



トライツブログセレクション (7)

営業マネージャ／営業企画のためのデータ分析

トライツコンサルティング株式会社

統計学の全体像を見てみよう

「統計学が最強の学問である」という本が出版されて早や6年。今では、滋賀大学や横浜市立大学など「データサイエンス」を学部名に冠する大学も増えてきています。また、最近話題の統計問題では、各省庁に統計の専門家が不足していることが一因だとも言われており、統計学／データサイエンスに関するスキルを持った人材の育成がより重要視されるようになってきています。このように、統計学は「最強」とまでは言えないにしてもとても重要な学問領域となっています。

一方で、B2Bの営業やマーケティングの分野でも、SFAやMAなどのシステムツールや売上データなどから、分析可能な様々なデータが生まれていますが、それらのデータが有効に分析されている企業はあまり多くありません。「統計学が役に立つ」と言われても、「なんだか難しそう」「普段の仕事での活用の仕方が分からない」という営業マネージャーや営業企画の方も多いのではないのでしょうか。

そこでまず、「統計学の全体像と、その中で営業マネージャー／営業企画が知っておきたい分野」をできるだけ平易に解説することからはじめてみましょう。



統計学を構成している6つの分野

統計学を一言で言うなら「データの中に含まれている特徴や規則性を見つけること」。この見つけたい特徴や規則性の種類によって、分野が分かれています。専門家の人には怒られるかも知れませんが、ここでは分かりやすさを最優先にして、統計学を以下の6つの分野に整理してみました。

1. 1つのデータから特徴を見つける「統計量」
2. 複数のデータから違いを見つける「比較」
3. 複数のデータの間の関係性を見つける「相関」
4. 複数のデータの関係性をもとに予測する「回帰」
5. 複数のデータの関係性をもとにグループ化する「分類」
6. 1～5の分析の確からしさを確かめる「検定」

それぞれがどういうものなのか、続けて見ていくことにしましょう。また、それぞれの分野の紹介の最後に具体的な分析手法を例示していますが、詳細は次回以降に説明しますので、今の時点では「そんなものがあるんだな」程度で軽く読み流してください。

6つの分野をざっくり確認してみよう

「統計量」は1つのデータに含まれている特徴を明らかにするものです。例えば、個人ごとの売上金額の平均値というのも統計量の1つです。統計量には他に、中央値、分散、標準偏差といったものがあります。この統計量についてはご存知の方が多いのではないのでしょうか。

「比較」は、複数のデータ同士の違いを見つける方法です。例えば、営業拠点ごとや、商品分類ごと、顧客区分ごとに受注単価を比較することで、受注単価の高い営業拠点や商品区分、顧客区分を見つけられるようになります。この「比較」で使う分析方法には、分散分析、多重比較といったものがあります。

「相関」は、複数のデータの間にある関係性を見つけるものです。ある企業では、顧客別に商品分類ごとの売上データを使い、商品分類ごとの相関関係を分析しました。相関関係が強い商品分類ほど一緒に売れやすいということなので、クロスセルの提案を考えるときに役立ちます。「相関」で使う分析方法は、相関係数、共分散分析というものがあります。

「回帰」は、一般には「もどに戻る」という意味がありますが、統計学では「今あるデータをもとに未知のデータを予測する」という意味で使います。営業として予測したいものと言えば、商談が受注するかどうかでしょう。ある企業では、「顧客の予算」「決裁者」「課題」などのキーワードをSFAのメモ欄に記入しているかどうか、つまり客先でそのような話をしているかどうかで、商談の受注/失注を予測するという分析をしています。「回帰」で使う分析方法は、回帰分析、重回帰分析、ロジスティック回帰分析、判別分析などがあります。

「分類」は、関係性の強さをもとに大量のデータをグループにまとめるものです。3つ目の「相関」の例のように一緒に売れやすい商品同士をグループ化したり、企業規模や業種といったデータから売れる商品が似ている企業グループ分類を作るなど、様々な活用方法があります。この「分類」で使う分析方法は、主成分分析に因子分析、クラスター分析、多次元尺度法、コレスポンデンス分析など多数あります。

最後の「検定」は、これまでに紹介した5つの分野の分析結果の確からしさを確かめるものです。「比較」の結果が本当に正しいと言えるのか、確かに「相関」関係があると言えるのか、ということを確認するのですが、この「検定」の方法はとてつもない数がありますので、ここでは紙幅の都合上例示を控えさせていただきます。

使われているのが「統計量どまり」で宝の持ち腐れな日本のB2B営業

ここまで6つの分野をご紹介してきましたが、実際に皆さんが仕事で使っているのは6つのうちいくつあったでしょうか。

これまで私が色々な営業組織で使われているExcelのデータを見てきたところ、その多くが売上金額を集計したり、平均値を求めたりといった「統計量」どまりでした。拠点別や四半期ごとの売上を「比較」しているとは言え、それは合計額や平均値といった「統計量」を見比べているだけで、「比較」分析には至っていないところが多いのです。

この「統計量」は1つのデータの特徴を調べるものなので、どうしても出てくる結果は「分かり切ったもの」になってしまいます。これに対し、「比較」「相関」「回帰」「分類」ではデータの中に隠れている特徴や規則性を探るものなので、思いもよらなかったアイデアが得られることがあり、そこにこそ統計分析の価値があります。

日本の B2B 企業の多くで SFA や MA、BI などのツールを上手く使いこなせていない理由の 1 つは、この「統計の使い方が、統計量どまり」というところにあると私は考えます。

多くのデータが目の前にあり、それを様々な切り口で分析できる機能を持ったシステムがあるのに、実際に使っている統計の手法が「統計量どまり」では、わかっていることを集計しただけの答えしか得られず、新たな気付きにつながりません。せっかくハイテク機能テンコ盛りの最新オープンレンジを買ったのに、冷ご飯を温めることにしか使っていないようなもので、「宝の持ち腐れ」になっている状態と言えます。

ちなみにここでの「宝」とは、分析ツールのことだけでなく、手間ひまかけて集めたデータのことでもあります。料理の仕方を知らないと、苦労して手に入れた産地直送の貴重な食材を活かせないのと同じです。

オープンレンジを使いこなし、貴重な食材を美味しくいただくためには、マニュアルやレシピ本を読んで勉強したり、食材のことを知る必要があります。それと同様に、データを活かしていくためには統計の知識を身につけることが不可欠なのです。

このシリーズで学ぶのは「統計量」「比較」「相関」「回帰」「分類」の 5 分野

そのため、次回以降の 3 回では「統計量」どまりにせず、「比較」「相関」「回帰」「分類」の 4 分野で実際の営業分析に使える手法をご紹介します。

ここで「あれ、検定はやらないの?」と思われるでしょうが、「検定」は対象外とします。と言うのも、営業で分析するときは「意味のある傾向」を見つけることが最重要となるので、その傾向がどこまで厳密に正確だと検定することにはあまり意味がありません。「検定」は分析手法が無数にあって学習するのに骨が折れるものですが、B2B 営業ではあまり使いどころがなく、コストパフォーマンスの低い分野なので、このシリーズでは割愛します。

今回は「統計量とその落とし穴」についてご紹介します

今回は、統計解析の基本中の基本である「統計量」についてご紹介します。「比較」「相関」「回帰」「分類」を理解するためには、この「統計量」についての理解は欠かせません。また、簡単だと思われがち「統計量」ですが、そこには多くの企業でやってしまっている根本的な間違いがありますので、その落とし穴についてもご説明します。

私は統計とは、データという貴重な食材を最大限に活かし、美味しくいただくための料理方法のようなものだと考えています。一緒に統計を学び、料理上手を目指しましょう。

統計の基本『統計量』とその落とし穴

まず、営業現場でよくあるシーンをケースとして解説していきます。あなたのこれまでの「常識」の間違いに気づくことになるかもしれません。



シーンその1：イベントのアンケートの集計で「平均」を使っていますか？

営業担当者：今回の展示会は、前回と比較するとお客様の反応が良かったですね。

課長：どうして？

営業担当者：いつもの来場者アンケートの結果をまとめたんですが、満足度平均が0.4ポイント高かったんですよ。

課長：おおそうか、それは良かったな。ごくろうさん。

このような会話はよくあることではないでしょうか。

展示会の来場者にアンケートを配布し、その満足度を「とても満足、やや満足、どちらでもない、やや不満足、とても不満足」という5つのどれかにチェックしてもらうというようなやり方は日常的に行われています。

そして、「とても満足」を5点とし、そこから4、3、2・・・「とても不満足」を1点として平均点を算出していることは少なくないと思いますが、実はこれが大問題なのです。

このデータは統計的には「順序尺度」というものになります。「やや満足」よりも「とても満足」の方の満足度が高いことは明確なのですが、それぞれの間隔は同じであるかどうかはわからないというのが特徴です。

仮に、展示会に来た人に一通りの説明が終わって、担当した営業担当者に「アンケートへのご記入お願いします」と言われ、引換に景品を貰えることもわかっていたとしたら・・・よほど悪い印象を持たない限り、「とても満足」「やや満足」のどちらかにチェックする人が多いのではないかと思います。

ただ、その状況にも関わらず、「やや不満足」「とても不満足」というチェックをつけているとしたら、そこには何か大きなマイナス要因があった可能性があります。従って、顧客心理という面から考えると、「やや満足」と「どちらでもない」の間隔よりも、「どちらでもない」と「やや不満足」という間隔の方がずっと広い可能性が高いのです。このようなシーンで無理やり「平均」を使うと、顧客のこの心理的な間隔の違いを全部ごちゃ混ぜにして計算することになり、顧客満足度の実態とは異なるものになってしまいます。そのため、順序尺度のデータでは「平均」は使えず、「最頻値」や「中央値」を見ることになります。

ここで「最頻値」と「中央値」についてですが、「最頻値」とは最も数が多いデータのことで、「中央値」とはデータを上から下まで並べたときの真ん中のデータのことで、先ほど例で挙げた展示会の満足度への回答件数が計11件あり、その内訳が「とても満足：1人、やや満足：4人、どちらでもない：3人、やや不満足：1人、とても不満足：2人」の場合、「最頻値」は一番回答数が多い「やや満足」になります。その一方で、上から数えても下から数えても6番目の「中央値」は「どちらでもない」です。「最頻値」は単純に最も数が多いものを見ているだけですが、「中央値」は全体のバランスを見ているので「とても不満足～やや不満足」と3名が回答していることを含んだものですので、より正確に実態を捉えていると言えます。

シーンその2：営業生産性を分析するときに安易に「平均」を使っていますか？

営業企画部長：自社の営業の生産性はどうなっているんだろう？打ち手を考えたいんだけど、組織ごとに違いはあるの？

メンバー：東日本営業部と西日本営業部で以前分析したときは、営業担当者一人当たりの売上高はほぼ同じでしたよ。

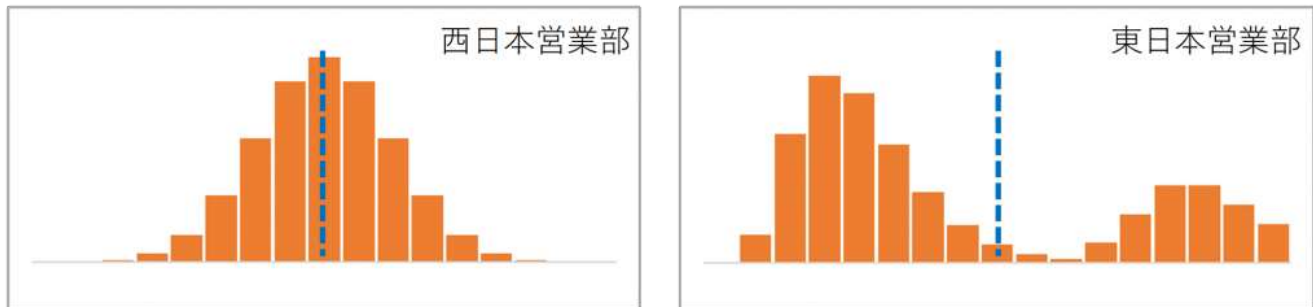
営業企画部長：それってそれぞれの部の売上高の平均ってこと？

メンバー：そうです。

営業企画部長：じゃあ、傾向はほぼ一緒か。まとめて実施できる施策を考えよう。

このように全国に営業組織を持っている会社であれば、拠点ごとに営業担当者個人の売上高や利益額を見比べるということをやっているところも多いでしょう。そのようなときに、手っ取り早いからと「平均」だけを見てしまっていると、大きな落とし穴が口を開いて待っているということがあります。

このシーンで出てきた「売上高」は、先ほどのアンケートとは異なり「比尺度」と言われるものです。数の大小という順序関係だけでなく、間隔にも意味があって四則演算することに何の問題もありませんので、統計量として「平均」がよくつかわれます。しかし、その「平均」がデータの実態を適切に表していない、ということがよくあるのです。



上の2つのグラフが、西日本営業部と東日本営業部の営業担当者一人当たり売上高の分布だとします。どちらも平均売上高はほぼ一緒ですが、分布のカタチが異なっています。西日本営業部は大勢のアベレージヒッターがいて、ずば抜けて沢山売る営業担当者もずば抜けて売れない営業担当者もそんなにいないチームです。一方で、東日本営業部は平均より高い売上を稼ぐ営業が西日本より多くいるものの、大半の営業担当者の売上は平均未満というチームです。

このようにデータの分布のカタチがかなり違うのに、「平均」の値はほぼ同じということが起こります。この2つの営業部を「最頻値」や「中央値」で見ると、西日本よりも東日本の方が全体的な売上が低い傾向にあるので、西日本営業部には「平均」付近に固まっている営業担当者に向けた施策が向いているでしょうし、東日本営業部には平均未満の営業メンバー向けを底上げするような施策が必要になってくるのが分かります。分布のカタチによっては、「平均」がデータの特徴を適切に表さないということがあるのです。

このような統計量の分析のミスマッチを防ぐために大事なのが、分析を始める前に分布のグラフを描いてみて、分布のカタチを確認するということだと私は考えます。分布の山がグラフの真ん中にあるのか、それとも偏っているのか、そもそも山の数は1つだけなのか、複数あるのか、そもそも山のないグラフなのかなど、分析のカタチをチェックしておくことで、データの実態を適切に表現するための統計量を正しく選べるようになります。

「いちいちグラフを書くのは面倒くさい」と思われるかも知れませんが、統計解析は料理と同じで下準備が味の決め手。下準備を怠るとその後どれだけ頑張っても不味くなってしまいう料理があるように、統計でも最初の一手間がトンデモな分析結果を出してしまわないために大事になってくるのです。

安易な「平均」の使用には要注意

今回ご紹介した統計量とは、データの特徴を知るための数値のこと。その統計量の代表格がこれまで2つのシーンで取り上げてきた「平均」です。「平均」は理解しやすく計算も簡単なのでついつい深く考えずに使ってしまうがち。しかし、今回見てきたように、アンケートの満足度などデータの種類や、左右に大きく偏って分布しているデータなどでは、「平均」を使うことでデータの特徴が正しく見えなくなる、という落とし穴に陥りかねません。

「平均」などの統計量を使うときは、自分がこれから分析しようとしているデータがどのような種類のデータなのかや、分布がどうなっているのかということ意識するのが何よりも大事なのです。



営業分析に役立つ『相関・回帰』の使い方

ここからは受注確率や売上高の分析に使える「相関」と「回帰」について、皆さんがお持ちの Excel での簡単な使い方を営業現場でよくあるシーンを使って解説していきます。これを読んだらきっと皆さんも自社のデータを使って分析してみたくなるはずです。

シーン 1：売れない商品 E の売り方を考える

営業会議での 1 シーン

課長 : 商品 E の売上が目標よりずいぶん低いし、案件もあまり無いようだ。

営業担当者 : 商品 E のニーズを誰が持っているか分からないですし、どうやって売ったらよいかも分からないので・・・

課長 : よその部署では売れているみたいだから、来週の営業会議で成功事例でも話してもらおうようにお願いするよ。

営業担当者 : いいですね。お願いします。(事例を聞いても業種や規模が違うし、どうせ手柄話で終わりだろうなあ)

どうやって売ったらよいか分からないのに、目標だけはしっかりと与えられている商品売ることほど大変なことはありません。そんな時に個別事情だけの販売成功事例を聞いても、自分の顧客には置き換えられず参考にならなかったということがよくあります。

そんな時に使えるのが「相関」分析です。例えば、売れている部署から成功事例ではなく売上データを借りてきて、「商品 E と一緒に売れている傾向の強い商品を見つける」ことができれば、その商品が売れているところに商品 E を紹介すればよいということになります。売れている部署から下のような顧客別の商品の売上データをもらったらしめたもの。Excel では「CORREL」という関数で相関係数を計算することができます。

	A	B	C	D	E	F	G	
1		年間売上高 (千円)						
2	顧客名	商品A	商品B	商品C	商品D	商品E		
3	・・・	1,500	0	0	2,653	0		
4	・・・	0	0	458	0	150		
5	・・・	1,600	7,147	0	1,281	0		
6	・・・	0	2,951	0	0	0		
7	・・・	0	11,207	0	5,866	0		
8	・・・	1,700	2,944	210	1,320	1,044		
9	・・・	442	1,000	1,000	1,000	6,498		
10	・・・	1,400	1,530	0	0	1,533		
11	・・・	0	2,900	1,292	0	3,413		
12	・・・	2,710	220	200	0	1,834		
13	相関係数	-0.0966	-0.30939	0.790858	=CORREL(E3:E12,\$F3:\$F12)			

相関係数はこう読もう

この表の読み方をご紹介します。

商品 A～商品 D の商品 E との相関係数を計算したところ、一番数字が大きかったのは商品 C (≈ 0.79) でした。そのため、商品 C を買っている顧客をリストアップして、商品 E を紹介するのが良さそうです。

ここで相関係数を読む上での目安をご紹介します。相関係数は、0 が無相関で、+1 に近づくほど相関関係が強くなります。一般的に、0.7 以上が強い相関、0.4 以上が中程度の相関、0.2 以上が弱い相関とされていますので、商品 C と E の売上には強い相関があることが分かります。

一方で、商品 B との相関係数はマイナスになっていますので、商品 B が多く売れている顧客では商品 E は売れないということになります。

この例では他商品の売上との相関を見ましたが、他にも顧客の規模や利益額、営業担当者の経験年数など様々な変数同士の相関関係を分析することが可能です。

相関分析を使いこなす 2 つのコツ

このように簡単に使えて便利な Excel の「相関」分析を使いこなすためにはいくつかコツがあります。ここでは 2 つご紹介しましょう。

1 つ目のコツは「見せかけの相関」に騙されないこと。

上の例では商品 A と商品 E の相関係数がほぼゼロになっていますがこれは見せかけ。商品 A と商品 B、商品 B と商品 E などの商品同士の色々な関係が商品 A と E の相関関係に実は影響を与えているということがあります。それらの影響を取り除いた商品 A と E の純粋な相関（偏相関）係数は、+0.36 とそこそこの強さの相関関係がありますので、実はセットで販売できる組合せだということがあります。

2 つ目のコツは「使うデータの尺度」に気を付けること。

Excel の関数で使える尺度は間隔尺度と比尺度なので、順序尺度が混ざったデータは使わないことです。

シーン 2：顧客の売上高を予測する

またまた営業会議の 1 シーン

課長 : 来年度の課の売上目標を私なりに割り振ってみましたので、確認してください。

営業担当者 : 私の顧客の X 社ですけど、まだ取引開始してから 1 年目なのに 3,000 万円も積まれているんですが・・・

課長 : その会社は年商もかなりの規模だから、食い込んでいける余地はきっとあるよ。

営業担当者 : 確かにそれはそうですが、大きな売上を作るには長期的な関係性が必要だと・・・

課長 : まあ、まずは改めて情報収集してみて、そこから調整の必要があるか考えようよ。

自分の担当する顧客がどれだけの売上になりそうかは、営業担当者にとって、また営業マネージャーにとっても悩みの種です。そんな時に「回帰」分析を使えば、過去の顧客データをもとに今後の予測を立てることができます。

Excel 関数「TREND」でカンタン回帰分析

	A	B	C	D	E	F	G	H
1	顧客	年商 (百万円)	営業利益率 (%)	設備投資比率 (%)	累計取引年数 (年)	年間売上高 (千円)		
2	...	67,518	2	6	9	7,470		
3	...	4,318	3	1	6	9,430		
4	...	943	14	12	2	30,000		
5	...	12,062	1	13	5	7,000		
6	...	2,091	0	25	3	12,900		
7	...	17,545	1	12	3	24,374		
8	...	30,036	3	12	10	1,000		
9	...	6,138	3	8	10	23,070		
10	...	19,990	36	21	10	121,070		
11	顧客X	22,586	10	18	1	=TREND(F2:F10,B2:E10,B11:E11,TRUE)		

この表は、顧客の「年商」や「営業利益率」「設備投資比率」「累計取引年数」というデータ（説明変数）をもとに、顧客Xの「年間売上高」（目的変数）がどのくらいまでいけそうかを分析しようというものです。Excelの「TREND」という関数を使うと33,760（3,376万円）となりますので、営業会議での課長の読みはなかなかいい線を行っていたことが分かります。

これだけでもなかなかのものですが、Excelのアドイン「分析ツール」を使うと、さらに細かい分析まで可能になります。

Excel アドイン「分析ツール - 回帰分析」の結果はこう読もう

	A	B	C	D	E	F	G	H	I
1	概要								
2									
3		回帰統計							
4	重相関 R	0.974581							
5	決定 R2	0.949808							
6	修正 R2	0.719693							
7	標準誤差	13080.84							
8	観測数	9							
9									
10		分散分析表							
11		自由度	変動	分散	観測された分散比	有意 F			
12	回帰	4	1.62E+10	4.05E+09	23.65439481	0.004802			
13	残差	5	8.56E+08	1.71E+08					
14	合計	9	1.7E+10						
15									
16		係数	標準誤差	t	P-値	下限 95%	上限 95%	下限 95.0%	上限 95.0%
17	切片	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
18	年商 (百万円)	-0.07284	0.262405	-0.2776	0.792429901	-0.74738	0.601689	-0.74738	0.601689
19	営業利益率 (%)	2836.583	443.1331	6.401199	0.001379279	1697.473	3975.693	1697.473	3975.693
20	設備投資比率 (%)	432.2362	501.8287	0.861322	0.428423879	-857.756	1722.228	-857.756	1722.228
21	累計取引年数 (年)	470.3221	1290.966	0.364316	0.73052066	-2848.21	3768.857	-2848.21	3768.857

この表が、Excel の分析ツールを使うと出てくる分析結果です。難しそうな言葉が並んでいますが、見るべき場所は赤枠で囲っている 3 か所です。

1 つ目は「重決定 R2」です。これは回帰分析の予測精度であり、作ったモデルで目的変数（「年間売上高」）がどれだけ説明できるかを示します。およそ 95% の精度で予測できているので、かなりよい分析になっています。

2 つ目は「有意 F」。上の「重決定 R2」の予測精度の数字が本当に確かなのかをチェックするための値で、一般に 0.05 より小さければ安全。今回のモデルでは 0.004 と非常に小さいので問題なしです。

最後は「P-値」。これは「年商」から「累計取引年数」までのそれぞれの説明変数を使うことの安全性を見るための数字です。これも小さいほど良く 0.05 が目安となります。上の例だと「年商」と「累計取引年数」を含めずに分析する方が安全なようなので、営業会議の場で「年商」を理由としていた課長も、「関係性（累計取引年数）」を理由としていた営業担当者も、実は的外れな会話をしていたということになります。

回帰分析を使いこなす 2 つのコツ

と、営業の分析をする上でかなり使える Excel の「回帰」。これを使っていく上でもやはりいくつか注意点があります。今回は特に大事なものを 2 つご紹介します。

1 つ目の注意点は予測したいデータ（目的変数）を売上高のような連続した数字データにすること。

今やっている商談が「受注」するか「失注」するか分析したいと言う場合は、受注/失注のデータを受注金額や数量に置き換える（失注の場合は見積金額を負の値にする）などの工夫をすれば OK です。

もう 1 つの注意点は、分析前後のデータのチェックです。

例えば先ほどの分析のデータの中に「従業員数」のように「年商」とかなり相関関係の強いものを追加して分析すると、多重共線性という問題が起こって回帰分析の結果が不安定になってしまいます。そのため、回帰分析をする前に説明変数の相関関係を計算しておき 0.9 を超えるデータは削除しておくということが必要です。

また、先ほど見たように「P-値」が高い変数がある場合は、それを除いてもう一度分析してみることがお勧めです。ちなみに「年商」と「累計取引年数」を除いて分析すると、重決定 R2 は 0.948 のまま、有意 F はさらに小さくなり、0.42 もあった「設備投資比率」の P-値も 0.11 にまで下がりますので、より良いモデルになると言えます。

「相関」「回帰」で大事なものはデータをイメージする力

今回ご紹介した「相関」と「回帰」を使いこなすにあたって、一番のポイントは「関係性が隠れているデータをイメージする力」ことです。例えば、売上高を予測するのにどんなデータがあったら良いだろうと考えること。営業担当者のスキルでも予測できるか？顧客企業の長期的な伸び率を見るのに新卒採用人数が使えないか？などとイメージを膨らませてみて、実際にいくつかのデータを集めて分析しそのイメージの確からしさをチェックする。このイメージ力があるかないかで、面白くて役に立つ分析結果が出るか、とこかで見たような一般的な分析結果どまりになるかが決まってくるのです。

私が以前やったことがあるのは、商談が受注するかどうかのデータに「SFA 上での上司とのメッセージのやり取り」というものがありました。SFA で報告した内容に対して、上司がどれくらいの頻度で、かつどれくらいの文字量でコメントしてい

るか、というものが商談の受注／失注と大きな相関関係があったのです。このように「関係性が隠れているデータをイメージする」ことの大事さと楽しさを感じながら、皆さんも今回ご紹介した「相関」と「回帰」を Excel で試してみてください。



データの視覚化に役立つ『分類』の使い方

今回のトライツブログは「営業マネージャー／営業企画のための統計活用入門」として、実際の営業現場で使える統計学の手法／ツールについてです。

四回目で最終回の今回は「分類」について。ここでいう「分類」とは、複数のデータの関係性をもとにグループ化しようというもの。この「分類」の使い方について、営業現場でよくあるシーンを題材に解説していきます。

シーン：複数ある商品の特徴をわかりやすく見せられない！

食品メーカーでの社内打合せのシーン

課長：今日のX社向けの商品説明会はどうだった？

営業担当：グラフを入れた資料を使って説明したんですけど「違いが分かりにくい」って言われちゃいました…

課長：この間試食してもらって官能評価をしたアレだろ？商品ごとに「甘味」「苦味」「コク」とかの点数をつけた

営業担当：そうです。点数をランキング形式やレーダーチャートにしてみたのですが、「ゴチャゴチャしてる」と…

課長：（資料を見て）確かにゴチャゴチャしているなあ。もう少しパッと見てわかるようにシンプルにならないもんかな…

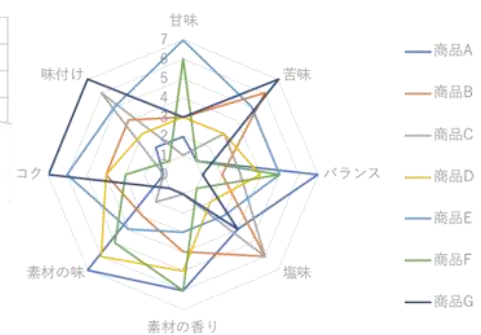
ここでこのメーカーの営業担当が持っていった資料を見てみましょう。

商品A-Gの官能評価結果（1～5点で評価、n=15）

製品名	甘味	苦味	バランス	塩味	素材の香り	素材の味	コク	味付け
商品A	3.0	3.0	3.3	3.6	2.2	2.7	3.3	3.0
商品B	3.3	3.7	2.4	3.7	1.9	1.9	3.7	3.4
商品C	2.7	3.1	2.7	3.7				4.2
商品D	3.3	3.1	3.1					
商品E	4.2	3.2						
商品F	4.1							

上記評価結果をランキング（1～7位）

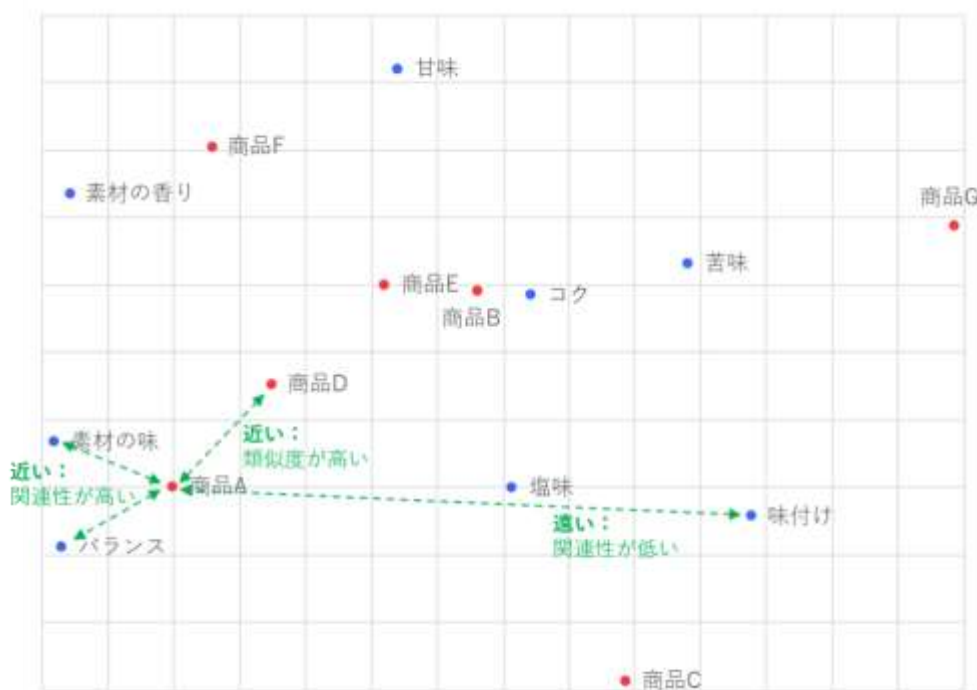
製品名	甘味	苦味	バランス	塩味	素材の香り	素材の味	コク	味付け
商品A	6	6	1	3	1	1	6	6
商品B	3	2	6	1	4	5	3	4
商品C	7	4	5	1				2
商品D	3	4	4					
商品E	1	3						
商品F	2							



点数をランキング形式にしてみたり、それをさらにレーダーチャートにしたりしてはいるのですが、顧客から「ゴチャゴチャしていてわかりにくい」と言われてしまうのも仕方ないですね。こんな時に役に立つのが統計の「分類」です。

コレスポネンス分析で商品と特性をマッピング

最初に顧客に提出した表のように「商品」と「官能評価の項目」など、複数の項目同士の関連性を一目でわかるように分析する手法がありますのでご紹介します。それはコレスポネンス分析というもので、項目同士の関連性を距離に置き換えて図示する分析手法です。「官能評価結果」の表を統計ソフトに読ませて分析方法を指定すると、簡単に分析結果が出てきます。



上はその分析結果ですが、見方に少しクセがあります。こういう図を見ると「縦軸と横軸は何?」と思われる方が多いと思いますが、この図の縦軸と横軸には意味がありません。この図の中で意味があるのはデータ同士の距離だけで、距離が近いものほど類似度／関連性が高いということを表しています。ですので、この図全体を回転させても表している意味合いは全く同じです。

上の図では、商品 A~G が赤丸で、「甘味」から「味付け」までの評価項目が青丸でマッピングされています。わかりやすいように商品 A と距離が近い評価項目と遠い評価項目の一部を緑色で補足してあります。このことから「商品 A は素材の味が強くまたバランスが取れており、味付けは弱めだ」ということがわかります。ほかにもこの一枚の図から

「商品 F は素材の香りと甘味が強く、苦味や塩味は比較的弱い」

「商品 E と商品 B はコクが強めて味の傾向が似ている」

など、さまざまなことが言えます。

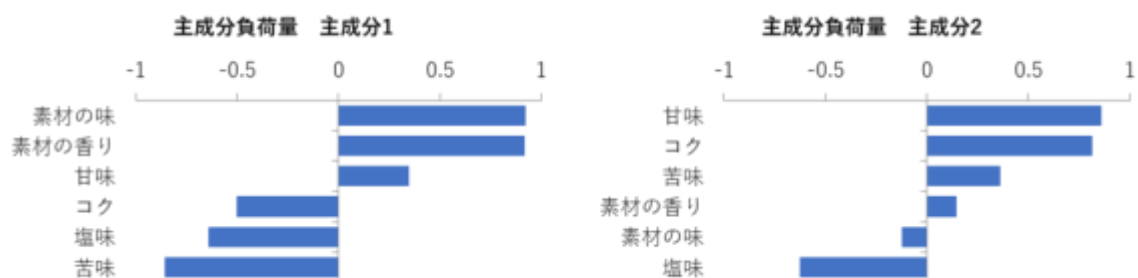
このコレスポネンス分析を使うと、例で示した「評価項目ごとの商品の分類」だけでなく、営業に役立つ様々な分析が可能になります。

「規模や業種などの顧客グループごとに、売れている商品グループを分類」

「過去購入されている商品グループごとに、追加で売れる商品グループを分類」
などのデータを視覚的に図示できるようになるのです。

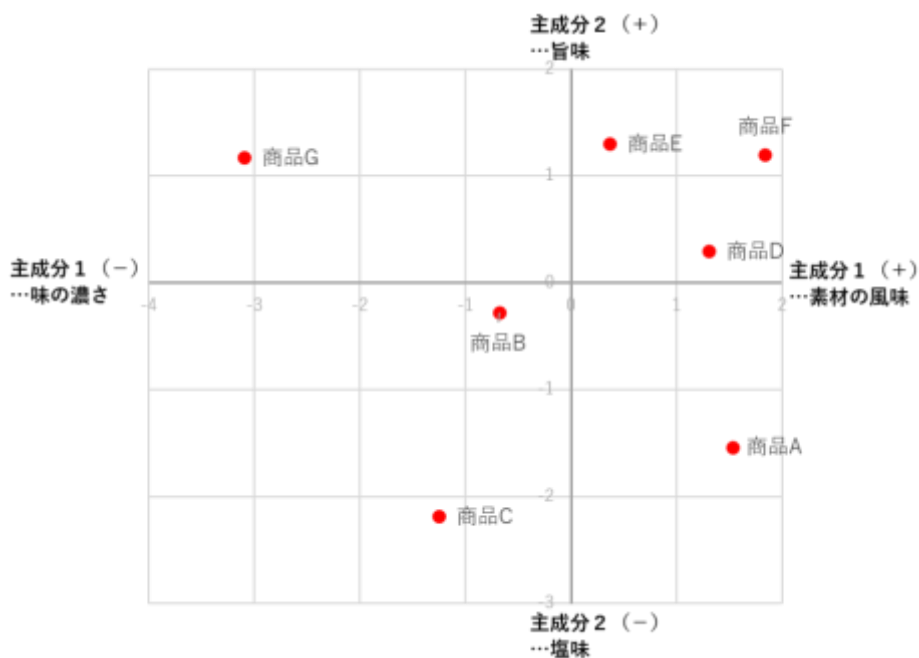
主成分分析で軸を整理してわかりやすくマッピング

とはいえ、上のマップでは「甘味」や「塩味」など評価項目が8つもあるので、冒頭の課長さんのように「もうちょっとシンプルにならないか?」と思われる方もいることでしょう。そんなときに使えるのが主成分分析です。主成分分析は、たくさんある項目（変数）をより少ない変数で効率的に表現できるように要約しようというもので、この要約した変数のことを「主成分」と呼びます。この分析で一般的なのが、たくさんの変数を2つの主成分にまとめて縦横二次元のマトリクスにするというもの。



上が主成分分析の結果です。主成分1がプラスだと素材の味や香りが強く、マイナスだと苦味や塩味などの味が濃くなります。2つ目の軸である主成分2は、プラスだと甘味とコクが強く（≒旨味があり）、マイナスだと塩味が強いということになります。また、もともとは8つあった評価項目が6つに減っていますが、これは相関係数が高すぎる項目が入ったままだと分析結果が正しくなくなってしまうために、それらの項目（「バランス」と「味付け」）を削除しているからです。

統計ソフトを使うと、上の2つの軸（主成分）で商品A~Gをマッピングしたものが自動的に出力されます。それに主成分1と2の意味合いを縦軸と横軸に書き足したのが次の図です。



このように主成分分析を使って評価項目をまとめることで、最初にレーダーチャートで図解していたときよりも評価項目がシンプルになり、さらに四象限ごとに商品の特性が明確になっていることがわかります。

右上の商品 D・E・F は「素材の風味が強く、旨味の強い商品」ですし、左下の商品 B・C は「味が濃く、塩味の強い商品」とシンプルに表現できるので、これなら顧客の X 社の方にもきっとわかってもらえることでしょう。

「分類」で分析するためのおすすめツール 2 つ

ここまでご紹介した分析手法は Excel の標準機能にはありませんので、無償または有償の分析ツールを使うこととなります。以前は統計解析のツールといえば難解で高額、さらに使い勝手が悪いものが多かったのですが、最近はそれも様変わりしています。ここでは営業マネージャー／営業企画の方にも使いやすい分析ツールを 2 つご紹介します。

とにかくコストをかけずに簡単に分析したいという方は「HAD」がおすすめです。完全に日本語で作られていますし、Excel のマクロとして使えるので、ほかの無償分析ツールと比べて取っつきやすさはピカイチです。ただし、個人で開発されたものをオープンにシェアしているという性質のため、機能の保証などはありませんので、そこだけ注意して使ってください。

多少なら費用を払ってもよいから安心して使えるものが欲しいという方には、株式会社社会情報サービスの「エクセル統計」がおすすめです。こちらも完全日本語対応でかつ Excel のマクロとして使えますし、さまざまな種類の分析手法が収められています。以前はこのような分析ソフトは 1 つで数十万円していたのですが、エクセル統計は 1 万円／年以内でライセンス購入が可能なので、かなり取っつきやすくなっています。ちなみに今回ご紹介したコレスポンデンス分析／主成分分析には、このエクセル統計を使っています。

「分類」で大事なのはデータを「正しくモデル化する力」と「解釈する力」

今回ご紹介した「分類」を使いこなすためのポイントは2つ。「正しくモデル化する力」と「解釈する力」です。

「正しくモデル化する力」とは、分析で得られた結果がもともとのデータをどれだけ漏れが少なくモデル化できているかということで、より漏れの少ない正確なモデルを作ることが大事です。正しくモデル化できているかを測る方法として、例えば先ほどの主成分分析では、以下のような表が分析結果として出力されます。

固有値表				
主成分	固有値	寄与率	累積寄与率	
1	3.201	53.34%	53.34%	
2	1.965	32.75%	86.10%	
3	0.497	8.28%	94.37%	
4	0.246	4.10%	98.48%	
5	0.080	1.33%	99.80%	
6	0.012	0.20%	100.00%	

この表でモデルの正しさを表しているのが「寄与率」と「累積寄与率」。この表によると、6つまである主成分のうち最初の2つだけで、もともとのデータの86.1%をモデルに含められていますので、なかなか出来の良いモデルになっていることがわかります。コレスポンデンス分析でも出力結果に「寄与率」と「累積寄与率」が表示されますので、どれだけ正しくモデル化できているかをチェックするようにしてください。

もう一つのポイントである「解釈する力」は、私たち分析者側に求められるものです。これまで見てきた「統計量」や「相関」「回帰」といった分析は、「この変数とこの変数の相関関係は高い」「顧客のこういう特性を見れば、売上高を予測できる」というように分析結果自体が明確な意味を持っています。

一方で、今回見てきた「分類」の分析は、乱暴な言い方をすると元の複雑なデータをわかりやすく表現するための分析なので、分析結果自体には新たな意味というものがありません。そのため、私たち分析者側で意味づけ（=解釈）してあげなければならないのです。マッピングの結果や主成分の要素を眺めてみて、「これは何を指しているのだろうか」ということを考えて言葉に変えていく「解釈する力」が「分類」には欠かせませんし、いくら「正しくモデル化する力」が高い分析であっても、実際のビジネスに役立つ解釈が得られないものであれば、もっと上手に解釈できるように分析し直す必要があるのです。

統計を活用してデータ分析をもっと面白く

ここまで「営業マネージャー／営業企画のための統計活用入門」として営業現場で使える統計の手法／ツールをご紹介してきました。その中で私が一番お伝えしたかったのが、「統計をうまく使えば普通のデータ分析はもっと面白いものになる」ということです。

私は統計の役割は、データや分析結果の正しさを「確かめる」ことと、パッと見ただけでは分からない関係性を「見つける」ことの2つにあると思っています。特に後者の「見つける」ができると、分析は面白いものになります。複雑なデータの中に埋没していて誰も気づいていない商品同士の売れ方の関連性や、売上や受注につながる隠れた要因を見つけられるのが統計の力。ぜひ4回の連載でご紹介した統計手法をご自分でも試してみて、データ分析の面白さを体感してください。

これからの営業担当者にとって、「統計」は絶対に武器になると思います。そして、もしご希望がありましたら、こちらでご紹介しているような内容での社内勉強会なども承ります。お気軽にご相談ください。

B2B 営業が今取り組むべき「探索型データ分析」とは

昨今、データサイエンスや AI の普及が日本の大企業の姿を変えつつあります。

今年に入ってから、証券業や銀行業、製薬業などで大幅な組織体制の見直しが発表されています。そして、それらの企業の多くが既存の営業体制や事務体制を見直して、データサイエンスや AI に明るい人材の強化を進めようとしていることは皆様もご存知の通りです。新聞などでよく使われる「データの世紀」という言葉が、まさに現実のものになってきているのです。

そのような中、データサイエンスや AI の普及に伴って「探索型データ分析」というキーワードが注目を浴びつつあります。そこで、「探索型データ分析と、今 B2B 営業がそれに取り組むべき理由」について考えていきたいと思います。



B2B 営業でよくある「検証型データ分析」

探索型データ分析の言葉の説明に入る前に、皆さんが仕事で目にしてこられた「B2B 営業でよくあるデータ分析」を思い返してみてください。おそらく、多くの方がイメージされているデータ分析とは、次の 2 つのタイプのものが多いのではないのでしょうか。

1. 半期末や四半期末といったタイミングで、当初立てた計画通りに売上が上がったのかを確認する
2. 計画と実績に差が出た場合に、実績数字を年間／半期／四半期／月次といった「期間」ごと、支店や営業所／部課などの「部署」ごと、「商品」ごとに、比較／分解する

これらの 2 つのデータ分析は、「あの顧客にはこの商品がいつのタイミングでこれだけ売れるはずだ」という計画（仮説）に対して、計画通りに進んだのかどうかを確認（検証）するというもので、「（仮説）検証型データ分析」と呼ばれます。売上見込などの仮説が当初から存在している、というのが「検証型データ分析」の最大の特徴です。

新たなアイデアを見つける「探索型データ分析」

その一方で、最近よく取り上げられるようになった「探索型データ分析」というのは、明確な仮説がない状況でその仮説を見つけるために行う分析のことを言います。

「売上が伸びている商品 A と、伸びていない商品 B との違いは何だろう？ 営業担当のスキルだろうか、それとも顧客の分類に何か違いがあるのだろうか？」

「受注する商談と、失注する商談との違いは何だろう？商談の進め方に何か違いがあるのだろうか？」

などというように、売上が伸びたり、受注したりする理由がよく分からないときに、営業担当者のスキルや、顧客の分類、商談の進め方などの様々な要素を組み合わせて分析し、今まで明らかになっていなかったアイデア（因果関係や構造）を見つけようというものです。

私は、今まで色々な B2B 営業組織の売上データや SFA データなどを使って探索型データ分析をしてきました。その中でも印象的でその後の営業改革につながった分析結果は以下のようなものです。

- 新規開拓での受注確率が最大となる提案書のページ構成
- SFA への入力データから見る、商談の受注確率向上につながる顧客情報の種類とキーワード
- 研修形式で実施した新規開拓営業プロジェクトで、商談の進捗に好影響を与える営業マネージャーの指導スタイル

分析を始めるときには「何か関係があるかもしれない」程度のぼんやりとした疑問に過ぎなかった考えが、分析を進めていくことで売上につながる因果関係へと姿を変えていくというのが、この探索型データ分析の醍醐味なのです。

B2B 営業組織が探索型データ分析に取り組めない理由

しかし、残念ながら日本の B2B 営業組織でこの探索型データ分析に取り組んでいるところはほとんどない、というのが実情だと私は考えます。

その理由として、

「営業現場にデータ分析が得意な人がいない」

「そもそも探索型データ分析をやったことも、見たこともない」

「価値あるアイデアが必ず見つかるとは確言できず、投資対効果が不明なので踏み出せない」

など、いろいろとありますが、最たるものだと私が考えるのが「ちょうどよい無知の人がいない」というものです。

探索型データ分析をするうえで必要になるのが、これまでの固定観念や過去の経験から当たり前だと思っていることを一度頭の中から追い出して、「もしかしたらこのデータと売上数字の間に何か関係があるかも!？」と一見するとバカバカしい発想でも試行錯誤を繰り返すことです。このときに固定観念や過去の経験にとらわれたままだと、自分の頭の中に既にある仮説を検証するだけの分析になってしまい、新しいアイデアを見つけることはできません。

その一方で、業界や業務に関する予備知識がまったくない外部の分析屋さんに依頼すると、使い道のない結果や、当たり前すぎる結果が探索型データ分析のレポートとして上がってくることがままあります。そうなると、ページ数は多いものの本当に使える価値のある考察はほんの数ページだけ、ということになり「手間と金額の割には・・・」となってしまいます。

そのため、探索型データ分析を実施するには、業界や業務に関する予備知識はありつつも、過去の経験や固定観念にとらわれずに試行錯誤を繰り返すことができる「ちょうどよい無知の人」を見つける必要があるのです。この「ちょうどよい無知の人」は営業現場や営業企画部門にはなかなかいません。どうしても日々の経験から固定観念が生まれてしまうので、目新しい切り口での分析にならないのです。個人的な経験ではありますが、SFAなどを運用・管理しているグループ会社のシステム担当、市場データを集めるときによく使う調査会社の分析担当者、などが探索型データ分析に適した「ちょうどよい無知の人」であることが多いように私は感じます。

B2B 営業を取り巻く環境が激変する今こそ探索型データ分析に取り組もう

SFA や MA など、B2B 営業の現場で様々なシステムツールが導入され、多くの商談データや顧客データを扱うことができるようになりました。また、多くのシステムツールで分析機能が強化されており、以前は統計解析の専門家でなければできなかったようなデータ分析を、今ではユーザーサイドでも簡単に実現できるようになっています。このように B2B 営業にも「データの世紀」が訪れている現在では、これらのデータを活用しないというのは非常にもったいないことです。

そして、探索型データ分析に今取り組むべきもう一つの理由は、B2B 営業が大きな転換点に差し掛かっているからです。顧客はますます Web を使って情報収集するようになってきており、顧客が購買活動の主導権を握りつつあります。また、MA やインサイドセールスなどによって、これまで対面でのコミュニケーションが主体だった B2B 営業が大きく変わりつつあります。これはつまり、今までに獲得した経験に基づく営業の成功パターンが通用しなくなる、ということです。これまでの固定観念にとらわれず、売上や受注につながる新しいアイデアを見つけることが激変する現在の B2B 営業組織に必要であり、今回ご紹介した探索型データ分析はそのための有効な切り口の 1 つなのです。

トライツコンサルティングでは、蓄積された SFA データを「探索型」で分析するなどいろいろな実績があります。一つのパターンができると、それを応用しながら発展させていくというのもデータ分析を浸透させていくやりかたのひとつだと思います。ご興味のある方はお気軽にご相談ください。

SFA データ分析で「売上予測」「受注確率向上のヒント発見」に挑戦してみよう

「統計学が最強の学問である」という本が出版されてから早7年。最初に書店でタイトルを目にしたときは、「そう言いたいのは分かるけど、ちょっと言い過ぎでは？」と思っていたのですが、AIや機械学習がこれだけ日常生活に浸透してきているのを見ると、まさにタイトル通りの時代になってきたと言わざるを得ません。

先日（2020年1月26日）の日経新聞の朝刊にも、「経済学者、企業の参謀に最新理論で効率化」という記事がありました。記事によると、統計学で今注目されている「因果推論」という理論を通販サイトで活かし、最適なサイト作りに役立っているとのこと。日本でも、統計学が実際のビジネスの意思決定に大きな影響を与えるようになってきているのです。

その一方で、B2B営業ではSFAが導入されるなどして、日々蓄積されるデータは格段に増えているものの、それを上手に分析して営業戦略や商談の進め方の意思決定に役立っている事例はあまりないようです。そこで、SFAに入力される営業データの統計的な活用方法「予測」と「発見（因果推論）」について、ご紹介したいと思います。



B2B 営業で皆が欲しがるデータ分析「予測」と「発見」

B2B 営業でデータ分析をしようというとき、どの営業組織の方と話をしても「それを知りたい！」と言われるのは、「今の状態だと、将来どうなるか（＝予測）」
「より成果を上げるために、何をしたらいいか（＝発見）」
の2つです。

営業マネージャーや営業企画の方は、定期的に報告される数字を見ながら期末の着地数字がどうなりそうかを予測して、上に報告することが求められていますし、それと同時に、より良い営業成績を上げるために何をすれば良いかを発見しようとしています。

この「予測」と「発見」をするための手法が「回帰分析」というものです。回帰分析について、まったく聞いたことがないという方もいらっしゃると思いますので、簡単にご説明します。

予測と発見のための「回帰分析」

回帰分析とは、売上や受注確率のような予測したいもの（目的変数、 y ）を、それに関係する要因（説明変数、 x ）で説明するもの。中学生の頃に

$$y = ax + b$$

というシンプルな式を勉強しましたが、回帰分析のモデルもこの式を使って

$$\text{目的変数} = a \times \text{説明変数} + b$$

というように表します。この式の意味は、説明変数に係数 a を掛けて定数 b を足すと目的変数の予測値になる、というもの。実際の分析では、説明変数が 1 つだけということは稀ですので、

$$\text{目的変数} = a_1 \times \text{説明変数}_1 + a_2 \times \text{説明変数}_2 + \dots + a_n \times \text{説明変数}_n + b$$

というように長い式になりますが、基本的な構造は「 $y = ax + b$ 」と同じです。

このような式（データ分析ではモデルと呼びます）を使った、「期末の売上を予測する」「受注確率の向上に関する要因を発見する」などの分析が「予測」と「発見」です。

同じ回帰分析という名前前の分析手法を使い、同じモデルで表現するのに、なぜ「予測」と「発見」と言葉が使い分けられているのでしょうか。それは、モデルの中で重要視しているものが異なるからです。

正確性を最優先する「予測」分析

「予測」で大事なものは目的変数（ y ）の値の正確性。より正確に予測できればよいため、説明変数（ x ）がブラックボックス的でも許される、という考え方です。極端な例を挙げると、営業チームごとの売上を正確に予測できるのであれば、説明変数の中に「営業担当者の昼食時間」のようにその関連性がイメージしにくいものが入っていてもかまわない、というのが「予測」におけるスタンスです。

この「予測」はデータの件数が多ければ多いほど、かつデータのバラつきが大きければ大きいほど精度が高くなる傾向がありますので、データの件数さえ揃えばそれほど技術や経験がなくてもある程度の精度で予測できるようになります。SFA を使ったデータ分析に取り組む第一歩としては、恰好の分析作業だと言えるでしょう。

もっともらしさを最優先する「発見」分析

かたや、「発見」で大事にされるのは説明変数（ x ）です。そのため、「売上を予測するモデルの中に、営業担当者の昼食時間が説明変数として入っている」というのは、新たな「発見」につながるようになってきます。

この「発見」の質を高めるには、データの件数を増やすよりもいろいろな説明変数を数多く考え、その中でもっともらしいものを探すことです。そのため、分析者には業務への理解と、それをもとに説明変数を設定できるだけの分析モデルを構築するスキル／経験が必要になります。

今流行りの「因果推論」とは？

例えば、失注した顧客 A に提案したときの対象者が、「担当者」だった場合、ドラえもんの「もしもボックス」を持っていない私たちは、顧客 A の「部長」「課長」に提案していたら受注していたかどうかを知ることはできないので、本当に提案対象者の役職が受注確率に因果関係を持っているかは分かりえません。

この、現実には起こっていないことに対して統計的に推論して仮の数値を求め、それを基に因果関係を確認するのが、冒頭の日経新聞の記事にあった「因果推論」です。極端な言い方をすると、分析によって「部長に提案していればどうなっていたか」をわかるようにしようということです。

あなたがやりたいのは予測？発見？

B2B 営業にとって身近な存在である SFA。その中には、売上などの「予測」や、受注確率向上のためのヒントの「発見」が可能になるデータが日々蓄積されています。これらのデータを定期的に棚卸し分析することで、現在の自分たちの営業がどこに向かっており、今後改善するために何が有効なのかを知ることができます。

ただこれまで使ってきた帳票を見やすくする、手間を掛けずに出力するというような枠にはまらず、どんな予測や発見がしたいのかを考え、少しずつ「データを使った予測や発見」を日々の仕事に取り入れていく。これが SFA を有効活用していくということなのだと考えます。

トライツコンサルティングでは、SFA などのデータを活用した「予測」や「発見」のサポートをしています。ご興味のある方はご相談ください。



トライツコンサルティングは

デジタル時代に合わせ、営業の「現場力」を高めるしくみづくりと人材育成を実現します

デジタルツール導入で営業現場の働き方改革を実現させようとして、苦勞している企業が増えています。

Webの発展によって、顧客の購買活動が大きく変化し、それに対応していくことも大きな課題になっています。

トライツは営業の「現場力」にフォーカスし、デジタル時代に合ったしくみづくりと、結果を出し続けられる営業人材育成を実現します。



デジタル時代の営業活動に変革

営業活動を通じて提供する価値が顧客から高く評価されるよう、人とデジタルの両方の顧客接点を変革し、短期間で結果を出します。



テクノロジーを営業の真の味方に

営業にとっても顧客にとってもメリットがあるテクノロジー活用を推進します。蓄積データを活用してPDCAを加速させ、継続的に進化させます。



次世代を担う営業人材を育成

イノベーションと一緒に取り組むことを通じて、5～10年先に営業部門を率いるリーダー、営業企画人材などの育成を行います。

トライツコンサルティング株式会社

東京都港区虎ノ門 3-8-25 日総第 23 ビル 9F

03 5733 4035

info@trights.co.jp