

観光地として不人気の埼玉県と
人気のある神奈川県
相違点の解明
—国内外の観光客の口コミデータによる検証—

法政大学 理工学部 創生科学科 呉研究室

田原 加奈子

0.目次

①目的と方法

1.背景と目的

2.先行研究と本研究
の違い

3.使用するデータの説明

②観光地と観光客の分析 (Excel使用)

4.観光地の比較

5.観光客の居住地
の比較

③観光要因の分析 (Excelとvms使用)

6.観光スポットの
時間的距離の比較

7.観光スポットの
評価の比較

8.観光スポットを
訪れた月の比較

④観光客の感想 (tms使用)

9.口コミのテキストマイニング

⑤複数のカテゴリを扱う分析 (vmsとR使用)

10.6カテゴリの
アソシエーション分析

11.「観光客の居住地」と
「観光地」の対応分析

⑥まとめ

12.埼玉県の問題と施策の提案

13.今後の課題

14.参考文献

15.付録

1.背景と目的

埼玉県は「コロナ後に観光で行きたい都道府県ランキング」[1]で最下位。

埼玉県は日本の首都である東京都とアクセスが容易にも関わらず、なぜここまで観光地として人気がないのか？

本研究では、トリップアドバイザー[2]の観光スポットに対する観光客の口コミから、同様に東京都とアクセスが容易であるが観光地として人気が高い神奈川県([1]のランキングで7位)と埼玉県の観光地を比較し、埼玉県が観光地として人気がない理由を探った。

トリップアドバイザーとは2018年、日本の「旅行・交通」カテゴリwebサイトのpcからの閲覧者数ランキングでは3位、口コミ掲示板の中では1位[3]の旅行関連の口コミ掲載サイト。

2. 先行研究と本研究の違い

大久保立樹・室町泰徳(2014)「旅行ガイドブックと口コミの言語解析による訪日外国人の観光地に対するイメージに関する研究」では旅行ガイドブック(lonely planet)と旅行口コミサイト(トリップアドバイザー)を用いてテキスト分析を行い、訪日外国人観光客の出身国別に東京都内の5観光地のイメージを抽出している。

本研究では、訪日外国人に加え国内からの観光客の居住地(都道府県)別に対しても分析を行った。また、埼玉県・神奈川県各50観光スポットを分析することで、双方の観光地としての特性を比較しながら分析を進めた。また、テキストマイニング、対応分析、アソシエーション分析といった多種多様な分析を用いることで多角的に埼玉県と神奈川県の観光地としての相違点を求めた。

3.使用するデータの説明

トリップアドバイザーの使用可能言語は28言語にも及ぶが今回は日本語、英語、中国語を使用。

2008年(サービス開始)から2020年6月18日の期間に投稿された埼玉県, 神奈川県の口コミ数が多かった上位50観光スポット(計100観光スポット)の51,617件の口コミを分析に用いた。

使用するカテゴリは

- ①観光客の居住地(国) ②観光客の居住地(都道府県) ③観光スポット
- ④観光地(観光スポットがある市町村) ⑤観光客が訪れた月
- ⑥観光客がつけた評価(5段階) ⑦口コミのテキストデータ(日本語のみ)

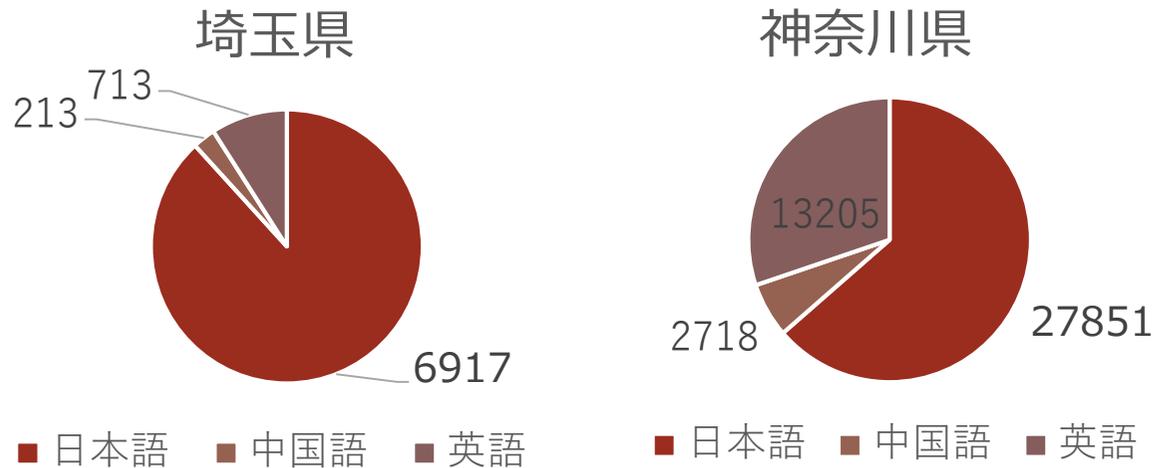
※本研究では「観光スポット」と、「観光地(観光スポットがある市町村)」は分けて使用した。

また、観光スポットのアクセスのしやすさを調べるためにJORUDAN(乗り換え案内サイト)[6]で東京駅から各観光スポットまでの移動時間を求めたものを使用した。

3.使用するデータの説明

3-1.言語の割合と口コミ数の比較

言語の割合



- ・埼玉県の観光スポット，神奈川県
の観光スポットともに日本語の口コミが
ほとんどを占めている。
- ・中国語の口コミが特に少ない。

口コミ数の比較

表1. .埼玉県， 神奈川県の観光スポットに対する口コミ数
約5.6倍

	埼玉県	神奈川県
全体の口コミ数	7843	43774
日本語で投稿され，国内から訪れた観光客の口コミ数	6772	27034

↑テキストマイニングで使用

約4倍

- ・トリップアドバイザー内でも埼玉県の観光
スポットより神奈川県の観光スポットに対す
る口コミが多い。

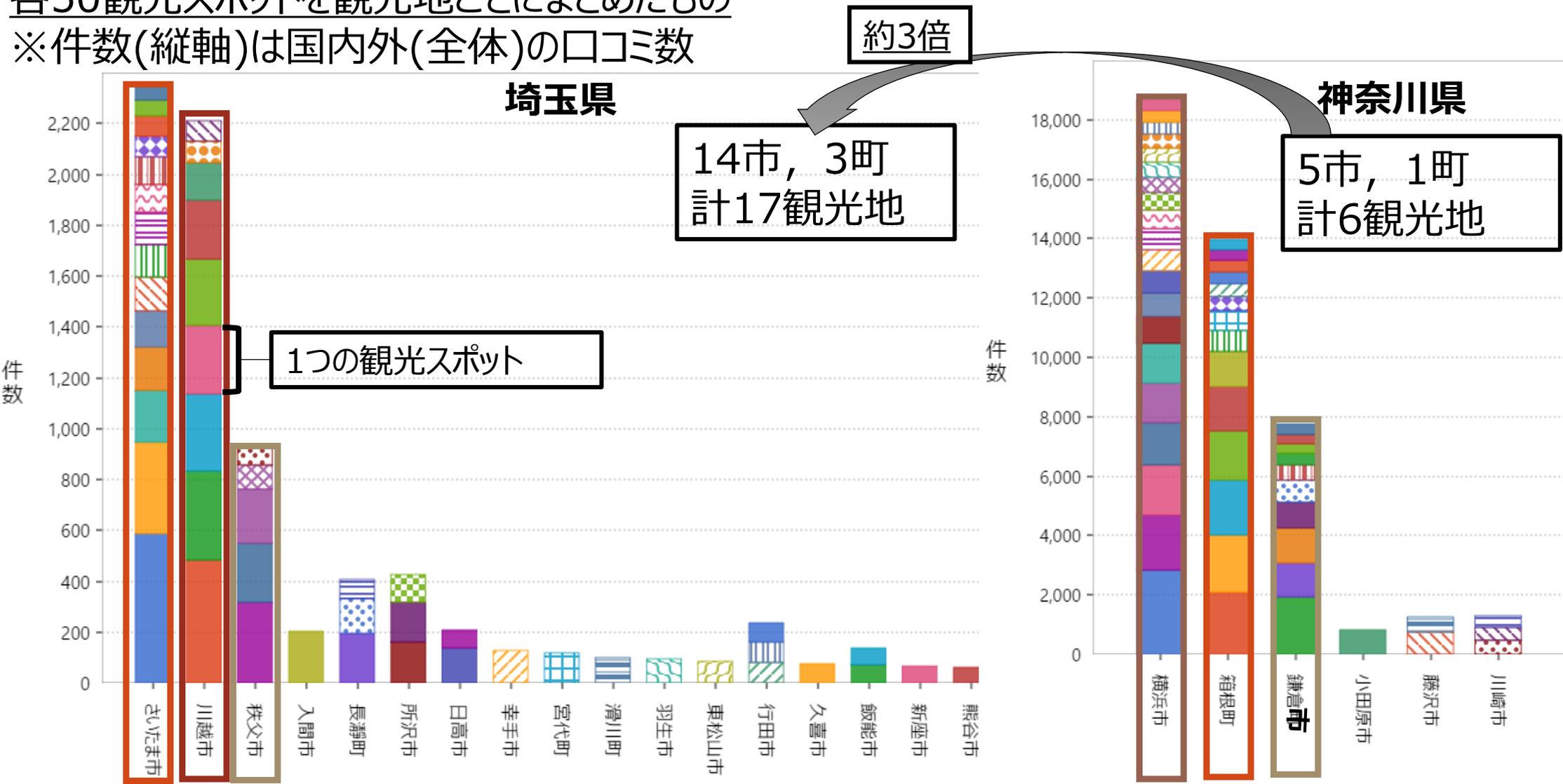
※以下では，口コミ数多い=観光客数が多いと表現する場合がある。

4. 観光地の比較

4-1. 観光地数の比較

Vmsの「データ可視化」を用いて
 今回使用した埼玉県・神奈川県
 の各50観光スポットを観光地ごとにまとめたもの
 ※件数(縦軸)は国内外(全体)の口コミ数

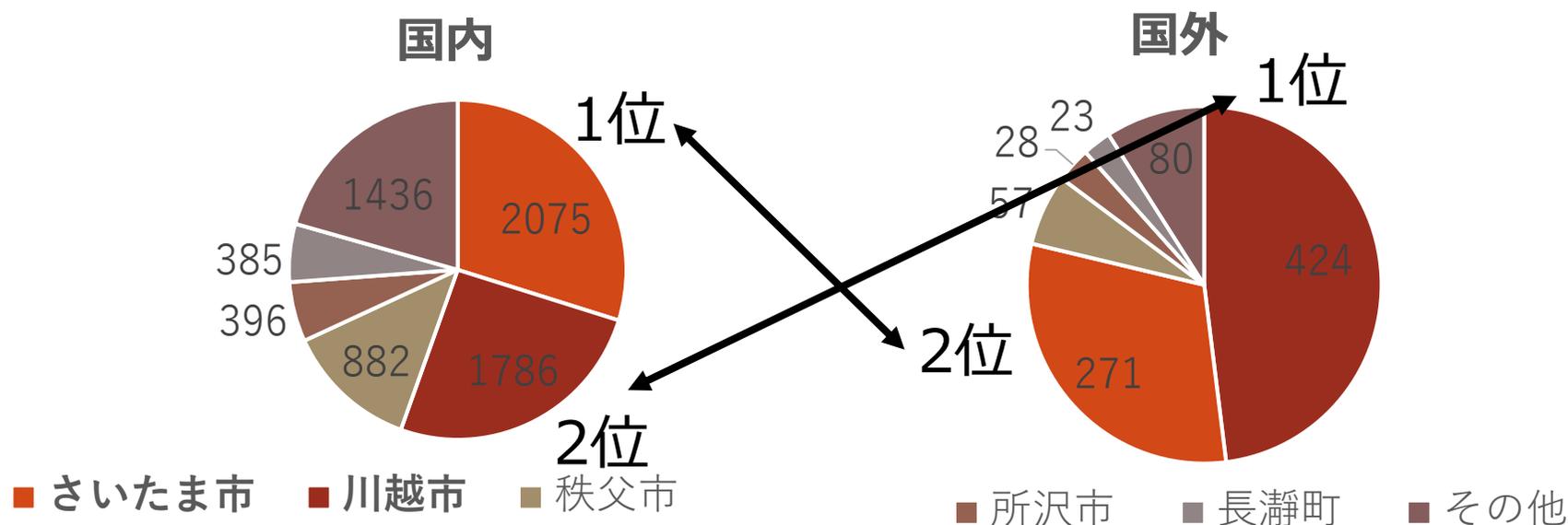
神奈川県は観光地より、埼玉県の観光地の方が多く
 埼玉県の観光スポットが分散していることが分かった。



4.観光地の比較

4-2.国内・国外別，観光客が訪れる埼玉県の観光地の割合

国内・国外では1,2位の順位が逆転

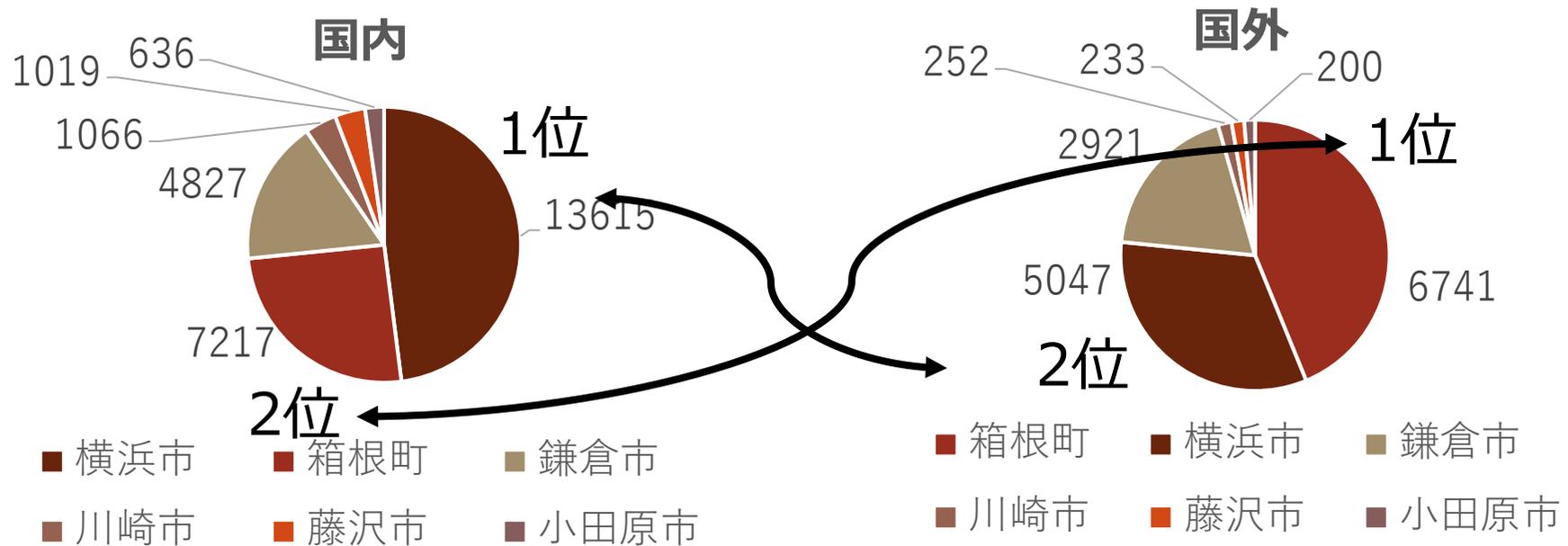


- ・埼玉県を国内から訪れる観光客は「さいたま市」に最も訪れており，2番目に「川越市」を訪れている。
- ・埼玉県を国外から訪れる観光客のうち半数近くが「川越市」を訪れており，2番目に「さいたま市」を訪れている。
- ・比較的埼玉県の観光地として有名な「秩父市」は国内から埼玉県を訪れる観光客の13%未満，国外から埼玉県を訪れる観光客の7%未満しか訪れておらず，「さいたま市」・「川越市」と比べると人気がないことが分かった。

4.観光地の比較

4-3.国内・国外別，観光客が訪れる神奈川県内の観光地の割合

国内・国外では1,2位の順位が逆転



- ・神奈川県にある6観光地の中で「横浜市」「箱根町」「鎌倉市」を主に国内・国外の観光客が訪れている。
- ・国内から神奈川県を訪れる観光客の半数近くが「横浜市」、国外から神奈川県を訪れる観光客の半数近くが「箱根町」を訪れている。

5.観光客の居住地の比較

5-1.観光客の主な居住地 国編

表2.埼玉県に訪れる観光客の
上位10位の居住地(国)

順位	国名	口コミ数
1	日本	6960
2	中国	213
3	アメリカ	146
4	オーストラリア	103
5	シンガポール	89
6	イギリス	54
7	マレーシア	32
8	インド	26
9	タイ	26
10	インドネシア	25

表3.神奈川県に訪れる観光客
の上位10位の居住地(国)

順位	国名	口コミ数
1	日本	28380
2	アメリカ	3486
3	オーストラリア	1992
4	イギリス	1823
5	中国	1734
6	シンガポール	1254
7	カナダ	719
8	マレーシア	522
9	タイ	394
10	インドネシア	342

表4.埼玉県・神奈川県を訪れる
観光客の居住地(国)の割合

比率(%)	埼玉県	神奈川県
日本	+7% 88.76	64.83 +24%
上位6位	+2% 96.48	88.34 +8%
上位10位	97.55	92.07

・埼玉県を訪れる観光客の9割近く、神奈川県を訪れる観光客の6割以上が国内からの観光客だった。このことから埼玉県は神奈川と比べ国内から訪れる観光客に依存し、国外から訪れる観光客を呼び寄せることが今後の課題という事が分かった。

・神奈川県の観光地に訪れる観光客の約9割、埼玉県の観光地に訪れる観光客の98%が口コミ数上位6カ国となっている。このことから双方、少数の国から訪れる観光客が全体の観光客の大多数を占めるという事が分かった。

5.観光客の居住地の比較

5-2.観光客の主な居住地 都道府県(国内)編

表5.埼玉県に訪れる観光客の上位10位の居住地(都道府県)

表6.神奈川県に訪れる観光客の上位10位の居住地(都道府県)

表7.埼玉県・神奈川県を訪れる観光客の居住地(都道府県)の割合

順位	都道府県名	口コミ数
1	埼玉県	2244
2	東京都	2215
3	神奈川県	610
4	千葉県	369
5	北海道	150
6	大阪府	137
7	群馬県	86
8	愛知県	85
9	茨城県	70
10	福岡県	61

順位	都道府県名	口コミ数
1	神奈川県	9157
2	東京都	8290
3	千葉県	1288
4	埼玉県	1188
5	大阪府	905
6	愛知県	841
7	北海道	766
8	静岡県	475
9	兵庫県	454
10	福岡県	435

比率(%)	a.埼玉県	b.神奈川県
a or b	32.24	> 32.27
(a or b)と東京都	64.07	> 61.48
上位4位(南関東)	78.13	> 70.20
上位10位	86.59	> 83.86

2倍

+14%

+9%

- ・双方、県内からの観光客が最も訪れており、全体の3割を占める。これに東京都から訪れる観光客を足し合わせると全体の6割を占め、この割合は若干神奈川県よりも埼玉県の方が大きい。
- ・埼玉県・神奈川県を訪れる上位4都道府県は南関東(埼玉県、神奈川県、東京都、千葉県)。これは埼玉県を訪れる8割近く、神奈川を訪れる観光客の7割を占める。
- ・埼玉県・神奈川県を訪れる上位10都道府県からの観光客は全体の85%近くを占め、残りの37都道府県から訪れる観光客は全体の15%にも満たない。このことから、埼玉県は神奈川県よりも南関東からの観光客に依存している事が分かる。

5.観光客の居住地の比較

5-3.補足

Q1.何か国の口コミがあるのか

→埼玉県には54カ国，神奈川県には100カ国の口コミがある。

埼玉県の約2倍

考察：埼玉県の観光地は神奈川県の観光地に比べ海外から認知されていない。

Q2. 埼玉県と神奈川県の人口差はどの程度なのか

5-2から埼玉県・神奈川県を訪れる観光客は県内から訪れる観光客が3割を占めるということが分かった。そのため，県内の人口が少ないから埼玉県を訪れる観光客が少ない(観光地として人気がない)のではないかと考察できる。

結果：埼玉県の人口は全国5位，（神奈川県は2位，埼玉県の約1.25倍の人口数）[7]
このことから，埼玉県は人口が少ないために埼玉県を訪れる観光客が少ない(観光地として人気がない)という事は考えづらいということが分かった。

観光地と観光客の分析のまとめ

4. 観光地の比較

- 埼玉県の方が神奈川県よりも観光地が分散している。
 - 双方、観光スポットが多い2,3の観光地に観光客が集中していることが分かる。
- 考察：埼玉県の観光地が多すぎて、観光スポットが少ない観光地に観光客が訪れづらい。
- 国内から訪れる観光客と国外から訪れる観光客では主に訪れる観光地は同じだが、訪れる観光地の割合が同じではないという事が分かった。

5. 観光客の居住地の比較

- 埼玉県・神奈川県を訪れる国内外の主な観光客の居住地は似ているが、埼玉県は国内から訪れる観光客に依存し、特に南関東から訪れる観光客に依存している。
- 埼玉県は人口が少ないため、埼玉県を訪れる観光客が少ない(観光地として人気がない)という事は考えづらいという事が分かった。

6. 観光スポットの時間的距離の比較

本研究では「埼玉県」と「神奈川県」は東京からアクセスがしやすいと仮定していたが「東京からのアクセス」を「東京駅から埼玉県，神奈川県各観光スポットまでの移動時間」として調べてみた。
 ※東京駅から各観光スポットまでの移動時間はJORUDAN(乗り換え案内サイト)で、同時刻発、飛行機・有料特急不可、一番早く到着する条件で検索

表8. 埼玉県・神奈川県別，東京駅から各観光スポットまでの平均移動時間

	埼玉県	神奈川県
東京駅から観光スポットまでの平均移動時間(分)	89.04	87.2

表8より埼玉県，神奈川県各観光スポットから東京駅までの平均移動時間の差はほとんどないという事が分かる。

また，東京駅から移動時間が短い観光スポットは，口コミ数が多い(観光地として人気)観光スポットではないかと考え，埼玉県・神奈川県各観光スポットの口コミ数(国内・国外)と観光スポットまでの移動時間の相関関係を求めた。

表9. 埼玉県・神奈川県別国内外の観光客の各観光スポットに対する口コミ数と各観光地までの移動時間の相関関係

	埼玉県	神奈川県
国内 「観光スポットの口コミ数」と「東京駅から観光スポットまでの移動時間」の相関関係	0.07	-0.123
国外 「観光スポットの口コミ数」と「東京駅から観光スポットまでの移動時間」の相関関係	-0.143	0.323

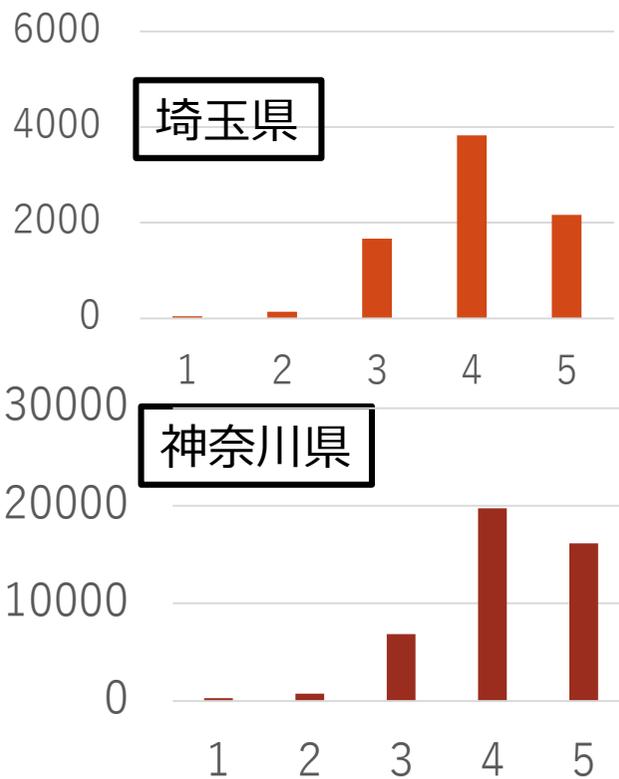
表9より、「観光スポットの口コミ数」と「東京駅から各観光スポットの移動時間」には相関関係がないことが分かる。

7.観光スポットの評価の比較

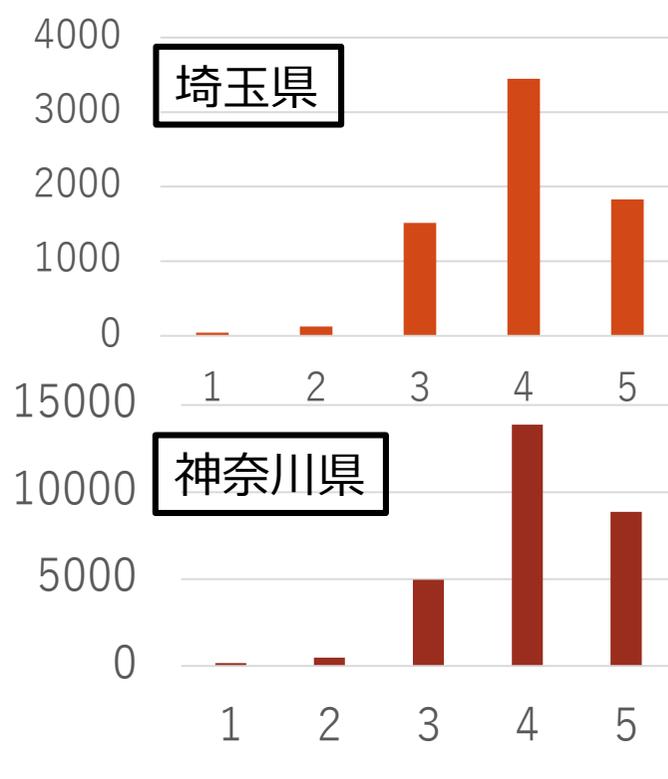
7-1.全体，国内，国外の埼玉県と神奈川県の評価の比較

右下の神奈川県の観光スポットに国外の観光客が訪れたケースのみ評価5が最も多い
それ以外は評価4が最も多く、続いて評価5，評価3が多い。

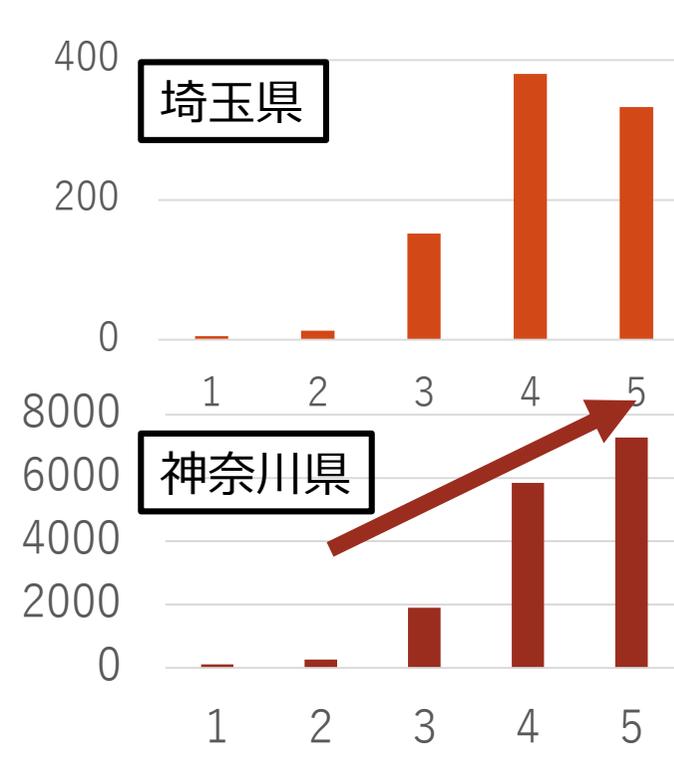
全体(国内外からの観光客)



国内外からの観光客



国外からの観光客



※縦軸が口コミ数，横軸が評価

7.観光スポットの評価の比較

7-2.全体, 国内, 国外の埼玉県と神奈川県の評価の平均の比較

全体	評価	埼玉県	神奈川県
	平均	4.01	4.16

国内	評価	埼玉県	神奈川県
	平均	3.99	4.08

国外	評価	埼玉県	神奈川県
	平均	4.16	4.29

いずれも神奈川県の評価の平均の方が少し高い(差は0.15以下)

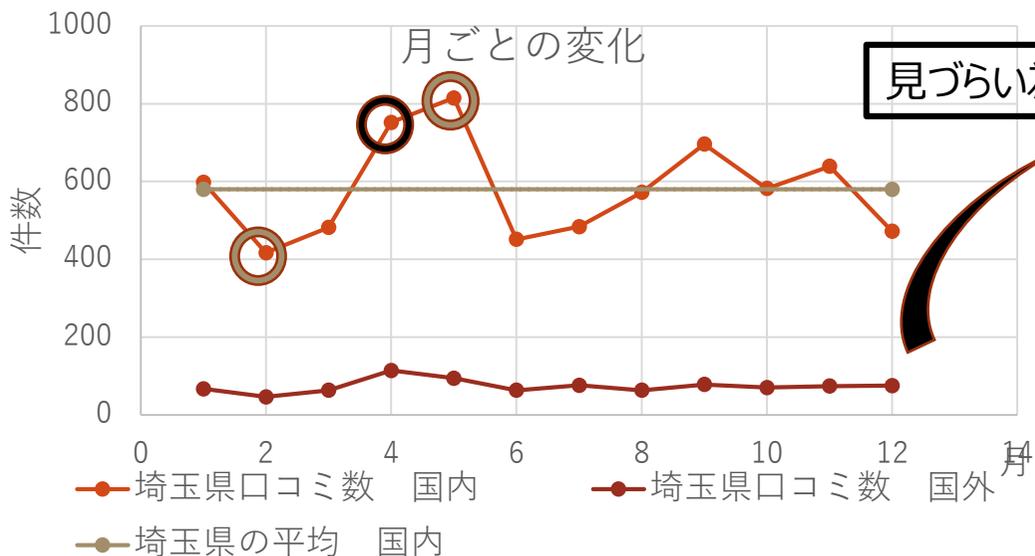
→考察：埼玉県が観光地として人気がない割には評価が高いと感じた。これについて言及するには他の都道府県も調べる必要がある。

また、国内から訪れる観光客よりも国外から訪れる観光客の評価の平均が高いことが分かった。

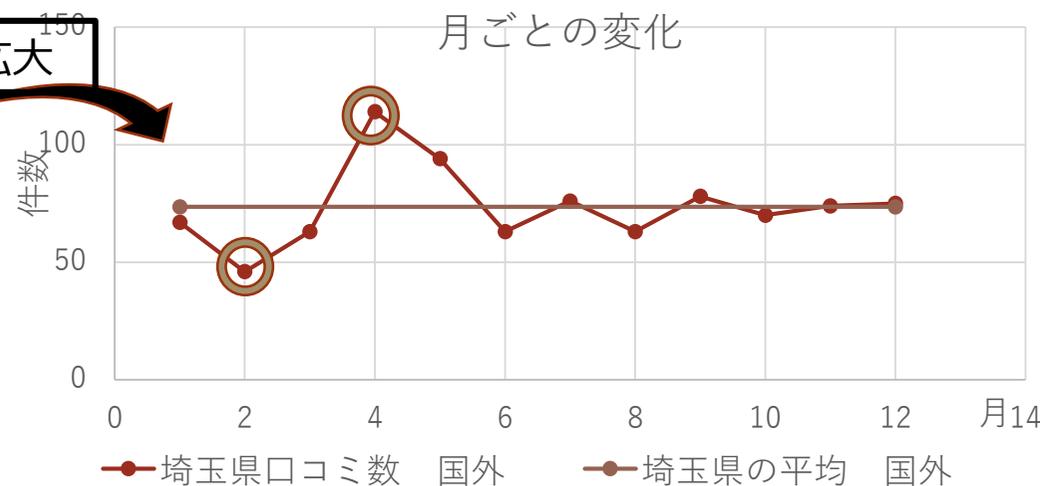
8. 観光客が訪れた月の比較

8-1. 国内外からの観光客が訪れた月 埼玉県・神奈川県

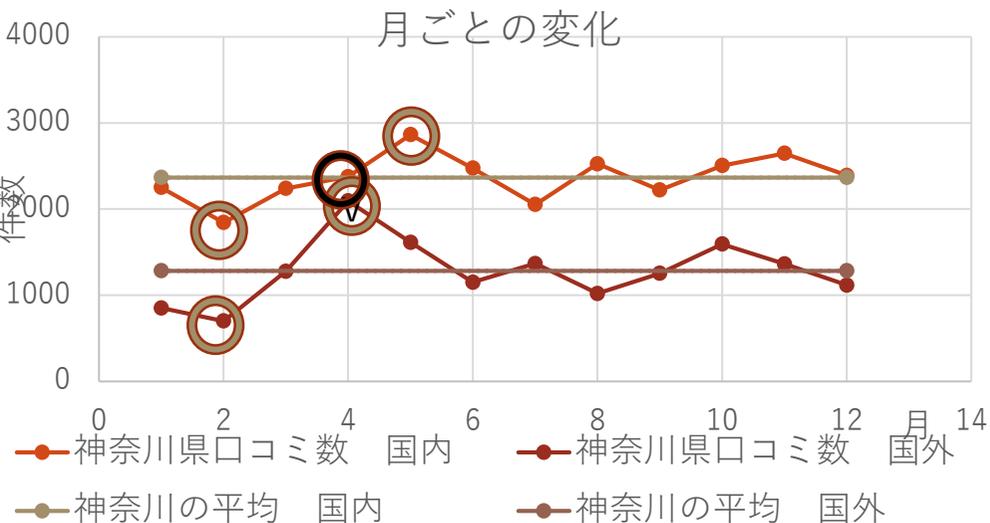
埼玉県 国内外から訪れる観光客の



埼玉県 国外から訪れる観光客の



神奈川 国内外から訪れる観光客の



・国内から訪れる観光客数は埼玉県・神奈川県ともに5月が最も多く、2月が最も少ない。
 ・国外から訪れる観光客数は埼玉県・神奈川県ともに4月が最も多く、2月が最も少ない。
 考察：国内から訪れる観光客はゴールデンウィークを利用して埼玉県・神奈川に訪れるのではないかと考えられる。

・4月は埼玉県を訪れる国内の観光客数が2番目に多い。(神奈川県は平均)

8.観光客が訪れた月の比較

8-2.月ごとの観光客数のばらつき(標準偏差)

全体

	埼玉県	神奈川県
標準偏差	134.5	< 528.0

埼玉県よりも神奈川県の方が月によって観光客数のばらつきが大きく、埼玉県の方が神奈川県よりも年間を通して安定的に観光客が訪れている事が分かる。

埼玉県

	埼玉県 国内	埼玉県 国外
標準偏差	121.0	> 16.4

国外から訪れる観光客数よりも国内から訪れる観光客数の方が月ごとのばらつきが大きく、国外から訪れる観光客の方が国内から訪れる観光客よりも年間を通して、安定的に訪れている事が分かる。

神奈川県

	神奈川 国内	神奈川県 国外
標準偏差	259.0	< 355.5

国内から訪れる観光客数よりも国外から訪れる観光客数の方が月によってばらつきが大きく、国内から訪れる観光客の方が国外から訪れる観光客よりも年間を通して、安定的に訪れている事が分かる。

観光要因の分析のまとめ

6. 観光スポットの時間的距離の比較

- 埼玉県、神奈川県各観光スポットから東京駅までの平均移動時間の差はほとんどないという事が分かる。
- 「東京駅から各観光地スポットまでの移動時間」と「観光スポットの口コミ数」には相関関係がなく、東京駅からアクセスがしやすい事が観光スポットの人気(口コミ数が多い)の大きな理由ではない事が分かる。

7. 観光スポットの評価の比較

- 評価5が最も多いケースは神奈川県を外国人が訪れたケース。
- 全体的に、国内から訪れる観光客よりも国外から訪れる観光客の方がよい評価をつける傾向がある。
- 埼玉県の観光地よりも神奈川県の観光地の方が評価が若干高い。

8. 観光客が訪れた月の比較

- 埼玉県、神奈川県の国内からの観光客は5月に最も訪れており、国外からの観光客は4月に最も訪れている。
- 埼玉県、神奈川県を訪れる国内外の観光客は2月が最も少ない。
- 神奈川県より埼玉県の方が観光客が年間を通して安定的に訪れている。
- 埼玉県は国内から訪れる観光客より国外から訪れる観光客の方が安定的に訪れている。
- 神奈川県は国外から訪れる観光客より国内から訪れる観光客の方が安定的に訪れている。

9. 口コミのテキストマイニング

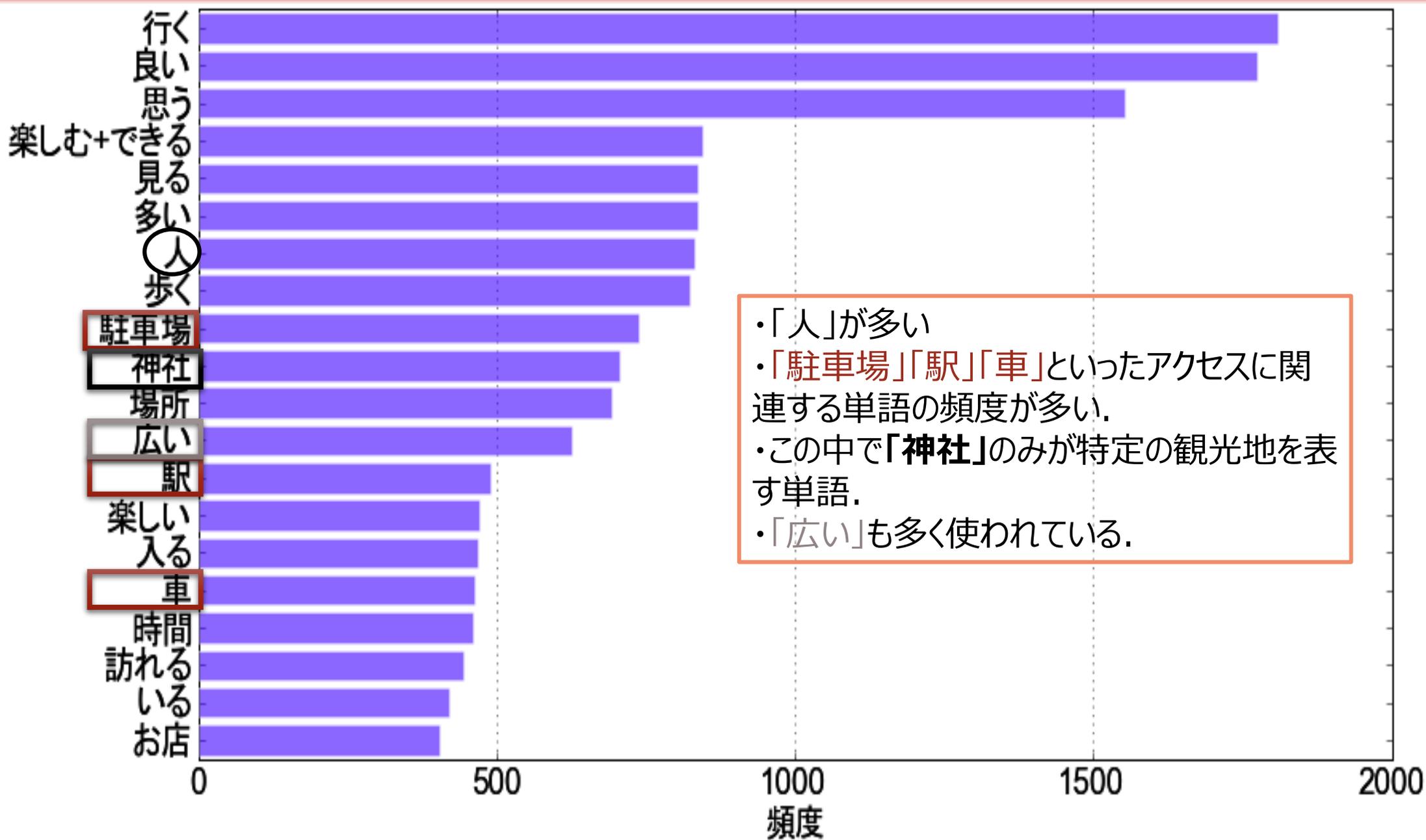
埼玉県・神奈川県観光スポットの特徴を見つけるために
日本語かつ居住地が国内(日本)の口コミのテキストデータを
NTTデータ数理システムのText Mining Studioを用いて

- ・単語頻度解析
- ・係り受け頻度解析

をそれぞれ埼玉県・神奈川県ごとに行った。

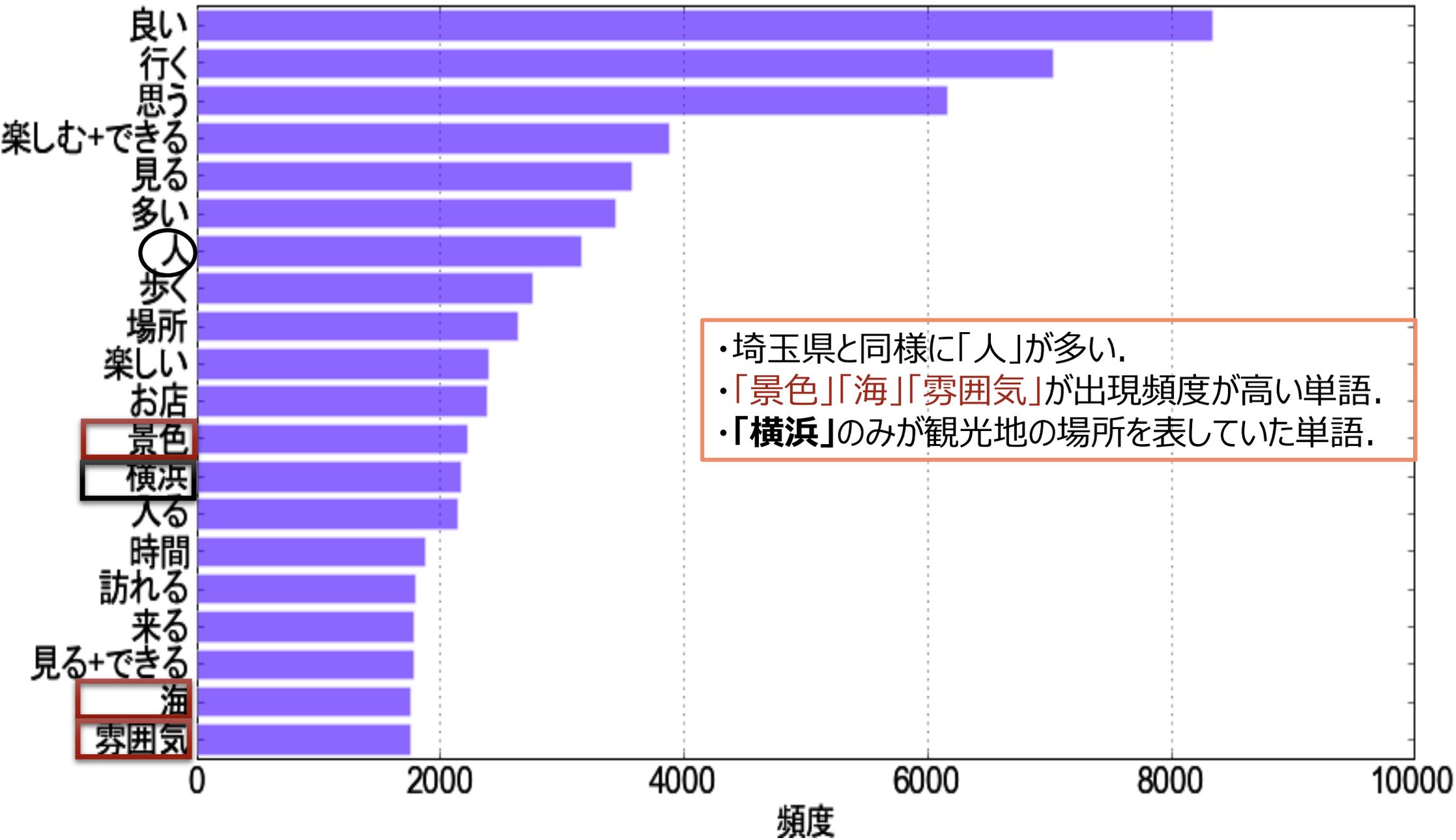
9. 口コミのテキストマイニング

9-1. 単語頻度解析 上位20件 埼玉県編



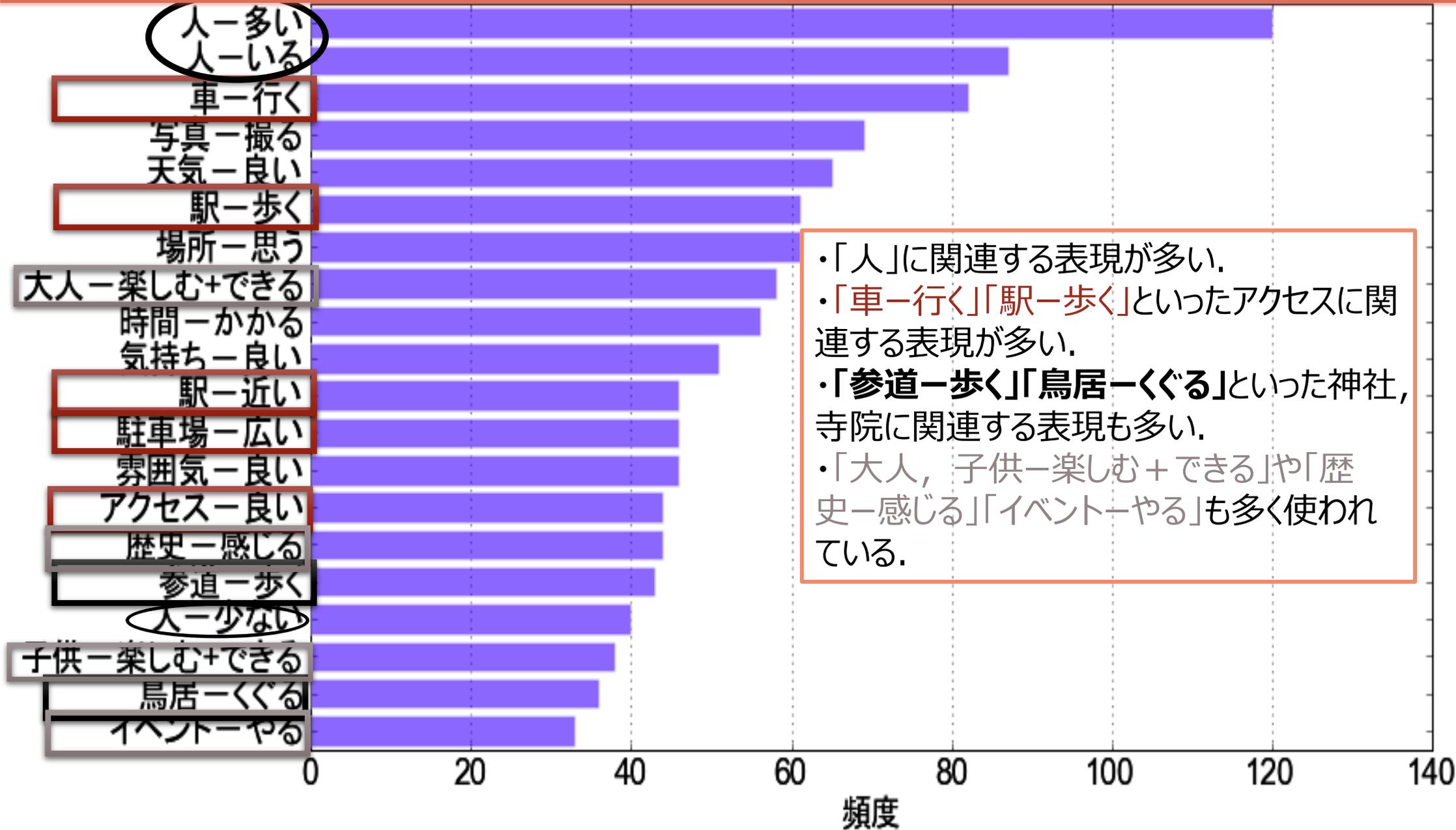
9. 口コミのテキストマイニング

9-2. 単語頻度解析 上位20件 神奈川県編



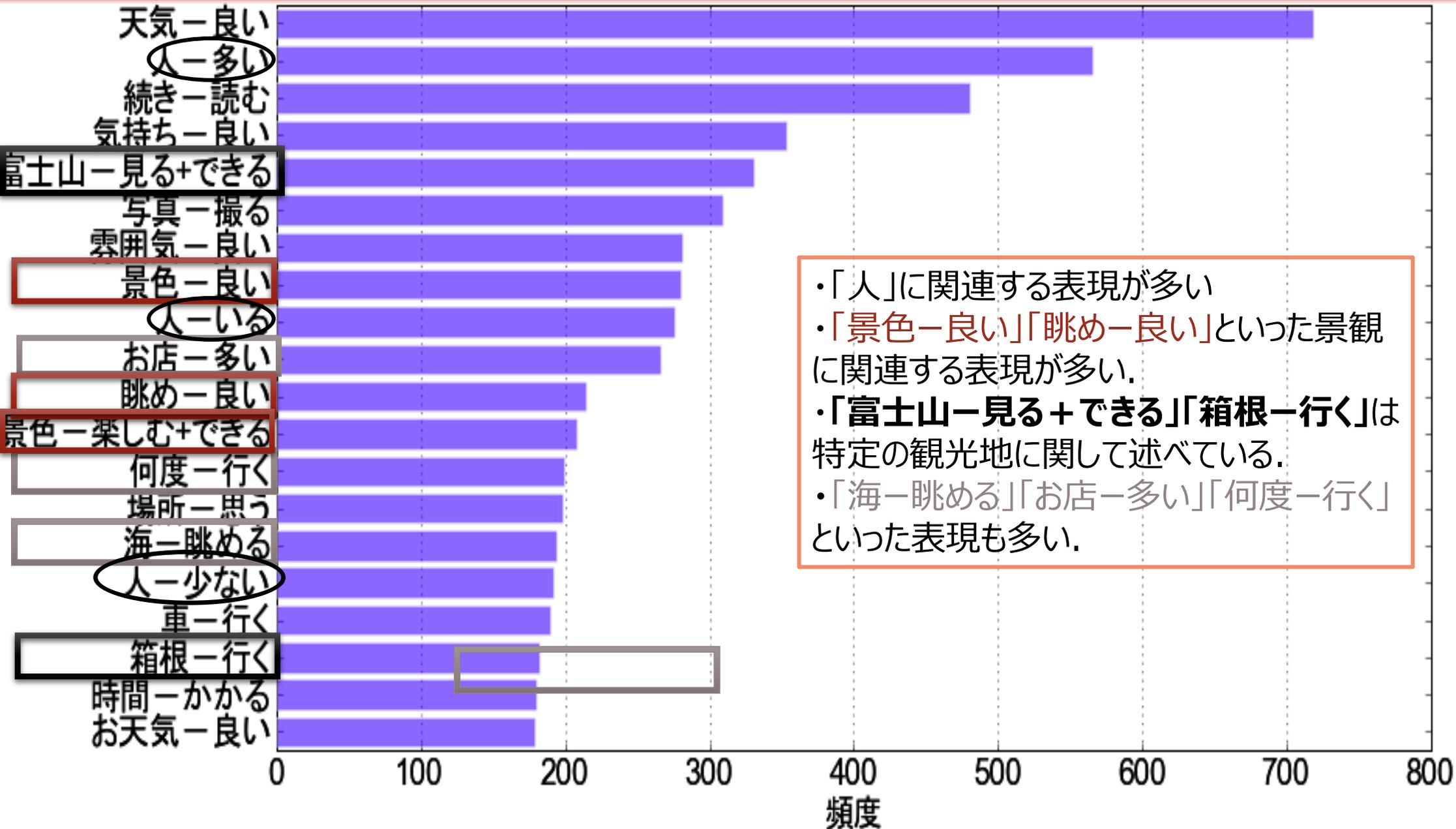
9. 口コミのテキストマイニング

9-3. 係り受け頻度解析 上位20件 埼玉県編



9. 口コミのテキストマイニング

9-4. 係り受け頻度解析 上位20件 神奈川県編



観光客の感想(口コミ)のまとめ

9. 口コミのテキストマイニング

○全体

「人」に関連するものが多い(混雑を気にしている表現が多かった).

○埼玉県

・アクセスに関連するものが特に多い.

また, 神社, 寺院に関連するものも多い.

考察：埼玉県の観光地数が多く、観光スポット同士が離れており、車や電車、バスを用いての移動が頻出するため。

○神奈川県

・景観に関する表現が多い.

10.6カテゴリのアソシエーション分析

埼玉県・神奈川県観光地の「観光スポット」、「観光地」と「観光客の居住地(国)」、「観光客の居住地(都道府県)」の関係や「訪れた月」と「評価」の関係といった普段気づかない事象間の関係を見つけるために、NTTデータ数理システムのVisual Mining Studioの「アソシエーション分析」を用いて分析する。（アソシエーション分析とは、アソシエーションルールと呼ばれる事象間のつながりの強さに関する規則を知識として発見する分析）

以下の4つの指標に沿って分析を進める[8]

①信頼度(A→B)

Aが発生した場合にBが発生する割合。本研究では60%以上のルールのみ。

②サポート(A→B)

全データのうちAとBが同時に発生する割合。本研究では10回以上起きるルールのみ

③lift(A→B)

$(\text{信頼度}(A \rightarrow B)) / (\text{Bが発生する割合})$ 本研究では全てliftが1以上のルールのみ

④ Conviction(A→B)

$(\text{Bが発生しない割合}) / (\text{信頼度}(A \rightarrow B))$ 本研究では全てConvictionが1以上のルールのみ

10.6カテゴリのアソシエーション分析

10-1. 全体像

埼玉県，神奈川県の観光スポット．国内，国外からの観光客を比較する為に
 A: **埼玉・国内**→埼玉県の観光スポットに国内から訪れた観光客
 B: **神奈川・国内**→神奈川県の観光スポットに国内から訪れた観光客
 C: **埼玉・国外**→埼玉県の観光スポットに国外から訪れた観光客
 D: **神奈川・国外**→神奈川県の観光スポットに国外から訪れた観光客
 計15個のアソシエーションルールが見つかった．

番号	前提→	結論	A.埼玉国内	B.神奈川国内	C.埼玉国外	D.神奈川国外
1	観光地or観光地の市町村	評価	A1	×	×	D1
2	観光地or観光地の市町村	月	A2	×	×	×
3	観光地or観光地の市町村	住所(国or都道府県)	A3	B3	×	×
4	住所(国or都道府県)	評価	A4	B4	C4	D4
5	住所(国or都道府県)	月	A5	×	×	×
6	住所(国or都道府県)	観光地or観光地の市町村	×	B6	C6	D6
7	評価	観光地or観光地の市町村	×	×	C7	D7

前提(訪れた観光地, 観光スポット)→結論(評価)

A1. 埼玉県内の観光スポットに国内から訪れた観光客

前提

1. 羽生パーキングエリア 上り

2. 羽生市の観光地

3. さきたま古墳公園

4. 埼玉県営大宮公園

結論

評価 4

この4観光地, 観光スポットを訪れた国内の観光客のうち60%以上が評価4をつけているという事が分かった。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	観光スポット-s32(羽生パーキングエリア 上り)	評価4	65.169	0.833	1.315	1.448
2	観光地-羽生市	評価4	65.169	0.833	1.315	1.448
3	観光スポット-s38(さきたま古墳公園)	評価4	63.514	0.675	1.282	1.383
4	観光スポット-s15(埼玉県営大宮公園)	評価4	62.805	1.480	1.267	1.356

前提(訪れた観光スポット)→結論(評価)

D1. 神奈川県内の観光スポットに国外から訪れた観光客

前提

1. 箱根彫刻の森美術館

2. 長谷寺

7. 横浜ランドマークタワー
展望フロアスカイガーデン

3. 円覚寺

4. 三溪園

5. ポーラ美術館

6. 報国寺(竹寺)

結論

評価5

A1より、埼玉県に訪れた国内の観光客のうち60%以上が評価5をつけている観光スポットはなかったが、神奈川県に訪れた外国人観光客の60%以上が評価5をつけている観光スポットが7つある事が分かった。

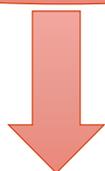
番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	観光スポット-y02(箱根彫刻の森美術館)	評価-5	78.768	7.061	1.667	2.484
2	観光スポット-y14(長谷寺)	評価-5	72.937	2.468	1.543	1.949
3	観光スポット-y48(円覚寺)	評価-5	69.048	0.377	1.461	1.704
4	観光スポット-y16(三溪園)	評価-5	68.285	1.371	1.445	1.663
5	観光スポット-y26(ポーラ美術館)	評価-5	67.500	0.702	1.428	1.623
6	観光スポット-y29(報国寺(竹寺))	評価-5	65.464	0.825	1.385	1.527
7	観光スポット-y25(横浜ランドマークタワー 展望フロア)	評価-5	60.606	1.039	1.282	1.339

前提(訪れた観光スポット)→結論(訪れた月)

A2. 埼玉県 of 観光スポットに国内から訪れた観光客

前提

5. 巾着田曼珠沙華公園



9月

結論

巾着田曼珠沙華公園を訪れる観光客の7割近くが9月に訪れていることが分かった。

[考察]

9月は曼珠沙華(彼岸花)の開花時期. そのために巾着田曼珠沙華公園を訪れる人の7割が9月に訪れるのではないかと考えられる。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	観光スポット-s20(巾着田曼珠沙華公園)	9月	67.442	1.250	6.744	2.764

前提(訪れた観光スポット)→結論(観光客の居住地)

A3. 埼玉県の観光スポットに
国内から訪れた観光客

全て公園

1. 埼玉県営大宮公園

3. 別所沼公園

2. けやきひろば

埼玉県から訪れた人

前提

結論

この3観光スポットに訪れた国内の観光客のうち、
6割以上が県内(埼玉県)から訪れた人。
県外からの観光客に人気がなく、地元向けの観
光スポットともいえる。

B3. 神奈川県内の観光スポットに
国内から訪れた観光客

4. 野毛山動物公園

神奈川県から訪れた人

前提

結論

この観光スポットに訪れた国内の観光客のうち、
6割以上が県内(神奈川県)から訪れた人。
県外からの観光客に人気がなく、地元向けの観
光スポットともいえる。

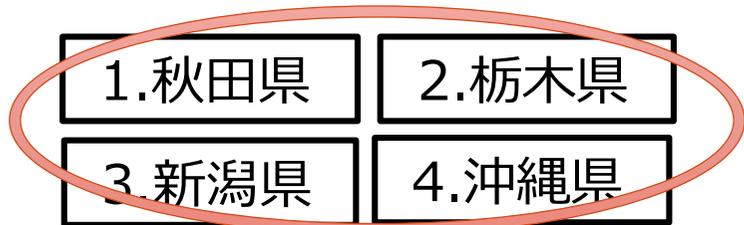
番号	前提：観光スポット	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	観光スポット-s15(埼玉県営大宮公園)	居住地-埼玉県	61.585	1.451	1.910	1.764
2	観光スポット-s24(けやきひろば)	居住地-埼玉県	61.538	1.149	1.909	1.762
3	観光スポット-s42(別所沼公園)	居住地-埼玉県	60.811	0.647	1.886	1.729
4	観光スポット-y32(野毛山動物園)	居住地-神奈川県	60.899	0.955	1.887	1.732

前提(観光客の居住地)→結論(評価)

A4. 埼玉県観光スポットに国内から訪れた観光客

B4. 神奈川県観光スポットに国内から訪れた観光客

前提



結論



この4都道府県から埼玉県を訪れる観光客のうち6割以上が評価4をつけている。

この4都道府県から神奈川県を訪れる観光客のうち6割以上が評価4をつけている。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-秋田県	評価4	73.333	0.158	1.480	1.892
2	居住地-栃木県	評価4	62.500	0.359	1.261	1.345
3	居住地-新潟県	評価4	60.870	0.201	1.228	1.289
4	居住地-沖縄県	評価4	60.000	0.302	1.211	1.261
5	居住地-大分県	評価4	82.353	0.049	1.683	2.894
6	居住地-山形県	評価4	61.702	0.102	1.261	1.334
7	居住地-青森県	評価4	60.714	0.180	1.241	1.300
8	居住地-富山県	評価4	60.563	0.152	1.238	1.295

前提(観光客の居住地)→結論(評価)

C4.埼玉県観光スポットに国外から訪れた観光客

前提

1.ドイツ

2.シンガポール

結論

評価5

評価4

ドイツから埼玉県を訪れる観光客のうち6割以上が評価5をつけている。

シンガポールから埼玉県を訪れる観光客のうち6割以上が評価4をつけている。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-ドイツ	評価-5	64.706	1.246	1.716	1.765
2	居住地-シンガポール	評価-4	60.674	6.116	1.410	1.449

前提(観光客の居住地)→結論(評価)

D4.神奈川県観光スポットに国外から訪れた観光客

前提

1.ハンガリー

2.メキシコ

3.ポーランド

4.ブラジル

9.アルゼンチン

5.スリランカ

6.コスタリカ

7.南アフリカ

結論

評価5

評価4

1~7の国から、神奈川県に訪れる観光客のうち6割以上が評価5をつけている。

アルゼンチンから神奈川県に訪れる観光客のうち6割以上が評価4をつけている。

埼玉県の観光スポットに対しては1か国のみ評価5をつけた人が6割以上の国があったが、神奈川県観光地に対して、7か国も評価5をつけた人が6割以上の国が見つかった。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-ハンガリー	評価-5	83.333	0.130	1.763	3.164
2	居住地-メキシコ	評価-5	77.586	0.292	1.642	2.353
3	居住地-ポーランド	評価-5	71.429	0.195	1.511	1.846
4	居住地-ブラジル	評価-5	68.421	0.253	1.448	1.670
5	居住地-スリランカ	評価-5	66.667	0.091	1.411	1.582
6	居住地-コスタリカ	評価-5	65.000	0.084	1.375	1.507
7	居住地-南アフリカ	評価-5	61.538	0.156	1.302	1.371
8	居住地-アルゼンチン	評価-4	64.286	0.117	1.694	1.737

前提(観光客の居住地)→結論(訪れた月)

A5. 埼玉県の観光スポットに国内から訪れた観光客

前提

1. 山形県

山形県から埼玉県に訪れる観光客の8割近くが12月に訪れている。

結論

12月

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-山形県	12月	78.571	0.158	11.586	4.350

前提(観光客の居住地)→結論(訪れた観光地)

B6.神奈川県観光スポットに国内から訪れた観光客

前提

1.秋田県

2.鹿児島県

3.長野県

4.島根県

結論

横浜市

箱根町

秋田県, 鹿児島県, 長野県から神奈川県に訪れる観光客の6割以上が横浜市を訪れている。

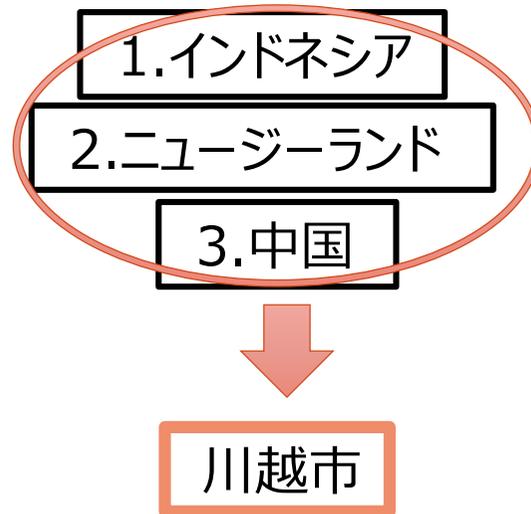
島根県から神奈川県に訪れる観光客の6割以上が箱根町を訪れている。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-秋田県	観光地-横浜市	69.388	0.120	1.446	1.700
2	居住地-鹿児島県	観光地-横浜市	64.286	0.159	1.340	1.457
3	居住地-長野県	観光地-横浜市	62.032	0.409	1.293	1.370
4	居住地-島根県	観光地-箱根町	61.111	0.039	2.403	1.918

前提(観光客の居住地)→結論(訪れた観光地)

C6. 埼玉県の観光スポットに国外から訪れた観光客

前提



結論

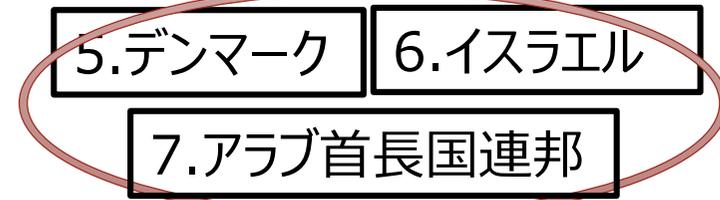
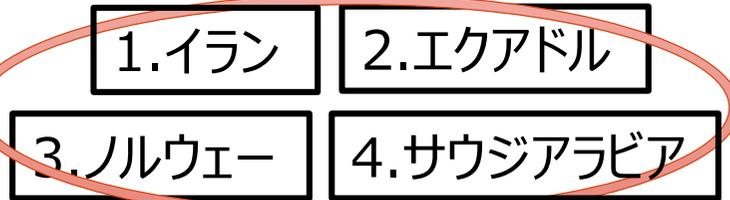
インドネシア, ニュージーランド, 中国から埼玉県に訪れる観光客の6割以上が川越市を訪れている。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-インドネシア	観光地-川越市	72.000	2.039	1.499	1.856
2	居住地-ニュージーランド	観光地-川越市	68.421	1.472	1.425	1.646
3	居住地-中国	観光地-川越市	60.465	14.723	1.259	1.315

前提(観光客の居住地)→結論(訪れた観光地)

D6.神奈川県観光スポットに国外から訪れた観光客

前提



結論

横浜市

箱根町

1~4の国から神奈川県を訪れる観光客の6割以上が横浜市を訪れている。

5~7の国から神奈川県を訪れる観光客の6割以上が箱根町を訪れている。

1.4.6.7は同じ中東だが訪れる地域が異なる。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	居住地-イラン	観光地-横浜市	91.667	0.071	2.796	8.066
2	居住地-エクアドル	観光地-横浜市	84.615	0.071	2.581	4.369
3	居住地-ノルウェー	観光地-横浜市	66.667	0.104	2.033	2.016
4	居住地-サウジアラビア	観光地-横浜市	63.158	0.078	1.926	1.824
5	居住地-デンマーク	観光地-箱根町	81.818	0.175	1.868	3.092
6	居住地-イスラエル	観光地-箱根町	68.153	0.695	1.556	1.765
7	居住地-アラブ首長国連邦	観光地-箱根町	66.667	0.221	1.522	1.686

前提(評価)→結論(訪れた観光地)

C7. 埼玉県観光スポットに
国外から訪れた観光客D7. 神奈川県観光スポットに
国外から訪れた観光客

前提

1. 評価3

2. 評価1

結論

川越市

箱根町

国外から埼玉県を訪れた観光客のうち、6割以上が川越市に評価3をつけている。

6-2より国外から埼玉県を訪れる観光客のうち「川越市」を訪れる観光客は全体の約1/2。そのため、他の埼玉県の観光地に比べ若干「川越市」に対して「評価3」をつける外国人観光客が多いという事が分かった。

国外から神奈川県を訪れた観光客のうち、約65%が箱根市に評価3をつけている。

6-3より国外から神奈川県を訪れる観光客のうち「箱根町」を訪れる観光客は全体の約25%。そのため、他の神奈川県の観光地に比べ「箱根町」に対して「評価1」をつける外国人観光客が多く、箱根町は外国人観光客から評価が低い観光地といえる。

番号	前提	結論	信頼度(%)	サポート(%)	Lift	Conviction
1	評価-1	観光地-箱根町	64.602	0.474	1.475	1.588
2	評価-3	観光地-川越市	61.184	10.532	1.274	1.339

11. 「観光客の居住地」と「観光地」の対応分析

対応分析は質的変数間の連関を視覚化することができる。

埼玉県と神奈川県「観光客の居住地(国内)」と「観光地」の連関を詳細に検討するために、対応分析(コレスポネンス分析)を実行した。

第2軸までの累積寄与率は84.6%であり、データがもつ情報のほとんどが平面上に集約されていると解釈することができる。今回は「観光客の居住地(国内)」のカテゴリに注目する。

	第1軸	第2軸	第3軸
固有値	0.2	0.037	0.009
累積寄与率(%)	71.2	84.6	87.8

ctrとcos2の説明[9]

ctr

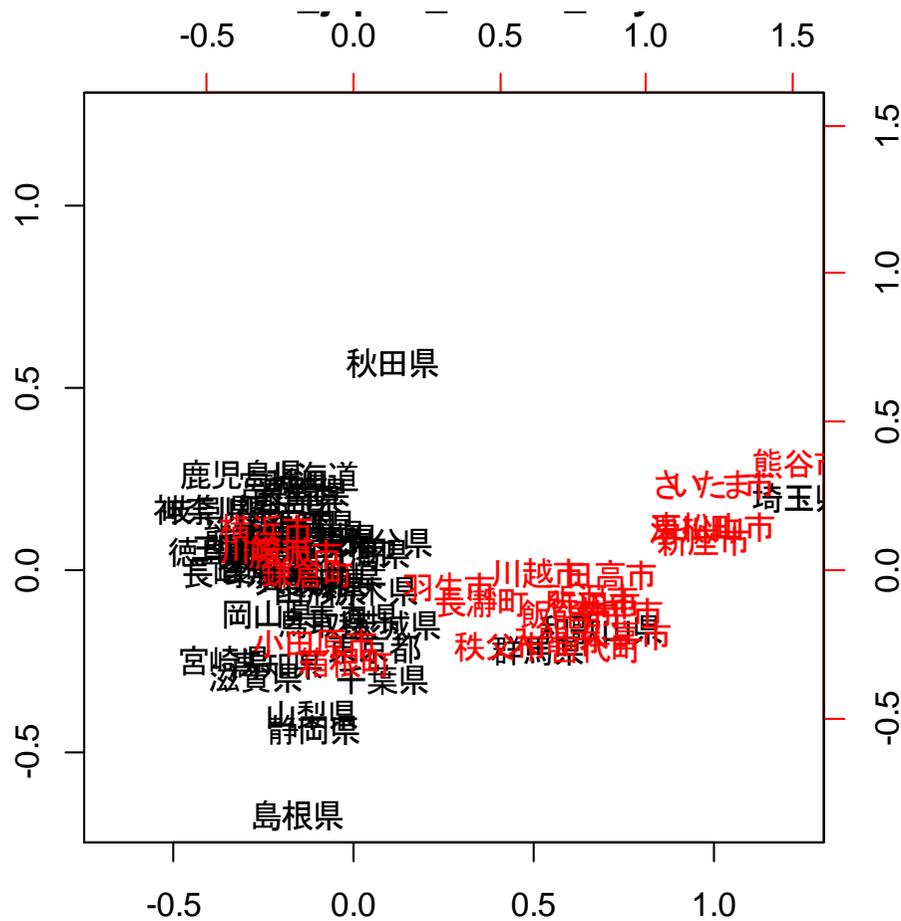
ctrは各カテゴリの軸への寄与率を表し、パーセンテージによる表示(各軸のctrの和は100)。この値が大きいカテゴリほど、その軸に強く寄与していると考えられる。

cos2

cos2は、軸が各カテゴリにどの程度寄与しているかを表す指標であり、平方相関と呼ばれる。平方相関が高いカテゴリほど軸によってよく説明されていると解釈できる。

11. 「観光客の居住地」と「観光地」の対応分析

11-1.Rで対応分析



プログラミング言語のRを用いて対応分析を行った所, 左の図のようになった.

この図(バイプロット)では47都道府県と23の市町村の計70個の要素をプロットしているため被ってしまい, 見づらいという問題が生じた.

11. 「観光客の居住地」と「観光地」の対応分析

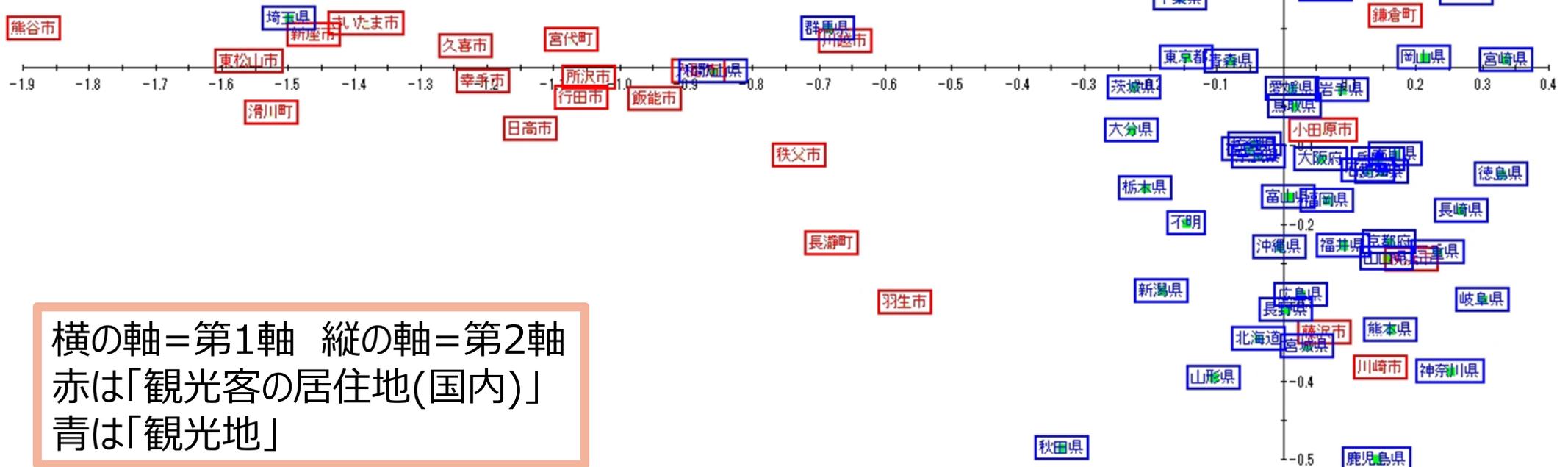
11-2. Visual Mining Studioで対応分析

NTTデータ数理システムのVisual Mining Studioで対応分析を行った所、下の図のようになった。

こちらでは同じく47都道府県と23の市町村の計70個の要素をプロットしているが、ほとんど被らずに見ることができる。

また、Visual Mining Studioでは

- 対応分析のためにクロス集計表を作成する必要がない
 - 簡単にプロットの色を変更することが出来る
- といった利点があった。そのため以後この図を使用する。



11. 対応分析

11-3. 観光客の居住地(国内)のカテゴリに注目 第1軸・第2軸

第1軸

number	都道府県	cos2(1軸)
17	埼玉県	0.972
26	神奈川県	0.809

number	都道府県	ctr(1軸)
17	埼玉県	73.847
26	神奈川県	20.386

ctr

- ・「埼玉県」が73%で高く1軸に寄与している.
- ・「神奈川県」も1軸に寄与している.
- ※ctrが10以上のものを表示

cos2

- ・「埼玉県」「神奈川県」は第1軸によってよく説明されている.
- ※cso2が0.7以上のものを表示

第2軸

number	都道府県	ctr(2軸)
37	東京都	33.704
26	神奈川県	21.321
17	埼玉県	10.97
30	千葉県	10.85

number	都道府県	cos2(2軸)
37	東京都	0.88

ctr

- ・「東京都」が最も2軸に寄与している.
- ・「神奈川県」「埼玉県」「千葉県」も2軸に寄与している.
- ※ctrが10以上のものを表示

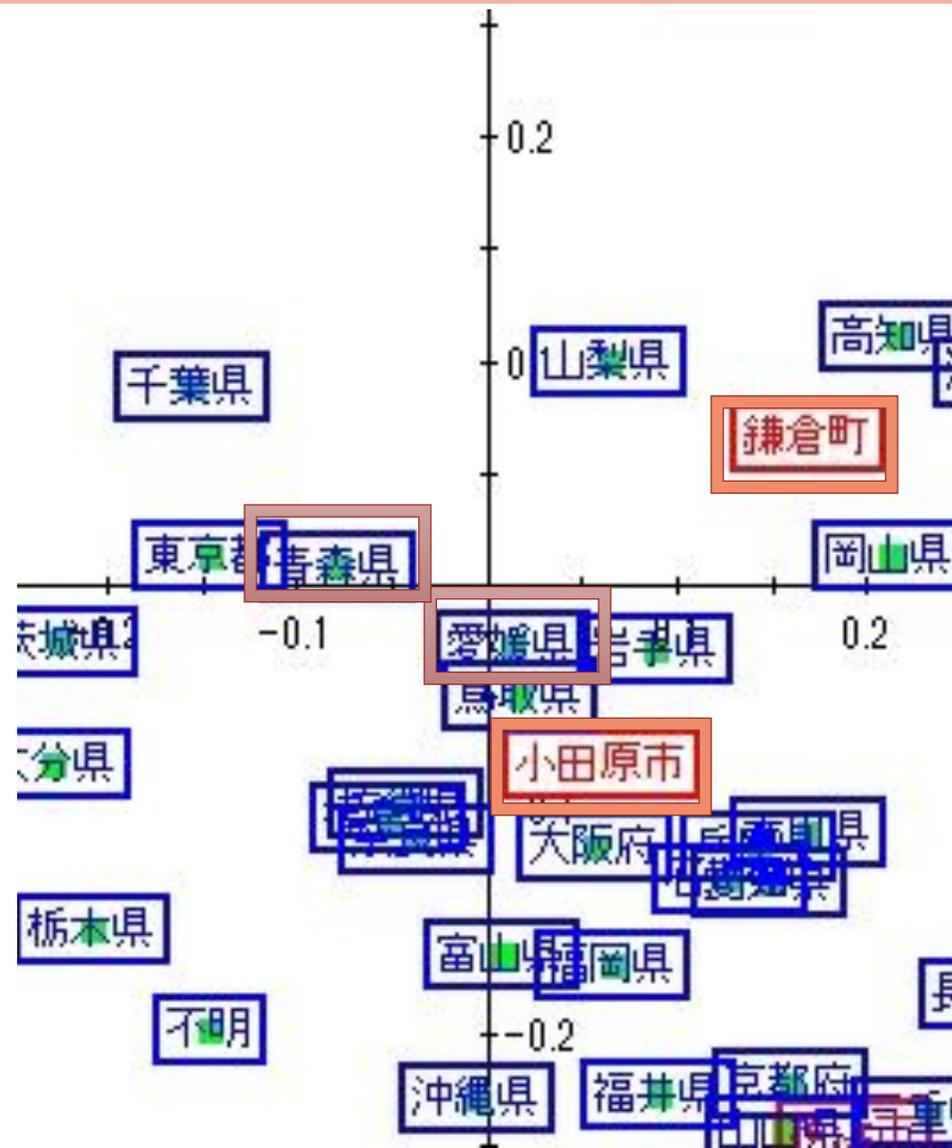
cos2

- ・「東京都」は第2軸によってよく説明されている.
- ※cso2が0.7以上のものを表示

11. 「観光客の居住地」と「観光地」の対応分析

11-6. 原点付近

原点に近くプロットされているものほど平均的であることを意味しており、「愛媛県」「青森県」が平均的な「観光客の居住地(国内)」, 「小田原市」「鎌倉市」が平均的な「観光地」といえる。



複数のカテゴリを扱う分析のまとめ

10. アソシエーション分析

- 埼玉県の観光地を訪れる観光客のうち60%以上が評価5をつけている観光地(スポット)はなかったが、神奈川の観光地を訪れる観光客のうち60%以上が評価5をつけている観光スポットは7つあった。
- ドイツ人の60%以上が埼玉県の観光スポットに対して評価5をつけている。
- 7カ国のそれぞれ60%以上が神奈川県の観光スポットに対して評価5をつけている。
- 他の埼玉県の観光地に比べ若干、「川越市」に対して「評価3」をつける外国人観光客が多い。
- 他の神奈川県の観光地に比べ「箱根町」に対して「評価1」をつける外国人観光客が多く、箱根町は外国人観光客から評価が低い観光地といえる。

11. 「観光客の居住地」と「観光地」の対応分析

- 埼玉県の観光地には埼玉県に住む人が訪れる傾向がある。
- 神奈川県の観光地には神奈川県に住む人が訪れる傾向がある。
- 「訪れる観光客」と「観光地」の対応分析で平均的な観光地は、「小田原市」「鎌倉市」。平均的な観光客の居住地は「愛媛県」「青森県」「岩手県」という事が分かった。

12. 埼玉県が観光地として人気がない理由と施策の提案

理由 1

埼玉県は国内から訪れる観光客に依存し、特に南関東から訪れる観光客に依存している。

施策の提案

- ・海外や他県に住む人に埼玉県の観光スポットを知ってもらう取り組み
(コストが安く済み、世界中の人から見てもらえるsnsを利用したPR活動等)

理由2

観光地数が多く、観光スポット同士が離れており、アクセスを気にした口コミが多いことから複数の観光地を回ることが難しい。

施策の提案：

- ・離れた観光地でもまわりやすいようにバスツアーを開催
- ・1つの観光スポットで1日満喫できるように、既存の大型観光スポット(東武動物公園、西武遊園地ゆうえんち)の宣伝

13. 今後の課題

- 今回は日本語で書かれた口コミのみテキストマイニングを行ったが、NTTデータ数理システムのText Mining Studioの英語アドオンでは英語で書かれた口コミもテキストマイニングできるため、利用したい。他にも、Rやpythonで分析を行うより、Vms, Tmsは手軽に見やすい分析を行う事が出来たため、今回使用しなかったツールも研究の内容を深めるために使用していきたい。

14. 参考文献

- [1] コロナ後に観光で行きたい都道府県ランキング【2020完全版】[ブランド総合研究所 ダイヤモンド・セレクト編集部](https://diamond.jp/articles/-/247551)<https://diamond.jp/articles/-/247551>
- [2] Tripadvisor <https://www.tripadvisor.jp/>
- [3] 「「旅行・交通」カテゴリwebサイトのpcからの閲覧者数」 公益社団法人日本観光振興協会
<https://www.nihonkankou.or.jp/home/userfiles/files/autoupload/kankousite2018.pdf>
- [4] 「旅行ガイドブックと口コミの言語解析による訪日外国人の観光地に対するイメージに関する研究」大久保立樹・室町泰徳(2014年10月)
https://www.jstage.jst.go.jp/article/journalcpj/49/3/49_573/pdf/-char/ja
- [5] 「日光，箱根における外国人観光者と日本人観光者の評価特徴分析」野瀬 元子，古屋 秀樹(2008年10月)
https://www.jstage.jst.go.jp/article/journalcpj/43.3/0/43.3_595/pdf/-char/ja
- [6] ジョルダン <https://www.jorudan.co.jp/>
- [7] 都道府県市区町村 ランキングデータ<https://uub.jp/rnk/rnk.cgi?T=p&S=j&B=20200401>
- [8] 「Visual Mining Studio 体験セミナー 基本操作 + 顧客分析」
株式会社NTTデータ数理システム データマイニング部(2019年)
- [9] 「Rによる多変量解析入門」OME社 川端一光，岩間徳兼，鈴木雅之

15.付録 今回使用した埼玉県の観光スポット口コミ数上位1~50位

順位	埼玉県の観光地	順位	埼玉県の観光地
1	鉄道博物館	26	東武動物公園
2	川越一番商店街	27	コクーンシティ
3	武蔵一宮 氷川神社	28	西武園ゆうえんち
4	時の鐘	29	大宮ソニックシティ
5	三峯神社	30	国営武蔵丘陵森林公園
6	菓子屋横丁	31	西武秩父仲見世通り
7	イオンレイクタウン	32	羽生パーキングエリア 上り
8	喜多院	33	こども動物自然公園
9	川越氷川神社	34	NACK5スタジアム大宮
10	羊山公園	35	川越まつり会館
11	秩父神社	36	道の駅ちちぶ
12	埼玉スタジアム2002	37	川越熊野神社
13	三井アウトレットパーク 入間	38	さきたま古墳公園
14	長瀬石畳	39	忍城址
15	埼玉県営大宮公園	40	宝登山
16	メットライフドーム	41	古代蓮の里
17	所沢航空記念公園	42	別所沼公園
18	川越城本丸御殿	43	百観音温泉
19	さいたま市大宮盆栽美術館	44	トーベ・ヤンソンあけぼの自然公園
20	巾着田曼珠沙華公園	45	高麗神社
21	宝登山神社	46	ムーミンバレーパーク
22	さいたまスーパーアリーナ	47	平林寺
23	権現堂桜堤	48	大宮盆栽村
24	調神社	49	妻沼聖天山歓喜院
25	けやきひろば	50	大宮公園小動物園

15.付録 今回使用した神奈川県観光スポット口コミ数上位1~50位

順位	神奈川県観光地	順位	神奈川県観光地
1	横浜中華街	26	ポーラ美術館
2	箱根彫刻の森美術館	27	横浜ランドマークタワー
3	芦ノ湖	28	横浜クルージング
4	高德院(鎌倉大仏)	29	報国寺(竹寺)
5	横浜みなとみらい21	30	新江ノ島水族館
6	箱根ロープウェイ	31	港の見える丘公園
7	横浜赤レンガ倉庫	32	野毛山動物園
8	大涌谷園地	33	よこはまコスモワールド
9	箱根登山電車(箱根登山鉄道)	34	箱根小涌園ユネッサン
10	カップヌードルミュージアム 横浜	35	横浜スタジアム
11	山下公園	36	川崎市 藤子・F・不二雄ミュージアム
12	横浜港大さん橋 国際客船ターミナル	37	ラゾーナ川崎プラザ
13	箱根神社 / 九頭龍神社新宮	38	強羅公園
14	長谷寺	39	日本郵船 氷川丸
15	鶴岡八幡宮	40	川崎大師 平間寺
16	三溪園	41	星の王子さまミュージアム
17	江ノ電	42	仙石原
18	小田原城	43	よこはま動物園ズーラシア
19	横浜・八景島シーパラダイス	44	建長寺
20	新横浜ラーメン博物館	45	箱根関所・箱根関所資料館
21	鎌倉小町通り	46	箱根 駒ヶ岳ロープウェイ
22	江ノ島	47	横浜ワールドポーターズ
23	横浜元町商店街	48	円覚寺
24	箱根ガラスの森美術館	49	銭洗弁財天宇賀福神社
25	横浜ランドマークタワー 展望フロア スカイガーデン	50	明月院(あじさい寺)