Estimating Destination Carrying Capacity: The Big Data approach

Oksana Tokarchuk
Free University of Bozen

Jacob Charles Barr
Free University of Bolzano

Claudia Cozzio
Free University of Bolzano

Estimating Destination Carrying Capacity: The Big Data approach

Introduction

Before the emergence of the COVID-19 crisis, forecasts of the continued tourism growth made overcrowding and related negative consequences for destination image and sustainability, the primary concerns for many destinations (Peeters et al., 2018). Although the emergency caused by the COVID-19 drastically reduced visitation of many renowned destinations, the hopes of "back to normal" are also associated with fears of large crowds invading touristic spots once the travel will be allowed. Overall, research shows that crowding leads to changes in the perception of a destination's liveability, appeal or economic viability (Bellini et al., 2017; Kuščer & Mihali, 2019). Accordingly, a deepening understanding of managing tourist flows and crowds appears essential for policymakers (Butler & Dodds, 2019).

For understanding, this the notion of carrying capacity is frequently used in the literature. The social carrying capacity of tourism is defined as the maximum number of tourists present at a destination without their activities being unacceptable to residents and without precluding tourists from appreciating the destination (McCool & Lime, 2001; Saveriades, 2000).

Individual emotional perceptions on social density in touristic areas have been analyzed in sociology and anthropology since late 1960 (Cohen, 1979; MacCannell, 1976; Smith, 1989). For example, Rosenow and Pulsipher (1979) associate seasonal visitor pressure, adverse tourist behaviour and environmental impacts with negative tourism outcomes, anticipating the underlying consequences of overtourism (Capocchi et al., 2019). Recently, in some destinations, vast tourism growth levels cause organized and active anti-tourist protests (Gössling et al., 2020).

Research measuring tourists' carrying capacity is scant. Only a few studies managed to provide a concrete carrying capacity measure (i.e. Rathnayake, 2015; Zhang et al., 2017). These studies are based on questionnaire-based data collection. This method suffers from several limitations. First of all, ad-hoc data collection requires considerable budgets and resources for its implementation. Given the budget limitations, it is challenging to conduct data collection with an appropriate sample for multiple years for the same destination. This limitation makes questionable external validity of the obtained estimates of carrying capacity and does not permit investigating the dynamics of destination evolution. Finally, questionnaires pose hypothetic situations, often tricky for interviewees to understand and come up with responses.

Yu and Egger (2021) propose a first study to explore tourists' perceptions when visiting crowded attractions in Paris. As tourists are more likely to share their travel experiences online (Varkaris & Neuhofer, 2017), they investigate tourists' feelings based on TripAdvisor online reviews. The findings illustrate that tourists feel most negatively about safety and security due to the crowds through topic modelling and sentiment analysis.

The present study aims at assessing social carrying capacity from the demand side (tourists), addressing the need for more objective evaluations. This research mainly relies on secondary data, combining sentiment scoring derived from TripAdvisor online reviews with official statistics about touristic flows.

This study adopts the NRC Emotion Lexicon which includes a list of words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, disgust) and two sentiments (negative, positive). While this method is commonly used to analyze public emotions

in certain situations (e.g. pandemic situation: Aslam et al., 2020; political elections: Fagbola & Thakur, 2019; online product reviews: Bose et al., 2020), its application in the tourism field appears novel. For this study the destination Berlin has been investigated.

Literature Review

Carrying capacity and overtourism

Initially adopted in animal ecology (Kessler, 1994), the notion of carrying capacity is proposed by range managers due to concerns on land use for grazing livestock (Liu & Borthwick, 2011). Accordingly, carrying capacity was defined as "the entire amount of a given animal species that the habitat can sustain without yielding irrecoverable damage on the ecosphere" (Wei et al., 2015, p. 65).

Successively, the concept has been adapted to the urban context, leading to a comprehensive definition of urban carrying capacity (Joardar, 1996; Oh et al., 2005). The urban carrying capacity is defined as the limit of urban development beyond which damages to natural resources, urban services, public perception may occur (Sarma et al., 2012). This notion's innate feature lies in its ever-changing nature, according to the reactions between technology utilization, human preferences, investments, and consumption patterns (Graymore et al., 2010).

In the tourism context, as early as 1960, initial discussions about the negative consequences of excessive tourism on the host communities and the natural environment emerge (Foster, 1964; Wager, 1964). These evolving discussions culminate in several theoretical concepts. Doxey's irridex model (Doxey, 1975) depicts residents' perceptions of tourism in their community and represents a first theoretical grounding of this phenomenon. Successively, Pizam's Attitude-Index (1978) measures residents' perceptions about the impact of tourism on different domains, attempting to examine the existence of tourism's adverse effects empirically. Rosenow and Pulsipher (1979) unveil other causes related to tourism's negative consequences, such as too many visitors, particularly during season peaks, rowdiness and noise as adverse visitors impact, crowdedness in touristic areas (e.g. city centres) with consequent destruction of natural resources. Later, Butler (1980) proposes a tourist life cycle, suggesting that the tourist destination life cycle goes through six stages (i.e. exploration, involvement, development, consolidation, stagnation, decline/rejuvenation).

Overall, this initial work underpins that tourism concentrations cause harm to the local environment and negative attitudes among residents.

Later, the discussion about the carrying capacity of a destination moves forward with the aim to investigate the number of tourists allowed in a specific touristic destination without causing negative consequences to the environment and residents' perceptions (Van der Borg et al., 1996). While social carrying capacity remains a popular concept to depict tourism's negative consequences, a precise definition is only provided in early 2000. The social carrying capacity of tourism is defined as the maximum number of tourists present at a destination without their activities being unacceptable to residents and without precluding tourists from appreciating the destination (McCool & Lime, 2001; Saveriades, 2000). The diffusion of social carrying capacity defines the emergence of the term overtourism, starting from 2017. When the number of tourists at a destination surpasses the social carrying capacity, the phenomenon of overtourism is observed (UNWTO, 2018). The UNWTO officially defines overtourism as "the impact of tourism on a

destination, or parts thereof, that excessively influences the perceived quality of life of citizens and/or quality of visitors' experiences in a negative way" (UNWTO, 2018).

Estimations of social carrying capacity

In the literature, social carrying capacity is perceived as an all-including notion able to measure the impact of tourism both on residents' lives and on the situations tourists are experiencing (Saveriades, 2000). Thus, the concept of social carrying capacity has two facets. From the demand side (quality of the touristic experience), it measures "the maximum level of use (in terms of numbers and activities) that can be absorbed by an area without an unacceptable decline in the quality of experience of visitors" (Saveriades, 2000, p. 149). In other words, the notion of social carrying capacity is closely related to the quality of experience that visitors accept before searching for alternative destinations. The supply side (residents) refers to the degree of tolerance of the host population to tourists' presence before detrimental effects on the society of the area occur (Gonzales et al., 2018).

Identifying social carrying capacity thresholds is problematic because it relies on subjective judgements, thus depending on individual preferences, attitudes and opinions (Koens et al., 2018). Consequently, multiple types of carrying capacity limits exist as residents' threshold may differ from the one for tourists. Both differ from the limits set by the environment (Jurado et al., 2012). However, proper research techniques (e.g. questionnaires, public surveys, interviews) allow to develop evaluative standards and obtain valuable insights into people's perceptions (Gonzales et al., 2018).

Concerning the supply side (residents), Tokarchuk et al. (2020) quantify the maximum value of tourism intensity before residents perceive a decline in their satisfaction with life. The use of a mix of primary and

secondary data to measure the threshold value of residents' social carrying capacity represents a novelty in the literature that traditionally relies on ad hoc data collection (e.g. survey, interviews). For example, Saveriades (2000) survey investigates the attitudes and perceptions of the indigenous population towards tourism on the east coast of the Republic of Cyprus, thus enabling the identification of the social carrying capacity threshold of the area and an optimum tourist-host contact ratio.

Regarding the demand side (tourists), Gonson et al. (2018) assess the social carrying capacity in popular coastal areas of New Caledonia, using data from visitors/boats' count survey and questionnaire-based surveys. The survey directly asks users what they will do in the case of high use level, and the number of encounters users would like to see simultaneously. This approach permits the evaluation of several thresholds' values (i.e. preferred number, acceptable number, intolerable number of meetings), enabling the assessment of social carrying capacity by gauging the observed numbers of boats/visitors estimated from the survey concerning the threshold values obtained from the interviews.

Zhang et al. (2017) empirically test the determinants of a theme park tourism social carrying capacity (i.e. waiting time per attraction, wait proportion, visited-attraction number) by surveying visitors at a theme park in China. For this study, the authors focus on the threshold that does not cause an unacceptable low visitor satisfaction level, introducing the notion of psychological carrying capacity. As prior research shows that theme parks' visitors complain about heavy attendance even in the presence of below capacity usage, a park's physical carrying capacity should

exceed its psychological carrying capacity (Zhang et al., 2012). Through one-way ANOVA, linear regression, and multi-group moderation tests, the results show that whereas waiting time is more important at a high attendance level than at a low level, the visited-attraction number is more important to visitors relatively high attendance level than at too high and low attendance levels.

Jacobsen et al. (2019) identify stimulus-overload theory (Schmidt & Keating, 1979) as a critical conceptual framework suited to understanding tourists' crowding perceptions within minor tourist hotspots. They measured approach/avoidance reactions in the context of cruise ship visitation, unveiling that self-organized tourists reveal a lower level of crowding tolerance compared to group tours.

Rathnayake (2015) investigates how crowding affects visitor satisfaction at a National Park located in Sri Lanka. It is the first study that estimates visitor satisfaction concerning crowding in monetary term. The findings indicate that visitor satisfaction as well as willingness to pay decrease with crowding.

Overall, quantifying and estimating social carrying capacity falls in the normative category, as visitors are asked directly about the maximum acceptable number of encounters (Manning et al., 1999). As this approach depends on the evaluation by respondents of meetings with other visitors during a specific period, more objective estimations arise from the literature.

While considerable research efforts have been dedicated to investigating the relationship between the crowd at the destination and tourists' satisfaction with stay, little progress has been made to estimate carrying capacity. Existing literature attempts to evaluate tourism carrying capacity for tourists by interviewing tourists at destination with the help of ad-hoc questionnaires (i.e. Rathnayake, 2015; Zhang et al., 2017). This approach allows individuating the overall satisfaction or dissatisfaction of tourists with a given level of the crowd present in the destination. It fails to estimate the offset point in which positive feelings of sharing destination experience with other tourists turns to negative impression with destination due to the crowd.

Gonson et al. (2018) estimated preferable, acceptable and intolerable levels of other tourists' presence. However, to replicate this approach, destination managers need to conduct ad-hoc data collection involving a representative sample of tourists to obtain a trustable measure of this indicator.

The present study aims to develop an alternative approach to measuring tourism carrying capacity for tourists. We propose a methodology based on big data analytics that provides a point estimate of tourists' carrying capacity. It allows to forgo costly data collection and obtain a more objective assessment of tourism carrying capacity.

Conceptual model development

3.1 Emotional responses to tourism intensity

While tourism intensity from residents' perspective has been theoretically explored in the form of emotional responses that trigger place-protective actions (García et al., 2015, McKercher et al., 2015), prior studies that investigate tourism intensity from the visitors' perspective appear scant. To date, research shows that tourist pressure from the visitors' side is a psychological construct

(Bell et al., 2011; Jurado et al., 2013) that depends on multifaced personal (e.g. motivations, cultural background) and situational (e.g. activity type, environmental characteristics) features (Alazaizeh et al., 2016; Klanjšček et al., 2018; Li et al., 2017). People might consider crowds stressful in specific contexts while appreciating intense flows in other environments (Jacobsen et al., 2019). Gössling et al. (2020) suggest that "tourism will not per se have negative outcomes; rather, it is possible that tourism can positively reinforce perceptions of a place. This, however, is less likely in scenarios where visitor numbers are high with respect to the local population or concentrated in small areas, where tourist behaviours interfere negatively with local customs and social norms, and where fewer people profit from tourism developments" (Gössling et al., 2020, p. 84).

Following this reasoning, we formulate the following hypothesis:

Hypothesis 1a: Tourists' positive emotions towards the destination are affected by tourism intensity

Hypothesis 1b: Tourists' negative emotions towards the destination are affected by tourism intensity

Research in consumer behaviour demonstrates that crowding perceptions affect consumer emotions in the retail setting. Higher levels of crowding are associated with lower levels of positive emotions and heightened negative emotions (Mahleit et al., 2000). Mehta et al. (2013) find that the relationship between crowding perception and consumers' emotions follows an inverted U-shape pattern. For low levels of perceived crowding, they observed increasing pleasure. For high levels of the crowd, a negative value for pleasure is observed. Mehta et al. (2013) conclude that there is an optimum crowding level that allows maximizing the positive emotions from store visitation.

Tokarchuk et al. (forthcoming) demonstrate an inverted U-shape influence of tourism intensity on residents' satisfaction with life. Following this approach, we suggest that tourists' emotions are affected by tourism crowding in a nonlinear fashion. In particular, positive emotions exhibit an inverted u-shape pattern in relationship with tourism intensity. At low levels of tourism intensity at the destination, an increase in tourists' number increases positive emotions related to the visit. Neuts and Nijkampas (2012) suggest that density experiences can activate positive reactions of friendliness and excitement, particularly for vacationers belonging to group tours. Simultaneously, the preferred level of visitation measured in Gonson et al. (2018) suggests that tourists will thoroughly enjoy their visit and thus, experience positive emotions, even if a certain number of other tourists will be present. However, negative sentiments caused by personal space violations, congestion and overuse of infrastructures increase with increased tourism intensity (Kim et al., 2016; Simancas et al., 2019).

Meanwhile, negative emotions follow a u-shape pattern. For low levels of tourism intensity, a decrease in the negative feelings will be observed. At high levels of crowding, tourists will demonstrate increasing negative emotions.

Accordingly, we formulate the following hypothesis:

Hypothesis 2a: Positive emotions are associated with tourism intensity following an inverted u-shape pattern.

Hypothesis 2b: Negative emotions demonstrate a u-shape pattern concerning tourism intensity.

Methodology

Study context

Berlin is a prolific European urban tourism destination that has risen quickly in numbers to occupy the third-most visited city in Europe in terms of bed nights with more than 34 million in 2019 after London and Paris (Statista, 2020). This pronounced growth has led researchers of overtourism to investigate this mass visitation's outworking and determine the effect of increasingly large numbers of visitors on residents' quality of life (Tokarchuk et al., forthcoming). To provide a novel perspective to the debate on whether and how Berlin policy-makers should moderate tourism in Europe's "Capital of Cool" (Novy, 2017), this paper explores how demand dynamics measured in online travel reviews can inform a better understanding of the resident quality of life. To this end, we collected for analysis the online reviews from the ten most popular Berlin attractions on TripAdvisor. These attractions are Berliner Dom, Brandenburg Gate, East Side Gallery, Friedrichstadt Palast, Memorial of the Berlin Wall, Museum Island, Pergamonmuseum, Reichtag Building, The Holocaust Memorial, Topography of Terror.

Tripadvisor Data collection and data treatment

Our analysis is conducted over the period from January 1 2013, to December 31 2019. All available online reviews from January 1 2013, until December 31, 2019, irrespective of the original language, were considered for analysis. Due to the monolingual processing libraries, previous studies tend to exclusively select English-language reviews (Yu & Egger 2020), missing a significant and valuable portion of the total sample. Altogether, 122,317 reviews, including the review text, date of experience, rating and review language, were factored into the analysis.

The review bodies were first pre-processed from their raw state through GoogleTranslate to establish a uniform lexicon and then cleaned by the R package TidyText (Silge & Robinson, 2021), removing punctuation, special characters and lowercasing letters. The textual analysis employed the NRC Word-Emotion Association Lexicon (Mohammad & Turney, 2013) to extract an emotional index based on the prevalence of emotionally charged sentiments in the reviews. Reviews were split by month, and a corresponding average emotional and sentiment score was ascribed to each month. Additionally, the original language variable per review was preserved as a factor, counted and recorded as a composite index. Through this procedure, we obtained two separate scores for positive and negative emotions. The trend of the two indices is presented in figure 1.

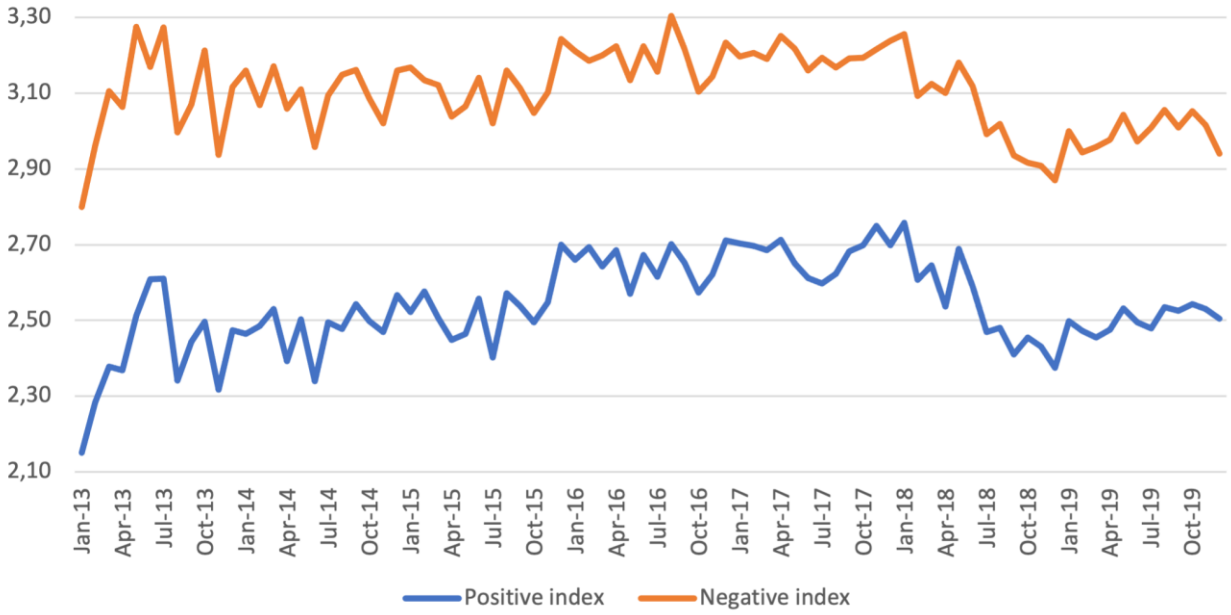


Figure 1. Monthly positive and negative index from TripAdvisor reviews for top 10 Berlin attractions

Berlin tourism data collection

We retrieved the data on tourists' presence in Berlin from the official statistical sources. We collected monthly tourists' beds in Berlin from January 1 2013, to December 31 2019. Following previous literature, we consider tourists nights per resident as the independent variable. The trend of tourists' presence in Berlin during the analyzed period is presented in Figure 2.

The incidence of tourists visiting Berlin has grown during the analyzed period from 26.9 mln registered in 2013 to 34.1 mln in 2019, a growth of 27%. As Figure 2 demonstrates, the tourists' presence in Berlin is seasonal. December, January and February are the months that correspond to the lowest number of tourists. January 2013, with 0.44 tourists nights per resident, registered the lowest incidence of tourists in Berlin during the analyzed period. The peak season goes from May to October, while August is the busiest month of the year. August 2019 was the months with the most tourists nights per resident. During the observed period, the growth of tourists' presence is mainly observed during the low season, with January and February demonstrating the highest growth, corresponding to 45% growth from 2013 to 2019. The tourism intensity grew the least in August, accounting for only 18% of growth from 2013 to 2019.

Method of analysis

The dependent variables used in the present study, the positive and negative indices, are continuous. For this reason, ordinary least square regression has been chosen to analyse the impact of tourists' presence on the sentiment score that is recovered from the tourists' reviews.

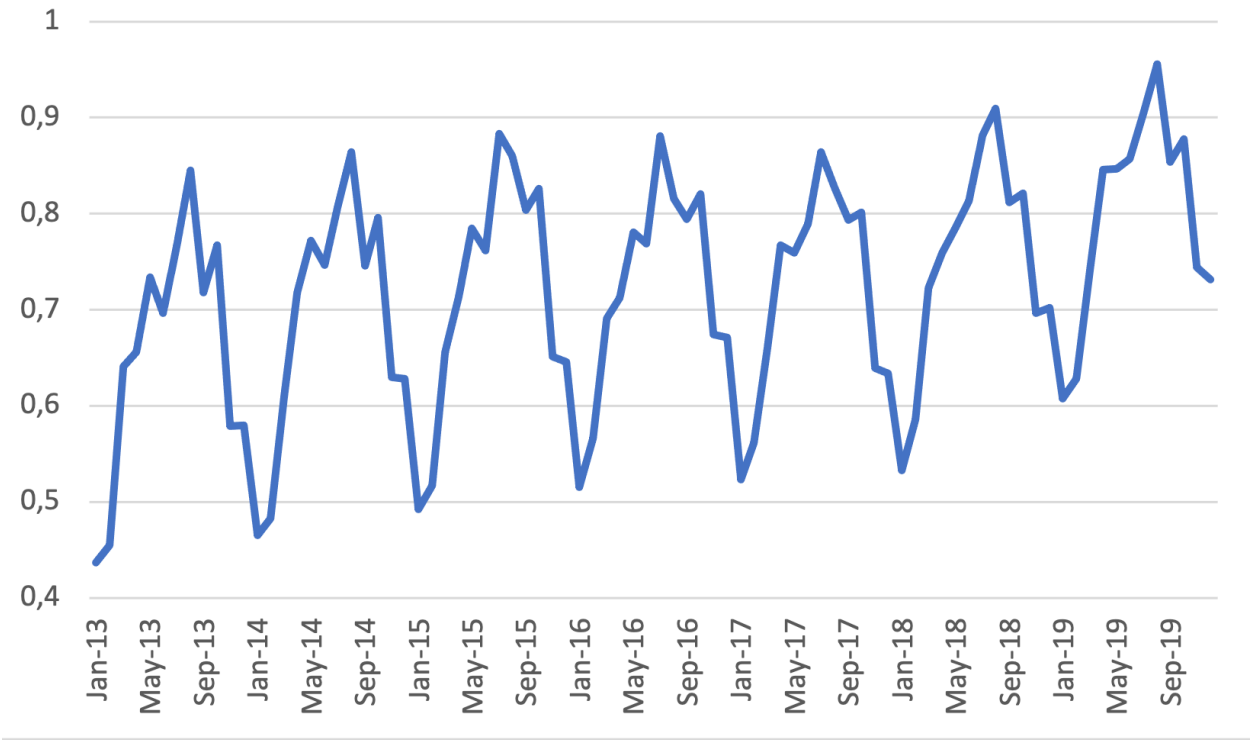


Figure 2. Monthly tourists' nights in Berlin

We hypothesize that there is a limit that tourists can tolerate other tourists' presence in the city. To test for this hypothesis, we follow Tokarchuk et al. (forthcoming) approach and assume that the impact of tourists' presence on the positive (negative) sentiment associated with visiting Berlin's attractions is nonlinear. Accordingly, we estimate the following model:

$$Positive\ index_t = Tour\_nights\_res_t^2 \beta_1 + Tour\_nights\_res_t \beta_2 + Dummy\_year_j + \varepsilon_t$$

$$Negative\ index_t = Tour\_nights\_res_t^2 \beta_1 + Tour\_nights\_res_t \beta_2 + Dummy\_year_j + \varepsilon_t$$

Where $Positive\ index_t$ is the value of index of positive emotions at month t, $Negative\ index_t$ is the value of index of negative emotions at month t, $Tour\_nights\_res_t^2$ is the quadratic term of tourist nights per resident in month t, $Tour\_nights\_res_t$ is the value of tourists nights per resident in month t, $Dummy\_year_j$ is a dummy for each year j from 2014 to 2019, ε_t is an error term.

Hypothesis 2a will be confirmed if coefficient β_1 is significant and negative, while coefficient β_2 is significant and positive. These values of coefficients correspond to an inverted parabola.

Hypothesis 2b will be confirmed if coefficient β_1 is significant and positive, while coefficient β_2 is significant and negative. These values of coefficients correspond to a parabola.

In case hypotheses 2a and 2b are confirmed, the threshold value can be calculated as the axis of the parabola

$$Threshold\ level = -\frac{\beta_1}{2\beta_2}$$

Results

Table 1 reports the results of the estimation. The coefficients corresponding to the linear and quadratic term of tourists nights per resident are significant both for positive index and negative index. This result confirms hypotheses 1a and 1b.

In the case of a positive index, the coefficient β_1 is negative and positive. This performance indicates that the relationship between tourists' positive emotions in their reviews on Tripadvisor and the crowdedness registered in the month of their visit, measured as tourists nights per resident, follow an inverted u-shape pattern. Accordingly, Hypothesis 2a is confirmed. Given the estimated values of the coefficients, the level of monthly tourists intensity that maximizes tourists' positive emotions is 0.67 tourists' nights per resident.

The sign of the coefficient β_1 corresponding to negative emotions index presents a positive sign, while coefficient β_2 is negative. This evidence suggests that negative emotions of tourists expressed in their reviews on Tripadvisor follow a u-shape pattern. The optimal value of tourism intensity that minimizes the negative emotions of tourists visiting Berlin is 0.62 tourists' nights per resident. This result confirms hypothesis 2b.

	Positive index	Negative index
Tourists nights per resident quadratic	-1.37** (0.5572)	0.55** (0.2417)
Tourists nights per resident	1.85** (0.7721)	-0.68** (0.3349)
Dummy_2014	0.07** (0.0328)	-0.05*** (0.0142)
Dummy_2015	0.11*** (0.0330)	-0.08*** (0.0143)
Dummy_2016	0.23*** (0.0332)	-0.12*** (0.0144)
Dummy_2017	0.26*** (0. 0331)	-0.14*** (0.0143)
Dummy_2018	0.12*** (0.0335)	-0.16*** (0.0145)
Dummy_2019	0.10** (0.0348)	-0.18*** (0.0151)
Constant	1.81*** (0.2622)	0.87*** (0.1137)
Number of observations, months	84	84
R^2	0.5750	0.7679

Table 1. Results of estimation

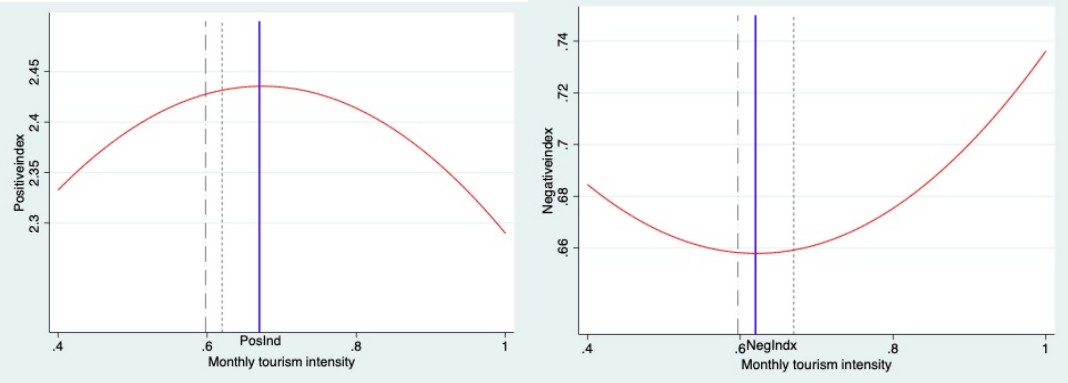
** correspond to significance level at 0.05, *** correspond to significance level at 0.001

The threshold for the negative index is lower than for the positive index. It indicates that positive feelings prevail even when negative emotions start increasing due to crowd related issues, like a

violation of personal space, queues or robbery. Tourists still enjoy the destination even if their visit is not perfect.

While the threshold is an essential indicator for measuring psychological carrying capacity, it is crucial to analyze the whole interval of possible values that the indices can take during the analyzed period. Figure 3a and figure 3b represent the relationship between positive and negative indices and monthly tourism intensity observed in Berlin during the analyzed period. Figure 3a suggests that the values of positive index observed for the lowest levels of crowdedness are similar to the index values corresponding to the highest levels of crowdedness. Although the positive index decreases after reaching the optimal value, with the increase in tourism intensity, the positive index's corresponding value does not go below the values of the index observed for the lowest levels of crowdedness. Only in August 2019, when tourism intensity in Berlin registered its peak of 0.95 tourists' nights per resident, this value got inferior to the value recorded in January 2013, corresponding to the lowest level of tourism intensity of 0.43 tourists' nights per resident. This finding suggests that tourists visiting Berlin enjoy the crowd.

At the same time, figure 3b demonstrates that for negative emotions increase in tourism intensity beyond the threshold level increases negative emotions. Accordingly, the peak months in 2016 and 2017, five months in 2018 and seven months in 2019, registered values of negative emotions' index considerably higher than the index observed for the lowest level of crowdedness in January 2013.



3a **3b**

Figure 3a. Positive index trend in relation to tourism intensity. The solid line corresponds to the optimal level of tourism intensity calculated for positive index, the short-dash line corresponds to the optimal level of tourism intensity calculated for negative index, the long-dash line corresponds to the optimal level of tourism intensity calculated for residents calculated based on results from Tokarchuk et al. (forthcoming).

Figure 3b. Negative index trend in relation to tourism intensity. The solid line corresponds to the optimal level of tourism intensity calculated for the negative index, the short-dash line corresponds to the optimal level of tourism intensity calculated for positive index, the long-dash line corresponds to the optimal level of tourism intensity calculated for residents calculated based on results from Tokarchuk et al. (forthcoming).

Finally, the present study allows a comparison between the levels of carrying capacity of residents observed in Tokarchuk et al. (forthcoming) and one of tourists. The analysis in Tokarchuk et al. is conducted at the district level involving annual tourism data. Residents' carrying capacity reached the threshold between 2011 and 2012. The yearly tourism intensity in Berlin in 2012 was 7.17

tourists' nights per resident in a year, which in terms of monthly tourism intensity corresponds to 0.59 tourists' nights per resident. Accordingly, residents appear to demonstrate less tolerance to the city's tourism intensity compared to what tourists seem to express towards the crowding in the destination.

Conclusion and Discussion

In the present study, we proposed an effective method of measuring destination carrying capacity on the demand side based on sentiment analysis of tourists' reviews on TripAdvisor. Based on the sentiment analysis technique, we constructed an index of positive and negative emotions. This index is then regressed on the tourism intensity in Berlin's city during the analyzed period from January 2013 to December 2019. Our results demonstrate that destinations can use the emotional index for measuring the carrying capacity of the destination. Our analysis indicates that positive emotions, including anticipation, joy, surprise, trust, resist the increasing crowding perception.

Moreover, the positive emotions' index suggests that Berlin's tourists require some crowd to fully enjoy the destination. On one side, this result may indicate that Berlin visitors inform themselves about the destination from TripAdvisor and form an expectation of a crowd. Consequently, their positive feelings are not harmed by the crowdedness because they were expecting to find it and were ready to face it. On the other hand, these findings can demonstrate that with time, larger crowds in the destination attract visitors that enjoy crowds. Thus, the evolution of a mass destination may be taking place.

Negative emotions like anger, disgust, fear and sadness, seem to follow a different pattern. The negative index reaches its minimum level at a lower value of tourism intensity than the positive index and increases drastically with the increase of tourists' nights per resident. This evidence underlines the negative consequences that overcrowding at the destination has on visitors' experience.

Destination managers should carefully consider the implications of the present study. While the positive index analysis suggests an optimistic view of the situation, the negative index's pattern presents warning signs that should be taken seriously. Berlin's negative index was growing fast in the last years. In 2019, the negative index values for high tourism intensity exceeded the negative index values corresponding to the lowest intensity for seven months, from April to October. This evidence suggests that at the present state of destination development, tourism intensity is reaching unacceptable levels, and further growth without intervention could damage the destination image. Destination managers need to understand the sources of this negativity. For instance, the same reviews on TripAdvisor can be analyzed with textual analysis to individuate the most frequent topics of complains and tourists' segments that are mostly affected by them. Based on this information, informed decisions on the type of intervention could be made.

Finally, for the first time in the literature, a comparison of carrying capacity on demand and supply sides has been performed. The results indicate a similar level of carrying capacity for residents and tourists. Residents' carrying capacity suggests lower levels of acceptance of tourists' crowds than tourists' levels seem to tolerate.

The innovative approach to measuring destination carrying capacity on tourists' side provides several advantages. First of all, it allows destination managers to obtain a point estimate of the carrying capacity and see its development in time. This measure can be built at any level of destination for which TripAdvisor reviews are collected. Consequently, it can be constructed for

the whole destination, as in the case presented in this analysis, or for its single sightseeing spots. Second, the use of this measure avoids costly ad-hoc data collection through interviews with tourists that is required for methods developed in other studies. This analysis allows to improve the guest experience as well as save budgets for more critical interventions. Finally, this method provides a more objective measure of tourists' sentiments. Tourists write their reviews to share their experience in the destination and not answer specifically designed questions to understand their behaviour.

We believe that the method developed in the present study will help advance academic research in the area of overtourism and, more largely, destination management. Implementation of this method opens for many applications, including confronting the carrying capacity of different destinations and studying their evolution over time.

The present study faces some limitations. The carrying capacity's estimate is based on reviews from TripAdvisor. People tend to share negative reviews on social media. This behavior might lead to overestimation of the negative index. On the other hand, automatic textual analysis is not able to distinguish scathing reviews. Thus, it is possible that some negative emotions in reviews expressed by sarcastic positive tone may not have been detected, which will overestimate the positive index. Keeping in mind these possible distortions in the index, the reliance of the analysis on many voluntarily published reviews makes its results precious for decision-makers and academics.

References (note – the references in this sample are for demonstration purposes only)

Alazaizeh, M. M., Hallo, J. C., Backman, S. J., Norman, W. C., & Vogel, M. A. (2016). Crowding standards at Petra Archaeological Park: a comparative study of McKercher's five types of heritage tourists. Journal of Heritage Tourism, 11(4), 364-381.

Aslam, F., Awan, T. M., Syed, J. H., Kashif, A., & Parveen, M. (2020). Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. Humanities and Social Sciences Communications, 7(1), 1-9.

Bell, C. M., Needham, M. D., & Szuster, B. W. (2011). Congruence among encounters, norms, crowding, and management in a marine protected area. Environmental Management, 48(3), 499-513.

Bellini, N., Go, F. M., & Pasquinelli, C. (2017). Urban tourism and city development: Notes for an integrated policy agenda. In Tourism in the City (pp. 333-339). Springer, Cham.

Bose, R., Dey, R. K., Roy, S., & Sarddar, D. (2020). Sentiment Analysis on Online Product Reviews. In Information and Communication Technology for Sustainable Development (pp. 559-569). Springer, Singapore.

Butler, R. W., & Dodds, R. (Eds.). (2019). Overtourism: Issues, realities and solutions. Berlin: De Gruyter.

Butler, R. W. (1980). The concept of a tourist area cycle of evolution: implications for management of resources. Canadian Geographer/Le Géographe canadien, 24(1), 5-12.

Capocchi, A., Vallone, C., Pierotti, M., & Amaduzzi, A. (2019). Overtourism: A literature review to assess implications and future perspectives. Sustainability, 11(12), 3303.

Cohen, E. (1979). Rethinking the sociology of tourism. Annals of tourism research, 6(1), 18-35.

Doxey, G. A. (1975). A causation theory of visitor–resident irritants, methodology and research inferences. The impact of tourism. In Proceedings of the Sixth Annual Conference Proceedings of the Travel Research Association, San Diego, CA, USA, 8–11 September 1975.

Fagbola, T. M., & Thakur, S. C. Lexicon-based Bot-aware Public Emotion Mining and Sentiment Analysis of the Nigerian 2019 Presidential Election on Twitter. International Journal of Advanced Computer Science and Applications, 10(10), 329-336.

Forster, J. (1964). The sociological consequences of tourism. International journal of Comparative sociology, 5(2), 217-227.

García, F. A., Vázquez, A. B., & Macías, R. C. (2015). Resident's attitudes towards the impacts of tourism. Tourism Management Perspectives, 13, 33-40.

Gonson, C., Pelletier, D., & Alban, F. (2018). Social carrying capacity assessment from questionnaire and counts survey: Insights for recreational settings management in coastal areas. Marine Policy, 98, 146-157.

Gonzalez, V. M., Coromina, L., & Gali, N. (2018). Overtourism: residents' perceptions of tourism impact as an indicator of resident social carrying capacity-case study of a Spanish heritage town. Tourism Review.

Gössling, S., McCabe, S., & Chen, N. C. (2020). A socio-psychological conceptualisation of overtourism. Annals of Tourism Research, 84, 102976.

Graymore, M. L., Sipe, N. G., & Rickson, R. E. (2010). Sustaining human carrying capacity: a tool for regional sustainability assessment. Ecological economics, 69(3), 459-468.

Jacobsen, J. K. S., Iversen, N. M., & Hem, L. E. (2019). Hotspot crowding and over-tourism: Antecedents of destination attractiveness. Annals of Tourism Research, 76, 53-66.

Joardar, S. (1996). Carrying capacity based planning for cities - Concept and procedure. Space,11(2), 7-16.

Jurado, E. N., Tejada, M. T., García, F. A., González, J. C., Macías, R. C., Peña, J. D., Gutiérrez, F. G., Fernández, G. F., Gallego, M. L., García, G. M., Gutiérrez, O. M., Concha, N. F., de la Rúa, R. F., Sinoga, J. R., & Becerra, F. S. (2012). Carrying capacity assessment for tourist destinations. Methodology for the creation of synthetic indicators applied in a coastal area. Tourism Management, 33(6), 1337-1346.

Jurado, E. N., Damian, I. M., & Fernández-Morales, A. (2013). Carrying capacity model applied in coastal destinations. Annals of Tourism Research, 43, 1-19.

Kessler, J. J. (1994). Usefulness of the human carrying capacity concept in assessing ecological sustainability of land-use in semi-arid regions. Agriculture, Ecosystems & Environment, 48(3), 273-284.

Kim, D., Lee, C. K., & Sirgy, M. J. (2016). Examining the differential impact of human crowding versus spatial crowding on visitor satisfaction at a festival. Journal of Travel & Tourism Marketing, 33(3), 293-312.

Klanjšček, J., Geček, S., Marn, N., Legović, T., & Klanjšček, T. (2018). Predicting perceived level of disturbance of visitors due to crowding in protected areas. PloS one, 13(6), e0197932.

Koens, K., Postma, A., & Papp, B. (2018). Is overtourism overused? Understanding the impact of tourism in a city context. Sustainability, 10(12), 4384.

Kuščer, K., & Mihalič, T. (2019). Residents' attitudes towards overtourism from the perspective of tourism impacts and cooperation - The case of Ljubljana. Sustainability, 11(6), 1823.

Li, L., Zhang, J., Nian, S., & Zhang, H. (2017). Tourists' perceptions of crowding, attractiveness, and satisfaction: a second-order structural model. Asia Pacific Journal of Tourism Research, 22(12), 1250-1260.

Liu, R. Z., & Borthwick, A. G. (2011). Measurement and assessment of carrying capacity of the environment in Ningbo, China. Journal of Environmental Management, 92(8), 2047-2053.

MacCannell, D. (1976). The tourist: A new theory of the leisure class. New York: Schocken Books

Machleit, K.A., Eroglu, S.A., & Mantel, S.P. (2000). Perceived retail crowding and shopping satisfaction: What modifies this relationship? Journal of Consumer Psychology, 9(1), 29–42.

Manning, R. E., Valliere, W. A., & Wang, B. (1999). Crowding norms: Alternative measurement approaches. Leisure sciences, 21(2), 97-115.

McCool, S. F., & Lime, D. W. (2001). Tourism carrying capacity: tempting fantasy or useful reality?. Journal of Sustainable Tourism, 9(5), 372-388.

McKercher, B., Wang, D., & Park, E. (2015). Social impacts as a function of place change. Annals of Tourism Research, 50, 52-66.

Mehta, R., Sharma, N. K., & Swami, S. (2013). The impact of perceived crowding on consumers' store patronage intentions: Role of optimal stimulation level and shopping motivation. Journal of Marketing Management, 29(7-8), 812-835.

Mohammad, S., Turney, Peter. (2013). Crowdsourcing a Word-Emotion Association Lexicon, Computational Intelligence, 29 (3), 436-465.

Neuts, B., & Nijkamp, P. (2012). Tourist crowding perception and acceptability in cities: An applied modelling study on Bruges. Annals of Tourism Research, 39(4), 2133-2153.

Novy, J. (2016). The selling (out) of Berlin and the de-and re-politicization of urban tourism in Europe's 'Capital of Cool'. Protest and resistance in the tourist city, 52.

Oh, K., Jeong, Y., Lee, D., Lee, W., & Choi, J. (2005). Determining development density using the urban carrying capacity assessment system. Landscape and Urban Planning, 73(1), 1-15.

Peeters, P., Gössling, S., Klijs, J., Milano, C., Novelli, M., Dijkmans, C., Postma, A. (2018). Research for TRAN committee - Overtourism: Impact and possible policy responses. Brussels: European Parliament, Policy Department for Structural and Cohesion Policies.

Pizam, A. (1978). Tourism's impacts: The social costs to the destination community as perceived by its residents. Journal of travel research, 16(4), 8-12.

Rathnayake, R. M. W. (2015). How does 'crowding'affect visitor satisfaction at the Horton Plains National Park in Sri Lanka?. Tourism Management Perspectives, 16, 129-138.

Rosenow, J. E., & Pulsipher, G. L. (1979). Tourism the good, the bad, and the ugly. Lincoln: Century Three Press.

Sarma, A. K., et al. (2012). Urban carrying capacity: Concept and calculation. Guwahati, Assam, India.

Saveriades, A. (2000). Establishing the social tourism carrying capacity for the tourist resorts of the east coast of the Republic of Cyprus. Tourism Management, 21(2), 147-156.

Schmidt, D. E., & Keating, J. P. (1979). Human crowding and personal control: An integration of the research. Psychological Bulletin, 86(4), 680.

Simancas Cruz, M., & Penarrubia Zaragoza, M. P. (2019). Analysis of the accommodation density in coastal tourism areas of insular destinations from the perspective of overtourism. Sustainability, 11(11), 3031.

Smith, V. L. (1989). Hosts and guests: The anthropology of tourism. Philadelphia: University of Pennsylvania Press

Statista, (2020). https://www.statista.com/statistics/314340/leading-european-city-tourism-destinations-by-number-of-bednights/, consulted on February, 22nd , 2021.

Tokarchuk, O., Gabriele, R., & Maurer, O. (2020). Estimating tourism social carrying capacity. Annals of Tourism Research, 102971.

UNWTO (2018). Overtourism? Understanding and Managing Urban Tourism Growth beyond Perceptions; UNWTO: Madrid, Spain, 2018.

Van der Borg, J., Costa, P., & Gotti, G. (1996). Tourism in European heritage cities. Annals of Tourism Research, 23(2), 306-321.

Varkaris, E., & Neuhofer, B. (2017). The influence of social media on the consumers' hotel decision journey. Journal of Hospitality and Tourism Technology, 8, 101-118.

Wagar, J. A. (1964). The carrying capacity of wild lands for recreation. Forest Science, 10(suppl_2), a0001-24.

Wei, Y., Huang, C., Lam, P. T., & Yuan, Z. (2015). Sustainable urban development: A review on urban carrying capacity assessment. Habitat International, 46, 64-71.

Yu, J., & Egger, R. (2021). Tourist Experiences at Overcrowded Attractions: A Text Analytics Approach. In Information and Communication Technologies in Tourism 2021 (pp. 231-243). Springer, Cham.

Zhang, Y., Li, X. R., Su, Q., & Hu, X. (2017). Exploring a theme park's tourism carrying capacity: A demand-side analysis. Tourism Management, 59, 564-578.

Zhang, Y., Su, Q., & Hu, X. (2012). Study on the capacity of Fantawild Adventure theme park based on queuing theory. Tourism Tribune, 27(1), 61-72.