#### Business Intelligence | Batch 13

15 March 2025





# Leveraging RFM & CLV for a Data-Driven Marketing Strategy at All-U-Need Mart

Hijir Della Wirasti





## **Essential Links**



Dataset	https://www.kaggle.com/datasets/ahmedmohamed2003/retail-store-sales-dirty-for-data-cleaning/data
Tableau	https://public.tableau.com/app/profile/hijir.della.wirasti5486/viz/OptimizingRetailGrowthRFMAnalysisCLVPredictionMarketingStrategy/Dashboard1
GitHub	https://github.com/hijirdella/All-U-Need-Mart_RFM-CLV-Marketing-Strategy
Portfolio	https://www.hijirdata.com/
LinkedIn	https://www.linkedin.com/in/hijirdella/
Email	hijirdw@gmail.com
Drive	https://drive.google.com/drive/folders/1kd8P2A 9RJrKmZeOWTdunVPFcJv3y 0x9?usp=sharing



**Hijir Della Wirasti**Business Intelligence





# **Table of contents**



01

Background & Data Processing

02

Quick Analysis

03

Data Visualization



Business Insight & Recommendation

**Business Intelligence | Batch 13** 15 March 2025



# **Background**











All-U-Need Mart struggles to optimize its marketing strategy due to a lack of insights into customer behavior, discount effectiveness, and long-term profitability. Without data-driven decisions, marketing efforts risk inefficiency, high customer churn, and poor personalization.



#### Goal:

How can Sales Analysis, RFM Segmentation, and CLV enhance customer retention, profitability, and marketing effectiveness?



## **Literature Review**

The implementation of data-driven strategies such as RFM Analysis (Recency, Frequency, Monetary), and Customer Lifetime Value (CLV) can significantly impact sales and customer loyalty at All-U-Need Mart. Below are some data and figures supporting the effectiveness of this approach:

#### Increase in Customer Loyalty and Revenue

- Eastwood: Achieved a 21% increase in email marketing profits after implementing RFM analysis.
- L'Occitane: Experienced a 25-fold increase in revenue per email through customer segmentation using the RFM model.
- Frederick's of Hollywood: Recorded a 6-9% conversion rate in their campaigns after adopting RFM analysis. (Desai et al. 2021)

#### **Reduction in Churn Rate**

- Companies utilizing churn prediction can reduce churn rates by up to 30% with more targeted strategies.
- By identifying at-risk customers early, businesses can implement retention strategies such as loyalty programs and exclusive discounts (Yulianti, 2018).

#### **Marketing Cost Efficiency**

- With **better customer segmentation**, businesses can allocate marketing budgets more efficiently, focusing on high-value customers.
- Using RFM and CLV analysis can reduce Customer Acquisition Cost (CAC) by up to 25%, as companies can concentrate on more profitable customers (Feaseo, 2022)

#### **Optimized Customer Lifetime Value (CLV)**

RFM-based CLV calculations and clustering techniques like K-Means can improve marketing efficiency by 30% and long-term profitability by 20% (<u>Laksono et al. 2023</u>).



## **Data Preprocessing**



	Column	Missing Values	Missing Percentage (%)
0	transaction_id	0	0.000000
1	customer_id	0	0.000000
2	category	0	0.000000
3	item	1213	9.646123
4	price_per_unit	609	4.842942
5	quantity	604	4.803181
6	total_spent	604	4.803181
7	payment_method	0	0.000000
8	location	0	0.000000
9	date	0	0.000000
10	discount	4199	33.391650

#### Missing Value Handling

#### Columns with <5% missing</li>

"price\_per\_unit", "quantity",

"total\_Spent" 
Imputed with median
(no major data loss, maintains revenue
accuracy).

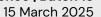
#### Columns with 5–20% missing

- "item" (9.65%) → Imputed using mode (most frequent item) instead of median because it's categorical.
- This ensures consistency, but we should check whether imputed values make sense for product categories.

#### Columns with >20% missing

- "discount" (33.39%) → Imputed with "False" assuming missing values mean no discount was applied.
- Key Benefit: Allows discount impact analysis to remain valid.







## 1. Customer Retention vs. One-Time Buyers



total_customers bigint		one_time_buyers bigint	â	repeat_customers bigint
2	5		0	12575

- $\bigvee$  Strong Retention  $\rightarrow$  Focus on increasing CLV.
- ✓ High Spenders → VIP perks & exclusive deals.
- Medium Spenders → Targeted discounts & upselling.
- Low Spenders → Loyalty points & free shipping.
- ✓ Al Recommendations & Cashback → Boost engagement & spending.



## 2. Customer Segmentation by Spending Behavior

	customer_id character varying (50)	total_spending numeric	customer_segment text
1	CUST_24	71056.00	High Spender
2	CUST_08	70172.50	High Spender
3	CUST_05	70012.50	High Spender
4	CUST_23	67979.00	High Spender
5	CUST_13	67858.00	High Spender
6	CUST_16	67849.00	High Spender
7	CUST_10	65325.50	High Spender
8	CUST_21	65320.00	Medium Spender
9	CUST_15	65070.50	Medium Spender
10	CUST_22	64662.00	Medium Spender
11	CUST_20	64462.50	Medium Spender
12	CUST_02	64325.00	Medium Spender
13	CUST_19	64052.00	Medium Spender

	customer_id character varying (50)	total_spending numeric	customer_segment text
14	CUST_07	63949.50	Low Spender
15	CUST_04	63829.00	Low Spender
16	CUST_09	63702.00	Low Spender
17	CUST_12	63530.00	Low Spender
18	CUST_11	63120.00	Low Spender
19	CUST_03	62872.50	Low Spender
20	CUST_18	62722.50	Low Spender
21	CUST_14	62481.00	Low Spender
22	CUST_17	61248.00	Low Spender
23	CUST_01	61118.50	Low Spender
24	CUST_06	60911.00	Low Spender
25	CUST_25	59976.50	Low Spender

## 3. Highest Revenue-Generating Items

	item character varying (100)	total_revenue numeric	total_quantity_sold numeric
1	Item_2_BEV	149000.50	7687.00
2	Item_25_FUR	25256.00	616.00
3	Item_25_EHE	23083.00	563.00
4	Item_25_BUT	21894.00	534.00

- **Item\_2\_BEV dominates sales** → Prioritize stock & promotions.
- Other high-revenue items have lower sales volume → Consider bundling or upselling strategies.



### 4. Weekend vs. Weekday Sales Performance

21172.00

	day_of_week numeric	total_revenue numeric
1	5	241357.50
2	0	234128.50
3	6	233088.00
4	2	229632.00
5	4	229471.00
6	3	228049.00
7	1	221879.00

Item 24 FUR

5

**Highest revenue on Day 5 (likely Friday)** → Ideal for flash sales & promotions.

536.00

Consistent revenue on other days → Maintain steady marketing efforts but boost weekend campaigns.

## 5. Customer Purchase Recency & Loyalty Analysis



	customer_id character varying (50)	last_purchase_date date	days_since_last_purchase integer	total_purchases bigint
1	CUST_17	2025-01-12	60	487
2	CUST_05	2025-01-14	58	544
3	CUST_23	2025-01-15	57	513
4	CUST_21	2025-01-15	57	498
5	CUST_22	2025-01-15	57	501
6	CUST_16	2025-01-16	56	515
7	CUST_20	2025-01-16	56	507
8	CUST_02	2025-01-16	56	488
9	CUST_03	2025-01-16	56	465
10	CUST_06	2025-01-16	56	481
11	CUST_18	2025-01-17	55	507
12	CUST_19	2025-01-17	55	487
13	CUST_09	2025-01-17	55	519

	customer_id character varying (50)	last_purchase_date date	days_since_last_purchase integer	total_purchases bigint
14	CUST_08	2025-01-17	55	533
15	CUST_14	2025-01-17	55	484
16	CUST_15	2025-01-17	55	519
17	CUST_07	2025-01-18	54	491
18	CUST_13	2025-01-18	54	534
19	CUST_01	2025-01-18	54	507
20	CUST_24	2025-01-18	54	543
21	CUST_04	2025-01-18	54	474
22	CUST_12	2025-01-18	54	498
23	CUST_11	2025-01-18	54	503
24	CUST_25	2025-01-18	54	476
25	CUST_10	2025-01-18	54	501

- Most customers made purchases
   54-60 days ago, indicating strong retention but a need for re-engagement.
- High total purchases suggest loyal customer behavior.

#### **Marketing Strategies**

- Reactivation Campaigns → Send reminders & special offers to customers inactive for 50+ days.
- **Loyalty Incentives** → Reward frequent buyers with exclusive discounts or VIP perks.
- Subscription Model → Offer auto-replenishment or membership benefits to ensure regular purchases.

Note: data analyzed on March 13, 2025

## 6. Customer Lifetime Value (CLV) Segmentation



	customer_id character varying (50)	lifetime_value numeric	total_transactions bigint	clv_segment text		customer_id character varying (50)	lifetime_value numeric	total_transactions bigint	clv_segment text
1	CUST_24	71056.00	543	High CLV	14	CUST_07	63949.50	491	Low CLV
2	CUST_08	70172.50	533	High CLV	15	CUST_04	63829.00	474	Low CLV
3	CUST_05	70012.50	544	High CLV	16	CUST_09	63702.00	519	Low CLV
4	CUST_23	67979.00	513	High CLV	17	CUST_12	63530.00	498	Low CLV
5	CUST_13	67858.00	534	High CLV	18	CUST_11	63120.00	503	Low CLV
6	CUST_16	67849.00	515	High CLV	19	CUST_03	62872.50	465	Low CLV
7	CUST_10	65325.50	501	High CLV	20	CUST_18	62722.50	507	Low CLV
8	CUST_21	65320.00	498	Medium CLV	21	CUST_14	62481.00	484	Low CLV
9	CUST_15	65070.50	519	Medium CLV	22	CUST_17	61248.00	487	Low CLV
10	CUST_22	64662.00	501	Medium CLV	23	CUST_01	61118.50	507	Low CLV
11	CUST_20	64462.50	507	Medium CLV	24	CUST_06	60911.00	481	Low CLV
12	CUST_02	64325.00	488	Medium CLV	25	CUST_25	59976.50	476	Low CLV
13	CUST_19	64052.00	487	Medium CLV	Tot	al rows: 25 of 25 Que	ery complete 00:00	0:00.280	

 $\label{eq:high-clv} \begin{array}{l} \textbf{High CLV Customers} \rightarrow \text{Major revenue drivers}, \\ \text{frequent \& high-value transactions}. \\ \textbf{Medium CLV Customers} \rightarrow \text{Potential to be converted} \\ \text{into High CLV with the right incentives}. \end{array}$ 

 $\mbox{\bf Low CLV Customers} \rightarrow \mbox{Need engagement \& retention strategies}.$ 

#### **Marketing Strategies**

- High CLV → VIP programs, early access to new products, premium offers.
- Medium CLV → Personalized discounts, loyalty points, targeted upselling.
- Low CLV → First-time purchase incentives, referral rewards, exclusive bundle deals.



	customer_id character varying (50)	past_purchases bigint	recent_purchases bigint	date
1	CUST_22	113	40	2024-09-07
2	CUST_19	101	50	2024-09-13
3	CUST_08	99	71	2024-09-05
4	CUST_13	98	61	2024-09-12
5	CUST_24	98	62	2024-09-10
6	CUST_11	96	57	2024-09-11
7	CUST_10	93	62	2024-09-13
8	CUST_05	93	62	2024-09-13
9	CUST_16	93	55	2024-09-11
10	CUST_09	92	64	2024-09-13

Total rows: 10 of 10 Query complete 00:00:00.294



- Win-Back Campaigns → Personalized emails with discounts or exclusive offers.
- Reactivation Discounts → Limited-time deals to encourage repeat purchases.
- Loyalty Perks → Reward them for coming back with points or VIP access.

Note: data analyzed on March 13, 2025

## 8. Discount Dependency Analysis



	customer_id character varying (50)	total_purchases bigint	discounted_purchases bigint	discount_dependency_percentage numeric
1	CUST_05	544	211	38.79
2	CUST_12	498	186	37.35
3	CUST_17	487	174	35.73
4	CUST_07	491	171	34.83
5	CUST_16	515	179	34.76
6	CUST_15	519	180	34.68
7	CUST_19	487	168	34.50
8	CUST_20	507	174	34.32
9	CUST_08	533	182	34.15
10	CUST_21	498	170	34.14
11	CUST_10	501	170	33.93
12	CUST_13	534	181	33.90
13	CUST_23	513	172	33.53

#### **Discount Dependency Trends**

- Some customers rely on discounts for over
   38% of their purchases, indicating a strong price sensitivity.
- Lower dependency customers (~28-30%) may still buy at full price.



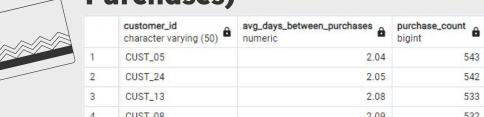
#### **Marketing Strategies**

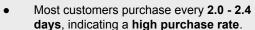
- Reduce discount reliance → Gradually decrease discounts while increasing value-based offers (bundles, exclusives).
- Personalized pricing → Offer discounts only to high-dependency customers to retain them while maximizing full-price sales from others.
- Loyalty Rewards → Encourage purchases without discounts through cashback, reward points, and exclusive perks.

	customer_id character varying (50)	total_purchases bigint	discounted_purchases bigint	discount_dependency_percentage numeric
14	CUST_11	503	168	33.40
15	CUST_01	507	169	33.33
16	CUST_18	507	167	32.94
17	CUST_02	488	160	32.79
18	CUST_09	519	170	32.76
19	CUST_22	501	163	32.53
20	CUST_04	474	154	32.49
21	CUST_06	481	154	32.02
22	CUST_25	476	151	31.72
23	CUST_24	543	168	30.94
24	CUST_14	484	144	29.75
25	CUST_03	465	133	28.60

Total rows: 25 of 25 Query complete 00:00:00.165

## 9. Purchase Frequency Patterns (Time Between **Purchases**)





Customers with lower average days between purchases are more engaged and ideal for retention efforts.



3	CUST_13	2.08	533						
4	CUST_08	2.09	532		customer_id		ava dava between nurebases		purchase_count
5	CUST_09	2.14	518		character varying (5	0)	avg_days_between_purchases numeric	â	bigint 6
6	CUST_15	2.15	518	14	CUST_22		2.2	22	500
7	CUST_16	2.16	514	15	CUST_21		2.2	23	497
8	CUST_23	2.17	512	16	CUST_12		2.2	24	497
9	CUST_20	2.19	506	17	CUST_02		2.2	25	487
10	CUST_01	2.20	506	18	CUST_07		2.2	27	490
11	CUST_18	2.20	506	19	CUST_17		2.2	28	486
12	CUST_10	2.22	500	20	CUST_14		2.2	29	483
13	CUST_11	2.22	502	21	CUST_19		2.2	29	486
Sub	scription Model → Off	er auto-replenishment or subscription		22	CUST_06		2.3	31	480
discounts for frequent buyers.  Timed Promotions → Send targeted deals right before their usual purchase cycle to maintain engagement.  Loyalty Program → Reward frequent purchases with exclusive perks or			_	23	CUST_25		2.3	34	475
			ıal	24	CUST_04		2.3	35	473
			25	CUST_03		2.3	38	464	
boni	bonus points.			Total rows: 25 of 25 Query complete 00:00:00.182					

543

542

E22

## 10. High-Value Customers at Risk of Churning



	customer_id character varying (50)	total_spent numeric	last_purchase_date date	days_since_last_purchase integer
1	CUST_05	70012.50	2025-01-14	58
2	CUST_23	67979.00	2025-01-15	57
3	CUST_16	67849.00	2025-01-16	56
4	CUST_08	70172.50	2025-01-17	55
5	CUST_10	65325.50	2025-01-18	54
6	CUST_24	71056.00	2025-01-18	54
7	CUST_13	67858.00	2025-01-18	54

Total rows: 7 of 7

Query complete 00:00:00.182

These customers have **spent significantly** but have not purchased recently (**54-58 days since last purchase**).

Losing them would impact revenue substantially.

#### **Marketing Strategies**

- Exclusive Win-Back Offers → Personalized discounts or VIP deals to re-engage.
- **Premium Loyalty Perks** → Special incentives (early access, exclusive bundles).
- Personalized Outreach → Email, SMS, or calls with tailored product recommendations.

Note: data analyzed on March 13, 2025

**Business Intelligence | Batch 13** 15 March 2025

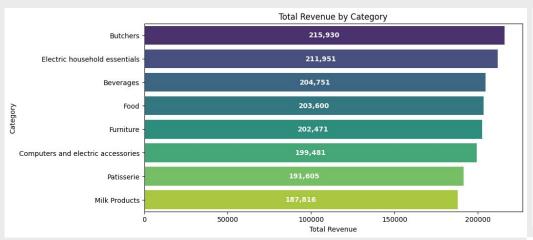


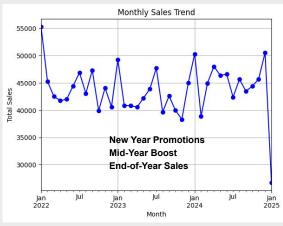
# Data Visualization





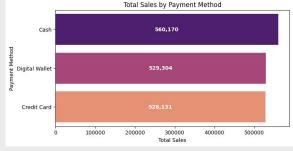
## Trends, Categories, and Payment Insights





#### **Key Insights & Marketing Strategy**

- Sales Trend: Sales fluctuate with seasonal peaks.
  - **Strategy:** Boost peak months with promotions, stabilize low months with targeted discounts.
- 2 Top Revenue Categories: Butchers, Electronics, and Beverages lead sales.
  - \* Strategy: Expand top categories & bundle low-performers with best-sellers.
- 3 Payment Methods: Cash dominates, but Digital Wallet & Credit Card are close.
  - **Strategy:** Offer **cashback & discounts** for digital payments to increase usage.

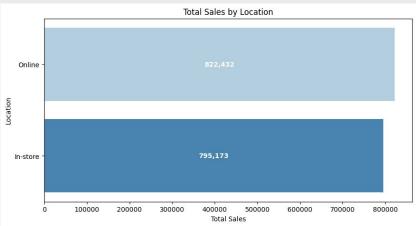


## **Location & Discount Impact Analysis**



**Discounts Have Minimal Impact** (546K vs. 535K).

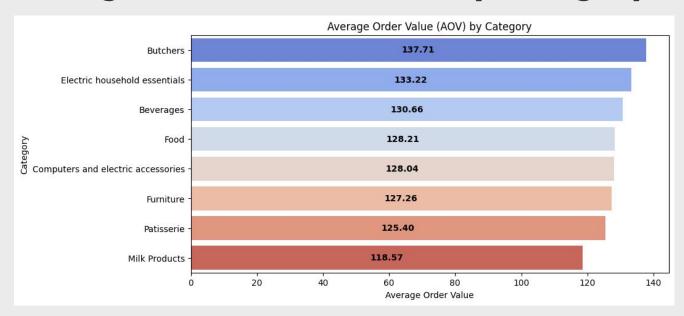
Poptimize promotions & focus on value-based incentives.



Online Sales Lead (822K vs. 795K in-store).

Boost online marketing & improve in-store experience.

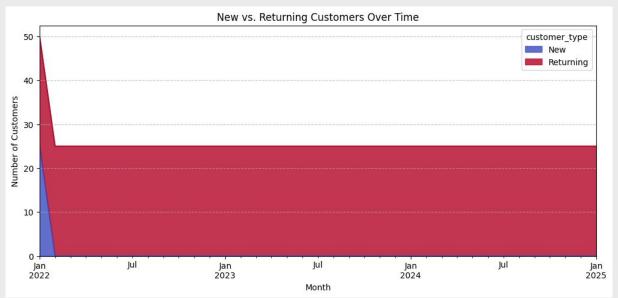
## **Average Order Value (AOV) by Category**



#### **Key Insights:**

- ☐ Highest AOV → Butchers (\$137.71), Electronics (\$133.22), and Beverages (\$130.66).
- **2** Lowest AOV → Milk Products (\$118.57).

## **New vs. Returning Customers Over Time**

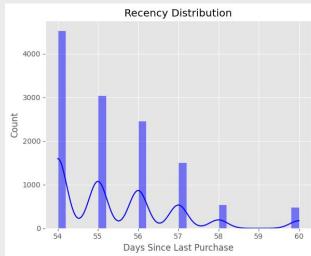


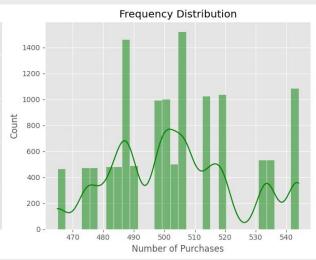
#### **Key Insights:**

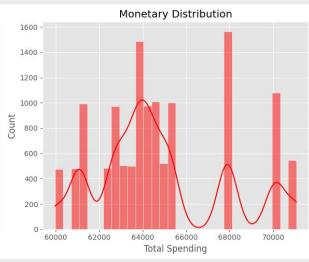
- Returning Customers Dominate
  - Strong customer retention, but fewer new customers over time.
- 2 New Customer Decline
  - Need to attract more first-time buyers.

- ★ Marketing Strategy:
- ✓ Targeted acquisition campaigns (ads, referrals, welcome discounts).
- Loyalty perks & exclusive deals to maintain returning customers.

# Distribution of Recency, Frequency, and Monetary Values







#### Recency (Days Since Last Purchase):

- Most customers made purchases 54-60 days ago, indicating strong repeat buying behavior.
- Strategy: Use reminders & special deals to maintain engagement.

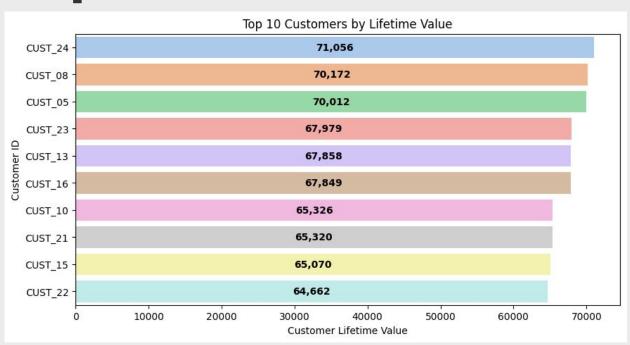
#### 2 Frequency (Number of Purchases):

- Customers frequently buy 470-545 times, showing high engagement.
- Strategy: Reward frequent buyers with loyalty perks & bulk purchase incentives.

#### 3 Monetary (Total Spending):

- Spending clusters around \$62,000 \$70,000, highlighting high-value customers.
- Strategy: Upsell premium products & introduce VIP loyalty programs.

# Top 10 CLV



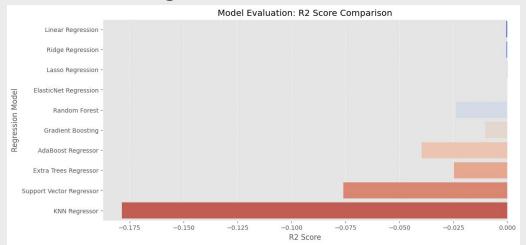
## **I**CLV is Based on Total Spending

 The CLV calculation directly reflects total spent rather than predictive modeling.

#### 2 Low Variability in CLV

Customers' CLV values do not differ significantly, making segmentation harder.

#### **Machine Learning for 3-Month CLV Prediction**



Мо	del Evaluation Results:				
	Model	MAE	MSE	RMSE	R2 Score
3	ElasticNet Regression	75.675945	8781.279591	93.708482	-0.000127
2	Lasso Regression	75.677738	8781.538598	93.709864	-0.000157
1	Ridge Regression	75.686408	8786.344102	93.735501	-0.000704
0	Linear Regression	75.686934	8786.510836	93.736390	-0.000723
5	Gradient Boosting	76.030122	8869.922072	94.180264	-0.010223
4	Random Forest	76.417563	8988.720612	94.808864	-0.023753
7	Extra Trees Regressor	76.447660	8997.637359	94.855877	-0.024769
6	AdaBoost Regressor	80.574862	9129.730664	95.549624	-0.039813
8	Support Vector Regressor	74.245381	9447.094423	97.196165	-0.075959
9	KNN Regressor	81.695905	10346.948286	101.719950	-0.178446

#### **Key Insights:**

- **□** ElasticNet Regression Performed Best
  - Achieved lowest MAE & MSE, showing potential for CLV prediction.
- 2 Data Quality Impacts Accuracy
  - CLV values show limited variation





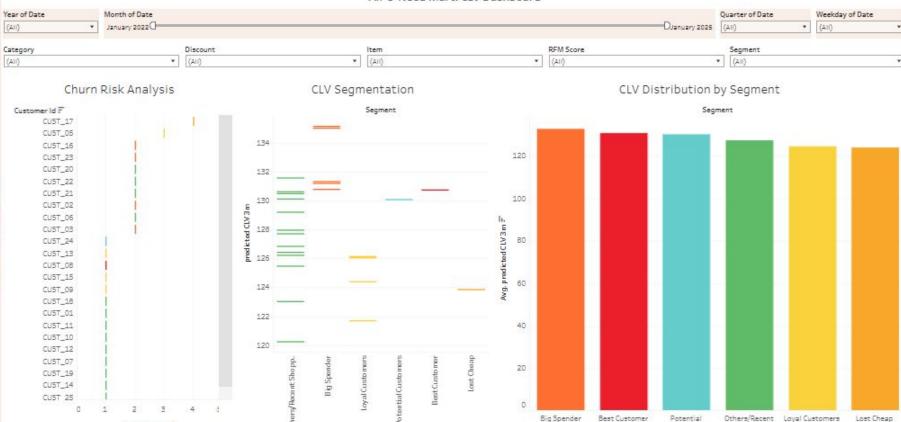
# Dashboard







#### All-U-Need Mart: CLV Dashboard



Customers

Shopper

Recency Quartile







# **Business Recommendation**



Recomm.	Insights	Actionable Items
Online sales outperform in-store.	Online sales outperform in-store.	Strengthen digital marketing efforts.     Optimize logistics for online deliveries.
Improve Customer Retention	Retention rate is high (100%), but churn risk needs monitoring.	Implement personalized promotions for repeat buyers.     Strengthen customer engagement.
Enhance RFM-Based Customer Segmentation	Big Spenders & Loyal Customers contribute most to sales.	Offer exclusive discounts & perks.     Re-engage at-risk customers with targeted campaigns.
Optimize Product Pricing & Promotions	High AOV (128.6), but discount impact is minimal on total sales.	Optimize discount strategy to avoid over-reliance.     Focus on high-value product bundles.
Leverage Best-Selling Categories	Butchers, Household Essentials, and Beverages lead revenue.	Highlight these categories in marketing.     Introduce loyalty points for frequent buyers.
Boost Low-Performing Products	Some products have lower sales contributions.	Bundle low-sellers with best-sellers.     Offer limited-time promotions to drive demand.
Maximize Peak Month Sales	Sales peak in January and July.	Launch seasonal campaigns.     Plan exclusive promotions during high-sales periods.

Business Intelligence | Batch 13

15 March 2025

# Thanks!

Do you have any questions? <a href="https://www.hijirdata.com/">https://www.hijirdata.com/</a>



https://www.linkedin.com/in/hijirdella/







## Resources



- Desai, S. K., & Kalyandurgmath, K. (2021). Strategy to increase lifetime value of a customer using RFM. International Journal of Creative Research Thoughts (IJCRT), 9(7), ISSN: 2320-2882.
- Yulianti, Yulianti. (2018). Metode Data mining Untuk prediksi Churn Pelanggan. Jurnal ICT Akademi Telkom Jakarta. 9. 46-52.
- Laksono, B. C., & Wulansari, I. Y. (n.d.).(2023). Pemodelan dan penerapan metode RFM pada estimasi nilai konsumen (Customer Lifetime Value) menggunakan K-Means Clustering Machine Learning. Politeknik Statistika STIS.



# **Customer Segmentation**

- 1. **Best Customer** (R=1, F=1, M=1)
  Highly engaged, frequent buyers with high spending. Prioritize loyalty programs and exclusive rewards.
- 2. **Potential Customers** (R=1, F=1, M=2) or (R=1, F=2, M=2) or (R=2, F=1, M=1) or (R=2, F=2, M=2) Emerging valuable customers. Encourage repeat purchases with personalized offers.
- 3. **Lost Cheap** (R=4, F=4, M=4 or 3) Inactive, low-value buyers. Low priority for retention, but could be re-engaged with discounts.
- 4. **Big Spender** (M=1)
  High-value customers regardless of frequency. Offer premium deals and upsell opportunities.
- Loyal Customers (F=1)
   Frequent buyers with consistent purchases. Strengthen relationships through VIP programs.
- Others/Recent Shopper (R=1 or R=2)
   New or occasional buyers. Encourage further engagement through follow-ups and promotions.
- Almost Lost (R=3)
   Customers showing decreased engagement. Use retention strategies like personalized reactivation campaigns.
- 8. **Lost Customers** (R=4) Inactive buyers. Re-engage with strong incentives like exclusive discounts or special deals.