```
import numpy as np
import pandas as pd
movies = pd.read_csv("/content/movies.dat", delimiter='::')
print(movies.head())
```

```
   0000008      Edison Kinetoscopic Record of a Sneeze (1894)  \
0       10                 La sortie des usines Lumière (1895)
1       12                       The Arrival of a Train (1896)
2       25  The Oxford and Cambridge University Boat Race ...
3       91                           Le manoir du diable (1896)
4      131                             Une nuit terrible (1896)

     Documentary|Short
0    Documentary|Short
1    Documentary|Short
2                  NaN
3          Short|Horror
4  Short|Comedy|Horror
<ipython-input-4-f439e3607e4a>:3: ParserWarning: Falling back to the 'python' engine because the 'c' engine does not support regex s
  movies = pd.read_csv("/content/movies.dat", delimiter='::')
```

```
movies.columns = ["ID", "Title", "Genre"]
print(movies.head())
```

```
     ID                                          Title  \
0   10                 La sortie des usines Lumière (1895)
1   12                       The Arrival of a Train (1896)
2   25  The Oxford and Cambridge University Boat Race ...
3   91                           Le manoir du diable (1896)
4  131                             Une nuit terrible (1896)
5  417                             A Trip to the Moon (1902)
6  439                       The Great Train Robbery (1903)
7  443          Hiawatha, the Messiah of the Ojibway (1903)
8  628                      The Adventures of Dollie (1908)
9  833                           The Country Doctor (1909)

                                    Genre
0                     Documentary|Short
1                     Documentary|Short
2                                   NaN
3                         Short|Horror
4                    Short|Comedy|Horror
5  Short|Action|Adventure|Comedy|Fantasy|Sci-Fi
6              Short|Action|Crime|Western
7                                   NaN
8                          Action|Short
9                           Short|Drama
```

```
ratings = pd.read_csv("/content/ratings.dat", delimiter='::')
print(ratings.head())
```

```
<ipython-input-7-4336c02674ee>:1: ParserWarning: Falling back to the 'python' engine because the 'c' engine does not support regex s
  ratings = pd.read_csv("/content/ratings.dat", delimiter='::')
   1  0114508  8  1381006850
0  2   499549  9  1376753198
1  2  1305591  8  1376742507
2  2  1428538  1  1371307089
3  3    75314  1  1595468524
4  3   102926  9  1590148016
```

```
ratings.columns = ["User", "ID", "Ratings", "Timestamp"]
print(ratings.head())
```

```
8     219311
7     203476
9     128749
6     118323
10    107284
5      68458
4      27779
3      15258
1      10663
2       9053
Name: Ratings, dtype: int64
```

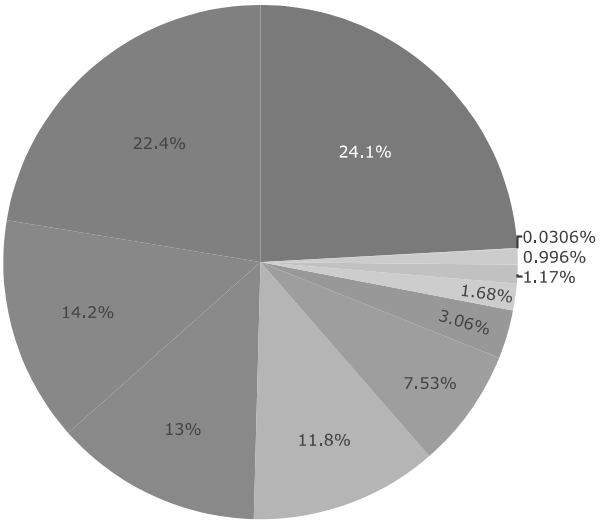
```
data = pd.merge(movies, ratings, on=["ID", "ID"])
print(data.head())
```

```
   ID                                          Title              Genre  \
0  10                 La sortie des usines Lumière (1895)  Documentary|Short
1  12                       The Arrival of a Train (1896)  Documentary|Short
2  25  The Oxford and Cambridge University Boat Race ...                NaN
```

```
3  91                      Le manoir du diable (1896)       Short|Horror
4  91                      Le manoir du diable (1896)       Short|Horror

    User  Ratings   Timestamp
0  70577       10  1412878553
1  69535       10  1439248579
2  37628        8  1488189899
3   5814        6  1385233195
4  37239        5  1532347349
```

```
ratings = data["Ratings"].value_counts()
numbers = ratings.index
quantity = ratings.values
import plotly.express as px
fig = px.pie(data, values=quantity, names=numbers)
fig.show()
```




```
#data2 = data.query("Ratings == 10")
#data2
print(data["Title"].value_counts().head(10))
```

```
Gravity (2013)                    3104
Interstellar (2014)               2948
1917 (2019)                       2879
The Wolf of Wall Street (2013)    2836
Joker (2019)                      2753
Man of Steel (2013)               2694
World War Z (2013)                2429
Iron Man Three (2013)             2417
Now You See Me (2013)             2379
Gone Girl (2014)                  2284
Name: Title, dtype: int64
```