

Customer Segmentation By Book Purchase History





Agenda



Problem
background

SMART & SWOT
Analysis

Competitor
Analysis

Data Analysis

Model
Evaluation

Insights

Problem Background



Problem Background

Background: the previous approach from our client is the "most popular" items recommendation system, based on customers' purchase history and the purchased items' relationship with the ecommerce website, that directs users to add items to their shopping carts. This method has successfully increased sales for this website in the past, but we have a better "Recommendation Engine" system that can apply on a broader customer pool (by clustering customers and apply different further approaches to different segments).

Data: Online Book Ecommerce Data.

Approach: Our "Recommendation Engine" AI solution for this ecommerce website is to implement the unsupervised clustering(K-Means) method, which creates customer segmentations (by giving different RFM scores, also based on their prior purchases and relationship with the ecommerce website), and evaluated the results (for further actions catering to different customer segments).

SMART & SWOT Analysis



SMART Framework

Specific:

Increase total revenue by 8% of the book selling website within the US.

Measurable:

Increase total units sale by 10%. In addition, increase sales by 12% for most engaged customers and 3%-5% for the rest of customers.

Actionable:

Develop an AI Recommendation Engine to create customer segmentation, send out promotions or product suggestions by email / text messages / notifications to establish great relationship with customers. Increase marketing spending to incentivize customers with low engagement or to encourage new customers to make purchase.

Relevant:

Significant improvements in the recommendation engine system would drive better accuracy to customer in different segments based on their prior purchases.

Time-bound:

Build, test, and implement the Recommendation Engine in 30 days; Drive to increase sales volume by 10% from the most engaged customer and 3% - 5% from the rest in 90 days.



SWOT Analysis

Strength

- Affordable prices, discounts
- Products cover wide range of book categories
- Brand recognition among existing customers

Weakness

- Resources spent on inventory management might erode profits
- Do not currently have a recommendation system, leads to a higher rate of missed opportunity
- Requires more investments in new technologies

Opportunity

- Create physical retail stores as a community experience
- Create global market, attract more users from overseas
- Implementing data analysis to better fit customer needs

Threat

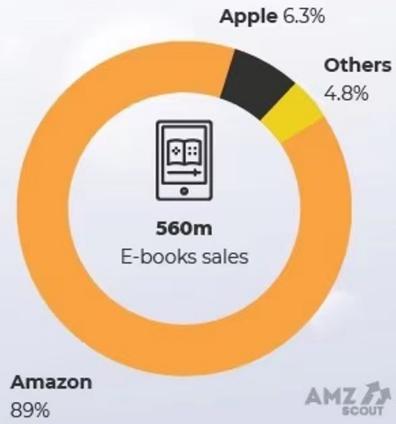
- Increase amount of eBook platforms and electronic version of books
- Uncertainties in the supply of new products, lead to high and low swings of sales
- Price differences with competitors

Competitor Analysis



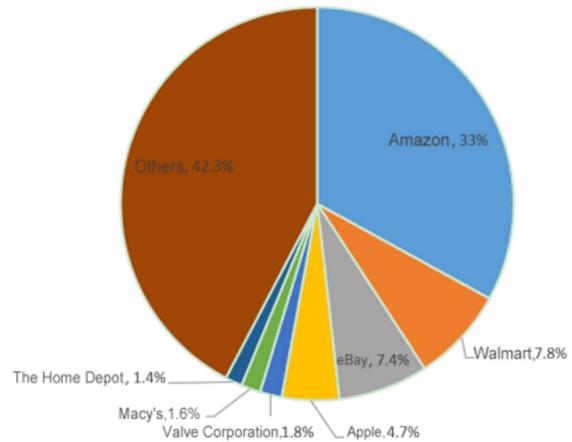
Competitor Analysis

**Amazon is largest book seller
in the US**



chinabrands

The E-commerce Market of US





Competitor Analysis

On-site personalized recommendations

- Item-to-Item Collaborative Filtering
- Improve personalized shopping experience
- Increase average order value and the amount of revenue generated from each customer
- Produces recommendations in real time, scales to massive data sets, and generates high quality personalized recommendations.
- McKinsey estimated that 35 percent of consumer purchases on Amazon come from product recommendations

The screenshot shows the Amazon.com homepage with a 'Recommended for You' section. The section features the Amazon logo and a heading 'Recommended for You'. Below this, a message states: 'Amazon.com has new recommendations for you based on items you purchased or told us you own.' Four product recommendations are displayed with their titles and purchase links:

Image	Title	Link
	<u>The Little Big Thing: 163 Ways to Pursue EXCELLENCE</u>	<u>[Blu-ray]</u>
	<u>Fascinate: Your 7 Triggers to Persuasion and Captivation</u>	<u>[Blu-ray]</u>
	<u>Sherlock Holmes [Blu-ray]</u>	
	<u>Alice in Wonderland [Blu-ray]</u>	



Competitor Analysis

Off-site personalized recommendations with Emails

- Divide customers into different segments: homepage browsers, category browsers, product browsers, Site search browsers, and cart abandonment browsers
- Only recommend products and brands that this person has viewed on their site or items they had added to their cart.
- Increase CTR by 300% or more

Amazon.com: Bestselling Canon Cameras [Newsletters](#) | X

★ [Amazon.com to me](#) show details May 30 (9 days ago) [Reply](#) ▾

amazon.com **More to Explore**

Customers who have shown an interest in point-and-shoot cameras might like to see this week's bestselling models.



Canon PowerShot A495 10.0 MP Digital Camera with 3.3x Optical Zoom and 2.5-Inch LCD (Blue)	Canon PowerShot A3000IS 10 MP Digital Camera with 4x Optical Image Stabilized Zoom and 2.7-Inch LCD	Canon PowerShot ELPH 300 HS 12 MP CMOS Digital Camera with Full 1080p HD Video (Black)	Canon PowerShot S95 10 MP Digital Camera with 3.8x Wide Angle Optical Image Stabilized Zoom and 3.0-Inch inch LCD
------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------

Data Analysis

Data Analysis- process

Step 3 Data visualization

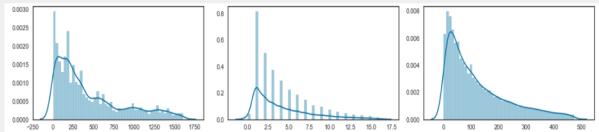
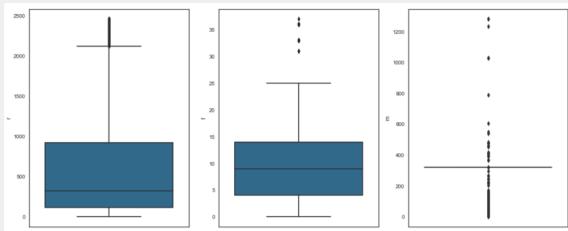
Specifically for the variable
'tof' - see next page

Step 1 Data preview

	id	r	f	m	tof	label
count	3.371300e+04	33713.000000	33713.000000	33713.000000	33713.000000	33713.000000
mean	8.234209e+06	508.427254	5.820603	251.924325	1296.088186	
std	4.426015e+06	542.440039	7.008049	3929.929826	818.976321	
min	9.140000e+02	0.000000	0.000000	0.000000	0.000000	
25%	4.391616e+06	117.000000	1.000000	39.899994	536.000000	
50%	7.967691e+06	293.000000	3.000000	102.629944	1311.000000	
75%	1.199823e+07	712.000000	7.000000	247.297485	2132.000000	
max	1.625264e+07	2460.000000	118.000000	532892.000000	2462.000000	

Process

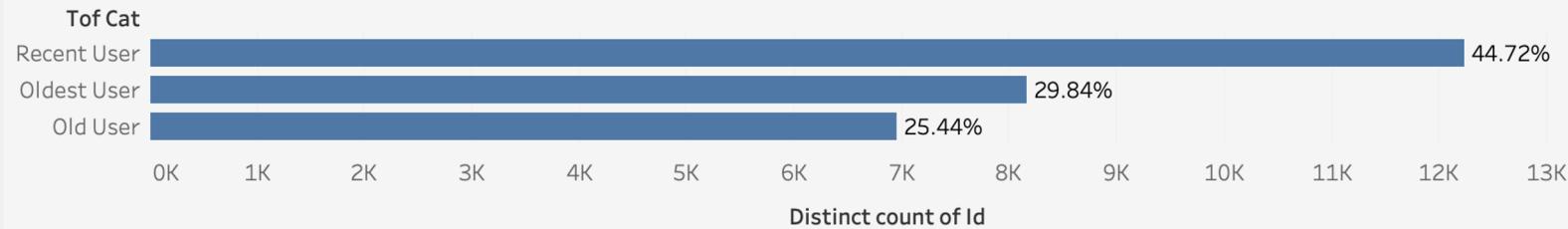
Step 2 Remove outliers



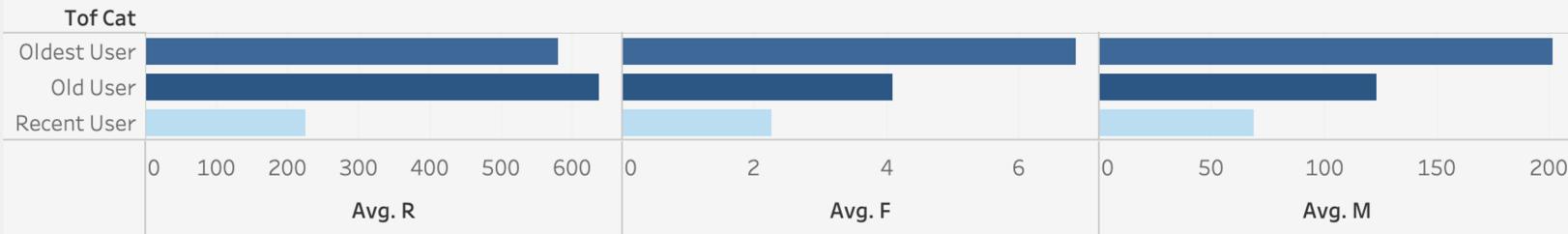


Data Analysis - tof

New/old user distribution



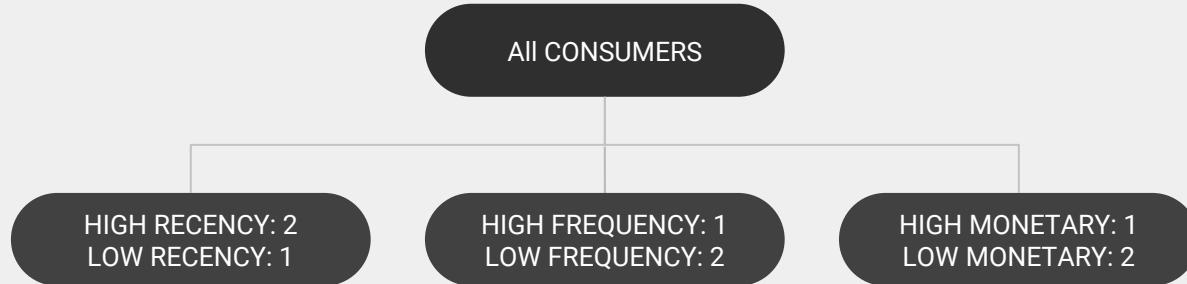
Tof category by Avg R, F, M



Model Evaluation



RFM Score



111: Champions (N=5794)

112: Loyal (N=714)

121: Potential Loyalist(N=1349)

122: New Customers(N=5854)

211: Promising (N=4790)

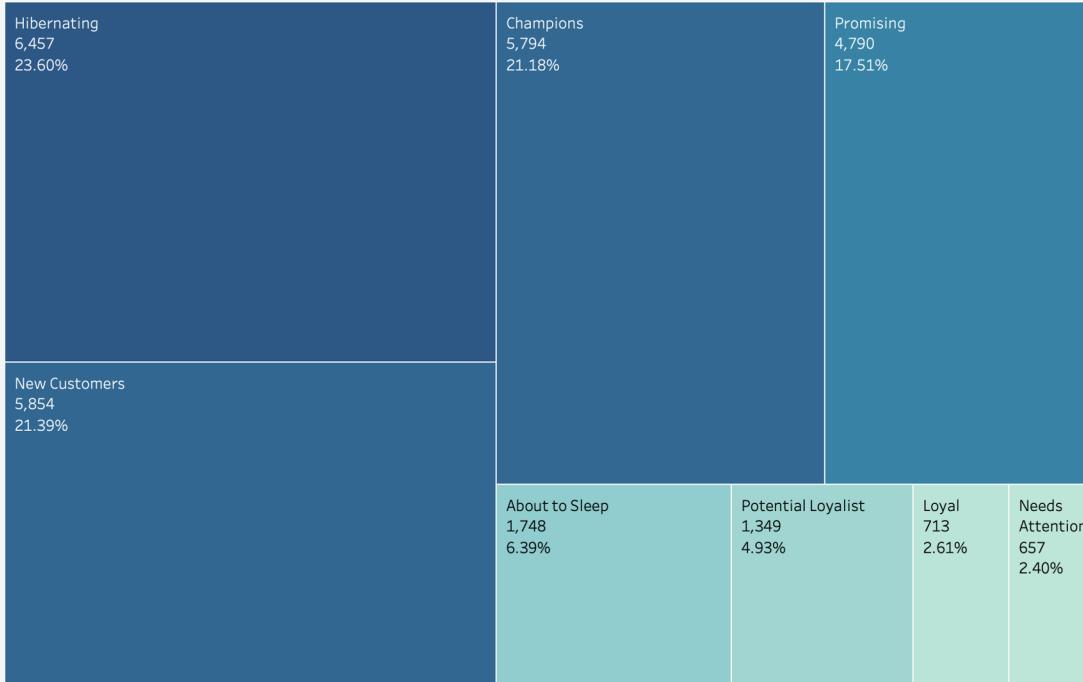
212: Needs Attention(N=657)

221: About to Sleep (N=1748)

222: Hibernating (N=6457)



RFM Score



Cons:

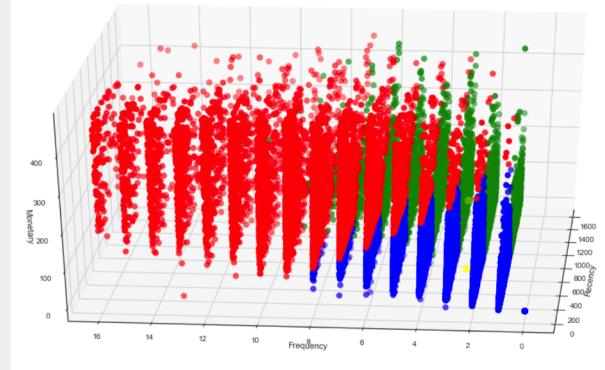
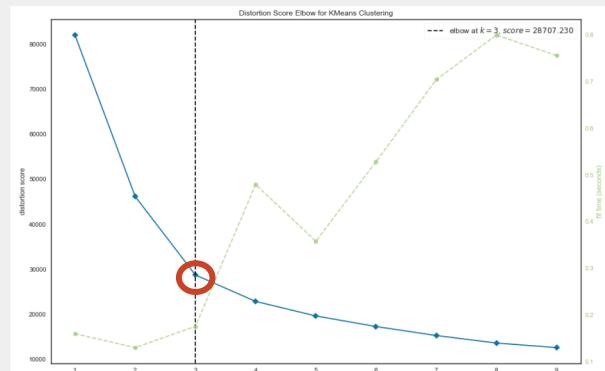
1. Too many segments
2. Hard to define these 8 clusters

K-Means

Step1: Normalize data

Step2: Choose optimal K with
Elbow method (k=3)

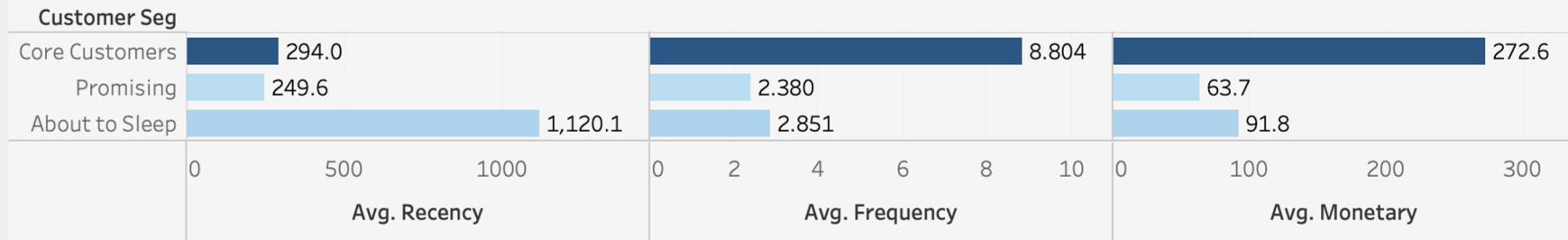
Step3: Label each record



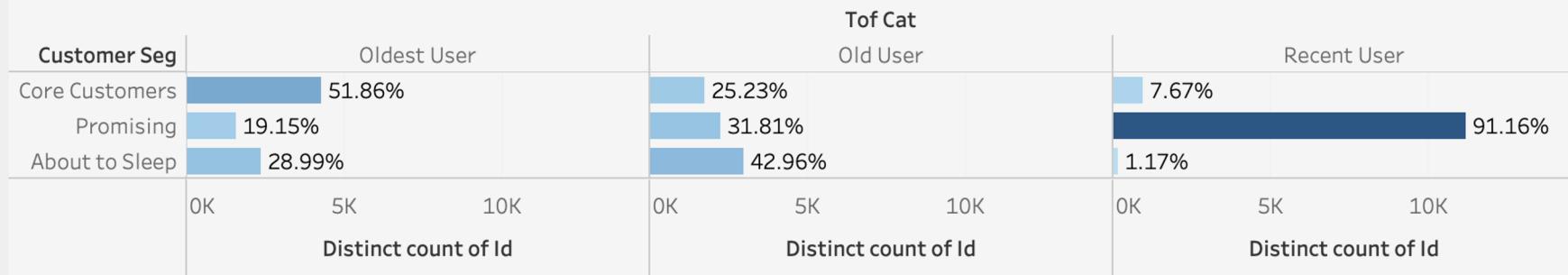


Segments at-a-glance

Segments by Avg. RFM



Segments by Tof Category





Segments At-a-glance



Core customers



Promising

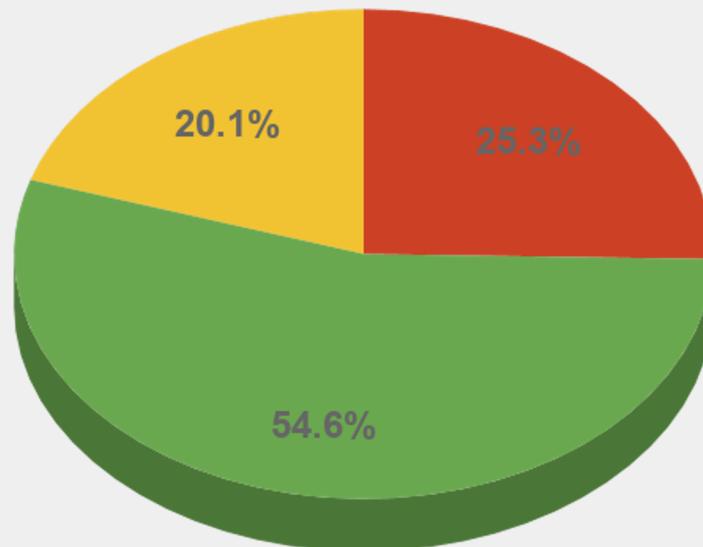


About to Sleep

Sample Size	5,500
Avg. recency	1120.1
Avg. frequency	2.851
Avg. monetary	91.8
Pct. of oldest users	28.99%

Pct. of old users	42.96%
Sample Size	14,932
Pct. of recent users	57.04%
Avg. recency	1.17%
	249.6
Avg. frequency	2.380
Avg. monetary	63.7

Pct. of oldest users
19.15%

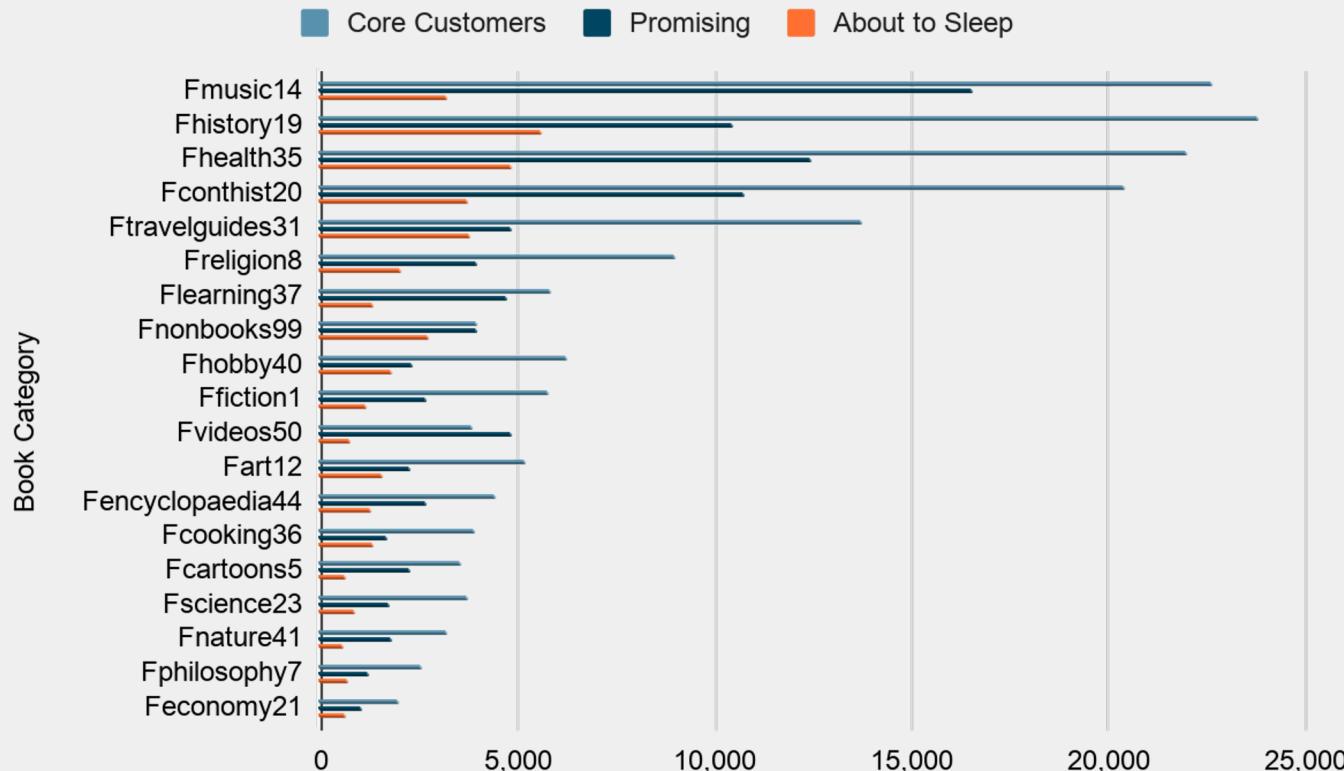


Sample Size	6,930
Avg. recency	294
Avg. frequency	8.804
Avg. monetary	272.6
Pct. of oldest users	51.86%

Pct. of old users
25.23%
Pct. of recent users
7.67%

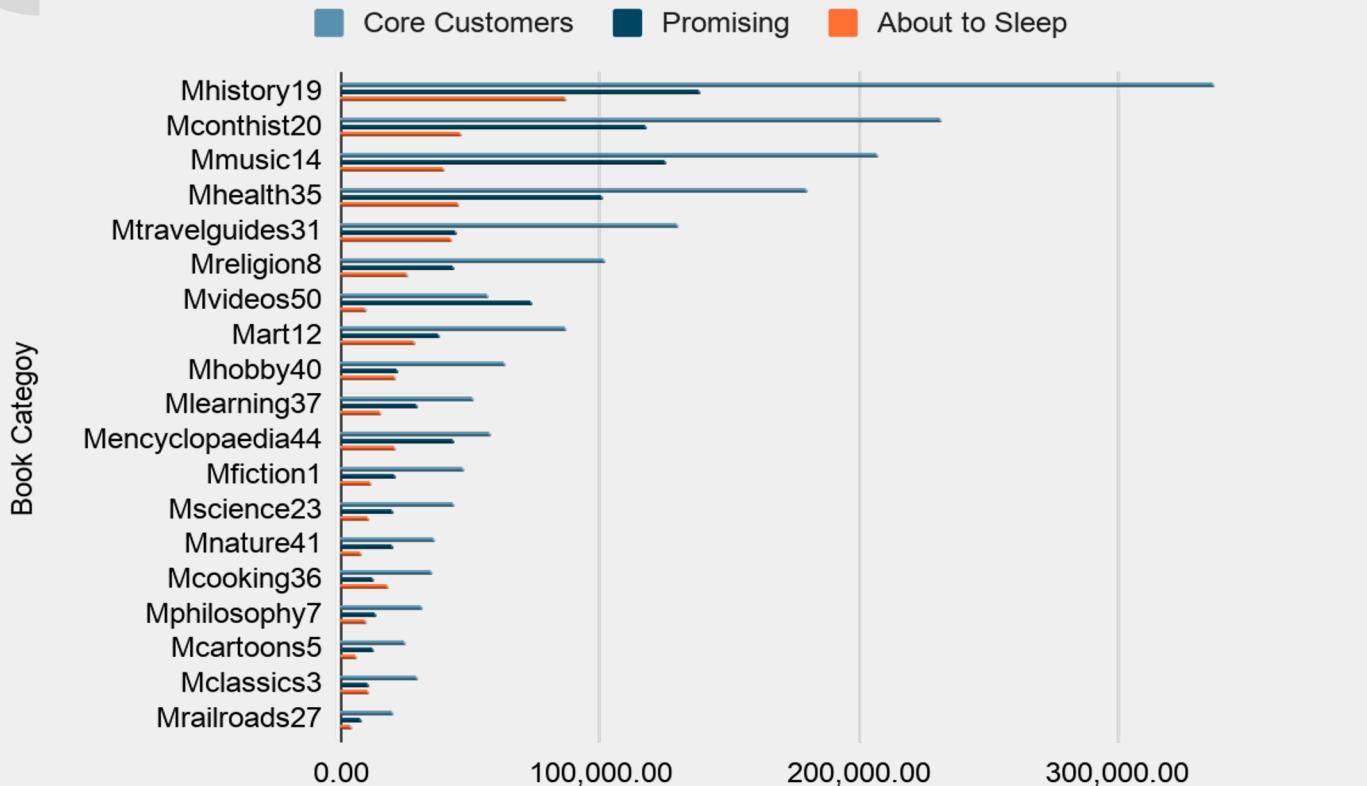


Frequency VS Book category





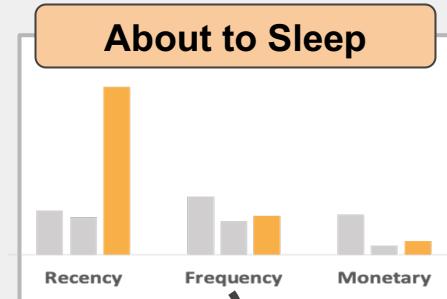
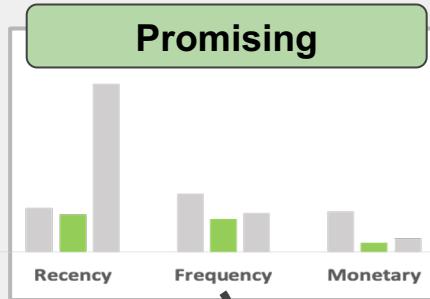
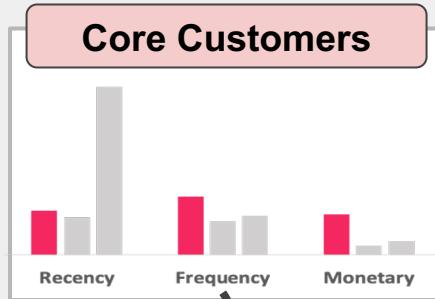
Monetary VS Book category



Insights



Action Plan



Feature

- Loyal (largest portion of oldest users)
- Active (High frequency & monetary, low recency)

Strategy

- Cannot lose

Action

- Recommend popular items

Feature

- New (largest portion of new users)
- Not very active (low frequency & monetary, low recency)

Strategy

- Gain interest

Action

- Promotion on preference

Feature

- Mainly old users (Almost no recent users)
- Inactive (Low frequency & monetary, high recency)

Strategy

- Win back

Action

- Promotion on membership
- Personalized Email recommendation



Advantages of New AI Solution

1. More Personalized & Flexible

2. Intelligent & Self-service Tool

3. More competitive

4. Economic Benefit

Thanks for listening!

