



Research Note

Booking.com: The unexpected scoring system

Juan Pedro Mellinas^{*}, Soledad-María Martínez María-Dolores, Juan Jesús Bernal García

Departament of Métodos Cuantitativos e Informáticos, Universidad Politécnica de Cartagena, Calle Real, 3, CP. 30201, Cartagena, Murcia, Spain

H I G H L I G H T S

- 14 papers proposing that Booking.com uses a 0–10 or 1–10 scale were detected.
- However, they are mistaken, as it actually uses a 2.5–10 scale.
- This error resulted in inaccuracies in the results and conclusions.
- Inflated scores could lead customers to wrong perceptions about hotel quality.
- Booking changed the denomination system to avoid call “Pleasant”, “Passable” or “OK” the worst rated hotels.

A R T I C L E I N F O

Article history:

Received 16 February 2014

Accepted 27 August 2014

Available online

Keywords:

Booking.com

Scores

Reviews

Scale

A B S T R A C T

Academic researchers in the hospitality industry found in Booking.com an excellent source of information, as it collects millions of hotel reviews in a rapid, inexpensive and convenient manner. They proposed that Booking.com uses a scoring system with a traditional scale of 0–10 or 1–10. However, they are mistaken, as it actually uses a 2.5–10 scale. This error may cause statistical inaccuracies when using this database for research. It also helps to inflate scores to the higher end of the scale.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The existence of millions of Internet user-generated opinions on specific issues provides researchers with valuable information from self-interviewed individuals. Ethnography researchers have paid specific attention to the Internet as a valid data source for academic research (Hine, 2000; Kozinets, 2002, 2006). Hine (2000) uses the term “virtual ethnography” to refer to this methodology, whereas Kozinets (2002) defines the term “netnography” to refer to a similar concept.

The case of hotel reviews is one of the clearest examples of this new reality. Websites such as TripAdvisor.com and Booking.com have been used in studies as a large database of traveller's reviews. We focused on the Booking.com scoring system to expose how such sites actually works to avoid errors made by researchers over recent years.

2. Literature

The existing literature shows that authors are stating that Booking.com uses a 0–10 scale (de Albornoz, Plaza, Gervás, & Díaz, 2011; Bjørkelund, Burnett, & Nørvåg, 2012; Estárico, Medina, & Marrero, 2012; Gal-Oz, Grinshpoun, & Gudes, 2010; Grinshpoun, Gal-Oz, Meisels, & Gudes, 2009) or a 1–10 scale (Chaves, Gomes, & Pedron, 2012; Costantino, Martinelli, & Petrocchi, 2012a, 2012b; Filieri & McLeay, 2014; Korfiatis & Poulos, 2013; Martínez María-Dolores, Bernal García, & Mellinas, 2012; Plata-Alf, 2013; Yacouel & Fleischer, 2012). An interesting compilation of online customer reviews (Trenz & Berger, 2013) contributes to propagate these theories, which cite one study (Chaves et al., 2012).

Only one of the 14 articles found referred to a minimum score of 2.5 in hotel reviews (Bjørkelund et al., 2012): “It seems that Booking.com does not have any reviews with scores below 2.5 ... scores are somewhat inflated toward the higher end of the scale ... 662,991 reviews from Booking.com, the lowest score from this set was 2.5”. The authors conclude that these data do not make any sense: “However, one would still assume that some very angry customers would rate all the subcriteria with the bottom score, at least when checking

^{*} Corresponding author. Tel.: +34 968 325780; fax: +34 968 325745.

E-mail addresses: losmellinas@yahoo.es (J.P. Mellinas), soledad.martinez@upct.es (S.-M. Martínez María-Dolores), juanjesus.bernal@upct.es (J.J. Bernal García).

How would you rate the hotel for the following?

	poor	fair	good	excellent
Hotel staff	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Services / facilities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Cleanliness of your room	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Comfort	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Value for money	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Location	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fig. 1. Booking.com satisfaction survey.

Source: Booking.com.

hundreds of thousands of reviews". They try to explain this anomaly: "It is therefore possible that Booking.com employ some sort of filtering mechanism, and ratings with all negative scores are regarded as spam."

3. Methods

To understand how the scoring system works, we wrote real reviews on the website. We stayed in two Spanish establishments in 2011 and four Canadian establishments in 2013, all of which were booked using Booking.com.

Every customer received an e-mail from Booking.com that asked for his/her opinion about the experience in the contracted hotel. The customer was asked to consider six aspects, but instead of a numerical scale (0–10 or 1–10), they had to choose between four grades for every aspect (Fig. 1).

We analysed a sample of 1440 Spanish coastal hotels that have 185,802 reviews (at least five reviews per hotel) to determine how the scores were statistically distributed. The average number of reviews per hotel was 129, and 89% of the establishments have 20 or more registered reviews. The data were extracted in November 2011 using public data obtained from Booking.com.

4. Data analysis

A few days after the survey was completed, the website published the review using numerical values. Table 1 shows the hotels that were reviewed, when and how the surveys were filled out and the results published by Booking.com.

The final score published by Booking.com was not given by the reviewer but was automatically calculated by the website from the rates assigned by the reviewer to the six aspects using a plane average, such that every aspect had the same weight (Costantino et al., 2012a; de Albornoz et al., 2011).

We observed that equivalents of the two scales should be the following:

Poor = 2.5, Fair = 5, Good = 7.5, Excellent = 10.

What apparently appeared to be a scale of 0–10 or 1–10 was actually a scale of 2.5–10. The system has been functioning in this manner since at least 2011 and until 2013.

Table 1

Booking.com completed survey and the published results.

Date	Hotel	Clean	Comfort	Location	Facilities	Staff	Value	Score
Oct '11	PRINCESA GALIANA TOLEDO	Exc	Exc	Good	Good	Good	Exc	8.8
Nov '11	VIK GRAN HOTEL COSTA SOL	Poor	Fair	Good	Fair	Fair	Fair	5
Aug '13	TRAVELODGE MONTREAL CENT	Exc	Good	Exc	Good	Good	Good	8.3
Aug '13	LAWRENCE COLLEGE BROCKVILLE	Poor	Good	Fair	Fair	Good	Fair	5.4
Aug '13	ALEXANDRA HOTEL TORONTO	Good	Good	Good	Good	Good	Good	7.5
Aug '13	DIPLOMAT INN NIAGARA FALLS	Good	Good	Good	Good	Exc	Exc	8.3

Source: Self elaboration and Booking.com.

Table 2

Sample distribution.

Denomination	Review score	Nr of hotels	% s/total
Poor	2.5–3.9	0	0%
Disappointing	4.0–4.9	1	0.07%
Passable	5.0–5.5	1	0.07%
OK	5.6–5.9	6	0.42%
Pleasant	6.0–6.9	98	6.81%
Good	7.0–7.9	697	48.40%
Very Good	8.0–8.5	476	33.06%
Fabulous	8.6–8.9	125	8.68%
Superb	9.0–9.4	34	2.36%
Exceptional	9.5–10	2	0.14%

Source: Compiled from data obtained from Booking.com.

Booking.com uses a subjective denomination for every score range. The results of our Spanish hotels database are shown on Table 2.

By 2014, Booking.com changed the denomination system. Scores under 7.1 did not show a denomination; they only showed a number (Fig. 2). Thus the "Poor", "Disappointing", "Passable", "OK" and "Pleasant" denominations have disappeared. These denominations make sense because it is not appropriate to score a hotel with a large number of negative reviews as "Pleasant". However, it does not prevent low-quality hotels from obtaining acceptable scores above 6.

This 2.5–10 scale largely explains why, in the vast majority of cases, the ratings of the hotels were above 7. If a customer was completely unsatisfied with some aspect of the hotel, then he had no option to mark a 0 or 1 because it was only possible to choose "Poor", which equals to 2.5. However, if a customer was very satisfied, then there was no option to choose an 8 or 9 because choosing "Excellent" automatically rates the aspect as a score of 10.

Using this scoring system there were practically no hotels with bad ratings. More than 93% of the hotels had a score of 7 or more points, which in a traditional 0–10 scale is commonly considered as a medium–high score. Other authors (Bjørkelund et al., 2012) had similar observations when using the same database: "... scores are somewhat inflated toward the higher end of the scale ..."

5. Conclusions

The 14 publications analysed in this study were wrong to assume that Booking.com uses a traditional scale of 0–10 or 1–10. Only one study (Bjørkelund et al., 2012) realized that there was something strange in this scale but failed to explain why the minimum score was 2.5. Usually this error had no relevant effect on a study's conclusions. However, sometimes when this database was used for statistical analysis, this error resulted in inaccuracies in the results and conclusions.

It appears that Booking.com is trying to inflate these scores, which ensure that nearly all of the hotels are scored as "medium" or



Fig. 2. Booking.com denominations.
Source: Booking.com.

“high” quality hotels. It is very difficult for an average Internet user to appropriately judge a hotel, as they would not suspect from the rating scale on the website that a seemingly well-rated hotel might actually be one of the worst hotels in the area. This scoring system does not appear to be illegal but could indicate a lack of honesty with customers. Moreover, recommending hotels with a better score will stimulate Booking.com sales.

Thus, this peculiar scale should not be a problem for an academic researcher, who understands how the system works. We hope that our contribution will help to prevent this error in future studies using Booking.com. Further investigations should confirm the effects of this scoring system by quantifying how it inflates values.

Acknowledgements

This research was partially supported by the Faculty of Business Studies at the Technical University of Cartagena (UPCT).

References

- de Alborno, J. C., Plaza, L., Gervás, P., & Díaz, A. (2011). A joint model of feature mining and sentiment analysis for product review rating. In P. Clough, C. Foley, C. Gurrin, H. Lee, & G. J. F. Jones (Eds.), *Proceedings of the 33rd European conference on Advances in information retrieval (ECIR'11)* (pp. 55–66). Berlin, Heidelberg: Springer-Verlag.
- Björkelund, E., Burnett, T. H., & Nørvåg, K. (2012). A study of opinion mining and visualization of hotel reviews. In *Proceedings of the 14th International Conference on Information Integration and Web-based Applications & Services* (pp. 229–238). ACM.
- Chaves, M. S., Gomes, R., & Pedron, C. (2012). Analysing reviews in the Web 2.0: small and medium hotels in Portugal. *Tourism Management*, 33(5), 1286–1287.
- Costantino, G., Martinelli, F., & Petrocchi, M. (2012a). Priorities-based review computation. In *AAAI Spring Symposium, 2012 1st Workshop on Intelligent Web Services Meet Social Computing* (Vol. 12, p. 04).
- Costantino, G., Morisset, C., & Petrocchi, M. (2012b). Subjective review-based reputation. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing* (pp. 2029–2034). ACM.
- Estárico, E. H., Medina, L. F., & Marrero, S. M. (2012). Una aproximación a la reputación en línea de los establecimientos hoteleros españoles. *Papers de Turisme*, 52, 63–88.
- Filieri, R., & McLeay, F. (2014). E-WOM and accommodation an analysis of the factors that influence travelers' adoption of information from online reviews. *Journal of Travel Research*, 53(1), 44–57.
- Gal-Oz, N., Grinshpoun, T., & Gudes, E. (2010). Sharing reputation across virtual communities. *Journal of Theoretical and Applied Electronic Commerce Research*, 5(2), 1–25.
- Grinshpoun, T., Gal-Oz, N., Meisels, A., & Gudes, E. (2009). CCR: a model for sharing reputation knowledge across virtual communities. In *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology* (Vol. 01, pp. 34–41). IEEE Computer Society.
- Hine, C. (2000). *Virtual ethnography*. London: Sage.
- Korfatis, N., & Poulos, M. (2013). Using online consumer reviews as a source for demographic recommendations: a case study using online travel reviews. *Expert Systems with Applications*, 40(14), 5507–5515.
- Kozinets, R. (2002). The field behind the screen: using netnography for marketing research in online communities. *Journal of Marketing Research*, 39, 61–72.
- Kozinets, R. (2006). Click to connect: netnography and tribal advertising. *Journal of Advertising Research*, 46, 279–288.
- Martínez María-Dolores, S. M., Bernal García, J. J., & Mellinas, J. P. (2012). Los hoteles de la región de Murcia ante las redes sociales y la reputación online. *Revista de Análisis Turístico*, 13, 1–10.
- Plata-Alf, D. (2013). Marketing communications in a virtual environment – opportunities and challenges for companies in the tourism sector. In A. Nalepka, & A. Ujwary-Gil (Eds.), *Business and non-profit organizations facing increased competition and growing customers' demands*. Wyższa Szkoła Biznesu-National-Louis University.
- Trenz, M., & Berger, B. (2013). Analyzing online customer reviews-an interdisciplinary literature review and research agenda. In *Proceedings of the 21st European Conference on Information Systems (ECIS), June 5-8 2013, Utrecht, Netherlands* (p/pp. 83).
- Yacouel, N., & Fleischer, A. (2012). The role of cybermediaries in reputation building and price premiums in the online hotel market. *Journal of Travel Research*, 51(2), 219–226.

Juan Pedro Mellinas is Ph.D. student at Department of Quantitative Methods and Computing (Faculty of Business Studies), Technical University of Cartagena (UPCT), Cartagena, Spain. Degree in administration and management (Universidad de Murcia). Master in Tourism (UPCT).



Soledad María Martínez María-Dolores is Professor at Department of Quantitative Methods and Computing (Faculty of Business Studies), Technical University of Cartagena (UPCT), Cartagena, Spain. Degree in Business and economy. Ph.D in administration and management.



Juan Jesús Bernal García is Professor at Department of Quantitative Methods and Computing (Faculty of Business Studies), Technical University of Cartagena (UPCT), Cartagena, Spain. Director of Department of Quantitative Methods and Computing.

