

-----Import Important Libraries-----

```
import pandas as pd
import numpy as np
```

-----Read Dataset-----

```
data = pd.read_csv('Downloads/books.csv')
data
```

	Unnamed: 0	User.ID	Book.Title	Book.Rating
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

-----Calculate lengths-----

```
len(data['User.ID'].unique())
```

2182

```
len(data['Book.Title'].unique())
```

9659

```
len(data['Book.Rating'].unique())
```

10

-----Drop Duplicates-----

```
data = data.drop_duplicates('User.ID').reset_index()
data
```

	index	Unnamed: 0	User.ID	Book.Title	Book.Rating
0	0	1	276726	Classical Mythology	5
1	1	2	276729	Clara Callan	3
2	3	4	276736	Flu: The Story of the Great Influenza Pandemic...	8
3	4	5	276737	The Mummies of Urumchi	6
4	5	6	276744	The Kitchen God's Wife	7
...
2177	9987	9988	162107	What's Bred in the Bone	7
2178	9988	9989	162109	The Flower in the Skull	10
2179	9989	9990	162113	The Cape Ann (Contemporary American Fiction)	8
2180	9990	9991	162121	The Cloister Walk	7
2181	9999	10000	162129	Kids Say the Darndest Things	6

2182 rows × 5 columns

-----Use Pivot-----

data1 =

data.pivot(index='User.ID',columns='Book.Title',values='Book.Rating').reset_index(drop=True)

data1

Book.Title	'48	'O Au No Keia: Voices from Hawaii'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	...	Your Oasis on Flame Lake	Yours, Mine And ...Ours (Silhouette Special Edition, No 682)	Yukon Ho!	Yuwipi: Vision and Experience in Oglala Ritual	Zeit im Wind.	Ar (Vi Sp Cla
0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
...
2177	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
2178	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
2179	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
2180	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
2181	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN

-----Indexing-----

data1.index = data['User.ID'].unique()

data1

Book.Title	'48	'O Au No Keia: Voices from Hawaii'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	...	Your Oasis on Flame Lake	Yours, Mine And ...Ours (Silhouette Special Edition, No 682)	Yukon Ho!	Yuwipi: Vision and Experience in Oglala Ritual	Zeit im Wind.	Ar (Vi Sp Cla
276726	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
276729	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
276736	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
276737	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
276744	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
...
162107	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
162109	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
162113	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
162121	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN
162129	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN	NaN

2182 rows × 2163 columns

-----Fill 0's-----

```
data1.fillna(0,inplace=True)
```

```
data1
```

Book.Title	'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	Spanish and English Idioms/2001 Modismos Espanoles E Ingleses: 2001 Modismos Espanoles E Ingleses	2061: Odyssey Three	253	...	Your Oasis on Flame Lake	Yours, Mine And ...Ours (Silhouette Special Edition, No 682)	Yukon Ho!	Yuwipi: Vision and Experience in Oglala Ritual	Zeit im Wind.	Ze the Arch (Vint Spiri Class
276726	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
276729	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
276736	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
276737	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
276744	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
...
162107	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
162109	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
162113	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
162121	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0
162129	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0

2182 rows x 2163 columns

-----Import Important libraries-----

```
from sklearn.metrics import pairwise_distances
```

```
from scipy.spatial.distance import cosine,correlation
```

-----Use pairwise distances-----

```
sim = 1-pairwise_distances(data1,metric='cosine')
```

```
sim
```

```
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.]])
```

-----Use dataframe-----

```
sim1 = pd.DataFrame(sim)
```

```
sim1
```

	0	1	2	3	4	5	6	7	8	9	...	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181
0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
...
2177	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
2178	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
2179	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0
2180	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
2181	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0

2182 rows × 2182 columns

-----Use iloc-----

sim1.iloc[0:5,0:5]

	0	1	2	3	4
0	1.0	0.0	0.0	0.0	0.0
1	0.0	1.0	0.0	0.0	0.0
2	0.0	0.0	1.0	0.0	0.0
3	0.0	0.0	0.0	1.0	0.0
4	0.0	0.0	0.0	0.0	1.0

-----Fill diagonal elements-----

np.fill_diagonal(sim,0)

sim1.iloc[0:5,0:5]

	0	1	2	3	4
0	0.0	0.0	0.0	0.0	0.0
1	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0

-----Read Data-----

```
data[(data['User.ID']==16) | (data['User.ID']==136)]
```

	index	Unnamed: 0	User.ID	Book.Title	Book.Rating
630	2412	2413	16	Random Acts Of Management:A Dilbert Book	9

-----Read first part-----

```
user1 = data[(data['User.ID']==16)]
user1
```

	index	Unnamed: 0	User.ID	Book.Title	Book.Rating
630	2412	2413	16	Random Acts Of Management:A Dilbert Book	9

-----Read 2nd part-----

```
user2 = data[(data['User.ID']==136)]
user2
```

index	Unnamed: 0	User.ID	Book.Title	Book.Rating
-------	------------	---------	------------	-------------

-----Read column-----

```
user1['Book.Title']
```

```
630    Random Acts Of Management:A Dilbert Book
Name: Book.Title, dtype: object
```

```
user2['Book.Title']
```

```
Series([], Name: Book.Title, dtype: object)
```

-----Merge Data-----

```
pd.merge(user1,user2, on = 'Book.Title',how='outer')
```

	index_x	Unnamed: 0_x	User.ID_x	Book.Title	Book.Rating_x	index_y	Unnamed: 0_y	User.ID_y	Book.Rating_y
0	2412	2413	16	Random Acts Of Management:A Dilbert Book	9	NaN	NaN	NaN	NaN