In [1]:
```python
import pandas as pd
import numpy as np
```

In [2]:
```python
data=pd.read_csv("C:\\Users\\Admin\\Downloads\\Assignment\\book A.csv",encoding='ISO-88
data
```

Out[2]:

	Unnamed: 0	UserID	BookTitle	BookRating
0	1	276726	Classical Mythology	5
1	2	276729	Clara Callan	3
2	3	276729	Decision in Normandy	6
3	4	276736	Flu: The Story of the Great Influenza Pandemic...	8
4	5	276737	The Mummies of Urumchi	6
...	...	...	...	...
9995	9996	162121	American Fried: Adventures of a Happy Eater.	7
9996	9997	162121	Cannibal In Manhattan	9
9997	9998	162121	How to Flirt: A Practical Guide	7
9998	9999	162121	Twilight	8
9999	10000	162129	Kids Say the Darndest Things	6

10000 rows × 4 columns

In [3]:
```python
data.sort_values('UserID')
```

Out[3]:

	Unnamed: 0	UserID	BookTitle	BookRating
2401	2402	8	Wings	5
2400	2401	8	The Western way: A practical guide to the West...	5
2399	2400	8	Ancient Celtic Romances	5
2402	2403	8	Truckers	5
2405	2406	8	The Art Of Celtia	7
...	...	...	...	...
2395	2396	278854	La crÃ³nica del PerÃº (CrÃ³nicas de AmÃ©rica)	7
2398	2399	278854	Celtic Mythology (Library of the World's Myths...	8
2393	2394	278854	A corrente de Trewis Scott	7
2394	2395	278854	As valkÃrias	7
2397	2398	278854	A Treasury of Irish Myth, Legend, and Folklore	6

10000 rows × 4 columns

In [4]:
```python
len(data.UserID.unique())
```

Out[4]: 2182

In [6]:
```python
data['BookRating'].value_counts()
```

Out[6]:
```
8     2283
7     2076
10    1732
9     1493
5     1007
6      920
4      237
3      146
2       63
1       43
Name: BookRating, dtype: int64
```

In [7]:
```python
len(data.BookTitle.unique())
```

Out[7]: 9659

In [8]:
```python
data['BookTitle'].value_counts()
```

Out[8]:
```
Fahrenheit 451                                            5
Charlie and the Chocolate Factory                        4
The Subtle Knife (His Dark Materials, Book 2)            4
Vanished                                                 4
Ender's Game (Ender Wiggins Saga (Paperback))            4
                                                        ..
Murder on St. Mark's Place (Gaslight Mysteries)          1
State of Grace                                           1
Valsalva's Maneuver: Mots Justes and Indispensable Terms 1
I love you, I hate you                                   1
Kids Say the Darndest Things                             1
Name: BookTitle, Length: 9659, dtype: int64
```

In [15]:
```python
data=data.drop_duplicates(['UserID'])
user_data=data.pivot(index='UserID',
                     columns='BookTitle',
                     values='BookRating')
```

In [16]:
```python
user_data
```

Out[16]:

BookTitle	'48	'O Au No Keia: Voices from Hawai'I's Mahu and Transgender Communities	101 Bright Ideas: Esl Activities for All Ages	101 Dalmatians	11th Hour	13 99 Euros	20,001 Names for Baby	2001 Spanish and English Idioms/2001 Modismos Espanoles E Ingleses: 2001 Modismos Espanoles E Ingleses	2061: Odyssey Three	25
UserID										
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	
278846	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
278849	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
278851	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
278852	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
278854	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

2182 rows × 2163 columns

In [17]:
```python
user_data.fillna(0,inplace=True)
```

In [18]:
```python
user_data
```

Out[18]:

BookTitle	'48	'O Au No Keia: Voices from Hawai'I's Mahu and Transgender Communities	101 Bright Ideas: Esl Activities for All Ages	101 Dalmatians	11th Hour	13 99 Euros	20,001 Names for Baby	2001 Spanish and English Idioms/2001 Modismos Espanoles E Ingleses: 2001 Modismos Espanoles E Ingleses	2061: Odyssey Three	253
UserID										
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

BookTitle	'48	'O Au No Keia: Voices from Hawai'I's Mahu and Transgender Communities	101 Bright Ideas: Esl Activities for All Ages	101 Dalmatians	11th Hour	13 99 Euros	20,001 Names for Baby	2001 Spanish and English Idioms/2001 Modismos Espanoles E Ingleses: 2001 Modismos Espanoles E Ingleses	2061: Odyssey Three	253
UserID										
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...
278846	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
278849	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
278851	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
278852	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
278854	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

2182 rows × 2163 columns

In [23]:
```python
from sklearn.metrics import pairwise_distances
from scipy.spatial.distance import cosine,correlation
```

In [24]:
```python
user_sim=1-pairwise_distances(user_data.values,metric='cosine')
```

In [25]:
```python
user_sim
```

Out[25]:
```
array([[1., 0., 0., ..., 0., 0., 0.],
       [0., 1., 0., ..., 0., 0., 0.],
       [0., 0., 1., ..., 0., 0., 0.],
       ...,
       [0., 0., 0., ..., 1., 0., 0.],
       [0., 0., 0., ..., 0., 1., 0.],
       [0., 0., 0., ..., 0., 0., 1.]])
```

In [26]:
```python
user_sim=pd.DataFrame(user_sim)
```

In [27]:
```python
user_sim.index=data.UserID.unique()
user_sim.columns=data.UserID.unique()
```

In [28]:
```python
user_sim.iloc[0:5, 0:5]
```

Out[28]:

	276726	276729	276736	276737	276744
276726	1.0	0.0	0.0	0.0	0.0
276729	0.0	1.0	0.0	0.0	0.0
276736	0.0	0.0	1.0	0.0	0.0
276737	0.0	0.0	0.0	1.0	0.0
276744	0.0	0.0	0.0	0.0	1.0

In [34]:
```python
user_sim.idxmax(axis=1)[0:5]
```

Out[34]:
```
276726    276726
276729    276729
276736    276736
276737    276737
276744    276744
dtype: int64
```

In [35]:
```python
data[(data['UserID']==8)|(data['UserID']==276726)]
```

Out[35]:

	Unnamed: 0	UserID	BookTitle	BookRating
0	1	276726	Classical Mythology	5
2399	2400	8	Ancient Celtic Romances	5

In [36]:
```python
user_1=data[data['UserID']==8]
```

In [37]:
```python
user_2=data[data['UserID']==276726]
```

In [39]:
```python
user_2.BookTitle
```

Out[39]:
```
0     Classical Mythology
Name: BookTitle, dtype: object
```

In [40]:
```python
user_1.BookTitle
```

Out[40]:
```
2399    Ancient Celtic Romances
Name: BookTitle, dtype: object
```

In [41]:
```python
pd.merge(user_1,user_2,on='BookTitle',how='outer')
```

Out[41]:

	Unnamed: 0_x	UserID_x	BookTitle	BookRating_x	Unnamed: 0_y	UserID_y	BookRating_y
0	2400.0	8.0	Ancient Celtic Romances	5.0	NaN	NaN	NaN

	Unnamed: 0_x	UserID_x	BookTitle	BookRating_x	Unnamed: 0_y	UserID_y	BookRating_y
1	NaN	NaN	Classical Mythology	NaN	1.0	276726.0	5.0

In [ ]: