

TDB 景気動向調査「企業の声」に関するテキストマイニング分析

「企業の声」による景気分析の可能性

～「人手不足」、「米中貿易摩擦」、「消費税率引き上げ」と各種 DI の関係～

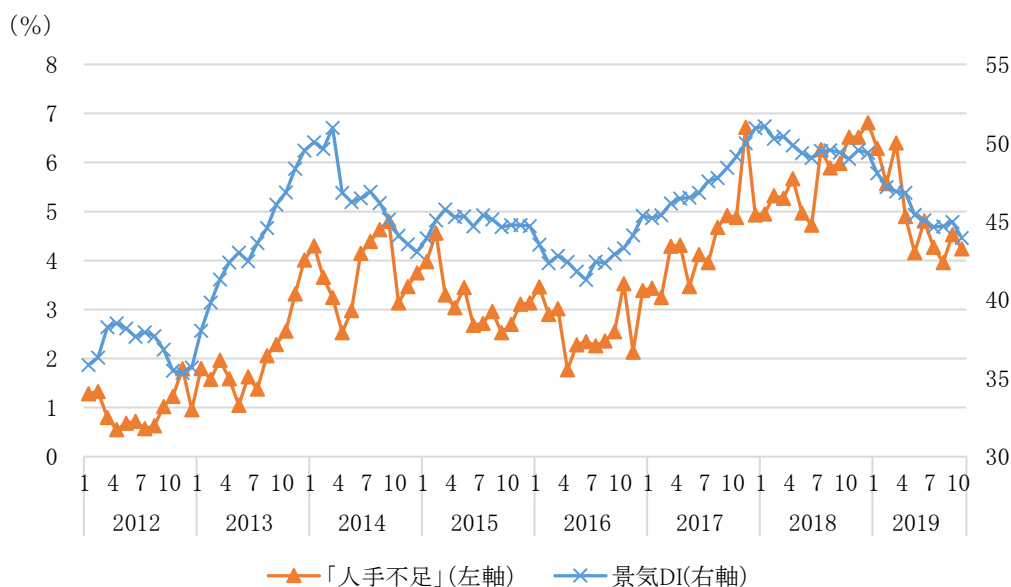
データソリューション企画部

杉原 翔太

【要約】

1. 景況感に影響を与えていると思われるトピックを「企業の声」から抽出するために、一定のルールで文章を分類するコーディングルールを作成し、それらと各種 DI との関係を確認した。過去1年間では、「米中貿易摩擦」がマイナスの方向に相関係数が大きく、「人手不足」はプラスの方向に大きくなっていった。2012年以降では、「人手不足」の相関係数が大きく、また、企業の業績を表していると考えられる語句も高い相関を示していた。
2. 近年景況感に影響を与えていると考えられる「人手不足」、「米中貿易摩擦」、「消費税率引き上げ」について、それぞれ動きを図表化して確認した。その結果、「人手不足」の頻度と、正社員の雇用過不足DIは比較的良く似た動きをしていることがわかった。2018年以降で、「米中貿易摩擦」と設備投資意欲DIはそれぞれ反対方向に変化していた。「消費税率引き上げ」に関して、2014年4月と2019年10月を比較したところ、前回よりも今回は「消費税率引き上げ」を意味する文章の登場頻度が低くなっていることがわかった。

「人手不足」に関する自由回答の頻度と景気DI



TDB 景気動向調査¹では毎月約1万社から、その月の景況感について回答を得ている。現在の景況感が「良い」か「悪い」か7段階で選択し、回答者の任意に基づき自由回答で、その理由を収集している。そうして集められた理由は、毎月の TDB 景気動向調査の発表とあわせて、「企業の声」としてインターネット上に公表している。

本レポートは、2012年1月から2019年10月調査の期間²を対象にし、その自由回答についてテキストマイニング分析³を行ったものである。テキストマイニングとは、大量のテキストデータを対象にしたデータマイニングのことである。文章をいくつかの単語、または特定のルールによって分類し、それらの出現頻度と傾向を分析する。

たとえば、

- 「企業の声」のなかで、“人手不足”について記述されている文章の出現頻度を月別、業界別に集計、その傾向をみる
- 特定の語句と共に用いられることが多い語句を抽出する

といった分析が考えられる。

1. 自由回答に頻出する語句

表1は、2018年11月調査から2019年10月調査までの1年間で、自由回答テキストに高い頻度で用いられている語句を抽出したものである。たとえば“受注”は、上記期間の全コメント20,543件のうち2,606件(12.7%)の回答で用いられていた。

消費税率引き上げに関する文章では“消費”、“増税”、“税”といった語句、人手不足に関する文章では、“不足”や“人手”などの語句が高頻度で用いられている。また製造業を中心に影響の大きいと思われる米中貿易摩擦に関しても、表1から“米”、“中”、“中国”、“貿易”といった語句をみると、現状の景況感の理由として高い頻度で使われているようにみえる。

また、“良い”、“悪い”、“多い”、“少ない”といった形容詞も、景況感を表す語句として頻繁に用いられていることが表1から推測される。

つぎに、頻出度の高い語句と共に用いられる(共起される)ことが多い語句の関係性をみる。図1は共起ネットワーク図と呼ばれ、語句間の共起関係を表している。たとえば、“受注”は“減少”と共起されるケースが多く、“受注”が含まれる文章の13.7%で“減少”が共起されている。

図1を細かくみていくと、消費税率引き上げに関する語句と考えられる“消費”、“増税”、“駆け込み”、“需要”が多く共起されていることがわかる。また、“オリンピック”と“建設”の間に共起関係がみられることから、オリンピックに関連した建設についての文章が、「企業の声」

¹ TDB 景気動向調査は2002年5月に開始した統計調査

² 上記の期間における、「企業の声」の回答数は16万9,152件にのぼる。1カ月平均でおよそ1,800件のテキストデータが蓄積されている。

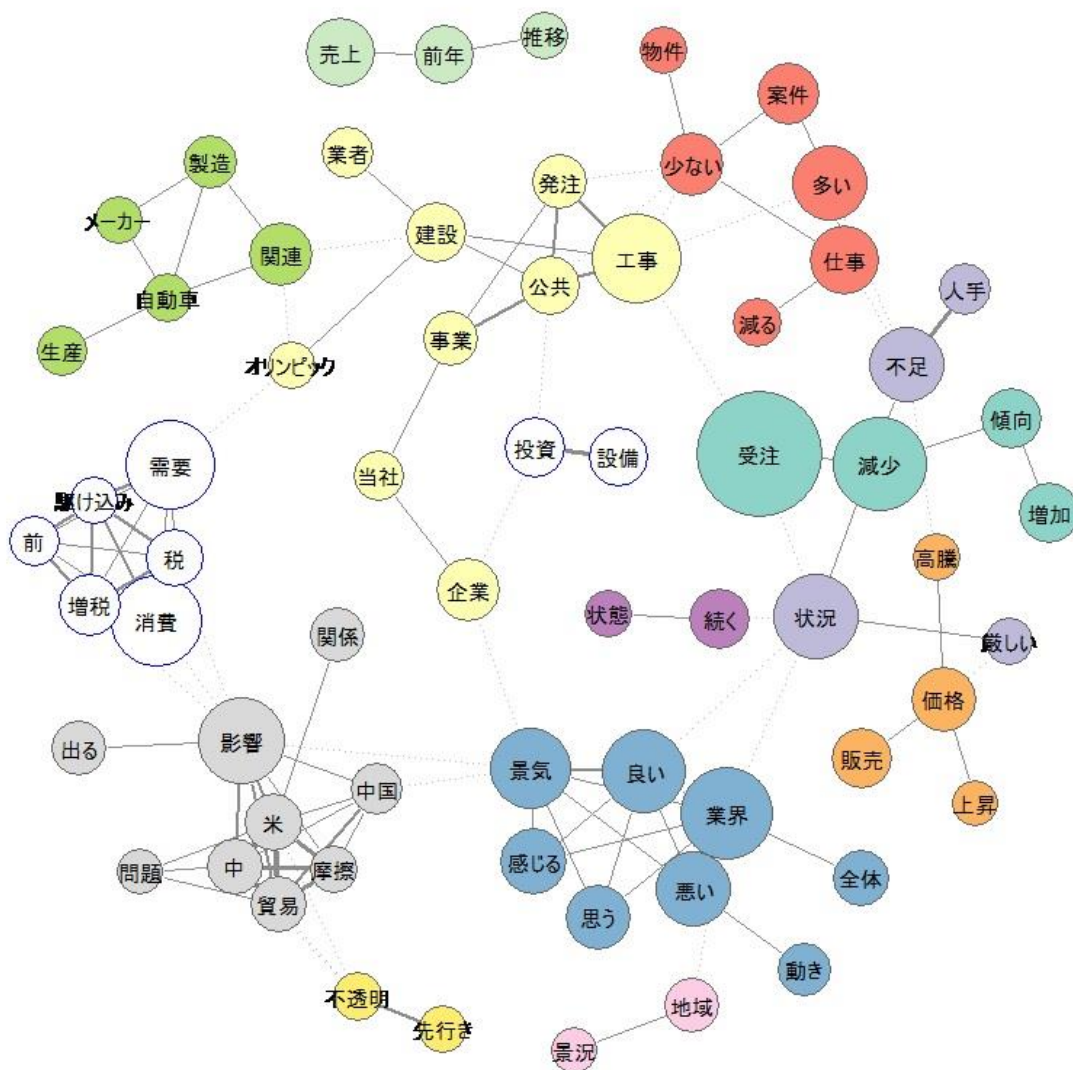
³ 本レポートでは、KH Coderというテキスト型データを統計分析するためのソフトを使っている。KH Coderは(<https://kncoder.net/>)より、ダウンロード可能である。

としてあげられていることもわかる。このように図1の共起ネットワーク図は、語句間の共起関係を視覚的にとらえることができる。

表1：自由回答テキストの上位頻出語句（2018年11月調査から2019年10月調査）

抽出語	コメント数	抽出語	コメント数	抽出語	コメント数	抽出語	コメント数	抽出語	コメント数
受注	2606	多い	1213	案件	798	前年	739	事業	597
減少	1652	不足	1206	価格	798	米	731	出る	593
業界	1635	悪い	1184	関連	793	税	724	発注	586
消費	1579	売上	1071	投資	777	設備	717	製造	581
需要	1549	仕事	998	傾向	776	全体	699	動き	574
影響	1526	感じる	953	増加	769	中	693	関係	562
状況	1395	思う	878	建設	765	貿易	672	人手	557
良い	1384	少ない	875	公共	759	増える	639	低迷	555
景気	1322	増税	825	販売	752	地域	624	中国	537
工事	1264	続く	811	企業	748	好調	612	昨年	536

図1：2018年11月調査から2019年10月調査の共起ネットワーク図



2. コーディングルールを用いた分析

つぎに自由記入形式のテキストデータを、コーディングルールによって分類した。コーディングルールを用いることで、ある一定のルールに該当する文章の出現頻度を計算することができる。たとえば、下記の文章は「消費税率引き上げ」を景況感の理由としてあげられているコメントを、最新の2019年10月調査から抜粋したものである。

- 「**駆け込み需要**一巡による売上減、どこまで水準が下がるか不明。」(小売)
- 「**増税**後の反動がある。」(卸売)
- 「本来は繁忙期だが、**消費増税**の影響か設備投資が少ない。」(建設)
- 「**軽減税率**対応のピークが過ぎたがまだ残案件・関係案件有り」(サービス)

上記は大まかにいえばすべて「消費税率引き上げ」についての文章であるが、コーディングルールを作成せず分類した場合、「駆け込み需要」、「増税」、「軽減税率」についてそれぞれ別の文章として分類される。コーディングルールを作成することで、これらをまとめて「消費税率引き上げ」に関する文章として分類することが可能となる。同様に「人手不足」や「米中貿易摩擦」についても、コーディングルールを作成することで、それぞれに関連するコメントを分類できる。

3. 景気DIとコーディングルールの相関関係

表2は、「米中貿易摩擦」、「人手不足」について、2018年11月から2019年10月までの自由回答をコーディングルールで分類し、その各月の登場頻度と景気DIの相関係数をとったものである。

これをみると、「米中貿易摩擦」は、景気DIと負の相関関係がみられた。2018年11月以降、景気DIは低下しているが、「米中貿易摩擦」に関する文章の頻度は反対の動きをみせ、増加している。「人手不足」は景気DIと正の相関がみられ、景気DIの低下にあわせて、その頻度が低下している。

表2：2018年11月調査から2019年10月調査における景気DIとの相関係数

語句	相関
米中貿易摩擦	-0.87
人手不足	0.92

表3では2012年1月以降のデータを用いて、コーディングルール⁴で分類した語句の頻度と、各月の景気DIで相関をとった結果のうち、相関係数の絶対値が大きいものを抜粋した。上記の結

⁴ 2018年以降のトピックである「米中貿易摩擦」や2014年4月、2019年10月の「消費税率引き上げ」については、2012年以降のデータとの相関関係でトピックとして取り上げられない期間が長いいため、表3の結果からは除いている。

果と同様に、「人手不足」が景気 DI と高い正の相関を示したほか、「好調」、「低迷」など業績を表していると思われる語句で相関係数の絶対値が大きくなっていることがわかる。

また、「人手不足」は、2012 年以降よりも直近 1 年間の方が高く、企業の景況感との関連が強まってきていることがうかがえる。

表 3：景気 DI との相関（2012 年以降）

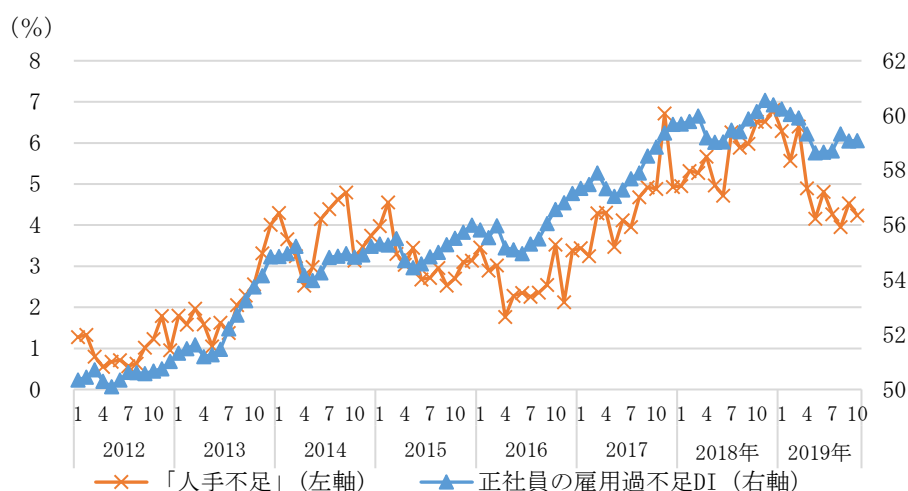
語句	相関	語句	相関
人手不足	0.82	低迷	-0.81
好調	0.81	輸出	-0.74
多い	0.79	復興	-0.72
堅調	0.70	競争	-0.71
忙しい	0.68	海外	-0.70
推移	0.67	国内	-0.69
順調	0.66	経済	-0.65
オリンピック	0.63	生産	-0.62
良い	0.62	悪化	-0.60
引き合い	0.61	先行き	-0.60

4. 「人手不足」と正社員の雇用過不足 DI

図 2 は、コーディングルールで「人手不足」と分類された文章の頻度 (%) を左軸に、TDB 景気動向調査の正社員の雇用過不足 DI（現在の正社員数の過不足感について指数化したもの。数値が大きいほど不足感が強い）を右軸に図示したものである。

これをみると、正社員の雇用過不足 DI と「人手不足」が文章中に登場する頻度は、総じて良く似た推移を辿っていることがうかがえる。2012 年 1 月から 2019 年 10 月までの期間において、「人手不足」と正社員の雇用過不足 DI の相関係数は 0.90 と非常に高い値になっている。

図 2：「人手不足」と正社員の雇用過不足 DI



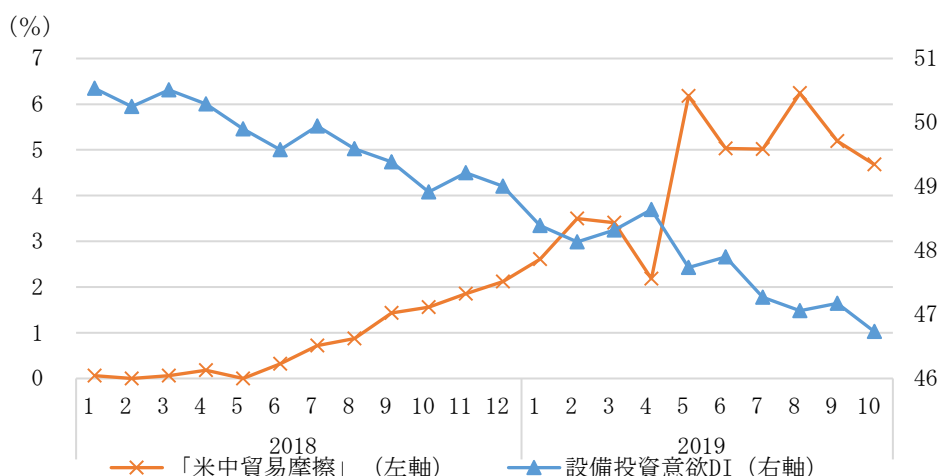
5. 「米中貿易摩擦」と設備投資意欲 DI

図3では「米中貿易摩擦」に分類された文章の頻度と、設備投資意欲DI（値が大きいほど設備投資に対する意欲が高い）を比較している。

2018年7月以降、「米中貿易摩擦」の回答頻度が徐々に高くなっており、とくに2019年5月に前月の2.2%から6.2%へと急上昇している。同月10日に、米政府は2千億ドル分の中国製品に課していた制裁関税の税率を10%から25%に引き上げ、8月には対中追加関税第四弾の発動を決定した。米中間での貿易摩擦が激しさを増す流れにあわせて、「米中貿易摩擦」の頻度は高くなっている。

一方、設備投資意欲DIは2018年以降低下傾向にある。米中貿易摩擦の影響で企業の設備投資意欲が減速しているという見方と、図3での「米中貿易摩擦」と設備投資意欲DIの動きはおおむね一致しているように見える。

図3：「米中貿易摩擦」と設備投資意欲DI



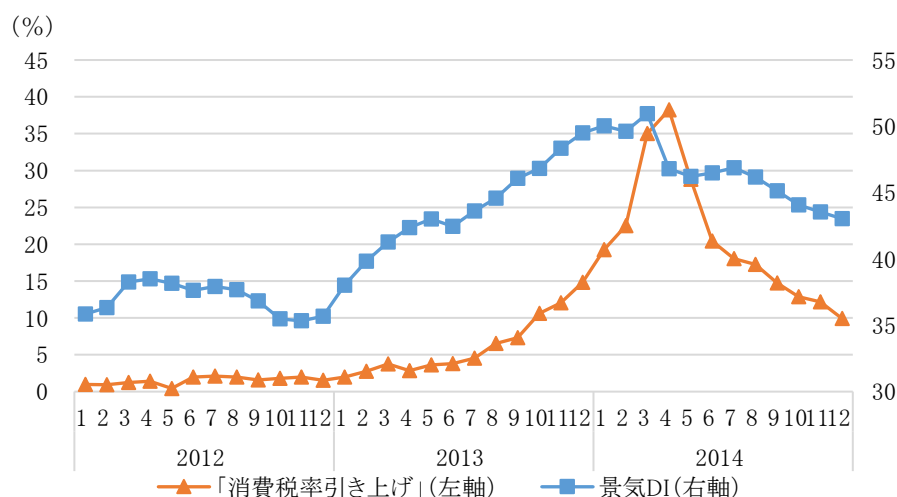
6. 「消費税率引き上げ」と景気DI

最後に、「消費税率引き上げ」と景気DIとの関係を2014年4月の5%から8%への引き上げ時、2019年10月の8%から10%への引き上げ時でそれぞれ確認する。

図4は、2012年1月調査から2014年12月調査までの期間で、コーディングルール「消費税率引き上げ」の頻度と、景気DIを時系列にそって図示したものである。2014年のときは、DIは2013年から2014年3月にかけて51.0まで増加し、その後4月に46.8に低下している。他方、「消費税率引き上げ」の頻度は2014年4月にかけて増加し、特に4月は自由回答の38.2%が「消費税率引き上げ」に関連する文章となっている。3月は『小売』、『運輸・倉庫』など5業界の景気DIが過去最高となるなど、駆け込み需要による景気押し上げがみられた。しかし、4月は『小売』の景気DIが前月比10.7ポイント減となるなど、駆け込み需要の反動減が幅広い業界に現れた。景気DIはリーマン・ショックがあった2008年12月（同4.1ポイント減）を超える、過去最大の悪化

幅（同 4.2 ポイント減）を記録した。「消費税率引き上げ」の動きは、消費税率を 5%から 8%へ引き上げた当時の景気 DI の動きと整合性がみられる。

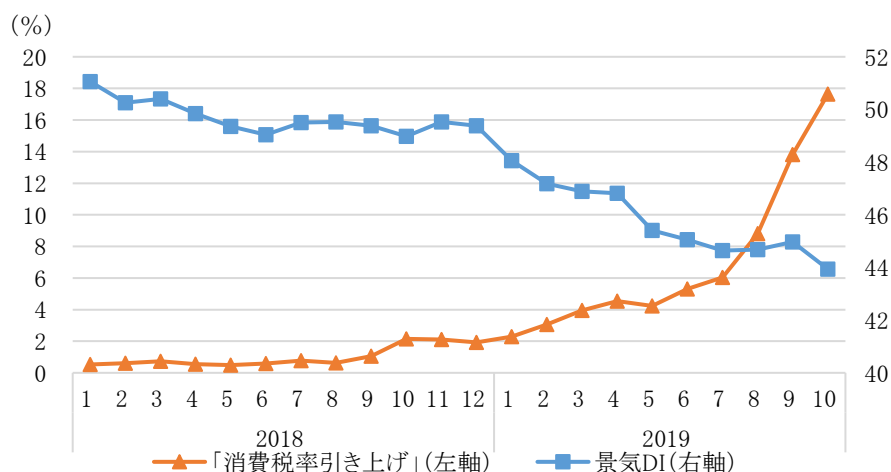
図 4 : 「消費税率引き上げ」と景気 DI (2014 年 4 月)



同様に、図 5 では 2019 年 10 月の引き上げを図示している。2014 年 4 月のときと比べると、「消費税率引き上げ」の頻度が 9 月 13.7%、10 月 17.6%と前回より低くなっている。また、景気 DI が税率引き上げ前から低下傾向にある点も前回と異なっている。9 月から 10 月で、景気 DI は 1.1 ポイント減となり、減少幅は 2014 年と比べると小幅であった。

「消費税率引き上げ」の頻度が低下した背景として、消費税の軽減税率制度、ポイント還元制度が一定の効果を発揮しているとも考えられる。一方、回答者の懸念材料が「米中貿易摩擦」や「人手不足」といった他の要因に割かれてしまったため、「消費税率引き上げ」については記入しなかった可能性も考えられる。

図 5 : 「消費税率引き上げ」と景気 DI (2019 年 10 月)



7.10 業界別クロス集計

表4は、2018年11月以降について「消費税率引き上げ」、「米中貿易摩擦」、「人手不足」の各コーディングルールにより10業界別に集計した結果である。「消費税率引き上げ」は影響の大きいと推測される『卸売』、『小売』で頻度が多く、「米中貿易摩擦」は同様に『製造』で多くなっている。

「人手不足」は『建設』、『運輸・倉庫』、『サービス』で頻度が多くなっている。帝国データバンクが2019年8月にリリースした「人手不足の解消に向けた企業の意識調査」によれば、この3業界での人手不足の影響は、「需要増加への対応が困難」が全体より高くなっている。上記の3業界では、やはり「人手不足」が景況感を抑制する要因になっているとみられる。

表4：10業種別クロス集計（2018年11月調査から2019年10月調査）

10業界	消費税率引き上げ	米中貿易摩擦	人手不足	コメント数
農・林・水産	2 (1.11%)	0 (0.00%)	9 (5.00%)	180
金融	15 (6.98%)	12 (5.58%)	8 (3.72%)	215
建設	211 (6.60%)	20 (0.63%)	267 (8.35%)	3,197
不動産	31 (4.75%)	12 (1.84%)	14 (2.15%)	652
製造	214 (3.88%)	400 (7.26%)	85 (1.54%)	5,510
卸売	456 (8.95%)	238 (4.67%)	116 (2.28%)	5,096
小売	123 (12.81%)	11 (1.15%)	15 (1.56%)	960
運輸・倉庫	49 (4.71%)	33 (3.17%)	159 (15.27%)	1,041
サービス	229 (6.35%)	77 (2.13%)	387 (10.73%)	3,608
その他	6 (7.14%)	3 (3.57%)	8 (9.52%)	84
合計	1,336 (6.50%)	806 (3.92%)	1,068 (5.20%)	20,543

※単位：件、カッコ内はコメント数に対する各コーディングルールの出現率を示している。

まとめ

本レポートでは、現状の景況感の理由として挙げられている『企業の声』について、テキストマイニング分析を実施した。

景況感に影響を与えていると思われるトピックをテキストから抽出するために、一定のルールで文章を分類するコーディングルールを作成し、それらと各種DIとの関係を確認した。過去1年間では、「米中貿易摩擦」がマイナスの方向に相関係数が大きく、「人手不足」はプラスの方向に大きくなっていった。2012年以降では、「人手不足」の相関係数が大きく、また、企業の業績を表していると考えられる語句（好調や低迷など）も高い相関を示していた。

また、近年、景況感に影響を与えていると考えられる「人手不足」、「米中貿易摩擦」、「消費税率引き上げ」について、それぞれ動きを図表化して確認した。その結果、「人手不足」の頻度と、正社員の雇用過不足DIは比較的良く似た動きをしていることがわかった。2018年以降で、「米中貿易摩擦」と設備投資意欲DIはそれぞれ反対方向に変化していた。「消費税率引き上げ」に関して、2014年4月と2019年10月を比較したところ、前回よりも今回は「消費税率引き上げ」を意味す

る文章の登場頻度が低くなっていることがわかった。

「企業の声」のテキストデータから抽出した指標と、DIなどの景気指標との間に明確な相関関係があるか、今後しっかりと検証する必要がある。もし明確な相関関係があると分かれば、TDB景気動向調査の結果を基に様々なテキストデータ（SNSなど）から景気DIを算出できるようになるだろう。「企業の声」のテキストマイニング分析は、地域別、業種別、時系列別での、各種DIの分析をサポートする指標など、さまざまな方面での利用が考えられる。

参考文献

樋口耕一（2014）『社会調査のための計量テキスト分析——内容分析の継承と発展を目指して』ナカニシヤ出版

山澤成康（2018）『計量テキスト分析による景気判断—コーディングルールや主成分を使った時系列分析—』 ESRI Discussion Paper Series No.345

【 問い合わせ先 】

株式会社帝国データバンク データソリューション企画部 産業データ分析課

杉原 翔太

03-5775-3163 keiki@mail.tdb.co.jp

帝国データバンクで毎月実施している TDB 景気動向調査にご協力いただける企業さまは、こちらから登録できます（スマートフォン等をお使いの方はQRコードからも可能です）。
<https://www.tdb-di.com/ent/rent.html>



当レポートの著作権は株式会社帝国データバンクに帰属します。著作権法の範囲内でご利用いただき、私的利用を超えた複製および転載を固く禁じます。