



# RECOMMENDER ENGINE

---

For Greg's Book Store

Vishal Parameswaran

# Introduction

Greg's Book Store is an historical bookstore, in the heart of Chicago. Established in 1860, when Bartholomew Greg and his brothers turned their voracious appetite for books into a thriving business. They currently have 5 bookstores around Chicago and are the primary book distributors for University of Chicago. Due to popular demand from local customers and due to the pandemic, they decided to setup their distribution network over the internet and have developed an Online Business Store. Their online presence has boosted their sales by a considerable margin, and they were able to spread their name across the country. Unfortunately, one of the biggest complaints they have been receiving is the lack of a recommender. Previously, their loyal customers remember, coming to the store and being recommended books by the staff. These recommendations were cherished by the loyal customers because it was an indicator of the bond the customers had with the store and its staff. Customers were also quick to point out that the staff would give loyal customers discounts for frequent purchases, and there were usually a lot of perks afforded to being frequent customers. Unfortunately, the current system is unable to identify loyal customers and unable to give suggestions on books. The owner, on receiving these complaints, contacted us to make a recommender system keeping the above points in mind. The CMO and the webmaster do not believe that a recommender system is needed.

Looking at the competition, paints a completely different picture. Big Wigs in the industry, like Amazon, Barnes & Noble and Higgin Botham's invest heavily in their recommender systems. They combined cover over 90% of the online Book Store sales.



# SMART Framework

**Business Problem:** The consulting team is to design and build a Recommendation system that uses machine learning to categorize customers based on their past shopping patterns and accordingly market products to them to increase sales and loyalty to the ecommerce website.

S

Implement a recommender system that will

- 1) Increase Sales
- 2) Increase Customer Retention
- 3) Decrease cart abandonment rate.

M

- 1) Increase overall sales over a two month time period by 40%
- 2) Lower Cart Abandonment rate by 80%
- 3) Identify Loyal Customers

A

- 1) Perform an Exploratory Data Analysis on the data provided.
- 2) Create an RFM and Recommender model
- 3) Create Personalized Suggestions based on Customer Segmentations

R

Competitors, invest heavily in recommendation algorithms, that are able to track multiple aspects of the customer on the website. This has resulted in customers being spoiled for choice, and are dependent on such algorithms.

T

There is a hard deadline: 50 Days from this presentation.

# The Team

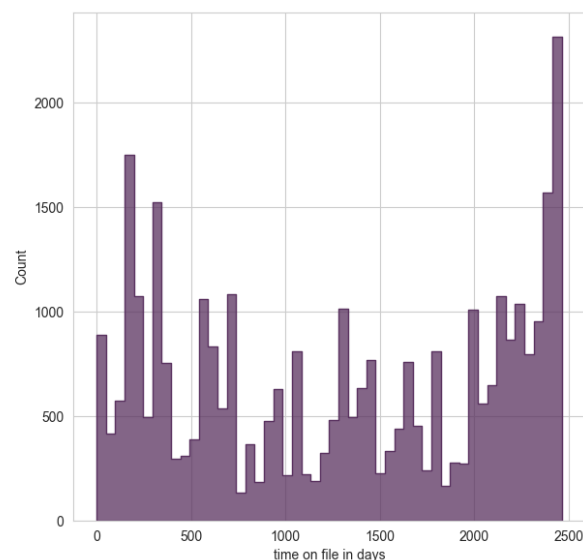
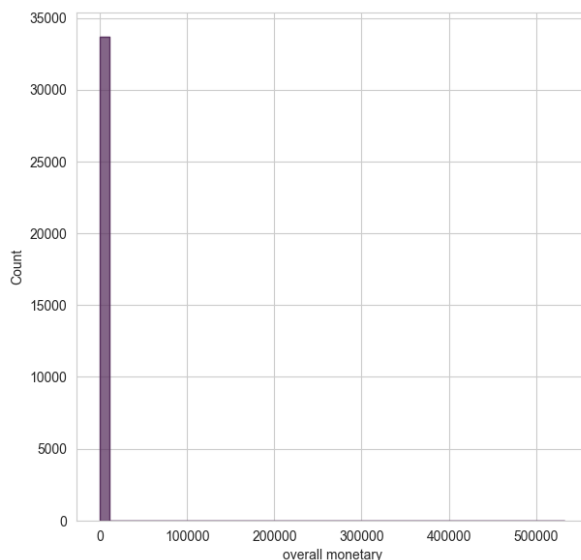
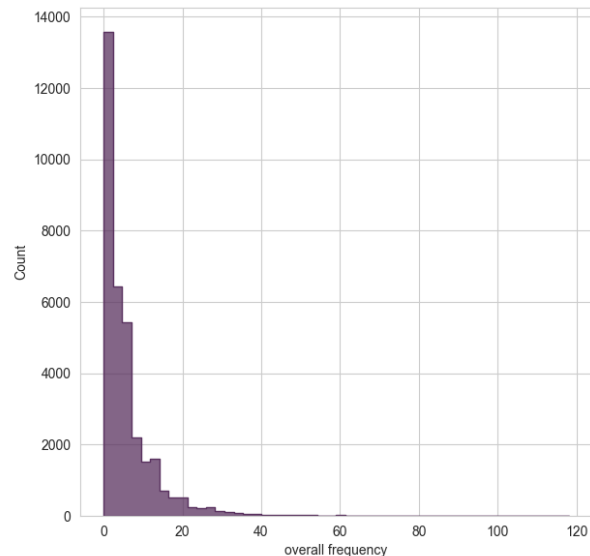
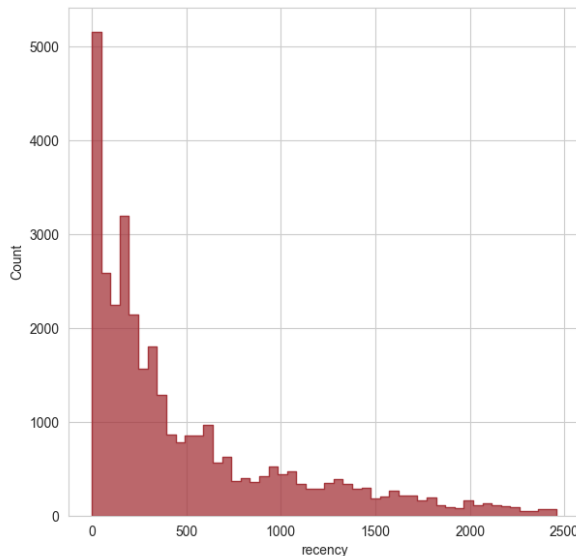
The team for this project consists of

- VP Marketing: Our lead for this project. The Shot Caller
- Customer Satisfaction Manager: As our customer, Our CSM will always be present for such meetings and is the PoC for the client.
- SVP Data Scientist: The Senior most data scientist, responsible for the development of the system.
- Data Visualization Specialist: As a visualization specialist, they will be able to effectively communicate the intent of the data.
- Marketing Analytics Manager: Conduct research on current and future marketing campaigns and check their effectiveness. Suggest future marketing campaigns based on learnings from previous ones.
- Performance Marketing SEM/SEO Analyst: Rank the current website, and offer suggestions on increasing their online visibility and effectiveness

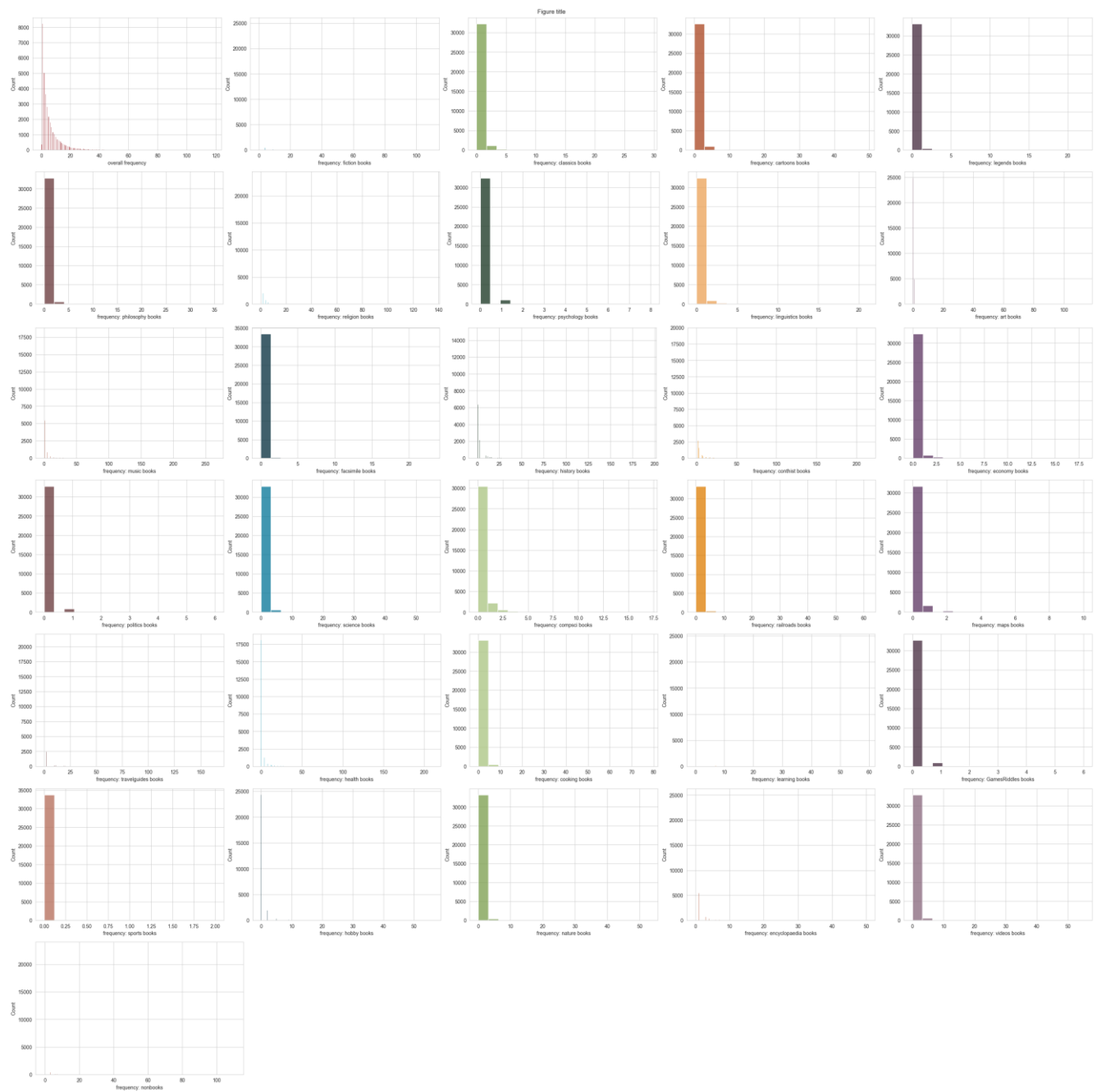
# Methodology

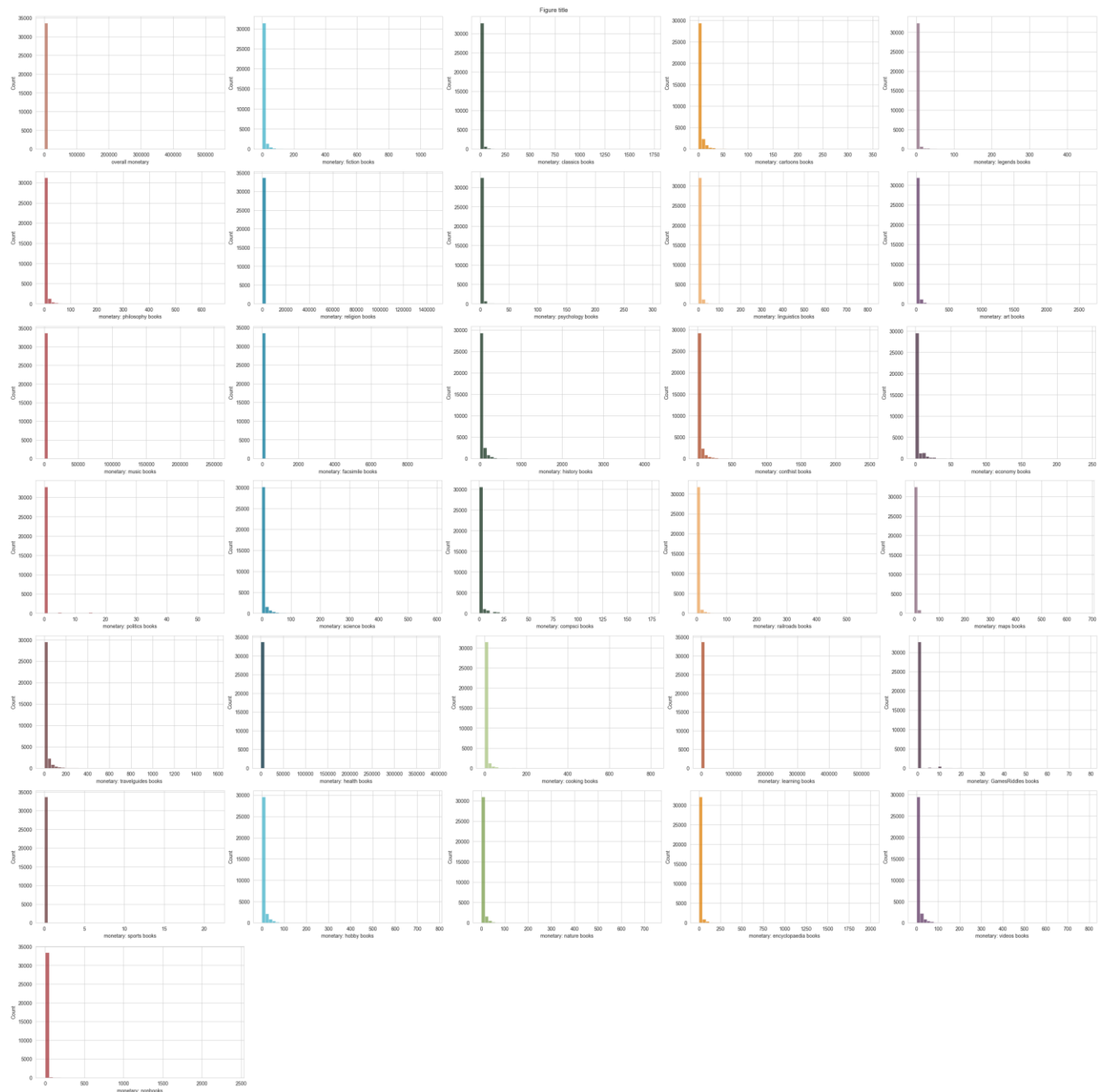
The Predictive Estimator Process, can be divided into 7 stages

1. **Data Ingestion:** Using Pandas (a free open-source python library used for data ingestion and data manipulation) the data was extracted from the provided input file as a Data Frame.
2. **Data Cleaning:** A few steps were involved in data cleaning
  - a. Dropped column 'logtarg' as it was majorly null values
3. **Data Preprocessing:** As all the data was numerical, there was no preprocessing required. The columns were renamed to be more understandable
4. **Data Visualizations:**
  - a. The relationships between the overall columns were shown, as well as the distributions for the overall variables.



b. The right skew of the Frequency and Monetary variables were also visualized.





5. Model Creation: The RFM model was created to rank the recency, frequency and the monetary values.
  - a. The recency was ranked on a scale of 3, with the quantiles at 10% and 50%. It was labeled as ["Recent", "Nudge", "Old"]
  - b. The frequency was ranked on a scale of 4, with the quantiles at 10%, 50% and 90%. It was labeled as ["low", "Average", "Promising", "Loyal"]
  - c. The monetary was ranked on a scale of 3, with the quantiles at 20% and 80%. It was labeled as ["Cheap", "Moderate", "Whale"]
  - d. To make it easier for categorization, the Labels, for R, F, M were concatenated.
    - i. This resulted in almost 45 different categorizations,

- ii. To reduce it further few more categories were made to encompass other categories

R	F	M	Segments
1	4	3	Star Customer
3	1, 2, 3, 4	3	Needs Urgent Attention
2	3,4	3	At Risk
1	3,4	2, 3	Potentially Big
1	2, 3, 4	2, 3	Potentially Loyal
1, 2	1, 2	1	New Customer
1, 2	1, 2, 3, 4	1, 2	Promising
2	1, 2, 3	2, 3	Needs Big Nudge
3	1, 2	1	Not Worth It
3	1, 2, 3, 4	2	Requires Activation



# Insights

With the help of an RFM Model, the initial challenge of identifying customer segments is completed. For further recommendations, a more detailed dataset, including previous purchases, and metadata of purchases would be required. Using the provide data, the RFM model will help the marketing team direct advertisements towards segments and will also afford them the ability to track the progress of the customer.

## FURTHER RECOMMENDATIONS

- Utilize the most popular recommender in the main page, to showcase recent releases and most popular books.
- Tracking customers more intently (While following the various Data Protection Laws) could help with creating more robust Machine Learning solutions.
- Based on the above given Segments, we can tune our marketing strategy.

Segments	Market Strategy
Star Customer	Highest tier in loyalty program
Needs Urgent Attention	Frequently send out promotions to entice and hook them
At Risk	Contact over call, or more promotions can help. Loyalty trials are also a potential avenue
Potentially Big	Offer huge incentives, like promotions and loyalty deals
Potentially Loyal	Personalized ads, to show you care.
New Customer	Standard promotions, wait for either promotion or regression in segments
Promising	Like potentially big, but at lesser frequency
Needs Big Nudge	Huge promotional offers
Not Worth It	Look at the customer likes and dislikes and check if customer has moved on to a competitor. If so, what attracted them?
Requires Activation	Like at risk.