

階層クラスタ分析を用いた中堅・中小企業のクラウド活用状況に基づく分類

調査設計/分析/執筆: 岩上由高

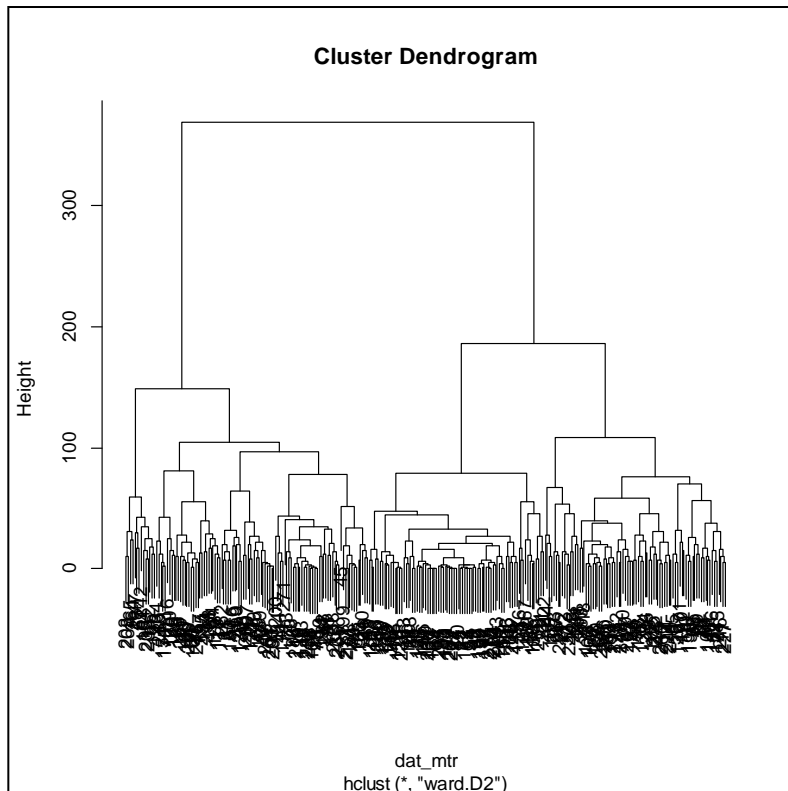
株式会社ノークリサーチ(本社〒120-0034 東京都足立区千住1-4-1 東京芸術センター1705:代表伊嶋謙二 03-5244-6691 URL: <http://www.norkresearch.co.jp>)は階層クラスタ分析を用いた中堅・中小企業のクラウド活用分類に関する調査結果を発表した。本リリースは「2015年版 中堅・中小企業におけるクラウド活用の実態と展望レポート」のデータを元に追加の集計/分析を行った結果をまとめたものである。(本リリースに掲載された結果は標準の調査レポートには含まれない)

調査対象企業: 年商500億円未満のクラウド導入済み企業(277社)
 調査対象職責: 企業経営もしくはITの導入/選定/運用作業に関わる立場
 調査対象地域: 日本全国
 調査対象業種: 組立製造業/加工製造業/建設業/流通業/卸売業/小売業/IT関連サービス業/サービス業
 調査実施時期: 2015年10月

<実データに基づくクラスタ分類は企業の課題を適切に反映したソリューション構築に有効>

- クラウドなどの比較的新しいIT活用の状況は企業属性だけでは十分に分類/把握できない
- グループ分けの事前情報が得られない場合には階層クラスタ分析が有効な手段の一つ
- クラスタ分類により「個別要件への対応力」や「運用費用の削減」などの課題が抽出される

一般に、企業におけるIT活用は年商などの企業属性に大きく依存する。だが、クラウドやスマートデバイスなどの比較的新しいIT活用に関しては「年商の高い企業では成功事例やIT投資額が多く、年商の低い企業では活用が遅れている」といった単純な解釈だけでは説明しきれない状況も垣間見える。



階層クラスタ分析におけるdendrogram(樹形図)の例
 出典: ノークリサーチ(2016年)

こうした場合には年商などの企業属性だけではなく、「クラウドの活用状況」といった実データに基づく企業の分類を行うことが有効なアプローチとなる。

左記に掲載した図は「導入済みの最も成功しているクラウドサービス」の活用状況データを元に年商500億円未満の企業(277社)を階層クラスタ分析の手法を用いて分類したdendrogram(樹形図)を表したものである。(詳細は後述)

このように、本リリースでは「2015年版 中堅・中小企業におけるクラウド活用の実態と展望レポート」のデータに対して階層クラスタ分析を適用し、実データに基づく企業分類を行うアプローチの一部を紹介している。

※本リリースに掲載された内容は調査レポートの標準には含まれない。

クラウドなどの比較的新しいIT活用の状況は企業属性だけでは十分に分類/把握できない

「2015年版 中堅・中小企業におけるクラウド活用の実態と展望レポート」では、企業が導入済みのクラウドサービスのうち「期待通りだった」あるいは「正しい選択だった」といった観点で見た場合に最も成功したと思われるものを一つ選び、それに関する評価を様々な観点から尋ねている。

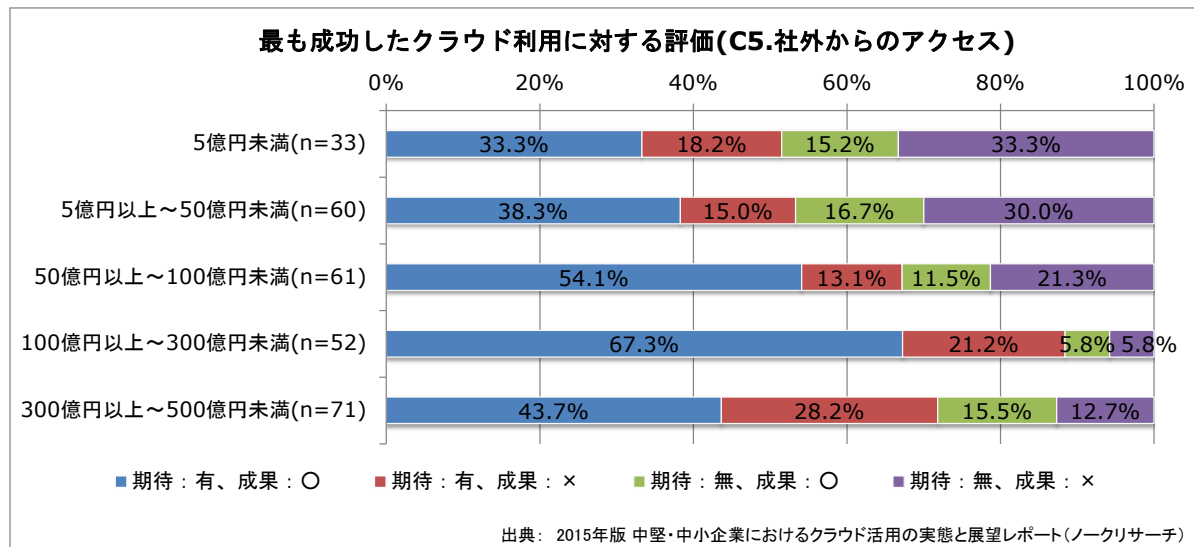
評価対象となる項目は以下の通りである。

- | | |
|-----------------|--|
| C1. 運用費用の削減: | 従来型のシステム形態と比べた場合、保守費用やバージョンアップ費用は軽減できたか？ |
| C2. 運用作業の削減: | 従来型のシステム形態と比べた場合、管理/運用の社内作業負担は軽減できたか？ |
| C3. セキュリティ強化: | ネットワークを経由したシステム利用を踏まえて、セキュリティは十分に確保されているか？ |
| C4. モバイル端末対応: | スマートフォンやタブレットなどのモバイル端末からもシステムを利用することができるか？ |
| C5. 社外からのアクセス: | オフィス内だけでなく、社外や移動中においてもシステムを利用することができるか？ |
| C6. 海外拠点展開: | 海外に新たな拠点を設けた場合も、そこから手軽にシステムを利用することができるか？ |
| C7. 国内拠点展開: | 国内に新たな拠点を設けた場合も、そこから手軽にシステムを利用することができるか？ |
| C8. 外部サービス連携: | 他社製も含めた他のサービスと必要に応じてデータをやりとりすることができるか？ |
| C9. 社内システム連携: | 社内に設置された既存システムと必要に応じてデータをやりとりすることができるか？ |
| C10. 機能面での豊富さ: | 従来型のシステム形態と比べた場合、機能の数や質は十分に備わっているか？ |
| C11. 個別要件への対応力: | 標準機能にない要件があった場合、何らかの方法で対応することができるか？ |

調査レポートでは上記の項目について、以下のような選択肢を設けて企業によるクラウドサービスの評価結果を尋ねている。

- | | |
|-------------|-------------------------------|
| 期待: 有、成果: ○ | 期待していた項目であり、実際に十分な成果が得られた |
| 期待: 有、成果: × | 期待していた項目だったが、十分な成果は得られなかった |
| 期待: 無、成果: ○ | 期待していなかった項目だが、予想に反して成果が得られた |
| 期待: 無、成果: × | 期待していなかった項目であり、予想通り成果は得られなかった |

以下のグラフは「C5. 社外からのアクセス」に関する企業の回答結果を年商別に集計したものである。「期待: なし、成果: ×」の回答割合については年商規模との相関が見られるが、「期待: 有、成果: ○」の回答割合は必ずしも年商規模とは比例していないことがわかる。このようにクラウド活用において成果を得られている企業層(=IT企業にとっては今後もIT投資の伸びが期待される企業層)は年商などの企業属性だけでは特定が難しいことがわかる。(次頁へ続く)



グループ分けの事前情報が得られない場合には階層クラスタ分析が有効な手段の一つ

前頁に述べた課題を解決するためには年商などの企業属性を用いた集計に加え、別のアプローチを加える必要がある。その一つが「階層クラスタ分析」である。階層クラスタ分析とはデータ間の「距離」に基づいて類似すると考えられるデータ同士をグルーピングする手法である。グルーピングに関する事前情報が与えられているわけではないので、教師なし学習によるデータの識別/分類に該当する。

階層クラスタ分析は

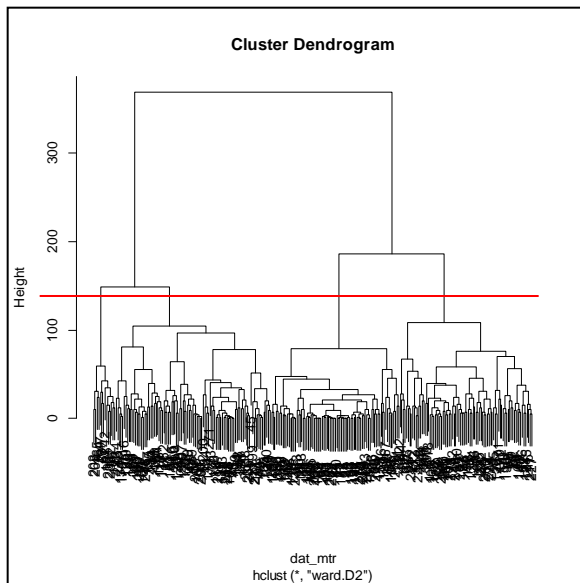
1. 「個々のデータ(サンプル)間の距離をどのように定義するか？」
2. 「クラスタ間の距離をどのように定義するか？」

によって得られる結果も変わってくる。

ここでは1.については「平方Euclid距離」(成分同士の差を二乗して合算したもの)を用い、2.としては「Centroid法(重心法)」(クラスタに属するデータ個数に基づく重心を求め、その重心間の距離をクラスタ値の距離として用い、距離の近いクラスタ同士を融合させていく方法)、および「Ward法」(ある2つのクラスタを融合させる場合、融合前と融合後のデータ分散変化が最小となるようにする方法)の2つを用いている。

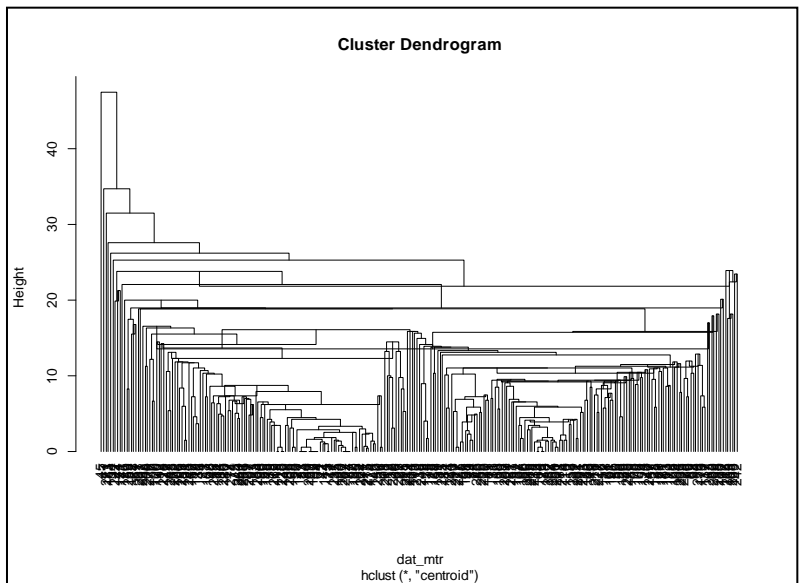
上記の1.と2.に基づいて分析を行っていくと、最初は互いに最も近い距離にある2つのデータによって構成される多数のクラスタから始まり、最終的に全データが属する1つのクラスタが形成される。この様子をグラフ化したものが以下のdendrogram(樹形図)である。

Ward法を用いた場合のdendrogram



出典：ノークリサーチ(2016年)

Centroid法を用いた場合のdendrogram



出典：ノークリサーチ(2016年)

一般的には「Ward法」の方が「Centroid法」を含めた他の方法と比べて感度の良いクラスタ分類が得られるとされている。実際、上記のdendrogram(樹形図)を見ても、「Ward法」による結果の方が整合性が取れているように見える。念のため、ここで用いた2つの手法の「Cophenetic相関係数」を確認しておく。上記で述べた分析手法は1.によって個々のデータ間の距離を算出した「距離行列」から、2.によってクラスタ間の距離を表す「Cophenetic行列」(同じ値を持つ行と列の組み合わせはdendrogram上で同じ高さで融合される)を生成する過程でもある。ここで「距離行列」と「Cophenetic行列」の相関係数を算出したものが「Cophenetic相関係数」であり、データ間の元々の距離関係と比べてどれだけ歪みが少ないか？を示す指標といえる。今回の場合「Ward法」の値は0.498、「Centroid法」の値は0.561である。「Cophenetic相関係数」の値が高い方が歪みは少ないことになるが、それが必ずしも分析結果の妥当性に直結するわけではない。また、分散を用いる「Ward法」は他の手法と比べて「Cophenetic相関係数」の値が低くなる傾向があるが、ここでの両者の差は著しいものではない。そこで、以降では「Ward法」による結果を採用し、その詳細について述べることにする。

クラスタ分類により「個別要件への対応力」や「運用費用の削減」などの課題が抽出される

dendrogram(樹形図)が得られた後は分類を行うクラスタ数を決定する必要がある。その際の明確な基準はなく、状況に応じて適宜判断する必要がある。ここでの最終的な目的は『クラウド活用を訴求する際に顧客となる企業を適切に分類し、それぞれに即したソリューション提案を行う』ことである。分類数が多すぎると、提案すべきソリューションの数も多くなってしまい現実的ではない。ここでは前頁の左側グラフで示されているように「Ward法」による結果から4つのクラスタを抽出してみることにする。

抽出された4つのクラスタが意味のある分類になっているかを確認するために分散分析を行っておく。「C1. 運用費用の削減」と年商区分(5億円未満/5~50億円/50~100億円/100~300億円/300~500億円)のそれぞれについて、4つのクラスタを水準とする分散分析を行った結果が以下の表である。

C1	自由度	偏差平方和
クラスタ間偏差	3	5949.3
クラスタ内偏差	273	4891.6

F値	F値を上回る確率
110.68	2.2×10^{-16}

クラスタ間偏差とクラスタ内偏差の比として得られるF値を超える確率は低く、同様の結果がC1以外の評価項目においても得られる。したがって、この4つのクラスタは意味のある分類であると考えられる。また、グラフは割愛するが年商区分とクラスタ分類の関係をプロットした結果を確認すると、この分類は年商区分とも異なっていることが確認できる。

では、実際にどのような分類となったのかを見ていくことにする。以下のグラフは4つのクラスタ(group1~group4)毎に「C1.運用費用の削減」「C5.社外からのアクセス」「C11.個別要件への対応力」の評価結果をBoxplot(箱ひげ図)として表したものである。(箱上部の縦点線上の横線:上限値、箱の上端横線:第3四分位点、箱内の横太線:中央値、箱の下端横線:第1四分位点、箱下部の縦点線下の横線:下限値、A:平均値、M:最頻値)

この結果を見ると、4つのクラスタには以下のような特徴があることがわかる。

クラスタ1(group1): C1、C5、C11の全ての項目において高い成果を得られている

クラスタ2(group2): C1、C5では成果が得られているが、C11における成果は十分ではない

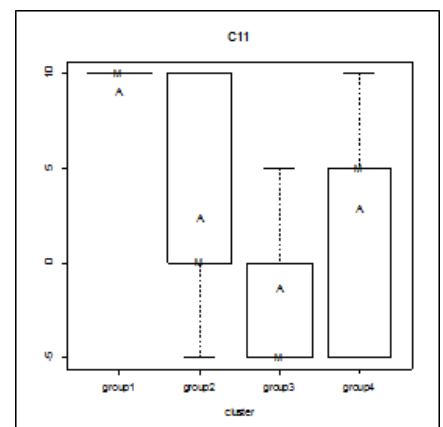
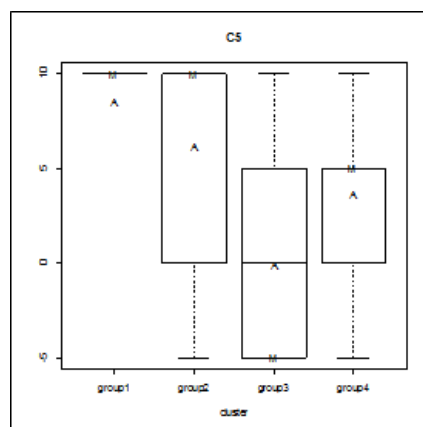
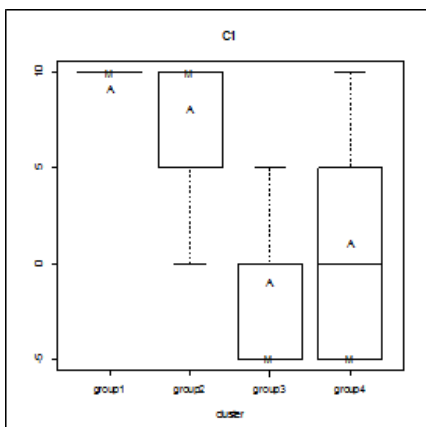
クラスタ3(group3): C1、C5、C11のいずれにおいても十分な成果を得られていない

クラスタ4(group4): C1では十分な成果が得られていないが、C5とC11では若干の成果が得られている

この結果を踏まえて、

クラスタ2に属する企業には「C11. 個別要件への対応力」の課題を解消するため、SaaSに加えてPaaSの活用を提案する
クラスタ4に属する企業には「C1. 運用費用の削減」の課題を解消するため、全体コストの洗い出しをまず行ってみる

などの取り組みを進めていくことが有効と考えられる。仮にこうした分類を行わなかったとすると、C1~C11の評価項目毎に課題を解決しなければならず、ソリューション構築の負担が大きくなってしまふ。IT活用が多様化しつつある昨今では、こうした実データに基づく企業分類を行うことによって、企業が抱える課題を的確に反映すると同時に労力を極力抑えたソリューション構築を行うことが重要と考えられる。



階層クラスタ分析の結果に基づくクラウド活用状況の整理(出典: ノークリサーチ 2016年)

各種調査レポートのご案内

「2015年版 中堅・中小企業におけるITアプリケーション利用実態と評価レポート」

13種類の業務アプリケーションにおける「導入社数シェア獲得/拡大の施策」や「ユーザ企業のニーズ」を網羅

レポート案内: http://www.norkresearch.co.jp/pdf/2015itapp_rep.pdf

各アプリケーション分野のサンプル/ダイジェストもホームページ上に掲載

価格: 180,000円(税別)

「2015年版 中堅・中小企業におけるIT投資の実態と展望レポート」

クラウドやスマートデバイスといった新たな潮流を受けて、中堅・中小企業のIT投資はどのように変わっていくのか？

レポート案内: http://www.norkresearch.co.jp/pdf/2015IT_usr_rep.pdf

ダイジェスト(サンプル):

ビッグデータ/ウェアラブル等の投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel1.pdf

基幹系システムの投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel2.pdf

情報系・顧客管理システムの投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel3.pdf

運用管理システムの投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel4.pdf

スマートデバイス関連の投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel5.pdf

PC関連の投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel6.pdf

ネットワーク関連の投資動向 http://www.norkresearch.co.jp/pdf/2015IT_usr_rel7.pdf

価格: 180,000円(税別)

「2015年版 中堅・中小企業におけるクラウド活用の実態と展望レポート」

クラウド形態で提供すべき業務システム、提供すべきでない業務システムは何か？その際に留意すべき事項は何か？

レポート案内: http://www.norkresearch.co.jp/pdf/2015cloud_usr_rep.pdf

ダイジェスト(サンプル):

クラウドへの移行状況と障壁および解決策 http://www.norkresearch.co.jp/pdf/2015cloud_usr_rel1.pdf

クラウドサービスのユーザ企業による評価 http://www.norkresearch.co.jp/pdf/2015cloud_usr_rel2.pdf

クラウドが販社/Sierの選択に与える影響 http://www.norkresearch.co.jp/pdf/2015cloud_usr_rel3.pdf

価格: 180,000円(税別)

「2015年版 中堅・中小企業におけるPC/スマートデバイス活用の実態と展望レポート」

スマートフォンやタブレットはPC市場を侵食していくのか？既存のPC販売とスマートデバイスへの転換は両立できるのか？

レポート案内: http://www.norkresearch.co.jp/pdf/2015mobpc_usr_rep.pdf

ダイジェスト(サンプル):

Windows 10の中堅・中小企業への影響 http://www.norkresearch.co.jp/pdf/2015mobpc_usr_rel1.pdf

PCとスマートデバイスの選択比率と方針 http://www.norkresearch.co.jp/pdf/2015mobpc_usr_rel2.pdf

スマートデバイス端末の形状やOSの選択 http://www.norkresearch.co.jp/pdf/2015mobpc_usr_rel3.pdf

スマートデバイス提案に有効な活用シーン http://www.norkresearch.co.jp/pdf/2015mobpc_usr_rel4.pdf

価格: 180,000円(税別)

本ドキュメントの無断引用・転載を禁じます。引用・転載をご希望の場合は下記をご参照の上、担当窓口にお問い合わせください。

引用・転載のポリシー: <http://www.norkresearch.co.jp/policy/index.html>

株式会社 ノークリサーチ 調査設計、分析、執筆: 岩上由高

東京都足立区千住1-4-1東京芸術センター1705

TEL 03-5244-6691 FAX 03-5244-6692

inform@norkresearch.co.jp <http://www.norkresearch.co.jp/>