netflix

February 27, 2024

```python
import pandas as pd
netflix = pd.read_csv(r"D:\Data\archive (6)\Netflix TV Shows and Movies.csv")
```

```python
netflix
```

```
[ ]:       index          id                              title   type  \
      0         0     tm84618                       Taxi Driver  MOVIE
      1         1    tm127384     Monty Python and the Holy Grail  MOVIE
      2         2     tm70993                     Life of Brian  MOVIE
      3         3    tm190788                       The Exorcist  MOVIE
      4         4     ts22164         Monty Python's Flying Circus   SHOW
      ...     ...         ...                               ...    ...
      5278   5278  tm1040816         Momshies! Your Soul is Mine  MOVIE
      5279   5279  tm1014599                         Fine Wine  MOVIE
      5280   5280  tm1045018                             Clash  MOVIE
      5281   5281  tm1098060                     Shadow Parties  MOVIE
      5282   5282   ts271048  Mighty Little Bheem: Kite Festival   SHOW

                                        description  release_year  \
      0      A mentally unstable Vietnam War veteran works …          1976
      1      King Arthur, accompanied by his squire, recrui…          1975
      2      Brian Cohen is an average young Jewish man, bu…          1979
      3      12-year-old Regan MacNeil begins to adapt an e…          1973
      4      A British sketch comedy series with the shows …          1969
      ...                                        ...           ...
      5278   Three women with totally different lives accid…          2021
      5279   A beautiful love story that can happen between…          2021
      5280   A man from Nigeria returns to his family in Ca…          2021
      5281   A family faces destruction in a long-running c…          2021
      5282   With winter behind them, Bheem and his townspe…          2021

            age_certification  runtime      imdb_id  imdb_score  imdb_votes
      0                     R       113   tt0075314         8.3    795222.0
      1                    PG        91   tt0071853         8.2    530877.0
      2                     R        94   tt0079470         8.0    392419.0
      3                     R       133   tt0070047         8.1    391942.0
      4                 TV-14        30   tt0063929         8.8     72895.0
      ...                 ...       ...         ...         ...         ...
```

```
5278            NaN    108  tt14412240        5.8       26.0
5279            NaN    100  tt13857480        6.9       39.0
5280            NaN     88  tt14620732        6.5       32.0
5281            NaN    116  tt10168094        6.2        9.0
5282            NaN      0  tt13711094        8.8       16.0

[5283 rows x 11 columns]
```

```
[ ]: netflix.drop(columns= ['index','id','imdb_id','description'] )
```

```
[ ]:                                   title    type  release_year  \
     0                          Taxi Driver   MOVIE          1976
     1      Monty Python and the Holy Grail   MOVIE          1975
     2                        Life of Brian   MOVIE          1979
     3                         The Exorcist   MOVIE          1973
     4          Monty Python's Flying Circus    SHOW          1969
     ...                                 ...     ...           ...
     5278         Momshies! Your Soul is Mine   MOVIE          2021
     5279                           Fine Wine   MOVIE          2021
     5280                               Clash   MOVIE          2021
     5281                       Shadow Parties   MOVIE          2021
     5282  Mighty Little Bheem: Kite Festival    SHOW          2021

           age_certification  runtime  imdb_score  imdb_votes
     0                     R      113         8.3    795222.0
     1                    PG       91         8.2    530877.0
     2                     R       94         8.0    392419.0
     3                     R      133         8.1    391942.0
     4                 TV-14       30         8.8     72895.0
     ...                 ...      ...         ...         ...
     5278                NaN      108         5.8        26.0
     5279                NaN      100         6.9        39.0
     5280                NaN       88         6.5        32.0
     5281                NaN      116         6.2         9.0
     5282                NaN        0         8.8        16.0

     [5283 rows x 7 columns]
```

```
[ ]: netflix.describe()
```

```
[ ]:              index  release_year      runtime   imdb_score    imdb_votes
     count  5283.000000   5283.000000  5283.000000  5283.000000  5.267000e+03
     mean   2641.000000   2015.879992    79.199886     6.533447  2.340719e+04
     std    1525.215067      7.346098    38.915974     1.160932  8.713432e+04
     min       0.000000   1953.000000     0.000000     1.500000  5.000000e+00
     25%    1320.500000   2015.000000    45.000000     5.800000  5.210000e+02
     50%    2641.000000   2018.000000    87.000000     6.600000  2.279000e+03
```

```
75%      3961.500000    2020.000000    106.000000      7.400000   1.014400e+04
max      5282.000000    2022.000000    235.000000      9.600000   2.268288e+06
```

[ ]: netflix.dtypes

[ ]: index   int64
     id   object
     title   object
     type   object
     description   object
     release_year   int64
     age_certification object
     runtime   int64
     imdb_id   object
     imdb_score   float64
     imdb_votes   float64
     dtype: object

[ ]: netflix.isna().sum()

[ ]: index   0
     id   0
     title   0
     type   0
     description   5
     release_year   0
     age_certification 2285
     runtime   0
     imdb_id   0
     imdb_score   0
     imdb_votes   16
     dtype: int64

[ ]: netflix['age_certification'].value_counts()

[ ]: age_certification
     TV-MA   792
     R   548
     TV-14   436
     PG-13   424
     PG   238
     TV-PG   172
     G   105
     TV-Y7   104
     TV-Y   94
     TV-G   72
     NC-17   13
```

```
Name: count, dtype: int64
```

[ ]: `netflix['imdb_score'].value_counts()`

[ ]:
```
imdb_score
6.6    201
6.8    199
6.5    193
6.2    192
7.4    190
       ...
2.0      1
1.9      1
2.4      1
1.6      1
1.8      1
Name: count, Length: 81, dtype: int64
```

[ ]: `netflix.shape`

[ ]: `(5283, 11)`

[ ]:
```
netflix['rating']= pd.qcut(netflix['imdb_score'] , 3, labels =␣
 ↪['bad','good','v_good'])
netflix
```

[ ]:
```
         index          id                                   title   type  \
0            0     tm84618                             Taxi Driver  MOVIE
1            1    tm127384     Monty Python and the Holy Grail  MOVIE
2            2     tm70993                           Life of Brian  MOVIE
3            3    tm190788                            The Exorcist  MOVIE
4            4     ts22164       Monty Python's Flying Circus   SHOW
...        ...         ...                                     ...    ...
5278      5278   tm1040816         Momshies! Your Soul is Mine  MOVIE
5279      5279   tm1014599                               Fine Wine  MOVIE
5280      5280   tm1045018                                   Clash  MOVIE
5281      5281   tm1098060                         Shadow Parties  MOVIE
5282      5282    ts271048  Mighty Little Bheem: Kite Festival   SHOW

                                       description  release_year  \
0        A mentally unstable Vietnam War veteran works …          1976
1        King Arthur, accompanied by his squire, recrui…          1975
2        Brian Cohen is an average young Jewish man, bu…          1979
3        12-year-old Regan MacNeil begins to adapt an e…          1973
4        A British sketch comedy series with the shows …          1969
...                                              ...           ...
5278  Three women with totally different lives accid…          2021
```

```
5279  A beautiful love story that can happen between…        2021
5280  A man from Nigeria returns to his family in Ca…        2021
5281  A family faces destruction in a long-running c…        2021
5282  With winter behind them, Bheem and his townspe…        2021

      age_certification  runtime       imdb_id  imdb_score  imdb_votes  rating
0                     R      113  tt0075314            8.3    795222.0  v_good
1                    PG       91  tt0071853            8.2    530877.0  v_good
2                     R       94  tt0079470            8.0    392419.0  v_good
3                     R      133  tt0070047            8.1    391942.0  v_good
4                 TV-14       30  tt0063929            8.8     72895.0  v_good
...                 ...      ...        ...            ...         ...     ...
5278                NaN      108  tt14412240           5.8        26.0     bad
5279                NaN      100  tt13857480           6.9        39.0    good
5280                NaN       88  tt14620732           6.5        32.0    good
5281                NaN      116  tt10168094           6.2         9.0     bad
5282                NaN        0  tt13711094           8.8        16.0  v_good

[5283 rows x 12 columns]
```

```
[ ]: netflix.drop(columns= ['index','id','imdb_id','description'] )
```

```
[ ]:                                    title    type   release_year  \
0                            Taxi Driver   MOVIE           1976
1          Monty Python and the Holy Grail   MOVIE           1975
2                          Life of Brian   MOVIE           1979
3                           The Exorcist   MOVIE           1973
4              Monty Python's Flying Circus    SHOW           1969
...                                   ...     ...            ...
5278           Momshies! Your Soul is Mine   MOVIE           2021
5279                            Fine Wine   MOVIE           2021
5280                                Clash   MOVIE           2021
5281                        Shadow Parties   MOVIE           2021
5282  Mighty Little Bheem: Kite Festival    SHOW           2021

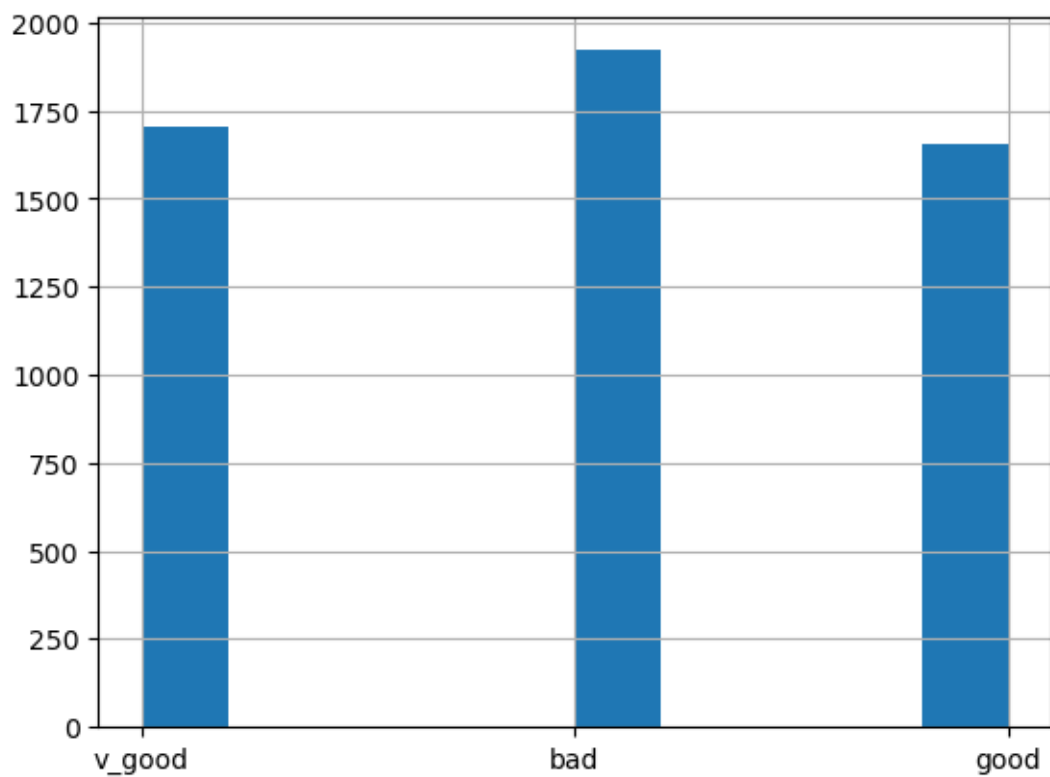      age_certification  runtime  imdb_score  imdb_votes  rating
0                     R      113         8.3    795222.0  v_good
1                    PG       91         8.2    530877.0  v_good
2                     R       94         8.0    392419.0  v_good
3                     R      133         8.1    391942.0  v_good
4                 TV-14       30         8.8     72895.0  v_good
...                 ...      ...         ...         ...     ...
5278                NaN      108         5.8        26.0     bad
5279                NaN      100         6.9        39.0    good
5280                NaN       88         6.5        32.0    good
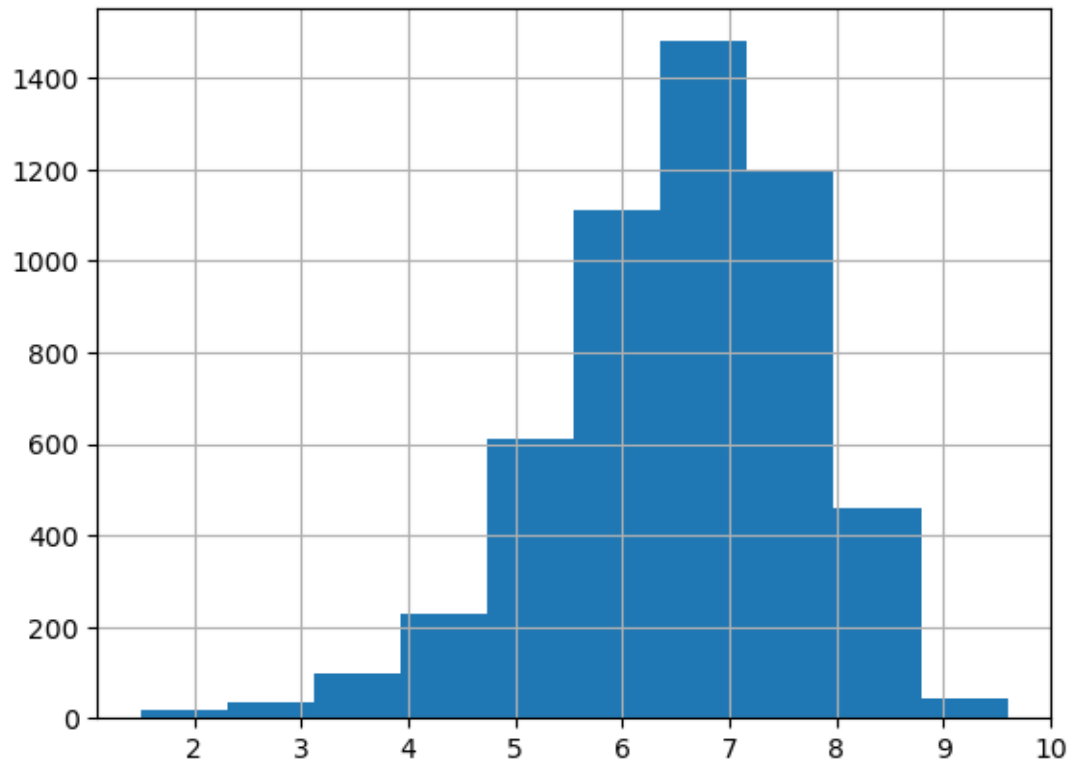5281                NaN      116         6.2         9.0     bad
5282                NaN        0         8.8        16.0  v_good
```

[5283 rows x 8 columns]

[ ]: netflix['rating'].hist()

[ ]: <Axes: >

[ ]: netflix['imdb_score'].hist()

[ ]: <Axes: >

```
[ ]: netflix['type'].value_counts()
```

```
[ ]: type
     MOVIE    3407
     SHOW     1876
     Name: count, dtype: int64
```

```
[ ]: netflix[(netflix['release_year'] >= 1959 ) & (netflix['release_year'] <= 1995 )␣
     ↪]
```

```
[ ]:      index        id                          title   type  \
     0        0   tm84618                    Taxi Driver   MOVIE
     1        1  tm127384  Monty Python and the Holy Grail   MOVIE
     2        2   tm70993                   Life of Brian   MOVIE
     3        3  tm190788                    The Exorcist   MOVIE
     4        4   ts22164      Monty Python's Flying Circus    SHOW
     ..     ...       ...                            ...     ...
     218    218  tm117840                   The Emigrant   MOVIE
     219    219  tm334788                          Yodha   MOVIE
     223    223  tm118256                    Yaar Gaddar   MOVIE
     225    225  tm314602                        Dushmani   MOVIE
     228    228  tm449775                        Shikari   MOVIE
```

```
                                              description  release_year  \
0      A mentally unstable Vietnam War veteran works …          1976
1      King Arthur, accompanied by his squire, recrui…          1975
2      Brian Cohen is an average young Jewish man, bu…          1979
3      12-year-old Regan MacNeil begins to adapt an e…          1973
4      A British sketch comedy series with the shows …          1969
..                                                  …            …
218    The biblical tale of Joseph is told from an Eg…          1994
219    Suraj Singh is addicted to drugs and as a resu…          1991
223    When his brother becomes involved in a deadly …          1994
225    In order to settle personal scores; two gang l…          1995
228    Shankar and Chanchal earn their living by doin…          1991

     age_certification  runtime    imdb_id  imdb_score  imdb_votes  rating
0                    R      113  tt0075314         8.3    795222.0  v_good
1                   PG       91  tt0071853         8.2    530877.0  v_good
2                    R       94  tt0079470         8.0    392419.0  v_good
3                    R      133  tt0070047         8.1    391942.0  v_good
4                TV-14       30  tt0063929         8.8     72895.0  v_good
..                 …        …          …           …           …       …
218                NaN      129  tt0110545         6.5       998.0    good
219                NaN      142  tt0267130         4.5       198.0     bad
223                NaN      147  tt0111780         3.6       142.0     bad
225                NaN      156  tt0112916         5.0       248.0     bad
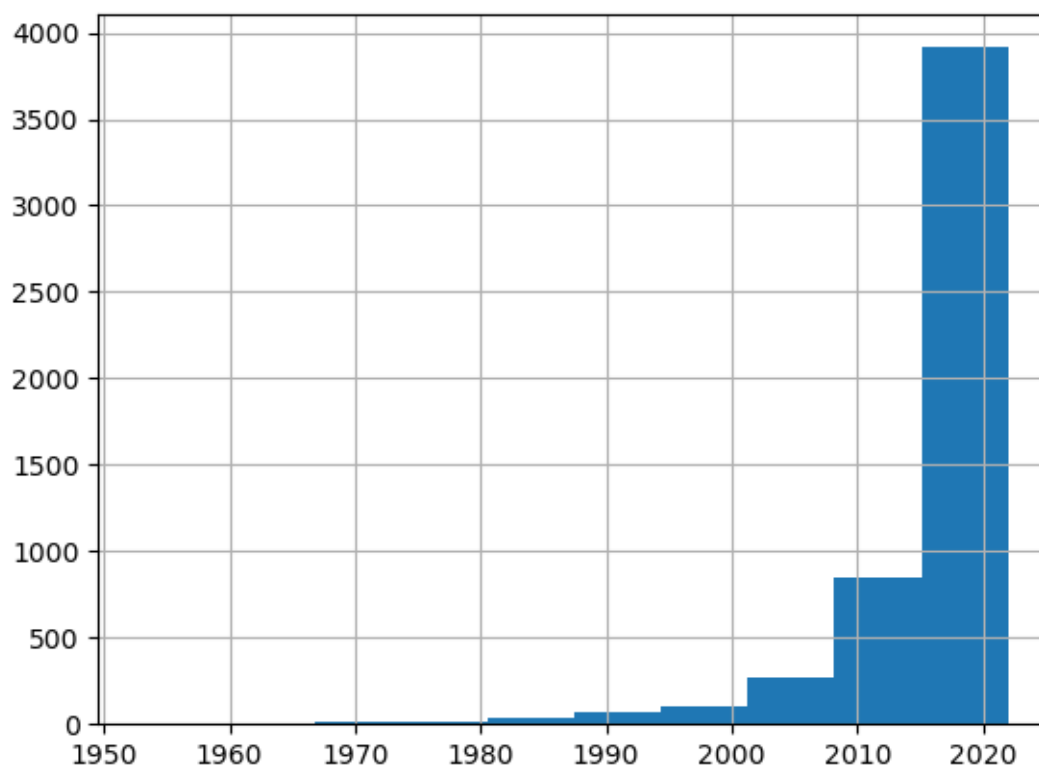228                NaN      162  tt0170416         5.9        66.0     bad

[151 rows x 12 columns]
```

[ ]: `netflix['release_year'].value_counts()`

```
[ ]: release_year
     2019    749
     2018    733
     2021    687
     2020    657
     2017    555
              …
     1974      1
     1959      1
     1962      1
     1978      1
     1985      1
     Name: count, Length: 66, dtype: int64
```

[ ]: `netflix['release_year'].hist()`

[ ]: <Axes: >

[ ]: netflix['age_certification'].replace(' ','sill')

[ ]: 0   R
     1   PG
     2   R
     3   R
     4   TV-14
             …
     5278   NaN
     5279   NaN
     5280   NaN
     5281   NaN
     5282   NaN
     Name: age_certification, Length: 5283, dtype: object

[ ]: netflix['age_certification'].value_counts()

[ ]: age_certification
     TV-MA   792
     R   548

```
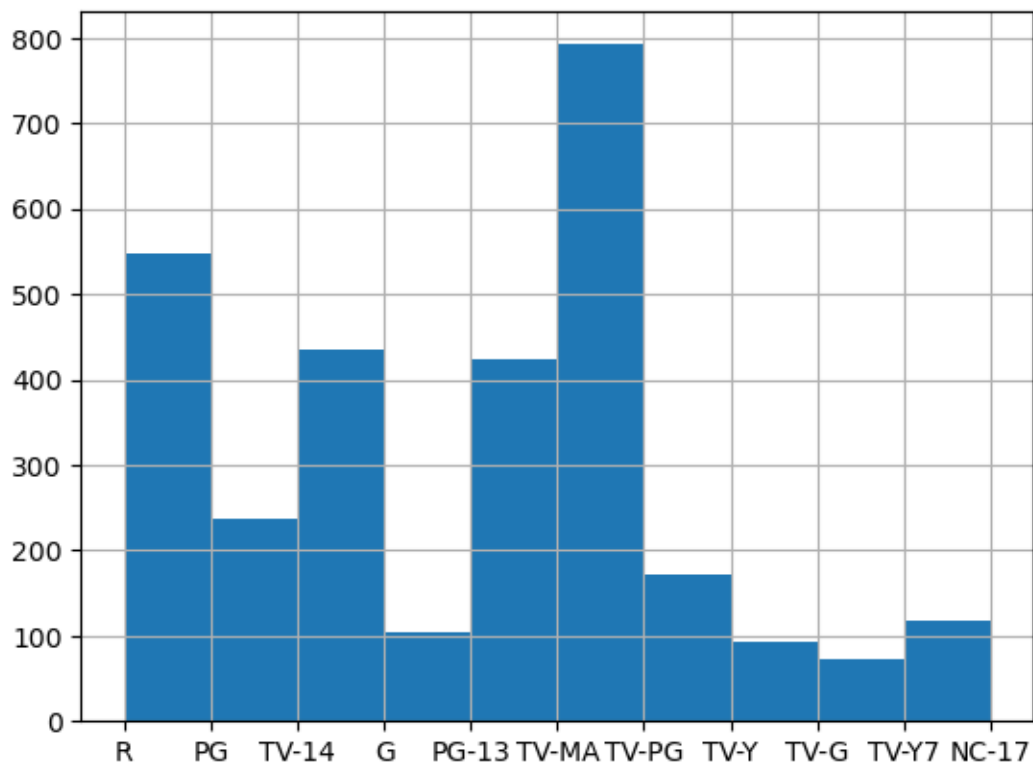TV-14    436
PG-13    424
PG       238
TV-PG    172
G        105
TV-Y7    104
TV-Y      94
TV-G      72
NC-17     13
Name: count, dtype: int64
```

[ ]: `netflix['age_certification'].hist()`

[ ]: `<Axes: >`

[ ]: `netflix['imdb_votes'].sum()`

[ ]: `123285696.0`

[ ]: `votes_per_film = netflix['imdb_votes'].sum() / 5283`
`votes_per_film`

```
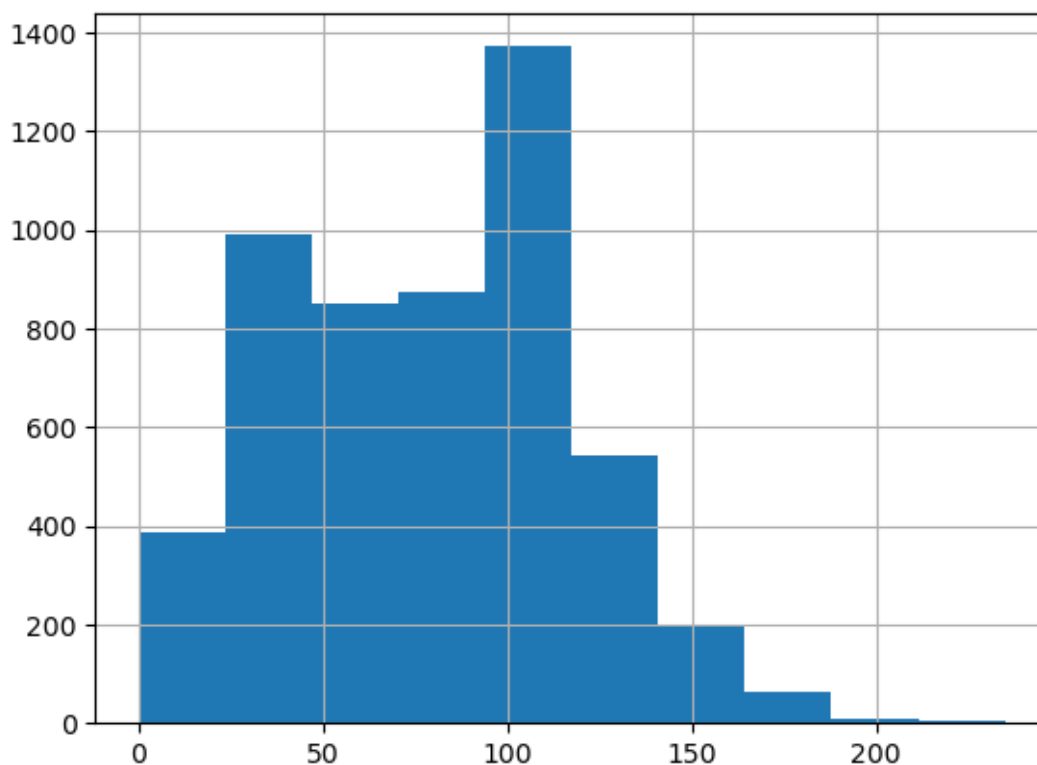[ ]: 23336.304372515617
```

```
[ ]: netflix['runtime'].sum()
```

```
[ ]: 418413
```

```
[ ]: votes_per_run = netflix['imdb_votes'].sum() / netflix['runtime'].sum()
     votes_per_run
```

```
[ ]: 294.6507302593371
```

```
[ ]: netflix['runtime'].hist()
```

```
[ ]: <Axes: >
```



```
[ ]: netflix.groupby('age_certification')[['imdb_score','imdb_votes']].sum()
```

```
[ ]:                    imdb_score  imdb_votes
     age_certification
     G                       671.3    808385.0
     NC-17                    83.2    209965.0
     PG                     1477.6   8512142.0
```

```
PG-13                      2732.5  27889552.0
R                          3464.0  42344416.0
TV-14                      3166.5  10466379.0
TV-G                        457.1    339340.0
TV-MA                      5596.2  19179596.0
TV-PG                      1189.7   1752578.0
TV-Y                        615.5     67809.0
TV-Y7                       719.0    731830.0
```

```
[ ]: netflix.groupby('release_year')[['imdb_score','imdb_votes','runtime']].sum()
```

```
[ ]:              imdb_score  imdb_votes  runtime
     release_year
     1953                6.8       231.0      105
     1954               14.9     43592.0      215
     1956                6.7       590.0      120
     1958                7.5      4385.0       77
     1959                6.7        78.0      142
     ...                 ...         ...      ...
     2018             4786.9   6845715.0    55327
     2019             4888.1   8729219.0    56888
     2020             4168.3   6507700.0    48696
     2021             4341.2   6639992.0    51404
     2022             1139.9   1124575.0    13133

     [66 rows x 3 columns]
```

```
[ ]: netflix.
     ↪groupby(['release_year','age_certification'])[['imdb_score','imdb_votes']].
     ↪sum()
```

```
[ ]:                                 imdb_score  imdb_votes
     release_year age_certification
     1964         G                         7.8     94121.0
     1966         PG-13                     7.3     16168.0
     1967         R                         7.7    111189.0
     1969         TV-14                     8.8     72895.0
     1971         R                         7.7    153463.0
     ...                                    ...         ...
     2022         TV-G                      5.6       176.0
                  TV-MA                   365.0    327191.0
                  TV-PG                    35.6      4732.0
                  TV-Y                     26.4       265.0
                  TV-Y7                    27.8      6436.0

     [288 rows x 2 columns]
```

```
netflix.groupby(['release_year','age_certification'])[['age_certification']].
 ↪count()
```

```
                               age_certification
release_year age_certification
1964         G                                1
1966         PG-13                            1
1967         R                                1
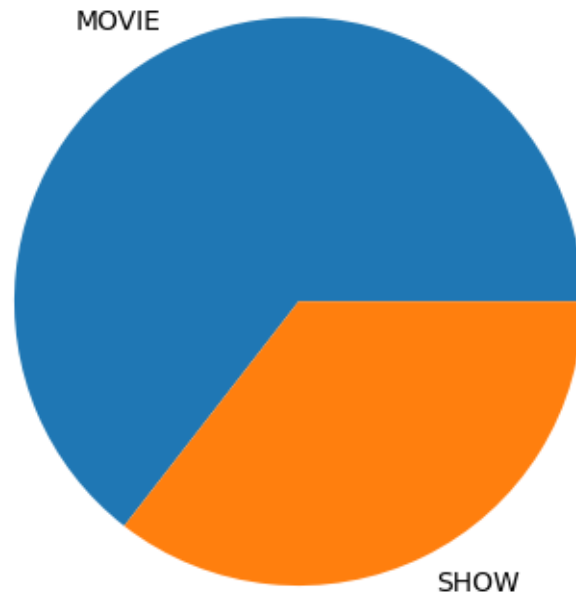1969         TV-14                            1
1971         R                                1
…                                           …
2022         TV-G                             1
             TV-MA                           57
             TV-PG                            6
             TV-Y                             4
             TV-Y7                            4

[288 rows x 1 columns]
```

```
netflix.groupby(['release_year','type'])[['type']].count()
```

```
                  type
release_year type
1953         MOVIE    1
1954         MOVIE    2
1956         MOVIE    1
1958         MOVIE    1
1959         MOVIE    1
…                   …
2020         SHOW   250
2021         MOVIE  411
             SHOW   276
2022         MOVIE   90
             SHOW    92

[106 rows x 1 columns]
```

```
import matplotlib.pyplot  as plt
plt.pie(x= netflix['type'].value_counts(), labels =netflix['type'].unique() )
```

```
([<matplotlib.patches.Wedge at 0x24a18202ed0>,
  <matplotlib.patches.Wedge at 0x24a18563610>],
 [Text(-0.48361873113742904, 0.9879842725939636, 'MOVIE'),
  Text(0.48361863863566334, -0.9879843178736127, 'SHOW')])
```

```
[ ]: import plotly.express as px
     px.histogram( netflix['age_certification'])
```

```
[ ]: px.box(x=netflix['age_certification'], y= netflix['imdb_score'])
```