Group Project 1

BAN 5763, Spring 2024

Customer Lifetime Value Analysis

SFRR Analytics

Shreyan Datta Chakraborty
Roe Djer Tan
Frankle Muchahary
Ritwik Roy Chowdhury

Why do we need to understand Customer Lifetime Value?

Average Churn Rate for B&M Retailers[1]

Cost of Acquiring New Customers vs Retaining Existing Customers[2]

Value generated in 2-5 Years timeframe vs Year 1[3]

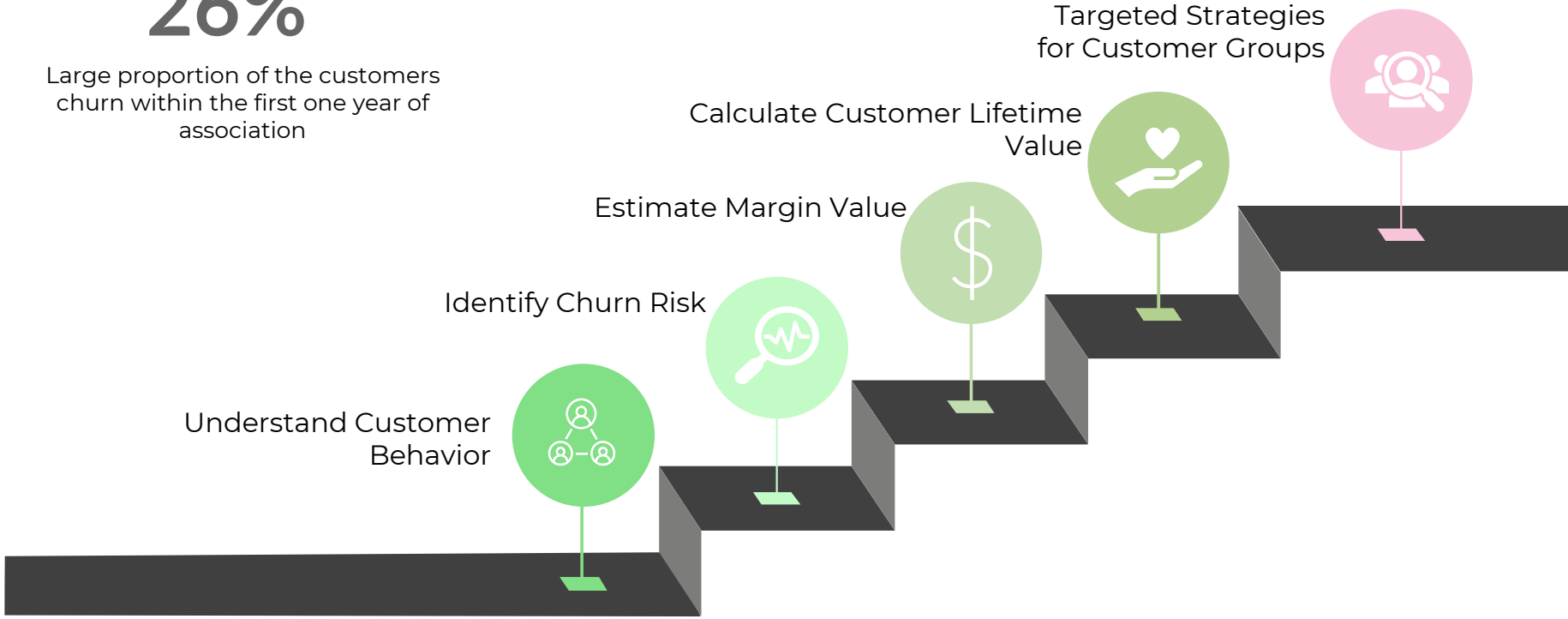
Can we estimate Potential Future Customer Value???

[1] https://www.zendesk.com/blog/customer-churn-rate/#:~:text=According%20to%20the%20Zendesk%20Customer,all%20resulting%20in%20customer%20churn
[2] https://www.forbes.com/sites/forbesbusinesscouncil/2022/12/12/customer-retention-versus-customer-acquisition/?sh=67b707a91c7d
[3] https://www.forbes.com/sites/sap/2019/09/12/why-and-how-companies-should-calculate-customer-lifetime-value/?sh=49937efd459d

Objective: What are we trying to solve?

26%

Large proportion of the customers churn within the first one year of association



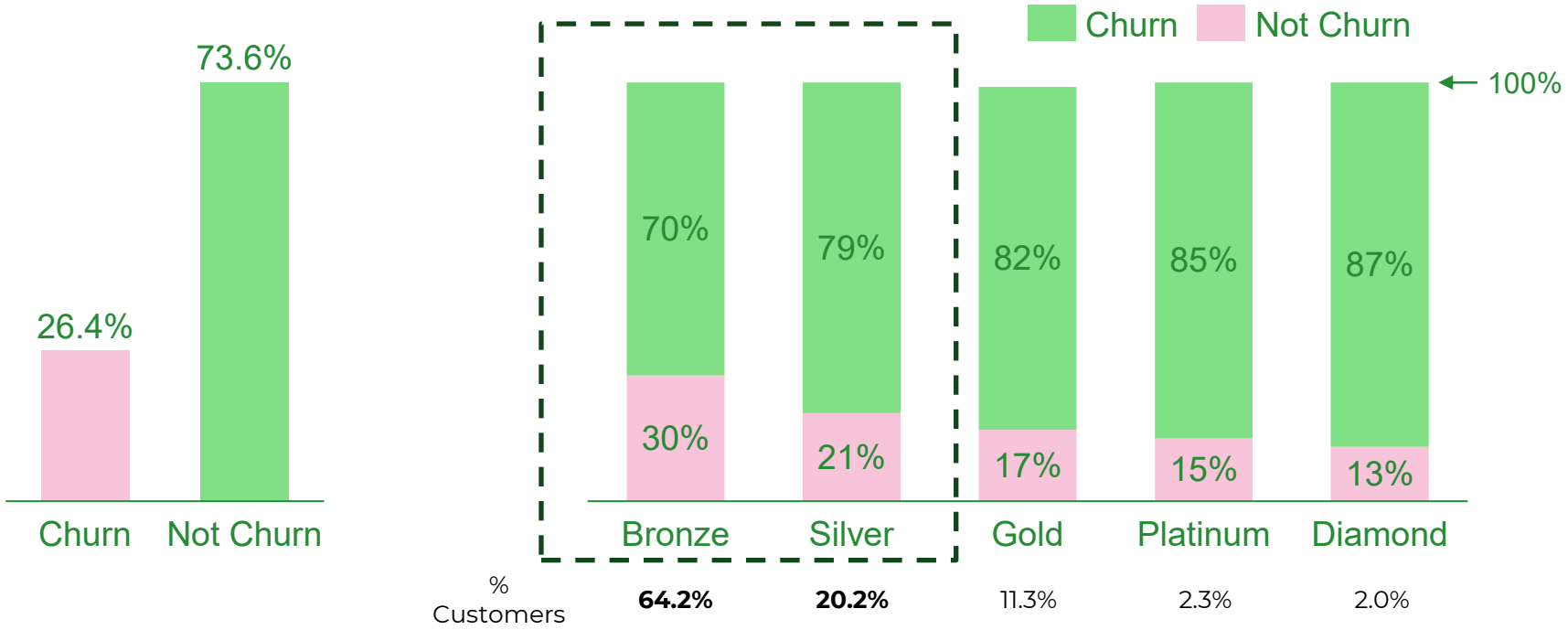
Executive Summary

 Complication	• Retail customer engagement is multifaceted & exhibit diverse behaviors - influencing their loyalty and spending • Retaining customers is crucial for creating long term engagement & value
 Varying Problems	• Customers are concentrated in the lower tiers and have highest rates of churn. • Different demographic groups have varying preferences and rate of churn • Segmenting customers into groups can help design targeted strategies
 Recommendations	• Immediate focus required for disengaged customers. • Tailored loyalty campaigns for high-value customer groups – long term value. • Higher Tier offerings need revamping to offer better value for money.

EXPLORATORY DATA ANALYSIS

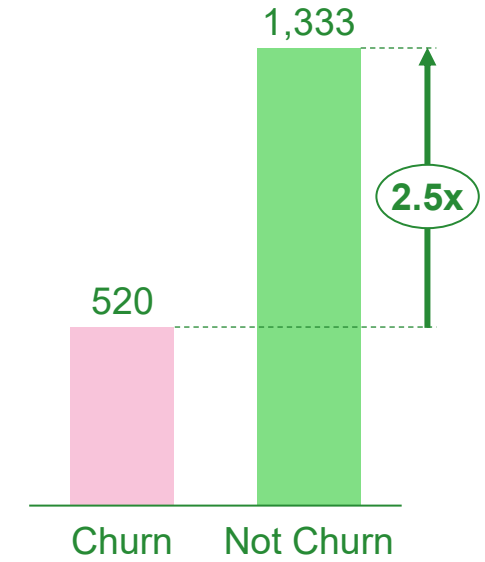
Year I Churn Rate – A look across the different subscription tiers

Customers from lower subscription tiers are at a much higher risk of churning. Customers are concentrated in these tiers

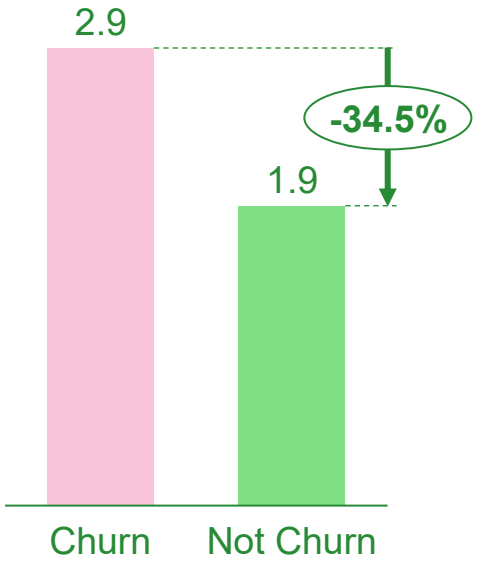


Retained customers generate up to 2.5 times more revenue across 3 years

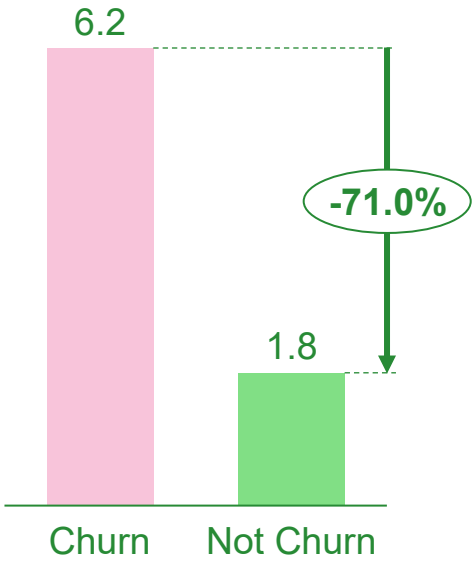
RFM Metrics will be crucial in determining Churn Risk. Significantly different b/c Churn and Not Churn



Average Purchase Amount – 3 Years

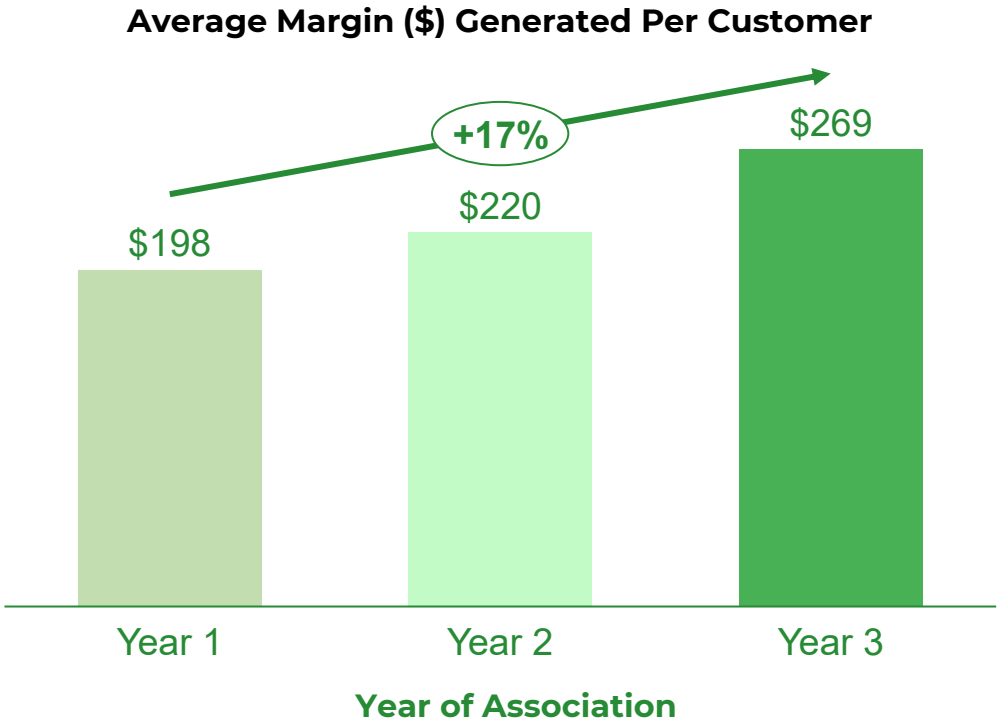


Average Interval B/W Purchases



Average Recency Months

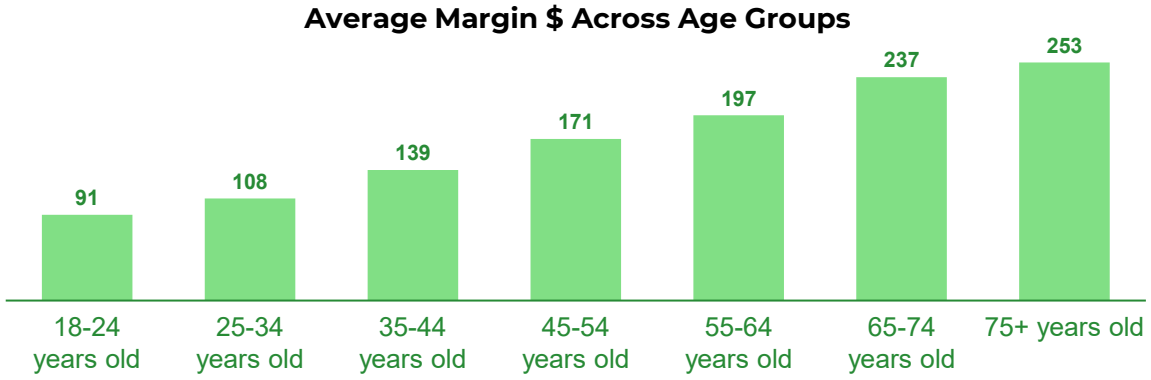
Customers generate more value on average in the corresponding years of association compared to the first year

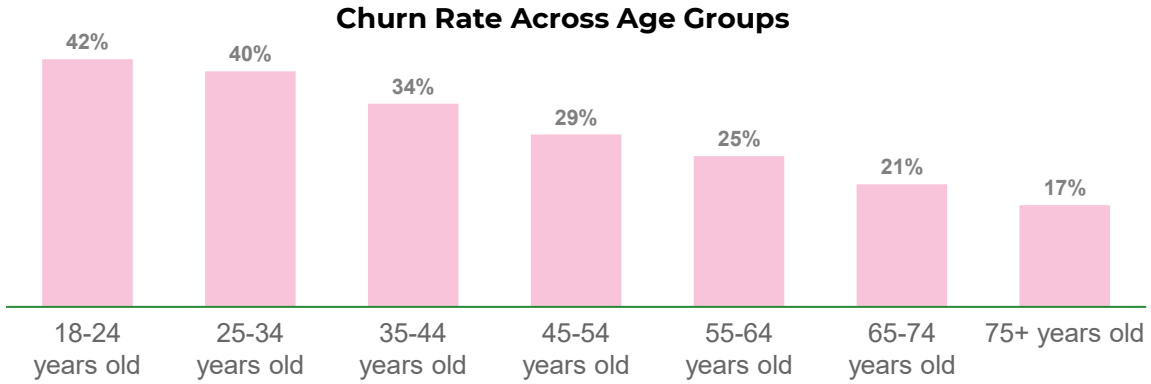


- **Average value increases across the years**
- **Creating loyal customers can generate 17% more margin in the long run**
- **Loyalty and Retention Campaigns are crucial**

Note: Same set of customers were used across all three years for a fair comparison. Newly acquired customers in year 2 and year 3 were ignored.

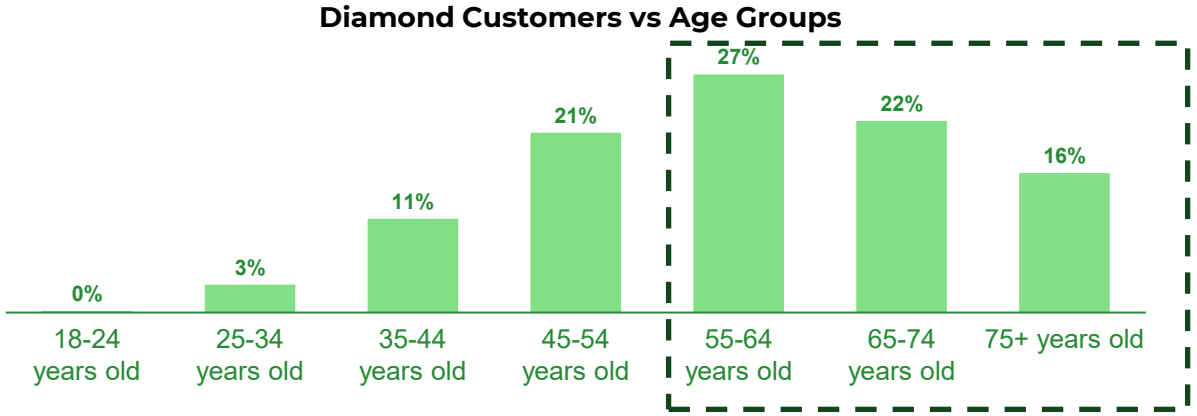
Older Age Groups: Higher Average Margin $ and Higher Retention Rates

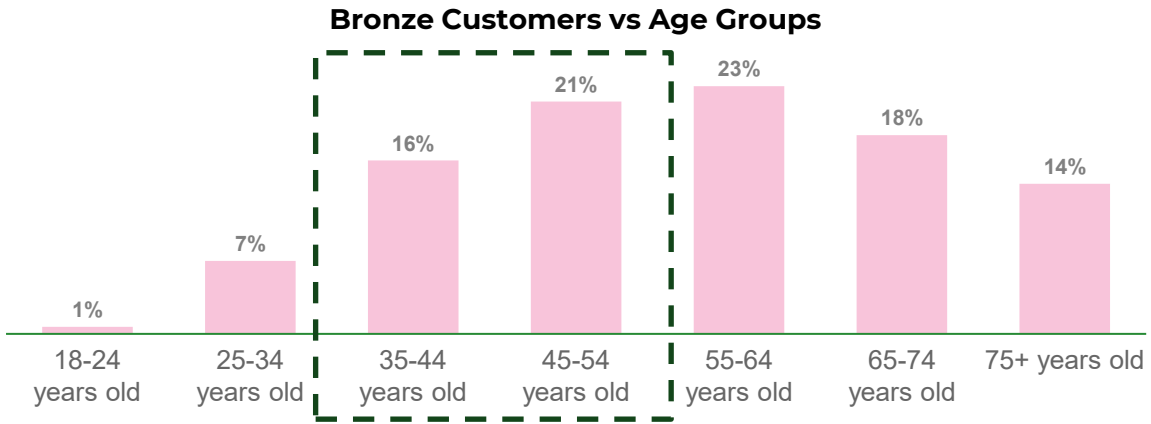




- **Older aged customers are more satisfied with the services**
- **Crucial to understand why younger aged customers are leaving**
- **Customized service offerings fine-tuned towards different age buckets might be helpful**

Older Age Groups: Higher proportion of upper tier customers





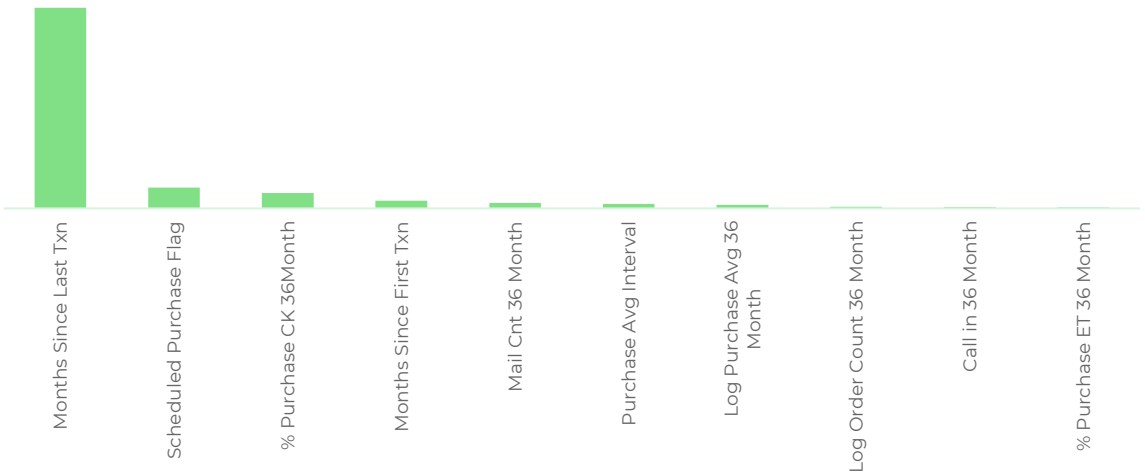
- **Larger proportion of Diamond Customers are older aged**
- **Lower Tier (Bronze, Silver) customers tend to be on the younger spectrum**

PREDICTING CHURN RATE

Churn Model – Performance Metrics

Model/Approach	Misclassification Rate	AUC
Logistic Regression	**10.0%**	**0.91**
Gradient Boosted Decision Trees	8.6%	0.94

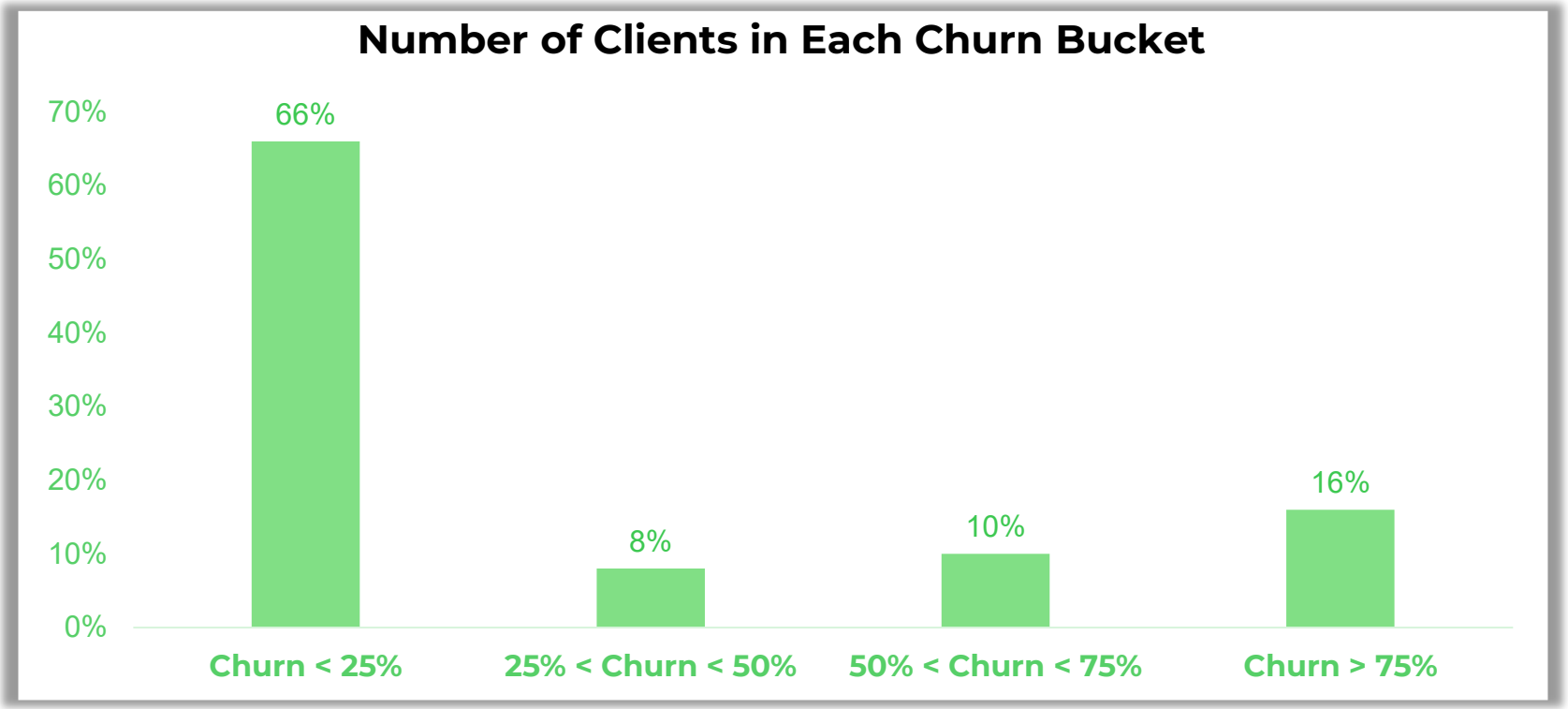
- GBDT - additional 1.4% customers identified correctly
- Can lead to better retention and cost savings during retention campaigns



- Recency is the most important metric for identifying risk of churn
- Automated Purchase Schedule also helps identify churn
- Mode of Payment also seems to have an impact on churn rate (CK, ET)

Approach for modeling detailed out in Appendix

Churn Probability Scores – Group Breakdown

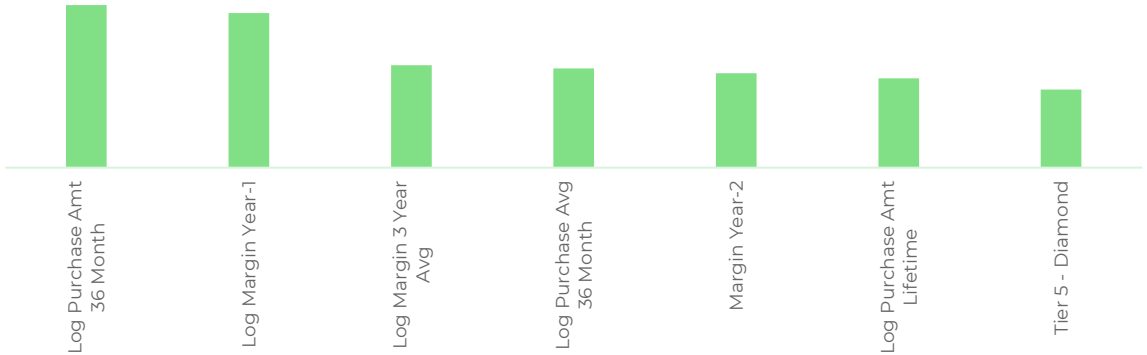


PREDICTING MARGIN VALUE

Margin Estimation Model – Performance Metrics

Model/Approach	R Squared	Test MSE
Logistic Regression	59.8%	1.4×10^{-4}
Random Forest	**72.12%**	**1.0×10^{-4}**
Gradient Boosted Decision Trees	53.12%	1.1×10^{-4}

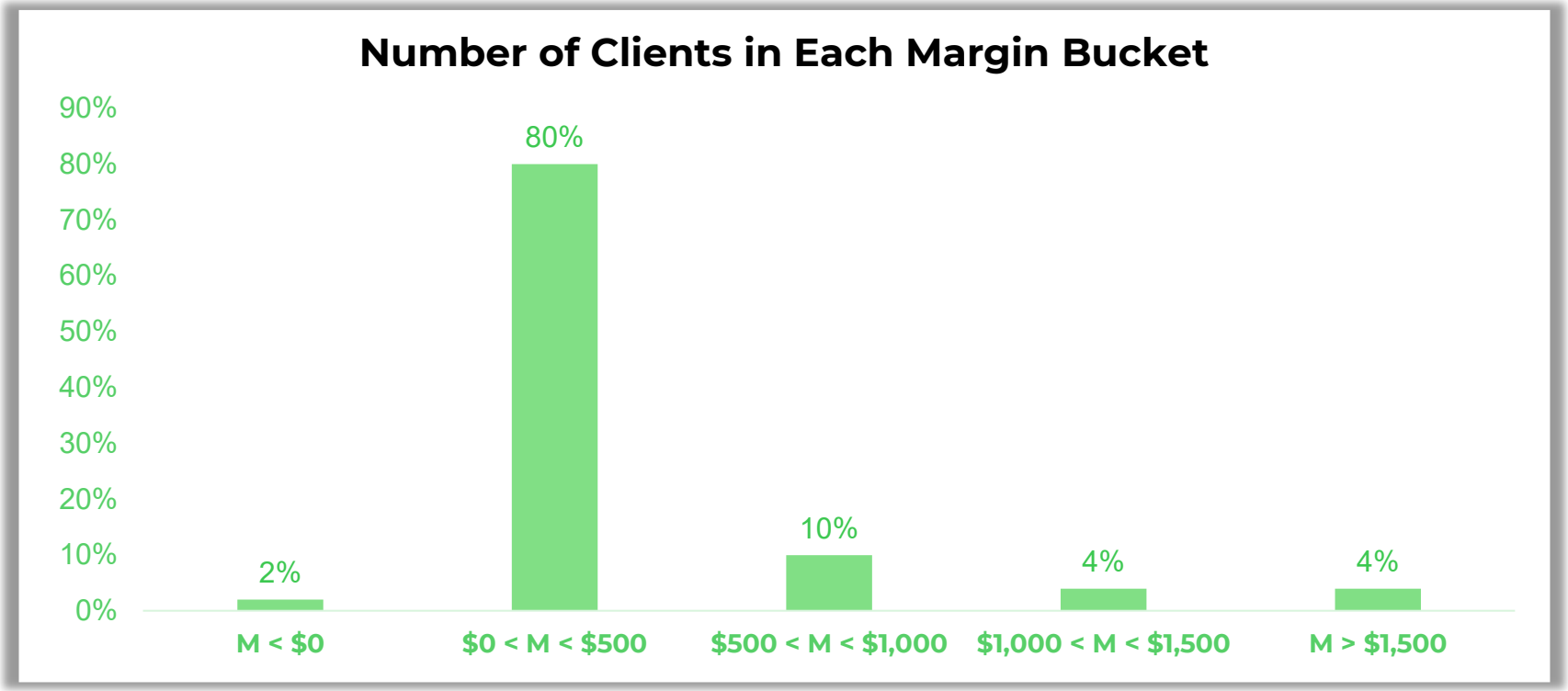
- RF model provides a good tradeoff between training time and model performance
- GBDT will likely require extensive parameter tuning for performance



Historical Customer Monetary Values are the best indicators for future margin estimates

Approach for modeling detailed out in Appendix

Estimated Margin – Group Breakdown



*M = Margin

ESTIMATING FUTURE CLV

Approach for estimating Customer Future Value (CFV)

Step 1

Churn Modeling & Margin Modeling

Obtained churn score and margin score

Step 2

Reverse Transformation

Reverse transformed margin score to dollar value

$$P_{margin} = 10^{\log\_margin} - 25{,}000$$

Step 3

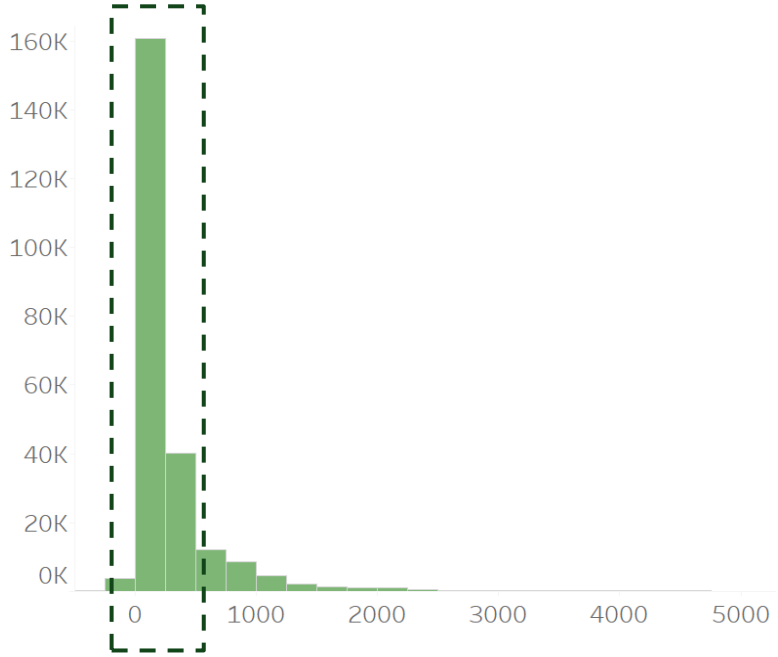
Calculate CFV

1-year, 2-year, and 3-year CFVs are calculated based on the following formula:

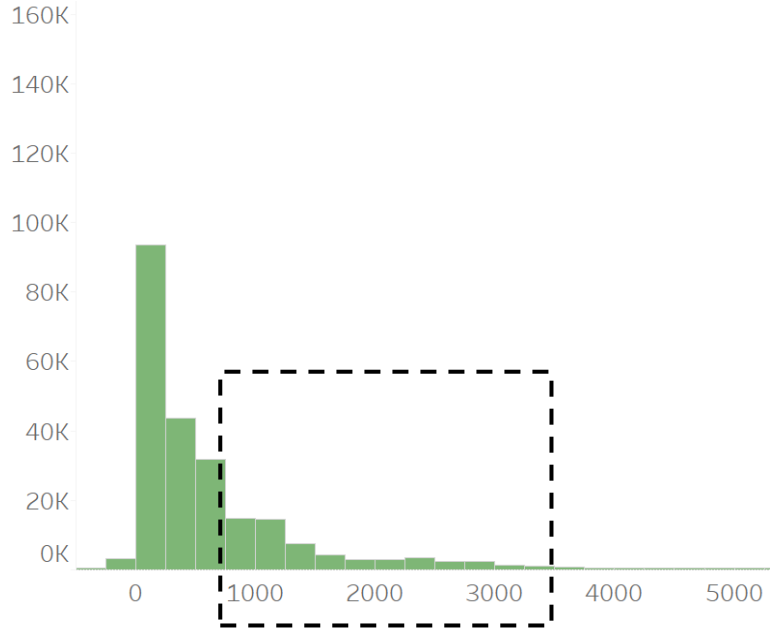
$$CLV_j = m_j \times \left(\frac{1 - c_j}{d + c_j}\right)$$

Customer Future Value: 1-Year Value vs 3-Year Value

Customer value increases as the length of relationship with the company increases. Increasing customer loyalty can help generate long-term value.



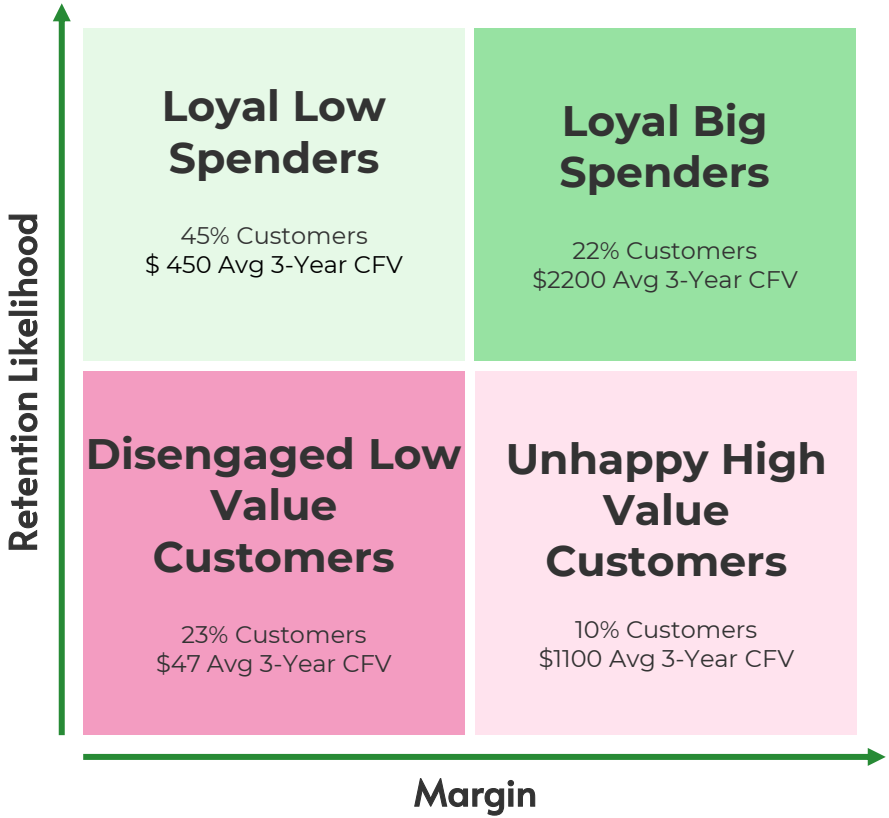
CFV Distribution – 1 Year



CFV Distribution – 3 Year

CUSTOMER SEGMENTATION

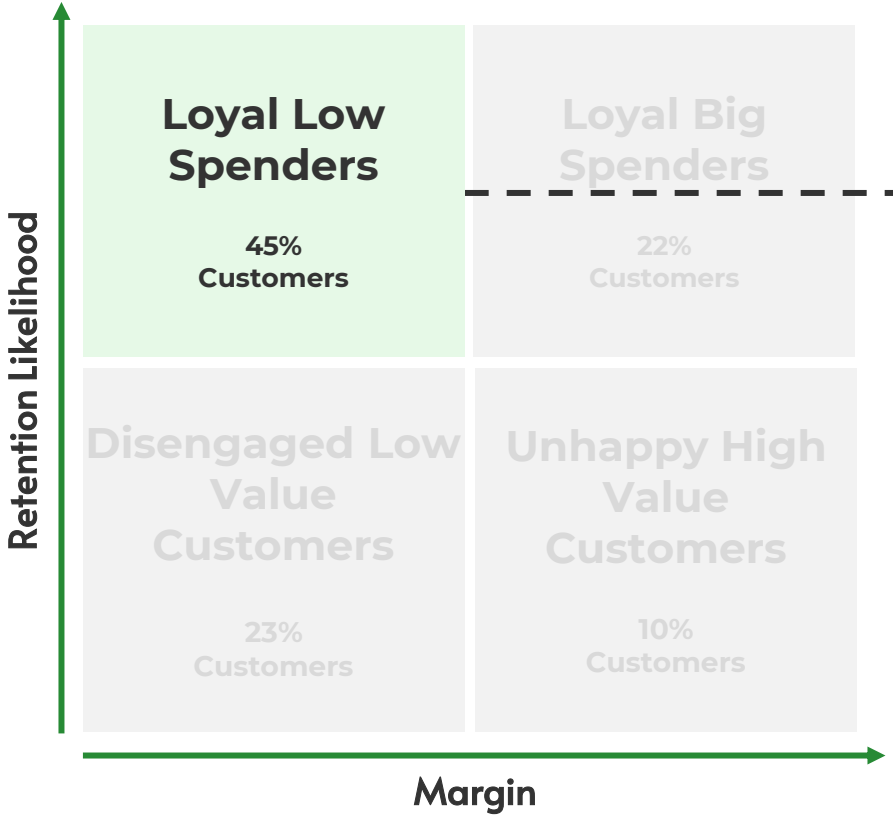
Customers can be segmented and prioritized based on Margin and Retention



Loyal Big Spenders: High Margin – High Retention (Group I)

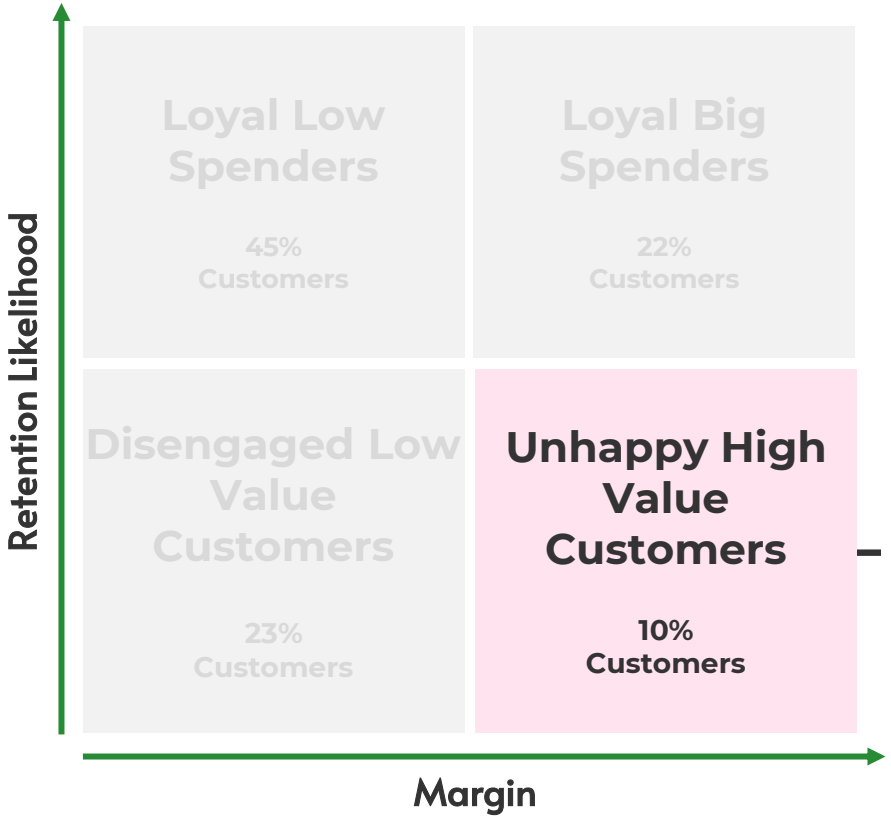
- Lowest Churn Rate (5%)
- Highest 3-Year CFV ($ 2200)
- High Income Group
- 75% Customers in Silver + Gold Tiers

Loyal Low Spenders: High Retention & Low Margin (Group 2)



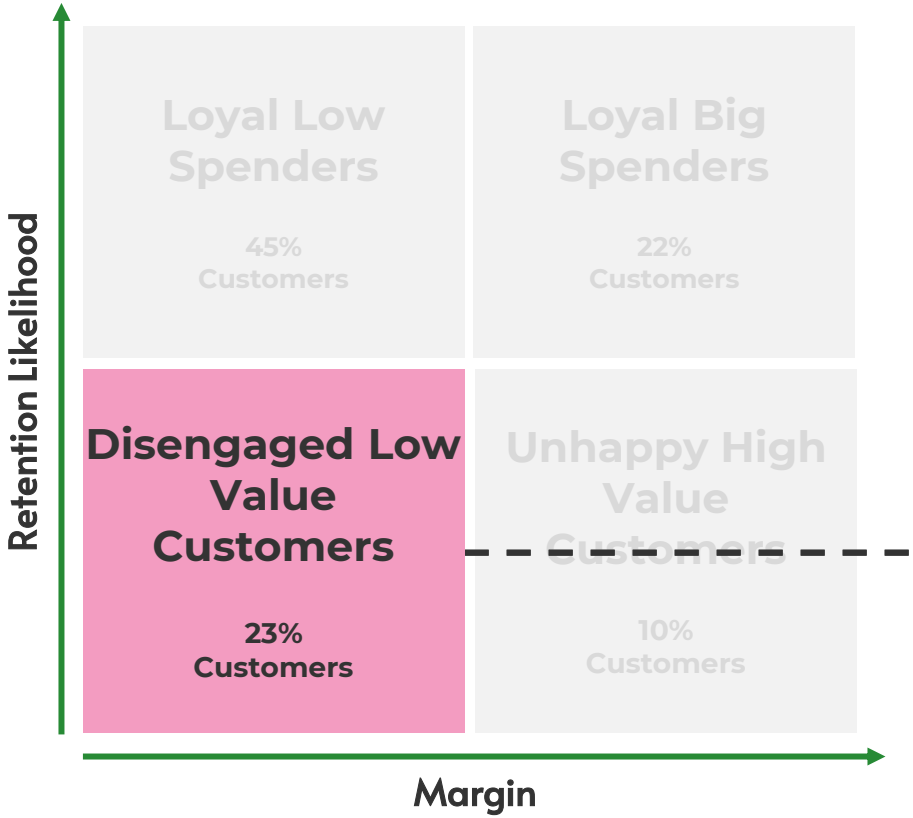
- Low Churn Rate (9%)
- Second Lowest 3-Year CFV ($ 450)
- Low Income Group
- 80% Customers in Bronze Tier

Unhappy High Value Customers: High Margin – Low Retention Rate (Group 3)



- High Churn Rate (38%)
- High Spenders. $1100 3-Year CFV
- Mostly middle-aged customers
- Highest % of Platinum + Diamond Customers

High-Risk Low Value Customers: Low Margin – Low Retention Rate (Group 4)



- Highest Churn Rate (76%)
- Mostly younger aged customers
- Similar Avg Purchase Amount to Loyal Low Spenders
- Last purchase was 7 months back on average

Customer Segments vs Subscription Tiers

- Opportunity for moving these customers to Platinum and Diamond
- Identify issues with Platinum and Diamond before running up-selling campaigns

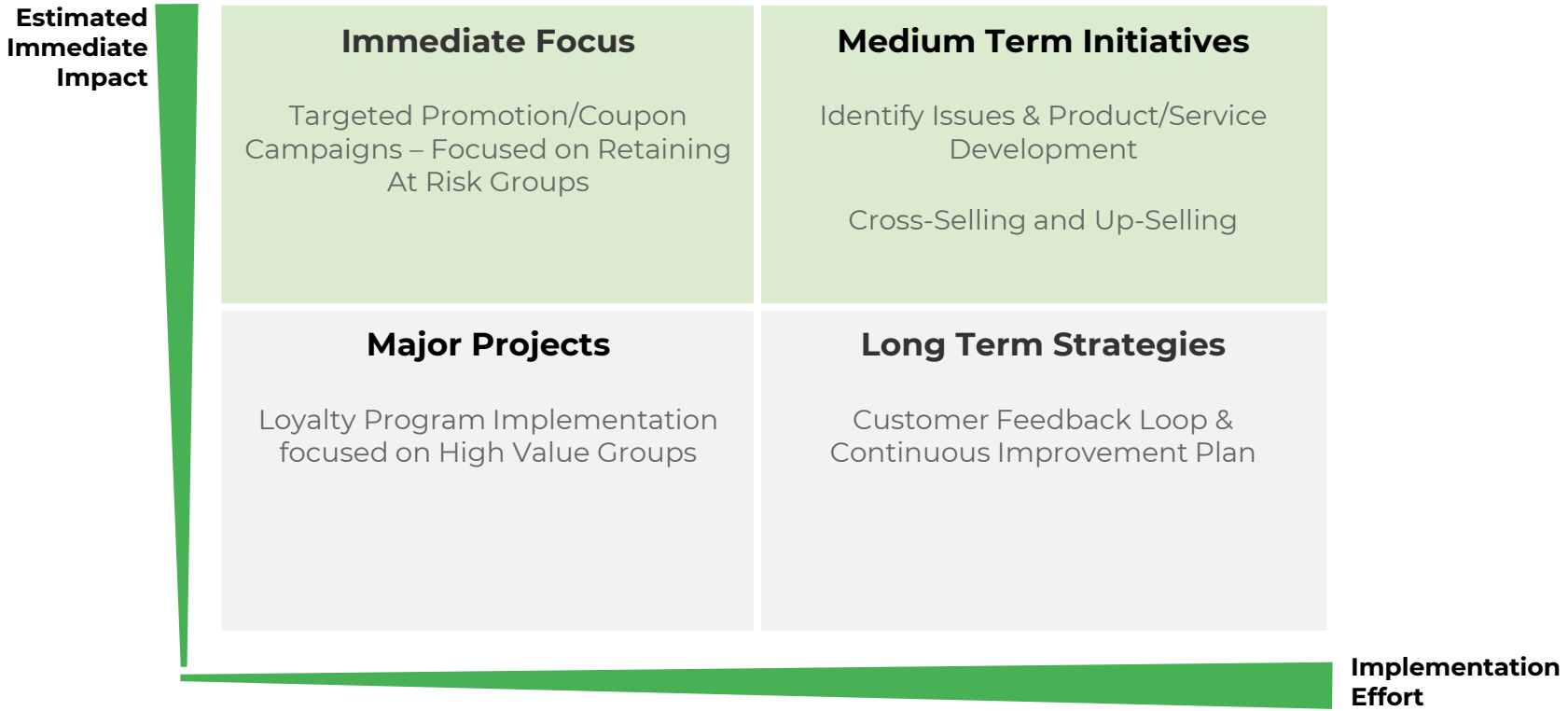
	Bronze	Silver	Gold	Platinum	Diamond
Loyal Big Spenders	15%	41%	35%	6%	4%
Loyal Low Spenders	81%	15%	4%	0%	0%
Unhappy High Value Customers	72%	8%	5%	5%	10%
Disengaged Low Value Customers	75%	17%	7%	1%	0%

- Opportunity for up-selling
- Move these customers to Silver & Gold Tiers through up-selling and cross-selling

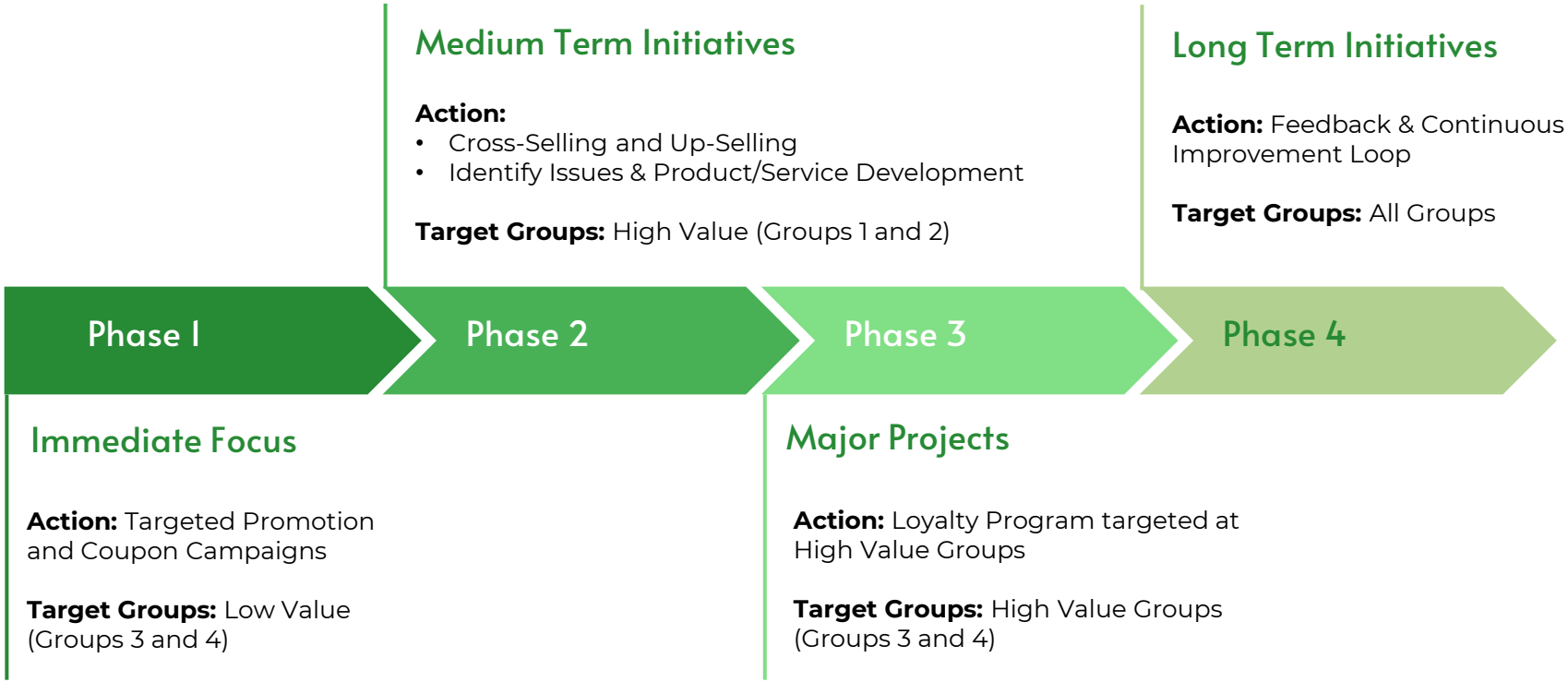
- Unhappy High Value customers have the highest proportion of Platinum and Diamond Customers.
- Group has high churn rate
- Higher tiers might have lower value for money

RECOMMENDATIONS

Recommended Actions – Estimated Immediate Impact vs. Implementation Effort

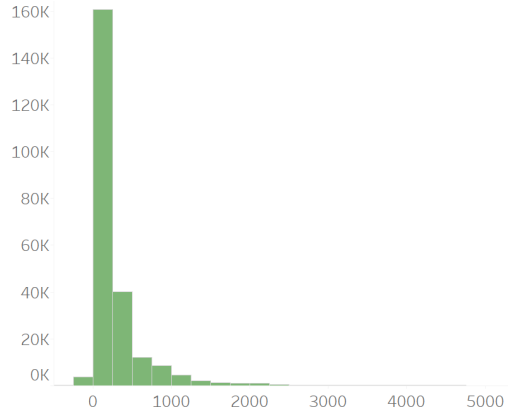


Phased Implementation Strategy & Groups to Focus for Each Phase

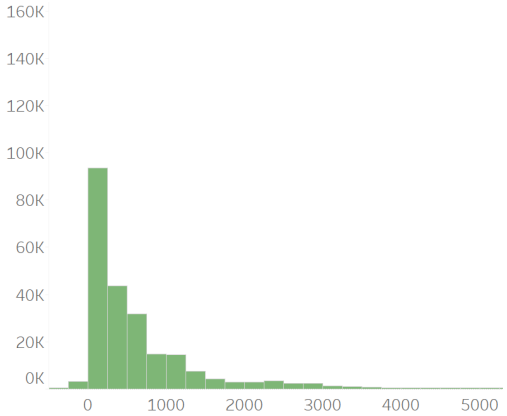


APPENDIX

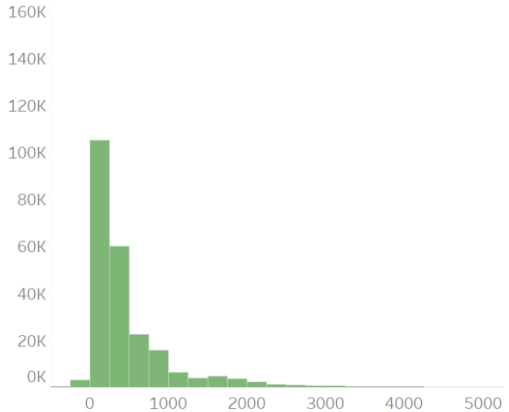
CFV Distribution Histogram



Year 1 CFV



Year 3 CFV



Year 2 CFV

Hyperparameters – Margin Model and Churn Model

Margin Model

1. Random Forest
n_estimators: 120,
min_samples_split: 6,
min_samples_leaf: 2,
max_features: 'sqrt',
max_samples: 0.5,
max_depth: 40

2. Gradient Boosted DT
n_estimators: 70,
min_samples_split: 2,
max_depth: 5,
learning_rate: 0.01,
loss: squared_error}

Churn Model

1. Gradient Boosting
n_estimators: 200,
min_samples_split: 2,
max_depth: 5,
Learning_rate: 0.05

Segmentation Profiling

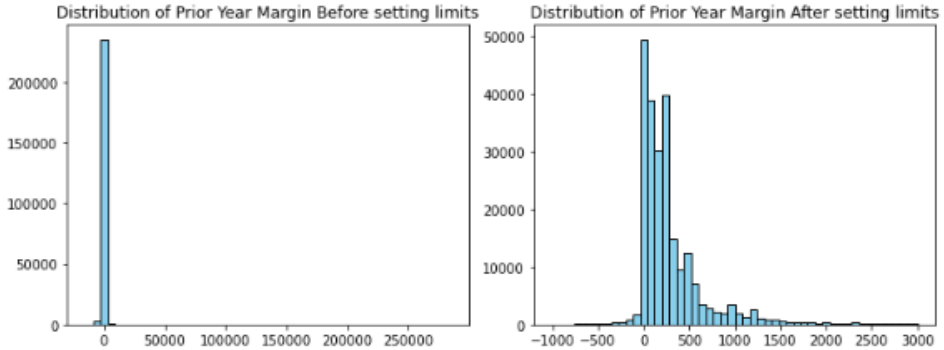
Cluster	Proportion of Client	Average Margin	Average Churn Rate	Average 3-Year CFV	Average Purchase Intervals	Average Purchase Amount in 36 Months	Average Months Since Last Transaction
Loyal Big Spender	21.54%	984.09	5%	$2,209.82	1.4	$3,030.66	1.31
Disengaged Low Value Customer	22.72%	131.05	76%	$47.84	2.71	$418.46	7.36
Unhappy High Value Customer	10.29%	573.83	38%	$1,136.63	30.1	$1,492.12	3.2
Loyal Low Spender	45.45%	207.41	9%	$453.3	1.99	$478.26	1.39

Segmentation Profiling

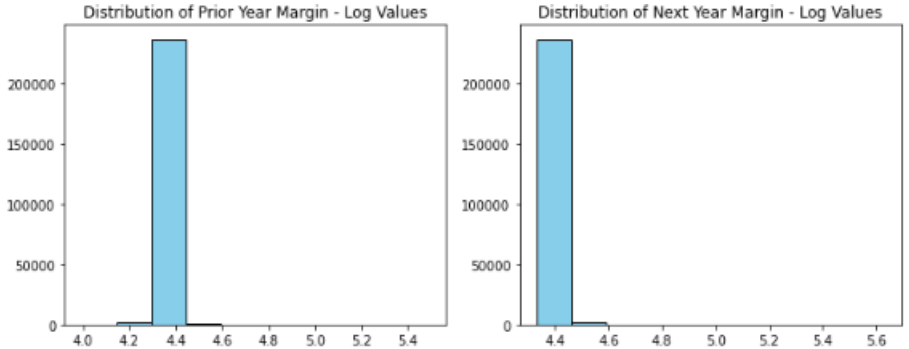
	Demog Inc Crs (group)	Demog Inc Crs (group)	Demog Inc Crs (group)
Cluster	$ 15000 - $ 29999, $ 30000 - $ 39999, Less than $15000	$ 40000 - $ 49999, $ 50000 - $ 74999, $ 75000 - $ 99999	$ 100000 - $149999, $ 150000 - $199999, $ 200000 or more
Loyal Big Spender	46.17%	42.90%	10.93%
Disengaged Low Value Customer	55.66%	36.98%	7.36%
Unhappy High Value Customer	55.06%	36.40%	8.54%
Loyal Low Spender	54.16%	38.50%	7.35%

Churn Model: Sample Untransformed vs Transformed Variables

Original Variables (Untransformed)



Transformed Independent Variables (Prior Year)

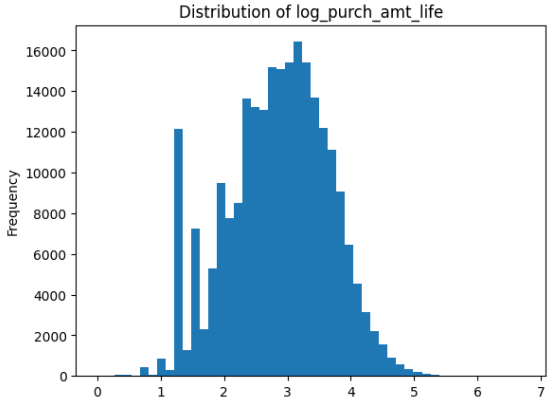


Transformed Target

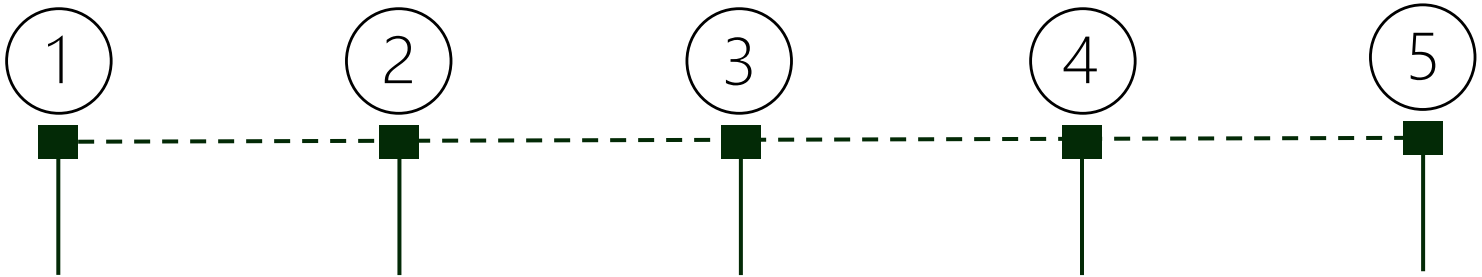
Margin Model: Sample Untransformed vs Transformed Variables

Original Variables (Untransformed)

Transformed Variables



Approach for Estimating Churn/Retention Rate



Data Cleaning

Remove Duplicate Account IDs

Remove high cardinality variables

Feature Engineering

Reduce cardinality (Recategorization)

Power Transformation

Standardization

Train-Test Split

30% Holdout Set

Feature Selection

Logistic Regression with L1 Regularization

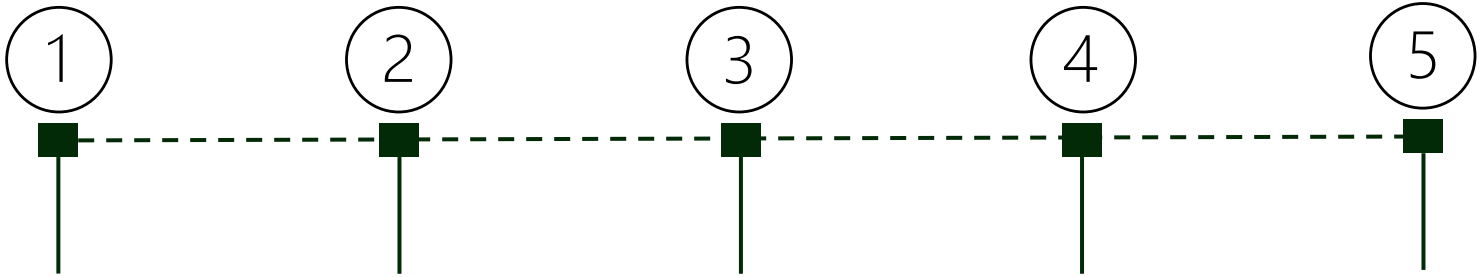
VIF Feature Reduction

Classification Modeling

Logistic Regression

Gradient Boosting

Approach for Estimating Margin



Data Cleaning

Remove Duplicate Account IDs

Remove high cardinality variables

Feature Engineering

Margin Calculation

Average Rolling Margins

Power Transformation

Train-Test Split

30% Holdout Set

Feature Selection

VIF Feature Reduction

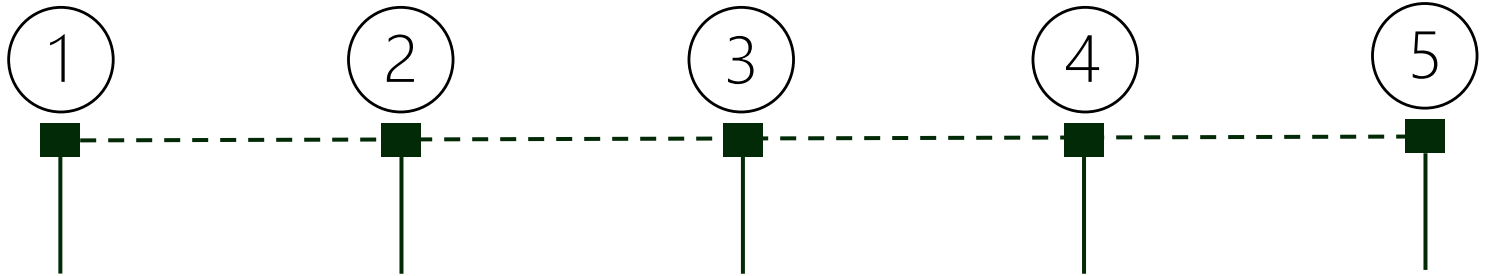
Regression Modeling

Linear Regression

Random Forest

Gradient Boosting

Approach for segmenting customers into actionable groups



Churn Risk

Estimate Churn Risk using Churn Classification Model

Margin Value

Estimate Future Margin Value using Margin Regression Model

Lifetime Value

Calculate Customer Lifetime Value (CFV)

Segmentation

K-Means clustering based on Churn Risk, Estimated Margin, and CFV Value

Profiling

Profile using sales and demographic metrics