1) Extract BX-CSV-Dump.zip (Dataset)

2) Run ESVD_1.ipynb in jupyter notebook.

In this notebook we have first extracted intersection of top 1000 most rated books and top 1500 most active users and applied SVD on zero ratings to get denser dataset.

(Note: the last loop which is used there to update the ratings in original matrix takes time about 1.5hrs, that's why we have decided to run it once a month to get updated denser rating csv file.) To see our updated csv extract denser_ratings.zip.

3) Run ESVD_2 for CF and recommendation of books.

In this notebook we applied SVD on the denser ratings matrix and predicted missing ratings for a given user (we've shown for user 8). We've shown both the top-rated books by user along with author names and also top recommended books for user after applying SVD.

At last step we've gave Final recommendation of 10 books. (6 from user's favourite authors and 4 other books not related to that author)