

スーパーマーケットにおける 優良顧客の 行動に関する分析

中央大学 理工学部経営システム工学科4年
和田昌樹

背景

人口減少、高齢化による食品市場の縮小、
インターネット通販の普及による食品購買手段の多様化



スーパーマーケットの来客数が減少する。売上げが低下する。



既存店舗のお客の取り合いが激化する。
今後、スーパーマーケットが生き残るには、

自社の優良顧客を把握することが課題となっている。

目的

スーパーのID-POSデータを用いて、優良顧客のモデルを構築し、優良顧客に何が寄与しているのかを調べる。
それを元に優良顧客の確保、定着のため、経営戦略の構築を行う。

POSデータの基礎集計

```
graph TD; A[POSデータの基礎集計] --> B[VMSを使った変数の作成]; B --> C[優良顧客のモデル構築]; C --> D[顧客に対する戦略の構築];
```

VMSを使った変数の作成

優良顧客のモデル構築

顧客に対する戦略の構築

使用データの説明

● POSデータの項目

店舗コード	月	日	時間帯	曜日
POSNo	レシート識別No	ID	商品コード	店舗コード
会員買上点数	非会員買上点数	会員買上金額	非会員買上金額	

● 商品分類マスタの項目

商品コード	商品名	部門コード	部門名称
大分類コード	大分類名称	中分類コード	中分類名称
小分類コード	小分類名称	品番コード	品番名称

● 今回使用する項目

日付	ID	時間帯	POSNo	明細識別No
レシート識別No	商品名	大分類名称	中分類名称	小分類名称
品番名称	会員買上点数	非会員買上点数	会員買上金額	非会員買上金額

● 対象店舗: あるスーパーマーケット1店舗

● 期間期間: 2015年6月～9月

● この店舗は24時間営業

● 日付はVMSの機能を使って、2015/〇〇/〇〇という表記の変数を作成。

優良顧客の把握するための 着目点

優良顧客は、購売
商品に特徴がある
のではないかと

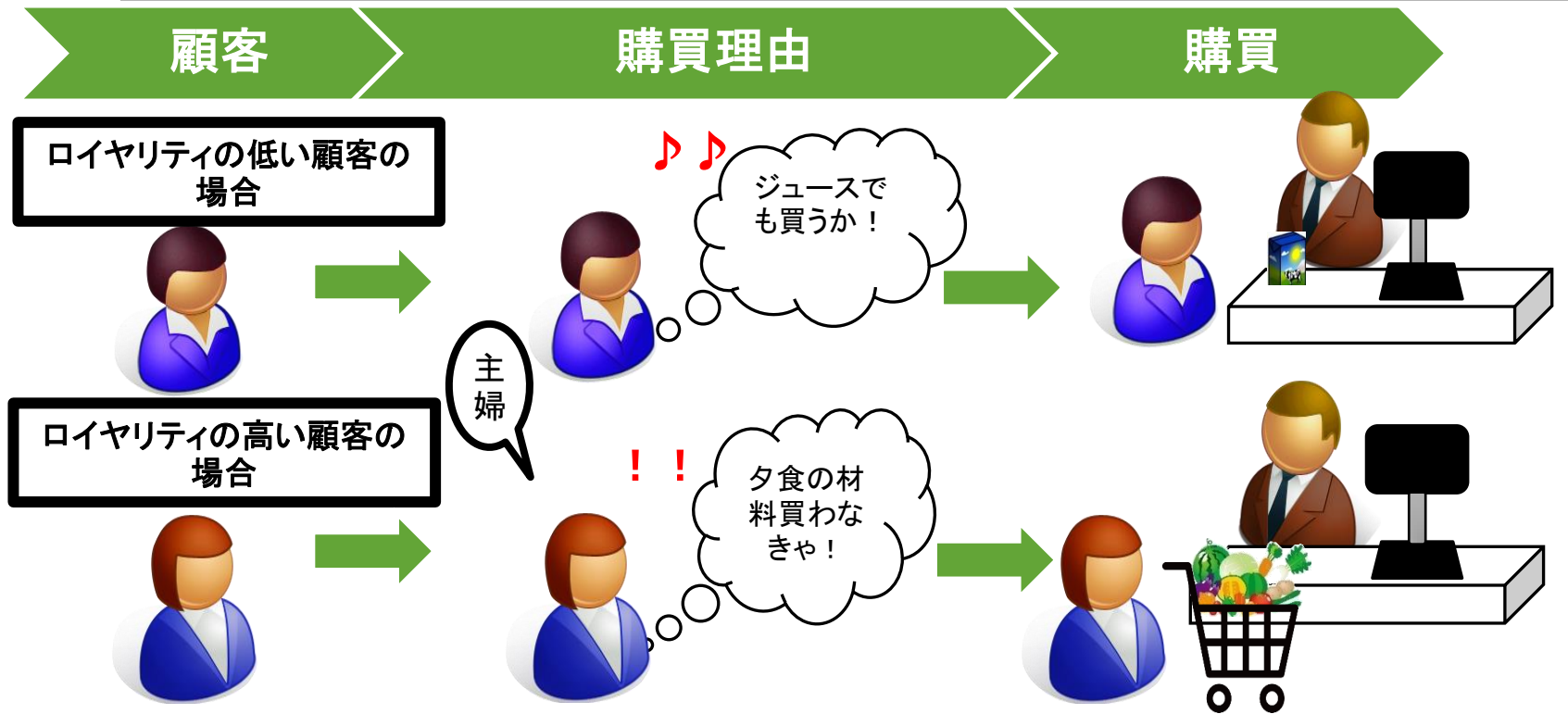


優良顧客は、購売
行動に特徴がある
のではないかと

①優良顧客の顧客行動に対する仮説

- 買う時間帯は決まっているのではないか？
⇒主婦などが夕食時に多くの商品を買うため。
- 混んでいるとき(特売など)に多く出沒するのではないか？
⇒優良顧客は、お店の情報を把握しているため。
- 一回当たりの購買金額が他と比べて高いのではないか？
⇒夕食などの材料をまとめ買いするため。
- 深夜に買うのではないか？
⇒お酒など単価の高いものを買うため。

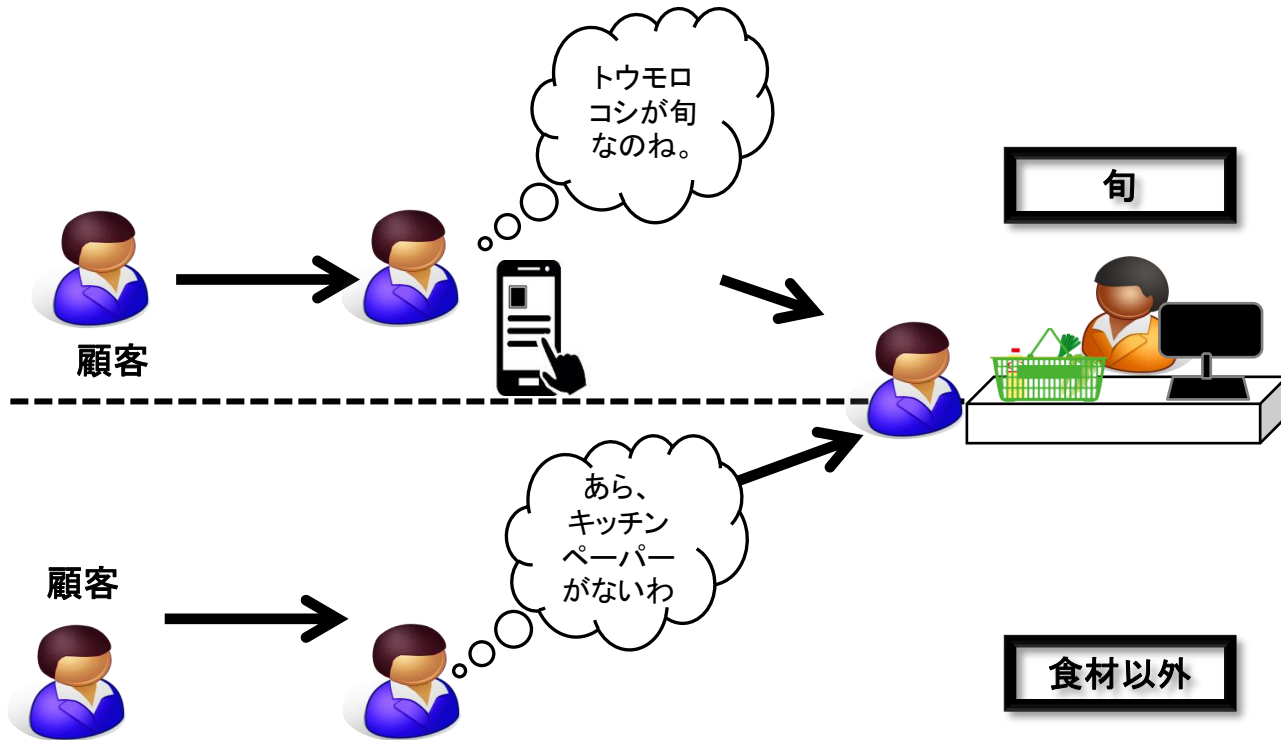
顧客種類別行動のイメージ



②優良顧客の購買品に対する仮説

- 旬の食べ物に反応して、よく買うのではないか？
- 旬の食べ物の時期を把握し、旬の初めに購入するのではないか？
⇒旬のものを買う料理好きは多くの買い物をするため。
- スーパーで食品以外の商品を、お気に入りのスーパーで買うのではないか？
⇒お気に入りのショップのポイントを貯めるため。
- 良く買われている人気商品を多く買うのではないか？
⇒食事を作る際の欠かせないものため。または、特売になっている商品は多く買うため。

優良顧客の購買のイメージ

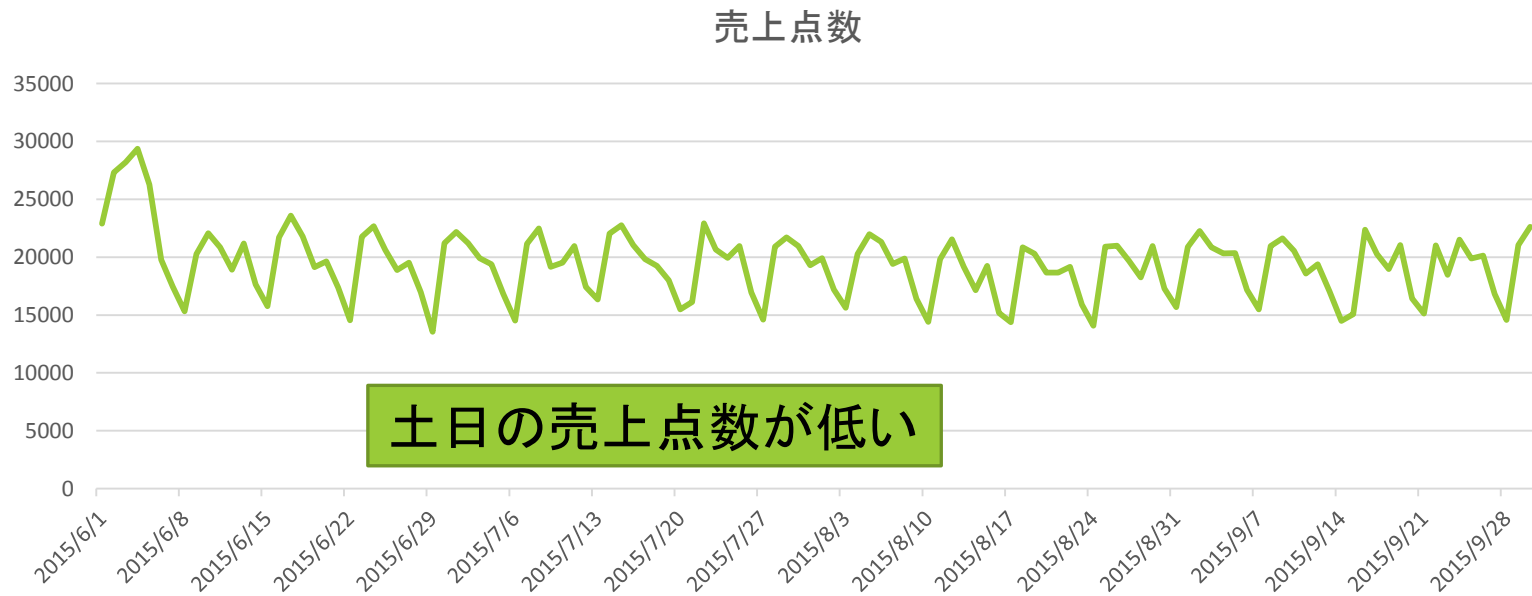


基礎集計

本研究ではVMSを用いて基礎集計、変数作成を行った。

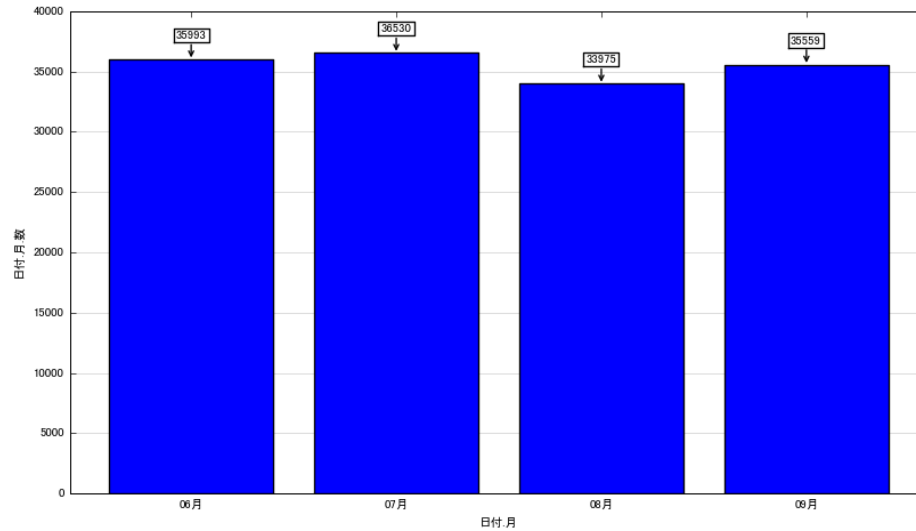
1. 一か月ごとに売上点数に大きな差がないか検証を行った。
2. 各変数を作るための顧客の行動分析を行った。

6月～9月の総合売上点数の推移



各月の来店回数

グラフ



大きな差はない

顧客の分類

顧客の分類をRFM分析を用いて試みた。

・RFM分析

Recency(最新購買日)、Frequency(購買頻度)、Monetary(購買金額)の頭文字を取った分析手法。この3つの要素の組み合わせで顧客にランク付けし、分割する。

今回は、この3つの要素を5段階評価でランク付けする。

VMSを用いてRFM分析を行った。

FscoreとMscoreは、大きすぎる値が基準にならないように、等数分割を行う。

RFM分析

R(最新購買日) 等間隔分割で行った。

ブロック	購買日	あてはまった人数
5	2015年9月6日～9月30日	6239人
4	2015年8月13日～9月5日	1262人
3	2015年7月20日～8月12日	766人
2	2015年6月26日～7月19日	658人
1	2015年6月1日～6月25日	590人

RFM分析

Fは購入回数における人数から、以下のように設定した。

F(購買頻度)

ブロック	購買回数	あてはまった人数
5	25回～	1907人
4	10回～24回	2117人
3	4回～9回	1994人
2	2回～3回	1513人
1	1回	2034人

RFM分析

M(購買金額)

ブロック	購買金額	あてはまった人数
5	25,723円～	1903人
4	9,985～25,713円	1903人
3	3,613～9,984円	1904人
2	1,001～3,612円	1905人
1	0～1,156円	1900人

優良顧客の定義

- 直近に購買したかは、優良顧客分類に関係がないと考えたので、Rは除外してF,Mだけで優良顧客を定義する。
- FとMの相関係数は0.824であった。

今回は、**FMのスコアが共に5である顧客**を優良顧客と定義する。

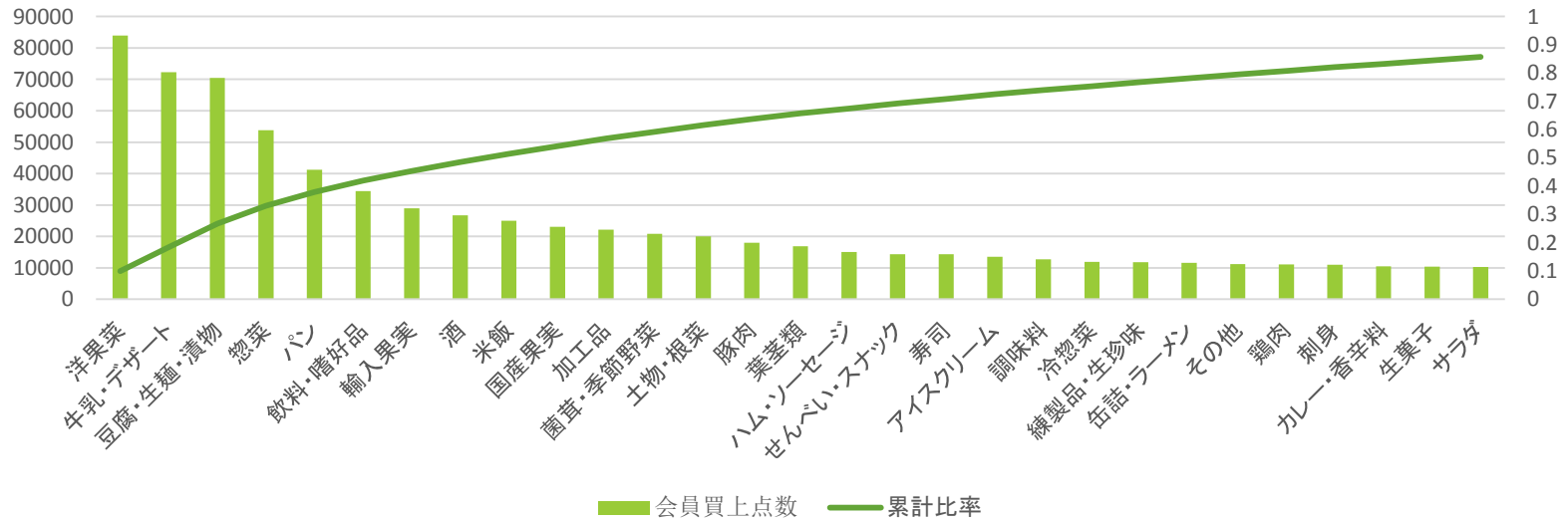


目的変数

属性: 優良顧客(1458人)を1、非優良顧客(8057人)を0とする。

6月～9月の中分類別売上点数

中分類の売上点数



人気商品の平均購入回数

購買品に関する仮説より

良く買われている食品

「牛乳・デザート」、「洋果菜」、「惣菜」、「豆腐・生麺・漬物」

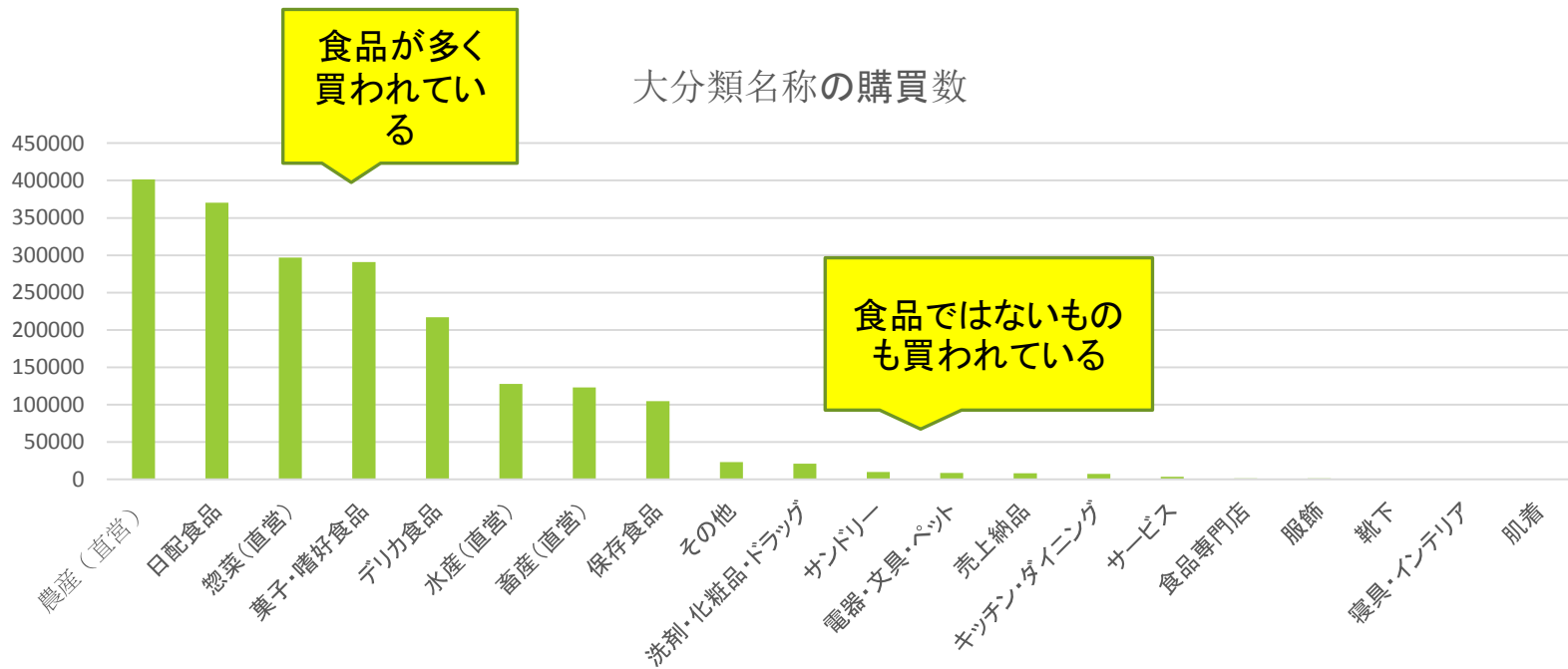
の4つを購入した回数を合計し、来店回数で割ったものを変数とする



変数

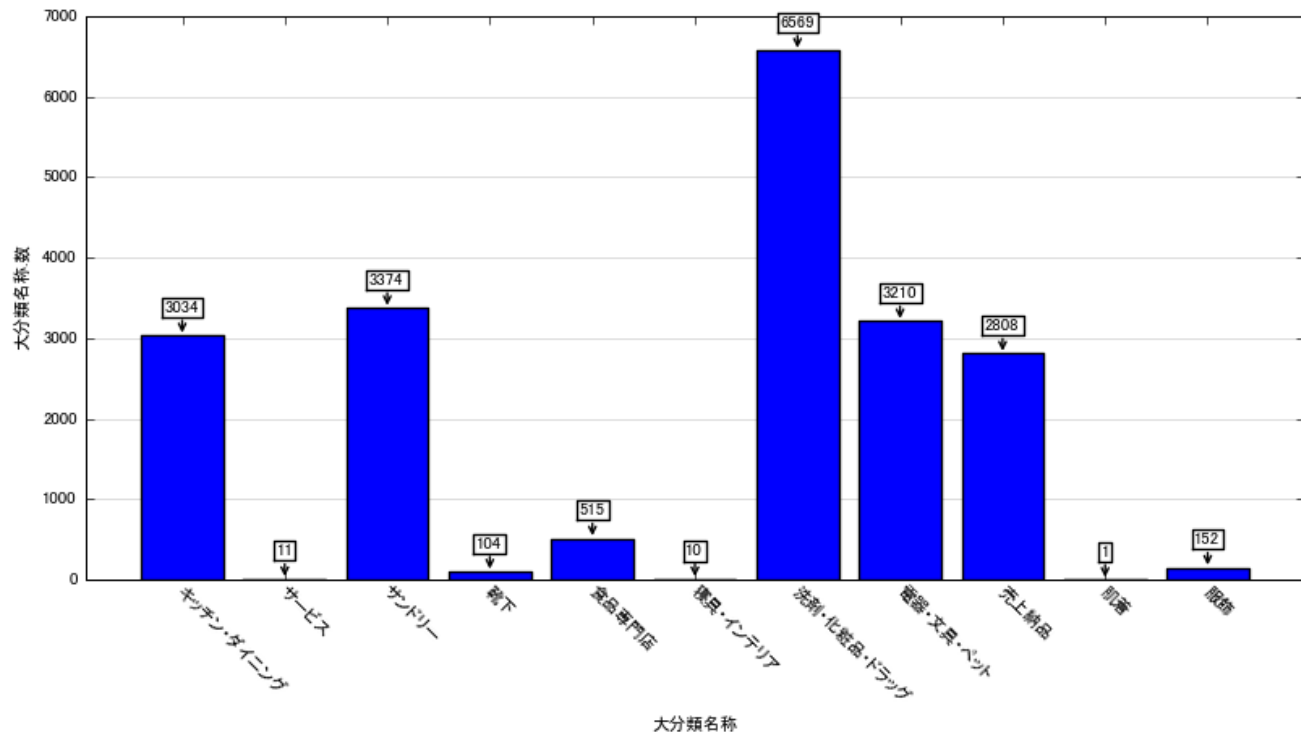
人気商品割合

6月～9月の大分類名称の POSデータ数



大分類の食品ではないもの(その他以下)

主婦(女性)が
買うものと、
その他の二つ
に分けられる



変数の構築

購買に対する仮説より、食品以外の購買に関する変数を作る。

食品以外の購買を「主婦の買うもの」、「主婦以外も買うもの」の二つに分け変数を作る。



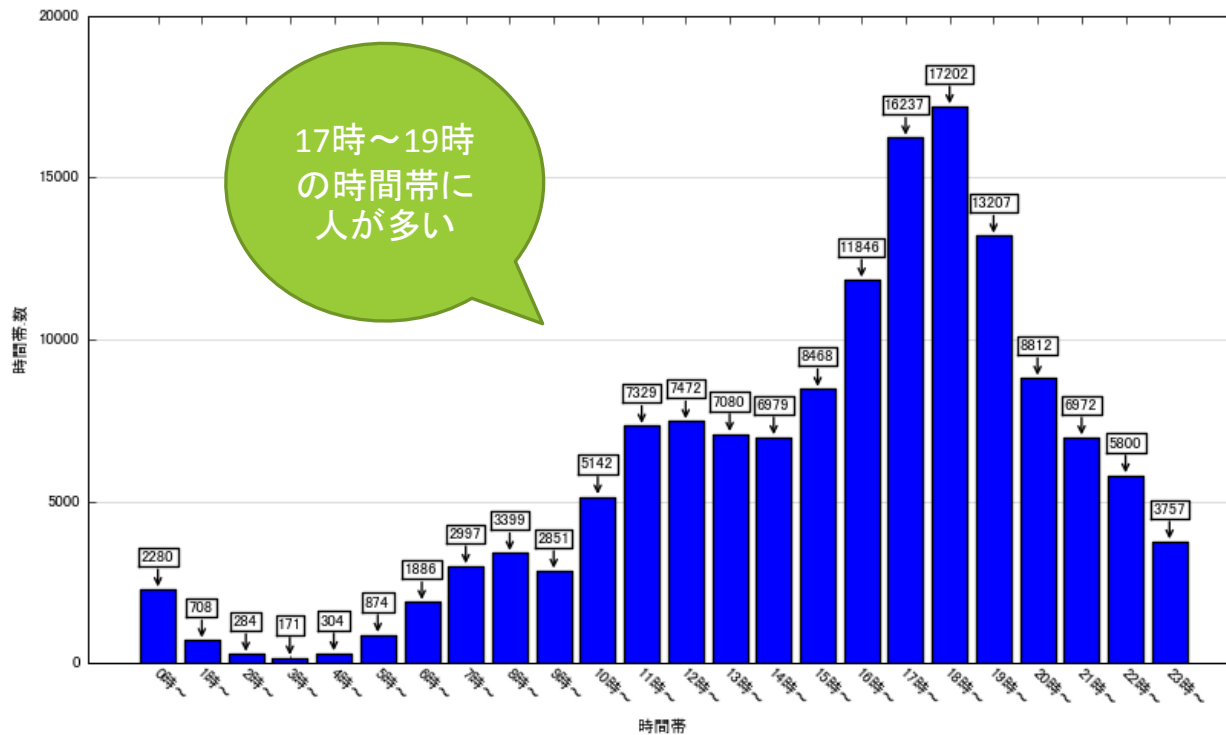
変数

主婦01: 大分類のキッチン・ダイニング、洗剤・化粧品・ドラッグ
に該当する商品を買った

特殊商品01: 大分類の肌着、靴下、電器・文具・ペット用品
に該当する商品を買った

6～9月の時間帯別来店数

グラフ



時間帯に関する変数①

優良顧客になるほど、お店のことが分かっているので、混んでるときに登場するのではないか？

累積比率のグラフから、混んでる時間を判断。

右図の時間帯に来店したら1とする。



変数名

混み01:混んでいる時間に来客

混んでいる時間帯

曜日	時間帯	シート識別No.件数
曜日	18時～	40767
火曜	18時～	39250
月曜	18時～	35439
火曜	17時～	35291
水曜	18時～	34022
木曜	18時～	33654
月曜	17時～	33267
火曜	19時～	33061
月曜	19時～	32267
金曜	18時～	31770
水曜	17時～	30411
水曜	19時～	

時間帯に関する変数②

優良顧客は、非優良顧客と違って、毎回決まった時間帯に買い物して
るのではないか。

深夜01: (1時～6時)に購入した

昼01: (13時～18時)に購入した

夜01: (19時～0時)に購入した

夜01: (19時～0時)に購入した

夏の旬の食べ物

仮説より、

夏の旬の食べ物(トウモロコシ、スイカ、枝豆、冷や麦、アイス)の買われたか、その個数

夏の旬の食べ物(トウモロコシ、スイカ、枝豆)の旬の初めに買っているかを変数にする。

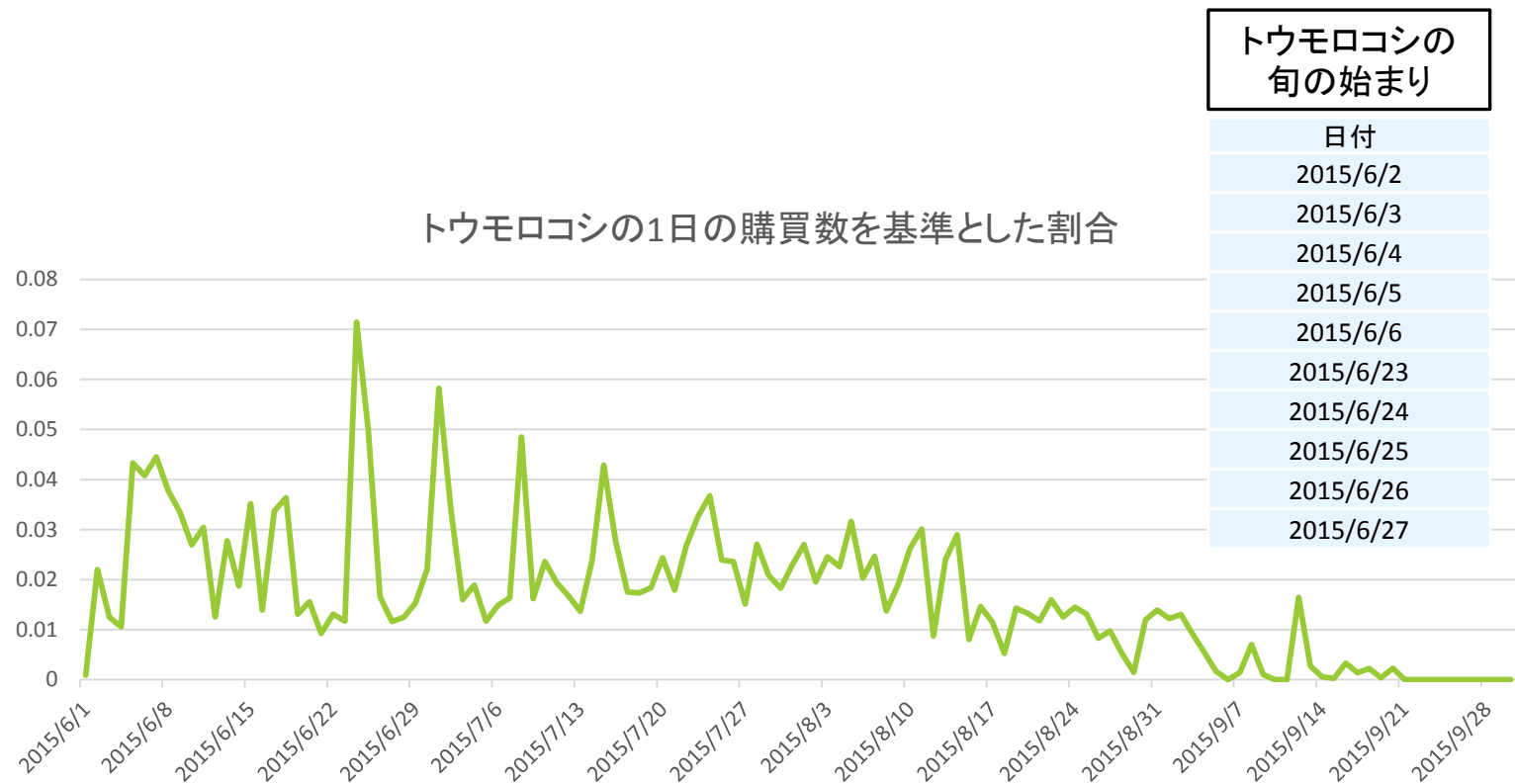


夏の旬購買数: 夏の旬を買った個数

夏旬購買01: 夏の旬を買ったか

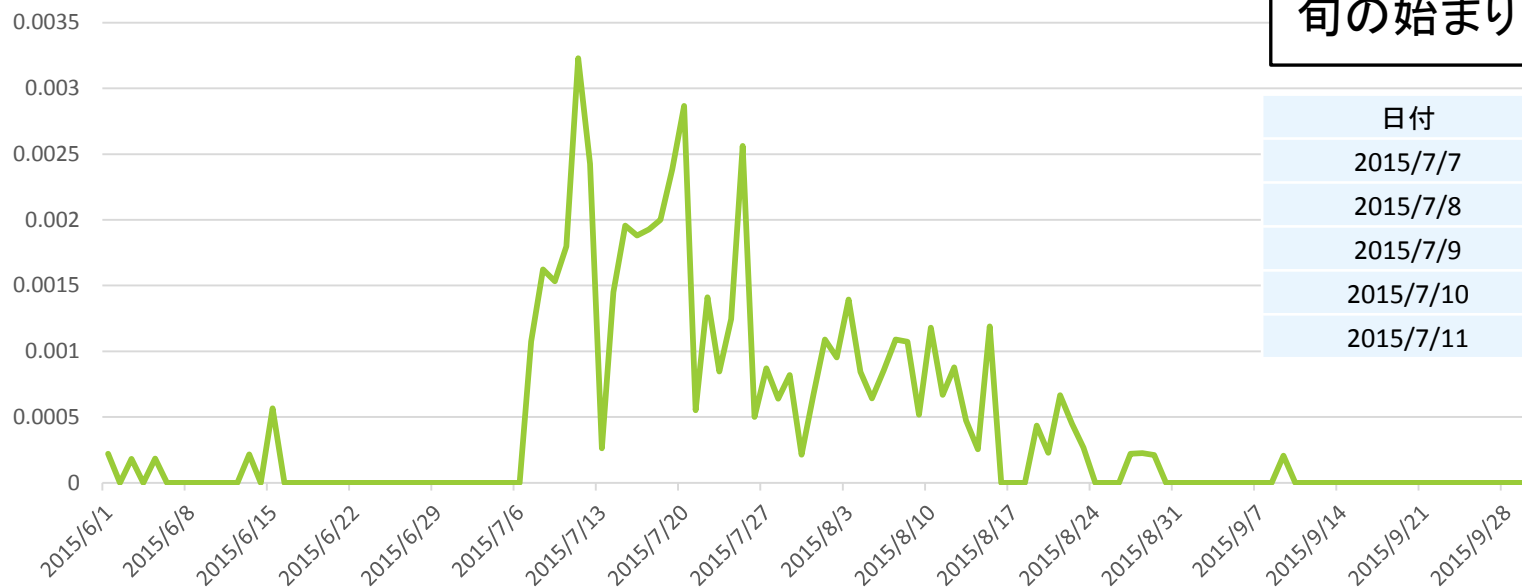
トレンド01: 旬の買われ初めに購入したか

トウモロコシの時系列



スイカの時系列

スイカの1日の購買数を基準とした割合



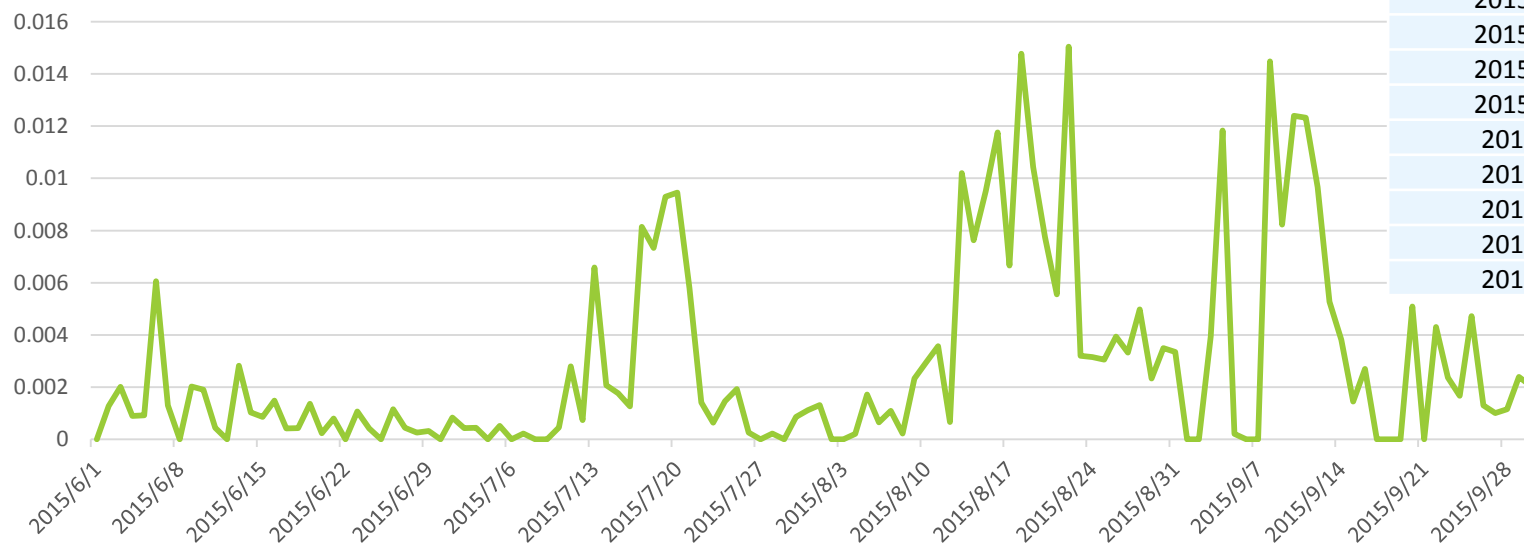
枝豆の時系列

枝豆の旬の始まり

日付

2015/6/2
2015/6/3
2015/6/4
2015/6/5
2015/6/6
2015/7/11
2015/7/12
2015/7/13
2015/7/14
2015/7/15
2015/8/12
2015/8/13
2015/8/14
2015/8/15
2015/8/16
2015/9/3
2015/9/4
2015/9/5
2015/9/6
2015/9/7

枝豆の1日の購買数を基準とした割合



優良顧客のモデルの作成

R-studioを使ってロジスティック回帰分析を使ってモデルを構築する。

ロジスティック回帰分析

$$\log \left(\frac{p}{1-p} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p$$

ロジット関数

$$\frac{p}{1-p} = \exp (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p)$$

オッズ

分析手順

- ① データを学習用データとテストデータに分ける。
- ② glm関数を用いてロジスティック回帰分析を行う。
- ③ ①、②を変数を制限して何回も行う。
- ④ ROCカーブ、AIC、正解率から考えて最も良いモデルを作成した。
- ⑤ 最も当てはまりの良いモデルのオッズ比を求める。

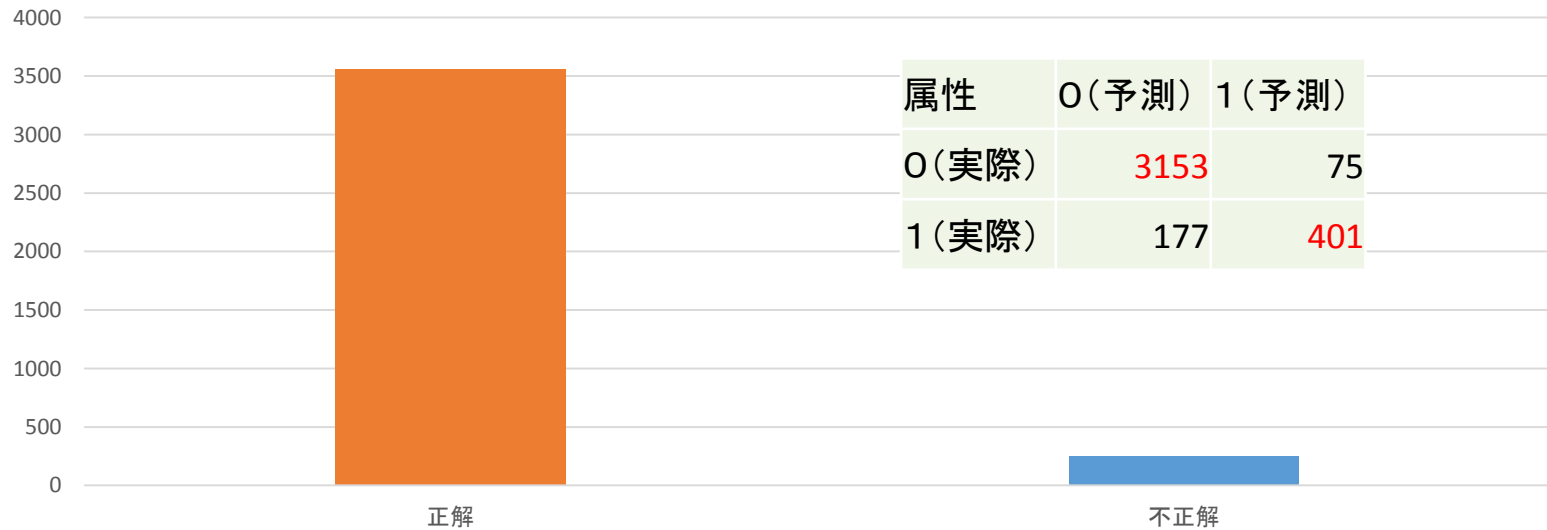
優良顧客の ロジスティクス回帰モデル

係数がロジスティック回帰分析の
係数(β)を示す。
すべての変数が有意といえる。

	係数	z値	P値
(Intercept)	-4.798	-13.21	0.1%有意
夏の旬購買数	0.095	5.63	0.1%有意
平均購買金額	-0.001	-9.71	0.1%有意
人気商品割合	0.374	25.04	0.1%有意
深夜01	0.642	3.44	0.1%有意
朝01	0.497	3.71	0.1%有意
夜01	0.321	2.16	5%有意
込み01	0.983	6.10	0.1%有意
主婦商品01	0.682	5.05	0.1%有意
特殊商品01	0.797	6.02	0.1%有意
性別コード	-0.339	-2.08	5%有意
トレンド01	0.401	2.13	5%有意

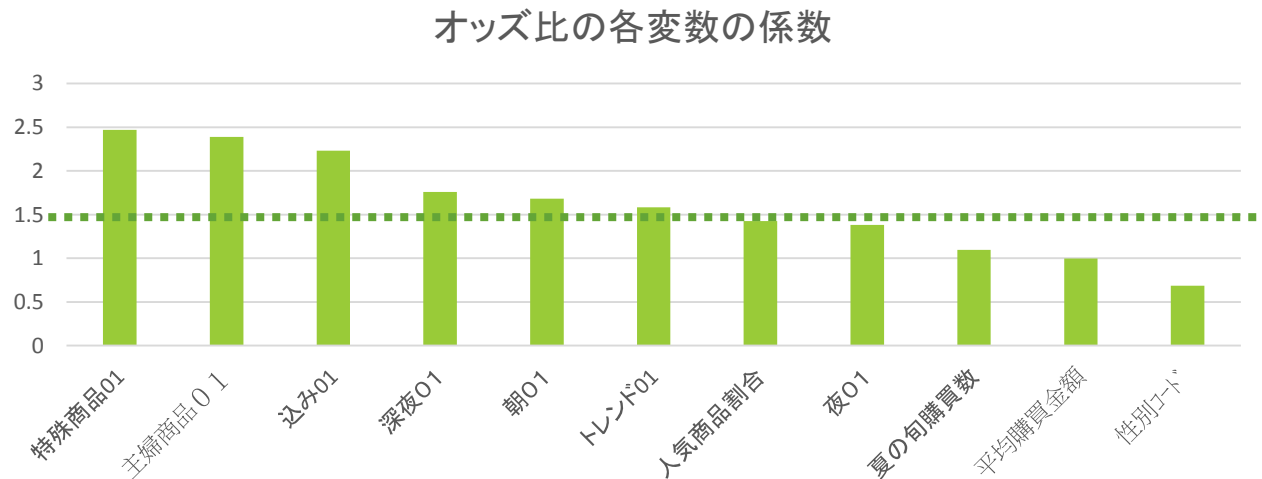
優良顧客の ロジスティック回帰モデル

ロジスティックモデルの正解数と不正解数



ロジスティック回帰モデルの結果

● オッズ比



● 優良顧客分類の予測の精度

93%になった。(データサンプルを10回変更させた平均)

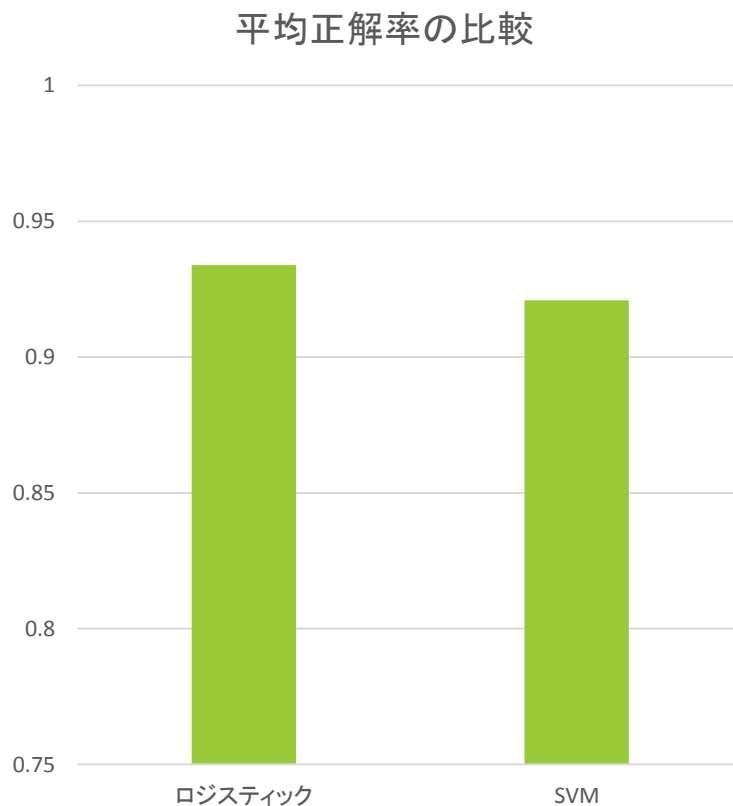
SVMとの正解率の比較

SVM(サポートベクターマシン)
分類と回帰問題を主とした教師あり機械学習方法である。

ロジスティックモデルとの予測結果の比較のために用いた。

今回、SVMは線形SVMを使用し、カーネル関数はガウシアンを用いた。変数はロジスティックモデルと同じものを用いている。

結果は、ロジスティックモデルの方が予測の精度が高かった。



優良顧客の考察

◆ 正解率が高いことから、モデルの当てはまりは良いといえる。オッズ比から優良顧客は、非優良顧客と比べて次のことが言える。

1. 混んでいるときに来る
(優良顧客になる確率を2.7倍に引き上げている)
2. キッチン用具、化粧品を買う
(優良顧客になる確率を1.9倍に引き上げている)
3. 旬のトレンドに敏感である
(優良顧客になる確率を1.5倍に引き上げている)
4. 人気商品をよく買う
(優良顧客になる確率を1.4倍に引き上げている。)



優良顧客は主婦が多い。

優良顧客の考察

5. 深夜にも来る

(優良顧客になる確率を1.9倍に引き上げている。)



優良顧客には深夜に来るような主婦でない人もいる。

6. 食品以外の商品を買う。

(優良顧客になる確率を2.2倍に引き上げている。)



ポイントのためにドラッグストアなどで
買わずに、この店舗を利用している。
⇒ポイントを大事にしている。

経営戦略1～優良顧客を離さない

● ポイントを使った戦略

優良顧客を飽きさせないように毎月のポイントデーなどを作る。商品ごとにポイントの割合を高める。

● 旬の食材をできるだけ早く取り寄せる

「買いたいののに旬の食材がない」ということを避ける。

● 平日の18時前後のサービスを増やす。その時間帯のレジの回転を上げる

試食やポイント還元ボーナス、特売などのサービス。回転数を上げ、優良顧客のストレスを軽減させる。

● 旬の食べ物をポイントで値下げできるようにする

ポイントを貯めて使う喜びを作る。

経営戦略2～優良顧客を獲得する

●旬の商品のアピール

トレンドを早めに把握して買ってもらおう。旬の食べ物の解禁を顧客から顧客へ伝播してもらおう。

●ポイントカードのアピール。

●ポイントの商品ごとだけでなく、来店回数に応じてポイント還元するようにする

来店回数と購買金額の相関が高いため。ちょっとした食品(飲料、お菓子など)や食品以外をほかのドラッグストアで買うのではなく、α店舗で買ってもらおう確率が増える。

●ポイントカード発行してからしばらくの期間はインセンティブを作る。

新規顧客にお店に数回着て、ポイントを貯めてもらおう。
ポイントを使う喜びを知ってもらおう。ポイントを貯めたし、この店を使うかという心理状態にさせる。

参考文献

- 金明哲：“Rによるデータサイエンス”
- 浅野熙彦：“最新マーケティング・サイエンスの基礎”
- 涌井貞美：“意味が分かる統計解析”