

テキストマイニングを用いた観光地活性化の提言

佳
作
(2)

学 年：3 年
学部学科：経営学部経営学科
氏 名：湯浅寿々丸

学 年：3 年
学部学科：経営学部経営学科
氏 名：笹岡 建吾

学 年：3 年
学部学科：経営学部経営学科
氏 名：高山 美来

要 約

めまぐるしく技術の発展が進む昨今、IT 革命によるインターネットの普及によって我々のライフスタイルは大きく変化してきた。その中で観光ビジネスに注目すると、宿泊予約サイトの登場があげられる。その中で各地域の温泉街にスポットを当て、近畿圏である兵庫県の有馬温泉と、日本温泉ランキング 1 位の群馬県にある草津温泉の違いについて、両温泉の口コミデータに対しテキストマイニング分析を用いることで明らかにしたうえで、地元有馬を活性化し、有馬がより良い観光地となるための案を提言する。

本研究では恣意的な分析になる可能性を含む「手作業」を排除するために、KH coder によるテキストマイニング分析である共起ネットワーク分析を用いた計量分析を用いて分析を行った。分析の結果、共起ネットワーク分析では両温泉の価格帯に大きな差異は無いにも関わらず、有馬温泉では価格に対し満足度が低くさらに坂道が多いなど、価格以外でもネガティブな内容が目立った。対して草津温泉は、土産屋が充実し食べ歩きに適した店が多いなど温泉以外にも力を入れることで、満足度が高いことが明らかとなった。そのうえで、両者共起の違いをもとに有馬の現状は外観や施設に手を加えることでさらなる観光地へと変貌を遂げる可能性を示唆している。

目 次

1 章	はじめに
2 章	有馬温泉と草津温泉の観光力
2-1	草津温泉と有馬温泉の外観
2-2	草津温泉と有馬温泉の観光力の比較
2-3	草津温泉と有馬温泉のレビューの比較
2-4	本研究の大まかな流れの枠組み
3 章	分析手法、分析結果
3-1-1	語句の抽出
3-1-2	共通項と期待値による検定
3-2	共起ネットワーク分析
3-2-1	草津温泉のお客様レビュー
3-2-2	有馬温泉の共起ネットワーク
3-3	考察
4 章	まとめ

1 章 はじめに

医療技術の発展や人工知能の発達が著しく進む昨今は、時代の流れを肌で感じることができる。この発展・発達はさまざまな革命によって革新的に進んできた。例えば、産業革命によって移動手段が馬から列車になり、連絡手段では電話が用いられた。このような変化は大小様々な規模で我々の身近におきている。中でも IT 革命によるインターネットの普及は、行政のあり方や我々のライフスタイルに多大な変化をもたらした。観光ビジネスに注目してみると、宿泊の予約サイトがあげられる。以前は宿泊の希望を電話で直接宿に伝える、または旅行会社に行き、そこで予約するという二択が主な手段であった。しかし宿泊予約サイトが登場してからは、自宅や移動先で手軽に宿泊先を検索し情報を入力するだけで予約が完了するため、直接電話で話す煩わしさもなく旅行会社に足を運ぶ手間もかからない。もちろん電話予約、旅行会社共に機能しているので、必要に応じて使い分けも可能だ。また多くの企業がアプリ内の広告などで観光地や宿泊先の情報が表示させているため、ネットワーク利用者の目に観光地や宿泊先のバナーが目にとまり、思い立って予約サイトにアクセスするというケースも少なくない。

現在、全国各地には観光スポットとよばれる名所が存在する。近畿県内で例をあげると、和歌山県のアドベンチャーワールドや大阪のユニバーサルスタジオジャパン、京都の清水寺などがあげられる。

本研究では、観光地の中でも、地元の活性化を狙いとして三大名湯である有馬温泉をとりあげ、同じく三大名湯であり、温泉ランキング 1 位である草津温泉と様々な観点から比較し、有馬がより良い観光名所となる為の案を提言する。

2 章 有馬温泉と草津温泉の観光力

本章では、分析対象となる二つの温泉地について、その概観を述べる。

2-1 草津温泉と有馬温泉の外観

草津温泉（くさつおんせん）は、群馬県吾妻郡草津町草津界隈に所在する温泉である。場所は群馬県北西部であり、県庁所在地の前橋市からは 70.6km ほど離れている。地図上の具体的な場所は図 1 に示されている通りである。

図1 草津温泉の位置



出典：白地図から筆者作成

日本を代表する名泉（名湯）の一つであり、室町時代の禅僧、歌人である万里九集と江戸時代初期の朱子学派偶学者である林羅山によって日本三名泉の一つに数えられた

同温泉の敷地の広さは、温泉に関連する宿泊施設をマップ上で線によって繋ぎ合わせて算出した結果 7,750㎡であった。また、当地にある宿泊施設数は草津温泉旅館協同組合のウェブサイトによると 15 施設存在している。宿泊施設は概ね高級な旅館が占めており、「湯宿 季の庭」、「望雲」、「奈良屋」などが高級旅館として有名である。これらの高級旅館の平均宿泊料金は平均では 56,066 円とかなり高額であった¹⁾。また、草津温泉の泉質は基本的には酸性泉であるが、場所によっては硫黄泉も見られる。この泉質の適応症状としては皮膚病、神経痛、糖尿病などがあり、主に効能目当てで行くとすれば「糖尿病ネットワーク」により糖尿病の患者は 70 歳前後が最も多いとされるのでそれらの年代の客層が多いことが伺える。「宿泊予約サイトじゃらん」の口コミの投稿件数は家族が最多であった。来客数は同温泉地の来客数の推移を見てみると、表 1 に要約されている。2007 年から 2016 年にかけて、来客数は平均で約 280 万人であった。また、この規模は長期的に見ると微増の傾向があった。「産経・WEST」によると有名な遊園地であるハウステンボスの 2015 年～

1) ここでの料金は各宿の「和室で朝・夕食付で貸切露天風呂付」という条件のもとでの料金の平均値である。

2016 年の入場客数が 307 万人であったが、この規模は非常に大きいと言える。

表1 草津温泉の来客数推移

草津温泉	単位（千人）
2007	2808
2008	2691
2009	2645
2010	2658
2011	2687
2012	2765
2013	2807
2014	2809
2015	2965
2016	3074

出典：群馬県観光客動態調査報告書（各年版）より筆者作成

有馬温泉（ありまおんせん）は、兵庫県神戸市北区有馬町に所在する温泉である。場所は兵庫県南東部であり、県庁所在地の神戸市からは 21km ほど離れている。地図上の具体的な場所は図 2 に示されている通りである。

図 2 有馬温泉の位置



出典：白図より筆者作成

草津温泉と同様で日本を代表する名泉（名湯）の一つであり、枕草子の日本三名泉にも数えられた。同温泉の敷地の広さを草津温泉と同様に算出したところ、結果 6,500㎡であった。また、当地にある宿泊施設数は 31 存在している。その中でも高級な宿泊施設として有名な「兵衛向陽閣」、「銀水荘 兆楽」、「有馬グランドホテル」の平均宿泊料金は 49,989 円であった。金額的には草津温泉のそれとあまり大きな差は見られなかった。また、有馬温泉の泉質は湧出場所により異なり、塩分と鉄分を多く含む含鉄塩化物泉、ラジウムを多く含む放射能泉、炭酸を多く含む炭酸水素塩泉の 3 種類がある。この泉質の適応症状としては炎症性および非炎症性のリマウチ疾患、硬直性脊椎症、高血圧症などがあり、おもに健康に興味があり山登りを好む人や団塊の世代である 67 歳～69 歳のという客層に人気があることが予想できる。また、「宿泊予約サイトじゃらん」の口コミによると草津温泉と同様に家族客からの口コミが最多であった。来客数は同温泉の来客数の推移は表 2 の通りである。2007 年から 2016 年にかけて、来客数は平均で 73 万人であった。草津温泉とは異なり趨勢的にはほぼ変わらず一定レベルの来客数が毎年存在していることがうかがえる。

表2 有馬温泉の来客数推移

有馬温泉	単位（千人）
2007	694
2008	679
2009	666
2010	735
2011	720
2012	747
2013	799
2014	723
2015	780
2016	782

出典：兵庫県観光客動態調査報告書より筆者作成

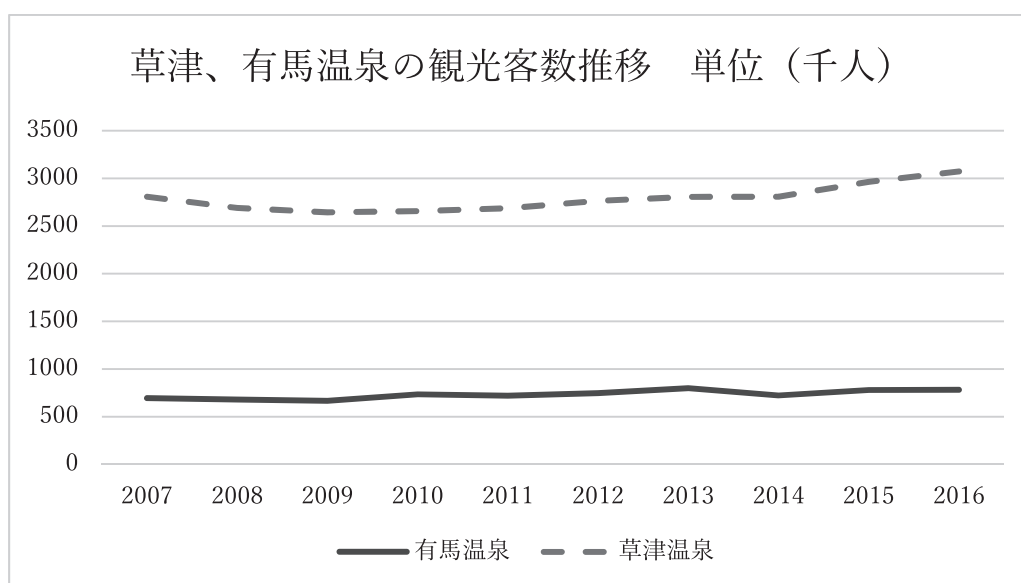
2-2 草津温泉と有馬温泉の観光客数の比較

草津温泉と有馬温泉は同じ日本三名湯ではあるが、宿泊施設数で比較すると草津温泉が 15 に対し、有馬温泉が 31 とほぼ倍の規模であった。交通の便で比較すると草津温泉は東京駅から鉄道、高速バス共に利用可能であり所要時間は平均して 3 時間 45 分ほどであった。有馬温泉も大阪駅から鉄道、高速バス共に利用可能であり所要時間は平均して 60 分ほどであった。面積で比較すると、草津温泉が 7,750㎡に対し、有馬温泉が 6,500㎡とほぼ

同規模であった。これらのことから、宿泊施設の立地密度は有馬温泉のほうが高いと言える。両温泉地の来客数の推移を比較した図2を見ると、4倍近い差が存在していることがわかる。

両温泉とも知名度は高く、外形的にはそれほどの違いがあるとは思えない。宿泊施設数の差があるとはいえ、なぜ来客数にここまでの差がついているのであろうか。草津温泉と有馬温泉を比較して有馬温泉の観光力に不足しているものは何かを検証するため、本稿では観光客のユーザーレビューから来客数の差について考察を行っていく。

グラフ1



2-3 草津温泉と有馬温泉のレビュー比較

草津温泉と有馬温泉のレビューについて以下に旅行に関する口コミサイトより引用したユーザーレビューの例を示す。これは、実際に「トリップアドバイザー」に投稿された草津温泉と有馬温泉を訪問した旅行客の口コミである。口コミの内容はあくまでその観光客の主観的な意見が述べられているが、それぞれの温泉をどのように感じたか読み取ることが可能である。

有馬温泉の口コミから読みとれる意見として、

- ・ 豊臣秀吉が訪問した温泉
- ・ 歴史があり情緒ある温泉
- ・ 坂道が多い
- ・ 宿泊費が高い

草津温泉の口コミから読みとれる意見として、

- ・夜のライトアップが美しい
- ・硫黄のにおいがする
- ・バイキングがおいしい
- ・幻想的であるといった内容が見られた。

例1 有馬温泉の口コミの例

太閤秀吉に所縁のある温泉街です。

神戸や大阪からそれほど遠くないので行きやすいです。

その場所柄が宿泊料はちょっと高めのイメージがあります。

昔ながらの温泉街にはお土産物屋さんが並び歩いていると楽しいです。

ただ坂道も多いので歩きやすい靴の方が楽だと思います。

日帰り入浴出来る公衆浴場もあるので日帰りでも温泉が楽しめます。

出典：トリップアドバイザーより一部引用

例2 草津温泉の口コミの例

日本三大名泉の1つ。行ったことはなくてもきっと名前はほとんどの人が聞いたことがある有名な温泉地。湯畑や湯もみショーは必見です。湯畑での硫黄の匂いや湯けむりで癒されました。

温泉の自然湧出量は日本一だそうで、ホテルのお風呂は源泉かけ流しでした。過剰に開発されていない所が、情緒があってとても落ち着きます。

出典：トリップアドバイザーより一部引用

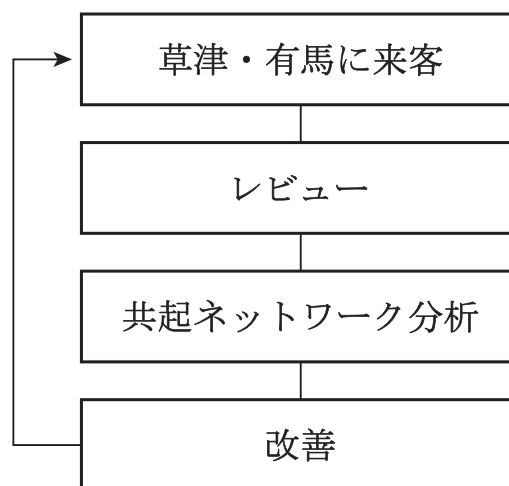
例1、例2から草津温泉と有馬温泉とでは、同じ三大温泉なのになぜこれほど人気度に差が生じるかという観点から「お客様目線、草津温泉は有馬温泉より魅力的である」という仮説を提唱する。客数の差は、立地の違いと街づくりの違いが起因していると考えられる。第二章の表から、有馬温泉の宿数31個に対して草津温泉の宿数が15個しかないことが分かっている。一見有馬温泉のほうが宿泊に適しているように見えるが、敷地面積に大きな差があり、有馬周辺と草津周辺の敷地面積を比較すると、草津周辺のほうが広い。これに宿数の情報を加えると、有馬温泉は敷地が狭いうえに宿が密集しているため、本来気をはけるべき景観を損ねているのではないかと考えられ、比べて草津温泉は、宿数が15と

有馬温泉より少ないが、広い敷地面積を生かし、宿数を抑えて景観に力を入れているのではないかと考えることができる。また草津温泉といえば硫黄の匂いやライトアップなどが印象的だが、対する有馬温泉には印象的なものがない。こういったお客様の心に残る景観や五感を刺激する何かが有馬温泉より草津温泉にはあるので、自然とお客様が草津温泉のほうが集まる、すなわち草津温泉のほうが魅力的に感じると仮説を立てることができる。以上の仮説を立証するために、次章では本研究の大まかな流れと枠組みについて述べる

2-4 本研究の大まかな流れと枠組み

本稿はアンケート分析から、有馬温泉の街おこしの実現可能性を論じる。分析のフレームワークは以下のとおりである（図3参照）

図3 分析フレームワーク



本研究の流れは、草津温泉と有馬温泉の来客者が宿泊予約サイト等に投稿したレビューをテキストマイニングによって分析し、その分析結果から明らかになった来客者の意見に基づいてその観光地がより活性化されるように改善していく、というものである。次章ではこれらの分析手法及び分析結果について具体的に述べていく。

3 章 分析手法、分析結果

第二章で草津温泉と有馬温泉を比較したところ、同じ三大名湯である二つの銭湯に来客数や敷地面積に差異があることに触れた。また3章で立てた仮説を立証するために、本章ではまずカイ二乗検定によって頻出する語句に関係性がみられるかを検証し確認した上で、共起ネットワーク分析を行い、示された結果から差異の要因を明らかにする。

3-1-1 語句の抽出

本研究ではカイ二乗検定を行うにあたって、トリップアドバイザーに書き込まれた草津温泉と有馬温泉のお客様レビューから両温泉の200レビューずつをサンプルサイズとして無作為に抽出した。そのデータにおいて頻出する名詞と形容動詞を上位10語ずつピックアップし比較した。(表3)

表3. 頻出名詞、頻出形容動詞上位10語昇順

草津温泉			有馬温泉		
単語	出現数	割合	単語	出現数	割合
湯	194	20.7%	温泉	342	24.7%
温泉	191	20.4%	湯	115	8.3%
行く	55	5.9%	有馬	101	7.3%
足湯	47	5.0%	多い	77	5.6%
来る	46	4.9%	日帰り	59	4.3%
見る	43	4.6%	銀	57	4.1%
思う	41	4.4%	金泉	55	4.0%
観光	38	4.1%	行く	55	4.0%
場所	35	3.7%	神戸	54	3.9%
土産	27	2.9%	泉	54	3.9%
人	26	2.8%	良い	51	3.7%
良い	24	2.6%	思う	50	3.6%
多い	23	2.5%	歩く	46	3.3%
店	23	2.5%	楽しめる	43	3.1%
たくさん	22	2.4%	金	40	2.9%
客	21	2.2%	炭酸	40	2.9%
言う	21	2.2%	店	38	2.7%
入る	20	2.1%	バス	36	2.6%
楽しめる	19	2.0%	人	36	2.6%
感じる	19	2.0%	旅館	35	2.5%
合計	935	100.0%	合計	1384	100.0%

3-1-2 共通項と期待値による検定

表3において、2つのデータの共通項を抽出し(表4)そのデータに対しカイ二乗検定を行った(表5)。その結果、有意水準は1%水準と極めて高い水準を示した。(p値=0.000000000000000000009356)このことは、両サンプルが独立の存在であると言える。つまり、両温泉地は少なくとも顧客アンケートから見るとその特徴・傾向は別物であり、一方

がもう一方の特徴を参考にして改善が図れる可能性があることを示している。しかし、カイ二乗検定だけでは、それぞれの温泉がどのような特徴を持つのかを具体的に明らかにすることはできない。そこで、共起ネットワーク分析を用いて詳細な関係性を次の章で明らかにする。

(表 4) 草津と有馬の共通項

共通項	草津温泉	有馬温泉	合計
湯	194	115	309
温泉	191	342	533
行く	55	55	110
思う	41	50	91
人	26	36	62
良い	24	51	75
多い	23	77	100
合計	554	726	1280

(表 5) カイ二乗検定

データと期待値との差		
	草津温泉	有馬温泉
湯	27.2	20.7
温泉	6.8	5.2
行く	1.1	0.9
思う	0.1	0.1
人	0.0	0.0
良い	2.2	1.7
多い	9.5	7.3
X二乗カイ:	82.74	
p値:	0.00000000000000000009356	

3-2 共起ネットワーク分析

詳細な関係性を明らかにするにおいて、人が手作業で分析を行うと恣意的な分析になる可能性が存在する。その場合客観性が損なわれた分析になるため、本研究では分析にフリーソフトウェアの KH Coder を用いた。KH Coder はデータを要約し提示する際に、手作業が省かれるため、分析者の持つ考えや理論からなる偏見を排除することができるため、本研究に客観性の確立を実現できた。

本研究では KH Coder の「共起ネットワーク」のコマンドを用いる。共起ネットワークとは選択したテキストに用いられた語において、出現パターンの似通った共起関係のあるネットワーク、つまりテキスト中の語と語の関係の強さを表したものである。本研究では、全体で登場数 10 以上の単語を取り上げ、Jaccard 係数を 0.15 以上の共起関係に設定し分析を行った。Jaccard 係数とは集合の類似性を表すものであり、単語間の共起関係が、全体に占める割合を示したものである。

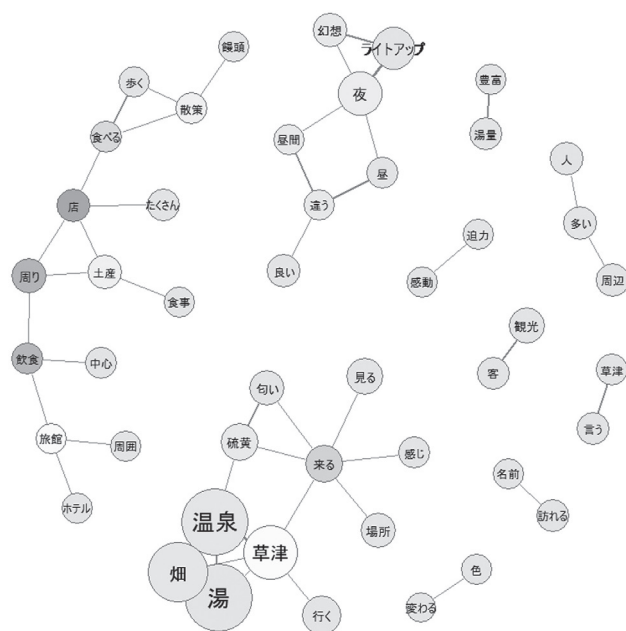
Jaccard 係数を扱うに当たって単独での出現数が多い単語ほど共起関係が現れづらい点が見られるため、Jaccard 係数をデフォルトの 0.2 から 0.15 に変更した。

以下では各温泉に宿泊したお客様がトリップアドバイザーに書き込んだユーザーレビュー(3.1 のサンプルと同じ)をもとに共起ネットワーク図を示しながら説明する。なお、図における円の大きさは単語の出現数を示していて、エッジ(円と円をつなぐ線)の太さは共起関係の強さを示している。ただし、円と円の距離は共起関係に一切関係はなく、意味をもたない。以下の文章で下線の単語は共起ネットワークに表示された語を示す。

3-2-1 草津温泉のお客様レビュー

草津温泉のお客様レビューを共起ネットワーク図にすると、図4のような結果となった。図4では大きく3つの共起関係がみられた。上部分では昼と夜の景観の差のまとまりがみられる。口コミには「夜の湯畑は昼とは違い幻想的でロマンチックでした」「夜は日没から23時頃までライトアップされているので、昼夜と違った景色が楽しめます」などの記述が多く見られた。下部分では、温泉の硫黄の匂いの関連がまとまっている。口コミには「硫黄の匂いが立ちこめ、まさに温泉地という印象」「硫黄の匂いが長旅の疲れに沁みます」「湯けむりと硫黄の匂い、湯量が豊富な事が解ります」などの記述が多く見られた。左部分には周辺の店で土産を見て回り食べ歩く、という内容の共起がみられる。「お土産店や飲食店、食べ歩きできるお店なども湯畑の周りに集まっており、賑やかな温泉街です！外湯巡りをしては湯畑周辺で食べ歩き散策をしました」「周りには土産物店やお饅頭屋さん、飲食店が立ち並んでいます」などの記述がみられた。加えて右部分に周辺に人が多いという記述があることから温泉以外にも定評があることが読み取れる。

図4 草津温泉レビューの共起ネットワーク

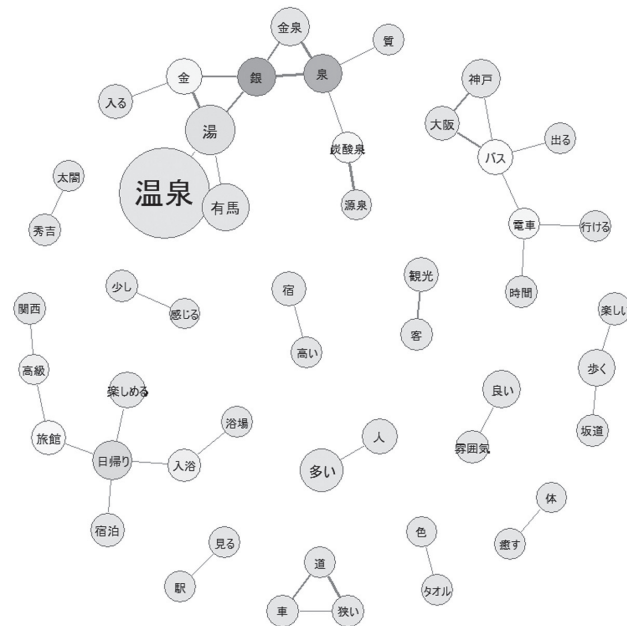


3-2-2 有馬温泉の共起ネットワーク

これに対して、有馬温泉の共起ネットワーク図は図5のようになった。こちらで特徴的なのは「日帰り」というワードである。口コミでは「日帰り入浴出来る公衆浴場もあるので日帰りでも温泉が楽しめます」などがあげられるのだが、一方「宿泊すると一人数万円かかるので、日帰り温泉でも充分楽しめる。」「ホテルに宿泊するには、もちろん駐車場には困りませんが、日帰りで遊びに行く場合、いつも駐車場に悩みます」というように宿泊

にも日帰りにもネガティブな内容も多く見られた。これに加え図5の中央部分で、宿が高いという共起から宿泊に対してネガティブな意見が口コミで多く上がっていることが読み取れる。その他で草津温泉と差異がみられる共起関係は「道」「狭い」「車」などがあげられることから、街づくりに問題がある可能性がよみとれる。

図5 有馬温泉レビューの共起ネットワーク



3-3 考察

上記の共起ネットワーク分析の結果から、以下のようなことが言えよう。

温泉以外の点において、有馬温泉では「坂道が多く道が狭い」と読み取れ、お年寄りの方が足を運び辛いと考えられる。対して草津温泉では「夜のライトアップが幻想的」「周辺の店が多くあり、食べ歩きに適す」と読みとれ、草津温泉の方が温泉以外にも力を入れていることがお客様のレビューの傾向として明らかだといえる。

有馬温泉と草津温泉は同価格帯であるが、共起ネットワークより有馬温泉では宿泊費が高いという共起関係が出ていることから、高価格の割に宿泊に対し満足感を得られていないことが読み取れる。対して草津温泉では、価格が高い、などのネガティブな共起が出ないことから、価格相応の満足度をお客様が得られていると推測できる。

4 章 まとめ

本研究は、地元有馬の来客数向上を課題とし、カイ二乗検定とテキストマイニング分析を行うことで、草津温泉と有馬温泉の観光客数の差の要因を考察し、提言を行ったものである。分析の結果から、3章で触れたように草津は広い面積に程よい量の宿泊施設を建設しており、外観が美しく保たれている。その上街並みに五感を刺激されるものがあることが共起ネットワークから読み取れ、街づくりに成功していると言える。対して有馬は狭い面積に対して多くの宿泊施設を建設しており、本来力を入れるべき外観が損なわれているため、草津に比べ街づくりに失敗していると言える。そこで有馬を改革するにあたって草津から学べることは、まず密集しすぎた街並みに手を加え、損なわれた外観を一変する必要がある。例えば、現在宿が立っている所のいくらかを特産品やお土産、食べ歩きに適した店に変えてみる。足が不自由な方や、ご年配の方々の来客数を上げるために案内観光の送迎バスを設置してみる、などが考えられる。宿数を減らせばその分外観に力を入れられる上に、送迎バスの停車位置に宿泊施設周辺を利用者も必然と増えるだろう。また、草津温泉以外にも他の人気観光地の外観や街づくりのモデルを比較することもよいだろう。

上記改善点を踏まえたうえで有馬の町並みや施設に手を加えることで、有馬がより良い観光地になれば本望である。

[参考文献]

- ・ Craft マップ
 < <http://www.craftmap.box-i.net/> > 2018 年 8 月 2 日アクセス
- ・ Google マップ
 < <https://www.google.co.jp/maps/@34.7440092,135.5350016,15z?hl=ja> > 2018 年 8 月 2 日アクセス
- ・ 名湯草津温泉—日本三大名湯の始まり
 < <http://www2l.cx/kusatsu/onsen/onsen.html> > 2018 年 8 月 2 日アクセス
- ・ 草津温泉旅館協同組合
 < <http://www.mlit.go.jp/common/001141879.pdf> > 2018 年 8 月 16 日アクセス
- ・ 宿泊予約サイト ジャらん
 < <https://www.jalan.net/> > 2018 年 8 月 16 日アクセス
- ・ 糖尿病ネットワーク
 < <http://www.dm-net.co.jp/> > 2018 年 8 月 3 日アクセス
- ・ 産経 WEST
 < <https://www.sankei.com/west/west.html> > 2018 年 8 月 3 日アクセス

- ・群馬県観光動態調査報告書（2007-2016）
＜ <http://kusatsu-shokokai.jp/wp-content/uploads/2018/02/kankou.pdf> ＞ 2018 年 9 月 13 日アクセス
- ・有馬温泉観光協会公式サイト
＜ <http://www.arima-onsen.com/> ＞ 2018 年 8 月 16 日アクセス
- ・兵庫県観光動態調査報告書（2007-2016）
＜ <https://web.pref.hyogo.lg.jp/sr16/documents/h28kankodotaichousa.pdf> ＞ 2018 年 9 月 13 日アクセス
- ・トリップアドバイザー
＜ <https://www.tripadvisor.jp/> ＞ 2018 年 9 月 22 日
- ・樋口耕一（2014）「社会調査のための軽量テキスト分析」株式会社ナカニシヤ出版
第 7 章 ,p.93-198.