

Book Recommender System ¶

Based on Ratings

Purpose

Create a book recommender system using data from the “Book Recommendation Dataset” (<https://www.kaggle.com/datasets/arashnic/book-recommendation-datasetUsing>). The system will allow users to input a book they like (limited to the titles within the data set) and recommends other book for them to add to their reading list.

```
In [1]: 1 # Import libraries
        2 import pandas as pd
        3 import numpy as np
        4 import matplotlib.pyplot as plt
        5 import seaborn as sns
        6
        7 # Hide warnings
        8 import warnings
        9 warnings.filterwarnings('ignore')
```

Data Import

```
In [2]: 1 # Import data from the Books.csv into a dataframe and then display
2 books = pd.read_csv (r'/Users/kimberlyadams/Documents/GitHub/Portf
3 books.head()
```

Out [2]:

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

```
In [3]: 1 # Import data from the Ratings.csv into a dataframe and then displ
2 ratings = pd.read_csv (r'/Users/kimberlyadams/Documents/GitHub/Portf
3 ratings.head()
```

Out [3]:

	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



Exploration and Cleanup

```
In [4]: 1 # Determine number of values in each column
        2 books.count(axis=0)
```

```
Out[4]: ISBN                271360
         Book-Title          271360
         Book-Author         271359
         Year-Of-Publication  271360
         Publisher            271358
         Image-URL-S          271360
         Image-URL-M          271360
         Image-URL-L          271357
         dtype: int64
```

```
In [5]: 1 # Determine number of unique values in each column
        2 books.nunique(axis=0)
```

```
Out[5]: ISBN                271360
         Book-Title          242135
         Book-Author         102023
         Year-Of-Publication    202
         Publisher            16807
         Image-URL-S          271044
         Image-URL-M          271044
         Image-URL-L          271041
         dtype: int64
```

Note that the counts of the ISBN and Book-Title are the same, but the unique values number is different. This indicates that a book title might have multiple ISBN numbers.

```
In [6]: 1 # Count how many Title-Author combinations are unique
        2 books.groupby(['Book-Title', 'Book-Author']).ngroups
```

```
Out[6]: 251184
```

Looks like there are 251,184 unique books in the dataset based on Title and Author combinations. I will add a column combining these columns so I can avoid duplicates later on. I can't just assume that each title is unique since multiple authors can write books with the same title and unlikely as it is, some authors might even share names.

```
In [7]: 1 # Concatenate Title and Author text into a new column
        2 books["TitleAuthor"] = books["Book-Title"] + " by " + books["Book-
```

```
In [8]: 1 # Merge the two dataframe based on the 'ISBN' column to get book a
2 BookRatings= pd.merge(books, ratings, on= 'ISBN')
3 BookRatings.head()
```

Out[8]:

	ISBN	Book- Title	Book- Author	Year-Of- Publication	Publisher	
0	0195153448	Classical Mythology	Mark P. O. Morford	2002	Oxford University Press	http://images.amazon.com/imag
1	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/imag
2	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/imag
			Richard			

My computer also started having issues with the larger dataset, so I also trimmed out users that had not rated at least 9 books so as to only view the "active" users. Normally I would NOT have done this step, but the computer crashed non-stop until I did.

```
In [9]: 1 # Creating a dictionary using value_counts()
2 items = BookRatings['User-ID'].value_counts().to_dict().items()
3
4 # Filtering only those rows where duplicate entries occur more tha
5 n = 9
6 BookRatingsActiveUsers = BookRatings[BookRatings['User-ID'].isin([
```

Many books have a rating of 0. Since most rating system will only let you rate as low as 1 star, I am going to assume that 0 means the user did not rate the book and thereby I am going to replace the 0s with null values.

```
In [10]: 1 # Replace 0 book rating scores with null.
2 BookRatings['Book-Rating'].replace(0, np.nan, inplace=True)
```

```
In [11]: 1 # Create new dataframe with active user ratings that are greater t
2 BookRatings = BookRatingsActiveUsers[BookRatingsActiveUsers['Book-
3 BookRatings.head()
```

Out[11]:

	ISBN	Book- Title	Book- Author	Year-Of- Publication	Publisher	
1	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/
3	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/
5	0002005018	Clara Callan	Richard Bruce Wright	2001	HarperFlamingo Canada	http://images.amazon.com/images/

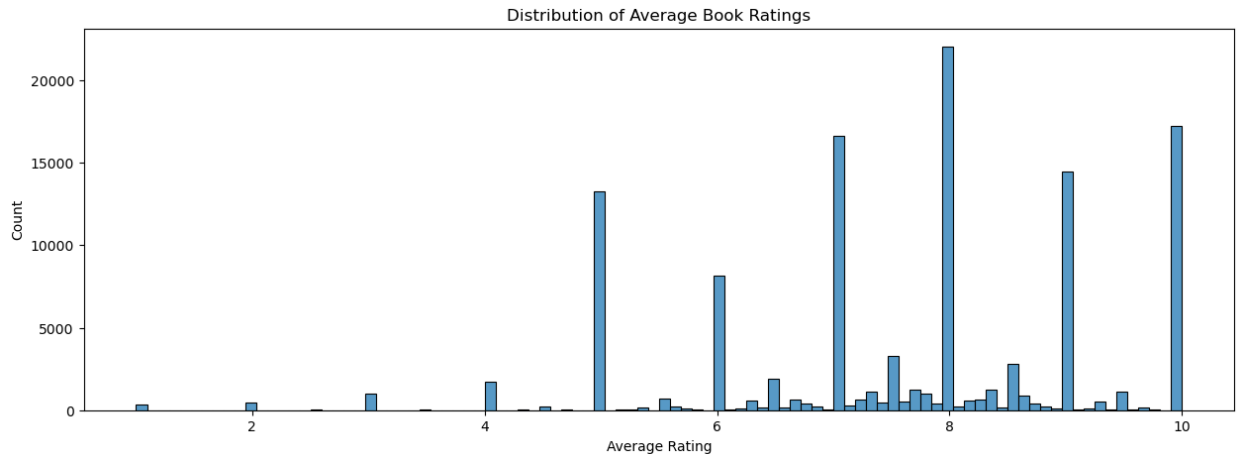
Looks like there are multiple listings now for each book so we might need to average the ratings to get an overall score and then remove the duplicates.

```
In [12]: 1 # Group the rows by the book's ISBN to get all the reviews for each
2 # Get the average rating for the book from all the user ratings
3 # Sort the resulting dataframe by rating
4 Summary = BookRatings.groupby('TitleAuthor').agg({'Book-Rating': 'mean'})
```

```
In [13]: 1 # Group the combined dataframe rows by the book's title/author combination
2 # Put count values into new column
3 Summary['RatingsNum'] = pd.DataFrame(BookRatings.groupby('TitleAuthor').count())
```

```
In [14]: 1 # Plot distribution of the ratings to see the shape of the data
2 plt.figure(figsize=(15,5))
3 sns.histplot(Summary['Book-Rating'])
4 plt.title("Distribution of Average Book Ratings")
5 plt.xlabel('Average Rating')
6 plt.ylabel('Count')
```

```
Out[14]: Text(0, 0.5, 'Count')
```

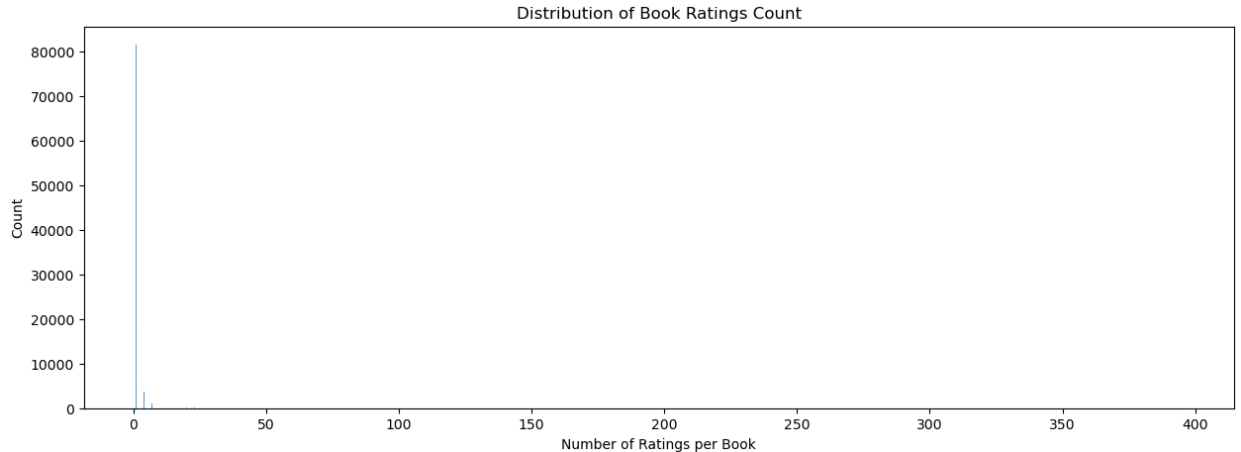


From this distribution we can see that most of the ratings are 5-10 so overall users are rating books they liked. There are a few lower ratings so occasionally users indicated they did not like a book, but these are rare.

In addition to the overall skew towards higher values, there is still a large number of 5 ratings that would potentially keep this from an otherwise normal distribution (when only looking at the higher values).

```
In [15]: 1 # Plot distribution of the number of ratings per book to see the s
2 plt.figure(figsize=(15,5))
3 sns.histplot(Summary['RatingsNum'])
4 plt.title("Distribution of Book Ratings Count")
5 plt.xlabel('Number of Ratings per Book')
6 plt.ylabel('Count')
```

Out[15]: Text(0, 0.5, 'Count')



That's a little hard to see anything. Let's look at the data itself.

```
In [16]: 1 # Allow all rows of dataframe to show
2 pd.set_option('display.max_rows', None)
3
4 # Display the number of books with each number of ratings
5 Summary.RatingsNum.value_counts().sort_index()
```

```
Out[16]: 1      81590
2      17492
3       6935
4       3671
5       2226
6       1513
7       1092
8        805
9        566
10       487
11       356
12       297
13       287
14       269
15       219
16       158
17       157
18       122
19       119
20       105
```

```
In [17]: 1 # What is the highest number of ratings on a single book?
2 print(max(Summary['RatingsNum']))
```

395

```
In [18]: 1 # What is the average number of ratings?
2 print(Summary[['RatingsNum']].mean(axis=0))
```

RatingsNum 2.412841
dtype: float64

From this plot and table we can see that most books are not rated and after that most have only 1 or 2 ratings. The book with the highest number of ratings has only 707. On average books have only 2.75 ratings so that isn't very high.

The recommender will need to take into account movies that just have a single rating may not be accurately portrayed by that rating as that is a single person's opinion. However, since there is not much data available on each book, any rating is better than none.



Recommendation

```
In [19]: 1 # Create a pivot table with each user's ratings for all books with
2 recommend = BookRatings.pivot_table(index='User-ID', columns='Title',
3 recommend.head())
```

Out[19]:

TitleAuthor	A Light in the Storm: The Civil War Diary of Amelia Martin, Fenwick Island, Delaware, 1861 (Dear America) by Karen Hesse	Ask Lily (Young Women of Faith: Lily Series, Book 5) by Nancy N. Rue	Dark Justice by Jack Higgins	Earth Prayers From around the World: 365 Prayers, Poems, and Invocations for Honoring the Earth by Elizabeth Roberts	Final Fantasy Anthology: Official Strategy Guide (Brady Games) by David Cassady	Flight of Fancy: American Heiresses (Zebra Ballad Romance) by Tracy Cozzens	Garfield Bigger and Better (Garfield (Numbered Paperback)) by Jim Davis	Pi B
	User-ID							

This table is mostly NaNs because not every user has seen every movie or at very least submitted a rating for it.

In [20]:

```
1 # Type in book and author combination from database to search for.
2 # Note: If this were a standalone app, the user would ideally have
3 search = 'The Da Vinci Code by Dan Brown'
4
5 # Lookup rating for the searched book
6 Rating_Lookup = recommend[search]
7
8 # Find similar book ratings from the `recommend` pivot_table based
9 # Create a dataframe with results as a Correlation column
10 correlation = pd.DataFrame(recommend.corrwith(Rating_Lookup), columns=
11
12 # Drop the empty values
13 correlation.dropna(inplace=True)
14
15 # Add the number of users who rated that book to the dataframe
16 correlation = correlation.join(Summary['RatingsNum'])
17
18 # Get recommendations by looking for high correlation values between
19 # Only consider books that over 25 users have read to weed out less
20 # Print out top 10 recommendations
21 correlation[correlation['RatingsNum']>25].sort_values(by='Correlation')
```

Out[20]:

	Correlation	RatingsNum
TitleAuthor		
Let Me Call You Sweetheart by Mary Higgins Clark	1.0	28
On the Road (Penguin 20th Century Classics) by Jack Kerouac	1.0	26
The Lion, the Witch, and the Wardrobe (The Chronicles of Narnia, Book 2) by C. S. Lewis	1.0	37
Foucault's Pendulum by Umberto Eco	1.0	30
The Da Vinci Code by Dan Brown	1.0	314
The Mulberry Tree by Jude Deveraux	1.0	41
Open House (Oprah's Book Club (Paperback)) by Elizabeth Berg	1.0	34
Shell Seekers by Rosamunde Pilcher	1.0	26
Speak by Laurie Halse Anderson	1.0	26
Milkrun by Sarah Mlynowski	1.0	30

```

In [21]: 1 # Running again with a different book to compare results
          2
          3 # Type in book and author combination from database to search for.
          4 # Note: If this were a standalone app, the user would ideally have
          5 search = 'A Christmas Carol by Charles Dickens'
          6
          7 # Lookup rating for the searched book
          8 Rating_Lookup = recommend[search]
          9
         10 # Find similar book ratings from the `recommend` pivot_table based
         11 # Create a dataframe with results as a Correlation column
         12 correlation = pd.DataFrame(recommend.corrwith(Rating_Lookup), colu
         13
         14 # Drop the empty values
         15 correlation.dropna(inplace=True)
         16
         17 # Add the number of users who rated that book to the dataframe
         18 correlation = correlation.join(Summary['RatingsNum'])
         19
         20 # Get recommendations by looking for high correlation values betwe
         21 # Only consider books that over 25 users have read to weed out les
         22 # Print out top 10 recommendations
         23 correlation[correlation['RatingsNum']>25].sort_values(by='Correlat

```

Out [21]:

	Correlation	RatingsNum
TitleAuthor		
Fried Green Tomatoes at the Whistle Stop Cafe by Fannie Flagg	-1.0	97

Overall, this recommendation system only seems to be recommending similarly rated books without taking anything else into account. This makes sense as this is the only data we fed the model. However, this approach returns a lot of titles that would likely not interest the user due to other characteristics of the book such as distasteful genres. Or in the case of A Christmas Carol, barely returned any suggestions at all.

I feel like there is a better way to take more details about the book into account.

▼ Based on Content

▼ Purpose

Create a book recommender system using data from the “Goodreads' Best Books Ever” dataset found on Kaggle (<https://www.kaggle.com/datasets/meetnaren/goodreads-best-books> (<https://www.kaggle.com/datasets/meetnaren/goodreads-best-books>)).

The system will allow users to input a book they like (limited to the titles within the data set) and recommends other books for them to add to their reading list. Recommendations are based on the similarity of the books' descriptions and thus is more tailored to individual tastes than a flat rating system. By factoring in the actual content of the book, we can avoid recommending something in a completely different genre that just happened to be highly rated.

▼ Data Import

```
In [22]: 1 # Import libraries
          2 import pandas as pd
          3 import matplotlib.pyplot as plt
          4 import seaborn as sns
          5
          6 from sklearn.feature_extraction import text
          7 from sklearn.metrics.pairwise import linear_kernel
```

```
In [23]: 1 # Import data from the book_data.csv into a dataframe and then display
2 books = pd.read_csv (r'/Users/kimberlyadams/Documents/GitHub/Portf
3 books.head()
```

Out[23]:

	book_authors	book_desc	book_edition	book_format	book_isbn	book_pages	book_
0	Suzanne Collins	Winning will make you famous. Losing means cer...	NaN	Hardcover	9.78044E+12	374 pages	
1	J.K. Rowling Mary GrandPré	There is a door at the end of a silent corrido...	US Edition	Paperback	9.78044E+12	870 pages	
2	Harper Lee	The unforgettable novel of a childhood in a sl...	50th Anniversary	Paperback	9.78006E+12	324 pages	
3	Jane Austen Anna Quindlen Mrs. Oliphant George...	«È cosa ormai risaputa che a uno scapolo in po...	Modern Library Classics, USA / CAN	Paperback	9.78068E+12	279 pages	
4	Stephenie Meyer	About three things I was absolutely positive.F...	NaN	Paperback	9.78032E+12	498 pages	



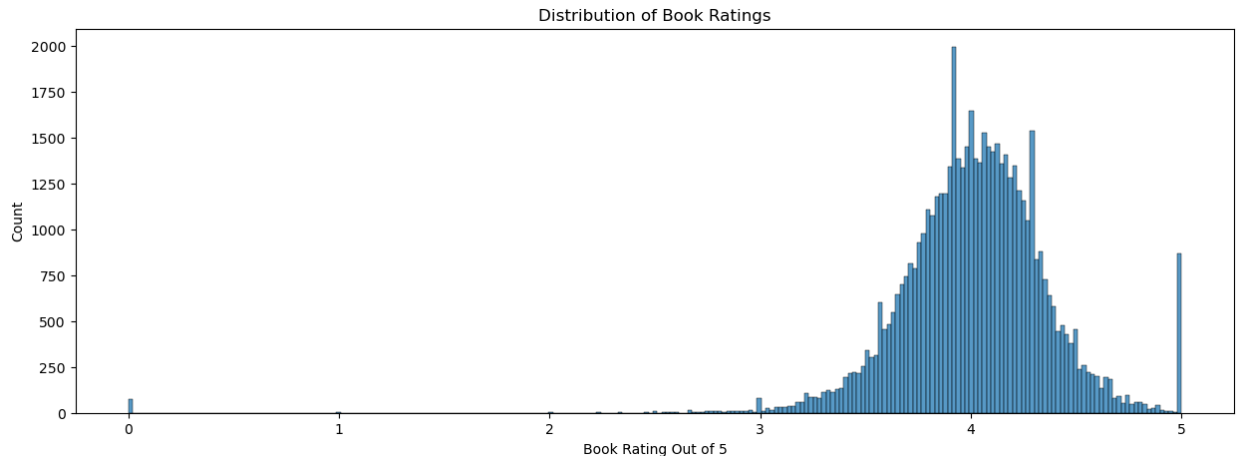
Data Exploration

```
In [24]: 1 # Determine number of unique values in each column
2 books.nunique(axis=0)
```

```
Out[24]: book_authors      27159
book_desc      51781
book_edition    2134
book_format     147
book_isbn       548
book_pages     1403
book_rating     259
book_rating_count 21860
book_review_count 6895
book_title     48483
genres         30094
image_url      53618
dtype: int64
```

```
In [25]: 1 # Plot distribution of the number of ratings per book to see the s
2 plt.figure(figsize=(15,5))
3 sns.histplot(books['book_rating'])
4 plt.title("Distribution of Book Ratings")
5 plt.xlabel('Book Rating Out of 5')
6 plt.ylabel('Count')
```

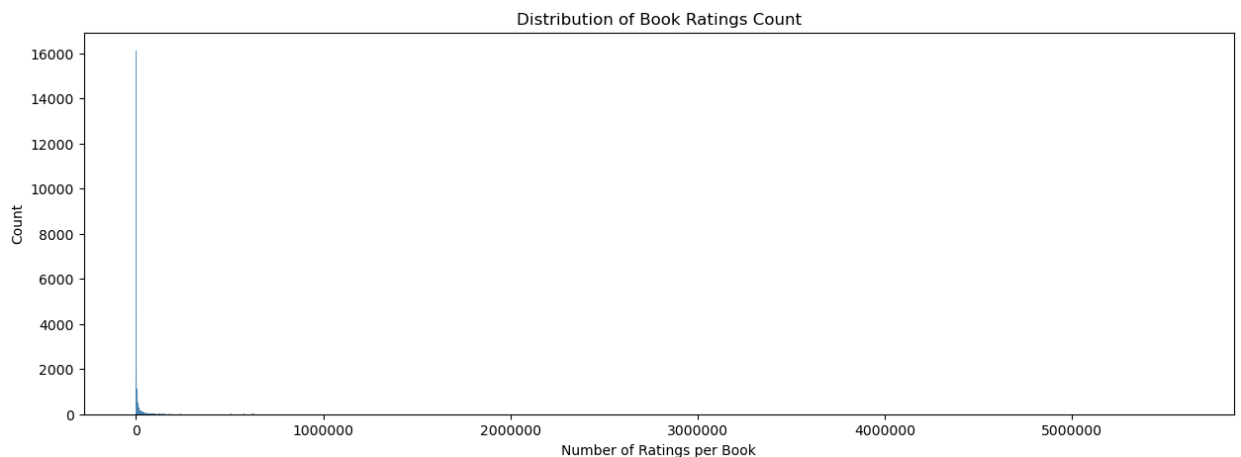
Out[25]: Text(0, 0.5, 'Count')



The graph shows a nicely normal distribution centered around 4 but it quickly tapers off to 3 and 5, effectively cutting off anything below a 3 rating. This shows that the users who rate books enjoyed them and that users are unlikely to put the effort into rating a book they did not enjoy. There is also a fun little spike at 5 which shows people love to rate their favorite books.

```
In [26]: 1 # Plot distribution of the number of ratings per book to see the s
2 plt.figure(figsize=(15,5))
3 sns.histplot(books['book_rating_count'])
4 plt.title("Distribution of Book Ratings Count")
5 plt.xlabel('Number of Ratings per Book')
6 plt.ticklabel_format(style='plain', axis='x')
7 plt.ylabel('Count')
```

Out[26]: Text(0, 0.5, 'Count')



When looking at the number of ratings each book had, a graph just was not showing much so a table is a better visual here. There are some books (like The Hunger Games) that were rated by over 550,000 users, while 75 books in the database have no ratings and 357 only have 1 rating. Such a difference made the scale of the graph so large, the data was invisible on it.

Overall there is fairly good showing of ratings for the books in the database. Based on the count earlier though, only 259 books in the database have a rating, so probably not going to be useful for this recommendation system as the data is too sparse to take into meaningful account.

```
In [27]: 1 # Allow all rows of dataframe to show
          2 pd.set_option('display.max_rows', None)
          3
          4 # Display the number of ratings and the number of books with that
          5 books.book_rating_count.value_counts().sort_index()
```

```
Out[27]: 0          75
          1        357
          2        373
          3        353
          4        278
          5        218
          6        209
          7        180
          8        187
          9        179
         10        157
         11        165
         12        162
         13        143
         14        156
         15        135
         16        110
         17        106
         18        101
         19         99
```

▼ Data Cleaning and Processing

```
In [28]: 1 # Trim dataframe to needed columns for book title, description, and
          2 books = books[["book_title", "book_authors", "book_desc", "genres"]]
```

```
In [29]: 1 # Count null values in each column.  
2 books.isna().sum()
```

```
Out[29]: book_title          0  
book_authors          0  
book_desc          1331  
genres          3242  
book_rating_count      0  
dtype: int64
```

```
In [30]: 1 # Drop rows with missing data  
2 books = books.dropna()
```

```
In [31]: 1 # Group same named books together to avoid duplicated results  
2 # Grab the description and rating count from first encountered line  
3 books = books.groupby("book_title", as_index=False).agg({"book_title":
```

```
In [32]: 1 # Re-determine number of unique values in each column after trimming  
2 books.nunique(axis=0)
```

```
Out[32]: book_title          44526  
book_authors          23552  
book_desc          44129  
genres          28423  
book_rating_count      19316  
dtype: int64
```

```
In [33]: 1 # Add quotes to beginning and end of book description  
2 books["book_desc"] = books["book_desc"].apply(lambda x: "'" + str(x) + "'")
```

▼ Set Up Model and Access Function

```
In [34]: 1 # Define characteristic feature and vectorizing model  
2 feature = books["book_desc"].tolist()  
3 tfidf = text.TfidfVectorizer(stop_words="english")
```

```
In [35]: 1 # Form matrix from model and calculate similarity  
2 tfidf_matrix = tfidf.fit_transform(feature)  
3 similarity = linear_kernel(tfidf_matrix, tfidf_matrix)
```

```
In [36]: 1 # Set book title as search parameter  
2 indices = pd.Series(books.index,  
3                     index=books['book_title']).drop_duplicates()
```

```
In [37]: 1 # Define a function to return up to 20 recommended titles
2 # Sort based on similarity to entered title
3 def book_recommendations(title, similarity = similarity):
4     index = indices[title]
5     similarity_scores = list(enumerate(similarity[index]))
6     similarity_scores = sorted(similarity_scores, key=lambda x: x[1])
7     similarity_scores = similarity_scores[0:20]
8     bookindices = [i[0] for i in similarity_scores]
9     recommendations = pd.DataFrame(books[['book_title', 'book_autho
10     return recommendations
```

▼ Perform Query and Return Results


```
In [38]: 1 # Enter book to search for recommendations within quotes
2 # Results will be displayed in a dataframe with corresponding genres
3 book_recommendations("The Da Vinci Code")
```

Out[38]:

	book_title	book_authors	genres
30961	The Da Vinci Code	Dan Brown	Fiction Mystery Thriller
2783	Angels and Demons / The Da Vinci Code	Dan Brown	Fiction Mystery Thriller Historical Historical...
34365	The Lost Symbol	Dan Brown	Fiction Mystery Thriller
22419	Oprindelse	Dan Brown	Fiction Thriller Mystery Thriller Mystery Thri...
2774	Angels & Demons	Dan Brown	Fiction Mystery Thriller
2775	Angels & Demons - Malaikat dan Iblis	Dan Brown Isma B. Koesalamwardi	Fiction Mystery Thriller
21792	O Código Da Vinci	Dan Brown Celina Cavalcante Falck- Cook	Fiction Mystery Thriller
31282	The Devil's Chord	Alex Archer Michele Hauf	Fantasy Fiction Fantasy Urban Fantasy Action A...
37127	The Smile	Donna Jo Napoli	Historical Historical Fiction Young Adult Hist...
10645	Evil in the Beginning	Gary Williams Vicky Knerly	Mystery Thriller Adventure Fiction
17844	Leonardo da Vinci	Walter Isaacson	Biography Nonfiction History Art Science
16681	King Dork	Frank Portman	Young Adult Fiction Humor Young Adult Teen Mus...
3790	Be Great!: 365 Inspirational Quotes from the W...	Daniel Willey	Classics
17846	Leonardo, the Terrible Monster	Mo Willems	Childrens Picture Books Childrens Childrens St...
8831	Digital Fortress	Dan Brown	Fiction Thriller Mystery Suspense
17845	Leonardo's Notebooks	Leonardo da Vinci H. Anna Suh	Art Nonfiction History Science Classics Biography
38082	The Uncanny	Sigmund Freud Adam Phillips David McLintock Hu...	Nonfiction Psychology Philosophy Philosophy Th...
40473	Unbreakable	Kami Garcia	Young Adult Fantasy Paranormal Fantasy Paranor...
33190	The Hourglass Door	Lisa Mangum	Young Adult Fantasy Romance Science Fiction Ti...
29410	The Aylesford Skull	James P. Blaylock	Science Fiction Steampunk Fantasy Science Fict...

```
In [39]: 1 # Running again with a different book to verify different results
2 book_recommendations("A Christmas Carol")
```

Out [39] :

	book_title	book_authors	genres
416	A Christmas Carol	Charles Dickens Joe L. Wheeler	Classics Fiction Holiday Christmas Fantasy Lit...
417	A Christmas Carol and Other Christmas Writings	Charles Dickens Michael Slater	Classics Fiction Holiday Christmas Short Stori...
418	A Christmas Carol, The Chimes and The Cricket ...	Charles Dickens Katharine Kroeber Wiley	Classics Fiction Holiday Christmas Literature
95	12 Stocking Stuffers	Beverly Barton Helen Bianchin Janelle Denison ...	Romance Contemporary Holiday Christmas Antholo...
30393	The Christmas Box	Richard Paul Evans	Holiday Christmas Fiction Holiday Inspirational
11309	Finding Noel	Richard Paul Evans	Holiday Christmas Fiction Romance Holiday
400	A Charlie Brown Christmas	Charles M. Schulz	Holiday Christmas Childrens Childrens Picture ...
4101	Belstarr The Lost Toymaker	David Jacks Daniel S. Morrow	Holiday Christmas Childrens Picture Books
30391	The Christmas Basket	Debbie Macomber	Holiday Christmas Romance Holiday Fiction Wome...
6321	Christine Kringle	Lynn Brittney	Holiday Christmas
21604	North Pole Reform School	Jaimie Admans	Young Adult Holiday Christmas Fantasy Romance ...
32142	The First Christmas Carol	Marianne Jordan	Holiday Christmas Christian Fiction Christian ...
34124	The Life and Times of Scrooge McDuck	Don Rosa	Sequential Art Comics Sequential Art Graphic N...
30394	The Christmas Box Collection: The Christmas Bo...	Richard Paul Evans	Holiday Christmas Fiction Romance
12878	Grace	Richard Paul Evans	Holiday Christmas Fiction Romance Holiday Youn...
16799	Kissing Under the Mistletoe	Bella Andre	Romance Romance Contemporary Romance Contempor...
21603	North Pole High: A Rebel Without a Claus	Candace Jane Kringle	Holiday Christmas Romance Young Adult Humor Fi...
30392	The Christmas Books, Volume 1: A Christmas Car...	Charles Dickens Michael Slater	Classics Fiction Holiday Christmas Short Stories
1674	A wartime Christmas	Carol Rivers	Historical Historical Fiction Holiday Christma...

The results of this model are much nicer than the previous approach, but it too has some drawbacks. The recommendations often include the title submitted and variations there of. This makes sense as there is nothing preventing it from doing so and the variations would be similar to the original from the descriptions.

References

Aman Kharwal. July 19, 2022. *Book Recommendation System using Python*.
<https://thecleverprogrammer.com/2022/07/19/book-recommendation-system-using-python/>
(<https://thecleverprogrammer.com/2022/07/19/book-recommendation-system-using-python/>)

Book Recommendation Dataset. (2024, February 9).
<https://www.kaggle.com/datasets/arashnic/book-recommendation-dataset>
(<https://www.kaggle.com/datasets/arashnic/book-recommendation-dataset>)

Jayashree Domala. April 8, 2021. *Movie Similarity Recommendations Using Python*.
<https://betterprogramming.pub/movie-similarity-recommendation-using-python-b98a2670a2ad> (<https://betterprogramming.pub/movie-similarity-recommendation-using-python-b98a2670a2ad>)

Naren. *Goodreads Best Books Ever*. Kaggle.
<https://www.kaggle.com/datasets/meetnaren/goodreads-best-books>
(<https://www.kaggle.com/datasets/meetnaren/goodreads-best-books>)

Sisodia, R. (2021a, December 11). Movie recommendation system. Medium.
<https://medium.com/@rahulsisodia06/movie-recommendation-system-c8113226c0aa>
(<https://medium.com/@rahulsisodia06/movie-recommendation-system-c8113226c0aa>)