MSDS 460 Final Project

Chekit Mehta, Corvus Lee, Paritosh Sharma, Yining Feng

Northwestern University

Team Introduction

Chekit Mehta

Sr. Manager at Motorola/Lenovo.

4th program; Joined Fall-2019. Pursuing specialization in AI.

Corvus Lee

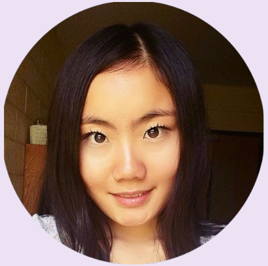
Analytics Specialist Solutions Architect at AWS

7th course; Joined in Spring-2019

Paritosh Sharma

Sr. Manager at Salesforce.

Joined the MSDS Program in Fall-2019. Planning to specialize in AI.



Yining Feng

UPS Capital International Marketing Analyst

This is my 7th course during the 4th Quarter of MSDS Program

Direct Marketing Campaign Optimization:

Multi-objective Approach with RFM Model

Problem Description

- Increasing marketing $ impact i.e. Maximizing ROI while minimizing the total $ amount spend
  - Total $ amount spend hence needs to be smarter
    - Per customer loyalty
    - Per new customer base creation
    - Per targeted possibility of customer spend
- RFM (recency, frequency, monetary) framework can help create highly effective direct marketing campaigns
  - Homogeneous segmentation of customer base into different purchasing behaviors
  - Creating direct marketing targets of these segmentation depending on organization goals
    - Based on profitability of segments
    - Generating more sale
    - Maximizing $ impact of targeted segment

Objective & Challenges of Work

The objective of this paper is to demonstrate the use of integer linear programming and goal programming approaches respectively to determine RFM customer segments that should be targeted to achieve profit maximization for a direct marketing campaign.

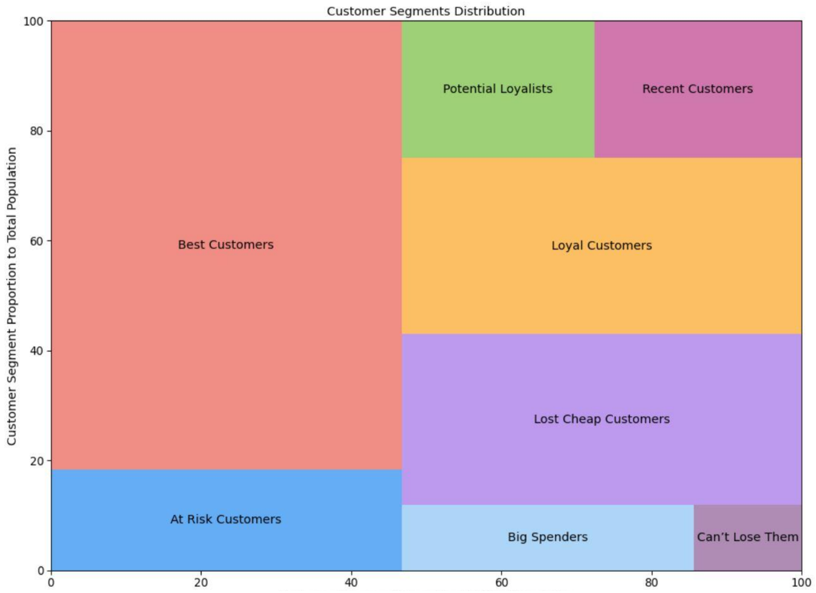
Objective & Challenges of Work

This study highlights a multi-objective optimization methodology based on a goal programming (GP) approach enabling profit maximization for direct marketers using RFM data. One unique characteristic of this (GP) model is the <u>inclusion of varying direct marketing objectives as well as corresponding budget constraints</u>.

Proposed Methodologies

- RFM Customer Segmentation
- Integer Linear Programming (ILP)
- Multi-objective Goal Programming

RFM Customer Segmentation



Customer Segment	RFM Rank	Percentage	Customer Segment Description
Best Customers	444	31.1%	Highest frequency as well as monetary value with least recency
Loyal Customers	344	13.7%	High frequency as well as monetary value with good recency
Potential Loyalists	434	8.3%	High recency and monetary value, average frequency
Big Spenders	334	7.3%	High monetary value but good recency and frequency values
At Risk Customers	244	10.7%	Customers shopping less often now who used to shop a lot
Can't Lose Them	144	5.7%	Customers shopped long ago who used to shop a lot
Recent Customers	443	9.9%	Customers who recently started shopping a lot but with less monetary value
Lost Cheap Customers	122	13.3%	Customers shopped long ago but with less frequency and monetary value

The treemap above shows the distribution of 8 customer segments. The area represents the proportion of each customer segment, while the color represents the range of RFM score level (from 3 to 12) of each customer segment, which is distinguished by characteristics summarized in the table above.

Notation

$i = 1..R$ index to indicate a specific recency group

$j = 1..F$ index to indicate a specific frequency group

$k = 1..M$ index to indicate a specific monetary group

$V_{jk} =$ Expected revenue from a customer in frequency group j and monetary group k

$p_{ijk} =$ Probability of purchase for customers in recency group i, frequency group j and monetary group k

$N_{ijk} =$ Number of customers in recency group i, frequency group j and monetary group k

$C =$ Average cost to reach a customer

$B =$ Total budget of the marketing campaign

Integer Linear Programming (ILP)

This methodology focuses on only one objective without any priority — to maximize the profit from future purchases. This serves as the best case scenario, as adding extra objectives and/or constraints will only have negative effects on the objective value.

$x_{ijk} = 1$ if the customer segment of recency i, frequency j, and monetary k is reached; 0 otherwise

The objective function is:

Maximize:

$$Z = \sum_{i=1}^{R} \sum_{j=1}^{F} \sum_{k=1}^{M} N_{ijk}(p_{ijk}V_{jk} - C)x_{ijk} \tag{1}$$

subject to:

$$\sum_{i=1}^{R} \sum_{j=1}^{F} \sum_{k=1}^{M} N_{ijk}Cx_{ijk} \leq B \tag{2}$$

with:

$$x_{ijk} = \{0, 1\} \quad i = 1..R \;\; j = 1..F \;\; k = 1..M \tag{3}$$

Multi-objective Goal Programming (GP)

The objective of this approach is to prioritize the direct marketing campaigns to "At Risk Customers (244)" and "Can't Lose Them (144)" customer segments characterized with a recency score of 1 or 2, a frequency score of 3 or 4, or a monetary value score of 3 or 4.

$x_{ijk} = 1$ if the customer segment of recency i, frequency j, and monetary k is reached; 0 otherwise

The objective functions are:

Maximize:

$$Z = \sum_{i=1}^{2}\sum_{j=1}^{F}\sum_{k=1}^{M} N_{ijk}(p_{ijk}V_{jk} - C)x_{ijk} \qquad (4) \text{ weight } W_1$$

$$Z = \sum_{i=1}^{R}\sum_{j=3}^{4}\sum_{k=1}^{M} N_{ijk}(p_{ijk}V_{jk} - C)x_{ijk} \qquad (5) \text{ weight } W_2$$

$$Z = \sum_{i=1}^{R}\sum_{j=1}^{F}\sum_{k=3}^{4} N_{ijk}(p_{ijk}V_{jk} - C)x_{ijk} \qquad (6) \text{ weight } W_3$$

subject to:

$$\sum_{i=1}^{R}\sum_{j=1}^{F}\sum_{k=1}^{M} N_{ijk}Cx_{ijk} \le B \qquad (7)$$

with:

$$x_{ijk} = \{0,1\} \quad i = 1..R \quad j = 1..F \quad k = 1..M \qquad (8)$$

Computation experiments

Recency: i
Frequency: j
Monetary: k

Assumptions:

- Quarterly promotional campaign
- Total Budget = $7,000
- Avg Cost to reach a customer = $3.50

Calculate expected revenue (V_{jk}) and probability of purchase (p_{ijk}) using different period of historical transactions

- Uncertainties for Monte Carlo Simulations, assume
  - V_{jk}: uniform distribution across ($0.5V_{ijk}$, $1.5V_{ijk}$)
  - p_{ijk}: normal distribution, with SD = 0.05

Results

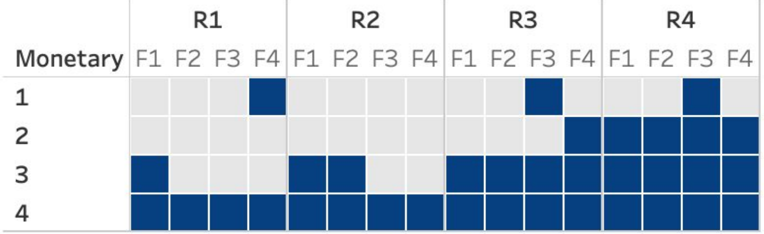
Recency / Frequency

Monetary	R1 F1	R1 F2	R1 F3	R1 F4	R2 F1	R2 F2	R2 F3	R2 F4	R3 F1	R3 F2	R3 F3	R3 F4	R4 F1	R4 F2	R4 F3	R4 F4
1	359	100	27	1	193	83	35	0	146	39	14	0	69	21	7	0
2	97	188	66	8	56	167	90	26	52	133	36	20	41	78	32	4
3	22	47	83	22	29	66	174	60	19	60	162	77	13	54	150	56
4	7	5	24	21	5	11	34	82	12	19	66	224	5	15	78	486

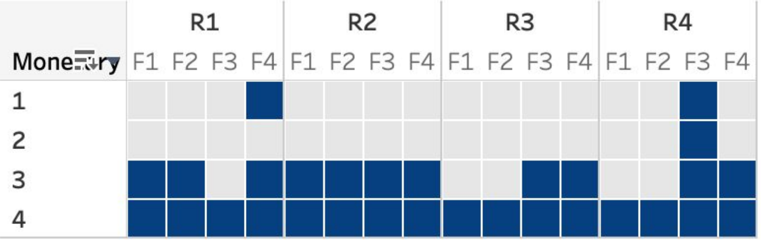
Customers count broken down by Recency and Frequency vs. Monetary. Color shows V*p.

Decision variables x: 1 (blue) or 0 (grey)

ILP result



GP result



Simulation result	ILP	GP
Mean	$1,478,382	$1,456,657
Probability that profit > $1.2M	86.6%	84%

Limitations and Improvements of Work

- Flexibility and practicability of methodologies
- Limitation of potential revenue forecast accuracy
- Limitation of probability of purchase prediction accuracy
- Inclusion of customers' web browsing data **+** transactional data for future research

Questions?