

ビッグデータの基本的な解析方法および事例

丸 山 友希夫

1. はじめに

IT 革命が起こり 20 年以上の月日が経ち、IT 社会、情報化社会なるものが形成され、IT は現代社会において必要不可欠な物となった。そして、近年においては、IT 社会の産物となるビッグデータが存在し、このビッグデータの解析方法に注目が集まっている。本プロジェクトのテーマにある ERP/SCM における在庫管理や生産管理において様々なデータや情報が作り出され、ビッグデータを形成することは言うまでもない。そこで、本論では、企業における ERP/SCM に関連する管理部門とビッグデータの関わり方を考慮し、ビッグデータの解析方法について述べる。そして、ビッグデータ分析の事例を示す。ビッグデータ分析の事例については、企業におけるマーケティング管理に関連すると考えられる事例を扱う。これは、昼食の満足に関するアンケート調査結果を用い、ビッグデータ解析方法の 1 つである因子分析を用いて、昼食に満足する要因を導出する事例を示す。

2. ERP/SCM とビッグデータと解析方法

本章では、本研究プロジェクトテーマにある ERP/SCM とビッグデータの関連について述べる。まず、生産管理において ERP (Enterprise Resources Planning) とは、企業資源を有効活用するための管理手法をさし、SCM (Supply Chain Management) とは、調達から販売までの流れを効率化するための管理手法をさす。例として、一般的な製造業における製品作製工程を図 1 に示す。



図 1. 製造業における製品作成工程

図1に示すように、製造業における一般的な製品作製工程は、開発・企画部門→設計部門→製造部門→品質管理部門→販売部門となる。この製品作成工程における設計部門において製造物が決まると原料等が調達され製造が始まる。したがって、原料等の調達を含め製造部門から販売部門までの管理方法（図1における斜線で囲まれた部分）がSCMということになる。そして、各部門については、開発・企画部門においては企画書が作成され、設計部門においては設計図が作成され、製造部門においては工程表が作成され、品質管理部門においては品質管理表が作成され、販売部門においては販売リストおよび配送リストが作成されるように成果物がデータまたは情報となって生じる。また、各部門においては人件費が発生し、販売部門においては製品を配送する輸送費が発生し、これら全て、すなわち企業資源の管理方法（図1における全体）がERPということになる。さらに、全体の流れを効率化するためには、各部門におけるデータまたは情報等を一元化する必要がある。これにより各部門間においてデータまたは情報の共有が可能となり、最終的には経営者が経営状況等を把握することが可能になる。つまり、全体の流れを効率化する一つ的手段として、図2に示すようにデータベースを導入しデータまたは情報を管理することが挙げられる。

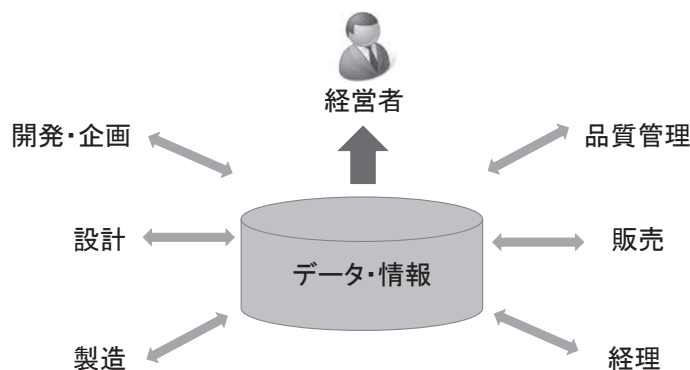


図2. データまたは情報の一元化

そして、このデータベース上のデータまたは情報がビッグデータとなる。一般的な企業のERPにおける経営管理項目については、予算管理、購買管理、在庫管理、人材管理、マーケティング管理、顧客管理等、様々な管理項目が存在する。そして、これらの管理を行うためにデータベース上のデータまたは情報を用いて解析・分析を行い、得られた情報をデータベースに追加する。例えば、マーケティング管理では、自社製品を効率よく売るために、販売部門において時系列分析等を用いてライバル社製品の売り上げ等の傾向を把握し、この傾向を基に製造部門へ製品作製数の提案をすることが可能となる。このように、ビッグデータの主な解析方法については、クロス集計分析、クラスター分析、因子分析、主成分分析、線形回帰分析、ロジスティック回帰分析、アソシエーション・バスケット分析等が使用される。

3. ビッグデータの解析事例

本章では、ビッグデータの解析事例として、昼食に満足する要因分析について挙げる。本事例の背景には、Suzanne ら（2013）によるランチ休憩におけるストレスの増加が創造性を減退させる可能性がある」と発表され、ここから、昼食の満足度がランチ休憩を有意義に過ごす 1 つの理由になると考えたことにある。そして、昼食に満足する要因についてアンケート調査結果を使用し、ビッグデータ解析の 1 つである因子分析を用いて昼食に満足する要因を導出することを試みた。本事例のフローを図 3 に示す。

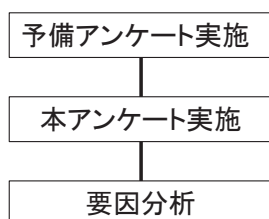


図 3. 昼食に満足する要因導出のフロー

図 3 に示すように、昼食に満足する要因を導出するために 2 回のアンケート調査を実施する。1 回目のアンケート調査は予備アンケート調査と位置づけし、2 回目のアンケート調査を本アンケート調査と位置づける。そして、本アンケート調査の結果を用いて、昼食に満足する要因を導出する。

(1) 予備アンケート調査の実施

予備アンケート調査においては、本アンケート調査項目を作成する際のキーワードを得ることが目的である。予備アンケート調査内容については、大別して下記する 2 つの内容について作成した。

- ・ 昼食で利用する店舗形態、メニュー種類等の動向調査
- ・ 昼食を取る際に重視する意識調査

予備アンケート調査項目は、回答者属性（性別、年齢等）以外に 4 項目を設定した。実施した調査の概要は次の通りである。

- ・ 調査期間：2016 年 4 月 12 日～4 月 19 日（1 週間）
- ・ 調査方式：対面式
- ・ 調査対象：10 代～40 代の男女 6 名

予備アンケート調査の結果については簡潔に示すが、昼食を取る際に重視する意識につい

ては、安さ、店舗までの距離、ボリューム、接客態度、栄養バランスの5つのキーワードを得た。

(2) 本アンケート調査の実施

予備アンケート調査結果から得た5つのキーワードを用い、これまで行ったアンケート調査項目作成の経験を基に、調査項目を新たに作成した。アンケート調査内容については、大別して下記する2つの内容について作成した。なお、本アンケート調査は学生を対象としている。

- ・ 普段の昼食に対する満足度の意識調査
- ・ 学食における昼食に対する満足度の意識調査

まず、1つ目のアンケート調査内容である“普段の昼食に対する満足度の意識調査”については、昼食を取ろうとする際に、学生がどのようなことを考えているかについて、単に昼食に満足する要因を調査することを目的としている。次に、2つ目のアンケート調査内容である“学食における昼食に対する満足度の意識調査”については、学生が実際に学食で昼食を取ろうとする際に、学生がどのようなことを考えているかについて、学食を限定とした昼食に満足する要因を調査することを目的としている。これら2つのアンケート調査において、1つ目のアンケート調査結果からは理想的な昼食に満足する要因を導出することと仮定し、2つ目のアンケート調査結果からは現実的な昼食に満足する要因を導出することと仮定する。したがって、昼食に満足する要因について理想と現実の差の傾向を導出する。アンケート調査項目は、回答者属性（性別、年齢等）以外に6項目（表1）を設定した。

表1. 本アンケート調査項目（回答者属性以外）

項目番号	項目内容
5	普段昼食として食べたいメニューを下記から選択してください
6	昼食の現状に満足していますか？
7	昼食を飲食店で利用する際に、次の項目についてどれくらい重視しますか？
8	講義期間中に最も利用する学食の施設を1つ選択してください
9	項目8で選択した施設を利用する際に、次の項目についてどれくらい重視しますか？
10	自由記述欄

上述した理想的な昼食に満足する要因に関するアンケート項目番号は7であり、現実的な昼食に満足する要因に関するアンケート項目番号は9である。また、これらの項目内容にある“次の項目”は同じ項目を設定した。これを表2に示す。そして、これら12項目について四者択一方式の評価方法（表3）を設定した。

表 2. 本アンケート調査項目 7 および 9 におけるサブ項目

項目番号	項目内容	項目番号	項目内容
1	安さ	7	品揃えの豊富さ
2	接客態度	8	クーポン
3	栄養バランス	9	ボリューム
4	素材の安全性	10	周りの評判
5	高級感	11	店の雰囲気
6	立地場所	12	ポイントカード

表 3. アンケート調査の評価方法

評価値	評価内容
4	大変重視にする
3	まあまあ重視にする
2	あまり重視しない
1	全く重視しない

また、調査項目 8 における施設を表 4 に示す。

表 4. 調査項目 8 における施設

項目番号	項目内容	項目番号	項目内容
1	第一食堂	8	パン屋（購買）
2	第二食堂	9	持参弁当
3	みのり寿司	10	出前
4	スチューデントセンター	11	学外の飲食店
5	アルテリーベ	12	コンビニエンスストア
6	弁当販売（第一食堂）	13	自宅
7	購買	14	食べない

実施した調査の概要は次の通りである。

- ・ 調査期間：2016 年 8 月 1 日～2016 年 8 月 30 日（1ヶ月間）
- ・ 調査方式：配布式
- ・ 調査対象：N工業大学の学生
- ・ 配布数　：150 部

(3) アンケート調査結果

今回のアンケート調査において 150 部を配付し、91 部（回答率：約 61％）を得た。この回答から、不備のあるもの（全ての調査項目に回答していない）やきちんと回答されていない（全て同じ回答である）ものを省き、有効回答数 85 部（有効回答率：約 61％）を得た。ここから示す調査結果については、この有効回答数 85 部の結果を用いる。まず、基本情報として性別、年齢について、それぞれ図 4 および図 5 に示す。図 4 に示すように、男性 81 人、女性 19 人の回答を得た。

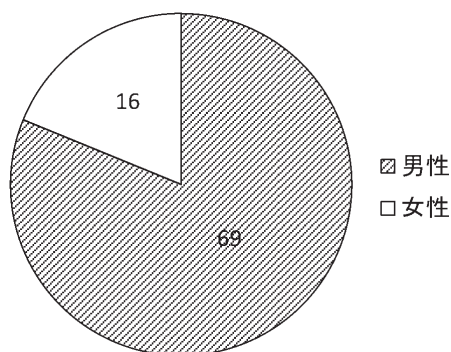


図 4. アンケート回答者の性別（単位：人）

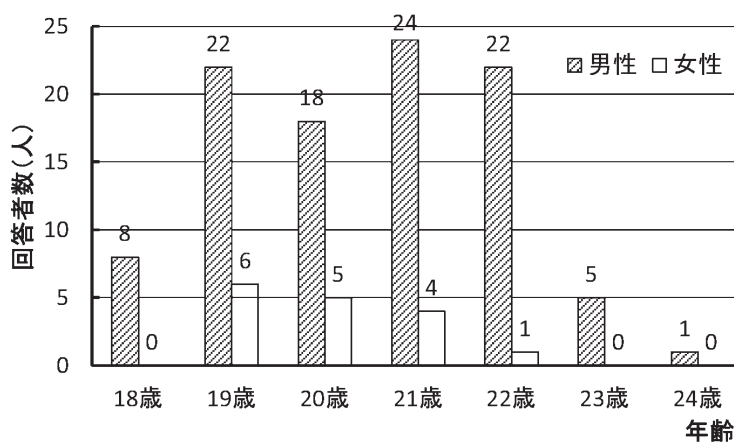


図 5. アンケート回答者の年齢

図 5 において、横軸は年齢を示し、縦軸は回答者数を示している。図 5 に示すように、食べ盛りの年齢であることがわかる。

(4) 昼食に満足する要因

まず、アンケート調査項目番号7の結果から“理想的な昼食に満足する要因”を導出する。はじめに、アンケート調査結果について表3に示したように数値化する。その結果の一部を表5に示す。

表5. 調査項目番号7の結果の数値化（一部）

No.	回答者	性別	年齢	所属学科	出身地	項目1	項目2	項目3	項目4	項目5	項目6	項目7	項目8	項目9	項目10	項目11	項目12
1	a-1	女性	20歳	A	埼玉県	3	2	2	2	2	3	3	3	3	3	3	2
2	a-2	女性	19歳	A	東京都	3	2	3	3	1	3	2	1	3	2	3	1
3	a-3	女性	21歳	C	高知県	4	2	1	2	2	3	3	1	3	2	3	1
4	a-4	女性	19歳	A	山梨県	4	3	3	4	2	3	3	2	2	3	3	2
5	a-5	男性	20歳	A	埼玉県	4	1	2	3	1	3	3	1	3	1	3	1
																	
81	m-1	男性	19歳	S	栃木県	3	2	3	2	1	2	2	1	3	2	2	1
82	m-2	女性	20歳		埼玉県	3	2	2	3	1	1	3	2	2	2	2	3
83	n-1	男性	20歳		C 東京都	3	3	2	2	2	3	3	2	3	3	3	2
84	n-2	男性	22歳	C	埼玉県	1	3	2	3	2	2	3	1	1	3	4	1
85	n-3	男性	22歳	C	埼玉県	2	2	2	3	2	2	2	2	3	4	2	1

 は未回答

そして、表5に示した結果を用いて因子分析を行った。因子分析の結果における固有値と寄与率を表6に示し、そのスクリープロットを図6に示す。表6および図6に示すように、固有値が1以上、累積寄与率が0.5以上となる因子数3を今回の分析から推定した。

表6. 理想的な昼食に満足する因子分析結果（固有値，寄与率）

No	固有値	寄与率	累積寄与率
1	3.076	0.256	0.256
2	1.592	0.133	0.389
3	1.422	0.119	0.507
4	1.156	0.096	0.604
5	0.962	0.080	0.684
6	0.883	0.074	0.758
7	0.761	0.063	0.821
8	0.563	0.047	0.868
9	0.506	0.042	0.910
10	0.429	0.036	0.946
11	0.376	0.031	0.977
12	0.275	0.023	1.000

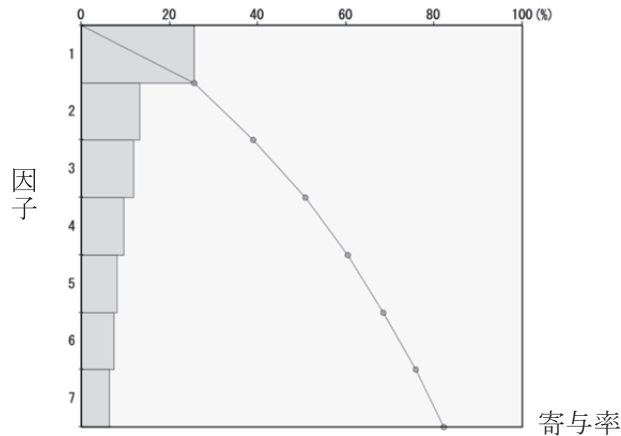


図 6. 理想的な昼食に満足する因子分析結果（スクリープロット）

さらに、推定した 3 因子について各項目番号の因子負荷量の結果について表 7 に示す。

表 7. 理想的な昼食に満足する因子負荷量

項目番号, 内容	因子1	因子2	因子3
④素材の安全性	-0.789	0.049	0.132
③栄養バランス	-0.568	0.133	-0.021
⑥立地場所	-0.421	0.138	0.215
⑫ポイントカード	-0.195	0.703	0.135
⑧クーポン	0.053	0.699	0.265
①安さ	-0.079	0.279	-0.108
⑩周りの評判	-0.003	0.029	0.736
⑪店の雰囲気	-0.392	-0.167	0.56
⑤高級感	-0.252	0.127	0.5
⑦品揃えの豊富さ	-0.287	0.282	0.378
②接客態度	-0.248	-0.088	0.367
⑨ボリューム	0.045	0.118	0.291

表 7 において、因子負荷量が正の数値は、プラスの方向に影響し、因子負荷量が負の数値はマイナスの方向に影響することを意味する。因子 1 に帰属する項目は、素材の安全性、栄養バランス、立地場所の 3 項目となり、これらの因子負荷量は負の数値である。したがって、素材の安全性は重視せず、栄養のバランスも重視せず、立地場所も重視しないことを意味する。したがって、因子 1 を“自己欲求”と命名する。同様に、因子 2 に帰属する項目からは、価格がキーワードとして挙げられる。したがって、因子 2 を“価格”と命名する。因子 3 に帰属する項目からは、店舗の環境がキーワードとして挙げられる。したがって、因子 3 を“店舗環境”と命名する。

次に、アンケート調査項目番号 9 の結果から“現実的な昼食に満足する要因”を導出する。表 5 で示したように、アンケート調査結果を数値化し、この結果を用いて因子分析を行った。

因子分析の結果における固有値と寄与率を表8に示し、そのスクリープロットを図7に示す。

表 8. 現実的な昼食に満足する因子分析結果 (固有値, 寄与率)

No	固有値	寄与率	累積寄与率
1	4.288	0.357	0.357
2	1.252	0.104	0.462
3	0.989	0.082	0.544
4	0.942	0.078	0.623
5	0.885	0.074	0.696
6	0.834	0.069	0.766
7	0.720	0.060	0.826
8	0.689	0.057	0.883
9	0.486	0.040	0.924
10	0.405	0.034	0.957
11	0.301	0.025	0.983
12	0.210	0.017	1.000

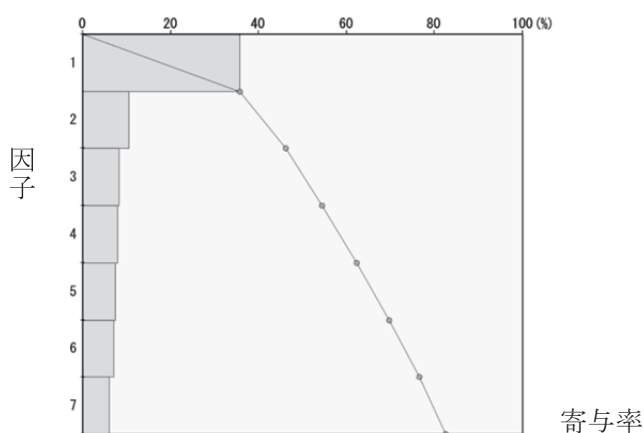


図 7. 現実的な昼食に満足する因子分析結果 (スクリープロット)

表 8 および図 7 に示すように、固有値は 1 以下となっているが 1 以上とみなし、累積寄与率が 0.5 以上となる因子数 3 を今回の分析から推定した。さらに、推定した 3 因子について各項目番号の因子負荷量の結果について表 9 に示す。表 9 に示すように、因子 1 に帰属する項目は、ボリューム、安さの 2 項目となり、コストパフォーマンスがキーワードとして挙げられる。したがって、因子 1 を“コストパフォーマンス”と命名する。同様に、因子 2 に帰属する項目からは、立地場所がキーワードとして挙げられる。したがって、因子 2 を“立地場所”と命名する。因子 3 に帰属する項目からは、店舗の環境がキーワードとして挙げられる。したがって、因子 3 を“店舗環境”と命名する。以上より、理想的な昼食に満足する要

因と現実的な昼食に満足する要因を表 10 にまとめる。

表 9. 現実的な昼食に満足する因子負荷量

項目番号, 内容	因子1	因子2	因子3
⑨ボリューム	0.378	0.047	0.106
①安さ	0.173	-0.009	-0.028
⑥立地場所	0.033	1.212	0.203
⑫ポイントカード	-0.246	0.063	0.895
⑧クーポン	-0.005	0.16	0.829
⑩周りの評判	0.38	0.174	0.584
⑤高級感	0.341	0.039	0.499
⑪店の雰囲気	0.475	0.312	0.496
②接客態度	0.415	0.035	0.447
④素材の安全性	0.338	0.065	0.431
⑦品揃えの豊富さ	0.353	0.214	0.391
③栄養バランス	0.182	0.148	0.329

表 10. 昼食に満足する因子の命名

	第1	第2	第3
理想的	自己欲求	価格	店舗環境
現実的	コストパフォーマンス	立地場所	店舗環境

表 10 に示すように、今回のアンケート調査結果においては、理想的な昼食に満足する要因と現実的な昼食に満足する要因が異なる傾向を示すことが分かる。特徴な点については、理想的な昼食に満足する要因において、価格が 2 番目の要因になっていることである。以前、他のプロジェクト研究において消費者購買行動に関する研究においても、食料品店を選ぶ要因についても価格は 1 番目の要因ではなかった。つまり、人々は購買する際には価格を一番気にしているわけではないことが、今回のアンケート調査結果においても同様の結果となった。また、今回の現実的な昼食については、大学における学食を対象にしたため、学食においてはコストパフォーマンスを求めていることが昼食に満足する 1 番の要因となる傾向となった。

このように、昼食の満足に関するアンケート調査結果を用いて因子分析を適応することにより、人々の意識傾向を導出する事が可能となり、この分析結果を基にさらなるマーケティング活動等に役立てることが可能となる。

4. まとめ

本論では、ビッグデータの基本的な解析方法を示し、その事例を示した。ビッグデータ解析の事例として、今回は昼食の満足に関するアンケート調査結果を用い、ビッグデータ解析の 1 つである因子分析を適応した。そして、理想的な昼食に満足する要因と現実的な昼食に

満足する要因について導出した。今回示した事例では、大学の学食における昼食を対象としたため、因子分析より導出した現実的な昼食に満足する要因については、学食に関連する要因が導出されたと考える。つまり顧客ニーズについて要因を導出することができたことにつながる。これは、ERP/SCM との関連においては、マーケティング管理、顧客管理に関連するものと考えられる。さらに、このマーケティング管理、顧客管理において、例えば新規店舗を展開する際に応用することが可能になるほか、在庫管理、購買管理、倉庫管理における、例えば食材の発注、食材の消費の予測等にも応用することが可能になることが考えられる。

今後の課題については、ERP/SCM における関連する管理間（例えばマーケティング管理と在庫管理）について効率の良い解析方法の提案を行っていくことが挙げられる。

参考文献

- 1) Suzanne Spenc et al. "The Impact of Food and Nutrient-Based Standards on Primary School Children's Lunch and Total Dietary Intake: A Natural Experimental Evaluation of Government Policy in England", PLOS One, Vol.8, pp.1-7 (2013).
- 2) 田村光太郎, 三浦航 「学生・技術者のためのビッグデータ解析入門」, 株式会社日本評論社 (2014).
- 3) HadoopTimes 編集部 「ビッグデータ分析の基礎」, <https://www.hadoop-times.com/bigdata-basic-analytics.html> (2016).