```
%matplotlib inline
import pandas as pd
import numpy as np
import matplotlib.pylab as plt
import seaborn as sns
from statsmodels import robust


from pathlib import Path
from scipy.stats import trim_mean

anime = pd.read_csv(r"C:\Users\HP\Downloads\Stats project\anime.csv")

print(anime.head(8))
```

```
   anime_id                                                name  \
0     32281                                     Kimi no Na wa.   
1      5114                   Fullmetal Alchemist: Brotherhood   
2     28977                                           Gintama°   
3      9253                                        Steins;Gate   
4      9969                                      Gintama&#039;   
5     32935  Haikyuu!!: Karasuno Koukou VS Shiratorizawa Ga...   
6     11061                             Hunter x Hunter (2011)   
7       820                               Ginga Eiyuu Densetsu   

                                               genre   type episodes  
rating  \
0               Drama, Romance, School, Supernatural  Movie        1  
9.37   
1  Action, Adventure, Drama, Fantasy, Magic, Mili...     TV       64  
9.26   
2  Action, Comedy, Historical, Parody, Samurai, S...     TV       51  
9.25   
3                                   Sci-Fi, Thriller     TV       24  
9.17   
4  Action, Comedy, Historical, Parody, Samurai, S...     TV       51  
9.16   
5             Comedy, Drama, School, Shounen, Sports     TV       10  
9.15   
6            Action, Adventure, Shounen, Super Power     TV      148  
9.13   
7                     Drama, Military, Sci-Fi, Space    OVA      110  
9.11   

   members  
0   200630  
1   793665  
2   114262  
3   673572  
4   151266  
```

```
5    93351
6   425855
7    80679

anime

       anime_id                                               name  \
0         32281                                     Kimi no Na wa.
1          5114                   Fullmetal Alchemist: Brotherhood
2         28977                                           Gintama°
3          9253                                        Steins;Gate
4          9969                                      Gintama&#039;
...         ...                                                ...
12289      9316       Toushindai My Lover: Minami tai Mecha-Minami
12290      5543                                        Under World
12291      5621                     Violence Gekiga David no Hoshi
12292      6133  Violence Gekiga Shin David no Hoshi: Inma Dens...
12293     26081                   Yasuji no Pornorama: Yacchimae!!

                                                   genre   type
episodes  \
0                   Drama, Romance, School, Supernatural  Movie
1
1      Action, Adventure, Drama, Fantasy, Magic, Mili...     TV
64
2      Action, Comedy, Historical, Parody, Samurai, S...     TV
51
3                                       Sci-Fi, Thriller     TV
24
4      Action, Comedy, Historical, Parody, Samurai, S...     TV
51
...                                                  ...    ...      .
..
12289                                             Hentai    OVA
1
12290                                             Hentai    OVA
1
12291                                             Hentai    OVA
4
12292                                             Hentai    OVA
1
12293                                             Hentai  Movie
1

       rating  members
0        9.37   200630
1        9.26   793665
2        9.25   114262
3        9.17   673572
4        9.16   151266
```

```
...       ...      ...
12289    4.15      211
12290    4.28      183
12291    4.88      219
12292    4.98      175
12293    5.46      142

[12294 rows x 7 columns]
```

Estimates of Location

```python
print("Mean rating =")
print(anime['rating'].mean())
```

```
Mean rating =
6.473901690981445
```

```python
print("median of  rating =")
print(anime['rating'].median())
```

```
median of  rating =
6.57
```

```python
print("mode of  episodes =")
print(anime['episodes'].mode())
```

```
mode of  episodes =
0    1
Name: episodes, dtype: object
```

```python
print(anime.head(10))
```

```
   anime_id                                               name  \
0     32281                                     Kimi no Na wa.
1      5114                   Fullmetal Alchemist: Brotherhood
2     28977                                           Gintama°
3      9253                                        Steins;Gate
4      9969                                      Gintama&#039;
5     32935  Haikyuu!!: Karasuno Koukou VS Shiratorizawa Ga...
6     11061                             Hunter x Hunter (2011)
7       820                               Ginga Eiyuu Densetsu
8     15335  Gintama Movie: Kanketsu-hen - Yorozuya yo Eien...
9     15417                           Gintama&#039;: Enchousen

                                               genre   type episodes
rating  \
0                Drama, Romance, School, Supernatural  Movie        1
9.37
1  Action, Adventure, Drama, Fantasy, Magic, Mili...     TV       64
```

```
9.26
2  Action, Comedy, Historical, Parody, Samurai, S...     TV       51
9.25
3                                   Sci-Fi, Thriller     TV       24
9.17
4  Action, Comedy, Historical, Parody, Samurai, S...     TV       51
9.16
5             Comedy, Drama, School, Shounen, Sports     TV       10
9.15
6            Action, Adventure, Shounen, Super Power     TV      148
9.13
7                     Drama, Military, Sci-Fi, Space    OVA      110
9.11
8  Action, Comedy, Historical, Parody, Samurai, S...  Movie        1
9.10
9  Action, Comedy, Historical, Parody, Samurai, S...     TV       13
9.11
```

```
   members
0   200630
1   793665
2   114262
3   673572
4   151266
5    93351
6   425855
7    80679
8    72534
9    81109
```

```python
print(trim_mean(anime['members'],0.1))
```

```
5589.94286295242
```

```python
sample_anime = anime.head(100)
```

```python
print(np.average(sample_anime['rating'],
weights=sample_anime['members']))
```

```
8.754444415324095
```

# Estimates of Variablity

```python
print(anime['members'].std())
```

```
54820.676924907515
```

```python
#Interquartile range is calculated as the difference of the 75% and
25% quantile.
```

```
print(anime['members'].quantile(0.75) -
anime['members'].quantile(0.25))
```

```
9212.0
```

```
#Median absolute deviation from the median
#method- 1
#print(abs(anime['members'] - anime['members'].median()).median() /
0.6744897501960817)
#method-2
print(robust.scale.mad(anime['members']))
```

```
2172.0122501107066
```

Estimates on Percentiles

```
print(anime['members'].quantile([0.05, 0.25, 0.5, 0.75, 0.95]))
```

```
0.05       58.0
0.25      225.0
0.50     1550.0
0.75     9437.0
0.95    93164.3
Name: members, dtype: float64
```

```
percentages = [0.05, 0.25, 0.5, 0.75, 0.95]
df = pd.DataFrame(anime['members'].quantile(percentages))
df.index = [f'{p * 100}%' for p in percentages]
print(df.transpose())
```

```
         5.0%  25.0%   50.0%   75.0%    95.0%
members  58.0  225.0  1550.0  9437.0  93164.3
```

Explore data distribution

Frequency table

```
animax = anime.head(2000)
binnedmembers = pd.cut(animax['members'], 20)
print(binnedmembers.value_counts())
```

```
(-644.548, 51046.4]      1190
(51046.4, 101723.8]       356
(101723.8, 152401.2]      156
(152401.2, 203078.6]       96
(203078.6, 253756.0]       66
(253756.0, 304433.4]       37
(304433.4, 355110.8]       33
(355110.8, 405788.2]       16
(405788.2, 456465.6]       11
(456465.6, 507143.0]       10
(557820.4, 608497.8]        8
(507143.0, 557820.4]        8
(608497.8, 659175.2]        5
(659175.2, 709852.6]        2
(709852.6, 760530.0]        2
(861884.8, 912562.2]        2
(760530.0, 811207.4]        1
(963239.6, 1013917.0]       1
(811207.4, 861884.8]        0
(912562.2, 963239.6]        0
Name: members, dtype: int64
```

```
binnedmembers.name = 'binnedmembers'
df = pd.concat([animax, binnedmembers], axis=1)
df = df.sort_values(by='members')

groups = []
for group, subset in df.groupby(by='binnedmembers'):
    groups.append({
        'BinRange': group,
        'Count': len(subset),
        'Genre': ','.join(subset.genre)
    })
print(pd.DataFrame(groups))
```

```
                 BinRange  Count  \
0     (-644.548, 51046.4]   1190
1     (51046.4, 101723.8]    356
2    (101723.8, 152401.2]    156
3    (152401.2, 203078.6]     96
4    (203078.6, 253756.0]     66
5    (253756.0, 304433.4]     37
6    (304433.4, 355110.8]     33
7    (355110.8, 405788.2]     16
8    (405788.2, 456465.6]     11
9    (456465.6, 507143.0]     10
10   (507143.0, 557820.4]      8
11   (557820.4, 608497.8]      8
12   (608497.8, 659175.2]      5
13   (659175.2, 709852.6]      2
14   (709852.6, 760530.0]      2
15   (760530.0, 811207.4]      1
16   (811207.4, 861884.8]      0
17   (861884.8, 912562.2]      2
18   (912562.2, 963239.6]      0
19  (963239.6, 1013917.0]      1
                                                Genre
0   Action, Fantasy, Historical, Martial Arts,Dram...
1   Action, Adventure, Samurai,Romance, School, Sh...
2   Comedy, Drama, Mystery, Romance, Slice of Life...
3   Action, Adventure, Drama, Fantasy, Historical,...
4   Action, Seinen,Action, Fantasy, Magic, Romance...
5   Comedy, School, Slice of Life,Action, Adventur...
6   Action, Comedy, Dementia, Mecha, Parody, Sci-F...
7   Drama, Fantasy, Psychological, Thriller,Fantas...
8   Action, Drama, Horror, Mystery, Psychological,...
9   Drama, Fantasy, Romance, Slice of Life, Supern...
10  Action, Comedy, School, Super Power,Action, Po...
11  Action, Adventure, Comedy, Mecha, Sci-Fi,Comed...
12  Action, Drama, Horror, Mystery, Psychological,...
13  Sci-Fi, Thriller,Action, Comedy, Martial Arts,...
14  Action, Mecha, Military, School, Sci-Fi, Super...
15  Action, Adventure, Drama, Fantasy, Magic, Mili...
16
17  Action, Adventure, Fantasy, Game, Romance,Acti...
18
19  Mystery, Police, Psychological, Supernatural, ...
```

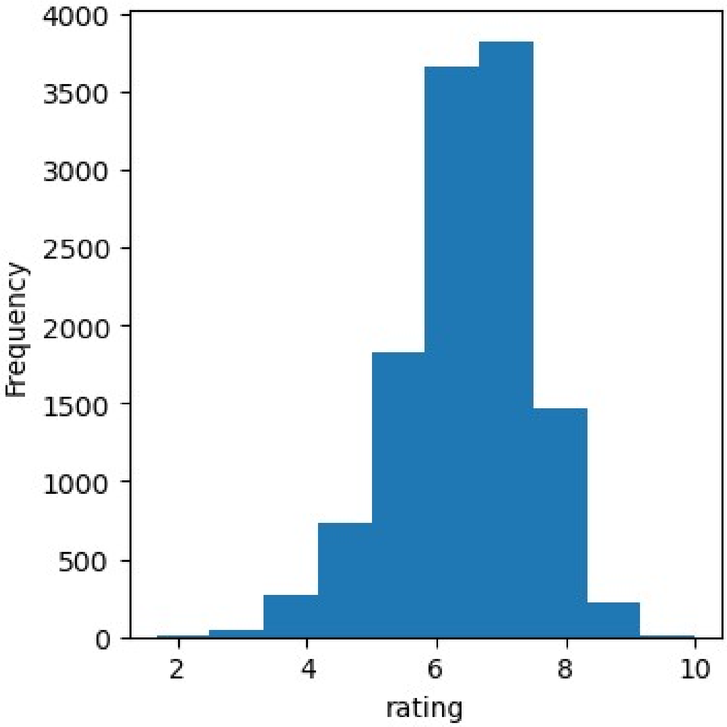
```
print(pd.DataFrame(groups))
```

```
                 BinRange  Count  \
0     (-644.548, 51046.4]   1190
1     (51046.4, 101723.8]    356
2    (101723.8, 152401.2]    156
3    (152401.2, 203078.6]     96
4    (203078.6, 253756.0]     66
5    (253756.0, 304433.4]     37
6    (304433.4, 355110.8]     33
7    (355110.8, 405788.2]     16
8    (405788.2, 456465.6]     11
9    (456465.6, 507143.0]     10
10   (507143.0, 557820.4]      8
11   (557820.4, 608497.8]      8
12   (608497.8, 659175.2]      5
13   (659175.2, 709852.6]      2
14   (709852.6, 760530.0]      2
15   (760530.0, 811207.4]      1
16   (811207.4, 861884.8]      0
17   (861884.8, 912562.2]      2
18   (912562.2, 963239.6]      0
19  (963239.6, 1013917.0]      1

                                                Genre
0   Action, Fantasy, Historical, Martial Arts,Dram...
1   Action, Adventure, Samurai,Romance, School, Sh...
```

```
2    Comedy, Drama, Mystery, Romance, Slice of Life...
3    Action, Adventure, Drama, Fantasy, Historical,...
4    Action, Seinen,Action, Fantasy, Magic, Romance...
5    Comedy, School, Slice of Life,Action, Adventur...
6    Action, Comedy, Dementia, Mecha, Parody, Sci-F...
7    Drama, Fantasy, Psychological, Thriller,Fantas...
8    Action, Drama, Horror, Mystery, Psychological,...
9    Drama, Fantasy, Romance, Slice of Life, Supern...
10   Action, Comedy, School, Super Power,Action, Po...
11   Action, Adventure, Comedy, Mecha, Sci-Fi,Comed...
12   Action, Drama, Horror, Mystery, Psychological,...
13   Sci-Fi, Thriller,Action, Comedy, Martial Arts,...
14   Action, Mecha, Military, School, Sci-Fi, Super...
15   Action, Adventure, Drama, Fantasy, Magic, Mili...
16
17   Action, Adventure, Fantasy, Game, Romance,Acti...
18
19   Mystery, Police, Psychological, Supernatural, ...
```

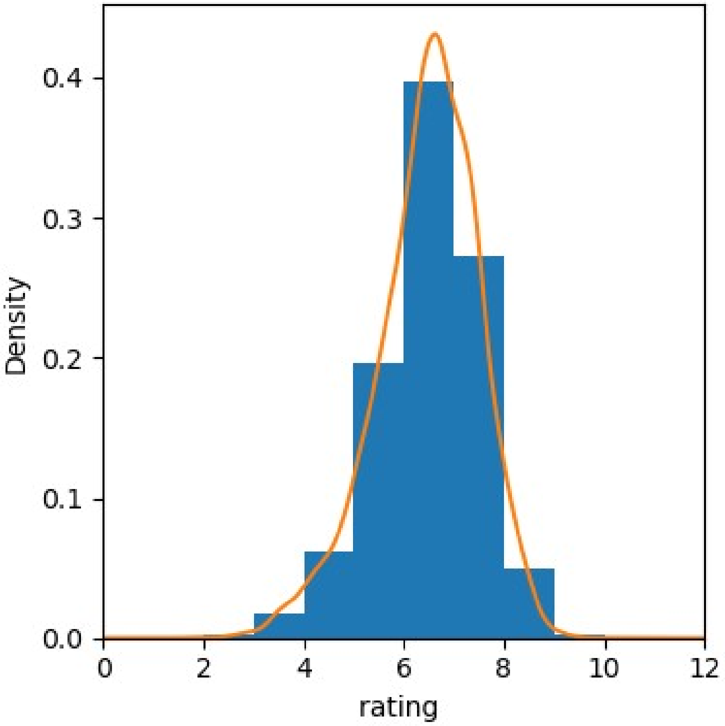
Histograms

```
ax = (anime['rating'] ).plot.hist(figsize=(4, 4))
ax.set_xlabel('rating')

plt.tight_layout()
plt.show()
```



density plot

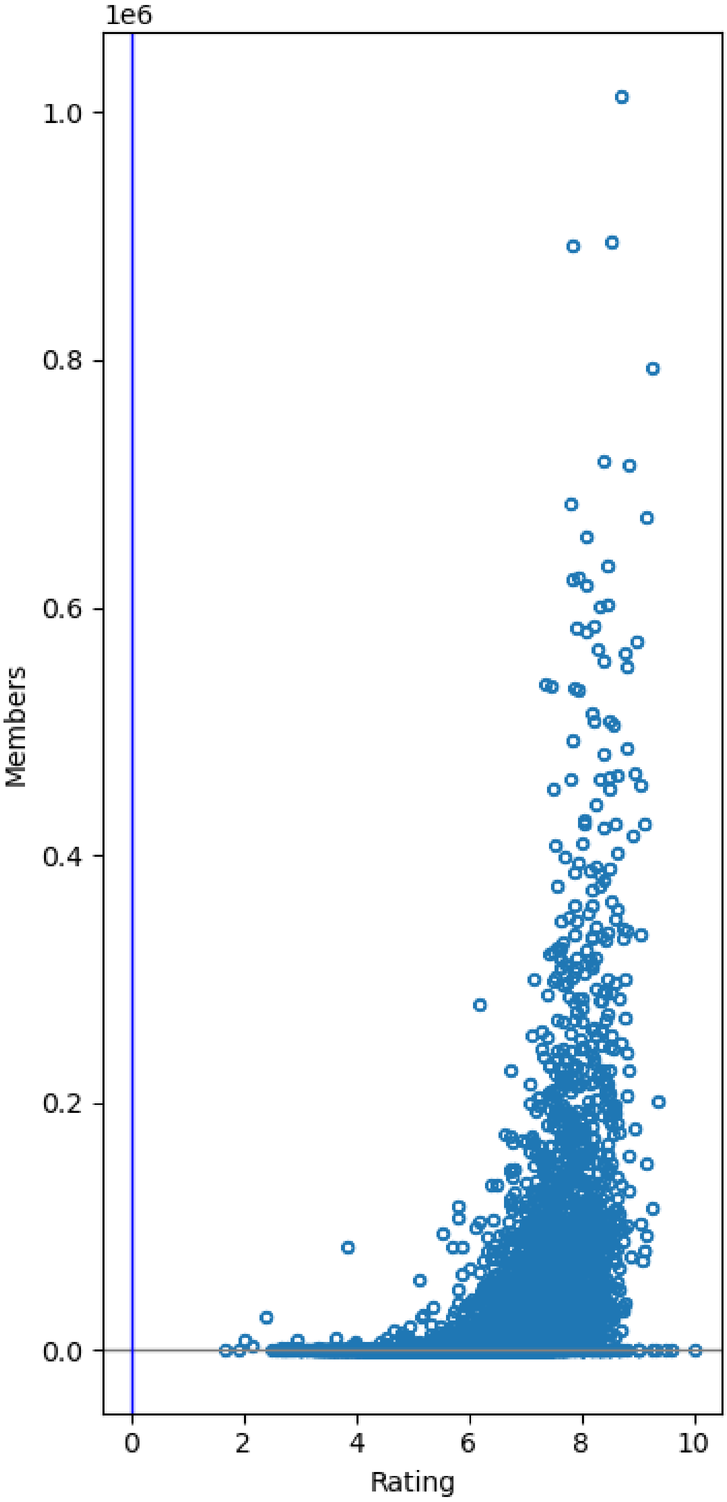
```
ax = anime['rating'].plot.hist(density=True, xlim=[0, 12],
                                    bins=range(1,12), figsize=(4, 4))
anime['rating'].plot.density(ax=ax)
ax.set_xlabel('rating ')

plt.tight_layout()
plt.show()
```



Exploring Binary and Categorical Data

Scatterplots

```

ax = anime.plot.scatter(x='rating', y='members', figsize=(4, 8),
marker='$\u25EF$')
ax.set_xlabel('Rating')
ax.set_ylabel('Members')
ax.axhline(0, color='grey', lw=1)
ax.axvline(0, color='blue', lw=1)

plt.tight_layout()
plt.show()
```

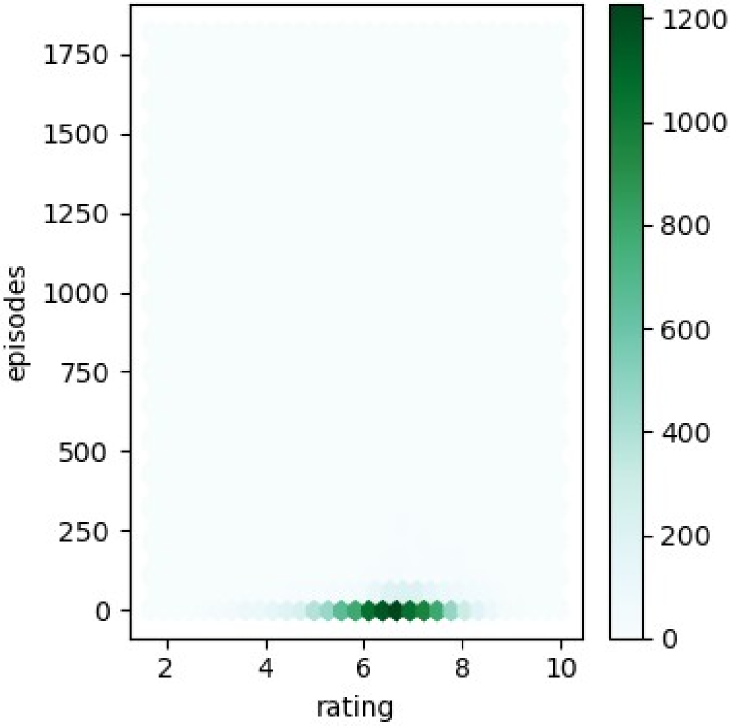


Binning

```
animes = anime.drop(anime[anime.episodes == 'Unknown'].index)
print(animes)
```

```
       anime_id                                               name  \
0         32281                                     Kimi no Na wa.
1          5114                   Fullmetal Alchemist: Brotherhood
2         28977                                           Gintama°
3          9253                                        Steins;Gate
4          9969                                      Gintama&#039;
...         ...                                                ...
12289      9316       Toushindai My Lover: Minami tai Mecha-Minami
12290      5543                                        Under World
12291      5621                     Violence Gekiga David no Hoshi
12292      6133  Violence Gekiga Shin David no Hoshi: Inma Dens...
12293     26081                   Yasuji no Pornorama: Yacchimae!!

                                                   genre   type
episodes  \
0                   Drama, Romance, School, Supernatural  Movie
1
1      Action, Adventure, Drama, Fantasy, Magic, Mili...     TV
64
2      Action, Comedy, Historical, Parody, Samurai, S...     TV
51
3                                       Sci-Fi, Thriller     TV
24
4      Action, Comedy, Historical, Parody, Samurai, S...     TV
51
...                                                  ...    ...      .
..
12289                                             Hentai    OVA
1
12290                                             Hentai    OVA
1
12291                                             Hentai    OVA
4
12292                                             Hentai    OVA
1
12293                                             Hentai  Movie
1

       rating  members
0        9.37   200630
1        9.26   793665
2        9.25   114262
3        9.17   673572
4        9.16   151266
...       ...      ...
12289    4.15      211
12290    4.28      183
12291    4.88      219
12292    4.98      175
12293    5.46      142
```

```
[11954 rows x 7 columns]

animes = animes.astype({'episodes':'float'})
ax = animes.plot.hexbin(x='rating', y='episodes',
                        gridsize=30, sharex=False, figsize=(4, 4))
ax.set_xlabel('rating')
ax.set_ylabel('episodes')

plt.tight_layout()
plt.show()
```

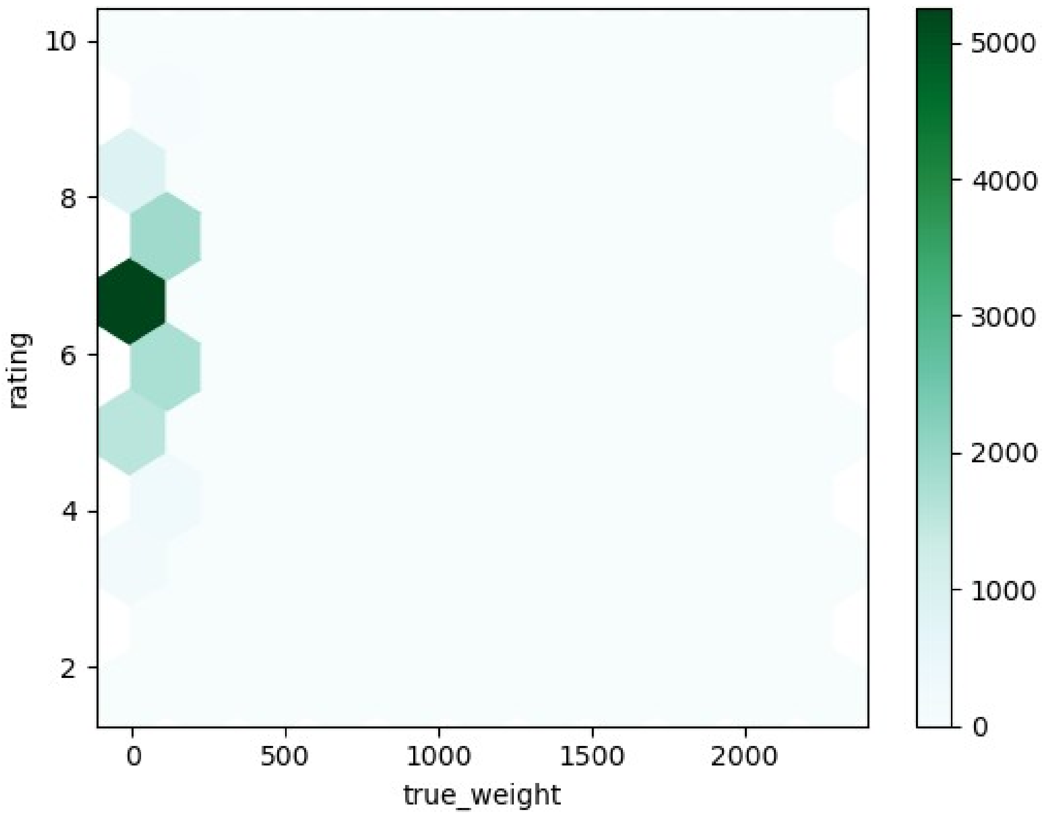


```
animes['true_weight'] = animes['episodes'] * animes['members'] /
100000
animes.head(10)

   anime_id                                              name  \
0     32281                                    Kimi no Na wa.
1      5114                  Fullmetal Alchemist: Brotherhood
2     28977                                          Gintama°
3      9253                                       Steins;Gate
4      9969                                     Gintama&#039;
5     32935  Haikyuu!!: Karasuno Koukou VS Shiratorizawa Ga...
6     11061                               Hunter x Hunter (2011)
7       820                              Ginga Eiyuu Densetsu
8     15335  Gintama Movie: Kanketsu-hen - Yorozuya yo Eien...
```

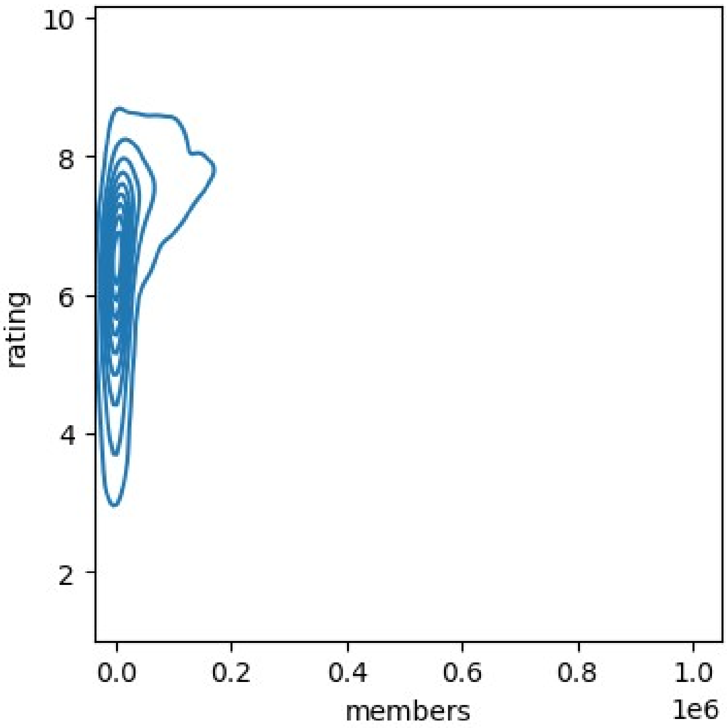
```
9     15417                           Gintama&#039;: Enchousen

                                               genre   type  episodes
rating  \
0                Drama, Romance, School, Supernatural  Movie       1.0
9.37
1  Action, Adventure, Drama, Fantasy, Magic, Mili...     TV      64.0
9.26
2  Action, Comedy, Historical, Parody, Samurai, S...     TV      51.0
9.25
3                                   Sci-Fi, Thriller     TV      24.0
9.17
4  Action, Comedy, Historical, Parody, Samurai, S...     TV      51.0
9.16
5              Comedy, Drama, School, Shounen, Sports     TV      10.0
9.15
6             Action, Adventure, Shounen, Super Power     TV     148.0
9.13
7                      Drama, Military, Sci-Fi, Space    OVA     110.0
9.11
8  Action, Comedy, Historical, Parody, Samurai, S...  Movie       1.0
9.10
9  Action, Comedy, Historical, Parody, Samurai, S...     TV      13.0
9.11

   members  true_weight
0   200630      2.00630
1   793665    507.94560
2   114262     58.27362
3   673572    161.65728
4   151266     77.14566
5    93351      9.33510
6   425855    630.26540
7    80679     88.74690
8    72534      0.72534
9    81109     10.54417
```

```
animes.plot(kind='hexbin',x='true_weight', y = 'rating' , gridsize =
10)
```

```
<AxesSubplot:xlabel='true_weight', ylabel='rating'>
```

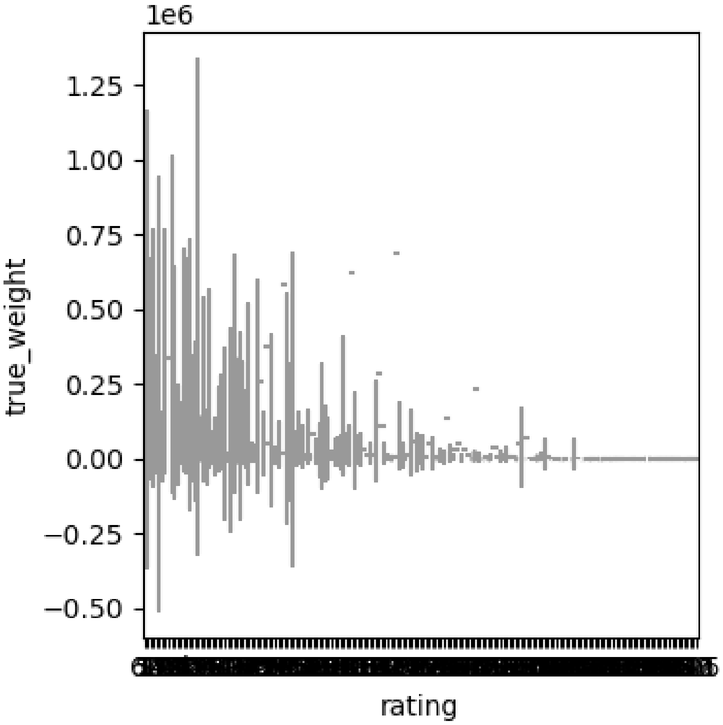


```
fig, ax = plt.subplots(figsize=(4, 4))
sns.kdeplot(data=animes.sample(10000), x='members', y='rating', ax=ax)
ax.set_xlabel('members')
ax.set_ylabel('rating')

plt.tight_layout()
plt.show()
```



Violen plot

```
fig, ax = plt.subplots(figsize=(4, 4))
sns.violinplot(data=anime, x='episodes', y='members',
               ax=ax, inner='quartile', color='white')
ax.set_xlabel('rating')
ax.set_ylabel('true_weight')

plt.tight_layout()
plt.show()
```


1e6
1.25
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
true_weight
rating