Research Note

Booking.com: The unexpected scoring system

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

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

HIGHLIGHTS

- 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 calling the worst rated hotels.

ARTICLE INFO

Article history:
Received 16 February 2014
Accepted 27 August 2014
Available online 12 March 2015

Keywords:
Booking.com
Scores
Reviews
Scale

ABSTRACT

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; Estérico, 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; Korfias & 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
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 or more registered reviews. The data were extracted in November 2011 using public data obtained from Booking.com.

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.

Table 1 shows the hotels with a large number of negative reviews as “Pleasant”. However, it does not prevent low-quality hotels from obtaining acceptable scores above 6.

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

<table>
<thead>
<tr>
<th>Date</th>
<th>Hotel</th>
<th>Clean</th>
<th>Comfort</th>
<th>Location</th>
<th>Facilities</th>
<th>Staff</th>
<th>Value</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oct ’11</td>
<td>PRINCESA GALIANA TOLEDO</td>
<td>Exc</td>
<td>Exc</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Exc</td>
<td>8.8</td>
</tr>
<tr>
<td>Nov ’11</td>
<td>VIK GRAN HOTEL COSTA SOL</td>
<td>Poor</td>
<td>Fair</td>
<td>Good</td>
<td>Fair</td>
<td>Fair</td>
<td>Fair</td>
<td>5</td>
</tr>
<tr>
<td>Aug ’13</td>
<td>TRAVELODGE MONTREAL CENT</td>
<td>Exc</td>
<td>Good</td>
<td>Exc</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>8.3</td>
</tr>
<tr>
<td>Aug ’13</td>
<td>LAWRENCE COLLEGE BROCKVILLE</td>
<td>Poor</td>
<td>Good</td>
<td>Fair</td>
<td>Good</td>
<td>Fair</td>
<td>5.4</td>
<td></td>
</tr>
<tr>
<td>Aug ’13</td>
<td>ALEXANDRA HOTEL TORONTO</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>Good</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Aug ’13</td>
<td>DIPLOMAT INN NIAGARA FALLS</td>
<td>Good</td>
<td>Good</td>
<td>Exc</td>
<td>Exc</td>
<td>8.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Booking.com.
دانلود مقاله

http://daneshyari.com/article/1011949

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات