

ホームセンターにおける顧客購買データの分析と その販売促進策への適用

2012SE020 古田大揮 2012SE051 猪狩敦也
2012SE106 河村晴香 2012SE269 渡邊ことこ

指導教員：鈴木敦夫

1 はじめに

近年、POS (Point Of Sales) システムを導入している小売店では、レシートデータの集積が進み、膨大なレシートを分析することで、各商品の販売日時、個数の予測ができるようになってきた。これらはデータサイエンスと呼ばれる一連の手法を生み出し、膨大なデータをこれらの手法で分析することで、顧客動向や販売動向などを分析できると期待されており、多くの小売業で利用法が模索されている。[4]

本研究で対象とするホームセンターは、オペレーションズ・リサーチ (以下 OR) を用いて売上増加や経費削減などの研究に取り組んでいる。例えば、以下のような研究が挙げられる。

- ID 付きレシートデータの分析 [1]
- 大規模レシートデータの分析と活用 [2]

本研究では、このホームセンターにおけるクレジットカードの ID 付きレシートデータを用いた分析を行い、販売促進策に活用をする。目的とする促進策に活用できる形でデータを加工し、ソフトウェアを利用してデータを抽出し、対象商品の絞り込み、各種の最適化モデル構築、DM(ダイレクトメール)の送付対象者の選定などを行った。

本研究で取り扱う施策は、以下の4つである。

1. カード会員の離反防止策
2. 園芸用品におけるポイント活用策
3. ペットシーツの販売促進策
4. ペットトリミング DM 結果分析

1. については、カード会員の減少を防ぎ、売上増加に繋げることを目的として、対象者となる会員に対して DM を送付した。この DM は、対象者に対して、10% のキャッシュバックキャンペーンを行うことで、顧客の来店を促し、そのカード会員の離反を防止することが狙いである。この施策については、既にホームセンターがキャンペーンの実施を行った。

2. については、クレジットカードを利用することで各商品の売価の 1% 分のポイントが付与される。今回、園芸用品において付与ポイントを、3 倍、5 倍、8 倍にすることで、園芸用品の売上増加に繋げることを目的としている。この際、どの商品に何倍のポイントを付与するのかを選定する。限られた予算を考慮しながら、売上が最大になるモデルを作成した。なお、このホームセンターでは、付与ポイントを 3 倍にするという施策の実施は行っていたが、5 倍、8 倍については今回が初めての試みとなる。通常日と 3 倍ポイント実施日のデータを比較し、各商品について何倍のポイン

トを付与するのが最適であるのかを、OR の考え方に基づき、分析する。

3. については、ペットシーツのポイント倍率を 5 倍にすることによって、ペットシーツの売上増加に繋げることを目的としている。カード会員を分析し、ポイント 5 倍キャンペーンの告知 DM を送付する対象者を絞り込み、対象者に対し、DM を送付し、キャンペーンを実施する。

4. については、昨年 1 月に送付した、トリミング割引を中心とするクーポンの利用の有無、DM 発送後の購買傾向に変化が見られたか分析を行う。この結果を基に、キャンペーン告知をする DM の会員の選定を行うモデルを作成し、更なる売上増加に繋げることを目的とする。

我々が分析した結果に基づいて、1. については、実施済みであり、その結果も得られている。3. についてはホームセンターでキャンペーンの実施中である。今後、2., 4. についても実施予定である。

2 使用データについて

本研究で使用するデータは、ホームセンターから提供された機密データであるため、本文中ではダミー値を使用する。

2.1 ID 付きレシートデータについて

本研究では、ホームセンターから提供された ID 付きレシートデータを用いる。レシートデータは、レシートの内容が記録されている。そのレシートデータから、レシート NO、店コード、売上日付、売上時刻、部門、JAN (商品識別番号)、商品漢字名、規格漢字名、数量、販売売価、売上金額、会員番号、年齢、性別を必要に応じて、適宜使用する。(JAN とは、各商品に与えられた固有のコードである。コードのみで商品の識別は可能であるが、商品漢字名、規格漢字名も用いて、詳しく購買傾向を確認する。) レシートデータは、2013 年 3 月から 2015 年 5 月 (約 2 年半) の期間のものを使用する。本研究では、レシート 1 枚を顧客が 1 回の買い物で購入した商品のデータとして考える。

3 カード会員離反防止策

3.1 背景

このホームセンターでは、クレジットカード付きの会員カードを設けている。そのクレジットカードを利用すると、利用額に応じてポイントが付与されるシステムである。しかし、毎月数百名が退会しているのが現状である。その会員の減少を防ぎ、新規会員の増加を図り、売上増加に繋げることが

今回の研究の目的である。また、元々来店頻度が高かったのにも関わらず、その後来店回数が減少傾向となり、離反傾向となってしまった会員の来店頻度の回復も目的とする。

手法として、会員別の離反指数を定義、算出し、離反傾向にある会員にキャンペーン案内の DM 送付を行う。離反指数とともに、直近の来店日、来店回数、平均売上金額などを考慮し、限られた販促費の中で、最も効果の高い施策を行う。キャンペーン内容としては、キャンペーン期間中に会員カードのクレジットで買い物をする、後日 10% キャッシュバックされるというものである。キャンペーンを利用した会員の 10% の離反阻止を目標とする。

3.2 用語の説明

・離反指数

クレジット会員カードを利用している人の中で、来店頻度が少なく、カード会員を退会しそうな人を示す。

0 を基準とし、値がマイナスであれば離反する可能性が低いことを表し、プラスであれば離反する可能性が高いことを表す。

離反指数を以下の式で与える。

$$\text{離反指数} = \frac{(\text{現在日} - \text{直近来店日}) - (\text{来店間隔の平均})}{(\text{来店間隔の標準偏差})}$$

この式は、分子においては、直近来店日から現在日の日数が平均来店間隔からどの程度かけ離れているかを表している。分母に、来店間隔の標準偏差を置くことで、標準偏差を基準とすることができ、分子が標準偏差の何倍離れているのかを指標で表される。このような式を作成し、離反指数を得た。

3.3 分析方法

2013 年 2 月 25 日から 2015 年 5 月 24 日（2 年 3 ヶ月間）のレシートデータを使用する。現在日を 2015 年 5 月 24 日とし、会員別に最終来店日（直近日）、平均来店間隔、標準偏差、（現在日 - 直近日）、離反指数、合計金額、来店回数、平均金額を算出する。

これらの式は以下に示す。

$$\text{平均来店間隔} = \frac{(\text{来店間隔合計})}{(\text{来店回数}) - 1}$$

$$\text{標準偏差} = \sqrt{\frac{\{\sum (\text{来店間隔} - \text{平均来店間隔})^2\}}{(\text{来店回数} - 1)}}$$

3.4 抽出および分析

レシートデータを分析した結果、全会員を対象としていたため、対象人数をさらに絞り込むことにした。抽出を二段階に分け、対象者の選定を行った。

● 抽出方法 1

全会員に対して、来店回数 5 回以上の顧客について離反指数を計算した。

離反指数の 2.0 を抽出条件の値とする。（離反指数 2.0

≒ 1.96 となり、これは 5% の指数であるので、平均よりも来なくなっていることを示す。）

この結果を基に、再度抽出を行う。

● 抽出方法 2

離反指数 2.0 以上であり、かつ来店回数が 24 回以上（月 1 回以上）の顧客について抽出を行う。

このデータを基に、以下の 4 つの抽出条件で DM 送付者の選定を行う。

● 抽出条件

- ・来店回数 24 回以上（約月 1 回来店）
- ・1 回あたり 2,000 円以上の購入
- ・離反指数 2.0 以上
- ・直近来店日 2014 年 10 月 1 日～2015 年 3 月 31 日

3.5 抽出結果

以下は、上記の抽出条件を基に、抽出した結果である。

- DM 送付対象会員数：約 3,000 名
- 販促費（DM 発送費）：約 3000a 円（1 件あたり a 円）

3.6 施策結果

予測値としては 15% の利用率であったが、結果としては、約 40%（約 1200 名）の利用率という予測値を大きく超える結果となった。また、客単価も普段を大きく上回る値となり、粗利および利益も予想を大きく上回る結果となった。以下の表 1、図 1 は、DM に反応し、10% キャッシュバックを利用した会員の離反指数別利用率の表とヒストグラムである。

3.7 考察

昨年、このホームセンターでは、クレジットカードの利用が無く、年会費が発生しそうな会員に対して、5% キャッシュバックキャンペーンを行った。その結果、DM 送付人数約 6000 名中、利用者数は約 850 名（利用率 13.6%）という結果になった。今年は、上記のような絞り込みを行った結果、DM 送付対象会員数は約 3000 名となり、結果として約 40%（約 1200 名）の利用率という予測値の 15% を遥かに上回る結果となった。また、クーポンを利用した会員を離反指数別にヒストグラムで表したところ、累計利用率が 35.5% を上回る会員を対象に絞り込むと、より効果が得られるのではないかと推測される。更に、離反指数 1.5 から 2.0 の会員を対象に DM を送付すれば、利用率がさらに上がり、利益が見込めるであろう。今後の課題としては、キャンペーンを利用した約 1200 名の会員が 9 月以降に来店頻度が上がり、継続した利用に繋がっているのかを検証することである。そして、ホームセンターが送付した DM に反応した顧客の 10% の離反者数減少を目指す。

表 1 離反指数別利用率表

離反指数	2	3	4	5	6	7	8	9	10	11	12	13	14
利用率	38.9	39.5	39.5	32.4	31.6	28.3	16.1	23.9	28.3	32.6	21.6	28.9	12.5
累計利用率	38.9	39.2	39.3	38.1	37.6	37.0	36.2	35.9	35.8	35.7	35.5	35.5	35.3

離反指数	15	16	17	18	19	20	21	22	23	24	25	26	27
利用率	13.6	33.3	10.0	0.0	25.0	10.0	42.9	100.0	100.0	33.3	33.3	33.3	0.0
累計利用率	35.2	35.2	35.1	35.0	35.0	34.9	35.0	35.0	35.1	35.1	35.1	35.1	35.1

離反指数	28	31	32	33	35	38	40	45	48	52	64	72
利用率	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	100.0	0.0	0.0
累計利用率	35.1	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0	35.0

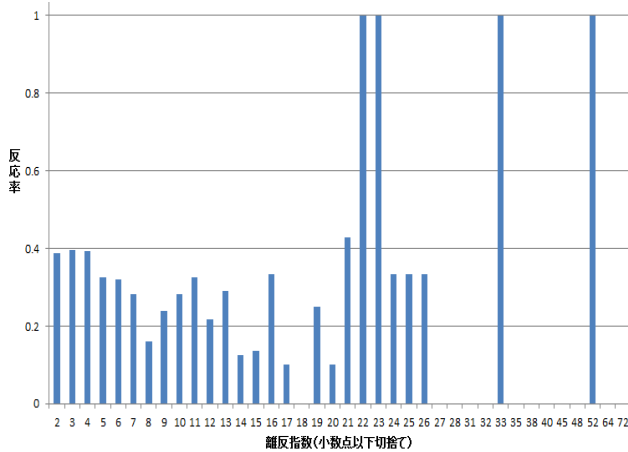


図 1 離反指数別利用率ヒストグラム

4 園芸ポイント活用策

4.1 背景

秋の園芸シーズンに向けて、園芸用品の売上増加を促す施策を行うこととなった。今回の研究は、園芸分野の各商品に対して、3倍、5倍、8倍のポイントを付与し、園芸用品の売上増加を目的とする。

ホームセンターは昨年、2014年9月から11月までの土曜日と日曜日（26日間で12日間）に全商品を対象として商品を購入した際に付与されるポイントを通常日の3倍にするというキャンペーンを実施した。我々はその結果を用いて、園芸用品に対する付与ポイントを通常日の3倍、5倍、8倍にする商品の選定を行った。

そこで、各商品に何倍のポイントを付与するのかが選定するモデルの作成を行った。通常日とは、各商品毎の売価の1%のポイントを付与する日を示す。

なお、このホームセンターでは、過去に、付与ポイントを3倍にするという施策の実施は行っていたが、5倍、8倍については今回が初めての試みとなる。

4.2 使用するデータ

2014年9月から11月までの3ヶ月間の土曜日、日曜日（計26日間）のレシートデータを使用する。園芸用品に分類される商品を購入しているレシートデータを抽出した。

4.3 分析方法

レシートデータを3倍ポイント実施日（12日間）と通常日（14日間）に分割し、数量の合計や金額の合計、1日あたりの数量、金額を算出する。

3倍ポイント実施日の分析結果と通常日の分析結果を比較し、3倍ポイント実施の効果を売上数量の伸び率（倍率）、売上金額の伸び率（倍率）を算出し、表す。

また、1枚のレシートデータにおいて、同時に購入されている商品の合計金額から、ある購入商品1つに対して、同時に購入されることが期待できる、その他の商品の合計金額を期待値によって表す。

4.4 分析結果

- 各項目の詳細を以下に示す。

数量伸び率：

$$\frac{(3倍ポイント実施日の1日あたりの売上数量)}{(通常日の1日あたりの売上数量)}$$

金額伸び率：

$$\frac{(3倍ポイント実施日の1日あたりの売上金額)}{(通常日の1日あたりの売上金額)}$$

同時購入金額：同時購入された商品の金額合計

レシート枚数：商品が含まれているレシート枚数

期待値：1回の購入で同時購入される金額の期待値

$$\sum (\text{平均売価} \times \text{同時購入される確率})$$

期待値+平均売価：1回の購入で使用される金額の期待値

通常日と比較すると、付与ポイント3倍キャンペーン実施日に金額伸び率、数量伸び率ともに最大25.6倍という結果を得た。また、商品Aは数量の伸び率と金額の伸び率、同時購入のレシート枚数、売価から期待値562.1円という結果を得た（表2）。

これらの結果を基に、最適付与ポイントの定式化を行う。

4.5 最適ポイントの定式化

抽出結果を基に、どの商品に何倍のポイントを付与するのが最適であるのかを分析するために、定式化を行う。

まずは、定数および変数を定義する。

- 定数

- ・ N ：商品の集合

- ・ M ：ポイント倍率パターン

$$M = \{0, 1, 2, 3\}$$

(0 = 1 倍, 1 = 3 倍, 2 = 5 倍, 3 = 8 倍)

・ n : ポイント倍率変更時 ($M=1, 2, 3$ のとき) の商品数の上限

・ a_{ij} : 商品 $i \in N$ をパターン $j \in M$ で 1 単位販売すると得られる利益

($a_{ij} = (\text{売価} \times \text{利益率}) - (\text{付与ポイント数})$)

・ P_{ij} : 商品 $i \in N$ をパターン $j \in M$ で販売したときのポイント数

(ex. 3 倍 \rightarrow 数量 \times 売価 \times 0.03)

● 変数

・ x_{ij} : 商品 $i \in N$ をパターン $j \in M$ で販売したとき,

$$x_{ij} = \begin{cases} 1 (\text{選ぶ}) \\ 0 (\text{選ばない}) \end{cases}$$

以上の定数および変数を基に, 定式化を行う. 定式化は以下の通りである.

4.6 定式化

Max.

$$\sum_{i \in N} \sum_{j \in M} a_{ij} x_{ij} \quad (1)$$

s.t.

$$\sum_{i \in N} \sum_{j \in M \setminus \{0\}} x_{ij} \leq n \quad (2)$$

$$\sum_{i \in N} \sum_{j \in M} P_{ij} x_{ij} \leq 1.0 \times 10^6 \quad (3)$$

目的関数は, 粗利を最大にする関数とし, 制約条件は, 倍増ポイントの対象とする商品数の上限および付与するポイント数の上限の 2 つとした. (ポイントの上限は, 仮の上限として 100 万ポイントを設定している.)

この定式化を, Excel 上の数理計画法最適化ソフトウェアである What'sBest!12.0 を用いて, 最適解を求めた.

まず, 5 倍, 8 倍ポイントについては, このホームセンターでは初めての実施であり, 数量のデータが存在しなかったため, 予想の数量を作成した. はじめに, 抽出商品を対象に, 通常日と 3 倍ポイント実施日の数量を比較し, それぞれの伸び率を算出した. 次に, その値を用いて, 3 倍から 5 倍, 5 倍から 8 倍でも同様に比例で伸びると仮定し, 5 倍, 8 倍ポイント実施時の予想の数量を作成した. その結果を示した (表 3).

各対象商品に対して, 最適な付与ポイントは, 表 4 の通りである. この表を作成することで, 各対象商品に対する最適ポイントを目で知ることができる.

期待される粗利, 3, 5, 8 倍ポイントを適用した商品数, 付与ポイント数は, 表 5 の通りである. (今回は, 付与ポイント数上限を 20 万ポイントで設定した.)

表 2 期待値表

商品情報			結果					
JAN	商品名	平均売価	数量伸び (倍率)	金額伸び (倍率)	同時購入合計金額	レシート枚数	期待値 (円)	期待値 + 平均売価 (円)
00001	A	284.0	25.6	25.6	68,023	121	562.1	846.1
00002	B	55.0	20.0	20.0	5,839	23	253.8	308.8
00003	C	284.0	17.5	17.5	10,076	20	503.8	787.8
00004	D	7,800.0	16.3	16.3	74,046	108	684.0	8,484.0
00005	E	189.0	16.3	16.3	59,498	131	456.9	645.9

表 3 抽出商品データ

商品名	売価	利益 (売価 \times 利益率)	(売価 \times 利益率) - 付与ポイント (1.35, 8 倍)					ポイント数 / 個 (売価 / 100)	数量1倍	数量3倍	数量5倍	数量8倍	数量伸び (倍率)
			1.35	1.11	1.02	0.87	0.77						
A	474	166	162	154	146	134	4	865	1381	2572	4791	1.86	
B	361	126	123	117	111	102	3	1256	2072	3682	6543	1.78	
C	139	49	48	46	44	41	1	1914	2698	5144	9131	1.78	
D	284	103	101	97	93	87	2	831	1192	1965	3241	1.65	
E	264	93	91	87	83	77	2	737	924	1292	1806	1.40	
F	398	139	136	130	124	115	3	1114	1323	1833	2540	1.39	

表 4 実行結果 (最適ポイント配置)

商品名	1 倍	3 倍	5 倍	8 倍	最適 P
A	0	0	1	0	5
B	0	0	1	0	5
C	0	0	1	0	5
D	0	0	0	1	8
E	1	0	0	0	1
F	0	1	0	0	3

表 5 実行結果 (粗利, 商品数, 付与ポイント数)

目的関数	aixij	1531687		
制約条件	xij			
		5	<=	6
	pixij			
		197629	<=	200000

商品数, 付与ポイント数の上限を変更することで, それぞれに応じた最大の粗利, 最適なポイント配置を得ることができる.

4.7 考察

3 倍ポイント実施時には, 通常日に比べ, 売上が伸びていた. 5 倍, 8 倍ポイントを実施することができれば, 更なる成果が望むことができると考えた. また, このシステムを用いれば, 状況に応じて商品, 付与ポイントも変更可能であるため, この考え方 (OR) を知らない人であっても, いかなる時も各対象商品毎の付与ポイントを決定することができる.

5 ペットシーツの販売促進策

5.1 背景

ペットシーツに関するキャンペーンを行うこととなった。内容は、11月1日から12月31日の2ヶ月間、ペットシーツを各種購入した際に付与されるポイントを通常の5倍にするというものだ。

このキャンペーンを実施する際、よりキャンペーンに効果をもたらしたいと考え、会員にキャンペーンを告知するDMを送付する。しかしながら、送付にコストがかかるため、会員の一部のみにしか送付できない。だが、告知のDMを送ることによって、店舗に足を運ぶ頻度が上がると考えられる。また、ペット用品の属する部門だけではなく、全部門に渡る部門の商品の購入に繋げることができ、売上増加にも大きな影響を与えられるとも考えられる。

したがって、より効果を上げるには、全会員の中からDMを送る一部の会員を選定するという絞り込みが重要であると考えられる。我々は、レシートデータを用いて絞り込みを行うことで、売上増加に効果的な絞り込みができると考えた。

この施策では、キャンペーンを告知するDMの一部の会員の選定をし、更なる売上増加に繋げる事を目的とする。

5.2 使用するレシートデータ

2014年3月1日から7月31日まで(5ヶ月間)のレシートデータを使用する。そのうち、「生態のいる店舗」である13店舗を対象とする。

また、ペット部門のうち、付与ポイント5倍の対象商品であるペットシーツを購入しているレシートを抽出した。

ペットシーツを利用する人の傾向として、室内で小型犬を飼っている場合が多いということに気づき、その観点に着目した。したがって、小型犬用ドッグフードを購入しているレシートデータを抽出した。

また、モデルを考える際や施策実施後の反応結果の分析を行う際に、より精密な分析ができるよう、5ヶ月間に来店回数が3回以上の会員のレシートに限定した。

5.3 分析方法

会員番号別に、傾向別購入確率や小型犬用ドッグフードとペットシーツの購入金額の合計、1回来店ごとの小型犬用ドッグフードとペットシーツの平均購入金額、小型犬ドッグフード、ペットシーツ別の購入金額の合計、合計来店回数、小型犬ドッグフード、ペットシーツ別の購入数量の合計を算出する。

これらの比較により、購入傾向を会員別に視える化することができる。

5.4 分析結果

抽出条件を基に絞り込みを行った結果、会員は1667名に絞り込むことができた。表6は分析結果である。

- 以下に項目の詳細を示す。

・傾向別購入確率：同時に小型犬用ドッグフード、ペットシーツの両方、小型犬用ドッグフードのみ、ペットシーツのみで購入する割合をそれぞれ確率で表したもの(来店回数10回のうち、両方購入した回数が5回、小型犬用ドッグフードのみを購入した回数が2回、ペットシーツのみを購入した回数が3回の場合、傾向別購入確率は左から50%、20%、30%となる。)

$$\text{傾向別購入確率} = \frac{\text{(来店した内の傾向別購入した回数)}}{\text{(来店回数)}}$$

表6 会員別購入傾向集計結果

会員番号	傾向別購入確率(%)			購入金額			来店回数	合計個数		
	両方	シーツ	フード	合計	平均	フード合計		シーツ合計	合計	フード
00001	0.0	100.0	0.0	49711	1775.4	0	49711	28	0	44
00002	0.0	100.0	0.0	33441	1337.6	0	33441	25	0	31
00003	30.0	45.0	25.0	27178	1132.4	6980	20198	24	11	4
00004	0.0	100.0	0.0	32286	1467.5	0	32286	22	0	30
00005	0.0	100.0	0.0	37938	1724.5	0	37938	22	0	35

小型犬用ドッグフードとペットシーツを、各1点以上購入した会員は513名であり、そのうち、小型犬用ドッグフードとペットシーツを毎回同時に購入した人は5名であった。これは全体の約0.3%とかなり少なく、この5名は特殊な購入傾向であることがわかった。

また、小型犬用ドッグフードのみ購入する会員は、ペットシーツのみ購入する会員に比べ、大幅に多いことが分かった。

5.5 考察

今後の課題の1つとしては、効果が最大となるDMの送付者を選定するモデルを作成することが挙げられる。

キャンペーン実施期間終了後は、結果の分析を行う予定である。DMを送った会員のうち、このキャンペーンにおける反応の差がどのような会員の傾向によるのかを中心に分析を行う。今回算出した傾向別購入確率や性別、年齢、購入金額などの会員の属性からグループ分けをし、より反応率が高いグループを抽出する。

この結果を分析することによって、今後のDM送付をより効率的に行うことができるため、効率よく利益を上げることができると考えられる。

6 ペットトリミングDM結果分析

6.1 背景

2015年1月に、トリミング割引を中心とする4枚のクーポンを1250名に送付した。

ホームセンター側は、今後もこの内容のクーポン送付を考えている。しかしながら、送付にコストがかかるため、会員の一部のみにしか送付できない。そこで、我々はその

クーポン利用の有無や、クーポン送付前と送付後の購買傾向の変化を分析することによって、今後のクーポン利用率、購入金額、来店頻度の向上に繋げることができると考えた。

したがって、より効果を上げるには、全会員の中からDMを送る一部の会員を選定するという絞り込みが重要であると考えられる。

この施策では、DMでキャンペーンを告知する一部の会員の選定をし、更なる売上増加に繋げる事を目的とする。

6.2 使用するレシートデータ

2013年2月1日から2015年8月31日まで(2年7ヶ月間)のレシートデータを使用する。また、2014年3月から8月において、犬用品を2000円から8000円の範囲で購入している会員のレシートを抽出した。

6.3 分析方法

年齢を20代から80代、性別を男女、の性別年代で計14分割し、それぞれの反応率、およびクーポン利用枚数の期待値を算出した。また、2014年3月から8月における犬用品買上金額2000円代から7000円代に6分割し、反応率を算出した。

2013年2月1日から2015年1月31日(2年間)を実施前、2013年2月1日から2015年8月2日(約6ヶ月間)を実施後(実施中を含む)とした。それぞれにおける、1ヶ月間における平均来店回数、1ヶ月間における平均買上金額、来店1回における平均買上金額を会員別に算出し、それを基に実施前、から実施後における伸び率を算出した。

実施前、実施後の期間別に、DMに対する反応者の性別年代の14分割、犬用品買上金額別の6分割それぞれにおいて、1ヶ月間における平均来店回数、1ヶ月間における平均買上金額、来店1回における平均買上金額を算出した。

6.4 分析結果

以上の分析方法を基に分析を行った結果、以下の表7、表8を得ることができた。

表7はクーポン利用者の各金額帯の来店回数、買上金額(1ヶ月あたり)、平均金額(来店1回あたり)をキャンペーン実施前と実施後に分類し、集計したものである。全ての金額帯の平均金額が増加していることがわかる。

表8はクーポン反応者を性別、年代で分類し、同様に、来店回数、買上金額、平均金額をキャンペーン実施前と実施後に分類し、集計したものである。

男女とも20代の反応者はいない。男性の反応者は50代が実施後の平均金額が最も高く、女性の反応者は40代が平均金額が最も高いという結果を得た。

これらの分析結果を用いて、モデルを作成する。

6.5 定式化

今回の分析では、結果の精度を上げるため、2つのモデルを作成した。1つ目のモデルの目的としては、1ヶ月1人あたりの買上金額における伸びの期待値を最大とすることであ

表7 金額別傾向分析表

利用者		2000円台 (26人)	3000円台 (22人)	4000円台 (18人)	5000円台 (23人)	6000円台 (11人)	7000円台 (20人)
来店回数	前	1.5	1.8	1.6	1.5	1.9	1.8
	後	1.7	1.7	1.5	1.6	2.1	2.0
買上金額	前	6742.4	5984.2	5008.7	5087.1	6089.8	6451.2
	後	7219.8	5975.7	5039.9	6068.3	5904.5	6491.7
平均金額	前	4563.4	3419.6	3078.3	3404.6	3255.1	3638.8
	後	4318.9	3603.4	3301.2	3861.1	2785.1	3259.4

表8 性別年代別傾向分析表

利用者男性		20代 (0人)	30代 (3人)	40代 (11人)	50代 (11人)	60代 (18人)	70代 (2人)	80代 (2人)
来店回数	前	0	1.1	1.6	1.6	1.8	2.9	1.6
	後	0	0.8	1.4	1.8	2.0	3.0	2.4
買上金額	前	0	4000	5000	10000	6000	12000	4500
	後	0	3000	6000	15000	10000	20000	10000
平均金額	前	0	3788	3068	6111	3420	4065	2734
	後	0	3857	4304	8390	5070	6667	4138

利用者女性		20代 (0人)	30代 (3人)	40代 (18人)	50代 (19人)	60代 (26人)	70代 (7人)	80代 (0人)
来店回数	前	0	1.1	1.3	1.8	1.7	1.9	0
	後	0	0.8	1.5	1.7	2.1	1.8	0
買上金額	前	0	4000	6000	6000	5000	5000	0
	後	0	6000	11000	12000	11000	8000	0
平均金額	前	0	3512	4500	3411	2882	2697	0
	後	0	7200	7480	6874	5340	4571	0

り、2つ目は、1ヶ月1人あたりの来店回数における伸びの期待値を最大とすることである。制約式としてはまず、DM送付には費用がかかるため、送付対象人数の上限を共に与えた。さらに、1つ目のモデルでは、送付対象者の1ヶ月あたりの来店回数における伸びの制約を、2つ目のモデルでは、送付対象者の1ヶ月あたりの買上金額における伸びの制約をそれぞれ加えた。

まずは、定数および変数を定義する。

- 定数

- ・ N : クーポン利用 ($i \in N$)

$$N = \{0, 1\}$$

- (クーポン利用なし $\rightarrow 0$, クーポン利用あり $\rightarrow 1$)

- ・ M : 犬部門利用金額 ($j \in M$)

$$M = \{0, 1, 2, 3, 4, 5\}$$

- (2000円台 $\rightarrow 0$, 3000円台 $\rightarrow 1$, ..., 7000円台 $\rightarrow 5$)

- ・ L : 性別年代分類 ($k \in L$)

$$L = \{0, 1, 2, 3, \dots, 13\}$$

- (20代男性 $\rightarrow 0$, 30代男性 $\rightarrow 1$, ..., 80代女性 $\rightarrow 13$)

- ・ $a_{ij}(a_{ik})$: クーポン利用の有無 $i \in N$, 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) のときの、クーポン利用後の1ヶ月間の全部門における買上金額

- ・ $b_{ij}(b_{ik})$: クーポン利用の有無 $i \in N$, 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) のときの、DM送付対象人数

- ・ $c_{ij}(c_{ik})$: クーポン利用の有無 $i \in N$, 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) のときの, DM 送付後の 1 ヶ月あたりの来店回数
- ・ $d_j(d_k)$: クーポン利用の有無 $i \in N$, 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) のときの, 利用率

$$d_j = \frac{b_{1j}}{b_{0j} + b_{1j}} (d_k = \frac{b_{1k}}{b_{0k} + b_{1k}}) \quad (4)$$

- ・ a'_j : 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) のときの, DM 反応者の非反応者に対する買上金額の伸び (差)

$$a'_j = a_{1j} - a_{0j} (a'_k = a_{1k} - a_{0k}) \quad (6)$$

- ・ c'_j : 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) のときの, DM 反応者の非反応者に対する来店回数の伸び (差)

$$c'_j = c_{1j} - c_{0j} (c'_k = c_{1k} - c_{0k}) \quad (7)$$

● 変数

- ・ $x_{ij}(x_{ik})$: 0,1 変数
クーポン利用の有無 $i \in N$, 犬部門利用金額 $j \in M$ (性別年代分類 $k \in L$) で,

$$x_{ij} = \begin{cases} 1 & (\text{選ぶ}) \\ 0 & (\text{選ばない}) \end{cases}$$

以上の定数および変数を基に, 定式化を行う. 定式化は以下の通りである.

● 最適化モデル 1

- 1 ヶ月 1 人あたりの買上金額の伸びの期待値最大の場合,

Max.

$$\sum_{i \in N} \sum_{j \in M} a'_j d_j x_{ij} \quad (8)$$

s.t.

$$\sum_{i \in N} \sum_{j \in M} b_{ij} x_{ij} \leq 1.0 \times 10^3 \quad (9)$$

(=送付対象者の人数制約式)

$$\sum_{i \in N} \sum_{j \in M} c'_j b_{1j} x_{ij} \geq 1.0 \times 10 \quad (10)$$

(=送付対象者の来店回数制約式)

$$x_{ij}(x_{ik}) = \begin{cases} 1 \\ 0 \end{cases} \quad (i \in N, j \in M) \quad (11)$$

(※性別年代で分類する場合: $j \rightarrow k$)

この定式化を, Excel 上の数理計画法最適化ソフトウェアである What'sBest12.0!を用いて最適解を求めた (表 9, 表 10). 今回制約式については, 仮の値として (8) の上限を 1000, (9) の下限を 10 とした.

● 最適化モデル 2

- 1 ヶ月 1 人あたりの来店回数における伸びの期待値最大の場合.

Max.

$$\sum_{i \in N} \sum_{j \in M} c'_j d_j x_{ij} \quad (12)$$

s.t.

$$\sum_{i \in N} \sum_{j \in M} a'_j b_{ij} x_{ij} \geq 5.0 \times 10^4 \quad (13)$$

(=送付対象者の買上金額制約式)

$$\sum_{i \in N} \sum_{j \in M} b_{ij} x_{ij} \leq 1.0 \times 10^3 \quad (14)$$

(=送付対象者の人数制約式)

$$x_{ij}(x_{ik}) = \begin{cases} 1 \\ 0 \end{cases} \quad (i \in N, j \in M) \quad (15)$$

(※性別年代で分類する場合: $j \rightarrow k$)

この定式化を, Excel 上の数理計画法最適化ソフトウェアである What'sBest12.0!を用いて最適解を求めた.(表 11, 表 12). 今回制約式については, 仮の値として (12) の上限を 50000, (13) の下限を 1000 とした.

表 9 1 ヶ月 1 人あたりの買上金額における伸びの期待値最大 (性別年代別)

		利用者(男)							
年代		20代	30代	40代	50代	60代	70代	80代	
金額差	akj-a0k	-4168.9	-5895	442.3	6732.4	3387.7	14450	6354.7	
買上金額	aik	0	3130.1	6260.7	13545	10353	19953	10361	
送付人数	bik	2	14	72	134	174	112	27	
利用人数	b1k	0	3	11	11	18	2	2	
来店回数差	c1k-c0k	-2.5	-1.5	-0.2	0	0.1	1.3	1.1	
来店回数	cik	0	0.8	1.4	1.8	2	3	2.4	
利用率	dk(=b1k/bik)	0.000	0.214	0.153	0.082	0.103	0.018	0.074	
xik		0	0	0	1	1	1	1	
		利用者(女)							
年代		20代	30代	40代	50代	60代	70代	80代	
金額差	akj-a0k	-1549.4	955.3	4450.3	7330.1	5597.2	2980	-3414	
買上金額	aik	0	5694.8	10915	12755	10618	7410.8	0	
送付人数	bik	6	34	113	201	233	114	16	
利用人数	b1k	0	3	18	19	26	7	0	
来店回数差	c1k-c0k	-0.5	-0.8	-0.1	0.1	0.7	0.3	-0.9	
来店回数	cik	0	0.8	1.5	1.7	2.1	1.8	0	
利用率	dk(=b1k/bik)	0.000	0.088	0.159	0.085	0.112	0.061	0.000	
xik		0	0	1	1	1	0	0	
dk(a1k-a0k)xik		3658.2							
bikxik		994	<=	1000					
b1k(c1k-c0k)xik		24.9	>=	10					

表 10 1 ヶ月 1 人あたりの買上金額における伸びの期待値最大 (金額別)

		利用者					
金額		2000円台	3000円台	4000円台	5000円台	6000円台	7000円台
金額差	a1j-a0j	2474.64	-387.99	-1054.74	80.93	-237.21	-550.44
買上金額	ajj	7219.77	5975.66	5039.88	6068.31	5904.52	6491.73
送付人数	bij	278	242	205	199	154	172
利用人数	b1j	26	22	18	23	11	20
来店回数差	c1j-c0j	0.16	-0.14	-0.17	-0.41	0.3	0.14
来店回数	cij	1.67	1.66	1.53	1.57	2.12	1.99
利用率	dj(=b1j/bij)	0.094	0.091	0.088	0.116	0.071	0.116
xij		1	0	0	0	1	1
dj(a1j-a0j)xij		150.493					
bijxij		604	<=	1000			
b1j(c1j-c0j)xij		10.26	>=	10			

表 11 1ヶ月1人あたりの来店回数における伸びの期待値最大(性別年代別)

		利用者(男)						
年代		20代	30代	40代	50代	60代	70代	80代
金額差	akj-a0k	-4169	-5895	442.3	6732.4	3387.7	14450	6354.7
買上金額	aik	0	3130.1	6260.7	13545	10353	19953	10361
送付人数	bik	2	14	72	134	174	112	27
利用人数	b1k	0	3	11	11	18	2	2
来店回数差	c1k-c0k	-2.5	-1.5	-0.2	0.1	0.1	1.3	1.1
来店回数	cik	0	0.8	1.4	1.8	2	3	2.4
利用率	dk(=b1k/bik)	0.000	0.214	0.153	0.082	0.103	0.018	0.074
xik		0	0	0	1	1	1	1
		利用者(女)						
年代		20代	30代	40代	50代	60代	70代	80代
金額差	akj-a0k	-1549	955.3	4450.3	7330.1	5597.2	2980	-3414
買上金額	aik	0	5694.8	10915	12755	10618	7410.8	0
送付人数	bik	6	34	113	201	233	114	16
利用人数	b1k	0	3	18	19	26	7	0
来店回数差	c1k-c0k	-0.5	-0.8	-0.1	0.1	0.7	0.3	-0.9
来店回数	cik	0	0.8	1.5	1.7	2.1	1.8	0
利用率	dk(=b1k/bik)	0.000	0.088	0.159	0.095	0.112	0.061	0.000
xik		0	0	1	1	1	0	0
b1k(a1k-a0k)xik		541548	>=	50000				
bikxik		994	<=	1000				
dk(c1k-c0k)xik		0.195						

表 12 1ヶ月1人あたりの来店回数における伸びの期待値最大(金額別)

		利用者					
金額		2000円台	3000円台	4000円台	5000円台	6000円台	7000円台
金額差	alj-a0j	2474.64	-387.99	-1054.7	80.93	-237.21	-550.44
買上金額	ajj	7219.77	5975.66	5039.88	6068.31	5904.52	6491.73
送付人数	bij	278	242	205	199	154	172
利用人数	b1j	26	22	18	23	11	20
来店回数差	clj-c0j	0.16	-0.14	-0.17	-0.41	0.3	0.14
来店回数	cij	1.67	1.66	1.53	1.57	2.12	1.99
利用率	dj(=b1j/bij)	0.094	0.091	0.088	0.116	0.071	0.116
xij		1	0	0	0	1	1
b1j(a1j-a0j)xij		50722.5	>=	50000			
bijxij		604	<=	1000			
dj(c1j-c0j)xij		0.05					

6.6 考察

送付対象人数が1250名、その内クーポンを1枚でも利用した会員は120名であった。我々の考えていた年代、性別、犬用品買上金額別で分割すると、1グループの人数が0名であるグループもあり、1人のデータが大きくグループの数値に影響してしまう。よって、細かく分割しすぎると制度の高い分析を行うことができないと判断した。したがって、今回の分析では、性別年代の14分割、犬用品買上金額の6分割に分け、それぞれ2つのモデルを作成した。

最適解を求めた結果(表9から表12)、性別年代別のモデルではどちらも男性は50代以上を、女性は40から60代の中高年を選択している。また、金額別のモデルではどちらも2000、6000、7000円台を選択している。条件とした金額帯の中間の範囲は選択せず、低い金額と高い金額の両極端という結果になった。

今後の課題としては、ホームセンターの更なる送付者の増加により、統計的に精密な購入傾向のデータができた際、その分析結果を用いたシステムの作成が挙げられる。また、年代、性別、犬用品買上金額別の計84分割で、どのグルー

プに送付するかを決定するシステムを作ることも挙げられる。

7 おわりに

本研究では、ホームセンターより提供された様々な情報が付随しているレシートデータを分析し、カード会員の離反防止策や販売促進策に繋げることを目的とした。そのためにレシートデータを抽出、分析し、表化、又はグラフ化することで会員の傾向や状況を視える化することに到達した。それを基に、会員の絞り込みや対象商品の選出を行い易くすることに繋げることができた。

また、ORの考え方にに基づき、理論的に制度の高い分析を行うことで、根拠を明確にすることができた。それにより、施策によっては実際に実施され、期待された結果以上の結果を報告されるところまで到達した。

今後の展望として、従来より限られた期間、人数で評価されてきた研究も更に大規模なデータを利用することで新しい発見ができると思う。

またそれに加え、データ量の不足などにより、数値で分析を行うことが難しい分野においては、アンケートを実施することで、市場のニーズを把握し、より精度の高い分析を行うことができると考えられる[3]。

本研究でも、期間や対象会員の増加によって、より有意義な結果を示せるのではないだろうか。また、研究課題として、今後の研究は蓄積されたノウハウを残し、今後の研究や施策に利用し易くすること、ホームセンターの会員数増加、対象期間長期化で更なる大規模なデータ利用が将来予想され、PCのスペック上昇やデータ分割法を更に考慮することをあげる。

参考文献

- [1] 森口元気：『ホームセンターのID付レシートデータの分析』。2014年度南山大学大学院理工学研究科修士論文、2015。
- [2] 村上史也、鈴木政樹：『大規模レシートデータの分析と活用』。2014年度南山大学情報理工学部情報システム数理学科卒業論文、2015。
- [3] 野々垣壇：『商品の最適価格決定問題について』。2010年度南山大学数理情報研究科数理情報専攻修士論文、2011。
- [4] 山本仁志、諏訪博彦、岡田勇、小川祐樹、和泉潔、磯崎直樹、服部進：『ID-POSの購買履歴情報に基づく購買人格の抽出と分析』。経営情報学会2011年春季全国研究発表大会要旨集、pp.1-3、2011。