiCaprio      2067042  292576195.0       74.0

          10         Elijah Wood      1661481  315544750.0       92.0

          13         Elijah Wood      1485555  342551365.0       87.0

...                          ...          ...          ...        ...

Thriller 700      Audrey Hepburn        27733   17550741.0       81.0

Western  12      Clint Eastwood        688390    6100000.0       90.0

         48          Henry Fonda       302844    5321508.0       80.0

         115     Clint Eastwood        232772   15000000.0       74.0

         691     Clint Eastwood         65659   31800000.0       69.0


              norm_rating
Genre
Action   2       1.000000
         5       0.928571
         8       0.857143
         10      0.857143
         13      0.785714
...                   ...
Thriller 700          NaN
Western  12      1.000000
         48      0.700000
         115     0.500000
         691     0.000000

[1000 rows x 11 columns]
```

```python
# groupby on multiple cols
duo = movies.groupby(['Director','Star1'])
duo
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x0000023A9C458410>
```

```python
#size
duo.size()
```

```
Director             Star1
Aamir Khan           Amole Gupte       1
Aaron Sorkin         Eddie Redmayne    1
Abdellatif Kechiche  Léa Seydoux       1
```

```
Abhishek Chaubey     Shahid Kapoor     1
Abhishek Kapoor      Amit Sadh         1
                                      ..
Zaza Urushadze       Lembit Ulfsak     1
Zoya Akhtar          Hrithik Roshan    1
                     Vijay Varma       1
Çagan Irmak          Çetin Tekindor    1
Ömer Faruk Sorak     Cem Yilmaz        1
Length: 898, dtype: int64
```

```
#get_group
duo.get_group(('Aamir Khan','Amole Gupte'))
```

```
        Series_Title Released_Year  Runtime  Genre  IMDB_Rating
Director  \
65  Taare Zameen Par          2007      165  Drama          8.4  Aamir
Khan

          Star1  No_of_Votes      Gross  Metascore
65  Amole Gupte       168895  1223869.0        NaN
```

```
# # find the most earning actor->director combo
duo['Gross'].sum()  # abhi sare actor director ka earning ek sath aa
jayega
```

```
Director              Star1
Aamir Khan            Amole Gupte         1223869.0
Aaron Sorkin          Eddie Redmayne    853090410.0
Abdellatif Kechiche   Léa Seydoux         2199675.0
Abhishek Chaubey      Shahid Kapoor     218428303.0
Abhishek Kapoor       Amit Sadh           1122527.0
                                            ...
Zaza Urushadze        Lembit Ulfsak        144501.0
Zoya Akhtar           Hrithik Roshan      3108485.0
                      Vijay Varma         5566534.0
Çagan Irmak           Çetin Tekindor    461855363.0
Ömer Faruk Sorak      Cem Yilmaz        196206077.0
Name: Gross, Length: 898, dtype: float64
```

```
duo['Gross'].sum().sort_values(ascending=True)
```

```
Director               Star1
Anders Thomas Jensen   Ulrich Thomsen     1.305000e+03
Thomas Jahn            Til Schweiger      3.296000e+03
Jaco Van Dormael       Jared Leto         3.600000e+03
Shane Meadows          Paddy Considine    6.013000e+03
Jeong-beom Lee         Won Bin            6.460000e+03
                                              ...
Werner Herzog          Klaus Kinski       1.124605e+09
Christopher Nolan      Christian Bale     1.242940e+09
Billy Wilder           William Holden     1.286779e+09
```

```
Anthony Russo         Joe Russo          2.205039e+09
Akira Kurosawa        Toshirô Mifune     2.999877e+09
Name: Gross, Length: 898, dtype: float64
```

```
duo['Gross'].sum().sort_values(ascending=True).head(1)
```

```
Director              Star1
Anders Thomas Jensen  Ulrich Thomsen    1305.0
Name: Gross, dtype: float64
```

```
# find the best(in-terms of metascore(avg)) actor -> genre combo
movies.groupby(['Star1','Genre'])['Metascore'].mean()
```

```
Star1                 Genre
Aamir Khan            Action        NaN
                      Adventure    84.0
                      Comedy       67.0
Aaron Taylor-Johnson  Action       66.0
Abhay Deol            Drama         NaN
                                   ...
Zbigniew Zamachowski  Comedy       88.0
Zooey Deschanel       Comedy       76.0
Çetin Tekindor        Drama         NaN
Éric Toledano         Biography    57.0
Ömer Faruk Sorak      Comedy        NaN
Name: Metascore, Length: 829, dtype: float64
```

```
movies.groupby(['Star1','Genre'])['Metascore'].mean().reset_index()  #
reset index karne se dataframe ban jayega
```

```
                    Star1      Genre  Metascore
0              Aamir Khan     Action        NaN
1              Aamir Khan  Adventure       84.0
2              Aamir Khan     Comedy       67.0
3    Aaron Taylor-Johnson     Action       66.0
4              Abhay Deol      Drama        NaN
..                    ...        ...        ...
824  Zbigniew Zamachowski     Comedy       88.0
825       Zooey Deschanel     Comedy       76.0
826        Çetin Tekindor      Drama        NaN
827         Éric Toledano  Biography       57.0
828      Ömer Faruk Sorak     Comedy        NaN

[829 rows x 3 columns]
```

```
movies.groupby(['Star1','Genre'])
['Metascore'].mean().reset_index().sort_values('Metascore',ascending=F
alse).head(1)
```

```
              Star1  Genre  Metascore
230  Ellar Coltrane  Drama      100.0
```

```python
# agg on multiple groupby
duo[['Runtime','IMDB_Rating','No_of_Votes','Gross','Metascore']].agg([
'min','max','mean'])
```

```
                                       Runtime             IMDB_Rating
\
                                           min  max   mean         min
max mean
Director             Star1

Aamir Khan           Amole Gupte           165  165  165.0         8.4
8.4  8.4
Aaron Sorkin         Eddie Redmayne        129  129  129.0         7.8
7.8  7.8
Abdellatif Kechiche Léa Seydoux            180  180  180.0         7.7
7.7  7.7
Abhishek Chaubey     Shahid Kapoor         148  148  148.0         7.8
7.8  7.8
Abhishek Kapoor      Amit Sadh             130  130  130.0         7.7
7.7  7.7
...                                        ...  ...    ...         ...  ..
.  ...
Zaza Urushadze       Lembit Ulfsak          87   87   87.0         8.2
8.2  8.2
Zoya Akhtar          Hrithik Roshan        155  155  155.0         8.1
8.1  8.1
                     Vijay Varma           154  154  154.0         8.0
8.0  8.0
Çagan Irmak          Çetin Tekindor        112  112  112.0         8.3
8.3  8.3
Ömer Faruk Sorak     Cem Yilmaz            127  127  127.0         8.0
8.0  8.0

                                       No_of_Votes
Gross  \
                                               min      max       mean
min
Director             Star1

Aamir Khan           Amole Gupte            168895   168895   168895.0
1223869.0
Aaron Sorkin         Eddie Redmayne          89896    89896    89896.0
853090410.0
Abdellatif Kechiche Léa Seydoux             138741   138741   138741.0
2199675.0
Abhishek Chaubey     Shahid Kapoor           27175    27175    27175.0
218428303.0
Abhishek Kapoor      Amit Sadh               32628    32628    32628.0
1122527.0
...                                            ...      ...        ...
```

```
...
Zaza Urushadze      Lembit Ulfsak        40382   40382   40382.0
144501.0
Zoya Akhtar         Hrithik Roshan       67927   67927   67927.0
3108485.0
                    Vijay Varma          31886   31886   31886.0
5566534.0
Çagan Irmak         Çetin Tekindor       78925   78925   78925.0
461855363.0
Ömer Faruk Sorak    Cem Yilmaz           56960   56960   56960.0
196206077.0

                                                                Metascore
\
                                            max         mean        min
max
Director            Star1

Aamir Khan          Amole Gupte       1223869.0    1223869.0        NaN
NaN
Aaron Sorkin        Eddie Redmayne  853090410.0  853090410.0       77.0
77.0
Abdellatif Kechiche Léa Seydoux       2199675.0    2199675.0       89.0
89.0
Abhishek Chaubey    Shahid Kapoor   218428303.0  218428303.0        NaN
NaN
Abhishek Kapoor     Amit Sadh         1122527.0    1122527.0       40.0
40.0
...                                         ...          ...        ...
...
Zaza Urushadze      Lembit Ulfsak      144501.0     144501.0       73.0
73.0
Zoya Akhtar         Hrithik Roshan    3108485.0    3108485.0        NaN
NaN
                    Vijay Varma       5566534.0    5566534.0       65.0
65.0
Çagan Irmak         Çetin Tekindor  461855363.0  461855363.0        NaN
NaN
Ömer Faruk Sorak    Cem Yilmaz      196206077.0  196206077.0        NaN
NaN


                                    mean
Director            Star1
Aamir Khan          Amole Gupte      NaN
Aaron Sorkin        Eddie Redmayne  77.0
Abdellatif Kechiche Léa Seydoux     89.0
Abhishek Chaubey    Shahid Kapoor    NaN
Abhishek Kapoor     Amit Sadh       40.0
...                                  ...
```

```
Zaza Urushadze      Lembit Ulfsak   73.0
Zoya Akhtar         Hrithik Roshan   NaN
                    Vijay Varma     65.0
Çagan Irmak         Çetin Tekindor   NaN
Ömer Faruk Sorak    Cem Yilmaz       NaN

[898 rows x 15 columns]
```

EXERCISE

```
ipl = pd.read_csv('deliveries.csv')
ipl

        match_id  inning         batting_team
bowling_team  \
0              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
1              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
2              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
3              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
4              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
...          ...     ...                  ...
...
179073     11415       2  Chennai Super Kings               Mumbai
Indians
179074     11415       2  Chennai Super Kings               Mumbai
Indians
179075     11415       2  Chennai Super Kings               Mumbai
Indians
179076     11415       2  Chennai Super Kings               Mumbai
Indians
179077     11415       2  Chennai Super Kings               Mumbai
Indians

        over  ball    batsman non_striker      bowler
is_super_over  ...  \
0          1     1  DA Warner    S Dhawan    TS Mills
0  ...
1          1     2  DA Warner    S Dhawan    TS Mills
0  ...
2          1     3  DA Warner    S Dhawan    TS Mills
0  ...
3          1     4  DA Warner    S Dhawan    TS Mills
0  ...
4          1     5  DA Warner    S Dhawan    TS Mills
```

```
0  ...
...      ...  ...       ...        ...        ...            ...
...
179073    20    2  RA Jadeja   SR Watson  SL Malinga
0  ...
179074    20    3  SR Watson   RA Jadeja  SL Malinga
0  ...
179075    20    4  SR Watson   RA Jadeja  SL Malinga
0  ...
179076    20    5  SN Thakur   RA Jadeja  SL Malinga
0  ...
179077    20    6  SN Thakur   RA Jadeja  SL Malinga
0  ...

        bye_runs  legbye_runs  noball_runs  penalty_runs  batsman_runs
\
0              0            0            0             0             0

1              0            0            0             0             0

2              0            0            0             0             4

3              0            0            0             0             0

4              0            0            0             0             0

...          ...          ...          ...           ...           ...

179073         0            0            0             0             1

179074         0            0            0             0             2

179075         0            0            0             0             1

179076         0            0            0             0             2

179077         0            0            0             0             0


        extra_runs  total_runs  player_dismissed dismissal_kind
fielder
0                0           0               NaN            NaN
NaN
1                0           0               NaN            NaN
NaN
2                0           4               NaN            NaN
NaN
3                0           0               NaN            NaN
NaN
4                2           2               NaN            NaN
```

```
NaN
...             ...         ...              ...            ...
...
179073            0           1              NaN            NaN
NaN
179074            0           2              NaN            NaN
NaN
179075            0           1        SR Watson        run out  KH
Pandya
179076            0           2              NaN            NaN
NaN
179077            0           0        SN Thakur            lbw
NaN

[179078 rows x 21 columns]
```

```
ipl.shape
```

```
(179078, 21)
```

```
ipl.head()
```

```
   match_id  inning         batting_team                 bowling_team
over  \
0         1       1  Sunrisers Hyderabad  Royal Challengers Bangalore
1
1         1       1  Sunrisers Hyderabad  Royal Challengers Bangalore
1
2         1       1  Sunrisers Hyderabad  Royal Challengers Bangalore
1
3         1       1  Sunrisers Hyderabad  Royal Challengers Bangalore
1
4         1       1  Sunrisers Hyderabad  Royal Challengers Bangalore
1

   ball    batsman non_striker    bowler  is_super_over  ...  bye_runs
\
0     1  DA Warner    S Dhawan  TS Mills              0  ...         0

1     2  DA Warner    S Dhawan  TS Mills              0  ...         0

2     3  DA Warner    S Dhawan  TS Mills              0  ...         0

3     4  DA Warner    S Dhawan  TS Mills              0  ...         0

4     5  DA Warner    S Dhawan  TS Mills              0  ...         0


   legbye_runs  noball_runs  penalty_runs  batsman_runs  extra_runs  \
0            0            0             0             0           0
1            0            0             0             0           0
```

```
2            0            0             0             4           0
3            0            0             0             0           0
4            0            0             0             0           2

   total_runs  player_dismissed dismissal_kind fielder
0           0               NaN            NaN     NaN
1           0               NaN            NaN     NaN
2           4               NaN            NaN     NaN
3           0               NaN            NaN     NaN
4           2               NaN            NaN     NaN

[5 rows x 21 columns]
```

```
# find the top 10 batsman in terms of run scored

# har batsman ne jitne ball kheli hai usko grouped kar lo
ipl.groupby('batsman')
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x0000023AA3FC7AA0>
```

```
ipl.groupby('batsman')['batsman_runs'].sum()  # har group ke batsman
ka run sum ho jayega abhi index ke according sort hai value ke
according kar lo
```

```
batsman
A Ashish Reddy     280
A Chandila           4
A Chopra            53
A Choudhary         25
A Dananjaya          4
                  ...
YV Takawale        192
Yashpal Singh       47
Younis Khan          3
Yuvraj Singh      2765
Z Khan             117
Name: batsman_runs, Length: 516, dtype: int64
```

```
ipl.groupby('batsman')
['batsman_runs'].sum().sort_values(ascending=False)
```

```
batsman
V Kohli        5434
SK Raina       5415
RG Sharma      4914
DA Warner      4741
S Dhawan       4632
               ...
IC Pandey         0
J Denly           0
```

```
P Raj              0
Sunny Gupta        0
L Ablish           0
Name: batsman_runs, Length: 516, dtype: int64
```

```
ipl.groupby('batsman')
['batsman_runs'].sum().sort_values(ascending=False).head(10)
```

```
batsman
V Kohli            5434
SK Raina           5415
RG Sharma          4914
DA Warner          4741
S Dhawan           4632
CH Gayle           4560
MS Dhoni           4477
RV Uthappa         4446
AB de Villiers     4428
G Gambhir          4223
Name: batsman_runs, dtype: int64
```

```
# find the batsman with max no of sixes
ipl['batsman_runs'] == 6   # ek boolean series mil jayega
```

```
0         False
1         False
2         False
3         False
4         False
          ...
179073    False
179074    False
179075    False
179076    False
179077    False
Name: batsman_runs, Length: 179078, dtype: bool
```

```
six = ipl[ipl['batsman_runs'] == 6]
six                                    # sare ball ka details mil
jayega jaha pe ki six laga hai
```

```
        match_id  inning         batting_team
bowling_team  \
10             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
47             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
75             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
89             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
```

```
91              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
...           ...     ...                  ...
...
178987     11415       2  Chennai Super Kings               Mumbai
Indians
179048     11415       2  Chennai Super Kings               Mumbai
Indians
179061     11415       2  Chennai Super Kings               Mumbai
Indians
179062     11415       2  Chennai Super Kings               Mumbai
Indians
179063     11415       2  Chennai Super Kings               Mumbai
Indians

        over  ball       batsman   non_striker       bowler
is_super_over  \
10         2     4     DA Warner      S Dhawan  A Choudhary
0
47         8     4  MC Henriques      S Dhawan      TM Head
0
75        13     2  Yuvraj Singh  MC Henriques  A Choudhary
0
89        15     3  Yuvraj Singh  MC Henriques    S Aravind
0
91        15     5  MC Henriques  Yuvraj Singh    S Aravind
0
...      ...   ...           ...           ...          ...
...
178987     6     4     SR Watson      SK Raina   SL Malinga
0
179048    16     1      DJ Bravo     SR Watson   SL Malinga
0
179061    18     2     SR Watson      DJ Bravo    KH Pandya
0
179062    18     3     SR Watson      DJ Bravo    KH Pandya
0
179063    18     4     SR Watson      DJ Bravo    KH Pandya
0

        ...  bye_runs  legbye_runs  noball_runs  penalty_runs
batsman_runs  \
10      ...         0            0            0             0
6
47      ...         0            0            0             0
6
75      ...         0            0            0             0
6
89      ...         0            0            0             0
```

```
6
91       ...          0            0            0             0
6
...      ...        ...          ...          ...           ...
...
178987   ...          0            0            0             0
6
179048   ...          0            0            0             0
6
179061   ...          0            0            0             0
6
179062   ...          0            0            0             0
6
179063   ...          0            0            0             0
6

        extra_runs  total_runs  player_dismissed dismissal_kind
fielder
10               0           6               NaN            NaN
NaN
47               0           6               NaN            NaN
NaN
75               0           6               NaN            NaN
NaN
89               0           6               NaN            NaN
NaN
91               0           6               NaN            NaN
NaN
...            ...         ...               ...            ...     ..
.
178987           0           6               NaN            NaN
NaN
179048           0           6               NaN            NaN
NaN
179061           0           6               NaN            NaN
NaN
179062           0           6               NaN            NaN
NaN
179063           0           6               NaN            NaN
NaN

[8170 rows x 21 columns]
```

```
six.groupby('batsman')# batsman ke basis pe group kar diye and abb
kisi bhi column ko extract karo suppose batsman colum ko hi  pakarte
hain
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at
0x0000023AA3B3BBC0>
```

```
six.groupby('batsman')['batsman']  # abb iski count nikal lo
```

```
<pandas.core.groupby.generic.SeriesGroupBy object at
0x0000023A9CFF9B80>
```

```
six.groupby('batsman')['batsman'].count()  # abb sort kar lo
```

```
batsman
A Ashish Reddy      15
A Choudhary          1
A Flintoff           2
A Hales              6
A Mishra             5
                  ...
Y Venugopal Rao     37
YK Pathan          161
YV Takawale          3
Yuvraj Singh       149
Z Khan               2
Name: batsman, Length: 336, dtype: int64
```

```
six.groupby('batsman')['batsman'].count().sort_values(ascending=False)
```

```
batsman
CH Gayle           327
AB de Villiers     214
MS Dhoni           207
SK Raina           195
RG Sharma          194
                  ...
CK Langeveldt        1
JDS Neesham          1
SK Trivedi           1
D Shorey             1
MN van Wyk           1
Name: batsman, Length: 336, dtype: int64
```

```
six.groupby('batsman')
['batsman'].count().sort_values(ascending=False).head(1)
```

```
batsman
CH Gayle    327
Name: batsman, dtype: int64
```

```
# find batsman with most number of 4's and 6's in last 5 overs

# hume unn sare ball ki need hi nahi hai jo ki 1 se 15 over tak feka
gya hai
temp=ipl[ipl['over']>15]
temp
```

```
        match_id  inning         batting_team
bowling_team  \
93             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
94             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
95             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
96             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
97             1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
...          ...     ...                  ...
...
179073     11415       2  Chennai Super Kings               Mumbai
Indians
179074     11415       2  Chennai Super Kings               Mumbai
Indians
179075     11415       2  Chennai Super Kings               Mumbai
Indians
179076     11415       2  Chennai Super Kings               Mumbai
Indians
179077     11415       2  Chennai Super Kings               Mumbai
Indians

        over  ball       batsman   non_striker      bowler
is_super_over  \
93        16     1  MC Henriques  Yuvraj Singh   YS Chahal
0
94        16     2  MC Henriques  Yuvraj Singh   YS Chahal
0
95        16     3  Yuvraj Singh      DJ Hooda   YS Chahal
0
96        16     4      DJ Hooda  Yuvraj Singh   YS Chahal
0
97        16     5  Yuvraj Singh      DJ Hooda   YS Chahal
0
...      ...   ...           ...           ...         ...
...
179073    20     2     RA Jadeja     SR Watson  SL Malinga
0
179074    20     3     SR Watson     RA Jadeja  SL Malinga
0
179075    20     4     SR Watson     RA Jadeja  SL Malinga
0
179076    20     5     SN Thakur     RA Jadeja  SL Malinga
0
179077    20     6     SN Thakur     RA Jadeja  SL Malinga
0
```

```
        ...  bye_runs  legbye_runs  noball_runs  penalty_runs  
batsman_runs  \
93      ...         0            0            0             0   
0   
94      ...         0            0            0             0   
0   
95      ...         0            0            0             0   
1   
96      ...         0            0            0             0   
1   
97      ...         0            0            0             0   
1   
...     ...       ...          ...          ...           ...   
...   
179073  ...         0            0            0             0   
1   
179074  ...         0            0            0             0   
2   
179075  ...         0            0            0             0   
1   
179076  ...         0            0            0             0   
2   
179077  ...         0            0            0             0   
0   

        extra_runs  total_runs  player_dismissed dismissal_kind  
fielder  
93               0           0               NaN            NaN  
NaN  
94               0           0      MC Henriques         caught  
Sachin Baby  
95               0           1               NaN            NaN  
NaN  
96               0           1               NaN            NaN  
NaN  
97               0           1               NaN            NaN  
NaN  
...            ...         ...               ...            ...  
...  
179073           0           1               NaN            NaN  
NaN  
179074           0           2               NaN            NaN  
NaN  
179075           0           1         SR Watson        run out    KH 
Pandya  
179076           0           2               NaN            NaN  
NaN  
179077           0           0         SN Thakur            lbw  
NaN  
```

```
[40400 rows x 21 columns]

temp[(temp['batsman_runs'] ==4) | (temp['batsman_runs'] ==6)]

        match_id  inning         batting_team
bowling_team  \
101            1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
105            1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
109            1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
114            1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
115            1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
...          ...     ...                  ...
...
179052     11415       2  Chennai Super Kings               Mumbai
Indians
179061     11415       2  Chennai Super Kings               Mumbai
Indians
179062     11415       2  Chennai Super Kings               Mumbai
Indians
179063     11415       2  Chennai Super Kings               Mumbai
Indians
179071     11415       2  Chennai Super Kings               Mumbai
Indians

        over  ball       batsman   non_striker       bowler
is_super_over  \
101       17     3  Yuvraj Singh      DJ Hooda     TS Mills
0
105       18     1      DJ Hooda  Yuvraj Singh  A Choudhary
0
109       18     5  Yuvraj Singh      DJ Hooda  A Choudhary
0
114       19     2  Yuvraj Singh      DJ Hooda     TS Mills
0
115       19     3  Yuvraj Singh      DJ Hooda     TS Mills
0
...      ...   ...           ...           ...          ...
...
179052    16     5     SR Watson      DJ Bravo   SL Malinga
0
179061    18     2     SR Watson      DJ Bravo    KH Pandya
0
179062    18     3     SR Watson      DJ Bravo    KH Pandya
0
```

```
179063    18     4     SR Watson      DJ Bravo    KH Pandya
0
179071    19     6     RA Jadeja     SR Watson    JJ Bumrah
0

        ...  bye_runs  legbye_runs  noball_runs  penalty_runs
batsman_runs  \
101     ...         0            0            0             0
4
105     ...         0            0            0             0
6
109     ...         0            0            0             0
4
114     ...         0            0            0             0
4
115     ...         0            0            0             0
6
...     ...       ...          ...          ...           ...
...
179052  ...         0            0            0             0
4
179061  ...         0            0            0             0
6
179062  ...         0            0            0             0
6
179063  ...         0            0            0             0
6
179071  ...         4            0            0             0
4

        extra_runs  total_runs  player_dismissed dismissal_kind
fielder
101              0           4               NaN            NaN
NaN
105              0           6               NaN            NaN
NaN
109              0           4               NaN            NaN
NaN
114              0           4               NaN            NaN
NaN
115              0           6               NaN            NaN
NaN
...            ...         ...               ...            ...     ..
.
179052           0           4               NaN            NaN
NaN
179061           0           6               NaN            NaN
NaN
179062           0           6               NaN            NaN
```

```
NaN
179063           0           6                NaN            NaN
NaN
179071           4           8                NaN            NaN
NaN

[7592 rows x 21 columns]

temp.groupby('batsman')
['batsman'].count().sort_values(ascending=False).head(1).index[0]

'MS Dhoni'

# find V Kohli's record against all teams
temp = ipl[ipl['batsman'] == 'V Kohli']
temp

        match_id  inning                 batting_team
bowling_team  \
2590          12       1  Royal Challengers Bangalore       Mumbai
Indians
2591          12       1  Royal Challengers Bangalore       Mumbai
Indians
2593          12       1  Royal Challengers Bangalore       Mumbai
Indians
2594          12       1  Royal Challengers Bangalore       Mumbai
Indians
2597          12       1  Royal Challengers Bangalore       Mumbai
Indians
...          ...     ...                          ...
...
177522     11345       2  Royal Challengers Bangalore  Sunrisers
Hyderabad
177523     11345       2  Royal Challengers Bangalore  Sunrisers
Hyderabad
177524     11345       2  Royal Challengers Bangalore  Sunrisers
Hyderabad
177525     11345       2  Royal Challengers Bangalore  Sunrisers
Hyderabad
177527     11345       2  Royal Challengers Bangalore  Sunrisers
Hyderabad

        over  ball  batsman     non_striker          bowler
is_super_over  \
2590       1     2  V Kohli        CH Gayle      TG Southee
0
2591       1     3  V Kohli        CH Gayle      TG Southee
0
2593       1     5  V Kohli        CH Gayle      TG Southee
0
```

```
2594       1     6  V Kohli        CH Gayle       TG Southee
0
2597       2     1  V Kohli        CH Gayle  Harbhajan Singh
0
...      ...   ...      ...             ...              ...
...
177522     1     7  V Kohli  AB de Villiers          B Kumar
0
177523     2     1  V Kohli  AB de Villiers          K Ahmed
0
177524     2     2  V Kohli  AB de Villiers          K Ahmed
0
177525     2     3  V Kohli  AB de Villiers          K Ahmed
0
177527     2     5  V Kohli  AB de Villiers          K Ahmed
0

        ...  bye_runs  legbye_runs  noball_runs  penalty_runs
batsman_runs  \
2590    ...         0            0            0             0
0
2591    ...         0            0            0             0
1
2593    ...         0            0            0             0
0
2594    ...         0            0            0             0
1
2597    ...         0            0            0             0
0
...     ...       ...          ...          ...           ...
...
177522  ...         0            0            0             0
1
177523  ...         0            0            0             0
4
177524  ...         0            0            0             0
6
177525  ...         0            0            0             0
1
177527  ...         0            0            0             0
0

        extra_runs  total_runs  player_dismissed dismissal_kind
fielder
2590             1           1               NaN            NaN
NaN
2591             0           1               NaN            NaN
NaN
2593             0           0               NaN            NaN
```

```
NaN
2594              0           1                NaN            NaN
NaN
2597              0           0                NaN            NaN
NaN
...             ...         ...                ...            ...       .
..
177522            0           1                NaN            NaN
NaN
177523            0           4                NaN            NaN
NaN
177524            0           6                NaN            NaN
NaN
177525            0           1                NaN            NaN
NaN
177527            0           0            V Kohli         caught  WP
Saha

[4211 rows x 21 columns]
```

```python
temp.groupby('bowling_team')['batsman_runs'].sum()
```

```
bowling_team
Chennai Super Kings         749
Deccan Chargers             306
Delhi Capitals               66
Delhi Daredevils            763
Gujarat Lions               283
Kings XI Punjab             636
Kochi Tuskers Kerala         50
Kolkata Knight Riders       675
Mumbai Indians              628
Pune Warriors               128
Rajasthan Royals            370
Rising Pune Supergiant       83
Rising Pune Supergiants     188
Sunrisers Hyderabad         509
Name: batsman_runs, dtype: int64
```

```python
temp.groupby('bowling_team')['batsman_runs'].sum().reset_index()
```

```
               bowling_team  batsman_runs
0       Chennai Super Kings           749
1           Deccan Chargers           306
2            Delhi Capitals            66
3          Delhi Daredevils           763
4             Gujarat Lions           283
5           Kings XI Punjab           636
6      Kochi Tuskers Kerala            50
7     Kolkata Knight Riders           675
```

```
8            Mumbai Indians           628
9             Pune Warriors           128
10         Rajasthan Royals           370
11   Rising Pune Supergiant            83
12  Rising Pune Supergiants           188
13      Sunrisers Hyderabad           509

# Create a function that can return the highest score of any batsman
ipl

        match_id  inning         batting_team
bowling_team  \
0              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
1              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
2              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
3              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
4              1       1  Sunrisers Hyderabad  Royal Challengers
Bangalore
...          ...     ...                  ...
...
179073     11415       2  Chennai Super Kings               Mumbai
Indians
179074     11415       2  Chennai Super Kings               Mumbai
Indians
179075     11415       2  Chennai Super Kings               Mumbai
Indians
179076     11415       2  Chennai Super Kings               Mumbai
Indians
179077     11415       2  Chennai Super Kings               Mumbai
Indians

        over  ball    batsman non_striker      bowler
is_super_over  ...  \
0          1     1  DA Warner    S Dhawan    TS Mills
0  ...
1          1     2  DA Warner    S Dhawan    TS Mills
0  ...
2          1     3  DA Warner    S Dhawan    TS Mills
0  ...
3          1     4  DA Warner    S Dhawan    TS Mills
0  ...
4          1     5  DA Warner    S Dhawan    TS Mills
0  ...
...      ...   ...        ...         ...         ...              ...
...
179073    20     2  RA Jadeja   SR Watson  SL Malinga
```

```
0  ...
179074    20     3  SR Watson   RA Jadeja  SL Malinga
0  ...
179075    20     4  SR Watson   RA Jadeja  SL Malinga
0  ...
179076    20     5  SN Thakur   RA Jadeja  SL Malinga
0  ...
179077    20     6  SN Thakur   RA Jadeja  SL Malinga
0  ...

        bye_runs  legbye_runs  noball_runs  penalty_runs  batsman_runs
\
0              0            0            0             0             0

1              0            0            0             0             0

2              0            0            0             0             4

3              0            0            0             0             0

4              0            0            0             0             0

...          ...          ...          ...           ...           ...

179073         0            0            0             0             1

179074         0            0            0             0             2

179075         0            0            0             0             1

179076         0            0            0             0             2

179077         0            0            0             0             0


        extra_runs  total_runs  player_dismissed dismissal_kind
fielder
0                0           0               NaN            NaN
NaN
1                0           0               NaN            NaN
NaN
2                0           4               NaN            NaN
NaN
3                0           0               NaN            NaN
NaN
4                2           2               NaN            NaN
NaN
...            ...         ...               ...            ...
...
179073           0           1               NaN            NaN
```

```
NaN
179074              0             2                 NaN            NaN
NaN
179075              0             1           SR Watson         run out   KH
Pandya
179076              0             2                 NaN            NaN
NaN
179077              0             0           SN Thakur            lbw
NaN

[179078 rows x 21 columns]
```

```
temp= ipl[ipl['batsman'] =='V Kohli']
temp.groupby('match_id')
['batsman_runs'].sum().sort_values(ascending=False).head(1).values[0]
```

```
113
```

```
def highest(batsman):
    temp = ipl[ipl['batsman'] == batsman]
    return temp.groupby('match_id')
['batsman_runs'].sum().sort_values(ascending=False).head(1).values[0]


highest('MS Dhoni')
```

```
89
```

```
highest('DA Warner')
```

```
126
```

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

```
                  Series_Title Released_Year  Runtime   Genre
IMDB_Rating  \
0    The Shawshank Redemption          1994      142   Drama
9.3
1               The Godfather          1972      175   Crime
9.2
2             The Dark Knight          2008      152  Action
9.0
3      The Godfather: Part II          1974      202   Crime
9.0
4                12 Angry Men          1957       96   Crime
9.0
..                        ...           ...      ...     ...
...
995    Breakfast at Tiffany's          1961      115  Comedy
7.6
```

```
996                          Giant          1956      201   Drama
7.6
997        From Here to Eternity          1953      118   Drama
7.6
998                     Lifeboat          1944       97   Drama
7.6
999                 The 39 Steps          1935       86   Crime
7.6

                 Director              Star1  No_of_Votes        Gross
\
0          Frank Darabont        Tim Robbins      2343110   28341469.0

1    Francis Ford Coppola      Marlon Brando      1620367  134966411.0

2       Christopher Nolan     Christian Bale      2303232  534858444.0

3    Francis Ford Coppola          Al Pacino      1129952   57300000.0

4            Sidney Lumet        Henry Fonda       689845    4360000.0

..                    ...                ...          ...          ...

995         Blake Edwards     Audrey Hepburn       166544  679874270.0

996        George Stevens  Elizabeth Taylor        34075  195217415.0

997        Fred Zinnemann     Burt Lancaster        43374   30500000.0

998      Alfred Hitchcock  Tallulah Bankhead       26471  852142728.0

999      Alfred Hitchcock      Robert Donat        51853  302787539.0


     Metascore
0         80.0
1        100.0
2         84.0
3         90.0
4         96.0
..         ...
995       76.0
996       84.0
997       85.0
998       78.0
999       93.0

[1000 rows x 10 columns]
```