

A woman with dark hair and glasses is looking over the top of an open book. The background is a blurred library with bookshelves. The image has a blue tint and a semi-transparent orange rectangle is overlaid on the left side.

# **Book Recommendation System**

**DSI14 Capstone**  
Jocelyn Pok

01

## **Use Cases**

Pervasiveness of  
recommendation systems

02

## **Problem Statement**

Independent retailers:  
adapting to changing retail  
landscape and external  
events

03

## **Exploratory Data Analysis**

Insights from the datasets

04

## **Models & Evaluation**

Content-based, Collaborative  
filtering and Hybrid  
techniques

05

## **Conclusion**

Limitations and future  
considerations

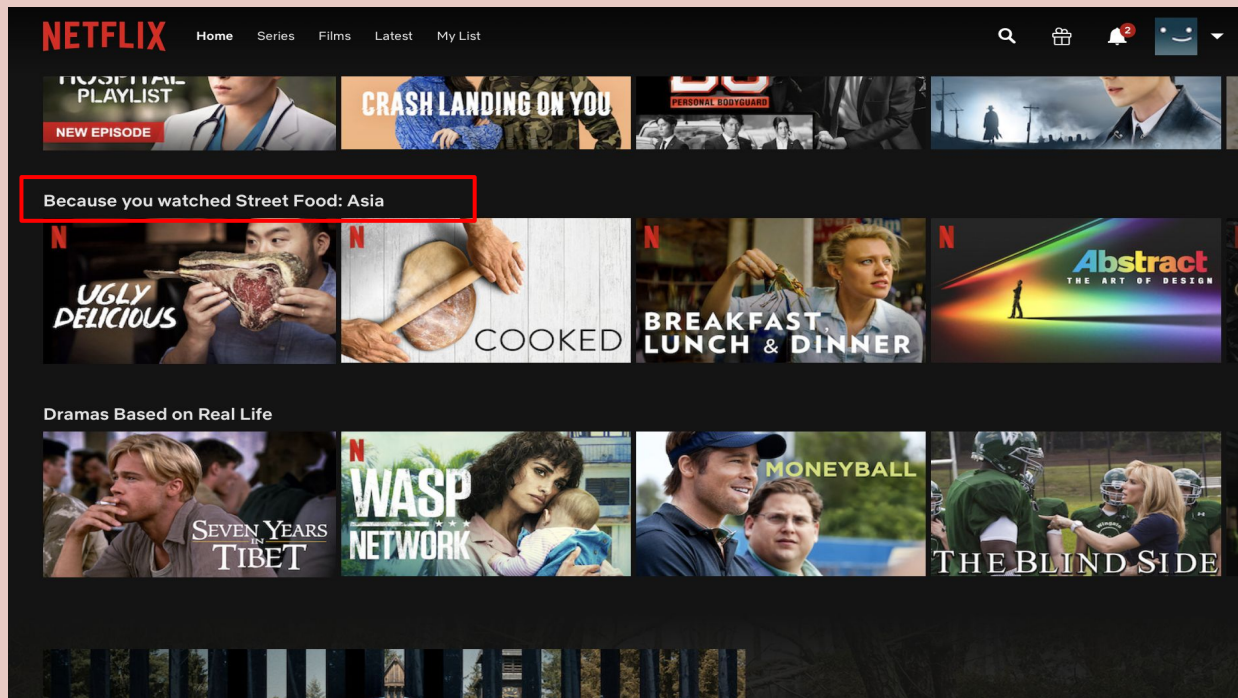
06

## **Demo**

# **Table of Contents**

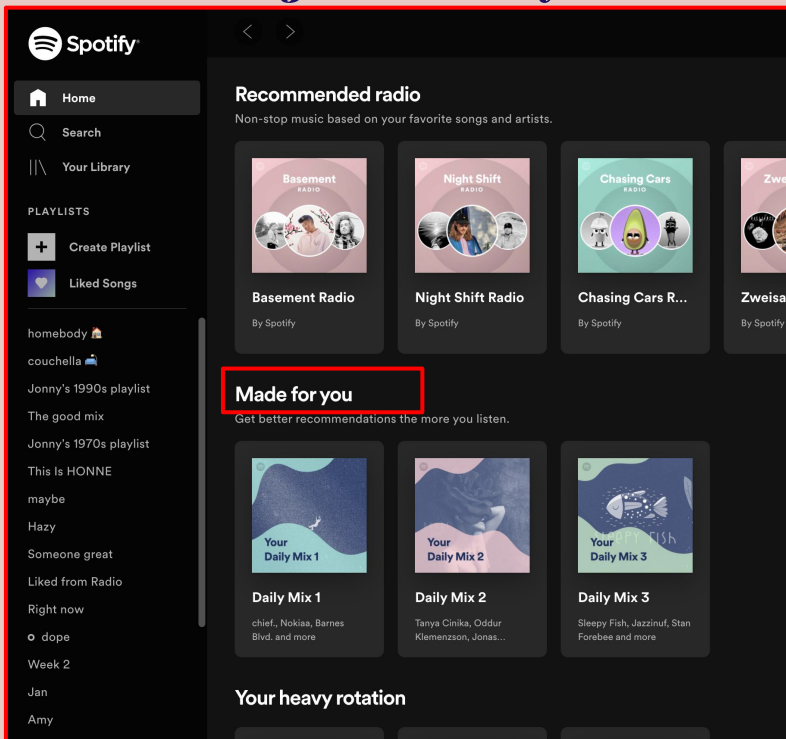
# Use case 1 : User-item

## “ Recommended just for you”



# Use case 1 : User-item

## “ Recommended just for you”





## Use case 2 : Item-item


### “ Similar products...”

**Customers who viewed this item also viewed**


Page 1 of 1




Nintendo Switch Lite - Gray  
★★★★★ 10,025  
124 offers from \$229.99




Nintendo Switch – Neon Red and Neon Blue Joy-Con - HAC 001...  
Nintendo  
★★★★★ 14,948  
Nintendo Switch  
32 offers from \$409.99




Nintendo Switch with Neon Blue and Neon Red Joy-Con HAC-001 w...  
Nintendo  
★★★★★ 1,322  
Nintendo Switch  
26 offers from \$549.96



Nintendo Switch Console Bundle- Pikachu & Eevee Edition with Pokemon...  
Nintendo  
★★★★★ 661  
Nintendo Switch  
33 offers from \$603.98



Nintendo Switch - Animal Crossing: New Horizons Edition - Switch  
Nintendo  
★★★★★ 643  
Nintendo Switch  
23 offers from \$696.96



## Amazon

According to Mckinsey & Company, 35% of Amazon's revenue is generated by its recommendation engine

## Netflix

According to Mckinsey & Company, 75% of what users watch on Netflix come from product recommendations

## Spotify

Possibly one of Spotify's most innovative uses of AI and recommendation systems is their popular Discover Weekly playlist. Known as Release Radar, this algorithmically powered tool updates personal playlists on a weekly basis so that users won't miss newly released music by artists they like.

The new recommendation system has helped Spotify increase its number of monthly users from 75 million to 100 million at a time, in spite of competition from rival streaming service Apple Music.

# Situation

The New York Times

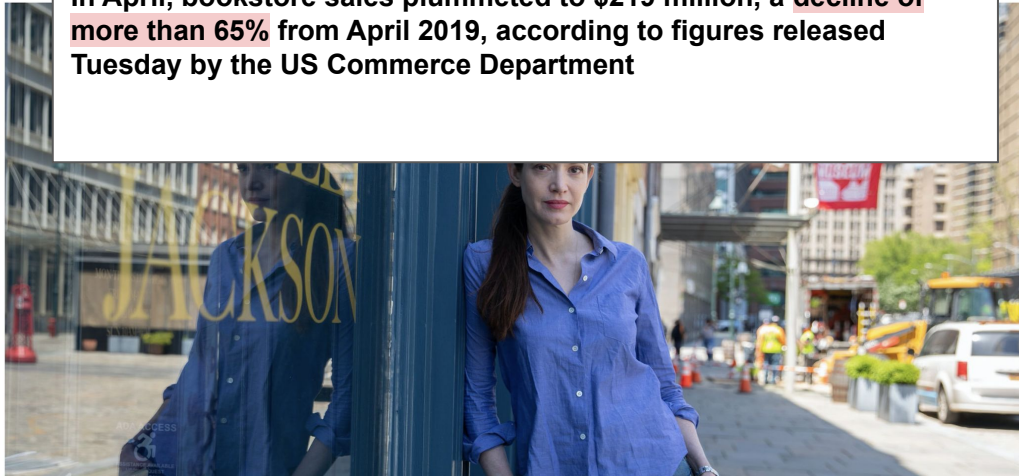
SUBSCRIBE NOW

Opinion

## What Could Kill My New York Bookstores?

It won't be Amazon or the coronavirus. It will be artificially high rents.

In April, bookstore sales plummeted to \$219 million, a decline of more than 65% from April 2019, according to figures released Tuesday by the US Commerce Department





# Problem Statement

---

“COVID-19 is not only a public health issue. It is also a serious economic, social and political problem” -PM Lee

Beyond the pandemic, independent bookstores that have joined the online retail ecosystem can explore implementing recommendation systems that can significantly boost sales and positive effects on the user experience as well, which translates into metrics that are of much importance to online businesses such as customer satisfaction and retention.

This project seeks to identify the best recommendation model that obtains the best top-N accuracy metrics for the dataset. The methods that will be explored in this project include : content-based, collaborative filtering and hybrid





# Exploratory Data Analysis

---

Lorem ipsum dolor sit amet consectetur  
adipiscing elit massa sociis mollis iaculis,  
arcu fermentum per ullamcorper id at tempor  
turpis volutpat curae, quisque vitae montes  
magna sodales purus fringilla mattis sem nunc.  
Elementum tempor scelerisque metus hac lacus,  
sem rhoncus consequat mollis turpis porta,  
vivamus bibendum porttitor suscipit.

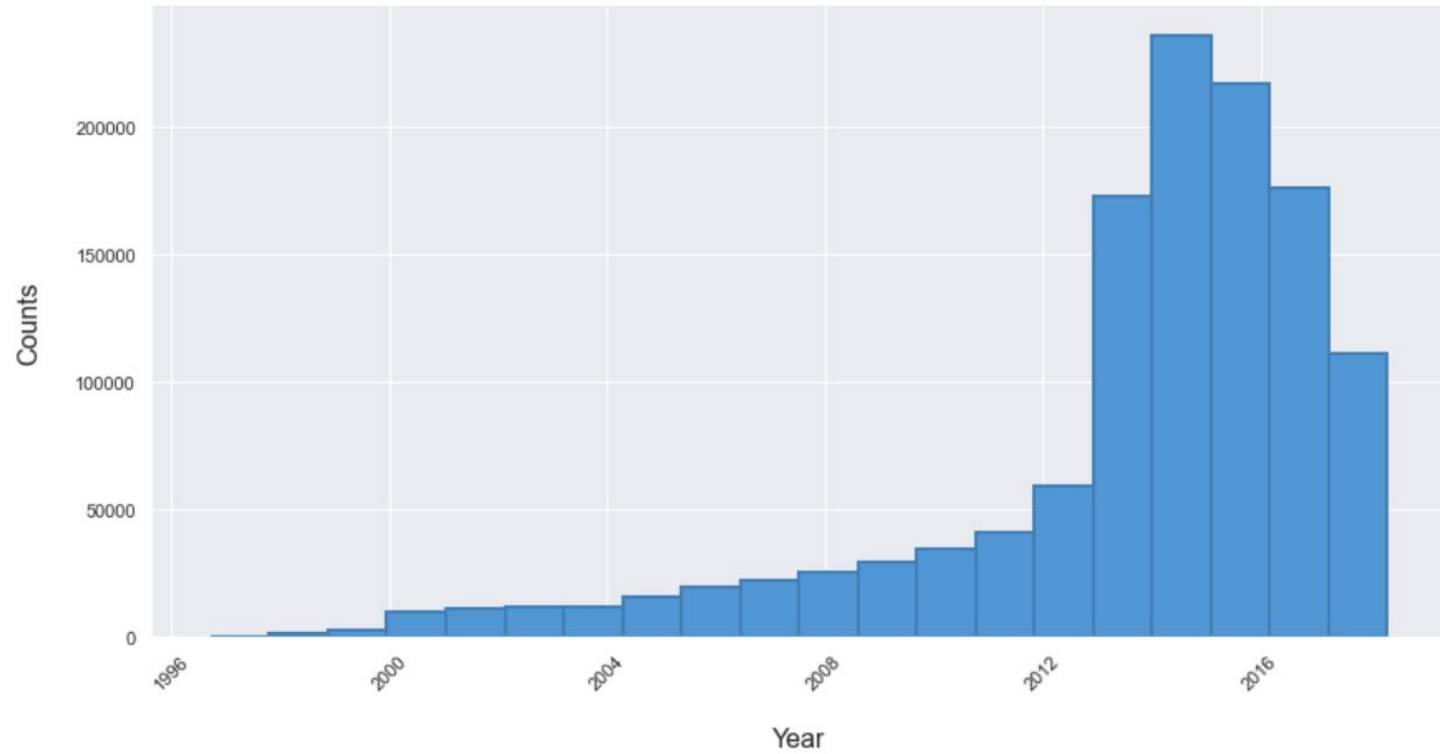
Tristique curae mollis porta sed fusce tempor  
lacus est mus magna iaculis, eu ultricies  
aliquet aptent diam accumsan varius auctor non.  
Dapibus purus varius potenti sociosqu pulvinar  
quis ut sagittis, per turpis fermentum litora  
magnis class tincidunt, faucibus auctor  
tristique inceptos erat vehicula eu. Platea  
metus auctor blandit id cubilia dapibus  
parturient dictum risus volutpat, rhoncus  
tristique bibendum praesent laoreet nulla  
ornare pharetra sociosqu penatibus, gravida mi  
aliquam malesuada condimentum eu imperdiet at  
nam.

Tristique curae mollis porta sed fusce tempor  
lacus est mus magna iaculis, eu ultricies  
aliquet aptent diam accumsan varius auctor non.  
Dapibus purus varius potenti sociosqu pulvinar  
quis ut sagittis, per turpis fermentum litora  
magnis class tincidunt, faucibus auctor  
tristique inceptos erat vehicula eu. Platea  
tristique inceptos erat vehicula eu. Platea  
metus auctor blandit id cubilia dapibus  
parturient dictum risus volutpat, rhoncus  
accumsan varius auctor non. Dapibus purus  
accumsan varius auctor quis ut  
varius potenti sociosqu pulvinar litora magnis  
sagittis, per turpis fermentum litora magnis  
class tincidunt, faucibus auctor tristique  
inceptos erat vehicula eu. Platea metus auctor  
blandit id cubilia dapibus parturient dictum  
inceptos erat vehicula eu. Platea metus auctor  
risus volutpat, rhoncus tristique  
praesent laoreetvolutpat, rhoncus tristique  
bibendum praesent laoreet nulla ornare pharetra  
sociosqu penatibus, gravida mi aliquam  
malesuada condimentum eu imperdiet at nam.

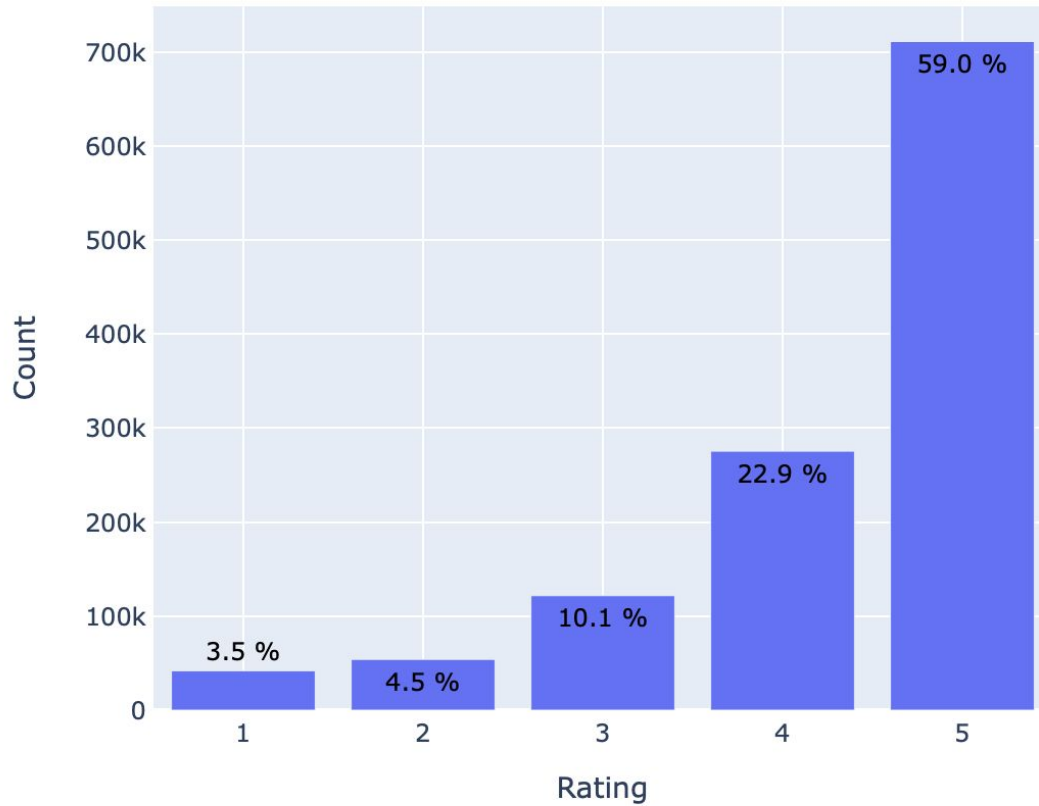
Lorem ipsum dolor sit amet consectetur  
adipiscing elit massa sociis mollis iaculis,  
arcu fermentum per ullamcorper id at tempor  
turpis volutpat curae, quisque vitae montes  
magna sodales purus fringilla mattis sem nunc.  
Elementum tempor scelerisque metus hac lacus,  
sem rhoncus consequat mollis turpis porta,  
vivamus bibendum porttitor suscipit.

Tristique curae mollis porta sed fusce tempor  
lacus est mus magna iaculis, eu ultricies  
aliquet aptent diam accumsan varius auctor non.  
Dapibus purus varius potenti sociosqu pulvinar  
quis ut sagittis, per turpis fermentum litora  
magnis class tincidunt, faucibus auctor  
tristique inceptos erat vehicula eu. Platea  
metus auctor blandit id cubilia dapibus  
parturient dictum risus volutpat, rhoncus  
tristique bibendum praesent laoreet nulla  
ornare pharetra sociosqu penatibus, gravida mi  
aliquam malesuada condimentum eu imperdiet at  
nam.

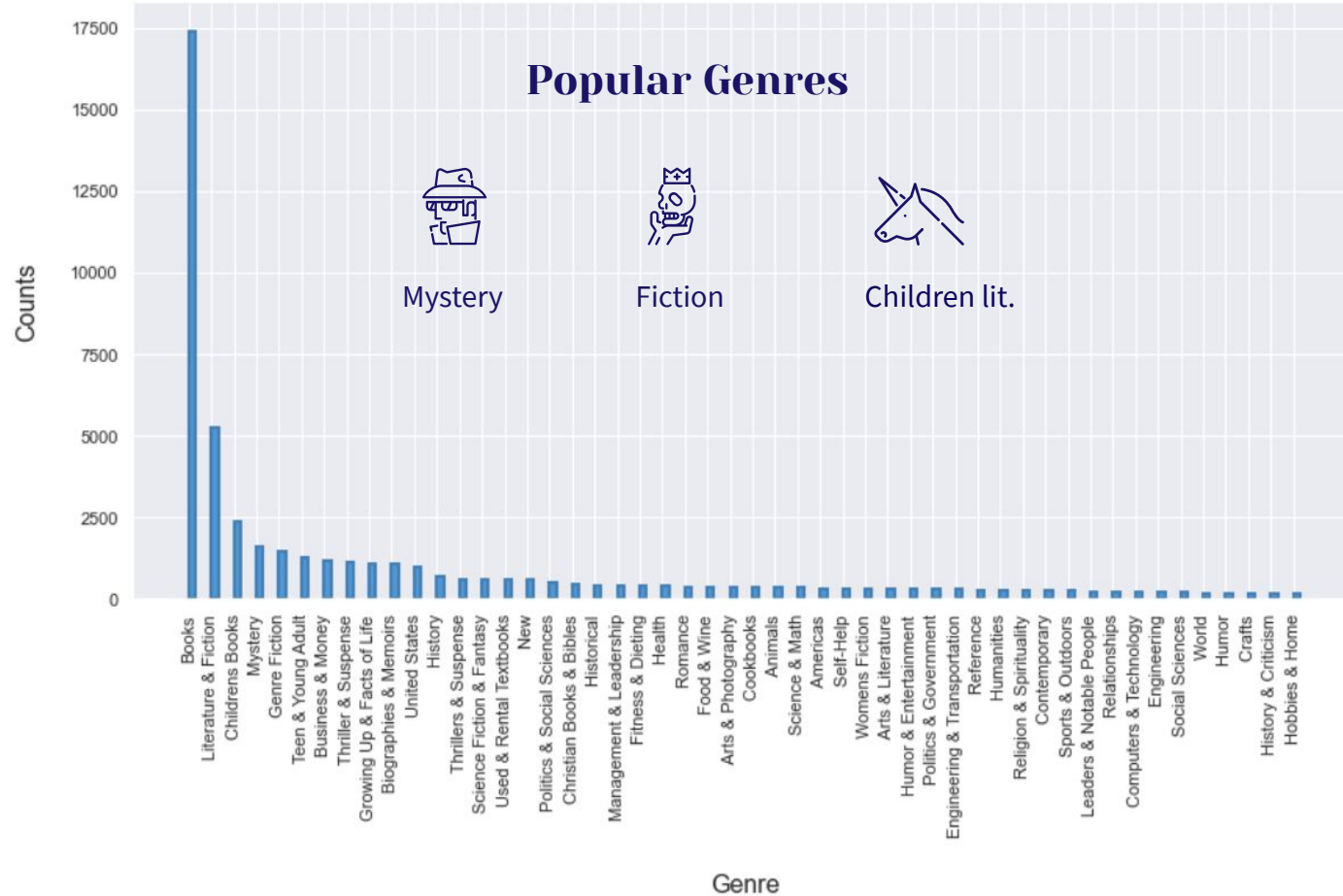
Reviews per year



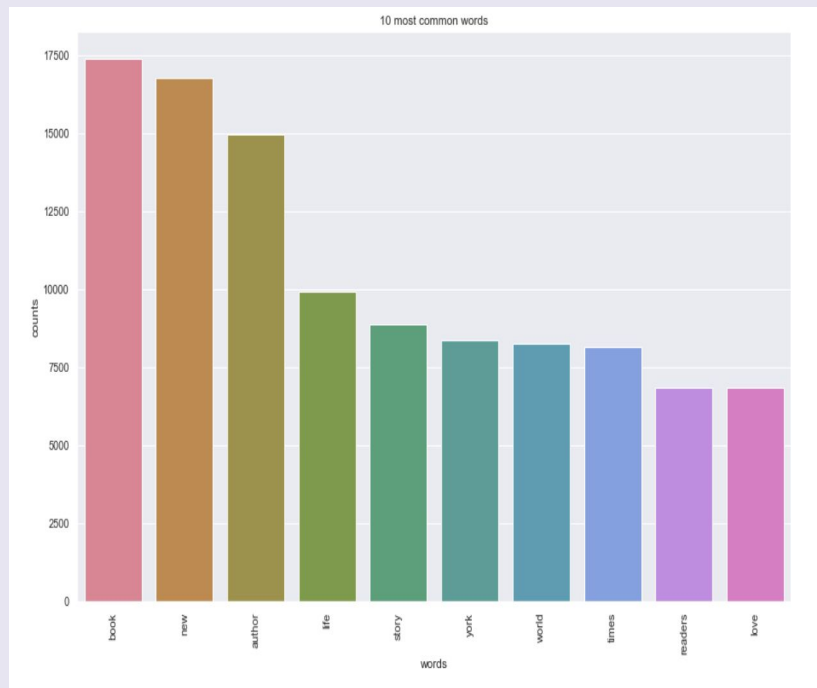
## Distribution Of 1204430 book-ratings



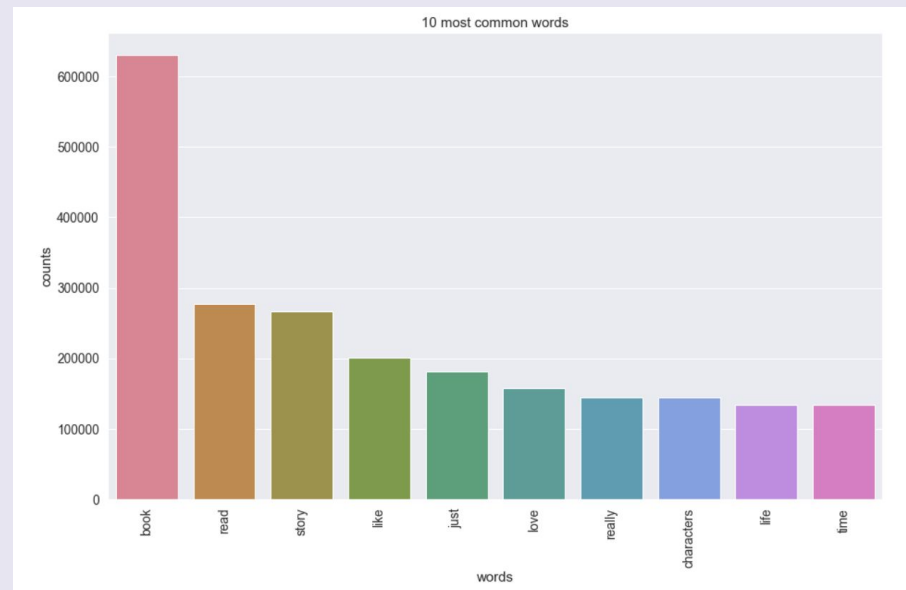
# Most Popular Book Genres



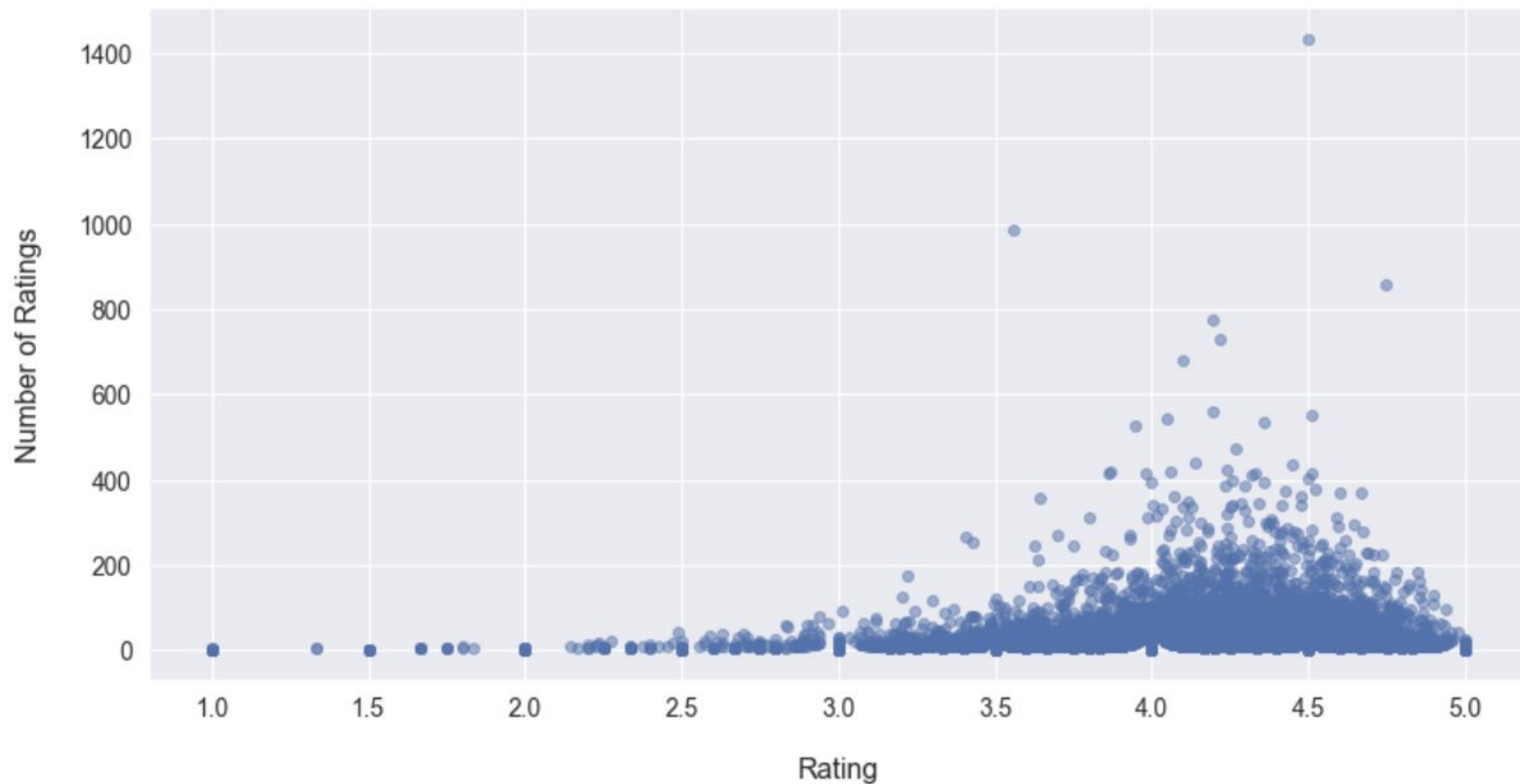
## Description



## Reviews



# Rating vs. Number of Ratings



# Model fit and Evaluation

---





# Measure model performance

## Top N-accuracy

- Provide users with a list of ranked list of N items they will be interested in, in order to encourage views and purchase

## Recall@N

- $\text{Hit@N count} / \text{total interacted items in test set}$

## Hit Rate

- To evaluate the top-N recommendations, if the user has rated one of the top-N recommended items, it is considered a “hit”

# Which model to use?

## Recommendation algorithms

### Collaborative filtering

- Matrix completion
- User/Item similarity

### Content-based Models

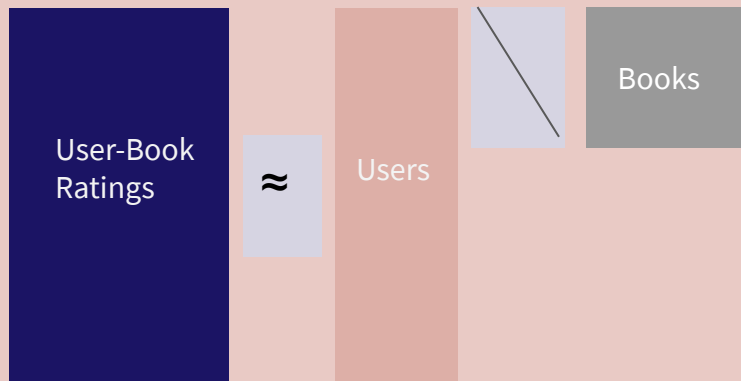
- Match based on user and book features

### Hybrid Models

# Modelling

## Recommendation Algorithms

### Collaborative Filtering



- Singular Value Decomposition(SVD)
- Decomposes user-book interactions matrix into a set of user and book vectors
- Use vectors to estimate user responses to unseen items

### Content-based Models



Extract features out of the book description to create useful representations

$$\begin{array}{|c|} \hline \text{Value of} \\ \text{"Crime"} \\ \text{to description for} \\ \text{"Mystery"} \\ \hline \end{array} = \begin{array}{|c|} \hline \text{Frequency} \\ \text{of "Crime" in} \\ \text{description} \\ \hline \end{array} \times \begin{array}{|c|} \hline \text{Inverse Freq} \\ \text{of "Crime" in} \\ \text{all} \\ \text{documents} \\ \hline \end{array}$$

- Vector for each document is represented as represented as the computed TF-IDF weights
- Find similar books by comparing their word vector

# Comparison of methods

	recall@5	recall@10
modelName		
Popularity	0.413076	0.542086
Collaborative Filtering	0.576140	0.710156
Content-Based	0.898594	0.911581
Hybrid	0.542114	0.542443



06

# Business Recommendations

Your subtitle goes here

The content-based method is well-suited for businesses new to the realm of recommendation systems but seek to incorporate an element of personalization. By identifying user's specific tastes and then provide them with relevant products will enhance the user's experience.

## **Limitations**

A content-based system is limited by the novelty in recommended results. Because it only looks at user's past history, it never jumps to other areas that users might like but never interacted before. Additionally, if the quality of the content does not contain enough information to discriminate the items precisely, this method will perform poorly



# Looking forward

## Explore

- More complex representations of data( eg word embeddings)
- Deep learning: Neural network architectures for collaborative filtering can be explored.

## Deployment

- **A/B testing**: Assign users to treatment(new model) and control(existing) groups.
- Evaluate outcomes; CTR ,sales of the books
- Complex models with lots of data take a long time to train
- Do **scheduled** batch prediction
- Save results to **in-memory datastores** for fast retrieval( Redis, Memcached)
- Enable **autoscaling** on API servers





# Demo



# Thanks



CREDITS: This presentation template was created by Slidesgo,  
including icons by Flaticon, and infographics & images by Freepik.

Please keep this slide for attribution.

