

BOOK RECOMMENDATION SYSTEM



GROUP MEMBERS	GITHUB
1. MAUREEN WAMBUGU	https://github.com/Mau-Wambugu
2. STEPHEN WAWERU	https://github.com/stendewa
3. LYDIA NJERI	https://github.com/lydiahsherry23
4. LILIAN MULI	https://github.com/mwikali24
5. PETER MAINA	https://github.com/Mr-PeterMaina

Business Understanding

- **Recommendation systems** are powerful tools that use machine learning algorithms to provide relevant suggestions to users based on behaviour patterns or user data.
- **A Book Recommendation System** is a recommendation system where we recommend similar books to readers based on their interests. Recommendation systems help drive engagement, increase sales, increases revenue and this in return brings in loyal clients as the customer experience is elevated promoting customer satisfaction. We have 2 main recommendation system models:
 - 1. Collaborative filtering
 - 2. Content-based Filtering

BUSINESS PROBLEM

- Over the past years, there has been rise in huge ecommerce and online services leading clients facing difficulty when searching for the right products.
- Clients looking to purchase books also face the same struggle when trying to match the right books with their taste and preferences.
- **The Business Problem** is to develop a recommendation system that recommends books that are tailored to our users preferences inorder to improve customer experience and engagement.

PROJECT OBJECTIVE

1. To build a book recommendation system that provides personalized suggestions to our users.
2. Improve sales by showcasing books a user is most likely to buy.
3. Offer relevant books to users inorder to improve customer retention.
4. Increase customer engagement

The analysis aims to achieve these objectives by considering the following analysis questions:

1. Which authors consistently receive higher ratings from users?
2. How does the year of publication affect the average ratings of books? Are newer books preferred over older ones?
3. How accurate are the recommendations generated by the collaborative filtering model especially in terms of precision?
4. How does class imbalance in ratings affect the performance of the recommendation model?

DATA SOURCE

- We used data obtained from [Kaggle](#) mined by Cai-Nicolas Ziegler
- It contains 3 CSV Files:
 1. Books.csv - contains information about books{ISBN;Title;Author;Year;Publisher}
 2. Ratings.csv - contains book ratings provided by users that range from 0 to 10. {{User-ID;ISBN;Rating}}
 3. Users.csv - contains information about the users {User-ID;Age}

STAKEHOLDERS

1. Customers
 - As the end user, they expect accurate book suggestions based on personal interests.
2. Marketing team
 - They would want to do targeted advertising on specific books and also promote personalized offers.
3. Data scientist
 - Interested in ensuring the recommendation system models are accurate and scalable.
4. Book Authors
 - They would be interested in knowing how their books are recommended inorder to learn and understand their readers taste and preferences.
5. Executive {CEO}
 - They would want to understand how recommendation systems impact revenue and customer retention comparing it to the budget allocated to the project.

METHODOLOGY

- Our project will focus on the CRISP-DM:
 1. Business Understanding
 2. Data Understanding
 3. Data Preparation
 4. Modeling
 5. Evaluation
 6. Deployment

DATA SOURCE

- We used data obtained from [Kaggle](#) mined by Cai-Nicolas Ziegler
- It contains 3 CSV Files:
 1. Books.csv - contains information about books{ISBN;Title;Author;Year;Publisher}
 2. Ratings.csv - contains book ratings provided by users that range from 0 to 10. {{User-ID;ISBN;Rating}}
 3. Users.csv - contains information about the users {User-ID;Age}

STAKEHOLDERS

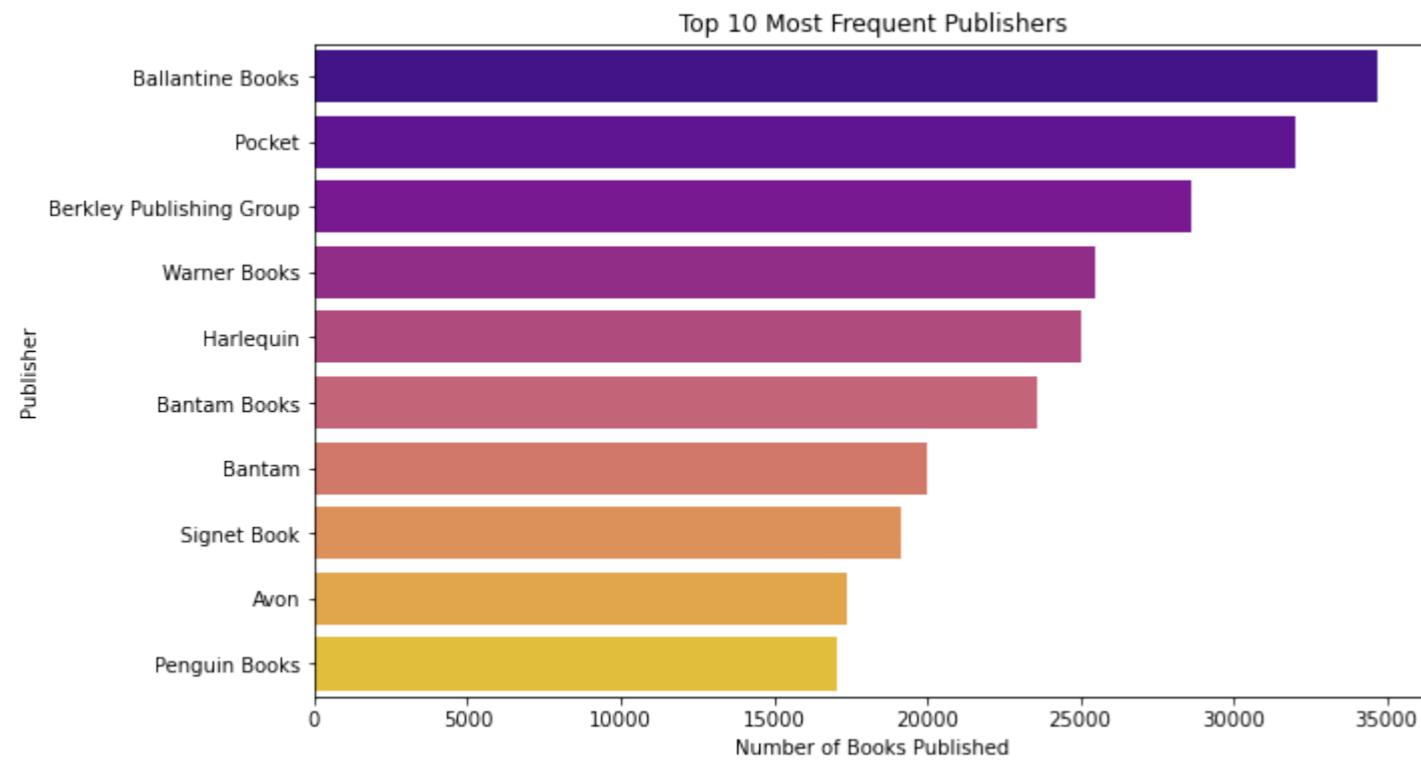
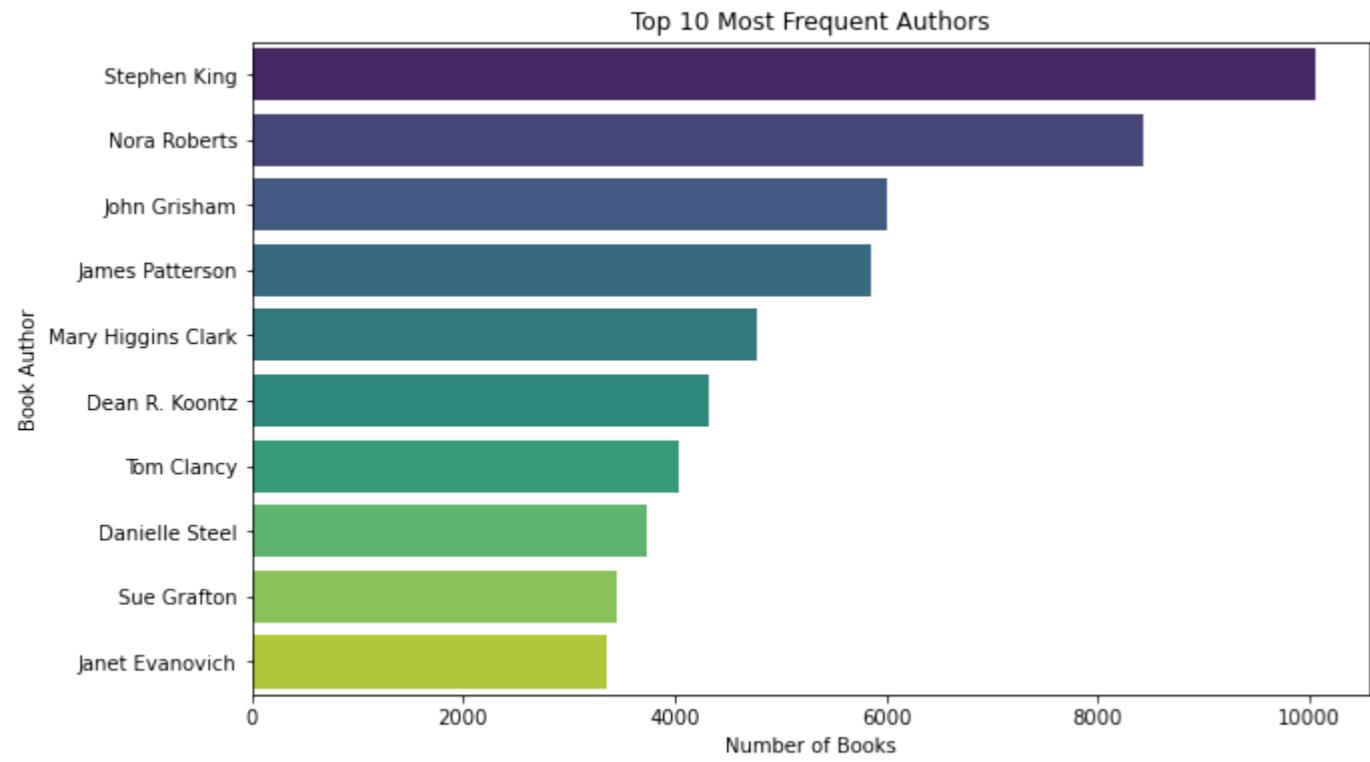
1. Customers
 - As the end user, they expect accurate book suggestions based on personal interests.
2. Marketing team
 - They would want to do targeted advertising on specific books and also promote personalized offers.
3. Data scientist
 - Interested in ensuring the recommendation system models are accurate and scalable.
4. Book Authors
 - They would be interested in knowing how their books are recommended inorder to learn and understand their readers taste and preferences.
5. Executive {CEO}
 - They would want to understand how recommendation systems impact revenue and customer retention comparing it to the budget allocated to the project.

METHODOLOGY

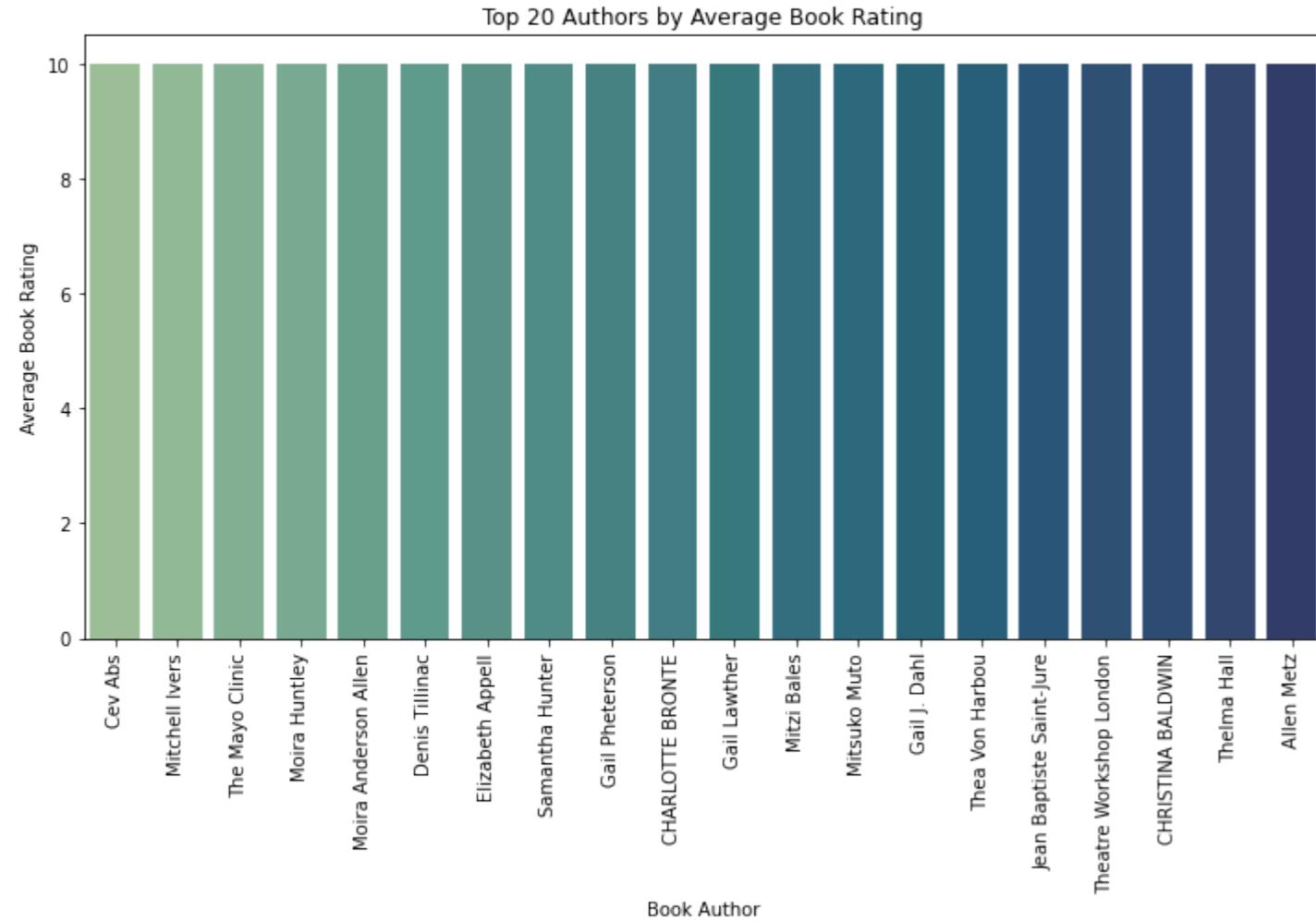
- Our project will focus on the CRISP-DM:
 1. Business Understanding
 2. Data Understanding
 3. Data Preparation
 4. Modeling
 5. Evaluation
 6. Deployment

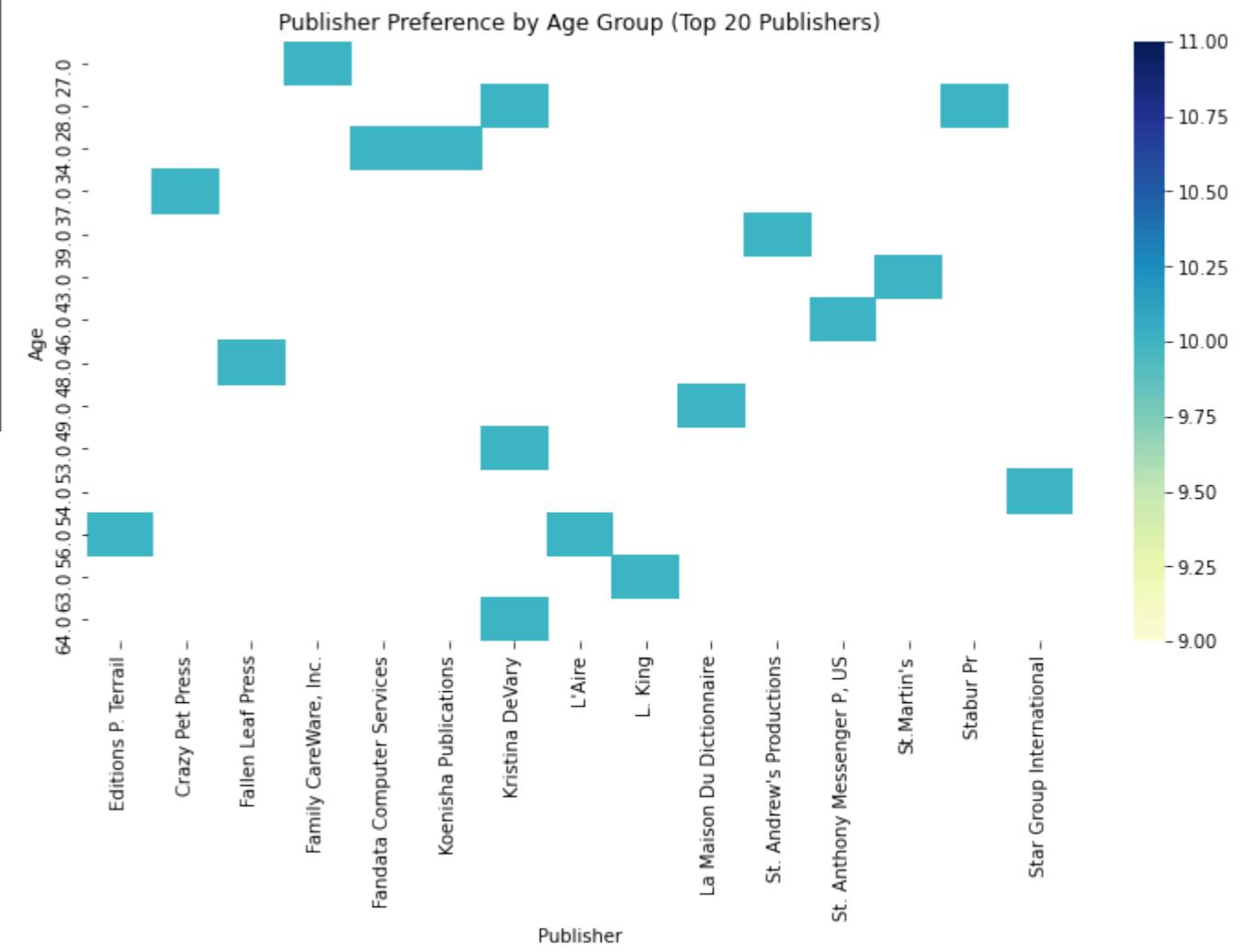
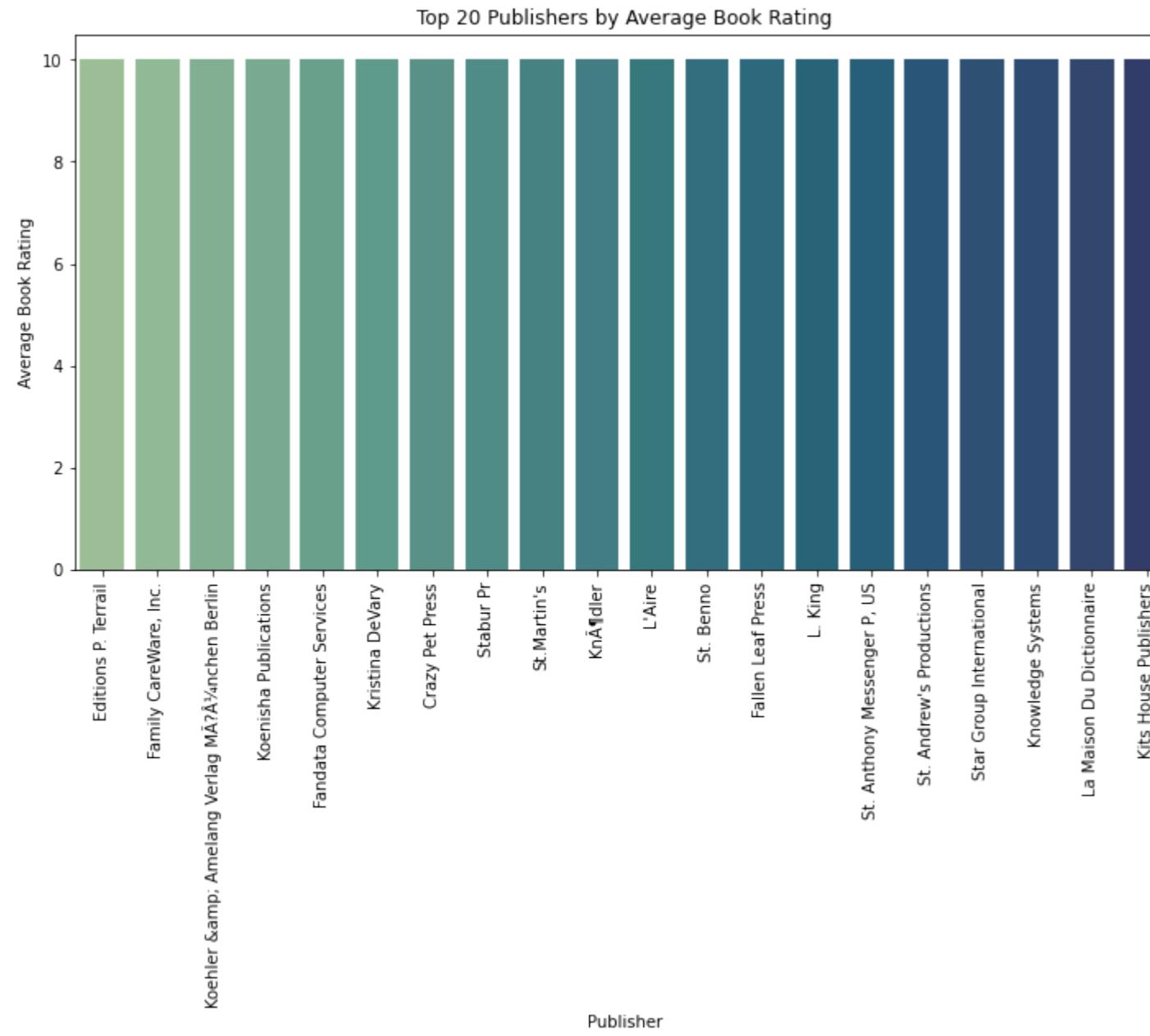
Explanatory Data Analysis

1. Univariate data analysis



2. Bivariate analysis





Modelling

In our analysis we used **collaborative filtering** to build a recommendation system.

In our evaluation we focused majorly on **precision score, recall score, f1 score, accuracy score**

Conclusion

1. Effectiveness of collaborative filtering
 - Collaborative filtering using **Singular Value Decomposition** {SVD} can effectively predict book ratings for users based on their historical preferences and those of similar users.
2. Model performance evaluation
 - The model has a reasonable predictive **accuracy score** since **RMSE= 3.4219** and **MAE = 2.7470**
 - However the model has a low **recall score** of **0.0817**, this indicates most relevant books are not captured by the model. Thus it requires further attention to enhance the model quality.
3. User - Item based collaborative filtering
 - **User-User method** effectively identifies similar users and their preferences.

- **Item based approach** using SVD provided personalized recommendations based on predicted rating for unseen books.
 - Our analysis shows that the same user has different recommendations based on the two different approaches.
4. Item - based approach
- Based on this approach it is evident that highly rated books are preferred by most users.

Recommendations

1. Enhanced personalization and feedback mechanism.
2. Utilize data for targeted marketing.
3. Business should feature highly rated books.
4. Use top rated labels to highlight highly rated books in marketing campaigns.