```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

import pandas_profiling
import warnings
warnings.filterwarnings('ignore')

rating_header = "UserID::MovieID::Rating::Timestamp".split("::")
user_header= "UserID::Gender::Age::Occupation::Zip-code".split("::")
movies_header = "MovieID::Title::Genres".split("::")

print(rating_header)
print(user_header)
print(movies_header)
```

```
['UserID', 'MovieID', 'Rating', 'Timestamp']
['UserID', 'Gender', 'Age', 'Occupation', 'Zip-code']
['MovieID', 'Title', 'Genres']
```

In [4]:

```
movies =
pd.read_csv("/Users/dakshgoel/Downloads/movies.dat",sep='::',names=movies_header)
ratings = pd.read_csv("/Users/dakshgoel/Downloads/ratings.dat",sep='::',names=rating_header)
user = pd.read_csv("/Users/dakshgoel/Downloads/users.dat",sep='::',names=user_header)


ratings = pd.read_csv("/Users/dakshgoel/Downloads/ratings.dat",sep='::',names=rating_header)

user = pd.read_csv("/Users/dakshgoel/Downloads/users.dat",sep='::',names=user_header)

movie_ratings = pd.merge(movies,ratings, on ="MovieID" )

movie_ratings
```

	MovieID	Title	Genres	UserID	Rating	Timestamp
0	1	Toy Story (1995)	Animation|Children's|Comedy	1	5	978824268

1	1	Toy Story (1995)	Animation|Children's|Comedy	6	4	978237008
2	1	Toy Story (1995)	Animation|Children's|Comedy	8	4	978233496
3	1	Toy Story (1995)	Animation|Children's|Comedy	9	5	978225952
4	1	Toy Story (1995)	Animation|Children's|Comedy	10	5	978226474
...	...	...	...	...	...	...
1000204	3952	Contender, The (2000)	Drama|Thriller	5812	4	992072099
1000205	3952	Contender, The (2000)	Drama|Thriller	5831	3	986223125
1000206	3952	Contender, The (2000)	Drama|Thriller	5837	4	1011902656
1000207	3952	Contender, The (2000)	Drama|Thriller	5927	1	979852537
1000208	3952	Contender, The (2000)	Drama|Thriller	5998	4	1001781044

```
movie_user_ratings=pd.merge(movie_ratings,user, on ="UserID" )
```

```
movie_user_ratings
```

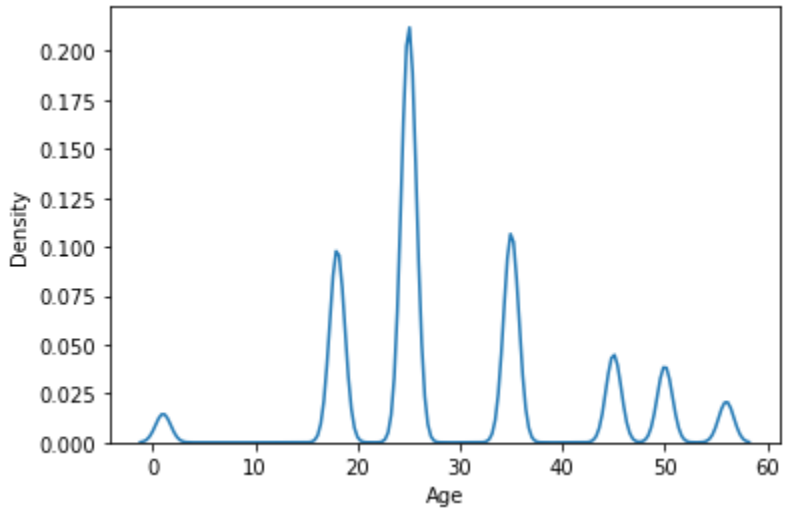
```
movie_user_ratings.isnull().sum()
```

```
MovieID       0
Title         0
Genres        0
UserID        0
Rating        0
Timestamp     0
Gender        0
```

```
Age           0
Occupation    0
Zip-code      0
dtype: int64
```

```
sns.distplot(movie_user_ratings["Age"],hist=False)
```

```
<AxesSubplot:xlabel='Age', ylabel='Density'>
```




```
movie_user_ratings[movie_user_ratings.UserID==2696]["Title"]
```

991035	Client, The (1994)
991036	Lone Star (1996)
991037	Basic Instinct (1992)
991038	E.T. the Extra-Terrestrial (1982)
991039	Shining, The (1980)
991040	Back to the Future (1985)
991041	Cop Land (1997)
991042	L.A. Confidential (1997)
991043	Game, The (1997)
991044	I Know What You Did Last Summer (1997)
991045	Devil's Advocate, The (1997)
991046	Midnight in the Garden of Good and Evil (1997)
991047	Palmetto (1998)
991048	Wild Things (1998)
991049	Perfect Murder, A (1998)
991050	I Still Know What You Did Last Summer (1998)
991051	Psycho (1998)
991052	Lake Pla['Client, The (1994)',

In [54]:

```
list(movie_user_ratings[movie_user_ratings.UserID==2696]["Title"])
```

```
 'Lone Star (1996)',
 'Basic Instinct (1992)',
 'E.T. the Extra-Terrestrial (1982)',
 'Shining, The (1980)',
 'Back to the Future (1985)',
 'Cop Land (1997)',
 'L.A. Confidential (1997)',
 'Game, The (1997)',
 'I Know What You Did Last Summer (1997)',
 "Devil's Advocate, The (1997)",
 'Midnight in the Garden of Good and Evil (1997)',
 'Palmetto (1998)',
 'Wild Things (1998)',
 'Perfect Murder, A (1998)',
 'I Still Know What You Did Last Summer (1998)',
 'Psycho (1998)',
 'Lake Placid (1999)',
 'Talented Mr. Ripley, The (1999)',
 'JFK (1991)']
```

In [ ]:

```
cid (1999)
991053                Talented Mr. Ripley, The (1999)
991054                                    JFK (1991)
Name: Title, dtype: object
```

```
movie_user_ratings[movie_user_ratings.Title=="Toy Story (1995)"]["Rating"]
0         5
53        4
124       4
263       5
369       5
         ..
575166    5
575214    5
575485    4
575589    4
575869    3
Name: Rating, Length: 2077, dtype: int64
```

```
movie_user_ratings[movie_user_ratings.Rating==5]["MovieID"].head(25)
```

```
0       1
1      48
2     150
4     527
8     595
17   1022
18   1028
19   1029
20   1035
22   1193
26   1270
27   1287
31   1836
33   1961
36   2028
40   2355
47   2804
49   3105
56     48
57    199
68    597
71    914
76   1035
78   1088
83   1380
Name: MovieID, dtype: int64
```

In [ ]:

We have to go for classification model.