

ソーシャルメディア上での「住みたい地域」イメージとその構成要素の
抽出方法に関する基礎的研究

～居住地選好に関する投稿ログと投稿者居住地データのテキストマイニングを通じて～

2012年度 株式会社数理システム 学生研究奨励賞 応募論文

東京大学大学院

工学系研究科 都市工学専攻

佐藤 遼

1. 研究の構成

研究の背景

地域活性化を進める上での地域イメージの
重要性

研究の目的

地域イメージとその構成要素の定量的な抽
出方法の検討

研究の方法

地域イメージの抽出方法

抽出用のキーワード

データの入手方法と種類

イメージの評価の指標

分析の手順

分析結果の解釈の視点

既往研究との比較、本研究の新規性

2. 全体的な傾向についての分析

住みたい地域ランキング

住みたくない地域ランキング

発言者率(イメージの強さ)の分布

イメージの良い地域と悪い地域の抽出

季節変動による影響

住みたい地域イメージとの関連ワード

居住地域の分布

3. 特定の地域の傾向についての分析

「東京」のイメージ構成要素とイメージ保有者の居住地域

「関東」のイメージ構成要素とイメージ保有者の居住地域

「京都」のイメージ構成要素とイメージ保有者の居住地域

「北海道」のイメージ構成要素とイメージ保有者の居住地域

「吉祥寺」のイメージ構成要素とイメージ保有者の居住地域

4. 結果のまとめ

本研究にて得られた知見

今後の課題

主要参考文献

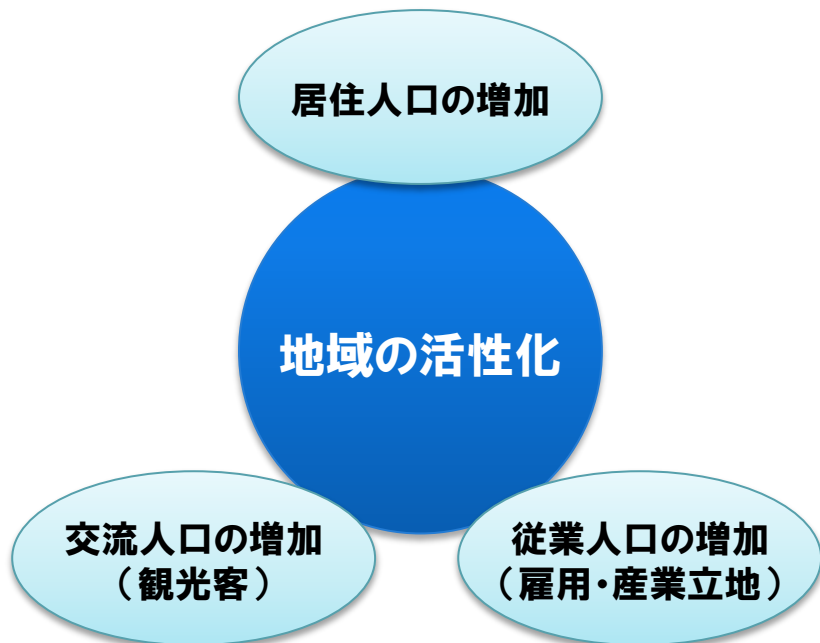


1. 研究の構成

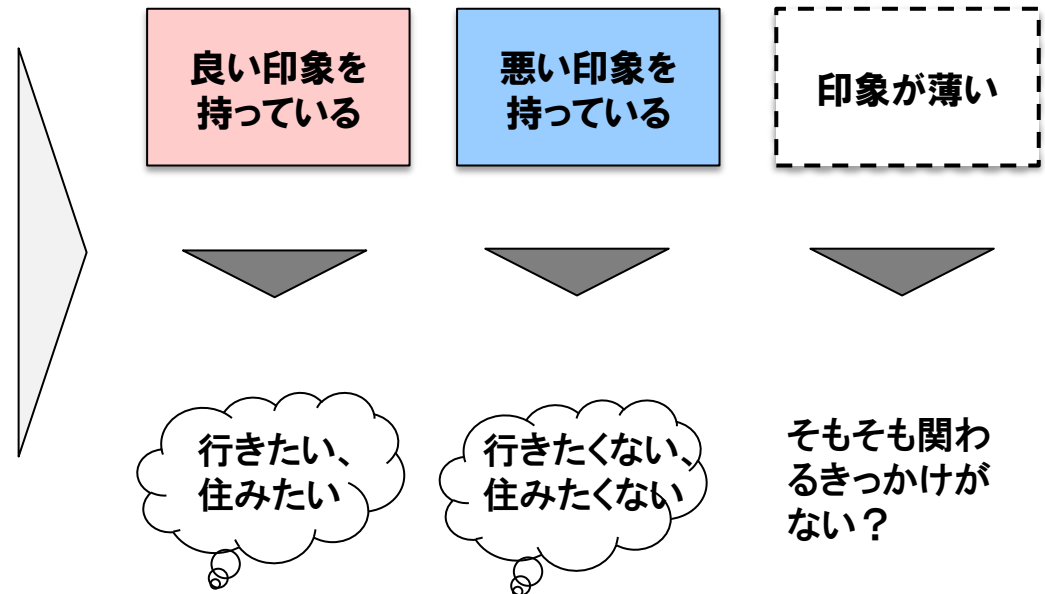
研究の背景:地域活性化を進める上での地域イメージの重要性

- 都市・地域の活性化を進める上で、その地域が持つイメージは、大きな影響力を持つ要素である。
- 人が集まるためには、その地域が持つ魅力が、良いイメージとして人々に伝わっている必要がある。

地域活性化のためには、地域に人と資本を呼ぶことが必要

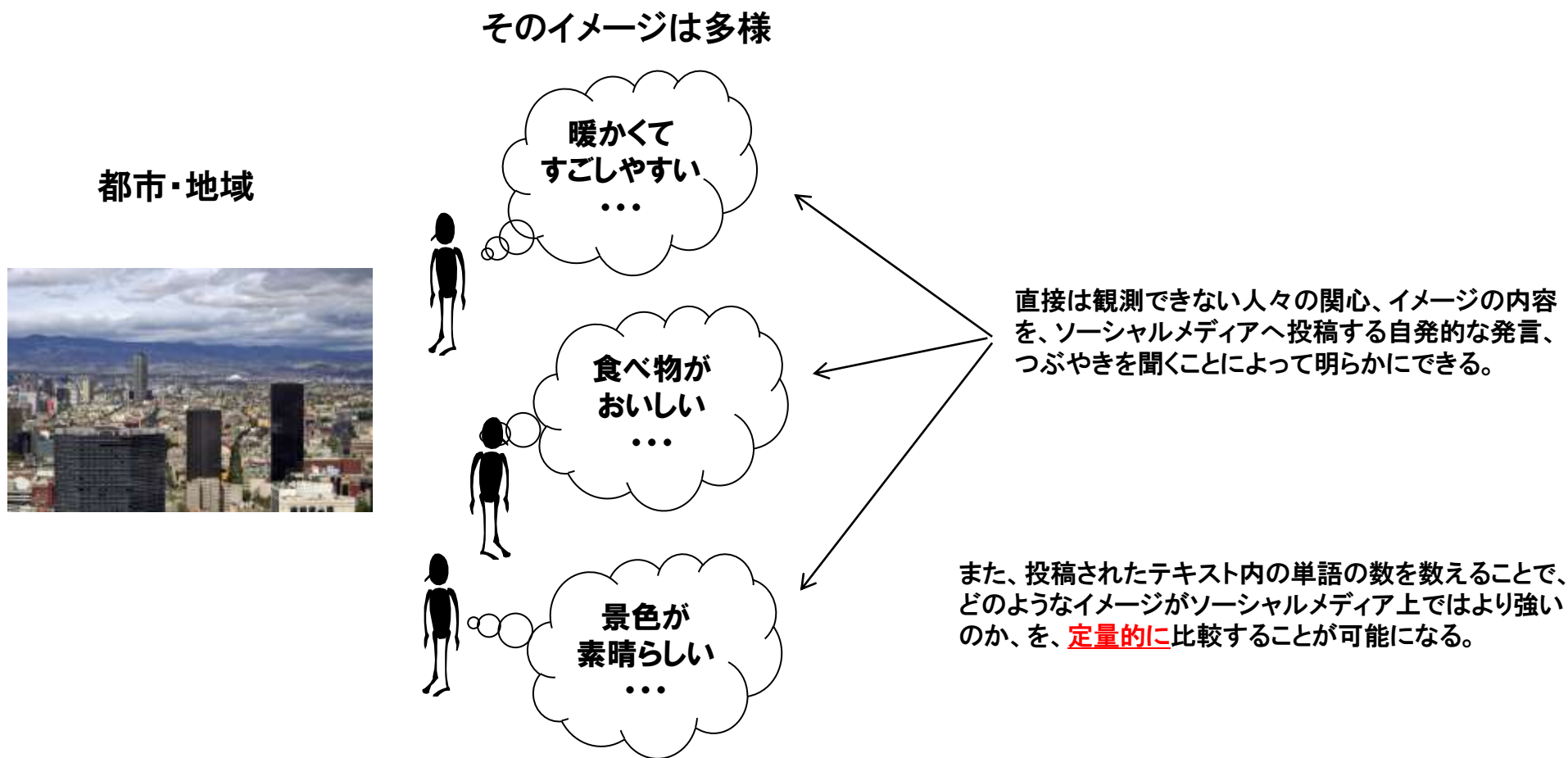


関心を持った地域、良い印象を持った地域に、人々は集まると考えられる



研究の目的: 地域イメージとその構成要素の定量的な抽出方法の検討

- 地域に対するイメージは一律ではなく、個々人が知る情報に応じて、多様に異なっていると考えられる。
- 本研究では、そのような多様な各地域のイメージと、その構成要素を、主要なソーシャルメディアの1つであるTwitterへの投稿ログデータを用いて、定量的に抽出する方法を検討した。



研究の方法:地域イメージの抽出方法

- Twitterの投稿ログから、特定のキーワードを含む発言を抜き出して、人々の自発的な発言からイメージを探った。
- 具体的には、“〇〇に住みたい／住みたくない”という発言を抜き出し、“〇〇”の部分のテキストと、その前後のテキスト、および発言者の居住地情報を分析することによって、イメージの対象地域とその関連情報を抽出した。

元のデータ形式(例)

日付	ユーザ名	発言(Tweet)	居住地
2012年11月12日	qchannel31	今年こそライブ見に行きたい！あー東京に住みたい。	福井県の山の中

テキストマイニングの対象 (※Text mining Studioの分かち書き機能、単語頻度分析機能、および辞書機能を使用)

①:「に住みたい」というキーワードの10文字前までのテキスト → 品詞が「固有名詞地名」の単語を抽出
(「住みたい地域」を意味する)

②:発言(Tweet)全体 → 品詞が「名詞(一般、固有名詞)」の単語を抽出
(「住みたい地域」との関連ワードを意味する)

③:発言者の居住地 → 品詞が「固有名詞地名」の単語を抽出
(「発言者の居住地」を意味する)

整形したデータ(例)

日付	ユーザ名	1. 住みたい地域	2. 住みたい地域との関連ワード	3. 居住地
2012年11月12日	qchannel31	東京	今年、ライブ、東京	福井

研究の方法:抽出用の検索キーワード

- 人々が「住みたい」と思う地域は「良い」イメージの地域、「住みたくない」と思う地域は「悪い」イメージの地域、という仮定をおき、データを作成した。
- より詳細には、下記の条件の検索キーワードにて、Twitterの発言の抽出を行った。

	「良い」地域イメージ	「悪い」地域イメージ
発言抽出用 検索キーワード	～に住みたい というフレーズを含む文	に住めない には住めない に住みたくない には住みたくない に 住みたくはない には住みたくはない に住みたいと思わない には 住みたいと思わない に住みたいとは思わない には住みたいとは思 わない のいずれかのフレーズを含む文

※以下の2つのキーワードのいずれかを含む文は、抽出対象から除外した (NGキーワード)

「RT」……ユーザの自発的な発言のみを対象とするため (他人の意見の引用はカウントしない)

※RTは、リツイート(他者の発言の引用)が含まれている発言に含まれているキーワード

「http」……Twitter内で完結している発言のみを対象とするため (リンク先の文章を含めて分析することが技術的に難しい)

※httpは、他サイトへのリンクが貼られている発言に含まれているキーワード

※集計は、「発言者単位」で行った

同じ発言者が複数回発言した場合は、最初の1回のみをカウントした。地域イメージは「人」単位で形成されているため、集計も「人」単位で行うのが良いと判断した。

また、特定の1人が大量に発言している場合(スパム、Botなど)の影響を軽減する効果もある。

研究の方法:データの入手方法と種類

- Twitterのデータ入手には、ホットリンク社のWebサービス「クチコミ@係長」を利用した。前スライドのキーワードにてホットリンク社のデータベースを検索し、データを抽出した。
- データ作成上の仕様により、目的に応じて2種類のデータを使い分けた。

	データ取得期間	データ特性	使用目的
データ1	2011年10月1日 ～2012年9月30日	・サンプル数が多い 住みたい地域:約6万件 住みたくない地域:約6千件 ・居住地属性なし	居住地以外の各分析において使用
データ2	2012年10月28日～ 2012年11月6日	・サンプル数が少ない 住みたい地域:約1万件 住みたくない地域:約1千件 →住みたくない地域については、分析を断念 ・居住地属性あり	居住地の分析においてのみ使用

※ホットリンク社、および「クチコミ@係長」の詳細については、下記HPを参照ください。

<http://www.hottolink.co.jp/>

研究の方法:イメージの評価の指標

- イメージの評価の指標として、「発言者率」と「好印象率」の2種類の指標を設けた。
- 「発言者率」は地域イメージの強さを測る指標として、「好印象率」はイメージの良さを測る指標として用いた。

指標名	意味	計算式
発言者率	イメージの強さの指標 (0%~100%) ※ 実際には最大でも数%程度	キーワードAの発言者率 = キーワードAの発言者数 / 全キーワードの発言者数 例: 発言の収集対象者が全部で100人いたときに、うち4人が「横須賀に住みたい」と発言している → 横須賀の発言者率 = $4/100 = 4\%$
好印象率	イメージの良さの指標 (0%~100%)	キーワードAの好印象率 = キーワードAに「住みたい」発言者率 / (キーワードAに「住みたい」発言者率 + キーワードAに「住みたくない」発言者率) 例:「横須賀に住みたい」人の発言者率が4%、 「横須賀に住みたくない」人の発言者率が1% → 横須賀の好印象率 = $4\% / (4\% + 1\%) = 80\%$ (高い値)

- ・はじめに、どのような地域が抽出されるのか、また属性の分布はどのようになっているのか、という全体的な傾向について、分析を行った。
- ・そののち、いくつかの特徴的な地域を対象として、関連ワードと居住地に関する詳細な分析を実施した。

全体的な傾向についての分析

- ・住みたい地域ランキングの作成(P.5①)
(イメージが良い地域の抽出)
- ・住みたくない地域ランキングの作成(P.5①)
(イメージが悪い地域の抽出)
- ・各地域のイメージの良さと強さの比較(P.5①)
- ・季節変動による影響の検討(P.5①)
- ・住みたい地域イメージとの関連ワードの抽出(P.5②)
- ・地域イメージの保有者の居住地の分布の把握(P.5③)



特定の地域の傾向について分析

- ・地域に特徴的な関連ワードの抽出(P.5①②)
TMSの「特徴語分析」機能を使用した。抽出の尺度としては補完類似度を用いた
- ・対象地域のイメージ保有者の居住地の分布の把握(P.5①③)
(※サンプル数が少ないため、結果の解釈は慎重に行う必要がある)

研究の方法:分析結果の解釈の視点

- 本分析の結果、抽出された地域名、関連ワードについては、多様な解釈が可能である。
- その視点の検討は非常に難しい問題であるが、本研究においては、以下のような視点での結果の解釈を試みた。

視点	具体例	視点の意義
強い地域イメージが構成されている圏域の所在地	国内か海外か？ 大都市圏内か、地方圏内か？	人々の関心が、どの方向に対してより集まっているのかを把握する 現時点で関心を引けていない地域は、今後、一層の努力が求められる。
強い地域イメージが構成されている圏域の単位(広さ)	国単位か？ 都道府県単位か？ 市町村単位か？ あるいは他の圏域か？	人々が、どのレベルの名称で、地域に対してより強い関心を持っているのかを把握する 例えば、観光PRを目的とした広報をする際に、どの地域名を使うのがより印象がよくなるのか、などを検討する材料になる
地域イメージが構成される要素の軸	地域の物的資源をもとに構成されるのか？ 人的資源をもとに構成されるのか？ など	地域イメージの核となる要素が何なのか、についての仮説を提示する
定量的な数値の分布についての解釈	どのような分布で近似されるか？ 現実の世界の、どのようなデータと分布が似ているか？	今後、研究を深めていく上で役立つ可能性がある基礎的な知見を蓄積する

既往研究との比較、本研究の新規性

- ・ 地域イメージについては、民間調査、学術研究などで既に多くの蓄積がある。
- ・ 本研究は、調査媒体としてソーシャルメディアを用いることにより、一定以上の規模の数の回答を得る点と、調査者が設定した質問に対する回答ではなく、人々の自発的な発言の中からイメージを抽出する点に新規性がある。

地域イメージの抽出の手法	特徴
アンケート質問票による調査	<ul style="list-style-type: none">・ 直接、聞きたい内容を質問して調査することができる。反面、対象者の自発的な発言を収集することは難しい。・ 調査コストがかかる。そのため、長期的に、他地域にわたって同じ観点で調査が続けることが難しい
マスメディアにおけるテキスト分析	<ul style="list-style-type: none">・ 調査したい内容の発言が載っているとは限らない。反面、寄稿者の自発的な発言からイメージを読み取ることができる・ 記事を書いた、専門家のイメージを継続的に、詳細に抽出することができる。逆に、一般の人々のイメージを抽出することは難しい
ソーシャルメディアにおけるテキスト分析 (本研究)	<ul style="list-style-type: none">・ 調査したい内容の発言が載っているとは限らない。反面、投稿者の自発的な発言からイメージを読み取ることができる・ 一般の人々のイメージを抽出することができる。ただし、ソーシャルメディアの利用者、かつ発言者に限られる。・ 口語調の会話を載せる媒体であるため、論理的な深い考察に基づく記述が載っている可能性は少ない<ul style="list-style-type: none">→ 地域イメージという直感的な要素の抽出においては、大きな問題にはならないと想定・ 他の調査よりも大量のサンプルデータを、長期にわたって得られる可能性がある



2. 全体的な傾向についての分析

住みたい地域ランキング

- ・はじめに、「～に住みたい」というフレーズの対象テキスト(P.5の①)を分析して、住みたい対象の地域を抽出した。
- ・住みたいという発言者が最も多かった地域は「東京」、次いで「関東」「京都」「大阪」「韓国」の順であった。
- ・上位100件内では、**海外の地域**が22件、**関東圏内の地域**が45件を占めていた。

順位	地域名	発言者数	発言者率
1	東京	3773	6.15%
2	関東	1092	1.78%
3	京都	1025	1.67%
4	大阪	993	1.62%
5	韓国	549	0.90%
6	北海道	540	0.88%
7	関西	516	0.84%
8	名古屋	495	0.81%
9	沖縄	475	0.77%
10	横浜	466	0.76%
11	福岡	326	0.53%
12	千葉	299	0.49%
13	日本	281	0.46%
14	神奈川	267	0.44%
15	ハワイ	260	0.42%
16	埼玉	235	0.38%
17	吉祥寺	227	0.37%
18	神戸	214	0.35%
19	アメリカ	207	0.34%
20	仙台	203	0.33%
21	札幌	201	0.33%
22	池袋	189	0.31%
23	下北沢 (下北沢+下北)	164	0.27%
24	鎌倉	162	0.26%
25	新宿	161	0.26%

順位	地域名	発言者数	発言者率
26	渋谷	134	0.22%
27	イタリア	129	0.21%
28	中野	122	0.20%
29	秋葉原 (秋葉原+秋葉)	121	0.20%
30	ヨーロッパ	120	0.20%
31	広島	113	0.18%
32	高円寺	113	0.18%
33	ドイツ	105	0.17%
34	奈良	97	0.16%
35	長野	94	0.15%
36	イギリス	89	0.15%
37	愛知	89	0.15%
38	フランス	88	0.14%
39	博多	86	0.14%
40	舞浜	80	0.13%
41	九州	79	0.13%
42	ブータン	69	0.11%
43	ニューヨーク (NY+ニューヨーク)	68	0.11%
44	スイス	63	0.10%
45	新大久保	63	0.10%
46	静岡	63	0.10%
47	兵庫	63	0.10%
48	台湾	62	0.10%
49	町田	62	0.10%
50	北欧	62	0.10%

順位	地域名	発言者数	発言者率
51	本州	61	0.10%
52	新潟	60	0.10%
53	パリ	59	0.10%
54	浦安	57	0.09%
55	大宮	56	0.09%
56	川崎	55	0.09%
57	代官山	55	0.09%
58	八王子	54	0.09%
59	湘南	52	0.08%
60	中目黒	52	0.08%
61	原宿	51	0.08%
62	恵比寿	50	0.08%
63	スペイン	49	0.08%
64	オーストラリア	48	0.08%
65	岡山	48	0.08%
66	三鷹	46	0.08%
67	カナダ	45	0.07%
68	目黒	45	0.07%
69	フィンランド	44	0.07%
70	上野	44	0.07%
71	中央線	44	0.07%
72	山手線	43	0.07%
73	江ノ島	42	0.07%
74	スウェーデン	41	0.07%
75	香川	41	0.07%

順位	地域名	発言者数	発言者率
76	滋賀	39	0.06%
77	梅田	39	0.06%
78	ソウル	37	0.06%
79	荻窪	37	0.06%
80	六本木	37	0.06%
81	蒲田	33	0.05%
82	世田谷	33	0.05%
83	東横線	33	0.05%
84	東北	33	0.05%
85	福島	33	0.05%
86	豊洲	33	0.05%
87	オランダ	32	0.05%
88	中央区 (中央)	32	0.05%
89	柏	32	0.05%
90	ロンドン	31	0.05%
91	熊本	31	0.05%
92	難波	31	0.05%
93	阿佐ヶ谷	30	0.05%
94	栄	30	0.05%
95	岐阜	30	0.05%
96	四国	30	0.05%
97	武蔵小杉 (小杉)	30	0.05%
98	長崎	30	0.05%
99	日吉	30	0.05%
100	品川	30	0.05%

※TMSの単語頻度分析機能を用いた。

データ件数:61314件

住みたくない地域ランキング

- 前スライドと同様に、住みたくない地域についても抽出した。
- 住みたくないという発言者が最も多かった地域も同じく「東京」、次いで「日本」「北海道」「関東」「大阪」の順であった。
- 上位20件内では、「**住みたい地域**」でも**上位100件に入っている地域**が15件を占めていた。

順位	地域名	発言者数	発言者率
1	東京	388	7.60%
2	日本	117	2.29%
3	北海道	85	1.66%
4	関東	84	1.64%
5	大阪	77	1.51%
6	アメリカ	39	0.76%
7	沖縄	37	0.72%
8	福島	34	0.67%
9	京都	30	0.59%
10	韓国	26	0.51%
11	関西	24	0.47%
12	千葉	23	0.45%
13	名古屋	23	0.45%
14	東北	22	0.43%
15	埼玉	19	0.37%
16	大津	19	0.37%
17	本州	16	0.31%
18	中国	15	0.29%
19	渋谷	14	0.27%
20	日本海	12	0.23%

※注

「住みたくない」という発言の割合のみが大きい地域は、テキストの原文を参照したところ、特定の悪い印象のイベントの存在が影響していた。

福島、東北：東日本大震災による福島原子力発電所の事故

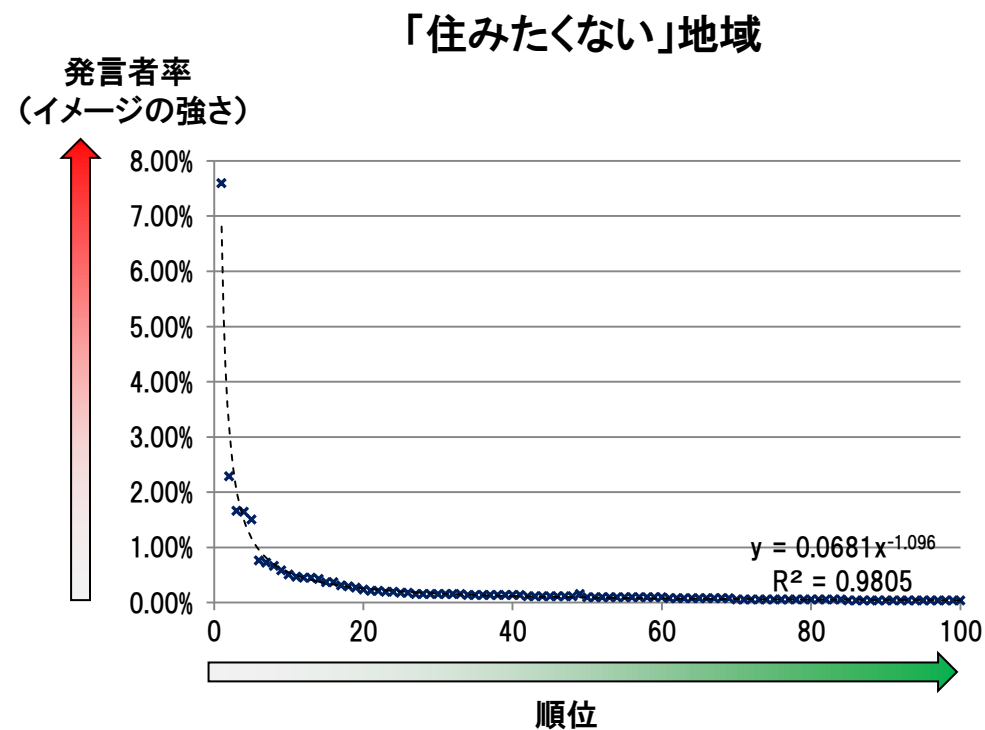
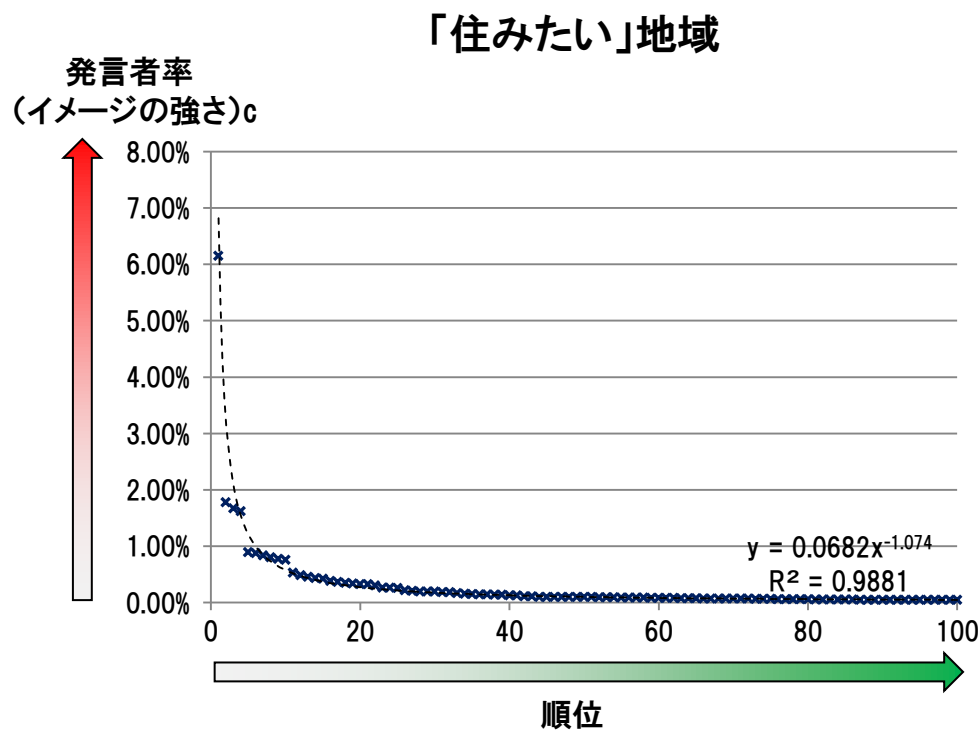
大津：いじめによる小学生の自殺事件

中国：領土問題をめぐる国際関係の悪化

日本海：大雪による影響

発言者率(イメージの強さ)の分布

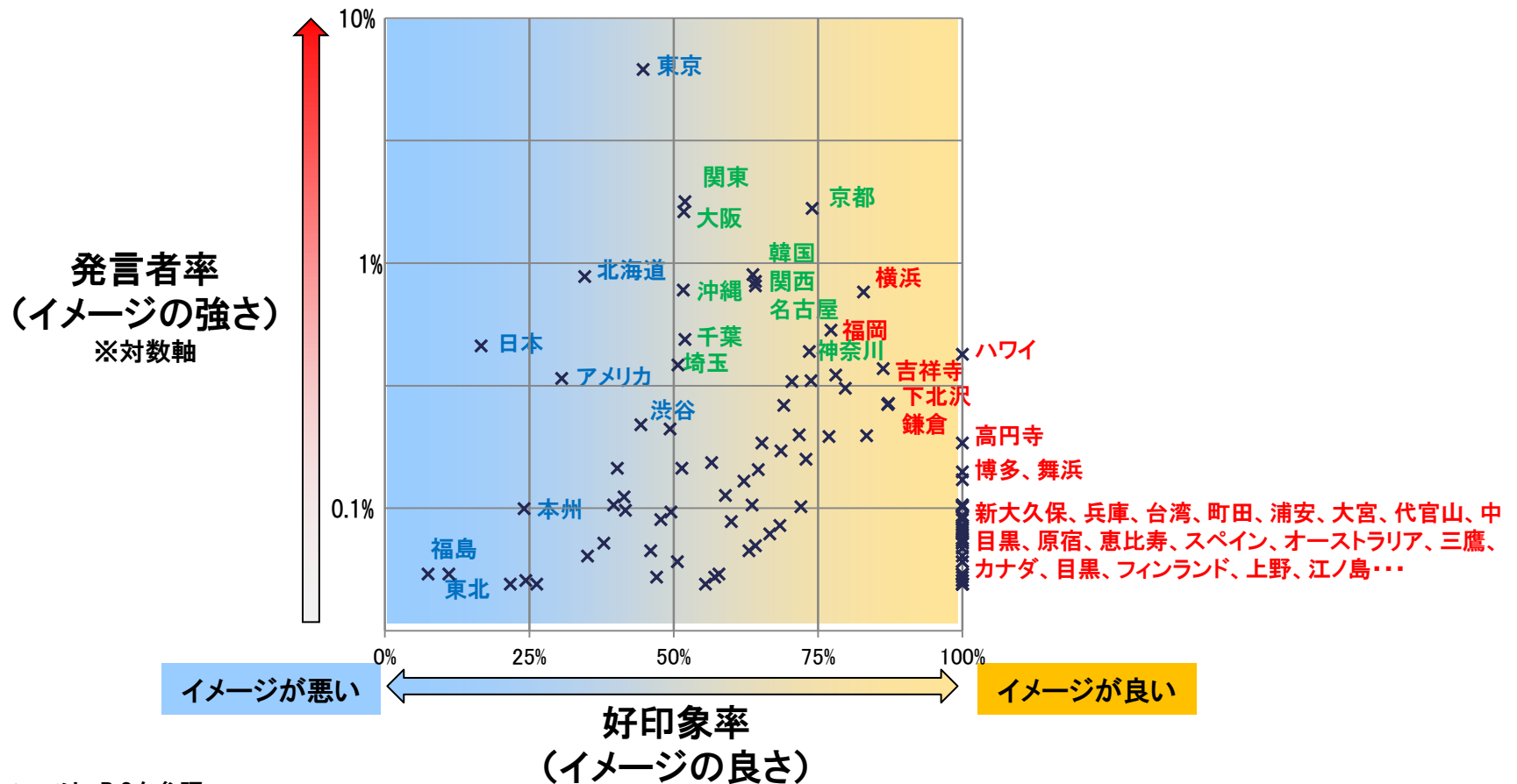
- 各地域の発言者率を順位別に並べると、「住みたい」地域、「住みたくない」地域ともに、累乗分布で近似される曲線が描かれた。
- 次スライドでは、これらを組み合わせて、各地域のイメージの強さと良さについて比較を試みる。



※近似曲線はExcel2007を用いて算出

イメージの良い地域と悪い地域の抽出

- X軸に好印象率(イメージの良さ)、Y軸に発言者率をとって、各地域を散布図にプロットした。
- **好印象率が75%を超えた地域**は、「横浜」「福岡」「ハワイ」「吉祥寺」「下北沢」「鎌倉」「高円寺」などであった。
- **好印象率が50%を下回った地域**は、「東京」「北海道」「日本」「アメリカ」「渋谷」などであった。

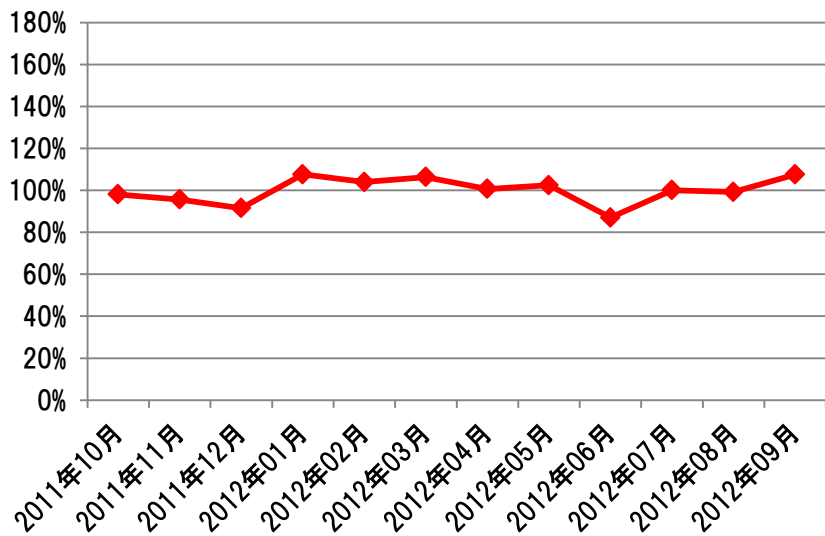


※指標の定義については、P.8を参照

季節変動による影響

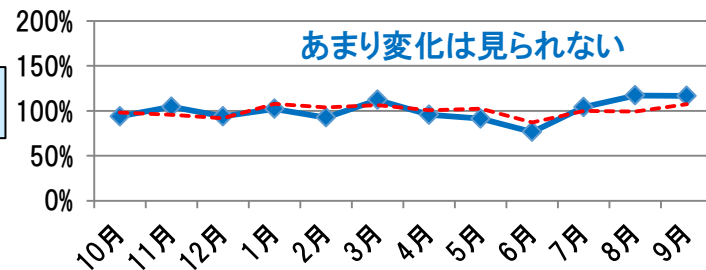
- ・ 時系列にそって比較したところ、季節、月によって、「住みたい」という全発言数には大きな変動は見られなかった。
- ・ 北海道、沖縄など、一部の地域では、特定の季節に発言数が多くなる傾向が見られた。地域イメージの形成に、特定の季節に由来する要素(夏の涼しさ、冬の暖かさなど)が影響しているのではないかと考えられる。

1日あたりの
「住みたい」全発言数
※年平均に対する比

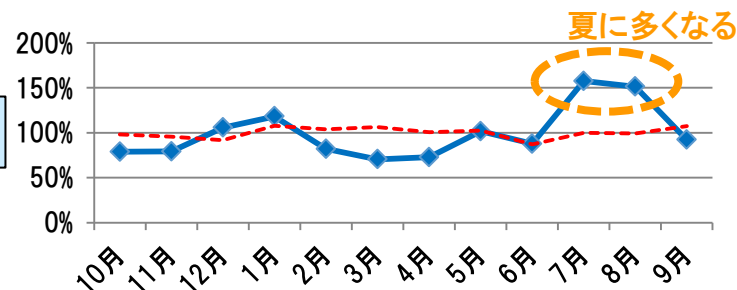


地域別に
比較

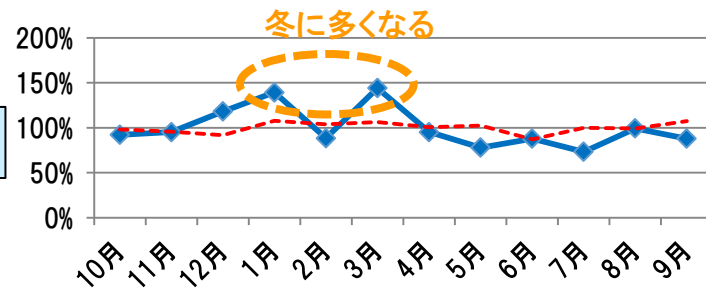
東京



北海道



沖縄



※点線は全発言数の推移(左グラフと同じ)

住みたい地域イメージとの関連ワード

- 次に、「住みたい」というテキストの前後を含めた発言の全体(P.5②)から、出現頻度の高い一般名詞を抽出した。
- 上位100件内では、**周辺の施設・環境に関する単語**が22件、**特定の圏域を示す単語**が10件、**人間関係に関する単語**が5件、などが抽出された。これらの要素などが、住みたい地域のイメージに影響を与えているものと考えられる。

順位	単語	発言者数	発言者率
1	家	5335	8.70%
2	人	2511	4.10%
3	部屋	2185	3.56%
4	都会	1818	2.97%
5	田舎	1545	2.52%
6	街	1043	1.70%
7	駅	835	1.36%
8	世界	833	1.36%
9	マンション	817	1.33%
10	学校	804	1.31%
11	自分	802	1.31%
12	国	783	1.28%
13	都内	781	1.27%
14	場所	777	1.27%
15	海	761	1.24%
16	夢	732	1.19%
17	お金	710	1.16%
18	電車	612	1.00%
19	住	537	0.88%
20	大学	527	0.86%
21	隣	518	0.84%
22	海外	515	0.84%
23	町	489	0.80%
24	本	462	0.75%
25	感じ	449	0.73%

順位	単語	発言者数	発言者率
26	友達	440	0.72%
27	実家	434	0.71%
28	図書館	412	0.67%
29	お風呂	379	0.62%
30	地元	379	0.62%
31	お店	377	0.61%
32	気持ち	377	0.61%
33	都心	371	0.61%
34	次	365	0.60%
35	家賃	358	0.58%
36	目	355	0.58%
37	最高	354	0.58%
38	距離	352	0.57%
39	本気	351	0.57%
40	会社	350	0.57%
41	トイレ	345	0.56%
42	雪	340	0.55%
43	近所	337	0.55%
44	本屋	332	0.54%
45	辺り	328	0.53%
46	山	315	0.51%
47	城	306	0.50%
48	バイト (バイト先)	303	0.49%
49	物件	301	0.49%
50	本当	286	0.47%

順位	単語	発言者数	発言者率
51	ライブ	284	0.46%
52	老後	283	0.46%
53	外国	279	0.46%
54	地域	279	0.46%
55	車	276	0.45%
56	イベント	274	0.45%
57	東	273	0.45%
58	カラオケ	270	0.44%
59	地方	268	0.44%
60	アパート	257	0.42%
61	二次元	246	0.40%
62	現実	243	0.40%
63	島	241	0.39%
64	一軒家	240	0.39%
65	交通	235	0.38%
66	猫	233	0.38%
67	浴線	230	0.38%
68	子	228	0.37%
69	土地	226	0.37%
70	子供	221	0.36%
71	ご飯	220	0.36%
72	徒歩	218	0.36%
73	環境	211	0.34%
74	森	210	0.34%
75	バス	207	0.34%

順位	単語	発言者数	発言者率
76	周辺	207	0.34%
77	付近	206	0.34%
78	泣	204	0.33%
79	外	202	0.33%
80	レベル	198	0.32%
81	頭	197	0.32%
82	理由	197	0.32%
83	星	194	0.32%
84	ガチ	189	0.31%
85	壁	187	0.30%
86	市内	186	0.30%
87	虫	186	0.30%
88	家族	185	0.30%
89	庭	183	0.30%
90	村	180	0.29%
91	親	177	0.29%
92	お菓子	175	0.29%
93	暮らし	175	0.29%
94	音	171	0.28%
95	西	171	0.28%
96	系	170	0.28%
97	心	170	0.28%
98	地震	169	0.28%
99	空気	166	0.27%
100	映画	164	0.27%

※TMSの単語頻度分析機能を用いた。

ただし、抽出された単語のうち、ひらがな、英数字、および記号のみから構成されていた単語については、抽出がうまくいかなかった結果と解釈し、結果から削除した

居住地域の分布

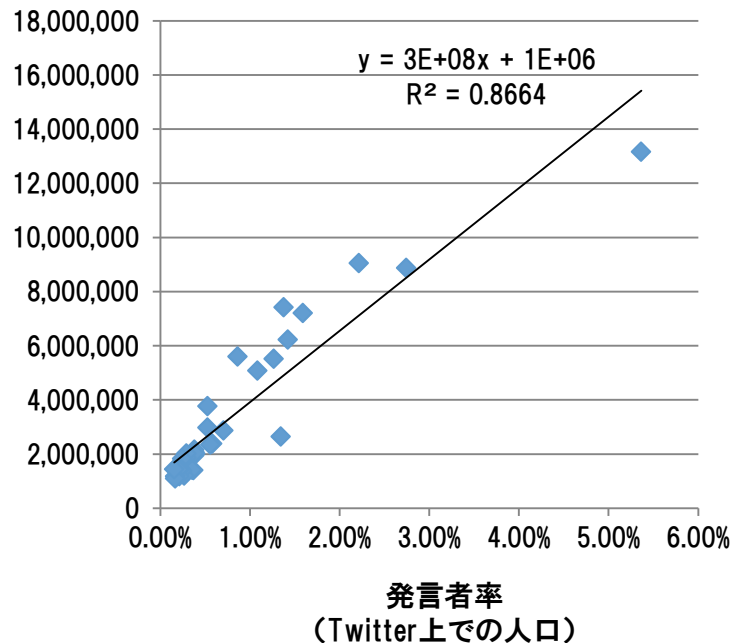
- 次に、「住みたい」という発言をした人の居住地域(P.5の③)を、テキストマイニングにより抽出した。
- 上位50件内では、都道府県名が34件、市町村名が8件、市町村の中の地域名が4件を占めていた。
- 都道府県別の発言者率については、2010年の国勢調査人口との間で線形関係が見られた。

順位	地域名	発言者数	発言者率
1	東京	479	5.36%
2	大阪	245	2.74%
3	神奈川	198	2.22%
4	埼玉	142	1.59%
5	千葉	127	1.42%
6	愛知	123	1.38%
7	京都	120	1.34%
8	北海道	113	1.27%
9	日本	108	1.21%
10	横浜	101	1.13%
11	福岡	97	1.09%
12	兵庫	77	0.86%
13	仙台	65	0.73%
14	名古屋	64	0.72%
15	広島	63	0.71%
16	札幌	56	0.63%
17	新潟	52	0.58%
18	宮城	50	0.56%
19	さいたま	47	0.53%
20	茨城	47	0.53%
21	関西	47	0.53%
22	静岡	47	0.53%
23	関東	44	0.49%
24	神戸	38	0.43%
25	群馬	35	0.39%

順位	地域名	発言者数	発言者率
26	岡山	34	0.38%
27	岐阜	34	0.38%
28	長野	34	0.38%
29	福島	34	0.38%
30	奈良	33	0.37%
31	滋賀	32	0.36%
32	新宿	32	0.36%
33	三重	26	0.29%
34	栃木	26	0.29%
35	渋谷	25	0.28%
36	世田谷	25	0.28%
37	大分	24	0.27%
38	熊本	22	0.25%
39	青森	22	0.25%
40	山口	20	0.22%
41	川崎	20	0.22%
42	東北	20	0.22%
43	石川	19	0.21%
44	池袋	18	0.20%
45	長崎	18	0.20%
46	八王子	17	0.19%
47	沖縄	15	0.17%
48	山形	15	0.17%
49	秋田	15	0.17%
50	愛媛	14	0.16%

国勢調査人口と発言者率の関係
(上位50位以内に入る都道府県)

2010年国勢調査人口
(単位:人)



※現実に行政界名、駅名などとして存在する地名と日本語で一致する単語のみを抽出した。(TOKYO、などは対象外)
※都道府県名と市町村名の両方に当てはまる地名(例:大阪府と大阪市)については、都道府県名として分類した。



3. 特定の地域の傾向についての分析

以下では、サンプルとして、東京、関東、京都、北海道、吉祥寺の5地域を取り上げ、P.5①②③のデータを組み合わせた、詳細な分析を行う。

「東京」のイメージ構成要素とイメージ保有者の居住地

- 「東京」に関連する単語としては、**イベント、ライブ等に関連する単語**が多く抽出された。
- イメージ保有者の居住地は、**関東以外の地域**が占める割合が大きくなっている。特徴的な単語の中にも、**大阪、愛知という遠方の地域名**が抽出されており、これらの地域とイメージの比較が行われていると考えられる。

特徴的な単語

単語	当該発言者数	全発言者数	補完類似度
東京	3784	4767	14029
イベント	105	274	352
大阪	177	1418	340
ライブ	68	284	200
近郊	44	74	158
付近	47	206	135
千葉	55	452	103
近辺	32	134	94
地方	39	268	87
埼玉	47	398	85
神奈川	44	396	73
お金	63	710	67
周辺	30	207	66
舞台	20	65	64
交通	31	235	63
画像	18	43	61
田舎者	14	36	47
東京ドーム	13	22	47
新幹線	17	85	46
グッズ	15	60	45
LIVE	15	63	44
愛知	21	156	44
トレンド	11	14	41
イベ（イベント）	11	33	36
東京タワー	11	34	35
高校	16	123	32
飛行機	12	62	32

居住地の分布

データ件数:618件(居住地)

順位	単語	発言者数	発言者率
1	大阪	27	3.3%
2	愛知	15	1.8%
3	京都	9	1.1%
4	北海道	9	1.1%
5	福岡	9	1.1%
6	名古屋	9	1.1%
7	神奈川	8	1.0%
8	千葉	7	0.9%
9	関西	7	0.9%
10	埼玉	6	0.7%
11	兵庫	6	0.7%
12	群馬	6	0.7%
13	長野	6	0.7%
14	茨城	5	0.6%
15	関東	5	0.6%
16	東京	4	0.5%
17	仙台	4	0.5%
18	札幌	4	0.5%
19	宮城	4	0.5%
20	福島	4	0.5%

※「特徴的な単語」は、TMSの特徴語抽出機能を用い、補完類似度が30以上で、2文字以上からなる単語を抽出した。

ただし、抽出された単語のうち、ひらがな、英数字、および記号のみから構成されていた単語については、抽出がうまくいかなかった結果と解釈し、結果から削除した

「関東」のイメージ構成要素とイメージ保有者の居住地

- 「関東」に関連する単語としては、「東京」と同様、**イベント、ライブ等に関連する単語**が多く抽出された。また、特徴的な単語が少ない傾向にあった。
- イメージ保有者の居住地も、「東京」のイメージ保有者の居住地とほぼ重なっている。

特徴的な単語

単語	当該発言者数	全発言者数	補完類似度
関東	1095	1315	7677
関西	57	725	306
イベント	34	274	205
地方	33	268	199
ライブ	18	284	89
交通	16	235	82
お金	24	710	72
楽しみ	12	145	66
地元	13	379	40
イベ	6	33	38
アニメ	7	99	36
LIVE	6	63	34
テレビ	7	133	31
新幹線	6	85	31

居住地の分布

データ件数:245件(居住地)

順位	単語	発言者数	発言者率
1	愛知	10	4.1%
2	大阪	9	3.7%
3	仙台	6	2.4%
4	名古屋	6	2.4%
5	兵庫	4	1.6%
6	宮城	4	1.6%
7	北海道	3	1.2%
8	福島	3	1.2%
9	京都	2	0.8%
10	福岡	2	0.8%
11	岐阜	2	0.8%
12	長野	2	0.8%
13	札幌	1	0.4%
14	新潟	1	0.4%
15	静岡	1	0.4%
16	神戸	1	0.4%
17	三重	1	0.4%
18	熊本	1	0.4%
19	青森	1	0.4%
20	山口	1	0.4%

「京都」のイメージ構成要素とイメージ保有者の居住地

- 「京都」に関連する単語としては、**近隣の地域に関連する単語**と、**観光資源に関連する単語**が多く抽出された。
- イメージ保有者の居住地は、**近隣の東海・近畿圏の地域**が占める割合が大きくなっている。

特徴的な単語

単語	当該発言者数	全発言者数	補完類似度
京都	1025	1268	8146
大阪	60	1418	308
奈良	14	150	95
神戸	15	297	84
老後	14	283	78
滋賀	8	75	55
抹茶	7	16	55
修学旅行	7	33	53
市内	9	186	50
京大	6	10	47
神社	6	44	43
大学	13	527	39
雰囲気	7	142	39
鴨川	5	14	39
歴史	5	27	37
思い出	5	32	37
岐阜	5	49	34
地元	10	379	34
感じ	11	449	33
仏閣	4	5	32
人生	6	136	32
関西	15	725	31

居住地の分布

データ件数:213件(居住地)

順位	単語	発言者数	発言者率
1	東京	7	3.3%
2	愛知	6	2.8%
3	大阪	5	2.3%
4	京都	5	2.3%
5	滋賀	4	1.9%
6	神奈川	3	1.4%
7	北海道	3	1.4%
8	横浜	3	1.4%
9	兵庫	3	1.4%
10	神戸	3	1.4%
11	新潟	2	0.9%
12	長野	2	0.9%
13	三重	2	0.9%
14	大分	2	0.9%
15	千葉	1	0.5%
16	仙台	1	0.5%
17	名古屋	1	0.5%
18	札幌	1	0.5%
19	奈良	1	0.5%
20	栃木	1	0.5%

「北海道」のイメージ構成要素とイメージ保有者の居住地

- 「北海道」に関連する単語としては、**地域の特色ある気候や食べ物に関連する単語**や、**北海道以外の観光地に関連する単語**が多く抽出された。北海道以外の観光地は、北海道とのイメージの比較対象になっていると考えられる。
- イメージ保有者の居住地については、明確な特徴が見られない結果となった。

特徴的な単語

単語	当該発言者数	全発言者数	補完類似度
北海道	541	729	5835
沖縄	31	600	283
札幌	15	287	137
雪かき	7	32	73
ゴキブリ	7	54	71
梅雨	4	16	42
老後	6	283	39
季節	4	53	39
県民	4	53	39
飛行機	4	62	38
フィンランド	4	65	38
食べ物	4	75	37
大阪	15	1418	33
ウニ	3	10	32
かに	3	18	31
憧れ	3	20	31
九州	4	140	31
ロシア	3	29	30
生まれ	3	31	30
修学旅行	3	33	30

居住地の分布

データ件数:76件(居住地)

順位	単語	発言者数	発言者率
1	東京	2	2.6%
2	大阪	2	2.6%
3	神奈川	2	2.6%
4	福岡	2	2.6%
5	京都	1	1.3%
6	横浜	1	1.3%
7	兵庫	1	1.3%
8	岡山	1	1.3%
9	大分	1	1.3%
10	熊本	1	1.3%

「吉祥寺」のイメージ構成要素とイメージ保有者の居住地

- 「吉祥寺」に関連する単語としては、**近隣の地域、路線名に関連する単語**が多く抽出された。地域名以外に特徴的な単語の数は少なかった。
- イメージ保有者の居住地は、「東京」イメージの場合と異なり、**東京を含む関東圏の地域**が占める割合が大きい。

特徴的な単語

単語	当該発言者数	全発言者数	補完類似度
吉祥寺	227	289	3553
吉祥	13	22	203
三鷹	11	70	169
中野	9	157	132
中央線	8	81	121
家賃	9	358	119
高円寺	8	140	117
下北	6	100	88
下北沢	6	109	87
高円	5	14	78
近辺	5	134	70
辺り	5	328	58
お店	5	377	55
地元	5	379	54
周辺	4	207	50
阿佐ヶ谷	3	38	45
学生	3	84	42
埼玉	4	398	38
徒歩	3	218	33
井の頭公園	2	13	31
お別れ	2	18	30

居住地の分布

データ件数:60件(居住地)

順位	単語	発言者数	発言者率
1	東京	7	12%
2	神奈川	3	5%
3	渋谷	3	5%
4	大阪	1	2%
5	埼玉	1	2%
6	千葉	1	2%
7	愛知	1	2%
8	新潟	1	2%
9	岐阜	1	2%
10	奈良	1	2%
11	池袋	1	2%
12	愛媛	1	2%
13	富山	1	2%



4. 結果のまとめ

- ・本研究にて得られた知見は、下記のとおりである。

全体的な傾向についての分析

- ・イメージの良い地域としては、海外の地域や関東圏内の地域が多く抽出された。
- ・イメージの悪い地域は、イメージの良い地域と同様の地域が抽出される傾向にあった。
ただし、一部、特定のイベントの影響により極端にイメージの悪い地域が見られた。
- ・イメージの良い地域、悪い地域の印象の強さは、累乗分布に従った。
- ・イメージが相対的に良い地域と、悪い地域を抽出することに成功した。
- ・北海道、沖縄など、一部の地域において、特定の季節に印象の強くなる傾向が見られた。
- ・地域イメージとの関連ワードとしては、周辺の施設・環境に関する単語、特定の圏域を示す単語、人間関係に関する単語、などが抽出された。



特定の地域の傾向についての分析

- ・地域に特徴的な関連ワードとしては、以下の種類の単語が抽出される傾向があった。
 - ・地域の特色や、観光資源に関する単語
 - ・北海道では気候や食べ物、京都では施設や食べ物、東京・関東ではライブやイベントに関する単語が抽出された
 - ・周辺の地域に関する単語
 - ・当該地域と、一体的にイメージを形成している可能性がある)
 - ・遠方の地域に関する単語
 - ・イメージの比較対象になっている可能性がある
- ・対象地域のイメージ保有者の居住地域の分布については、対象地域の近隣地域である場合(京都、吉祥寺)と、対象地域から離れた地域である場合(東京、関東)に分かれた。
- ・地域によっては、明確な傾向がつかめない場合もあった。(単にサンプル数の不足の可能性もあり)

今後の課題

- ・本研究では、限られたサンプル数のデータを用いて、一部の基礎的な分析を試みたにすぎない。
- ・今後の課題として、主に下記の3点が挙げられる。特に、他のデータや事象との関係について分析することにより、より有用な知見を得られる可能性があるものと考えられる。

テーマ	課題
分析精度の向上	<ul style="list-style-type: none"> ・データのサンプル数増加による、統計的有意性の検証 ・地名辞書の整備による、結果の体系的整理 類義語の整理(東京とTOKYOとトーキョー、など) 地名の階層構造の整理(関東→東京→新宿→新宿御苑、など)
他のデータとの比較、組合せによる、本研究の分析結果の有用性検証	<ul style="list-style-type: none"> ・「住みたくない」地域の発言者と、その居住地域とのクロス分析 ・「住みたい」以外の、地域イメージと結びつきうるキーワードを用いた分析 例えば、「行きたい」「行った」「宿泊した」対象の地域を抽出する、など ・Twitter以外のソーシャルメディアデータ(ブログ、Facebook等)を用いた分析 ・ソーシャルメディア以外のメディアでの記述との比較分析 新聞・雑誌記事、政策文書など ・実際の移動履歴に関するデータとの比較分析 パーソントリップ調査データ、観光移動データなど
本研究の分析結果を用いた、イメージの変化をもたらすイベントの効果分析	<ul style="list-style-type: none"> ・時系列分析による、イメージの向上・低下が発生した地域、時点の抽出 および、イメージの変化の要因の分析 メディアによる取り上げの効果測定 地域の自治体、企業等による取り組み、イベントの効果測定、など

ケヴィン・リンチ(1968)「都市のイメージ」岩波書店

山崎ほか(2004)「地域イメージの表現手法に関する研究—司馬遼太郎『街道をゆく』における文章構成の分析から」
都市計画論文集No.39-3 pp.97-102

和田ほか(2009)「地域ブランド・マネジメント」有斐閣

津田ほか(2011)「雑誌『旅』にみる温泉地に対するイメージの変遷に関する研究」都市計画論文集No.46 pp.607-612

ブランド総合研究所(2012)「第7回「地域ブランド調査2012」」
http://tiiki.jp/news/05_research/survey2012/1243.html

リクルート(2012)「2012年版 関東編 住みたいまちランキング」
http://suumo.jp/edit/sumi_machi/2012/kanto/?vos=nsualap20110328002

総務省統計局(2011)「平成22年国勢調査」
<http://www.stat.go.jp/data/kokusei/2010/index.htm>