

テキストマイニングによる観光地の定量的評価分析

○越智香那恵, 桂信太郎, 井形元彦 (高知工科大学)

1. はじめに

観光庁は「国・地方公共団体・民間事業者等が連携し、訪日外国人旅行者の受入環境の整備・充実を総合的に推進することで、全国的に訪日外国人旅行者が安心して快適に、移動・滞在・観光することができる環境を提供し、訪日外国人旅行者の訪問を促進するとともに、満足度を高め、リピーターの増加を図ることを目指して」(同庁 HP)、観光地域づくりを推進している。また、観光を通じた地域の活性化を図るために「観光関係者のみならず、地域住民も含めた地域の幅広い関係者が連携し、住んでよし、訪れてよしの魅力ある地域づくり」(同庁 HP) を目指している。全国、全世界には、各所に観光地が存在しているが、近年の世界的な IT 化の進展と、質量ともに情報データベースが増大しているため、観光地の口コミ情報が入手しやすくなっている。情報の質的なバラツキは否めないが、量的に一定数以上確保できる文字データを解析することで、一定レベル以上の観光地の現状把握と評価分析を行うことができるようになってきている。本稿では、樋口耕一氏の開発した分析ソフト KHCoder を採用し、先行研究をもとに、テキストマイニングによる観光地の定量的評価分析を試みる。

2. 調査概要

本研究の狙いは、観光地の印象について口コミ情報をもとにテキストマイニング手法で分析し、観光地の特徴の強化ポイントを明らかにし、今後の観光地の整備や PR 活動に反映する等で、地域活性化につなげようとするものである。本稿では、高知県の代表的な観光地である高知城(県中部)、モネの庭(県東部)、足摺岬(県西部)を取り上げ、口コミデータの量的評価分析を試みる。解析は、学術研究で援用されている KH-Coder を使用し、計量テキスト分析を行う。まず、語彙の抽出を行い、抽出語間の特徴について階層的クラスタ分析を通じて解析する。次に、抽出語の用いられ方を分析するために、KWIC コンコーダンスを実施する。さらに、ネットワーク分析、主成分分析、対応分析、潜在的意味解析などを行う。

(1) 頻出語彙解析

①高知城訪問者口コミ全体 114 サンプル

KH-Coder を使用して、頻出語彙の抽出を行った。図 1 に関して、名詞では、上位には、眺めや景観に関連する場所を示す語彙が多い。例えば「天守閣」「景色」「眺め」などである。サ変名詞でも「一望」などの単語が上位にある。また形容動詞も「綺麗」「きれい」が上位に抽出されている。また、図 2 に関連して、形容詞では、上位に「良い」「多い」「素晴らしい」「美しい」「楽しい」「気持ちいい」などポジティブな表現の語彙抽出が目立つ。また、高知県が注力しているはずの「おもてなし」に関しては、名詞 B の 1 つでしか抽出できず、県の観光政策との認識のズレがある可能性が指摘できる。また、未知語というカテゴリで 8 つ出てきており、訪問者がポジティブな気持であった可能性が高いと思われる。

名詞	サ変名詞	形容動詞	固有名詞	動詞	形容詞	副詞	名詞B	名詞C
天守閣	48 観光	21 急	17 高知駅前	行く	53 良い	34 少し	8 おすすめ	5
階段	46 一望	15 綺麗	15 はりまや橋	登る	29 多い	13 初めて	4 おかげ	1
景色	19 現存	14 大変	9 ビリ	見る	26 小さい	6 少々	3 おもてなし	1
市内	19 ガイ	12 残念	8 高知城前	思う	24 少ない	6 丁度	3 ごはん	1
最高	18 案内	10 立派	6 大阪城	歩く	12 近い	5 特に	3 つぶし	1
眺め	16 散歩	10 きれい	5 姫路城	残る	11 素晴らしい	5 結構	2 どなり	1
天守	16 散策	9 好き	5	感じる	9 美しい	5 比較的	2 なかなか	1
歴史	13 展示	8 素敵	4	出来る	8 暑い	4 本当に	2 もと	1
市場	11 旅行	5 にぎやか	2	入る	8 大きい	4 もう一度	1	
石垣	10 お願	4 貴重	2	入れる	8 楽しい	3 もう少し	1	
				上る	7 気持ちいい	3 ヒヤヒヤ	1	
				訪れる	7 興味深い	3 案の定	1	

図 1 頻出語彙 (筆者作成)

②高知城訪問者男女別(61:53)の分析

次に、このサンプルを男女別に分析する。図 2 に示す抽出語彙では、男性と女性の名詞、サ変名詞、形容動詞、固有名詞の頻出語彙について示している。男女とも名詞の最上位には「天守閣」がある。また男性の場合「石垣」が「階段」の上位にあり、女性はその逆に位置

している。これは、男性は「石垣」そのものや眺望を目的として来訪しているのに対して、女性はそれよりも「階段」をのぼることに対するネガティブな思い（「しんどい」もしくは「きつい」など）が先行している可能性が出ているのではないだろうか。また女性には「アイス」の単語が上位に出ており、これも女性の来訪目的や印象を示す語彙として特徴的である。

名詞	サ変名詞	形容動詞	固有名詞	名詞	サ変名詞	形容動詞	固有名詞		
天守閣	41 一望	9 急	8 はりまや橋	2	天守閣	43 一望	14 急	9 リ	1
天守	15 観光	8 綺麗	5 姫路城	1	階段	30 観光	9 綺麗	7	
石垣	14 現存	8 重要	3		眺め	9 展示	6 立派	6	
階段	12 散歩	6 大変	3		市内	8 案内	4 大変	4	
市内	12 散策	5 必要	3		石垣	8 昼寝	4 便利	3	
眺め	12 駐車	5 立派	3		アイス	7 ガイド	3 きれい	2	
歴史	12 整備	4 きれい	2		リン	7 禁止	3 好き	2	
最高	9 注意	4 格別	2		無料	7 散歩	3 小さめ	2	
景色	6 旅行	4 気軽	2		歴史	7 出店	3 静か	2	
見晴らし	6 運動	3 好き	2		景色	6 イメージ	2 爽やか	2	
					市場	6 意味	2 必要	2	
					ボランティア	5 幻想	2 さわやか	1	
					ライトアップ	4 工事	2 キレイ	1	

図2 高知城／男性の名詞等の頻出語彙（左）女性の名詞等の頻出語彙（右）

次に、男女別の名詞、サ変名詞、形容動詞、固有名詞を分析した。紙面の都合でプレゼン当日に示す。ここで特徴的なのは、動詞の頻出語彙である。男性の場合、「思う」「見る」「感じる」「楽しめる」などの感覚的・主観的な動詞が上位に抽出されている。また女性は「上る」「入る」「歩く」「たどり着く」「上がる」などの、肉体的・行動に関する動詞が、上位に抽出されている。これは、男性に関しては、高知城を楽しもうと来訪しているケースが多いが、女性は高知城を楽しもうというよりも、登ったりする行為そのものがきついなど、見たり楽しむ行為よりも上位に来ていることが伺える。

③高知城来訪者年代別の分析

名詞	サ変名詞	形容動詞	名詞	サ変名詞	形容動詞	固有名詞		
天守閣	26 一望	14 きれい	11	天守閣	30 現存	16 立派	8 姫路城	2
階段	20 観光	9 急	11	天守	27 観光	9 重要	6 はりまや橋	1
市内	10 散歩	5 綺麗	6	歴史	13 案内	6 好き	4 キツ	1
ライトアップ	8 展示	5 素敵	5	眺め	9 展示	6 静か	3 桂浜	1
最高	7 サービス	4 立派	5	無料	6 案内	2 結構	2 高知城前	1
眺め	7 アクセス	2 いろいろ	2	シンボル	5 運動	2 すてき	1	
場所	5 花見	2 気軽	1	場所	5 花見	2 気軽	1	
迫力	5 見学	2 古風	1	迫力	5 見学	2 古風	1	
歴史	5 入場	2 困難	1	歴史	5 入場	2 困難	1	
アイス	4 疲労	2 自然	1	アイス	4 疲労	2 自然	1	

図3 高知城／29歳以下の名詞、サ変名詞、形容動詞の頻出語彙（筆者作成）

次に、このサンプルを年代別に分析した。抽出語彙では、29歳以下と60代以上の名詞、サ変名詞、形容動詞、固有名詞の頻出語彙について示している。60歳以上の抽出語彙では、歴史、国宝という言葉が多いのが一つの特徴である。また29歳以下は「無料」などが抽出されており、次に観光に来る人へ向けてロコミを発信している可能性がある。抽出された動詞が特徴的で、29歳以下は「行く」「登る」「見る」が上位にあり、60歳代以上は「思う」「訪れる」「楽しめる」などが比較的上位にある。これは若年層が肉体的に行動して高知城を体感しているのに対して、60代以上は、楽しんだり、歴史に思いをはせたりすることに主眼がある可能性が高い。29歳以下は高知城を体感し、60歳以上は高知城の歴史を見に来ている可能性がある。

3. 小括

本稿では、語彙抽出を中心に記したが、抽出語間の特徴についての階層的クラスター分析、抽出語の用いられ方を分析するためのKWICコンコーダンス、ネットワーク分析、主成分分析、対応分析、潜在的意味解析などを行っている。また、他の観光地でも同様の分析を行っており、これらの比較も行いながら、県庁などの支援機関へのヒアリングとすり合わせを行う予定である。さらに、全国各地における主要観光地についての分析を進めていくことを考えている。

【参考文献】

- [1]高知県庁観光振興部観光政策課「平成27年県外観光客入込動態調査報告書第1章県外観光客入込調査」
- [2]樋口耕一『社会調査のための計量テキスト分析』ナカニシヤ出版(2014)