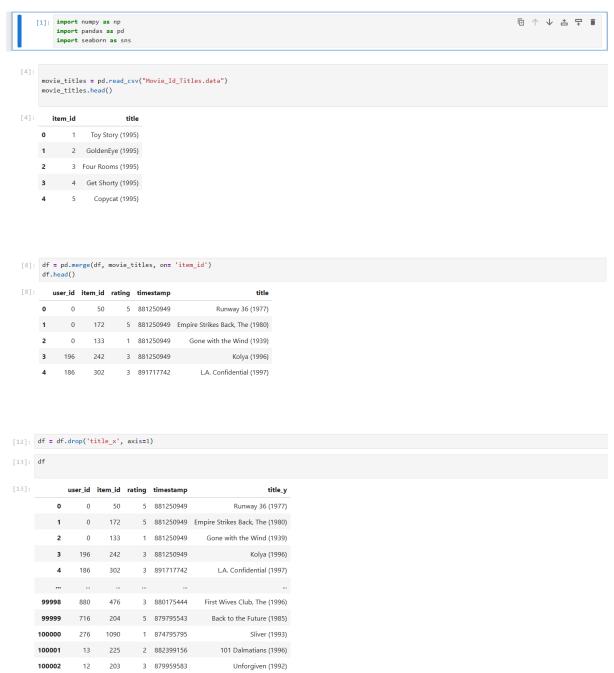
## Shubham Shashikant Bawane **Film Recommendation System** using numpy, pandas, seaborn

## Language python



```
[2]: columns_names = ['user_id', 'item_id', 'rating','timestamp'] df = pd.read_csv('u.data',sep = '\t', names= columns_names)
   [3]: df.head()
              user_id item_id rating timestamp
                0
                          50
                                     5 881250949
          1 0 172 5 881250949
          2 0 133 1 881250949
          3 196 242 3 881250949
              186 302
                                  3 891717742
 [14]: import matplotlib.pyplot as plt
         import seaborn as sns
sns.set_style('white')
         %matplotlib inline
 [15]: df.groupby('title_y')['rating'].mean().sort_values(ascending=False).head()
 [15]: title_y
         title_y
Aiqing wansui (1994)
Entertaining Angels: The Dorothy Day Story (1996)
Saint of Fort Washington, The (1993)
They Made Me a Criminal (1939)
Prefontaine (1997)
                                                                          5.0
5.0
5.0
5.0
         Name: rating, dtype: float64
    [16]: df.groupby('title_y')['rating'].count().sort_values(ascending=False).head()
    [16]: title_y
Runway 36 (1977)
Contact (1997)
                                               584
509
508
            Fargo (1996)
Return of the Jedi (1983)
Liar Liar (1997)
Name: rating, dtype: int64
                                               507
485
[18]: ratings['num of ratings'] = pd.DataFrame(df.groupby('title_y')['rating'].count())
        ratings
[18]:
                                                       rating num of ratings
                                            title_y
                        'Til There Was You (1997) 2.333333
                                   1-900 (1994) 2.600000
                          101 Dalmatians (1996) 2.908257
                                                                             109
```

**12 Angry Men (1957)** 4.344000

 Zeus and Roxanne (1997)
 2.166667

 unknown
 3.444444

 Á köldum klaka (Cold Fever) (1994)
 3.000000

Young Poisoner's Handbook, The (1995) 3.341463

**187 (1997)** 3.024390

... Young Guns II (1990) 2.772727

125

41

44 41

```
[17]: ratings = pd.DataFrame(df.groupby('title_y')['rating'].mean())
ratings

title_y

'Til There Was You (1997) 2.333333

1-900 (1994) 2.600000

101 Dalmatians (1996) 2.908257

12 Angry Men (1957) 4.344000

187 (1997) 3.024390

Young Guns II (1990) 2.772727

Young Poisoner's Handbook, The (1995) 3.341463

Zeus and Roxanne (1997) 2.166667

unknown 3.444444

Å köldum klaka (Cold Fever) (1994) 3.000000
```

```
[27]: sns.scatterplot(data = ratings , x = "num of ratings", y = "rating")
```



