Get Your Book

A recommender system for you to find your preferred book

Rafael Cabral

Overview

Get Your Book has over 3000 books for you.

By choosing a book, you will receive recommendations for other books according to its title, genre, author and description.
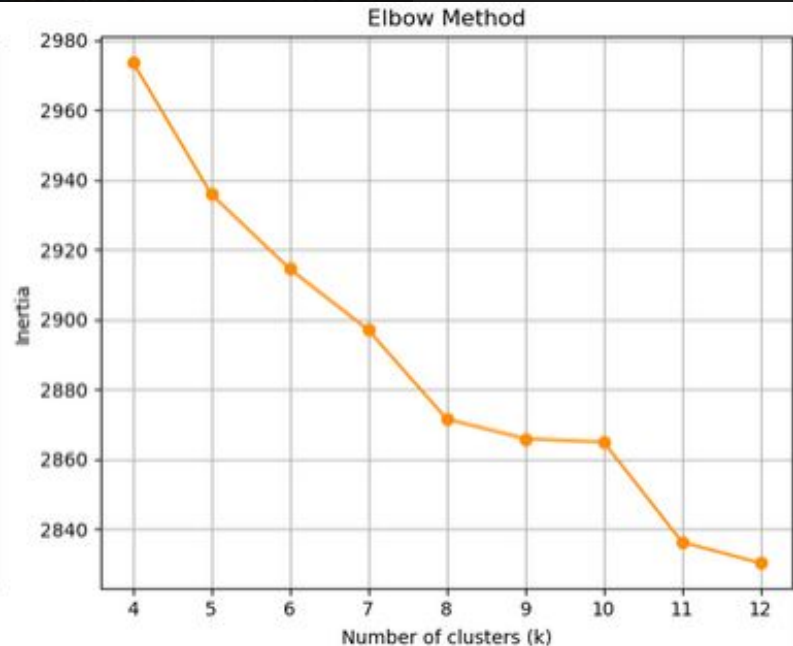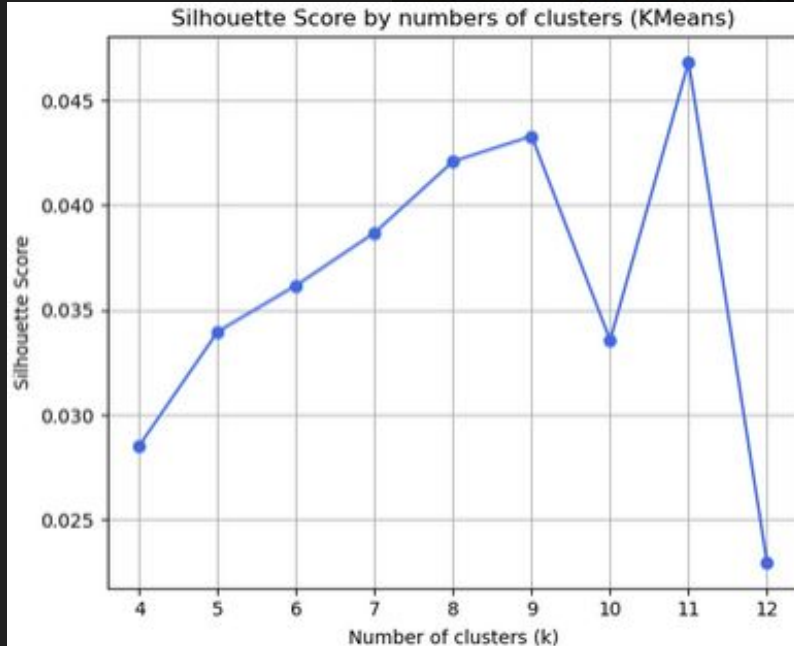
Data overview and cleaning

- Gutenberg.org: 2394 books by scraping;
- Google API: 1010 books;
- Books.csv, from kaggle: to get genre(Gutenberg has none);
- Open Library API: to get the covers by isbn13;
- TF-IDF: to filter out common words like "the", "is", "a" to make the clustering using 'description' in a better way;

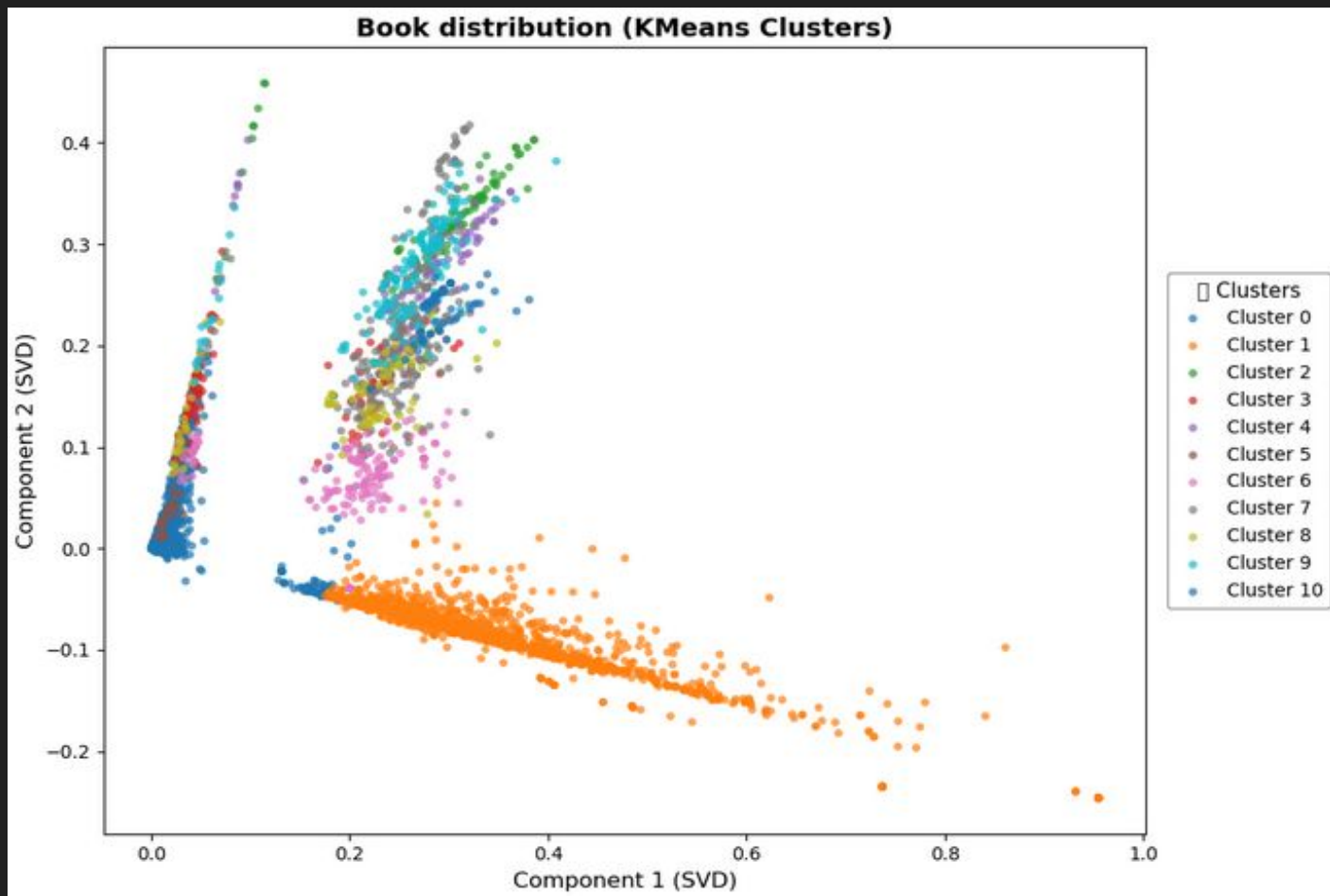After remove duplicates I got 3347 books.

Silhouette Score

Clusters

For greater accuracy and smaller groups, I used k = 11(as Silhouette suggested).

Book distribution (KMeans Clusters)

Clusters:
- Cluster 0
- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4
- Cluster 5
- Cluster 6
- Cluster 7
- Cluster 8
- Cluster 9
- Cluster 10

Thank you

http://localhost:8501/

Rafael Cabral