

# GOODREADS

## EDA & BASELINE MODELING

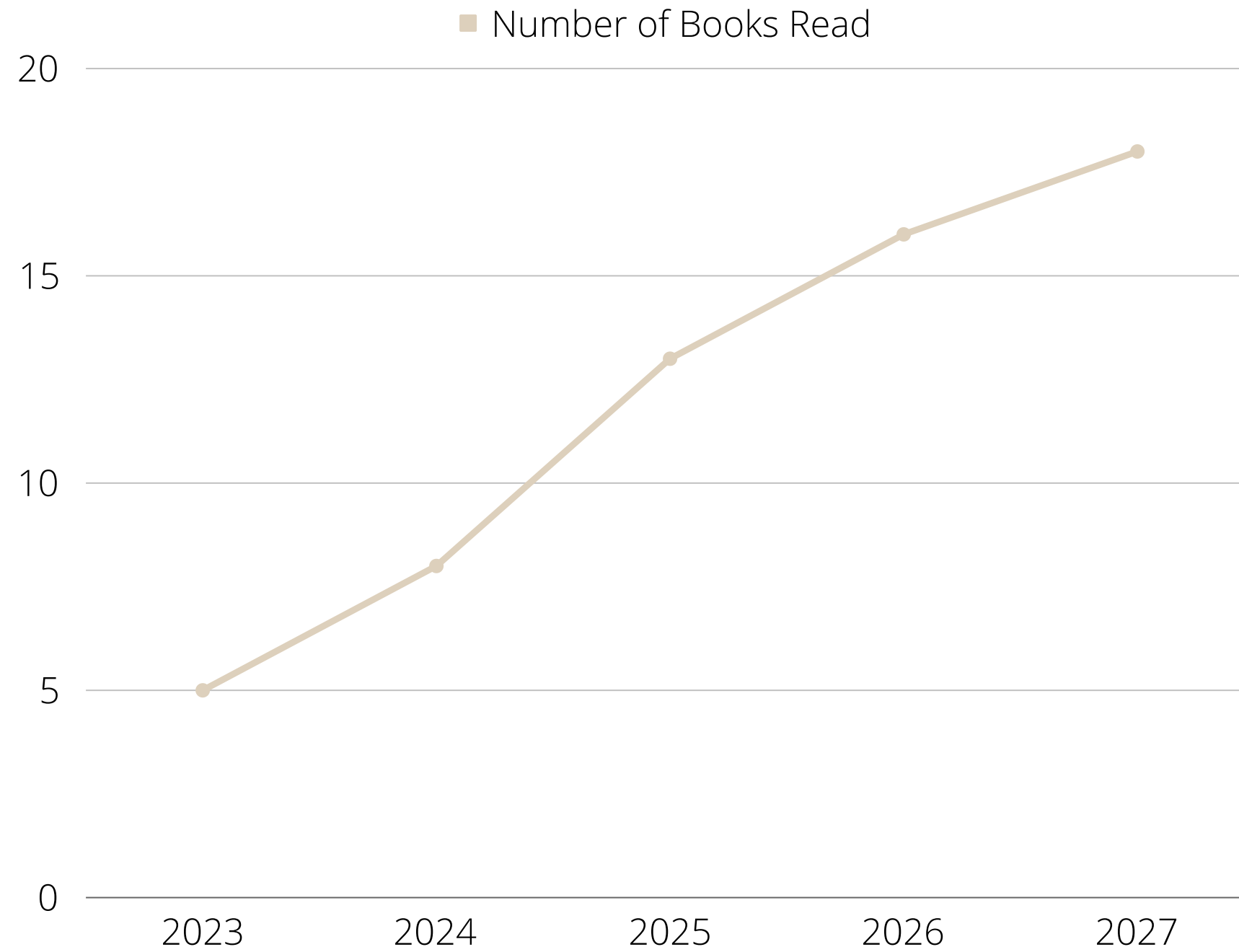
BY QUINCI BIRKER





# Problem & Solution

- Problem: If you pick up a book you don't like, it can put you off reading for a long time.
- Solution: Personalized book recommender system



# Project Impact

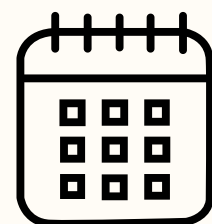
- Overall goal: Make reading more popular
- Target: Occasional and avid readers
- Impact: Increase revenue for bookstores

# Dataset Overview w/ Preprocessing

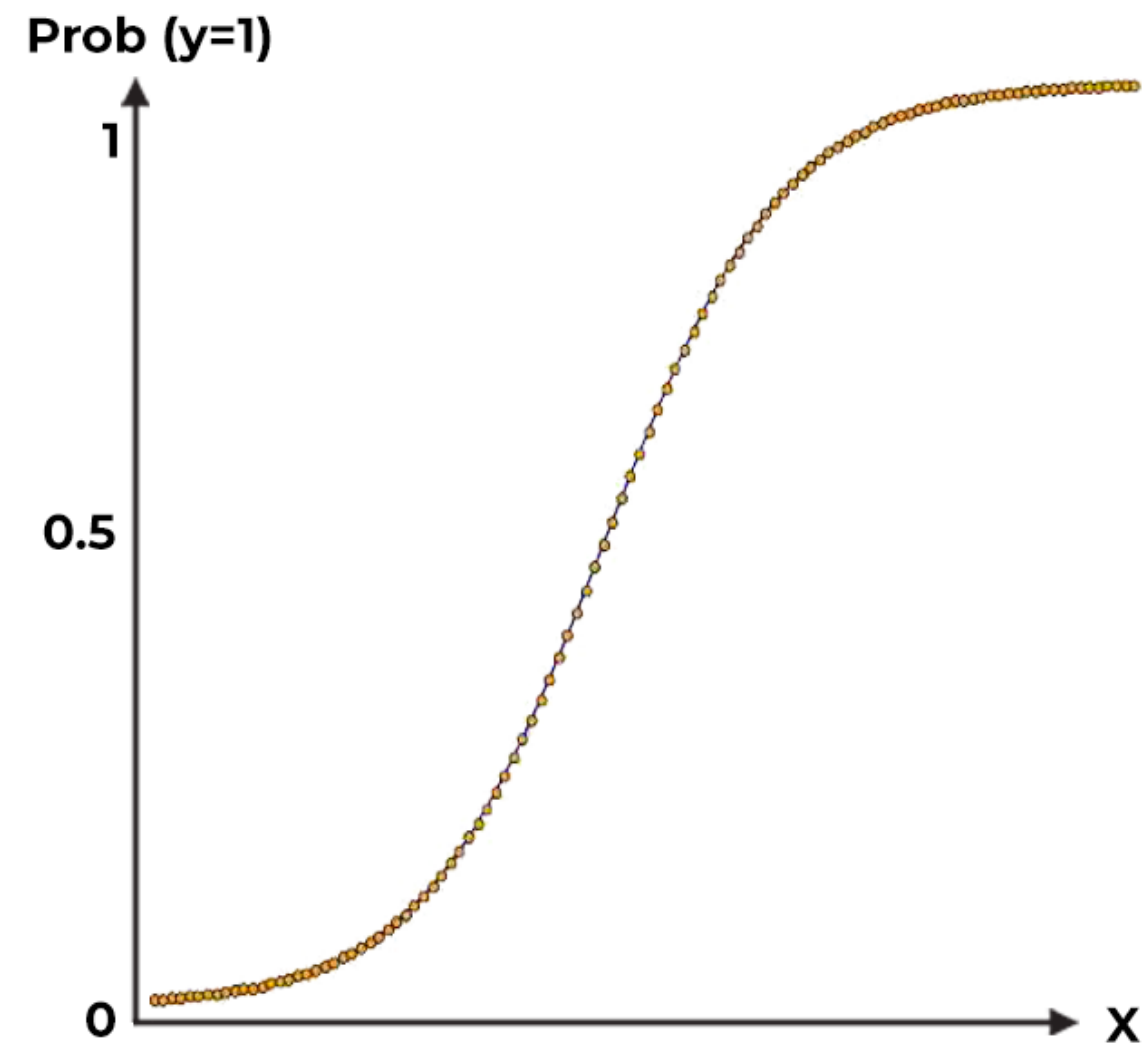
- 52,400 rows → 41,300 rows
- 25 columns → 45 columns

	pages	numRatings	bbeScore	year	month	series_bool	genre_count	Adult	Adventure	Audiobook	Chick Lit	Childrens
35783	22.0	666	89	1974	1	0	52	0	0	0	0	1
39773	176.0	161	83	1983	1	0	95	0	0	0	0	0
51074	339.0	2535	24	2014	5	1	171	0	0	0	1	0

- Target variable: Rating Score
  - 1 - 0.52
  - 0 - 0.48
- Independent variables:

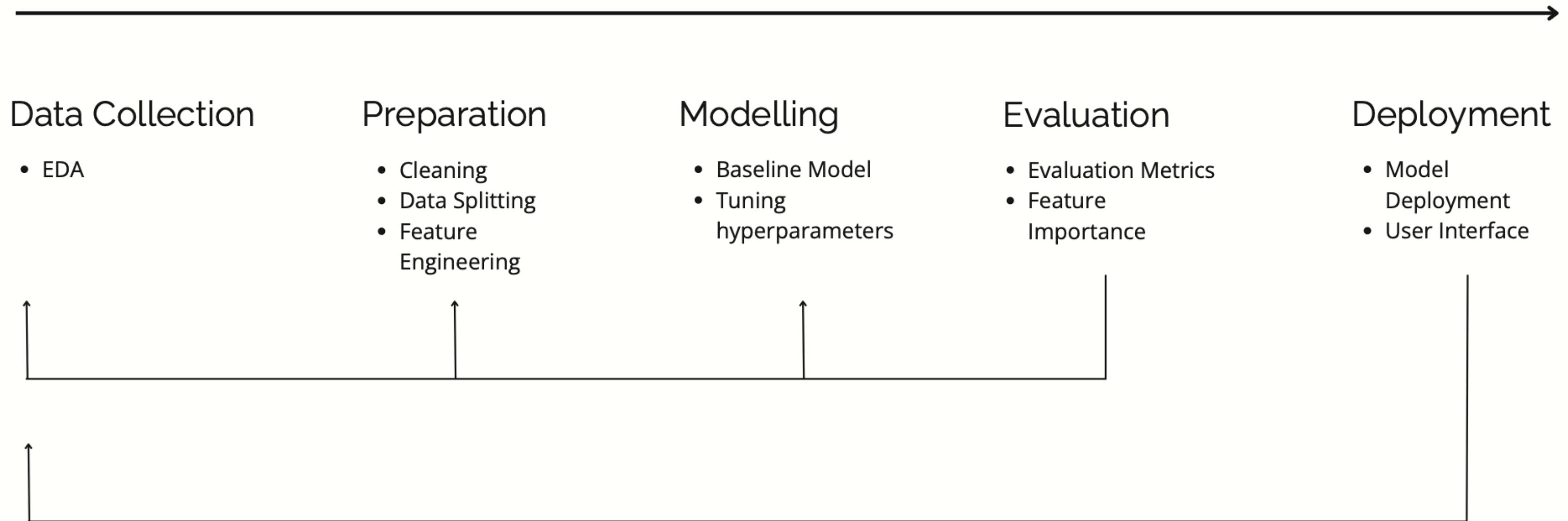


# Baseline Model



Average Cross-Validation Accuracy: 54%

# Process Overview



*Thank-you!*

