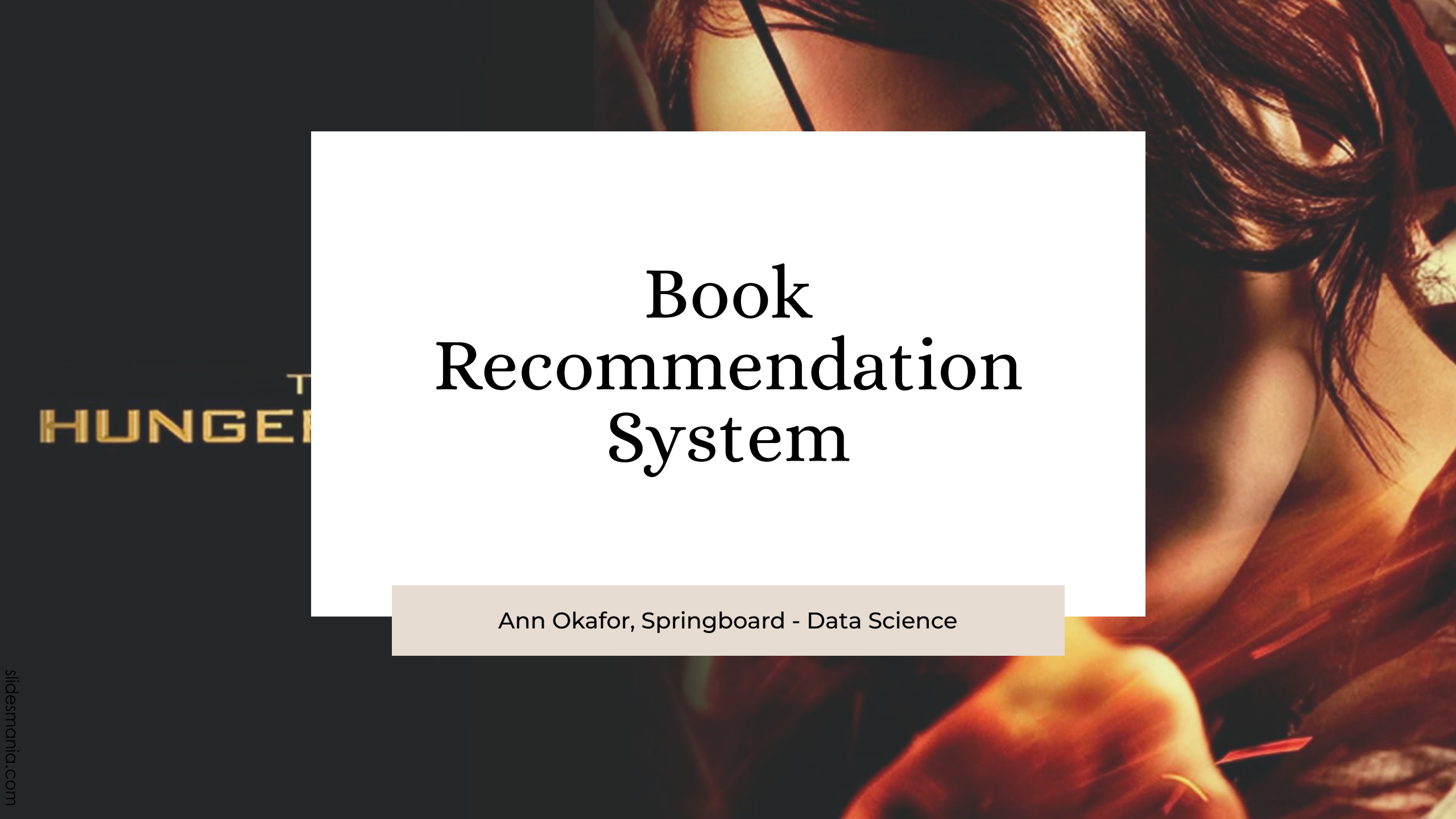


Book Recommendation System

Ann Okafor, Springboard - Data Science

Table of Contents

01

Project Overview

This project provides insights into the development of a book recommendation system using machine learning and data mining techniques.

Project Overview

- Challenge: Finding relevant books in a vast sea of options
- Solution: Book recommendation system using machine learning and data mining techniques
- Explored Recommendation Models: Content-based filtering, KNN (collaborative filtering), and random forest
- Model Selection: Random forest identified as the most accurate and reliable

02

Key Points

The key points of this project include exploring three recommendation models, data preprocessing and feature engineering, model selection, and evaluation of the recommendation models.

Key Points

- Three recommendation models explored: content-based filtering, collaborative filtering (KNN), and random forest.
- Random forest model identified as the most accurate and reliable.
- Evaluation conducted using a test set of user preferences.
- Project showcases the potential of leveraging data analysis, machine learning, and user preferences.

03

Appendix

Recommendation Approaches

Content-based filtering:

Focuses on book data and identifies similarities based on content

Collaborative filtering (KNN):

Examines user behavior and recommends books based on similar users

Random forest model:

Combines multiple features (titles, authors, genres, ratings) for accurate recommendations

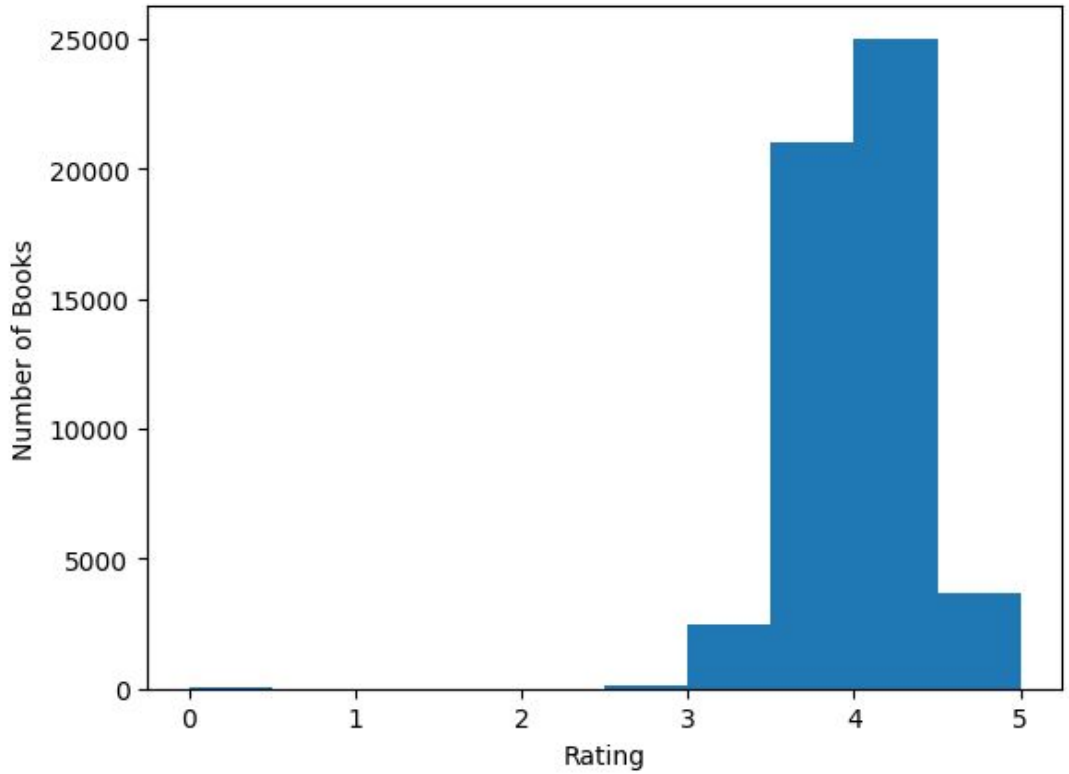
Data Collection and Preprocessing

- Book attributes (titles, authors, genres, ratings, descriptions) gathered from various sources
- Handling missing values, cleaning text data, transforming categorical variables into numeric representations
- Gain insights into the dataset and understand attribute distribution

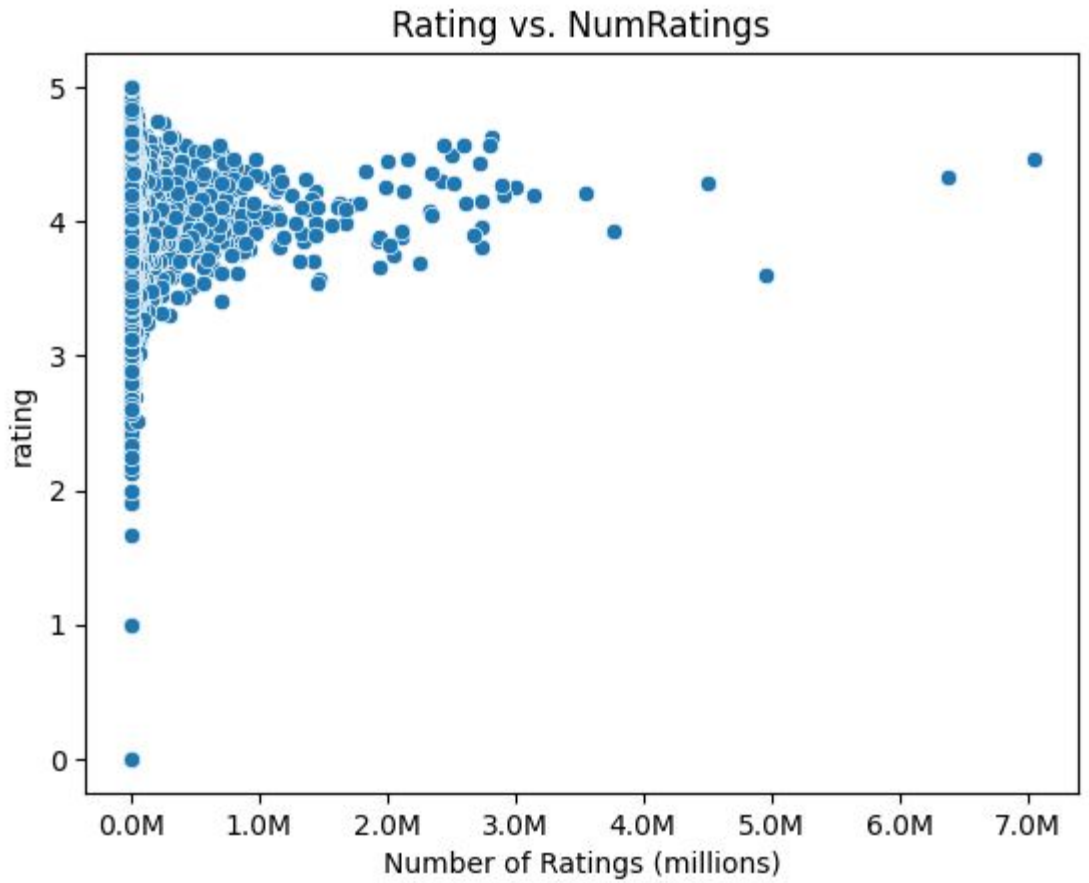
Distribution of Ratings

Lowest - 0

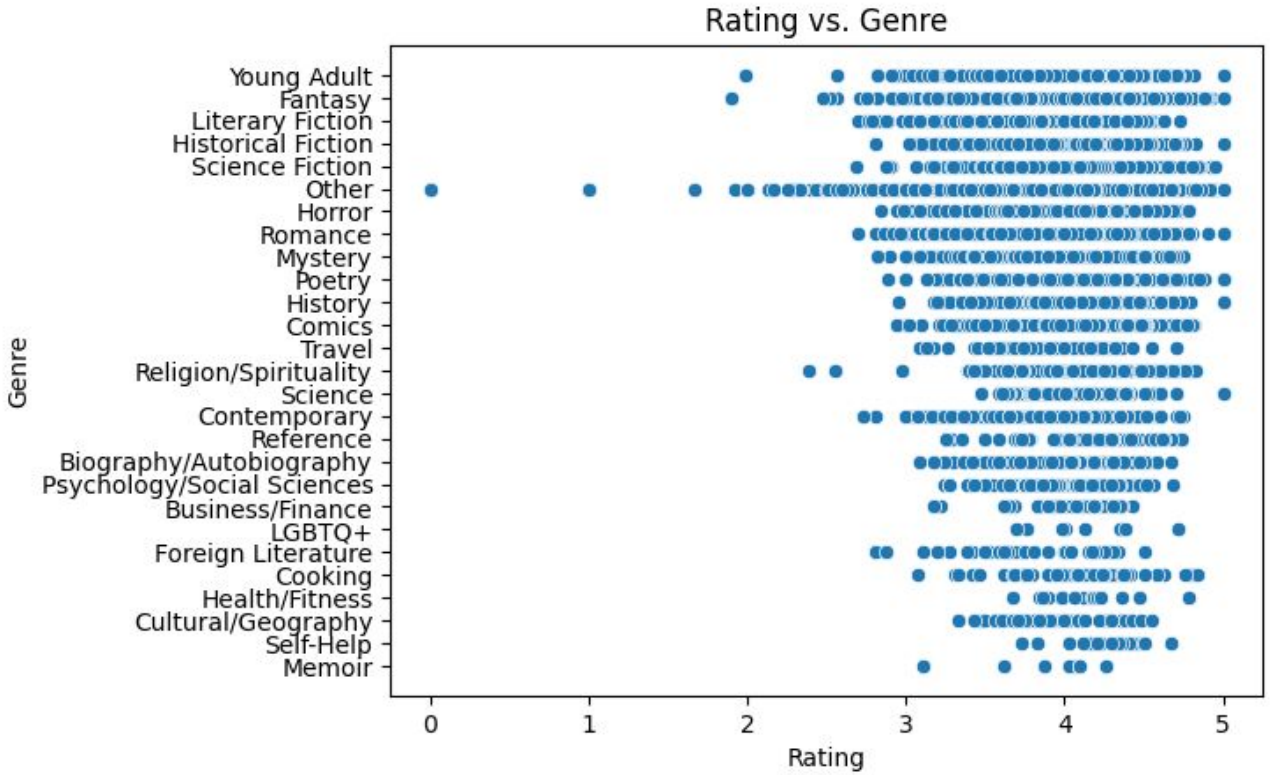
Highest - 5



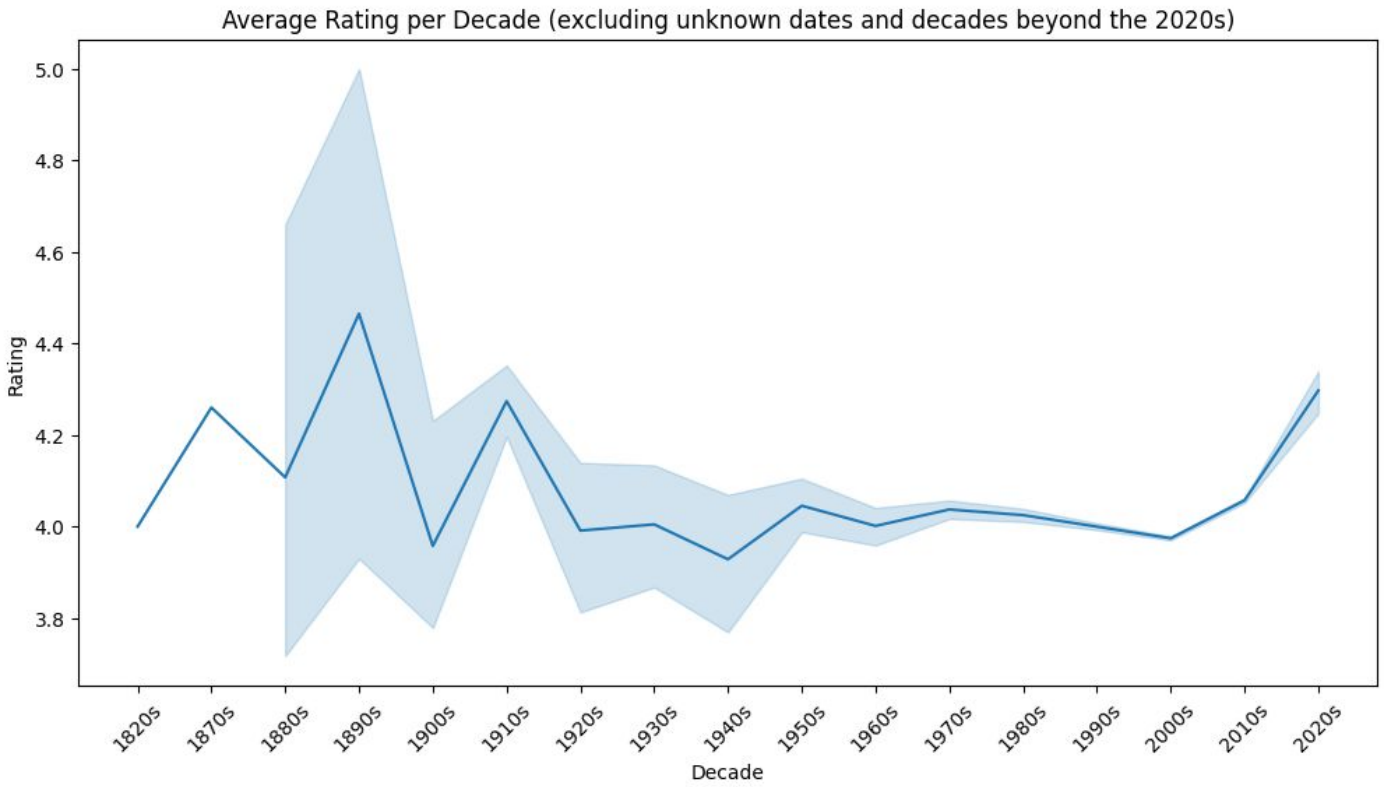
Rating vs. NumRatings



Rating vs. Genre



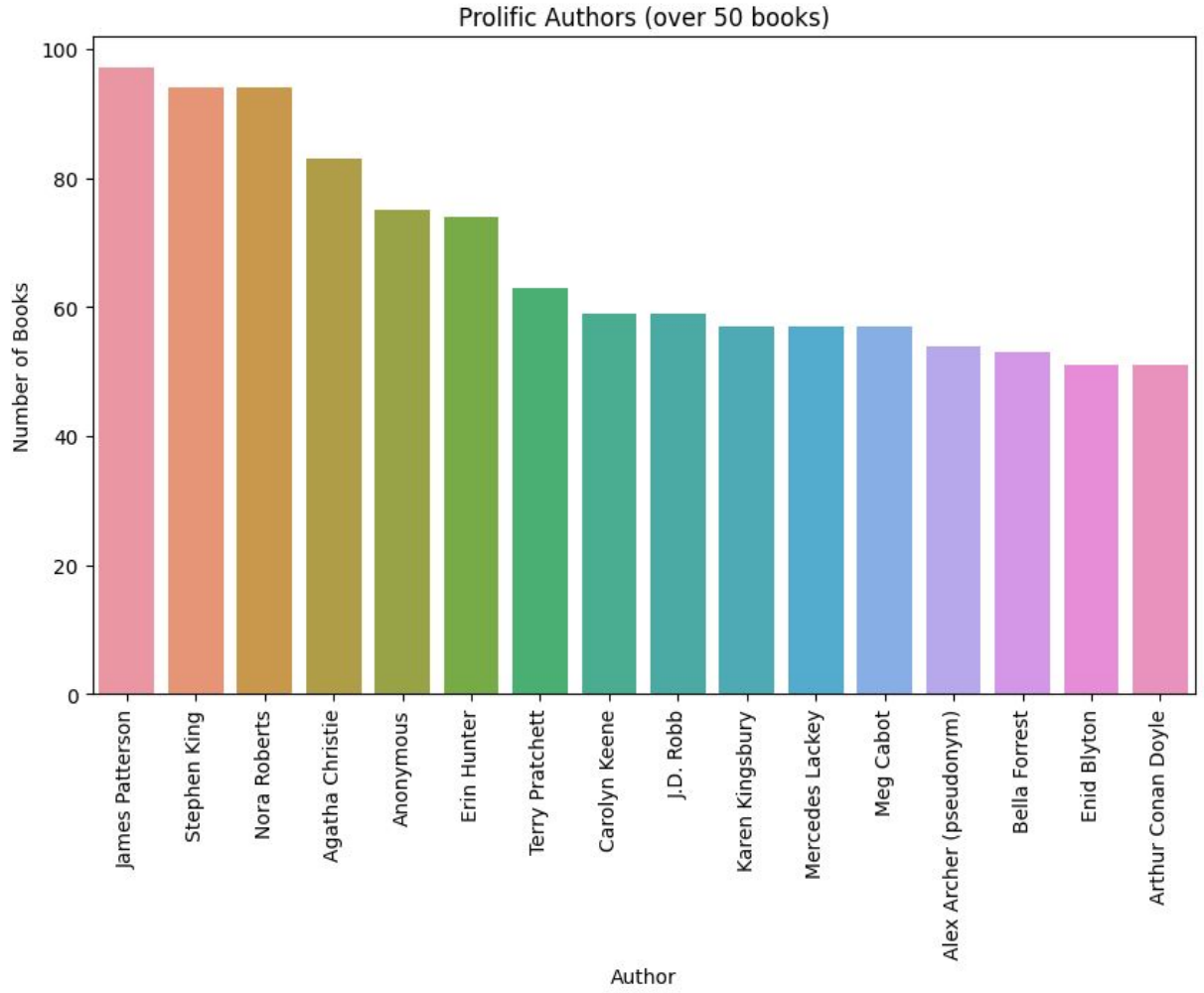
Average Rating per Decade



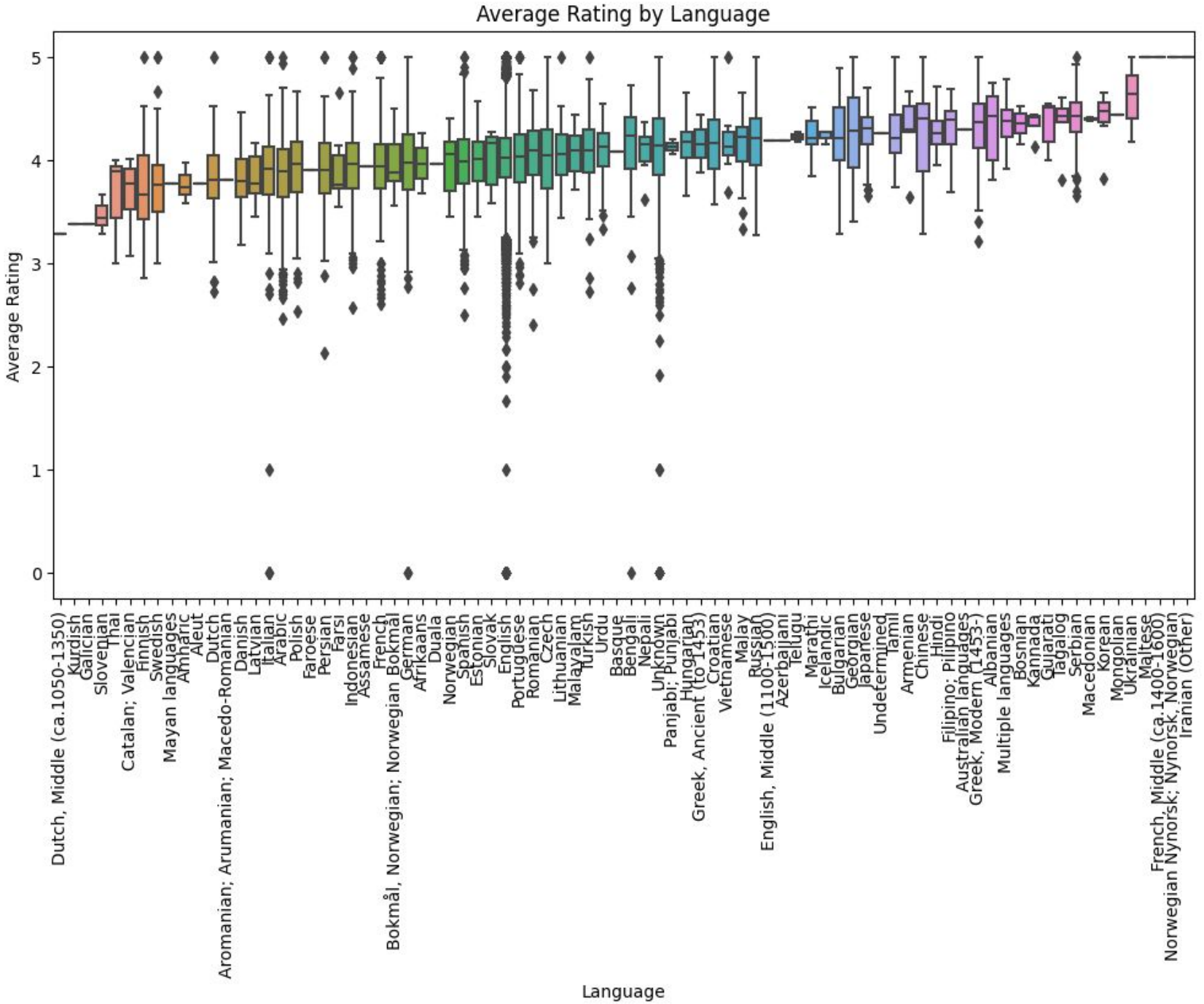
Outliers (years beyonce 2020s)
have been removed

Title World Cloud

Prolific Authors (over 50 books)



Average Rating by Language



Description Word Cloud

Content-based Filtering Model

- TF-IDF vectorization: Convert book descriptions into numerical representations
- Similarity calculation: Cosine similarity scores between books based on TF-IDF features
- Top 5 recommendations: Books with highest similarity scores

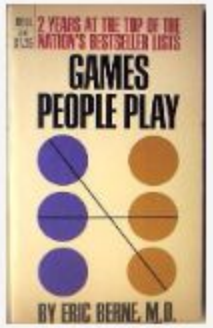
Recommended books for 'The Hunger Games':

SAMPLER ONLY: Catching Fire (The Hunger Games, #2)
Author: Suzanne Collins
Liked Percent: 97.0

Mockingjay
Author: Suzanne Collins
Liked Percent: 93.0

Catching Fire
Author: Suzanne Collins
Liked Percent: 97.0

The Hunger Games Trilogy Boxset
Author: Suzanne Collins
Liked Percent: 98.0



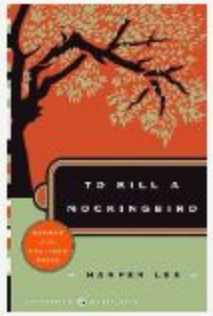
Games People Play
Author: Eric Berne
Liked Percent: 90.0

Collaborative Filtering (KNN) Model

- User-item matrix: Represents user preferences and book interactions
- Cosine similarity calculation: Identifies similar users based on preference vectors
- Recommendations: Books interacted with by similar users but not the target user

Recommended books for 'The Hunger Games':

Harry Potter and the Order of the Phoenix
Author: J.K. Rowling
Liked Percent: 98.0



To Kill a Mockingbird
Author: Harper Lee
Liked Percent: 95.0

Pride and Prejudice
Author: Jane Austen
Liked Percent: 94.0

The Book Thief
Author: Markus Zusak
Liked Percent: 96.0

Animal Farm
Author: George Orwell
Liked Percent: 91.0

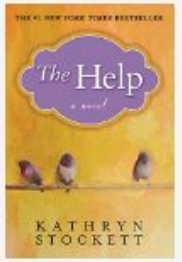
Random Forest Model

- Combination of book attributes: Titles, authors, genres, ratings, numerical features
- Random forest regressor: Trained on feature matrix and 'likedPercent' target variable
- Recommendations: Cosine similarity scores used to identify similar books

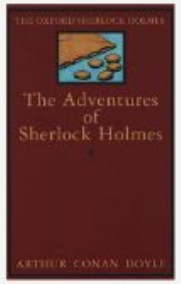
Recommended books for 'The Hunger Games':

Harry Potter and the Order of the Phoenix
Author: J.K. Rowling
Liked Percent: 98.0

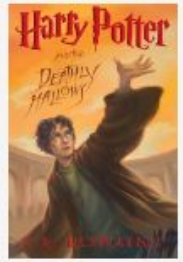
J.R.R. Tolkien 4-Book Boxed Set: The Hobbit and The Lord of the Rings
Author: J.R.R. Tolkien
Liked Percent: 98.0



The Help
Author: Kathryn Stockett
Liked Percent: 98.0



The Adventures of Sherlock Holmes
Author: Arthur Conan Doyle
Liked Percent: 98.0



Harry Potter and the Deathly Hallows
Author: J.K. Rowling
Liked Percent: 98.0

Conclusion

- Project success: Personalized book recommendations using different techniques
- Potential for improvement: Evaluation helps identify areas for enhancement
- User experience: Diverse approaches cater to different preferences and enhance book discovery
- Future possibilities: Foundation for advanced recommendation systems in various domains

Editable Icons

slidesMania

Free themes and templates for
Google Slides or **PowerPoint**

NOT to be sold as is or modified!

Read FAQ on slidesmania.com

Do not remove the slidesmania.com text on the sides.

Sharing is caring!