

# 口コミのテキストマイニングから見た日本人と 外国人の観光行動の差異

11910176 松井陽史

# I はじめに

新型コロナの水際対策の緩和によって、2022年12月に訪日外国人人数が137万人に達し、感染拡大前の2020年2月以来、2年10か月振りに100万人を超えた。

観光行動に着目しても、日本人と外国人それぞれに特有の指向があると考えられる。両者の観光行動におけるニーズについて傾向を把握することは、観光地がより適切な政策を作るうえで有意義である。

# I はじめにー先行研究ー

## Web上のビッグデータを研究に活用する意義

馬 (2017)

従来の社会統計学的な手法で実施されてきた観光地調査は、多大な時間とコストがかかり、大規模な調査が難しかった。一方でUGC (User Generated Contents) は、ユーザの意見を大規模かつタイムリーに収集分析することが可能である。

保田 (2003)

SNSや口コミといったWeb上の情報を活用する手法として、テキストマイニングがある。膨大なテキストデータを様々な観点から分析可能で、社会調査の自由回答データ、新聞記事、口コミなど幅広いテキストに応用されている。

# I はじめにー先行研究ー

## テキストマイニングを観光行動分析に適用した研究

安原・劉（2021）

日本人と外国人の日本庭園への訪問体験を取り上げ、共通点、相違点、評価傾向を分析。日本人は外国人観光客の多さが記憶に残り、外国人は庭園固有の歴史性に関心を有した可能性などを指摘した。



口コミの分析が観光体験の差異を可視化するのに有効であることを示した。

# I はじめにー先行研究ー

## 日本人と外国人の観光行動の差異をインタビュー調査により分析した研究

市川ほか（2016）

白川郷に来訪した日本人と外国人双方のツーリストを対象にして、移動や滞在の形態などの特性、ツーリストが抱く世界遺産白川郷のイメージを、それぞれインタビュー調査を通じて分析し、以下のような差異を明らかにした。

日本人：消費よりも景観や散策に充てる時間が多く、自家用車による短時間の滞在 など

外国人：宿泊傾向が強く、近隣の観光地との周遊観光の一環としての来訪が多い など

# I はじめにー研究目的ー

①

市川ら（2016）にみられるように、日本人と外国人のロコミデータを、共通する属性毎に分類し、両者の観光行動の傾向を把握。

②

ロコミのテキストマイニングで観光行動を分析した場合、インタビュー調査で行った場合と結果が異なるのか検証する。



日本人と外国人の観光行動の差異を、ロコミデータのテキストマイニング分析を通じて明らかにしたうえで、従来のインタビュー調査から得られた結果と比較し、観光行動分析へのテキストマイニングの有用性を示すことを目的とする。

## Ⅱ 研究対象・方法

### 対象地域

#### **白川郷合掌造り集落**

所在地：岐阜県大野郡白川村

1976年 重要伝統的建造物群保存地区に選定

1995年 白川郷・五箇山の合掌造り集落として世界文化遺産に登録

2019年 白川村への観光入込客数が、過去最高の215万人（外国人は101万人）となる

## Ⅱ 研究対象・方法

### 分析対象

「tripadvisor」上の、「白川郷合掌造り集落」及び「The Historic Villages of Shirakawa-go Gassho Style Houses」に投稿された口コミデータ

日本語版：931件、英語版：669件

### 対象とした訪問時期

日本語版：2009年6月～2022年6月

英語版：2011年7月～2021年2月



## Ⅱ 研究対象・方法

### 分析に使用したソフト

「KH coder3.Beta.o6d」

文章に現れる単語の使用頻度や傾向を把握し、その後各単語の階層化クラスター分析、共起ネットワーク分析を実施

※KH coderとは

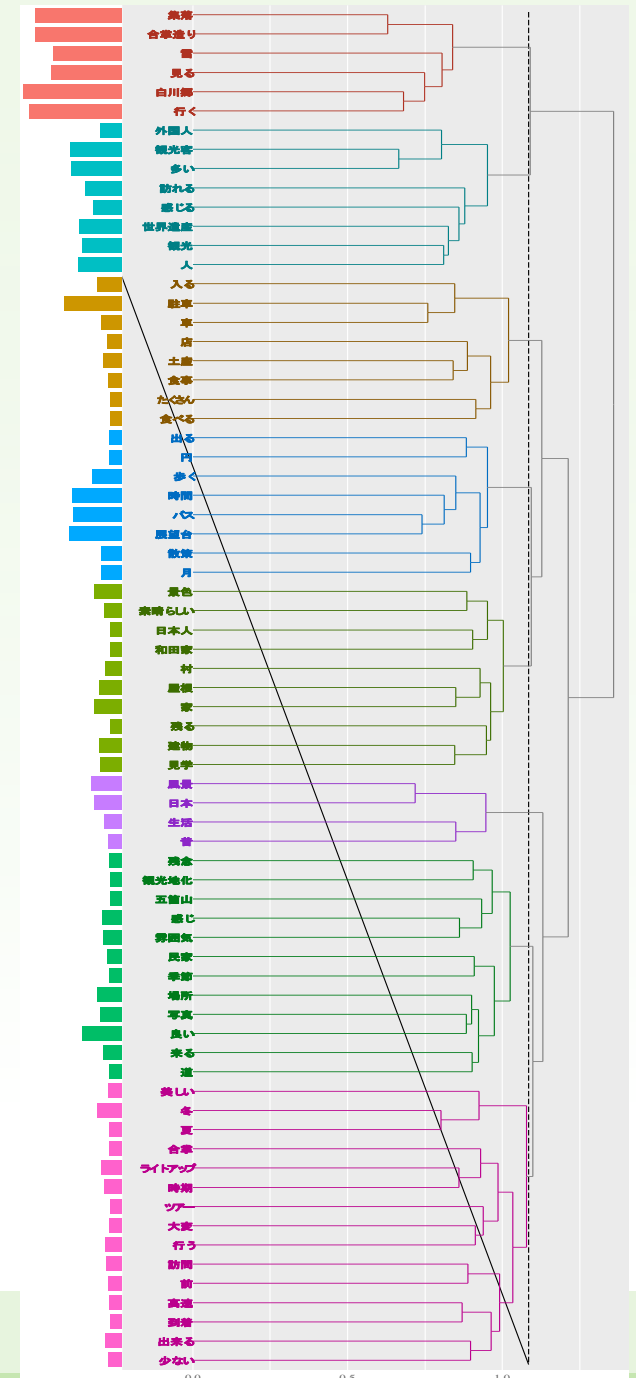
・・・テキスト型データのテキストマイニングのためのフリーソフトウェア。出現頻度の高い単語を調べたり、出現傾向を調べたりできるなど、多変量解析により有用な情報を取り出すことが可能。

# Ⅲ 結果

## 階層的クラスタ分析 日本語版

- 「観光客」「多い」「外国人」といった語がみられる、「観光客の多さ」のクラスターの出現頻度が高い
- 「駐車」「土産」「食事」などが含まれる、「消費活動」が3番目に高い
- 「歩く」「展望台」「散策」などが含まれる、「探索・散策」が4番目に多い

※最も多いクラスターである「対象地の概観」は、目的地の明示などの目的で非常に多くの口コミに画一的に使用されており、特徴が見られなかったため分析対象から外す

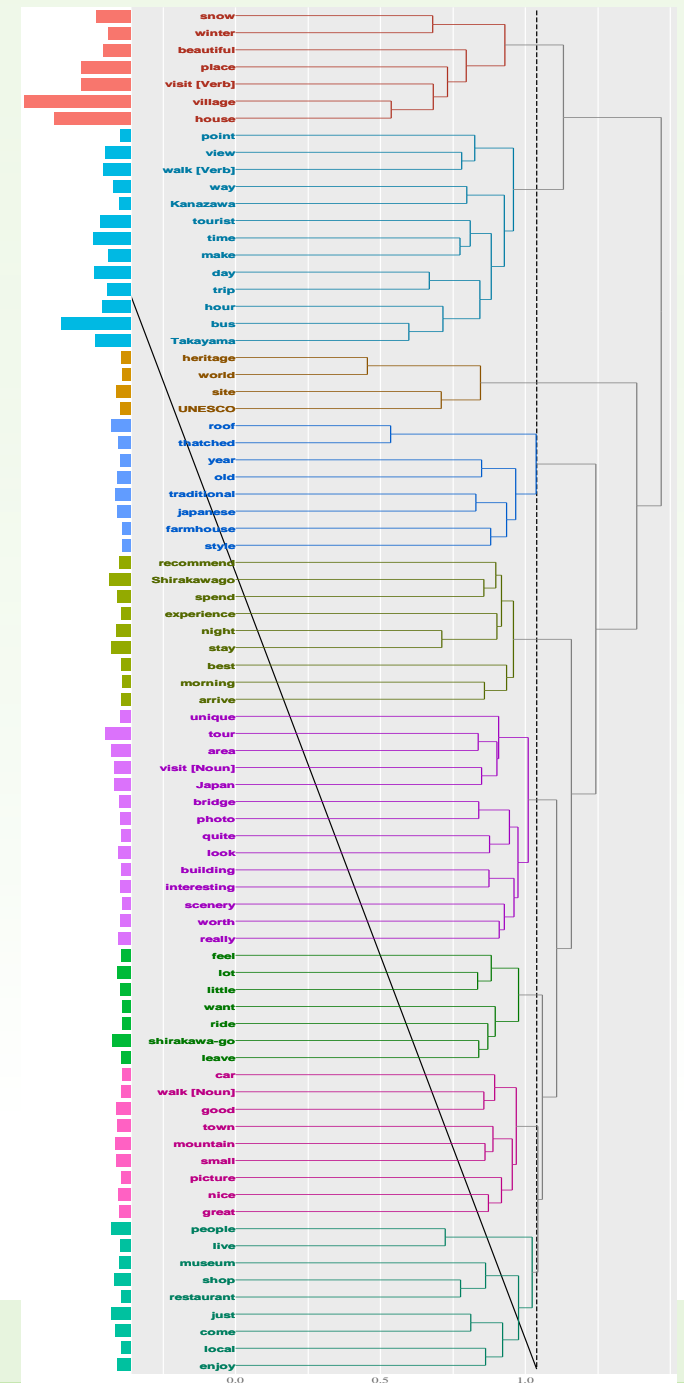


右図：図1 日本語の口コミの階層的クラスタ分析結果

### Ⅲ 結果

## 階層的クラスタ分析 英語版

- 「village」「house」「snow」など、「集落・合掌造りの景観」のクラスターの出現頻度が最も高い
- 「bus」「Takayama」「time」などが含まれる、「アクセス」が次いで高い
- 「heritage」「site」「UNESCO」などにみられる、「世界遺産」が3番目に高い



右図：図2 英語のロコミの階層的クラスタ分析結果

### Ⅲ 結果

## 階層的クラスタ分析

- ・日本人と外国人のクラスタを比較すると、上位のクラスタを中心に、頻出語は異なる
- ・日本人にのみ出現したクラスタに、「消費活動」「散策・探索」「観光地化」などがある
- ・外国人にのみ出現したクラスタに、「アクセス」「宿泊体験」などがある

表1 日本人及び外国人のロコミに頻出した単語のクラスタ  
図1及び図2より作成

頻出したクラスタ（多い順）	日本人	外国人
1	対象地の概観	集落・合掌造りの景観
2	観光客の多さ	アクセス
3	消費活動	世界遺産
4	散策・探索	建築
5	景観・建築	宿泊体験
6	原風景	ツアー・地域
7	観光地化	移動
8	季節差	写真
9		消費活動

# Ⅲ 結果

## 階層的クラスタ分析：読み取れること

### 日本語版

- 「世界遺産」「観光客」「外国人」などが抽出された、「観光客の多さ」のクラスタでは、多くの外国人観光客が訪れる世界遺産として認識されている。
- 「食事」「土産」「店」「車」などが抽出された、「消費活動」のクラスタは、食事や土産の購入などが楽しめて、車で訪れる人が多い観光地であると認識されている。

### 英語版

- 「village」や「house」が、「beautiful」や「visit」と関連して使用されている「集落・合掌造りの景観」のクラスタは、美しい景観を見るために訪れるべき場所であると認知されている。
- 「Takayama」などの地名のほか、「bus」「trip」などが使用されている「アクセス」のクラスタからは、外国人が高山など周辺の観光地との周遊の一環として白川郷を位置づけていることが分かる。

# Ⅲ 結果

## 共起ネットワーク分析 日本語版

- 大都市圏、関東からの来訪者に特徴的な語として、「駐車」「時間」が多い
- 関西からの来訪者に「写真」「景色」「風景」などの語が多い
- 九州からの来訪者に「ツアー」、北海道からの来訪者に「出る」が使用されている

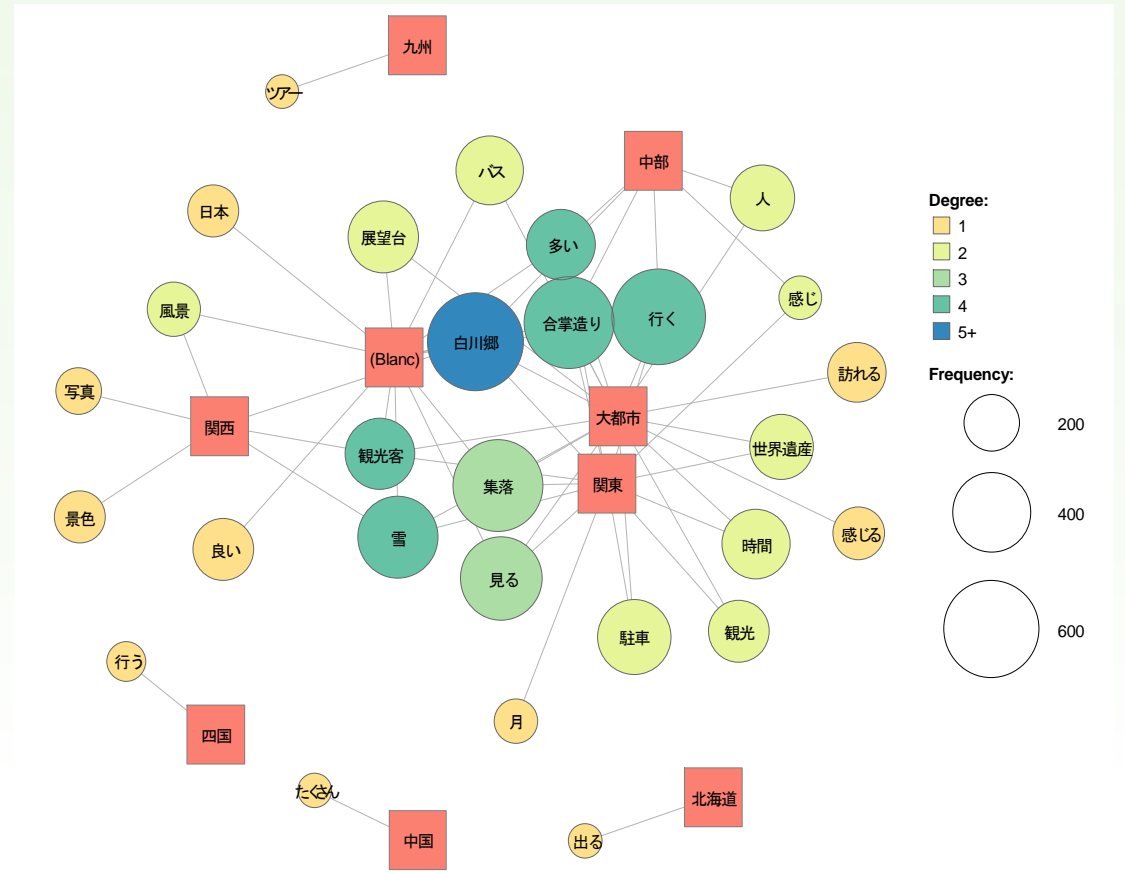


図3 日本語の共起ネットワーク分析結果

# Ⅲ 結果

## 共起ネットワーク分析 英語版

- アジアからの来訪者が、「snow」「winter」「day」などの語を多く使用している
- ヨーロッパからの来訪者は、「old」「interesting」「Japanese」などの語を特に使用している
- 北米からの来訪者には、「town」「Japan」「roof」、オセアニアからの来訪者では、「stay」「way」が特徴的である

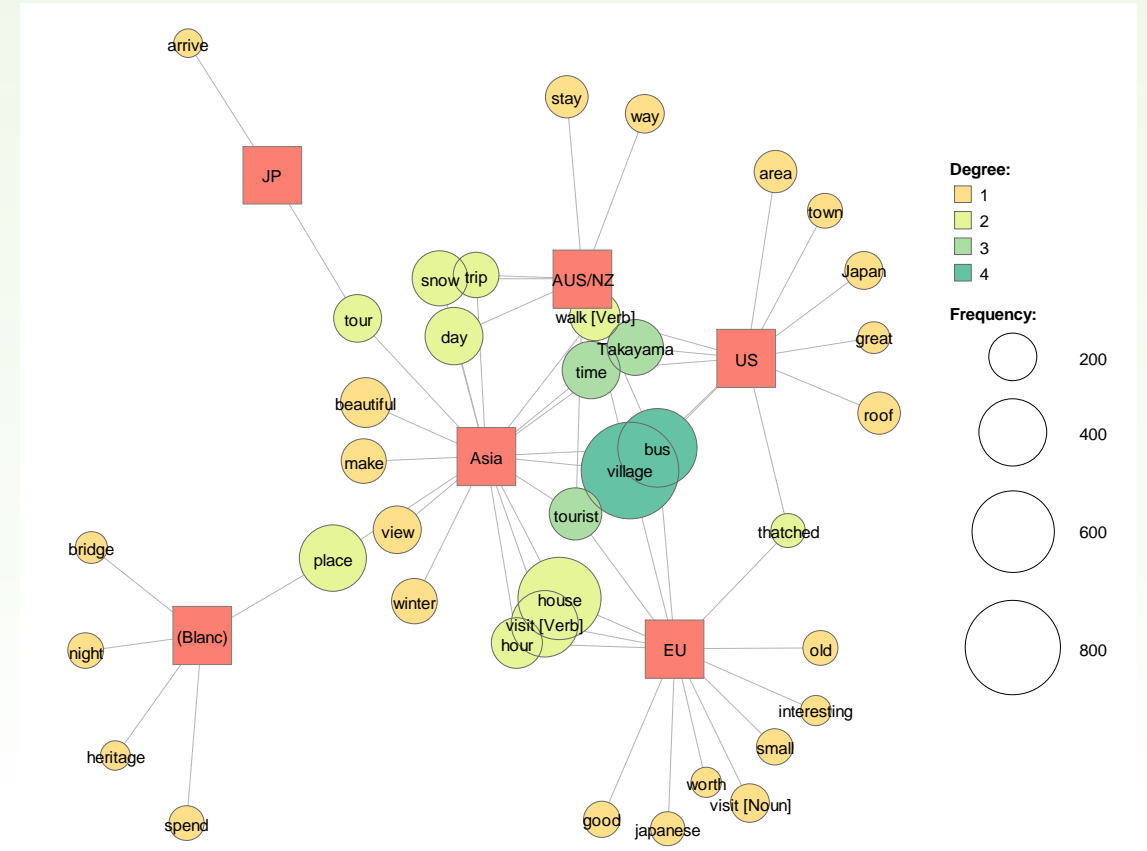


図4 英語の口コミの共起ネットワーク分析結果

# Ⅲ 結果

## 共起ネットワーク分析：読み取れること

### 日本語版

- ・関東、大都市圏と「駐車」の結びつきが強いことから、自家用車での来訪者が多い。
- ・関西と「景色」「風景」「写真」が結びついており、関西からの来訪者が景色や風景の写真を撮るなど、景観を特に意識している。

### 英語版

- ・最も頻出していた「集落・合掌造りの景観」はヨーロッパ・北米の来訪者が中心となって使用している。彼らの間で、白川郷が景観を楽しむ観光地として認識されている。
- ・2番目に多かった「アクセス」は、各地域の来訪者が多く使用しており、外国人がツアーなどを利用して周遊して観光をしていることが分かる。
- ・3番目に多かった「建築」は、北米からの来訪者が多く使用している。



# Ⅲ 考察

## 先行研究との比較：日本人

市川ら（2016）によると...

- ・ メディアが表象する農山村らしい白川郷のイメージと、人の多さとのギャップに違和感や落胆を覚えつつも、合掌造りや自然を観光地として評価する傾向にある
- ・ 短時間の滞在で定番スポットを周遊し、買い物や食事には費用をかけない傾向



出現頻度が上位のクラスターに、「観光客の多さ」、「探索・散策」、「景観・建築」などがある。日本人の口コミのテキストマイニングでは、インタビュー調査と類似した結果が得られたと言える。

# Ⅲ 考察

## 先行研究との比較：外国人

市川ら（2016）によると...

- ・ 宿泊傾向が強く、日本固有の建築や生活様式を評価している
- ・ 白川郷を周遊観光の立寄り地として位置付けている
- ・ アジア人観光客は北陸の観光地などとの周遊を目的とした、団体のツアーによる来訪が特に多い。欧米人観光客は日本食や農村景観の体験を動機とした、個人での来訪が多い。



「集落・合掌造りの景観」のクラスターが最も多く出現し、次いで白川郷への「アクセス」が多かったほか、「宿泊体験」は外国人のクラスターにのみ出現した。ツアーについては、地域を問わず、広く言及されていたが、景観や建築については、ヨーロッパや北米からの来訪者が多く言及していた。

外国人の口コミのテキストマイニングでもインタビュー結果と類似した結果が出た。

## IV 結論

白川郷を事例として日本人と外国人の観光行動の差異を明らかにするために、膨大なテキストデータを扱うことができるテキストマイニングの手法を用いたところ、市川ら（2016）と類似した結果が得られた。

観光行動研究において、従来のインタビュー調査によって得られた知見を、テキストマイニングによる分析で補強することが可能であるということの一例を示した。

## 参考文献

馬 強 2017. 観光情報学の最前線ー観光の分散化と個人化を促進する集合知活用情報技術ー. 情報処理 58-3:220-226.

保田明夫 2003. テキストマイニングの技術と応用性. 薬学図書館 48-4: 247-252.

安原有紗・劉銘 2021. 口コミの言語解析からみた日本人と外国人の浜離宮庭園に対する観光体験の相違. 環境情報科学 学術研究論文集 35: 227-232.

市川康夫・羽田司・松井圭介 2016. 日本人・外国人ツーリストの観光特性とイメージにみる白川郷の世界遺産観光. 人文地理学研究 36,11-28.

# 参考文献

「昨年12月の訪日客、100万人超え コロナ前の5割超えまで回復」『朝日新聞』2023年1月18日付.

<https://www.asahi.com/articles/ASR1L5DC2R1LULFA00T.html>

白川村の観光統計

<https://www.vill.shirakawa.lg.jp/1260.htm>

トリップアドバイザーメディアセンター トリップアドバイザーについて  
<https://tripadvisor.mediaroom.com/JP-about-us>

KH coderとは？

<https://web.pref.hyogo.lg.jp/tjk04/documents/3sankoushiryou1.pdf>

# 参考文献

Tripadvisor「白川郷合掌造り集落」(口コミ931件)

[https://www.tripadvisor.jp/Attraction\\_Review-g1119912-d1407426-Reviews-or40-The\\_Historic\\_Villages\\_of\\_Shirakawa\\_go\\_Gassho\\_Style\\_Houses-Shirakawa\\_mura\\_On.html](https://www.tripadvisor.jp/Attraction_Review-g1119912-d1407426-Reviews-or40-The_Historic_Villages_of_Shirakawa_go_Gassho_Style_Houses-Shirakawa_mura_On.html)

Tripadvisor「The Historic Villages of Shirakawa-go Gassho Style Houses」(口コミ670件)

[https://www.tripadvisor.com/Attraction\\_Review-g1119912-d1407426-Reviews-The\\_Historic\\_Villages\\_of\\_Shirakawa\\_go\\_Gassho\\_Style\\_Houses-Shirakawa\\_mura\\_Ono\\_gun.html](https://www.tripadvisor.com/Attraction_Review-g1119912-d1407426-Reviews-The_Historic_Villages_of_Shirakawa_go_Gassho_Style_Houses-Shirakawa_mura_Ono_gun.html)