

Book Recommender Engines



By Jen Hill



The Issue



A bookstore is considering having a recommender engine added to their website. They'd like a look at what goes into building an engine, plus a couple examples of the different types available.



The Agenda

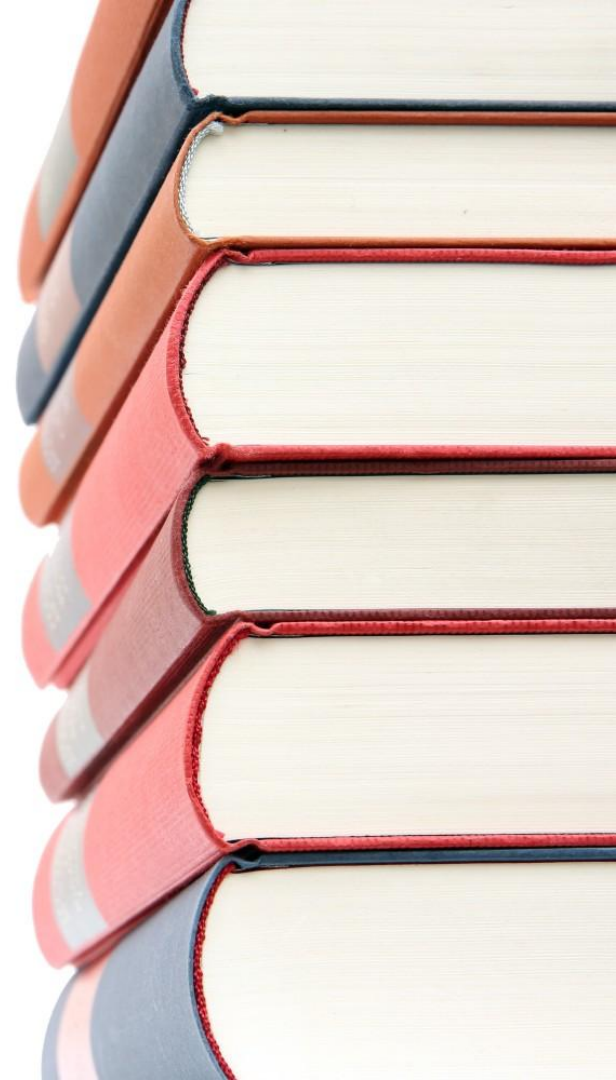


Collaborator Engine

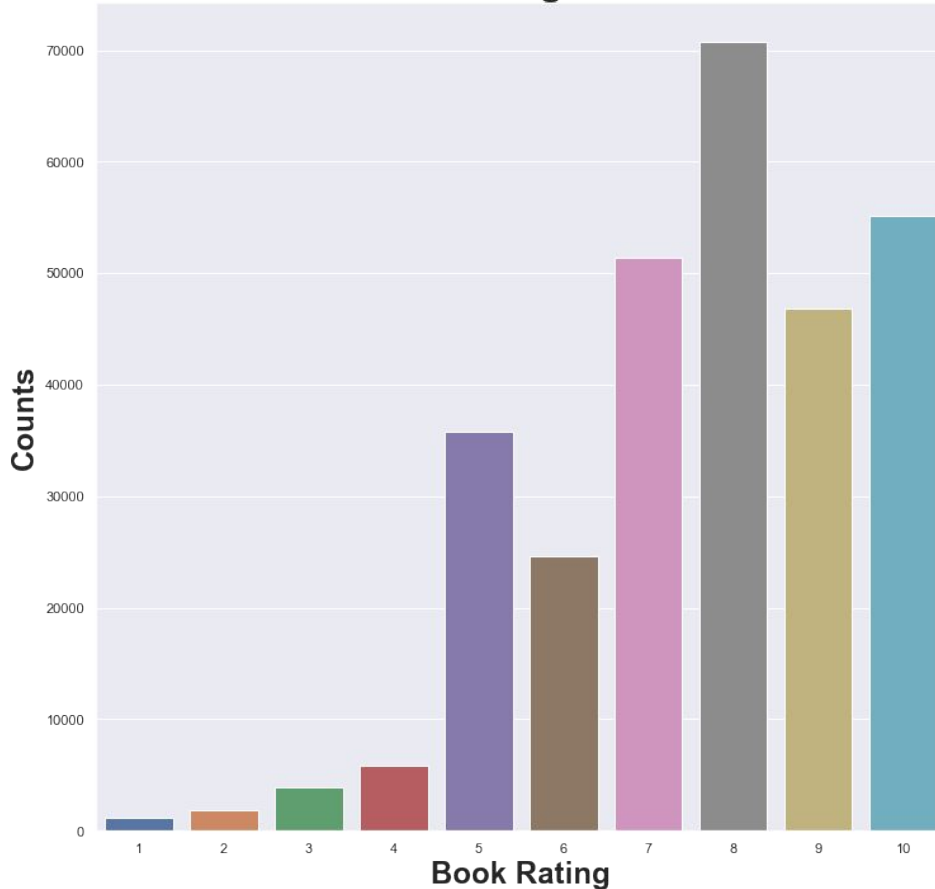
- Data Cleaning
- EDA
- Preprocessing
- Model
- Evaluation

Content Engine

- Data Cleaning
- EDA
- Preprocessing
- Model
- Evaluation



Book Rating Counts



Exploratory Data Analysis



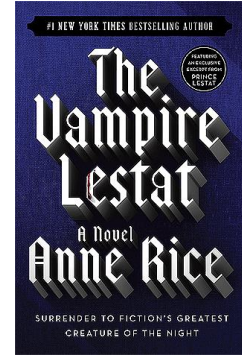
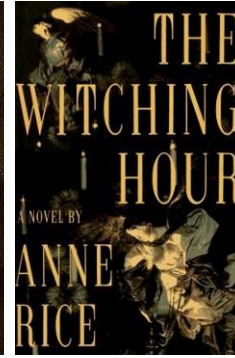
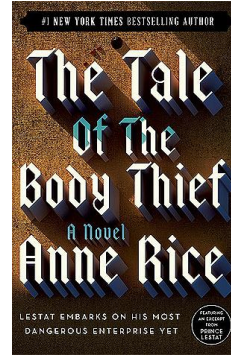
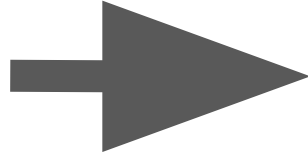
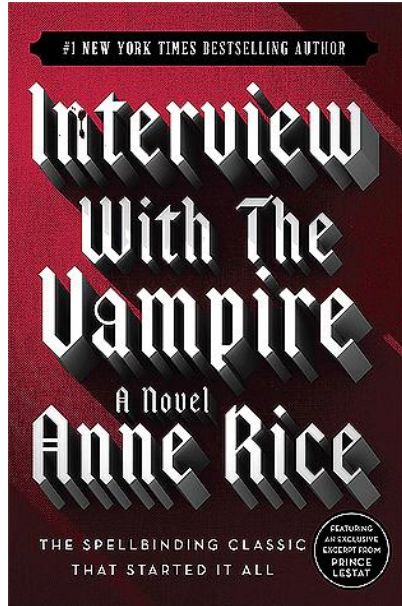
- Of the approximately 123k books rated, most were given a score of 8.
- Few books were rated below 5.

Top Rated Titles

Title	Total Ratings
The Lovely Bones	5188
The Da Vinci Code	3814
The Red Tent	2890
Harry Potter and the Sorcerer's Stone	2546
The Secret Life of Bees	2404

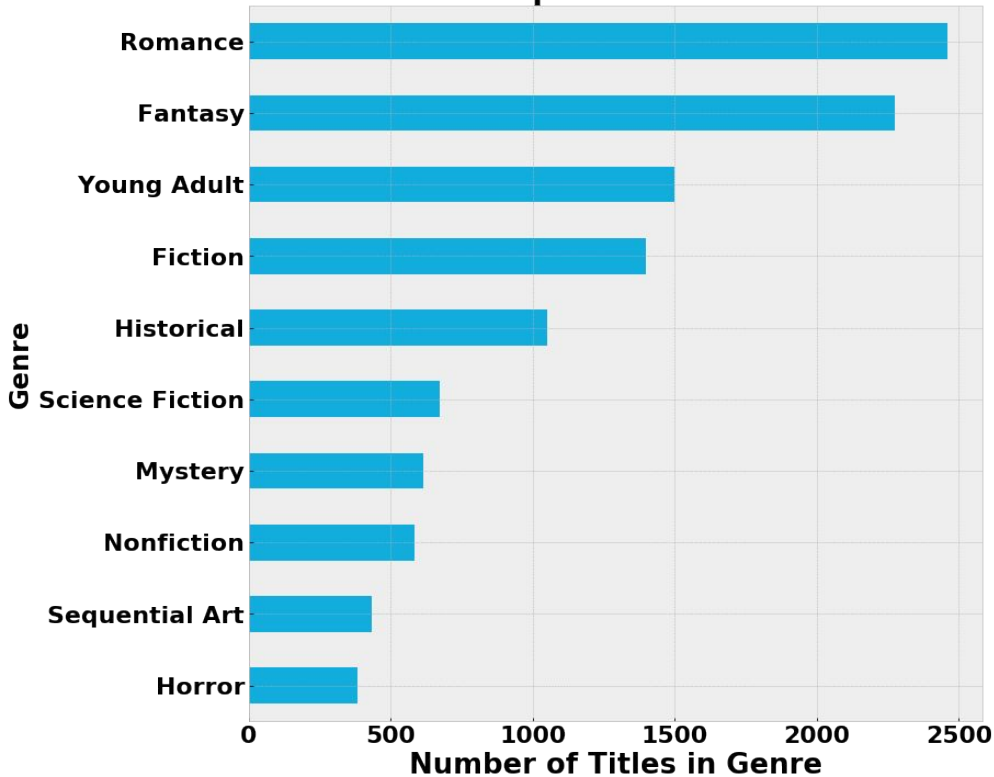
- This group of users rated The Lovely Bones much higher than all other titles

Recommender



- **The Tale of the Body Thief**
- Seinlanguage
- Candide
- Best American Poetry 1993
- Risen
- The Only Astrology Book You'll Ever Need
- **The Witching Hour**
- **The Vampire Lestat**

Top 10 Genres

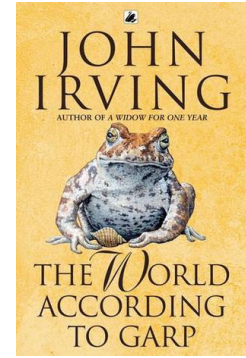
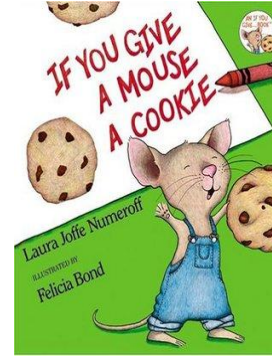
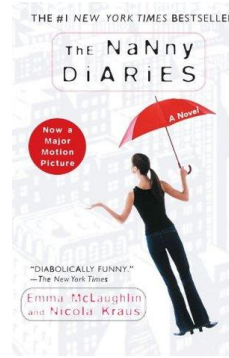
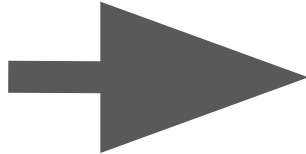
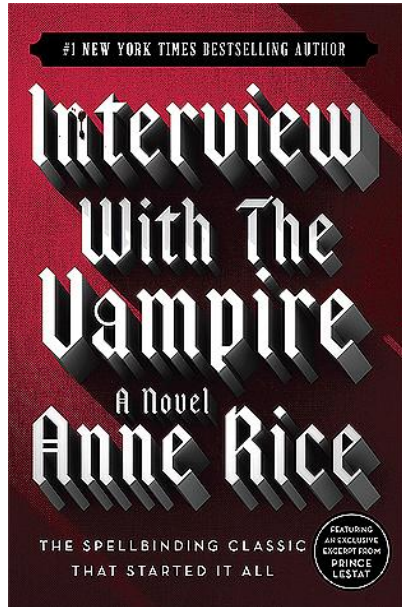


Exploratory Data Analysis



- The most popular books in the Goodreads data set are from the romance and fantasy genres.

Recommender



- The Nanny Diaries
- **If You Give a Mouse a Cookie**
- The World According to Garp
- The Bourne Identity
- Ella Enchanted
- The Purpose Driven Life: What on Earth Am I Here for?
- Breakfast of Champions
- Fables, Vol. 1: Legends in Exile

Conclusion



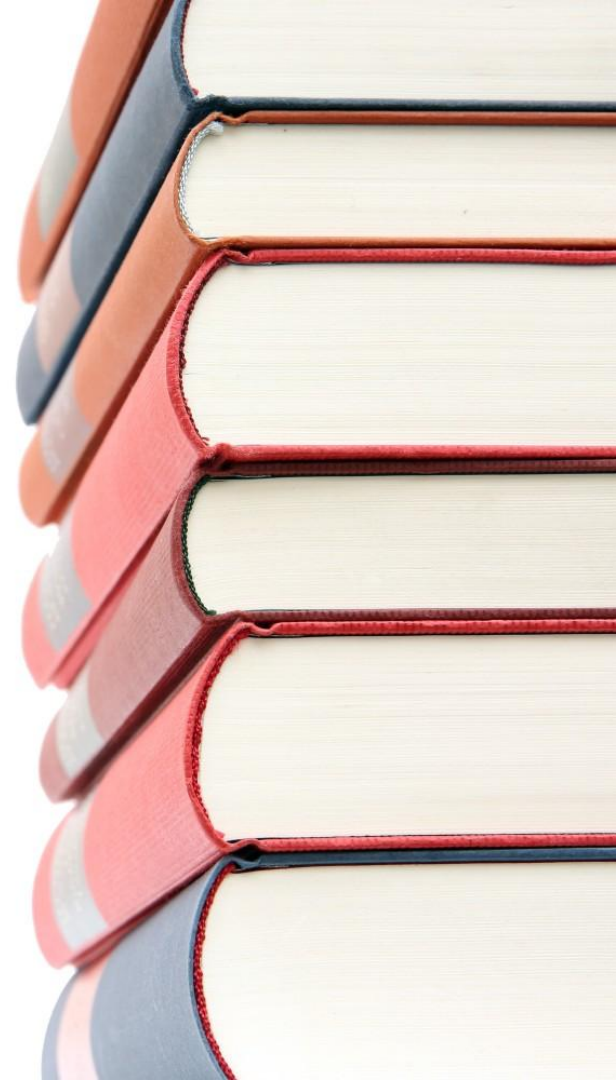
- **Recommendation:** Collaborator engine because it gave slightly better recommendations based on subject matter.
- **Alternate:** If user ratings can't be collected, a content-based can be created but I suggest adding more descriptive content to the data.



Next Steps



1. New Data: Ideally Bookstore Data
2. Address Foreign Titles
3. Test Extra Author Related Info
4. Explore Front-End Options for User Access



Thank You

