

# 中国人の旅行行動・評価特性の分析 －コンテンツツーリズムと旅行口コミ行動を対象として－

東洋大学国際観光学部 古屋秀樹

東洋大学大学院国際観光学専攻 梁嘉偉,宋紫龍

- 発表の構成
  - 日本における観光の現状
  - 訪日中国人旅行者の旅行記の分析
    - ビッグデータによるニーズ・周遊の把握,特異データの抽出
  - 態度がアニメ旅行行動に及ぼす影響分析
    - 財・サービスの特性に加えて,旅行者の態度による差異の把握

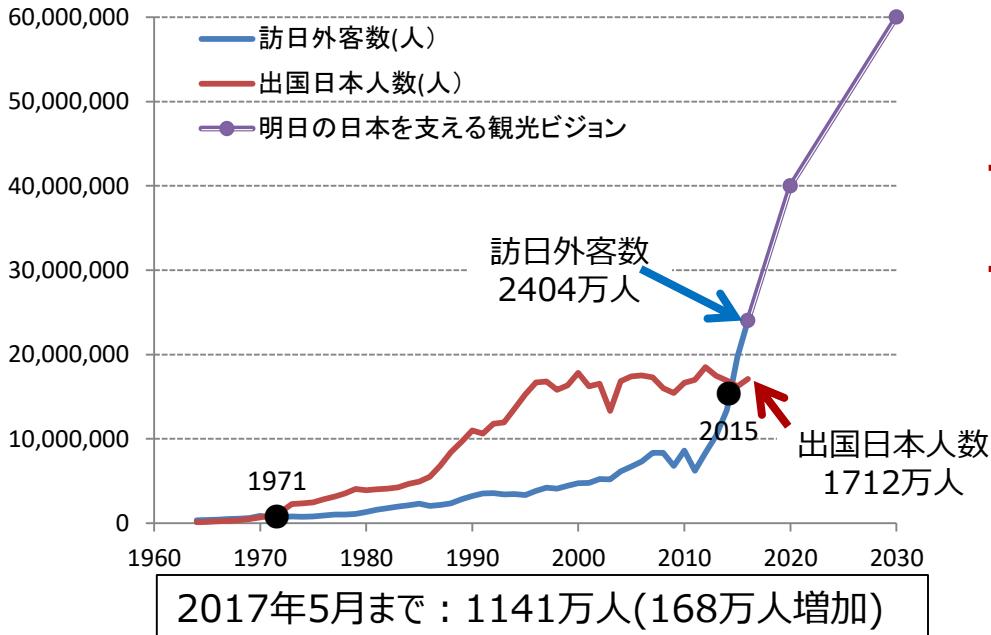
# 对于中国人的旅行行动・评价特性的分析 —以文化体验旅游与旅游口碑评价为对象—

东洋大学国际观光学部 古屋秀树

东洋大学研究生院国际观光学专业 梁嘉伟，宋紫龙

- 发表的构成
  - 日本观光的现状
  - 访日中国人旅行者游记的分析
    - 通过大数据对需求・游览实态的把握，特殊数据的抽出
  - 态度对动漫文化旅游的旅游行动的影响
    - 包含项目・服务的特性，通过旅行者的态度对旅游行动的差异的把握

# 訪日外国人旅行者数のトレンド



出典：訪日外国人旅行者統計,JNTO

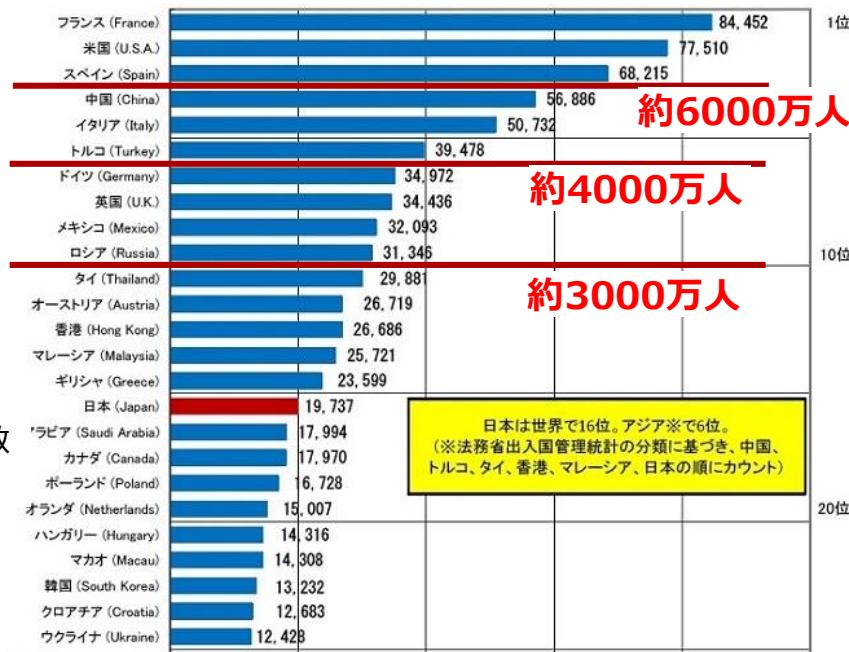
## 2016年実績

**訪日外客数：2404万人(22%増加。1千万人増/2年)**

**目標 2020年:4000万人,2030年: 6000万人**

明日の日本を支える観光ビジョン構想会議

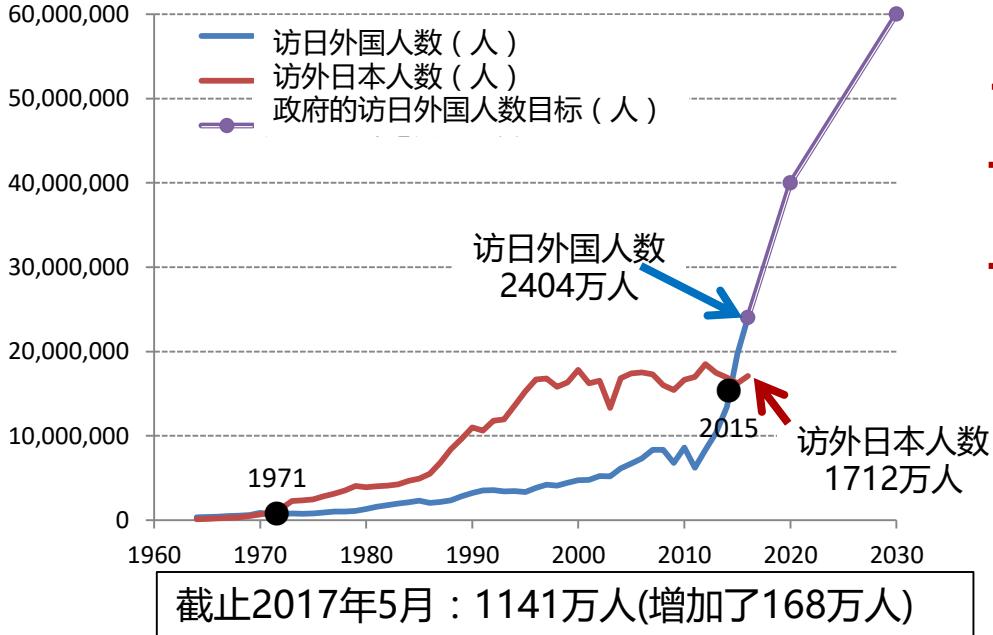
## 世界各国・地域への外国人訪問者数



出典：JNTO HP

# 访日外国人旅行者数的趋势

## 到世界各国·地区访问的外国人数



出处：访日外国人旅行者统计,日本政府观光局 ( JNTO )

### 2016年の実際成果

访日外国人数 : **2404万人** (22%增加 . 1千万人増/2年)

目标 2020年:4000万人, 2030年: 6000万人



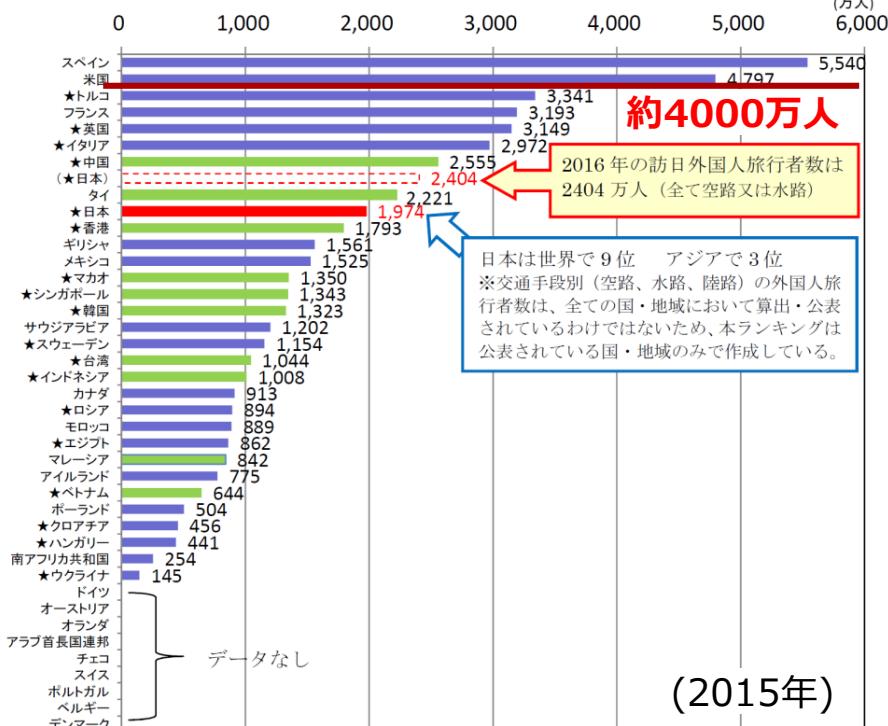
出处：日本政府观光局 ( JNTO ) 官方网站

支持未来日本的观光战略 构想会议

# 訪日外国人旅行者数,国籍等の構成割合

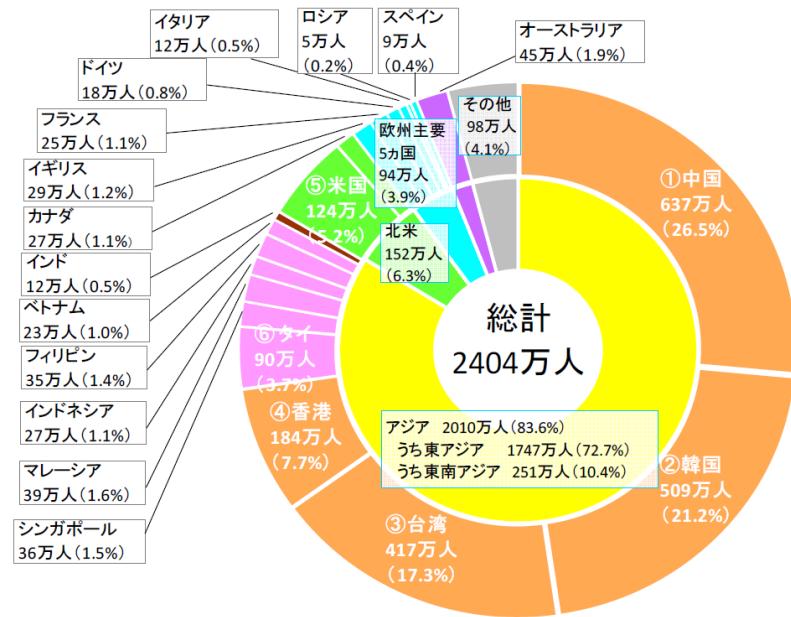
- 中国:前年比27.6%増で全市場で初の600万人台.
- 韓国(27.2%増),台湾(13.3%増),香港(20.7%増)を加えた  
**東アジア4市場:73%-前年比23.1%増.**

## 空路又は水路による外国人旅行者受入数



出典：観光白書

## 訪日外国人旅行者の内訳(2016年)

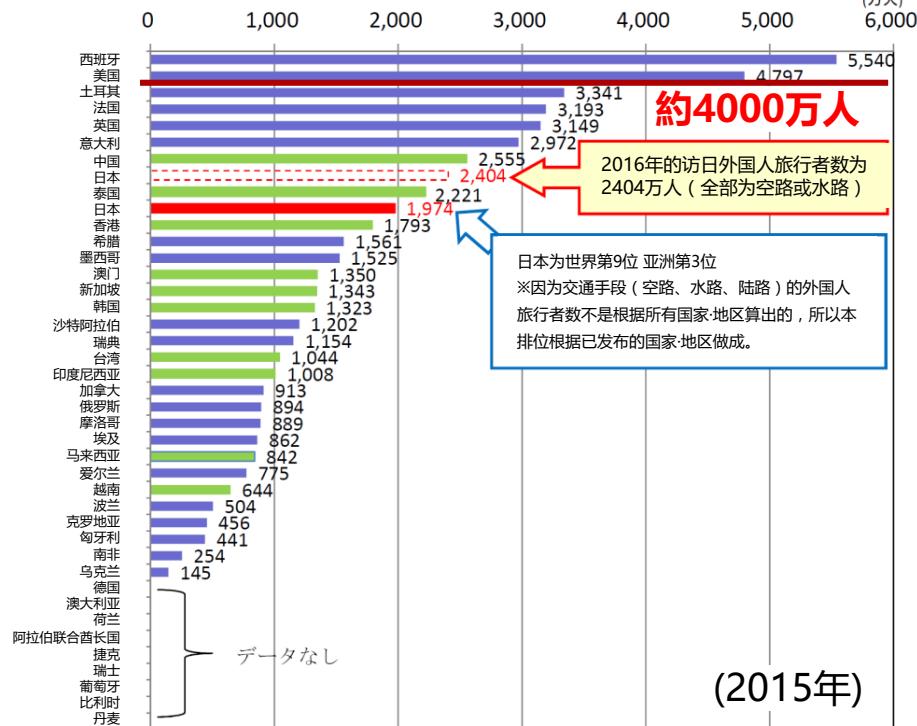


出典：観光白書

# 访日外国人旅行者数，国籍等的构成占比

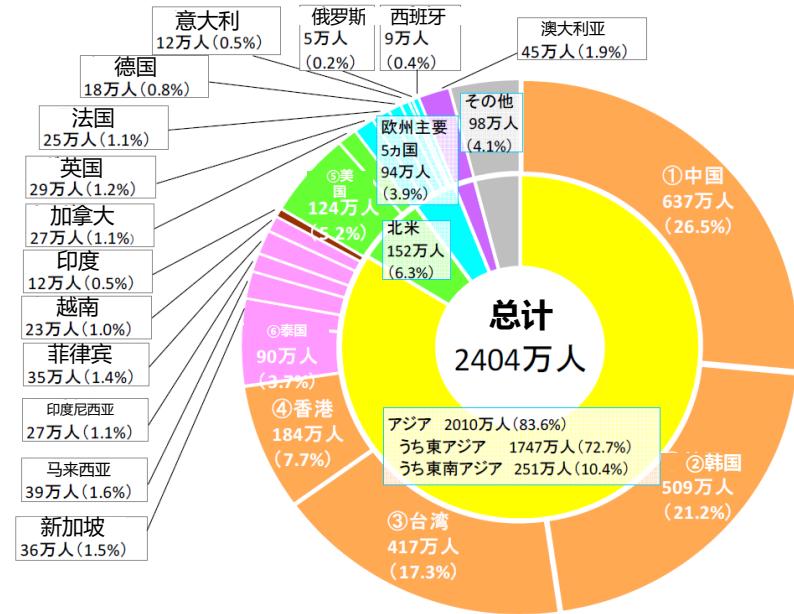
- 中国：与去年相比增加了27.6%。整个访日旅游市场中首次达到600万人次。
- 加上韩国(27.2%增),台湾(13.3%增),香港(20.7%增)  
**东亚4市场:73%-与去年相比增加23.1%。**

## 通过空路或水路入国的外国人旅行者数



出处：观光白书

## 访日外国人旅行者的构成 ( 2016年 )



出处：观光白书

# 旅行消費・経済効果

旅行消費額25.5兆円

	規模	平均費用(円/人)	消費総額(兆円)
日本人宿泊観光旅行	1.39回/人・年 (2.28泊/人・年)	49,234円	16兆円 (64%)
日帰り観光旅行	1.53回/人・年	15,602円	4.6兆円(18%)
海外旅行(国内分)			1.3兆円(5.3%)
訪日外国人旅行	1973.7万人	155,896円	3.3兆円(13%)

建設業・  
鉄道

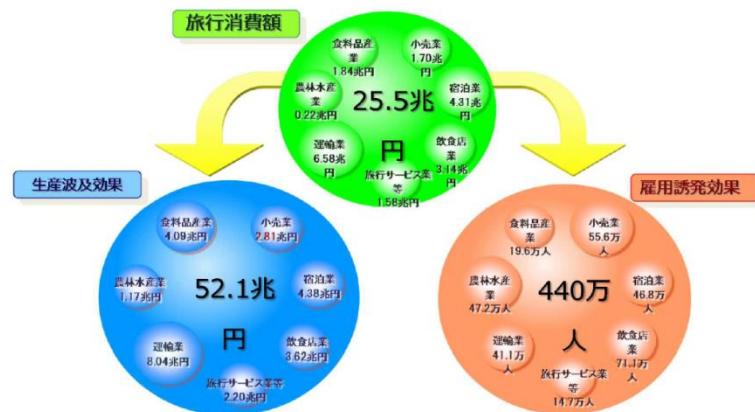
飲食業・  
製糸業

- 訪日外国人旅行・費目別消費額

- 買物代(5.9万円),宿泊料金(4.2万円),飲食費(3.2万円)

## 訪日外国人1人当たり旅行支出(2016年) 旅行消費による経済効果(2015年)

国籍・地域	総額	訪日外国人1人当たり旅行支出						平均泊数	
		宿泊料金	飲食費	交通費	娯楽・サービス費	買物代	その他		
全国籍・地域	155,896	-11.5%	42,182	31,508	17,838	4,725	59,323	320	10.1
韓国	70,281	-6.5%	22,090	17,847	7,505	3,042	19,562	234	4.5
台湾	125,854	-11.1%	33,634	26,611	14,126	4,137	47,122	224	7.4
香港	160,230	-7.0%	41,501	34,570	17,328	4,412	62,389	29	6.3
中国	231,504	-18.4%	44,126	38,943	19,917	5,014	122,895	609	11.8
タイ	127,583	-15.3%	34,532	25,858	16,529	4,821	45,414	430	9.9
シンガポール	163,210	-12.9%	58,462	34,903	24,891	3,242	40,428	1,285	8.0
マレーシア	132,353	-12.0%	41,047	27,753	19,799	3,671	40,023	60	11.7
インドネシア	136,619	-7.2%	47,117	22,586	27,171	5,051	34,665	28	15.0
フィリピン	112,228	-11.3%	33,192	26,505	13,419	4,959	34,093	59	28.3
ベトナム	186,138	-4.5%	49,314	54,182	19,556	4,125	58,883	77	35.4
インド	144,275	-2.7%	61,354	27,379	19,713	2,784	32,971	74	22.8
英国	181,795	-13.7%	80,131	41,220	30,011	6,408	24,000	24	12.6
ドイツ	171,009	-0.0%	78,849	36,899	29,715	5,068	20,350	128	14.0
フランス	189,006	-9.7%	75,462	40,799	34,590	7,830	30,299	26	16.0
イタリア	198,000	-2.0%	78,597	45,299	38,900	9,315	25,596	293	12.8
スペイン	224,072	-1.4%	92,211	49,103	44,168	6,482	32,052	57	14.0
ロシア	190,874	+4.6%	61,845	31,004	21,812	7,258	64,889	4,067	21.1
米国	171,418	-2.4%	70,707	41,137	27,856	5,427	26,111	180	14.1
カナダ	154,977	-9.2%	55,366	36,319	28,509	5,184	29,326	273	12.8
オーストラリア	246,866	+6.7%	99,802	51,202	40,169	17,957	37,587	150	13.2
その他	183,083	+2.8%	67,698	41,889	29,633	6,505	37,333	24	18.9



出典：観光白書

出典：観光白書

# 旅行消费・经济效益

旅行消费额25.5兆日元

	规模	平均费用(日元/人)	消费总额(兆日元)
日本人 多日观光旅行	1.39回/人·年 (2.28泊/人·年)	49,234日元	16兆日元 (64%)
一日游观光旅行	1.53回/人·年	15,602日元	4.6兆日元(18%)
海外旅行(国内部分)			1.3兆日元(5.3%)
访日外国人旅行	1973.7万人	155,896日元	3.3兆日元(13%)

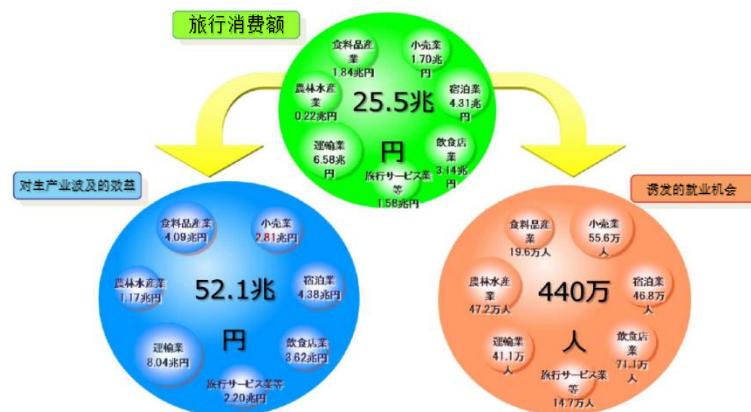
建设业·  
铁道

饮食业·  
缫丝业

- 访日外国人旅行・按消费目的分的消费额
  - 购物费用(5.9万日元), 住宿费用(4.2万日元), 饮食费用(3.2万日元)

## 访日外国人平均1人的旅行支出 (2016年) 旅行消费带来的经济效益 (2015年)

国籍・地域	総額 (前年比)	訪日外国人1人当たり旅行支出 (円/人) (泊)					
		宿泊料金	飲食費	交通費	娯楽 サービス費	買物代	その他
全国籍・地域	155,896	-11.5%	42,182	31,508	17,838	4,725	59,323
韩国	70,281	-6.5%	22,090	17,847	7,505	3,042	19,562
台湾	125,854	-11.1%	33,634	26,611	14,126	4,137	47,122
香港	160,230	-7.0%	41,501	34,570	17,328	4,412	62,389
中国	231,504	-18.4%	44,126	38,943	19,917	5,014	122,895
タイ	127,583	-15.3%	34,532	25,858	16,529	4,821	45,414
シンガポール	163,210	-12.9%	58,462	34,903	24,891	3,242	40,428
マレーシア	132,353	-12.0%	41,047	27,753	17,999	3,671	40,023
インドネシア	136,619	-7.2%	47,117	22,586	27,171	5,051	34,665
フィリピン	112,228	-11.3%	33,192	26,505	13,419	4,959	34,093
ベトナム	186,138	-4.5%	49,314	54,182	19,556	4,125	58,883
インド	144,275	-2.7%	61,354	27,379	19,713	2,784	32,971
英国	181,795	-13.7%	80,131	41,220	30,011	6,408	24,000
ドイツ	171,009	-0.0%	78,849	36,899	29,715	5,068	20,350
フランス	189,006	-9.7%	75,462	40,799	34,590	7,830	30,299
イタリア	198,000	-2.0%	78,597	45,299	38,900	9,315	25,596
スペイン	224,072	-1.4%	92,211	49,103	44,168	6,482	32,052
ロシア	190,874	+4.6%	61,845	31,004	21,812	7,258	64,889
米国	171,418	-2.4%	70,707	41,137	27,856	5,427	26,111
カナダ	154,977	-9.2%	55,366	36,319	28,509	5,184	29,326
オーストラリア	246,866	+6.7%	99,802	51,202	40,169	17,957	37,587
その他	183,083	+2.8%	67,698	41,889	29,633	6,505	37,333



出处：观光白书

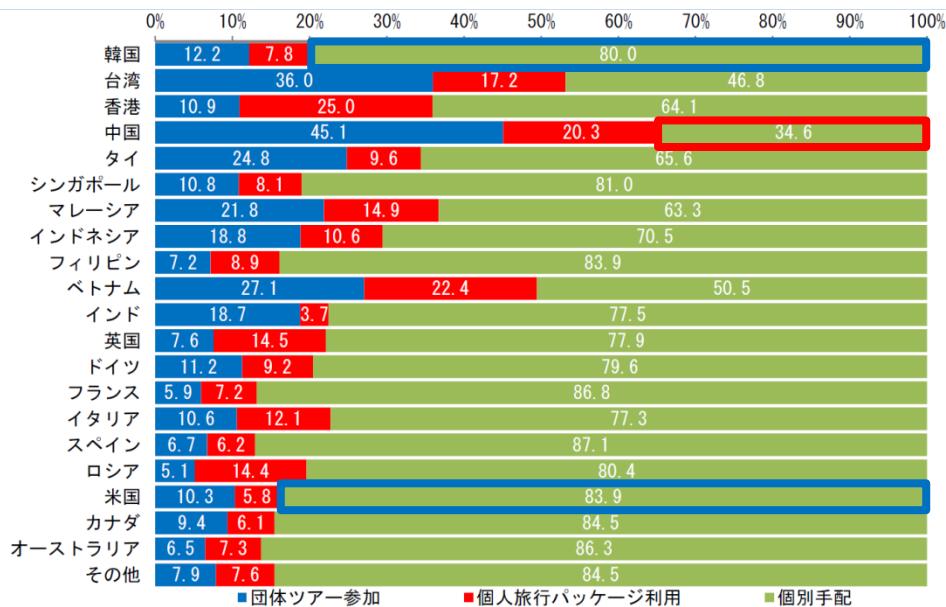
出处：观光白书

# 訪日外国人旅行者の属性

## 訪日外国人旅行者の来日回数



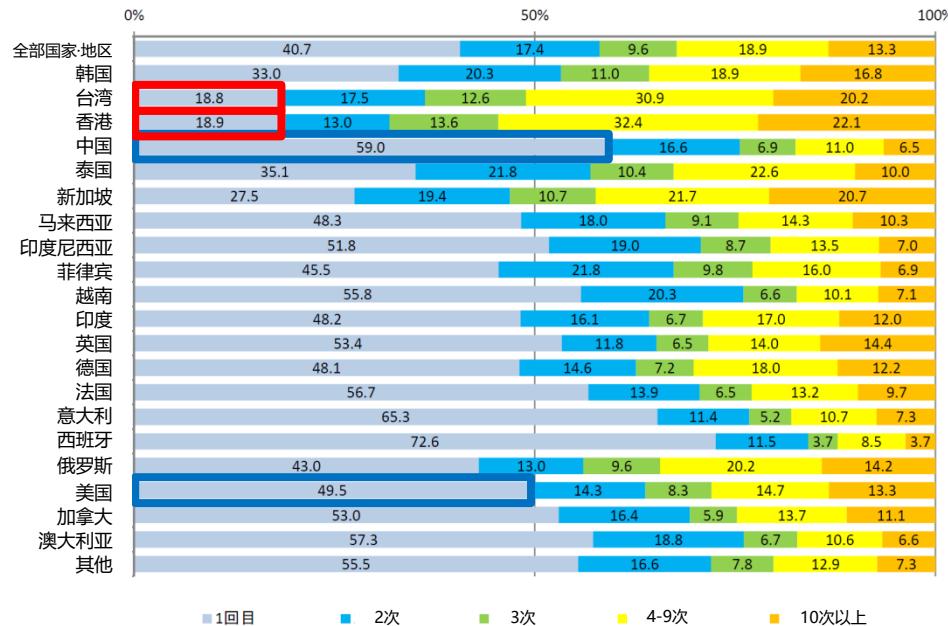
## 訪日外国人旅行者の旅行形態 (2016)



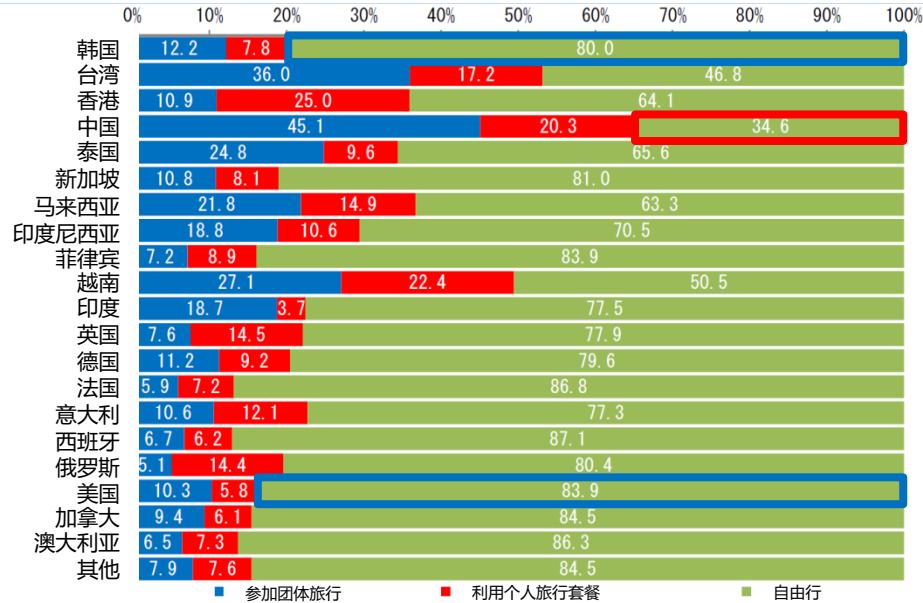
出典：訪日外国人旅行者消費動向調査・観光白書

# 访日外国人旅行者的属性

## 访日外国人旅行者的访日次数



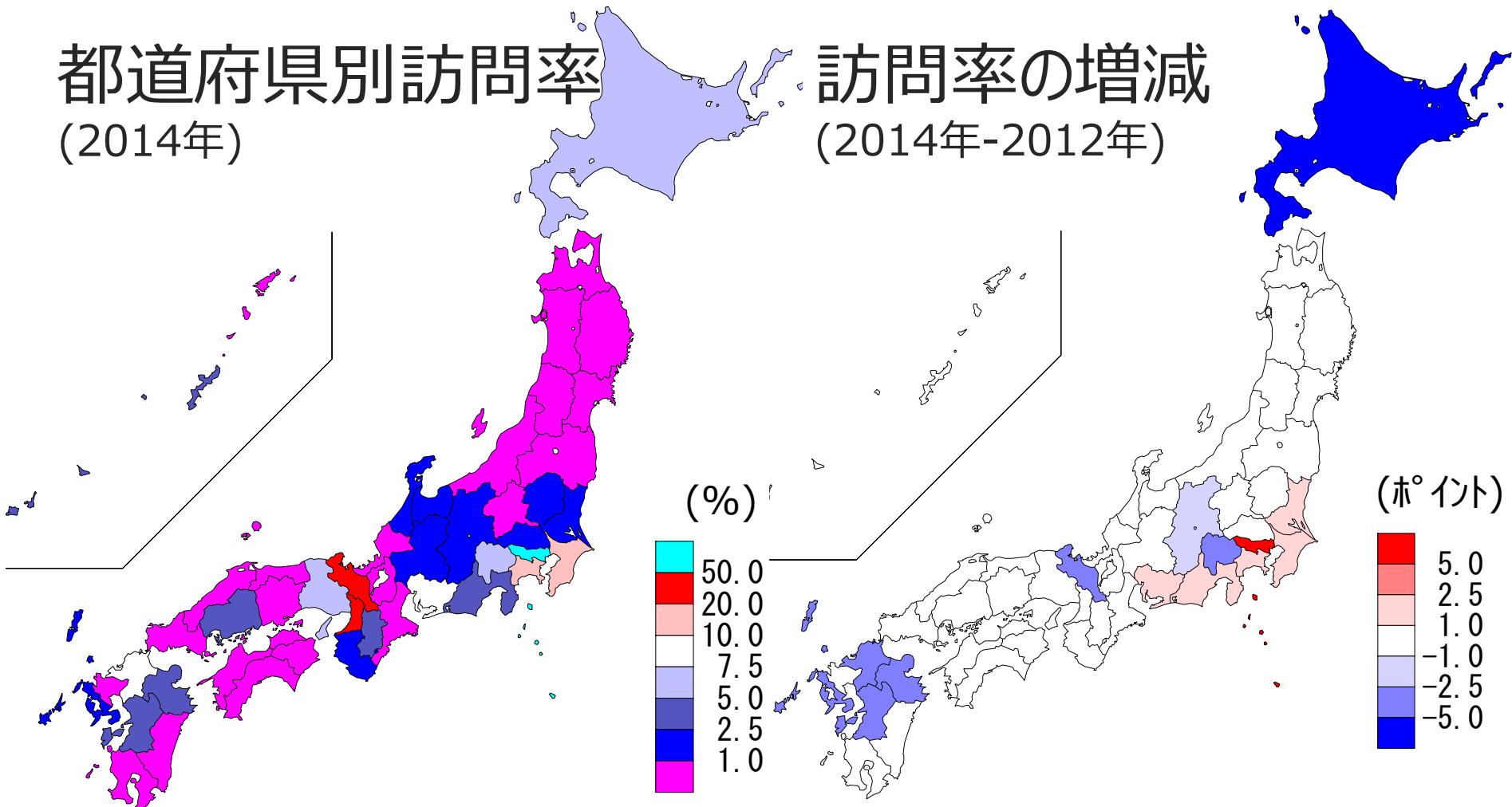
## 访日外国人旅行者的旅行方式 ( 2016 )



出处：访日外国人旅行者消费动向调查·观光白书

# 都道府県別訪問率 (2014年)

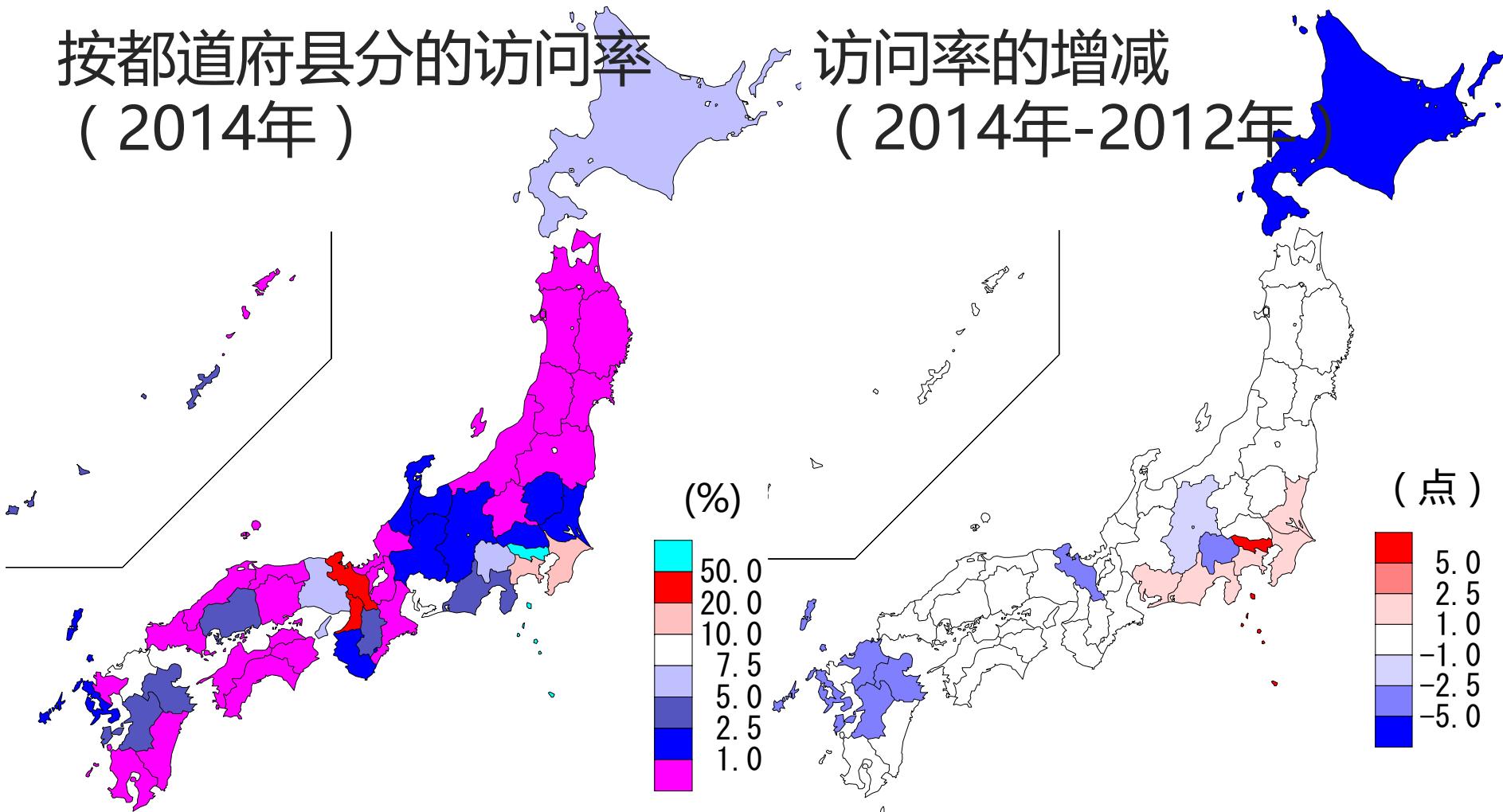
# 訪問率の増減 (2014年-2012年)



- 平均訪問都道府県数 : 2.37
- 訪問地組合せ（都道府県単位） 全1074パターン
  - ①東京都のみ : 17%, ②北海道のみ : 10%, ③沖縄県のみ : 9%,
  - ④大阪府のみ : 6%, ⑤大阪府・京都府 : 6%

# 按都道府县分的访问率 ( 2014年 )

# 访问率的增减 ( 2014年-2012年 )



- 平均的访问都道府县数 : 2.37
- 访问地的组合 ( 都道府县单位 ) 全部1074类型
  - ① 东京都一地 : 17% , ② 北海道一地 : 10% , ③ 冲绳县一地 : 9% , 12
  - ④ 大阪府一地 : 6% , ⑤ 大阪府·京都府 : 6%

数据 : 访日外国人消费动向调查 ( 观光厅 )

# 地方部への誘客に向けて-交通施策・推進体制

## 外国人旅行者の特徴

(S.シャウエッカ-氏(ジャパンガイド編集長),観光文化233号)

- 食：寿司、ラーメン⇒リピート層：冬のカニを北陸で
- 土産：包丁,酒←自国での**日本食レストランの経験**
- 体験：暮らすような旅行をしてみたい→体験・動機付け
  - 交流体験を紹介するプラットフォーム
    - asoview (日本最大級のレジャー・体験予約サイト)
    - NAGOMI VISIT (交流体験をコーディネイトするNPO)

asoview



NAGOMI  
VISIT

界中の旅人と食卓を囲むホームビジット

VISITは、世界中から日本にやってくる旅行者を我が家に招き、普段のおうちごはんと一緒に食べながら2~3時間交流をする、ホーム

って普段と違うおもてなし料理を作つて迎えるわけではなく、毎日おうちで食べている、普通の家庭料理を用意します。

数時間。しかも、いつものごはんと一緒に食べるだけ。たったこれだけのことだけど、自宅にいながらにして世界中の様々な人と過ごすスタイルになるはずです。

らず、シンプルに。いつもの生活中に、グローバルな体験、取り入れてみませんか？



# 针对向地方部诱客-交通措施・推進体制

## 外国人旅行者的特征

( S. Schauwecker(日本指南主编), 观光文化233号)

- 美食：寿司、拉面→多次访问者：北陆冬天的螃蟹
- 特产：菜刀, 酒←在自己国内的日本料理店的经验
- 体验：希望尝试在目的地生活一样的旅行→体验·动机的形成
  - 介绍交流体验的平台
  - asoview ( 日本最大的度假 · 体验预约网站 )
    - NAGOMI VISIT ( 包含交流体验的公益组织)

asoview



とっておきのワクワクを見つけよう  
北海道から沖縄まで、16,198件のレジャー・遊び・体験を紹介、予約できます

着物・浴衣レンタル

人気のレジャー・体験・遊びプラン続々追加中！！掲載プラン数 16,198件

NEW 【兵庫・カヌー】基本を教わる  
NEW 【沖縄・サーフィン】ショート  
NEW 【沖縄・北谷・SUP】最大10  
NEW 【千葉・七宝焼】カラフルな  
【千葉・陶芸体験】まっさら

NAGOMI VISIT

界中の旅人と食卓を囲むホームビジット

VII VISITは、世界中から日本にやってくる旅行者を我が家に招き、普段のおうちごはんと一緒に食べながら2~3時間交流をする、ホーム

って普段と違うおもてなし料理を作つて迎えるわけではなく、毎日おうちで食べている、普通の家庭料理を用意します。

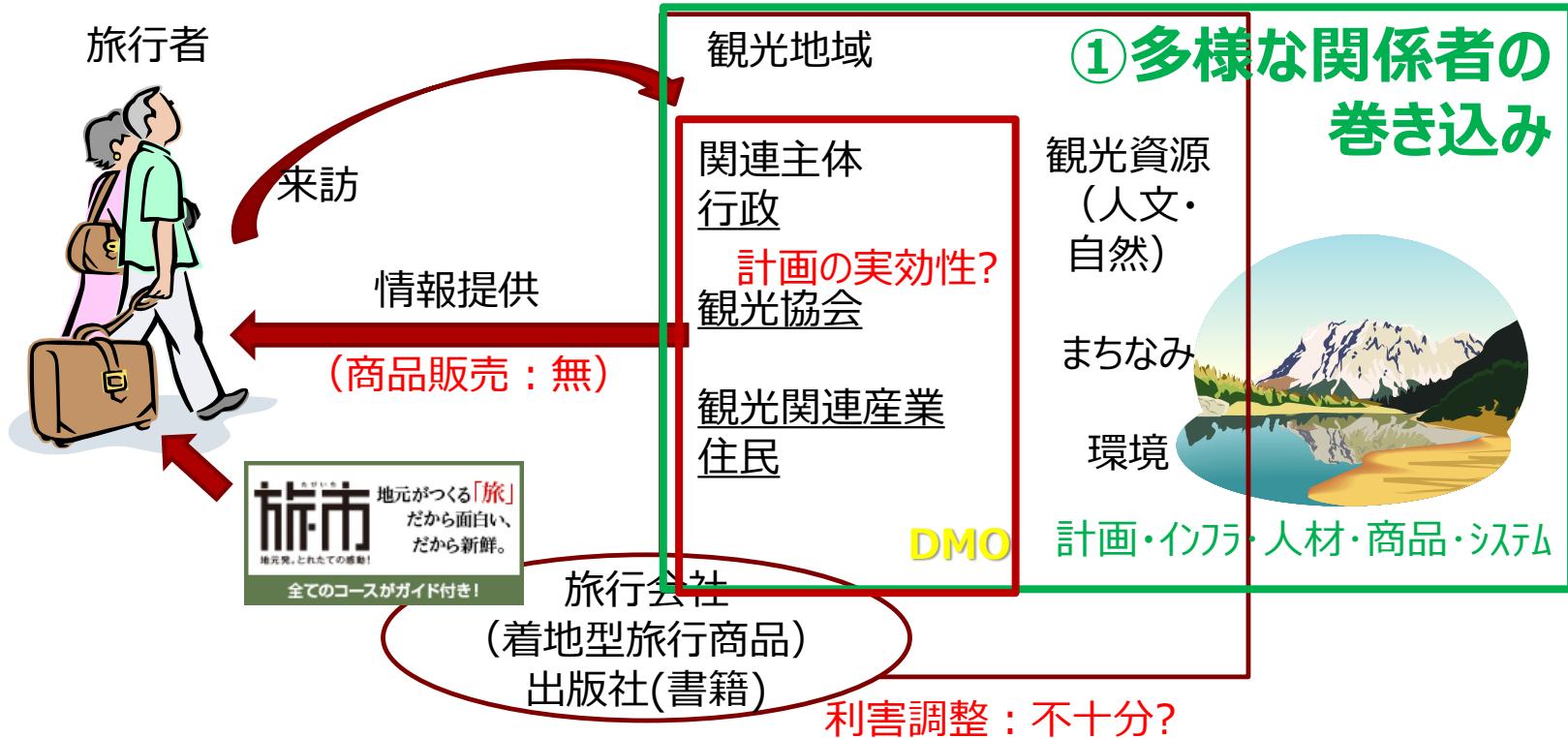
数時間。しかも、いつものごはんと一緒に食べるだけ。たったこれだけのことだけど、自宅にいながらにして世界中の様々な人と過ごすスタイルになるはずです。

らず、シンプルに。いつもの生活中に、グローバルな体験、取り入れてみませんか？



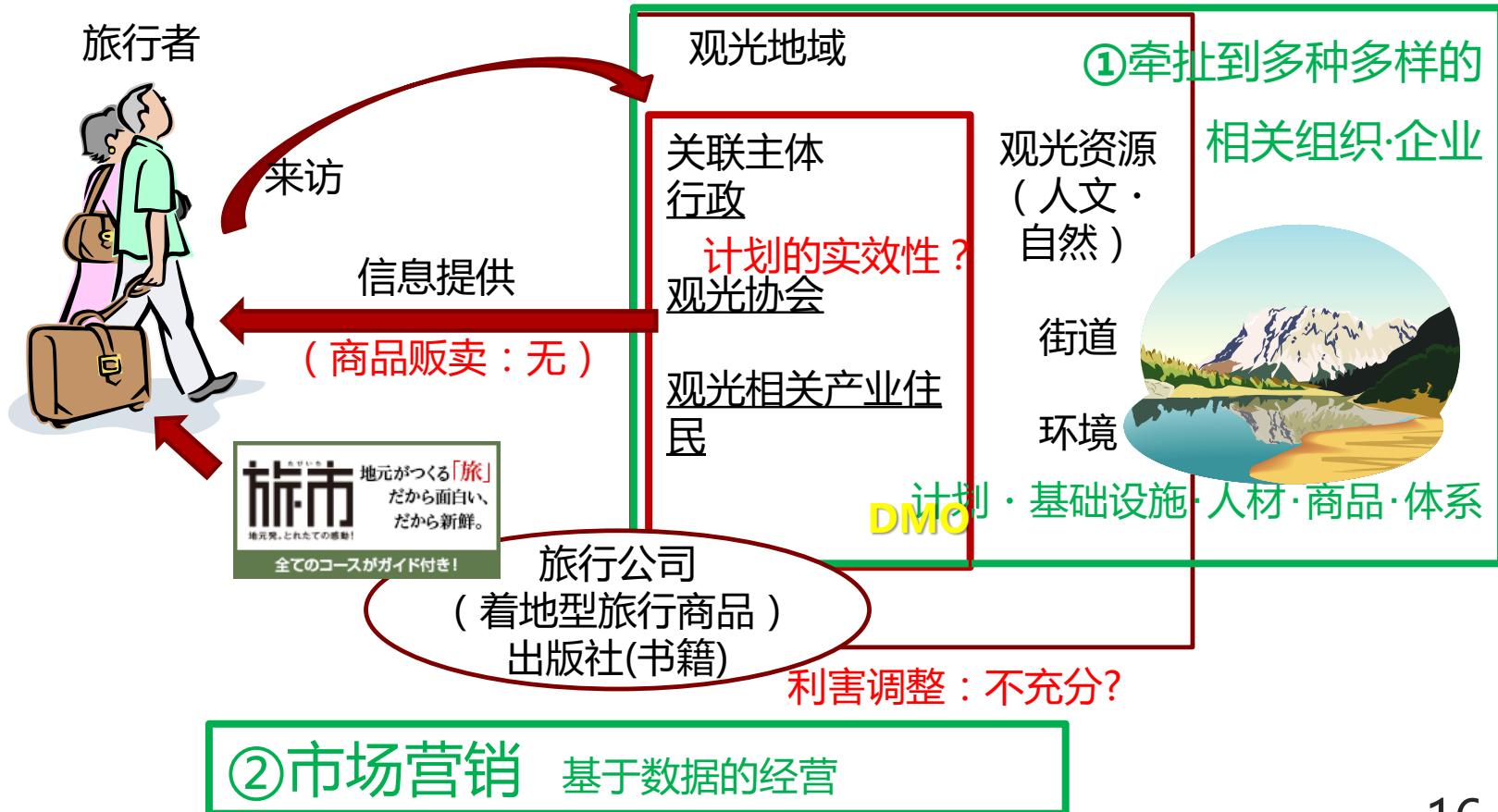
# 観光振興のための体制整備：DMO

- ・ プラニングやマーケティングに加えて，(Destination Management/Marketing Organization)
- ・ 地域の観光関係主体や観光資源の効果的な管理・調整等を行う組織.



# 为了观光振兴的体制整备：DMO

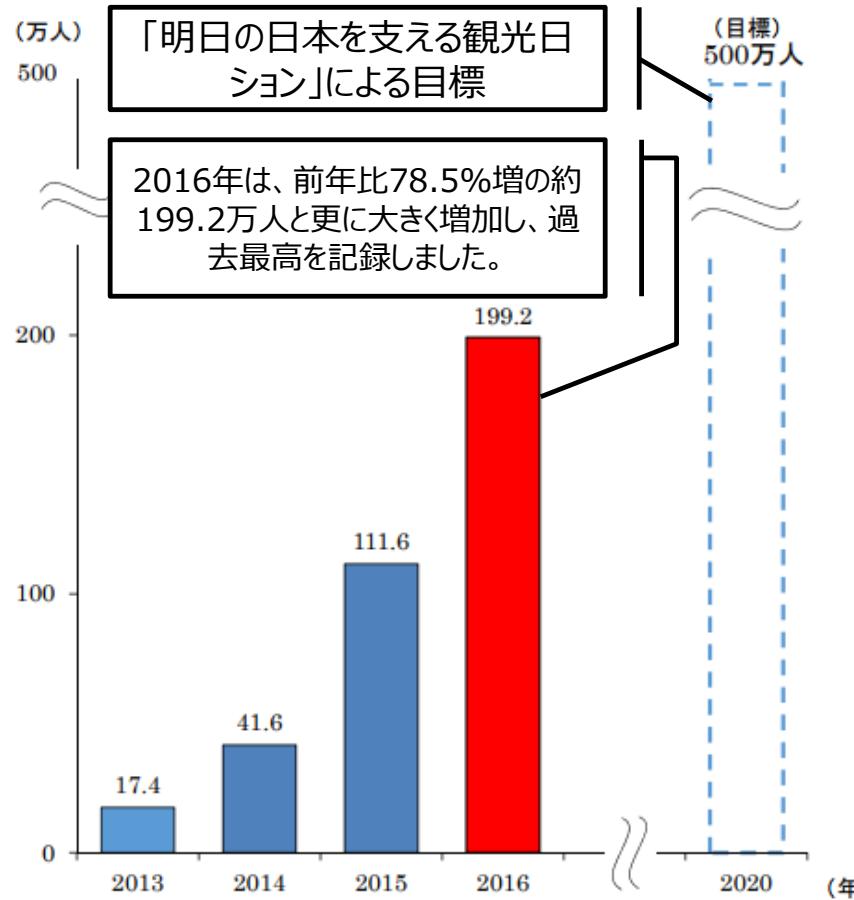
- 包含策划及市场营销，(Destination Management/Marketing Organization)
- 实行地方的观光相关主体或观光资源的有效管理・调整等的组织



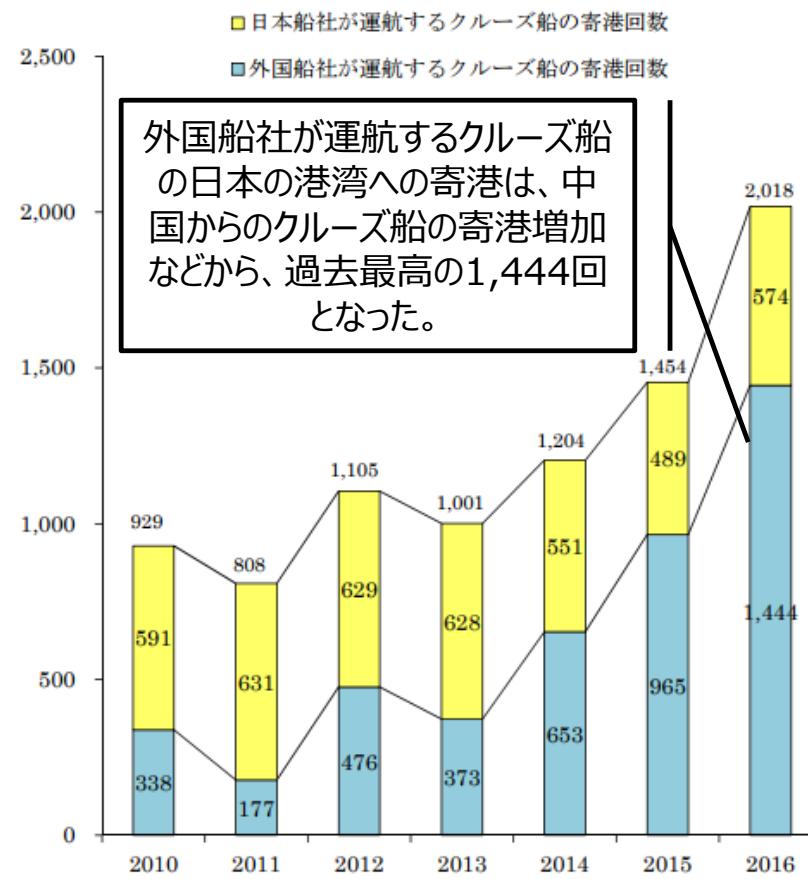
③ 相关业界的调整・结构创成, 宣传与收益确保

# クルーズ観光の増加

クルーズ船による外国人入国者数  
(2013-2016)



日本の港湾へのクルーズ船の寄港回数  
(2013-2016)

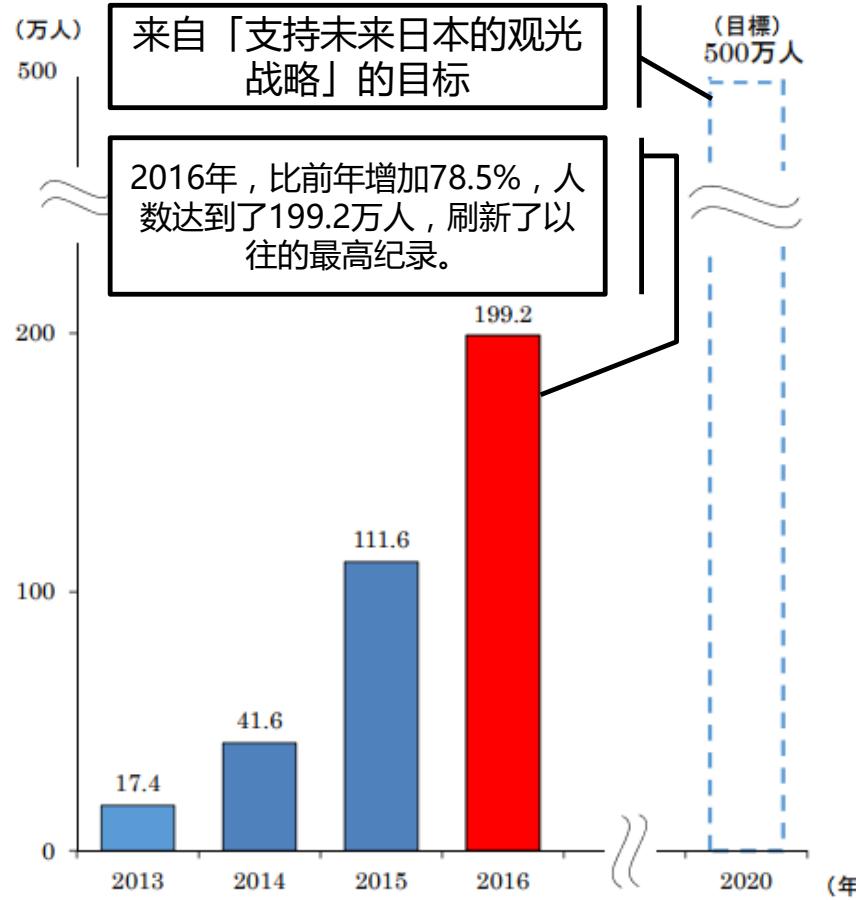


出典：国土交通省「2016年の訪日クルーズ旅客数とクルーズ船の寄港実績（速報値）」

クルーズ船による外国人入国者数は年毎上昇し、新たな日本を訪問する手段となっている。外国のクルーズ船が日本の港湾に寄港する回数年毎上昇し、特に中国籍のクルーズの増加に伴い、過去最高となった。

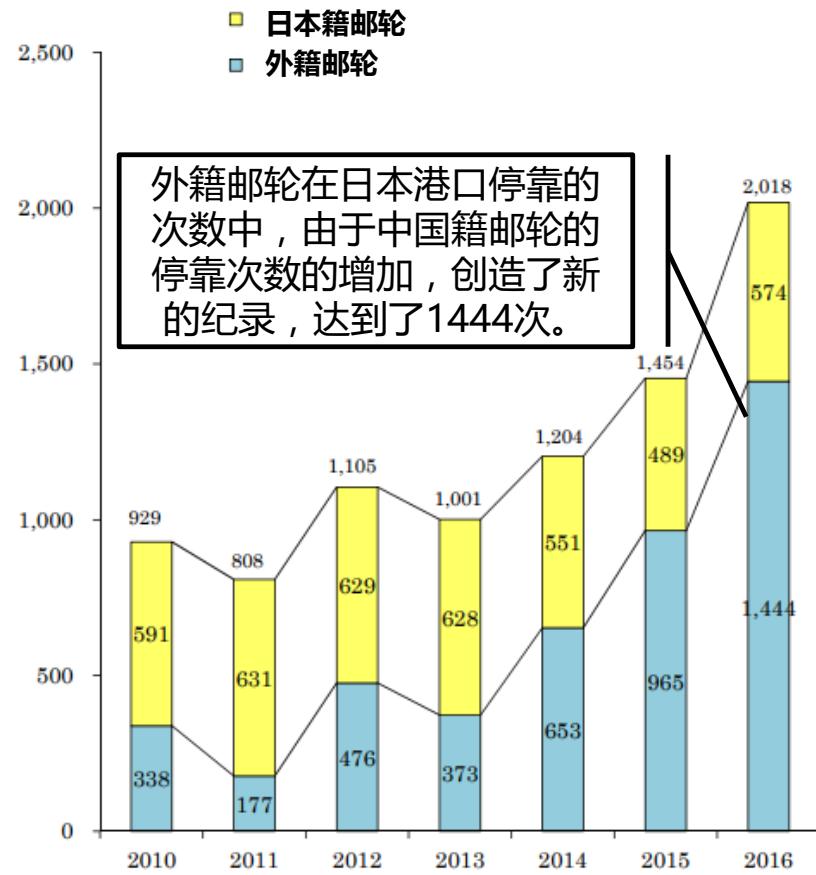
# 邮轮观光的增加

通过邮轮进入日本的外国人旅行者数  
(2013-2016)



出处：国土交通省《2016年访日邮轮旅客数及邮轮到港实际情况（速报值）》

在日本的港湾停靠的邮轮的回数  
(2013-2016)



经由邮轮进入日本国境的旅行者数量每年增加，邮轮也成为了新的访日交通工具。

外籍邮轮在日本港口停靠次数每年增加，特别中国籍邮轮增加，外籍邮轮停靠次数创新高

# 旅行記上の旅行情報の抽出による 訪日中国人旅行者の旅行行動に関する研究

東洋大学大学院国際観光学専攻博士前期課程1年 宋紫龍

東洋大学国際観光学部 古屋秀樹

## 発表の構成：

- 本研究の概要
- 分析データ
- 分析方法
- 分析結果
- まとめと今後の課題



# 通过对游记中旅游相关信息抽出的 访日中国人游客旅行行动相关的研究

东洋大学研究生院国际观光学专业研究生1年 宋紫龙

东洋大学国际观光学专业 古屋秀树

## 发表的构成：

- 本研究的概要
- 分析数据
- 分析方法
- 分析结果
- 总结及今后的课题



# 本研究の概要

## □ 背景 :

- ・ 訪日中国人は日本インバウンド旅行市場の重要な部分である
- ・ 旅行ポータルサイト上の情報が多くなり、役に立つようになっている
- ・ 中国における旅行記の使用者が増えている

## □ 目的 :

- ・ 旅行記の実態分析、収集及び特徴の抽出方法の検討
- ・ 訪日中国人旅行者の個人属性別の旅行嗜好を把握し、新たな旅行行動の形を発掘すること

## □ 意義 :

- ・ 訪日中国人旅行者の旅行行動を把握し、旅行者の満足度の醸成や再訪意向の高まりが達成することに対して、資すること。

## □ 研究方法 :

- ・ クローラーのWebページ上の旅行記の収集、形態素分析、トピックモデル

## □ 本研究の位置づけ :

- ・ トピックモデルの分析方法を用い、訪日中国人旅行者の訪問地、訪問観光スポット、旅行内容を対象とし、訪日中国人旅行者の実態分析

## □ キーワード :

観光情報学、自然言語処理、形態素、旅行行動、パターン化

# 本研究的概要

## □ 背景：

- 目前，中国人访日游客是日本入境旅游市场中的重要组成部分
- 旅行网站中的旅游相关信息不断增加，可参考度也在不断增加
- 在中国，使用游记的人正在不断增加

## □ 目的：

- 对于游记实际状况的分析·收集以及特征抽出的方法的探讨
- 把握访日中国人旅行者的按个人属性区别的偏好，发掘出访日旅行新形态

## □ 意义：

- 把握访日中国人旅行者的旅行行动现状，对于提高满足度以及提高再访意向做出贡献。

## □ 研究方法：

- 通过爬虫对于网页中的游记收集、关键词抽取（中文分词）、主题模型（LDA）

## □ 本研究的定位：

- 使用主题模型，以访日中国人旅行者的访问地点，访问的景点，访问的内容为对象进行的访日中国人旅行者旅行行动的现状分析

## □ 关键词：

观光情报学，自然语言处理，分词，旅行行动，归类

# 分析データ

本研究の位置づけにより、「訪日中国人旅行者の訪問地、訪問観光スポット、旅行内容」が本研究の分析対象となっており、このようなデータの探しと収集は本研究の課題・肝心である。

The screenshot shows the homepage of the Mafengwo website. At the top, there is a navigation bar with links for '首页' (Home), '目的地' (Destination), '旅游攻略' (Travel Guide), '自由行商城' (Freelance Travel Mall), '酒店' (Hotels), '社区' (Community), and search functions. Below the navigation, there is a breadcrumb trail: '目的地 > 日本 > 日本游记'. The main content area features three travel forum posts:

- 走你走过的风景，算不算重** (走你走过的风景，算不算重) by 无忧无虑的余姑娘 (无忧无虑的余姑娘) on 17-08-06 09:35 (1802 replies). The post discusses the atmosphere of a traditional Chinese opera stage.
- 独行日本，探寻东京周边绝密美景** (·独行日本，探寻东京周边绝密美景) by LUKEINSHOT (LUKEINSHOT) on 17-08-06 10:20 (1306 replies). The post is about solo travel in Japan, specifically around Tokyo.
- 【你的东京，我的京都，她的大阪】我们** (【你的东京，我的京都，她的大阪】我们) by 弹琵琶的小虎牙 (弹琵琶的小虎牙) on 17-08-06 (344 replies). The post is a collaborative travel thread involving three cities.

On the right side of the page, there is a sidebar with a download link for the 'Japan Travel Guide' app and a section for '日本问答' (Japanese Questions and Answers).

- 旅行記の中で、一人旅行者の旅行経験が豊富的に記載されている。
- 「蚂蜂窩\*」は中国における最大級の旅行ポータルサイトであり、中国旅游研究院と連携しており、データの真実性を保証できる。

## 課題

多くの数の旅行記を収集することは大変

出处：「蚂蜂窩」网站

手動ではなく、クローラーによる自動的なWebページ上の情報を取得する方法を用いた。

蚂蜂窩\*：「蚂蜂窩」の日本語の意味は「蜂の巣」であり、中国のトリップアドバイザと想定できる。

# 分析数据

本研究将「访日中国人旅行者的访问地点，访问的景点，访问的内容」作为研究对象。如何对这类的数据进行收集，便成为了本研究的**重点以及难点**。

The screenshot shows the homepage of the Mafengwo travel website. At the top, there's a navigation bar with links for '首页' (Home), '目的地' (Destinations), '旅游攻略' (Travel Guides), '自由行商城' (Independent Travel Mall), '酒店' (Hotels), '社区' (Community), and search functions. Below the navigation, there's a banner for '日本游记' (Japan Travel Diaries). The main content area displays three travel posts:

- 走你走过的风景，算不算重** (Walked through your landscape, is it not heavy?) by 无忧无虑的余姑娘 (无忧无虑的余姑娘) with 1802 replies. It features a photo of people at a food stall.
- 独行日本，探寻东京周边绝密美景** (Solo travel in Japan, exploring secret beauty around Tokyo) by LUKEINSHOT (LUKEINSHOT) with 1306 replies. It features a photo of a street scene in Japan.
- 【你的东京，我的京都，她的大阪】我们** (Our Tokyo, Your Kyoto, Her Osaka) by 弹琵琶的小芽牙 (弹琵琶的小芽牙) with 344 replies. It features a photo of a woman in a kimono.

At the bottom left, there's a note: "出典：「蚂蜂窝」のHPから抜粋" (Excerpted from the Mafengwo website).

- 游记中包含着大量的旅行者的旅行经验。
- 蚂蜂窝为中国**最大的旅游互动交流相关网站**，而且与中国旅游研究院合作，**数据真实可靠**。

难点

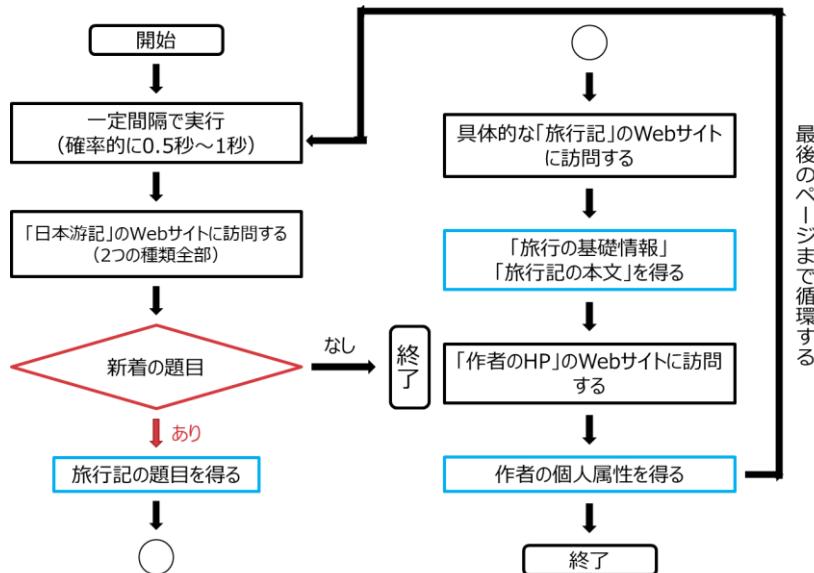
大量收集网页中的游记非常的困难

最终本研究选定了使用爬虫对网页上的信息进行收集的方法，而非手动收集。

蚂蜂窝\*：「蚂蜂窝」的意思为蜜蜂的巢穴。此网站可看成中国的猫途鹰。

# 分析データ

図1. クローラーによる旅行記収集の流れ



出典：筆者より作成

## □ データの自動収集時期：

4月上旬～7月中旬

(各旅行記の出発時間は1999年8月から2017年7月迄)

## □ データの自動収集のサンプル数：

旅行記の総計：10606篇

## □ データの自動収集の基本項目：

旅行記の題目、居住地、性別、出発時間、同行人物、滞在日数、コスト、旅行記の本文

図2. 収集したデータ

Database Structure										
Browse Data										
Edit Pragmas										
Execute SQL										
表:	crawlerDataBase_article	新建记录	删除记录	过滤	过滤	过滤	过滤	过滤	过滤	过滤
id	articleName	authorPos	authorSex	startTime	people	duringDay	cost	articleInfo	keyWords	过滤
1	「从成都到多 一横行列车…	None	Male	2016-08-08	一个人	15 天	12000RMB	八千里铁道…	「特色」、「便当」、「便宜」、「宽敞」、…	
2	「一场完全不攻 略走的『赏』」	上海	Male	2017-03-29	带小孩	7 天	None	2016年07月25 日，我向自己…	「便宜」、「宽敞」、…	
3	「冬日晚归内海 的那片蓝」	上海	Female	2016-12-02	一个人	8 天	10000RMB	插播小广告…	「轻松」、「原爆」、…	
4	「今晚报马斯 塞」	上海	None	2017-02-02	带小孩	7 天	6000RMB	今天的行程，当…	「特色」、「激动」、…	
5	「炎炎的太阳们 —香蕉椰子…」	None	Female	2014-10-23	带小孩	9 天	None	关西在哪里…	「豪华」、「度假」、…	
6	「降落人间的花 海天堂—让…」	None	None	2017-05-01	和朋友	8 天	None			
7	「(一人旅…尋 跡)【北九洲】 【冲绳】	日本	Male	2015-05-03	一个人	4 天	2500RMB	有时候，一个…	「便当」、「完美」、…	
8	「(冲绳)【冲绳 —香蕉椰子…】	None	Female	2014-10-15	带小孩	5 天	None	为什么星空…	「不要」、「甜点」、…	
9	「日本，既是异 国又是往昔」	北京	Male	2017-02-06	其它	8 天	10000RMB	早年NBA流…	「充满」、「不同」、…	
10	「13天扫遍所有 日料TOP5名…」	None	Male	2016-05-26	其它	14 天	None	友情提示：先看…	「一群」、「灿烂」、…	
11	「日本东北大众 冰雪之旅」	北京	Male	2017-01-23	小两口	9 天	13000RMB	计划这次的…	「美丽」、「便当」、…	
12	「(在北海道的後 期)【北海道】 【13天连假】	香港	Male	2016-10-01	小两口	5 天	10000RMB	『菊花夕拾』…	「独特」、「浪漫」、…	
13	「13天连假， 感受美好的…」	日本	Male	2016-12-27	一个人	13 天	None	大学时便知…	「好奇」、「崇拜」、…	
14	「(北海花园)漫 画我的北海道…」	重庆	Male	2016-07-15	和朋友	13 天	11000RMB	这一次的漫…	「最美」、「最好」、…	
15	「想在日本便 重游一遍」	None	None	2017-05-01	和朋友	8 天	None			
16	「日本东北大众 之旅」	上海	Male	2016-07-17	和朋友	6 天	None	青森县不同…	「有声」、「清晰」、…	
17	「避开了人，桥人, 名古屋南边…」	上海	None	2016-10-14	带小孩	5 天	8000RMB	日本人都爱…	「特色」、「宽敞」、…	
18	「日本东山急乐 园鬼屋倒影…」	上海	None	2017-01-01	和朋友	10 天	None	Fujio!富士急…	「恶搞」、「恐怖」、…	
19	「次元壁被破！」	江苏省	Female	2016-05-21	小两口	9 天	None	2016年5月,…	「方便」、「有趣」、…	

出典：筆者より作成

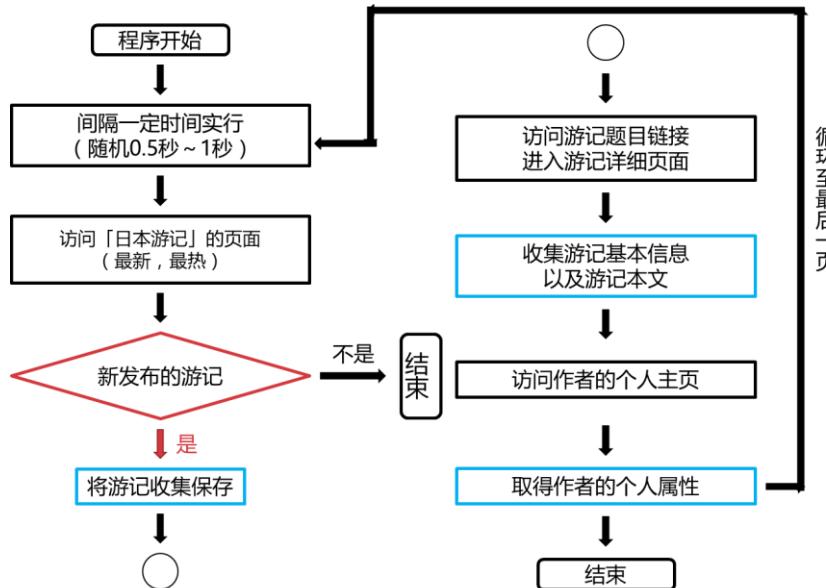
表1. 形態素の分析結果

種類	数(個)
名詞	96386
地名又は観光スポット名	2495
合計	98881

出典：筆者より作成

# 分析数据

图1. 通过爬虫收集游记的流程



出处：由作者制作

## □ 自动收集数据的时间段：

4月上旬~7月中旬

(各游记的出发时间为1999年8月到2017年7月)

## □ 样本数量：

游记总计：10606篇

## □ 收集到的基本项目：

游记题目，作者居住地，性别，出发时间，同行人，在日滞在天数，花费，游记本文

图2. 收集到的数据

Database Structure										
表: crawlerDataBase_article										
id	articleName	authorPos	authorSex	startTime	people	duringDay	cost	articleInfo	keyWords	过滤
1	['从京都到佐多...一横行列车…']	None	Male	2016-08-08	一个人	15 天	12000RMB	八千里铁道…	['特色', '便当...', '便宜', '宽敞...', '便宜', '方便...', '精致', '最高...', '高松', '原宿...', '插播广告...', '今天的行程', '仁不让的要…', '豪华', '浪漫...', '精致', '美好...', '']	
2	['一场完全没有攻略走的“赏...']	上海	Male	2017-03-29	带小孩	7 天	None	2016年07月25日,我向自己...']	['精致', '最高...', '精妙', '']	
3	['冬日晚游内海的那片蓝...']	上海	Female	2016-12-02	一个人	8 天	10000RMB	插播广告...	['高松', '原宿...', '插播广告...', '今天的行程', '仁不让的要…', '豪华', '浪漫...', '精致', '美好...', '']	
4	['今晚住在马斯...']	上海	None	2017-02-02	带小孩	7 天	6000RMB	今天的行程,当...']	['特色', '便当...', '便宜', '宽敞...', '便宜', '方便...', '精致', '最高...', '高松', '原宿...', '插播广告...', '今天的行程', '仁不让的要…', '豪华', '浪漫...', '精致', '美好...', '']	
5	['从西雅图到...']	None	Female	2014-10-23	带小孩	9 天	None	关西在哪里...']	['豪华', '浪漫...', '精致', '最高...', '高松', '原宿...', '']	
6	['【攻略】...']	None	None	2017-05-01	和朋友	8 天	None			
7	['【一人旅】...']	日本	Male	2015-05-03	一个人	4 天	2500RMB	有时候,一个...']	['便当', '有名...', '正好', '完整...', '不厚', '甜品...', '']	
8	['【穷游】...']	None	Female	2014-10-15	带小孩	5 天	None	为什么星冲...']	['豪华', '浪漫...', '精致', '最高...', '高松', '原宿...', '']	
9	['日本...又是往昔...']	北京	Male	2017-02-06	其它	8 天	10000RMB	早年NBA流...']	['充满', '不同...', '一腔', '灿烂...', '']	
10	['13天扫遍所有...']	None	Male	2016-05-26	其它	14 天	None	友情提示:先看...']	['好看', '浓郁...', '小巧玲珑...', '简单', '厚切...', '']	
11	['日本东北小众...']	北京	Male	2017-01-23	小两口	9 天	13000RMB	计划这次的...']	['美丽', '便当...', '好看', '著名...', '友好', '舒适...', '']	
12	['在日本...的伎...']	香港	Male	2016-10-01	小两口	5 天	10000RMB	『菊花夕拾』...']	['独特', '浪漫...', '别致', '']	
13	['13天关西漫...']	日本	Male	2016-12-27	一个人	13 天	None	大学时便知...']	['好看', '靓丽...', '宽敞', '无聊...', '']	
14	['【北上花间】...']	重庆	Male	2016-07-15	和朋友	13 天	11000RMB	这一次的漫...']	['最美', '最好...', '宽敞', '精致...', '']	
15	['想在日本便...']	None	None	2017-05-01	和朋友	8 天	None			
16	['日本东北小众...']	上海	Male	2016-07-17	和朋友	6 天	None	青森县不同...']	['有名', '清晰...', '透彻', '']	
17	['【避开人...']	上海	None	2016-10-14	带小孩	5 天	8000RMB	日本人都爱...']	['特色', '宽敞...', '享受', '隐私...', '']	
18	['日本富士急乐...']	上海	None	2017-01-01	和朋友	10 天	None	Fujio!富士急...']	['慈祥', '恐怖...', '微博', '阴暗...', '']	
19	['次元壁被破...']	江苏省	Female	2016-05-21	小两口	9 天	None	2016年5月,...']	['方便', '有爱...', '']	

出处：由作者制作

表1. 中文分词结果

种类	数量 (个)
名词	96386
地名, 景点名	2495
合计	98881
出典：筆者より作成	26

# 分析方法

古屋(2016)により、単独のトピックならびにすべての被験者が同一の分布を有することを仮定する潜在クラスは、**制約の強いモデル**であると考えられる。また、モデル推定においてパラメータ推定の事前確率を想定していないことから、データによる**過学習の恐れ**がある。



統計的潜在意味解析の1つの方法であり、教師データなしの機会学習の1つであるトピックモデルを用いた。

図3.2 250篇の旅行記本文に対するトピックモデル分析によるワールドクラウド図



出典：筆者より作成

複数の単語の共起性によって創造された「潜在的意味」から、データ上、実際に共起した表層的（字面の）なパターンだけではなく、「潜在的な旅行嗜好」を発掘することもできる。



- ・ 新たな訪日中国人旅行者の旅行実態
  - ・ 訪日中国人旅行者の潜在的な旅行嗜好

# 分析方法

古屋于2016年在使用潜在类别模型进行旅游行动分析时提到，单独的主题以及所有被分析者都有统一分布的潜在类别分析是有**强烈制约性的模型**。而且，由于模型的参数推定没有对事前概率进行考虑，容易产生**过学习现象**。

本研究使用作为统计性潜在意思分析的其中一种，无监督学习的机械学习的其中一种的**主题模型**进行分析。

图3.对于2250篇游记进行主题模型分析后得到的词云图



出处：由作者制作

从多个单词的共起性被创造出的潜在的意思来看，实际上，并非从字面意思进行归类，而是试图发掘潜在旅行偏好。

- 访日中国人旅行者的旅行新形态
  - 访日中国旅行者的潜在旅游偏好

# 分析方法

## トピックモデル(LDA-Latent Dirichlet Allocation)

統計的潜在意味解析の1つの方法であり、教師データなしの機会学習の1つと位置づけられる。文書中の単語の順序を無視し、トピック別に単語と出現頻度のペアの集合をモデル化する方法である。

$$\text{公式} : P(\text{word}|\text{article}) = \sum_{topic} P(\text{word}|topic) \times P(topic|\text{article})$$

1. For トピック  $k = 1, \dots, K$ 
  - (a) 単語分布を生成  $\phi_k \sim \text{Dirichlet } (\beta)$
2. For 文書  $d = 1, \dots, D$ 
  - (a) トピック分布を生成  $\theta_d \sim \text{Dirichlet } (\alpha)$
  - (b) For 単語  $n = 1, \dots, N_d$ 
    - i. トピック生成  $z_{dn} \sim \text{Categorical}(\theta_d)$
    - ii. 単語を生成  $w_{dn} \sim \text{Categorical}(\phi_{z_{dn}})$

図4. トピックモデルの生成過程

出典：岩田具治（2015）：「トピックモデル」

トピック分布  $\theta_d$  と 単語分布集合  $\phi$  が与えられた際の文書  $w_d$  の確率：

$$p(w_d | \theta_d, \Phi) = \prod_{n=1}^{N_d} \sum_{k=1}^K p(z_{dn} = k | \theta_d) p(w_{dn} | \phi_k) = \prod_{n=1}^{N_d} \sum_{k=1}^K \theta_{dk} \phi_{kw_{dn}}$$

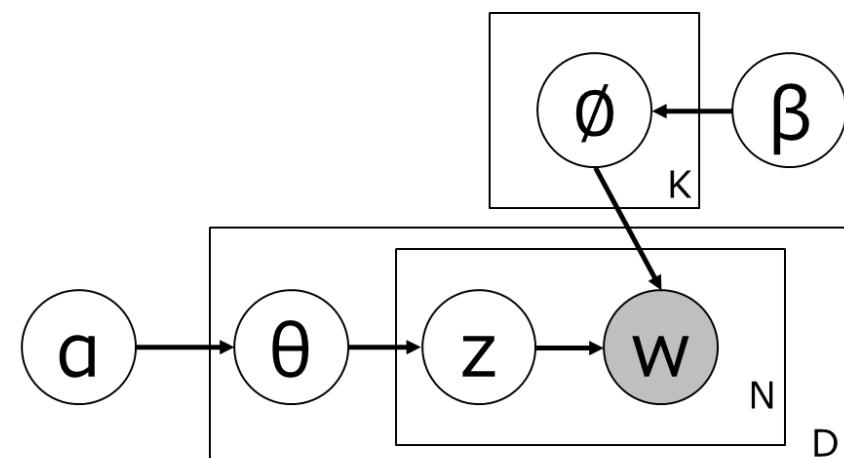


図5. トピックモデルのグラフィカルモデル

出典：佐藤一誠（2015）：「トピックモデルによる統計的潜在意味解析」

# 分析方法

## 主题模型(LDA-Latent Dirichlet Allocation)

主题模型是统计性潜在意思分析的其中一种，无监督学习的机械学习的其中一种。无视文章中的单词的顺序，对每个主题之中的单词以及出现频度组合的集合进行模型化的方法。

$$\text{公式} : P(\text{word}|\text{article}) = \sum_{\text{topic}} P(\text{word}|\text{topic}) \times P(\text{topic}|\text{article})$$

1. For 主题  $k = 1, \dots, K$

( a ) 生成单词的分布  $\phi_k \sim \text{Dirichlet}(\beta)$

2. For 文章  $d = 1, \dots, D$

( a ) 生成主题的分布  $\theta_d \sim \text{Dirichlet}(\alpha)$

( b ) For 单词  $n = 1, \dots, N_d$

i. 主题生成  $z_{dn} \sim \text{Categorical}(\theta_d)$

ii. 单词生成  $w_{dn} \sim \text{Categorical}(\phi_{z_{dn}})$

图4.主题模型的生成过程

出处：岩田真治（2015）：《主题模型》

被赋予主题分布  $\theta_d$  与单词分布集合  $\phi$  的时候文章  $w_d$  的概率：

$$p(w_d | \theta_d, \Phi) = \prod_{n=1}^{N_d} \sum_{k=1}^K p(z_{dn} = k | \theta_d) p(w_{dn} | \phi_k) = \prod_{n=1}^{N_d} \sum_{k=1}^K \theta_{dk} \phi_{kw_{dn}}$$

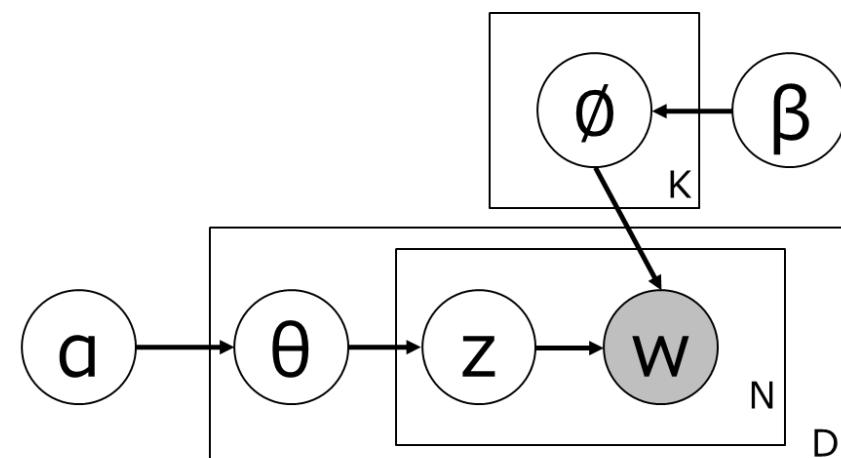


图5.主题模型的概率图模型

出处：佐藤一诚：《通过主题模型的统计潜在意思解析》

# 分析方法

## トピックモデル(LDA-Latent Dirichlet Allocation)

Latent  
潜在的

Dirichlet  
ディリクレ分布

Allocation  
配置

### □ ディリクレ分布

ディリクレ分布 (Dirichlet distribution) は、連続型の確率分布である。

多項分布の「共役事前分布 (conjugate prior)」である。

カテゴリ分布や多項分布のパラメータ  $\phi = (\phi_1, \dots, \phi_n)$  は  $\phi_v \geq 0, \sum_{v=1}^V \phi_v = 1$  を満たすべきトロリの確率分布としてディリクレ分布

$$\text{Dirichlet}(\phi|\beta) = \frac{\Gamma(\sum_{v=1}^V \beta_v)}{\prod_{v=1}^V \Gamma(\beta_v)} \prod_{v=1}^V \phi_v^{\beta_v - 1}$$

がある。ここで  $\beta = (\beta_1, \dots, \beta_n)$  はディリクレ分布のパラメータで、 $\beta_v > 0$  である。

「単語分布を生成  $\phi_k$ 」と「トピック分布を生成  $\theta_d$ 」はディリクレ分布に従い、  
ディリクレ分布はトピックモデルにとって、肝心ある。

# 分析方法

## 主题模型(LDA-Latent Dirichlet Allocation)

Latent

潜在的

Dirichlet

狄利克雷分布

Allocation

配置法

### □ 狄利克雷分布

狄利克雷分布 ( Dirichlet distribution ) 是连续型的概率分布。

是多项分布的共轭先验 ( conjugate prior ) 。

作为满足类别分布或多项分布的参数  $\boldsymbol{\phi} = (\phi_1, \dots, \phi_n)$  は  $\phi_v \geq 0, \sum_{v=1}^V \phi_v = 1$  的矢量  $\boldsymbol{\phi}$  的概率分布的狄利克雷分布

$$\text{Dirichlet}(\boldsymbol{\phi}|\beta) = \frac{\Gamma(\sum_{v=1}^V \beta_v)}{\prod_{v=1}^V \Gamma(\beta_v)} \prod_{v=1}^V \phi_v^{\beta_v-1}$$

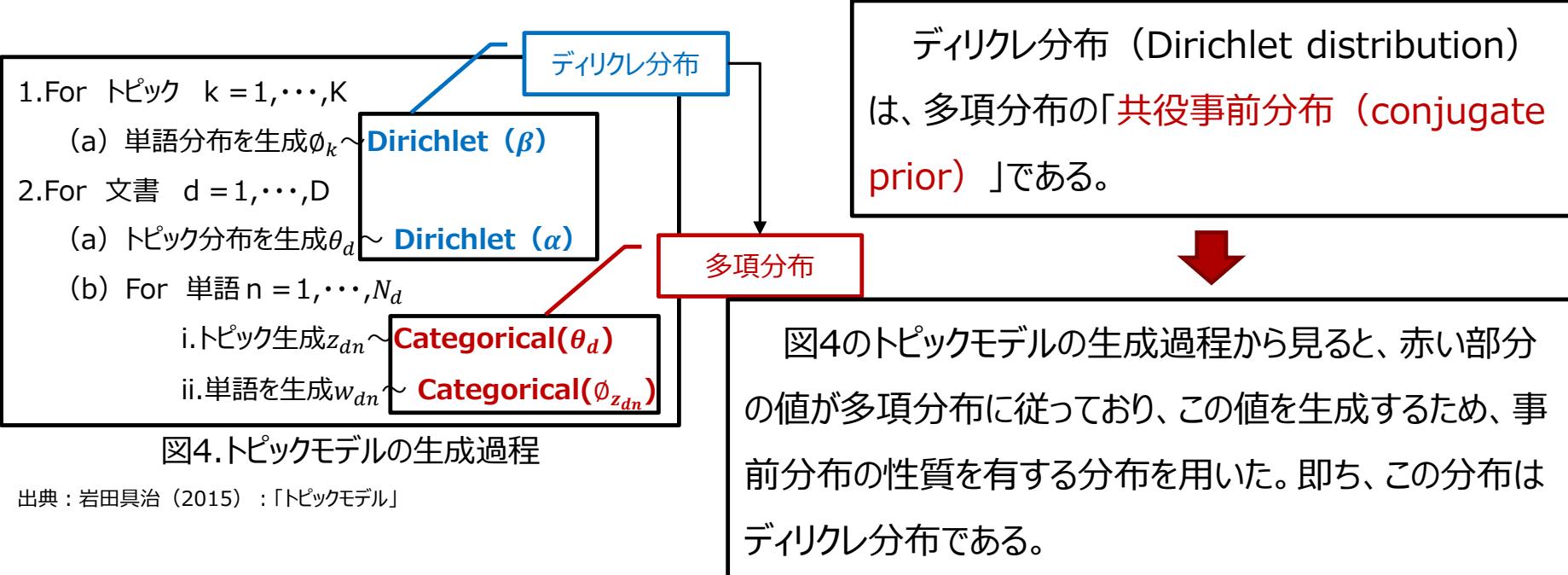
此处  $\beta = (\beta_1, \dots, \beta_n)$  是狄利克雷分布的参数， $\beta_v > 0$ 。

「生成单词分布的  $\boldsymbol{\phi}_k$ 」と「生成主题分布的  $\boldsymbol{\theta}_d$ 」全部遵循狄利克雷分布，  
对于主题模型来说狄利克雷分布是非常重要的组成部分。

# 分析方法

## トピックモデル(LDA-Latent Dirichlet Allocation)

ディリクレ分布を使う理由：



出典：岩田真治（2015）：「トピックモデル」

過学習を避けるために事前分布としてディリクレ分布を設定すること、データ数に対してパラメータが多い場合や比率が小さいセルが多い場合に偏った結果が導かれる危惧があることを避けるためである。

# 分析方法

## 主题模型(LDA-Latent Dirichlet Allocation)

使用狄利克雷分布的原因：

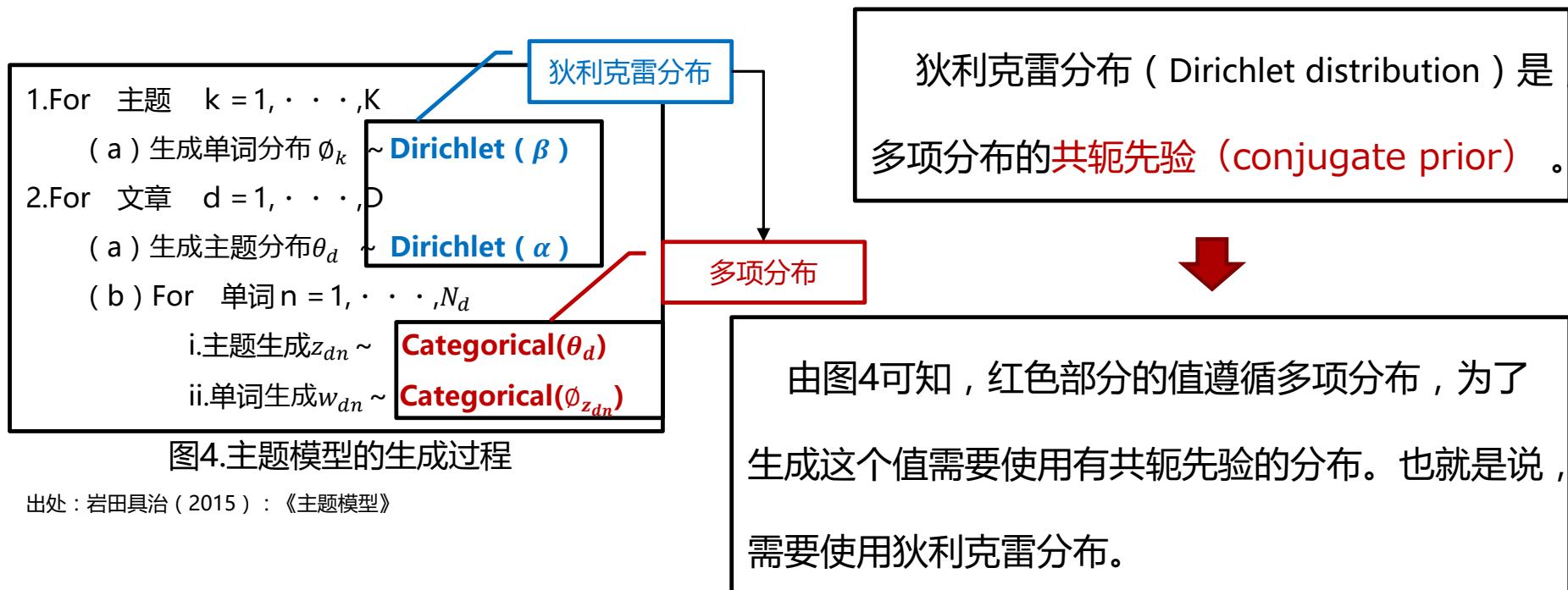


图4.主题模型的生成过程

出处：岩田真治（2015）：《主题模型》

为了防止过学习，需要作为共轭先验分布的狄利克雷分布。对于数据量的参数多的情况或者比率比较小的单元比较多的情况，可避免导致结果偏差的情况。

# 分析方法

## トピックモデルの数を評価する方法

確率モデルの性能を評価する尺度とし、テキストデータに対するパープレキシティ (perplexity) が使われる。パープレキシティは分岐数または選択肢の数を表しており、確率の逆数で定義されている。低いパープレキシティはテキストデータを高い精度で予測できるよい確率モデルであることを示す。したがって、筆者の理解する限り、パープレキシティはモデルに従って正解を選ぶためのある種の難しさを表す値である。

- Perplexityを算出する公式：

$$\text{perplexity}(W^{test}|M) = \left( -\frac{\sum_{d=1}^{D^{test}} \log p(W_d^{test}|M)}{\sum_{d=1}^{D^{test}} N_d^{test}} \right)$$

トピックモデルの尤度：

$$p(w_d|M) = \prod_{n=1}^{N_d} \sum_{k=1}^K \theta_{dk} \phi_{kw_{dn}}$$

本研究は10606個のサンプルに対し、1個トピックから50個トピックのテストを行い、  
トピックモデルへの評価をした。

# 分析方法

## 评价适度的主题数的方法

对于文章数据的困惑度 ( perplexity ) 可作为评价概率模型的性能的标准。困惑度被定义为分歧数，表示选项的数量的值以及概率的逆数。困惑度的值越低，就越能说明文章数据被以高精度分析。所以，作者认为，困惑度表示为了选择出正确的解的难度的值。

- 计算困惑度的公式：

$$perplexity(W^{test}|M) = \left( -\frac{\sum_{d=1}^{D^{test}} \log p(W_d^{test}|M)}{\sum_{d=1}^{D^{test}} N_d^{test}} \right)$$

主题模型的似然 ( Likelihood ) :

$$p(w_d|M) = \prod_{n=1}^{N_d} \sum_{k=1}^K \theta_{dk} \phi_{kw_{dn}}$$

本研究对10606篇文章进行了1个到50个的主题数实验  
从而对本模型进行了评价

# 分析結果

## □ 分析による得られた項目：

- **PPL :**

最適なトピック数を推定するため、評価する値である。

- **全文書のトピック別構成比率 :**

全ての文書におき、1つトピックの構成比率を示している。値が高ければ高いほど、文書の共通の点を示している。それに対して、値が低ければ低いほど、文書の特別な特徴を示している。

- **各トピックの中の各単語の構成比率 :**

各トピックが表している主題を判断するための重要な参考値である。

- **各文書のトピック別構成比率 :**

文書の作者を個人属性によって分けたグループごとの嗜好を分析するための重要な参考値である。

# 分析结果

## □ 通过分析得到的结果的项目：

- **PPL：**

判断是否为**最为适合的主题数**的评价值。

- **所有文章中各个主题的构成比率：**

所有文章中，各个主题的构成比率。比例越高，越可显示全部文章的**共通点**。值越低就越能说明文章组合的**特征**。

- **各个主题中的各个单词的构成比：**

判断各个主题的**主要内容**的重要参考值。

- **各个文章的各个主题的构成比率：**

通过作者的个人属性进行交叉分析时，重要的参考值。

# 分析結果

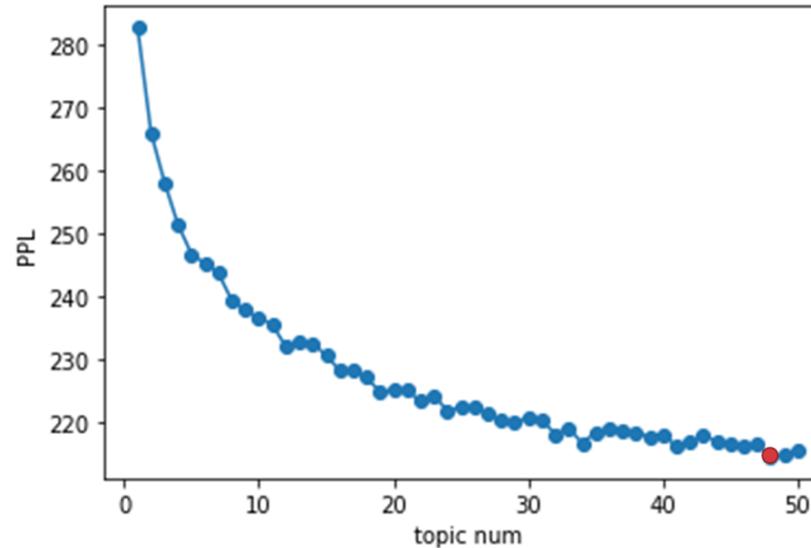
分析結果の概観：

最適なトピック数：48個

各トピックの単語の数：15個

単語の品詞と数： $n = 536$ 個（訪問地256個、旅行コンテンツ299個）

図5. トピック数別のPPL



出典：筆者より作成

トピックを1～50を設定して各々20回推定を行い、PPLの平均値を算出した。（図5）

その結果、PPLの値はトピック数の増加にともない減少し、48個のトピックの時に最小となった。

# 分析結果

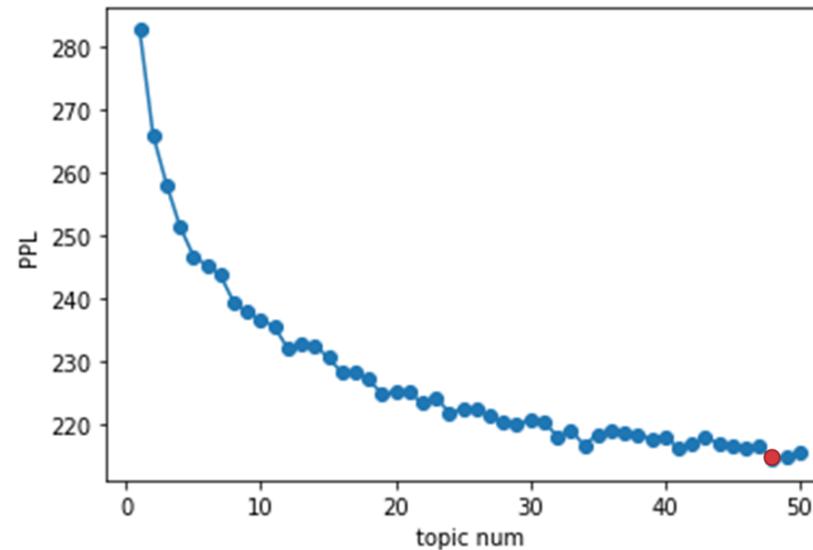
分析結果の概観：

**最适合的主题数**：48个

**各个主题的单词数**：15个

**单词的数量**：n = 536个（地名：256个、旅游内容：299个）

图5. 各个主题数のPPL



出处：由作者制作

将主题数设定为1到50，每个主题数重复分析20次，算出PPL的平均值。（图5）

结果：PPL伴随主题数的增加而减少，**48个主题**的时候PPL达到了最小。

# 分析結果

表2. 48個のトピックの抽出

トピック	トピック1	トピック2	トピック3	トピック4	トピック5	トピック6	トピック7	トピック8	トピック9	トピック10	トピック11	トピック12
各トピックの構成比率	12.00%	10.75%	8.76%	6.74%	4.43%	3.94%	3.90%	3.64%	3.59%	3.56%	3.11%	2.56%
W1構成比率	2.70%	3.30%	18.00%	1.40%	1.60%	3.20%	7.50%	7.40%	11.00%	1.90%	4.20%	1.80%
W1	感觉	日元	酒店	价格	神社	寿司	清水寺	新宿	浅草	建筑	道頓堀	证明
W2構成比率	2.60%	2.50%	3.80%	1.20%	1.40%	2.60%	5.70%	2.90%	6.30%	1.50%	3.70%	1.80%
W2	地方	机场	机场	药妆店	世界	味道	伏见稻荷大社	银座	台场	历史	大阪城公园	海贼王
W3構成比率	1.70%	2.40%	2.30%	1.10%	1.10%	2.50%	4.50%	2.60%	4.20%	1.30%	3.40%	1.40%
W3	时间	时间	时间	机场	游客	拉面	金阁寺	涩谷	东京塔	时代	奈良公园	主题乐园
W4構成比率	1.10%	2.20%	1.90%	1.10%	1.00%	2.00%	3.80%	2.00%	3.20%	1.00%	3.20%	1.20%
W4	照片	交通	行李	购物	文化	鳗鱼	二条城	明治神宫	银座	神社	通天阁	复印件
W5構成比率	1.00%	2.10%	1.70%	0.90%	1.00%	1.70%	2.60%	1.90%	3.00%	1.00%	3.10%	1.10%
W5	民宿	地铁	房间	味道	建筑	牛肉	八坂神社	原宿	上野公园	地方	市场	时间
W6構成比率	1.00%	1.40%	1.50%	0.80%	0.80%	1.70%	2.00%	1.60%	2.70%	1.00%	2.80%	1.10%
W6	小时	价格	行程	百货	历史	餐厅	春日大社	表参道	筑地市场	时间	清水寺	护照
W7構成比率	0.90%	1.30%	1.40%	0.80%	0.70%	1.60%	1.90%	1.30%	1.90%	0.90%	2.70%	1.00%
W7	行程	行程	小时	商场	地方	套餐	神社	池袋	新宿	景点	伏见稻荷大社	酒店
W8構成比率	0.90%	1.20%	1.10%	0.80%	0.70%	1.60%	1.60%	1.30%	1.80%	0.80%	2.20%	1.00%
W8	行李	景点	飞机	感觉	街道	本店	奈良公园	地铁站	明治神宫	参观	商店	旅行社
W9構成比率	0.80%	1.20%	1.10%	0.60%	0.70%	1.50%	1.60%	1.20%	1.70%	0.80%	2.10%	0.90%
W9	小伙伴	地方	地方	商店	电影	美食	东大寺	成田	皇居	地区	梅田蓝天大厦	办理
W10構成比率	0.80%	1.00%	1.10%	0.60%	0.60%	1.30%	1.50%	1.10%	1.50%	0.70%	2.10%	0.90%
W10	朋友	建议	早餐	日元	传统	地址	花见小路	地铁	地铁	公园	机场	照片
W11構成比率	0.70%	1.00%	0.80%	0.60%	0.60%	1.20%	1.40%	0.90%	1.40%	0.60%	1.80%	0.70%
W11	排队	行李	感觉	衣服	国家	口感	稻荷	筑地市场	成田	游客	地铁站	材料
W12構成比率	0.60%	0.90%	0.80%	0.60%	0.60%	1.10%	1.30%	0.90%	1.20%	0.60%	1.80%	0.60%
W12	味道	小时	晚餐	人民币	体验	河豚	鸭川	羽田机场	观音堂	庭园	楼主	机票
W13構成比率	0.60%	0.80%	0.70%	0.60%	0.50%	1.00%	1.20%	0.90%	1.20%	0.60%	1.80%	0.60%
W13	天气	机票	价格	地方	故事	日元	鸟居	歌舞伎	东京天空树	庭院	大阪城公园天守阁	游记
W14構成比率	0.60%	0.80%	0.70%	0.60%	0.50%	1.00%	1.20%	0.90%	1.10%	0.60%	1.50%	0.50%
W14	地图	民宿	机票	时间	鸟居	口味	三年坂二年坂	六本木	上野	幕府	拉面	银行
W15構成比率	0.50%	0.80%	0.60%	0.50%	0.50%	0.90%	1.00%	0.80%	1.10%	0.60%	1.40%	0.50%
W15	计划	住宿	车站	拉面	场景	食材	建筑	东京都	地铁站	文化	四天王寺	行程

# 分析结果

表2. 48个主题的抽出结果

主题	主题1	主题2	主题3	主题4	主题5	主题6	主题7	主题8	主题9	主题10	主题11	主题12
各个主题的构成比率	12.00%	10.75%	8.76%	6.74%	4.43%	3.94%	3.90%	3.64%	3.59%	3.56%	3.11%	2.56%
W1构成比率	2.70%	3.30%	18.00%	1.40%	1.60%	3.20%	7.50%	7.40%	11.00%	1.90%	4.20%	1.80%
W1	感觉	日元	酒店	价格	神社	寿司	清水寺	新宿	浅草	建筑	道顿堀	证明
W2构成比率	2.60%	2.50%	3.80%	1.20%	1.40%	2.60%	5.70%	2.90%	6.30%	1.50%	3.70%	1.80%
W2	地方	机场	机场	药妆店	世界	味道	伏见稻荷大社	银座	台场	历史	大阪城公园	海贼王
W3构成比率	1.70%	2.40%	2.30%	1.10%	1.10%	2.50%	4.50%	2.60%	4.20%	1.30%	3.40%	1.40%
W3	时间	时间	时间	机场	游客	拉面	金阁寺	涩谷	东京塔	时代	奈良公园	主题乐园
W4构成比率	1.10%	2.20%	1.90%	1.10%	1.00%	2.00%	3.80%	2.00%	3.20%	1.00%	3.20%	1.20%
W4	照片	交通	行李	购物	文化	鳗鱼	二条城	明治神宫	银座	神社	通天阁	复印件
W5构成比率	1.00%	2.10%	1.70%	0.90%	1.00%	1.70%	2.60%	1.90%	3.00%	1.00%	3.10%	1.10%
W5	民宿	地铁	房间	味道	建筑	牛肉	八坂神社	原宿	上野公园	地方	市场	时间
W6构成比率	1.00%	1.40%	1.50%	0.80%	0.80%	1.70%	2.00%	1.60%	2.70%	1.00%	2.80%	1.10%
W6	小时	价格	行程	百货	历史	餐厅	春日大社	表参道	筑地市场	时间	清水寺	护照
W7构成比率	0.90%	1.30%	1.40%	0.80%	0.70%	1.60%	1.90%	1.30%	1.90%	0.90%	2.70%	1.00%
W7	行程	行程	小时	商场	地方	套餐	神社	池袋	新宿	景点	伏见稻荷大社	酒店
W8构成比率	0.90%	1.20%	1.10%	0.80%	0.70%	1.60%	1.60%	1.30%	1.80%	0.80%	2.20%	1.00%
W8	行李	景点	飞机	感觉	街道	本店	奈良公园	地铁站	明治神宫	参观	商店	旅行社
W9构成比率	0.80%	1.20%	1.10%	0.60%	0.70%	1.50%	1.60%	1.20%	1.70%	0.80%	2.10%	0.90%
W9	小伙伴	地方	地方	商店	电影	美食	东大寺	成田	皇居	地区	梅田蓝天大厦	办理
W10构成比率	0.80%	1.00%	1.10%	0.60%	0.60%	1.30%	1.50%	1.10%	1.50%	0.70%	2.10%	0.90%
W10	朋友	建议	早餐	日元	传统	地址	花见小路	地铁	地铁	公园	机场	照片
W11构成比率	0.70%	1.00%	0.80%	0.60%	0.60%	1.20%	1.40%	0.90%	1.40%	0.60%	1.80%	0.70%
W11	排队	行李	感觉	衣服	国家	口感	稻荷	筑地市场	成田	游客	地铁站	材料
W12构成比率	0.60%	0.90%	0.80%	0.60%	0.60%	1.10%	1.30%	0.90%	1.20%	0.60%	1.80%	0.60%
W12	味道	小时	晚餐	人民币	体验	河豚	鸭川	羽田机场	观音堂	庭园	楼主	机票
W13构成比率	0.60%	0.80%	0.70%	0.60%	0.50%	1.00%	1.20%	0.90%	1.20%	0.60%	1.80%	0.60%
W13	天气	机票	价格	地方	故事	日元	鸟居	歌舞伎	东京天空树	庭院	大阪城公园天守阁	游记
W14构成比率	0.60%	0.80%	0.70%	0.60%	0.50%	1.00%	1.20%	0.90%	1.10%	0.60%	1.50%	0.50%
W14	地图	民宿	机票	时间	鸟居	口味	三年坂二年坂	六本木	上野	幕府	拉面	银行
W15构成比率	0.50%	0.80%	0.60%	0.50%	0.50%	0.90%	1.00%	0.80%	1.10%	0.60%	1.40%	0.50%
W15	计划	住宿	车站	拉面	场景	食材	建筑	东京都	地铁站	文化	四天王寺	行程

# 分析結果

表2. 48個のトピックの抽出

トピック	トピック13	トピック14	トピック15	トピック16	トピック17	トピック18	トピック19	トピック20	トピック21	トピック22	トピック23	トピック24
各トピックの構成比率	2.47%	2.25%	2.11%	2.00%	1.70%	1.51%	1.48%	1.36%	1.31%	1.27%	1.10%	0.95%
W1構成比率	12.30%	7.30%	13.80%	8.70%	8.70%	17.10%	7.20%	16.10%	5.00%	19.90%	5.30%	21.70%
W1	嵐山	环球	乐园	旅馆	富士山	镰仓	北海道	樱花	餐厅	名古屋	芦之湖	河口湖
W2構成比率	4.50%	7.00%	3.50%	3.20%	3.10%	10.80%	7.20%	3.00%	2.20%	5.30%	5.20%	11.00%
W2	风景区	影城	孩子	日式	忍野八海	江之島	小樽	新宿御苑	邮轮	白川乡	新宿	富士山
W3構成比率	3.00%	5.30%	2.80%	2.30%	2.50%	4.40%	6.70%	2.40%	2.20%	3.30%	4.70%	8.30%
W3	火车	项目	项目	房间	浅草	镰仓高校前站	函馆	樱花树	房间	名古屋城	大涌谷	富士
W4構成比率	2.80%	4.30%	2.50%	1.60%	2.20%	2.50%	2.30%	2.20%	2.00%	2.20%	2.70%	4.50%
W4	天龙寺	排队	小朋友	早餐	银座	长谷寺	洞爷湖	哲学之道	量子	城崎温泉一条街	殿场	新宿
W5構成比率	1.50%	3.20%	2.20%	1.30%	1.90%	1.80%	2.30%	1.80%	1.90%	2.00%	2.40%	2.00%
W5	红叶	时间	排队	酒店	酒店	鹤冈八幡宫	白色	樱花季	时间	白川乡合掌村	缆车	富士急游乐园
W6構成比率	1.40%	2.10%	1.90%	1.30%	1.90%	1.70%	2.30%	1.80%	1.50%	1.70%	2.20%	1.90%
W6	南禅寺	过山车	时间	晚餐	导游	新宿	恋人	醍醐寺	码头	热田神宫	富士山	酒店
W7構成比率	1.10%	1.80%	1.10%	1.20%	1.80%	1.60%	2.20%	1.70%	1.40%	1.60%	1.90%	1.10%
W7	东福寺	门票	儿童	泡汤	机场	高校	工厂	千鸟	甲板	飞驒高山	本站	秀峰
W8構成比率	1.00%	1.70%	1.00%	0.90%	1.80%	1.40%	2.00%	1.60%	1.40%	1.50%	1.80%	1.00%
W8	时间	园区	度假区	体验	金阁寺	灌篮高手	酒店	公园	游轮	合掌	小田	车站
W9構成比率	1.00%	1.30%	0.90%	0.80%	1.70%	1.20%	1.50%	1.60%	1.10%	1.30%	1.40%	0.80%
W9	嵯峨野	蜘蛛侠	玩具	传统	皇居	电车	地狱谷	目黑川	美金	高全市	箱根神社	八音盒
W10構成比率	1.00%	1.20%	0.80%	0.70%	1.40%	1.10%	1.50%	1.10%	1.00%	1.20%	1.30%	0.70%
W10	清水寺	体验	餐厅	旅店	大阪城公园	明月院	函馆山	白川	房卡	名铁	桃源	殿场
W11構成比率	1.00%	1.10%	0.80%	0.70%	1.30%	1.00%	1.20%	1.00%	0.90%	1.20%	1.20%	0.70%
W11	枫叶	感觉	儿子	住宿	富士山五合目	神社	小樽运河	时间	免税店	飞弹	雕刻	大巴
W12構成比率	0.90%	1.10%	0.70%	0.60%	1.20%	1.00%	1.10%	1.00%	0.90%	1.10%	1.10%	0.60%
W12	庭院	电影	世界	浴衣	公园	八幡	拉面	吉野山	皇家	岐阜县	酒店	天气
W13構成比率	0.80%	1.00%	0.70%	0.60%	1.00%	1.00%	1.00%	0.90%	0.90%	1.00%	1.10%	0.60%
W13	景点	黄人	门票	客人	时间	湘南海岸	函馆朝市	上野公园	阳台	牛肉	海盗船	缆车
W14構成比率	0.70%	1.00%	0.70%	0.60%	0.90%	0.90%	0.90%	0.90%	0.80%	1.00%	1.10%	0.50%
W14	高台寺	世界	爸爸	车站	地方	地方	机场	地方	排队	黑部	电车	音乐盒
W15構成比率	0.70%	0.90%	0.70%	0.50%	0.80%	0.80%	0.90%	0.90%	0.80%	0.90%	1.10%	0.50%
W15	春日大社	工作人员	体验	时间	成田	镰仓小町通	夜景	平野神社	收费	高山阵屋	小王子博物馆	河口湖天上山公园

# 分析结果

表2. 48个主题的抽出结果

主题	主题13	主题14	主题15	主题16	主题17	主题18	主题19	主题20	主题21	主题22	主题23	主题24
各个主题的构成比率	2.47%	2.25%	2.11%	2.00%	1.70%	1.51%	1.48%	1.36%	1.31%	1.27%	1.10%	0.95%
W1构成比率	12.30%	7.30%	13.80%	8.70%	8.70%	17.10%	7.20%	16.10%	5.00%	19.90%	5.30%	21.70%
W1	岚山	环球	乐园	旅馆	富士山	镰仓	北海道	樱花	餐厅	名古屋	芦之湖	河口湖
W2构成比率	4.50%	7.00%	3.50%	3.20%	3.10%	10.80%	7.20%	3.00%	2.20%	5.30%	5.20%	11.00%
W2	风景区	影城	孩子	日式	忍野八海	江之岛	小樽	新宿御苑	邮轮	白川乡	新宿	富士山
W3构成比率	3.00%	5.30%	2.80%	2.30%	2.50%	4.40%	6.70%	2.40%	2.20%	3.30%	4.70%	8.30%
W3	火车	项目	项目	房间	浅草	镰仓高校前站	函馆	樱花树	房间	名古屋城	大涌谷	富士
W4构成比率	2.80%	4.30%	2.50%	1.60%	2.20%	2.50%	2.30%	2.20%	2.00%	2.20%	2.70%	4.50%
W4	天龙寺	排队	小朋友	早餐	银座	长谷寺	洞爷湖	哲学之道	量子	城崎温泉一条街	殿场	新宿
W5构成比率	1.50%	3.20%	2.20%	1.30%	1.90%	1.80%	2.30%	1.80%	1.90%	2.00%	2.40%	2.00%
W5	红叶	时间	排队	酒店	酒店	鹤冈八幡宫	白色	樱花季	时间	白川乡合掌村	缆车	富士急游乐园
W6构成比率	1.40%	2.10%	1.90%	1.30%	1.90%	1.70%	2.30%	1.80%	1.50%	1.70%	2.20%	1.90%
W6	南禅寺	过山车	时间	晚餐	导游	新宿	恋人	醍醐寺	码头	热田神宫	富士山	酒店
W7构成比率	1.10%	1.80%	1.10%	1.20%	1.80%	1.60%	2.20%	1.70%	1.40%	1.60%	1.90%	1.10%
W7	东福寺	门票	儿童	泡汤	机场	高校	工厂	千鸟	甲板	飞驒高山	本站	秀峰
W8构成比率	1.00%	1.70%	1.00%	0.90%	1.80%	1.40%	2.00%	1.60%	1.40%	1.50%	1.80%	1.00%
W8	时间	园区	度假区	体验	金阁寺	灌篮高手	酒店	公园	游轮	合掌	小田	车站
W9构成比率	1.00%	1.30%	0.90%	0.80%	1.70%	1.20%	1.50%	1.60%	1.10%	1.30%	1.40%	0.80%
W9	嵯峨野	蜘蛛侠	玩具	传统	皇居	电车	地狱谷	目黑川	美金	高山市	箱根神社	八音盒
W10构成比率	1.00%	1.20%	0.80%	0.70%	1.40%	1.10%	1.50%	1.10%	1.00%	1.20%	1.30%	0.70%
W10	清水寺	体验	餐厅	旅店	大阪城公园	明月院	函馆山	白川	房卡	名铁	桃源	殿场
W11构成比率	1.00%	1.10%	0.80%	0.70%	1.30%	1.00%	1.20%	1.00%	0.90%	1.20%	1.20%	0.70%
W11	枫叶	感觉	儿子	住宿	富士山五合目	神社	小樽运河	时间	免税店	飞弹	雕刻	大巴
W12构成比率	0.90%	1.10%	0.70%	0.60%	1.20%	1.00%	1.10%	1.00%	0.90%	1.10%	1.10%	0.60%
W12	庭院	电影	世界	浴衣	公园	八幡	拉面	吉野山	皇家	岐阜县	酒店	天气
W13构成比率	0.80%	1.00%	0.70%	0.60%	1.00%	1.00%	1.00%	0.90%	0.90%	1.00%	1.10%	0.60%
W13	景点	黄人	门票	客人	时间	湘南海岸	函馆朝市	上野公园	阳台	牛肉	海盗船	缆车
W14构成比率	0.70%	1.00%	0.70%	0.60%	0.90%	0.90%	0.90%	0.90%	0.80%	1.00%	1.10%	0.50%
W14	高台寺	世界	爸爸	车站	地方	地方	机场	地方	排队	黑部	电车	音乐盒
W15构成比率	0.70%	0.90%	0.70%	0.50%	0.80%	0.80%	0.90%	0.90%	0.80%	0.90%	1.10%	0.50%
W15	春日大社	工作人员	体验	时间	成田	镰仓小町通	夜景	平野神社	收费	高山阵屋	小王子博物馆	河口湖天上山公园

# 分析結果

表2. 48個のトピックの抽出

トピック	トピック25	トピック26	トピック27	トピック28	トピック29	トピック30	トピック31	トピック32	トピック33	トピック34	トピック35	トピック36
各トピックの構成比率	0.94%	0.94%	0.86%	0.80%	0.75%	0.61%	0.58%	0.57%	0.57%	0.55%	0.55%	0.54%
W1構成比率	18.30%	2.10%	8.80%	1.20%	33.20%	4.20%	3.60%	8.00%	5.50%	5.20%	12.40%	17.10%
W1	神戸	時間	市場	茶道	日本环球影城	直島	车站	熊本	仙台	邮轮	冲绳	宇治
W2構成比率	4.90%	1.50%	8.50%	1.10%	3.50%	3.10%	2.60%	2.90%	3.00%	2.90%	5.10%	7.20%
W2	东大寺	環球	大阪海游馆	客人	伏见稻荷大社	小豆島	列车	別府	松島	長崎市	星野	姫路城
W3構成比率	3.90%	1.20%	2.10%	1.00%	3.00%	1.80%	2.30%	2.40%	2.90%	2.00%	5.10%	5.70%
W3	春日大社	機場	海游馆	豆腐	奈良公园	兼六园	铁道	地震	仙台市	由布院小镇	水族馆	平等院
W4構成比率	3.70%	1.10%	1.10%	0.90%	2.50%	1.60%	1.90%	2.40%	1.70%	1.60%	4.00%	2.10%
W4	异人馆	車站	麻麻	庭院	清水寺	金泽市	北荣町柯南小镇	布院	八幡	公主	度假村	宇治川
W5構成比率	2.90%	1.10%	1.00%	0.80%	1.60%	1.60%	1.60%	2.10%	1.50%	1.60%	3.80%	1.90%
W5	奈良公园	時候	天保山摩天轮	房间	环球	栗林公园	纪念馆	屋久島	新干线	豪斯登堡	国际	右卫
W6構成比率	2.30%	1.10%	0.80%	0.70%	1.50%	1.60%	1.50%	1.80%	1.40%	1.50%	3.30%	1.70%
W6	神戸港	公園	味道	豆子	商店	高松市	公園	熊本市	牛舌	国际	美国村	貴船神社
W7構成比率	2.10%	1.00%	0.70%	0.70%	1.50%	1.60%	1.40%	1.70%	0.90%	1.30%	3.20%	1.60%
W7	唐招提寺	地方	拉面	体验	道顿堀	丰岛	电车	櫻島	车站	太宰府	大酒店	好古园
W8構成比率	1.50%	0.90%	0.70%	0.60%	1.40%	1.50%	1.30%	1.30%	0.90%	1.30%	1.40%	1.60%
W8	牛肉	電車	木屋	咖啡	哈利波特主题游乐园	瀬户内海	琵琶湖	地狱	地方	公园	公园	神社
W9構成比率	1.40%	0.90%	0.70%	0.60%	1.40%	1.30%	1.00%	1.30%	0.70%	1.30%	1.10%	1.20%
W9	牛排	地鐵	时间	时间	影城	作品	鸟取市	部长	时间	福冈县	海豚	本店
W10構成比率	1.40%	0.70%	0.60%	0.60%	1.00%	1.20%	1.00%	1.30%	0.60%	1.00%	1.10%	1.10%
W10	夜景	行程	三千院	感觉	梅田蓝天大厦	艺术	铁路	鹿儿岛市	圆觉寺	門司	石垣島	宇治上神社
W11構成比率	1.20%	0.70%	0.50%	0.60%	0.90%	1.20%	0.90%	1.20%	0.60%	0.90%	1.00%	0.90%
W11	若草山	市場	鴨川	酒店	大阪城公园	时间	神社	角館	山形	抒情	潛水	桂离宮
W12構成比率	1.20%	0.70%	0.50%	0.60%	0.90%	0.90%	0.80%	1.20%	0.60%	0.90%	0.90%	0.80%
W12	法隆寺	酒店	小路	用餐	三年坂二年坂	近江町	火车	豪斯登堡	日元	人吉市	琉球	参观
W13構成比率	1.20%	0.60%	0.50%	0.50%	0.80%	0.90%	0.80%	1.20%	0.60%	0.90%	0.90%	0.80%
W13	建筑	日元	动物	餐厅	春日大社	金泽21世纪美术馆	金刀比罗宮	列车	丰田汽车博物馆	哥拉巴公园	商业街	宇治桥
W14構成比率	1.00%	0.60%	0.50%	0.50%	0.80%	0.90%	0.80%	1.10%	0.60%	0.90%	0.90%	0.70%
W14	六甲山	马笼宿	排队	日式	大阪城公园天守閣	国际	妖怪	十和田湖	山形市	稻佐山	万座毛	三室戸寺
W15構成比率	1.00%	0.50%	0.50%	0.40%	0.70%	0.80%	0.70%	1.10%	0.50%	0.80%	0.70%	0.70%
W15	兴福寺	景點	鲸鲨	果子	排队	高尾山	伊豆河津	奥入瀬溪	建长寺	太宰府天満宮	冲绳首里城公園	贵船

# 分析结果

表2. 48个主题的抽出结果

主题	主题25	主题26	主题27	主题28	主题29	主题30	主题31	主题32	主题33	主题34	主题35	主题36
各个主题的构成比率	0.94%	0.94%	0.86%	0.80%	0.75%	0.61%	0.58%	0.57%	0.57%	0.55%	0.55%	0.54%
W1构成比率	18.30%	2.10%	8.80%	1.20%	33.20%	4.20%	3.60%	8.00%	5.50%	5.20%	12.40%	17.10%
W1	神户	時間	市场	茶道	日本环球影城	直岛	车站	熊本	仙台	邮轮	冲绳	宇治
W2构成比率	4.90%	1.50%	8.50%	1.10%	3.50%	3.10%	2.60%	2.90%	3.00%	2.90%	5.10%	7.20%
W2	东大寺	環球	大阪海游馆	客人	伏见稻荷大社	小豆岛	列车	别府	松岛	长崎市	星野	姬路城
W3构成比率	3.90%	1.20%	2.10%	1.00%	3.00%	1.80%	2.30%	2.40%	2.90%	2.00%	5.10%	5.70%
W3	春日大社	機場	海游馆	豆腐	奈良公园	兼六园	铁道	地震	仙台市	由布院小镇	水族馆	平等院
W4构成比率	3.70%	1.10%	1.10%	0.90%	2.50%	1.60%	1.90%	2.40%	1.70%	1.60%	4.00%	2.10%
W4	异人馆	車站	麻麻	庭院	清水寺	金泽市	北荣町柯南小镇	布院	八幡	公主	度假村	宇治川
W5构成比率	2.90%	1.10%	1.00%	0.80%	1.60%	1.60%	1.60%	2.10%	1.50%	1.60%	3.80%	1.90%
W5	奈良公园	時候	天保山摩天轮	房间	环球	栗林公园	纪念馆	屋久岛	新干线	豪斯登堡	国际	右卫
W6构成比率	2.30%	1.10%	0.80%	0.70%	1.50%	1.60%	1.50%	1.80%	1.40%	1.50%	3.30%	1.70%
W6	神户港	公園	味道	豆子	商店	高松市	公园	熊本市	牛舌	国际	美国村	贵船神社
W7构成比率	2.10%	1.00%	0.70%	0.70%	1.50%	1.60%	1.40%	1.70%	0.90%	1.30%	3.20%	1.60%
W7	唐招提寺	地方	拉面	体验	道顿堀	丰岛	电车	樱岛	车站	太宰府	大酒店	好古园
W8构成比率	1.50%	0.90%	0.70%	0.60%	1.40%	1.50%	1.30%	1.30%	0.90%	1.30%	1.40%	1.60%
W8	牛肉	電車	木屋	咖啡	哈利波特主题游乐园	濑户内海	琵琶湖	地狱	地方	公园	公园	神社
W9构成比率	1.40%	0.90%	0.70%	0.60%	1.40%	1.30%	1.00%	1.30%	0.70%	1.30%	1.10%	1.20%
W9	牛排	地鐵	时间	时间	影城	作品	鸟取市	部长	时间	福冈县	海豚	本店
W10构成比率	1.40%	0.70%	0.60%	0.60%	1.00%	1.20%	1.00%	1.30%	0.60%	1.00%	1.10%	1.10%
W10	夜景	行程	三千院	感觉	梅田蓝天大厦	艺术	铁路	鹿儿岛市	圆觉寺	门司	石垣岛	宇治上神社
W11构成比率	1.20%	0.70%	0.50%	0.60%	0.90%	1.20%	0.90%	1.20%	0.60%	0.90%	1.00%	0.90%
W11	若草山	市場	鸭川	酒店	大阪城公园	时间	神社	角馆	山形	抒情	潜水	桂离宫
W12构成比率	1.20%	0.70%	0.50%	0.60%	0.90%	0.90%	0.80%	1.20%	0.60%	0.90%	0.90%	0.80%
W12	法隆寺	酒店	小路	用餐	三年坂二年坂	近江町	火车	豪斯登堡	日元	人吉市	琉球	参观
W13构成比率	1.20%	0.60%	0.50%	0.50%	0.80%	0.90%	0.80%	1.20%	0.60%	0.90%	0.90%	0.80%
W13	建筑	日元	动物	餐厅	春日大社	金泽21世纪美术馆	金刀比罗宫	列车	丰田汽车博物馆	哥拉巴公园	商业街	宇治桥
W14构成比率	1.00%	0.60%	0.50%	0.50%	0.80%	0.90%	0.80%	1.10%	0.60%	0.90%	0.90%	0.70%
W14	六甲山	马笼宿	排队	日式	大阪城公园天守阁	国际	妖怪	十和田湖	山形市	稻佐山	万座毛	三室户寺
W15构成比率	1.00%	0.50%	0.50%	0.40%	0.70%	0.80%	0.70%	1.10%	0.50%	0.80%	0.70%	0.70%
W15	兴福寺	景點	鲸鲨	果子	排队	高尾山	伊豆河津	奥入瀧溪	建长寺	太宰府天满宫	冲绳首里城公园	贵船

# 分析結果

表2. 48個のトピックの抽出

トピック	トピック37	トピック38	トピック39	トピック40	トピック41	トピック42	トピック43	トピック44	トピック45	トピック46	トピック47	トピック48
各トピックの構成比率	0.53%	0.37%	0.36%	0.35%	0.33%	0.32%	0.29%	0.19%	0.17%	0.15%	0.14%	0.09%
W1構成比率	7.10%	6.60%	2.80%	7.40%	9.50%	6.40%	32.20%	3.90%	7.10%	12.40%	5.80%	3.90%
W1	热海	日光	教堂	大会	有马温泉	宮島	博物館	静岡市	甲子	皇家	奶奶	江戸东京博物館
W2構成比率	4.40%	2.80%	1.90%	6.20%	6.00%	2.70%	16.80%	1.60%	6.20%	9.60%	2.10%	3.60%
W2	横浜市	日光市	軽井沢町	花火	高野山	兔子	藤子	神社	新大久保	游輪	老猪	国立西洋美術館
W3構成比率	2.60%	2.30%	1.70%	5.10%	2.70%	2.30%	6.60%	1.00%	5.70%	3.40%	2.00%	2.40%
W3	修善寺	日光東照宮	別墅	美瑛	六甲山	广島市	三鷹之森吉卜力美术馆	小仓城	巨蛋	水手	足利花卉公園	国立科学博物館
W4構成比率	2.60%	1.70%	1.70%	2.80%	1.70%	2.10%	5.90%	0.90%	5.40%	2.10%	1.70%	2.20%
W4	中华街	诹访湖	森林	富田农場	伊勢市	后乐园	川崎市	富士宮市	球场	忍者	振袖	太郎
W5構成比率	2.50%	1.70%	1.50%	2.40%	1.70%	2.10%	3.00%	0.80%	3.20%	1.80%	1.60%	2.20%
W5	半島	东照宮	光之教堂	烟花	缆车	严島神社	胖子	河内藤园	棒球	画村	日立海滨公园	总司
W6構成比率	2.30%	1.60%	1.50%	2.10%	1.60%	1.90%	1.20%	0.80%	2.50%	1.60%	1.20%	1.80%
W6	下田	中禅寺湖	农場	四季彩之丘	三段壁	天桥立	东京迪士尼乐园	嬉野	演唱会	外滩	富士市	新撰組
W7構成比率	2.20%	1.50%	0.90%	2.00%	1.50%	1.40%	1.20%	0.70%	2.00%	1.40%	1.20%	1.70%
W7	城崎海岸	鬼怒川温泉	东川町	浴衣	新宮	鸟居	漫画	佐贺市	鞍馬山	筆者	艺伎	土方
W8構成比率	1.90%	1.40%	0.70%	1.80%	1.30%	1.20%	1.20%	0.60%	1.60%	1.20%	1.10%	1.50%
W8	清水	日光江戸村	萤火虫	车站	熊野古道	广島县	机器猫	屋城	体育场	客房	报国寺	御岳升仙峡
W9構成比率	1.90%	1.30%	0.70%	1.30%	1.30%	1.20%	0.90%	0.60%	1.50%	1.00%	1.10%	1.50%
W9	热海市	效果	瀑布	北海道	伊勢神宮	尾道市	门票	时间	图说	小袖	次郎	壬生寺
W10構成比率	1.80%	1.00%	0.70%	1.30%	1.20%	1.20%	0.80%	0.60%	1.40%	0.60%	1.00%	1.30%
W10	伊豆市	二荒山神社	牧場	美瑛町	六甲山牧场	仓敷美观	登户	伊豆泰迪熊博物馆	神田明神社	图片	法师	甲府市
W11構成比率	1.70%	1.00%	0.70%	1.30%	1.10%	1.10%	0.80%	0.60%	1.30%	0.50%	1.00%	1.30%
W11	公園	面膜	小屋	烟火	和歌山县	姫路城	地狱谷野猿公苑	千光寺	东寺	大馆	舞妓	冰川丸
W12構成比率	1.60%	0.90%	0.60%	1.30%	1.10%	1.10%	0.80%	0.60%	1.20%	0.50%	0.90%	1.20%
W12	三島市	肌肤	会津若松市	自行车	奥之院	牡蛎	主题	狂欢节	生驹	体验	静酱	近藤
W13構成比率	1.50%	0.90%	0.60%	1.00%	1.10%	1.10%	0.60%	0.50%	0.80%	0.50%	0.90%	1.10%
W13	静岡县	皮肤	地址	会场	和歌山市	弥山	尘表阁	大谷	日币	大道	摄影师	谷根千
W14構成比率	1.30%	0.90%	0.60%	0.90%	1.00%	1.00%	0.60%	0.50%	0.80%	0.40%	0.70%	1.00%
W14	大室山	产品	季节	富良野花田	和歌山城	广島城	作品	交通	東京都立小石川后乐园	文字	婚纱	前川
W15構成比率	1.20%	0.80%	0.60%	0.90%	0.80%	1.00%	0.40%	0.50%	0.80%	0.40%	0.70%	0.80%
W15	櫻桃小丸子主题博物馆	松本市	石之教堂	民宿	日币	野島	胖虎	嬉野温泉	球迷	价格	小野	冲田

# 分析结果

表2. 48个主题的抽出结果

主题	主题37	主题38	主题39	主题40	主题41	主题42	主题43	主题44	主题45	主题46	主题47	主题48
各个主题的构成比率	0.53%	0.37%	0.36%	0.35%	0.33%	0.32%	0.29%	0.19%	0.17%	0.15%	0.14%	0.09%
W1构成比率	7.10%	6.60%	2.80%	7.40%	9.50%	6.40%	32.20%	3.90%	7.10%	12.40%	5.80%	3.90%
W1	热海	日光	教堂	大会	有马温泉	宫岛	博物馆	静冈市	甲子	皇家	奶奶	江户东京博物馆
W2构成比率	4.40%	2.80%	1.90%	6.20%	6.00%	2.70%	16.80%	1.60%	6.20%	9.60%	2.10%	3.60%
W2	横滨市	日光市	轻井泽町	花火	高野山	兔子	藤子	神社	新大久保	游轮	老猪	国立西洋美术馆
W3构成比率	2.60%	2.30%	1.70%	5.10%	2.70%	2.30%	6.60%	1.00%	5.70%	3.40%	2.00%	2.40%
W3	修善寺	日光东照宫	别墅	美瑛	六甲山	广岛市	三鹰之森吉卜力美术馆	小仓城	巨蛋	水手	足利花卉公园	国立科学博物馆
W4构成比率	2.60%	1.70%	1.70%	2.80%	1.70%	2.10%	5.90%	0.90%	5.40%	2.10%	1.70%	2.20%
W4	中华街	诹访湖	森林	富田农场	伊势市	后乐园	川崎市	富士宫市	球场	忍者	振袖	太郎
W5构成比率	2.50%	1.70%	1.50%	2.40%	1.70%	2.10%	3.00%	0.80%	3.20%	1.80%	1.60%	2.20%
W5	半岛	东照宫	光之教堂	烟花	缆车	严岛神社	胖子	河内藤园	棒球	画村	日立海滨公园	总司
W6构成比率	2.30%	1.60%	1.50%	2.10%	1.60%	1.90%	1.20%	0.80%	2.50%	1.60%	1.20%	1.80%
W6	下田	中禅寺湖	农场	四季彩之丘	三段壁	天桥立	东京迪斯尼乐园	嬉野	演唱会	外滩	富士市	新撰组
W7构成比率	2.20%	1.50%	0.90%	2.00%	1.50%	1.40%	1.20%	0.70%	2.00%	1.40%	1.20%	1.70%
W7	城崎海岸	鬼怒川温泉	东川町	浴衣	新宫	鸟居	漫画	佐贺市	鞍马山	笔者	艺伎	土方
W8构成比率	1.90%	1.40%	0.70%	1.80%	1.30%	1.20%	1.20%	0.60%	1.60%	1.20%	1.10%	1.50%
W8	清水	日光江户村	萤火虫	车站	熊野古道	广岛县	机器猫	屋城	体育场	客房	报国寺	御岳升仙峡
W9构成比率	1.90%	1.30%	0.70%	1.30%	1.30%	1.20%	0.90%	0.60%	1.50%	1.00%	1.10%	1.50%
W9	热海市	效果	瀑布	北海道	伊势神宫	尾道市	门票	时间	图说	小袖	次郎	王生寺
W10构成比率	1.80%	1.00%	0.70%	1.30%	1.20%	1.20%	0.80%	0.60%	1.40%	0.60%	1.00%	1.30%
W10	伊豆市	二荒山神社	牧场	美瑛町	六甲山牧场	仓敷美观	登户	伊豆泰迪熊博物馆	神田明神神社	图片	法师	甲府市
W11构成比率	1.70%	1.00%	0.70%	1.30%	1.10%	1.10%	0.80%	0.60%	1.30%	0.50%	1.00%	1.30%
W11	公园	面膜	小屋	烟火	和歌山县	姬路城	地狱谷野猿公苑	千光寺	东寺	大馆	舞妓	冰川丸
W12构成比率	1.60%	0.90%	0.60%	1.30%	1.10%	1.10%	0.80%	0.60%	1.20%	0.50%	0.90%	1.20%
W12	三岛市	肌肤	会津若松市	自行车	奥之院	牡蛎	主题	狂欢节	生驹	体验	静酱	近藤
W13构成比率	1.50%	0.90%	0.60%	1.00%	1.10%	1.10%	0.60%	0.50%	0.80%	0.50%	0.90%	1.10%
W13	静冈县	皮肤	地址	会场	和歌山市	弥山	尘表阁	大谷	日币	大道	摄影师	谷根千
W14构成比率	1.30%	0.90%	0.60%	0.90%	1.00%	1.00%	0.60%	0.50%	0.80%	0.40%	0.70%	1.00%
W14	大室山	产品	季节	富良野花田	和歌山城	广岛城	作品	交通	东京都立小石川后乐园	文字	婚纱	前川
W15构成比率	1.20%	0.80%	0.60%	0.90%	0.80%	1.00%	0.40%	0.50%	0.80%	0.40%	0.70%	0.80%
W15	樱桃小丸子主题博物馆	松本市	石之教堂	民宿	日币	野岛	胖虎	嬉野温泉	球迷	价格	小野	冲田

# 分析結果

表3. トピック別の場所、コンテンツ、構成比率

トピック	訪問地	コンテンツ	構成比率	トピック	特定都市	総合的な観光	構成比率
トピック1	-	日本の日常生活体験	12.24%	トピック25	特定都市	総合的な観光	0.96%
トピック2	-	滞在する際の交通手段	10.97%	トピック26	-	滞在する際の交通手段	0.96%
トピック3	-	滞在する際の交通手段	8.93%	トピック27	特定都市	テーマパーク	0.87%
トピック4	-	購買行動	6.87%	トピック28	-	グルメ	0.82%
トピック5	-	歴史文化	4.52%	トピック29	ゴールデンルート	総合的な観光	0.77%
トピック6	-	グルメ	4.01%	トピック30	地方	総合的な観光	0.62%
トピック7	ゴールデンルート	歴史文化	3.97%	トピック31	地方	アニメ	0.59%
トピック8	特定都市	大都市遊覧	3.72%	トピック32	九州	総合的な観光	0.58%
トピック9	ゴールデンルート	総合的な観光	3.66%	トピック33	地方	総合的な観光	0.58%
トピック10	-	歴史文化	3.63%	トピック34	地方	クルーズ	0.56%
トピック11	特定都市	総合的な観光	3.18%	トピック35	沖縄	リゾート	0.56%
トピック12	-	旅行する手続きに関する内容	2.61%	トピック36	特定都市	歴史文化	0.55%
トピック13	特定都市	季節風景	2.52%	トピック37	地方	総合的な観光	0.54%
トピック14	特定都市	テーマパーク	2.30%	トピック38	地方	総合的な観光	0.38%
トピック15	-	親子旅行	2.15%	トピック39	地方	総合的な観光	0.36%
トピック16	-	温泉旅館	2.04%	トピック40	北海道	イベント参加	0.36%
トピック17	ゴールデンルート	総合的な遊覧	1.74%	トピック41	地方	総合的な観光	0.33%
トピック18	地方	アニメ	1.54%	トピック42	地方	総合的な観光	0.33%
トピック19	北海道	総合的な遊覧	1.51%	トピック43	地方	アニメ	0.29%
トピック20	特定都市	季節風景	1.38%	トピック44	地方	総合的な観光	0.19%
トピック21	-	クルーズ	1.34%	トピック45	-	スポーツ観戦	0.17%
トピック22	地方	総合的な観光	1.30%	トピック46	-	クルーズ	0.16%
トピック23	特定都市	総合的な観光	1.12%	トピック47	地方	季節風景	0.15%
トピック24	特定都市	総合的な観光	0.97%	トピック48	地方	歴史文化	0.09%
合計							100.00%

# 分析结果

表3. 各个主题的访问地，内容及构成比率

主题	访问地	内容	构成比率
主题1	-	日本的日常生活体验	12.24%
主题2	-	日本滞在时的交通手段	10.97%
主题3	-	日本滞在时的交通手段	8.93%
主题4	-	购物行动	6.87%
主题5	-	历史文化	4.52%
主题6	-	美食	4.01%
主题7	黄金线路	历史文化	3.97%
主题8	特定城市	大都市游览	3.72%
主题9	黄金线路	综合性游览	3.66%
主题10	-	历史文化	3.63%
主题11	特定城市	综合性游览	3.18%
主题12	-	旅游相关手续	2.61%
主题13	特定城市	季节风景	2.52%
主题14	特定城市	主题乐园	2.30%
主题15	-	亲子游	2.15%
主题16	-	温泉旅馆	2.04%
主题17	黄金线路	综合性游览	1.74%
主题18	地方	动漫相关内容	1.54%
主题19	北海道	综合性游览	1.51%
主题20	特定城市	季节风景	1.38%
主题21	-	邮轮	1.34%
主题22	地方	综合性游览	1.30%
主题23	特定城市	综合性游览	1.12%
主题24	特定城市	综合性游览	0.97%

主题25	特定城市	综合性游览	0.96%
主题26	-	日本滞在时的交通手段	0.96%
主题27	特定城市	主题乐园	0.87%
主题28	-	美食	0.82%
主题29	黄金线路	综合性游览	0.77%
主题30	地方	综合性游览	0.62%
主题31	地方	动漫相关内容	0.59%
主题32	九州	综合性游览	0.58%
主题33	地方	综合性游览	0.58%
主题34	地方	邮轮	0.56%
主题35	冲绳	度假	0.56%
主题36	特定城市	历史文化	0.55%
主题37	地方	综合性游览	0.54%
主题38	地方	综合性游览	0.38%
主题39	地方	综合性游览	0.36%
主题40	北海道	活动参加	0.36%
主题41	地方	综合性游览	0.33%
主题42	地方	综合性游览	0.33%
主题43	地方	动漫相关内容	0.29%
主题44	地方	综合性游览	0.19%
主题45	-	赛事观战，演唱会参加	0.17%
主题46	-	邮轮	0.16%
主题47	地方	季节风景	0.15%
主题48	地方	历史文化	0.09%
合计			100.00%

# 分析結果

表4. 個人属性別の構成比率と項目内の平均値の差

項目	代表トピック	性別		住所							
		男性	女性	北京	上海	東北	華北	華中	華南	西部	
訪問場所	ゴールデンルート	7,9,17,29	-0.18%	0.18%	-0.38%	-1.98%	1.15%	0.76%	-0.76%	-0.44%	1.66%
	特定都市	8,11,13,14,20,23,24,25,27,36	-0.14%	0.14%	1.27%	0.51%	-2.04%	-0.36%	0.09%	1.18%	-0.65%
	北海道	19,40	-0.10%	0.10%	0.18%	0.03%	-0.45%	-0.09%	-0.20%	0.65%	-0.11%
	沖縄	35	-0.04%	0.04%	0.11%	0.10%	-0.13%	0.35%	-0.05%	-0.20%	-0.19%
	九州	32	-0.01%	0.01%	0.12%	0.46%	0.02%	-0.22%	-0.09%	0.00%	-0.29%
	地方部	18,22,30,31,33,34,37,38,39,41,42,43,44,47,48	-0.16%	0.16%	0.18%	2.65%	0.68%	-0.82%	-0.91%	-0.99%	-0.80%
コンテンツ	歴史文化	5,7,10,36,48	-0.76%	0.76%	0.27%	-0.81%	-0.12%	-0.47%	-1.12%	0.44%	1.80%
	総合的遊覧	9,11,17,19,22,23,24,25,29,30,32,33,37,38,39,41,42,44	-0.63%	0.63%	0.00%	1.37%	1.52%	0.19%	-1.47%	-0.64%	-0.95%
	大都市遊覧	8	0.12%	-0.12%	0.73%	-0.08%	0.14%	-0.88%	-0.41%	0.30%	0.20%
	購買行動	4	0.40%	-0.40%	-0.42%	-1.33%	1.54%	2.12%	-0.21%	-0.99%	-0.70%
	グルメ	6,28	0.09%	-0.09%	0.60%	1.03%	-0.44%	-0.33%	0.03%	0.14%	-1.02%
	季節風景	13,20,47	-0.02%	0.02%	-0.05%	0.11%	-0.78%	-0.36%	-0.38%	1.09%	0.36%
	テーマパーク	14,27	0.01%	-0.01%	0.66%	-0.32%	0.04%	0.74%	0.27%	-0.93%	-0.45%
	アニメ	18,31,43	-0.01%	0.01%	0.25%	0.00%	0.18%	-0.17%	-0.43%	0.02%	0.16%
	親子旅行	15	0.15%	-0.15%	0.75%	-0.09%	1.10%	0.26%	-0.24%	-1.03%	-0.73%
	クルーズ	21,34,46	-0.02%	0.02%	-0.56%	0.35%	0.50%	0.51%	0.65%	-0.66%	-0.79%
	温泉旅館	16	0.03%	-0.03%	0.05%	0.43%	-0.56%	-0.55%	0.04%	0.71%	-0.13%
	リゾート	35	-0.04%	0.04%	0.11%	0.10%	-0.13%	0.35%	-0.05%	-0.20%	-0.19%
	イベント参加	40	0.03%	-0.03%	0.02%	0.15%	-0.10%	-0.10%	0.01%	-0.06%	0.09%
	スポーツ、コンサート	45	0.00%	0.00%	0.12%	-0.02%	-0.10%	-0.11%	-0.05%	0.12%	0.04%

# 分析结果

表4. 按个人属性分的各个项目内的比率与平均值的差

项目	代表的主题	性别		居住地							
		男性	女性	北京	上海	东北	华北	华中	华南	西部	
访问地点	黄金线路	7,9,17,29	-0.18%	0.18%	-0.38%	-1.98%	1.15%	0.76%	-0.76%	-0.44%	1.66%
	特定城市	8,11,13,14,20,23,24,25,27,36	-0.14%	0.14%	1.27%	0.51%	-2.04%	-0.36%	0.09%	1.18%	-0.65%
	北海道	19,40	-0.10%	0.10%	0.18%	0.03%	-0.45%	-0.09%	-0.20%	0.65%	-0.11%
	冲绳	35	-0.04%	0.04%	0.11%	0.10%	-0.13%	0.35%	-0.05%	-0.20%	-0.19%
	九州	32	-0.01%	0.01%	0.12%	0.46%	0.02%	-0.22%	-0.09%	0.00%	-0.29%
	地方	18,22,30,31,33,34,37,38,39,41,42,43,44,47,48	-0.16%	0.16%	0.18%	2.65%	0.68%	-0.82%	-0.91%	-0.99%	-0.80%
内容	历史文化	5,7,10,36,48	-0.76%	0.76%	0.27%	-0.81%	-0.12%	-0.47%	-1.12%	0.44%	1.80%
	综合性游览	9,11,17,19,22,23,24,25,29,30,32,33,37,38,39,41,42,44	-0.63%	0.63%	0.00%	1.37%	1.52%	0.19%	-1.47%	-0.64%	-0.95%
	大都市游览	8	0.12%	-0.12%	0.73%	-0.08%	0.14%	-0.88%	-0.41%	0.30%	0.20%
	购物行动	4	0.40%	-0.40%	-0.42%	-1.33%	1.54%	2.12%	-0.21%	-0.99%	-0.70%
	美食	6,28	0.09%	-0.09%	0.60%	1.03%	-0.44%	-0.33%	0.03%	0.14%	-1.02%
	季节风景	13,20,47	-0.02%	0.02%	-0.05%	0.11%	-0.78%	-0.36%	-0.38%	1.09%	0.36%
	主题乐园	14,27	0.01%	-0.01%	0.66%	-0.32%	0.04%	0.74%	0.27%	-0.93%	-0.45%
	动漫相关内容	18,31,43	-0.01%	0.01%	0.25%	0.00%	0.18%	-0.17%	-0.43%	0.02%	0.16%
	亲子游	15	0.15%	-0.15%	0.75%	-0.09%	1.10%	0.26%	-0.24%	-1.03%	-0.73%
	邮轮	21,34,46	-0.02%	0.02%	-0.56%	0.35%	0.50%	0.51%	0.65%	-0.66%	-0.79%
	温泉旅馆	16	0.03%	-0.03%	0.05%	0.43%	-0.56%	-0.55%	0.04%	0.71%	-0.13%
	度假	35	-0.04%	0.04%	0.11%	0.10%	-0.13%	0.35%	-0.05%	-0.20%	-0.19%
	活动参加	40	0.03%	-0.03%	0.02%	0.15%	-0.10%	-0.10%	0.01%	-0.06%	0.09%
	赛事观战，演唱会参加	45	0.00%	0.00%	0.12%	-0.02%	-0.10%	-0.11%	-0.05%	0.12%	0.04%

# 分析結果

表5. 個人属性別の構成比率と項目内の平均値の差

項目	代表トピック	訪問季節				同行者					
		春	夏	秋	冬	一人	恋人、夫婦	親子	友達	家族	
訪問場所	ゴールデンルート	7,9,17,29	0.58%	0.13%	0.15%	-0.87%	-1.19%	0.65%	-1.56%	0.37%	1.73%
	特定都市	8,11,13,14,20,23,24,25,27,36	2.11%	-1.89%	1.90%	-2.12%	-0.69%	2.64%	-3.43%	1.11%	0.36%
	北海道	19,40	-1.30%	0.51%	-0.89%	1.68%	0.37%	0.15%	-0.42%	-0.07%	-0.03%
	沖縄	35	-0.16%	0.08%	0.02%	0.07%	-0.28%	-0.07%	0.59%	-0.22%	-0.02%
	九州	32	-0.05%	-0.13%	0.10%	0.09%	0.33%	-0.12%	0.04%	-0.12%	-0.13%
	地方部	18,22,30,31,33,34,37,38,39,41,42,43,44,47,48	-0.64%	0.57%	-0.09%	0.16%	4.36%	-1.12%	-1.89%	-0.25%	-1.10%
コンテンツ	歴史文化	5,7,10,36,48	-0.28%	0.77%	0.64%	-1.13%	3.17%	-0.84%	-3.30%	-0.03%	1.01%
	総合的遊覧	9,11,17,19,22,23,24,25,29,30,32,33,37,38,39,41,42,44	-0.96%	-1.67%	-0.03%	2.65%	1.72%	0.29%	-2.70%	-0.21%	0.91%
	大都市遊覧	8	-0.11%	0.06%	-0.04%	0.09%	1.12%	0.39%	-1.37%	0.30%	-0.44%
	購買行動	4	0.19%	-0.23%	-0.30%	0.34%	-1.19%	0.49%	-0.71%	0.35%	1.07%
	グルメ	6,28	0.67%	-0.68%	-0.21%	0.22%	0.42%	0.82%	-1.70%	0.60%	-0.14%
	季節風景	13,20,47	2.12%	-1.60%	1.58%	-2.09%	0.46%	0.87%	-1.53%	0.13%	0.07%
	テーマパーク	14,27	-0.08%	0.17%	0.13%	-0.22%	-1.22%	0.77%	0.54%	0.37%	-0.47%
	アニメ	18,31,43	-0.02%	0.62%	-0.27%	-0.33%	1.61%	-0.18%	-0.61%	0.10%	-0.92%
	親子旅行	15	-0.17%	0.56%	-0.18%	-0.21%	-2.54%	-1.92%	7.74%	-1.95%	-1.34%
	クルーズ	21,34,46	-0.26%	1.35%	-0.46%	-0.63%	-1.25%	-1.06%	0.67%	-0.60%	2.24%
	温泉旅館	16	-0.32%	-0.60%	-0.04%	0.95%	-0.02%	0.10%	0.29%	-0.36%	-0.01%
	リゾート	35	-0.16%	0.08%	0.02%	0.07%	-0.28%	-0.07%	0.59%	-0.22%	-0.02%
	イベント参加	40	-0.38%	0.94%	-0.25%	-0.31%	0.42%	-0.06%	-0.22%	0.02%	-0.15%
	スポーツ、コンサート	45	-0.03%	0.14%	-0.02%	-0.09%	0.31%	-0.14%	-0.11%	0.01%	-0.08%

# 分析结果

表5. 按个人属性分的各个项目内的比率与平均值的差

项目	代表的主题	访问季节				同行人				
		春	夏	秋	冬	单人	小两口	带孩子	朋友	家庭
访问地点	黄金线路	7,9,17,29	0.58%	0.13%	0.15%	-0.87%	-1.19%	0.65%	-1.56%	0.37%
	特定城市	8,11,13,14,20,23,24,25,27,36	2.11%	-1.89%	1.90%	-2.12%	-0.69%	2.64%	-3.43%	1.11%
	北海道	19,40	-1.30%	0.51%	-0.89%	1.68%	0.37%	0.15%	-0.42%	-0.07%
	冲绳	35	-0.16%	0.08%	0.02%	0.07%	-0.28%	-0.07%	0.59%	-0.22%
	九州	32	-0.05%	-0.13%	0.10%	0.09%	0.33%	-0.12%	0.04%	-0.12%
内容	地方	18,22,30,31,33,34,37,38,39,41,42,43,44,47,48	-0.64%	0.57%	-0.09%	0.16%	4.36%	-1.12%	-1.89%	-0.25%
	历史文化	5,7,10,36,48	-0.28%	0.77%	0.64%	-1.13%	3.17%	-0.84%	-3.30%	-0.03%
	综合性游览	9,11,17,19,22,23,24,25,29,30,32,33,37,38,39,41,42,44	-0.96%	-1.67%	-0.03%	2.65%	1.72%	0.29%	-2.70%	-0.21%
	大都市游览	8	-0.11%	0.06%	-0.04%	0.09%	1.12%	0.39%	-1.37%	0.30%
	购物行动	4	0.19%	-0.23%	-0.30%	0.34%	-1.19%	0.49%	-0.71%	0.35%
	美食	6,28	0.67%	-0.68%	-0.21%	0.22%	0.42%	0.82%	-1.70%	0.60%
	季节风景	13,20,47	2.12%	-1.60%	1.58%	-2.09%	0.46%	0.87%	-1.53%	0.13%
	主题乐园	14,27	-0.08%	0.17%	0.13%	-0.22%	-1.22%	0.77%	0.54%	0.37%
	动漫相关内容	18,31,43	-0.02%	0.62%	-0.27%	-0.33%	1.61%	-0.18%	-0.61%	0.10%
	亲子游	15	-0.17%	0.56%	-0.18%	-0.21%	-2.54%	-1.92%	7.74%	-1.95%
	邮轮	21,34,46	-0.26%	1.35%	-0.46%	-0.63%	-1.25%	-1.06%	0.67%	-0.60%
	温泉旅馆	16	-0.32%	-0.60%	-0.04%	0.95%	-0.02%	0.10%	0.29%	-0.36%
	度假	35	-0.16%	0.08%	0.02%	0.07%	-0.28%	-0.07%	0.59%	-0.22%
	活动参加	40	-0.38%	0.94%	-0.25%	-0.31%	0.42%	-0.06%	-0.22%	0.02%
	赛事观战，演唱会参加	45	-0.03%	0.14%	-0.02%	-0.09%	0.31%	-0.14%	-0.11%	0.01%

# 分析結果

## クロス集計の考察

### 性別とトピックのクロス集計

男性の方は購買行動や大都市遊覧などの従来の旅行行動の傾向がある。それに対して、女性の方はある1つの地域に深く体験する嗜好が分かった。

### 居住地とトピックのクロス集計

沿岸部に居住している訪日中国人旅行者は体験するなどの旅行嗜好があり、一方、内陸に住んでいる方は購買行動やゴールデンルートにの総合的遊覧の訪問の傾向があることが分かった。また、沿岸部の方がクルーズのコンテンツが分かった。

### 季節とトピックのクロス集計

従来の訪日中国人旅行行動に対する認識と異なり、ゴールデンルートに関するトピックは少ない。代わりに、ゴールデンルートの中の1つの場所に深く体験や地方への体験などの傾向が明らかにした。特に、「春の時期に神奈川県におけるアニメーリズム及び花見」、「夏の時期に群馬県と栃木県における花見」のような新たな訪日旅行の嗜好が辿られた。

### 同行者とトピックのクロス集計

同行者人数の増加につれ、訪問地とコンテンツの主題はゴールデンルートの特徴に傾いていることが分かった。また、特に家族旅行のグループは、クルーズのコンテンツを持っている構成比率が他のと比べ、非常に高い。

# 分析结果

## 交叉分析的考察

### 性别与主题的交叉分析

男性偏向于购物，都市中游览等**传统的旅行行动**。

女性更喜欢在一个地方**深度旅游**。

### 居住地与主题的交叉分析

居住在沿海地区的中国人更注重于**在日体验旅游**，另一方面，生活在内陆的中国人对于**购物，黄金线路上的传统旅游**的倾向更多。值得注意的是，沿海地区居住的中国人，参加**邮轮旅游**的比例更高。

### 季节与主题的交叉分析

与通常的对于中国人访日行动的认识不同，黄金线路相关的主题书比较少。取而代之的是换金线路中其中一地的深度游以及日本地方的体验游。特别是「春天在神奈川县的动漫相关旅游以及赏花」「夏天在群马县和栃木县的赏花」的新的访日行动。

### 同行人与主题的交叉分析

随着同伴的**人数的增加**，访问地和内容的主题也向传统的**黄金线路**倾向。

另外，**家族出游**的群体，参加**邮轮游**的比率与其它群体相比异常的高。

# 分析結果

## 旅行実態の例1：京都府における異なる旅行形式

### • 962番旅行記の旅程：

一日目：西本願寺 – 三年坂 – 清水寺 – 八坂神社

二日目：金閣寺 – 二条城 – 京都御所

三日目：東本願寺 – 京都国立博物館

西本願寺は浄土真宗本願寺派の本山で、1994年に世界文化遺産に登録されました。実は金閣寺の本名が鹿苑寺です。建築物の外は金箔で包まれていますから、金閣寺で呼ばれています。

区域内に豊かな観光スポットは宝物です。思い切りに歴史の海洋の中で楽しめる。

### • 194番旅行記の旅程：

嵐山 – 天龍寺 – 永觀堂 – 南禅寺 – 清水寺

この一生の中に見るべき美しさで、心酔させる京都の紅葉です。

紅葉の美しさは人に呼吸できなくなる、動かない程です。物凄く美しいです。

### • 1946番旅行記の旅程：

中村藤吉（本店） – 宇治川 – 宇治橋 – 平等院表参道 – 平等院

町並みには本格的な日本式な家があって、宇治の土産である抹茶もあります。ここで最も本格的な抹茶ライスやお菓子などを食べられる。また、様々な抹茶料理のレストランも多いです。

伊藤久右衛門の本店へ、食べます！

宇治の抹茶、美味しいすぎる！食べ甲斐があります。



出典：筆者より作成

同じ京都府において、異なる3つの旅行形式があることが分かった。

# 分析结果

## 旅行形态的例1：位于京都府的不同的旅行形式

### • 962号游记的旅程

第一天：西本愿寺-三年坂-清水寺-八坂神社

第二天：金阁寺-二条城-京都御所

第三天：东本愿寺-京都国立博物馆

西本愿寺是净土真宗本愿寺派的大本山。1994年列为联合国世界文化遗产。

金阁寺正式名称其实是鹿苑寺，因为建筑物外面包有金箔，故又名金阁寺。

区域内丰富的景点和遗迹可以说是无尽的宝藏，可以尽情的在历史的海洋里探索。

### • 194号游记的旅程：

嵐山-天龙寺-永观堂-南禅寺-清水寺

一生必去欣赏的美，醉人于心的京都红叶

红叶美的让人窒息，美的让你挪不动步子，一步一景，美轮美奂

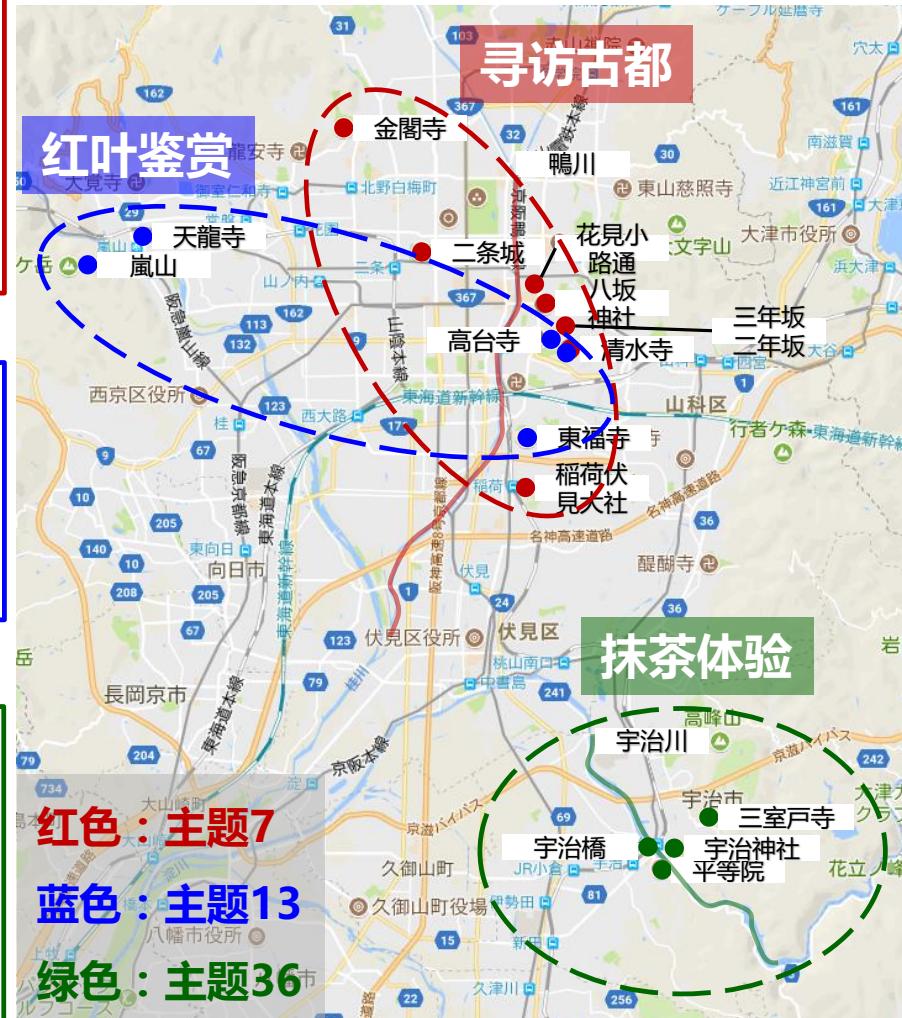
### • 1946号游记的旅程：

中村藤吉（本店）品尝抹茶-宇治川-宇治桥-平等院表参道-平等院

平等院表参道是一条日本文化气息浓厚的商业街，街道两旁是纯正的日式小院，当然少不了宇治的特产-抹茶，在这里能体验最正宗的抹茶冰激凌与和果子，还有以抹茶为创意的各式料理店。

前往伊藤久右卫门本店。吃！

宇治的抹茶，好吃！值得品味哦



出处：由作者制作

同样是京都府，却有着3种完全不同的旅行形态

# 分析結果

## 旅行実態の例2：大都市と地方の組み合わせ

### • 555番旅行記の旅程：

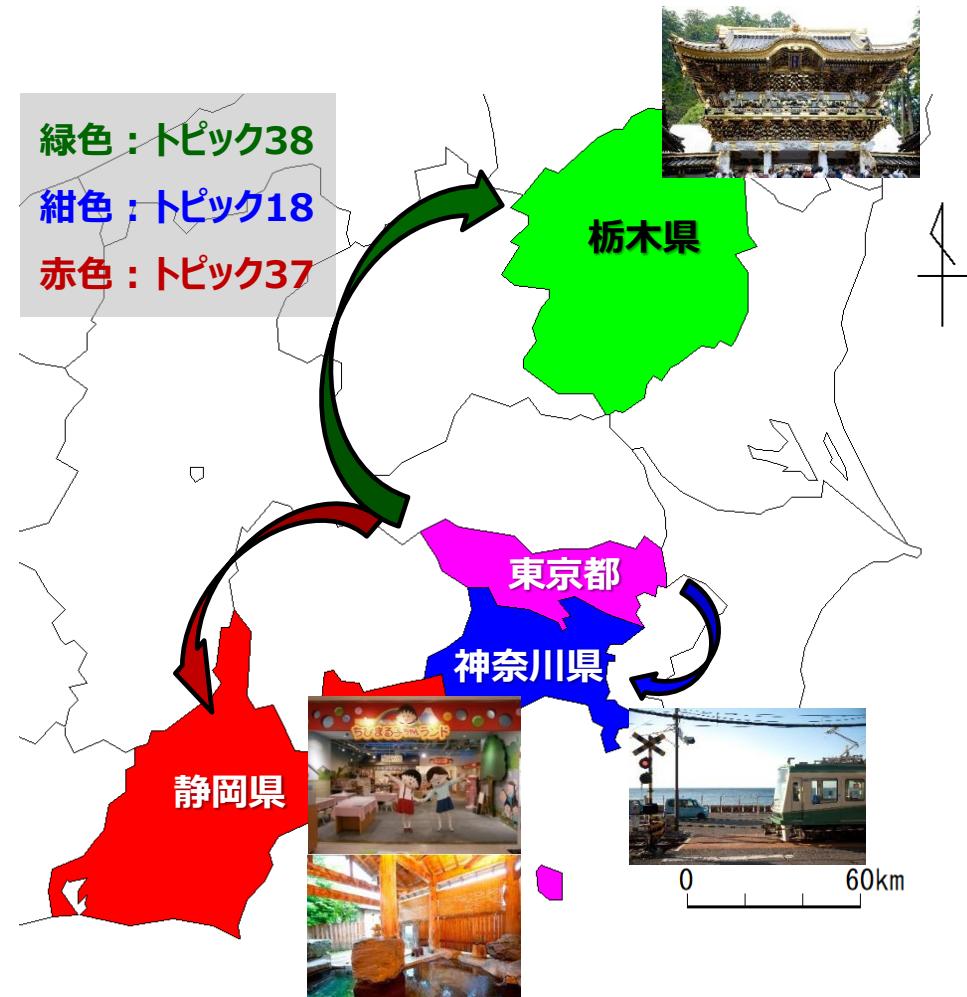
東京：浅草 – 東京タワー – 浅草寺 – 歌舞伎町 – 浅草寺商店街  
栃木県：宇都宮 – 日光東照宮 – 中禅寺湖 – 江戸村 – 鬼怒川  
鬼怒川は温泉郷です。この川の両側に多くの温泉旅館が並んでいます。  
徳川家康が天下を統一して、徳川幕府を建立し、家康が死んだあと江戸幕府の守護新（東照神君）として東照宮に供奉されています。  
今回の日光旅は繁忙期だけど、パーフェクトです。皆は全部早起きして、旅行者が多すぎる状況を避けました。次回、閑散期の時に必ず日光に戻って、もう一度日光を詳しく巡ります！

### • 864番旅行記の旅程：

東京：新宿 – 渋谷 – 銀座 – 六本木 – 代官山 – 東京ディズニーシー – サンリオピューロランド  
神奈川県：鎌倉市 – 鎌倉高校前 – 鎌倉高校前の踏切 – 鎌倉大仏 – 鶴岡八幡宮 – 湘南海岸  
スラムダンク！ そうだ！ スラムダンクの故郷へ見に行きます！  
この海岸（湘南海岸）は漫画の最も重要なロケ地です。今回の必ず行くところです。

### • 9444番旅行記の旅程：

東京 – 東京タワー – 熱海 – 伊豆 – 東伊豆（温泉）  
旅館が自分で作った日本酒があります。旅行を予約するとこのような日本酒も予約できます。  
この旅館を泊まると、貸切湯を入浴できます。唯一の写真を撮れる温泉です。  
• 1039番旅行記の旅程：  
東京 – 清水市（ちびまるこちゃんランド） – 三島市（「ごめんね青春」のロケ地巡礼） – 熱海  
駅に着いたら、まず案内所に地図を取ります。ここで「ごめんね青春」の聖地巡礼の地図があります。心の形のストラップも売っています。



出典：筆者より作成

単なる東京のような大都市を訪問するだけではなく、新たな3つの大都市 + 地方の旅行形式を発見した。

# 分析结果

## 旅行形态的例2：大城市与地方的组合

### • 555号游记的旅程：

东京：浅草-东京塔-浅草寺-新宿歌舞伎町-浅草寺商店街

栃木县：宇都宫-日光东照宫-中禅寺湖-江户村-鬼怒川

鬼怒川是温泉乡，许多温泉饭店都盖在这条河的两边。

德川家康统一天下，为德川幕府建立两百多年的基业，他死后被称为江户幕府的守护神‘东照神君’所以供奉他的神社就叫东照宫。

此次的日光行非常完美，虽然是旅游旺季，但因为大家都愿意早起赶路，并没有碰到很多塞车塞人的状况。改日当挑非旺季，好好把日光再游一遍

### • 864号游记的旅程：

东京：新宿-涉谷-银座-六本木-代官山-迪士尼海洋-三丽鸥乐园

神奈川县：镰仓-镰仓高校前-镰仓高校前铁路闸口-镰仓大佛-鹤冈八幡宫-湘南海岸

灌篮高手！对，我要去看灌篮高手的故乡！

这边海湾作为漫画中，最重要的外景地，也是我的必游之地。

### • 9444号游记的旅程：

东京-东京塔-热海-伊豆-东伊豆（温泉）

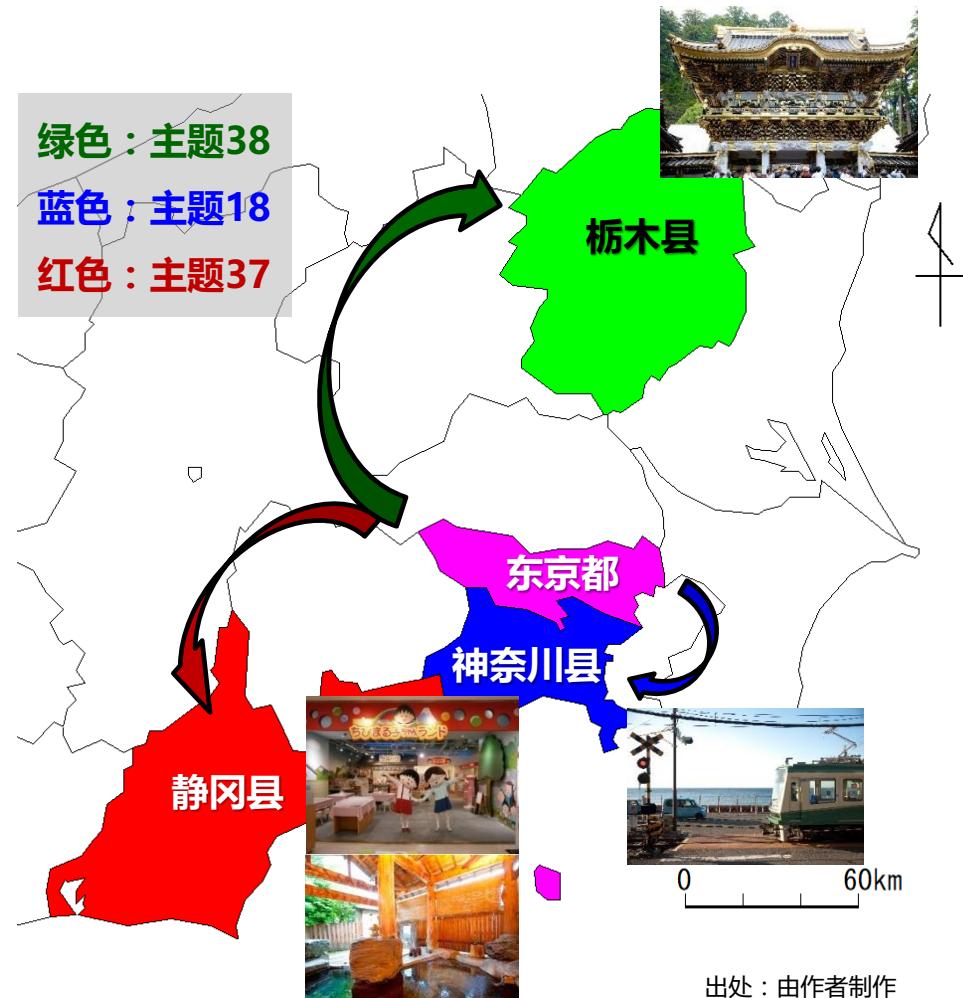
酒店自酿的清酒，几乎每家温泉酒店都会有自家酿酒，在订房间的时候可以预约

住酒店可以享受一次私人温泉，8点就起来泡泡泡～也是唯一一个可以使用手机拍照的温泉哟

### • 1039号游记的旅程：

东京-清水市（小丸子博物馆）-三岛市（《对不起青春》拍摄地）-热海

到了先去车站外面的观光案内所拿地图，这里有《对不起青春》所有的巡游地点图，还有心形吊环的手机链卖哦。



出处：由作者制作

不是仅仅访问东京，进行大城市中的游览  
而是发现了3个新的「大城市 + 地方」的旅游形式

# 分析結果

## 旅行実態の例3：クルーズ

### ・ 6949番旅行記の旅程：

乗船 – チェックイン – 船上のグルメを食べる – 船上の施設を楽しむ  
– 長崎に上陸 – 長崎平和公園 – 長崎中華街 – 免税店 – 船に戻る

- 室内プール、日光浴などは甲板の14階にあります。最もおすすめしたいのはロイヤルカリビアンが旅行者達に用意しているショーです。毎晩催されています。
- 長崎での旅程は平和公園、中華街、2つの免税店です。
- 最もおすすめしたいのは、毎晩登場するショーです。物凄く面白いです！
- 91メートルの景観鑑賞のキャビンに乗って、巨大な船の全体と海の美しい風景見えます。物凄く美しいです。



出典：筆者より作成

- ゴールデンルートを離れ、長崎、広島、高知のような日本の地方に訪問できた。
- クルーズ船の魅力：クルーズのそのもの、日本に上陸後の楽しめる項目を豊富することは今のクルーズ旅の課題となっている。
- 「短期間」、「団体旅行」、「買物を主体」

# 分析结果

## 旅行形态的例3：邮轮

### • 6949号游记的旅程：

登船-登记入住-品尝船上美食-体验船上的娱乐设施-长崎和平公园-长崎中华街-免税店-登船继续巡游

- 室内游泳池、日光浴场、水上冲浪都在**甲板**14层。尤其要大力推荐的就是皇家加勒比为游客精心准备的精彩**演出**，每晚都有，精彩纷呈！
- **长崎**的行程是平和公园、中华街、2个免税店（1个电器多点、1个药妆为主）。
- 尤其要大力推荐的就是皇家加勒比为游客精心准备的**精彩演出**，每晚都有，精彩纷呈！
- 登上91米高空的“北极星”观景舱，巨大的船体和360度海景可以尽收眼底！



出处：由作者制作

- 脱离了**日本旅游的黄金线路**，中国人的脚步到达了**长崎**，**广岛**，**高知**这样的**日本的地方区域**
- 邮轮的魅力：**邮轮本身**，登入日本后，如何**丰富在日游览项目**，成为了目前邮轮游的**课题**
- 「**短时间**」、「**团体旅行**」、「**购物为主**」

# まとめと今後の課題

## □ まとめ :

- ① 本研究はクローラーによる自動的なWebページ上の旅行記の収集を行い、**情報学が観光学への活用**を実現できた。
- ② 10606篇の旅行記の中のおよそ10万種類・280万個の形態素に対して、トピックモデルを行い、訪日中国人旅行者の**旅行行動を把握**し、**新たな旅行行動の発掘**を試みた。

## □ 今後の課題 :

- ① データの上限を**設定おらず**、クローラーによる**情報収集を引き続く**こと。
- ② 長期的に**バスケット分析**を行い、LDAの方が優れていること。
- ③ 個人属性を活用し、科学的、根拠的にサンプルをグループで分け、トピックとの**クロス集計を優れること**。

# 总结及今后的课题

## □ 总结：

- ① 本研究实现了爬虫对于游记的自动收集。**情报学在观光学方面的活用得以实现。**
- ② 对1万篇游记中的约10万种类，280万个单词进行了主题模型分析，对访日中国旅行者旅行行动进行了把握，尝试挖掘出了旅行行动的新形式。

## □ 今后的课题：

- ① 数据数量**不设上限**，**持续**进行爬虫的游记收集。
- ② 进行长期的分析整理，**优化LDA分析程序。**
- ③ 活用个人属性，科学的有根据的对样本进行分组，**优化与主题的交叉分析。**