



---

# 卒業研究報告書

令和6年度

---

研究題目

BtoBの契約型コンサルティング企業における  
顧客特性と販売情報の分析

---

指導教員 上野秀剛 准教授

---

氏名 由井陽都

---

令和7年2月28日 提出

奈良工業高等専門学校 情報工学科

# BtoBの契約型コンサルティング企業における

## 顧客特性と販売情報の分析

上野研究室 由井陽都

ほとんどの企業は顧客情報や販売履歴を何らかの形でデータとして蓄積し、売り上げの推移や顧客特性と売り上げの関係を把握するために分析を実施する。顧客特性の分析に関する研究は盛んに行われており、多様な業種やビジネスモデルの特徴に適した分析手法を提案する研究や、新たなアプローチを試みる研究もある。しかし、顧客情報や販売履歴の適切な分析手法が未提案のビジネスモデルも存在する。また、蓄積したデータを十分に活用できていない企業も存在する。そこで、本研究ではデータ活用が不十分かつ適切な分析手法が未提案のBtoBの契約型コンサルティング企業の顧客情報や販売履歴を分析し、データ理解を進めることで同様のビジネスモデルを展開する企業にとって有益な知見を提供することを目的とする。分析のために、RQ1) 月単位の新規契約数の要素（長期トレンド・季節性・残差）の推移、RQ2) 規格ごとの問い合わせ件数と成約率の差、RQ3) 顧客特性（従業員数・業種・エリア区分）ごとの規格別での契約数の相対的な差、RQ4) 契約金額と関係する顧客特性（従業員数・業種・エリア区分）の関係性、以上のRQ4つを問いとして設定する。RQ1の分析では最初契約日が2019年5月～2024年4月までの新規契約に関するデータを対象として、年・月ごとに契約数を集計し、STL分解を用いて契約数を長期トレンドと季節性、残差に分け、月単位の新規契約数の要素の推移と内訳を明らかにする。STL分解の結果、長期トレンド成分から上昇傾向が停滞傾向に切り替わった時期が明らかになり、ビジネスが停滞気味であることを示唆した。季節性成分からは、月ごとに契約数に差があることや、月ごとの傾向が変化していることが明らかになった。RQ2の分析では、問い合わせに関する表から新規契約に関する問い合わせを対象として、規格ごとに問い合わせ数と成約率を分析する。問い合わせ件数が最も多かったのは規格Pで、最も低いのは規格Eだった。成約率が最も高かったのは規格I、低かったのは規格P、規格Eだった。問い合わせ数が多く成約率が低い規格Pについては、他の規格に比べて改善による効果が高い可能性が明らかになった。RQ3の分析では新規契約に関するデータを対象とし、顧客特性ごとに規格別でクロス集計を実施し、カイ二乗検定結果のp値が0.05を下回った結果を有意な偏りがあるとする。従業員数区分では主要な規格で有意な結果は見られなかった。エリア区分と産業分類の結果はほとんどの規格で有意な結果がみられ、契約数の相対的な差が明らかになり、偏りの傾向が似ていることから顧客企業の産業分類と所在地の関連が示唆された。RQ4の分析では新規契約に関するデータを対象とし、顧客特性を独立変数、累計金額を従属変数として重回帰分析を実施することで各顧客特性が累計金額に与える影響の大きさと切片の大きさを明らかにする。各変数の係数の絶対値の大きさから、各変数が累計金額に与える影響の大きさが明らかになり、係数の符号からは累計金額に与える影響がプラスとマイナスのいずれであるかが明らかになった。

# 目次

1	はじめに	1
2	関連研究	3
3	準備	5
3.1	STL分解	5
3.2	重回帰分析	5
4	実験	7
4.1	データセット	7
4.2	月ごとの新規契約数	8
4.3	規格ごとの新規契約の成約率	8
4.4	顧客特性と規格別の契約数	9
4.5	契約金額と顧客特性の関係	10
5	結果と考察	12
5.1	STL分解	12
5.1.1	長期トレンド	12
5.1.2	季節性	13
5.1.3	残差	14
5.1.4	RQ1への回答	15
5.2	規格ごとの成約率	16
5.3	顧客特性と規格別の契約数	17
5.3.1	従業員数区分	17
5.3.2	エリア別	17
5.3.3	産業分類	18
5.3.4	RQ3への回答	19
5.4	顧客特性と契約金額	19
6	おわりに	22
	謝辞	24
	参考文献	25
	付録	26

# 1 はじめに

企業は顧客情報や販売履歴をデータとして蓄積し、売り上げの状況や推移の特徴、顧客特性と売り上げの関係性などを把握するために分析を実施する。また、蓄積したデータをもとに将来利益の予測や効果的な経営判断、経営戦略の策定が行われており、様々な分析手法が提案されている[1, 2, 3]。代表的な分析手法の1つにCLV(Customer Lifetime Value)分析がある。CLV分析とは、顧客との関係が続く間にどれほどの利益をもたらすかを算出する手法であり、顧客関係管理(Customer Relationship Management)の1つの指標となる。CLVの算出には様々な定義や予測モデルが存在し、ビジネスモデルに適した算出式や予測モデルを開発・提案する研究も多い[3, 4]。

多くの企業は何らかの形で顧客情報や売り上げのデータを蓄積しているが、一方で、蓄積したデータを分析して経営判断や経営戦略の策定に有効活用できていない会社も存在する。また、先述の通り様々なビジネスモデルに適したデータ活用の研究は多いが、まだ研究で適切なデータ活用の提案がされていないビジネスモデルも存在する。顧客関係管理に関する研究において、小売り業などのBtoC企業や非契約型を対象とする研究に比べてBtoB企業や契約型を対象とする研究は少ない[2, 5]。さらに、BtoBかつ契約型のコンサルティング企業を対象とする顧客関係管理に関する知見を提供する研究は行われていない。ビジネスモデルや業種が異なると、知見も異なると考えられる。そこで、データ活用が十分でない、かつ適切な分析の算出や予測モデルの研究がないビジネスモデルの企業のデータを分析することで当該企業だけでなく同様のビジネスモデルを展開する企業の経営判断や経営戦略の策定に有益な知見を提供することができる。

本研究はデータ理解が不十分な企業の顧客情報や販売履歴のデータを分析し、データ理解を進めることで利益向上に有益な知見を提供することを目的とする。また、蓄積されたデータ中のノイズとなるデータか否かを判別することで、今後のより詳細な分析や機械学習への応用にもつながる。

本研究では、共同研究を実施している企業から提供いただいたデータを整理・前処理した上で分析・考察をする。この企業は企業を顧客としたコンサルティングを業務としており、顧客企業がISO認証規格を取得するための支援を商品としている。主要な商品は4種類の規格のコンサルティングであり、それぞれの規格について新規に取得することと、取得した規格を維持することの2つに分けられる。契約は期間を定めて結ばれる。また、1度契約した顧客が、契約満了後に取得した規格の維持や別規格に関する契約を追加で結ぶこともある。顧客企業は、企業規模の大小や様々な業種など多様な属性を持つ。提供データの分析結果から考察とデータ理解を進め、データ理解に基づいて経営判断や経営戦略策定に有用な知見を明らかにする。

具体的には以下の4つの問い (Research Question) を立て、分析と考察を実施し、データ理解と経営判断や経営戦略策定に有用な知見を得る。

RQ1 月単位の新規契約数の要素（長期トレンド・季節性・残差）はどのような内訳・推移となっているのか

RQ2 主要4規格別の問い合わせ件数とその成約率にどのような差が存在するか

RQ3 顧客特性（従業員数・業種・エリア別）ごとの主要4規格別で契約数が相対的に多い特性は何か

RQ4 契約金額に関する顧客特性の間にどのような関係性が存在するのか

以下、2章では分析に用いるSTL分解やカイ二乗検定、CLVの関連研究について説明し、3章で本研究の各問に対して用いるデータセットや実施する分析実験について示す。4章では本実験の結果と、結果に基づいた考察を述べる。5章では本研究のまとめと今後の発展を述べる。

## 2 関連研究

顧客分析に関する研究は、対象とするビジネスモデルや目的の多様性からこれまでに数多く行われている。顧客分析の多くにおいて重要視されるのは優良顧客の存在である。優良顧客を重要視する理由として、集団の評価や報酬は全体の2割ほどの構成員に集中するとされるパレートの法則が挙げられる。Pradnyaらはインドネシアの美容クリニックの1店舗を対象としてCLVを算出し、D.A.Kandeilらによって開発・提案されたRFMモデル[5]を拡張した提案手法のLRFMによるCLVの算出結果に基づいたクラスタリングで顧客をセグメント化している[6]。機械学習を用いた顧客分析を行った研究も存在する。高島らは自動車市場を対象としてエージェントベースモデリングによって優良顧客の特徴を抽出した[7]。伊藤らはKnowledge Graph Attention Networkをユーザの属性情報を考慮するモデルへ拡張し、顧客分析を実施した[8]。

契約型サービスを対象とした研究は少数ながら存在する。小野は契約型サービスを対象として顧客離脱と顧客維持に着目して議論を述べ、満足や信頼など、企業と顧客の関係性の構成概念を用いた収益性を直線的で因果連鎖的な方程式とは異なった理論や方法が必要であることを定性的アプローチや理論研究の発展から示した[2]。蓮本らはプラットフォームビジネスを対象として、非負値行列因子分解を用いて顧客の購買パターンを抽出し、CLVの予測モデルを提案し、実用性を示した。さらにRFMS指標、非負値行列因子分解で抽出された購買パターン、CLVの相関行列をヒートマップで可視化することでRFMS指標がCLVの値へ与える影響の大きさとCLVの予測精度が向上したことを確認した[3]。福本らは既存顧客に働きかけてCLVを高めることを保険業界における契約者が契約を最適なものに最新化することとして、成約率予測モデルを構築するとともに適切な維持介入についてDMの再送期間を判別可能とした[9]。

経営戦略の策定や経営判断の指標に関する研究は、CLVが広く知られており、また、CLVに関する研究も盛んである。CLVはRFM(Recency, Frequency, Monetary)モデルをはじめとした様々な方法で予測・算出される。CLVの算出やCLVの予測モデルの開発・評価を行う研究は、多様なビジネスモデルを対象に多く行われている。阿部はRFMデータについて、RFM自体は顧客特性を直接表す指標ではないことに注意が必要としたうえで、データの観測終了時点や観測期間中の顧客の離脱タイミングによってRecencyやFrequencyが変化することを問題点であると指摘し、さらに、RFMがCLVへ与える影響を百貨店のメンバー顧客のデータから分析した[10]。JuliaらはCLVの概念について、近年のインターネットの普及に伴い直接の購買がない顧客が他の顧客や見込み顧客に影響を与えた場合にもたらされる利益が存在することに着目した。ヨーロッパのスポーツに関するオンラインソーシャルネットワークの利用者を対象として分析した結果から、CLVを拡張し

た CLNV(Customer Lifetime Network Value)を提案している [4].

これらの既存研究は、顧客特性の特徴を捉えることで企業の利益を向上させることや経営の指針を定めることに役立つことを目的とする点において本研究と類似する。しかし、BtoBかつ契約型ビジネスのコンサルティング企業を研究対象とする本研究と研究対象のビジネスモデルが異なるため知見が異なり、既存研究の知見をそのまま使用できない可能性がある。そこで本研究では、BtoBの契約型コンサル業の企業の顧客情報や販売履歴のデータを対象とした、顧客特性の分析手法を提案するための足掛かりとして顧客情報や販売履歴のデータ理解を進め、有用な知見を得る。

## 3 準備

### 3.1 STL分解

時系列データの分析手法の1つにSTL分解がある[1]。STL分解は時系列データを入力とし、長期トレンド・季節性・残差に分解した結果を出力する。

STL分解はLoess平滑化を繰り返し用いることで長期トレンド成分と季節性成分の推定し、最後に元データの時系列データから差し引くことで残差成分を計算する。Loess平滑化はすべてのデータ点を順番に、注目する点とその近傍の点に対して重み付き最小二乗法を用いることで局所的な回帰による平滑化値・回帰線を求める。具体的なステップを次に示す。

1. 最初のLoess平滑化で初期トレンド成分を推定する。
2. 元データから初期トレンドを削除した後、季節周期ごとにLoess平滑化を用いて季節性成分を推定する。
3. 元データから季節性成分を削除してLoess平滑化を用いることで長期トレンド成分を推定する。
4. 季節性成分と長期トレンド成分を元データから差し引いて残差成分を計算する。

例として数年間の月ごとの平均気温を入力とした場合を考えると、長期トレンドは右肩上がりの上昇傾向が見られ、平均気温が全体的に上がってきていることがわかり、地球温暖化の影響で全体的に気温が上昇していると考えられる。季節性からは7,8,9月は気温が高く12,1,2月は気温が低いといった周期的なパターンがわかる。残差からは長期トレンドと季節性で説明しきれなかった値が出力され、記録的な寒波が2月に何度も訪れ、例年に比べ大きく気温が低いなど、突発的なイベントや外的要因による異常と言える値がわかる。

本研究では新規契約数の60ヶ月にわたる推移を対象にSTL分解を行うことで新規契約数の変化に影響する要素を把握する。

### 3.2 重回帰分析

重回帰分析とは、目的変数（従属変数）に対して複数の説明変数（独立変数）が与える影響の向きと大きさを分析する手法である。入力として目的変数と複数の説明変数を与えることで、各説明変数に対する係数と切片が出力される。簡単な例として、目的変数にはテストの点数や売上金額、説明変数には勉強時間や性別、自宅と学校の距離、最寄り駅と店舗の距離などが挙げられる。例に挙げた性別のような数値でない説明変数はダミー変数化することで用いられる。重回帰分析の結果として、3.2.1式に示す形の式が導かれる。

$$\begin{aligned}
\text{目的変数} &= \text{説明変数}_1 \times \text{係数}_1 \\
&+ \text{説明変数}_2 \times \text{係数}_2 \\
&+ \text{説明変数}_3 \times \text{係数}_3 \\
&\dots \\
&+ \text{切片}
\end{aligned}
\tag{3.2.1}$$

各係数は、その説明変数が目的変数の対して与える影響の大きさと向きを表す。切片は説明変数では説明されず残った値を表す。

重回帰分析の内部処理では、目的変数の実測値・回帰式によって導かれる予測値・重回帰式を作成した際の回帰係数・説明変数の実測値をそれぞれベクトル化・行列化し、予測値と実測値の差が最少になるように係数を導く。

重回帰分析の結果から、要因分析として係数の大小から各要素の影響を把握できる。また、得られた式を用いる事で目的変数の予測分析も可能となる。

## 4 実験

企業を顧客とするコンサルティングを商品とする企業の顧客情報・販売履歴などのデータを対象に分析実験を実施する。1章に示した本研究の問い (Research Question) を再掲する。

RQ1 月単位の新規契約数の要素（長期トレンド・季節性・残差）はどのような内訳・推移となっているのか

RQ2 主要4規格別の問い合わせ件数とその成約率にどのような差が存在するか

RQ3 顧客特性（従業員数・業種・エリア別）ごとの主要4規格別で契約数が相対的に多い特性は何か

RQ4 契約金額に関係する顧客特性の間にどのような関係性が存在するのか

以下、4章の各節で上記の4つの問いそれぞれについて、データセット、前処理、分析手順を説明する。なお、本研究の分析実験に用いるデータセットはいずれも csv ファイルとして保存されており、前処理や分析には python コードを用いて分析を実施する。

### 4.1 データセット

本研究では BtoB の契約型コンサルティング業の企業から提供された顧客企業の情報や販売履歴、契約関連の問い合わせに関する情報を記録したデータセットを用いる。データセット中の顧客情報に関するデータ群の項目は顧客企業の社名や最初契約日、認証を取得する理由、担当事業所、顧客企業の産業分類など多数に及び、4,640 件のデータがある。問い合わせに関するデータ群の項目は新規／運用やステータス、お問い合わせ内容（規格）など多数に及び、12,991 件のデータがある。新規契約に関するデータ群の項目は従業員数や顧客企業の所在地、契約内容など多数に及び、1,114 件のデータがある。

本研究で共同研究を実施する企業は、顧客企業が規格認証を取得することに対するコンサルティングを商品としており、主要な商品は4種類（規格Q, 規格E, 規格I, 規格P）である。規格QはISO9001（品質マネジメントシステム）、規格EはISO14001（環境マネジメントシステム）、規格IはISMS・ISO27001（情報セキュリティマネジメントシステム）、規格PはPマーク（個人情報保護マネジメントシステム）の認証取得コンサルティングサービス商品の呼称である。それぞれの規格について認証取得のためのコンサルティングのサービスである新規契約と、認証取得後の運用をコンサルティングするサービスの追加契約が存在する。また、主要4規格以外の規格認証は“その他”として記録されている。

## 4.2 月ごとの新規契約数

RQ1“月単位の新規契約数の要素（長期トレンド・季節性・残差）はどのような内訳・推移となっているのか”を明らかにするために、データセットから新規契約に関する項目を取り出し分析に用いる。顧客情報が記録された表から、最初契約日が2019年4月～2024年5月の2429件のデータを対象として、最初契約日（年月日）から年と月を取り出す。また、最初契約日が空欄となっているデータを除いて分析に用いる。契約数を年・月ごとに集計して得られる集計結果[YYYYMM:xx件]に対してSTL分解を行う。

本実験ではpythonのstatsmodelsライブラリを用いる。本実験におけるSTL分解の目的は一月単位での新規契約数の成分分解であるため、分解を行う周期として「period = 12」を渡して実行する。

STL分解によって得られる長期トレンド成分は企業のビジネスの成長・衰退の推移や市場の需要の増減を表す。本研究で対象とする一月単位の新規契約数の時系列データにおいては、長期トレンドを契約数の基準値として、全体的な契約数の増減を表していると解釈する。季節性成分は元の時系列データの周期的なパターンを表す。本研究で対象とするデータにおいては、新規契約が多い月・少ない月を表すと解釈する。残差成分は元の時系列データから長期トレンドと季節性で説明できずに残った契約数として解釈する。

## 4.3 規格ごとの新規契約の成約率

RQ2“主要4規格別の問い合わせ件数とその成約率にどのような差が存在するか”を明らかにするために、データセットから各規格の新規契約に関する問い合わせの記録を取りだし、分析する。本研究で用いるデータにおいて問い合わせとは、問い合わせフォームや電話など何らかの形で顧客からの契約に関わるアクションを意味する。本研究で用いるデータは顧客からのアクション及びアクションに対する共同研究先企業側の記録を1レコードとしたデータである。問い合わせに関するデータ群から、すべての12991件のデータを対象として、各規格のフラグの有無と受注状況から規格別の問い合わせ数、成約数を集計する。問い合わせには認証規格を新規に取得する契約である「新規」と、認証規格取得後の規格運用に関する契約である「運用」の2種類があり、本研究では新規顧客獲得に関する知見として新規契約の分析をするため、新規のみを分析の対象とする。1件の問い合わせで複数規格に該当する場合、該当する規格すべてにフラグが立っている。本実験の分析上では複数規格でもそれぞれで1件ずつとして集計する。以下に分析対象とするデータに対する処理の手順を説明する。

1. 問い合わせに関するデータ群から上記で説明した、分析対象とするデータを取り出す。

2. 4規格それぞれの新規契約にフラグが立っているレコード件数を数える．各規格のレコード件数を問い合わせ数とする．

3. 前手順の問い合わせ数から受注状況のステータスが受注以外のレコード件数を除いた数を各規格の成約数とする．

上記手順で得られる4規格のそれぞれの「問い合わせ数」と「成約数」から成約率を求める．

$$\text{成約率} = \frac{\text{成約数}}{\text{問い合わせ数}} \quad (4.3.1)$$

#### 4.4 顧客特性と規格別の契約数

RQ3“顧客特性（従業員数・業種・エリア別）ごとの主要4規格別で契約数が相対的に多い特性は何か”を明らかにするために，新規に契約した顧客企業の情報を分析する．本研究の対象データに限らず，BtoBの契約型コンサルティング企業に共通する顧客の特徴を表す項目を選ぶ必要があることから，1)顧客企業の従業員数，2)住所に基づいたエリアの重要度，3)産業分類の3つを顧客特性を表す指標として用いる．従業員数は顧客企業の従業員数が記録された列から，0～20，21～50，51～100，101～300，301～．の5つの区分に分けた．エリア重要度は顧客企業の所在地に対して共同研究先企業が独自に判断して区分分けした．産業分類は顧客企業の業種を日本標準産業分類の分類番号に基づいて記録されている．本研究では記録されている産業分類は分析実験に用いるには分類番号が過多であるため，複数の産業分類番号を以下の16個にまとめて用いる．産業分類番号のまとめ方は共同研究先の業務経験に基づき社員の方に判断してもらった．元々の分類番号との対応関係を付録1に示す．

- 01. 建築/土木
- 02. 食品
- 03. 製造業
- 04. 印刷
- 05. 水道/ガス/電気/通信
- 06. 運輸物流
- 07. 商社/メーカー/小売
- 08. 自動車

- 09. 金融/保険
- 10. 不動産/ビルメン
- 11.ITシステム/SES
- 12. 広告
- 13. 土業
- 14. 産廃処理
- 15. 医療
- 99. そのほか

3つの顧客特性それぞれについて契約した規格ごとにクロス集計を行い、結果に対してカイ二乗検定を行うことで有意な偏りがみられるか確かめる。クロス集計にはpythonのpandasライブラリで提供されているcrosstabメソッドを、カイ二乗検定はscipyライブラリで提供されているchi2.contingencyメソッドを用いる。カイ二乗検定の結果でp値が0.05を下回ったときは有意な偏りと見なし、集計の内訳を出力する。有意な偏りは顧客特性と契約規格のそれぞれの組み合わせにおいて契約が多い組合せと少ない組合せについて調べる。

#### 4.5 契約金額と顧客特性の関係

RQ4“契約金額に関係する顧客特性の間にどのような関係性が存在するのか”を明らかにするために新規に契約した顧客企業の情報を分析する。RQ3と同じく、本研究の共同研究先企業に限らず顧客特性を表すと考えられることから、1)顧客企業の従業員数、2)住所に基づいたエリアの重要度、3)産業分類の3つと、新規契約の累計金額を用いる。従業員数は1人当たりの影響を示すため、RQ4の分析では区分分けせず数値として扱う。産業分類はRQ3での前処理と同じく大分類化した16分類を扱う。数値データではないエリア区分と産業分類はダミー変数化して用いる。累計金額について、新規契約のデータ群で会社名が重複した16社を同社名それぞれで累計金額を合算し1レコードにして扱う。また、本分析で用いる項目のいずれかに欠損があるレコードは削除して分析に用いない。顧客特性を表す1)顧客企業の従業員数、2)住所に基づいたエリアの重要度、3)産業分類の3つを独立変数(説明変数)、新規契約の累計金額を従属変数(目的変数)として重回帰分析を実施する。

重回帰分析は変数とする項目のデータを行列・ベクトルとして差が小さくなるような係数を計算する。本実験では重回帰分析をpythonで提供されているsklearnライブラリを用いて実行する。重回帰分析の実行結果から得られた係数と切片

から式4.5.1に示すような式を立てる。係数の大小が各独立変数の顧客特性が従属変数の累計金額に与える影響の大小を表し、切片は独立変数で説明しきれなかった値となる。

$$\begin{aligned} \text{新規契約累計金額} &= \text{従業員数} \times \text{係数}_1 \\ &+ \text{業種} \times \text{係数}_2 \\ &+ \text{エリア別} \times \text{係数}_3 \\ &+ \text{切片} \end{aligned} \tag{4.5.1}$$

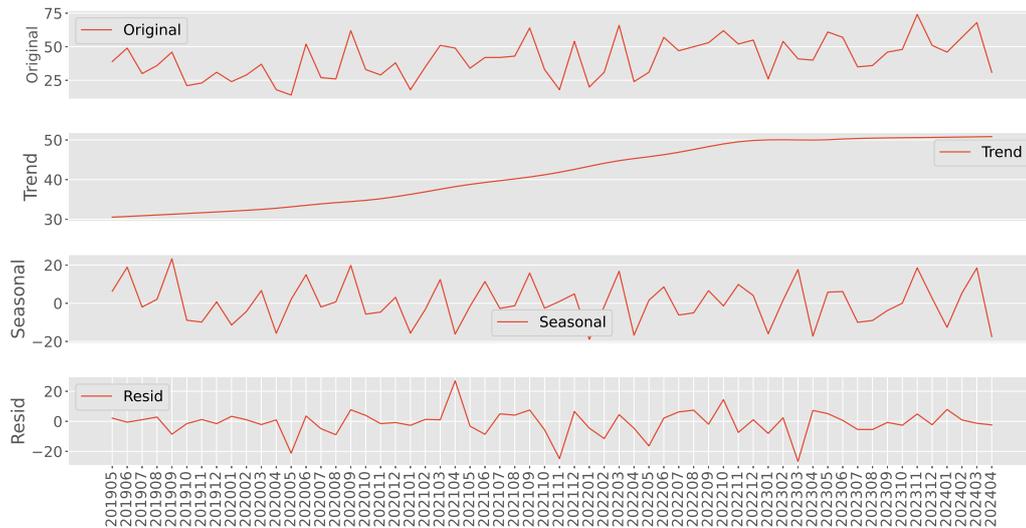


図1 STL分解の結果

## 5 結果と考察

### 5.1 STL分解

2019年5月から2024年4月までの5年間の新規契約数を年月ごとに集計し、STL分解を実行した。STL分解の結果を図1に示す。図は4つのグラフからなり、上から順に元データ、長期トレンド、季節性、残差を表す。すべてのグラフで縦軸は月あたりの新規契約の件数[件]、横軸は年月[yyyymm]を表す。以下3項に分けて長期トレンド、季節性、残差についてそれぞれの結果と考察を述べる。

#### 5.1.1 長期トレンド

図2に長期トレンドの成分を示す。分析対象の期間当初(2019年5月)には月あたりの新規契約件数が30.51件だったのに対し、3年7ヶ月後の2022年12月には49.87件まで増加しており、ビジネスが順調に成長しているといえる。一方で、2023年1月以降は横ばいで推移しており、分析対象の最後(2024年4月)に50.84件とほとんど増加していない。このことから2023年1月以降はビジネスの成長が停滞していると解釈できる。

新規契約件数の増加が近年停滞していることはデータ提供元の企業でも感じられており、長期トレンド成分を用いたビジネスの成長の推移が妥当であることが確かめられた。また、現状維持やさらなる成長のための経営的な判断が必要であることが明らかになった。具体的な判断・分析の例として、長期トレンド成分の増加が停滞した2022年と2023年の間で起こったことや前後で変化したことがあるか調査することが挙げられる。ビジネスの成長が停滞した原因を特定するこ

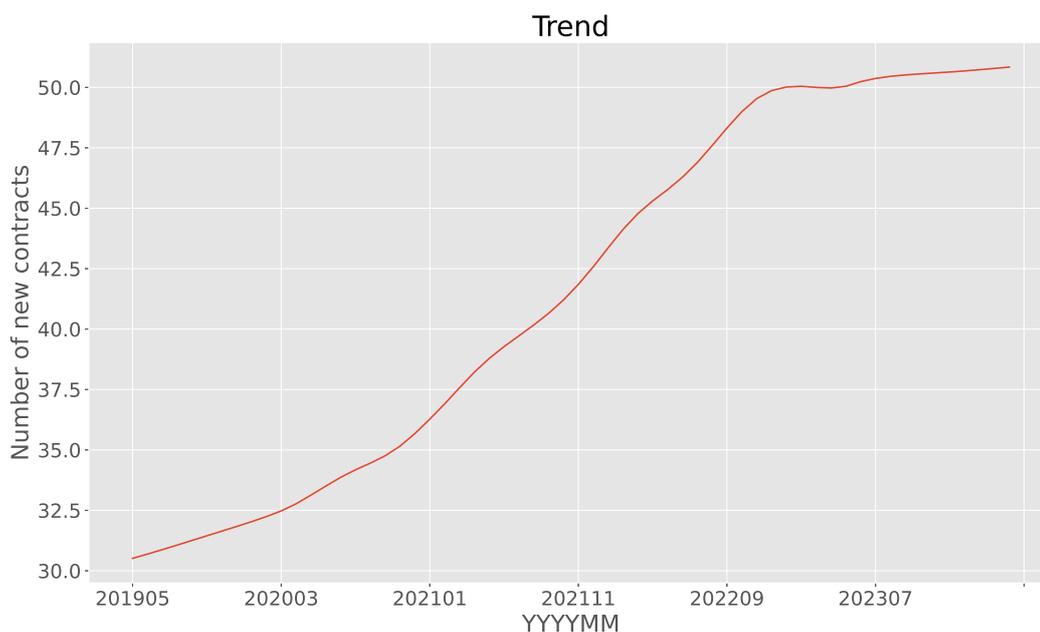


図2 長期トレンド

とで現状を打破するために効果的な改善策が明らかになり、新規契約数を再び増加させられる可能性がある。

### 5.1.2 季節性

図3に季節性の成分から求めた1月～12月の平均と標準偏差を示す。縦軸は新規契約数、横軸は月を表す。3月、6月、9月は+10件以上と新規契約数が多く、反対に1月、4月は-15件近く、新規契約数が少ない。

3月に多い理由として、顧客企業が4月からコンサルティングサービスを開始するために3月中に契約を結ぶことが考えられる。6月に多い理由として、データ提供元の企業は1月から年度が始まり、6月は上半期の末となり社員の業務にも一層力が入ることが影響していることが考えられる。9月に多い理由として、3月に多い理由と同様に4月に年度が始まる顧客企業が10月以降の下半期からコンサルティングサービスを開始するために9月に契約を結ぶことが考えられる。1月に少ない理由として、稼働日数が少なく一般的に企業のパフォーマンスが下がることが考えられる。4月に少ない理由として、3月に多い理由と反対に4月からコンサルティングサービスを開始するために契約を結ぶ顧客が多いことから契約が少ないと考えられる。

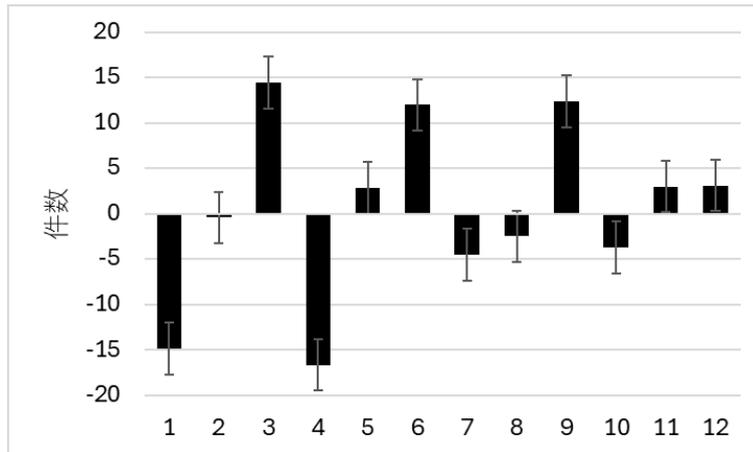


図3月ごとの新規契約件数

季節性成分の特徴が恒常的なものか、変化して現れたものかを明らかにするために、各月の季節性成分の経年変化を分析する。図4(a)に安定してマイナスの月、図4(b)に減少傾向の月、図4(c)に増加傾向の月、図4(d)に変化が小さい月を示す。各図の縦軸は新規契約数、横軸は年を表す。1月と4月（図4(a)）は安定してマイナスの値であり、季節の特徴として新規契約数が少ないと考えられる。6月と9月（図4(b)）は年ごとに値が大きく減少しているが、長期トレンドが除外された季節変動であることから、6月と9月に実施されている施策や外部要因のマイナス効果が年々増加しているか、プラス効果が減少していると考えられる。同様に3月と11月（図4(c)）の年ごとの減少は、3月と11月に実施されている施策や外部要因の効果であると考えられる。他の月（図4(d)）については上記の月に比べて大きな特徴は見られなかった。

先述の季節性成分の考察と合わせて、1月は提供元企業、4月は顧客企業の年度が始まる月に関係するため、安定して少ない傾向は妥当であると考えられる。一方で3月や6月9月は提供元および顧客企業の年度の節目となる時期に関わると考えたが、季節性成分は大きく変動しており、今後の季節性成分の特徴がどのように変化するかを追って分析し周期的特徴を把握することで適切な月間の施策を実施するなどの対応が可能になると考えられる。さらに、月別の季節性成分の平均値では大きな値を示さなかった11月の季節性成分が大きく増加しており、増加の原因を探り、得られた知見を他の月の契約数の増加に応用するなどの活用ができると考えられる。

### 5.1.3 残差

図5に残差成分を示す。残差は長期トレンドと季節性で説明されなかった余りであり、特定の月にのみ生じた変動である可能性がある。残差の絶対値が大きい

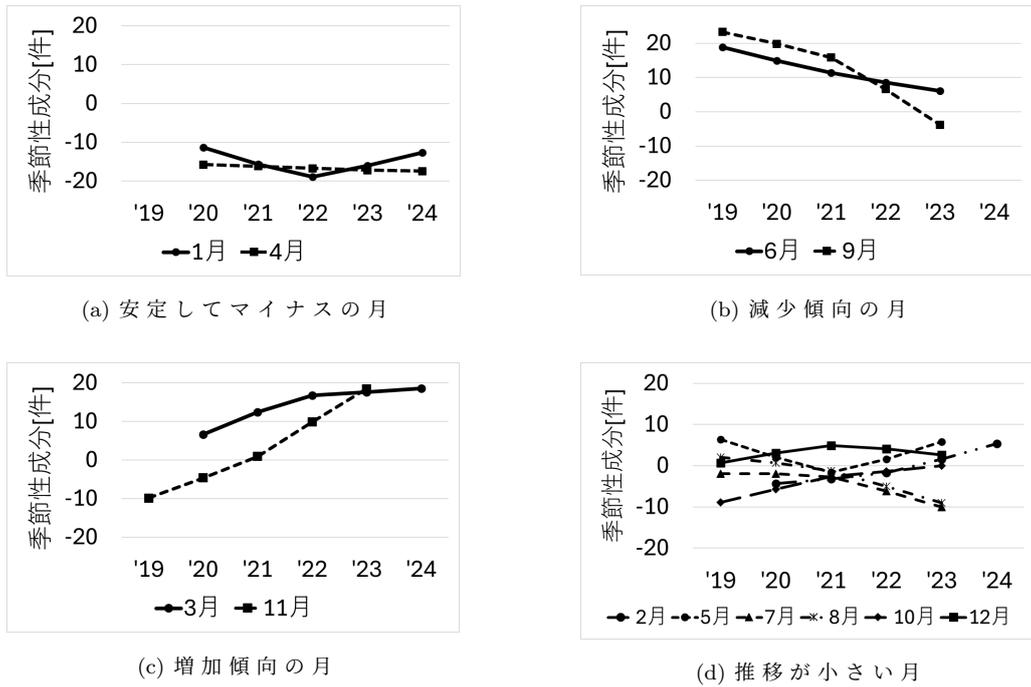


図4 特徴別季節成分の推移

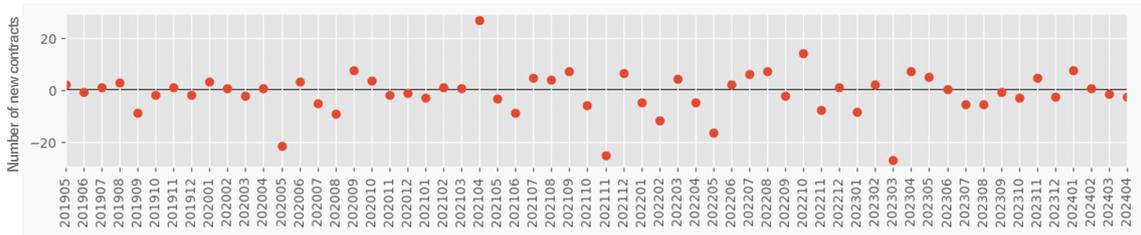


図5 残差成分

月に着目すると、2021年4月(26.95)、2020年5月(-21.19)、2021年11月(-24.81)、2023年3月(-26.66)でいずれも20を超えている。これらの月には新規契約数を増やすための特別な施策や突発的な外的影響などが発生した可能性があり、残差成分が大きいタイミングで起こった出来事などを探ることで新規契約数への影響を把握できる。また、今後似たような突発的な出来事が起こった時の影響を予測し、適切に対応するために有益な知見となる。

#### 5.1.4 RQ1への回答

STL分解の結果を基に、RQ1（月単位の新規契約数の要素（長期トレンド・季節性・残差）はどのような内訳・推移となっているのか）に回答する。

### RQ1 への回答

月単位の新規契約数の要素の内訳は、長期トレンド成分が約30~50件、季節性成分は影響が大きい月で+10件、-15件、他の月は絶対値5件以内、残差成分はほとんどが絶対値で5件以内、4つの月で20件を超え、長期トレンドが内訳を大きく占めることがわかった。新規契約数の要素の推移として、長期トレンドは分析対象期間当初の約30件から2022年12月の約50件まで上昇し、2023年以降は約50件で横ばいの推移となった。季節性成分は、1月と4月は安定して約-15件を示し、3月は大きい上昇傾向とともに平均で+10件以上を示した。6月と9月は大きい減少傾向を示しながらも平均では3月と並んで+10件以上となった。また、11月は大きい上昇傾向を示したものの平均値では絶対値で5件以下となった。

## 5.2 規格ごとの成約率

規格ごとの成約率を表1に示す。問い合わせ数の差に着目すると、もっとも多いP規格(2,417件)ともっとも少ないE規格(542件)で1,875件と大きな差がある。同様に成約率についても、もっとも高いI規格(36.7%)ともっとも低いE規格(29.3%)で7.4%の差がある。分析対象の企業における契約1件は最低でも48万円であることから、仮に問い合わせ数・成約率がともに最少の規格Eが同じ問い合わせ数でも最高成約率の36.7%まで上がるだけで約40件増加し、単純計算で1920万円以上の累計金額の上昇が発生するため無視できない。

規格ごとの問い合わせ数と成約率に着目すると、規格Pは問い合わせ数が2,417件と最大である一方で、成約率は29.4%と低い。規格の特性を考慮する必要があるが、成約率が他と比べて低く改善の余地があると考えられ、また、問い合わせ数が多いことから成約率を改善したときの効果が大きい。そのため、他の規格と比べて優先的に改善を行うことで高い効果に繋がると考えられる。規格Qは問い合わせ数が1,377件と少なく問い合わせ数が最大の規格Pと1,040件の差がある一方で、成約率は33.3%と高い。規格の特性を考慮する必要があるが、問い合わせ数に改善の余地があると考えられ、また、成約率が高いことから問い合わせ数の改善による効果が大きい。規格Pは成約率の改善が高い効果に繋がると考えられるのに対して規格Qは問い合わせ数の改善が高い効果に繋がると考えられる。

規格ごとの成約率の分析結果を基にRQ2(主要4規格別の問い合わせ件数とその成約率にどのような差が存在するか)に回答する。

### RQ2 への回答

問い合わせ数は最大で1,875件、成約率は最大で7.4%の差があった。契約1件当たりの金額を加味するとかなり大きな差であり、結果から規格Pおよび規格Qの改善による効果が高いことがわかった。

表1 規格ごとの問い合わせ数と成約率

	問い合わせ数	成約数	成約率
P規格	2,417	711	29.4%
Q規格	1,377	458	33.3%
E規格	542	159	29.3%
I規格	2,242	822	36.7%

### 5.3 顧客特性と規格別の契約数

新規契約を結んだ顧客の特性ごとの規格別で契約数が多い・少ないに偏りがある特性を分析した結果と結果に基づく考察を述べる。以降3項でそれぞれ従業員数を区分化した項目、顧客企業の所在地をエリア分けした項目、産業分類を大まかに分類しなおした項目のそれぞれに関する分析結果と考察を述べる。各規格のクロス集計の結果に対してカイ二乗検定のp値が0.05を下回った場合のみ有意な偏りがあるとみなし、結果を示したうえで考察を述べる。クロス集計結果は、割合と契約数の両方を示し、偏りがいない場合の契約数を同時に示し、比較する。

#### 5.3.1 従業員数区分

従業員数区分と規格別でクロス集計を行い、カイ二乗検定を行った結果、“その他”でのみ有意( $p=0.009$ )な分布の偏りが見られた。表2に“その他”のクロス集計表を示す。また、付録3に有意差が見られなかった主要4規格のクロス集計表を示す。表の“契約”は契約されていること、“未”は契約されていないことを表す。“契約”と“未”のそれぞれのセルについて左の値は契約数[件]を示し、右の値は“契約”と“未”の2項目での割合を示す。従業員数区分において主要な4つの規格すべてで契約数に有意な偏りは見られなかった。

従業員数区分と規格毎のクロス集計では、主要4規格とのクロス集計結果に有意な偏りがみられなかったことから、顧客企業の従業員数は新規契約の規格の偏りに影響しないと考えられる。契約規格の種類には後述の顧客企業の産業分類が大きく関わることや、規格取得の理由は信頼獲得・入札条件であることから、顧客企業自身の従業員数は契約規格の種類に関係しないことが有意な偏りが見られなかった理由として考えられる。

#### 5.3.2 エリア別

顧客企業の所在地をエリア分けした項目と各規格をクロス集計した結果、規格Q( $p < 0.000$ )、規格E( $p = 0.008$ )、規格I( $p < 0.000$ )で有意な偏りが見られた。表3に有意な偏りの見られた3規格のクロス集計表を示す。“規格Q”と“規格E”におい

表2 従業員数区分とその他のクロス集計

従業員数	その他			
	契約		未	
0-20	13	2.5%	499	97.5%
21-50	16	5.9%	257	94.1%
51-100	13	9.2%	129	90.9%
101-300	7	6.1%	107	93.9%
301-	7	16.7%	35	83.3%
合計	56	5.2%	1,027	94.8%

表3 エリア区分と規格ごとのクロス集計

エリア	規格Q				規格E				規格I			
	契約		未		契約		未		契約		未	
遠方	63	28.1%	161	71.9%	24	10.7%	200	89.3%	61	27.2%	163	72.8%
最重要	68	10.1%	604	89.9%	31	4.6%	641	95.4%	315	46.9%	357	53.1%
重要	47	28.1%	120	71.9%	24	14.4%	143	85.6%	57	34.1%	110	65.9%
その他	1	25.0%	3	75.0%	0	0.0%	4	100.0%	1	25.0%	3	75.0%
未記入	2	13.3%	13	86.7%	2	13.3%	13	86.7%	9	60.0%	6	40.0%
合計	181	16.7%	901	83.3%	81	7.5%	1001	92.5%	443	40.9%	639	59.1%

ではエリア区分が「遠方エリア」「重要エリア」である顧客は契約が多く、「最重要エリア」である顧客は契約が少ない。反対に“規格I”においてはエリア区分が「最重要エリア」である顧客は契約が多く、「遠方エリア」「重要エリア」である顧客は契約が少ない。

エリア区分と各規格のクロス集計結果で3つの規格で有意な偏りが見られたことから、顧客企業の所在地が契約する規格の種類に関係することから、営業や広告などの戦略を地域によって適切に変えることでより高い効果を得られる戦略の策定に有益な知見となる。また、規格Qと規格Eに同じ傾向が見られ、規格Iには反対の傾向が見られたことから、同傾向の規格を顧客に薦めるなど、顧客特性に応じた戦略策定が可能になると考えられる。

### 5.3.3 産業分類

顧客企業の産業分類と各規格をクロス集計した結果、規格Q( $p < 0.000$ )、規格E( $p = 0.000$ )、規格I( $p < 0.000$ )、規格P( $p < 0.000$ )、その他( $p < 0.004$ )で有意な偏りが見られた。紙面に収まらないため、ここでは割合のみを示し、件数を含むクロス集計表は付録2に掲載する。表4に有意な偏りの見られたすべての規格のクロス集計結果の割合を示す。

“規格Q”と“規格E”において、産業分類が「建築/土木」、「製造業」に分類される企業の契約が多く、「ITシステム/SES」、「土業」、「そのほか」に分類される企業の

契約が少ない。“規格I”の契約において、産業分類が「ITシステム/SES」、「そのほか」に分類される企業の契約が多く、「建築/土木」、「製造業」に分類される企業の契約が少ない。“規格P”の契約において、産業分類が「運輸物流」、「土業」、「そのほか」に分類される企業の契約が多く、「建築/土木」、「製造業」「データ無し」に分類される企業の契約が少ない。

5.3.2説で述べたエリア区分でのクロス集計結果と同様に、規格Qと規格Eにおける顧客特性と契約数の偏りの傾向が同じで、規格Iは反対の傾向を示しており、企業の所在地と業種に関係があり、契約する規格の種類に影響していることが明らかになった。さらに、エリア区分と規格Pのクロス集計に有意な偏りがみられなかった結果と合わせて、規格Pの契約は産業分類が運輸物流と土業の顧客が多いが、運輸物流と土業の地域性が低いことが原因でエリア区分によるクロス集計結果では有意な偏りがなかったと考察できる。すべての規格について有意な偏りが見られ、p値が限りなく小さいことから顧客企業の業種は契約する規格の種類に大きく影響していると考えられる。

エリア区分と同様に、さらに規格Pの結果を踏まえて、顧客企業の特性を把握することで、より契約が結ばれそうな規格を薦めるなど効果的な営業活動や戦略策定に有益な知見となると考えられる。また、規格Pについて産業分類が「運輸物流」と「土業」に分類される企業の契約が多いことと、「運輸物流」と「土業」の特性上地域性が低いことが原因でエリア区分と規格別のクロス集計で有意な偏りが見られなかったと考えられる。

#### 5.3.4 RQ3への回答

顧客特性と規格別の契約数の分析結果を基に、RQ3（顧客特性（従業員数・業種・エリア別）ごとの主要4規格別で契約数が相対的に多い特性は何か）に回答する。

##### RQ3への回答

“規格Q”と“規格E”において契約数が相対的に多い顧客特性はエリア区分が「遠方エリア」「重要エリア」、産業分類が「建築/土木」、「製造業」であった。“規格I”において契約数が相対的に多い顧客特性はエリア区分が「最重要エリア」、産業分類が「ITシステム/SES」、「そのほか」であった。“規格P”は産業分類でのみ有意な偏りが見られ、契約数が相対的に多い顧客特性は「運輸物流」と「土業」であった。従業員数区分は主要4規格のいずれにも有意な偏りが見られなかった。

#### 5.4 顧客特性と契約金額

表5に契約金額を目的変数、顧客特性を表す変数を説明変数として回帰分析を行った際の係数と切片の値を示す。

表4 産業分類と規格ごとのクロス集計

産業分類	規格Q		規格E		規格I		規格P	
	契約	未	契約	未	契約	未	契約	未
1	80.6%	19.4%	47.2%	52.8%	5.6%	94.4%	11.1%	88.9%
2	16.7%	83.3%	16.7%	83.3%	16.7%	83.3%	66.7%	33.3%
3	88.6%	11.4%	22.7%	77.3%	4.6%	95.5%	4.6%	95.5%
4	18.8%	81.3%	6.3%	93.8%	12.5%	87.5%	68.8%	31.3%
5	57.1%	42.9%	42.9%	57.1%	14.3%	85.7%	28.6%	71.4%
6	4.8%	95.2%	0.0%	100.0%	19.1%	81.0%	76.2%	23.8%
7	20.0%	80.0%	16.0%	84.0%	16.0%	84.0%	64.0%	36.0%
8	0.0%	100.0%	0.0%	100.0%	33.3%	66.7%	66.7%	33.3%
9	55.6%	44.4%	0.0%	100.0%	33.3%	66.7%	11.1%	88.9%
10	16.7%	83.3%	16.7%	83.3%	0.0%	100.0%	83.3%	16.7%
11	3.6%	96.4%	0.0%	100.0%	60.6%	39.4%	43.4%	56.6%
12	0.0%	100.0%	0.0%	100.0%	22.2%	77.8%	66.7%	33.3%
13	9.8%	90.2%	2.8%	97.2%	42.0%	58.0%	51.8%	48.3%
14	0.0%	100.0%	50.0%	50.0%	12.5%	87.5%	37.5%	62.5%
15	16.7%	83.3%	0.0%	100.0%	66.7%	33.3%	16.7%	83.3%
99	10.7%	89.3%	3.1%	96.9%	44.7%	55.4%	49.1%	50.9%
データ無し	15.6%	84.4%	9.6%	90.4%	39.8%	60.2%	36.8%	63.2%
合計	16.7%	83.3%	7.6%	92.4%	40.9%	59.1%	42.2%	57.8%

「01. 建築/土木」「05. 水道/ガス/電気/通信」「09. 金融/保険」の係数は100,000を超えており、係数が他の係数と比べて大きいと言える。特に「09. 金融/保険」の係数は約230,000と突出して大きい。反対に「04. 印刷」「12. 広告」「15. 医療」の係数は-100,000前後とマイナスの係数で他の係数と比べて大きいと言える。係数の絶対値が大きい変数は累計金額に大きく影響していると解釈でき、係数の正負の符号が累計金額への影響がプラスかマイナスのいずれかを示している。一方で、「◆遠方エリア」の係数は3,052.5、「☆重要エリア」の係数は-2,461.6と係数の絶対値が他の係数と比べて小さく、特に「従業員数(1人あたり)」の係数は119.0と非常に小さかった。係数の絶対値が小さい変数は累計金額への影響が小さいと解釈できる。

切片の値は約540,000となり支配的に大きいと言える。切片は、独立変数では説明できずに残った値を表していることから、契約の内容にかかわらず切片の値程度の金額で契約が結ばれると解釈できる。データ提供元の企業では480,000円が新規契約の累計金額の基準となると感じられており、60,000円分のずれはあるが切片に対する解釈と値の大きさの妥当性が確かめられた。

重回帰分析の結果は、顧客特性が新規契約の累計金額にどれほど影響するか、プラスとマイナスのどちらに影響するかを明らかにした。係数の絶対値が大きかった顧客特性に該当する企業は複数の規格の契約を結ぶことが多いなど、係数

表5重回帰分析結果の変数と係数の対応表

説明変数	係数
従業員数(1人あたり)	119.0
01. 建築/土木	102,709.1
02. 食品	46,555.3
03. 製造業	52,322.8
04. 印刷	-143,783.4
05. 水道/ガス/電気/通信	134,121.6
06. 運輸物流	-16,641.3
07. 商社/メーカー/小売	-32,180.2
08. 自動車	-37,563.0
09. 金融/保険	230,934.5
10. 不動産/ビルメン	-66,110.8
11. ITシステム/SES	40,133.7
12. 広告	-95,629.1
13. 土業	-32,094.3
14. 産廃処理	-27,599.7
15. 医療	-113,368.8
99. そのほか	-41,806.5
◆ 遠方エリア	3,052.5
★ 最重要エリア	25,150.5
☆ 重要エリア	-2,461.6
その他	-25,741.4
切片	544,242.5

の絶対値が大きくなった原因を明らかにすることで、契約の金額交渉や他規格の契約を推奨するなど有益な知見となる。見込み顧客の顧客特性と似た特性を持つ顧客の契約金額から利益の予測を立てられる可能性も考えられる。

重回帰分析の結果を基に、RQ4（契約金額に関係する顧客特性の間にどのような関係性が存在するのか）に回答する。

#### RQ4への回答

契約金額と顧客特性の間関係性は総じて産業分類と深い関係があると言える。係数の絶対値が100,000を超える変数は産業分類の6つであり、プラスとマイナス3つずつであった。従業員数は契約金額に対してほとんど影響がなく、エリア区分も4つの内2つの係数が5,000に満たず影響が小さいと言える。

## 6 おわりに

本研究はBtoBの契約型コンサルティング業務を行っている企業の利益向上に有益な知見を得ることを目的に、顧客情報や販売履歴のデータを分析した。分析にあたって、1) 新規契約数の要素の内訳と推移、2) 規格ごとの問い合わせ件数と成約率の差、3) 顧客特性ごとの規格別で契約数の相対的な比較、4) 契約金額に係る顧客特性の関係性以上の4つの問いを立て、分析を行った。分析では共同研究を実施しているBtoBの契約型コンサルティングを業務とする企業から提供いただいた、契約についての顧客情報や販売履歴のデータを用いた。

RQ1に対してSTL分解を用いて新規契約数を長期トレンド・季節性・残差の3つの成分に分けた結果、上昇傾向が停滞に切り替わった時期や、月ごとに契約数に差があること、月ごとの傾向が変化していることが明らかになった。RQ2に対して新規契約に関する問い合わせ数と成約率を分析し、問い合わせ数は多いが成約率が低い規格Pは、他の規格よりも改善による効果が高い可能性が明らかになった。RQ3に対して顧客特性ごとに規格別にクロス集計を行ったうえでカイ二乗検定を実施し、p値が0.05を下回った結果を有意な偏りとした結果、従業員数区分の主要4規格に有意な偏りは見られず、エリア区分と産業分類のほとんどの規格に有意な偏りがみられ、エリア区分と産業分類による契約規格の偏りが明らかになり、顧客企業の所在地と産業分類の関係が示唆された。RQ4では新規契約の累計金額を従属変数とし、従業員数・エリア区分・産業分類を独立変数として重回帰分析を行った結果、各変数の累計金額に対して与える影響の大きさと切片の値が明らかになった。

本研究の今後の課題として、顧客情報や販売履歴のデータ収集方法の改善や、本研究で実施したような分析をリアルタイムに自動で実施・結果を表示するようなシステム化する手法を提案することが挙げられる。本研究で扱ったような顧客情報や売り上げに関するデータは通常の業務を進める中で業務システム上に溜まっていく。分析者が本研究の知見を活かしてデータを前処理、分析を実施、結果を確認とプロセスを踏んでいるとどうしても分析に使用したデータと実際に進んでいる業務に時間の差が生じてしまう。そこで、リアルタイムに自動で分析を実施・結果を表示するようなシステムを構築できれば、より最新のデータを使用したリアルタイムの知見を得ることができる。分析とのタイムラグが減ると、ビジネスの成長トレンドや顧客特性による契約の偏りの変化が明らかになり、分析結果を踏まえた業務の改善効果の向上や改善を可視化が可能になることで企業の利益向上に繋がると考えられる。

その他、本研究の今後の発展として、重回帰分析で利用する独立変数の選定を見直すことによる精度向上や、本研究の知見に基づいてCLV分析の手法の提案が挙げられる。重回帰分析の精度に影響する要素として独立変数の数や組み合

わせが代表的である。本研究で独立変数に使用した従業員数1人当たりの係数は他の変数に比べて著しく小さく、累計金額への影響が小さいこと、または産業分類・エリア区分との組み合わせの相性が良くないことが考えられる。改善の方法として従業員数の変数を削除する他に、提供データの顧客情報に関する項目に「大手」か否かを記録する項目を用いる方法も考えられる。他の変数がダミー変数化して用いているのに対して従業員数のみ数値データとして用いていることや、RQ3の結果から示されたようにエリア区分と産業分類の関係が深いことから、重回帰分析の精度向上のために独立変数の中でも特に従業員数の変数を見直すことで精度が上がると考えられる。さらに、重回帰分析の結果から得られる係数と切片に、未契約の見込み顧客の顧客特性を変数に代入することで契約の累計金額を予測することができ、企業にとっての顧客の価値を把握することができる。顧客特性から顧客の価値を把握することはCLV分析と同じ考え方であり、データ提供元企業のようなBtoBかつ契約型コンサルティング企業に適切なCLV分析の手法を明らかにすることに繋がる。また、本研究の分析で明らかになった利益に関係する顧客情報や顧客情報同士の関係の理解に基づいてデータセットから売り上げに関係しないまたは影響が小さい項目を除去することで、重回帰分析の独立変数の候補や組み合わせを機械学習によって選定する機械学習や、見込み顧客の情報から契約金額や企業へもたらす利益の予測モデルを構築するなど、より詳細で妥当性の高い分析が可能になると考えられる。重回帰分析の精度向上やCLV分析の適切な手法を明らかにすることで顧客の価値を明らかにし、より顧客それぞれに対応した経営戦略や経営的な判断、顧客ロイヤルティの向上を目指す、さらなる利益向上が期待できる。

## 謝辞

本研究を行う上で、データの提供および分析結果に関する議論にご協力いただいた共同研究先企業の皆様に心から感謝を申し上げます。また、研究や論文の執筆にあたって常日頃からご指導いただきました上野准教授に厚く御礼申し上げます。査読教員である岩田准教授には、大変貴重な意見をいただきましたことに感謝を申し上げます。

## 参考文献

- [1] R. B. Cleveland, W. S. Cleveland, J.E.McRae and I. Terpenning: “STL: A Seasonal-Trend Decomposition Procedure Based on Loess,” *Journal of Official Statistics*, Vol.6, No.1, pp.3–73 (1990).
- [2] 小野 讓司:“契約型サービスにおける顧客関係”, *マーケティングジャーナル*, Vol.28, No.2, pp.15–27 (2008).
- [3] 蓮本 恭輔, 雲居 玄道, 後藤 正幸:“非負値行列因子分解を用いたプラットフォームビジネスにおける顧客生涯価値予測”, *情報処理学会論文誌*, Vol.60, No.7, pp.1283–1293 (2019).
- [4] J. Klier, M. Klier, F. Probst, L. Thiel: “Customer Lifetime Network Value,” in *Proceedings of the 35th International Conference on Information System (ICIS)* (2014).
- [5] D. A. Kandeil, A. A. Saad, and S. M. Youssef: “A two-phase clustering analysis for B2B customer segmentation,” in *Proceedings of the 2014 International Conference on Intelligent Networking and Collaborative Systems (INCoS 2014)*, pp. 221–228 (2014).
- [6] P. P. Pramono, I. Surjandari, and E. Laoh: “Estimating Customer Segmentation based on Customer Lifetime Value Using Two-Stage Clustering Method,” in *2019 16th International Conference on Service Systems and Service Management (ICSSSM)* (2019).
- [7] 高島 大輔, 高橋 真吾, 大野 高裕:“エージェントベースモデリングによる優良顧客の特徴分析”, *経営情報学会誌*, Vol.15, No.1, pp.1–13 (2006).
- [8] 伊藤 史世, 張 志穎, 雲居 玄道, 後藤 正幸:“Knowledge Graph Attention Networkに基づく購買行動の多様性を考慮した顧客分析手法の提案”, *情報処理学会論文誌*, Vol.63, No.1, pp.205–217 (2022).
- [9] 福本 信吾, 毛利 研, 深澤 信也, 伊藤 悠希, 池田 千枝, 菅野 敬明:“過去企画へのオーバーフィッティングを回避した成約率予測モデルの構築”, *情報科学技術フォーラム (FIT)*, Vol.19 (2020).
- [10] 阿部 誠:“RFMデータを用いた顧客生涯価値の算出—既存顧客の維持介入と新規顧客の獲得—”, *マーケティングジャーナル*, Vol.34, No.1 (2014).

## 付録

表6 付録1 産業分類の対応表

統合後	統合前
01. 建築/土木	05 金属鉱業, 07 原油・天然ガス鉱業, 09 総合工事業, 10 職別工事業, 11 設備工事業
02. 食品	12 食料品製造業, 50 飲食料品卸売業, 56 飲食料品小売業, 60 一般飲食店, 61 その他の飲食店
03. 製造業	13 飲料・たばこ・飼料製造業, 14 繊維工業, 15 衣服・その他の繊維製品製造業, 16 木材・木製品製造業, 17 家具・装備品製造業, 18 パルプ・紙・紙加工品製造業, 20 化学工業, 21 石油製品・石炭製品製造業, 22 プラスチック製品製造業, 23 ゴム製品製造業, 24 なめし革・同製品・毛皮製造業, 25 窯業・土石製品製造業, 26 鉄鋼業, 27 非鉄金属製造業, 28 金属製品製造業, 29 一般機械器具製造業, 30 電気機械器具製造業, 31 輸送用機械器具製造業, 32 精密機械器具製造業, 34 その他の製造業
04. 印刷	19 出版・印刷・同関連産業
05. 水道/ガス/電気/通信	35 電気業, 47 電気通信業
06. 運輸物流	46 郵便業
07. 商社/メーカー/小売	48 各種商品卸売業, 49 繊維・衣服等卸売業, 53 その他の卸売業, 54 各種商品小売業, 55 織物・衣服・身の回り品小売業, 58 家具・じゅう器・家庭用機械器具小売業, 59 その他の小売業
08. 自動車	57 自動車・自転車小売業, 77 自動車整備業
09. 金融/保険	63 中小企業等金融業, 67 補助的金融業, 金融附帯業, 69 保険業
10. 不動産/ビルメン	70 不動産取引業, 71 不動産賃貸業・管理業, 73 駐車場業
11. ITシステム/SES	82 情報サービス・調査業
12. 広告	83 広告業
13. 土業	86 その他の事業サービス業
14. 産廃処理	87 廃棄物処理業
15. 医療	88 医療業, 89 保健衛生, 90 社会保険, 社会福祉
99. そのほか	72 洗濯・理容・浴場業, 74 その他の生活関連サービス業, 76 娯楽業, 78 機械・家具等修理業, 79 物品賃貸業, 80 映画・ビデオ制作業, 81 放送業, 84 専門サービス業, 85 協同組合, 91 教育, 92 学術研究機関, 94 政治・経済・文化団体, 95 その他のサービス業, 97 国家公務, 98 地方公務, 99 分類不能の産業

表7付録2.1 産業分類と規格Q, 規格Eのクロス集計

産業分類	規格Q				規格E			
	契約		未		契約		未	
01. 建築/土木	29	80.6%	7	19.4%	17	47.2%	19	52.8%
02. 食品	1	16.7%	5	83.3%	1	16.7%	5	83.3%
03. 製造業	39	88.6%	5	11.4%	10	22.7%	34	77.3%
04. 印刷	3	18.8%	13	81.3%	1	6.3%	15	93.8%
05. 水道/ガス/電気/通信	4	57.1%	3	42.9%	3	42.9%	4	57.1%
06. 運輸物流	1	4.8%	20	95.2%	0	0.0%	21	100.0%
07. 商社/メーカー/小売	5	20.0%	20	80.0%	4	16.0%	21	84.0%
08. 自動車	0	0.0%	3	100.0%	0	0.0%	3	100.0%
09. 金融/保険	5	55.6%	4	44.4%	0	0.0%	9	100.0%
10. 不動産/ビルメン	1	16.7%	5	83.3%	1	16.7%	5	83.3%
11.ITシステム/SES	9	3.6%	242	96.4%	0	0.0%	251	100.0%
12. 広告	0	0.0%	9	100.0%	0	0.0%	9	100.0%
13. 土業	14	9.8%	129	90.2%	4	2.8%	139	97.2%
14. 産廃処理	0	0.0%	8	100.0%	4	50.0%	4	50.0%
15. 医療	1	16.7%	5	83.3%	0	0.0%	6	100.0%
99. そのほか	17	10.7%	142	89.3%	5	3.1%	154	96.9%
データ無し	52	15.6%	282	84.4%	32	9.6%	302	90.4%
合計	181	16.7%	902	83.3%	82	7.6%	1001	92.4%

表8 付録2.2 産業分類と規格I, 規格Pのクロス集計

産業分類	規格I				規格P			
	契約		未		契約		未	
01. 建築/土木	2	5.6%	34	94.4%	4	11.1%	32	88.9%
02. 食品	1	16.7%	5	83.3%	4	66.7%	2	33.3%
03. 製造業	2	4.6%	42	95.5%	2	4.6%	42	95.5%
04. 印刷	2	12.5%	14	87.5%	11	68.8%	5	31.3%
05. 水道/ガス/電気/通信	1	14.3%	6	85.7%	2	28.6%	5	71.4%
06. 運輸物流	4	19.1%	17	81.0%	16	76.2%	5	23.8%
07. 商社/メーカー/小売	4	16.0%	21	84.0%	16	64.0%	9	36.0%
08. 自動車	1	33.3%	2	66.7%	2	66.7%	1	33.3%
09. 金融/保険	3	33.3%	6	66.7%	1	11.1%	8	88.9%
10. 不動産/ビルメン	0	0.0%	6	100.0%	5	83.3%	1	16.7%
11. ITシステム/SES	152	60.6%	99	39.4%	109	43.4%	142	56.6%
12. 広告	2	22.2%	7	77.8%	6	66.7%	3	33.3%
13. 土業	60	42.0%	83	58.0%	74	51.8%	69	48.3%
14. 産廃処理	1	12.5%	7	87.5%	3	37.5%	5	62.5%
15. 医療	4	66.7%	2	33.3%	1	16.7%	5	83.3%
99. そのほか	71	44.7%	88	55.4%	78	49.1%	81	50.1%
データ無し	133	39.8%	201	60.2%	123	36.8%	211	63.2%
合計	443	40.9%	640	59.1%	457	42.2%	626	57.8%

表9 付録3.1 従業員数区分と規格Q, 規格Eのクロス集計

従業員数区分	規格Q				規格E			
	契約		未		契約		未	
0-20	81	15.8%	431	84.2%	32	6.2%	480	93.8%
21-50	11	9.7%	103	90.4%	7	6.1%	107	93.9%
51-100	53	19.4%	220	80.6%	22	8.1%	251	91.9%
101-300	10	23.8%	32	76.2%	7	16.7%	35	83.3%
301-	26	18.3%	116	81.7%	14	9.9%	128	90.1%
合計	181	16.7%	902	83.3%	82	7.6%	1001	92.4%

表 10 付録 3.2 従業員数区分と規格 I, 規格 P のクロス集計

従業員数区分	規格 I				規格 P			
	契約		未		契約		未	
0-20	238	46.5%	274	53.5%	201	39.3%	311	60.7%
21-50	49	43.0%	65	57.0%	54	47.4%	60	52.6%
51-100	88	32.2%	185	67.8%	129	47.3%	144	52.8%
101-300	12	28.6%	30	71.4%	16	38.1%	26	61.9%
301-	56	39.4%	86	60.6%	57	40.1%	85	59.9%
合計	443	40.9%	640	59.1%	457	42.2%	626	57.8%