INDENG 243 Project - Module 3

GoodReads: Book Analytics & Recommendation System User Interface

---

Agenda

- System Introduction
- Visualization & Exploratory Analysis
- Recommendation System User Interface
- Summary & Future Plans

Part 1) System Introduction

In order to provide better user experience with our Book Analytics and Recommendation system, our group decided to develop an interactive web-based user interface and deploy it online such that every user can access our system and enjoy the functions to help build their next to-read book lists.

You can directly access our system online through the following link:
GoodReads: Book Analytics & Recommendation System website

** In case you have any problems accessing the website, it might because our web page meets the resource limit and need to reboot it for new users. In that case, please contact us by email to let us know, thank you.*

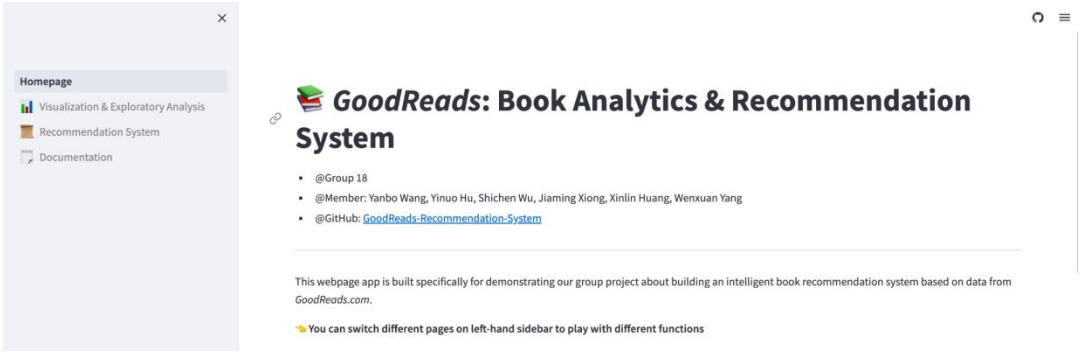


Fig 1.1: Snapshot of system welcome page

1.1 System Pages Overview

The system can be broken down mainly into three parts (see Fig 1.2):

1) **Visualization and Exploratory Analysis** of book information and reader reviews data from the source website *GoodReads.com;*

2) **Recommendation System** which will provide a list of book recommendations for book lovers given certain user input, and demonstrates helpful snapshots of recommended books' information;

3) **Documentation** about the technical details on building the website and the recommendation models.

You can switch on the left-hand sidebar on the welcome page to play with different functions as you like.

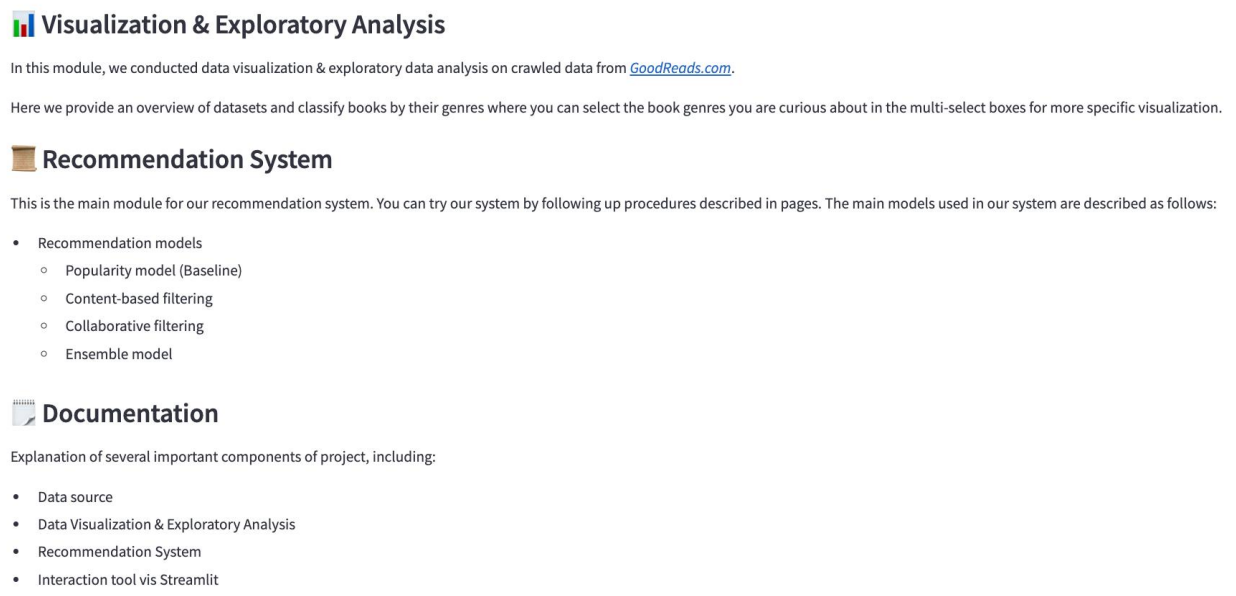


Fig 1.2: System pages overview

1.2 Target Audience & Impacts

Our system provides services for several categories of audience and help make beneficial impacts to their reading or research experiences, including:

1) **Researchers and professionals** with technical background and expertise who wish to explore more reading patterns and statistical trends of book lovers for research use. For this group, they can investigate our "Visualization & Exploratory Analysis" page to view clear visualizations of the datasets and insightful analytics from various perspectives to comprehensively grasp readers' preferences.

2) **Non-technical book lovers** who simply wish to build their next to-read book lists. For this group, they can use our "Recommendation System" page and follow the prompts for several questions about their past reading history, and then our system will output a list of recommended books for them, which is easier to use and highly accurate to cater for different users' reading tastes.

Part 2) Visualization & Exploratory Analysis

In order to better assist book enthusiasts in understanding the information contained in our database system, we conducted a visualization of the data in Module 1 and integrate this part in our interactive website. We invite you to explore our Jupyter notebook for more information on visualization and exploratory analysis.

2.1 Overview of All Books

To gain a broad understanding of books, we firstly developed an overview visualization of important book metrics and explored their relationships on the "Visualization & Exploratory Analysis" page.

Here are some topics we conclude in this part:

- Average Rating v.s. Genres
- Rating distribution among genres
- Rating/Review Number v.s. Genres
- Awards across genres

The following are some visualization examples:

- The bubble chart helps users to learn more about relationships between the ***review number*** and ***rating*** for books of different genres.

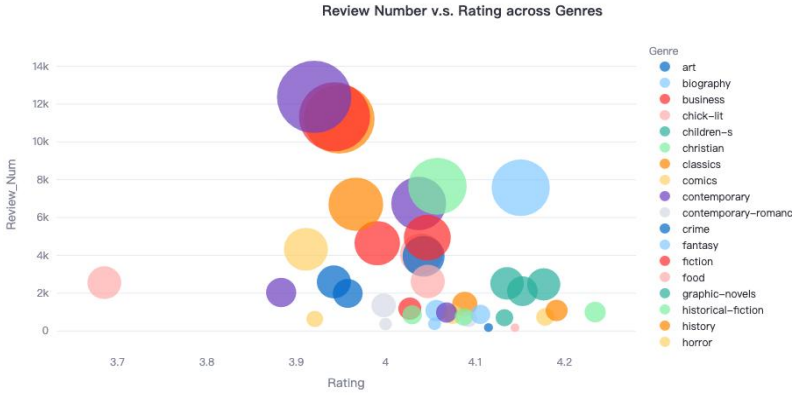


Fig 2.1: Review Number v.s. Rating across Genres

- To visualize genre distributions in individual awards, our system enables users to choose the awards they are curious about and give back the results, which is friendly for user interactions.

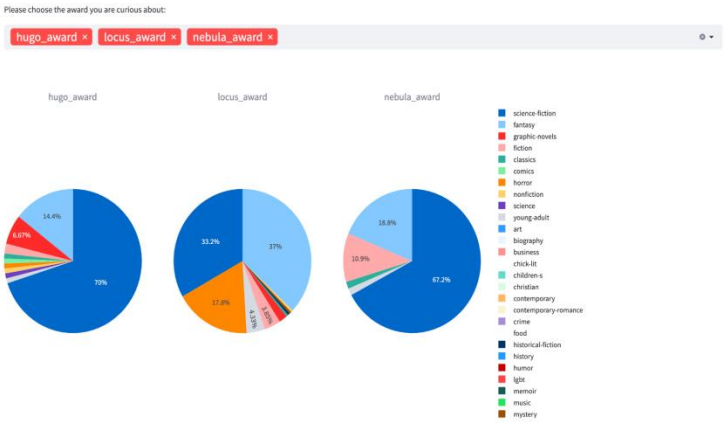


Fig 2.2: Genre distribution in individual awards

2.2 Books of Selected Genres

In this section, we will look into the visualization and exploratory analysis by genres. By using this interactive visualization page, readers will easily comprehend the distribution of ratings and the prevailing trend in the genres that they find interesting.

In order to start the analytical process, we provide a selection box for users to select a book genre of interest. A dashboard will show the general rating information of this genre (see Fig 2.2), including the average rating, average rating number, average review number and the rating distribution (1~5 stars percentage) of the selected genre.

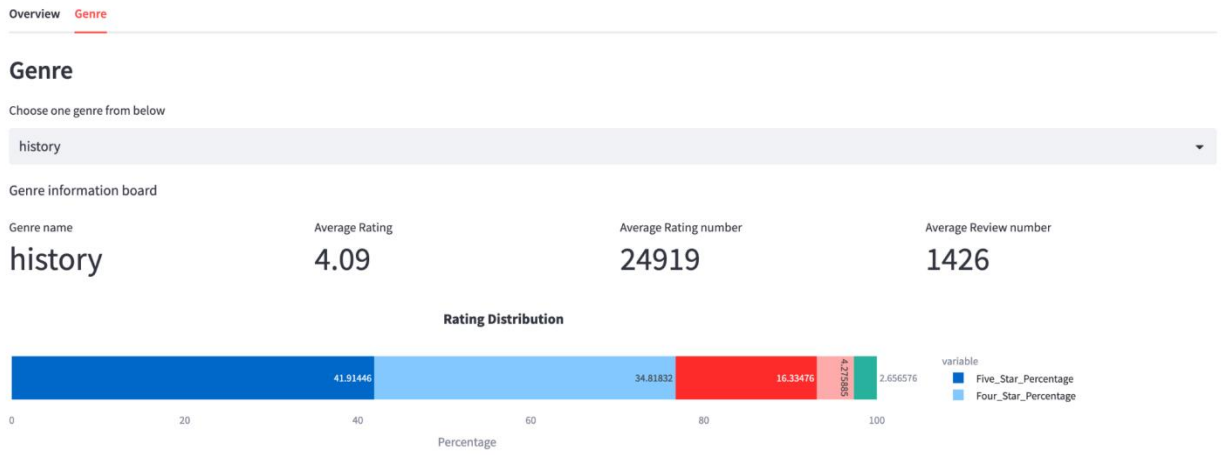


Fig 2.2: book information dashboard of selected genre

As winning awards signifies the official recognition of a book's merit, we compile a list of the total number of awards won by books in this genre and arrange them in descending order. Users can refer to the most commonly awarded prizes in the genre of interest as a guide to select well-regarded books.

Award_Name	Num
pulitzer_prize	27
national_book_critics_circle_award	22
national_book_award_finalist	15
los_angeles_times_book_prize	13
cundill_history_prize	10
national_book_award	10
ambassador_book_award	8
mark_lynton_history_prize	8
samuel_johnson_prize_for_nonfiction	7
lionel_gelber_prize	7
hesselltiltman_prize	7
duke_of_westminster_medal_for_military_literature	7
wolfson_history_prize	6

Fig 2.3: Award & Number of times winning in selected genre

We also enable users to refer to visualizations for statistics about a specific book in their preferred genre. This feature enables users to track a book's trend by examining the total number and time-based statistics of added to shelves, ratings, reviews, and want-to read counts (see Fig 2.4).

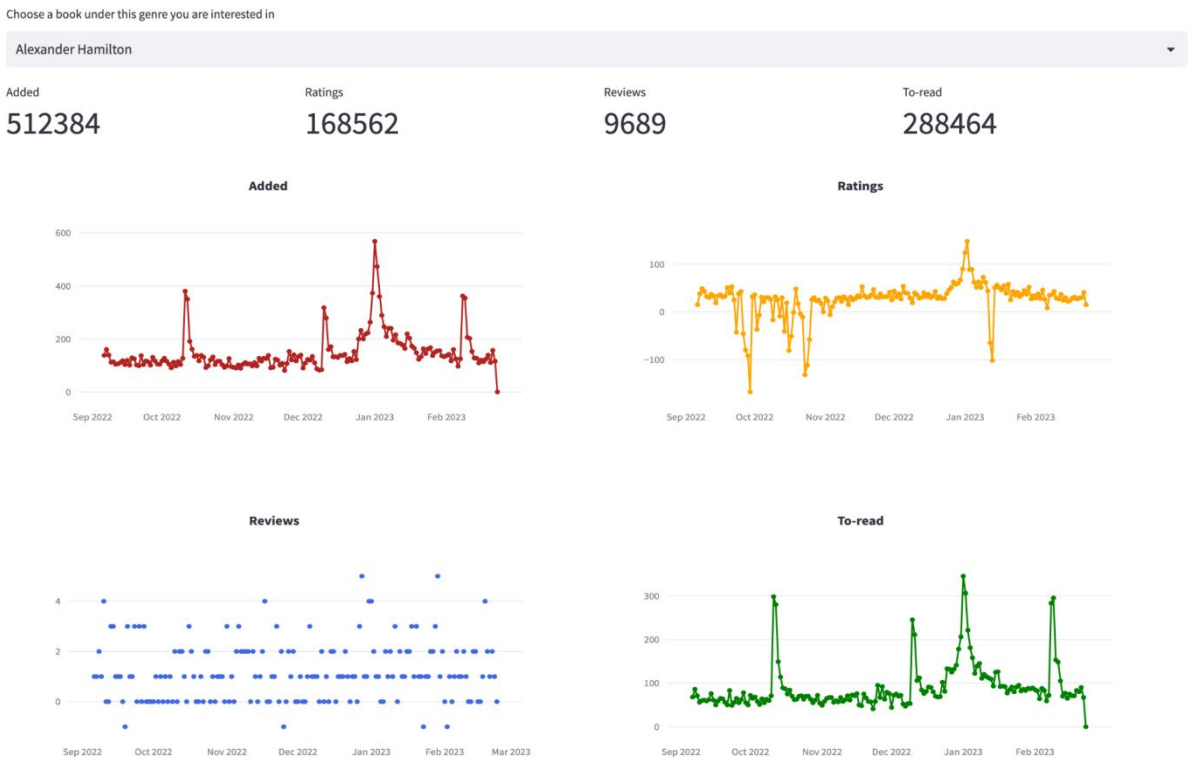


Fig 2.4: Statistics of a specific book

Part 3) Application of Recommendation System

The recommendation system page of our website is an interactive platform built for a non-technical audience. Based on users' input, it will be able to offer personalized book recommendations. This platform can be directly adopted in a variety of real-life situations such as bookstores.

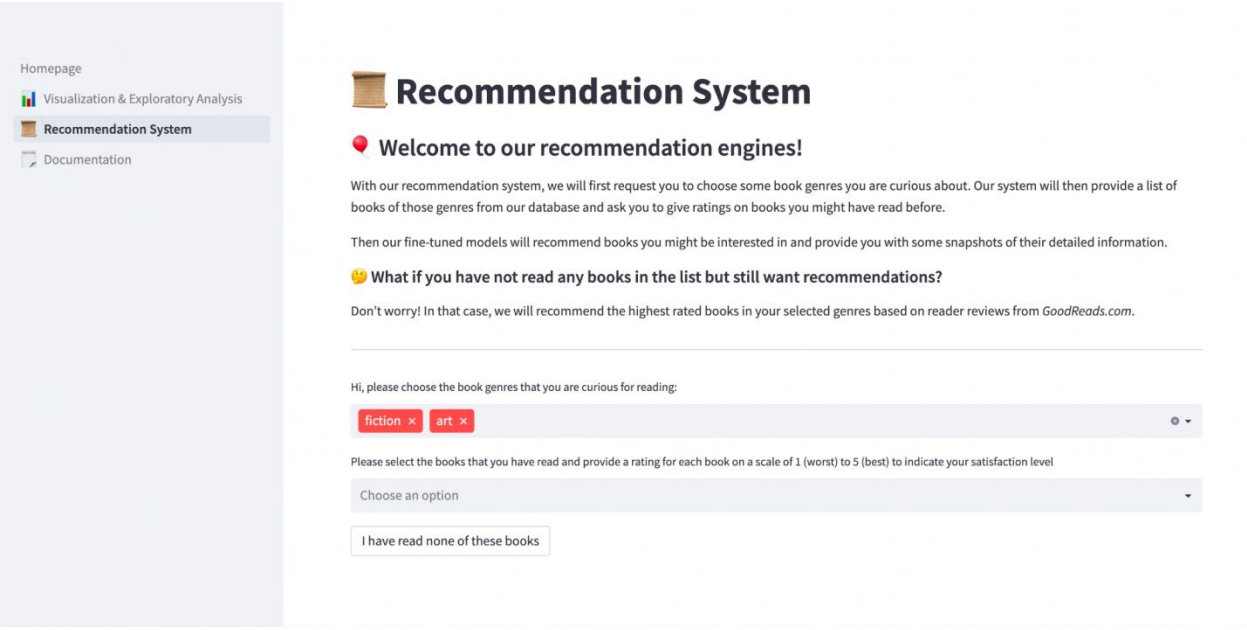


Fig 3.1: Interactive recommendation system page

3.1 Collect User Information

Our recommendation system will first ask users which book genres they are interested in and then the system will display the books of those selected genres in our database and ask users whether they have read any of them.

- **If the user has come across none of them before**, he/she can click the button "**I have read none of these books**" In this case, the system will recommend the top ten books with the highest ratings in selected genres.

- **If the user has already read some of the books shown**, they can select the ones they have read and give ratings from 1 (worst) to 5 (best). This information will be the input for back-end recommendation models and help output more personalized recommendations.

For instance, suppose we want to get recommendations to explore more books in **science fiction**.

- **Step 1**: Select **science fiction** from the first drop-down box asking for genres;
- **Step 2**: Suppose we have read ***The Threebody Problem*** and ***Way Station***, so we should select these two in the second drop-down box;

Here, we can surely take a glance at all the books to find them, but that could be time-consuming. ***Instead, our recommendation system allows you to quickly find the books by simply typing part of the title.***

As illustrated in Fig 3.2, if we type "way", *Way Station* will pop up there.

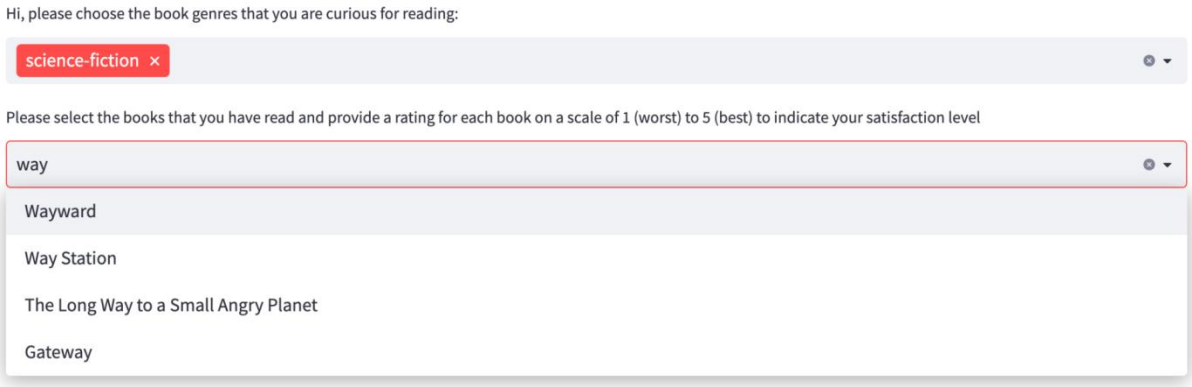


Fig 3.2: Select book genres & Quick search for books

- **Step 3**: Then we can just rate each of the books from 1 to 5;

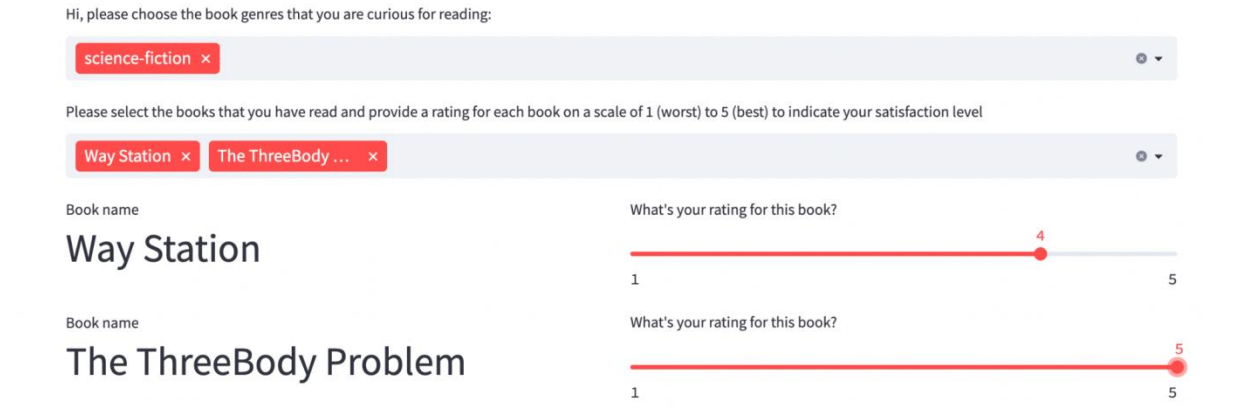
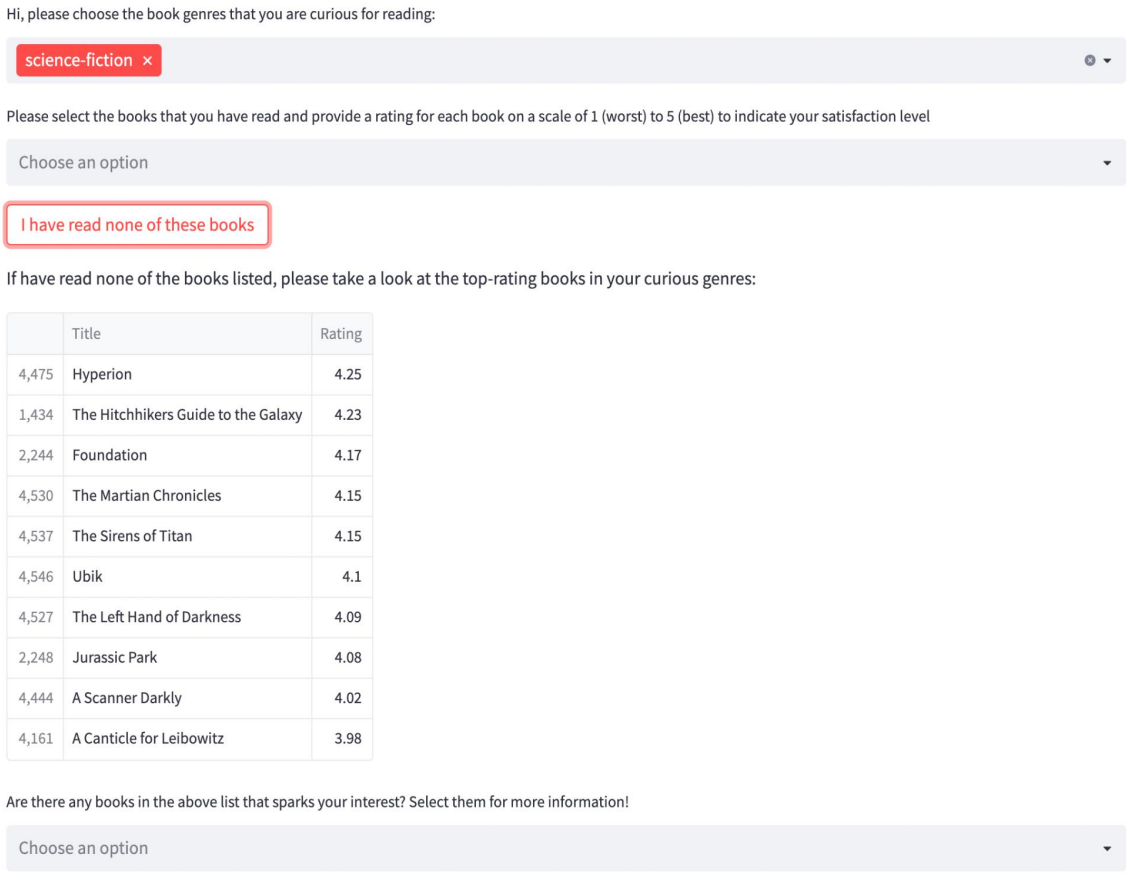


Fig 3.3: Rate books read before

But what if some readers have not read any of the books provided?

Under such cases, users can simply click the button "**I have read none of these books**"under the second drop-down box illustrated by Fig 3.4 (As detailed later, we can select books from the list to get more detailed information).



	Title	Rating
4,475	Hyperion	4.25
1,434	The Hitchhikers Guide to the Galaxy	4.23
2,244	Foundation	4.17
4,530	The Martian Chronicles	4.15
4,537	The Sirens of Titan	4.15
4,546	Ubik	4.1
4,527	The Left Hand of Darkness	4.09
2,248	Jurassic Park	4.08
4,444	A Scanner Darkly	4.02
4,161	A Canticle for Leibowitz	3.98

Fig 3.4: Recommendations if users read none of the books provided

3.2 Choose Recommendation Approach

After collected user's reading history, we provided three recommendation approaches as follows:

1) Based on your previous reading list and ratings: (**Content-based model**)
2) Based on readers with similar reading tastes: (**Collaborative filtering model**)
3) Based on the above two perspectives: (**Hybrid model**)

In a real-life situation, we will provide recommendations by taking results from these three models together. The users can choose to use different recommendation approaches by clicking the corresponding buttons (see Fig 3.5).

* Here we provide difference choices, because in a real-life situation, different users might have different preferences to find their next to-read books. Some users might want to listen to other users' advice and some might be more stick to his own tastes. To cater to different needs, we offer these choices to increase the flexibility of our system.

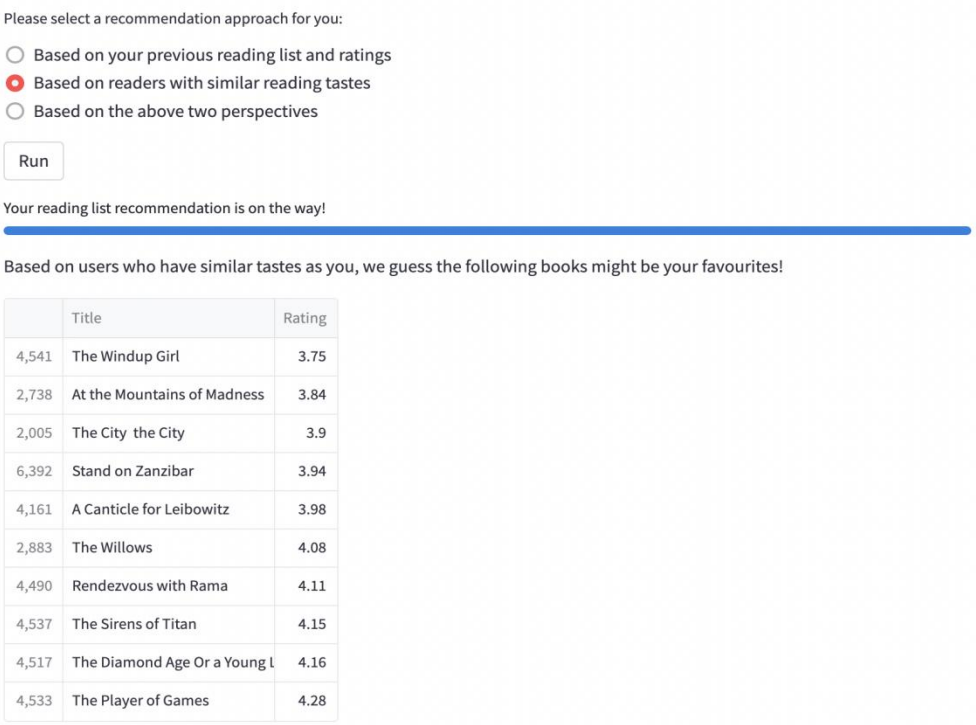


Fig 3.5: Recommendation approaches and the result list demo

For example, if we choose "**Based on readers with similar reading tastes**" and click "**Run**", the system will return the recommendation list of top 10 books that models guess the user is more probably like to read. Then the user can choose any books he is interested in for more detailed information:

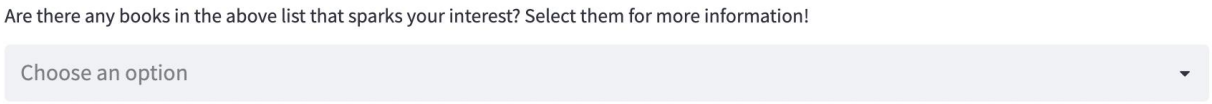


Fig 3.6: Drop-down box for choosing books the user has interest in

For example, he/she chooses ***the Player of Games*** and ***the Willows***:

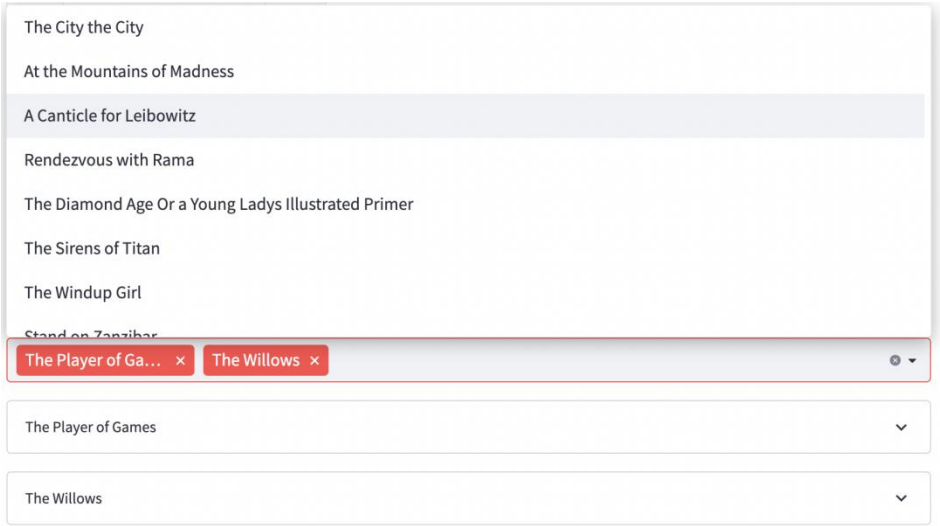


Fig 3.7 Two books chosen

If the user clicks each box with book name (Fig 3.7), it will show the information overview of each book, including the book basics, statistics, book tags and selected book reviews (see Fig 3.8):

- **Book basics**: Title, Author, Rating (also distribution), Review/Rating number;
- **Book statistics**: Times of Added to Shelves, Rating number, Reviews number, Times marked as To-read;
- **Book tags & selected reviews**: the top representative keywords and reviews extracted by NLP techniques.

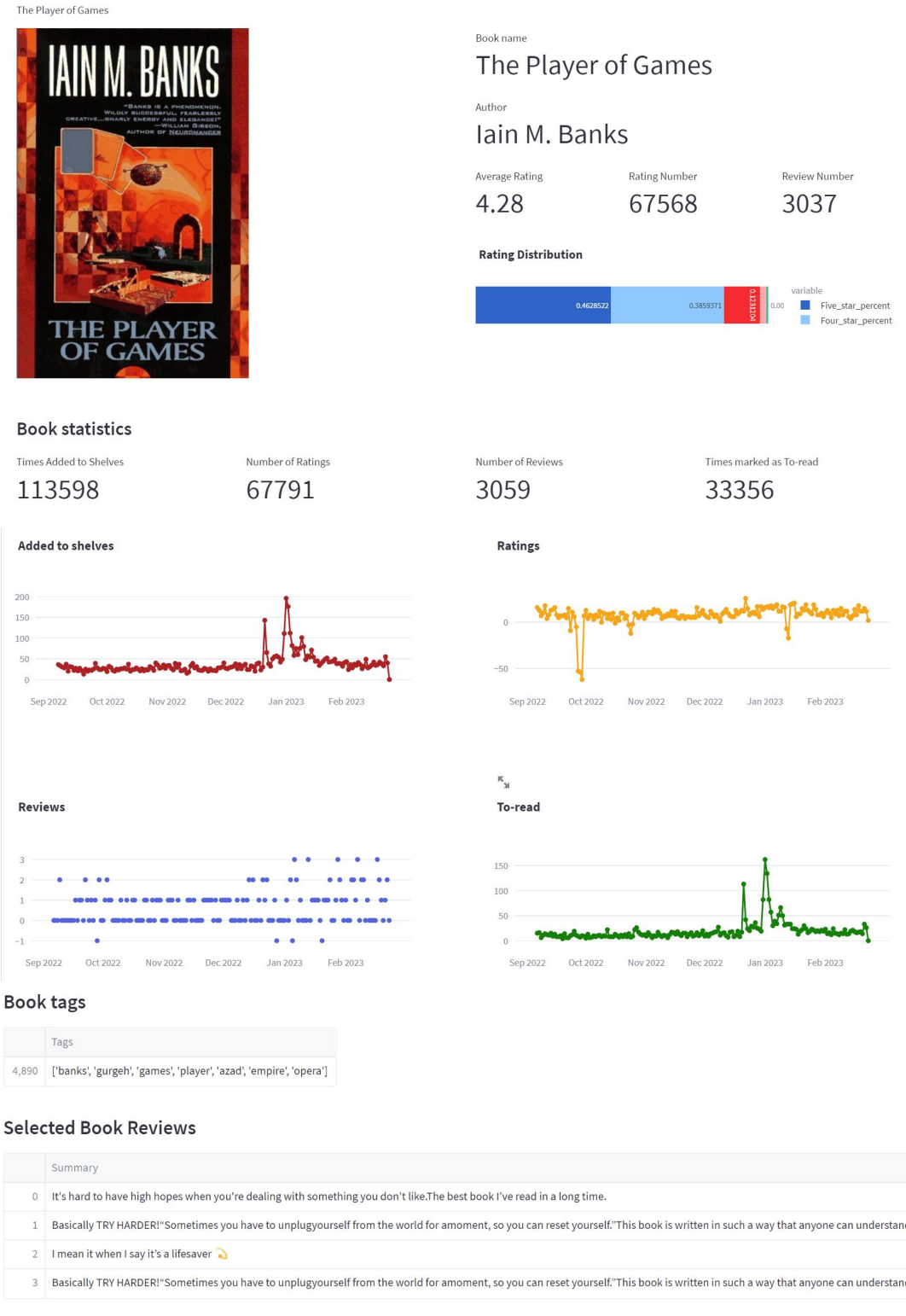


Fig 3.8 Book basic information, statistics, tags & selected reviews

Part 4) Summary & Future plans

In this module, we developed an online interactive user interface to truly help potential users more easily access our recommendation system and generate their next to-read book lists. We hope this website can at least contribute to users' better reading experiences in the following bullet points:

- Save time wasting in web searching for books cater to personal reading tastes;
- More accurate recommendation taking past reading history into account;
- User-friendly easy-use system to generate recommendations at few clicks;
- Automated information extraction to grasp a whole picture of books recommended, precisely targeting the next to-read books.

As a next step, we will consider to build a more fine-tuning system by:

- Broader our current book database and crawl more data from helpful websites to support analytics;
- Pursue a higher recommendation accuracy by considering more sophisticated models and retrieving more user interactions to provide more precise recommendations;
- Improve the appearance and functions of our website to build better experiences.