

ProxLib: What to read next?

A book recommender system

By Angelica DeGaetano

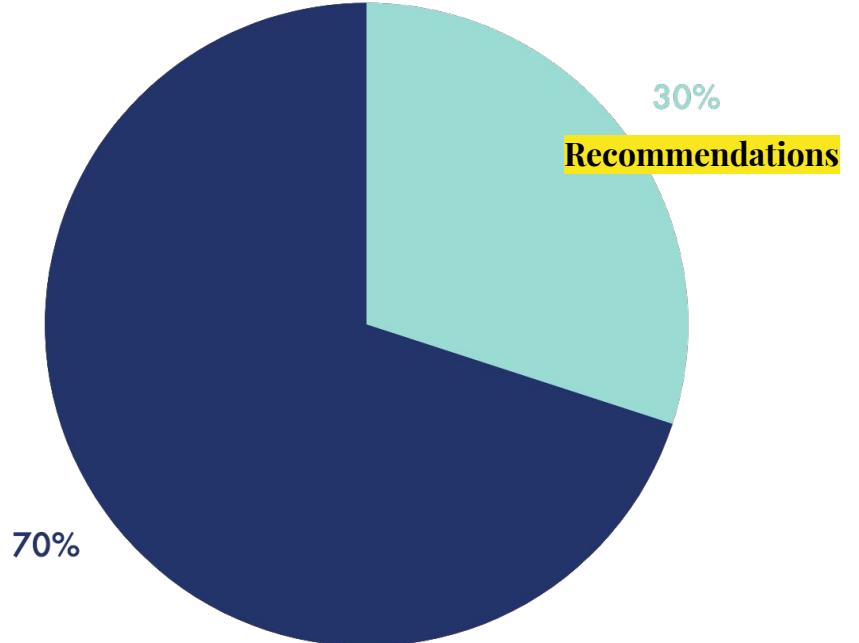


Addressing Choice-Overload with Recommendations



Photo by Susan Yin on Unsplash

E-Commerce Revenue Worldwide





Books:

484



Reviewers:

28,959



didn't like



OK



liked



really liked

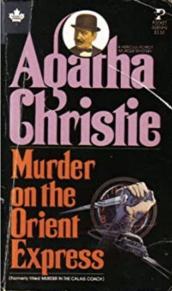
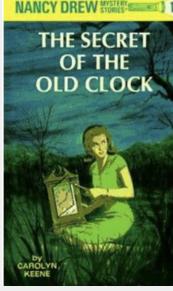
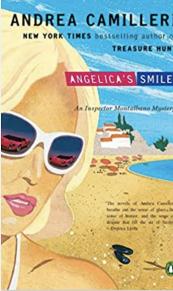
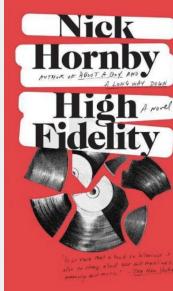


amazing

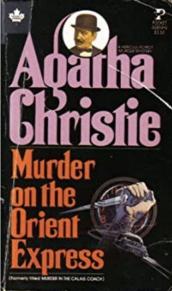
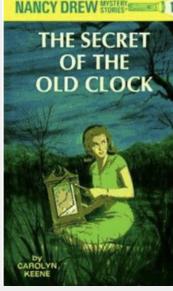
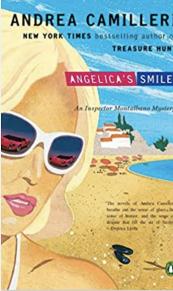
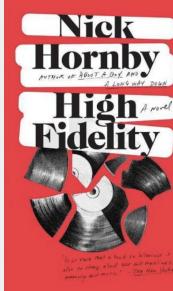
Ratings:

54,387

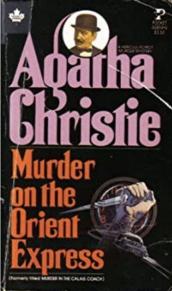
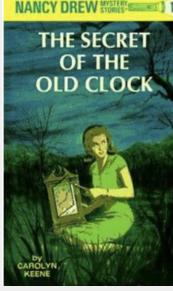
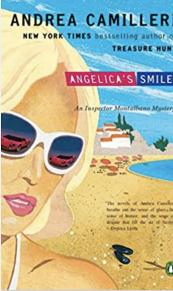
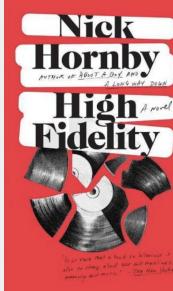
Because of Others Like You--Collaborative Filtering

	1 	2 	3 	4 	5 
Ana	★★★★		★★★★★	★	
Chris	★★★	★★★★	★★★★		
Sarah	★			★★★★★	★★★★★
Alex		★★★★★	★★★★★	★★	★

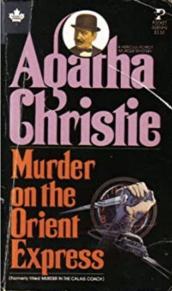
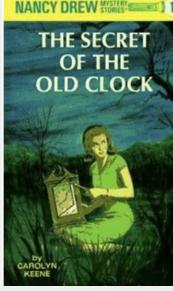
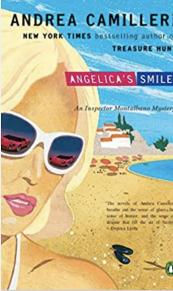
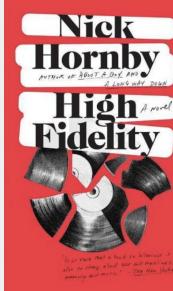
Because of Others Like You--Collaborative Filtering

	1 	2 	3 	4 	5 
Ana	★★★★	?	★★★★★	★	?
Chris	★★★	★★★★	★★★★	?	?
Sarah	★	?	?	★★★★★	★★★★★
Alex	?	★★★★★	★★★★★	★★	★

Because of Others Like You--Collaborative Filtering

	1 	2 	3 	4 	5 
Ana	★★★★	?	★★★★★	★	?
Chris	★★★	★★★★	★★★★	?	?
Sarah	★	?	?	★★★★★	★★★★★
Alex	?	★★★★★	★★★★★	★★	★

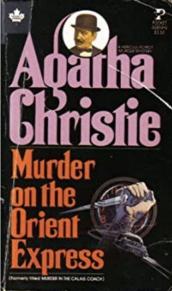
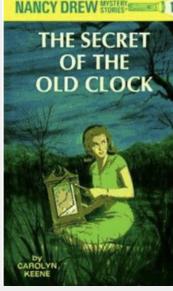
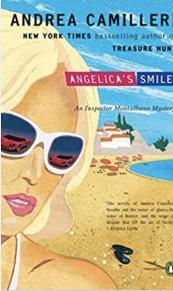
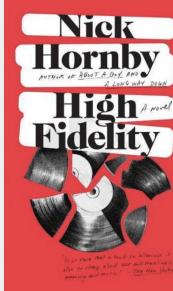
Because of Others Like You--Collaborative Filtering

	1 	2 	3 	4 	5 
Ana	★★★★	?	★★★★★	★	?
Chris	★★★	★★★★	★★★★★	?	?
Sarah	★	?	?	★★★★★	★★★★★
Alex	?	★★★★★	★★★★★	★★	★

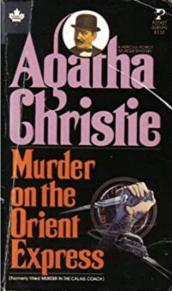
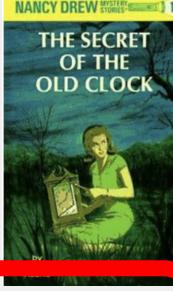
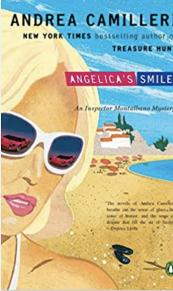
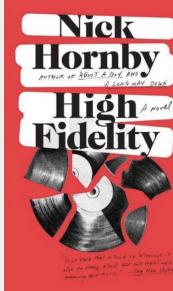
Because of Others Like You--Collaborative Filtering

Ana	★★★★	?	★★★★★	★	?
Chris	★★★	★★★★	★★★★★	?	?
Sarah	★	?	?	★★★★★	★★★★★
Alex	?	★★★★★	★★★★★	★★	★

Because of Others Like You--Collaborative Filtering

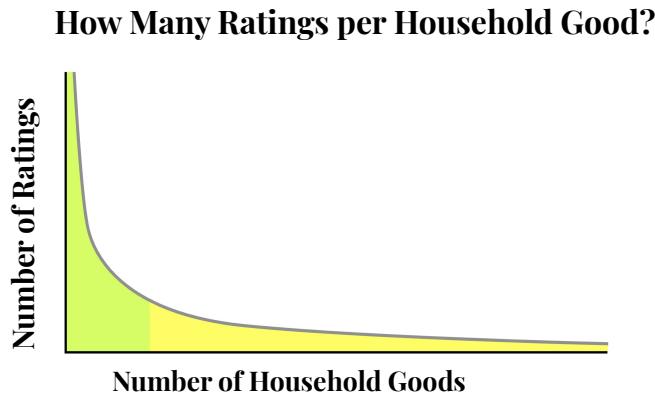
	1 	2 	3 	4 	5 
Ana	★★★★	?	★★★★★	★	?
Chris	★★★	★★★★	★★★★★	?	?
Sarah	★	?	?	★★★★★	★★★★★
Alex	?	★★★★★	★★★★★	★★	★

Because of Others Like You--Collaborative Filtering

	1 	2 	3 	4 	5 
Ana	★★★★	★★★★	★★★★★	★	?
Chris	★★★	★★★★	★★★★★	?	?
Sarah	★	?	?	★★★★	★★★★★
Alex	?	★★★★★	★★★★★	★★	★

Rating Distribution

Long Tail Problem
Example

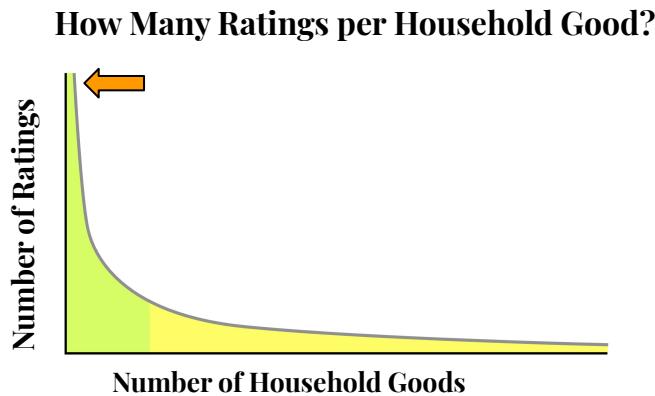


This Dataset

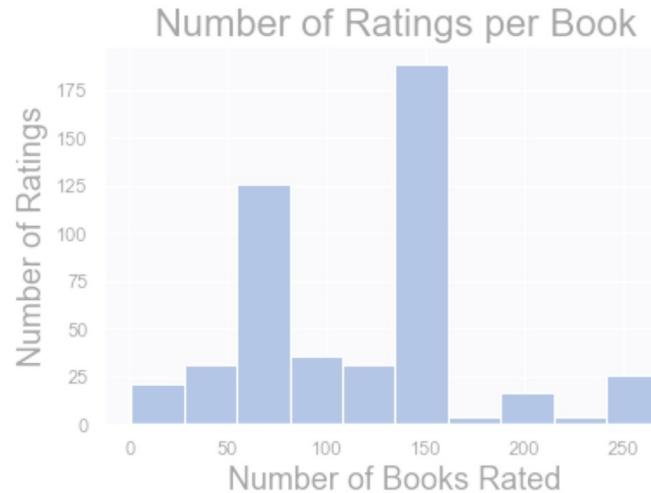


Rating Distribution

Long Tail Problem
Example

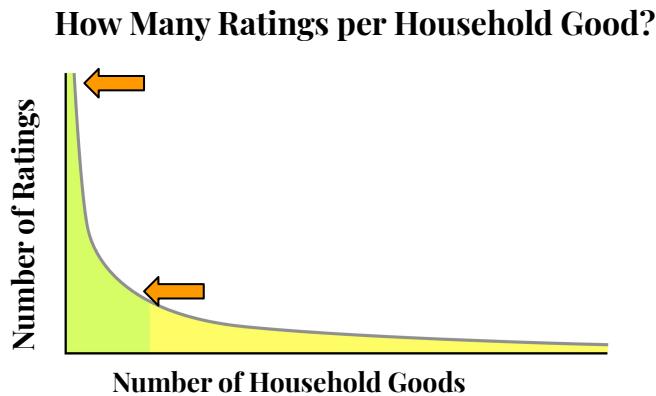


This Dataset

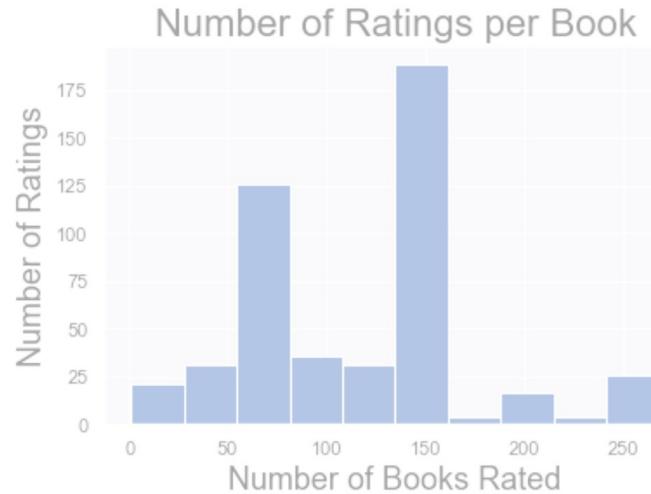


Rating Distribution

Long Tail Problem
Example

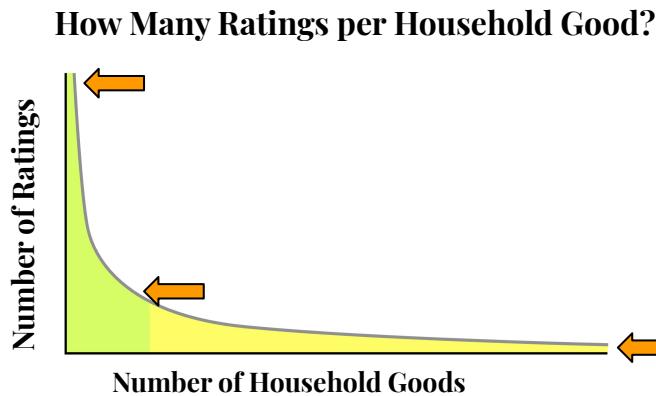


This Dataset

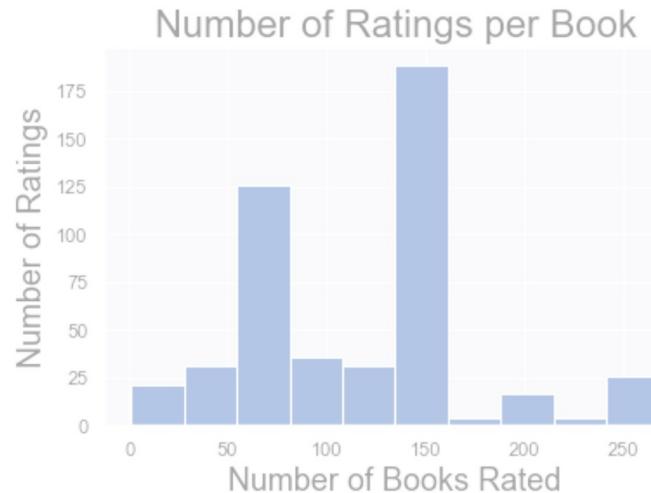


Rating Distribution

Long Tail Problem
Example

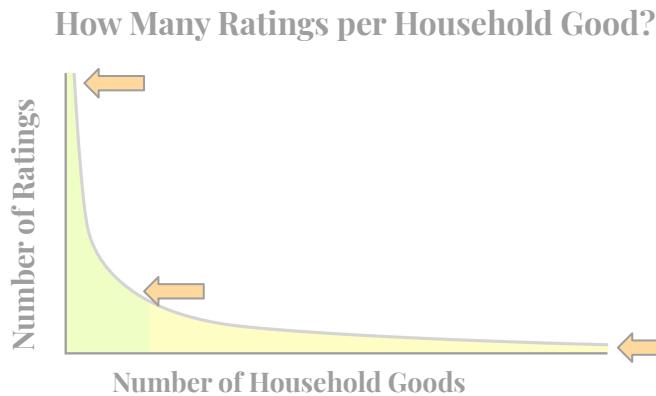


This Dataset



Rating Distribution

Long Tail Problem
Example

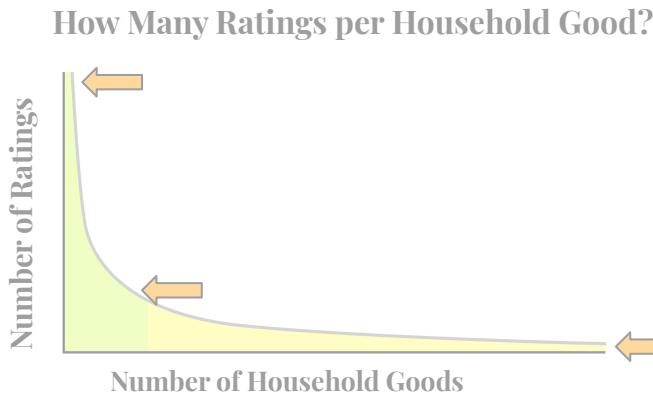


This Dataset

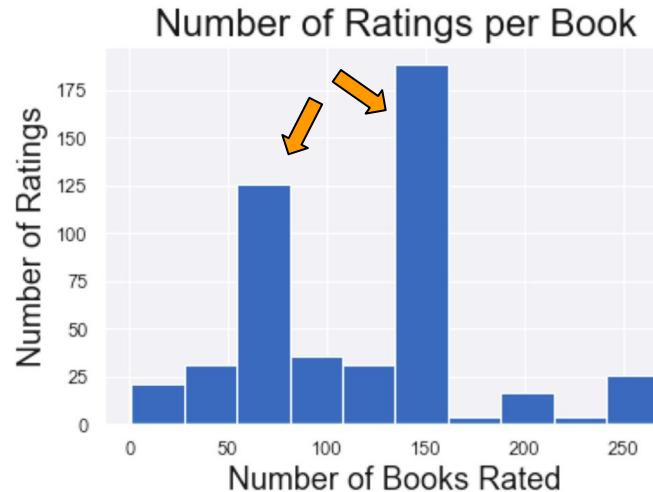


Rating Distribution

Long Tail Problem
Example

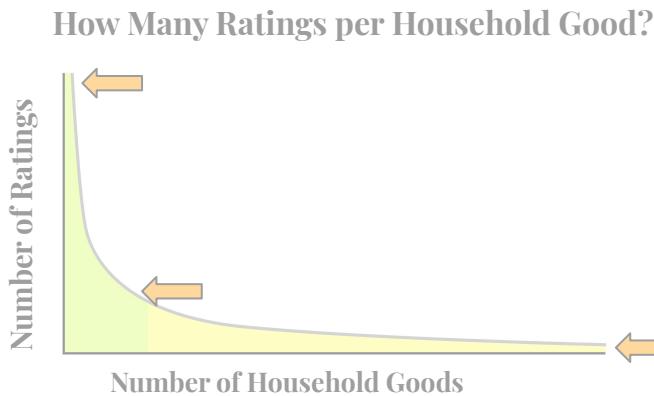


This Dataset



Rating Distribution

Long Tail Problem
Example



This Dataset





Dataset Sparsity: 99.60%

Model:  ALS

Best model RMSE: 1.2


1
didn't like


2
OK


3
liked


4
really liked


5
amazing



The Model

Dataset Sparsity: 99.60%

Model:  ALS

Best model RMSE: 1.2



On average, the model's prediction is off by:
1.2 stars

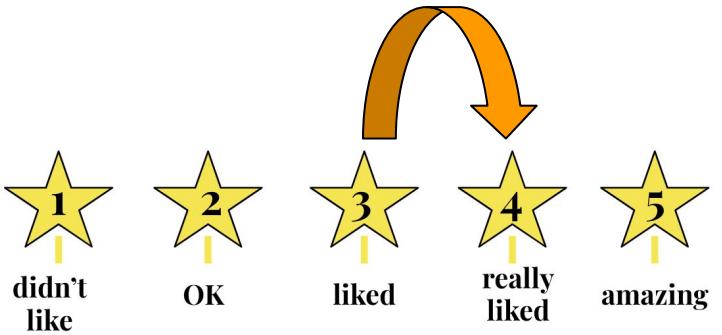


The Model

Dataset Sparsity: 99.60%

Model:  ALS

Best model RMSE: 1.2



On average, the model's prediction is off by:
1.2 stars

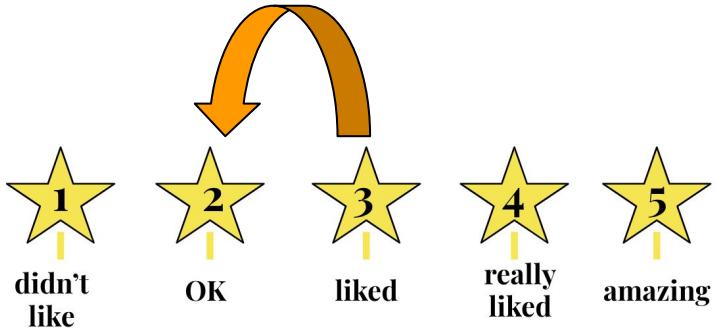


The Model

Dataset Sparsity: 99.60%

Model:  ALS

Best model RMSE: 1.2



On average, the model's prediction is off by:
1.2 stars

The Road to Further Recommendations

- ❑ Scrape more data
- ❑ Natural Language Processing on text reviews
- ❑ Content-based filtering
- ❑ Deployment of model



Photo by Laura Käpfer on Unsplash

Contact



github.com/angelicacodes



linkedin.com/in/degaetanoma



AngelicaCodes@gmail.com



Cover: Photo by [Patrick Tomasso](#) on [Unsplash](#)

Title: Photo by [Nick Fewings](#) on [Unsplash](#)

Book clipart: books by [Rocket000](#)

Group clipart: People by [OpenClipart](#)

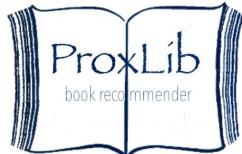


Photo by [Jared Craig](#) on [Unsplash](#)