

製品分類による顧客セグメンテーションの方法

— Customer Segmentation Using Existing Product Category Schemes —

SAS インスティテュートジャパン

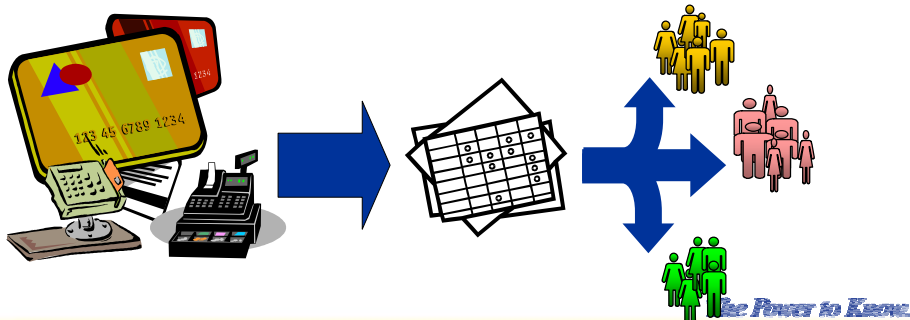
カスタマーサービス本部 データサイエンスグループ

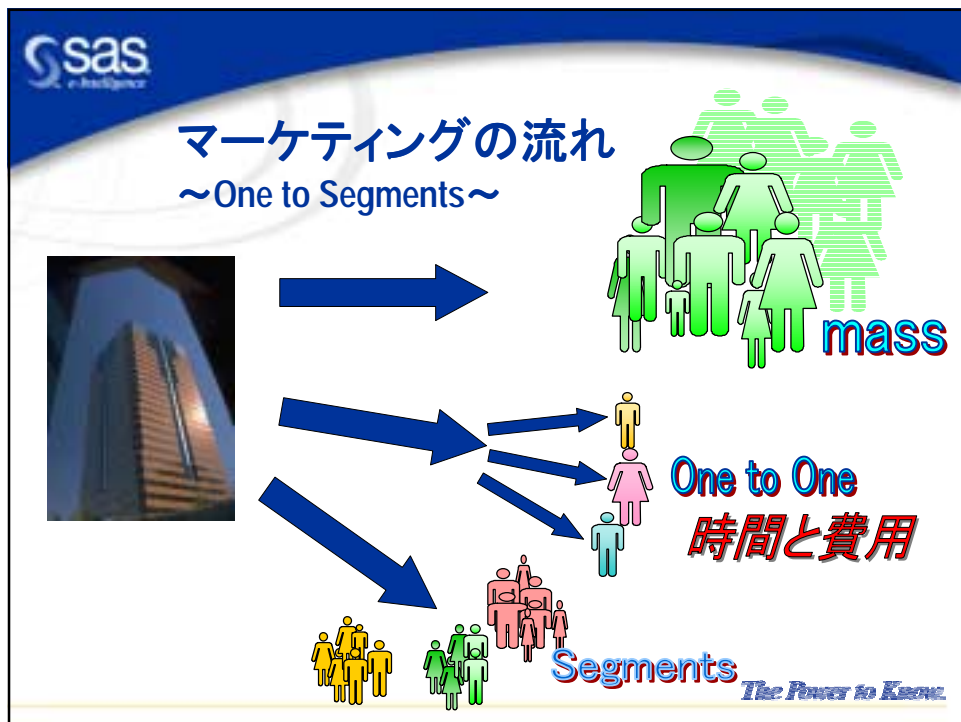
中西 由紀

Copyright © 2000, SAS Institute Inc. All rights reserved.

本日の内容

企業にある顧客の取引履歴データ(POSデータ)を、
製品分類を用いて数値化し、顧客セグメンテーションを生成する方法をご紹介します。





SAS家電データの加工

SAS家電



顧客ID	買上日	時間	製品
1	2001/7/25	13:20	100210330
2	2001/7/25	14:00	100200340
3	2001/7/25	18:00	200540480
4	2001/7/26	13:20	200440550
5	2001/7/26	14:00	300320290
6	2001/7/27	13:20	300339864
7	2001/7/27	14:00	100200280
8	2001/7/27	18:00	200440550
9	2001/7/27	21:00	200440550
:	:	:	:

製品の上3桁・・製品群
4桁～6桁・・メーカー
7桁～9桁・・製品

ソニー製デスクトップPC

CUST_ID	PROD_ID	PROD_CAT
3	1	SY
3	10	SY
4	3	SC
1	4	HT
4	4	HT
3	5	HY
1	6	HY
2	6	HY
3	6	HY
4	7	HST
1	2	HFN
1	8	HCN
3	9	KYA
5	9	KYA
5	11	KFA
5	12	KCA

カテゴリー1	カテゴリー2	カテゴリー3
ハードウェア(H)	キャノン(C)	AV機器(A)
AV/家電(K)	富士通(F)	デスクトップPC(D)
周辺機器(S)	シャープ(S)	ノートPC(N)
	東芝(T)	プリンター(P)
	ソニー(Y)	スキャナー(S)
		電話(T)
		テレビ(V)

The Power to Know.

SAS家電データの加工

カテゴリーごとに個数を集計する

CUST_ID	PROD_ID	PROD_CAT
3	1	SY
3	10	SY
4	3	SC
1	4	HT

CUST_ID	C1_H (ハードウェア個数)	C1_K (AV/家電機器個数)	C1_S (周辺機器個数)
1	4	0	0
2	1	0	0
3	2	1	2
4	2	0	1
5	0	3	0

カテゴリー1(製品群)

CUST_ID	C2_C	C2_F	C2_S	C2_T	C2_Y
1	1	1	0	1	1
2	0	0	0	0	1
3	0	0	0	0	5
4	1	0	1	1	0
5	1	1	0	0	1

カテゴリー2(メーカー)

CUST_ID	C3_A	C3_D	C3_N	C3_P	C3_S	C3_T	C3_V
1	0	1	2	0	0	0	1
2	0	1	0	0	0	0	0
3	1	1	1	1	1	1	0
4	0	0	0	0	1	1	1
5	3	0	0	0	0	0	0

カテゴリー3(製品)

The Power to Know.

加工データを解釈しよう

カテゴリー1での
顧客ID1、2のデータの
傾向を見てみよう。

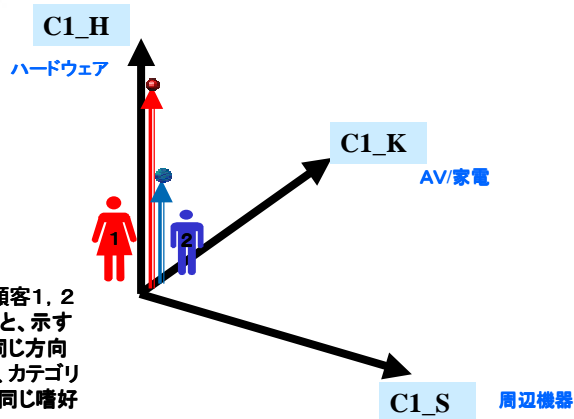
CUST_ID	C1_H	C1_K	C1_S
1	4	0	0
2	1	0	0
3	2	1	2
4	2	0	1
5	0	3	0

ハードウェア

AV/家電

周辺機器

ユークリッド空間で、顧客1、2
のベクトルを見てみると、示す
点は異なるが、両者同じ方向
を向いている。つまり、カテゴ
ー1に関しては、両者同じ嗜好
であるといえる。



The Power to Know.

SAS Version8.1 を使用してのクラスタ分析 ～クラスタ分析手法の種類～

PROC FASTCLUS

非階層的方法,k-means法
履歴データ等、大規模なデータの場合に使用を薦める
外れ値の検出も可能

PROC CLUSTER

階層的方法
非階層的方法より柔軟性がある
100以下のOBSの場合に使用を薦める

Step1
クラス中数を絞る

Step2
最終クラスタの生成

The Power to Know.

SAS Version8.1 を使用してのクラスタ分析 ～FAST CLUS の留意点～

PROC FASTCLUS の特徴

初期シード値を設定し、その点を基に仮クラスタを生成し、変数間の距離を計り、最終クラスタを決定する。よって、クラスタの結果は、初期シード点に大きく影響を受ける。

結果の検証

検証方法

1. データを分割し、クラスタリングを実行、結果を比較する。
2. 初期シード点を数回変えて実行し、結果を比較する。

The Power to Know.

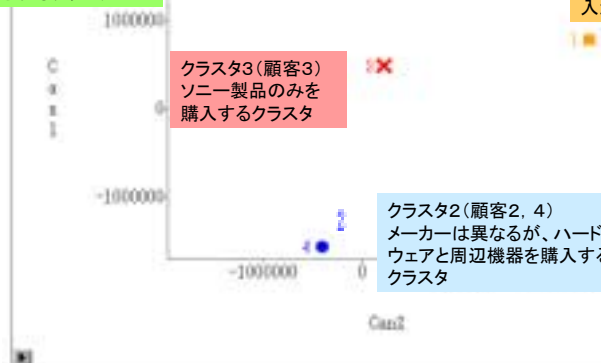
SAS Version8.1 を使用してのクラスタ分析 ～SAS家電データのクラスタリング結果～

クラスタ4 (顧客5)
AV・家電に偏った購入
が見られるクラスタ

クラスタ1 (顧客1)
ハードウェアに偏った購入
が見られるクラスタ

クラスタ3 (顧客3)
ソニー製品のみを
購入するクラスタ

クラスタ2 (顧客2, 4)
メーカーは異なるが、ハード
ウェアと周辺機器を購入する
クラスタ



The Power to Know.

SAS Version8.1 を使用してのクラスタ分析 ～収益性を考える～

クラス	顧客数	顧客比率	収益(単位100万円)	収益/顧客数	収益率
ハードウェア	2,090	39.4%	600	0.29	21.3%
ハードウェア&周辺機器	893	16.8%	550	0.61	44.9%
ソニー製品	1,125	21.2%	450	0.4	29.4%
AV/家電機器	1,200	22.6%	80	0.06	4.4%
合計	5,308	100.0%	1,680	1.36	100.0%

ターゲット顧客は

顧客比率？



収益率？

The Power to Know.

まとめ

顧客セグメンテーションを行うにあたり、製品分類を用いるのは1つの有効な手段である。

収益に見合ったターゲットセグメントを見つけ、販売促進活動を行いましょう。



The Power to Know.

