

User Generated Big Data Analysis of Customer Ratings of Beaches in Andaman and Nicobar Islands of India



Anagha Prem, Sangeetha Gunasekar, Dileep G Menon

Abstract: *In this study, we analyse user-generated content of beach tourist attractions across Andaman and Nicobar Islands in India to find the factors that influence tourist experience in terms of customer satisfaction in island beach destinations. The study is conducted to identify the specific attributes of beach tourism that has positive and negative influence over the high rating provided on online review portals. The study is conducted over the reviews obtained from online reviews over Trip Advisor. Text analytics and logistic regression is used to find the various attributing factors that affect the customer satisfaction in the form of rating. The online data is collected and the unique words are identified with respect to the frequency of occurrence of the unique words in the reviews and taken for analysis. The results indicate that there is more influence of factors such as cleanliness of the beach, the accommodation facilities available and the various beach activities in inducing better satisfaction among the tourists than the scenic beauty of the place. The study has policy implications as the tourism development activities undertaken by the government of Andaman & Nicobar Islands must focus on the positive influencing factors and focus should be given on the factors such as better restaurant facilities, crowd management and shopping facilities in order to enhance the tourist experience at Andaman & Nicobar Islands.*

Index Terms- *Beach Tourism, Online Review Analysis*

I. INTRODUCTION

Academic research has focused on island tourism since several years [1] [2]. While themes covering economic impact of foreign revenue earned [3][4], sustainability of island tourist destinations [5], coastal scenery assessments [6], cruise experience [7];[8], eco-tourism and resulting conflicts between human settlements, increasing tourist footfalls and ecological sustainability of islands [9] have been the focus of several studies in the literature, not many have analysed the reactions of tourists visiting different tourist attractions across an island destination using user-generated content analysis [10].

Revised Manuscript Received on December 30, 2019.

* Correspondence Author

Anagha Prem*, School of Business, Amrita Vishwa Vidyapeetham Coimbatore, India

Sangeetha Gunasekar, School of Business, Amrita Vishwa Vidyapeetham Coimbatore, India

Dileep G Menon, School of Business, Amrita Vishwa Vidyapeetham Coimbatore, India

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](http://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Tourism industry with a revenue of US\$733 billion and employment at 8% of global workforce, is expected to see over 1.6 billion international tourists by 2020 (UNWTO, 2008). Beaches are the most popular tourist attractions in islands. Many studies highlight the importance of beaches as a tourist destination in the literature [10]; [11].

The studies pertaining to Andaman & Nicobar Islands of India are scant in the literature. One of the few studies that have focused on these islands [12] estimate the carrying capacity of beaches at Andaman & Nicobar Islands and find that while the islands tourist carrying capacity is well within the limits, accommodation and other infrastructure needs to be improved to at least match the effective carrying capacity of these islands.

Others like [13] find these islands to have a high potential for ecotourism. The Andaman & Nicobar Islands of India are bestowed with beautiful coastline and sandy beaches and are a growing attraction for tourists across the world. The present study use the user generated content (online reviews from Tip Advisor) of tourist attractions across Andaman & Nicobar Islands to understand the factors that influence tourist experience in terms of customer satisfaction.

These islands of India are endowed with some of most beautiful white-sand beaches, mangrove forests, a large number of rare species of flora and fauna [14].

It is a growing tourist destination of India comprising of over 572 islands, islets and rocks, of which only 38 are inhabited. Most of these islands have forest cover with substantial parts under reserve and protected forest area that are earmarked as primitive tribal reserves.

Contact with local tribes is restricted. The islands hold the potential to grow into one of the most attractive tourist destinations of India.

The inflow of tourists increased over the years as clearly seen in Figure 1 with domestic tourists showing stronger increasing growth trends.

This increasing trend of domestic as well as foreign tourism can be capitalised for the growth of the tourism industry and there by the economy.

The data was sourced from Basic Statistics 2016-17, Directorate of Economic and Statistics, Andaman and Nicobar Administration.

logistic regression is used for the estimation. The dependent variable is customer satisfaction defined as:
 $y = 1$, if customer rates the experience as “Excellent” (Rating =5);
 $y = 0$, otherwise

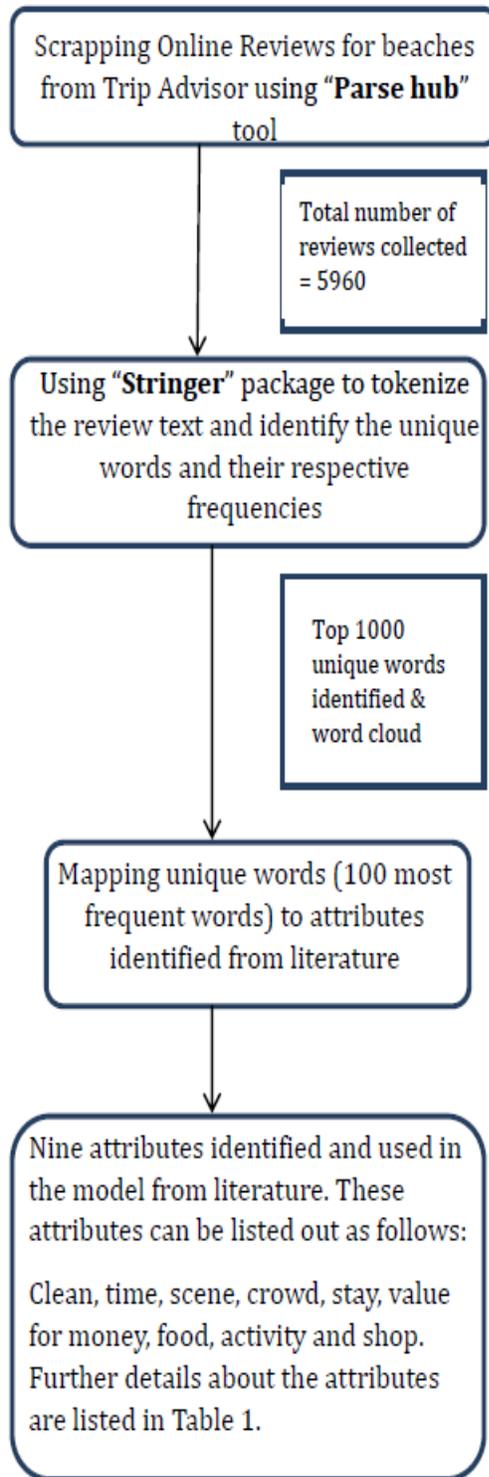


Fig 3: Research Methodology

Table 1: Words included under each attribute

S No:	Attribute	Words included
1	Clean	"clean,cleaned,cleaning,cleanly,cleanness,clear,crystal,crystal,crystals,neat,pristine"
2	Time	"sunset,'sunset',sunsets,time,timely,times,timing,timings,evening,evenings,morning,mornings, sunrise,sunrise,sunrises"
3	Scene	"forest,forests,scenic,coconut,coconut,coconuts,greeneries,greenery,tree,trees,view,viewed,viewing,views"
4	Crowd	"people,peoples,crowd,crowded,crowding,crowds,quite,quiet"
5	Stay	"room,rooms,resort,resorts,stay,stayed,staying,stays"
6	Value for money	"worth,worthed"
7	Food	"food,foodings,foods,restaurant,restaurant',restaurants,restaurant,restaurents"
8	Activity	"walk,walked,walking,walks,swim,swimmed,swimming,swims,ride,rides,riding,sport,sporting,sports"
9	Shop	"shop,shopped,shopping,shop,pings,shops,'shops"

A. Logistic Regression

Logistic regression estimation is a special case of linear regression models where the Y variable is categorical. It helps compute the probability of occurrence of an event. Dependent variable is log of odds ratio with equation forms as:

$$f(E(y)) = \alpha + \beta x_1 + \gamma x_2 + \epsilon$$

Here, f() is the link function, E(y) is the expectation of target variable and $\alpha + \beta x_1 + \gamma x_2 + \epsilon$ is the linear predictor (α, β, γ to be predicted) and ϵ is the error term. The role of link function is to 'link' the expectation of y to linear predictor. In the study the dependent variable of customer satisfaction is dichotomous, as only the highest rating is taken into consideration. The unique words are grouped into constructs specified by the attributes and these are taken as the independent variables.

User Generated Big Data Analysis of Customer Ratings of Beaches in Andaman and Nicobar Islands of India

The following logistic regression model was used to test the model

$$\begin{aligned} \text{Customer Satisfaction} &= \beta_1 + \beta_2 \text{clean} + \beta_3 \text{time} + \beta_4 \text{scene} \\ &+ \beta_5 \text{crowd} + \beta_6 \text{stay} \\ &+ \beta_7 \text{valueformoney} + \beta_8 \text{food} \\ &+ \beta_9 \text{activity} + \beta_{10} \text{shop} + \varepsilon \end{aligned}$$

Where ε = standard error.

The above model was estimated using logistic regression and the results of the estimate are given in Table 2

Table 2: Results of logistic regression analysis

Independent Variable	Coefficient	Odds Ratio	P value
Clean	.762	2.142	0.000*
Time	.026	1.026	0.643
Scene	.078	1.081	0.185
Crowd	-.282	.754	0.000*
Stay	.144	1.155	0.051**
Value for money	-.175	.839	0.144
Food	-.314	.730	0.001*
Activity	.183	1.200	0.006*
Shop	-.413	.661	0.000*
Notations of *, **,*** imply significance at 10%, 5% and 1% level for p respectively			

V. RESULTS

According to the model, various variables seem to impact the customer rating given the beach fronts of Andaman. From the table we see that, people who find the beaches clean tend to rate higher than the scenic beauty offered or the time of the visit. The restaurant facilities and shopping facilities available to the customers tend to have significantly negative effect on the review rating provided to the beaches. The behaviour of the people as well as the intensity of the crowd seems to have significantly negative influence as well. On the other hand the quality of accommodation and stay provided at the beach tends to prompt the customers to give a higher rating. The various beach activities such as swimming etc. tends to have significantly a positive effect on the higher rating of the beaches. Contrary to normal assumptions value for money does not seem to significantly contribute to the higher rating provided by customers.

VI. CONCLUSION

In the hospitality literature, studies have documented the use of websites and social media in providing useful information in terms of review of experiences shared by travellers for

destination, property, facilities and restaurants [24]. The reviews and opinions which were recorded in traditional form is now digitalized and shared globally. Online reviews of a destination not only attract or detract a traveller, but also develop consumer expectations [25]. The present study used online reviews to analyse the various attributes that impact customer satisfaction for beach tourism.

Our conclusion from the study indicates that the attributes contributing towards the higher rating given by customers are factors such as cleanliness, various beach activities and the resort facilities provided in the island which act as positive influence over the tourist satisfaction, whereas the poor state of restaurant facilities, the crowd and the lesser opportunities for shopping tends to have negative effect over the same. Therefore, government and private business dependent on beach tourism should focus on the above factors in order to enhance the tourist experience at Andaman & Nicobar Islands.

VII. LIMITATIONS

The scope of the study in this article has been limited to reviews taken from Trip Advisor. The study conducted poses the limitation of the focus being given to only 15 beaches of Andaman & Nicobar Islands. The study can also be broadened by doing a large-scale research with data collected from an extensive range of social media and other forums. Further study can be conducted with a broader sample and area to be focused upon.

REFERENCES

- Lockhart, D. G. (1996). Island tourism reviewed. *Tourism Management*, 17(5), 382-383.
- Sharpley, R. (2012). Island tourism or tourism on islands?. *Tourism Recreation Research*, 37(2), 167-172.
- Stauvermann, P. J., & Kumar, R. R. (2016). Economics of tourism & growth for small island countries. *Tourism Management*, 55, 272-275
- Roudi, S., Arasli, H., & Akadiri, S. S. (2018). New insights into an old issue—examining the influence of tourism on economic growth: evidence from selected small island developing states. *Current Issues in Tourism*, 1-21.
- Rodríguez, J. R. O., Parra-López, E., & Yanes-Estévez, V. (2008). The sustainability of island destinations: Tourism area life cycle and teleological perspectives. The case of Tenerife. *Tourism Management*, 29(1), 53-65.
- Rangel-Buitrago, N., Correa, I. D., Anfuso, G., Ergin, A., & Williams, A. T. (2013). Assessing and managing scenery of the Caribbean Coast of Colombia. *Tourism Management*, 35, 41-58
- DiPietro, R. B., & Peterson, R. (2017). Exploring cruise experiences, satisfaction, and loyalty: the case of Aruba as a small-island tourism economy. *International Journal of Hospitality & Tourism Administration*, 18(1), 41-60.
- Ozturk, U. A., & Gogtas, H. (2016). Destination attributes, satisfaction, and the cruise visitor's intent to revisit and recommend. *Tourism Geographies*, 18(2), 194-212.
- Ho, J. A., Chia, K. W., Ng, S. L., & Ramachandran, S. (2017). Problems and stakeholder responsibilities in island tourism: The case of Tioman Island in Malaysia. *Journal of Hospitality & Tourism Research*, 41(4), 445-474.
- López-del-Pino, F., & Grisóla, J. M. (2018). Pricing Beach Congestion: An analysis of the introduction of an access fee to the protected island of Lobos (Canary Islands). *Tourism Economics*, 24(4), 449-472

11. Chen, C. L., & Bau, Y. P. (2016). Establishing a multi-criteria evaluation structure for tourist beaches in Taiwan: A foundation for sustainable beach tourism. *Ocean & coastal management*, 121, 88-96.
12. Phillips, M. R., & House, C. (2009). An evaluation of priorities for beach tourism: Case studies from South Wales, UK. *Tourism Management*, 30(2), 176-183.
13. Bera, S., Majumdar, D. D., & Paul, A. K. (2015). Estimation of Tourism Carrying Capacity for Neil Island, South Andaman, India. *Journal of Coastal Sciences*, 2(2), 46-53.
14. Thulasimala, D., & Devdass, P. (2010). Island eco-tourism: A case study of Andaman islands, India. *Indian Journal of Science and Technology*, 3(12), 1247-1252
15. Sircar, P. K. (2004). *The Primitive Tribes of Andaman & Nicobar Islands*. Akansha Publishing House.
16. Fang, B., Ye, Q., Kucukusta, D., & Law, R. (2016). Analysis of the perceived value of online tourism reviews: Influence of readability and reviewer characteristics. *Tourism Management*, 52, 498-506.
17. Sotiriadis, M. D., & Van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research*, 13(1), 103-124
18. He, W., Tian, X., Tao, R., Zhang, W., Yan, G., & Akula, V. (2017). Application of social media analytics: a case of analyzing online hotel reviews. *Online Information Review*, 41(7), 921-935.
19. Lu, X., Ba, S., Huang, L., & Feng, Y. (2013). Promotional marketing or word-of-mouth? Evidence from online restaurant reviews. *Information Systems Research*, 24(3), 596-612.
20. Papathanassis, A., & Knolle, F. (2011). Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tourism Management*, 32(2), 215-224.
21. Cave, J., & Brown, K. G. (2012). Island tourism: destinations: an editorial introduction to the special issue. *International Journal of Culture, Tourism and Hospitality Research*, 6(2), 95-113.
22. Cui, W., Wu, Y., Liu, S., Wei, F., Zhou, M. X., & Qu, H. (2010, March). Context preserving dynamic word cloud visualization. In *2010 IEEE Pacific Visualization Symposium (PacificVis)* (pp. 121-128). IEEE.
23. Lohmann, S., Heimerl, F., Bopp, F., Burch, M., & Ertl, T. (2015, July). Concentri cloud: Word cloud visualization for multiple text documents. In *2015 19th International Conference on Information Visualisation* (pp. 114-120). IEEE.
24. Heimerl, F., Lohmann, S., Lange, S., & Ertl, T. (2014, January). Word cloud explorer: Text analytics based on word clouds. In *2014 47th Hawaii International Conference on System Sciences* (pp. 1833-1842). IEEE.
25. Litvin, S. W., & Hoffman, L. M. (2012). Responses to consumer-generated media in the hospitality marketplace: An empirical study. *Journal of Vacation Marketing*, 18(2), 135-145.
26. De Bruyn, A., & Lilien, G. L. (2008). A multi-stage model of word-of-mouth influence through viral marketing. *International journal of research in marketing*, 25(3), 151-163.

AUTHORS PROFILE



Anagha Prem, final year MBA student at Amrita School of Business Coimbatore, pursuing marketing and human resource management after graduating in computer science and engineering.



Dr. Sangeetha Gunasekar is currently working as the Associate Professor at Amrita School of Business, Coimbatore, India.



Dileep G Menon is a research scholar at Amrita School of Business, Coimbatore, India.