

The background features a dark blue field with white and light blue geometric elements. These include thin white lines forming various angles and paths, some ending in small white circles. A series of vertical white lines are positioned near the top center. In the upper right, there are concentric light blue circles and a horizontal white bar. Other light blue lines and shapes are scattered in the lower right and bottom right areas.

Book recommendation system

Diego Fuentes

► Overview of the subject

Over 1m Books ratings containing general Book Data, and Users Demographics.

Problem statement

With countless books available, finding one that matches individual tastes can be tedious. Existing systems often miss the mark on capturing a reader's interests.

Opportunity

A Book recommendation system that have the potential to deliver personalized suggestions that will resonate deeply with readers' preferences. By using factors such as author, publication year, genre, and other user ratings.

► Model Result

Readers

Better matches between reader preferences and recommendations, increasing user engagement and retention.

Business

Enhanced recommendations could boost sales by matching users with books they're more likely to purchase.

► Content Based Recommendations & Collaborative filtering



Book Title



Book Author



Book Category



Publisher



Language



Year of Publication



User Id



Age

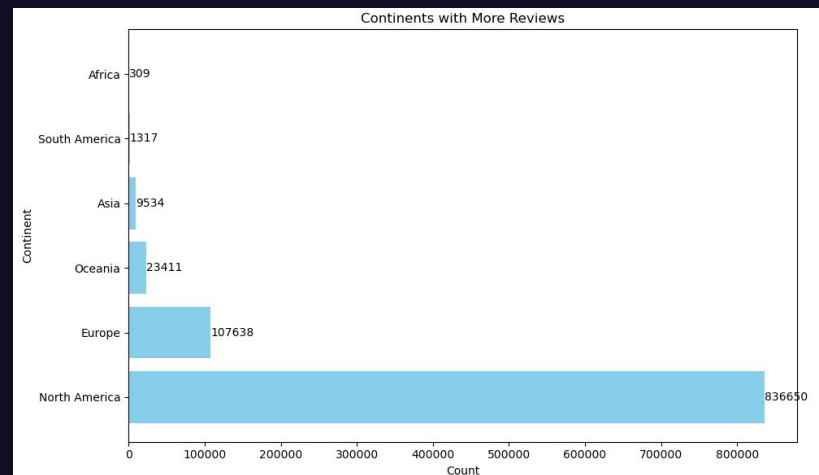
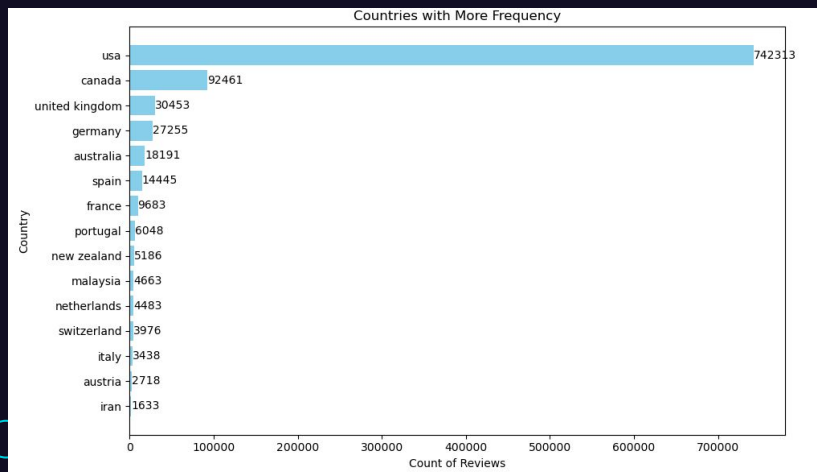


Country - Continent

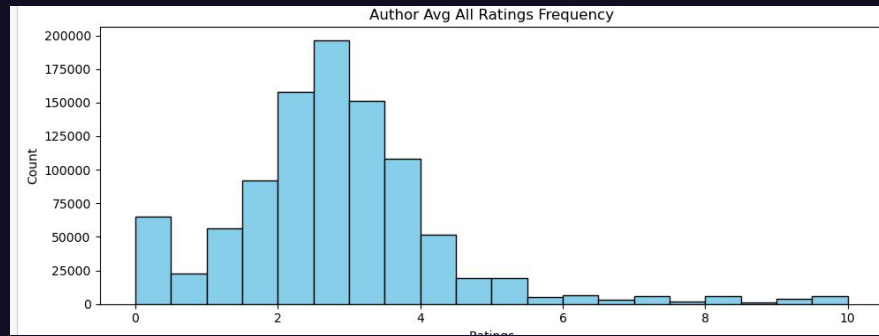
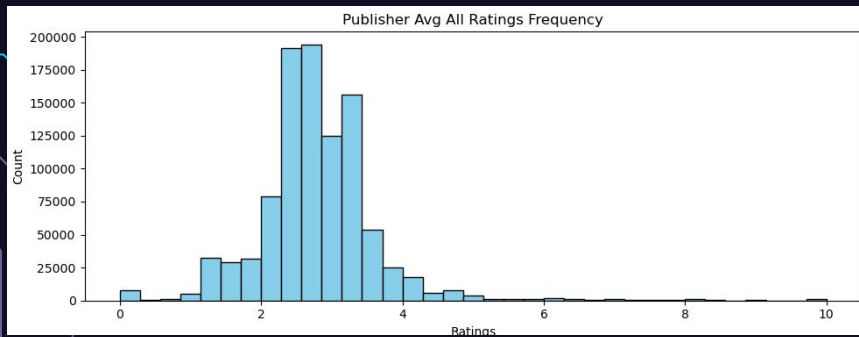
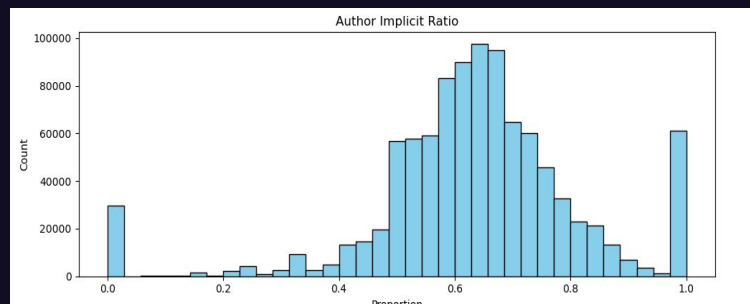
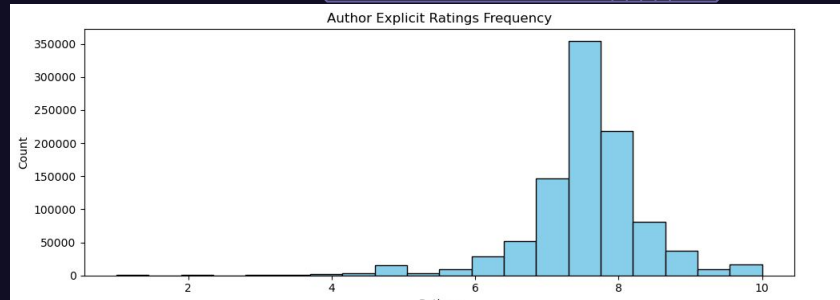
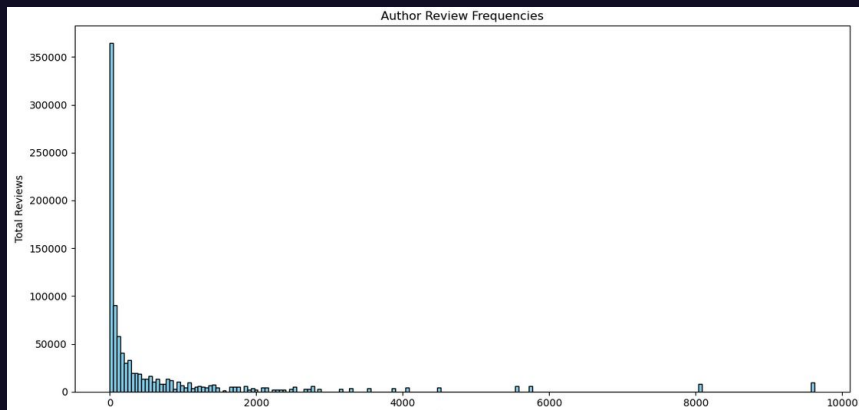


Ratings

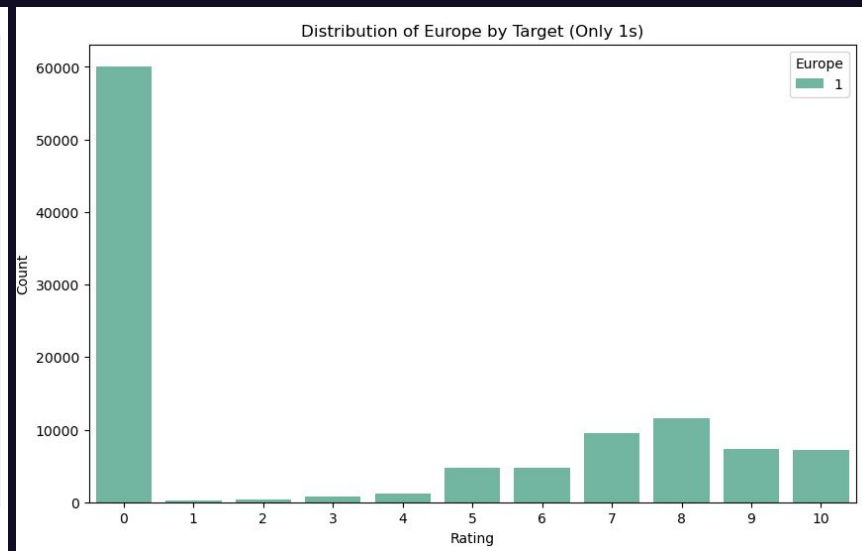
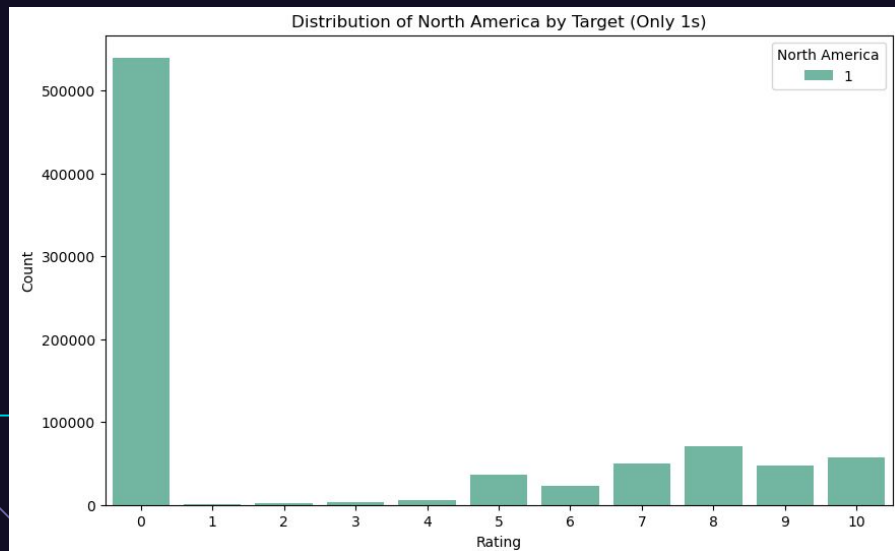
► Preprocessing - Countries Encoding



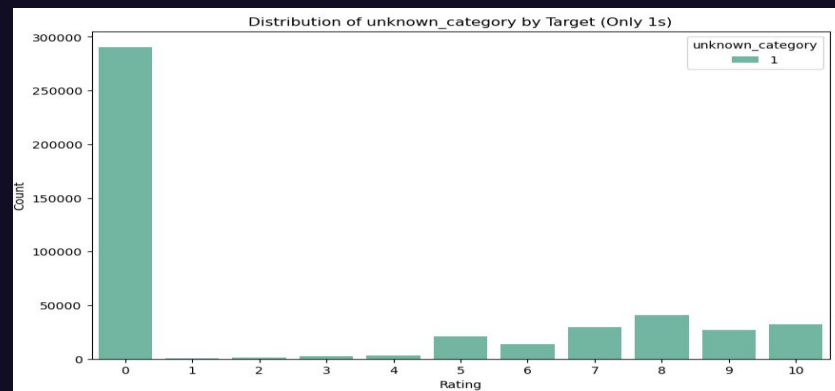
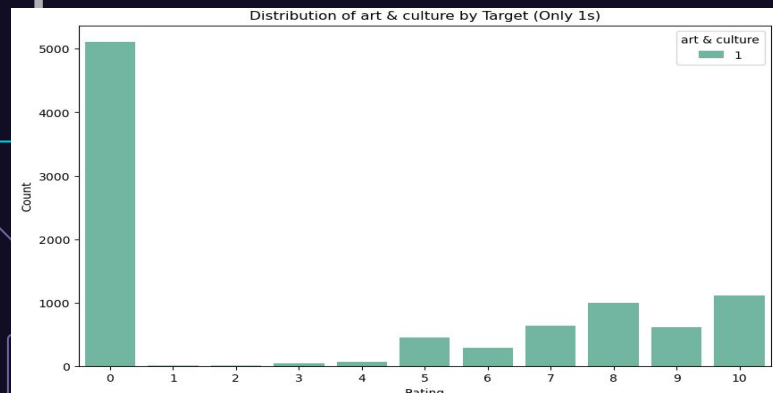
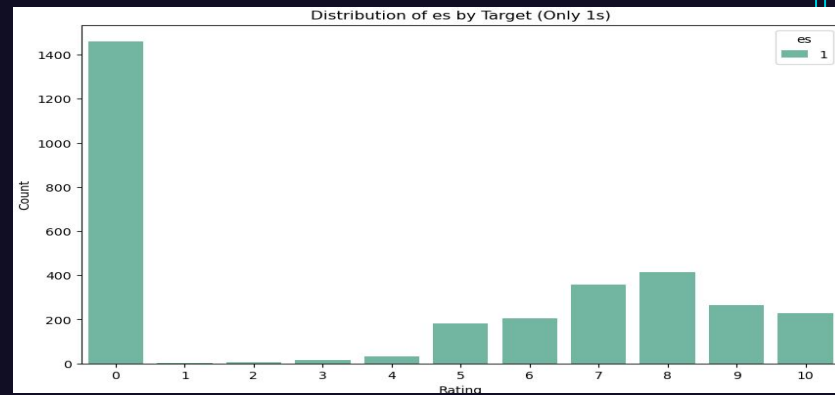
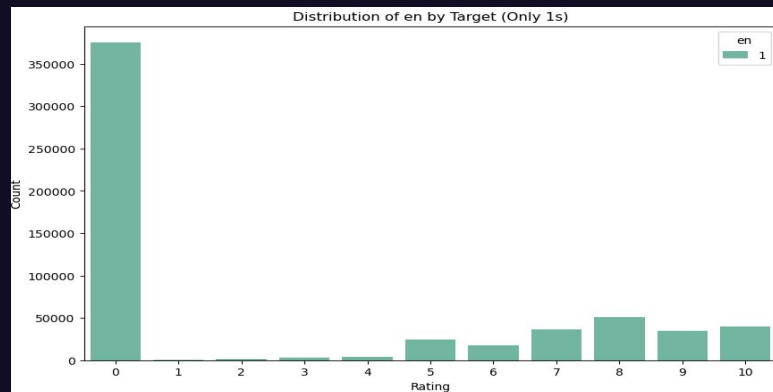
► Authors and Publishers



► Post preprocessing eda



► Categoricals



► Models

The performance of the models (Logistic Regression, KNN, and Decision Tree) is relatively similar in terms of results, with all achieving moderate accuracy (~68%) and comparable AUC values (0.70-0.71).

Logistic

Best performance predicting implicit or explicit ratings with a 68.43% accuracy on the test data.

KNeighbours

65% predicting ratings 0-10
68.20% predicting implicit or explicit on test data.

Decision Tree

66% predicting ratings 0-10
68% predicting implicit or explicit

► Content Based Recommender System

```
# Test the recommender
similar_movies = content_recommender("Harry Potter and the Chamber of Secrets (Book 2)", unique_titles, similarities, vote_threshold=10)
similar_movies.head(10)
```

	book	similarity	Number of reviews	Avg Rating
19691	Harry Potter and the Chamber of Secrets (Book 2)	1.000000	98	8.724490
43852	The Chamber	0.584344	71	7.323944
19708	Harry Potter and the Sorcerer's Stone (Harry P...	0.570851	85	8.917647
19696	Harry Potter and the Goblet of Fire (Book 4)	0.524118	73	9.287671
19706	Harry Potter and the Sorcerer's Stone (Book 1)	0.497518	60	8.916667
19699	Harry Potter and the Order of the Phoenix (Boo...	0.496432	64	9.062500
19703	Harry Potter and the Prisoner of Azkaban (Book 3)	0.466559	99	8.898990
17542	From Potter's Field	0.340878	30	7.933333
48839	The Murder Book	0.136553	32	7.906250
43337	The Book of Ruth (Oprah's Book Club (Paperback))	0.116924	39	7.153846