

```
[1]: #import libraries
import numpy as np
import pandas as pd

In [2]: #import datasets
books = pd.read_csv('books.csv');
users = pd.read_csv('users.csv');
ratings = pd.read_csv('ratings.csv');

F:\ANACONDA\lib\site-packages\IPython\core\interactiveshell.py:3444: DtypeWarning: Columns (3) have mixed types.Specify dtype option on import or set low_memo
ry=False.
exec(code_obj, self.user_global_ns, self.user_ns)

In [3]: #prints first five data of books dataset
books.head()

Out[3]:
```

	ISBN	Book-Title	Book-Author	Year-Of-Publication	Publisher	Image-URL-S	Image-URL-M	Image
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/P/0195153448.0...	http://images.amazon.com/images/P/0195153448.0...
1	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/P/0002005018.0...	http://images.amazon.com/images/P/0002005018.0...	http://images.amazon.com/images/P/0002005018.0...
2	0060973129	Decision in Normandy	Carlo D'Este	1991	HarperPerennial	http://images.amazon.com/images/P/0060973129.0...	http://images.amazon.com/images/P/0060973129.0...	http://images.amazon.com/images/P/0060973129.0...
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/P/0374157065.0...	http://images.amazon.com/images/P/0374157065.0...
4	0393045218	The Mummies of Urunchi	E. J. W. Barber	1999	W. W. Norton & Company	http://images.amazon.com/images/P/0393045218.0...	http://images.amazon.com/images/P/0393045218.0...	http://images.amazon.com/images/P/0393045218.0...

```
In [4]: #prints first five data of users dataset
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, ykoun territory, russia	NaN
3	4	porto, v.n.gaid, portugal	17.0
4	5	farnborough, hants, united kingdom	NaN

```
In [5]: #prints first five data of ratings dataset
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 total rows present in dataset of books,ratings and users
print(books.shape)
print(ratings.shape)
print(users.shape)

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

In [7]: #print missing values in books dataset
books.isnull().sum()

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

In [8]: #print missing values in users dataset
users.isnull().sum()

User-ID      0
Location     0
Age         110762
dtype: int64

Out[8]:
```

	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 [9]: #print missing values in ratings dataset
ratings.isnull().sum()

User-ID      0
ISBN         0
Book-Rating  0
dtype: int64

Out[9]:

0

In [10]: #if any duplicates present in dataset
books.duplicated().sum()

0

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

0

Out[11]:

0

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

0

Out[12]:

0
```

## Popularity Based Recommender System

```
In [13]: #BUILDING OUR MODEL
#MODEL SHOWS THE TOP 50 HIGHEST RATING BOOKS AND ALSO RECOMMEND SIMILAR BOOKS

In [14]: ##merging the datasets of ratings and books
ratings_with_name = ratings.merge(books,on='ISBN')

In [15]: ratings_with_name

Out[15]:
```

	User-ID	ISBN	Book-Rating	Book-Title	Book-Author	Year-Of-Publication	Publisher	Image-URL-S	Image-URL-M
0	276725	034545104X	0	Flesh Tones: A Novel	M. J. Rose	2002	Ballantine Books	http://images.amazon.com/images/P/034545104X.0...	http://images.amazon.com/images/P/034545104X.0...
1	2313	034545104X	5	Flesh Tones: A Novel	M. J. Rose	2002	Ballantine Books	http://images.amazon.com/images/P/034545104X.0...	http://images.amazon.com/images/P/034545104X.0...
2	6543	034545104X	0	Flesh Tones: A Novel	M. J. Rose	2002	Ballantine Books	http://images.amazon.com/images/P/034545104X.0...	http://images.amazon.com/images/P/034545104X.0...
3	8680	034545104X	5	Flesh Tones: A Novel	M. J. Rose	2002	Ballantine Books	http://images.amazon.com/images/P/034545104X.0...	http://images.amazon.com/images/P/034545104X.0...
4	10314	034545104X	9	Flesh Tones: A Novel	M. J. Rose	2002	Ballantine Books	http://images.amazon.com/images/P/034545104X.0...	http://images.amazon.com/images/P/034545104X.0...
...	...	...	...	...	...	...	...	...	...
1031131	276688	0517145553	0	Mostly Harmless	Douglas Adams	1995	Random House Value Pub	http://images.amazon.com/images/P/0517145553.0...	http://images.amazon.com/images/P/0517145553.0...
1031132	276688	1575660792	7	Gray Matter	Shirley Kennett	1996	Kensington Publishing Corporation	http://images.amazon.com/images/P/1575660792.0...	http://images.amazon.com/images/P/1575660792.0...
1031133	276690	0590907301	0	Triplet Trouble and the Class Trip (Triplet Tr...	Debbie Dadey	1997	Apple	http://images.amazon.com/images/P/0590907301.0...	http://images.amazon.com/images/P/0590907301.0...
1031134	276704	0679752714	0	A Desert of Pure Feeling (Vintage Contemporaries)	Judith Freeman	1997	Vintage Books USA	http://images.amazon.com/images/P/0679752714.0...	http://images.amazon.com/images/P/0679752714.0...
1031135	276704	0806917695	5	Perplexing Lateral Thinking Puzzles: Scholasti...	Paul Sloane	1997	Sterling Publishing	http://images.amazon.com/images/P/0806917695.0...	http://images.amazon.com/images/P/0806917695.0...

1031135 rows x 10 columns

```
In [16]: #shows all the books and number of votes of every book
num_rating_df = ratings_with_name.groupby('Book-Title').count()['Book-Rating'].reset_index()
num_rating_df.rename(columns={'Book-Rating':'num_ratings'},inplace=True)
num_rating_df

Out[16]:
```

	Book-Title	num_ratings
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	À?À?piraten.	2
241067	À?À?nger mit Produkt X. Roman.	4
241068	À?À?sterlich leben.	1
241069	À?À?stlich der Berge.	3
241070	À?À?hique en toc	2

241071 rows x 2 columns

```
In [17]: #finding the average of every book
avg_rating_df = ratings_with_name.groupby('Book-Title').mean()['Book-Rating'].reset_index()
avg_rating_df.rename(columns={'Book-Rating':'avg_ratings'},inplace=True)
avg_rating_df

Out[17]:
```

	Book-Title	avg_ratings
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, ...	0.000000
4	Beyond IBM: Leadership Marketing and Finance ...	0.000000
...	...	...
241066	À?À?piraten.	0.000000
241067	À?À?nger mit Produkt X. Roman.	5.250000
241068	À?À?sterlich leben.	7.000000
241069	À?À?stlich der Berge.	2.666667
241070	À?À?hique en toc	4.000000

241071 rows x 2 columns

```
In [18]: ##merging the dataframe num_rating_df and avg_rating_df
popular_df = num_rating_df.merge(avg_rating_df,on='Book-Title')
popular_df

Out[18]:
```

	Book-Title	num_ratings	avg_ratings
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	0.000000
4	Beyond IBM: Leadership Marketing and Finance ...	1	0.000000
...	...	...	...
241066	À?À?piraten.	2	0.000000
241067	À?À?nger mit Produkt X. Roman.	4	5.250000
241068	À?À?sterlich leben.	1	7.000000
241069	À?À?stlich der Berge.	3	2.666667
241070	À?À?hique en toc	2	4.000000

241071 rows x 3 columns

```
In [19]: #consider the books whose rating is greater than or equals to 250 and sort the values
popular_df=popular_df[popular_df['num_ratings']>=250].sort_values('avg_ratings',ascending=False).head(50)

In [20]: #deletes all duplicate books and show the books whose rating is greater than 250 with Book-Title,Book-Author,Image-URL-M,num_ratings,avg_ratings
popular_df=popular_df.merge(books,on='Book-Title').drop_duplicates('Book-Title')[['Book-Title','Book-Author','Image-URL-M','num_ratings','avg_ratings']]

In [21]: popular_df

Out[21]:
```

	Book-Title	Book-Author	Image-URL-M	num_ratings	avg_ratings
0	Harry Potter and the Prisoner of Azkaban (Book 3)	J. K. Rowling	http://images.amazon.com/images/P/0439136390.0...	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	Harry Potter and the Chamber of Secrets (Book...	J. K. Rowling	http://images.amazon.com/images/P/0439064872.0...	556	5.183451
16	The Hobbit : The Enchanting Prelude to The Lor...	J.R.R. TOLKIEN	http://images.amazon.com/images/P/0345339681.0...	281	5.007117
17	The Fellowship of the Ring (The Lord of the Ri...	J.R.R. TOLKIEN	http://images.amazon.com/images/P/0453339703.0...	368	4.948370
26	Harry Potter and the Sorcerer's Stone (Harry P...	J. K. Rowling	http://images.amazon.com/images/P/059035342X.0...	575	4.895652
28	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
39	To Kill a Mockingbird	Harper Lee	http://images.amazon.com/images/P/0446310786.0...	510	4.468726
47	The Da Vinci Code	Dan Brown	http://images.amazon.com/images/P/0385504209.0...	898	4.642539
53	The Five People You Meet in Heaven	Mitch Albom	http://images.amazon.com/images/P/0718668871.6.0...	430	4.551163
55	The Catcher in the Rye	J.D. Salinger	http://images.amazon.com/images/P/0316769487.0...	449	4.545657
62	The Lovely Bones: A Novel	Alice Sebold	http://images.amazon.com/images/P/0316666343.0...	1295	4.468726
62	1984	George Orwell	http://images.amazon.com/images/P/0451524934.0...	284	4.454225
73	Prodigal Summer: A Novel	Barbara Kingsolver	http://images.amazon.com/images/P/0060959037.0...	253	4.450593
73	Nowhere	Neil Gaiman	http://images.amazon.com/images/P/0380789019.0...	265	4.449057
78	The Secret Life of Bees	Sue Monk Kidd	http://images.amazon.com/images/P/0142001740.0...	774	4.075269
84	Stupid White Men ...and Other Sorry Excuses fo...	Michael Moore	http://images.amazon.com/images/P/006392452.0...	283	4.356880
85	Tuesdays with Morrie: An Old Man, a Young Man...	MITCH ALBOM	http://images.amazon.com/images/P/0385484518.0...	493	4.354970
89	The Red Tent (Bestselling Backlist)	Anita Diamant	http://images.amazon.com/images/P/0312195516.0...	723	4.334716
89	The Hitchhiker's Guide to the Galaxy	Douglas Adams	http://images.amazon.com/images/P/0451294936.0...	268	4.328358
98	Nickel and Dimed: On (Not) Getting By in America	Barbara Ehrenreich	http://images.amazon.com/images/P/0805063894.0...	335	4.289552
100	Into the Wild	Jon Krakauer	http://images.amazon.com/images/P/0385486804.0...	252	4.273810
103	Fahrenheit 451	Ray Bradbury	http://images.amazon.com/images/P/0257208626.0...	409	4.264059
116	Girl with a Pearl Earring	Tracy Chevalier	http://images.amazon.com/images/P/0451294934.0...	526	4.218631
117	Outlander	DIANA GABALDON	http://images.amazon.com/images/P/0440222915.0...	283	4.173145
123	Where the Heart Is (Oprah's Book Club (Paperba...	Billie Letts	http://images.amazon.com/images/P/0446672211.0...	585	4.105983
124	Seabiscuit: An American Legend	LAURA HILLENBRAND	http://images.amazon.com/images/P/0449005615.0...	275	4.098182
126	Life of Pi	Yann Martel	http://images.amazon.com/images/P/0151008116.0...	664	4.088855
130	The Bean Trees	Barbara Kingsolver	http://images.amazon.com/images/P/0060915544.0...	369	4.087404
133	A Child Called 'It': One Child's Courage to S...	Dave Pelzer	http://images.amazon.com/images/P/1558743669.0...	265	4.086792
135	ANGELA'S ASHES	Frank McCourt	http://images.amazon.com/images/P/0142001740.0...	279	4.075269
137	Good in Bed	Jennifer Weiner	http://images.amazon.com/images/P/0743418174.0...	490	4.055102
139	Fast Food Nation: The Dark Side of the All-Ame...	Eric Schlosser	http://images.amazon.com/images/P/0060938455.0...	321	4.037383
140	American Gods	Neil Gaiman	http://images.amazon.com/images/P/0380789035.0...	302	4.006623
141	Skipping Christmas	JOHN GRISHAM	http://images.amazon.com/images/P/0312924585.0...	322	4.006211
147	Left Behind: A Novel of the Earth's Last Days ...	Tim Lahaye	http://images.amazon.com/images/P/0842329129.0...	318	4.003145
149	The Golden Compass (His Dark Materials, Book 1)	PHILIP PULLMAN	http://images.amazon.com/images/P/037582345X.0...	336	4.000000
157	The Color Purple	Alice Walker	http://images.amazon.com/images/P/0671617028.0...	314	3.964968
163	Silence of the Lambs	Thomas Harris	http://images.amazon.com/images/P/0385504209.0...	256	3.960938
166	About a Boy	Nick Hornby	http://images.amazon.com/images/P/1573223703.0...	262	3.980763
173	Seven Up (A Stephanie Plum Novel)	Janet Evanovich	http://images.amazon.com/images/P/0312265840.0...	278	3.888489
175	The Alchemist: A Fable About Following Your Dream	Paulo Coelho	http://images.amazon.com/images/P/0062502174.0...	266	3.875940
177	Hard Eight : A Stephanie Plum Novel (A Stephan...	Janet Evanovich	http://images.amazon.com/images/P/0312983867.0...	269	3.825279
182	Suzanne's Diary for Nicholas	James Paterson	http://images.amazon.com/images/P/0316964943.0...	457	3.820569
186	Lord of the Flies	William Gerald Golding	http://images.amazon.com/images/P/0399501487.0...	259	3.818533
192	A Prayer for Owen Meany	John Irving	http://images.amazon.com/images/P/035631792.0...	413	3.796610
194	The Vampire Lestat (Vampire Chronicles, Book II)	ANNE RICE	http://images.amazon.com/images/P/0345313860.0...	301	3.777409
195	White Oleander : A Novel (Oprah's Book Club)	Janet Fitch	http://images.amazon.com/images/P/0316284955.0...	356	3.772472

## Collaborative Filtering Based Recommender System

```
In [22]: #segregation of users who have given minimum of 200 votes to any book
x = ratings_with_name.groupby('User-ID').count()[['Book-Rating']] > 200
book_readers = x[x].index

In [23]: #users with minimum 200 votes
filtered_rating = ratings_with_name[ratings_with_name['User-ID'].isin(book_readers)]

In [24]: #books having ratings greater than 50
y = filtered_rating.groupby('Book-Title').count()[['Book-Rating']] >=50
famous_books = y[y].index

In [25]: #all famous books i.e. top 50 books
final_ratings = filtered_rating[filtered_rating['Book-Title'].isin(famous_books)]

In [26]: #pivot table
pt=final_ratings.pivot_table(index='Book-Title',columns='User-ID',values='Book-Rating')

In [27]: #filling the missing values with 0
pt.fillna(0,inplace=True)

In [28]: pt

Out[28]:
```

	User-ID	254	2276	2766	2977	3363	4017	4385	6251	6323	6543	...	271705	273979	274004	274061	274301	274308	275970	277427	277639	278418
Book-Title																						
1984	9.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	...	10.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1st to Die: A Novel	0.0	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
2nd Chance	0.0	10.0	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
4 Blondes	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	0.0
A Bend in the Road	0.0	0.0	7.0	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
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Year of Wonders	0.0	0.0	0.0	7.0	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
You Belong To Me	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	0.0
Zen and the 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	0.0	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	0.0	0.0	0.0
VOI" Is for Outlaw"	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	8.0	0.0	0.0	0.0	0.0	0.0

706 rows x 810 columns

```
In [29]: #import cosine similarity to find the similar books by calculating euclidean distance
from sklearn.metrics.pairwise import cosine_similarity

In [30]: similarity_scores = cosine_similarity(pt)

In [31]: similarity_scores.shape

(706, 706)

Out[31]:

In [32]: #this function will recommend 10 similar books
def recommend(book_name):
    #index fetch
    index=np.where(pt.index==book_name)[0][0]
    similar_items=sorted(list(enumerate(similarity_scores[index])),key=lambda x:x[1],reverse=True)[1:11]

    for i in similar_items:
        print(pt.index[i][0])

In [33]: # shows the list of 10 recommended books
recommend('Message in a Bottle')

Nights in Rodanthe
The Mulberry Tree
A Walk to Remember
River's End
Nightmares & Dreams
Sanctuary
All I Really Need to Know
Midnight Bayou
Hard Eight : A Stephanie Plum Novel (A Stephanie Plum Novel)
Dark Angel

In [34]: # Book present in index 545
pt.index[545]

"The Handmaid's Tale"
```