ISSS621 – Data Science for Business

IMPROVING CUSTOMER LIFETIME VALUE

A case study of a European bank.

Group G1 - 3
Alicia YEE Mong Shin
CHEN Liuying
Cheryl YONG Li Ru
Federico Jose Sabile RODRIGUEZ
HUANG Lan
Sheryl Ann TAN Yi-Shi

AGENDA

01 Introduction

02 Dataset and Business Objectives

03 Recommendations: Customer Retention – Reducing Churn Rate

04 Recommendations: Customer Expansion – Increasing Number of Products Sold

05 Recommendations: Customer Expansion – Increasing Credit Card Spend

06 Closed Loop Data Eco-System

07 Data Governance

08 Conclusion

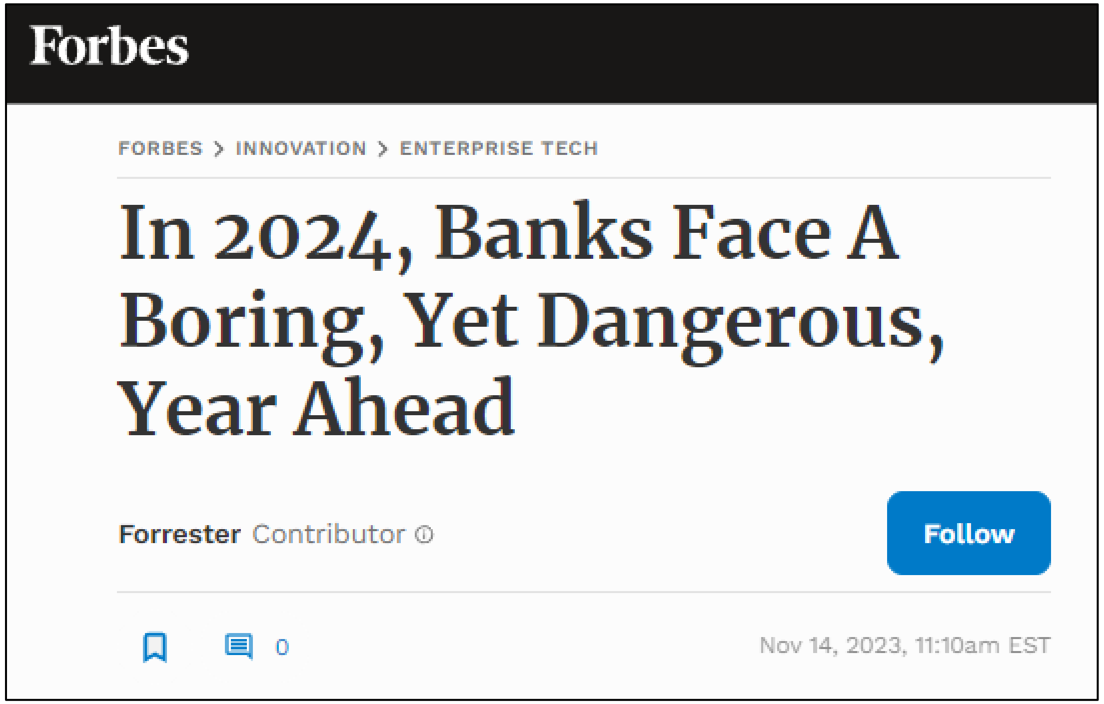
INTRODUCTION

European bank facing difficulties in ensuring its sustainability.

Tasked to create a solution based on data science thinking.

INTRODUCTION

AN OVERVIEW OF THE BANKING INDUSTRY



Forbes

FORBES > INNOVATION > ENTERPRISE TECH

In 2024, Banks Face A Boring, Yet Dangerous, Year Ahead

Forrester Contributor

Follow

0

Nov 14, 2023, 11:10am EST

Banks face challenges in achieving **sustained and profitable growth** due to:

1. **High Competition.**
   Traditional banks, credit unions and fintech companies offer similar services. Forbes highlighted the importance of designing products and experiences that would have a massive impact on customers.

2. **Switching Costs Have Lowered.**
   It is easier for customers to switch banks. Hence, banks will need to think of how they demonstrate value to customers if they want to grow deposits and drive loyalty.

How to ensure long-term sustainability?

WHAT IS CUSTOMER LIFETIME VALUE?

CUSTOMER LIFETIME VALUE

A measure of the total revenue generated by a customer over the entire course of their relationship with a business.

Importance of CLV to Banks

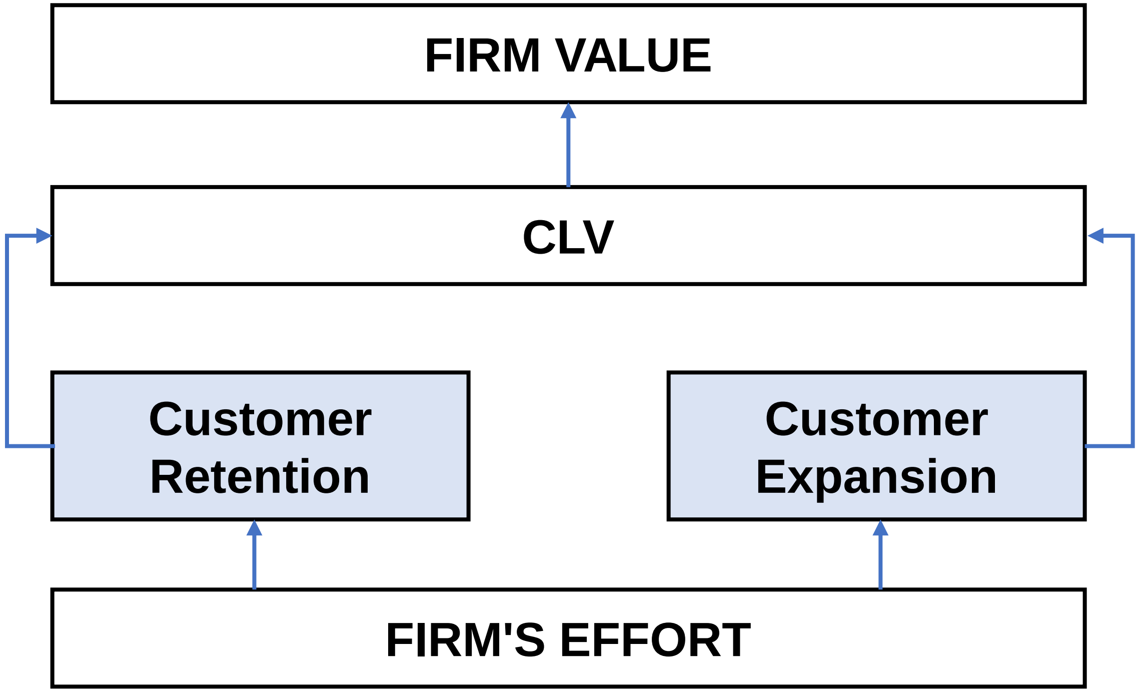
Predicts Long-Term Profitability
Allows banks to determine which customers are likely to be the most profitable, therefore guiding resource allocation.

Improves ROI of Targeted Strategies
Allows banks to tailor marketing and customer service efforts more effectively, thus maximizing ROI.

Ensures Financial Sustainability
Repeated purchases from existing customers ensures a steady cash inflow for banks, which can be invested to ensure growth.

INTRODUCTION

KEY DRIVERS OF CUSTOMER LIFETIME VALUE



CLV Conceptual Framework by Gupta et al. (2006):

Emphasizes the importance of several strategies in managing customer relationships and optimizing CLV.

KEY DRIVERS OF CUSTOMER LIFETIME VALUE

CUSTOMER RETENTION

Increasing the customer retention would increase CLV as customers would make repeated purchases over a longer duration.

CUSTOMER EXPANSION

- Involves increasing the value of customer through **up-selling, cross-selling, enhancing the purchase frequency or amount spent.**
- This would increase CLV due to the increase in average revenue generated for each customer.

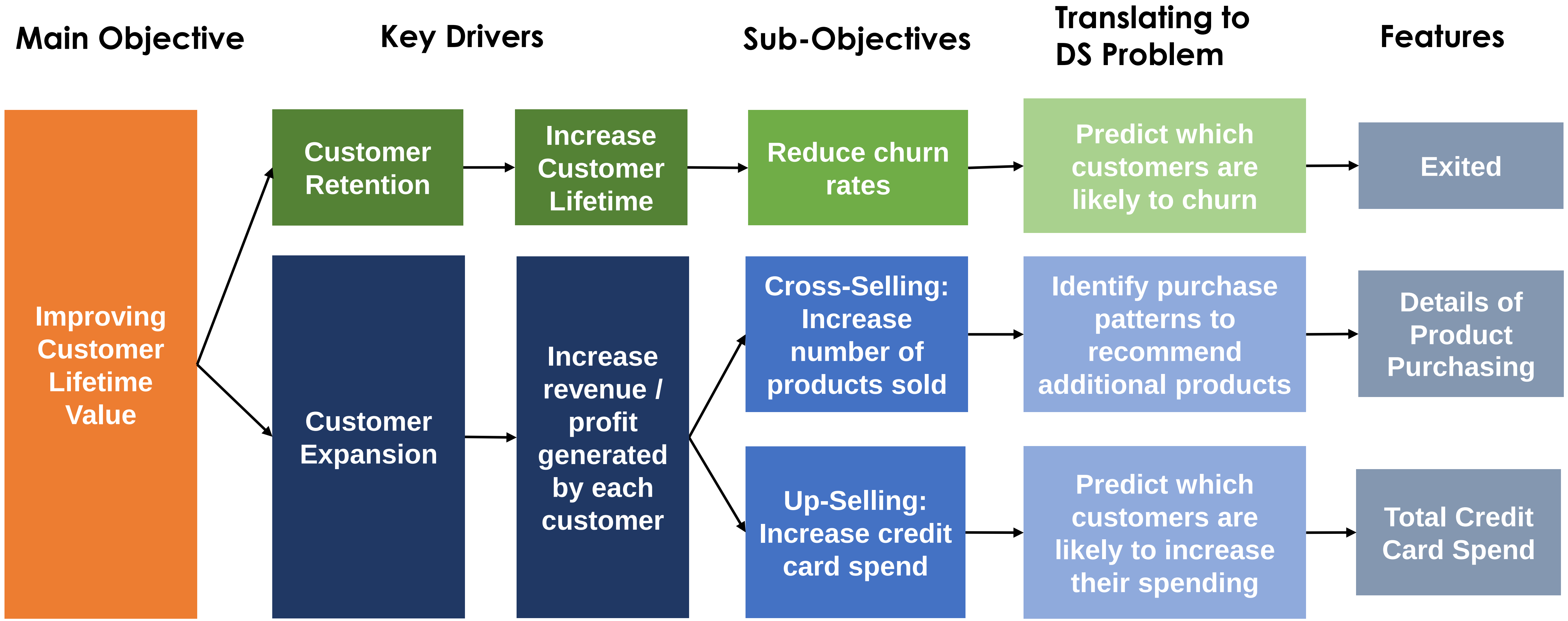
DATASET | # OVERVIEW OF DATASET

- **10,000** rows and **39** columns

Classification	Feature Names
CLV Metrics	Exited (Churn), Number of products purchased, Purchasing details of 8 products, Total credit card expenditure
Demographics	Geography, Gender, Age, Education, Marital Status, Job, Estimated Income
Customer Behaviors	Tenure, Referral code, Net banking, Most frequent interactions with the bank, Number of credit card transactions (per month), Point Earned, Card limit, Credit score, cc_balance, Purchase, One off purchase, Installments purchases, Cash advance, PCT purchase, PCT one off purchase, PCT instalments, PCT cash advance, Cash advance TRX, Purchase TRX, Credit limits, Total_cc_expenditure, Payments, Minimum payments, Min. payment excess shortfall, Complain, Complain products, Satisfaction Score

DATASET

LINKING BACK TO THE BUSINESS OBJECTIVE



DATA PREPROCESSING

Feature Engineering

- **Customer Value**
    - BrokerageAccounts*5000 + TrustFunds*2000 + MutualFunds*6000 + RetirementAccounts*3000 + PropertyInsurance*2000 + LifeInsurance*1000 + HealthInsurance*1000 + Annuities*4000 + CreditCardSpending*12*12%

RECOMMENDATIONS

CUSTOMER RETENTION: REDUCE CHURN RATES

Naïve Approach	Heuristic Approach	Data Science Approach
Apply to All Strategy Giving Rebates to all customers who have churned, therefore potentially aiming to increase conversions.	**Segmented Targeting** Visually segment the demographics of customers who have churned. Target specific segments with cash rebates.	**Machine Learning Models** Create machine learning models that leverage on historical data to identify patterns and trends to predict churn, leading to more accurate and targeted interventions.

RECOMMENDATIONS

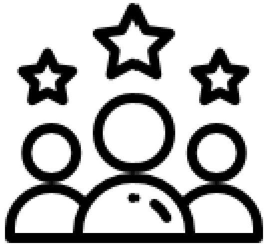
CUSTOMER RETENTION: REDUCE CHURN RATES

Naïve Approach – Apply to All Strategy

Giving Rebates to all customers who have churned, therefore potentially aiming to increase conversions.

Resources wasted on customers who will not stay with the bank (no conversion)

Budget Constrains: Cost maybe be high over a period of years



Without data-driven insights, opportunities to target high-risk segments are missed.

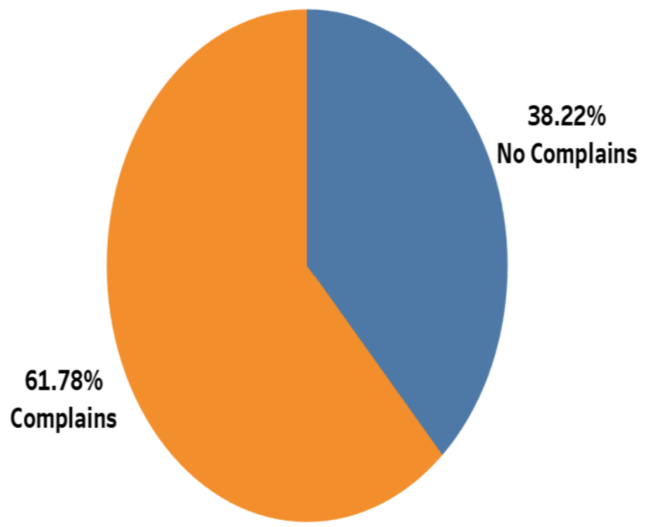
RECOMMENDATIONS

CUSTOMER RETENTION: REDUCE CHURN RATES

Heuristic Approach – Segmented Targeting

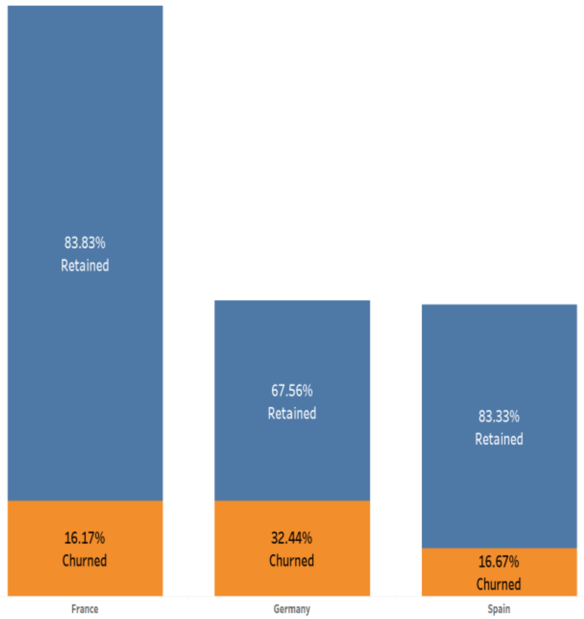
- Visually segment the demographics of customers who have churned.
- Target specific segments with cash rebates.
- Examples: Immediately target customers who have lodged a complaint or countries with higher churned rates.
- **Shortfall**: Not easily scalable. As the customer base grows, it becomes increasingly difficult to maintain accurate and effective segmentation, reducing the overall efficiency and effectiveness of business strategies.

% of Complaints per Churned Case



61.78% of churned cases were preceded by a lodged complaint.

Country Distribution



Germany had the highest portion of the customer churned (32%).

RECOMMENDATIONS

CUSTOMER RETENTION: REDUCING CHURN RATES

Data Science Approach : Logistic Regression Model

1.Data Processing

One hot encoding of categorical variables

2.Feature Selection

Drop correlated/unrelated variables

3. Feature Engineering

SMOTE to handle imbalance classes

4. Modelling

Logistic regression model (split 70% train and 30% test data)

5. Model evaluation

Accuracy, precision, F1 score, recall and confusion matrix

Leverages historical data to identify patterns and trends that predict churn, leading to more accurate and targeted interventions.

Highly scalable and can handle large datasets efficiently

```
Accuracy: 0.8206533192834563
Precision: 0.8186506231198969
Recall: 0.8161953727506427
F1 Score: 0.8174211542587427
Confusion Matrix:
[[1989  422]
 [ 429 1905]]
Classification Report:
              precision    recall  f1-score   support

           0       0.82      0.82      0.82      2411
           1       0.82      0.82      0.82      2334

    accuracy                           0.82      4745
   macro avg       0.82      0.82      0.82      4745
weighted avg       0.82      0.82      0.82      4745
```

Top Features with high coefficients
1.Complain Product – Retail Banking
2.Most frequent interactions with the bank
3.Fraud transaction detection
4.Referral Code
5.Satisfaction Score

RECOMMENDATIONS

CUSTOMER RETENTION: REDUCE CHURN RATES

Business Strategy – Improve retail banking services

Improve the retail banking services by improving the quality of the retail branches

Improve Comfort Quality: Enhance the visual appeal and functionality of bank branches.

Improve Service Quality : Provide frontline employees with the training and authority needed to swiftly address frequent complaints.

RECOMMENDATIONS

CUSTOMER RETENTION: REDUCING CHURN RATES

Quantifying our solutions – Approach and Assumptions

Naïve Approach

- **Costs**
  - Rebates of $300 would be given to all churn customers
- **Revenue**
  - Assumption: If the rebate given is more than 10% of their monthly estimated income, they will be converted to stay.
  - Revenue would be the sum of the customer value.
- **Profit (for 1 Year)**
  - Revenue – Costs
  - $8M

Heuristic Approach

- **Costs**
  - Rebates of $300 would be given to churn customers who lodged a complain or are from Germany
- **Revenue**
  - Assumption: If the rebate give is more than 10% of their monthly estimated income, they will be converted to stay.
  - Revenue would be the sum of the customer value.
- **Profit (for 1 Year)**
  - Revenue – Costs
  - $7.4M

Data Science Approach (Highest profit)

- **Costs**
  - Assumption: Cost of enhancing retail banking would be a one time fixed cost of $3M.
- **Revenue**
  - Assumption: 40% of the customers who churned would be enticed to stay
- **Profit (for 1 Year)**
  - Revenue – Costs
  - $11M

CUSTOMER RETENTION: REDUCING CHURN RATES

Quantifying our solutions

Comparing the methods

Total number of customers who churned: 2093

	Approach	Customers_Stayed	Cost	Revenue	Profit	5-Year Cost	5-Year Revenue	5-Year Profit
0	Data Science	580	3000000	14033316.64	11033316.64	3000000	70166583.2	55166583.2
1	Heuristics	340	559800	8035107.04	7475307.04	2799000	40175535.2	37376535.2
2	Naive	366	627900	8627610.56	7999710.56	3139500	43138052.8	39998552.8

Short Run (1 Year):
Data Science approach cost the most, but still has the highest revenue and profit
Long Run (5 Years):
Data Science approach does not incur additional cash rebates cost unlike heuristics and naïve approach.
Revenue and profit of data science approach in the long run is significantly higher.
Conclusion: Data Science would be the best approach

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Naïve Approach

- **General Cross-Selling**
  - Offer the most popular products or upgrades to all customers by every point of contact.
- **Shortfall**
  - Doesn’t consider the customers’ needs -> low take-up and conversion rate.
  - Budget constraints due to indiscriminate targeting.

Heuristic Approach

- **Rule-Based Cross-Selling:**
  - Pinpoint customers who have less than 4 products, and cross-sell products that they do not own from the database.
  - Correlation matrix -> cross-sell based on basic patterns and relationships between products.
- **Shortfall**
  - Lack of depth -> Correlation matrix misses non-linear patterns.

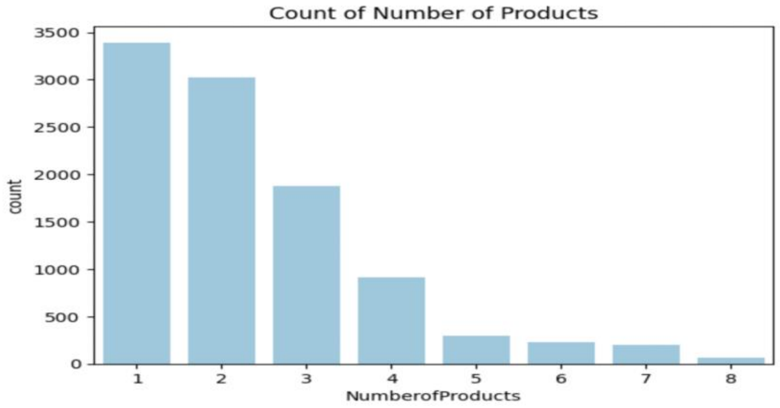
Data Science Approach

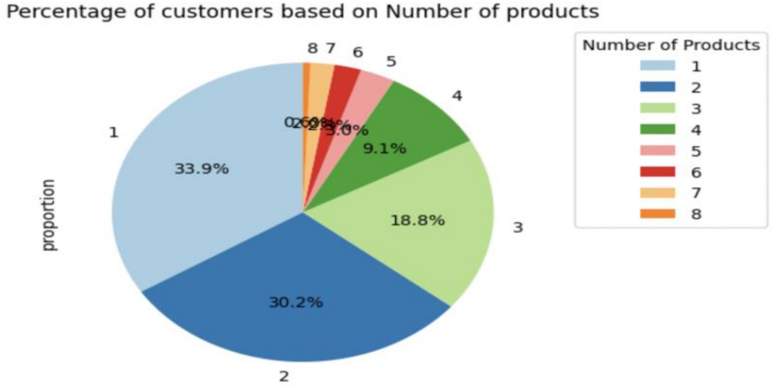
- **Association Rule Mining**
  - Identify frequent itemsets an-d strong association rules.
    - Apriori Algorithm
- **Outcome**
  - Recommendation system
    - Enhance the bank's digital banking platform
    - Additional feature to the banking's webpage
    - Personalised browser -> Recommendations based on the products the customer do not have based on product associations.

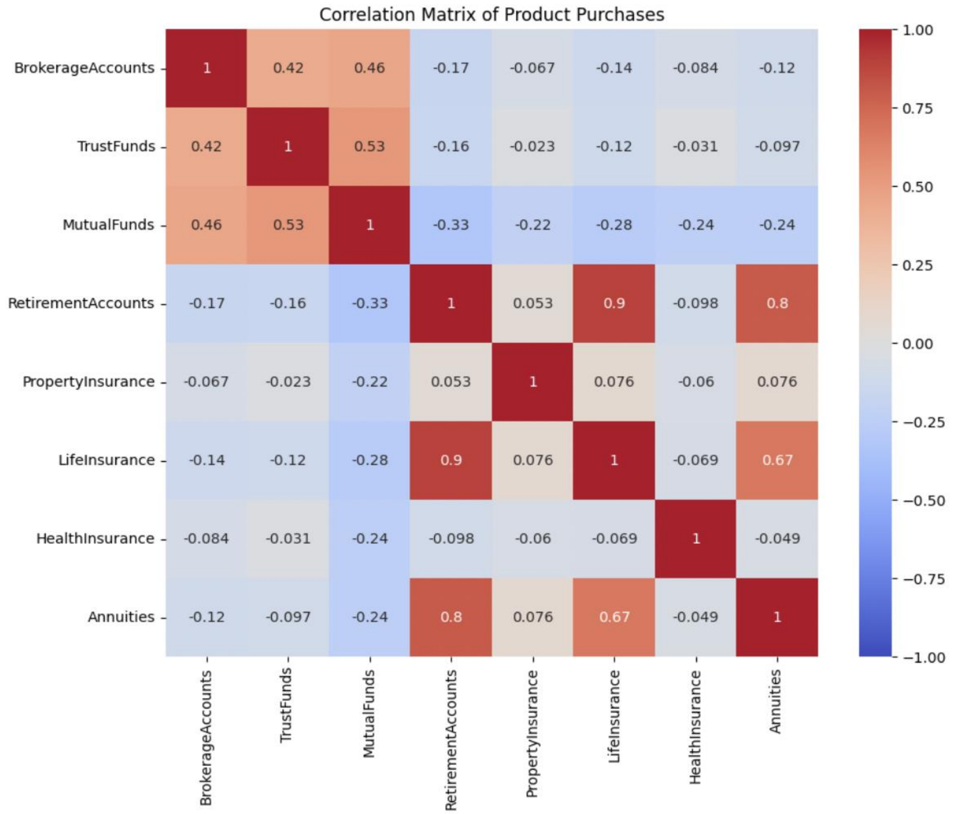
RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Proportions and No. of Products per Customer | Relationship between Products







Univariate Analysis:
- More than 80% customers have less than 4 products.
- Highest proportion of customers (33.9%) have only 1 product.
- Higher proportion of customers belonging to less than 4 products -> increase product sales.

Multivariate Analysis:
- Brokerage Accounts, Mutual Funds and Trust Funds are **moderately** correlated.
- Retirement Accounts, Life Insurance and Annuities are **highly** correlated.

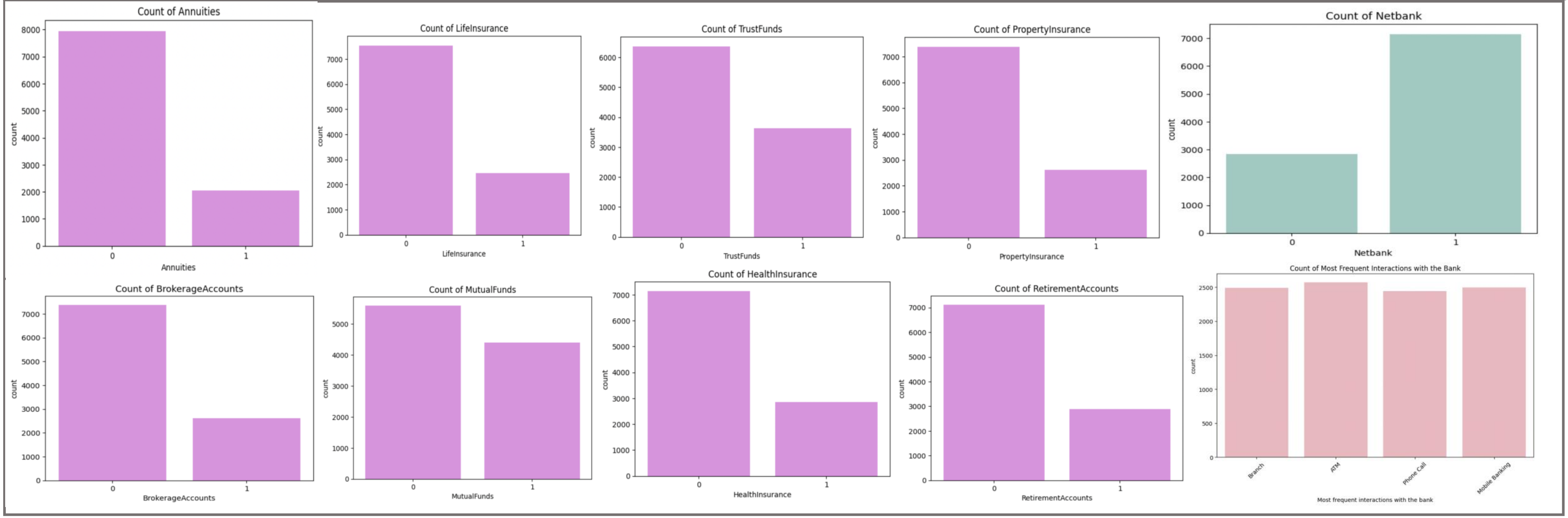
Recommendations based on Heuristic Approach:
1. **Target customers with less than 4 products**
2. **Cross-sell products frequently purchased together**

Can we do better?
E.g. What can we do to *specifically analyze the associations?*

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Frequency of customers based on each product | Platform for banking



1. For all financial products listed, the number of customers who do not have the product is higher than those who do -> there is a significant potential market for cross-selling all financial products to existing customers.
2. **Netbanking & Mobile banking** (large segment of customers) -> webpage good platform for recommendation systems

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Process of Association Mining (Apriori Algorithm)

Step 1:
Generate Frequent Itemsets
How frequent is the item/itemset?

Support

1. Generate 2- itemsets; support count >= **minsup = 0.01**
2. Candidate Generation; create 3-itemsets from merging)
3. Candidate Pruning; removes itemsets < minsup)
4. Iterative Process

$$\text{Lift}(\{Milk, Diaper\} \rightarrow \{Beer\}) = \frac{\text{Support}(\{Milk, Diaper, Beer\})}{\text{Support}(\{Milk, Diaper\}) \times \text{Support}(\{Beer\})}$$

Figure: Lift calculation formula

Step 2:
Rule Generation
Strength of association

Confidence

Generate high confidence rules and prune itemsets (confidence < **minconf = 0.7**)

Lift (Post-processing metric)

Calculate lift values, prune itemsets, lift< **minlift = 1**:
e.g Lift > 1, A and B are positively correlated. Namely, the 2 item sets lift the likelihood of each another (positively correlated)

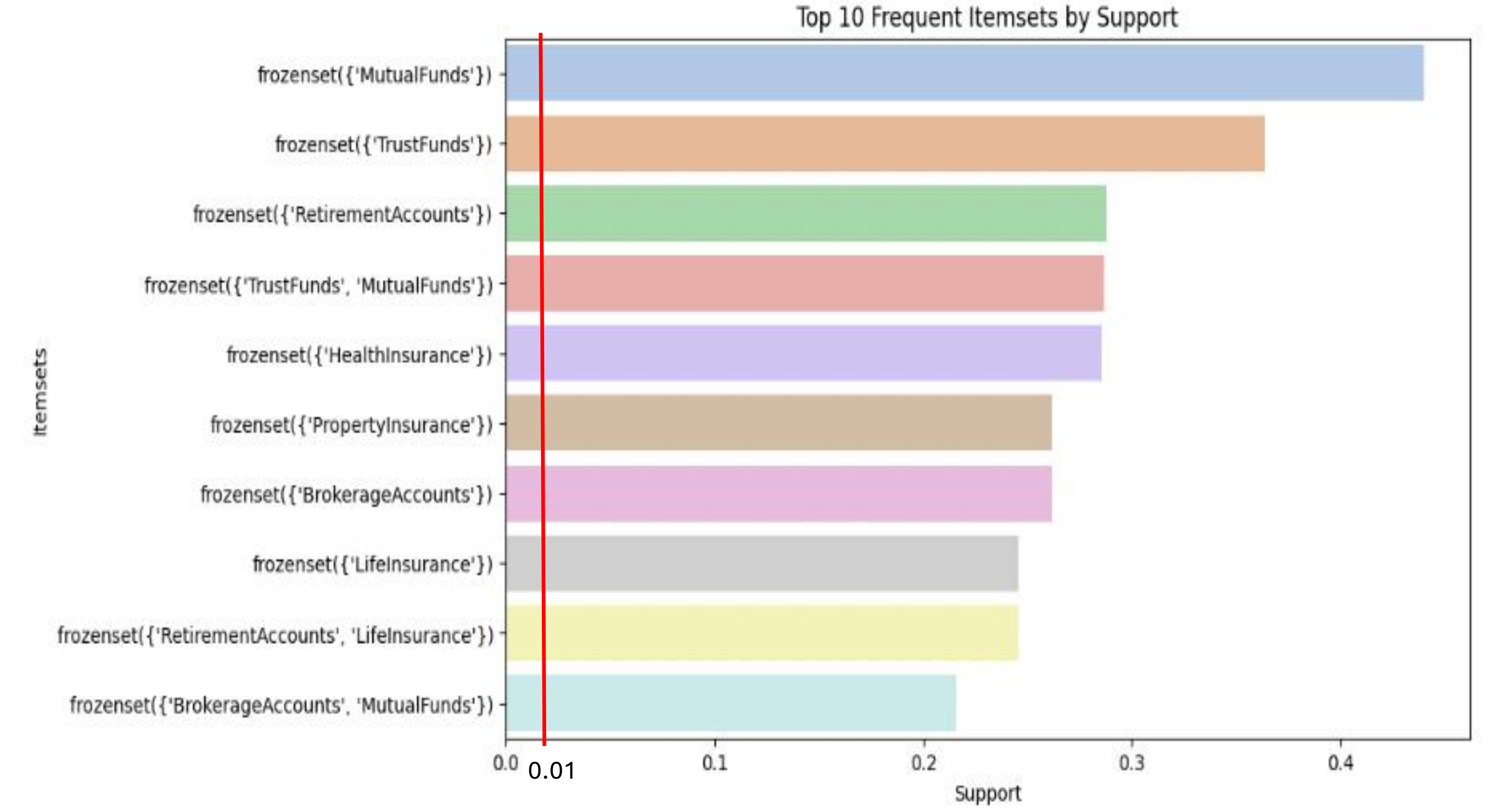
```
Top 10 Strong Associations:
Rule: ['LifeInsurance'] -> ['RetirementAccounts']
Support: 0.2456, Confidence: 1.0000, Lift: 3.4746

Rule: ['Annuities'] -> ['RetirementAccounts']
Support: 0.2055, Confidence: 1.0000, Lift: 3.4746

Rule: ['RetirementAccounts'] -> ['LifeInsurance']
Support: 0.2456, Confidence: 0.8534, Lift: 3.4746

Rule: ['RetirementAccounts'] -> ['Annuities']
Support: 0.2055, Confidence: 0.7140, Lift: 3.4746

Rule: ['BrokerageAccounts'] -> ['MutualFunds']
Support: 0.2155, Confidence: 0.8244, Lift: 1.8737

Rule: ['TrustFunds'] -> ['MutualFunds']
Support: 0.2864, Confidence: 0.7879, Lift: 1.7907
```

Figure: Association Mining Results

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Results of Associative Mining (Frequent Itemsets; Support)



Results:

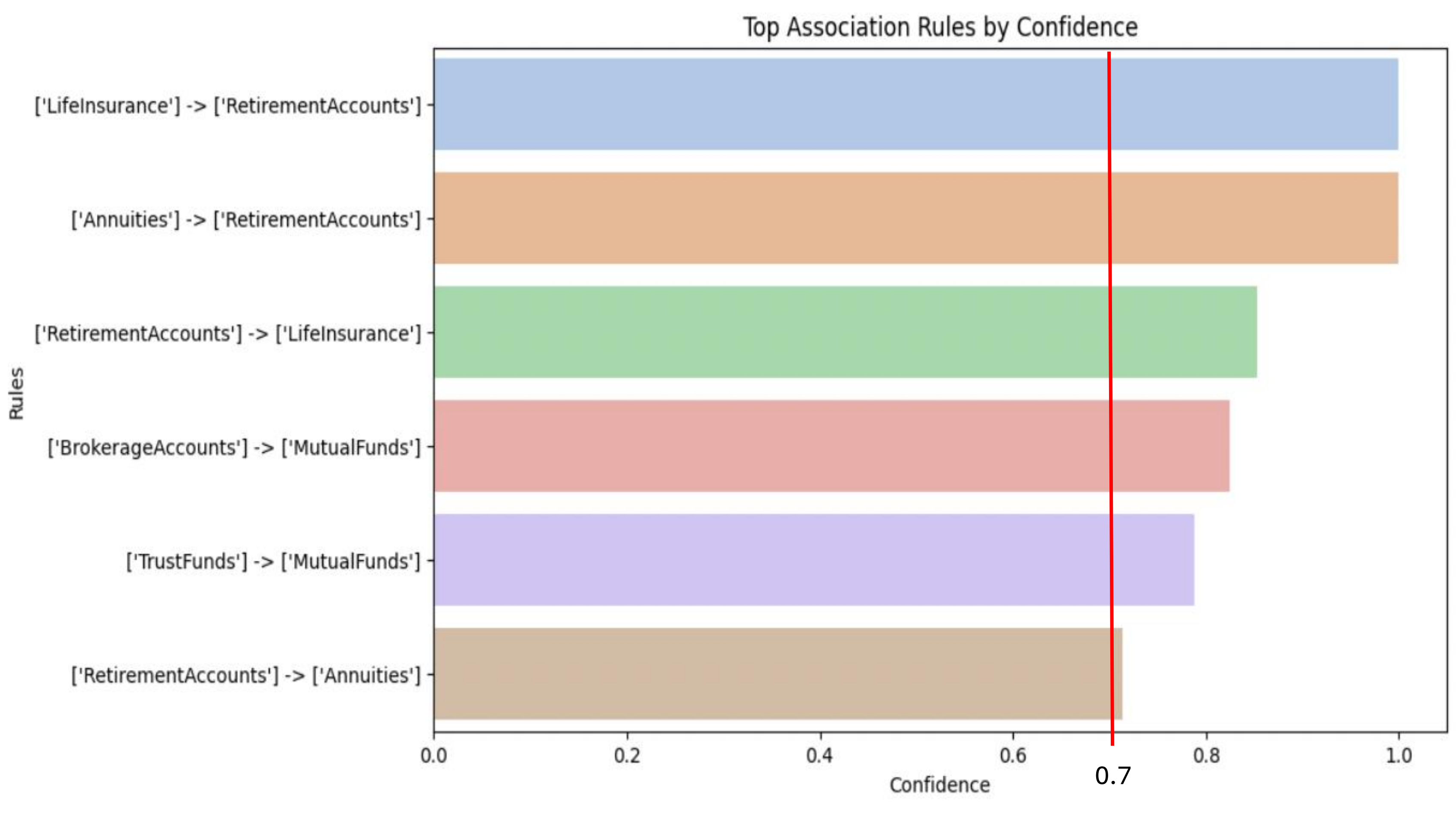
- The itemset **{'MutualFunds'}** has the highest support, meaning it appears in the largest proportion of transactions.
- Highly frequent 2-itemsets: **{'TrustFunds', 'MutualFunds'}, {'RetirementAccounts', 'LifeInsurance'}, {'BrokerageAccounts','LifeInsurance'}**
- These itemsets are common across transactions, indicating popular combinations of items.
- These itemsets are proceeded to be evaluated on strength of assocaitions based on forming rules.

What are the strength of associations in the itemsets?

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Results of Associative Mining (Confidence)



Results:

- Rule **{'LifeInsurance'} -> {'RetirementAccounts'}** has the highest confidence.

When {LifeInsurance} is bought, {RetirementAccounts} is also bought with high probability.

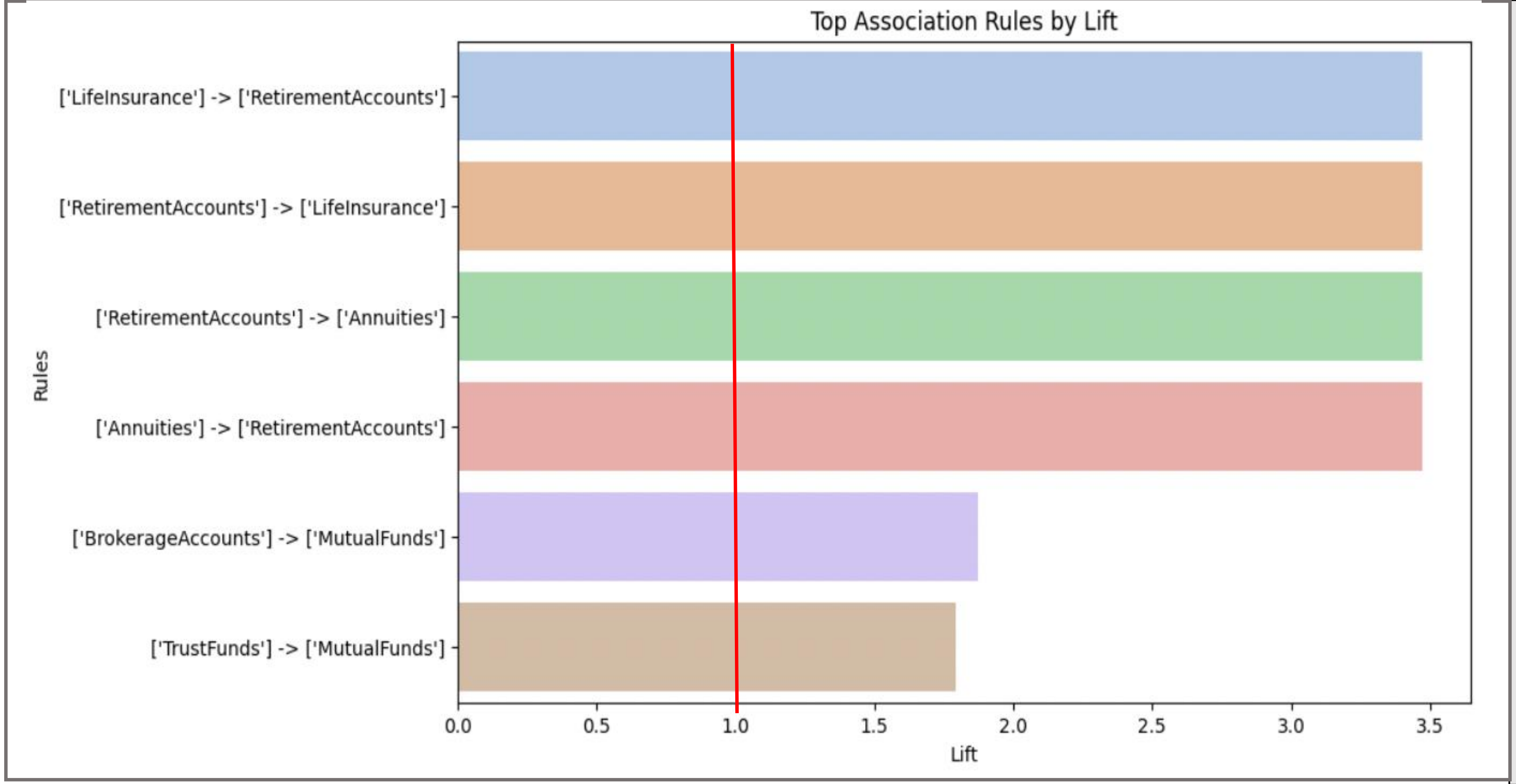
- Rule {'Annuities'} -> {'RetirementAccounts'} high confidence.
- Other notable rules include {'RetirementAccounts'} -> {'LifeInsurance'} and {'BrokerageAccounts'} -> {'MutualFunds'}.

High confidence indicates strong predictive relationships between the antecedent and consequent itemsets.

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Results of Associative Mining (Lift)



Results:

- The rule **{'RetirementAccounts'} -> {'LifeInsurance'}** has the highest lift, indicating that these two items are much more likely to be bought together than expected by chance.
- The rule **{'LifeInsurance'} -> {'RetirementAccounts'}** has a high lift value.
- Other notable rules include **{'RetirementAccounts'} -> {'Annuities'}** and **{'Annuities'} -> {'RetirementAccounts'}.**

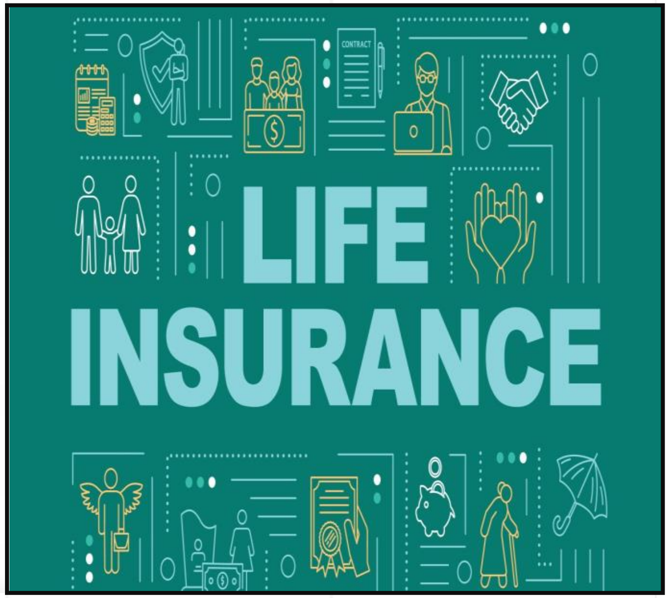
RECOMMENDATIONS | CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Example of Recommendation System from DS Approach

Related to items you've viewed See more

Recommended for you, Thomas

Annuities (*Click to see more)



Life Insurance (*Click to see more)

Data Science Goal:
Create a recommendation system as an additional feature for the personalized website for our customers.

Strongly associated products:

1. **Retirement Accounts, Annuities**
2. **Retirement Accounts, Life In surance**
3. **Brokerage Accounts, Mutual Funds**
4. **Mutual Funds, Trust Funds**

RECOMMENDATIONS

CUSTOMER EXPANSION: IMPROVE CROSS-SELLING

Quantifying our solutions

Naïve Approach

- **Costs**
  - Social Media Advertising,
  - Print Advertisements (e.g newspaper, posters)
  - Roadshows
  - Assumption of $24,000/Yr
- **Revenue**
  - Assumption of 30% conversion rate for our **popular products** (Trust and Mutual Funds)
  - Assumed Revenue/Yr for each product
- **Profit (for 2 Years)**
  - Revenue – Costs
  - $168,430,000

	BrokerageAccounts	TrustFunds	MutualFunds	RetirementAccounts	PropertyInsurance	LifeInsurance	HealthInsurance	Annuities	Total Products Purchased
Purchased	2614	3635	4400	2878	2617	2456	2854	2055	23509
Did not purchase	7386	6365	5600	7122	7383	7544	7146	7945	NA
After Naïve approach									
Purchased	2614	4726	5720	2878	2617	2456	2854	2055	25920
		Increased 30%							
Total Costs	$ 48,000.00								
Total Revenue	$ 168,478,000.00								
Profit	$ 168,430,000.00								

Heuristic Approach

- **Costs**
  - Email Campaigns,
  - Phone call
  - Text advertisements
  - Assumption of $12,000/Yr
- **Revenue**
  - Assumption of 50% conversion rate from products **moderately, strongly correlated** with each other
  - Assumed Revenue/Yr for each product
- **Profit (for 2 Years)**
  - Revenue – Costs
  - $180,713,000

	BrokerageAccounts	TrustFunds	MutualFunds	RetirementAccounts	PropertyInsurance	LifeInsurance	HealthInsurance	Annuities	Total Products Purchased
Purchased	1718	2424	3371	1691	1567	1342	1939	1028	15080
Not Purchased	6578	5872	4925	6605	6729	6954	6357	7268	NA
After Heuristic Approach									
Purchased	1718	3636	5057	1690	1567	2013	2909	1542	20131
		Increased 50%					Increased 50%		
Overall									
Total purchased (original data)	2614	3635	4400	2878	2617	2456	2854	2055	23509
Customers with 4 or more products	896	1211	1029	1187	1050	1114	915	1027	8429
Total purchased (after Heuristic approach)	2614	4847	6086	2877	2617	3127	3824	2569	28560
Total Costs	$ 24,000.00								
Total Revenue	$ 180,737,000.00								
Profit	$ 180,713,000.00								

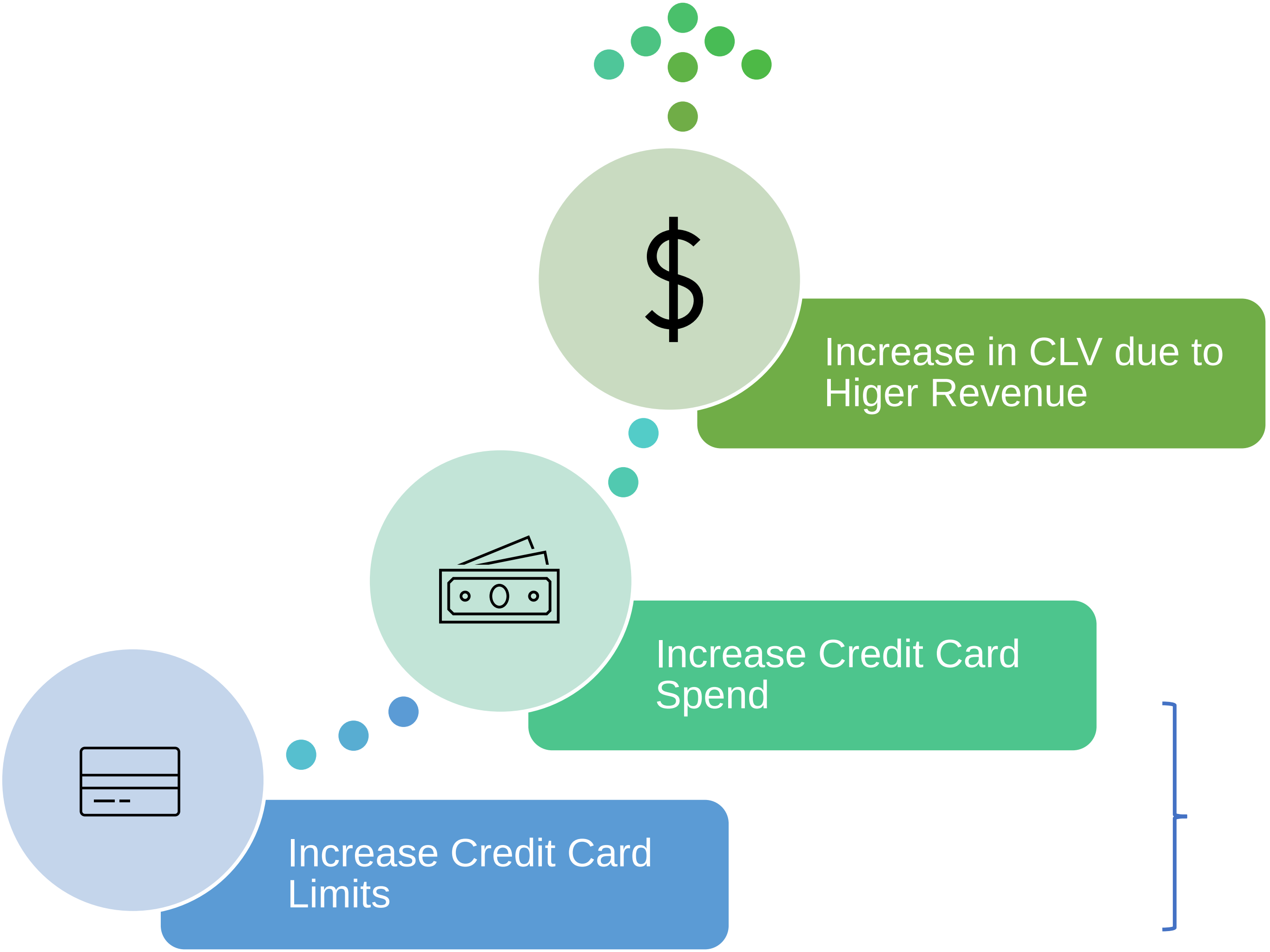
Data Science Approach (Highest profit)

- **Costs**
  - Assumption of $50,000 one-time payment for recommendation system creation
  - Assumption of $10,000/Yr maintenance costs
- **Revenue**
  - Assumption of 70% conversion rate from products **strongly associated** with each other
  - Assumed Revenue/Yr for each product
- **Profit (for 2 Years)**
  - Revenue – Costs
  - $240,676,000

	BrokerageAccounts	TrustFunds	MutualFunds	RetirementAccounts	PropertyInsurance	LifeInsurance	HealthInsurance	Annuities	Total Products Purchased
Purchased	2614	3635	4400	2878	2617	2456	2854	2055	23509
Did not purchase	7386	6365	5600	7122	7383	7544	7146	7945	NA
After DS Approach									
Purchased	4444	6180	7480	4893	2617	4175	2854	3494	36136
		Increased 70%				Increased 70%		Increased 70%	
Total Costs	$ 70,000.00								
Total Revenue	$ 240,746,000.00								
Profit	$ 240,676,000.00								

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND



Increase in credit card spending would increase revenue for a bank due to:

1. **Higher Transaction Fees**: Every time a customer uses a credit card, the merchant pays a fee to the bank that issued the card.
2. **Annual Fees**: Banks would earn from annual fees from credit cards.

The bank is considering **increasing credit card limits by 30%** to incentivize credit card expenditure. Providing more credit for customers would enhance a customer's purchasing power.

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND

01

- **Task**: Increase credit limits by 30% to incentivize customer spending.

02

- **Naïve Approach**: Offer 30% increase on credit limits to all customers.
- This serves as a baseline for comparison and provides a general understanding on the general impact of credit limit increases on spending.

Measuring the Effectiveness of Approach:

$$\Delta \text{ Total Credit Card Expenditure} = \textit{Predicted Credit Card Expenditure } \underline{\textit{after}} \textit{ Credit Limit Increase} - \textit{Predicted Credit Card Expenditure } \underline{\textit{before}} \textit{ Credit Limit Increase}$$

Impact of 30% Increase in Credit Limits Across All Customers:

Increase in Credit Card Expenditure	$478,693
% Increase	6.88%

HOWEVER:

	CustomerId	Expenditure Before Increase	Expenditure After Increase	Absolute Increase in Expenditure
0	6252	3317.250842	4389.054077	1071.803235
1	4684	2118.813041	3012.894168	894.081127
2	1731	713.245337	331.938466	-381.306871
3	4742	2162.750312	4501.434224	2338.683912
4	4521	2014.386387	1980.829210	-33.557177

Some customers would decrease their credit card expenditure despite the credit limit increase. We would need to exclude these customers when offering the 30% credit limit increase.

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND

01

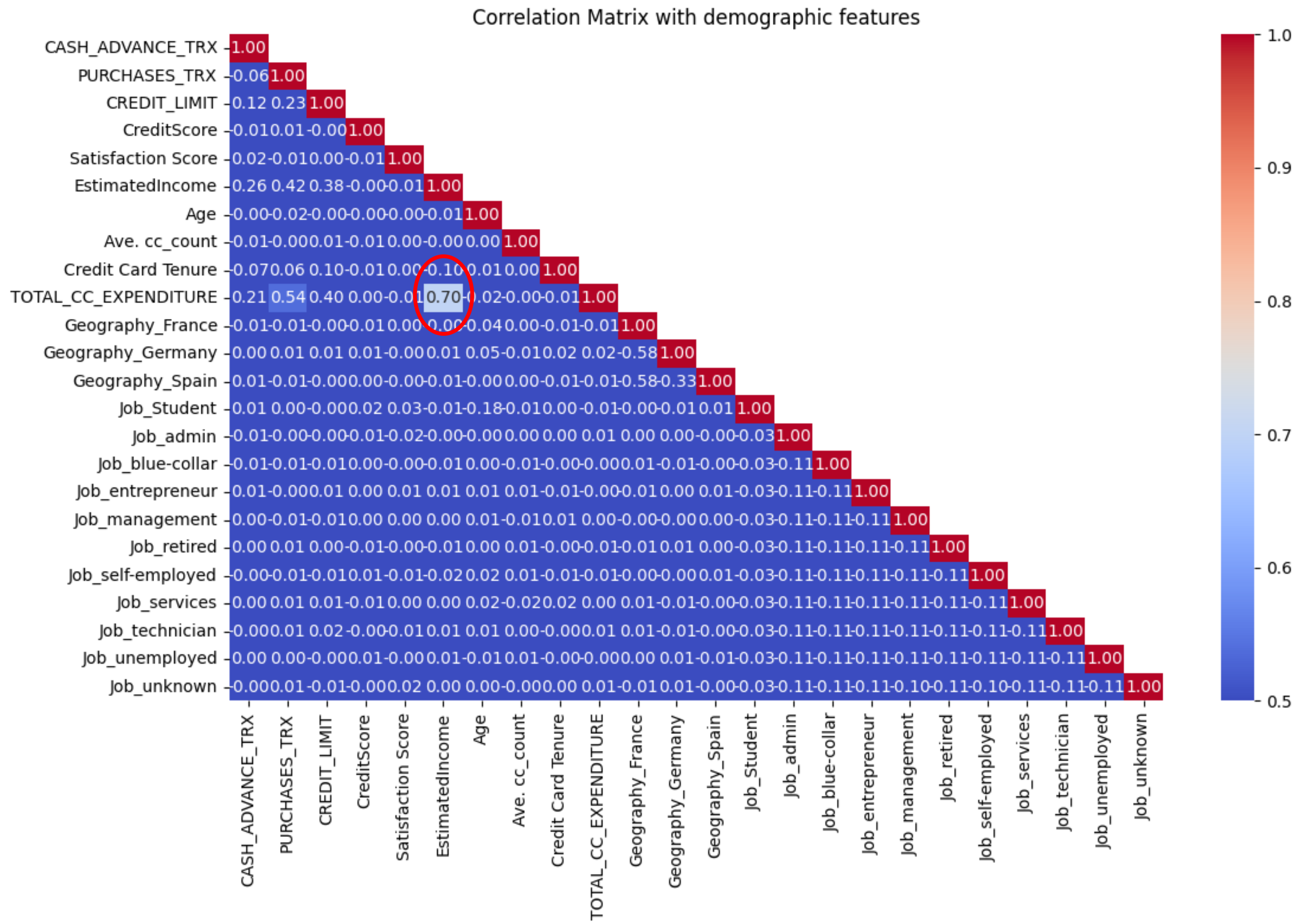
- **Task**: Increase credit limits by 30% to incentivize customer spending.

02

- **Naïve Approach**: Offer 30% increase on credit limits to <u>all</u> customers.
- This serves as a baseline for comparison and provides a general understanding on the general impact of credit limit increases on spending.
- Some customers may have a decrease in credit card expenditure after the credit limit increase.

03

- **Heuristic Approach**: Offer 30% increase on credit limits to selected customers.
- Visually inspect the relationship between selected features and credit card expenditure and select the feature that has the strongest relationship as a criteria to select customers.



There is a strong positive correlation between Estimated Income and Credit Card Expenditure, implying a relationship between a customer's estimated income and their credit card expenditure.

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND

01

- **Task**: Increase credit limits by 30% to incentivize customer spending.

02

- **Naïve Approach**: Offer 30% increase on credit limits to all customers.
- This serves as a baseline for comparison and provides a general understanding on the general impact of credit limit increases on spending.
- Some customers may have a decrease in credit card expenditure after the credit limit increase.

03

- **Heuristic Approach**: Offer 30% increase on credit limits to selected customers.
- Visually inspect the relationship between selected features and credit card expenditure and select the feature that has the strongest relationship as a criteria to select customers.

- Given the correlation, we can consider offering the increase in credit limits to high income customers.
- **Proposal:** Offer the 30% increase in credit limits to customers whose estimated income is above the 50th percentile.

Impact of 30% Increase in Credit Limits for High Income Customers:

Increase in Credit Card Expenditure	$562,807
% Increase	9.52%

HOWEVER:

Expenditure Before Increase	Expenditure After Increase	Absolute Increase in Expenditure
6866.205645	8587.753125	1721.547480
3454.301651	2870.775011	-583.526640
5922.920000	6091.665828	168.745828
5188.114999	6970.780124	1782.665125
4001.531566	7314.617873	3313.086307

Some customers are still predicted to have a decrease in credit card expenditure.

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND

01
- **Task**: Increase credit limits by 30% to incentivize customer spending.

02
- **Naïve Approach**: Offer 30% increase on credit limits to <u>all</u> customers.
- Some customers may have a decrease in credit card expenditure after the credit limit increase.

03
- **Heuristic Approach**: Offer 30% increase on credit limits to selected customers.
- More targeted approach would increase total credit card expenditure. However, there are several customers that may decrease their credit card expenditure.

04
- **Data Science Approach**: Build a predictive model (regression model) that would predict the change in customer's CC expenditure behaviour towards credit limit changes.
- Offer the 30% increase based on their predicted behaviour.

- From the Naïve and Heuristic Approaches, we can see that there may be complex interactions in predicting a customer's credit card expenditure.

- Hence, we propose to model their behaviour through a regression model, which helps to identify and quantify the relationships between credit card expenditure and various predictors as well as capture these complex interactions.

- **Y:** Total Credit Card Expenditure
- **X:** Predictors involving credit card spending behaviour such as number of purchase transactions, number of cash advances transactions, credit score, etc.

- **Proposal:** If the customer is expected to increase their credit card expenditure, we would offer the 30% credit limit increase. Otherwise, no credit limit increase would be offered.

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND

01

- **Task**: Increase credit limits by 30% to incentivize customer spending.

02

- **Naïve Approach**: Offer 30% increase on credit limits to all customers.
- Some customers may have a decrease in credit card expenditure after the credit limit increase.

03

- **Heuristic Approach**: Offer 30% increase on credit limits to selected customers.
- More targeted approach would increase total credit card expenditure. However, there are several customers that may decrease their credit card expenditure.

04

- **Data Science Approach**: Build a predictive model (regression model) that would predict the change in customer's CC expenditure behaviour towards credit limit changes.
- Offer the 30% increase based on their predicted behaviour.

Impact of 30% Increase in Credit Limits Based on Customer Credit Card Expenditure Behaviour

Increase in Credit Card Expenditure	$1,666,238
% Increase	23.94%

- When compared to the results from the Naïve and Heuristic Approaches, the Data Science Approach would result a higher increase in credit card expenditure.
- Unlike the Naïve and Heuristic Approaches, this model would specifically exclude customers that are predicted to have a decrease in credit card expenditure after a credit limit increase.

	Approach	Total Increase in Expenditure	Percentage Increase in Expenditure
0	Naive	4.786936e+05	6.879128
1	Heuristic	5.628076e+05	9.521975
2	Data Science	1.669239e+06	23.988014

RECOMMENDATIONS

CUSTOMER EXPANSION: INCREASE CREDIT CARD SPEND

Naïve Approach

- Offer a 30% increase in credit limits to all customers, regardless of their background and purchasing behaviour.
- **Shortfall**: This approach does not take into account complex interactions in variables affected credit card expenditure. This would result in certain customers decreasing their overall credit card expenditure after the credit limit increase.

Heuristic Approach

- Based on the correlation matrix, income is positively correlated to credit card expenditure. Hence, we should adopt a more targeted approach and offer the 30% credit limit increase to customers with higher income.
- **Shortfall**: Simplified approaches may miss opportunities for optimization and improvement that more comprehensive methods could identify.

Data Science Approach

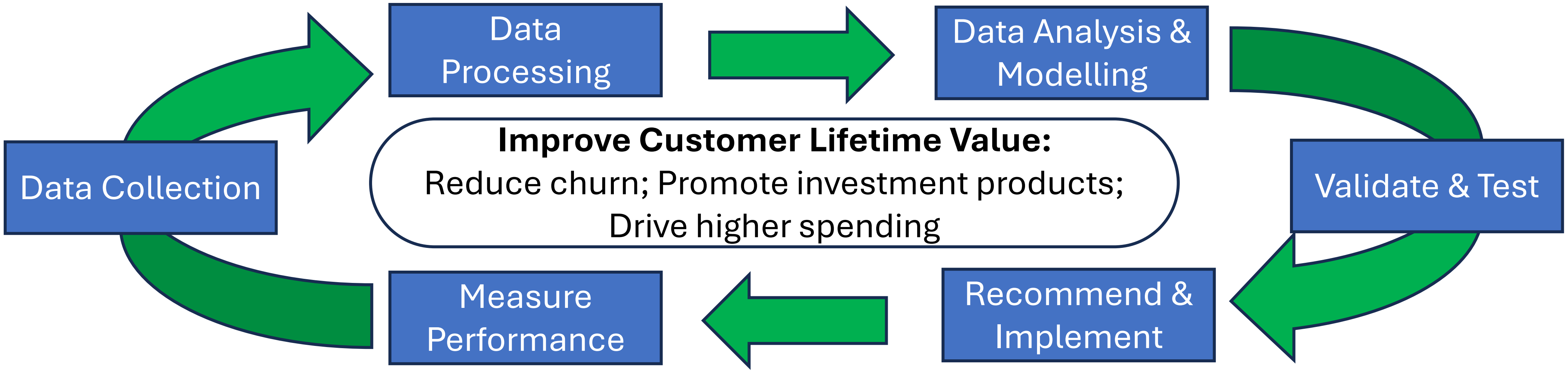
- Use **multiple linear regression** which would provide insight on how multiple factors affect customer spending behaviour. We can offer the credit limit increase to the customer based on their predicted behaviour.
- Able to exclude customers who are predicted to decrease their credit card expenditure from the credit limit increase offer.

Summary of Results

	Naïve	Heuristic	Data Science
Increase in Credit Card Expenditure	$487,968	$566,830	$1,666,128
Increase in Revenue (Assuming the banks earn a 2% from CC Expenditure from merchants)	$9,759	$11,337	$33,323

RECOMMENDATIONS CLOSED-LOOP DATA ECO-SYSTEM

Ensuring a system is in place to capture data and validate and refine the model/solution



Data Collection

- **Seamless capture and processing of current/existing data**
- **Identifying and acquiring <u>additional data</u> points**

Measurement and Feedback

- **Identifying and tracking <u>key performance metrics</u>**
- **Identifying and tracking <u>secondary indicators</u>**

Solution/Action Refinement

- **<u>Testing</u> before rollout (new or updated programs)**
- **Periodic review and update vs continuous <u>review and update</u>**

RECOMMENDATIONS — CLOSED-LOOP DATA ECO-SYSTEM

COLLECTION SYSTEM for the RIGHT INTERNAL and EXTERNAL DATA

Internal: Data can be captured and processed across all points of contact with customer

Account Opening and Registration
- Customer information
- Locale – e.g., city

Usage and Payment
- Balance, consumption, products, offers
- Frequency and timing of activity

Branch, App and Online Interaction
- Visits, emails, CS chats
- Survey
- Waiting time

External: Enhanced with additional information to increase understanding of customers and their needs

External Environment
- Economic climate – e.g., are certain products more popular
- Political climate

Trends
- Investment products
- Card types
- Effective Offers, Tie-ups

RECOMMENDATIONS

CLOSED-LOOP DATA ECO-SYSTEM

IDENTIFYING and COLLECTING PERFORMANCE METRICS

Primary KPIs

- Overall indicator:
  - CLV → quantification of profit generated per customer
- Individual Objectives:
  - Churn rate/retention
  - Investment Product subscription + value
  - Average Card Consumption

Secondary Indicators

- Action Specific Metrics (Feedback)
  - Conversion or Redemption rate
  - Clickthrough, read emails, dropped calls
- Secondary KPIs or factors that affect Primary KPIs:
  - Complaints / Satisfaction
  - Additional factors via research and expert opinion

RECOMMENDATIONS CLOSED-LOOP DATA ECO-SYSTEM

TESTING and REFINING the MODEL and SOLUTIONS

Testing and Validation

- Experimentation
  - Controlled experiments to validate models and verify assumptions (e.g., expected lift or conversion rates)
  - A/B testing to optimize approach or communication to customers
- Testing
  - New features or programs

Review and Refinement

- Periodic/Infrequent vs Continuous/Frequent
- Churn and CC program – annual (high cost, high impact)
- Cross-selling program – min quarterly (catch emerging trends)
- Continuous refinement for communications and quality improvement

POTENTIAL GOVERNANCE ISSUES & POSSIBLE SOLUTIONS

BIAS

Issue	Solutions

False Causality
• Implicit assumption that customers are more profitable if they purchase more products from the bank.
• However, it might be the case that already profitable customers have greater capacity and interest to purchase multiple products.
    • Profitability may be driven by other factors.
• Not all customers will become more profitable by owning more products.
	1. **Further Analysis**: Conduct **customer segmentation** to understand why certain customers own multiple products and look into factors such as income and spending patterns.
	
	1. **Testing**: **Run A/B tests or other experiments** to determine whether increasing product ownership causes an increase in profitability.
	    • Example: Selectively offer additional products to a subset of customers and compare the outcomes to a control group.

Cobra Effect • Increasing credit limits to incentivize spending might lead to over-leveraging, eventually increasing default rates.	1. **Closed Data Feedback Loop**: Obtain data about **customer's defaults rate, repayment behaviour, and current debt levels**. This enables the bank to assess the customer's risk of default. • After identifying the list of customers to offer the 30% credit limit increase based on our solution, further assess each shortlisted customer's risk of default and offer the credit limit increase accordingly.

POTENTIAL GOVERNANCE ISSUES & POSSIBLE SOLUTIONS

BIAS

Issue	Solutions

Sampling Bias
• Implicitly assumed that the dataset is representative of the bank's customer base (i.e 50% of the bank's customers are from France, while the remaining is split between Germany and Spain.
• However, if this is not the actual proportion, this would lead to an:
    • Over-representation of French customers: their preference, behaviours and issues would influence the analysis outcome.
    • Under-representation of German and Spanish customers.
	We would first need to **verify the actual proportion** of the bank's customer segments:
	
	1. Ensure representative sampling through stratified sampling techniques to ensure that the dataset proportionally represents all significant segments of the bank's customer base.
	
	2. Alternatively, if stratified sampling is not possible, **apply weights to the data** to account for over-represented and under-represented groups.

POTENTIAL GOVERNANCE ISSUES & POSSIBLE SOLUTIONS

PRIVACY

Issue	Solutions
• The dataset includes confidential information such as personal data, etc. • There is **a lack of data anonymization** – no masking was performed on customer identifiers. • Unauthorised access or data breaches can lead to **severe reputational damage and legal consequence** from breaching EU data protection regulations.	1. Develop policies and procedures, outlining how personal and confidential data should be handled. 2. When sending the data between departments (e.g. sending to the marketing department for targeted campaigns), ensure that there is **asymmetric encryption** due to the confidential nature of the data. **3. Hash customer identifiers (**e.g. customer number) so that the data cannot be deciphered by parties with unauthorised access.

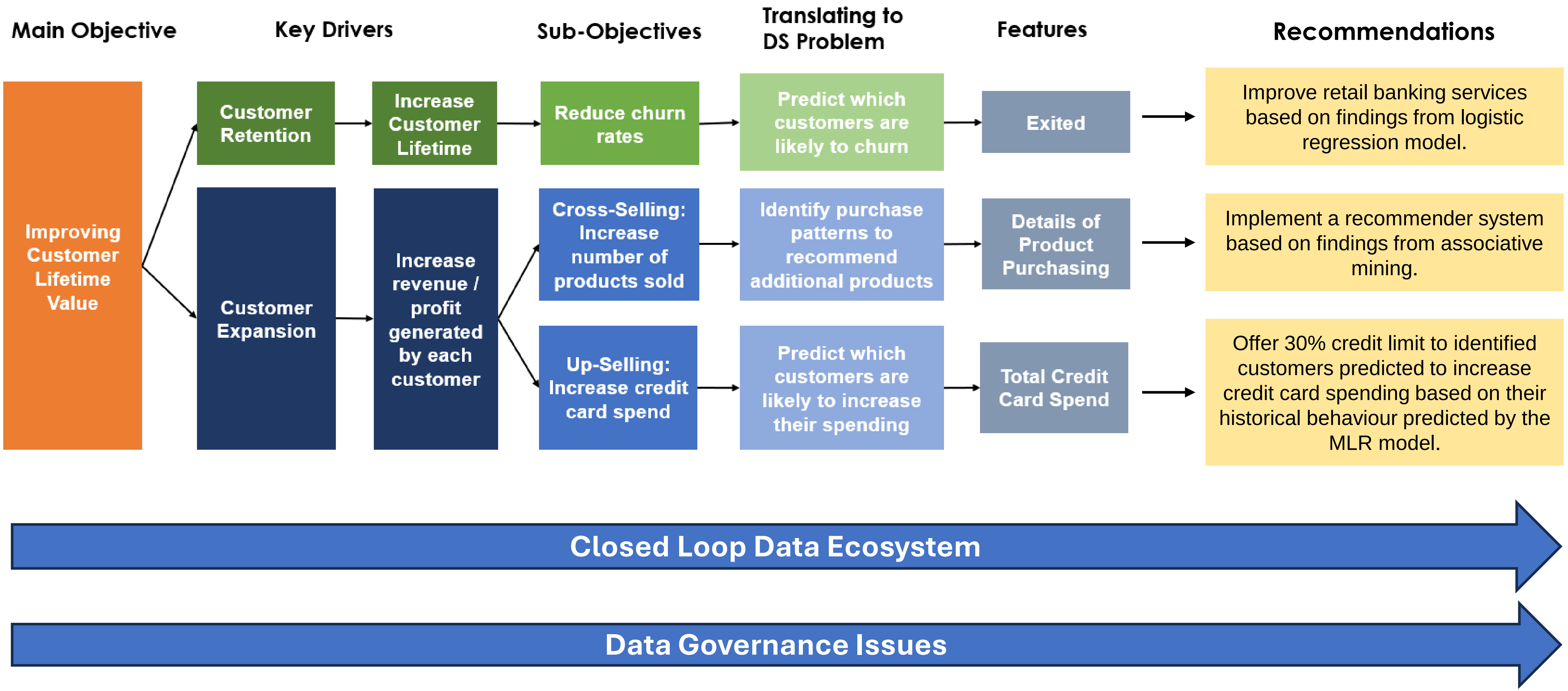
GOVERNANCE

POTENTIAL GOVERNANCE ISSUES & POSSIBLE SOLUTIONS

FAIRNESS

Issue	Solutions
CUSTOMERS • Fair lending regulations require banks to provide equal access to credit and financial services (including credit limit increases) **without discrimination** on gender, race, age. • Such regulations also render targeted marketing campaigns illegal when it poses risks of discrimination for statutorily protected classes (EU Equality Directives).	1. **Exclude features** that could introduce discrimination into the model, e.g. exclude using gender as a feature. 2. **Use objective and non-discriminatory criteria** such as payment history, credit utilization to evaluate customers for credit limit increases.
COMPANIES • Apart from increasing credit limits to drive credit card spending, the bank can consider **sharing customer data with retail merchants** and vice versa. For example, retail merchants can offer promotions for customers who owns the bank's credit cards. • Need to ensure that **both parties benefit equitably** so that they are more willing to share the data.	1. **Revenue Sharing Models**: Merchants can consider paying a fee for access to customer insights by the bank. • Example: They can pay a commission to the bank for each sale made through targeted offers driven by the shared data. 2. **Performance-Based Rewards**: Alternatively, banks can pay merchants for their customer data. The payment should be computed in proportion to the increase in credit card transactions at the merchant's store or growth in average transaction value.

CONCLUSION


Main Objective
Key Drivers
Sub-Objectives
Translating to DS Problem
Features
Recommendations
Improving Customer Lifetime Value
Customer Retention
Increase Customer Lifetime
Reduce churn rates
Predict which customers are likely to churn
Exited
Improve retail banking services based on findings from logistic regression model.
Customer Expansion
Increase revenue / profit generated by each customer
Cross-Selling: Increase number of products sold
Identify purchase patterns to recommend additional products
Details of Product Purchasing
Implement a recommender system based on findings from associative mining.
Up-Selling: Increase credit card spend
Predict which customers are likely to increase their spending
Total Credit Card Spend
Offer 30% credit limit to identified customers predicted to increase credit card spending based on their historical behaviour predicted by the MLR model.
Closed Loop Data Ecosystem
Data Governance Issues


Thank you.