

# A Novel Recommendation System: A Page Turner



## **Team DSC:**

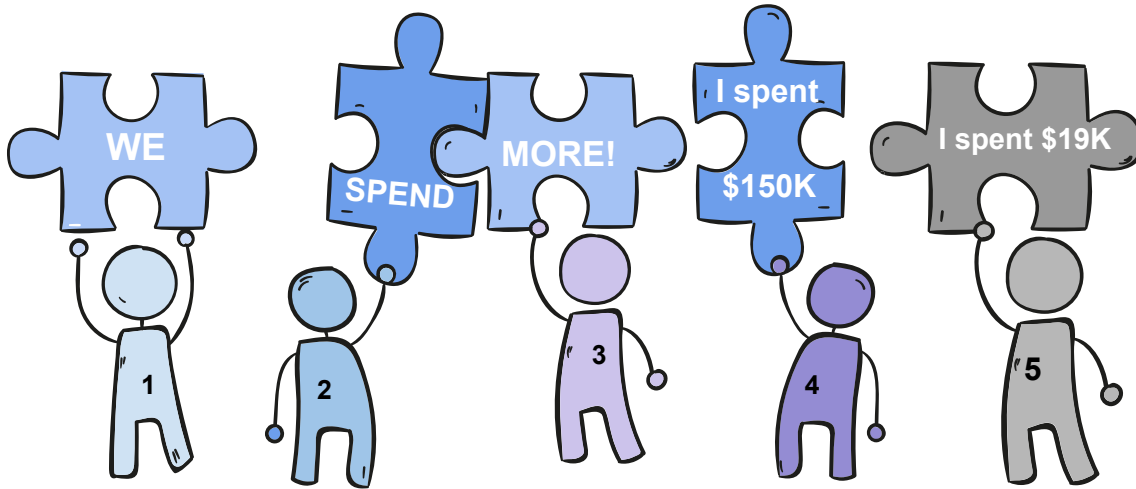
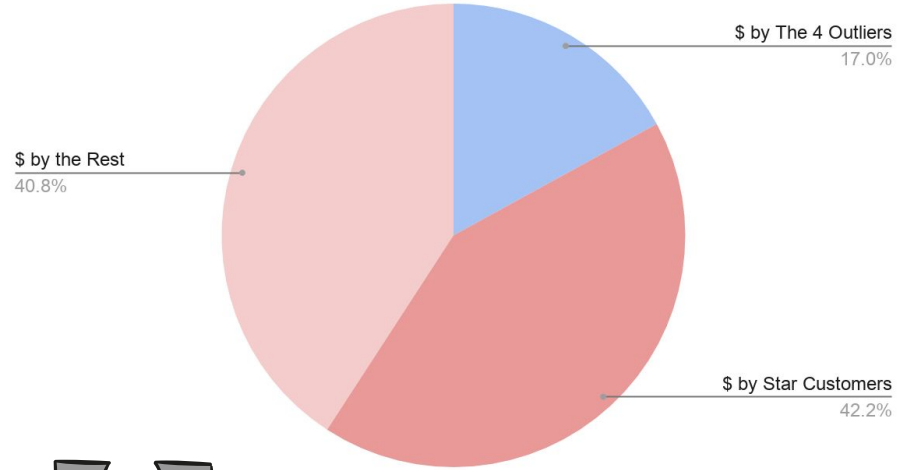
Jianghong (Nimo) Man, Zoila Joyo,  
Meenu Ravi, Xiaochun (Beca) Wang,  
Teresa Wang, Maggie Wu

# Executive Summary

- The goal of the project is to create **a recommendation engine** that improves the store's revenue. By testing the engine this month and implementing it the next month, we expect this engine to increase **\$100,000 annual revenue**
- The **top 15% customers** on the RFM score generate more than 50% of total revenue, so it is most efficient to focus on the them
- The RFM model is based on customers RFM score quantiles, through which we segment the top 15% customers into **4 groups**: The Best customers, the Big-Spenders, the Loyals, and the High Potentials
- By giving customized recommendations, the new system will boost both unit sales and net revenue

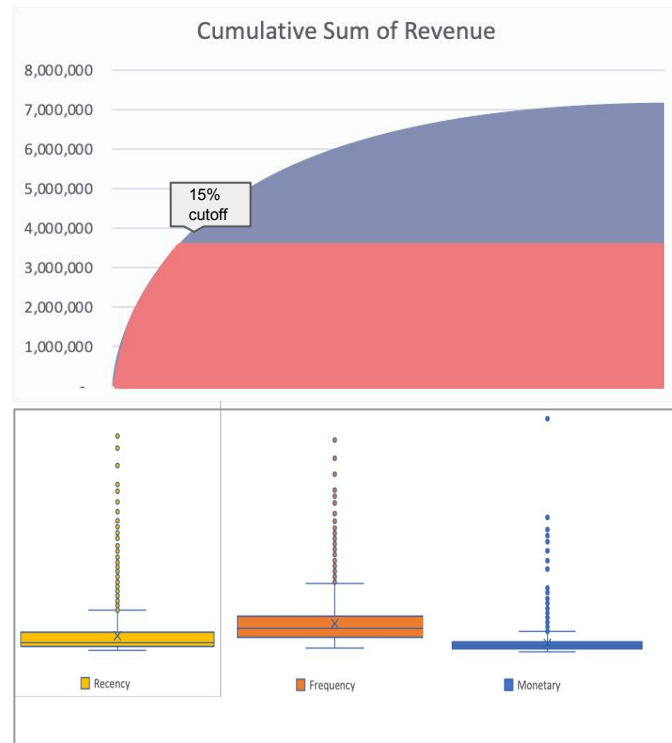
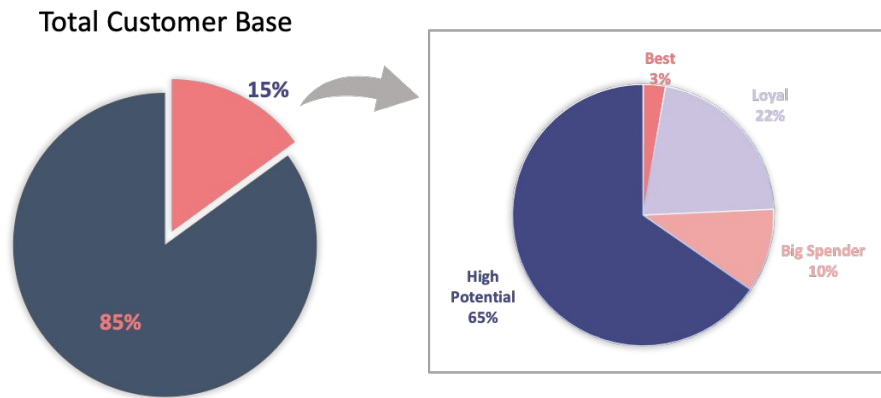
# The Outliers

- Generates 17% of total revenue
- 4th largest client spent almost 8 times more than the 5th.
- Potential b2b clients
- Designate outreach personnel to target their specific demands.








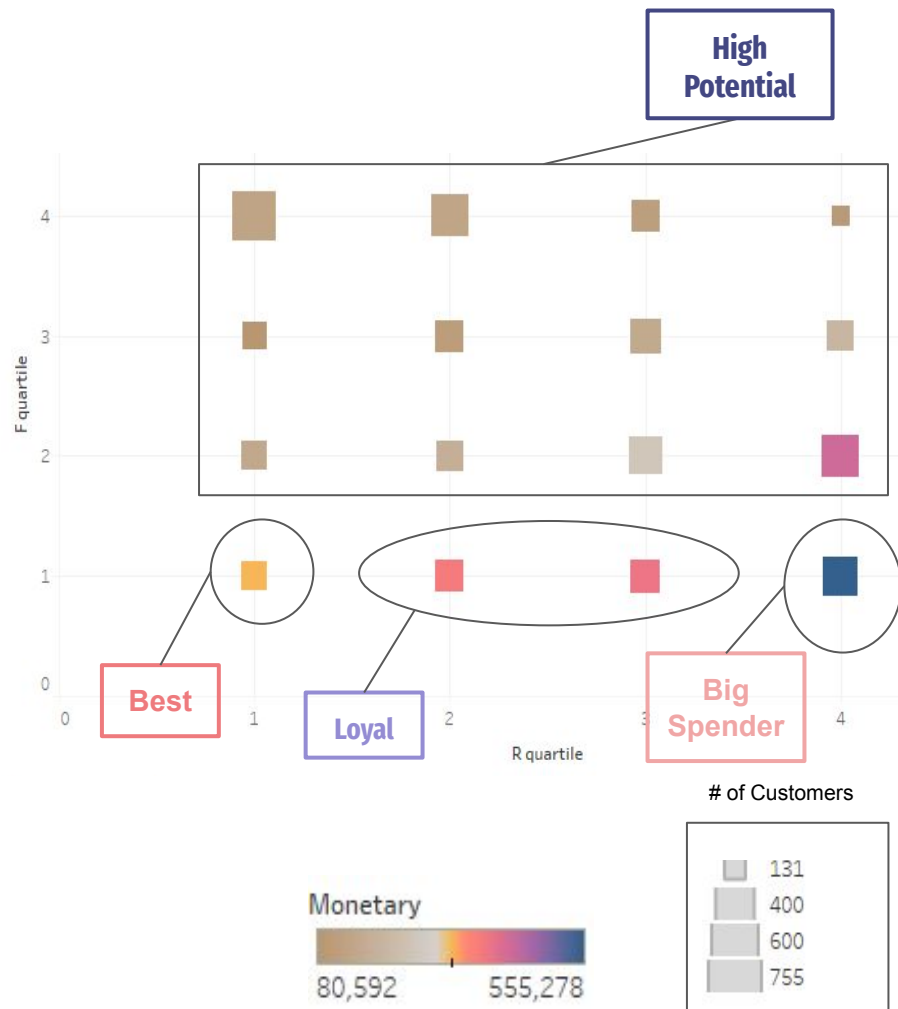
# 15/50 Split

- 15% of customers generated \$3,260,906, which is 50.8% of total revenue
- Top 15% customers give us a customer base of 5,000 ppl
- Recommendation system will focus on the top 15% customers



# The RFM Model

-  **1** Preprocessing: Remove outliers(4), frequency=0
-  **2** Calculate RFM Score:  $\log(F) * M/R$
-  **3** Order and slice **top 15%**
-  **4** Scale R/F/M scores: quantile
-  **5** Customer Segmentation: 3-digit score



# Customer Profiles by Segments

## Best Customers (R = 1, F = 1, M = 1)

### Don

- Recency = 10 days before
- Frequency = 38
- Monetary = \$1,757
- Time on file = 5yr + 8 mth

## Loyal (F = 1)

### Marco

- Recency = 3.5 mth ago
- Frequency = 30
- Monetary = \$1,154
- Time on file = 5yr + 8 mth

## Big Spenders (M = 1)

### Greg

- Recency = 4.5 mth ago
- Frequency = 15
- Monetary = \$1,208
- Time on file = 5yr + 3 mth

## High Potentials (others)

### Reshma

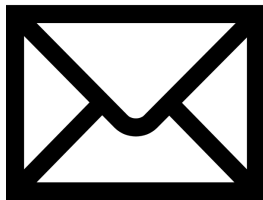
- Recency = 1.5 mth ago
- Frequency = 10
- Monetary = \$352
- Time on file = 4 yr + 1 mth



# Recommendation Strategies by Segments

## Best Customers (R = 1, F = 1, M = 1)

- Book Recommendation based on their previous preference
- **Weekly Communication**



## Big Spender (M = 1)

- Top 3 on their **monetary list**, constantly update the new info
- Monthly Communication

## Loyal (F = 1)

- Top 3 categories on their **frequency list**
- **Books on price reduction**
- Bi-weekly Communication



## High Potential (others)

- Best sellers
- **10% Discount for next purchase within 2 months**
- Bi-weekly Communication

# Data Preparation Summary



1

Calculate Mean RFM, and ToF Score for each segment



2

Calculate Average money each customer spent per purchase: **M/F**



3

Calculate how many times customers buy books per year:  
 **$12/(\text{tof}/\text{f}/30)$**

| Total 15%: 5002                        | Best    | Loyal   | Big Spender | High Potentials |
|--|---------|---------|-------------|-----------------|
| Mean (R)                               | 10.37   | 110.46  | 139.92      | 46.14           |
| Mean (F)                               | 37.64   | 30.16   | 14.56       | 10.41           |
| Mean (M)                               | 1757.2  | 1154.35 | 1207.87     | 351.64          |
| Mean(tof)                              | 2072.86 | 2072.86 | 1927.26     | 1504.56         |
| Avg \$ per purchase                    | 46.68   | 38.27   | 82.96       | 33.78           |
| Frequency customers buy books per year | 6.52    | 5.24    | 2.72        | 2.49            |



## Previous System:

Recommendation based on popular books

|       | Best Case Increased Revenue | Worst Case Increased Revenue |
|-------|-----------------------------|------------------------------|
| Total | \$50,444                    | \$37,833                     |

- **Method:** Rank book category based on the highest frequency, select the top 5 ones and recommend them to all 5002 customers.

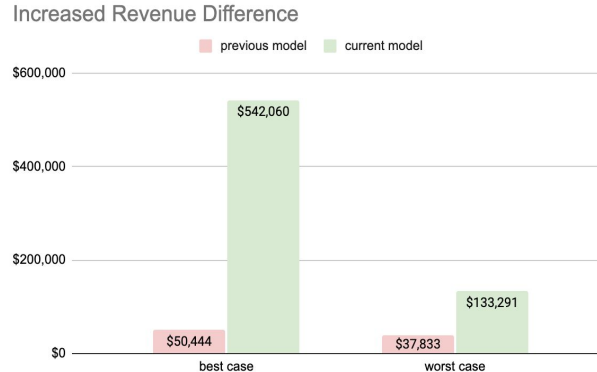
- **Disadvantage:**
  - Customers are not segmented, not all those would like to buy the recommendations.
  - Assume 20% will buy in the best case and 15% will buy in the worst case scenario



Difference by:

Best case: **\$491,615**

Worst case: **\$95,457**



## Our System:

Recommendation based on segmentation

|                | Best Case Increased Revenue | Worst Case Increased Revenue |
|----------------|-----------------------------|------------------------------|
| Best           | \$40,682                    | \$9,993                      |
| Big Spender    | \$38,655                    | \$55,382                     |
| Loyal          | \$326,573                   | \$14,445                     |
| High Potential | \$136,148                   | \$53,469                     |
| Total          | <b>\$542,059</b>            | <b>\$133,290</b>             |

- **Advantages**
  - 15% selected customers are selected to different categories and have different recommendations.

# Loyalty Program Proposal

2022



## Tailored Rewards

Customers earn 10 points for every \$1 they spend online. Then, they can put these points towards future purchases.

## Omni-Channel Communication

Using Email Marketing, we send a weekly newsletter and personalized messaging.

## Charity Program

Customers can choose to donate their change to **Teach For America** (NPO that confronts educational inequity).

### Checkout

PayPal

Hi, Zoila  
\$147 Until Next Reward!

#### Rewards & Offers (1)

You have 1,035 of 2,500 points. Earn up to 900 when this order ships!  
[See Details](#)

\$10 \$10 Rewards Cash\*  
Exp. 01/14/2021 [Exclusions apply](#)

Redeem



Welcome to Book Riot's Events Newsletter, hosted by me, María Cristina. We're looking ahead at some of the bookish ways you can spend your time in the next couple weeks, and I'm sure there's at least one item here that can tempt you to put on pants and go out (in that order, please). Clear your calendars on the following dates, my reading friends.

|  |                 |
|--|-----------------|
| <input type="checkbox"/> Round up for charity (optional) ⓘ | \$0.96          |
| Taxes  | \$4.20          |
| Total cost   | <b>\$68.04</b>  |
| Total savings  | <b>\$175.81</b> |
| <a href="#">Proceed to Checkout &gt;</a>                   |                 |

# Thank You!



# Appendix A: Operation Strategies

- **Privacy:** Recommend books in 3 categories instead of 1
  - Make the system **not too “smart”**
  - **More options** provided to customers



- **Website design:**
  - 1 - click buying option
  - Easy to navigate
  - AB Testing



# Appendix B: 15% data (first 20 rows displaying)

| id    | Recency  | Frequency | Monetary | Time on File (Days) | rfmScore | R_quartile  | F_quartile | M_quartile | Tof_quartile | RFM_Segment |
|-------|----------|-----------|----------|---------------------|----------|-------------|------------|------------|--------------|-------------|
| 11115 | 7219121  | 6         | 90       | 8032.539062         | 2057     | 6024.149492 | 1          | 1          | 1            | 111         |
| 21167 | 13729756 | 16        | 118      | 18582.531250        | 648      | 5540.712257 | 1          | 1          | 1            | 111         |
| 25960 | 2294222  | 6         | 65       | 7496.742188         | 2449     | 5215.717526 | 1          | 1          | 1            | 111         |
| 12923 | 8017166  | 4         | 25       | 4527.640625         | 1738     | 3643.478238 | 1          | 1          | 1            | 111         |
| 13649 | 8146691  | 2         | 26       | 2165.271484         | 1564     | 3527.331764 | 1          | 1          | 1            | 111         |
| 10034 | 7022603  | 6         | 43       | 4569.496094         | 2393     | 2864.464873 | 1          | 1          | 1            | 111         |
| 5840  | 4238567  | 2         | 68       | 1319.323242         | 2447     | 2783.447293 | 1          | 1          | 1            | 111         |
| 2575  | 2919338  | 2         | 54       | 1312.811523         | 2456     | 2618.392111 | 1          | 1          | 1            | 111         |
| 25577 | 137111   | 2         | 33       | 1315.105469         | 2370     | 2299.138108 | 1          | 1          | 1            | 111         |
| 6419  | 4412834  | 4         | 25       | 2681.558594         | 2456     | 2157.901033 | 1          | 1          | 1            | 111         |
| 8184  | 5411734  | 4         | 15       | 3057.689453         | 2056     | 2070.094135 | 1          | 2          | 1            | 121         |
| 703   | 227390   | 4         | 74       | 1693.238281         | 2448     | 1821.951945 | 1          | 1          | 1            | 111         |
| 28173 | 5473330  | 6         | 30       | 3165.509766         | 2419     | 1794.420588 | 1          | 1          | 1            | 111         |
| 6518  | 4429141  | 9         | 61       | 3747.314453         | 2460     | 1711.637450 | 1          | 1          | 1            | 111         |
| 9084  | 5919274  | 2         | 24       | 1067.684570         | 1634     | 1696.579519 | 1          | 1          | 1            | 111         |
| 31647 | 12031941 | 1         | 17       | 593.389648          | 709      | 1681.199470 | 1          | 2          | 2            | 122         |
| 9711  | 6946887  | 5         | 73       | 1951.649414         | 2346     | 1674.694531 | 1          | 1          | 1            | 111         |
| 9156  | 5952220  | 2         | 16       | 1199.541016         | 1763     | 1662.916946 | 1          | 2          | 1            | 121         |
| 1083  | 611948   | 3         | 35       | 1359.800781         | 2378     | 1611.521691 | 1          | 1          | 1            | 111         |
| 9373  | 6245080  | 5         | 77       | 1777.997070         | 2434     | 1544.654663 | 1          | 1          | 1            | 111         |

C:\Users\meenu\OneDrive\Consulting\Project>

Link to code folder with .xlsx file and .py file: [code](#)

# Appendix C: Data Summary

Total customers: 33351 spent

15% of the top rfm scored (using formula) customers: 5002

Total number of customers in “Best” (R=1, F=1, M=1): 138

Total number of customers in “Loyal” (F=1): 1077

Total number of customers in “Big Spenders” (M=1): 518

Total number of “others” and “New”: 3269->This is people who do not have 111, X1X,XX1 + people with 1 as tof quantile

|                 |  |
|-----------------|--|
| total customers | 5002   |
| f               | how many times the customer buy books in total |
| m               | money spent in total                           |
| tof             | time on file in days                           |

|                  | # customers | mean r | mean f | mean m     | Mean(tof) | Avg \$ per purchase (M/F) | Every x month customer make purchase = tof/F/30 | Frequency customers buy books per year = 12/Every x month customer buy books |
|------------------|-------------|--------|--------|------------|-----------|---------------------------|---|--|
| the best         | 138         | 10.37  | 37.64  | \$1,757.20 | 2072.86   | \$46.68                   | 1.84  | 6.54   |
| the big spenders | 518         | 139.92 | 14.56  | \$1,207.87 | 1927.26   | \$82.96                   | 4.41  | 2.72   |
| the loyal        | 1077        | 110.46 | 30.16  | \$1,154.35 | 2072.86   | \$38.27                   | 2.29  | 5.24   |
| high potentials  | 3269        | 46.14  | 10.41  | \$351.64   | 1504.56   | \$33.78                   | 4.82  | 2.49   |

## Appendix D: Financial Assumption

[illegible]