

## 二色擬似カラー表示を用いたPOSデータの可視化システムの作成

日大生産工(院) ○石川 了 日大生産工 吉田 典正

## 1. まえがき

効果的な経営データの解析は、経営活動の意思決定を補助するために重要である。経営に関するデータの一つにPOSデータがある。POSデータとはPoint-Of-Sales (POS) システムにより集計された店舗の販売履歴データである。

本研究では二色擬似カラー表示<sup>9)</sup>の手法を用いてPOSデータの可視化<sup>10)</sup>を行うシステムを作成し、POSデータの全体の傾向と詳細な値を同時に読み取れることを確認する。

## 2. 使用データ

本研究で使用するデータは流通経済研究所による2012年のある地方のスーパーマーケット18店舗の菓子類(キャンディー・キャラメル、チョコレート、チューインガム、ビスケット・クッキー、ファミリーアイスの5分類、12078種)の売上と店舗への来客数のデータである。

購買データの内容は、表1のように年月日、店舗コード、JANコード、購買点数、購買金額である。

客数データの内容は、表2のように年月日、店舗コード、来客数である。

表1 購買データ

年月日	店舗コード	JANコード	購買点数	購買金額
20120830	P05930	4901940022823	1	93
20120604	P27929	4969244502885	1	105
20120606	P75985	4902750956117	1	168
20121005	P24006	49812930	1	65
20121026	P23129	49404609	3	297
20120530	P40745	4901351019689	2	296
⋮				

表2 客数データ

年月日	店舗コード	客数
20120731	P45146	2912
20120707	P30730	1808
20120218	P36326	474
20121129	P72013	2626
20121001	P14726	2965
20121108	P40745	2537
20121031	P38325	2643
20120330	P23125	2170
20120425	P40330	1095
⋮		

## 3. 二色擬似カラー表示

二色擬似カラー表示<sup>9)</sup>とは、上下の帯の色の組合せとその長さ割合で数値を表す可視化手法である。二色擬似カラー表示は使用方法を学習することにより、折れ線グラフなどのExcelなどで利用できるグラフと比較し、非常に多量のデータを効率的に表示できる。

$n$  個のエリア  $A_i$  と各エリアの上限値  $a_{i+1}$  と下限値  $a_i$  の値と  $n+1$  個の色とその順序を設定し、スケールを作る。

図1のスケールの  $A_2(a_2$  と  $a_3$  の間) に所属する数値データを  $t$  とするとき、全体の長さを  $L$  とし、

$$y = \frac{t - a_2}{a_3 - a_2} \times L \quad (1)$$

によって計算できる  $y$  を用いて上下の帯の長さを決定する。

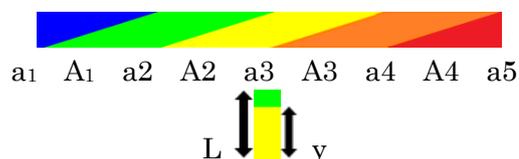


図1 スケール(上)と可視化例(下)

A Visualization System of POS Data using Two-Tone Pseudo-Coloring

Satoru ISHIKAWA and Norimasa YOSHIDA

#### 4. システム概要

本研究で作成したシステムでは、2012 年のある地方のスーパーマーケット 18 店舗の菓子類の売上、来客数、経過日数を横軸としたときの菓子類の売上の最小二乗法の傾き、来客数を横軸としたときの菓子類の売上の最小二乗法の傾きの 4 種類を二色擬似カラー表示で図 2 のように可視化する。

可視化画面では 1 週間および 1 ヶ月を判別しやすくするため、日曜と月末に空白、奇数月の上部に黒線を加える。

図 3 のユーザーインターフェース(UI)を実装し、可視化内容とスケールの値を変更するかを選択し、図 2 の表示結果をリアルタイムで変化させる。

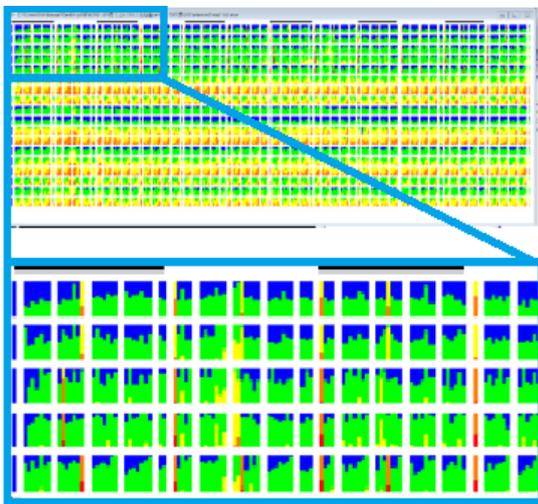


図 2 可視化画面

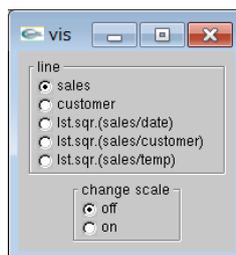


図 3 UI の画面

#### 5. 可視化結果

図 4 は、今回使用した地方の 18 店舗の全菓子類の売上金額を可視化したものである。左から日付を 1 月 1 日とし、スケールの色を低い方から青、緑、着、オレンジ、赤の順に、各エリアの範囲を 0 円、5 万円、10 万円、20 万円、100 万円の間としている。

図 4 から次のようなことが読み取れる。店舗

g、l、o にオレンジの部分が多いため売上金額が一定して 10 万円以上であることが読み取れる。逆に、店舗 j に青の部分が多いため売上金額が一定して 5 万円以下であることが読み取れる。

以下の特徴があることが分かる。

最初に、店舗 a、b、c、d、e で、月の初旬や中旬に、売上金額が大きい日があることが見て取れる。特に店舗 d、e では売上金額が 50 万円以上の日が複数あり、最も高い日は 90 万円ほどの日がある。各店舗の購買データから一定期間ごと、または毎月の特定の日付に販売している製品の単価を下げ、割引をしていることが分かった。このような店舗を、セール型と呼ぶことにする。

次に、店舗 f、g、h、k、l、m、n、o、p、q、r で、週末に売上金額が大きい日があるのが読み取れる。特に店舗 g、l、o は日曜日の売上金額が平日の約二倍の週がある。購買データからは原因がわからなかったが、客数データから週末に来客数が大きく増えていることがわかった。そのため、購買者一人あたりの購買金額が増えたのではなく、単純に菓子類の購買者数が増えたために売上が上がっていると考えられる。このような店舗を週末売上型と呼ぶことにする。

他に店舗 i、j は、日にちによってあまり大きな変化のないこと分かる。店舗 i では 6 月に売上金額が高くなっているが、購買データを調べると特定の商品が多く売れていることがわかった。

また、図 4 の多くの店舗が 2 月中旬および 12 月下旬に売上金額が高くなっていることが読み取れる。これらはそれぞれバレンタインデーとクリスマスの前後である。

前者はバレンタインデー付近なのでチョコレート類に注目する。図 5 は 2 月中旬にセールが行われていない図 4 の週末売上型の 3 店舗 f、g、h のあるメーカーのチョコレート類 (281 種) の売上金額の合計を 4 ヶ月毎に区切り可視化したものである。他の期間と比べるとバレンタインデーの当日の 14 日および前後の期間にチョコ類の売上金額がより高くなっていることが読み取れる。

後者はクリスマスパーティなどの準備とその後の年末セールなどが行われていたため菓子類全体の売上金額がより高くなっていると考えられる。

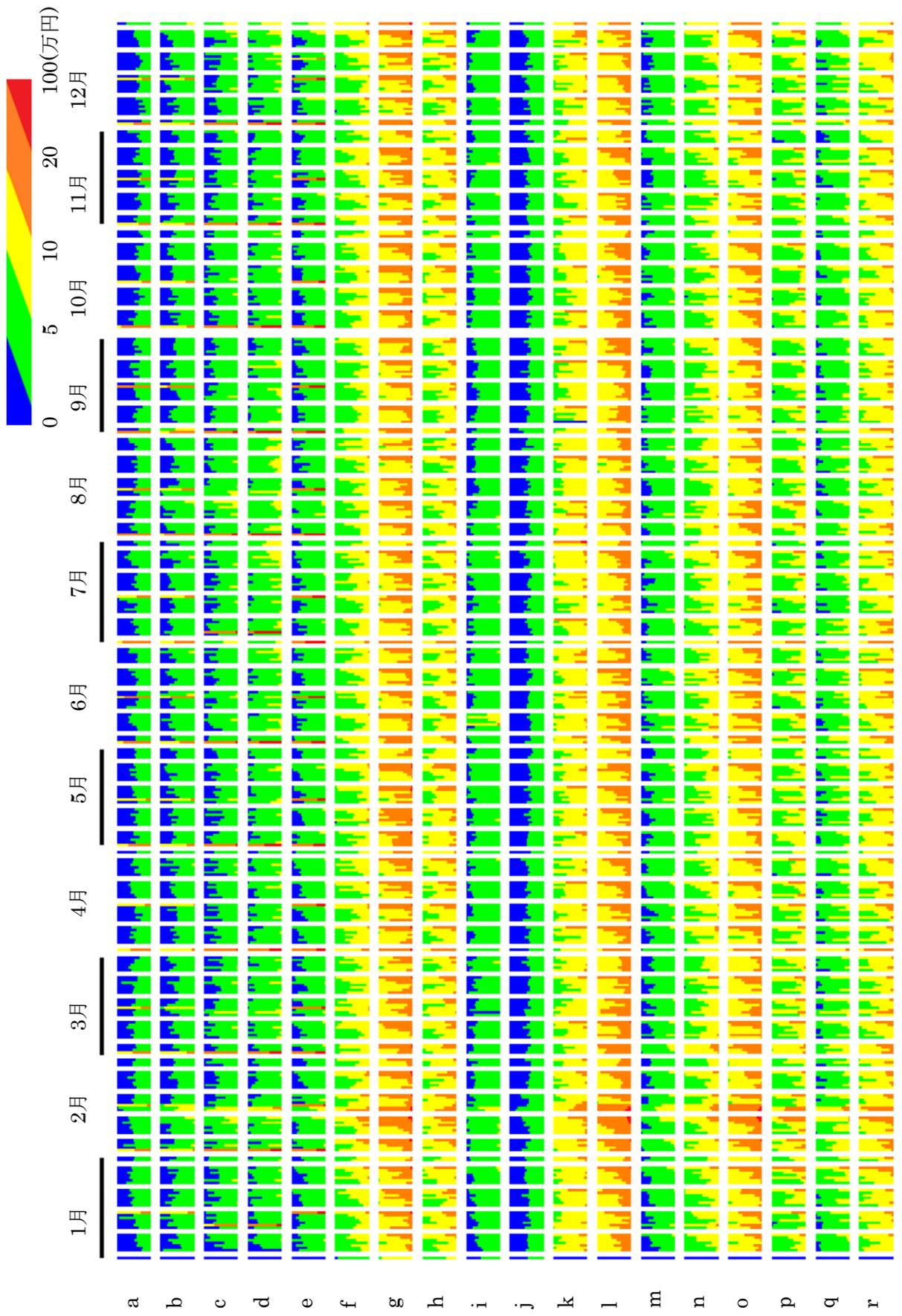


図 4 18 店舗の全菓子類の売上金額可視化

店舗名

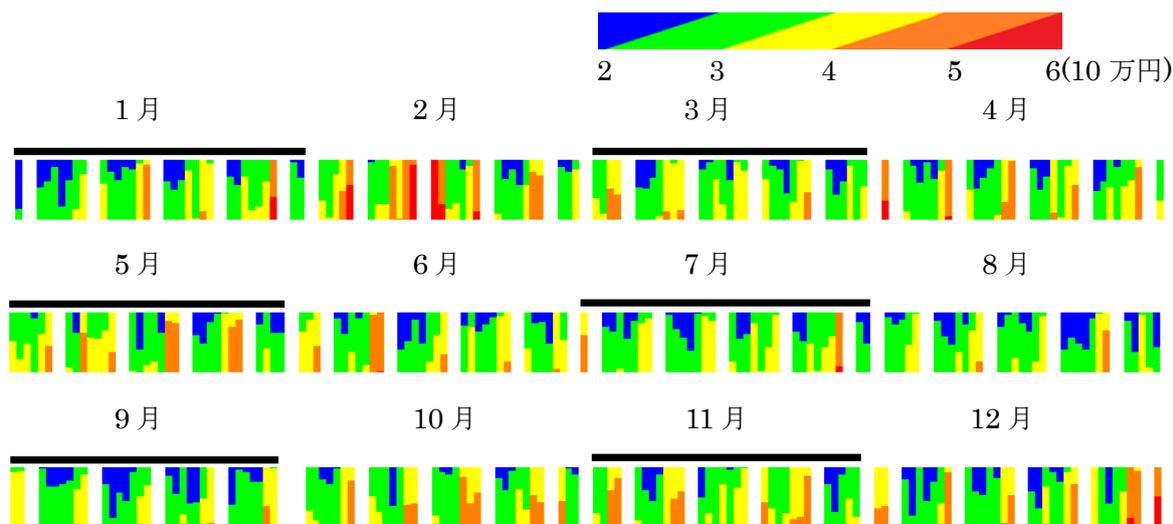


図5 店舗f・g・hのチョコレート類の売上金額可視化

## 6. まとめ

本研究では、作成したシステムで流通経済研究所より取得したある地方18店舗の菓子類のPOSデータに対して、二色擬似カラー表示を用いて店舗や種類、メーカー、製品ごとに4種類の可視化を行った。

可視化結果から二つの売上金額の傾向と特定の日に売上金額が高くなることがわかった。

現在、可視化を行ったのは特定の一地方のみだが、今後は、他の地方に関するPOSデータを用いて可視化できる店舗の数の増加や、可視化結果の切り替え機能の改良、より分かりやすく詳細な可視化、より良い配置、UIの探求を行う。

また、店舗についての情報および菓子類以外の購買情報も店舗の売上金額の傾向を予測するのに必要だと考えられる。

### 「参考文献」

- 1) Paolo Federico, Stephan Hoffmann, Alexander Rind, Wolfgang Aigner, Silvia Miksch, Qualizon Graphs: Space-Efficient Time-Series Visualization with Qualitative Abstractions, In Proceedings of the Working Conference on Advanced Visual Interfaces - AVI 2014.
- 2) Stephen, Few. Time on the Horizon. Visual Business Intelligence Newsletter, Jun/Jul 2008.  
[http://www.perceptualedge.com/articles/visual\\_business\\_intelligence/time\\_on\\_the\\_horizon.pdf](http://www.perceptualedge.com/articles/visual_business_intelligence/time_on_the_horizon.pdf)

- 3) Jeffrey Heer, Nicholas Kong, and Maneesh Agrawala, Sizing the Horizon: The Effects of Chart Size and Layering on the Graphical Perception of Time Series Visualizations, ACM Human Factors in Computing Systems (CHI), pp1303-1312, 2009.
- 4) Waqas Javed, Bryan McDonnell, Niklas Elmquist, Graphical Perception of Multiple Time Series, Visualization and Computer Graphics, pp927-934, 2010.
- 5) Milos Krstajic, Daniel A. Keim Visualization of Streaming Data: Observing Change and Context in Information Visualization Techniques, 2013 IEEE International Conference on Big Data, pp41-47, 2013.
- 6) GLUI User Interface Library, <http://glui.sourceforge.net>.
- 7) OpenGL Graphics Library, <https://www.opengl.org>.
- 8) William Playfair, [http://en.wikipedia.org/wiki/William\\_Playfair](http://en.wikipedia.org/wiki/William_Playfair).
- 9) Takafumi Saito, Hiroko Nakamura Miyamura, Mitsuyoshi Yamamoto, Hiroki Saito, Yuka Hoshiya, Takumi Kaseda, Two-Tone Pseudo Coloring : Compact Visualization for One-Dimensional Data, IEEE Symposium on Information Visualization, pp.173-180, 2005.
- 10) 常木 翔太, POSシステムの構築とPOSデータの可視化に関する研究, 日本大学生産工学部 マネジメント工学科卒業論文, 2013年度.