

株式会社 XXX様

データ・マイニングのご参考 「Webサイト分析の例」

サンプルデータを用いて、次の例を分析しています。

- ・現状の把握 ―― 現状分析
- ・発注・在庫管理 ―― 売上げ予測
- ・マーケティング ―― 顧客管理、トレンド

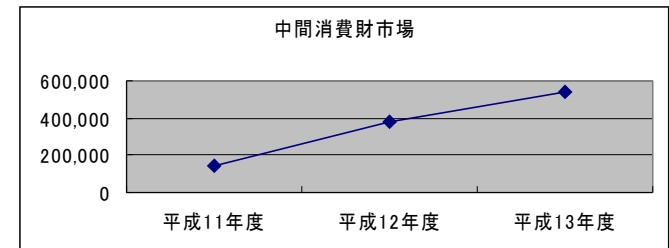
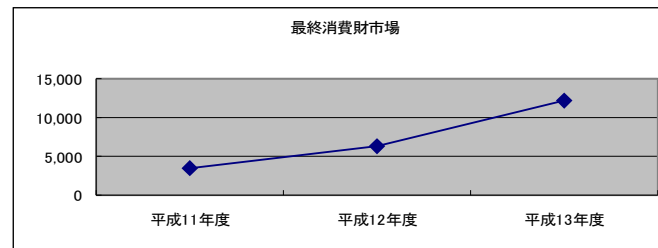
2004年5月
Sunlight co.,ltd.

「インターネットコマース」の市場規模

(平成14年度 通信白書より)

	平成11年度	平成12年度	平成13年度
最終消費財市場	3,500	6,233	12,218
中間財市場	144,000	381,000	539,000

単位:億円



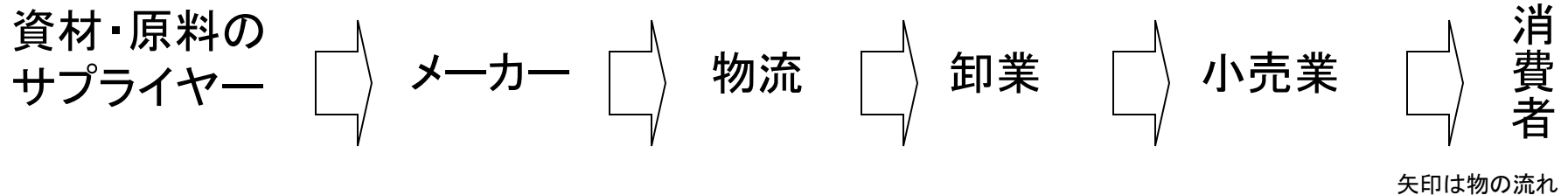
対前年度成長率

	平成12年度	平成13年度
最終消費財市場	78.09	96.02
中間財市場	164.58	41.47

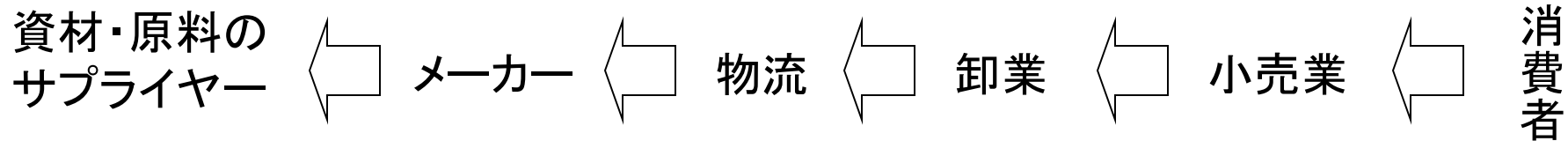
優良WebサイトとSCMの共通性

- ・リードタイムの短縮
- ・在庫の適正化
- ・物流コストの削減

SCM(SupplyChainManagement)とプレイヤー



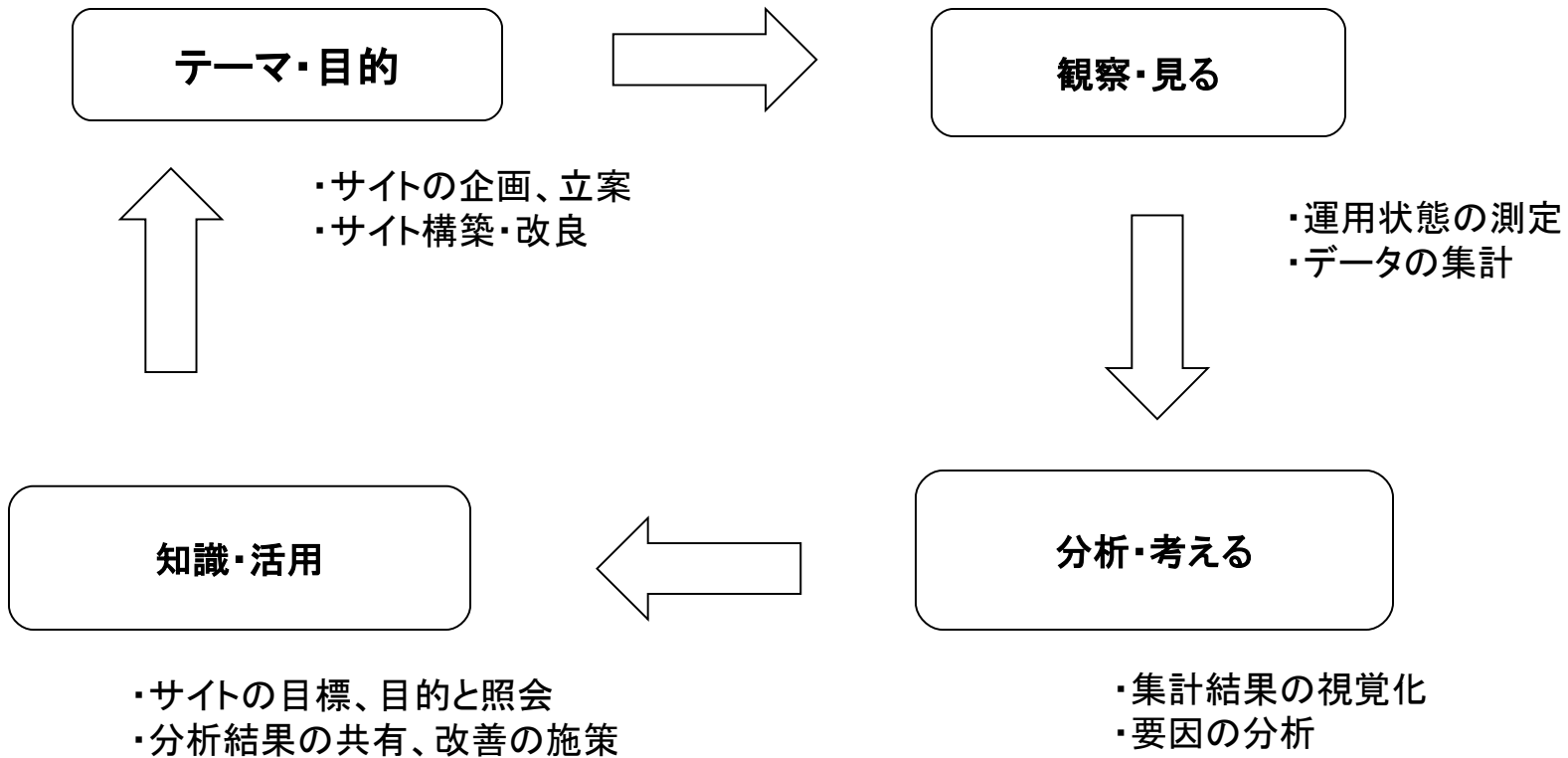
CRM(Customer Relationship Management)と情報の流れ



CRMのポイント

- ・顧客満足度(不満足度)
- ・ビジネスコスト

データ分析は、循環するプロセスの一翼



分析に用いた月間売上データ

日付	売上げ額	サイトアクセス数	販売個数	商品A売上個数	商品A価格	その他商品売上げ個数	その他商品平均額	キャンペーン(割引額)
11月1日	50400	1710	18	6	3000	12	2700	1
11月2日	68800	2090	22	8	3000	14	3200	1
11月3日	69400	1710	18	11	3000	7	5200	1
11月4日	65800	1995	21	7	3000	14	3200	1
11月5日	71400	2185	23	11	3000	12	3200	1
11月6日	110200	4275	45	14	3000	31	2200	1
11月7日	46100	1425	15	4	3000	11	3100	1
11月8日	65100	2185	23	10	3000	13	2700	1
11月9日	69000	1995	21	9	3000	12	3500	1
11月10日	49800	1425	15	3	3000	12	3400	1
11月11日	96000	3040	32	11	3000	21	3000	1
11月12日	127000	3705	39	14	3000	25	3400	1
11月13日	134700	4560	48	17	3000	31	2700	1
11月14日	59500	2090	22	9	3000	13	2500	1
11月15日	44000	1140	12	4	3000	8	4000	1
11月16日	44200	1235	13	0	0	13	3400	1
11月17日	58800	1710	18	6	3000	12	3400	1
11月18日	91200	2755	29	8	3000	21	3200	1
11月19日	77000	2185	23	3	3000	20	3400	1
11月20日	57800	1520	16	2	3000	14	3700	1
11月21日	75600	1995	21	0	0	21	3600	1
11月22日	60100	1615	17	4	3000	13	3700	1
11月23日	53200	1805	19	0	3000	19	2800	1
11月24日	34500	760	8	1	3000	7	4500	1
11月25日	206850	7220	76	31	3000	45	4500	0.7
11月26日	156240	6080	64	25	3000	39	3800	0.7
11月27日	197330	7220	76	27	3000	49	4100	0.7
11月28日	64500	2375	25	4	3000	21	2500	1
11月29日	54400	1805	19	6	3000	13	2800	1
11月30日	22800	570	6	2	3000	4	4200	1

月間サマリー

月間サイトアクセス数	76,380
月間売上額数	¥2,381,720
平均単価	¥2,962
月間売上個数	804
商品A売上個数	257
その他商品売上個数	547
商品A比率	31.97%
顧客数	572
コンバージョン率	0.75%

売上げ額に影響を与えた主要因

サイトアクセス数
商品A価格
その他の商品の売上げ個数

売上げ額の説明式 [信頼度97%]

$$\text{売上げ額} = 19,994 + 27.9 * \text{サイトアクセス数} - 1.7 * \text{商品A価格} - 0.7 * \text{その他商品売上げ個数}$$

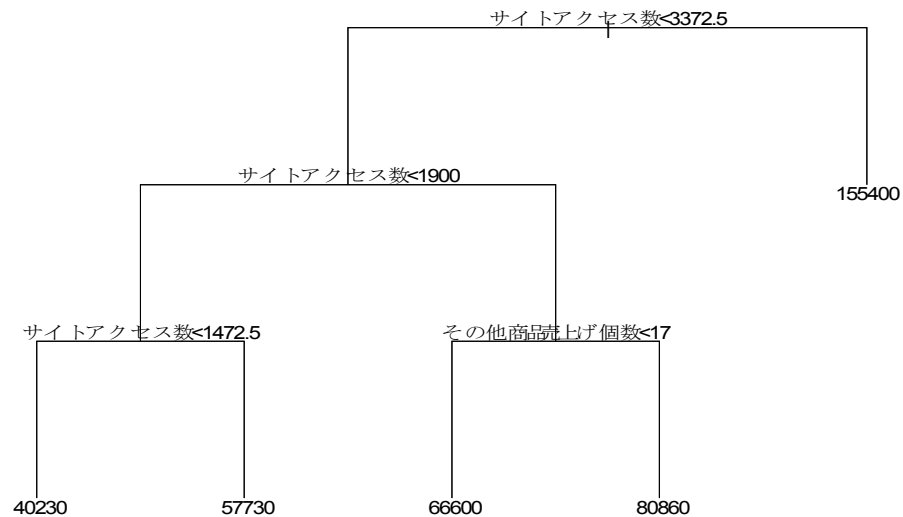
月間の売上げ額の異常性なし

(今月の売上げ額:238万円)

年間平均	271万円
標準偏差	32万円
信頼率80%の 正常範囲	230~312万円

*参考

月間売上げ要因分析(決定木)



*参考

月間売上げ要因分析
(全ての変数を用いた重回帰分析)

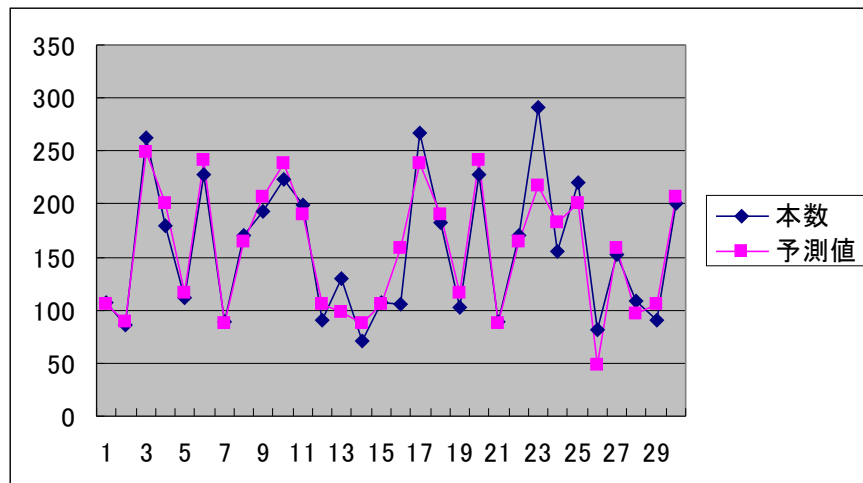
回帰統計				
重相関 R	0.995122159			
重決定 R2	0.990268112			
補正 R2	0.987171602			
標準誤差	5033.167642			
観測数	30			
	係数	標準誤差	t	P-値
切片	-112348.8887	25854.50694	-4.345427624	0.000259484
サイトアクセス数	-221.7504079	94005558.13	-2.35891E-06	0.999998139
販売個数	-8777.142857	3773547932	-2.32597E-06	0.999998165
商品A売上個数	32709.15914	11284109148	2.89869E-06	0.999997713
商品A価格	-1.372300499	1.438291163	-0.954118702	0.350390419
その他商品売上げ個数	32848.10668	11284109148	2.91101E-06	0.999997704
その他商品平均額	10.6610706	1.863504086	5.720980533	9.3793E-06
キャンペーン(割引額)	83855.66538	21715.15511	3.861619452	0.000844722

月間売上げ要因分析
(必要な変数に絞った重回帰分析)

概要				
回帰統計				
重相関 R	0.986544			
重決定 R2	0.973269			
補正 R2	0.970185			
標準誤差	7673.134			
観測数	30			
	係数	標準誤差	t	P-値
切片	19994.82	7029.39	2.84446	0.008556
サイトアクセス数	27.8694	3.628957	7.679728	3.77E-08
商品A価格	-1.67262	2.022657	-0.82694	0.415795
その他商品売上げ個数	-385.385	577.7216	-0.66708	0.5106

ある清涼飲料水の月間売上げ本数と実績と予測

日付	本数	単価	天気	曜日	予測本数
1	107	318	晴れ	木	106
2	86	318	晴れ	金	90
3	263	295	曇り	土	249
4	179	295	曇り	日	200
5	111	318	曇り	月	116
6	228	295	曇り	火	242
7	89	318	晴れ	水	87
8	171	308	晴れ	木	164
9	193	298	晴れ	金	206
10	224	295	晴れ	土	239
11	199	295	晴れ	日	190
12	91	318	晴れ	月	105
13	130	318	晴れ	火	97
14	71	318	晴れ	水	87
15	107	318	晴れ	木	106
16	106	308	曇り	金	158
17	267	295	晴れ	土	239
18	183	295	晴れ	日	190
19	102	318	曇り	月	116
20	228	295	曇り	火	242
21	89	318	晴れ	水	87
22	171	308	晴れ	木	164
23	291	298	曇り	金	217
24	155	295	雨	土	182
25	220	295	曇り	日	200
26	81	318	雨	月	49
27	153	298	雨	火	158
28	109	318	曇り	水	97
29	90	318	晴れ	木	106
30	201	298	晴れ	金	206



予測に用いた重回帰式 [信頼度87%]

		単価	天気	曜日
売上げ本数 =	1852.6	- 5.8*単価	+ 0(雨) 66.7(曇) 56.3(晴)	+ 0(日) 49.1(月) 41.1(火) 30.1(水) 49.8(木) 33.6(金) 48.7(土)
影響度指数		5.8	66.7	49.8

分析に用いた顧客データ

氏名	平均金額	合計金額	来店回数	購入数
上田	¥127,000	¥127,000	1	1
岡本	¥3,600	¥21,600	4	6
都築	¥28,400	¥85,200	2	3
山口	¥36,200	¥108,600	1	3
垂水	¥8,900	¥44,500	3	5
越智	¥41,000	¥82,000	1	2
水田	¥71,000	¥71,000	1	1
森	¥8,200	¥41,000	2	5
山本	¥6,700	¥13,400	2	2
浅野	¥51,100	¥102,200	1	2
伊藤	¥36,100	¥108,300	3	3
大津	¥7,100	¥14,200	2	2
戸瀬	¥98,000	¥294,000	1	3
中谷	¥24,100	¥72,300	2	3
柴田	¥97,000	¥194,000	1	2

クラスタリングの結果

氏名	平均金額	合計金額	来店数	購入数
上田	¥127,000	¥127,000	1	1
水田	¥71,000	¥71,000	1	1
越智	¥41,000	¥82,000	1	2
浅野	¥51,100	¥102,200	1	2
柴田	¥97,000	¥194,000	1	2
都築	¥28,400	¥85,200	2	3
山口	¥36,200	¥108,600	1	3
伊藤	¥36,100	¥108,300	3	3
戸瀬	¥98,000	¥294,000	1	3
中谷	¥24,100	¥72,300	2	3
岡本	¥3,600	¥21,600	4	6
垂水	¥8,900	¥44,500	3	5
森	¥8,200	¥41,000	2	5
山本	¥6,700	¥13,400	2	2
大津	¥7,100	¥14,200	2	2

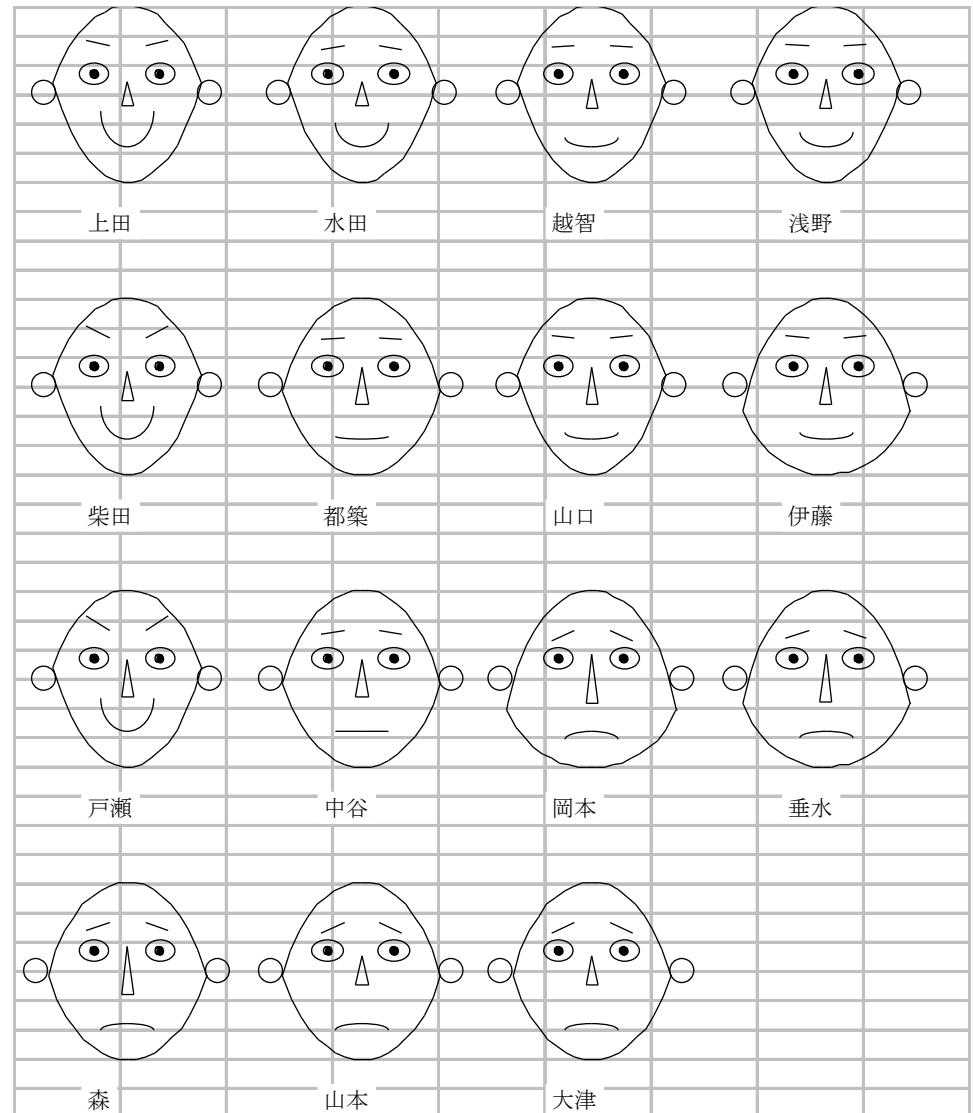
優良顧客だが、リピートになっていない		
一般的な優良顧客		
一般的な顧客		

*** 参考**

チャーノフの顔グラフ

- 眉の角度 — — — 合計金額
- 口の曲率 — — — 平均金額
- 鼻の長さ — — — 購入数
- 顔半分の — — — 来店数
- 楕円の離芯率

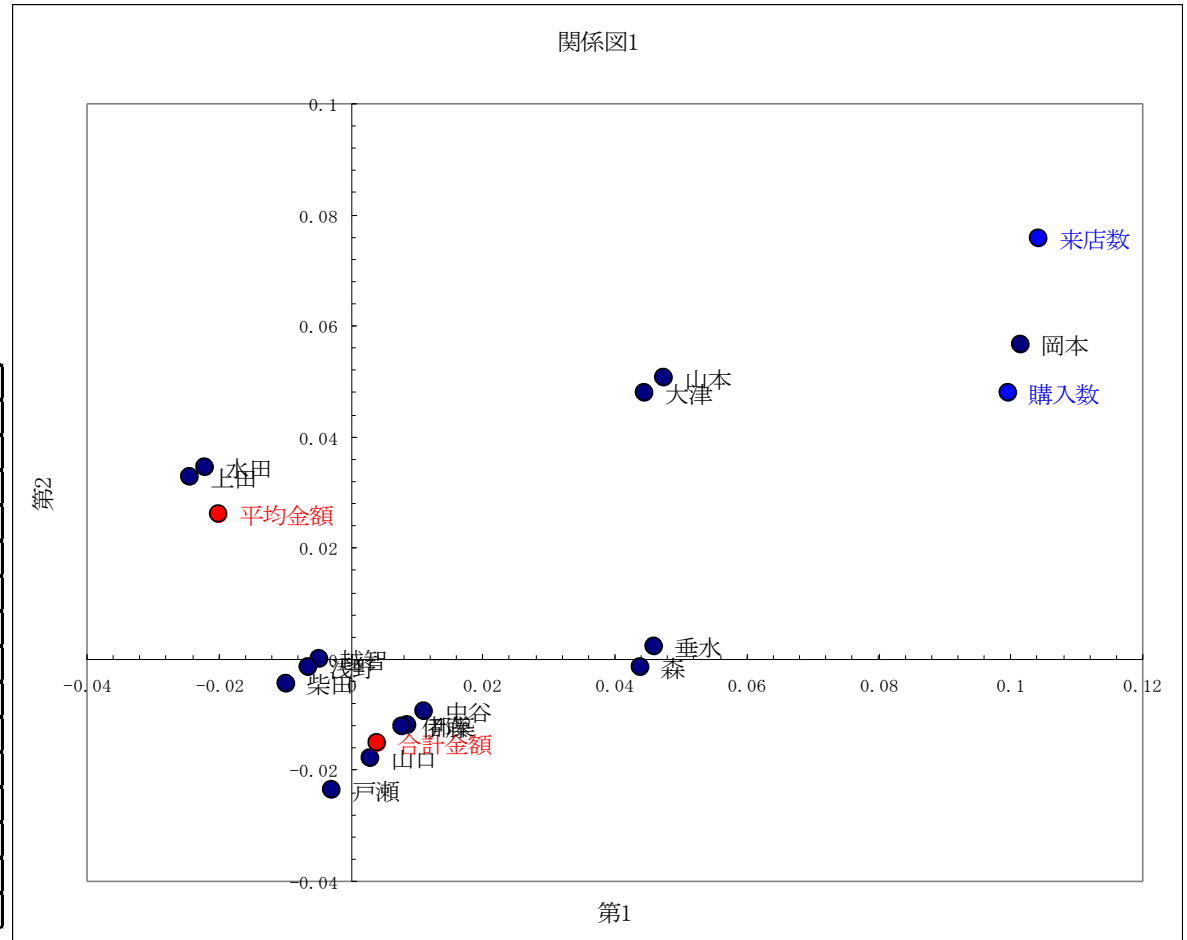
氏名	平均金額	合計金額	来店数	購入数
上田	¥127.000	¥127.000	1	1
水田	¥71.000	¥71.000	1	1
越智	¥41.000	¥82.000	1	2
浅野	¥51.100	¥102.200	1	2
柴田	¥97.000	¥194.000	1	2
都築	¥28.400	¥85.200	2	3
山口	¥36.200	¥108.600	1	3
伊藤	¥36.100	¥108.300	3	3
戸瀬	¥98.000	¥294.000	1	3
中谷	¥24.100	¥72.300	2	3
岡本	¥3.600	¥21.600	4	6
垂水	¥8.900	¥44.500	3	5
森	¥8.200	¥41.000	2	5
山本	¥6.700	¥13.400	2	2
大津	¥7.100	¥14.200	2	2



* 参考

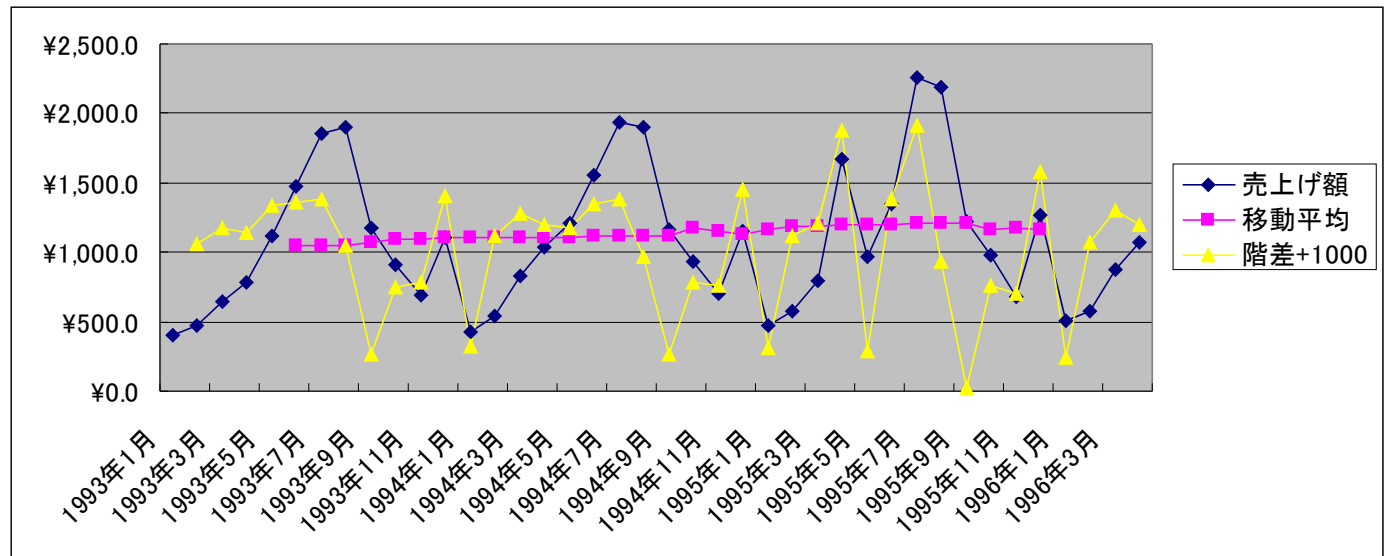
散布図
(双対尺度法での分析結果)

氏名	平均金額	合計金額	来店数	購入数
上田	¥127.000	¥127.000	1	1
水田	¥71.000	¥71.000	1	1
越智	¥41.000	¥82.000	1	2
浅野	¥51.100	¥102.200	1	2
柴田	¥97.000	¥194.000	1	2
都築	¥28.400	¥85.200	2	3
山口	¥36.200	¥108.600	1	3
伊藤	¥36.100	¥108.300	3	3
戸瀬	¥98.000	¥294.000	1	3
中谷	¥24.100	¥72.300	2	3
岡本	¥3.600	¥21.600	4	6
垂水	¥8.900	¥44.500	3	5
森	¥8.200	¥41.000	2	5
山本	¥6.700	¥13.400	2	2
大津	¥7.100	¥14.200	2	2



トレンドの分析に用いたデータ(左)と時系列グラフ(下)

	売上げ額
1993年1月	¥408.0
1993年2月	¥471.8
1993年3月	¥650.8
1993年4月	¥786.9
1993年5月	¥1,118.9
1993年6月	¥1,477.5
1993年7月	¥1,854.9
1993年8月	¥1,903.3
1993年9月	¥1,169.9
1993年10月	¥913.9
1993年11月	¥695.6
1993年12月	¥1,101.1
1994年1月	¥428.4
1994年2月	¥547.2
1994年3月	¥831.3
1994年4月	¥1,033.8
1994年5月	¥1,205.6
1994年6月	¥1,554.5
1994年7月	¥1,932.4
1994年8月	¥1,896.7
1994年9月	¥1,160.3
1994年10月	¥938.4
1994年11月	¥697.2
1994年12月	¥1,153.6
1995年1月	¥468.2
1995年2月	¥580.4
1995年3月	¥793.6
1995年4月	¥1,673.3
1995年5月	¥964.9
1995年6月	¥1,345.4
1995年7月	¥2,258.3
1995年8月	¥2,193.5
1995年9月	¥1,217.4
1995年10月	¥979.7
1995年11月	¥683.4
1995年12月	¥1,265.3
1996年1月	¥512.6
1996年2月	¥578.9
1996年3月	¥876.2
1996年4月	¥1,074.6



* 移動平均はトレンドを表しています。
 * 階差は、トレンドの成分を取り除いたデータです。

全体としてのトレンドは、5%程度の上向きの勾配。

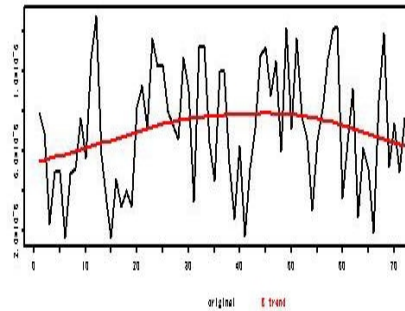
単位: 百万円

*参考

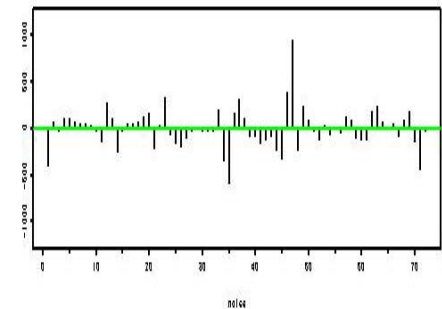
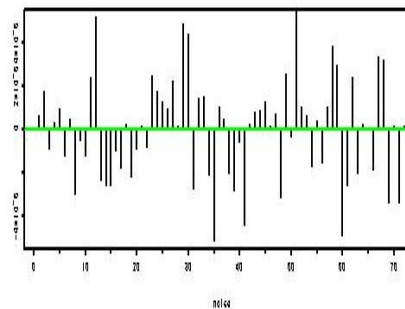
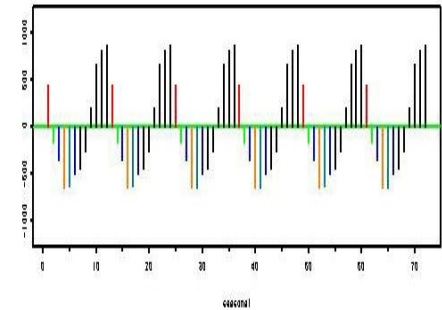
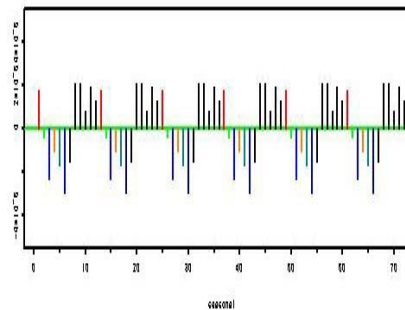
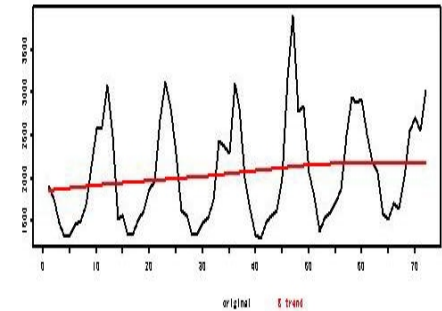
市場とのトレンド比較

右図は、製品Aの出荷数を対比したものです。上からトレンド、季節性、ノイズ成分を示しています。トレンドを見ると、市場全体ではピークが下がってきていますが自社の販売数は、まだ伸びています。市場全体よりトレンドの上昇が強ければ、占有率が上がっていることを示します。自社のノイズ成分が少ないことから、安定した要因で出荷できていると推移できます。在庫リスクは、トレンドの差の要因がポイントになると考えられます。

市場全体の出荷数(推移)



自社の販売数(推移)



株式会社 XXX様

データ・マイニングのご参考 「Webサイト分析の例」

サンプルデータを用いて、次の例を分析しています。

- | | | |
|----------|-----|-----------|
| ・現状の把握 | --- | 現状分析 |
| ・発注・在庫管理 | --- | 売上げ予測 |
| ・マーケティング | --- | 顧客管理、トレンド |

2004年5月
Sunlight co.,Ltd.

- ・分析対象のWebサイトは、商用のWebサイトをイメージしています。
- ・ビジネスとしてのWebサイトは、ユーザビリティやSEO(Search Engine Optimize)などが研究されています。これらによる「ユーザへの訴求力」と計画的(戦略的)運営の為のデータ分析「データマイニング」が必要とされています。
- ・ここでは、データ分析の例を簡単に紹介させて頂いております。

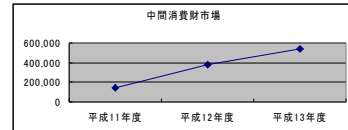
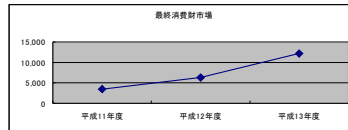
－Webサイト分析(参考例)－

「インターネットコマース」の市場規模

(平成14年度 通信白書より)

	平成11年度	平成12年度	平成13年度
最終消費財市場	3,500	6,233	12,218
中間財市場	144,000	381,000	539,000

単位:億円



対前年度成長率

	平成12年度	平成13年度
最終消費財市場	78.09	96.02
中間財市場	164.58	41.47

- 2 -

- UCHIDA -

・商用Webサイトの市場への理解として、インターネットコマースの市場規模を確認してみます。

・「Webサイト」が最終消費財市場と中間財市場の両方を担えるものと考えています。

・中間財市場では、全銀手順やJCA、SNAやHULFTなど既存のプロトコルが中心的役割を担っているかと思います。将来的にEDI(Electronic Data Interchange)は、

XML(SOAP)などが利用され、Webサイトが担う割合が大きくなるものと考えられます。

・将来予測について

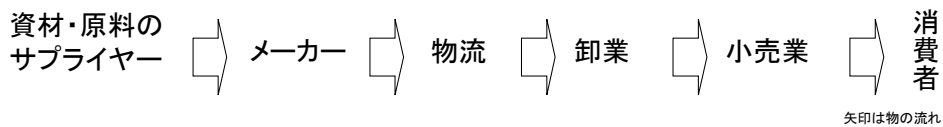
郵政事業庁は、平成12年度の通信白書で市場規模の予測をしています。

それによれば、最終消費財市場は平成17年度で7兆1,289億円、中間財で103兆4億円としています。

優良WebサイトとSCMの共通性

- ・リードタイムの短縮
- ・在庫の適正化
- ・物流コストの削減

SCM(SupplyChainManagement)とプレイヤー



- ・よいWebサイトとSCMには、多くの共通点があります。
- ・高度にシステム化されたWebサイトは、在庫確認や配達の状態なども確認できます。
これらは、SCMの要素を内在していると言えます。
- ・SCMへの期待には、生産や物流のリードタイムを短縮し、白い矢印部分の流れを向上させる事が挙げられます。
生産ロットを確保し、無駄なく早く製品を届けられるプロセスの構築が、SCMの構築と言い換えることの可能です。
- ・物は矢印の方向に流れますが、情報は逆向きに流れます。
これは、物を流す動機は消費者側にあると考える事が可能になります。

—Webサイト分析(参考例)—

CRM(Customer Relationship Management)と情報の流れ



CRMのポイント

- ・顧客満足度(不満足度)
- ・ビジネスコスト

- 4 -

- UCHIDA -

- ・SCMに例を見ても、ビジネスの動機は消費者側にあると考える事ができます。

・顧客不満足度

ご存知のように、顧客は「何か感じがよくない」と黙って去って行ってしまいます。

理由を説明してはもらえませんので、チェックを怠っていると何時までも知らずに顧客を失う事になります。

Webの上では、ユーザビリティがこれへの解答の一つになっています。

・顧客満足度

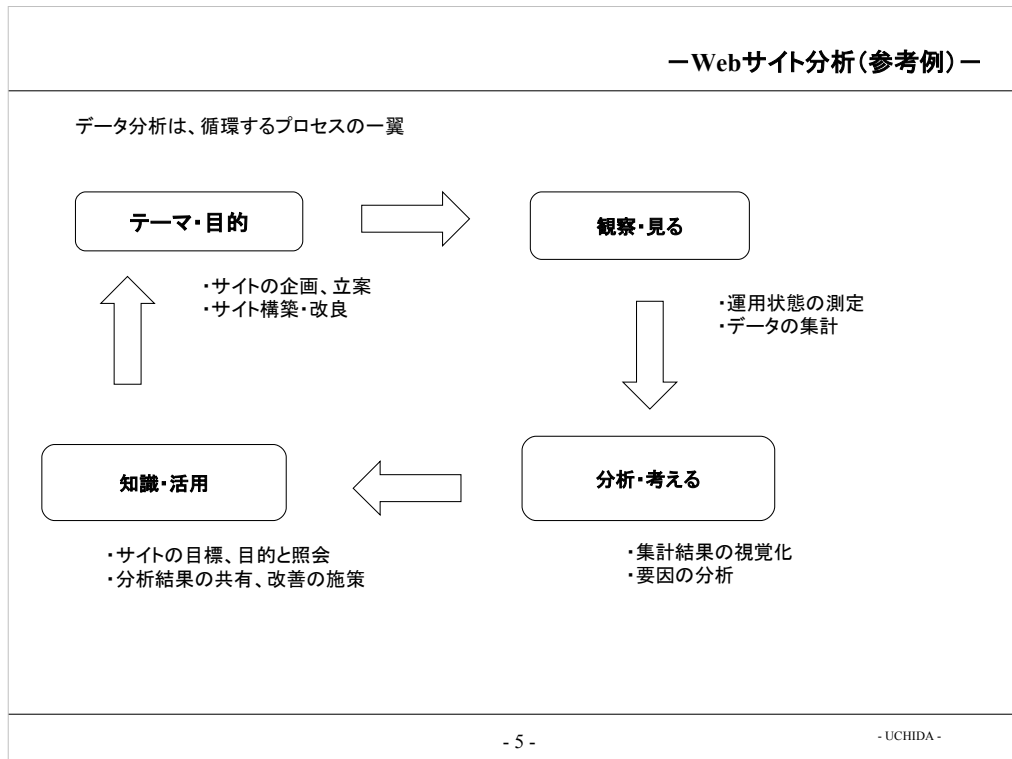
「売上の大半は、ごく少数の得意客で占められている」とする考え方があります。

獲得した顧客をセグメンテーション化して、お得意様にはコストかかけても、より良いサービスを提供しようという考え方です。

セグメンテーション化(クラスタリング)は、顧客の属性や好みでお勧め商品や休眠からの掘り起こしなどにも、利用されています。

・ビジネスコスト

費用だけでなく活動に対する効果を考える上でも、ビジネスの源流とも言え



• データ分析は、客観的に物事を捉えたり、経営的判断を支援するものと言えます。

テーマを持ち、現状を観察し分析することで知識を得て、行動指針(目的)とするプロセスが、経営や運営を支援致します。

• データ分析には、統計手法とデータマイニング手法を用います。
 厳密な境界線は無いように思いますが、個人的な理解を示します•

• 統計は、記述統計と推測統計に大別されます。
 記述統計は、平均値や標準偏差など主にデータ集計を指します。
 推測統計は、仮説検証など主に確立的な推測に利用されています。

• データマイニングは、大量データから何かを発掘する意味ですが、決定木やニューラルネットワークなど、
 統計のより進化した手法の実践(システム化)を指しているようです。

現状の把握 [現状分析]

-Webサイト分析(参考例)-

分析に用いた月間売上データ

日付	売上げ額	サイトアクセス数	販売個数	商品A売上個数	商品A価格	その他商品売上個数	その他商品平均額	キャンペーン(割引率)
11月1日	50400	1710	18	6	3000	12	2700	1
11月2日	68800	2090	22	8	3000	14	3200	1
11月3日	69400	1710	18	11	3000	7	5200	1
11月4日	65800	1995	21	7	3000	14	3200	1
11月5日	71400	2185	23	11	3000	12	3200	1
11月6日	110200	4275	45	14	3000	31	2200	1
11月7日	46100	1425	15	4	3000	11	3100	1
11月8日	65100	2185	23	10	3000	13	2700	1
11月9日	69000	1995	21	9	3000	12	3500	1
11月10日	49800	1425	15	3	3000	12	3400	1
11月11日	96000	3040	32	11	3000	21	3000	1
11月12日	127000	3705	39	14	3000	25	3400	1
11月13日	134700	4560	48	17	3000	31	2700	1
11月14日	59500	2090	22	9	3000	13	2500	1
11月15日	44000	1140	12	4	3000	8	4000	1
11月16日	44200	1235	13	0	0	13	3400	1
11月17日	58800	1710	18	6	3000	12	3400	1
11月18日	91200	2755	29	8	3000	21	3200	1
11月19日	77000	2185	23	3	3000	20	3400	1
11月20日	57800	1520	16	2	3000	14	3700	1
11月21日	75600	1995	21	0	0	21	3600	1
11月22日	60100	1615	17	4	3000	13	3700	1
11月23日	53200	1805	19	0	3000	19	2800	1
11月24日	34500	760	8	1	3000	7	4500	1
11月25日	206850	7220	76	31	3000	45	4500	0.7
11月26日	156240	6090	64	25	3000	39	3800	0.7
11月27日	197330	7220	76	27	3000	49	4100	0.7
11月28日	64500	2375	25	4	3000	21	2500	1
11月29日	54400	1805	19	6	3000	13	2800	1
11月30日	22800	570	6	2	3000	4	4200	1

- 6 -

- UCHIDA -

- ・売上げが、月間で280万円程度、年商で3000万円程度の個人サイトをイメージした11月度の月間データです。
- ・主力商品Aが、販売個数・額とも全体の30%程度
- ・販売だけでなく、情報の発信や交換の場として人気を持っているので、サイトアクセス数は若干多めになっています。
- ・商品を絞り込んでいるサイトでは、ある程度ユーザ層や人気商品を見込んでいると思われます。
ここでは、見込み(戦略)がどの程度、有効であったか、想定外の問題の有無などの分析が必要になります。

サイト運営の戦略と分析目的

- ・) 商品Aが主力であり、Aに強いことが自サイトの特徴でもある。
- ・) 将来を見据えるとAだけに頼れないが、自サイトの特徴を失いたくない
→ 商品Aで、50~30%程度の売上比率をキープしたい
- ・) 早急に手を打たなくてはならない問題はあるか

現状の把握 [現状分析]		-Webサイト分析(参考例)-	
月間サマリー		月間サイトアクセス数	76,380
		月間売上額数	¥2,381,720
		平均単価	¥2,962
		月間売上個数	804
		商品A売上個数	257
		その他商品売上個数	547
		商品A比率	31.97%
		顧客数	572
		コンバージョン率	0.75%
売上げ額に影響を与えた主要因		サイトアクセス数	
		商品A価格	
		その他の商品の売上げ個数	
売上げ額の説明式 [信頼度97%]			
		$\text{売上げ額} = 19,994 + 27.9 * \text{サイトアクセス数} - 1.7 * \text{商品A価格} - 0.7 * \text{その他商品売上げ個数}$	
		- 7 -	- UCHIDA -

・月間集計と要因の分析

ここでは、月間の集計と、要因分析を重回帰で行っています。

・ご存知の通り、重回帰分析は次のような説明式になります。

$$\text{結果} = \text{固定数} + \text{要因A} * \text{影響度A} + \text{要因B} * \text{影響度B} + \dots + \text{要因X} * \text{影響度X}$$

・この式のバランスに注目する事により、市場の変化や問題の発見につながります。

現状の把握 [現状分析]

—Webサイト分析(参考例)—

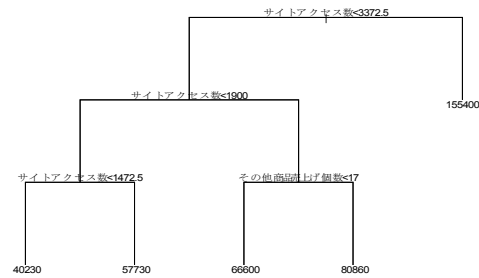
月間の売上げ額の異常性なし

(今月の売上げ額:238万円)

年間平均	271万円
標準偏差	32万円
信頼率80%の 正常範囲	230~312万円

***参考**

月間売上げ要因分析(決定木)



- 8 -

- UCHIDA -

- 全体像として売上げデータを評価する場合、ある閾値を設けることにより現状が正常なのか、アクションが必要かが分かり易くなります。
- 参考は、売上げを決定木で分析した例です。
売上げを説明する場合、サイトアクセス数でほとんどを分類しています。
中間的な売上額 66600と80860を分ける場合にその他売上個数が17以上か否かで判断しています。

決定木は要因を変数として扱い、値が幾つの時、結果に影響を及ぼすか否かを把握するのに優れています。

(決定木は、質的変数の方が向いているかも知れません)

現状の把握 [現状分析]

-Webサイト分析(参考例)-

*参考

月間売上げ要因分析
(全ての変数を用いた重回帰分析)

回帰統計				
重相関 R	0.995122159			
重決定 R2	0.990288112			
補正 R2	0.987171602			
標準誤差	5033.167642			
観測数	30			
	係数	標準誤差	t	P-値
切片	-112348.8887	25854.50694	-4.345427624	0.000259484
サイトアクセス数	-221.7504079	9400558.13	-2.35891E-06	0.999998139
販売個数	-8777.142857	3773547932	-2.32597E-06	0.999998165
商品A売上個数	32709.15914	11284109148	2.89969E-06	0.999997713
商品A価格	-1.372308499	1.438291163	-0.954118702	0.350390419
その他商品売上げ個数	32848.10688	11284109148	2.91101E-06	0.999997704
その他商品平均額	10.6610706	1.863504086	5.720980533	9.3793E-06
キャンペーン(割引額)	83855.66538	21715.15511	3.861619452	0.000844722

月間売上げ要因分析
(必要な変数に絞った重回帰分析)

概要				
回帰統計				
重相関 R	0.986544			
重決定 R2	0.973269			
補正 R2	0.970185			
標準誤差	7673.134			
観測数	30			
	係数	標準誤差	t	P-値
切片	19994.82	7029.39	2.84446	0.008556
サイトアクセス数	27.8694	3.628957	7.679728	3.77E-08
商品A価格	-1.67262	2.022657	-0.82694	0.415795
その他商品売上げ個数	-385.385	577.7216	-0.66708	0.5106

- 9 -

- UCHIDA -

- 要因分析の計算過程です。

Excelで行っていますので、CSVなどを生成する仕組みを作ることにより、自動化が可能になってきます。

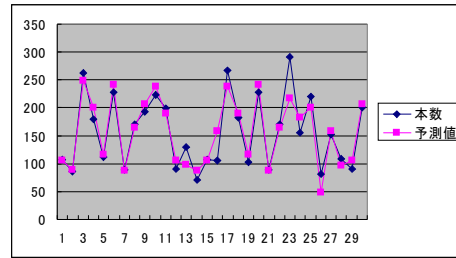
但し、多重共線性や式や変数の当てはまりの良さなど、Excelでは出来ない事も多いので、ここでは手動で調整しています。

発注・在庫管理 [売上げ予測]

-Webサイト分析(参考例)-

ある清涼飲料水の月間売上げ本数と実績と予測

日付	本数	単価	天気	曜日	予測本数
1	107	318	晴れ	木	106
2	86	318	晴れ	金	90
3	263	295	曇り	土	249
4	179	295	曇り	日	200
5	111	318	曇り	月	116
6	228	295	曇り	火	242
7	89	318	晴れ	水	87
8	171	308	晴れ	木	164
9	193	298	晴れ	金	206
10	224	295	晴れ	土	239
11	199	295	晴れ	日	190
12	91	318	晴れ	月	105
13	130	318	晴れ	火	97
14	71	318	晴れ	水	87
15	107	318	晴れ	木	106
16	106	308	曇り	金	158
17	267	295	晴れ	土	239
18	183	295	晴れ	日	190
19	102	318	曇り	月	116
20	228	295	曇り	火	242
21	89	318	晴れ	水	87
22	171	308	晴れ	木	164
23	291	298	曇り	金	217
24	155	295	雨	土	182
25	220	295	曇り	日	200
26	81	318	雨	月	49
27	153	298	雨	火	158
28	109	318	曇り	水	97
29	90	318	晴れ	木	106
30	201	298	晴れ	金	206



予測に用いた重回帰式 [信頼度87%]

	単価	天気	曜日
売上げ本数 =	1852.6	- 5.8*単価 +	0 (雨) 66.7(曇) 56.3(晴)
			0(日) 49.1(月) 41.1(火) 30.1(水) 49.8(木) 33.6(金) 48.7(土)
影響度指数	5.8	66.7	49.8

- 10 -

- UCHIDA -

- ・時系列データは、ARIMAモデルなど、代表的な分析手法がありますが、Excelでは少し単純な分析しかできませんので、ここでは、数量化一類を用いて回帰式を出しています。
- ・決定係数からは、87%は単価、天気、曜日で 売上げ本数が決まると言えます。言い方を変えれば 13%は別の要因があると言えます。
- ・ビールやアイスなどは、気温や価格、または曜日などで売上げが決まると言われていますが、近所で運動会などのイベントがあると様子が変わってきます。

自動発注などに応用する場合には、はずれ値の取り扱いと13%のノイズ分の調整が必要になります。

- ・予測は、表計算ではなく専門のソフトで予測することが多いようですが、時系列データには、トレンド、周期性の他に ノイズ成分の影響が大きく誤差が生じがちになります。
- ・予測がどの程度の信頼幅であるかを把握することが、運用のポイントになります。

マーケティング [顧客分析]

-Webサイト分析(参考例)-

分析に用いた顧客データ

氏名	平均金額	合計金額	来店回数	購入数
上田	¥127,000	¥127,000	1	1
岡本	¥3,600	¥21,600	4	6
都築	¥28,400	¥85,200	2	3
山口	¥36,200	¥108,600	1	3
垂水	¥8,900	¥44,500	3	5
越智	¥41,000	¥82,000	1	2
水田	¥71,000	¥71,000	1	1
森	¥8,200	¥41,000	2	5
山本	¥6,700	¥13,400	2	2
浅野	¥51,100	¥102,200	1	2
伊藤	¥36,100	¥108,300	3	3
大津	¥7,100	¥14,200	2	2
戸瀬	¥98,000	¥294,000	1	3
中谷	¥24,100	¥72,300	2	3
柴田	¥97,000	¥194,000	1	2

クラスタリングの結果

氏名	平均金額	合計金額	来店数	購入数
上田	¥127,000	¥127,000	1	1
水田	¥71,000	¥71,000	1	1
越智	¥41,000	¥82,000	1	2
浅野	¥51,100	¥102,200	1	2
柴田	¥97,000	¥194,000	1	2
都築	¥28,400	¥85,200	2	3
山口	¥36,200	¥108,600	1	3
伊藤	¥36,100	¥108,300	3	3
戸瀬	¥98,000	¥294,000	1	3
中谷	¥24,100	¥72,300	2	3
岡本	¥3,600	¥21,600	4	6
垂水	¥8,900	¥44,500	3	5
森	¥8,200	¥41,000	2	5
山本	¥6,700	¥13,400	2	2
大津	¥7,100	¥14,200	2	2

優良顧客だが、リピートになっていない		
一般的な優良顧客		
一般的な顧客		

- 2対8の法則、パレートの法則と言われるように顧客に適合したサービスやPRが必要とされています。
- 実態を把握するには、まずクラスタリングをしてみます。

マーケティング [顧客分析]

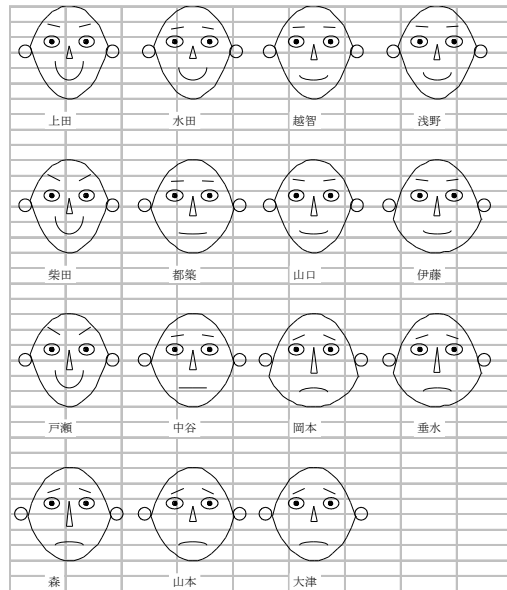
-Webサイト分析(参考例)-

* 参考

チャーターの顔グラフ

眉の角度 ---- 合計金額
 口の曲率 ---- 平均金額
 鼻の長さ ---- 購入数
 顔半分の ---- 来店数
 楕円の離芯率

氏名	平均金額	合計金額	来店数	購入数
上田	¥127,000	¥127,000	1	1
水田	¥71,000	¥71,000	1	1
越智	¥41,000	¥82,000	1	2
浅野	¥51,100	¥102,200	1	2
柴田	¥97,000	¥194,000	1	2
都築	¥28,400	¥85,200	2	3
山口	¥36,200	¥108,600	1	3
伊藤	¥36,100	¥108,300	3	3
戸瀬	¥98,000	¥294,000	1	3
中谷	¥24,100	¥72,300	2	3
岡本	¥3,600	¥21,600	4	6
垂水	¥8,900	¥44,500	3	5
森	¥8,200	¥41,000	2	5
山本	¥6,700	¥13,400	2	2
大津	¥7,100	¥14,200	2	2



- 12 -

- UCHIDA -

眉が凛々しく、口が笑っているほど、購入金額が大きかったり高額商品を購入しています。

鼻が長く、あごが立派になるほど、来店数や購入数が多く、店に密着していると言えます。

口が笑っていないくて、あごが立派だと、手が掛かる割に……と言われ兼ねません。

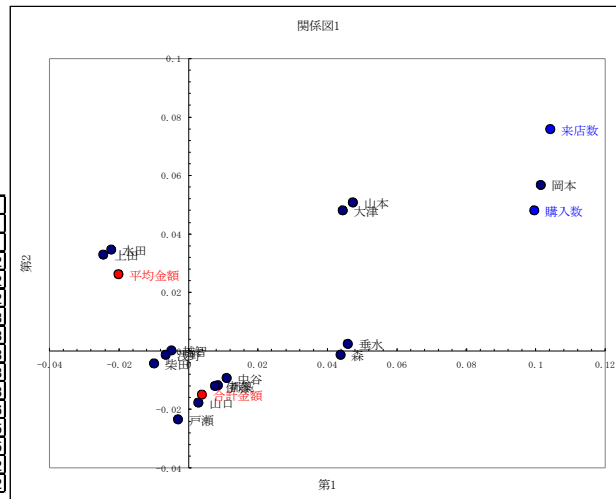
個人的には、顔グラフは遊びに近いですが、数字より人を惹きつけやすいですし、

複合的な見方にも適していると思います。

* 参考

散布図
(双対尺度法での分析結果)

氏名	平均金額	合計金額	来店数	購入数
上田	¥127,000	¥127,000	1	1
水田	¥71,000	¥71,000	1	1
梶智	¥41,000	¥82,000	1	2
法野	¥51,100	¥102,200	1	2
柴田	¥97,000	¥194,000	1	2
都築	¥28,400	¥85,200	2	3
山口	¥36,200	¥108,600	1	3
伊藤	¥36,100	¥108,300	3	3
戸瀬	¥98,000	¥294,000	1	3
中谷	¥24,100	¥72,300	2	3
岡本	¥3,600	¥21,600	4	6
垂水	¥8,900	¥44,500	3	5
藤	¥8,200	¥41,000	2	5
山本	¥6,700	¥13,400	2	2
大津	¥7,100	¥14,200	2	2



散布図上で距離に近いほど、同じ特性を持った仲間と言えます。

5つにも分けられますし、横方向で3つ、縦方向で2つなど、いろいろな分類が可能です。

直感的に分類する場合に優れています。

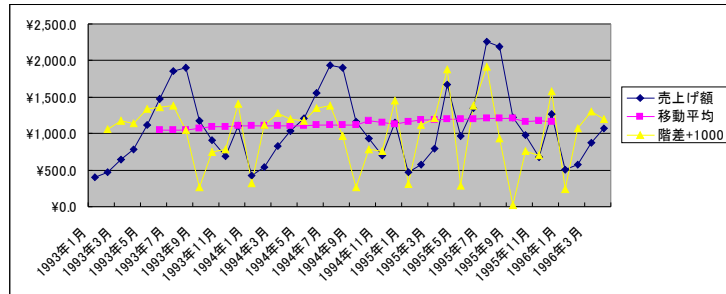
マーケティング [トレンド]

-Webサイト分析(参考例)-

	売上げ額
1993年1月	¥408.0
1993年2月	¥471.8
1993年3月	¥650.8
1993年4月	¥786.9
1993年5月	¥1,118.9
1993年6月	¥1,477.5
1993年7月	¥1,854.9
1993年8月	¥1,903.3
1993年9月	¥1,169.9
1993年10月	¥913.9
1993年11月	¥695.6
1993年12月	¥1,101.1
1994年1月	¥428.4
1994年2月	¥547.2
1994年3月	¥831.3
1994年4月	¥1,033.8
1994年5月	¥1,205.6
1994年6月	¥1,554.5
1994年7月	¥1,932.4
1994年8月	¥1,896.7
1994年9月	¥1,160.3
1994年10月	¥938.4
1994年11月	¥697.2
1994年12月	¥1,153.6
1995年1月	¥468.2
1995年2月	¥580.4
1995年3月	¥793.6
1995年4月	¥1,673.3
1995年5月	¥964.9
1995年6月	¥1,345.4
1995年7月	¥2,258.3
1995年8月	¥2,193.5
1995年9月	¥1,217.4
1995年10月	¥979.7
1995年11月	¥683.4
1995年12月	¥1,265.3
1996年1月	¥512.6
1996年2月	¥578.9
1996年3月	¥876.2
1996年4月	¥1,074.6

単位:百万円

トレンドの分析に用いたデータ(左)と時系列グラフ(下)



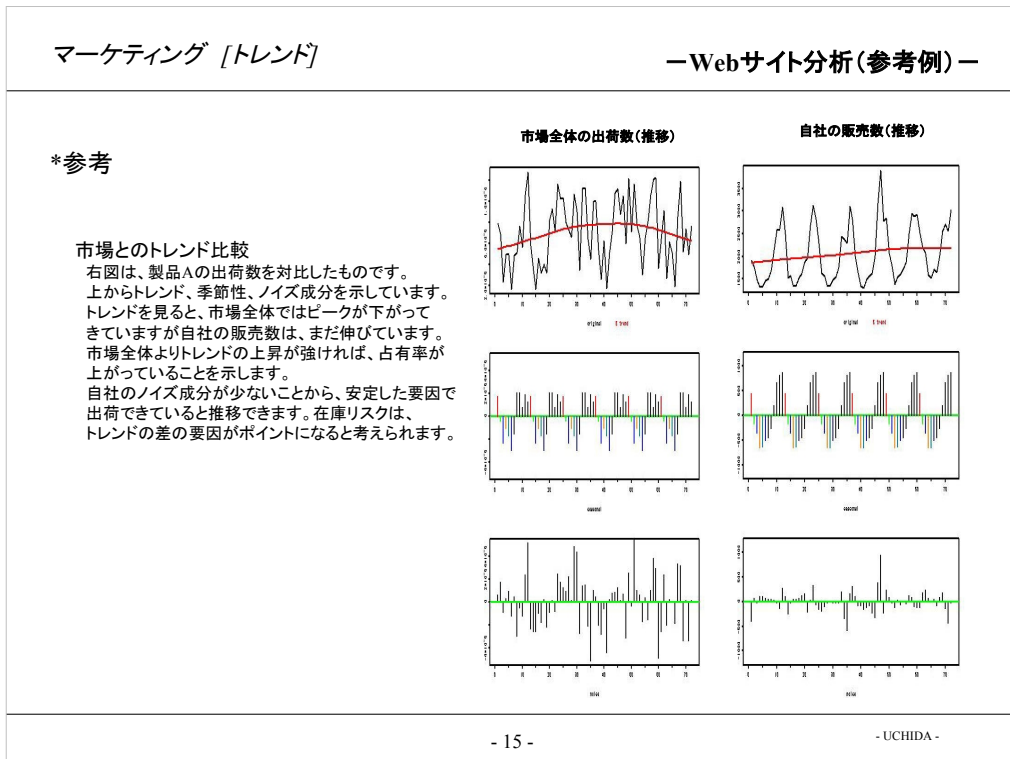
- * 移動平均はトレンドを表しています。
- * 階差は、トレンドの成分を取り除いたデータです。

全体としてのトレンドは、5%程度の上向きの勾配。

発注・在庫の分析例で少し触れましたが予測を行う場合、時系列分析で行う事が多いようです。

季節性などがある売上げデータは、生データをそのまま見ても事業として上向いているのか否か見てきません。

そこでトレンドだけを取り出したりすることが必要になります。



時系列データ分析には、ARIMAモデルなど代表的な手法がありますが、経済のデータの例のように、トレンドや周期性で説明できないノイズ成分が多くあり、精度を高めることが難しくなっています。

取り扱っているデータがどの程度ノイズ成分を持っているかを把握することが、
、
予測値を評価する上でも重要になってきます。

また、業界全体や同業他社との比較を行う場合には、トレンドや周期性が重要になります。