Lab Assignment 7

Objective: To create a book recommender which can be used to recommend similar books to the reader based on his/her interest. The books recommendation system is used by online websites which provide ebooks like google play books, open library, good Read's, etc.

1. Load dataset book_ratings, books and users
2. Data Pre-processing:
    - Impute the missing values present in categorical columns and numerical columns
    - Create a new dataframe "Combine_Book_Ratings by merging books csv file with their ratings csv file using pd.merge( )
    - Calculate total rating count for the books based on the book_ratings columns, using groupby ( ) function.
    - Filter out the lesser known books using total rating count. You need to take only those books for which total rating count is greater than or equal to 50.
    - Also filter the users based on the location. Take into consideration the users of "usa" and "Canada".
3. Convert your table into a 2D matrix using userid and movie title and fill the values with movie ratings. Fill the null values with 0 for the cases where user has not given a rating to a book. Hint: Use pivot function where index="Book-Title",columns="User-ID",values="Book-Rating".
4. Then transform the values(ratings) of the matrix dataframe into a scipy sparse matrix (use: from scipy.sparse import csr_matrix) for more efficient calculations.
5. Define model using sklearn.neighbors to find out 5 most similar books for the given book based on the title.
6. Test your model using by giving any random book_title and see the books suggestions returned by your model.
7. *For Fun:* Scrap books data from goodreads.com. Construct a content-based filtering book recommendation system.