Do Tourists from Different Countries Interpret Travel Experience with the Same Feeling? Sentiment Analysis of TripAdvisor Reviews

Luyu Wang^(✉) and Andrei P. Kirilenko

University of Florida, P.O. Box 118208, Gainesville, FL 32611-8208, USA
{luyuw.93, andrei.kirilenko}@ufl.edu

Abstract. National parks attract millions of tourists to enjoy the beauty of nature. The opinions and feelings expressed by tourists in their reviews through social media significantly impact other visitors' tourism-related decisions. Notably, tourists from different countries visiting the same park may express different sentiments and post different experiences. It is not clear if those differences could be attributed to the differences in sentiment analysis software for different languages, or they reflect existing variability in culturally defined tourists' sentiments. To address this question, this study analyzed 27,177 TripAdvisor Grand Canyon, US reviews from visitors arriving from ten different countries with the goal of identification of sentiment differences. We found that while all reviews tend to be positive, there are significant regional differences with European and Japanese tourists routinely expressing lesser satisfaction from their visit. We also found differences in the sentiment expressed in different regions of the same country, such as the north and south of Italy. Overall, we suggest that social media reflects the real differences in the sentiment of visitors coming from different origins.

Keywords: Tourist experience · User-generated content (UGC) · TripAdvisor · Social media

1 Introduction

Tourism is growing dramatically in the early 21st century. National parks and protected areas are identified as major attractions for both domestic and international visitors. Visitor feedbacks on travel experiences in social media are essential sources in trip planning; hence understanding the unstructured user-generated content is crucial for park managers to improve visitors' experience. With the easily accessible social media data, researchers have adopted the information technology approaches, such as text mining, to analyze unstructured user-generated content [1]. Sentiment analysis, a popular natural language processing method, helps the industry to understand the polarity of tourist's reviews and identify management failures. Thus, park management can better accommodate tourists' needs and expectations, such as language assistance, food selection, in-park lodging, and many others.

Many tourists travel to foreign destinations to experience different ways of living, traditions, and customs [2]. These tourists from different countries differ in travel

behaviors and service expectations [3]. The scholars confirm that cultural differences have a great influence on tourist's travel experience [4]. Tourists from different countries differ in their preferences on means of transportation, travel arrangements, activities, and travel styles [5,6]. For example, in a study on preferences and sentiment characteristics among Chinese and other international tourists based on user reviews from Chinese social media [7], the Chinese tourists were found to be more likely to express critical and diverse sentiments in their reviews about Australian destinations than tourists from other countries. Similarly, a study of the TripAdvisor cruise tour reviews found the differences in the sentiments expressed by the North American and European tourists [8] with Americans interpreting their cruise experience more positively and with a more subjective and intimate tone. Meanwhile, it is not entirely clear if the observed differences in sentiment can be attributed to the real differences in tourists' expectations or just reflect the differences in emotion expression in national languages.

In the majority of these studies, researchers focus on comparisons between two countries or regions. However, the tourism industry is becoming more and more internationalized with many destinations receiving tourists from all over the world. Today's tourism businesses emphasize developing a better understanding of the cultural diversity of international tourists from many countries. The purpose of this study is to understand and compare the attitudes of tourists from multiple countries to the same destination. The data are unstructured TripAdvisor reviews written by the Grand Canyon National Park travelers from the top 10 countries in terms of the visitor numbers.

2 Area of Study

The study focuses on the top six attractions of the Grand Canyon National Park, Arizona, United States (Fig. 1): South Rim, Bright Angel Trail, North Rim, South Kaibab Trail, and Rim Trail. The 4,926 km² Grand Canyon park, established in 1919, is a UNESCO World Heritage Site and one of the world's top 10 desired destinations. The park is also an important economic driver of the region, supporting 11,800 jobs. In 2019, nearly 6 million park visitors spent \$891 million in communities located within 60 miles of the park [9].

Fig. 1. Study area: Grand Canyon, USA [10]

3 Data and Methodology

TripAdvisor reviews of the top six attractions of the Grand Canyon National Park in Arizona, United States, were collected through web site scraping. In total, 30,237 reviews written between February 2008 and March 2020 were collected. These top six attraction reviews represent 75% of the Grand Canyon National Park’s total reviews. The reviewers’ place of living (at a city level) was determined from their self-report place of residence and transformed into the latitude and longitude using the Geopy Python software. The geographical coordinates were reverse-geolocated into countries using Google Geocode API. Overall, the location of 27,177 reviewers (89.9%) was determined, resulting in a list of 164 countries of origin. Table 1 shows the number of reviews from the top 10 countries representing 76.8% of the overall collected reviews.

Table 1. Country distribution of reviews with reported locations.

Ranking	Country	Count	%	Ranking	Country	Count	%
1	United States	14,473	53.3	6	France	826	3.0
2	United Kingdom	2,175	8.0	7	Australia	799	2.9
3	Italy	1,246	4.6	8	Germany	581	2.1
4	Canada	1,241	4.6	9	Japan	495	1.8
5	Brazil	997	3.7	10	Spain	399	1.5
Total	All countries	27,177	100				

The number of international tourists in the collected data is overwhelming, given its remote location with only half of the tourists being domestic. Between the foreign tourists, the UK visitors represent nearly one-fourth. While there are tourists from other English-speaking countries such as Canada and Australia, there are also speakers of other European languages and of Japanese, creating a multitude of cultural and linguistic data for analysis.

The overall methodology is as follows. First, the reviews written in non-English languages were translated from multiple languages to English using Google Cloud Translate API base on Google’s pre- trained machine learning models. Second, reviews from the top ten countries were pre-processed using the standard data cleaning methodology [11]. Then, the cleaned data was used to extract the sentiments from tourists’ stated experience. Sentiment analysis was performed by Vader (Valence Aware Dictionary for Sentiment Reasoning) software from the NLTK library using Python. Vader is a lexicon-based sentiment classifier that considers the context of the sentences. For each review, Vader generates four values: a neutrality score, a positivity score, a negativity score, and the overall compound sentiment score. Each of the scores ranges from −1 to 1. From those metrics, we adapted the compound scores to express tourists’ overall evaluation of their experience in the park. Finally, the compound

sentiment scores were used to find the locations of the sentiment hot and cold spots using the ESRI ArcGIS hotspot analysis tool. Each sentiment point was analyzed within the context of neighboring sentiment scores based on a certain neighborhood search threshold. Hence, the hot and cold spots identified locations with consistently high or low review scores.

4 Results

The compound scores were adopted to express tourists’ overall sentiment of their experience in the park (Table 2). The most positive feedback comes from Brazil tourists (M = 0.73), followed by the US and Canada. European tourists have less positive attitude with the lowest sentiment score coming from France (M = 0.59). Japanese tourists have the least positive feedback (M = 0.52). The sentiment scores are consistent with the star ratings but less positive with over 88.3% of tourists rated their experience as excellent.

Table 2. Compound sentiment scores of 10 countries.

Ranking	Country	Mean	SD	Ranking	Country	Mean	SD
1	Brazil	0.73	0.36	6	UK	0.66	0.39
2	US	0.72	0.32	7	Italy	0.61	0.41
3	Australia	0.71	0.34	8	Spain	0.61	0.36
4	Canada	0.69	0.38	9	France	0.59	0.39
5	Germany	0.67	0.34	10	Japan	0.52	0.41
Average	All countries	0.69					

The hotspot analysis (Fig. 2) reveals that the US, Brazil, and Australia are the statistically significant hotspots indicating that tourists from these countries consistently have high sentiment scores (M = 0.69). Contrasting, the tourists from European countries and Japan have fewer positive opinions making statistically significant clusters of low sentiment scores.

Fig. 2. Hotspot analysis of compound scores among the countries and regions. Notice the differences between the same country regions in the UK and Italy.

The analysis at a higher resolution reveals more intricate regional differences (Fig. 3). In Europe, Germany is a statistically significant hotspot contrasting the cold spot in France. In the UK, England is a hot spot while Scotland is a cold spot. Similarly, Northern Italy is a cold spot while Southern Italy is a hot spot.

Fig. 3. Hotspot analysis of compound scores among the European countries. Notice the differences between the same country regions in the UK and Italy.

5 Conclusion

We found significant differences in the attitudes of visitors from different countries visiting our area of study. The most positive sentiments were expressed by Brazilian tourists, consistent with other observations [12]. Similarly, the US reviews tend to be positive as the American tourists frequently use adjectives such as spectacular, awesome, and amazing. This is consistent with North Americans being more emotionally charged and expressive than Europeans [8,13]. The lowest sentiments were provided by tourists coming from the European countries and Japan.

Note that the reviews' star ratings do not necessarily correspond with the text sentiment [14]: even when the star rating is high, the sentiment score may reflect multiple topics of dissatisfaction in the overall positive tourist experience. The expression of those dissatisfaction topics is influenced by the distinct cultural background of the tourist reviewing the travel experience. The expression style differences may cause less positive sentiment detected by European tourists. In a similar vein, European tourists show fewer amount of sentiment-bearing words with a more objective tone [8]. The Japanese are the most unique tourist group [15] expressing the least positive sentiment. Japanese people focus on detail, aesthetics, quality, and

service [16]. Because of that, Japanese tourists are more demanding and have higher service expectations, which may align with the level of service provided in the US.

The differences in expressed sentiments are lower between the tourists coming from the same regions as compared to the between-region differences resulting in a pattern of hot and cold spots which mark the regions with consistently more positive and less positive tourist reviews. Notably, those patterns are frequently crossing the borders suggesting that they reflect cultural differences rather than an artifact in sentiment analysis software processing texts originating from different languages. On the other hand, some linguistically similar but culturally diverse countries such as England and Scotland or South and Northern Italy exhibit both the hot and cold spots. Overall, we suggest that social media reflects the real differences in the sentiment of visitors coming from different countries and regions. The sentimental difference across different countries may be due to cultural differences such as expression styles and service expectations.

One limitation is that the text analysis relied on machine translation. We did not check the translation quality explicitly. The literature however suggests that the interrater percentage agreement in human vs. Google Cloud translation vary between 85% and 97% for 9 major languages [17]. This hints that the quality of machine translation already exceeds the ability of humans or computers to recognize the emotions in the written text [18] and hence was deemed adequate for the purpose of this study. Another limitation is using only one park in this pilot study. In the full study under progress, we are applying this methodology to study multiple natural parks around the globe.

Notice that this study did not provide confidence intervals nor p-values for the findings. One reason for that is that data represents the entire population of published reviews and not a sample. Meanwhile, even for the “bid data” samples the traditional tests of statistical significance become meaningless since for large N the p-values tend to either zero or one. An excellent discussion of the alternative measures was provided in [19].

References

1. Stabb S et al (2002) Intelligent systems for tourism. *IEEE Intell Syst* 17(6):53–66
2. Yiamjanya S, Wongleedee K (2014) International tourists’ travel motivation by push-pull factors and the decision making for selecting Thailand as destination choice. *Int J Soc Behav Educ Econ Bus Ind Eng* 8(5):1348–1353
3. Li XR, Lai C, Harrill R, Kline S, Wang L (2011) When east meets west: an exploratory study on Chinese outbound tourists’ travel expectations. *Tour Manag* 32(4):741–749
4. Reisinger Y, Turner L (2003) *Culture. Cross-Cultural Behaviour in Tourism: Concepts and analysis*. Elsevier Science Limited, Oxford, pp 3–33
5. Nutsugbodo RY, Amenumey EK, Mensah CA (2018) Public transport mode preferences of international tourists in Ghana: implications for transport planning. *Travel Behav Soc* 11:1–8
6. Wong CKS, Kwong WYY (2004) Outbound tourists’ selection criteria for choosing all-inclusive package tours. *Tour Manag* 25(5):581–592
7. Liu Y, Huang K, Bao J, Chen K (2019) Listen to the voices from home: an analysis of Chinese tourists’ sentiments regarding Australian destinations. *Tour Manag* 71:337–347

8. Buzova D, Sanz-Blas S, Cervera-Taulet A (2019) Does culture affect sentiments expressed in cruise tours' eWOM? *Serv Ind J* 39(2):154–173
9. Thomas CC, Koontz L (2020) 2019 National Park Visitor Spending Effects - Economic Contributions to Local Communities, States, and the Nation. https://www.nps.gov/nature/customcf/NPS_Data_Visualization/docs/NPS_2019_Visitor_Spending_Effects.pdf
10. Staff (2020) Where is the Grand Canyon? <https://www.mygrandcanyonpark.com/park/where-is-the-grand-canyon>
11. Marine-Roig E, Clavé SA (2015) Tourism analytics with massive user-generated content: a case study of Barcelona. *J Destin Mark Manag* 4(3):162–172
12. Rede Nacional Do Esporte (2016) Tourists give Olympic Games thumbs up and want to visit Brazil again. <https://rededoesporte.gov.br/en/news/tourists-give-olympic-games-thumbs-up-and-want-to-visit-brazil-again>
13. Hardt D, Wulff J (2012) What is the meaning of 5*'s? An investigation of the expression and rating of sentiment. In KONVENS, pp 319–326
14. Ghose A, Ipeirotis PG (2010) Estimating the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. *IEEE Trans Knowl Data Eng* 23(10):1498–1512
15. Özdemir C, Yolal M (2017) Cross-cultural tourist behavior: an examination of tourists' behavior in guided tours. *Tour Hosp Res* 17(3):314–324
16. Reisinger Y, Turner L (1999) A cultural analysis of Japanese tourists: challenges for tourism marketers. *Eur J Mark* 33:1203–1227
17. Jackson JL et al (2019) The accuracy of google translate for abstracting data from non-english-language trials for systematic reviews. *Ann Internal Med* 171(9):677–679
18. Kirilenko AP, Stepchenkova SO, Kim H, Li X (2018) Automated sentiment analysis in tourism: comparison of approaches. *J Travel Res* 57(8):1012–1025
19. Lin M, Lucas HC Jr, Shmueli G (2013) Research commentary—too big to fail: large samples and the p-value problem. *Inf Syst Res* 24(4):906–917

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.