

IV 区内事業所と区民の自由記入回答データの計量テキスト分析

1 本章の目的

本章では、区内事業所を対象とした「港区の飲食業・小売業の実態と事業展開の方向性に関する調査」及び区民を対象とした「港区における買い物環境・食事行動に関する実態調査」の自由記入回答のテキストデータを用いて、区内事業所及び区民の意見の傾向を抽出するため計量テキスト分析を行う。

調査票における自由記入欄への回答は、回答者が比較的強い意見を持つ場合に記入されていることが多いと考えられる。それらの意見は有効な施策を検討するための基礎資料となる可能性があると考えられるが、これまでの港区の調査では、自由記述の回答内容を目視で解釈するのみの利用にとどまっていることが多かった。本章では、調査票における自由記入欄の回答の傾向をテキストデータとして整理し可視化することを試みる。

(本章で明らかにすること)

- 区内事業所が実施意思のある区民向け買い物支援サービスについてのテキストの特徴
- 区民が希望する「その他」の買い物支援サービスについてのテキストの特徴

(節の構成)

- 2 分析に用いるテキスト回答データ
- 3 分析の方法と基本属性データ
- 4 事業所調査のテキストデータの分析
- 5 区民調査のテキストデータの分析
- 6 テキスト分析の結果の概要

以下では、第2節で分析に用いる調査票の設問を紹介し、第3節でテキスト分析の方法を概説したのちに、第4節で事業所調査データ、第5節で区民調査データのテキスト分析を行う。第6節では、テキスト分析で示された結果の概要を示していく。

2 分析に用いるテキスト回答データ

本節では、分析に用いる調査票の自由記入式の設問の内容をみていく。

(1) 事業所調査のテキストデータ

事業所調査のテキストデータには、以下の設問を用いる。

図表4-1 事業所調査の設問

問15. 今後、区が区民の利便性向上に向けて、区民への買い物支援サービスに取り組む区内事業者（補助等の一定の支援を実施する）としたら、貴事業所は新たに取り組みたいと思いますか。取り組みたいと思う対象、形態について当てはまるものに○をつけ、その内容を以下の欄にご自由にご記入ください。（対象・形態の○はいくつでも）

取り組む意思	ア. ある	イ. ない
取り組む意思のある方		
対象	a. 子ども b. 若年層 c. 壮年層 d. 高齢層 e. 全年齢層 f. その他 ()	
形態	1. 店舗販売 2. デリバリー 3. 移動店舗販売 4. インターネット販売業者を通じた販売 5. その他 ()	
内容	(ご自由にお書きください):	

事業所調査の問15では、区民向け買い物支援サービスについて、実施意思を尋ねている。選択式の設問として取組実施意思の有無、実施意思のあるサービスの対象とする年齢層及び販売形態があり、自由記入式の設問で具体的なサービスの内容を尋ねている。同自由記入テキストデータを分析に用いる。

図表4-2 事業所調査の設問の単純集計

取組の実施意思(n=692)		対象(n=311)		形態(n=288)	
ある	45.66	子ども	10.61	店舗販売	75.69
ない	54.34	若年層	22.51	デリバリー	27.43
		壮年層	27.33	移動店舗販売	6.6
		高齢層	18.33	インターネット販売	27.78
		全年齢層	54.98	その他	2.43
		その他	1.29		
テキスト回答数					
全体				84	
営業年数20年以上				37	
延床面積第二・五分位以下				38	

図表4-2は、事業所調査の問15の単純集計の結果を示している。45.66%の事業所が区民向け買い物支援サービスへの実施意思があり、対象とする年齢層は「全年齢層」が54.98%と最大の値を示し、販売形態は「店舗販売」が75.69%と最大の値を示している。

分析に用いる自由記入式のサービスの内容については84件のテキスト回答があった（「なし」等を除く）。なお、第4節以降でサンプルを制限した図表を示すため、営業年数と延床面積を限定した場合のテキスト回答数も示している。

(2) 区民調査のテキストデータ

区民調査のテキストデータは、以下の設問を用いる。

図表4-3 区民調査の設問

問9. 区における行政サービスとして、導入してほしい買い物支援サービスはございますか。次の1～5のあてはまるものを選び（〇はいくつでも）、具体的な内容を空欄にお書きください。

	内容
1. 特にない	
2. 移動販売車の導入 (導入希望の場所、品類等を右の欄にご記入ください)	
3. 買い物代行サービスの充実 (充実を希望する品類等を右の欄にご記入ください)	
4. スーパーの配送サービス等に関する情報提供 (希望する情報提供方法や品類等を右の欄にご記入ください)	
5. その他 (具体的な内容を右の欄にご記入ください)	

区民調査の問9では、区民が導入を希望する買い物支援サービスについて尋ねている。選択式の設問の「その他」欄に自由記入式の設問があり、同項目では、他の選択肢以外で希望するサービスの具体的な内容を自由記入式で尋ねている。同自由記入テキストデータを分析に用いる。

図表4-4 区民調査の設問の単純集計

導入してほしい買い物支援サービス(n=944)	
特にない	70.02
移動販売車の導入	8.90
買い物代行サービスの充実	6.36
スーパーの配送サービス等に関する情報提供	11.55
その他	10.49
テキスト回答数	
全体	91
50歳未満	46
所得第三・五分位以下	53
18歳以下の子どもがいる世帯	28

図表4-4は、区民調査の問9の単純集計の結果を示している。「特にない」の回答が最大の70.02%を示しており、「その他」は10.49%となっている。

分析に用いる自由記入式の、その他に希望する買い物支援サービスの内容については、91件のテキスト回答があった（「なし」等を除く）。なお、第4節以降でサンプルを制限した図表を示すため、年齢、所得及び世帯構成を限定した場合のテキスト回答数も示している。

3 分析の方法と基本属性データ

本節ではテキスト分析の概要を説明したのちに、サンプルの制限に用いる事業所属性、区民属性の情報を説明していく。

(1) テキストデータの分析方法

テキストデータの分析には、KH Coder 3 (<https://kncoder.net/>) を利用する。分析においては、テ

キストデータを単語として意味を持つ最小単位まで分解をする形態素解析を行い、各単語の出現頻度をみていく。さらに、出現パターンが類似する単語同士を線で繋いだ図として表現した図表（共起ネットワーク）をみていく（樋口・中村・周 2023）。

なお、補足的に行った事業所調査と区民調査における「区政や調査に対する考えや自由意見」のテキストデータのネットワーク分析の結果については、本章末尾の資料アを参照されたい。

（2）サンプルの制限に用いる属性変数

次節では、テキストデータの出現単語についての共起ネットワークの図から単語同士のつながりをみていく。その際に、全体サンプルの結果とサンプルを制限した共起ネットワークの結果を示していく。サンプルの制限に利用する基本属性の変数の内容は次のとおりである。

図表4-5 分析に用いる属性変数の構築方法

変数名	変数の値
(事業所属性の変数)	
営業年数	Q28 20年以上=1、20年未満=0
延床面積	Q15 延床面積の実数値を5分位化(小さい順から2割程度ずつに分位化) Q16 第二・五分位以下(下位40%)=1、それ以外を0
(区民属性の変数)	
年齢	Q1 50歳未満=1、50歳以上=0
所得	Q15 2021年の世帯所得を世帯人数の平方根で除した値を5分位化(小さい順から2割程度ずつに分位化) Q16 第三・五分位以下(下位60%)=1、第四・五分位以上=0
世帯構成	Q18 同居する18歳以下の自身の子どもがいる=1、それ以外の世帯=0

サンプルの制限に用いる事業所の基本属性は、営業年数及び延床面積である。このうち、全体のサンプルに加え、営業年数20年以上、延床面積第二・五分位以下のサンプルに制限した結果を示していく。

区民の基本属性は、年齢、所得、世帯構成である。このうち、全体のサンプルに加え、年齢50歳未満、18歳以下の子どもあり世帯及び所得第三・五分位以下にサンプルを制限した結果を示していく。

4 事業所調査のテキストデータの分析

本節の分析では、事業所調査のテキストデータにおける頻出単語の傾向をみたのちに、共起ネットワークから単語同士のつながりをみていく。

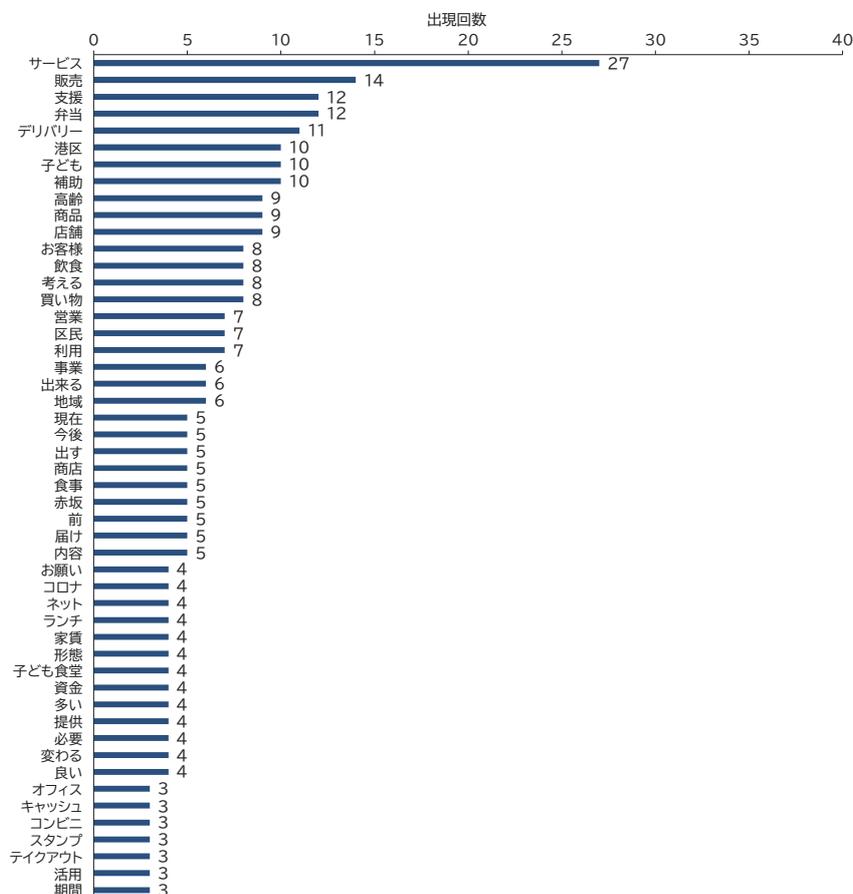
なお、基本属性ごとの出現単語の違いを可視化するため、補足的に行った対応分析の図表については本章末尾の資料イを参照されたい。

（1）頻出単語

「サービス」の出現回数が27回と最も多く、「販売」「支援」「弁当」「デリバリー」は10回を超える出現回数となっている

分析に用いる事業所のテキストデータでの頻出単語は次のとおりである。

図表4-6 事業所調査のテキストデータにおける頻出単語（上位50単語）



図表4-6は、形態素解析によって事業所調査のテキストデータから抽出した単語を出現する回数が多い順に並べたものである。出現回数は「サービス」が27回と最も多く、次いで「販売」「支援」「弁当」「デリバリー」「港区」「子ども」「補助」の単語が多くなっており、いずれも10回を超える出現回数となっている。

実際の回答テキスト内容を見ると、買い物サービス券や飲食サービス券、割引サービス券の発行、宅配サービスの実施などで「サービス」の出現が多くなっている。また、弁当販売や資金面の補助といった内容で、「弁当」「販売」「補助」の出現がみられる。

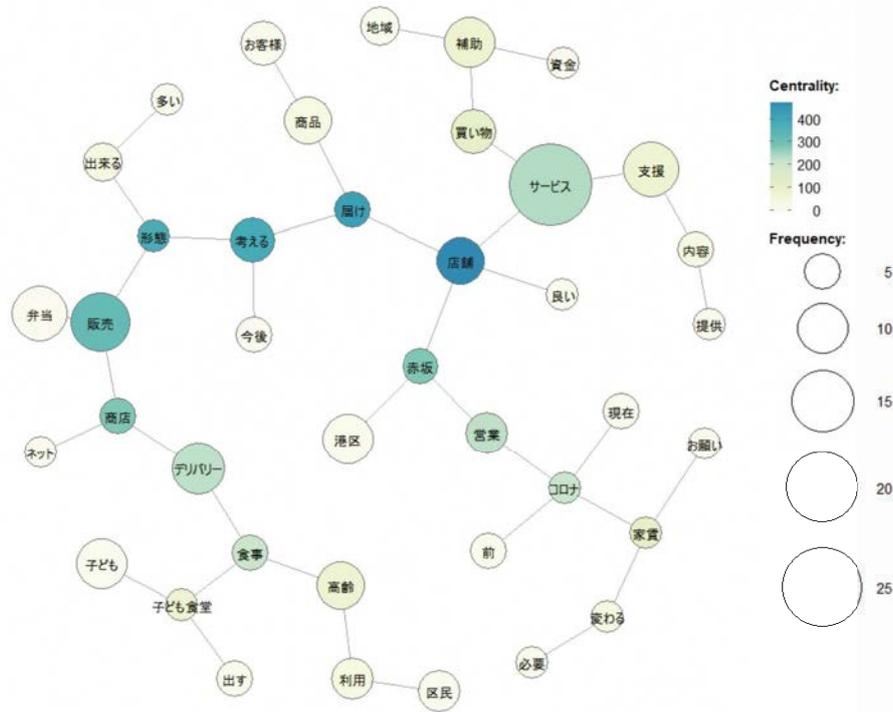
以下の共起ネットワーク分析では、出現回数が多い語のうち、施策のニーズの把握につながると考えられる「販売」と「デリバリー」に着目していく。

(2) 共起ネットワーク

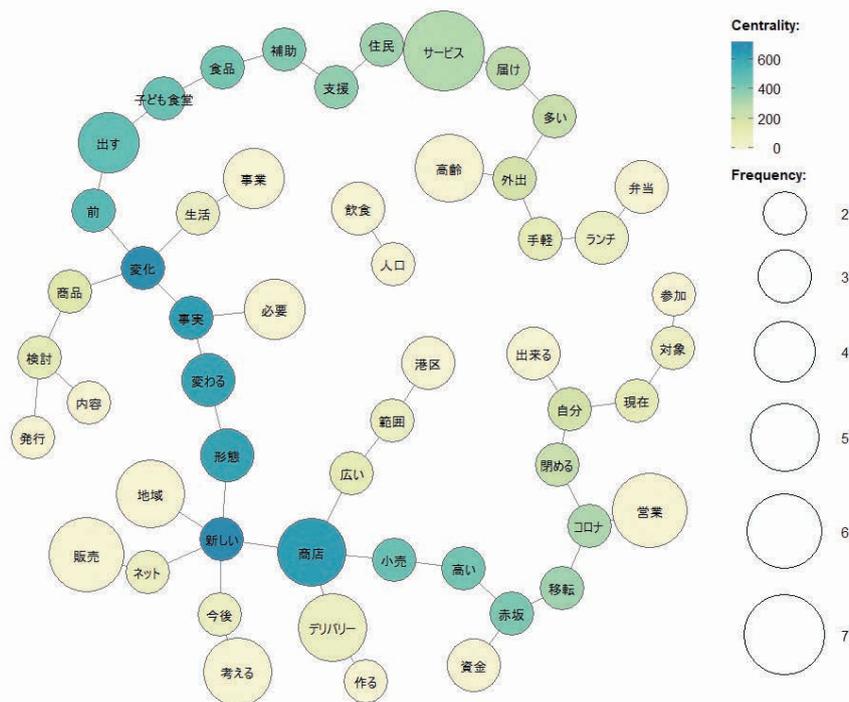
- ・「販売」は「弁当」「商店」「形態」との共起関係がみられ、「デリバリー」は「商店」「食事」との共起関係がみられる
- ・営業年数が20年以上の事業所サンプルに限定すると、「販売」と「ネット」とが共起関係にある

続いて、テキストデータにおける出現単語同士のつながりを示す共起ネットワークの結果を示していく。共起ネットワークの図の円の大きさは単語の出現頻度を表し、単語同士のつながりは線で示される。以下では、全サンプルの結果を示したのちに、サンプルに制限した結果を示し、「販売」と「デリバリー」の出現パターンに着目していく。

図表4-7 実施意思のある買い物支援サービスの記述における出現単語の共起ネットワーク（全体）



図表4-8 実施意思のある買い物支援サービスの記述における出現単語の共起ネットワーク（営業年数20年以上）



「ネット」が示されている。実際の回答テキスト内容を見ると、インターネットによる商品の販売についての意見が挙げられている。

5 区民調査のテキストデータの分析

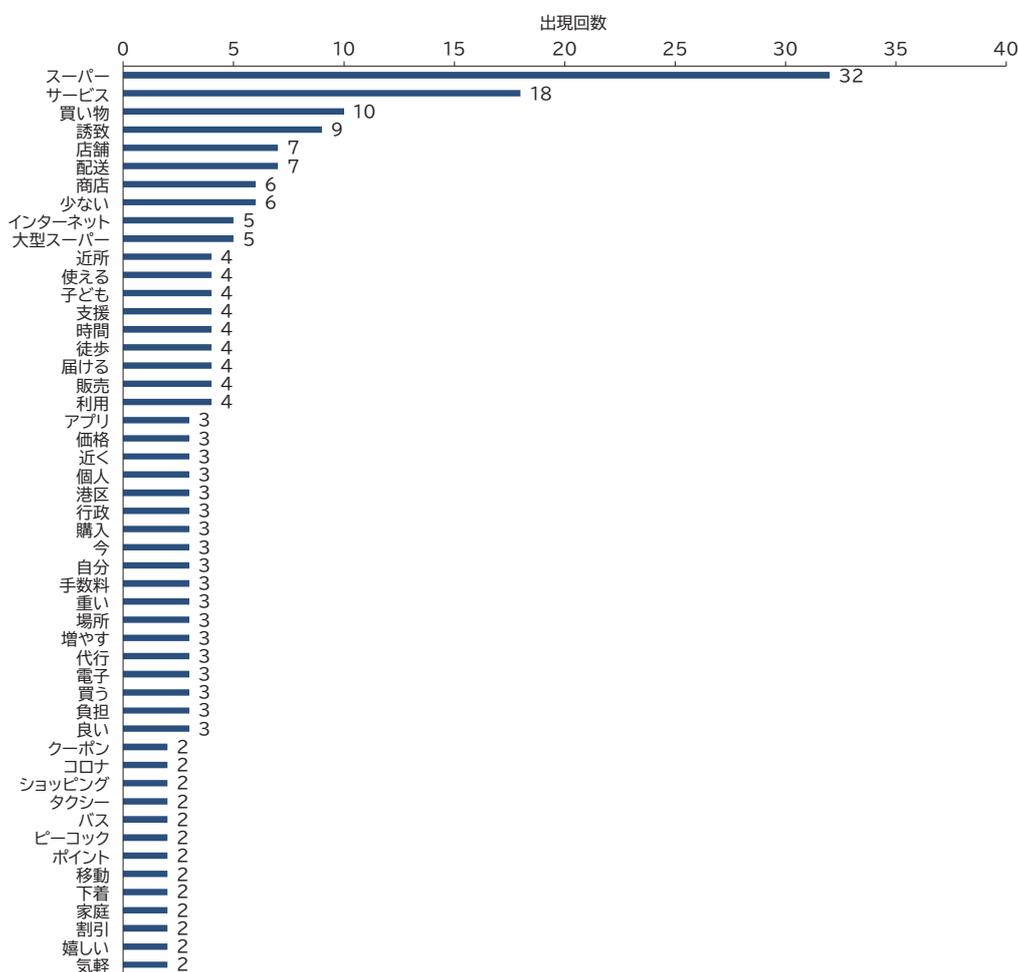
本節の分析では、区民調査のテキストデータにおける頻出単語の傾向をみたのちに、共起ネットワークから単語同士のつながりをみていく。

(1) 頻出単語

「スーパー」の出現回数が32回と最も多く、「サービス」「買い物」「誘致」は9回以上の出現回数となっている

分析に用いる区民のテキストデータでの頻出単語は次のとおりである。

図表4-10 区民調査のテキストデータにおける頻出単語（上位50単語）



図表4-10は、形態素解析によって区民調査のテキストデータから抽出した単語を出現する回数が多い順に並べたものである。出現回数は「スーパー」が32回と最も多く（「大型スーパー」を合わせると37回）、次いで「サービス」が18回、「買い物」が10回、「誘致」が9回となっている。「店舗」「配送」「商店」「少ない」などの語も出現が多くなっている。

全体で示される出現の有無や単語同士のつながりの傾向は、サンプルを限定すると変化する傾向にある。所得が第三・五分位以下の区民（下位60%）のサンプルに限定した図表4-14では、「大型スーパー」と共起関係にある語として「徒歩」「使える」「良い」が示されている。実際の回答テキスト内容を見ると、徒歩圏内の大型スーパーの誘致についての意見が挙げられている。

6 テキスト分析の結果の概要

ア 事業所調査データの結果

- ・区民向けの買い物支援サービスへの実施意思がある事業所は45.66%あり、実施意思がある買い物支援サービスの対象とする年齢層は「全年齢層」の54.98%が最大、実施形態は「店舗販売」の75.69%が最大であった。
- ・具体的に取り組みたい買い物支援サービスの自由回答データについてのテキスト分析を行ったところ、「サービス」「販売」「支援」「弁当」「デリバリー」などの単語の出現回数が多い傾向にあった。
- ・施策のニーズの把握につながると考えられる「販売」「デリバリー」の共起ネットワークにおける出現パターンに焦点を当てるため共起ネットワークを作成すると、「販売」と「弁当」「商店」「形態」とのつながりがみられた。「デリバリー」は、「商店」「食事」とのつながりがみられた。
- ・事業所属性に基づいてサンプルを制限すると、全体とは一部異なる単語のつながりみられた。例えば、営業年数が20年以上の事業所に限定すると、「販売」は「ネット」とのつながりがみられた。

イ 区民調査データの結果

- ・導入を希望する買い物支援サービスが「特にない」とした区民は70.02%おり、「スーパーの配送サービス等に関する情報提供」「移動販売車の導入」「買い物代行サービスの充実」を選択した割合はいずれも1割程度となっている。
- ・希望する「その他」の具体的な買い物支援サービスの自由回答データについてのテキスト分析を行ったところ、「スーパー」「サービス」「買い物」「誘致」などの単語の出現回数が多い傾向にあった。
- ・施策のニーズの把握につながると考えられる「スーパー」と「インターネット」の共起ネットワークにおける出現パターンに焦点を当てるため共起ネットワークを作成すると、「スーパー」と「誘致」とのつながりがみられた。「インターネット」は、「販売」「利用」「負担」とのつながりがみられた。
- ・区民属性に基づいてサンプルを制限すると、全体とは一部異なる単語のつながりみられた。例えば、18歳以下の子どもがいる区民に限定すると、「インターネット」は出現せず、所得が第三・五分位以下（下位60%）の区民に限定すると、「スーパー」は出現せず、「大型スーパー」のみの出現がみられた。

引用文献・引用資料

樋口耕一（2019）「計量テキスト分析における対応分析の活用——同時布置の仕組みと読み取り方を中心に」『コンピューター&エデュケーション』Vol.47、pp.18-24

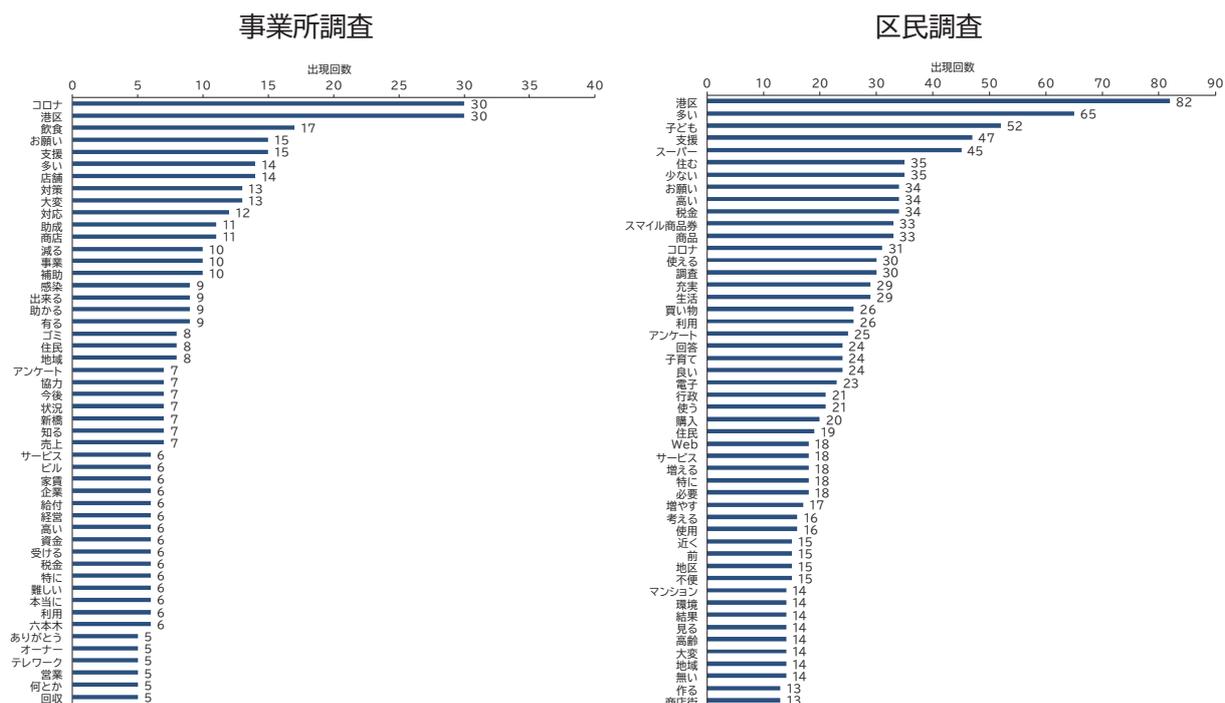
樋口耕一・中村康則・周景龍（2023）『動かして学ぶ！はじめてのテキストマイニング——フリー・ソフトウェアを用いた自由記述の計量テキスト分析』ナカニシヤ出版

資料ア 自由意見のテキストデータの分析

事業所調査と区民調査の調査票の末尾には、「本調査や港区政について何かご意見等ございましたら、ご自由にお書きください」という設問がある。事業所調査では407件、区民調査で142件の自由記入があった（「なし」等を除く）。

同設問のテキストデータについての頻出単語と共起ネットワークは以下のとおりである。

図表4-15 区政や調査に対する考えのテキストデータでの頻出単語（上位50単語）



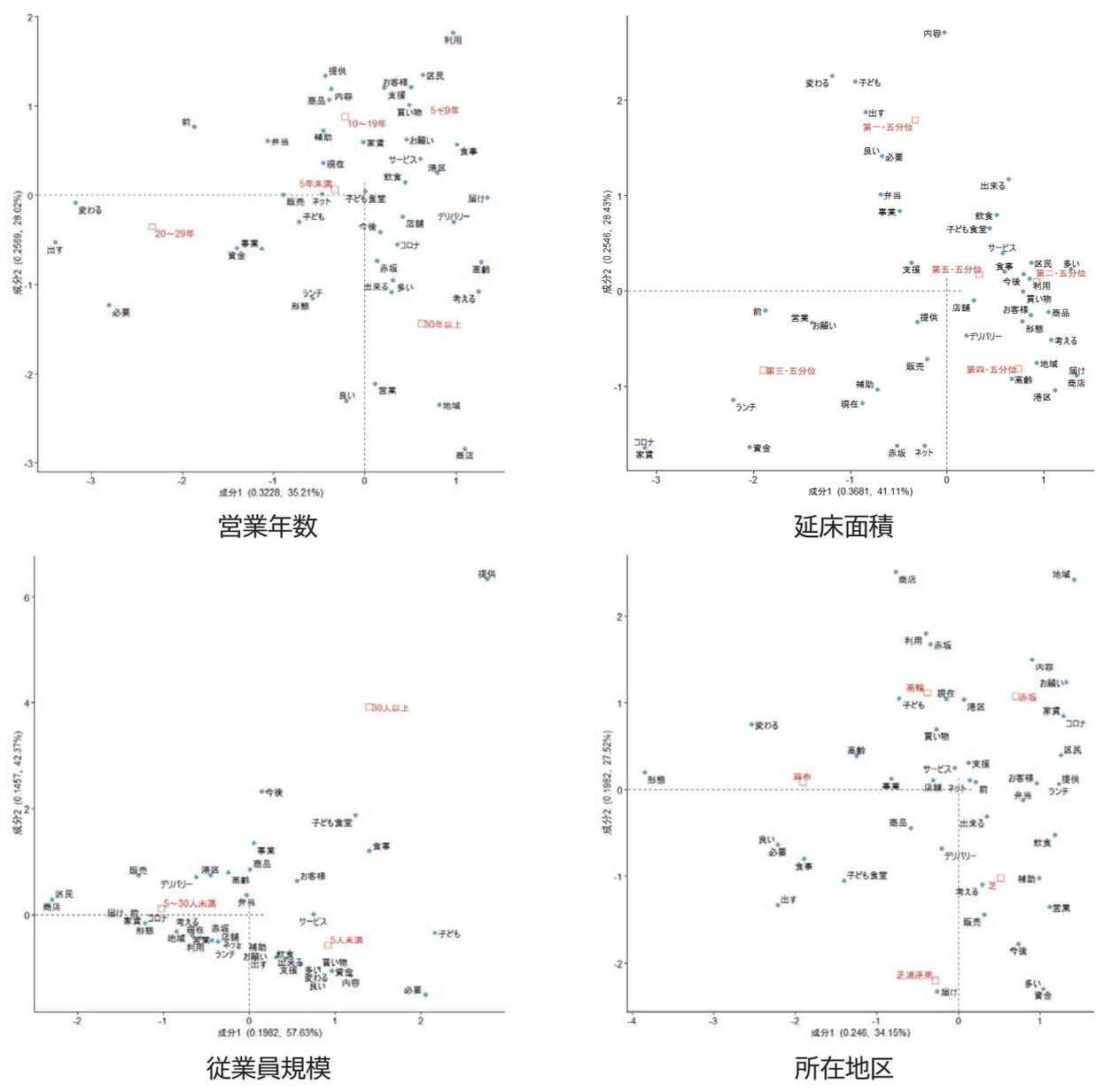
とができる（樋口 2019）。対応分析における同時布置図から読み取ることができる主な傾向は、次のとおりである。

- ①各群の布置：近接する群同士の出現単語の類似性
- ②原点からみた方向：各群に特徴的な出現単語

区民属性の年齢を例にあげると、①については、対応分析の同時布置図において距離が近い年齢層は自由記入の回答の出現単語が類似していることが示される。②については、対応分析の同時布置図において、原点からみて各年齢層の方向に布置する単語が各年齢層の回答に多く出現した単語であることが示される。

■事業所調査のテキストデータの対応分析

図表4-18 対応分析による単語と事業所属性の同時布置



図表4-18は、対応分析を基にした、出現単語と事業所属性の同時布置を示している。対応分析を行った事業所属性は、営業年数、延床面積、従業員規模及び所在地区である。

①群間の布置

図表内における各群の位置から出現単語の類似性をみていく。各基本属性について、近接する群は以下のとおりである。

- 営業年数：「5年未満」「5～9年」「10～19年」
- 延床面積：「第二・五分位」「第五・五分位」
- 従業員規模：なし
- 所在地区：「高輪」「赤坂」

これらの各群は、それぞれが比較的に近い位置にあり、出現単語が類似している傾向がうかがえる。

②原点からみた方向

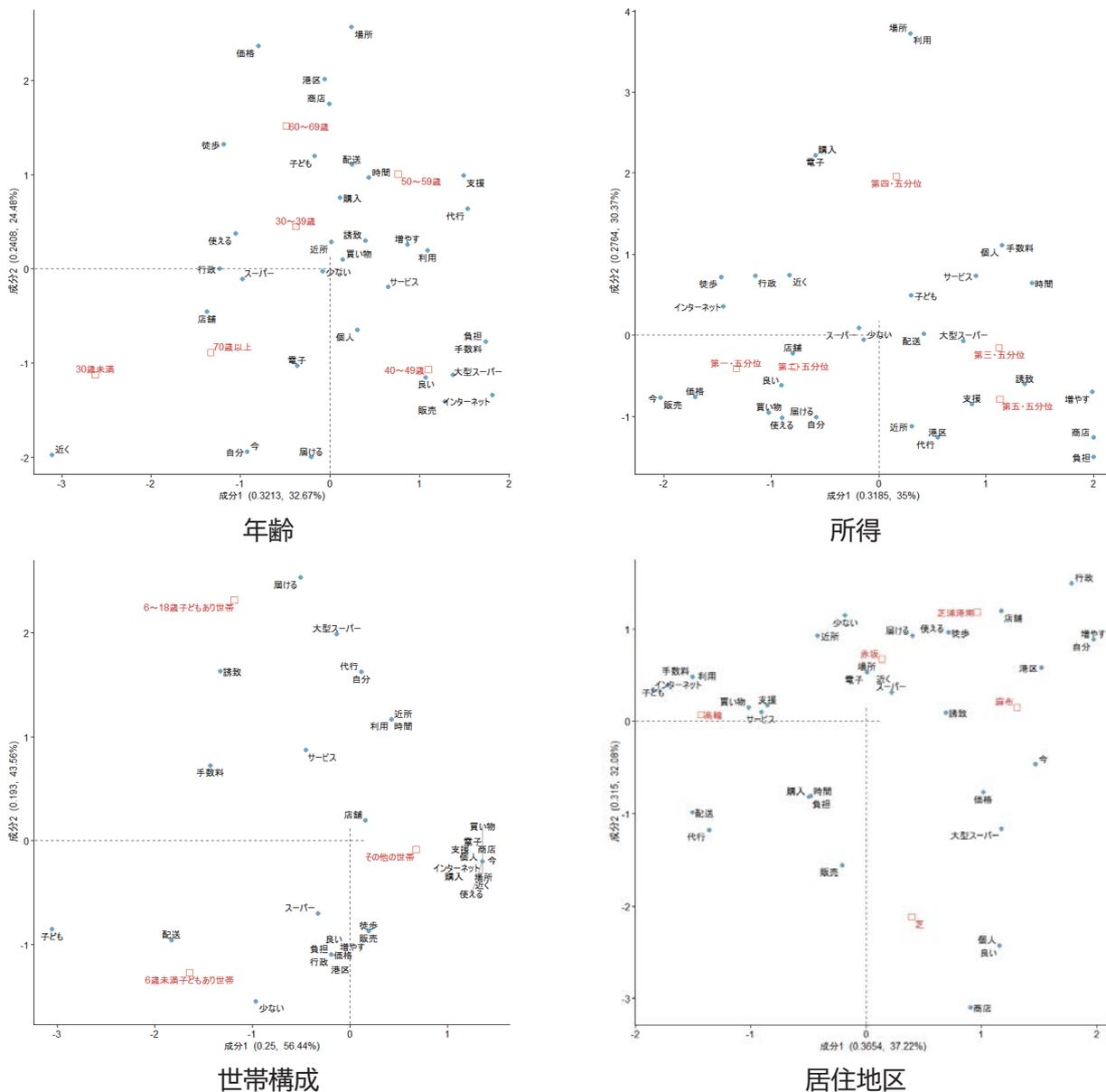
原点からみた各群の方向にある単語から、次の単語の出現頻度が高い傾向が示されている（主な語を列挙）。

- 営業年数
 - 5年未満：なし
 - 5～9年：「お願い」「サービス」
 - 10～19年：「内容」「提供」「商品」
 - 20～29年：「出す」
 - 30年以上：「今後」「多い」「地域」「商店」「出来る」
- 延床面積
 - 第一・五分位：なし
 - 第二・五分位：「利用」「今後」「多い」
 - 第三・五分位：「コロナ」「家賃」「ランチ」
 - 第四・五分位：「店舗」「高輪」「港区」「デリバリー」
 - 第五・五分位：「サービス」「食事」「区民」
- 従業員規模
 - 5人未満：「必要」
 - 5～30人：「区民」「商店」
 - 30人以上：「提供」
- 所在地区
 - 芝：「今後」「多い」「資金」「考える」
 - 麻布：「事業」「形態」「店舗」
 - 赤坂：「ネット」「支援」「内容」「地域」
 - 高輪：「サービス」「買い物」「利用」「赤坂」「現在」「子ども」「商店」
 - 芝浦港南：「届け」「デリバリー」

なお、隣り合う2つの群の間に布置する単語（例えば、営業年数の場合、「5～9年」と「10～19年」の間にある単語は「買い物」）は、両方の群の回答に比較的多く出現した単語となっている。

■区民調査のテキストデータの対応分析

図表4-19 対応分析による単語と区民属性の同時布置



図表4-19は、対応分析を基にした、出現単語と区民属性の同時布置を示している。対応分析を行った区民属性は、年齢、所得、世帯構成、居住地区である。

①群間の布置

図表内における各群の位置から出現単語の類似性をみていく。各基本属性について、近接する群は以下のとおりである。

- 年齢：「30～39歳」「50～59歳」「60～69歳」
- 所得：「第一・五分位」と「第二・五分位」、「第三・五分位」と「第五・五分位」
- 世帯構成：なし

➤ 居住地区：「麻布」「赤坂」「芝浦港南」

これらの各群は、それぞれが比較的に近い位置にあり、出現単語が類似している傾向がうかがえる。

②原点からみた方向

原点からみた各群の方向にある単語から、次の単語の出現頻度が高い傾向が示されている（主な語を列挙）。

➤ 年齢

30歳未満：「店舗」「少ない」

30～39歳：「徒歩」

40～49歳：「良い」「販売」

50～59歳：「買い物」「誘致」「近所」

60～69歳：「近所」「価格」「子ども」

70歳以上：なし

➤ 所得

第一・五分位：「店舗」「価格」「今」

第二・五分位：「少ない」「良い」

第三・五分位：「大型スーパー」「配送」

第四・五分位：「利用」「場所」

第五・五分位：「支援」「商店」「負担」

➤ 世帯構成

6歳未満の子どもあり：なし

6～18歳未満の子どもあり：「サービス」「誘致」

その他：「個人」「インターネット」「支援」

➤ 居住地区

芝：「商店」

麻布：「誘致」

赤坂：「近く」「場所」「スーパー」

高輪：「サービス」「買い物」「支援」

芝浦港南：「スーパー」「徒歩」「店舗」

なお、隣り合う2つの群の間に布置する単語は、両方の群の回答に比較的多く出現した単語となっている。