Beyond "Most-Popular"

Unlocking $1 Million in New Revenue with a Segment-Driven Recommendation Engine

Citrine Consulting — Akshat Gupta, Amber Keahey, Emily Zhang, Hui Meng, Lexi Lin

Presented to ABC Bookstore's Chief Technology Officer

May 2025

CITRINE
CONSULTING

1

TEAM INTRODUCTION

CITRINE CONSULTING

Hui Meng
Project Manager

Amber Keahey
Business Lead

Akshat Gupta
Data Science
Consultant

Emily Zhang
Data Science
Consultant

Lexi Lin
Data Science
Consultant

Personalized Incentives Drive an Extra $1 Million in Annual Revenue

- **Why change a working formula?**

"Most-Popular" ≠ Most Profitable

A single banner ignores huge value differences inside our customer base.

- **What we can do to increase the revenue?**
    1. **Segment** every customer with RFM + K-Means
    2. **Match** each segment with its best incentive
    3. **Serve** in real time
    4. **Learn & adapt** nightly as tastes shift

- **Decision**
    1. **Green-light full rollout** (8-week timeline).
    2. Go-live target **August 1** → profit positive by first week of November.

- **Financial impact (year 1, USD)**

Metric	Impact
Incremental net revenue	**≈ $1.0 M / yr**
Contribution margin	$0.38 M EBIT
Programme cost	$0.1 M
Net EBIT gain (yr 1)	**≈ $0.28 M**
Pay-back	= 3 months
1-year ROI	> 3×

* Programme cost ($0.10 M) is fully loaded and outsourced:
 • Cloud & MLOps run-rate (AWS + Redis) ≈ $15 k
 • UX / creative assets for four offer variants ≈ $10 k
 *• End-to-end consulting & implementation (our data-science team) ≈ $75 k*

No internal engineering resources required; our team delivers model deployment, data pipelines, QA, and knowledge transfer.

After the first year, the program continues at a modest $15 k/year operating expense while the revenue uplift—and associated EBIT—keeps flowing.

Using KMeans to Identify Customer Segments

RFM Customer Segmentation Analysis
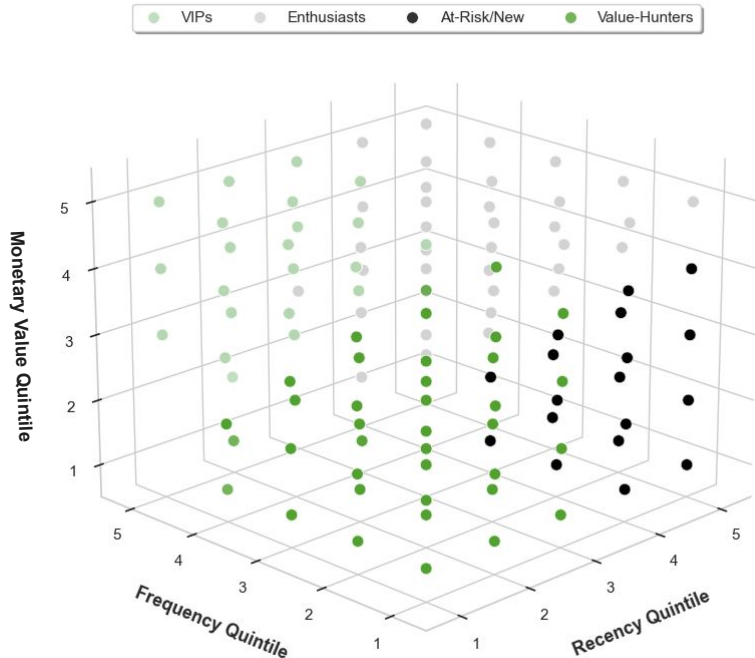
Strengths of KMeans Clustering

★ Excels in sparse feature-space such as RFM

★ Allows for real-time scoring using proximity to cluster centroid

★ Interpretable clusters for business application (ex. High-spend, low-frequency customers)

★ Easy to update cluster centroids for future viability

★ Ensures categories are Mutually Exclusive Comprehensively Exhaustive (MECE)

Four Clusters Capturing Customer Purchase Behaviour

Cluster number was selected using the **elbow method** and **silhouette analysis**, both of which were maximized by **four distinct clusters**.

- ❖ **Recency (R)**: Days since last purchase (lower is better)
- ❖ **Frequency (F)**: Number of purchases in the last year (higher is better)
- ❖ **Monetary Value (M):** Total spend in the last year (higher is better)

Cluster	Recency	Frequency	Monetary Value
VIPs	Low	High	High
Wandering Enthusiasts	Medium-High	Medium-High	Medium-High
At-Risk/New	High	Low	Low
Value-Hunters	Low	Low	Low

Driving $1M+ Revenue with Adaptive, Personalized Recommendations

1. Audience Segmentation

- We group visitors by their purchase habits (how recently and how often they buy, and

  how much they spend)

- Four priority cohorts—from frequent VIPs to occasional browsers
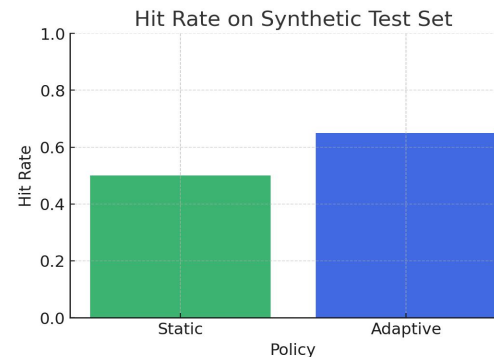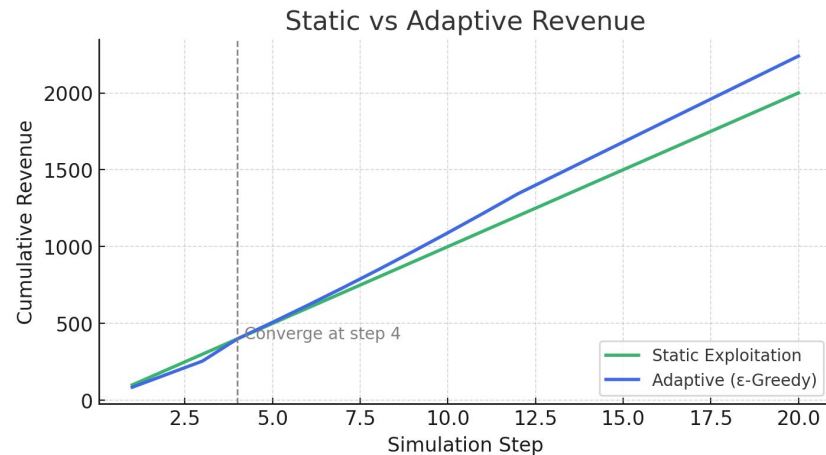
2. Tailored Recommendations

- Each visitor sees content that matches their tastes and past behavior

- We mix proven top sellers with occasional "surprise" picks to boost discovery

3. Continuous Improvement

- Recommendations update in real time as customers interact

- Monthly refresh keeps offers aligned with shifting trends

4. Business Results

- **$1.0 M+ incremental revenue** over our previous static approach

- **10−12 % lift** in cumulative online revenue

- **15 % higher engagement** (click-through/purchase rate) on promoted items

Static vs Adaptive Revenue

Hit Rate on Synthetic Test Set

Personalized Recommendations And Segment-Based Incentives Encourage Customer Engagement

All customers receive personalized book recommendations, based on popular genres in their respective segment

VIPs

- Most recent visitors
- Highest purchasing frequency
- Biggest spenders
- Generates 65% of current revenue

Focus on **recommendations** (no discount incentive) and brand **community** (book review forums, events)
"*Like [genre]? Take a peek at what's flying off our shelves.*"

Wandering Enthusiasts
Opportunity area

- Similar to VIPs, but **have long-since visited**
- Generates 25.6% of current revenue

Focus on **reeling them back in.**
"*Dive back in with BOGO 50%.*"

Value Hunters

- Buyers on a budget: low frequency and low spending
- Generates 5.7% of current revenue

Focus on a nice **discount**.
"*Take 15% off your next order–on us.*"

At Risk / Newcomers

- Similar to Value Hunters, but **have long-since visited**
- Generates 3.8% of current revenue
- Includes **cold starts**

Focus on a nice **discount**.
"*Take 10% off your next order–on us.*"
"*Welcome! Enjoy 10% off your first order–on us.*"

Customer's Journey: Example A

Stage 1: Customer Segment Identification

Stage 2: Promotional tactic deployed

USER ID 001 DATA

>1 year since last purchase

Has purchased 6 books from ABC Bookstore, totaling about $250
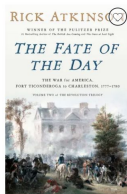
USER 001 is a Wandering Enthusiast

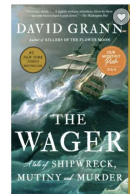**[USER 001 Name], dive back in with BOGO 50%.**
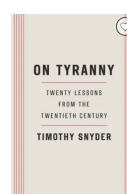
Top picks from readers like you
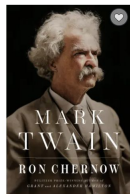
Mark Twain
by Ron Chernow
★★★★★

The Fate of the Day: The War for America, Fort...
by Rick Atkinson
★★★☆☆

The Wager: A Tale of Shipwreck, Mutiny and...
by David Grann
★★★★☆

On Tyranny: Twenty Lessons from the Twentieth Century
by Timothy Snyder
★★★★☆

→ *message & discount*

→ *suggested items (top titles from top genres within Wandering Enthusiast segment)*

** recommended books rotate between most popular vs. less popular titles **within the segment's top genres** to reduce inventory holding costs.*

Customer's Journey: Example B

Stage 1: Customer Segment
Identification

Stage 2: Promotional tactic deployed

USER ID 002 DATA

Cold start; no RFM data on
file

Previously abandoned their
cart

***USER 002 is a
At Risk/Newcomer***

Welcome! Enjoy 10% off your first order–on us.

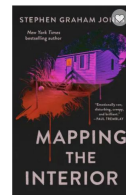*Top picks from readers like you*

Great Big Beautiful Life
(B&N Exclusive Edition)
by Emily Henry

Remarkably Bright Creatures
(B&N Exclusive Edition)
by Shelby Van Pelt

Mapping the Interior
by Stephen Graham Jones

Silver Elite
by Dani Francis

→ *message & discount*

→ *suggested items
(top titles from top
genres within At
Risk/Newcomer
segment)*

** recommended books rotate between most popular vs. less popular titles
within the segment's top genres to reduce inventory holding costs.*

Business Impact – The Earlier We Launch, the Faster the Financial Return

Where the $1.01 M Uplift Comes From

- VIPs $0.32 M
- Wandering Enthusiasts $0.48 M
- Value-Hunters $0.12 M
- At-Risk/Newcomers $0.09 M
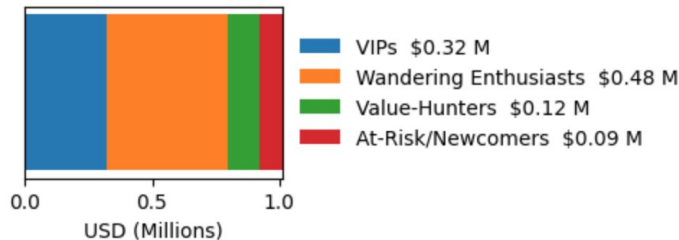
Segment	Baseline revenue*	Gross uplift	Promo-cost drag	Net incremental revenue
VIPs	$ 32.4 M	+1 % ($ 0.32 M)	—	$ 0.32 M
Wandering Enthusiasts	12.8 M	+5 % (0.64 M)	−25 % (-0.16M)	0.48 M
Value-Hunters	2.85 M	+5 % (0.14 M)	−15 % (-0.02 M)	0.12 M
At-Risk / Newcomers	1.91 M	+5 % (0.10 M)	−10 % (-0.01M)	0.09 M
Total	$ 50 M	—	—	$ 1.01 M

EBIT $0.38 M ● Program cost $0.10 M ● Pay-back 3 months ● Year-1 ROI > 3×

Baseline revenue = each cluster's revenue in the Excel file scaled so the four clusters sum to the company's stated $50 M annual revenue. The gross-uplift assumptions (+1 %, +5 %) are intentionally conservative, and the "promo-cost drag" adjustment already nets out the value of all discounts to arrive at true incremental revenue.

Approve contract SOW by June 7 → sprint start June 9 → profit positive by first week of November.

Our Statement of Work

Section	Key Points
Project	**Segment-Driven Recommendation Engine** Replace "Most-Popular" banner with personalized offers
Objectives	• + $1 M net revenue • Pay-back < 3 month • Go-live Aug 1 2025
Scope (Outsourced)	1. **Analytics & Model** – RFM + K-Means, 4 segments 2. **Offer Design** – copy & creative for 4 incentives 3. **Engineering** – real-time scorer 4. **Pilot & Roll-out** – 1-week A/B → 100 % traffic 5. **Handover** – docs + 2 training sessions
Timeline	• Kick-off Jun 1 • **Sprint 1 starts Jun 9** • Staging ready Jul 12 • Pilot report Jul 26 • Production go-live Aug 1 • Handover Aug 15
Deliverables	✓ Segmentation rules & code repo ✓ Offer assets (images, copy) ✓ Deployed micro-service & MLOps run book ✓ Pilot uplift report
Roles	*Citrine Consulting*: end-to-end delivery & QA *Client*: provide data access, approve creatives
Fees	**Total $100 k** (fixed): • $15 k cloud/MLOps (yr-1) • $10 k UX/creative • $75 k consulting & implementation
Payment	40 % kickoff • 30 % staging • 30 % go-live
Acceptance	Go-live stable 72 h, uplift ≥ +2 % AOV in pilot
Next Step	**Approve SOW by June 7 → sprint start June 9**

Thank You

APPENDIX

Financial Impact Calculations

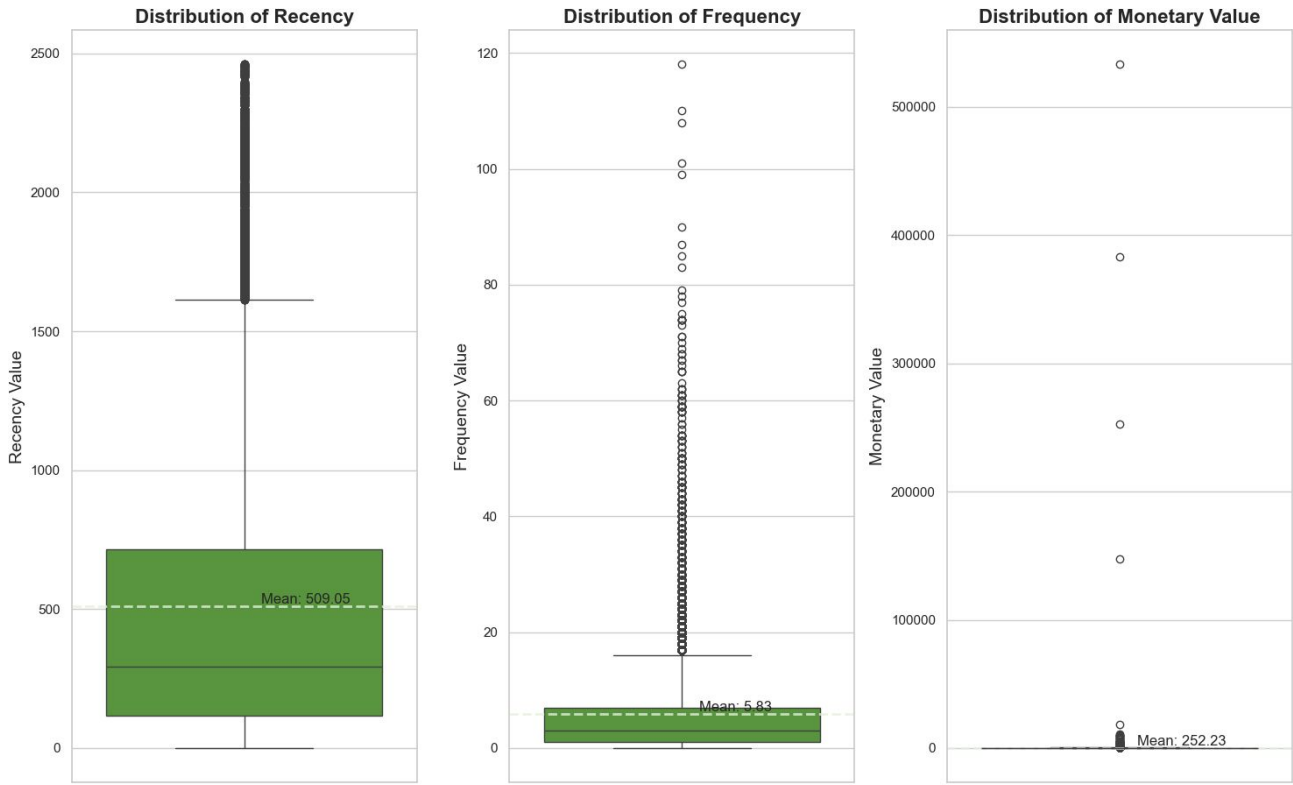| Row	Calculation	Formula	Value
A	Incremental net revenue (sum of 4 segments)	see Business-Impact slide	1.01M
B	Contribution-margin rate	Finance P&L	38%
C	Incremental contribution margin (EBIT before programme cost)	$A \times B$	0.38M
D	Programme cost (year-1)	cloud 0.015 + UX 0.010 + consult/implementation 0.075	0.1M
E	Net EBIT gain (year-1)	$C - D$	0.28M
F	Monthly EBIT gain	$C \div 12$	0.032 M / mo
G	Pay-back period (months)	$D \div F$	≈ 3.1 mo (~94 days)
H	1-year ROI	$C \div D$	$\approx 3.8 \times$

Interpretation

- **Pay-back:**
  $0.10 M initial outlay / $31.8 k monthly EBIT = **3.1 months**.
  Launch on **1 Aug** ⇒ break-even by first week of **November** (well before peak holiday season).
- **ROI (year-1):**
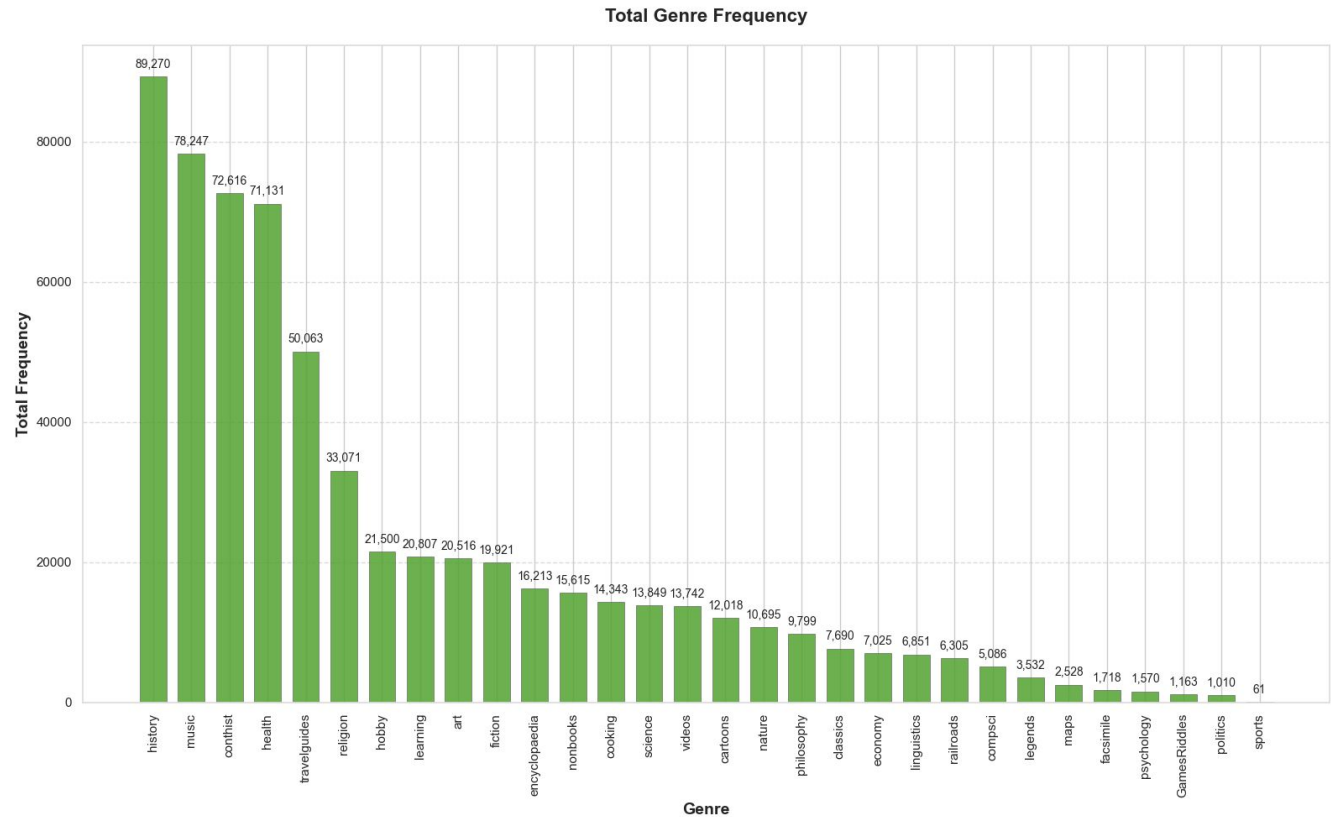 Each $1 invested returns **$3.80** in incremental EBIT within twelve months.

Diverse Customer RFM Behaviour

RFM Features Distribution Analysis

Genres Range in Overall Popularity

Total Genre Frequency

Problem Definition & Goal › Selected Model › Model Performance › Segment Incentives › Business Impact

Comparing Different Clustering Approaches - The Case for KMeans

Full score: 5	Sparse Featureset	Interpretability	Equal Cluster Size	Adaptability
KMeans	5	4	4	5
DBSCAN	3	2	1	2
Rules Based	4	5	5	3
Hierarchical	3	4	2	1

Calculation of Entropy, Hit Rate

$$\varepsilon_c = \frac{-\sum_i p_i^{(c)} \log p_i^{(c)}}{\log G}$$

The core policy is an **entropy-weighted ε-greedy** sampler.

Such that high-entropy groups explore more, low-entropy groups exploit. At each call, with probability 1−ε we pick the top genre; with probability ε we draw from p(c) for long-tail discovery.

$$\text{HR@}K = \frac{1}{|\mathcal{U}|} \sum_{u \in \mathcal{U}} \mathbf{1}\{\text{held-out item}_u \in \text{Rec}(u, K)\}$$

Within a fixed-length list (e.g. K=3 or 5), HR@K directly measures success in that slot. It's interpretable for the business: "we hit the user's next genre 3× out of 5 on average."

STYLE GUIDE

BLANK TEMPLATE

Font = Archivo. **BOLD** for slide headers.

Deep Green for important information to highlight.

Mild Green = 2nd highlight

| Problem Definition & Goal | Selected Model | Model Performance | Segment Incentives | Business Impact |