

# Sentiments and Opinions From Twitter About Peruvian Touristic Places Using Correspondence Analysis

Luis Cajachahua<sup>1</sup> and Indira Burga<sup>2</sup>

<sup>1</sup> Universidad Nacional de Ingeniería, Lima, Perú  
lcajachahua@gmail.com  
<https://www.linkedin.com/in/lcajachahua/>

<sup>2</sup> Universidad de Ingeniería y Tecnología, Lima, Perú  
indira.burga@gmail.com  
<https://www.linkedin.com/in/iburga/>

**Keywords:** Sentiment Analysis, Text Mining, Correspondence Analysis, Perú, Twitter, Tourism

Tourism in Perú has become very important, since there is a growing number of tourists arriving each year. This paper focus in understand what do English-speaking tourists consider when they visit Perú. We obtained all the tweets written in english during the year 2016, filtered by a list of touristic places visited.

In total, more than 192 thousand tweets were collected. We performed different analysis to describe the data, including correspondence analysis, a statistical technique which is normally applied to categorical data. The goal was to understand the sentiments and opinions expressed in those tweets.

## 1 Introduction

Tourism is an important source of economic growth in Perú. In 2015, 3.5 million of tourists visited Perú, leaving more than USD 4,151 million that correspond to the 3.75% of the gross domestic product. For this reason, all the government tourism bureaus, specially Promperú, are working hard to keep those numbers growing.

For that reason, we downloaded 192,525 tweets, published during 2016, considering only english language. After a careful data preparation and feature extraction, we applied some sentiment analysis techniques to tag each tweet using the following eight sentiments: anger, fear, sadness, disgust, surprise, anticipation, trust and joy, proposed by Plutchik. Using correspondence analysis, we discovered some associations between touristic places visited and sentiments expressed in the extracted tweets.

## 2 Conclusions

Considering the insights found in this study, we can conclude:

- The number of users tweeting about touristic places is interesting, but only 2.5% of them are active users, posting 12 or more tweets per year.
- The most commented place in Perú, by far, is Machu Picchu. It is mentioned more than words like Cusco, or even Perú. It was declared a New Wonder around the world in an international contest in 2007.
- The site most valued by English-speaking tweeters is Aguas Calientes, with more than 95% of comments classified as positive. Lago Titicaca is the worst rated, since it does not reach 70% of positive comments.
- We can notice a great influence of negative incidents in Twitter, due to the viral effect.
- This study was a good way to explore the opinions about Perú, but we have to keep in mind some limitations, like representability or social network popularity in our country.
- Twitter activity is the highest in July, during the northern hemisphere countries summer vacation time. And in October, when some incidents occurred in the most visited place, Cusco.

**Published in:** Cajachahua, L., Burga, I.: Sentiments and Opinions from Twitter About Peruvian Touristic Places Using Correspondence Analysis. Proceedings of the 4th Annual International Symposium on Information Management and Big Data, pp. 178–189. (2017). <http://ceur-ws.org/Vol-2029/paper17.pdf>