



Recommendation Engine Improvements

Team GUGU

- ◆ Xianqian(Sarah) Yang
- ◆ Kejia Ran
- ◆ Minchuan(Miko) Liu
- ◆ Biao Zhang
- ◆ Yuanyang(Jerry) Chang
- ◆ Yue Sun

Our Core Team

Data Team

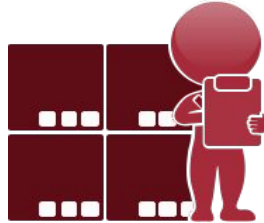


Sr. Data Scientist



Website Analyst

Marketing Team



VP Marketing



Marketing Manager

Financial & Customer Relation Team



Ecommerce Financial Manager



Customer Satisfaction Manager



AGENDA

01. Executive Summary

02. Data Analytics

03. Marketing Plan

04. Next Step

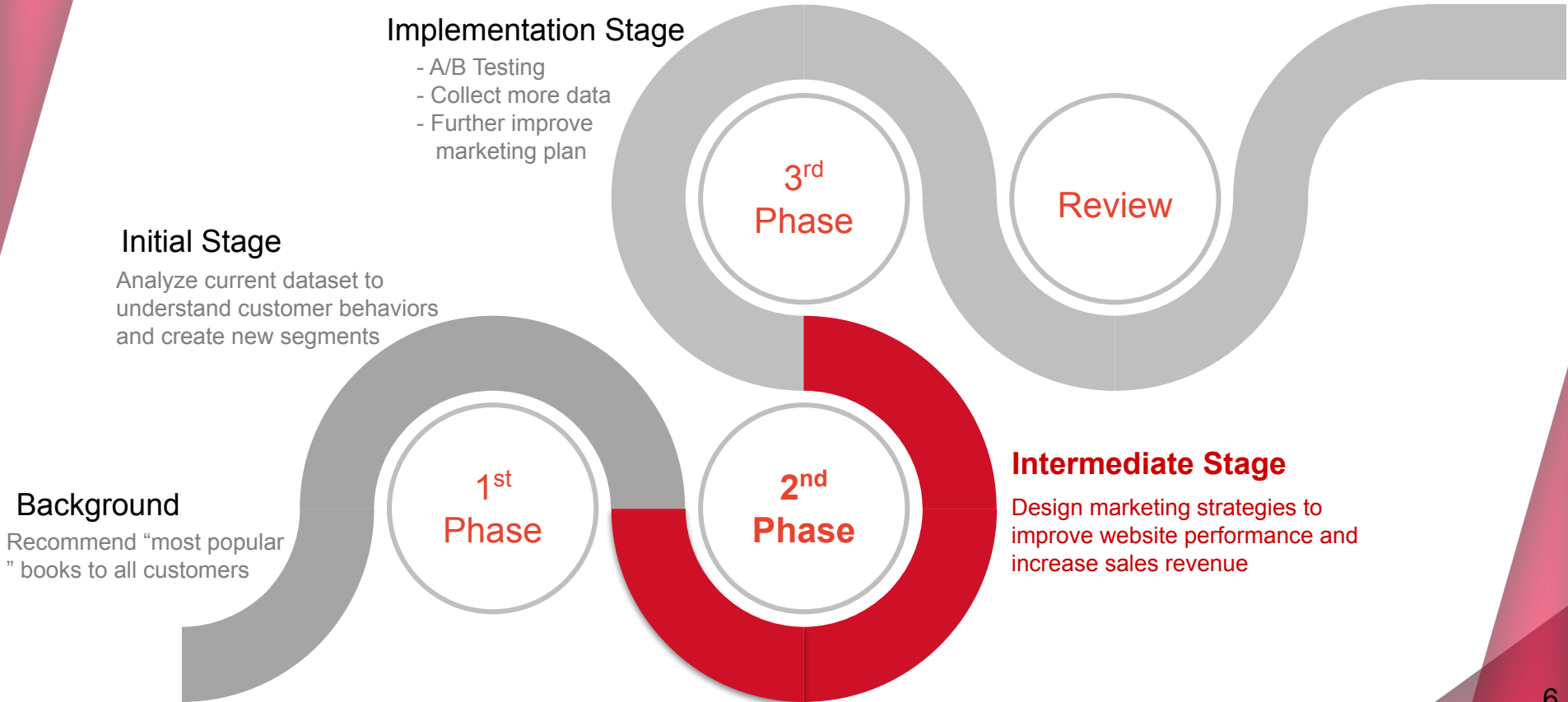
PART 01

Executive Summary

Executive Summary

- We would like to improve the existing engine by customizing book recommendations based on consumption behaviors.
- This is achieved by hierarchical clustering, RFM analysis and customer insight analysis.
- As a result, we can boost
 - About **26%** sales from the top **15%** customers
 - About **12%** sales from the remaining **85%** customers

Project Timeline

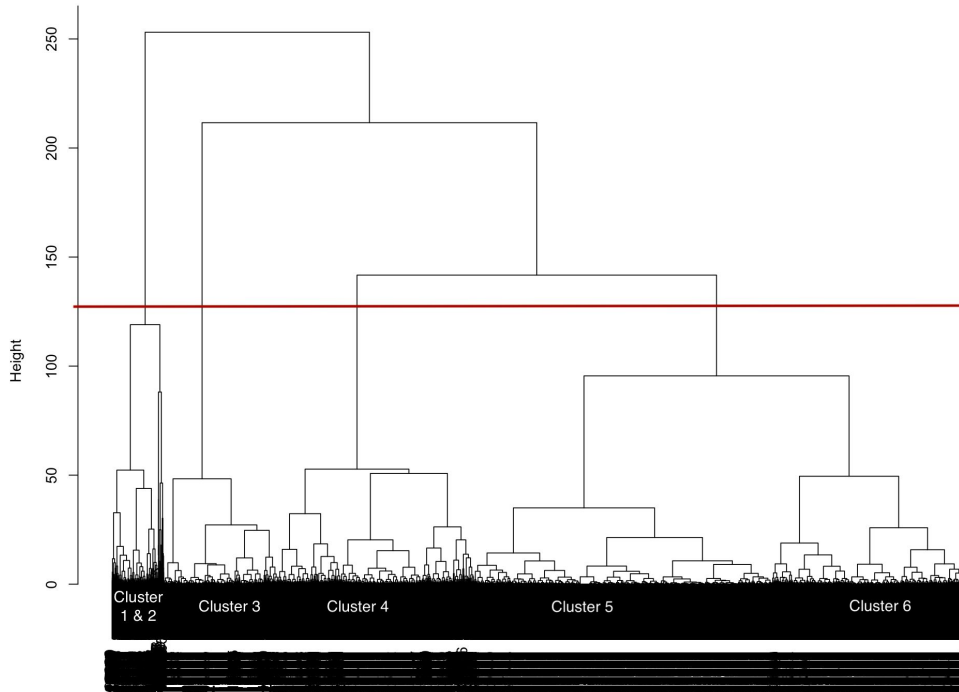


PART 02

Data Analytics

Algorithm and Design Methodology

Cluster Dendrogram



d
hclust (*, "ward.D2")

Design Methodology:

Cluster → Assign RFM Score →
Match Traits → Customize Book
Recommendations

Hierarchical Clustering:

- Ward's Method:
Variance based, no need to specify
numbers of clusters

RFM Analysis:

- Ranges cut off by the mean value of
corresponding clusters
- Assign score (1-6) to each group
(e.g. $r = 6$ for the most recent
customer group)
- Match clients to proper groups based
on their consumption behaviors

Customer Segmentation

Segment	Description	Segment%	FAV Book (Based on F* counts)
Champions	Bought recently, buy often and spend the most	3.28%	history19
Loyal Customers	Spend good money. Responsive to promotions	11.30%	history19
New Customers	Bought more recently, but not often	20.94%	music14
Promising	Recent shoppers, but haven't spent much	5.51%	music14
Need Attention	Above average recency, frequency & monetary values	2.31%	health35
About to Sleep	Below average recency, frequency & monetary values	22.51%	music14
At Risk	Spent big money, purchased often but long time ago	11.72%	history19
Can't Lose Them	Made big purchases and often, but long time ago	0.02%	history19
Lost	Low spenders, low frequency, purchased long time ago	15.62%	history19
Others	incorrect data OR enterprise level customer	6.78%	music14

- Champions/Loyal Customers received the highest scores in each category (e.g. rfm = 666)
- Customers from the middle segments (~40%) have exhibited different book preferences than the most popular item
- Others: we need further investigation into their consumption pattern

Data Visualization

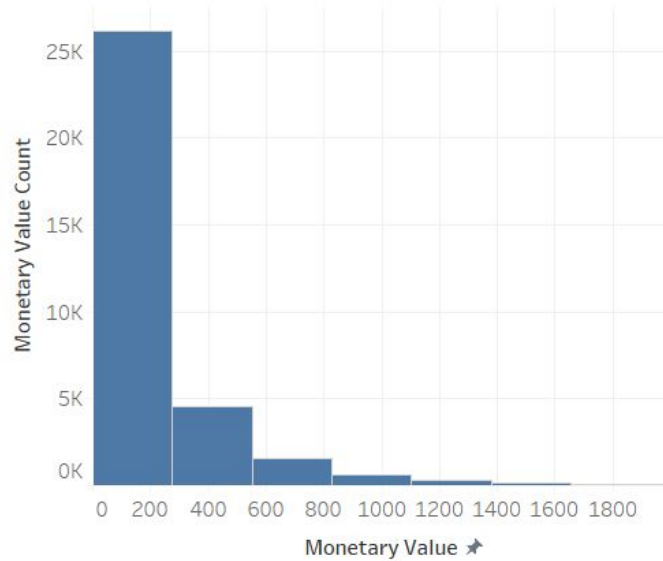
Segment, Median of RFM

Champions 655	Need Attention 433	Promising 411	Others 344
Loyal Customers 644			
New Customers 631	About to Sleep 311		Can't lose 155
	At Risk 234		Lost 112

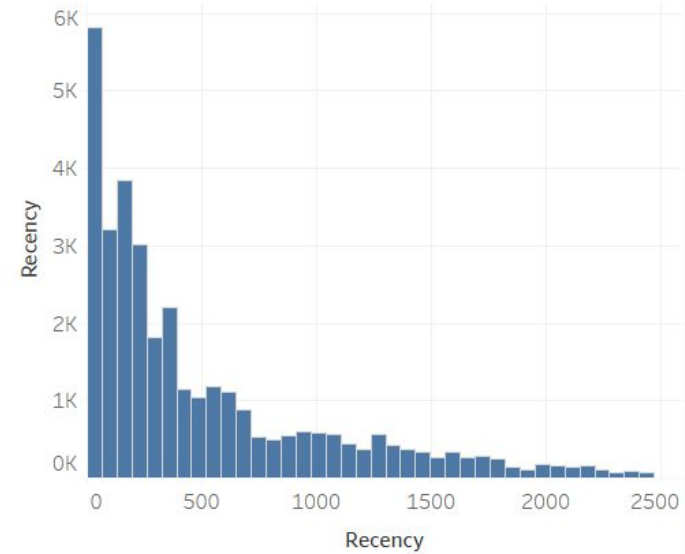
Segment, Customer Counts

About to Sleep: 7,589	Lost: 5,264	At Risk: 3,952	Loyal	
New Customers: 7,059				
	Others: 2,285	Promising: 1,857		

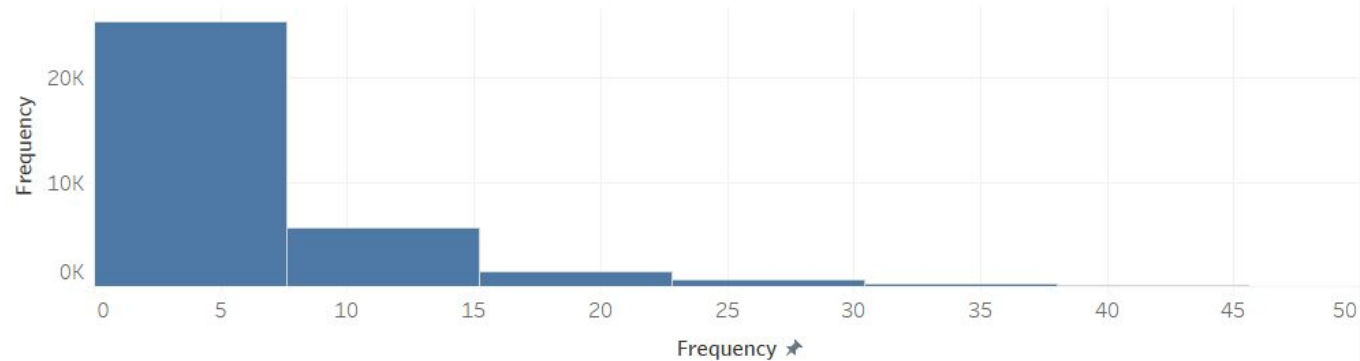
Count vs. Monetary Value



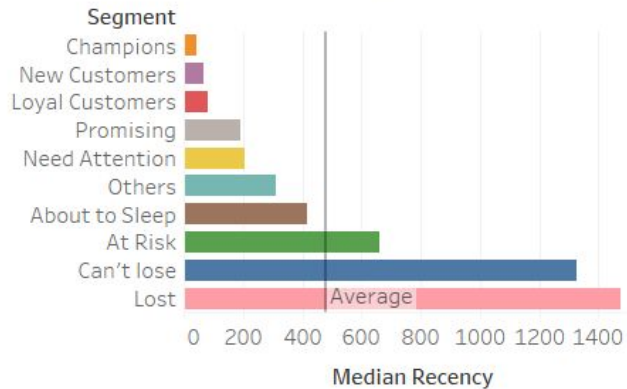
Count vs. Recency



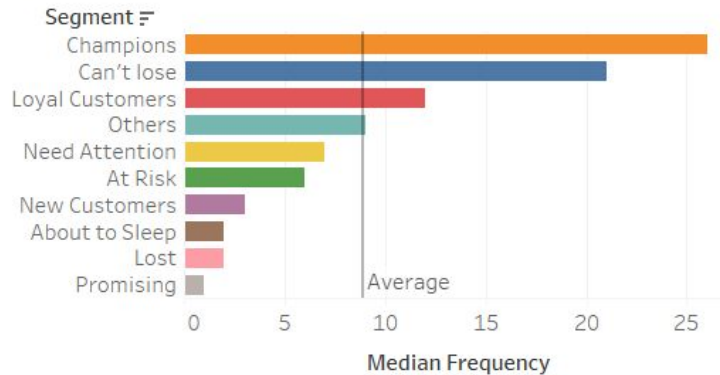
Count vs. Frequency



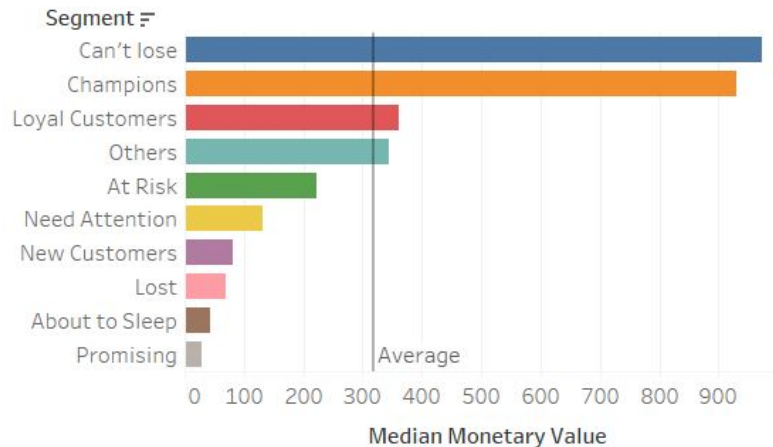
Median Recency by Segment



Median Frequency by Segment



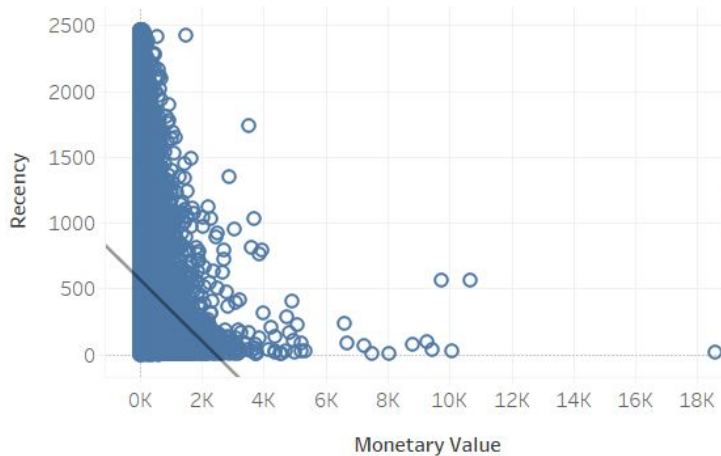
Median Monetary Value by Segment



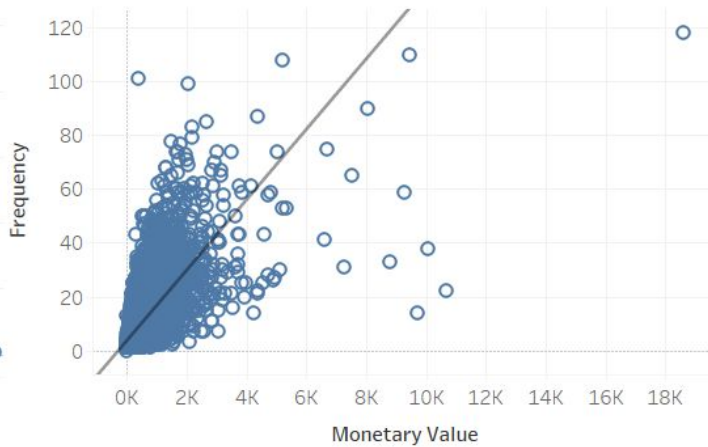
Segment

- Champions
- New Customers
- Loyal Customers
- Promising
- Need Attention
- Others
- About to Sleep
- At Risk
- Can't lose
- Lost

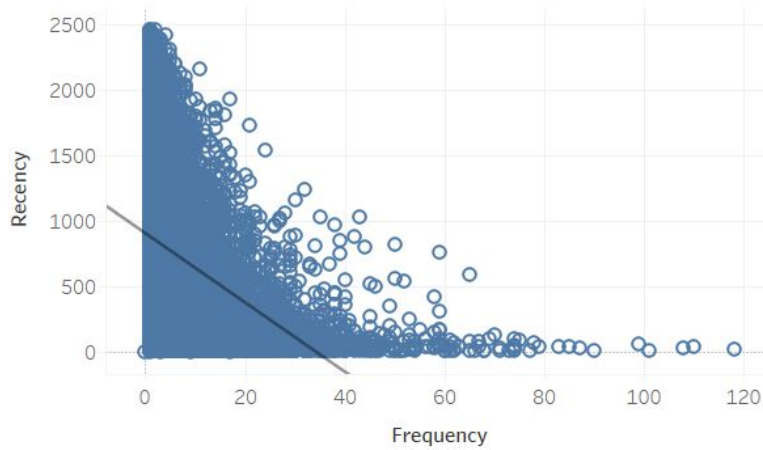
Recency vs Monetary Value



Frequency vs Monetary Value



Recency vs Frequency



PART 03

Marketing Plan

Marketing Plan

Classify Segments

01



02

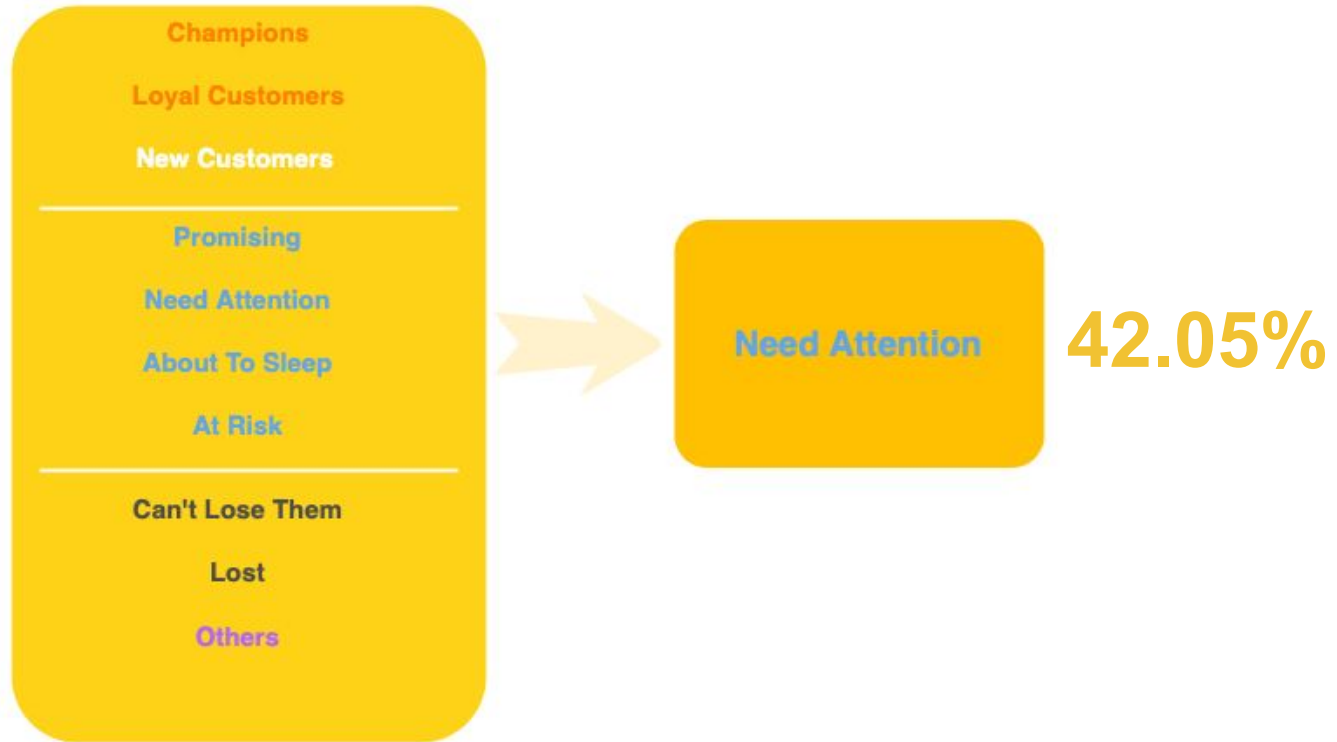
Stimulate Consumption

Book Recommendation

Segment	Description	Segment%	FAV Book (Based on F* counts)
Champions	Bought recently, buy often and spend the most	3.28%	history19
Loyal Customers	Spend good money. Responsive to promotions	11.30%	history19
New Customers	Bought more recently, but not often	20.94%	music14 ★
Promising	Recent shoppers, but haven't spent much	5.51%	music14 ★
Need Attention	Above average recency, frequency & monetary values	2.31%	health35
About to Sleep	Below average recency, frequency & monetary values	22.51%	music14 ★
At Risk	Spent big money, purchased often but long time ago	11.72%	history19
Can't Lose Them	Made big purchases and often, but long time ago	0.02%	history19
Lost	Low spenders, low frequency, purchased long time ago	15.62%	history19
Others	incorrect data OR enterprise level customer	6.78%	music14

To improve the current engine by recommending book types based on consumers' behaviors

Merged Segments



Customer Relationship



Customer Survey

- Understand Customer's feedback and needs
- To improve products and services



Weekly Email Newsletter

- Special Events Announcements
- Holiday Promotion



Exclusive Membership

- Offer exclusive discount
- Early Access to VIP events

Membership Program

	SILVER	GOLD	PLATINUM	DIAMOND
Starting Points	0	100	300	500
Exclusive Discount	✗	3%	5%	10%
Free Shipping	✗	> \$75	> \$75	> \$75
Experiences				
First Access to Events	✗	✗	✓	✓
Exclusive Events*	\$100/ticket	\$75/ticket	\$50/ticket	✓
Local Book Club	✓	✓	✓	✓
Referral				
Refer-A-Friend	\$5/person	\$5/person	\$5/person	\$5/person

Economic Impact

$$\text{Monetary Value} = 0.019704 * \text{Recency} + 39.123813 * \text{Frequency} - 24.897294$$

For **Top 15%** Segment,

R ↓ 30%, F ↑ 25%, Monetary Value will **increase by 25.68%**

For **Need Attention** Segment,

R ↓ 20%, F ↑ 18%, Monetary Value will **increase by 18.88%**

Overall, we can increase the remaining 85% group's sales by **12%**

PART 04

Next Step



Questions



Are we able to get cost-related data for the existing customers?



Besides paperback books, do we offer E-books?



Thank you

- ◆ Xianqian (Sarah) Yang
- ◆ Kejia Ran
- ◆ Minchuan (Miko) Liu
- ◆ Biao Zhang
- ◆ Yuanyang(Jerry) Chang
- ◆ Yue Sun

Appendix I - Membership Program

	SILVER	GOLD	PLATINUM	DIAMOND
Starting Points	0	100	300	500
Saving				
All Category Discount	×	×	×	10%
Specified Category Discount	×	3%	5%	×
Used Book Resale	×	3%	5%	3%
Free Shipping	×	> \$75	> \$75	> \$75
Experience				
Best Reader name/picture display on website	×	×	×	✓
Early Access - Shop at special sale prices —before everyone else!	×	×	✓	✓
Author Discussion/ Signature admission (2 times free ticket)	\$100	\$75	\$50	✓
Local reading group	✓	✓	✓	✓
Vote for the Annual Reader Choice Awards	✓	✓	✓	✓
Special Birthday Offer (Birthday Email with coupon)	✓	✓	✓	✓
Referral				
Refer-A-Friend(With > \$25 purchase) (Max referral limit: 10 / year)	\$5 each	\$5 each	\$5 each	\$5 each
Expiration Policy	Never expired			
Membership Expiration	Expire after half year for no corresponding amount of transaction			

Appendix II - Segmentation Ranking

Segment	Description	Segment%	r	f	m
Champions	Bought recently, buy often and spend the most	3.28%	5-6	5-6	5-6
Loyal Customers	Spend good money. Responsive to promotions	11.30%	4-6	4-6	4-6
New Customers	Bought more recently, but not often	20.94%	4-6	1-4	1-4
Promising	Recent shoppers, but haven't spent much	5.51%	4-5	1-4	1-2
Need Attention	Above average recency, frequency & monetary values	2.31%	3-5	3-5	2-3
About to Sleep	Below average recency, frequency & monetary values	22.51%	2-3	1-3	1-3
At Risk	Spent big money, purchased often but long time ago	11.72%	1-2	2-6	2-6
Can't Lose Them	Made big purchases and often, but long time ago	0.02%	1	5-6	5-6
Lost	Low spenders, low frequency, purchased long time ago	15.62%	1	1-4	1-4
Others	incorrect data OR enterprise level customer	6.78%			

Appendix III - Revenue Projection

Segment	Segment%	R ↓ by	F ↑ by	Revenue Increase %
Top 15%	14.58%	30%	25%	78.57%
New Customers	20.94%	25%	18%	57.33%
Need Attention	42.05%	20%	15%	43.75%
Can't Lose Them	15.64%	5%	5%	10.53%
Others	6.78%	3%	3%	6.19%

Appendix IV - Reference

PowerPoint Template:

https://docs.google.com/presentation/d/17eL3rL4SZxrDjqWSzM07-_Orcuaf66hPWyizzLAbugQ/edit?usp=sharing

RFM Method:

<https://www.r-bloggers.com/2019/07/customer-segmentation-using-rfm-analysis/>