

ビジネス顕微鏡を用いた業績向上に貢献する 指標探索アプリケーションの開発

Development of Index Search Application for Improving Business Performance Using Business Microscope

佐藤 信夫†
Nobuo Sato

福間 晋一†
Shinichi Fukuma

矢野 和男†
Kazuo Yano

1. まえがき

本研究は、ヒト・モノ・カネに関連するビッグデータから業績向上施策に結びつく指標を統計的に整理して提示する技術の実現を目的とする。ビジネス顕微鏡[1][2]から収集される行動データと、売り上げ等の業績データを用いて、売り上げに寄与する従業員の振る舞いを探索することである。

その第1段階として、目的指標を軸とし膨大な説明指標を分類したネットワークダイアグラムを表示することを特徴とした指標探索アプリケーションを検討した。その特徴は単相関と直線回帰によって有益指標とその期待効果の関係を表示することである。有益指標を探索する際に必要な要件として、「膨大な数の説明指標を分類・整理して見せること」「各説明指標が目的指標にもたらす影響を日本語かつ具体的な単位で表示すること」「類似する説明指標による効果の見込みを並べて表示すること」と定義し、実験の結果、これらを満たしていることを確認した。さらに、本アプリケーションを用いて店舗客単価向上施策を立案・実施したところ、客単価が15%向上したことを確認した。

本報告の構成として、第2章では我々のコンセプトの紹介、第3章では初期検討に必要な要件検討について記載し、第4章ではそのアプリケーションの実装方法について記載する。第5章の実験では本アプリケーションを用いて店舗に関するデータから業績向上に有効な指標を探索できるかを確認する。

2. コンセプト

我々の目指すところは、行動データを含んだ社会産業に関わるデータを入力とし、業績向上モデルを推定するシステムを実現することである。このシステムを活用するための原則を以下の3つとした。

- 原則1 目的： 向上すべきアウトカム(業績)を明確にする
 原則2 データ： アウトカム(業績)に関連するデータをヒト・モノ・カネに渡り広く取得する
 原則3 発見： 仮説に頼らず、コンピュータに業績向上策をデータからモデル化させる

原則1の目的について、アウトカムが明確になっていないと、解析した結果から改善施策を提案したとしても、それが実行されないことが多い。明確でないということは、

クライアント内でコンセンサスが取れていないといえる。よって、そこから導き出された改善施策は must(絶対に実施)ではなく want(できれば実施)になってしまい、多忙な業務の中では改善施策は後回しにされてしまうからである。原則2のデータについて、ヒト・モノ・カネに渡り広く取得すると示したが、具体的には、顧客情報、業務情報、設備情報、財務情報である。1つでも欠けていると、解析した結果に曖昧さが残ってしまい、クライアントに明確な説明ができなくなるからである。原則3の発見について、クライアントが知らなかったことを、ビッグデータから導き出す方が価値は高いと考えている。

3. プロトタイプの要件検討

3.1 目標設定

2章では、顧客の課題から改善施策を提示することを最終ゴールとしているが、本報告では、初期検討として、顧客の課題を軸として膨大なデータを体系的に整理して示すことまでを目標とした。まず、今回開発するプロトタイプの要件と期待する効果について表1にまとめた。

表1 機能要件と期待する効果

No	機能要件	期待する効果
機能要件 I	膨大な数の説明指標を分類・整理して見せること	目的変数に影響する要因の大局を理解し、説明指標の候補を絞り込める
機能要件 II	各説明指標が目的指標にもたらす影響を日本語かつ具体的な単位で表示すること	ピンポイントで幾つかの説明指標の候補を絞り込める
機能要件 III	類似する説明指標による効果の見込みを並べて表示すること	選んだ指標が他より有益だと、ストーリーを持って顧客に説明できる

3.2 ビジネス顕微鏡による活動センシング

ビジネス顕微鏡は、センサ技術を用いて企業内のコミュニケーションや活動状況を測定・解析するための計測システムである。ワーカーはHi バッジ(名札型センサ端末)(図1)を装着した状態で通常の業務(図2)を行なう。



図1 ビジネス顕微鏡 Hi バッジ

† (株)日立製作所 中央研究所, Central Research Laboratory, Hitachi, Ltd.

Hi バッジにて取得されるセンサデータは加速度、赤外線、温度であり、データ取得周期は 10 秒間である。加速度センサではワーカーの体の動きを 3 軸加速度で計測し、また、赤外線センサではワーカー同士が一定の距離内に近づくと、お互いの Hi バッジが通信して対面したことを検知する。さらに、これらのセンサは 100msec 以内で時刻同期しているため、複数センサデータの組み合わせによる分析が可能であり、対面中の体の動き等も計測することができる。



図2 業務風景

3.3 開発内容

3.1 章のコンセプトに基づきユーザが指定する業績向上などの目的指標と改善の仮説となる説明指標との関係を、図3に示す様な階層表示にて表現するアプリケーションを開発する。そして、それぞれの階層で表現したい事柄については表2にまとめた。

まず、ユーザに業績向上に関する特徴量を目的指標として選択させる。その目的指標と高い相関を持つ特徴量をビックデータの特徴量の中から説明指標として抽出する。さらに、相関をもつ説明指標を羅列しただけでは、指標数が膨大になった場合にユーザが理解しづらいため、図3の様に成分毎に説明指標をグループ化して表示する。

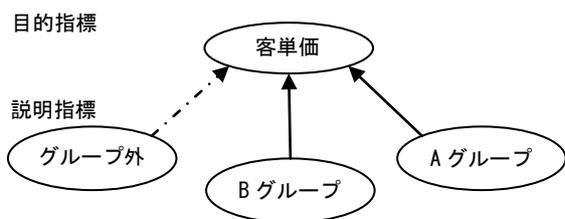


図3 出力画面イメージ

表2 表現内容

名称	表現内容
目的指標	<ul style="list-style-type: none"> 向上させたいものを目的指標とする ユーザが目的指標を任意に選択できるようにする 目標指標が複数の成分(グループ)によって、構成されていることが分かるようにする
説明指標	<ul style="list-style-type: none"> 同じ成分がある説明指標をグループとしてまとめる。現場によって、実行できる施策が異なるため、複数から選択できるようにする 目的指標を複数のグループで構成させる。細分化されているように表示する グループに属さない説明指標は、グループ外として録しておく

データからモデルを作成する既存の手法として、グラフィカルモデル[3]がある。今回は一番シンプルな単相関と直線回帰を用いることで、これを実現することを試みる。目的指標と相関値が高いものを説明指標として選択し、その説明指標を相関値が高いもので、グループを構成する。さらに、目的指標と説明指標との直線回帰の残差を用いることで、複数のグループに分割させる方法を用いる。

特徴量を3つ以上扱う場合の相関分析では、単相関ではなく偏相関を用いることが一般的であるが、単相関を選択した。その理由として、偏相関を用いた場合には、ユーザが不要な特徴量を削除することが、全体に影響を与えてしまうことがある。従って、グラフ構造の概観が削除するたびに大きく変化し、ユーザインタフェースとしてユーザを混乱させる恐れが高いからである。

4. プロトタイプの実装

4.1 概要

利用者が指定した目的指標をクエリとして、特徴量の中から有効な説明指標を抽出する Web アプリケーションを実装する。システム構成を図4に示す。

4階層から構成されており、1. アプリケーション部、2. 演算部、3. 特徴量部、4. 前処理部となっている。ユーザがクライアントで目的指標を選択すると、3にて登録されている特徴量データを用いて、2で分析を行ない、その結果をユーザに出力する流れとなる。図4での矢印であるが、実線の矢印はイベント処理、破線の矢印はバッチ処理を示している。なお、本報告では3. 特徴量部まで作成を事前に作成しており、1. アプリケーション部と2. 演算部についてのみ説明する。

図5はアプリケーション表示画面である。表3は図5の各機能の説明である。次節から本アプリケーションの主要処理であるモデル生成処理について述べる。

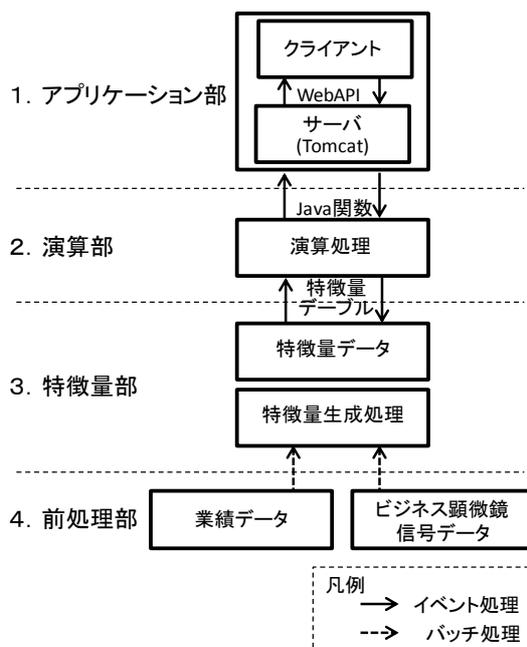


図4 システム構成

表3 アプリケーション機能概要説明

機能名	概要	処理技術	機能要件
A. 条件選択	解析対象, 期間や種類等を選択し, 解析に用いる条件を選択する	-	-
B. 目的指標選択	条件選択された数多くの特徴量の中から, 目的指標を選択する. 選択方法は2種類(カテゴリー分類におけるリストボックスとキーワード検索)がある	-	-
C. 説明指標選択	出力される数多くの説明指標を絞りこみための機能である. 選択方法は目的指標選択と同じである	-	機能要件Ⅱ
D. ネットワークダイアグラム表示	目的指標とそれに相関する説明指標との関係をモデル化した結果をネットワークダイアグラムで表示する	モデル生成	機能要件Ⅰ
E. リスト表示	モデル化された各説明指標をリスト表示したものである		
F. 実態と効果説明表示	目的変数にどのような効果が見込めるかの理解するために, 実態と効果説明を表示したものである	-	機能要件Ⅲ

4.2 モデル生成処理

図6で示すように目的指標 Y が複数の説明指標のグループから構成されているモデルを生成する. 各グループは説明指標の集合となっている. それぞれの要素の説明は表4に記載する.

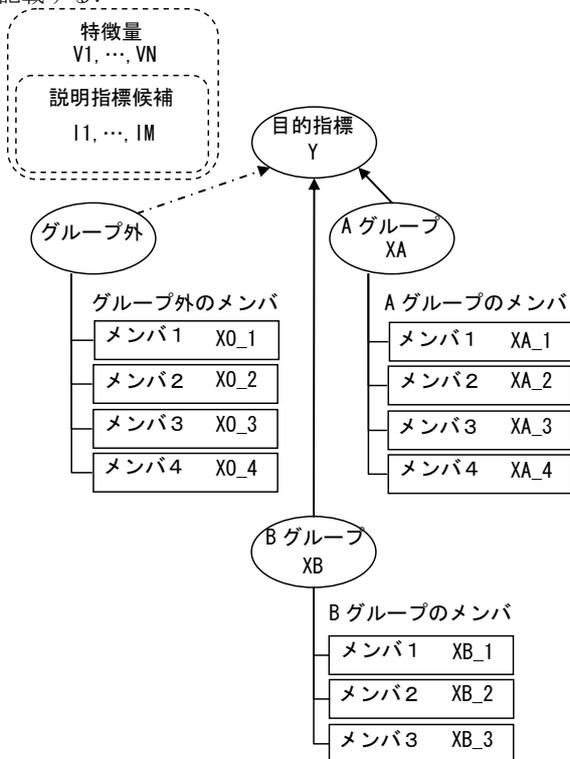


図6 モデル概要図

次に, 実際の処理の流れについて説明する. ①から⑥の手順によって処理していく. 相関係数のしきい値を r_{th} , 目的指標 Y と説明指標候補 I1 の相関係数を $r(Y, I1)$ と記載する. 処理を行なっている際に, 指標が何処に属するかを

管理する必要があり, 表5に示す属性フラグ対応表を用いて管理していく. これには1, 0の値を格納する.

表4 モデル概要説明

名称	意味
特徴量	条件選択によって選ばれたデータである
説明指標候補	特徴量の中から説明指標として用いることができる特徴量である
目的指標	ユーザによって指定された特徴量である
Aグループ	目的指標を軸とした時に, ある成分の要素である. Aグループの長となる説明変数が存在する
Aグループのメンバ	Aグループ長の説明変数と相関のある説明変数である
Bグループ	目的指標を軸とした時に, Aグループとは異なる成分の要素である. Bグループの長となる説明変数が存在する
Bグループのメンバ	Bグループ長の説明変数と相関のある説明変数である
グループ外	グループに割り当てられることが出来なかった説明変数を1つにまとめたものである. 他のグループと違い, メンバ内で相関があるものではないため, メンバ長となる説明指標は存在しない
グループ外のメンバ	グループに割り当てられることが出来なかった説明変数である

表5 属性フラグ対応表の例

特徴量 ID	1	2	..	N-1	N
特徴量フラグ	1	1	..	1	1
説明指標候補フラグ	1	0	..	0	0
グループ外のメンバフラグ	0	0	..	1	0
Aグループフラグ	1	0	..	0	0
Aグループのメンバフラグ	0	1	..	0	1
Bグループフラグ	0	0	..	0	0
Bグループのメンバフラグ	0	0	..	0	0

①特徴量の選択

ユーザが指定した条件選択によって今回の解析に用いる特徴量を選択し, 特徴量毎に番号(V1, ..., VN)を割り振る. 表5に示すように, 解析の対象となった特徴量 ID に対して, 説明指標候補とグループ外のメンバに1を代入する.

②説明指標対象候補の選択

ユーザが指定した目的指標選択の目的指標を Y とし, ①で選択した特徴量(V1, ..., VN)が説明変数の条件を満たしていれば, 説明指標候補とし, 指標毎に番号(I1, ..., IM)を割り振る. 条件を満たしていない場合には, 表5の該当の特徴量 ID に対して, 説明指標候補とグループ外のメンバに0を代入する.

説明変数の条件とは, 目的指標 Y と特徴量(I1, ..., IM)の相関値 r を求め, 条件選択で設定されたしきい値 r_{th} 以上となる場合とした. その際, 前処理として相関値を絶対値化しておく. もし, 相関値を求めることができなかった場合には, 説明指標候補に0を代入する.

③ Aグループ選択

目的指標 Y との相関値が絶対値が最も高い説明指標を Aグループの長となる説明指標 XA とする。表 5 の該当の特徴量 ID に対して、Aグループに 1 を代入し、かつ、説明指標候補とグループ外のメンバに 0 を代入する。

④ Aグループのメンバ選択

Aグループの説明指標 XA との相関値がしきい値 r_{th} 以上の説明指標対象候補 (I_1, \dots, I_M) を選択し、Aグループのメンバ (XA_1, \dots, XA_4) とする。その際に、Aグループのメンバ (XA_1, \dots, XA_4) は目的変数 Y との相関値を求め、ある一定以上の相関値が得られない場合には Aグループのメンバ (XA_1, \dots, XA_4) に含めない。今回はその基準を 0.02 とした。表 5 における該当の特徴量 ID に対して、Aグループのメンバに 1 を代入し、かつ、グループ外のメンバに 0 を代入する。

⑤ 残差計算

目的指標 Y と Aグループの説明指標 XA との直線回帰を求める。直線回帰式は $Y=XA*a+b$ となる。 a は回帰係数、 b は切片である。この式を変形させ、残差は $\varepsilon=Y-(a+b*XA)$ となる。目的指標 Y と Aグループの説明指標 XA に含まれているサンプル点毎にこの計算を行なう。

⑥ 繰り返し(終了条件)

⑤で求めた残差を新しい目的指標とし、②～⑤を繰り返す。繰り返していく間に、②で求めている相関値でしきい値 r_{th} 以上の説明指標対象候補 (I_1, \dots, I_M) がなくなった場合に処理が終了となる。

その他の処理

- ・グループが 1 つだけの時の処理

グループが 1 つしか出力されなかった場合には、その説明指標を取り除いて処理をもう一度行なう。その理由は、目的指標と説明指標が 1 対 1 対応の関係になってしまい、我々が重視している目的指標が複数の説明指標で構成されているという表示にはならないためである。

- ・全グループのメンバに共通の説明指標が含まれている

時の処理

全グループのメンバに共通の説明指標が含まれている場合には、その説明指標を全グループから削除し、グループ外のメンバとして再登録をする。その理由は、共通の説明指標では、我々が重視している目的指標が複数の説明指標で構成されているという表示にはならないためである。

5. 実験

実案件において本アプリケーションを用いることで分析者が膨大な特徴量から目的指標（顧客購入金額）に関連する有益な説明指標を絞り込み、その結果の意味を顧客に説明できるかを確認する必要がある。表 1 に示した期待する効果の 3 点が実現されているかを実験によって確認した。以下に、実験条件を示す。

被験者	2名
実験用データ	店舗における顧客・店員の行動データ、POS データ(特徴量 6000 種類)
実験方法	分析者のアプリケーション中を観察、実験後に分析者にヒアリング

なお、本案件の結果を踏まえて、店舗における客単価向上施策を立案・実施したところ、客単価が 15% 向上した。詳細は割愛する。

機能要件 I 膨大な数の説明指標を分類・整理して見せること

図 5 のアプリケーションにおいて、この案件では予め 6000 個の特徴量が格納されており、目的指標を顧客購入金額と設定し、計算実行を押した。その結果、図 5 の D. ネットワークダイアグラムでは、7 個のグループとグループ外の計 8 個に分類された。そして、それぞれのグループ内で複数の説明指標が表示されており、6000 個という数多い特徴量の中から、目的指標を軸に説明指標が整理されていることを確認した。さらに、図 5 の D. ネットワークダイアグラムをグループで表示させている説明指標に着目すると、Aグループでは日用品の購入金額の様に当たり前と感ずるものが表示されている。しかし、Bグループは店員の顧客との会話について抽出されており、顧客との会話に有益な説明指標の候補として見なすことができる。よって、表 1 の機能要件 I の期待する効果として挙げた、目的変数に影響する要因の大局を理解し、説明指標の候補を絞り込めることが実現されているため、機能要件 I を満たしていることを確認した。

機能要件 II 各説明指標が目的指標にもたらす影響を日本語かつ具体的な単位で表示すること

実際に効果がある説明指標を有益指標として導き出せるかを確認する。図 5 の C. 説明指標選択を用いて、絞り込みを行なうための入力を行なった。絞りこみを行なう際には、どのような観点で絞り込みを行なうかを考える必要がある。

- ・ 店員全体の滞在時間
- ・ 店員全体の積極性
- ・ 店員全体の会話量
- ・ 店員全体の 1 人作業時間
- ・ 副店長(店員)の滞在時間
- ・ 顧客接客時間
- ・ 顧客立ち止まり時間

等で絞りこみを行ない有益指標の候補を見つけた。その数例を紹介すると、以下の様なものがある。

- ・ 店員情報-全店員-場所(滞在)-第二マグネット①($r=0.25$) 1 秒毎 (3.19%増) に 14.5 円増加
- ・ 店員情報-全店員-会話-顧客($r=0.39$) 1 秒毎 (0.509%増) に 4.6 円増加
- ・ 店員情報-全店員-積極性・活動量-一人作業エネルギー和($r=0.32$) 1 秒毎 (4.34%増) に 38.2 円増加
- ・ 店員情報-全店員-積極性・活動量-会話リズム積極時間($r=0.27$) 1 秒毎 (0.211%増) に 1.43 円増加

ここで、説明指標とその見込める効果が言葉で記載されているため、絞込探索を用いて有益指標を探索・判断することができた。よって、表 1 の機能要件 II の期待する効果

として挙げた、ピンポイントで具体的に幾つかの説明指標の候補を絞り込めるが実現しているため、機能要件Ⅱを満たしていることを確認した。

機能要件Ⅲ 類似する説明指標による効果の見込みを並べて表示すること

本アプリケーションはピンポイントで有益指標を見つけるため、全体を俯瞰するようにデータを提示した方が、顧客を説得しやすくなる。例えば、店員の第二マグネット①への滞在時間という指標に言及する場合、他の場所に滞在した場合の効果もあわせて紹介することが望ましい。

図5のC.説明指標選択を用いて、絞り込みを行なう。店員全体の滞在時間と設定し、第二マグネット①以外のエリアの情報を抽出した結果をグラフ化したものが図7である。図7は、定員のエリア滞在によって売上向上するスポットの一覧をしめしており、これを用いて、被験者は、その中で一番効果の見込みが高いのは第二マグネット①ということが説明することができた。顧客へ第二マグネット①についてだけ述べるよりは、他のエリアと比べて第二マグネット①が一番よいと述べることで、有益指標の良さを伝えるストーリーを作成することができた。よって、表1の機能要件Ⅲの期待する効果として挙げた、選んだ指標が他より有益だと、ストーリーを持って顧客に説明できることが実現しているため、機能要件Ⅲを満たしていることを確認した。

これにより、機能要件Ⅰ、Ⅱ、Ⅲの3つの要件を満足したため、分析者が目的指標に関連する有益な説明指標を見つけることを本アプリケーションが支援できたことを確認した。

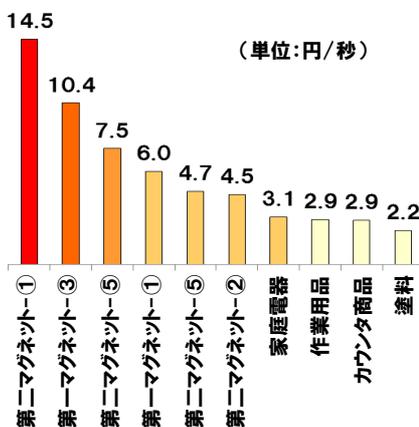


図7 顧客単価に対する高感度スポット

6. 考察

著者らは、本アプリケーションを用いずにモデル探索した場合に要した時間は1週間であったが、このアプリを用いることで、1時間程度に削減した。この背景には、所望の分析に対して、簡易に操作することができるため、結果出力のターンアラウンドタイムが短くなったのが理由と考えられる。

データとの関わり方の変化については、従来は仮説を作る専門家(コンサルタント)と分析する専門家(分析者)の作業分担が明確に分かれており、仮説を作成する人が顧客ヒ

アリングによって仮説を作り出すことに専念し、分析する人は仮説とデータから結果を出すことに専念していた。そのため、分析者は仮説が変わると1から作り直す必要があり、かなりの負担が課せられていた。本アプリケーションを用いることで、予め登録してある特徴量内において、インタラクティブに、顧客が納得しそうなストーリーを考えながらデータを分析していく。もし、仮説が変わったとしても、1から作り直す必要がなくなるため、負担が軽くなる。また、作業分担の必要もなくなり、コスト削減が期待できる。

施策提案については、一般的に、改善施策を顧客に受け入れられることが難しいと言われている。なぜならば、顧客側の組織状況や施策予算から制約条件が多いためである。本アプリケーションを使用することで、複数の視点の改善施策を立案できるため、改善施策が受け入れられる可能性が高くなると考えられる。

7. おわりに

我々の研究のゴールは、社会産業に関わるビックデータを入力とし、業績向上のためのモデル推定を行なう人工知能を実現することである。その初期検討として、扱いに困る大量データを自動的に整理することで、向上施策に結びつく有益指標を発見しやすくする支援技術の検討を行なった。我々は本報告で業績向上指標探索アプリケーションのプロトタイプを実装した。その特徴は単相関と直線回帰によって有益指標とその期待効果の関係を表示することである。有益指標を探索する際に必要な要件を、「膨大な数の説明指標を分類・整理して見せること」「各説明指標が目的指標にもたらす影響を日本語かつ具体的な単位で表示すること」「類似する説明指標による効果の見込みを並べて表示すること」と定義し、これらを満たしたことを確認した。さらに、本アプリケーションを用いて店舗客単価向上施策を立案・実施したところ、客単価が15%向上し、その有用性を確認した。

今回は、初期検討として、シンプルな単相関と直線回帰を用いて、目的指標を軸とし、大量データを自動的に整理することで、向上施策に結びつく有益指標を発見しやすくする支援技術の検討を行なった。今後は、改善施策の提案など、より高度な機能を追加していく。

謝辞

本研究を進めるにあたり、森脇氏、辻氏には、種々の御助言を頂き、また、大久保氏、早川氏には、ビジネス顕微鏡の機器を開発して頂き深く感謝いたします。(株)日立ハイテクノロジーズには、種々のサポートをして頂き深く感謝いたします。

参考文献

- [1] 荒宏視, 佐藤信夫, 矢野和男, 戸川望, 柳澤政生, 大附辰夫, “常時着用型センサ「ビジネス顕微鏡」による組織変革,” 電子情報通信学会 VLSI 設計技術研究会, Vol. 109, No. 462, pp. 43-47, 2010
- [2] 矢野和男, 荒宏視, 森脇紀彦, 栗山裕之, “人間行動の計測 - 機会発見の豊かな社会を創る,” 日立評論, Vol.91, No.4, pp. 30-33, 2009
- [3] 宮川雅巳, グラフィカルモデリング (統計ライブラリー), 朝倉書店, ISBN 4-254-12657-3, 1997

A.条件選択

STEP1:対象を選んでください
店舗: CH店

STEP2:期間と種類を選んでください
種類: 顧客別 期間: 店舗前半

STEP3:時間単位を選んでください
時間単位: なし

STEP4:目的変数を選んでください
目的変数 **B.目的指標選択**
顧客POS-購入金額-店舗全体

しきい値: 0.2

計算実行

キャンセル 自動掘下げ 2 階層まで

D.ネットワークダイアグラム表示

結果

C.説明指標選択

指標検索

指定したカテゴリや指標を画面中から探し出します

カテゴリ: キーワード

分類/指標: 未選択

探す 検索解除

設定書き出し読み込み

表示設定

[データをダウンロード](#)

E.リスト表示

顧客POS-購入金額-店舗全体

説明指標

Aグループ (全3指標) r値 ▲▼ 影響度 ▲▼ 元に戻す

- 顧客POS-購入金額-日用品 (r=0.4) 1円毎 (0.214%増) (こ) 1.1円増加 解説
- 顧客滞在時間-第一マグネット① (r=0.21) 1秒毎 (7.64%増) (こ) 13.3円増加 解説
- 顧客滞在時間-地点A (r=0.23) 1秒毎 (1.9%増) (こ) 9.15円増加 解説

Bグループ (全30指標) r値 ▲▼ 影響度 ▲▼ ▼続きを表示(残り25指標) 元に戻す

- 店員情報-全店員-会話-顧客 (r=0.36) 1秒毎 (0.509%増) (こ) 4.6円増加 解説
- 顧客接客-(滞在時間内)接客された時間-店舗全体 (r=0.25) 1秒毎 (6.65%増) (こ) 22円増加 解説
- 店員情報-全店員-場所(滞在)-第二マグネット① (r=0.25) 1秒毎 (3.19%増) (こ) 14.5円増加 解説

目的指標: 顧客POS-購入金額-店舗全体

説明指標: 店員情報-全店員-場所(滞在)-第二マグネット①

サンプル数: 304 p値: 0.000001527

1 実態

- 第二マグネット①の実態は最小値の場合は0.00000秒、最大値の場合は529秒で、平均は31.4秒でした。全体の8割は、2.25秒から77.1秒の間に分布しております。
- 第二マグネット①の長い場合(平均以上)と短い場合(平均以下)で比べると、店舗全体は実態として、2,976円と1,905円であり、1,072円違いました。
- より詳しく調べると、第二マグネット①が10秒増える毎に、店舗全体が145円増えるという比例関係が見られ、これは第二マグネット①の529秒以内で確認されました。
- 「店舗全体」全体の変動の27%を、この第二マグネット①が説明します。

2 店舗全体向上の期待効果

- 上記の実態を踏まえると、第二マグネット①を同じ平均10%向上させることが出来れば(31.4秒→34.5秒)、店舗全体は45.5円増加が期待されます。

F.実態と効果説明表示

- 店員情報-全店員-歩行・静止・歩行判定時の行動エネルギー和 (r=0.28) 1毎 (1.39%増) (こ) 12.3円増加 解説
- 基本情報-顧客の積極性・活動量-会話エネルギー活動時間 (r=0.22)

図5 アプリケーション表示画面