

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

```
In [2]: books=pd.read_csv('Books.csv')
users=pd.read_csv('Users.csv')
ratings=pd.read_csv('Ratings.csv')
```

C:\Users\KIRANS~1\AppData\Local\Temp\ipykernel_4556\3397335393.py:1: DtypeWarning: Columns (3) have mixed types. Specify dtype option on import or set low_memory=False.
books=pd.read_csv('Books.csv')

```
In [3]: books.head()
```

```
Out[3]:
```

	ISBN	Book-Title	Book-Author	Year-Of-Publication	Publisher	Image-URL-S
0	0195153448	Classical Mythology	Mark P. O. Morford	2002	Oxford University Press	http://images.amazon.com/images/P/0195153448.0... http://images.amazon.com/images/l
1	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/P/0002005018.0... http://images.amazon.com/images/l
2	0060973129	Decision in Normandy	Carlo D'Este	1991	HarperPerennial	http://images.amazon.com/images/P/0060973129.0... http://images.amazon.com/images/l
3	0374157065	Flu: The Story of the Great Influenza Pandemic...	Gina Bari Kolata	1999	Farrar Straus Giroux	http://images.amazon.com/images/P/0374157065.0... http://images.amazon.com/images/l
4	0393045218	The Mummies of Urumchi	E. J. W. Barber	1999	W. W. Norton & Company	http://images.amazon.com/images/P/0393045218.0... http://images.amazon.com/images/l

```
In [4]: users.head()
```

```
Out[4]:
```

	User-ID	Location	Age
0	1	nyc, new york, usa	NaN
1	2	stockton, california, usa	18.0
2	3	moscow, yukon territory, russia	NaN
3	4	porto, v.n.gaiia, portugal	17.0
4	5	farnborough, hants, united kingdom	NaN

```
In [5]: ratings.head()
```

```
Out[5]:
```

	User-ID	ISBN	Book-Rating
0	276725	034545104X	0
1	276726	0155061224	5
2	276727	0446520802	0
3	276729	052165615X	3
4	276729	0521795028	6

```
In [6]: print(books.shape)
print(ratings.shape)
print(users.shape)
```

```
(271360, 8)
(1149780, 3)
(278858, 3)
```

```
In [7]: books.isnull().sum()
```

Out[7]: ISBN 0
Book-Title 0
Book-Author 1
Year-Of-Publication 0
Publisher 2
Image-URL-S 0
Image-URL-M 0
Image-URL-L 3
dtype: int64

In [8]: users.isnull().sum()

Out[8]: User-ID 0
Location 0
Age 110762
dtype: int64

In [9]: ratings.isnull().sum()

Out[9]: User-ID 0
ISBN 0
Book-Rating 0
dtype: int64

In [10]: books.duplicated().sum()

Out[10]: 0

In [11]: users.duplicated().sum()

Out[11]: 0

In [12]: ratings.duplicated().sum()

Out[12]: 0

In [13]: rating_with_name=ratings.merge(books,on='ISBN')

In [14]: num_rating_df=rating_with_name.groupby('Book-Title').count()['Book-Rating'].reset_index()
num_rating_df.rename(columns={'Book-Rating':'rating'},inplace=True)
num_rating_df

Out[14]:

	Book-Title	rating
0	A Light in the Storm: The Civil War Diary of ...	4
1	Always Have Popsicles	1
2	Apple Magic (The Collector's series)	1
3	Ask Lily (Young Women of Faith: Lily Series, ...	1
4	Beyond IBM: Leadership Marketing and Finance ...	1
...
241066	Ä?Ä?lpiraten.	2
241067	Ä?Ä?rger mit Produkt X. Roman.	4
241068	Ä?Ä?sterlich leben.	1
241069	Ä?Ä?stlich der Berge.	3
241070	Ä?Ä?thique en toc	2

241071 rows × 2 columns

```
In [15]: avg_rating_df=rating_with_name.groupby('Book-Title').mean()['Book-Rating'].reset_index()
avg_rating_df.rename(columns={'Book-Rating': 'average-rating'}, inplace=True)
avg_rating_df
```

Out[15]:

	Book-Title	average-rating
0	A Light in the Storm: The Civil War Diary of ...	2.250000
1	Always Have Popsicles	0.000000
2	Apple Magic (The Collector's series)	0.000000
3	Ask Lily (Young Women of Faith: Lily Series, ...	8.000000
4	Beyond IBM: Leadership Marketing and Finance ...	0.000000
...
241066	Ä?Ä?Ipiraten.	0.000000
241067	Ä?Ä?rger mit Produkt X. Roman.	5.250000
241068	Ä?Ä?sterlich leben.	7.000000
241069	Ä?Ä?stlich der Berge.	2.666667
241070	Ä?Ä?thique en toc	4.000000

241071 rows × 2 columns

```
In [16]: popular_df = num_rating_df.merge(avg_rating_df,on='Book-Title')
popular_df
```

Out[16]:

	Book-Title	rating	average-rating
0	A Light in the Storm: The Civil War Diary of ...	4	2.250000
1	Always Have Popsicles	1	0.000000
2	Apple Magic (The Collector's series)	1	0.000000
3	Ask Lily (Young Women of Faith: Lily Series, ...	1	8.000000
4	Beyond IBM: Leadership Marketing and Finance ...	1	0.000000
...
241066	Ä?Ä?Ipiraten.	2	0.000000
241067	Ä?Ä?rger mit Produkt X. Roman.	4	5.250000
241068	Ä?Ä?sterlich leben.	1	7.000000
241069	Ä?Ä?stlich der Berge.	3	2.666667
241070	Ä?Ä?thique en toc	2	4.000000

241071 rows × 3 columns

```
In [17]: popular_df=popular_df[popular_df['rating']>=200].sort_values('average-rating',ascending=False).head(30)
```

```
In [18]: popular_df=popular_df.merge(books,on='Book-Title').drop_duplicates('Book-Title')[['Book-Title','Book-Author','Image-URL-M']]
```

```
In [19]: popular_df
```

Out[19]:

	Book-Title	Book-Author	Image-URL-M	rating	average-rating
0	Harry Potter and the Prisoner of Azkaban (Book 3)	J. K. Rowling	http://images.amazon.com/images/P/0439136350.0...	428	5.852804
3	Harry Potter and the Goblet of Fire (Book 4)	J. K. Rowling	http://images.amazon.com/images/P/0439139597.0...	387	5.824289
5	Harry Potter and the Sorcerer's Stone (Book 1)	J. K. Rowling	http://images.amazon.com/images/P/0590353403.0...	278	5.737410
9	Harry Potter and the Order of the Phoenix (Boo...	J. K. Rowling	http://images.amazon.com/images/P/043935806X.0...	347	5.501441
13	Ender's Game (Ender Wiggins Saga (Paperback))	Orson Scott Card	http://images.amazon.com/images/P/0312853238.0...	249	5.409639
17	Harry Potter and the Chamber of Secrets (Book 2)	J. K. Rowling	http://images.amazon.com/images/P/0439064872.0...	556	5.183453
20	The Hobbit : The Enchanting Prelude to The Lor...	J.R.R. TOLKIEN	http://images.amazon.com/images/P/0345339681.0...	281	5.007117
21	The Fellowship of the Ring (The Lord of the Ri...	J.R.R. TOLKIEN	http://images.amazon.com/images/P/0345339703.0...	368	4.948370
30	Harry Potter and the Sorcerer's Stone (Harry P...	J. K. Rowling	http://images.amazon.com/images/P/059035342X.0...	575	4.895652
32	The Two Towers (The Lord of the Rings, Part 2)	J.R.R. TOLKIEN	http://images.amazon.com/images/P/0345339711.0...	260	4.880769
43	To Kill a Mockingbird	Harper Lee	http://images.amazon.com/images/P/0446310786.0...	510	4.700000

Art of Motorcycle Maintenance: An Inquiry into Values	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
Zlata's Diary: A Child's Life in Sarajevo	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
Zoya	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
"Is for Outlaw"	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	8.0	0.0	0.0	0.0	0.0	0.0	0.0	0

1983 rows × 1113 columns

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

```
In [27]: from sklearn.metrics.pairwise import cosine_similarity
```

```
In [28]: similarity_score=cosine_similarity(pt)
```

```
In [29]: similarity_score.shape
```

Out[29]: (1983, 1983)

```
In [30]: def recommend(book_name):
# index fetch
index = np.where(pt.index==book_name)[0][0]
similar_items = sorted(list(enumerate(similarity_score[index])),key=lambda x:x[1],reverse=True)[1:5]

data = []
for i in similar_items:
    item = []
    temp_df = books[books['Book-Title'] == pt.index[i[0]]]
    item.extend(list(temp_df.drop_duplicates('Book-Title')['Book-Title'].values))
    item.extend(list(temp_df.drop_duplicates('Book-Title')['Book-Author'].values))
    item.extend(list(temp_df.drop_duplicates('Book-Title')['Image-URL-M'].values))

    data.append(item)

return data
```

```
In [31]: #recommend('Message in a Bottle')
#recommend('The Notebook')
#recommend('1984')
recommend('Animal Farm')
```

Out[31]: [['Hamlet',
'William Shakespeare',
'http://images.amazon.com/images/P/067172262X.01.MZZZZZZZ.jpg'],
['Women in His Life',
'Barbara Taylor Bradford',
'http://images.amazon.com/images/P/0345345738.01.MZZZZZZZ.jpg'],
['Before and After',
'Rosellen Brown',
'http://images.amazon.com/images/P/0440216540.01.MZZZZZZZ.jpg'],
['Poland',
'James A. Michener',
'http://images.amazon.com/images/P/0394533887.01.MZZZZZZZ.jpg']]

```
In [32]: pt.index[545]
```

Out[32]: 'Flowers for Algernon (Bantam Classic)'

```
In [33]: import pickle
pickle.dump(popular_df,open('popular.pkl','wb'))
```

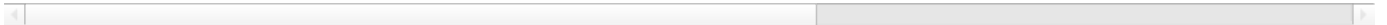
```
In [34]: books.drop_duplicates('Book-Title')
```

Out[34]:

ISBN	Book-Title	Book-Author	Year-Of-Publication	Publisher	Image-URL-S
------	------------	-------------	---------------------	-----------	-------------

0	0195153448	Classical Mythology	Mark P. O. Morford	2002	Oxford University Press	http://images.amazon.com/images/P/0195153448.0...	http://images.amaz
1	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/P/0002005018.0...	http://images.amaz
2	0060973129	Decision in Normandy	Carlo D'Este	1991	HarperPerennial	http://images.amazon.com/images/P/0060973129.0...	http://images.amaz
3	0374157065	Flu: The Story of the Great Influenza Pandemic...	Gina Bari Kolata	1999	Farrar Straus Giroux	http://images.amazon.com/images/P/0374157065.0...	http://images.amaz
4	0393045218	The Mummies of Urumchi	E. J. W. Barber	1999	W. W. Norton & Company	http://images.amazon.com/images/P/0393045218.0...	http://images.amaz
...
271354	0449906736	Flashpoints: Promise and Peril in a New World	Robin Wright	1993	Ballantine Books	http://images.amazon.com/images/P/0449906736.0...	http://images.amaz
271356	0525447644	From One to One Hundred	Teri Sloat	1991	Dutton Books	http://images.amazon.com/images/P/0525447644.0...	http://images.amaz
271357	006008667X	Lily Dale : The True Story of the Town that Ta...	Christine Wicker	2004	HarperSanFrancisco	http://images.amazon.com/images/P/006008667X.0...	http://images.amaz
271358	0192126040	Republic (World's Classics)	Plato	1996	Oxford University Press	http://images.amazon.com/images/P/0192126040.0...	http://images.amaz
271359	0767409752	A Guided Tour of Rene Descartes' Meditations o...	Christopher Biffle	2000	McGraw-Hill Humanities/Social Sciences/Languages	http://images.amazon.com/images/P/0767409752.0...	http://images.amaz

242135 rows × 8 columns



```
In [35]: pickle.dump(pt,open('pt.pkl','wb'))
pickle.dump(books,open('books.pkl','wb'))
pickle.dump(similarity_score,open('similarity_score.pkl','wb'))
```

In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js